

T E C H N I C A L D E E P D I V E

Anki Vector

A LOVE LETTER TO THE LITTLE DUDE

AUTHOR

RANDALL MAAS

OVERVIEW

This book explores how the Anki Vector was realized in hardware and software.



Copyright © 2019-2020 Randall Maas.
All rights reserved.

drawing by Steph Dere

RANDALL MAAS has spent decades in Washington and Minnesota. He consults in embedded systems development, especially medical devices. Before that he did a lot of other things... like everyone else in the software industry. He is also interested in geophysical models, formal semantics, model theory and compilers.

You can contact him at randym@randym.name.

LinkedIn: <http://www.linkedin.com/pub/randall-maas/9/838/8b1>

Table of Contents

<u>ANKI VECTOR.....</u>	<u>1</u>
<u>A LOVE LETTER TO THE LITTLE DUDE</u>	<u>1</u>
<u>PREFACE</u>	<u>1</u>
1. ORGANIZATION OF THIS DOCUMENT.....	1
1.1. ORDER OF DEVELOPMENT	3
1.2. VERSION(s)	3
1.3. CUSTOMIZATION AND PATCHING	4
1.4. CODE NAMES OR VECTOR VS VICTOR	4
<u>CHAPTER 1.....</u>	<u>5</u>
<u>OVERVIEW OF VECTOR.....</u>	<u>5</u>
2. OVERVIEW.....	5
2.1. COMPELLING CHARACTER	5
2.2. FEATURES	6
3. PRIVACY AND SECURITY	8
4. COZMO	8
5. ALEXA INTEGRATION	9
<u>PART I.....</u>	<u>11</u>
<u>ELECTRONICS DESIGN.....</u>	<u>11</u>
<u>CHAPTER 2.....</u>	<u>13</u>
<u>ELECTRONICS DESIGN DESCRIPTION</u>	<u>13</u>
6. DESIGN OVERVIEW.....	13
6.1. POWER SOURCE AND DISTRIBUTION TREE.....	16
7. REFERENCES & RESOURCES	17
<u>CHAPTER 3.....</u>	<u>18</u>
<u>HEAD-BOARD ELECTRONICS DESIGN DESCRIPTION.....</u>	<u>18</u>
8. THE HEAD-BOARD (THE MAIN PROCESSOR BOARD)	18
8.1. THE APQ8009 PROCESSOR	19
8.2. SPEAKERS	19
8.3. CAMERA.....	19
8.4. THE LCD	20
8.5. POWER MANAGEMENT	20
8.6. TRIM, CALIBRATION SERIAL NUMBERS AND KEYS.....	20
8.7. MANUFACTURING TEST CONNECTOR/INTERFACE.....	21
9. REFERENCES & RESOURCES	21

CHAPTER 4.....	22
BACKPACK & BODY-BOARD ELECTRONICS DESIGN DESCRIPTION.....	22
10. THE BACKPACK BOARD.....	22
10.1. BACKPACK CONNECTION	23
10.2. ELECTRO-STATIC DISCHARGE (ESD) PROTECTION	23
10.3. OPERATION	23
11. THE BODY-BOARD	25
11.1. POWER MANAGEMENT.....	26
11.2. ELECTRO-STATIC DISCHARGE (ESD) PROTECTION	29
11.3. STM32F030 MICROCONTROLLER.....	29
11.4. SENSING	30
11.5. OUTPUTS.....	33
11.6. COMMUNICATION.....	34
11.7. CONNECTORS & TEST POINTS	34
12. REFERENCES & RESOURCES	36
CHAPTER 5.....	38
ACCESSORY ELECTRONICS DESIGN DESCRIPTION.....	38
13. CHARGING STATION	38
14. CUBE	39
14.1. OVER THE AIR FIELD UPDATES	40
14.2. REFERENCES & RESOURCES	40
PART II.....	41
BASIC OPERATION	41
CHAPTER 6.....	43
ARCHITECTURE	43
15. OVERVIEW OF VECTOR'S COMMUNICATION INFRASTRUCTURE	43
15.1. APPLICATION SERVICES ARCHITECTURE	44
15.2. EMOTION MODEL, BEHAVIOUR ENGINE, ACTIONS AND ANIMATION ENGINE	46
16. STORAGE SYSTEM	46
16.1. ELECTRONIC MEDICAL RECORDS (EMR)	47
16.2. OEM PARTITION FOR OWNER CERTIFICATES AND LOGS	49
17. SECURITY AND PRIVACY	49
17.1. ENCRYPTED COMMUNICATION	50
17.2. ENCRYPTED FILESYSTEM.....	50
17.3. THE OPERATING SYSTEM	50
17.4. AUTHENTICATION	50
18. CONFIGURATION AND ASSET FILES	51
18.1. CONFIGURATION FILES	51
19. SOFTWARE-HARDWARE LAYERS	52
19.1. THE BODY BOARD INPUT/OUTPUT.....	52
19.2. THE LCD DISPLAY.....	52
19.3. THE CAMERA	53

20. REFERENCES & RESOURCES	53
<u>CHAPTER 7.....</u>	55
STARTUP	55
21. STARTUP	55
21.1. QUALCOMM'S PRIMARY AND SECONDARY BOOTLOADER	55
21.2. ANDROID BOOTLOADER (ABOOT)	56
21.3. REGULAR SYSTEM BOOT.....	57
21.4. ABNORMAL SYSTEM BOOT	59
22. SHUTDOWN	59
22.1. TURNING VECTOR OFF (INTENTIONALLY)	60
22.2. UNINTENTIONALLY	60
22.3. GOING INTO AN OFF STATE	60
23. REFERENCES & RESOURCES	60
<u>CHAPTER 8.....</u>	62
POWER MANAGEMENT.....	62
24. POWER MANAGEMENT	62
24.1. BATTERY MANAGEMENT	62
24.2. RESPONSES, SHEDDING LOAD / POWER SAVING EFFORTS	63
24.3. ACTIVITY LEVEL MANAGEMENT.....	65
25. CHARGING	65
<u>CHAPTER 9.....</u>	66
BASIC INPUTS AND OUTPUTS	66
26. BUTTON, TOUCH AND CLIFF SENSOR INPUT.....	66
26.1. TOUCH SENSING INFORMATION	67
26.2. TIME OF FLIGHT PROXIMITY SENSOR	67
27. BACKPACK LIGHTS CONTROL	67
<u>CHAPTER 10.....</u>	69
INERTIAL MOTION SENSING	69
28. MOTION SENSING	69
28.1. ACCELEROMETER AND GYROSCOPE	69
28.2. TILTED HEAD	70
28.3. SENSING MOTION.....	70
28.4. SENSING INTERACTIONS.....	70
29. REFERENCES AND RESOURCES	71
<u>PART III.....</u>	73
COMMUNICATION	73
<u>CHAPTER 11.....</u>	75

COMMUNICATION	75
30. OVERVIEW OF VECTOR'S COMMUNICATION INFRASTRUCTURE	75
31. INTERNAL COMMUNICATION WITH PERIPHERALS	76
31.1. COMMUNICATION WITH THE BODY-BOARD	76
31.2. SERIAL BOOT CONSOLE	77
31.3. USB	77
32. BLUETOOTH LE	77
33. WIFI.....	79
33.1. FIREWALL.....	80
33.2. WIFI CONFIGURATION	80
34. NETWORK COMMUNICATION.....	80
34.1. COMMUNICATING WITH MOBILE APP AND SDK.....	80
34.2. WEB-VIZ, A VISUAL CHARACTERIZATION TOOL	82
35. CLOUD SERVERS	83
36. REFERENCES & RESOURCES	83
CHAPTER 12.....	84
BLUETOOTH LE COMMUNICATION PROTOCOL.....	84
37. COMMUNICATION PROTOCOL OVERVIEW	84
37.1. SETTING UP THE COMMUNICATION CHANNEL	86
37.2. FRAGMENTATION AND REASSEMBLY	87
37.3. ENCRYPTION SUPPORT	88
37.4. THE RTS LAYER	89
37.5. FETCHING A LOG	90
38. MESSAGE FORMATS	91
38.1. APPLICATION CONNECTION ID.....	92
38.2. CANCEL PAIRING	93
38.3. CHALLENGE	94
38.4. CHALLENGE SUCCESS.....	95
38.5. CLOUD SESSION	96
38.6. CONNECT.....	97
38.7. DISCONNECT.....	98
38.8. FILE DOWNLOAD	99
38.9. LOG.....	100
38.10. NONCE	101
38.11. OTA UPDATE.....	102
38.12. RESPONSE	103
38.13. SDK PROXY.....	104
38.14. SSH.....	105
38.15. STATUS	106
38.16. WIFI ACCESS POINT.....	107
38.17. WIFI CONNECT	108
38.18. WIFI FORGET	109
38.19. WIFI IP ADDRESS	110
38.20. WIFI SCAN	111
CHAPTER 13.....	112

THE HTTPS BASED API	112
39. OVERVIEW OF THE SDK HTTPS API	112
39.1. SDK MESSAGE GROUPINGS	112
40. COMMON ELEMENTS.....	114
40.1. ENUMERATIONS	114
40.2. STRUCTURES.....	115
41. ACCESSORIES AND CUSTOM OBJECTS	117
41.1. ENUMERATIONS	117
41.2. EVENTS	121
41.3. CREATE FIXED CUSTOM OBJECT.....	125
41.4. DEFINE CUSTOM OBJECT	126
41.5. DELETE CUSTOM OBJECTS.....	129
42. ACTIONS AND BEHAVIOUR	130
42.1. ENUMERATIONS	130
42.2. EVENTS	131
42.3. STRUCTURES.....	132
42.4. BEHAVIOR CONTROL AND ASSUME BEHAVIOR CONTROL	135
42.5. CANCEL ACTION BY ID TAG	137
42.6. LOOK AROUND IN PLACE	137
43. ALEXA.....	138
43.1. ENUMERATIONS	138
43.2. EVENTS	138
43.3. ALEXA AUTH STATE.....	139
43.4. ALEXA OPT IN	139
44. ANIMATION	140
44.1. STRUCTURES.....	140
44.2. LIST ANIMATIONS	140
44.3. LIST ANIMATION TRIGGERS	141
44.4. PLAY ANIMATION	141
44.5. PLAY ANIMATION TRIGGER	142
45. ATTENTION TRANSFER	143
45.1. EVENTS	143
45.2. GET LATEST ATTENTION TRANSFER.....	144
46. AUDIO	145
46.1. ENUMERATIONS	145
46.2. EVENTS	146
46.3. APP INTENT.....	148
46.4. AUDIO FEED (FROM THE MICROPHONES).....	149
46.5. AUDIO PROCESSING MODE	150
46.6. EXTERNAL AUDIO STREAM PLAYBACK	151
46.7. MASTER VOLUME	153
46.8. SAY TEXT	154
47. BATTERY.....	155
47.1. ENUMERATIONS	155
47.2. BATTERY STATE	155
48. CONNECTION.....	156
48.1. EVENTS	156
48.2. EVENT STREAM	159

48.3. PROTOCOL VERSION	160
48.4. SDK INITIALIZATION	161
48.5. USER AUTHENTICATION.....	162
48.6. VERSION STATE	163
49. CUBE	164
49.1. ENUMERATIONS	164
49.2. EVENTS	165
49.3. CONNECT CUBE	166
49.4. CUBES AVAILABLE	166
49.5. DISCONNECT CUBE.....	167
49.6. DOCK WITH CUBE.....	168
49.7. FLASH CUBE LIGHTS.....	169
49.8. FORGET PREFERRED CUBE.....	169
49.9. PICKUP OBJECT	170
49.10. PLACE OBJECT ON GROUND HERE	171
49.11. POP A WHEELIE.....	172
49.12. ROLL BLOCK	173
49.13. ROLL OBJECT	174
49.14. SET CUBE LIGHTS	175
49.15. SET PREFERRED CUBE.....	176
50. DIAGNOSTICS.....	177
50.1. CHECK CLOUD CONNECTION	177
50.2. UPLOAD DEBUG LOGS.....	178
51. DISPLAY	179
51.1. EVENTS	179
51.2. DISPLAY IMAGE RGB	179
51.3. ENABLE MIRROR MODE	180
51.4. SET EYE COLOR	180
52. FACES	181
52.1. ENUMERATIONS	181
52.2. STRUCTURES.....	182
52.3. EVENTS	182
52.4. CANCEL FACE ENROLLMENT.....	184
52.5. ENABLE FACE DETECTION.....	185
52.6. ERASE ALL ENROLLED FACES	186
52.7. ERASE ENROLLED FACE BY ID	186
52.8. FIND FACES	187
52.9. REQUEST ENROLLED NAMES.....	188
52.10. SET FACE TO ENROLL.....	189
52.11. UPDATE ENROLLED FACE BY ID	190
53. FEATURES & ENTITLEMENTS	191
53.1. ENUMERATIONS	191
53.2. GET FEATURE FLAG	192
53.3. GET FEATURE FLAG LIST	193
53.4. UPDATE USER ENTITLEMENTS	194
54. IMAGE PROCESSING	195
54.1. ENUMERATIONS	195
54.2. EVENTS	196

54.3. CAMERA FEED.....	197
54.4. CAPTURE SINGLE IMAGE	198
54.5. ENABLE IMAGE STREAMING.....	198
54.6. ENABLE MARKER DETECTION.....	199
54.7. ENABLE MOTION DETECTION	200
54.8. IS IMAGE STREAMING ENABLED.....	200
55. INTERACTIONS WITH OBJECTS	201
55.1. STRUCTURES.....	201
55.2. DRIVE OFF CHARGER	202
55.3. DRIVE ON CHARGER.....	202
55.4. GO TO OBJECT	203
55.5. TURN TOWARDS FACE	204
56. JDOCS.....	205
56.1. ENUMERATIONS	205
56.2. STRUCTURES.....	205
56.3. EVENTS	206
56.4. PULL JDOCS	206
57. MOTION CONTROL.....	207
57.1. DRIVE STRAIGHT.....	207
57.2. DRIVE WHEELS.....	208
57.3. GO TO POSE	209
57.4. MOVE HEAD.....	210
57.5. MOVE LIFT.....	210
57.6. SET HEAD ANGLE.....	211
57.7. SET LIFT HEIGHT	212
57.8. STOP ALL MOTORS	213
57.9. TURN IN PLACE	214
58. MOTION SENSING AND ROBOT STATE.....	215
58.1. STRUCTURES.....	215
58.2. EVENTS	217
59. ON BOARDING	218
59.1. ENUMERATIONS	218
59.2. EVENTS	219
59.3. ONBOARDING INPUT	220
59.4. ONBOARDING STATE	223
59.5. ONBOARDING COMPLETE REQUEST	223
59.6. ONBOARDING WAKE UP REQUEST	223
59.7. ONBOARDING WAKE UP STARTED REQUEST.....	224
60. PHOTOS	225
60.1. STRUCTURES.....	225
60.2. EVENTS	225
60.3. DELETE PHOTO.....	226
60.4. PHOTO	226
60.5. PHOTOS INFO	227
60.6. THUMBNAIL	228
61. SETTINGS AND PREFERENCES	229
61.1. STRUCTURES.....	229
61.2. UPDATE SETTINGS.....	229

61.3. UPDATE ACCOUNT SETTINGS	230
62. SOFTWARE UPDATES	231
62.1. ENUMERATIONS	231
62.2. START UPDATE ENGINE	231
62.3. CHECK UPDATE STATUS.....	231
62.4. UPDATE AND RESTART.....	232
63. HISTORICAL ODDITIES	232
CHAPTER 14.....	233
THE WEB VISUALIZATION PROTOCOL	233
64. WEBSOCKET OVERVIEW	233
64.1. SETTING UP THE COMMUNICATION CHANNEL	234
65. WEBSOCKETS JSON MESSAGES.....	235
65.1. RECEIVED EVENTS	235
65.2. POSTED EVENTS	243
CHAPTER 15.....	244
THE CLOUD SERVICES	244
66. CONFIGURATION	244
67. JDOCS SERVER	245
67.1. JDOCS INTERACTION.....	245
67.2. DELETE DOCUMENT	245
67.3. ECHO TEST	246
67.4. READ DOCUMENTS	246
67.5. READ DOCUMENT ITEM	247
67.6. WRITE DOCUMENT.....	247
68. NATURAL LANGUAGE PROCESSING.....	248
68.1. PARAMETERS FOR THE CLOUD INTENTS	248
69. LOG UPLOADER.....	250
70. CRASH UPLOADER	250
71. DAS MANAGER	251
72. REFERENCES AND RESOURCES	251
PART IV	253
ADVANCED FUNCTIONS	253
CHAPTER 16.....	255
AUDIO INPUT.....	255
73. AUDIO INPUT.....	255
73.1. THE MICROPHONES AND CONVERSION TO AUDIO SAMPLES	256
73.2. SPATIAL AUDIO PROCESSING	258
73.3. NOISE REDUCTION.....	259
73.4. DETECTING ACTIVITY.....	259
73.5. BEAT DETECTION	260
73.6. RECORDING TO A FILE.....	262

73.7. VOICE ACTIVITY DETECTOR AND WAKE WORD.....	262
73.8. CONNECTIONS WITH VIC-GATEWAY AND SDK ACCESS	264
74. CLOUD SPEECH RECOGNITION	265
74.1. INTENT PARAMETERS	266
74.2. INTENT MAPPING CONFIGURATION FILE	268
75. REFERENCES AND RESOURCES	269
CHAPTER 17.....	270
IMAGE PROCESSING.....	270
76. CAMERA OPERATION	270
76.1. CAMERA OPERATION	271
76.2. CAMERA CALIBRATION	271
76.3. CORRECTION.....	272
76.4. VISION MODES	273
76.5. ILLUMINATION LEVEL SENSING.....	274
76.6. VISUAL MOTION DETECTION	274
77. THE CAMERA POSE: WHAT DIRECTION IS CAMERA POINTING IN?.....	275
78. MARKERS	276
78.1. THE INITIAL PREPARATION STEPS.....	277
78.2. DETECT AND ANALYZE SQUARES	277
78.3. DECODING THE SQUARES	278
78.4. REVAMPING SIZE AND ORIENTATION.....	278
78.5. INFERRING KNOWLEDGE ABOUT OBJECTS	278
79. FACE AND FACIAL FEATURES RECOGNITION.....	279
79.1. FACE DETECTION	279
79.2. FACE IDENTIFICATION AND TRAINING	280
79.3. COMMUNICATION INTERFACE	280
80. TENSORFLOW LITE, DETECTING HANDS, PETS... AND THINGS?	281
80.1. DETAILS ON TENSORFLOW LITE	281
80.2. OTHER IDEAS THAT WEREN'T FULLY REALIZED AND FUTURE POTENTIAL	283
81. PHOTOS/PICTURES	284
81.1. COMMUNICATION INTERFACE	284
82. CONFIGURATION FILES	285
82.1. VISION CONFIG	285
82.2. SCHEDULE MEDIATOR CONFIGURATION FILES	291
82.3. PHOTOGRAPHY CONFIGURATION FILES	291
83. RESOURCES & RESOURCES	292
CHAPTER 18.....	294
MAPPING & NAVIGATION	294
84. MAPPING OVERVIEW	294
85. MAP REPRESENTATION	294
85.1. QUAD-TREE MAP REPRESENTATION BASICS.....	295
85.2. THE MAP'S STARTING POINT	295
85.3. HOW THE MAP IS SENT FROM VECTOR TO SDK APPLICATIONS.....	296
86. MEASURING THE DISTANCE TO OBJECTS.....	296

86.1. FILTERING	296
86.2. INTERNAL DATA STRUCTURES	298
87. BUILDING THE MAP	300
87.1. MAPPING CLIFFS AND EDGES.....	300
87.2. TRACKING OBJECTS.....	301
87.3. BUILDING A MAP WITH SLAM	301
88. NAVIGATION AND PLANNING.....	302
89. RESOURCES & RESOURCES	302
CHAPTER 19.....	303
ACCESSORIES	303
90. ACCESSORIES IN GENERAL.....	303
90.1. DOCKING	303
91. HOME & CHARGING STATION	303
91.1. DOCKING	303
92. COMPANION CUBE.....	304
92.1. COMMUNICATION.....	304
92.2. TAP DETECTION	305
93. CUSTOM ITEMS	305
93.1. A FIXED, UNMARKED OBJECT (CUBE-SHAPED).....	305
93.2. CUSTOM WALL DEFINITION	306
93.3. CUSTOM CUBE DEFINITION	306
93.4. CUSTOM BOX DEFINITION.....	307
93.5. COMMUNICATION.....	307
PART V	309
ANIMATION.....	309
CHAPTER 20.....	311
ANIMATION.....	311
94. ANIMATION TRIGGERS AND ANIMATION GROUPS	311
94.1. FILES.....	312
94.2. TRIGGER MAP CONFIGURATION FILES	313
94.3. ANIMATION GROUP FILES	313
95. ANIMATIONS	314
95.1. ANIMATION TRACKS	314
95.2. ANIMATION FILES.....	314
95.3. ANIMATION NAMES MANIFEST	315
96. SDK COMMANDS TO PLAY ANIMATIONS	315
CHAPTER 21.....	316
LIGHTS ANIMATION	316
97. LIGHTS ANIMATION OVERVIEW	316
98. CUBE SPINNER GAME	316
99. BACKPACK LIGHTS ANIMATION	318

99.1. TRIGGER MAP CONFIGURATION FILES	318
99.2. THE BACKPACK LIGHTS PATTERN	318
100. CUBE LIGHTS ANIMATION	319
100.1. TRIGGER MAP CONFIGURATION FILES	319
100.2. CUBE ANIMATIONS	319
CHAPTER 22.....	321
VIDEO DISPLAY & FACE	321
101. OVERVIEW OF THE DISPLAY	321
101.1. ORIGIN.....	321
101.2. RENDERING SYSTEM	322
102. IMAGE LAYOUT, COMPOSITION, AND SPRITE SEQUENCES	322
102.1. MAPPING ANIMATION TRIGGER NAMES TO LAYOUT AND SPRITE SEQUENCES	322
102.2. LAYOUT FILE	323
102.3. IMAGE MAP FILE.....	324
102.4. SPRITE SEQUENCES.....	325
102.5. DISPLAYING TEXT ON THE SCREEN	325
103. PROCEDURAL FACE	326
103.1. THE RENDERING OF INDIVIDUAL EYES	327
103.2. THE PROCESS OF DRAWING THE PROCEDURAL FACE	328
104. COMMANDS	328
CHAPTER 23.....	330
AUDIO PRODUCTION	330
105. SPEAKER.....	330
106. SOUND EFFECTS FROM AUDIO FILES AND PROCEDURES	331
106.1. SOUND PLUGINS.....	332
106.2. AUDIO PIPELINE	333
106.3. PARAMETRIC OR PROCEDURAL SOUND GENERATION	334
106.4. EQUALIZER	334
106.5. THE CONFIGURATION.....	334
106.6. THE SOUND FILES	335
106.7. MAPPING AUDIO EVENT AND SOUND NAMES TO ID NUMBERS.....	336
107. TEXT TO SPEECH.....	336
107.1. THAT DISTINCT ROBOTIC VOICE QUALITY.....	337
107.2. THE CONFIGURATION AND LOCALIZATION FILES	338
107.3. CUSTOMIZATION.....	340
108. COMMANDS	341
109. REFERENCES AND RESOURCES	341
CHAPTER 24.....	343
MOTION CONTROL	343
110. MOTION CONTROL.....	343
110.1. FEEDBACK	344
110.2. MOTOR CONTROL	344

110.3. BURN OUT PROTECTION.....	344
110.4. NO PINCHING FINGERS!	345
110.5. GETTING THE LIFT AND HEAD POSITIONS JUST RIGHT.....	345
110.6. DIFFERENTIAL DRIVE KINEMATICS.....	345
111. MOTION CONTROL COMMANDS	346
CHAPTER 25.....	348
ANIMATION FILE FORMAT.....	348
112. ANIMATION BINARY FILE FORMAT.....	348
112.1. OVERVIEW OF THE FILE FORMAT	348
112.2. RELATIONSHIP WITH COZMO	348
113. STRUCTURES.....	349
113.1. ANIMCLIPS.....	349
113.2. ANIMCLIP	349
113.3. AUDIOEVENTGROUP	349
113.4. AUDIOPARAMETER	350
113.5. AUDIOSTATE	350
113.6. AUDIOSWITCH	350
113.7. BACKPACKLIGHTS.....	351
113.8. BODYMOTION	351
113.9. EVENT.....	352
113.10. FACEANIMATION	352
113.11. HEADANGLE	353
113.12. LIFTHEIGHT	353
113.13. KEYFRAMES.....	354
113.14. PROCEDURALFACE.....	355
113.15. RECORDHEADING.....	356
113.16. SPRITEBOX	357
113.17. ROBOTAUDIO	358
113.18. TURNTORECORDEDHEADING	358
PART VI	359
HIGH LEVEL AI.....	359
CHAPTER 26.....	361
BEHAVIOR	361
114. OVERVIEW	361
115. ACTIONS AND BEHAVIORS.....	361
115.1. ACTIONS AND THE ACTION QUEUES.....	361
115.2. BEHAVIORS	361
115.3. PATH PLANNING AND OTHER SMART THINGS TO SUPPORT US	363
115.4. DECIDING ON THE BEHAVIOR TO USE	363
115.5. INITIATING THE BEHAVIOR.....	363
115.6. MANAGING THE ACTIVE AND PAUSED BEHAVIORS	364
CHAPTER 27.....	365

EMOTION MODEL	365
116. OVERVIEW	365
117. EMOTIONS, AND STIMULATION	365
117.1. STIMULATION	365
117.2. THE EMOTION MODEL	366
117.3. INTERACTION WITH THE BEHAVIOR ENGINE.....	366
117.4. MOOD MANAGER CONFIGURATION	367
117.5. MOOD CONFIGURATION.....	368
118. REFERENCES & RESOURCES	369
CHAPTER 28.....	370
BEHAVIOR TREE	370
119. OVERVIEW	370
120. BEHAVIOR TREE.....	370
120.1. TIMERS.....	371
120.2. CONFIGURATION.....	372
120.3. BEHAVIOR NODE.....	372
120.4. CONDITION NODES	373
121. A LOOK AT SOME INTERESTING BEHAVIORS.....	376
121.1. SHOVING STUFF OFF OF THE TABLE.....	376
121.2. POUNCING	377
121.3. REACTING TO SOUND	377
121.4. DANCING	379
122. USER CONDITIONS	382
123. REFERENCES & RESOURCES	383
PART VII	385
MAINTENANCE	385
CHAPTER 29.....	387
SETTINGS, PREFERENCES, FEATURES, AND STATISTICS	387
124. THE ARCHITECTURE	387
125. WIFI CONFIGURATION	388
126. THE OWNER ACCOUNT INFORMATION	388
127. PREFERENCES & ROBOT SETTINGS.....	389
127.1. ENUMERATIONS	389
127.2. ROBOTSETTINGSCONFIG	391
128. OWNER ENTITLEMENTS	392
129. VESTIGAL COZMO SETTINGS	392
130. FEATURE FLAGS	392
130.1. CONFIGURATION FILE	393
130.2. COMMUNICATION INTERFACE TO THE FEATURES	393
131. ROBOT LIFETIME STATISTICS & EVENTS	393
132. REFERENCES & RESOURCES	393
CHAPTER 30.....	394

THE SOFTWARE UPDATE PROCESS.....	394
133. THE ARCHITECTURE	394
134. THE UPDATE FILE.....	394
134.1. MANIFEST.INI.....	395
134.2. HOW TO DECRYPT THE OTA UPDATE ARCHIVE FILES.....	397
135. THE UPDATE PROCESS	397
135.1. STATUS DIRECTORY.....	397
135.2. PROCESS.....	398
136. RESOURCES & RESOURCES	399
CHAPTER 31.....	400
DIAGNOSTICS	400
137. OVERVIEW	400
137.1. THE SOFTWARE INVOLVED	400
138. SPECIAL SCREENS AND MODES	402
138.1. CUSTOMER CARE INFORMATION SCREEN.....	402
138.2. VECTORS' DEBUG SCREEN (TO GET INFO FOR USE WITH THE SDK).....	402
138.3. DISPLAY FAULT CODES FOR ABNORMAL SYSTEM SERVICE EXIT / HANG.....	402
138.4. RECOVERY MODE	402
138.5. "FACTORY RESET"	403
139. BACKPACK LIGHTS	403
140. DIAGNOSTIC COMMANDS	403
141. LOGS	404
141.1. GATHERING LOGS, ON DEMAND.....	404
142. CONSOLE FILTER	405
143. USAGE STUDIES AND PROFILING DATA	406
143.1. EVENT TRACING	406
143.2. PROFIILING AND LIBOSSTATE.....	407
143.3. EXPERIMENTS	408
144. REFERENCES & RESOURCES	409
REFERENCES & RESOURCES	410
145. CREDITS.....	410
146. REFERENCE DOCUMENTATION AND RESOURCES.....	410
146.1. ANKI.....	410
146.2. OTHER	411
146.3. QUALCOMM	411
APPENDICES	413
APPENDIX A.....	415
ABBREVIATIONS, ACRONYMS, GLOSSARY	415
APPENDIX B.....	421
TOOL CHAIN	421
147. REFERENCES & RESOURCES	423

APPENDIX C	424
ALEXA MODULES	424
APPENDIX D	426
FAULT AND STATUS CODES	426
148. REFERENCES AND RESOURCES	432
APPENDIX E	433
FILE SYSTEM	433
APPENDIX F	438
BLUETOOTH LE SERVICES & CHARACTERISTICS.....	438
149. CUBE SERVICES	438
149.1. CUBE'S ACCELEROMETER SERVICE	439
150. VECTOR SERVICES SERVICE	439
150.1. VECTOR'S SERIAL SERVICE	440
APPENDIX G	441
SERVERS & DATA SCHEMA	441
APPENDIX H	443
FEATURES	443
APPENDIX I	447
PHRASES AND THEIR INTENT	447
APPENDIX J	452
EMOTION EVENTS.....	452
APPENDIX K	454
DAS TRACKED EVENTS AND STATISTICS	454
151. DAS TRACKED EVENTS AND STATISTICS	454
151.1. BASIC INFORMATION	454
151.2. POWER MANAGEMENT EVENTS AND STATISTICS	455
151.3. MOTOR AND IMU STATISTICS AND EVENTS.....	456
151.4. COMMUNICATION RELATED EVENTS POSTED TO DAS	457
151.5. UPDATE-RELATED EVENTS POSTED TO DAS	458
151.6. BEHAVIOUR, FEATURE, MOOD, AND ENGINE RELATED EVENTS POSTED To DAS	459
151.7. ROBOT LIFETIME STATISTICS & EVENTS	459
APPENDIX L	461

PLEO.....	461
151.8. SALES.....	462
151.9. RESOURCES	462

Preface

The Anki Vector is a charming little robot – cute, playful, with a slightly mischievous character. It is everything I ever wanted to create in a bot. Sadly, Anki went defunct shortly after its release. Almost a year later Anki’s software and designs were purchased by Digital Dream Labs, who are presently developing plans for future support and development.

This book is my attempt to understand the Anki Vector and its construction; it is not authoritative and based on speculation. Speculation informed by Anki’s SDKs, blog posts, patents and FCC filings; by articles about Anki, presentations by Anki employees; by PCB photos, and hardware teardowns from others; by a team of people (Project Victor) analyzing the released software; and by experience with the parts, and the functional areas. When you do find errors (and typos), please contact me (my email is on the second page.)

1. ORGANIZATION OF THIS DOCUMENT

- PREFACE. This introduction describes the organization of the chapters and appendices.
- CHAPTER 1: OVERVIEW OF VECTOR’S ARCHITECTURE. Introduces the overall design of the Anki Vector robot.

PART I: ELECTRICAL DESIGN. This part provides an overview of the design of the electronics in Vector and his accessories:

- CHAPTER 2: VECTOR’S ELECTRONICS DESIGN. An overview of the Vector’s electronics design.
- CHAPTER 3: HEAD-BOARD ELECTRONICS DESIGN. A detailed look at the electronics design of Vector’s main processing board.
- CHAPTER 4: BACKPACK & BODY-BOARD ELECTRONICS DESIGN. A detailed look at the electronics design of Vector’s backpack and motor driver boards.
- CHAPTER 5: ACCESSORY ELECTRONICS DESIGN. A look at the electronics design of Vector’s accessories.

PART II: BASIC OPERATION. This part provides an overview of Vector’s software design.

- CHAPTER 6: ARCHITECTURE. A detailed look at Vector’s overall software architecture.
- CHAPTER 7: STARTUP. A detailed look at Vector’s startup, and shutdown processes
- CHAPTER 8: POWER MANAGEMENT. A detailed look at Vector’s architecture for battery monitoring, changing and other power management.
- CHAPTER 9: BUTTON & TOUCH INPUT AND OUTPUT LEDs. A look at the push button, touch sensing, surface proximity sensors, time of flight proximity sensing, and backpack LEDs.
- CHAPTER 10: INERTIAL MOTION SENSING

PART III: COMMUNICATION. This part provides details of Vector’s communication protocols. These chapters describe structure communication, the information that is exchanged, its encoding, and the sequences needed to accomplish tasks. Other chapters will delve into the functional design that the communication provides interface to.

- CHAPTER 11: COMMUNICATION. A look at Vector's communication stack.
- CHAPTER 12: BLUETOOTH LE. The Bluetooth LE protocol that Vector responds to.
- CHAPTER 13: SDK PROTOCOL. The HTTPS protocol that Vector responds to.
- CHAPTER 14: WEB-VISUALIZATION PROTOCOL. The web-sockets protocol(s) that Vector provides for debugging in development builds.
- CHAPTER 15: CLOUD. A look at how Vector syncs with remote services.

PART IV: ADVANCED FUNCTIONS. This part describes items that are Vector's primary function.

- CHAPTER 16: AUDIO INPUT. A look at Vector's ability to hear spoken commands, and ambient sounds.
- CHAPTER 17: IMAGE PROCESSING. Vector vision system is sophisticated, with the ability to recognize marker, faces, and objects; to take photographs, and acts as a key part of the navigation system.
- CHAPTER 18: MAPPING & NAVIGATION. A look at Vector's mapping and navigation systems.
- CHAPTER 19: ACCESSORIES. A look at Vector's home (charging station), companion cube and custom objects.

PART V: ANIMATION. Vector uses animations – “sequence[s] of highly coordinated movements, faces, lights, and sounds” – “to demonstrate an emotion or reaction.” This part describes how the animation system works.

- CHAPTER 20: ANIMATION. An overview how Vector's scripted animations represents the “movements, faces, lights and sounds;” and how they are coordinated.
- CHAPTER 21: LIGHT ANIMATION. An overview of the backpack and cube light animation.
- CHAPTER 22: DISPLAY & PROCEDURAL FACE. Vector displays a face to convey his mood and helps forms an emotional connection with his human.
- CHAPTER 23: AUDIO PRODUCTION. A look at Vector's sound effects and how he speaks
- CHAPTER 24: MOTION CONTROL. At look at how Vector's moves.
- CHAPTER 25: ANIMATION FILE FORMAT. The format of Vector's binary animation file

PART VI: HIGH-LEVEL AI.

- CHAPTER 26: BEHAVIOR. A look at Vectors behaviors, and emotions.
- CHAPTER 27: EMOTION/MOOD MODEL. At Vector's emotions, where they come from and how they impact the sounds and choices he makes.
- CHAPTER 28: BEHAVIOR TREES. A look at how the behaviors are selected and their settings.

PART VII: MAINTENANCE. This part describes items that are not Vector's primary function; they are practical items to support Vector's operation.

- CHAPTER 29: SETTINGS, PREFERENCES, FEATURES AND STATISTICS. A look at how Vector syncs with remote servers
- CHAPTER 30: SOFTWARE UPDATES. How Vector's software updates are applied.
- CHAPTER 31: DIAGNOSTICS. The diagnostic support built into Vector, including logging and usage statistics.

REFERENCES AND RESOURCES. This provides further reading and referenced documents.

APPENDICES: The appendices provide extra material supplemental to the main narrative. These include tables of information, numbers and keys.

- APPENDIX A: ABBREVIATIONS, ACRONYMS, & GLOSSARY. This appendix provides a gloss of terms, abbreviations, and acronyms.
- APPENDIX B: TOOL CHAIN. This appendix lists the tools known or suspected to have been used by Anki to create, and customize the Vector, and for the servers. Tools that can be used to analyze Vector
- APPENDIX C: ALEXA MODULES. This appendix describes the modules used by the Alexa client
- APPENDIX D: FAULT AND STATUS CODES. This appendix provides describes the system fault codes, and update status codes.
- APPENDIX E: FILE SYSTEM. This appendix lists the key files that are baked into the system.
- APPENDIX F: BLUETOOTH LE PROTOCOLS. This appendix provides information on the Bluetooth LE interfaces to the companion Cube, and to Anki Vector.
- APPENDIX G: SERVERS. This appendix provides the servers that the Anki Vector and App contacts.
- APPENDIX H: FEATURES. This appendix enumerates the Vector OS “features” that can be enabled and disabled; and the AI behavior’s called “features.”
- APPENDIX I: PHRASES. This appendix reproduces the phrases that Vector keys off of.
- APPENDIX J: EMOTION EVENTS. This appendix provides a list of the emotion events that Vector internally responds to.
- APPENDIX K: DAS EVENTS. This appendix describes the identified DAS events
- APPENDIX L: PLEO. This appendix gives a brief overview of the Pleo animatronic dinosaur, an antecedent with many similarities.

Note: I use many diagrams from Cozmo literature. They're close enough

1.1. ORDER OF DEVELOPMENT

A word on the order of development; the chapters are grouped in sections of related levels of functionality and (usually) abstraction.

Most chapters will description a vertical slice or stack of the software. The higher levels will discuss features and interactions with other subsystems that have not been discussed in detail yet. For instance, the section on the basic operation of Vectors hardware includes layers that link to the behavior and communication well ahead of those portions. Just assume that you’ll have to flip forward and backward from time to time.

The communication interface is held to its own section with the relevant interactions, commands, structures and so on.

1.2. VERSION(S)

The software analyzed here is mostly version 1.5 and 1.6 of Vector’s production software, as well as some of version 1.7 of the development version. There are incremental differences with each version; I have not always described the pieces that only apply to a specific version. Version 1.6

was the last release to customers as Anki ceased operation. This release includes more software elements that are unused, but are nonetheless telling.

1.3. CUSTOMIZATION AND PATCHING

What can be customized – or patched – in Vector?

- The software in the main processor may be customizable; that will be discussed in many areas of the rest of the document
- The body-board firmware is field updatable, and will take expertise to construct updates.
- The cube firmware can be updated, but that appears to be the hardest to change, and not likely to be useful.

1.4. CODE NAMES OR VECTOR VS VICTOR

Vector's working name during development – aka code name – was Victor. Early products used ad hoc code names. After the development of Cozmo, Anki used NATO phonetic alphabet code words for their products:

Product	Code Word	Description
	Bingo	A larger, two-wheeled self-balancing robot that was more dog-like in inspiration. It would have a larger battery, depth-sensing camera (instead of time of flight sensing), could traverse floors, etc. The software was based on Vector's.
		The large version (called Big Bingo) was requested by the investors for use in security related applications. The smaller, home unit is referred to as Mini Bingo, and initial prototypes were ~15 cm tall.
		<i>Note: Bravo is the correct NATO alphabet codeword, so that rule of thumb isn't 100%</i>
Cozmo	Cozmo	Cozmo a predecessor to Vector. Named after the pet Pomeranian dog (Cosmo) of Patrick DeNeale, an early employee.
Fast and Furious	Foxtrot	Part of the Anki Drive car racing products.
Overdrive	Overdrive	Part of the Anki Drive car racing products.
Drive	Rush Hour	Part of the Anki Drive car racing products.
Vector	Victor	The name Vector was selected both for its similarity to Victor, its uniqueness (e.g., not already trademarked), and working well as a trigger word across many accents and locales.
	Whiskey	This was intended to be a lower cost Cozmo, with less memory, less expensive plastics, only a single cube.

CHAPTER 1

Overview of Vector

Anki Vector is a cute, palm-sized robot; a buddy with a playful, slightly mischievous character. This chapter provides an overview of Vector:

- Overview of Vector and his features
- Privacy and Security
- Ancestry: Cozmo
- Alexa Built-in

2. OVERVIEW

Vector is an emotionally expressive, life-like, animatronic robot pet that people connect with and feel affection for.

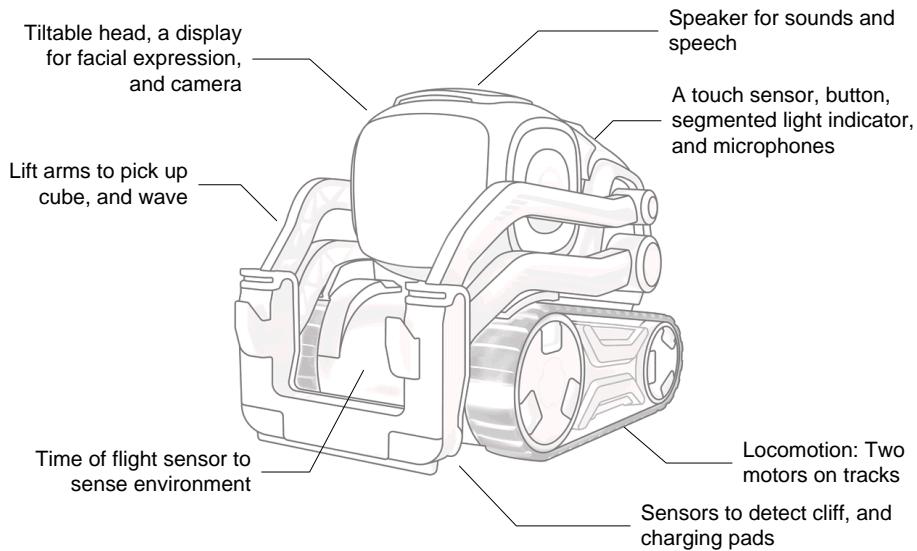


Figure 1: Vector's main features

2.1. COMPELLING CHARACTER

Anki's hallmark is that creating compelling, life-like robot characters, with film-style animations. What does that mean?

- A character has identifiable traits, and moods, something that we can empathize with.
- A compelling character *tries* but doesn't always succeed. As Pixar said, "we admire a character trying more than for their successes"
- He can sense the environment and has some awareness of what they and others are doing...
- He knows that he succeeded – or didn't – and that affects his mood.. So a character has moods, emotions and that affects what it does and how it does it.

- A living thing is never entirely at rest or silent; even when sleeping it moves a little and makes little sounds
- Movements vary and are never quite the same. When they look repetitive, the break the illusion. This is true for choices, reactions and other behaviors too.
- There are little motions, sounds and body's affect that anticipate what a character is thinking and going to do

Vector has a wide variety of behaviors, little motions (animations), and even some emotions that give him a personality. He can express emotions thru expressive eyes (on an LCD display), raising and lower his head, sounds, wiggling his body (by using his treads), or lifting his arms... or shaking them. He can sense surrounding environment, interact and respond to it. He can recognize his name¹, follow the gaze of a person looking at him, and seek petting.²

2.2. FEATURES

Although cute, small, and affordable,³ Vector's design is structured like many other robots.

He has a set of operator inputs:

- A touch sensor is used detect petting
- Internal microphone(s) to listen, hear commands and ambient activity level
- A button that is used to turn Vector on, to cause him to listen – or to be quiet (and not listen), to reset him (wiping out his personality and robot-specific information).
- He can detect his arms being raised and lowered.⁴

He has a set of indicators/annunciators:

- Segmented lights on Vector's backpack are used to indicate when he is on, needs the charger, has heard the wake word, is talking to the Cloud, can't detect WiFi, is booting, is resetting (wiping out his personality and robot-specific information).
- An LCD display, primarily to show eyes of a face. Robot eyes were Anki's strongest piece of imagery. Vector smiles and shows a range of expressions with his eyes.
- Speaker for cute sounds and speech synthesis

He has other means to express affect as well:

- His head can be tilted up and down to represent sadness, happiness, etc.
- His arms flail to represent frustration
- He can use his treads to shake or wiggle, usually to express happiness or embarrassment

He has environmental sensors:

- A camera is used to map the area, detect and identify objects and faces.
- Fist-bump and being lifted can be detected using an internal *inertial measurement unit* (IMU)

¹ Vector can't be individually named.

² Admittedly this is a bit hit and miss.

³ Although priced as an expensive toy, this feature set in a robot is usually an order of magnitude more expensive, with less quality.

⁴ and possibly a pat on his head?

- A forward facing “time of flight” proximity sensor aids in mapping and object avoidance
- Ground sensing proximity sensors that are used to detect cliffs at the edge of his area and to following lines when he is reversing onto his charger.

His internal sensing includes:

- Battery voltage, charging; charging temperature
- IMU for orientation position (6-axis)
- Encoders provide feedback on motor rotation

His other articulation & actuators are:

- Vector drives using two independent treads to do skid-steering
- Using his arms Vector can lift or flip a cube; he can pop a wheelie, or lift himself over a small obstacle.
- Vector can raise and lower his head

Communication (other than user facing):

- Communication with the external world is thru WiFi and Bluetooth LE.
- Internally RS-232 (CMOS levels) and USB

Motion control

- At the lowest level can control each of the motors speed, degree of rotation, etc. This allows Vector to make quick actions.
- Combined with the internal sensing, he can drive in a straight line and turn very tightly.
- Driving is done using a skid-steering, kinematic model
- To do all this, the motion control takes in feedback from the motor encoder, IMU-gyroscope. May also use the image processing for SLAM-based orientation and movement.

Guidance, path planning

- Vector plans a route to his goals – if he knows where his goal is – along a path free of obstacles; he adapts, moving around in changing conditions.
- A*, Rapidly-Expanding Random Tree (RRT), D*-lite
- Paths are represented as arcs, line segments, and turn points

Mapping and Navigation:

- Maps are built using *simultaneous location and mapping* (SLAM) algorithms, using the camera and IMU gyroscope movement tracking, time of flight sensor to measure distances, and particle system algorithms to fill in the gaps.
- The maps are represented uses quad-tree (position, pose)

Behaviour system:

- Variety of behaviors animations
- Goals, linking up to the guidance system to accomplish them
- A simple emotion model to drive selection of behaviours

Emotion model. Dimensions to emotional state

- Happy (also referred to as his default state)
- Confident
- Social
- Stimulated

Vision. This is one of Anki's hallmarks: they used vision where others used beacons. For instance, iRobot has a set of IR beacons to keep the robots out of areas, and to guide it to the dock. Mint has an IR beacon that the mint robots use to navigate and drive in straight lines. Although Vector's companion cube is powered, this is not used for localization. It has markers that are visually recognized by Vector.

- Illumination sensing
- Motion sensing
- Links to Navigation system for mapping, (SLAM etc)
- Recognizing marker symbols in his environment
- Detecting faces and gaze detection allows him to maintain eye contact

3. PRIVACY AND SECURITY

Vector's design includes a well thought out system to protect privacy. This approach protects the following from strangers gaining access to:

- Photos taken by Vector
- The image stream from the camera
- The audio stream from the microphone — if it had been finished being implemented
- Information about the owner
- Control of the robot's movement, speech & sound, display, etc.

Vector's software is protected from being altered in a way that would impair its ability to secure the above.

4. COZMO

We shouldn't discuss Vector without mentioning the prior generation. Vector's body is based heavily on Cozmo; the mechanical refinements and differences are relatively small. Vector's software architecture also borrows from Cozmo and extends it greatly. Many of Vector's behaviours, senses, and functions were first implemented in Cozmo (and/or in the smartphone application). One notable difference is that Cozmo did not include a microphone.

Cozmo includes a wide variety of games, behaviours, and ~940 animation scripts. Cozmo's engine is reported to be "about 1.8 million lines of code, the AI, computer vision, path planning, everything."⁵ This number should be discounted somewhat, as it likely includes many large 3rd party modules... Nonetheless, it represents the scale of work to migrate Cozmo's code base for reuse in Vector.

⁵ https://www.reddit.com/r/IAmA/comments/7c2b5k/were_the_founders_of_anki_a_robotics_and_ai/

Not all of Cozmo's functionality was ported to Vector at one time. Instead, key features and behaviours were incrementally brought to Vector in its regular software updates. It is likely the intent was to follow-up with much more in future updates (and perhaps on a faster schedule than they were able to deliver).

5. ALEXA INTEGRATION

Vector includes Amazon Alexa functionality, but it is not intimately integrated. Vector only acts like an Echo Dot. By using the key word "Alexa," Vector will suppress his activity, face and speaking, and the Alexa "echo dot" functionality takes over. Vector has no awareness of Alexa's to-do list, reminders, messages, alarms, notifications, question-and-answers, and vice-versa.

The most likely reason for including Alexa is the times: everything had to include Alexa to be hip, or there would be great outcry. Including Alexa may have also been intended to provide functionality and features that Anki couldn't, to gain experience with the features that Amazon provides, and (possibly) with the intent to more tightly integrate those features into Anki products while differentiating themselves in other areas.

Alexa clearly took a lot of effort to integrate, and a lot of resources:

"[Alexa Voice Service] solutions for Alexa Built-in products required expensive application processor-based devices with >50MB memory running on Linux or Android"⁶

Alexa's software resources consume as much space Vector's main software. And the software is not power efficient. Even casual use of Alexa noticeably reduces battery life, and (anecdotally) increases the processor temperature.

See Appendix C for a list for a list of the Alexa modules.

⁶ <https://aws.amazon.com/blogs/iot/introducing-alexa-voice-service-integration-for-aws-iot-core/>

Alexa's SDK and services have continued to evolve. New Alexa SDKs allow simpler processors and smaller code by acting as little more than a remote microphone.

[This page is intentionally left blank for purposes of double-sided printing]

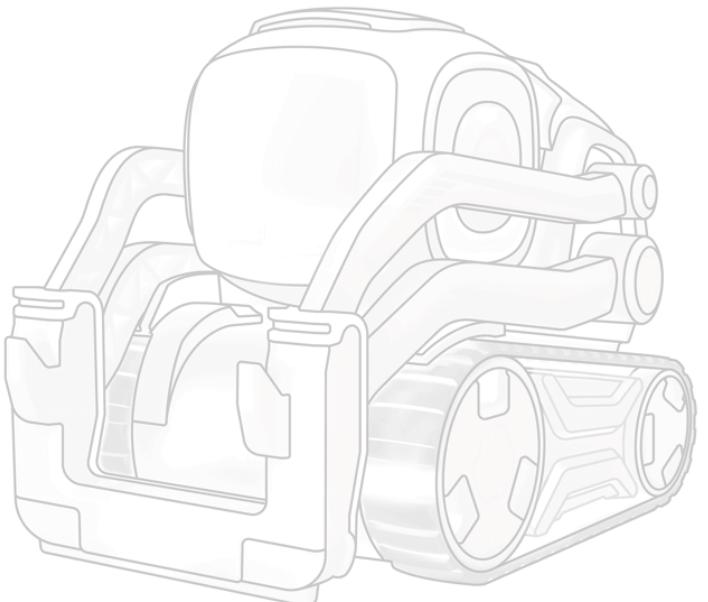
PART I

Electronics Design

This part provides an overview of the design of the electronics in Vector and his accessories

- VECTOR'S ELECTRONICS DESIGN. An overview of the Vector's electronics design.
- HEAD-BOARD ELECTRONICS DESIGN. A detailed look at the electronics design of Vector's main processing board.
- BACKPACK & BODY-BOARD ELECTRONICS DESIGN. A detailed look at the electronics design of Vector's backpack and motor driver boards.
- ACCESSORY ELECTRICAL DESIGN. A look at the electrical design of Vectors accessories.

Note: In previous versions called the circuit board in the bottom half the “*base-board*”. It is now referred to as “body-board” to match Anki’s naming



[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 2

Electronics Design Description

This chapter describes the design of Vector's electronics:

- Design Overview, outlining the main subsystems
- Power distribution

Subsequent chapters will examine in detail the design of the subsystems

6. DESIGN OVERVIEW

Vector's design includes numerous sensors to sense and interact with his environment, other to interact with people and express emotion and behaviour.

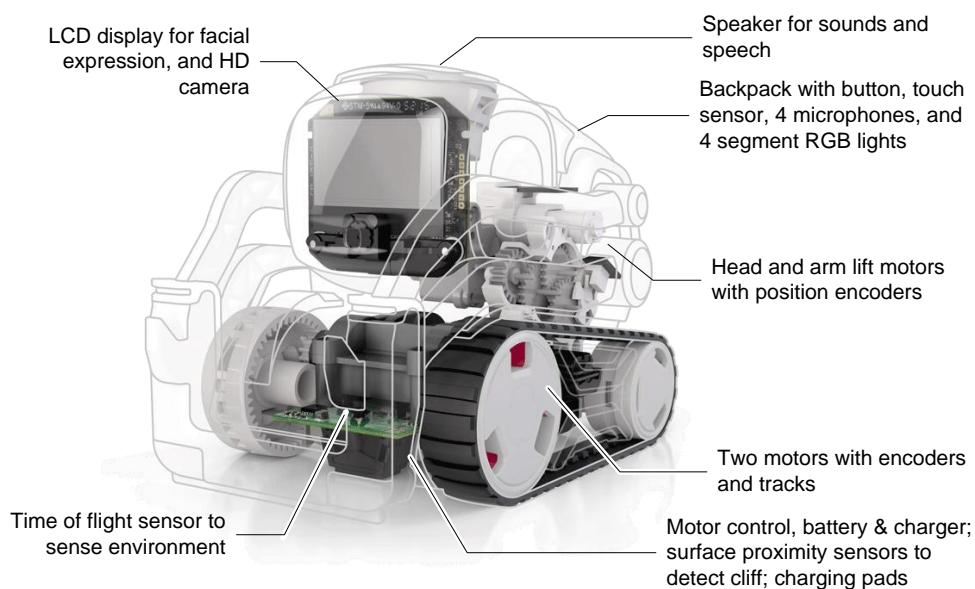


Figure 2: Vector's main elements

Vector's functional elements are:

Element	Description
backpack	The top of Vector, where he has a button, segmented lights, and a touch sensor.
battery	An internal battery pack is used as Vector's source of energy.
button	A momentary push button is used to turn Vector on, to cause him to listen – or to be quiet (and not listen) – to reset him (wiping out his personality and robot-specific information).
camera	Vector uses an HD camera to visualize his environment, and recognize his human companions.
charging pad	Two pads on the bottom are used to replenish the energy in the battery pack from the dock.
LCD display	An IPS LCD, with an active area is 23.2mm x 12.1mm. It has a resolution of 184 x 96 pixels, with RGB565 color.
microphones	There are 4 internal far-field microphone(s) to listen to commands and ambient activity level. Employs beam forming to localize sounds.
motors & encoders	There are four motors each with single-step optical encoders to measure their position and approximate speed. One motor controls the tilt of the head assembly. Another controls the lift of his arms. Two are used to drive him in a skid-steering fashion.
segmented RGB lights	There are 4 LEDs used to indicate when he is on, needs the charger, has heard the wake word, is talking to the Cloud, can't detect WiFi, is booting, is resetting (wiping out his personality and robot-specific information).
speaker	A speaker is used to play sounds, and for speech synthesis
surface proximity sensors	4 infrared proximity sensors are used to detect the surface beneath Vector – and to detect drop offs ("cliffs") at the edge of his driving area, and to follow lines.
time of flight sensor	A time of flight sensor is used to aid in mapping (by measuring distances) and object avoidance.
touch sensor	A touch allows Vector to detect petting and other attention.

Table 2: Vector's main elements

Vector has 6 circuit boards

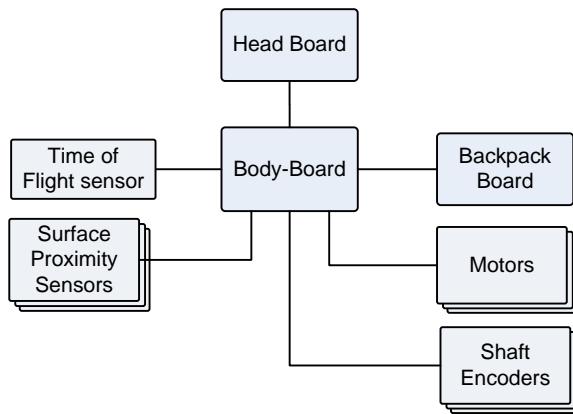


Figure 3: Circuit board topology

The main two boards are the head-board where the major of Vector's processing occurs, and the body-board, which drives the motors and connects to the other boards.

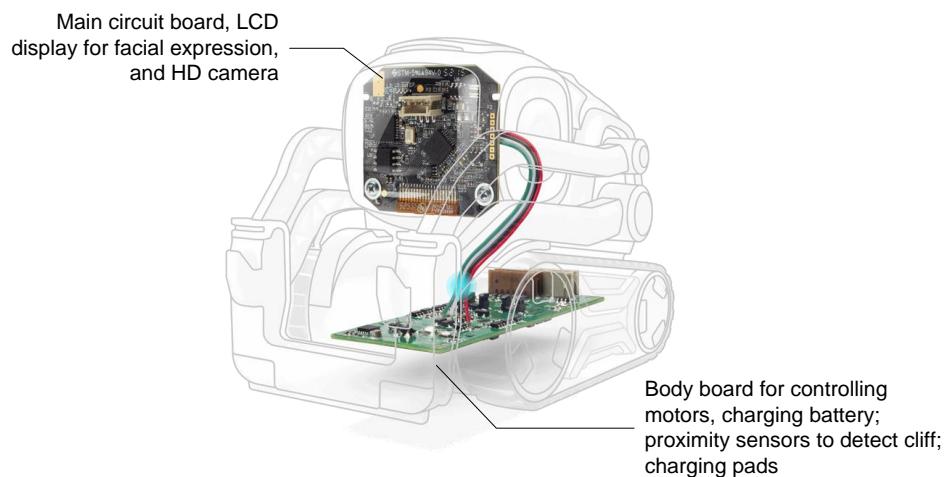


Figure 4: Vector's main microcontroller circuit boards

The table below summarizes the boards:

Circuit Board	Description
backpack board	The backpack board has 4 RGB LEDs, 4 MEMS microphones, a touch wire, and a button. This board connects to the body-board.
body-board	The body board drives the motors, provides power management, and the battery charger.
encoder-boards	The two encoder boards have single opto-coupler encoders each. The encoder is used to monitor the position of the arms and head, either as driven by the motor, or by a person manipulating them.
head-board	The head board includes the main processor, flash & RAM memory storage, an IMU, and a PMIC. The WiFi and Bluetooth LE are built into the processor. The camera and LCD are attached to the board, thru a flex tape. The speaker is also attached to this board.
time of flight sensor board	The time of flight sensor is on a separate board, allowing it to be mounted in Vector's front.

Table 3: Vector's circuit boards

6.1. POWER SOURCE AND DISTRIBUTION TREE

Vector is powered by a rechargeable battery pack, and the energy is distributed by the body-board:

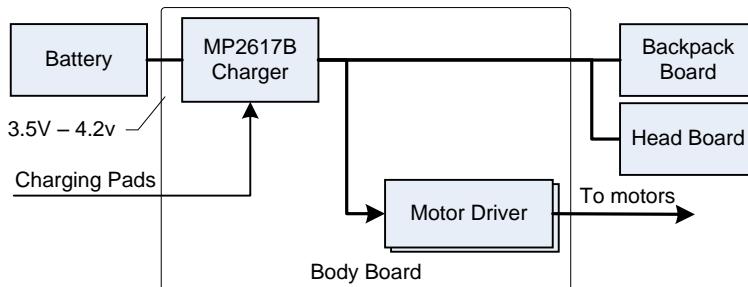


Figure 5: Power distribution

When the charging pads are energized – when Vector is in the charging dock – the system is powered by the external power source.

Excessive current demand – such as from a stalled motor – can trigger a system brown-out and shutdown.

6.1.1 Battery

Vector battery is a single-cell 3.7v 320mAh “toy safe” lithium-ion polymer battery. The battery is connected to the body-board. The pack is not a “smart” battery – it only has positive and negative leads, lacking an onboard temperature sensor or battery management system (BMS).

Battery heat is the most significant source of battery “aging” – its effective service life. High recharge rates internally heat the cells, causing them to deteriorate. Vector’s battery thinness gives it a high surface area to volume ratio allowing it shed heat much faster, greatly reducing the internal heating from charging and heavy loads. The battery is physically separated from the body-board, isolating it from the heat generated in the charging, power distribution and motor driver circuits. This increases the battery service life.

Vector takes care to thermally manage the battery, to promote a longer life. The software monitors the body board temperature (as a proxy of the battery temperature). When the temperature gets above one or more thresholds (e.g. 50C), Vector can slow down or stops his activities and charging to allow the battery cool.

The battery has a low internal resistance, reducing the internal heating and allowing it to usefully deliver higher currents without resulting in a brown-out. “Vector has brief but high (2A) peak currents when doing certain computations or flipping himself with his lift.”

Anki engineers certainly desired easy-to-replace batteries, and larger batteries. But there were challenges. Battery replacement requires more parts and design features. A larger battery would allow longer play time between charges, but they often have higher internal resistance (thus more prone to brown out). So it would have taken finding one with good thermal characteristics (i.e. didn’t get too hot), was toy safe despite holding more charge and chemicals, and so on. Ultimately schedule prevented finding a suitable larger battery.

6.1.2

Battery management

The MP2617B is a central element to managing the battery. It acts as a battery charger, a power switch and power converter for the whole system.

- When Vector is going into an *off* state – such as running too low on power, going into a ship state before first use, or has been turned off by a human companion – the MP2617B charger and power converted can be signaled to turn off
- When Vector is turned off the boards are not energized. The exception is that the high side of the push button is connected to the battery. When closed, the signals the MP2617B to connect the battery to the rest of the system, powering it up.
- The MP2617B is also responsible for charging the battery. There are two pads that mate the dock to supply energy to charge the battery.

In many rechargeable lithium ion battery systems there is a coulomb counter to track the state of charge. Vector does not have one. The need for recharge is triggered solely on the battery voltage.

7.

REFERENCES & RESOURCES

Anki, *Lithium single-cell battery data sheet*

https://support.anki.com/hc/article_attachments/360018003653/Material%20Safety%20Data%20Sheet_April%202018.pdf

CHAPTER 3

Head-board Electronics Design Description

This chapter describes the electronic design of Vector's head-board:

- Detailed design of the head-board

8. THE HEAD-BOARD (THE MAIN PROCESSOR BOARD)

The head-board handles the display, playing sounds, communication, and all of Vector's real processing. It is powered by a quad-core Arm-A7 Qualcomm APQ8009 microprocessor. The processor also connects to Bluetooth LE and WiFi transceivers, an HD camera, LCD display, speakers and an IMU.

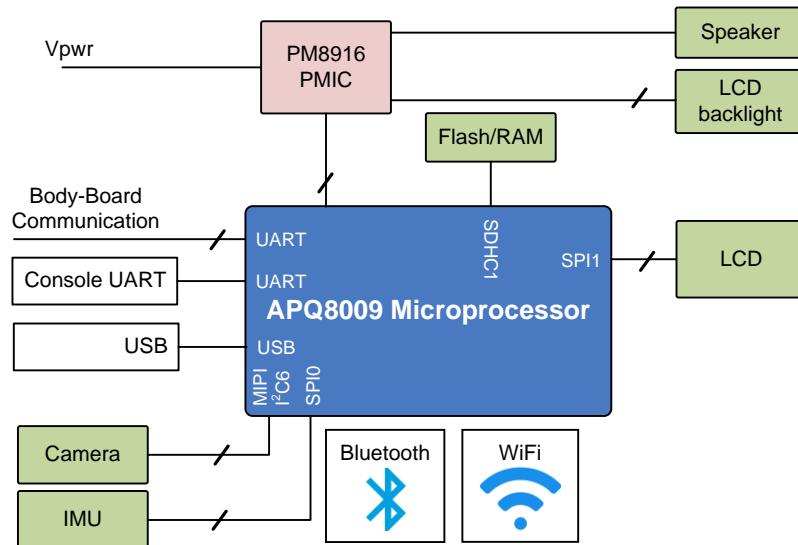


Figure 6: Head-board
block diagram

The head-board's functional elements are:

Element	Description
Bluetooth LE transceiver	A Bluetooth LE transceiver is built into the package
camera	Vector uses a 720P camera to visualize his environment and recognize his human companions.
flash/RAM (eMMC)	Flash and RAM are provided by single external package, a Kingston 04EMCP04-NL3DM627 mixed memory chip with 4 GB flash and 512MB RAM.
inertial measurement unit (IMU)	The headboard includes a 6-axis IMU – gyroscope and accelerometer – used for navigation and motion control.
LCD backlight	There are two LEDs used to illuminate the LCD display.
LCD display	An IPS LCD, with an active area is 23.2mm x 12.1mm. It has a resolution of 184 x 96 pixels, with RGB565 color.
microprocessor	The head-board is based on a Qualcomm APQ8009 (Snapdragon 212). The processor is a quad-core Arm A7 (32-bit) CPU.
power management IC (PMIC)	The PM8916 power management IC provides voltage regulation for the processor, flash/RAM and other parts; it also provides audio out to the speaker and controls the LCD backlight.
speaker	A speaker is used to play sounds, and for speech synthesis
WiFi transceiver	An 802.11AC WiFi transceiver is built into the processor package

Table 4: The head-boards functional elements

8.1. THE APQ8009 PROCESSOR

The head-board is based on the Qualcomm “Snapdragon 212” APQ8009 SOC. It is a quad-core processor; each core is a 32-bit ARM Cortex A7. It also includes a DSP (“Hexagon 536”), and GPU (Adreno 304). It also includes WiFi and Bluetooth LE transceivers. The processor has interfaces to external memory, for the camera (using MIPI), the display, and the audio playback.

The APQ8009 processor is a sibling to the MSM8909 processor employed in cell phones, where APQ is short for “application processor Qualcomm” and MSM is short for “mobile station modem.” The difference is that the later includes some form of modem, such as HPSA, CDMA, or LTE. Both designators are used in software code-bases employed with Vector. The most likely reason is the naming of registers, drivers, and other useful software didn't carefully limit the use of MSMxxxx references to just the processors with modems.

The flash & RAM are connected to the processor on SDHC1. The device tree file shows that during development Vector's also supported an SD card slot on SDHC2.

The processor dynamically adjusts its clock frequency, within an allowed region. The processor can be configured to limit its speed.

8.2. SPEAKERS

The speaker is driven at 16bits, single channel, with a sample rate of 8000-16025 samples/sec.

8.3. CAMERA

Vector has a 720p camera with a 120° field of view. The camera is calibrated at manufacturing time. The camera vertical sync (frame sync) is connected to the interrupt input on the IMU to synchronize the samples.

GPIO	Description
26	Camera interface clock
48	Camera reset
83	Camera power enable (from PM8916 PMIC)
94	Camera standby

Table 5: The camera controls

8.4. THE LCD

Vector's LCD is a backlit IPS display assembly made by Truly. The processor is connected to the LCD via SPI. Two LEDs are used to illuminate the LCD. The backlight is PWM controlled by the PM8916 PMIC.

LCD display

The prior generation, Cozmo, used an OLED display for his face and eyes. This display had the strengths of high contrast and self-illumination. However, OLEDs are susceptible to burn-in and uneven dimming or discoloration of overused pixels. Anki addressed this with two accommodations. First it gave the eyes regular motion, looking around and blinking. Second, the LCD's illuminated rows were regularly alternated to give a retro-technology interlaced row effect, like old CRTs.

US Patent 20372659

Vector's IPS display gives a smoother imagery – Cozmo's OLED was simply black and white. The LCD is also much less susceptible to burn-in, at the expense of higher power. Vector's LCD can also develop dead lines (or pixels) that grow in number until the display is non-functional. Some units have a defective LCD, where the glass is not properly sealed. This allows moisture in, causing progressive damage to the LCD. It is also speculated that these lines come from shocks to the head, causing breaks in the LCD connections.

8.5. POWER MANAGEMENT

The PM8916 PMIC is responsible for providing power and managing most of the power. The headboard is capable of being the highest power consumer in Vector. By limiting the clock rate of the processor, the power use can be capped.

The headboard can be put into a lower power state by reducing the clock rate of processor and using its sleep features; dimming or turning off the LCD, and reducing the camera frame rate (or turning it off). The APQ8009 processor has many sophisticated power controls, but these were not fully realized in Vector's software.

8.6. TRIM, CALIBRATION SERIAL NUMBERS AND KEYS

Each Vector has a set of per unit calibrations:

- The camera is calibrated
- The IMU is calibrated
- The motor position is calibrated, this is performed with each startup

There are per unit keys, MAC addresses and serial numbers

- Each processor has its own unique key, used to with the Trust Zone
- The WiFi and Bluetooth have assigned, unique MAC addresses.

- Each Vector has an assigned serial number

8.7. MANUFACTURING TEST CONNECTOR/INTERFACE

It is a common practice to include at least one interface on a product for use during manufacture. This is used to load software and firmware, unique ids – WiFi MACs, serial number – to perform any calibration steps and to perform run-up checks that the device functions / is assembled correctly. It is intended to be a fast interface that doesn't cause yield fallout. Typically (but there are exceptions) this is not radio based, as they can interfere or have fiddly issues.

The USB interface is used to load firmware. The microprocessors include a built-in boot-loader (ABOOT), which includes support for loading firmware into the device's flash.

For the other functions, there are three possibilities

- There is a UART, that provides a boot console, but does not accept input
- There is a USB connector that probably is used to load firmware.
- The WiFi, once MAC addresses have been loaded into the unit

9. REFERENCES & RESOURCES

Kingston Technology, *Embedded Multi-Chip Package 04EMCP04-NL3DM627-Z02U*, rev 1.2, 2016
<https://cdn.discordapp.com/attachments/573889163070537750/595223765206433792/04EMCP04-NL3DM627-Z02U - v1.2.pdf>

Qualcomm, APQ8009 Processor
<https://www.qualcomm.com/products/apq8009>

Qualcomm, *PM8916/PM8916-2 Power Management ICs Device Specification*, Rev C, 2018 Mar 13
https://developer.qualcomm.com/qfile/29367/lm80-p0436-35_c_pm8916pm8916_power_management_ics.pdf

CHAPTER 4

Backpack & Body-board Electronics

Design Description

This chapter describes the electronic design of the Anki Vector's supplemental boards:

- Detailed design of the backpack-board
- Detailed design of the body-board
- Power characteristics

10. THE BACKPACK BOARD

The backpack board is effectively daughter board to the body-board. It provides extra IO and a couple of smart peripherals:

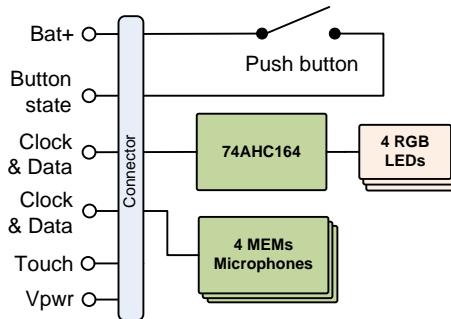


Figure 7: Backpack board block diagram

The table below summarizes the functional elements of the backpack board:

Elements	Description	Table 6: Backpack board functional elements
74AHC164	A SPI-like GPIO expander. This is used to drive the RGB LEDs.	
microphones	There are 4 far-field MEMS PDM microphones. The microphones are accessed via SPI, in an output only mode. These are designated MK1, MK2, MK3, MK4	
push button	A momentary push button is connected to the battery terminal, allowing a press to wake Vector, as well as signal the processor(s).	
RGB LEDs	There are 4 RGB LEDs to make up a segmented display. Each segment can be illuminated individually (in a time multiplexed manner) or may share a colour configuration with its counterparts.	
touch sensor	A touch-sensing wire (and passive components)	

10.1. BACKPACK CONNECTION

The backpack connection includes:

- Power and ground connections. This includes connection to the battery rail.
- The touch wire as an analog signal to the body-board
- A quasi digital signal out from the momentary push button
- A SPI-like clock, *two* master-in-slave-out (MISO) signals for the microphones
- A SPI-like clock and master-out-slave-in (MOSI) for the 74AHC164 LED controller

10.2. ELECTRO-STATIC DISCHARGE (ESD) PROTECTION

The touch sensing uses an insulated wire separated from the external touch plate. There is no direct path of conduction from the external plate for ESD. The separation reduces transient voltage to levels that the electronics can suppress.

10.3. OPERATION

The touch sensor conditioning and sensing is handled by the body-board. The touch sense wire is merely an extension from the body-board through the backpack board.

The push-button is wired to the battery. When pressed, the other side of the push button signals both body-board microcontroller, and (if Vector is off) the charger chip to connect power. The theory of operation will be discussed further in the body-board section below.

The 74AHC164 serial-shift-register is used as a GPIO expander. It takes a clock signal and serial digital input, which are used to control up to 8 outputs. The inputs determine the state of 8 digital outputs used to control the RGB LEDs.

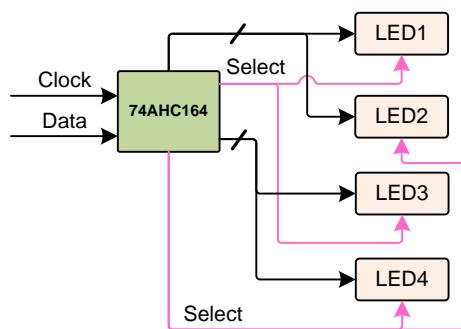
Each of the 4 MEMS microphones take a clock signal, and provide a serial digital output. The body-board reads all four microphones by simultaneously. (This will be discussed in the body-board section).

10.3.1

The LED controls

8 outputs are not enough to drive 4 RGB LEDs (each with 3 inputs) simultaneously with independent colors. While 3 of the LEDs often the same colour , they can have independent colors.

There are two possible topologies that can multiple the RGB signals on the 74AHC164 directing different RGB configurations to each light.⁷ The first possibility is two lights are driven at a time. LEDs 1 & 2 share the same red, green, and blue signals, but their low side is connected to separate GPIO lines, acting as a LED select. LEDs 3 & 4 are the same – sharing red, green, and blue. The even LEDs would share the same select line, and the odd LEDs would share the same. This is the simplest.



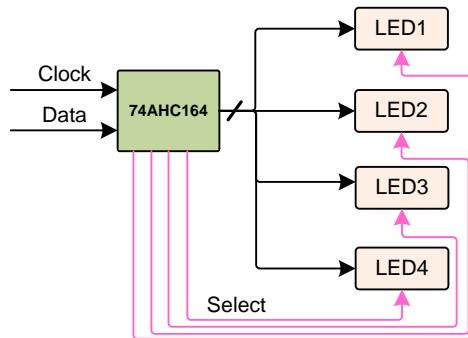
Backpack LED control
scheme
corrections by Melanie
T

Figure 8: Possible light topology on backpack board

The process of illuminating the lights would be:

1. The firmware would send the RGB signals for LEDs 1 and 3, enabling them and disabling LEDs 2 and 4.
2. Delay
3. Repeat for LEDs 2 and 4

The second possibility is that each LED's red signal goes to the same signal on 74AHC164; similar for green and blue. However each LED's low side is connected to separate signals on 74AHC164.



Another possible light topology on backpack board

This approach takes more work. The process of illuminating the lights in this configuration would be:

1. The RGB color and light 1 signal enables are sent, illuminating the first light
2. Then the RGB color and light 2 signal enables are sent – but the first light signal is disabled – illuminating the second light

⁷ I'd need to physically examine a backpack board. This is the limit of examining the available photos

- This is repeated for each of the other lights.

With either approach, if the switching between the LED's is done quickly enough – in a short time interval – the off period isn't visible. LED's don't immediately turn off, rather their brightness decays over a short period. And the human eye doesn't perceive short flickers. Although the lights are “pulse width modulated” – they are turned off a portion of the time, dimming them – current limiting resistors may have been set to achieve the desired maximum brightness for the fastest multiplexing time.

The body-board controller can dim the brightness of the LEDs further by choosing larger numbers of time slots to not illuminate a light.

11. THE BODY-BOARD

The body board is a battery charger, smart IO expander, and motor controller. It connects the battery to the rest of the system and is responsible for charging it. It is based on an STM32F030 which acts as second processor in the system.

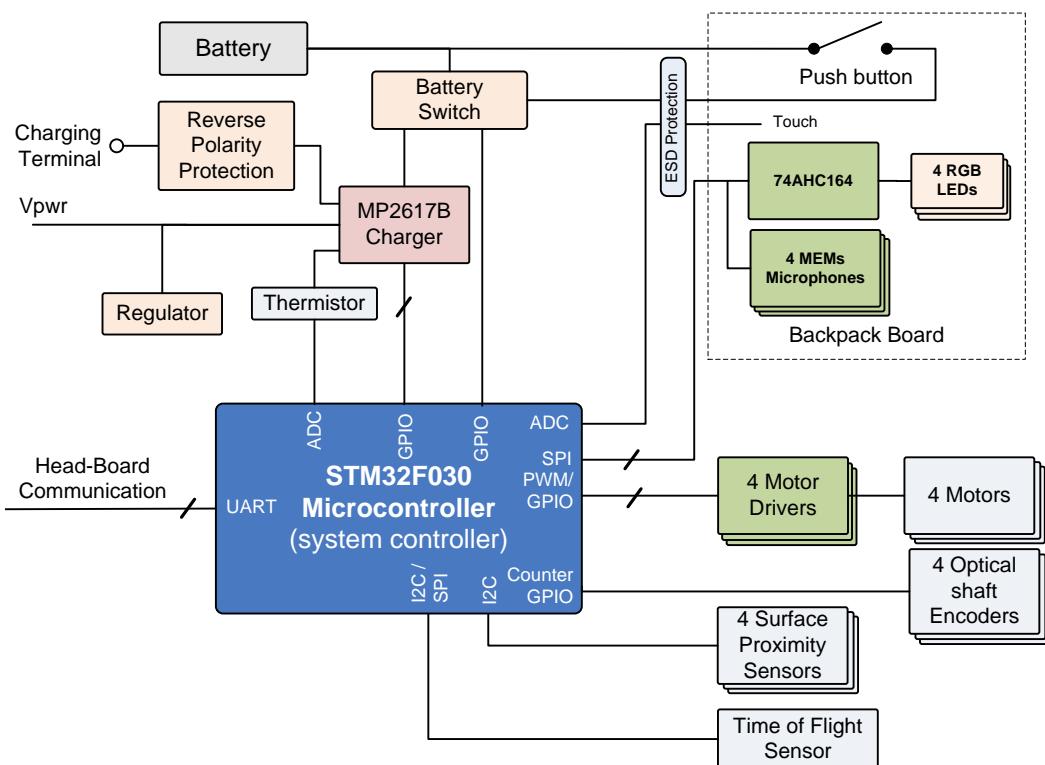


Figure 10: Body-board block diagram

The functional elements of the body-board are:

Table 7: The body-board functional elements

Element	Description
battery	An internal, rechargeable battery pack (3.7v 320 mAh)
battery switch	Used to disconnect the battery to support off-mode (such as when stored) and to reconnect the battery with a button press.
charging pad	Two pads on the bottom are used to replenish the energy in the battery pack from the dock.
motor driver	There are four motor drivers, based on an H-bridge design. This allows a motor to be driven forward and backward.
motors	There are four motors with to measure their position and approximate speed. One motor controls the tilt of the head assembly. Another controls the lift of his arms. Two are used to drive him in a skid-steering fashion.
MP2617B charger	The Monolithic Power Systems MP2617B serves as the battery charger. It provides a state of charge to the microcontroller. It also directs power from the charging pads to the rest of the system while the robot is on the charging dock.
optical shaft encoder	A Sharp GP1S092HCPIF opto-coupler, in conjunction with a slotted disc on a motor's shaft, is used to measure the amount a shaft has turned, and its speed.
regulator	A 3.3v used to supply power to the microcontroller and logical components.
reverse polarity protection	Protects the circuitry from energy being applied to the charging pads in reverse polarity, such as putting Vector onto the charging pads in reverse.
STM32F030 microcontroller	The “brains” of the body-board, used to drive the motors, and RGB LEDs; to sample the microphones, time of flight sensor, proximity sensor, temperature, and the touch sense;, and monitoring the battery charge state. It communicates with the head-board.
surface proximity sensors	4 infrared proximity sensors are used to detect the surface beneath Vector – and to detect drop offs (“cliffs”) at the edge of his driving area and to follow lines.
theristor	A temperature sense resistor used measure the battery pack temperature; it is used to prevent overheating during recharge.
VL53L0x time of flight sensor	A ST Microelectronics VL53L0x time of flight sensor is used to measure distance to objects in front of Vector. This sensor is connected by I ² C.

11.1. POWER MANAGEMENT

The battery charging is based on a MP2617B IC, which also provides some protection functions. There is no Coulomb counter; the state of charge is based solely on the battery voltage.

11.1.1 Protections

The charging pads have reverse polarity protection.

The MP2617B has an over-current cut off. If the current exceeds ~5A (4-6A), the battery will be disconnected from the system bus. Such a high-current indicates a short. There is no fuse.

The MP2617B has a low voltage cut off. If the battery voltage drops below ~2.4 (2.2-2.7V) the battery will be disconnected from the system bus (TBD) until the battery voltage rises above ~2.6V (2.4-2.8V).

The MP2617B has a temperature sense. If the temperature exceeds a threshold, charging is paused until the battery cools. The temperature sense is not on the battery. It is likely on the circuit board, or possibly top of the battery retention.

11.1.2

Battery connect/disconnect

To preserve the battery there is a need to isolate the battery from the rest of the system when in an off state. If there is minute current draw, the battery will irreversibly deplete while in storage even before the first sale. This constraint shapes the battery disconnect-reconnect logic. The schematic below shows one way to do this:

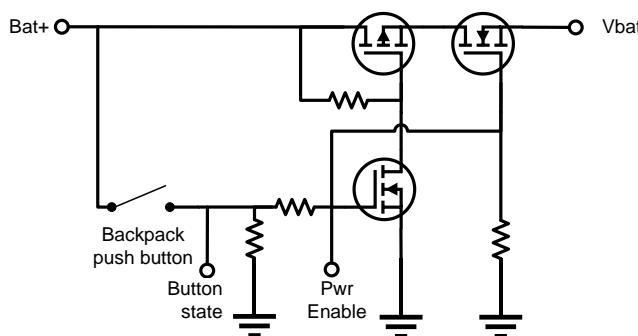


Figure 11: A representative battery connect switch

Two MOSFETS (a PFET and NFET)⁸ act as a switch. These are in a single package, the DMC2038LVT. (This part is also used in the motor drivers.)

- When the system is in an off state, the MOSFETS are kept in an off state with biasing resistors. The PFET's gate is biased high with a resistor. The NFET gate is biased low, to ground. There is no current flow. Two MOSFETS are needed due to internal body diodes. The PFET body diode would allow current to flow from the battery (from the source to the drain). However, this current is blocked by the NFET body diode, which has a different polarity
- The push button can wake the system. When the button is closed, the battery terminal (Bat+) is connected to the gate of the NFET, turning it on. A second NFET is also energized, pulling the PFET gate to ground, turning it on as well. When the button is open, Bat+ is not connected to anything, so there is no leakage path draining the battery.
- To keep the system energized when the button is open, the STM32F030 MCU must drive the Pwr Enable line high, which has the same effect as the button closed. The gate threshold voltage is 1V, well within the GPIO range of the MCU.
- The MCU can de-energize the system by pulling Pwr Enable line low. The switches will open, disconnect the battery.
- The MCU needs to be able to sense the state of the button while Pwr Enable is pulled high. The MCU can do this by sampling the Button State signal. This signal is isolated from Pwr Enable by a large resistor and pulled to ground by smaller resistor. This biases the signal to ground while the button is open.

This circuit also provides reverse polarity protection. It will not close the switch if the battery is connected backwards.

11.1.3

Charging

The charging station pads are connected to a MP2617B charger IC thru a reverse polarity protection circuit. The reverse polarity protection⁹ is a DMG2305UX PFET in a diode

charging station pads

⁸ Q11 and/or Q12

⁹ Q14

configuration. This approach has much lower losses than using an equivalent diode.

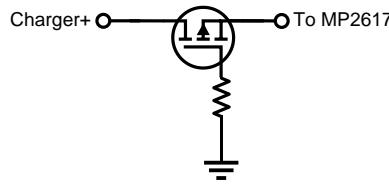


Figure 12: A representative PFET based reversed polarity protection

The MP2617B internally switches the charger input voltage to supply the system with power, and to begin charging the battery. This allows the charger to power the system whenever the robot is in the charging station, even when the battery is depleted, or disconnected.

The presence of the dock power, and the state of MP2617B (charging or not) are signaled to the microcontroller.

The charger goes through different states as it charges the battery. Each state pulls a different amount of current from the charging pads and treats the battery differently.

supplying power from the charging station

charging states

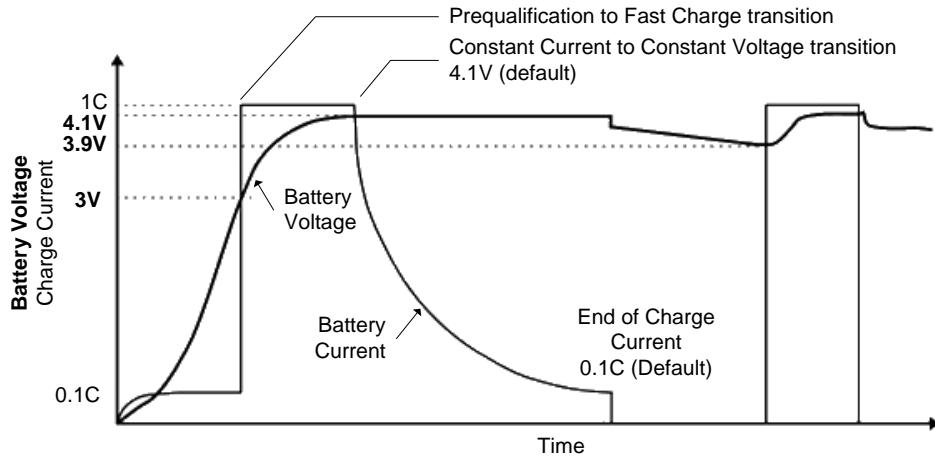


Figure 13: Charging profile (adapted from Texas Instruments)

The basic idea is that the charger first applies a low current to the battery to bring it up to a threshold; this is called *prequalification* in the diagram. Then it applies a high current, call *constant current*. Once the battery voltage has risen to a threshold, the charger switches to *constant voltage*, and the current into the battery tapers off. I refer to the data sheet for more detail.

constant current
constant voltage

The MP2617B measures the battery temperature using a thermistor. If the temperature exceeds a threshold, charging is paused until the battery cools. The microcontroller also samples this temperature.

The MP2617B supports limiting the input current, to accommodate the capabilities of external USB power converters. There are four different possible levels that the IC may be configured for: 2A is the default limit, 450mA to support USB2.0 limits, 825mA to support USB3.0 limits, and a custom limit that can be set by resistors. The input limit appears to be set for either default (up to ~2A input), or a programmable input.

input current limits

Commentary. In my testing, using a USB battery pack charging pulls up to 1A during the constant current, then falls off to 100mA-200mA during constant voltage, depending on the

Higher charge rates
are acceptable

head-board's processing load. Stepped down to the ~4V battery the applied current at peak is approximately 1A.¹⁰

With larger batteries this would be too high. Battery cells are normally charged at no more than a "1C" rate – e.g. the battery maximum charge rate "should" be 320mA at max. Vector's battery can be charged at a rate higher than 1C. Heat is what damages batteries. This battery's low internal resistance doesn't produce as much heat; and its large surface to volume ratio lets it shed heat.

11.1.4

Brown-out

The motor stall current is enough to cause Vector to brown-out and shut down unexpectedly.

This indicates two possible mechanisms:

- If the system browns out the STM32F030, the MCU will no longer hold the power switch closed, and the system power will be disconnected.
- If the current exceeds a threshold, the MP2617B will disconnect power to the system. This threshold is very high – ~5A – and is unlikely to ever be encountered in operation.

motor stall & brown out effects

Commentary: It may be interesting to modify either the MCU's Vdd to have a larger retaining capacitor, or to add a current limiting mechanism for the motors, such as an inline resistor.

11.1.5

Reducing power

The sensors – the encoders, cliff sensors, and time of flight sensor – have power controls. This allows them to be turned off to reduce power consumption. The time of flight sensor's sampling and communication interval can be controlled to greatly reduce power consumption, while still providing measurements. The other sensors can be duty cycled to maintain a lower power use, but still detect activity (but not measure it accurately).

11.2. ELECTRO-STATIC DISCHARGE (ESD) PROTECTION

The body-board employs a Vishay GMF05, TVS diode (U4) for electro-static discharge (ESD) protection, likely on the pushbutton and touch input.

11.3. STM32F030 MICROCONTROLLER

The body-board is controlled by a STM32F030C8T6 microcontroller (MCU), in a LQFP48 package. This processor essentially acts as a smart IO expander and motor controller. The microcontroller is also referred to as the *system controller*

The MCU's digital inputs:

- 4 opto-couplers used as shaft encoders, one for each motor (left, right, head, lift)
- Momentary push button
- 4 IR proximity sensor used to detect cliffs and lines
- 2 charger state

The MCU's digital outputs:

- 4 motors enable

¹⁰ Other reports suggest up to 2As into the battery, possible with the use of high-power USB adapters intended to support tablet recharge.

As a preventative measure, I have a current limiter between my USB power adapter and Vector's charging dock. 1Ω on the USB power. I tried 1Ω -14Ω; these should have limited the current to 1A and 500mA respectively. Instead, Vector would only pull 40mA - 370mA; in many cases, not enough to charge.

- 4 motors direction
- charger enable
- 3 chip selects

The MCU's analog inputs:

- Touch
- Battery voltage
- Charging pad voltage
- Temperature sensor (picks off the thermistor used by the MP2617)

The communication:

- 2 SPI to LED outputs. Uses an clock and data line to send the state to the LEDs.
- 5 SPI from microphones – an SPI MCLK to clock out, a timer divider (in and out), and 2 MISO to receive state of the data from the microphones.
- I2C for communication with the time of flight sensor
- UART, for communication with the head board

Note: The microcontroller does not have an external crystal¹¹ and uses an internal RC oscillator instead.

11.3.1

Manufacturing test connector

The body-board does include pads that appear to be intended for programming and test at manufacturing time.

11.3.2

Firmware updates

The firmware is referred to as “syscon”. The microcontroller includes a boot loader, allowing the firmware to be updated by the head-board. The firmware can be updated in OTA software releases.

Future changes to the body-board firmware will require expertise. The STM32F030 firmware can be analyzed by analysis of the syscon.dfu file (or be extracted with a ST-Link) and disassembled. By recreating the firmware source codes, the patches replace a key instruction here and there with a jump to the patch, created in assembly (most likely) code to fix or add feature, then jump back.

Emulation (such as QEMU-STM32), ST-link (\$25) and a development environment will be required to debug and modify the firmware initially. The development environment ranges from free to several thousand dollars, the later being the more productive tools.

11.4. SENSING

11.4.1

Time of Flight sensor

The MCU interfaces with a ST Microelectronics VL53L0x time of flight sensor, which can measure the distance to objects in front of vector. It “has a usable range 30mm to 1200mm away (max useful range closer to 300mm for Vector) with a field of view of 25 degrees.”

Anki SDK

¹¹ as far as I can see

These sensors work by timing how long it takes for a coded pulse to return. The time value is then converted to a distance. Items too close return the pulse faster than the sensor can measure. The measured distance is available to the microcontroller over I²C.

11.4.2

Proximity sensing

Has 4 IR proximity sensors that are used to detect drops offs (“cliffs”) and to follow lines. The exact model hasn’t been identified, but the Everlight EAAPMST3923A2 is a typical proximity sensor. The sensor is an LED and IR detector pair. The sensor reports, via I²C, the brightness sensed by the detector. This are often pulsed, to reject to sunlight; and use a configurable threshold to reduce sensitivity to ambient light.

11.4.3

Touch sensing

The touch sensing works by alternating pulsing and sampling (with the ADC) the touch wire. The samples will vary “by various environmental factors such as whether the robot is on its charger, being held, humidity, etc.”

Anki SDK

11.4.4

PDM Microphones

The body-board is responsible to driving and sampling the 4 PDM MEMs microphones. The communication with the backpack board to accomplish this is unique: the four microphones are read at a time, using a shared SPI clock and two separate data lines.

SPI communication with 4 microphones simultaneously

The microphones take a clock signal as input, and always drive one bit per clock; they have no chip select. Two microphones can share a single data line. We’ll refer to them as “left” and “right” here.

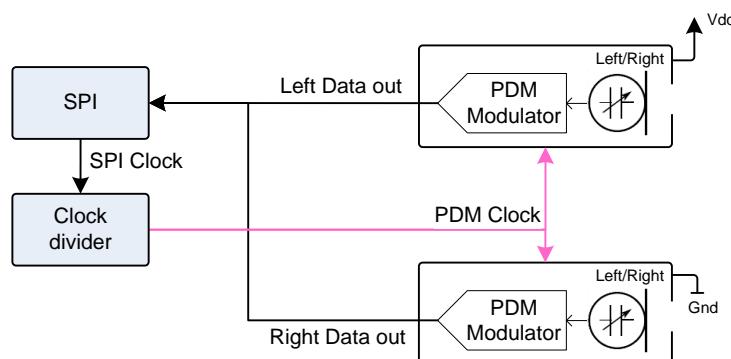


Figure 14: Sampling two microphones with a single SPI master (adapted from ST Microelectronics)

Pulling the left microphone’s “left/right” signal low will configure it to emit the data bit while the PDM clock is low. It does not drive the data line when the clock is high. Similarly, pulling the right microphone’s “left/right” signal high on will cause it to drive the data bit while the PDM clock is high.

SPI, however, only receives data bits on the clock's falling transition— not the rising edge. The trick is to run the SPI clock at twice the frequency of the PDM clocks, so that the SPI clock's first transition low is for the left microphone bit, and the second transition low is for the right microphone. This is done by dividing the SPI clock by two to produce the PDM clock to the microphones:

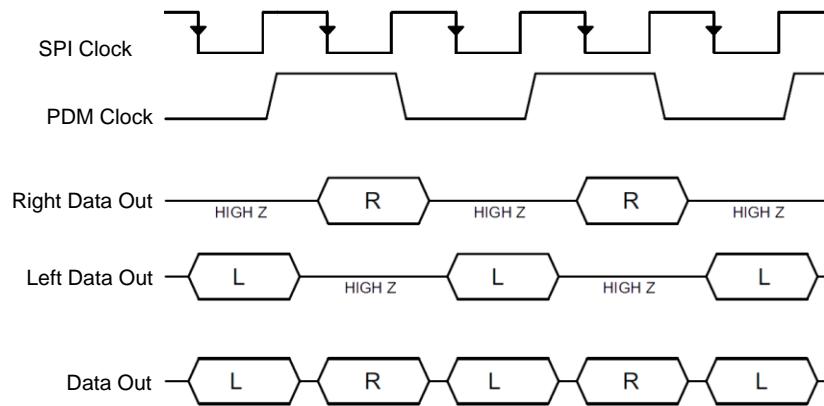


Figure 15:
Microphone clock and
signals

The received data bits (in each byte) will alternate between the left and right microphones, and will need to be separated and converted by firmware. The SPI peripheral along with a DMA can be configured to clock in large batches of bytes into a buffer for further processing.

Dividing the clock by two can be performed by a timer built into the STM32. The SPI clock signal is connected to the input of an STM32 timer (TIMxCHIN). The timer is configured to use an external input clock source, and generate an output after a divide by two. The output of the timer (TIMxCHOUT) can then be used as the clock for the PDM microphones (and 74AHC164 GPIO expander).

The clock rates have a limited range on the body board. PDM MEMS microphones clock rates must be in the range 1 MHz to 3.25MHz. (The products are pretty consistent about this range.) The SPI clock rate is 2x that PDM's clock, so the SPI clock rate must be in the range of 2MHz to 6.5MHz. The ST processor's clock is 48MHz, and its SPI clock must be this frequency divided by a power of two. This means there are only two possibilities: A 32:1 divider gives an SPI clock frequency of 6 MHz, and A 16:1 divider gives a clock rate of 3 MHz.

This approach can be extended to sample all four microphones, by coordinating with a second SPI peripheral:

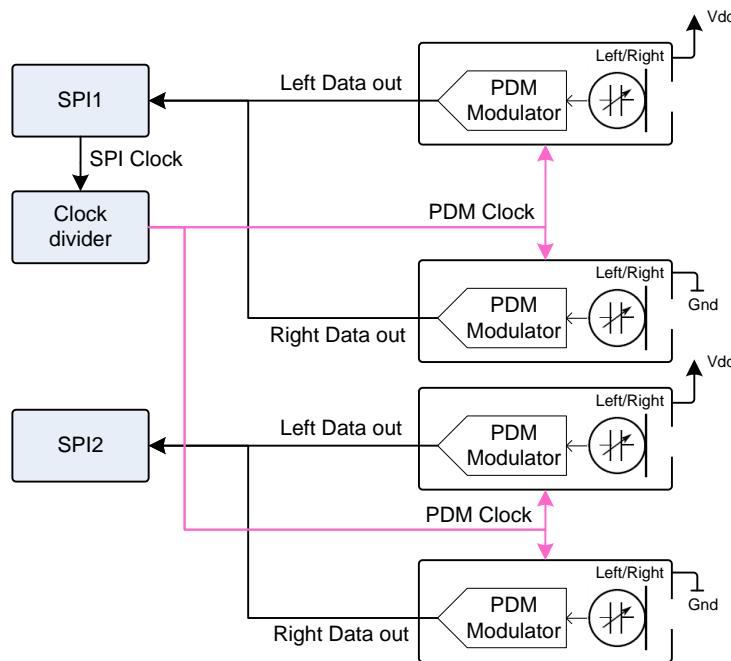


Figure 16: Sampling four microphones with two SPI masters (adapted from ST Microelectronics)

11.5. OUTPUTS

11.5.1 Light control

An earlier section (see section 10.3.1 *The LED controls*) described how the 74AHC164 receives its GPIO settings from a serial interface, and uses these to illuminate the LED segments within 4 RGB LEDs.

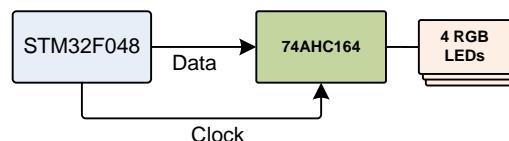


Figure 17: SPI interface to the 74AHC164 and RGB LEDs

The 74AHC164 does not share a clock with the PDM's microphones. The data and clock could be bit-banged – it's only 8 bits and there are little timing requirements – or they could be driven by an SPI peripheral.

Note: care must be taken so that an extra clock edge isn't received by the 74AHC164. (For instance, during body board initialization.) There is no synchronization to indicate the first bit of the 8 bits sent to the 74AHC164.

11.5.2

Motor Driver and control

Each motor driver is an H-bridge, allowing a brushed-DC motor to turn in either direction.

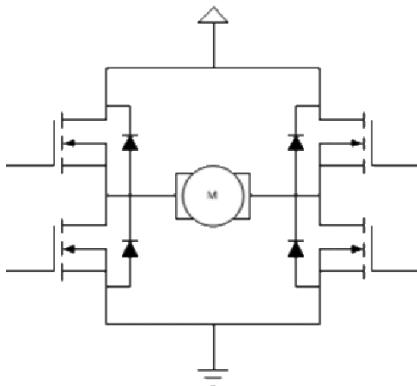


Figure 18: Motor driver H-bridge

Each side of the H-bridge based on the DMC2038LVT, which has a P-FET and N-FET in each package. Two of these are needed for each motor.

The MCU (probably) independently controls the high side and low side to prevent shoot thru. This is done by delaying a period of time between turning off a FET and turning on a FET.

The motors can be controlled with a control loop that takes feedback from the optical encoder to represent speed and position. The firmware must take care to prevent burn out if they have been stalled at full power for 15 seconds or more.

11.6. COMMUNICATION

The body-board communicates with the head-board via RS-232 3.3V (3 Mbits/sec¹²). As the MCU does not have a crystal, there may be communication issues from clock drift at extreme temperatures; since Vector is intended for use at room temperature, the effect may be negligible.

The firmware can be updated over the serial communication by the head-board.

11.7. CONNECTORS & TEST POINTS

The body-board has the following connectors:

Alexander Entinger

- Connector to the head-board
- Connector to the head motor & encoder
- Connector to the lift motor & encoder
- Connector to the time of flight sensor

¹² Value from analyzing the RAMPOST and vic-switchboard programs. Melanie T measured it on an oscilloscope and estimated it to be 2Mbps.

11.7.1 The head-motor connector

P1 Head motor connector (P1)

Pin#	Label	Test point	Cable Color	Description
1	CAI	HENCK	Brown	
2	E2	HENCB	Yellow	
3	E1	HENCA	Green	
4	V _{DD}		White	Head encoder voltage source
5	Motor -		Black	Motor connection
6	Motor +		Red	Motor connection

Table 8: Head Motor Connector (P1) pin map

11.7.2 The lift-motor connector

Lift motor connector (P2)

Pin#	Label	Test point	Cable Color	Description
1	CAI	LENCK	Brown	
2	E2	B	Yellow	
3	E1	A	Green	
4	V _{DD}		White	Lift encoder voltage source
5	Motor -		Black	Motor connection
6	Motor +		Red	Motor connection

Table 9: Lift Motor Connector (P2) pin map

11.7.3 The time of flight connector

Front (Time of Flight) sensor connector

Pin#	Label	Test point	Cable Color	Description
1	V _{DD}	V _{DD} (TP)	Red	Sensor voltage source
2	SCL1		Yellow	I ² C serial clock
3	SDA1		Green	I ² C serial data in/out
4	GND	GND (DC)	Black	Sensor ground reference

Table 10: Front Sensor (time of flight) Connector pin map

11.7.4 The debug connector

Debug connector

Pin#	Label	Description
1	VX	External power supply (connected with charger connector positive.)
2	BAI	Internal power supply for head
3	TX	UART transmit; connects to STM32F030C8T6 Pin #12 (PA2 = USART2_TX)
4	SWCLK	Single wire debug clock signal
5	SWDIO	Single wire debug bi-directional data signal
6	NRST	Processor reset (reset is transition from low to high).
7	GND	Ground

Table 11: Debug Connector pin map

11.7.5 Other body-board test points

Other test points

Test Point	Layer	Description
V _{dd}	Bottom	Body board MCU power supply
BODY_TX	Bottom	RS232 sent from the body board's MCU
SCL2	Top	I ² C serial clock
SDA2	Top	I ² C serial data in/out

Table 12: Head Motor Connector (P1) pin map

12. REFERENCES & RESOURCES

Amitabha, *Benchmarking the battery voltage drain in Anki Vector and Cozmo*, 2018 Dec 31
<https://medium.com/programming-robots/benchmarking-the-battery-voltage-drain-in-anki-vector-and-cozmo-239f23871bf8>

Diodes, Inc, *74AGC164 8-Bit Parallel-Out Serial Shift Registers*, Rev 2, 2015 Aug
<https://www.diodes.com/assets/Datasheets/74AHC164.pdf>

Diodes Inc, *DMG2305UX P-Channel Enhancement Mode MOSFET*
<https://www.diodes.com/assets/Datasheets/DMG2305UX.pdf>

Diodes, Inc, *DMC2038LVT Complementary Pair Enhancement Mode MOSFET*
https://www.diodes.com/assets/Datasheets/products_inactive_data/DMC2038LVT.pdf

Entinger, Alexander; *Anki Vector base-board connector*
<https://github.com/aenting/anki-vector-baseboard>

Everlight EAAPMST3923A2

Monolithic Power, *MP2617A, MP2617B 3A Switching Charger with NVDC Power Path Management for Single Cell Li+ Battery*, Rev 1.22 2017 Jun 29
https://www.monolithicpower.com/pub/media/document/MP2617A_MP2617B_r1.22.pdf

Panda, a data sheet for a similar single-cell lithium battery
<https://panda-bg.com/datasheet/2408-363215-Battery-Cell-37V-320-mAh-Li-Po-303040.pdf>

Sharp *GP1S092HCPIF Compact Transmissive Photointerrupter*, 2005 Oct 3
https://datasheet.lcsc.com/szlcsc/Sharp-Microelectronics-GP1S092HCPIF_C69422.pdf

- ST Microelectronics, *STM32F030x8*, Rev 4, 2019-Jan
<https://www.st.com/resource/en/datasheet/stm32f030c8.pdf>
- ST Microelectronics. *AN5027 Application Note: Interfacing PDM digital microphones using STM32 MCUs and MPUs*, Rev 2, 2019 July
https://www.st.com/resource/en/application_note/dm00380469-interfacing-pdm-digital-microphones-using-stm32-mcus-and-mpus-stmicroelectronics.pdf
- ST Microelectronics. *Touch sensing*
https://www.st.com/content/ccc/resource/technical/document/application_note/group0/ed/0d/4d/87/04/1d/45/e5/DM00445657/files/DM00445657.pdf/jcr:content/translations/en_DM00445657.pdf
<https://www.st.com/en/embedded-software/32f0-touch-lib.html>
<https://hsel.co.uk/2016/05/22/stm32f0-software-capacitive-touch/>
<https://github.com/pyrohaz/STM32F0-SoftTouch>
- ST Microelectronics. *Tutorial for MEMS microphones*, Rev2, 2017 Feb
https://www.st.com/resource/en/application_note/dm00103199-tutorial-for-mems-microphones-stmicroelectronics.pdf
- ST Microelectronics. *VL53L0X World's smallest Time-of-Flight ranging and gesture detection sensor*, Rev 2, 2018 Apr
<https://www.st.com/en/imaging-and-photonics-solutions/vl53l0x.html>
<https://www.st.com/resource/en/datasheet/vl53l0x.pdf>

CHAPTER 5

Accessory Electronics Design Description

This chapter describes the electronic design of the Anki Vector accessories:

- The charging station
- The companion cube

13. CHARGING STATION

The charging station is intended to provide energy to the Vector, allowing it to recharge.

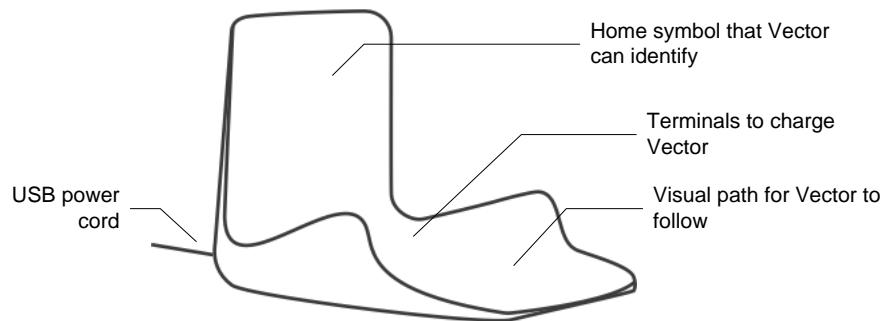


Figure 19: Charging station main features

The charging station has a USB cable that plugs into an outlet adapter or battery. The adapter or battery supplies power to the charging station. The base of the station has two terminals to supply +5V (from the power adapter) to Vector, allowing him to recharge. The terminals are offset in such a way to prevent Vector from accidentally being subject to the wrong polarity. Vector has to be backed into charging station in mate with the connectors. Vector face-first, even with his arms lifted, will not contact the terminals.

The charging station has an optical marker used by Vector to identify the charging station and its pose (see chapter 19).

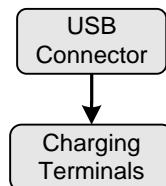


Figure 20: Charging station block diagram

14. CUBE

This section describes the companion cube accessory. The companion cube is a small toy for Vector play with. He can fetch it, roll it, and use it to pop-wheelies. Each face of the cube has a unique optical marker used by Vector to identify the cube and its pose (see chapter 17).

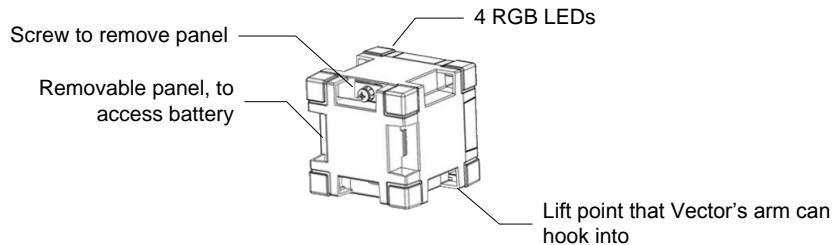


Figure 21: Cube's main features

Although the companion cube is powered, this is not used for localization or pose. The electronics are only used to flash lights for his human companion, and to detect when a person taps, moves the cube or changes the orientation.

The cube has holes near the corners to allow the lift to engage, allowing Vector to lift the cube. Not all corners have such holes. The top – the side with the multicolour LEDs – does not have these. Vector is able to recognize the cubes orientation by symbols on each face, and to flip the cube so that it can lift it.

The electronics in the cube are conventional for a small Bluetooth LE accessory:

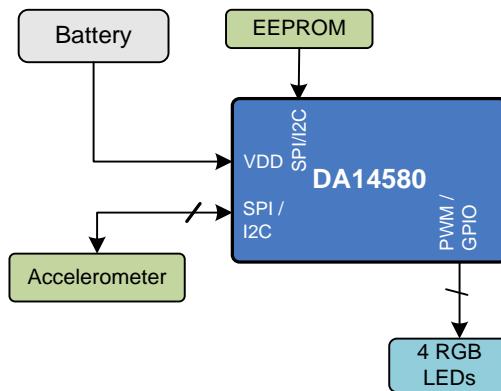


Figure 22: Block diagram of the Cube's electronics

The Cube's electronic design includes the following elements:

Element	Description
accelerometer	Used to detect movement and taps of the cube.
battery	The cube is powered by a 1.5 volt N / E90 / LR1 battery cell. ¹³
crystal	The crystal provides the accurate frequency reference used by the Bluetooth LE radio.
Dialog DA14580	This is the Bluetooth LE module (transmitter/receiver, as well as microcontroller and protocol implementation).
EEPROM	The EEPROM holds the updatable application firmware.
RGB LEDs	There are 4 RGB LEDs. They can flash and blink. Unlike the backpack LEDs, two LEDs can have independent colors.

Table 13: The Cube's electronic design elements

The communication protocol is given in Appendix F.

14.1. OVER THE AIR FIELD UPDATES

The DA14580 has a minimal ROM boot loader that initializes hardware, moves a secondary boot loader from “One Time Programmable” ROM (OTP) into SRAM, before passing control to it. The firmware is executed from SRAM to reduce power consumption. The secondary boot loader loads the application firmware from I2C or SPI EEPROM or flash to SRAM and pass control to it.

If the application passes control back to the boot loader – or there isn't a valid application in EEPROM –a new application can be downloaded. The boot loader uses a different set of services and characteristics to support the boot loading process.

14.2. REFERENCES & RESOURCES

Dialog, *SmartBond™ DA14580 and DA14583*

<https://www.dialog-semiconductor.com/products/connectivity/bluetooth-low-energy/smartbond-da14580-and-da14583>

Dialog, *DA14580 Low Power Bluetooth Smart SoC, v3.1, 2015 Jan 29*

Dialog, *UM-B-012 User manual DA14580/581/583 Creation of a secondary bootloader, CFR0012-00 Rev 2, 2016 Aug 24*

https://www.dialog-semiconductor.com/sites/default/files/um-b-012_da14580_581_583_creation_of_a_secondary_boot_loader_v3.2.pdf

Dialog, *Application note: DA1458x using SUOTA, AN-B-10, Rev 1, 2016-Dec-2*

https://www.dialog-semiconductor.com/sites/default/files/an-b-010_da14580_using_suota_0.pdf

¹³ The size is similar to the A23 battery, which will damage the cube's electronics.

PART II

Basic Operation

This part provides an overview of Vector's software design.

- THE SOFTWARE ARCHITECTURE. A detailed look at Vector's overall software architecture and main modules.
- STARTUP. A detailed look at Vector's startup and shutdown processes
- POWER MANAGEMENT. A detailed look at Vector's architecture for battery monitoring, changing and other power management.
- BASIC INPUT AND OUTPUT. A look at push button, touch sensing, surface proximity sensors, time of flight proximity sensing, and backpack LEDs.
- INERTIAL MOTION SENSING



[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 6

Architecture

This chapter describes Vector's software architecture:

- The architecture
- The emotion-behaviour system
- The communication infrastructure
- Internal support

15. OVERVIEW OF VECTOR'S COMMUNICATION INFRASTRUCTURE

Vector's architecture has a structure something like:

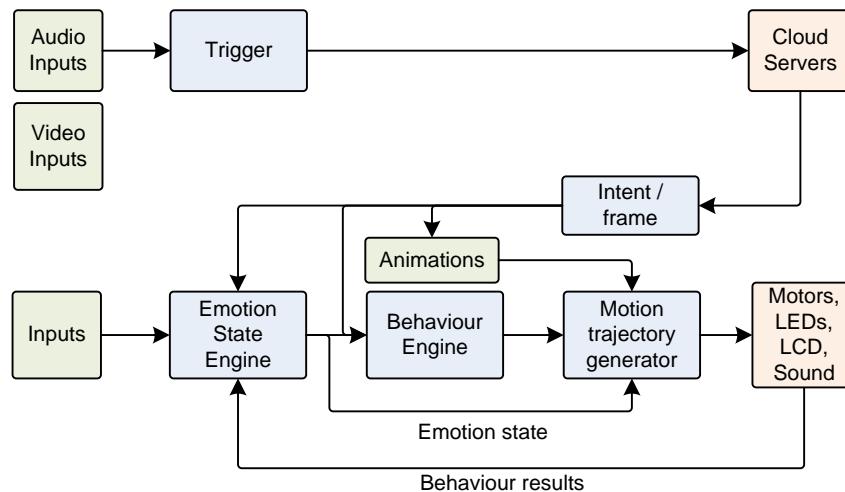


Figure 23: The overall functional block diagram

Fast control loops — to respond quickly — are done on the Vector's hardware. Other items, processing heavy including (but not limited to) speech recognition, natural language processing, and training for faces are sent to the cloud.

Vector is built on a version of Yocto Linux. Anki selected this for a balance of reasons: some form of Linux is required to use the Qualcomm processor, the low up front (and royalty) costs, the availability of tools and software modules. The Qualcomm is a multi-processor, with four main processing cores and a GPU. Vector runs a handful of different application programs, in addition to the OS's foundational service tasks and processes.

explored in Casner,
and Wiltz

15.1. APPLICATION SERVICES ARCHITECTURE

Vector's software is divided into the following services:

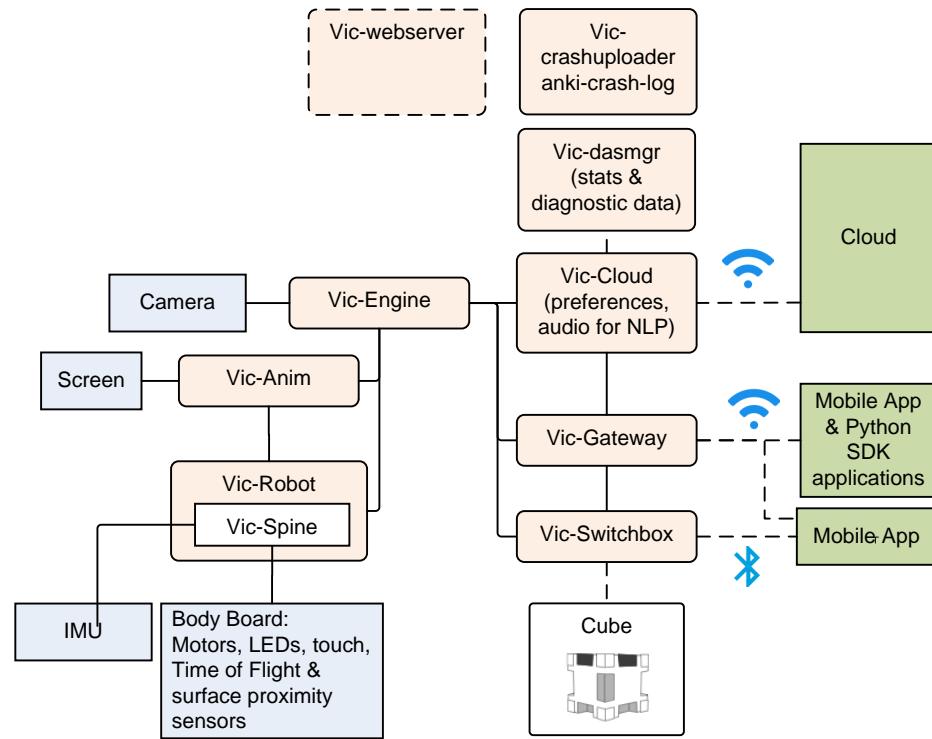


Figure 24: The overall communication infrastructure

These services are:

Services	Speculated purpose
vic-animate	Probably plays multi-track animations (which include motions as well as LCD display and sound) config file: /anki/etc/config/platform_config.json /anki/data/assets/cozmo_resources/webserver/webServerConfig_anim.json
vic-bootAnim	LCD and sound animations during boot
vic-cloud	Connects to the cloud services for natural language
vic-crashuploader anki-crash-log	A service that sends logs (especially crash logs and mini-dumps) to remote servers for analysis.
vic-dasmgr	Gathering data on processor and feature usage, possibly intended to serve as a foundation for gathering data when performing experiments on settings and features.
vic-engine	The behaviour / emotion engine. Hooks into the camera face recognizer.
vic-gateway	Responsible for the local API/SDK services available as gRPC services on https.
vic-robot	Basic power management. Resets watchdog timer. Internally has “vic-spine” that manages the sensors.
vic-switchboard	Supports the Bluetooth LE communication interface, including the mobile application protocol (see Chapter 12). Routes messages between the other services? Manages the access keys

Table 14: Vector services & processes

Within each vic- server processes, there are one or more event-driven communication threads. A thread has the following basic structure:

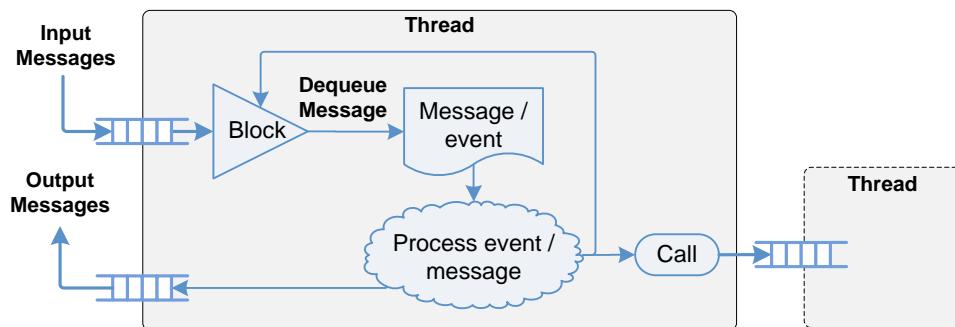


Figure 25: Basic communication thread structure

The communication threads have some form of input message queue. On Vector these include

- a socket, between processes
- A serial interface with the body board
- A web-socket
- Other, inter-thread message queue

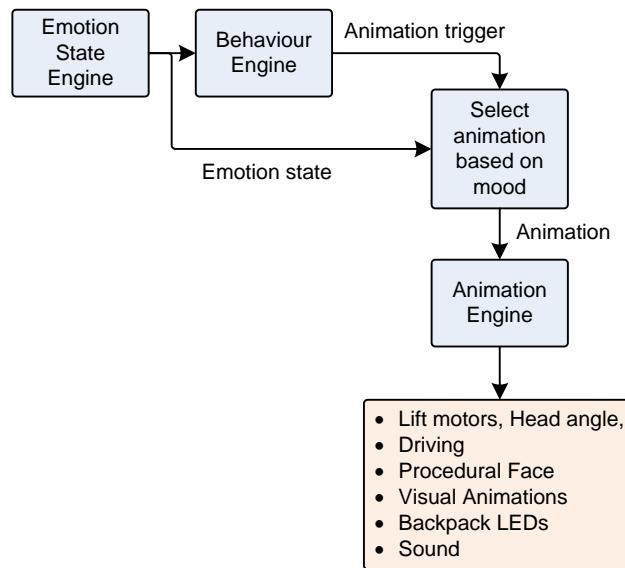
The communication thread blocks on one or more message queue events. It is woken when there is an incoming event/message, or there has been an error or timeout while wait. When it wakes, it dequeues the event, takes action and goes back to waiting. It may post messages (or other signals) to other threads, possibly indirectly as a result of framework/library/system calls.

Within a server process, convenient C++ data structures are used. The vic- servers also use CLAD, and JSON data structures, and include many helper procedures to convert between the two. It appears that a process interprets and generates a JSON data structure. To communicate with another process, it converts the JSON to a CLAD (since it is a contiguous span of bytes), sends that to the other process; the other process reverses the process, converting it JSON and using that interpret the message.

15.2. EMOTION MODEL, BEHAVIOUR ENGINE, ACTIONS AND ANIMATION ENGINE

Vector's high-level AI is organized around an *emotion model*, and a *behaviour engine* that drives goals, responses and other actions.

Figure 26: The behaviour-animation flow



There are many similar terms, but there are subtle distinctions between them:

- An *AI Feature* is the high level behaviors as a person would experience. There are about 70 of these. Note the name shouldn't be confused with a *feature flag* or *feature toggle*; that is a different concept, for software elements that are not ready yet, but included in the code base.
- A *behavior* is “a complex task [that] may include combinations of animation, path planning or other functionality. Examples include” driving to the charger, set the lift height, etc. An AI Feature takes at least one behavior to carry out; it often takes many. The current emotional state can influence which behavior is selected, and affect how it is carried out. Intents (response to voice interaction) can initiate behaviors. Behaviors can initiate actions.
- An *action* is like a mini-behavior, with some differences. Multiple actions can run at a time – so long as they don’t use the same resources– but only one behavior can run at a time. Actions can wait in a queue.
- An *animation* is a scripted motion, sound, light pattern, and facial animation (or picture on the display) that Vector carries out. Behaviors and actions can initiate animations. The animation engine selects the specific animation, from a pool of alternatives, based on context and current emotional state. An animation can’t use the sensors, so it can’t adapt to the environmental conditions. For instance, to drive up to a hand (or a cube) requires the time of flight sensor; so an action is required.

16. STORAGE SYSTEM

Vector’s system divides the storage into many regions, primarily based on whether the region is modifiable (and when), and which subsystem manages the data. Appendix E describes the flash

partitions and file system structure. See chapter 7 for a description of the partitions used for system start up and restore.

Most of the partitions on the flash storage are not modifiable – and are checked for authenticity (and alteration). These partitions hold the software and assets as delivered by Anki (and Qualcomm) for a particular release of the firmware. They are integrity checked as part of the start procedure. (See Chapter 7 for a description.)

Data that is specific to the robot, such as settings, security information, logs, and user data (such as pictures) are stored in modifiable partitions. Some of this data is erased when the unit is “reset” to factory conditions

These are described below.

16.1. ELECTRONIC MEDICAL RECORDS (EMR)

Vector’s “Electronic Medical Record” (EMR) partition holds the following information:

Offset	Size	Type	Field	Description
0	4	uint32_t	<i>ESN</i>	Vector’s electronic serial number (ESN). This is the same serial number as printed on the bottom of Vector. Serial numbers starting with 00 are engineering units.
4	4	uint32_t	<i>HW_VER</i>	Hardware revision code
8	4	uint32_t	<i>MODEL</i>	The model number of the product
12	4	uint32_t	<i>LOT_CODE</i>	The manufacturing lot code
16	4	uint32_t	<i>PLAYPEN_READY_FLAG</i>	The manufacturing fixture tests have passed, it ok to run play pen tests.
20	4	uint32_t	<i>PLAYPEN_PASSED_FLAG</i>	Whether or not Vector has passed the play pen tests.
24	4	uint32_t	<i>PACKED_OUT_FLAG</i>	
28	4	uint32_t	<i>PACKED_OUT_DATE</i>	(In unix time?)
32	192	uint32_t[4]	<i>reserved</i>	
224	32	uint32_t[8]	<i>playpen</i>	
256	768	uint32_t[192]	<i>fixture</i>	

This information is not modified after manufacture; it persists after a device reset or wipe.

16.1.1 FAC Mode

Vector has a “FAC” mode, used in the factory to test and calibrate the robot. When in FAC mode, the display has a red background, with either the letters “FAC” or one two two digits displayed (these are likely the testing stage to be performed), and his backpack lights have an unusual color pattern – red, green, and blue.

This mode is never intended to be seen outside of the factory, so little is known. Only a couple of units have been found in this mode; one after it had been intentionally damaged, and its

Table 15: Electronic Medical Record (EMR)

Figure 27: The LED pattern when in FAC mode



calibration & EMR data were corrupted or inaccessible.¹⁴ In all likelihood, the software checks to see its EMR to see if it has been released; if not, it enters the FAC mode at whatever the “next” stage is according to the EMR. At that point Vector would expect to be placed into manufacturing test fixtures, such as the playpen.

16.1.2 Manufacturing Lot Codes

A manufacturing lot code is an identifier that used to track the components, and robot subassemblies that were used in robots, as well as the date they were made. “If there’s a problem in a particular batch of components (or maybe the people working at the factory that day), we can identify which robots were affected.”

“A lot code is 4 numbers. A typical lot code is 2 18 36 201.

- “2 is the factory. All Vectors were made at factory 2.”
- 18 is the last two digits of the year, 2018.
- “36 means week 36 of 2018 - that’s first week of September.
- “201 means ‘Standard Edition Vector, US-only version’”

The robots were made “in big batches in July/August, and they didn’t start coming back [to customer service] until January/February,” when Anki would “put the fixes into the next big batch the upcoming year.”

16.1.3 Playpen Data

The playpen is a testing area with ramps, barriers, camera targets at a variety of angles, cube and a charging station. Vector is put into one during manufacture to check his sensors, camera calibration, motor function, microphone and a check over his overall functions. The playpen tests involve many checks to ensure that his head is assembled and attached correctly, as well as that his lower body is assembled correctly. These use almost of all of his functions: that he can correctly navigate, detect cliffs, see and count dots, see markers (getting their type and size correct), dock, and charge.

The images that Vector sees during these tests are kept with unit. This way, if the unit is returned later with a vision-related problem, the images from the manufacturing are there to see if . as part of the mfg record for analysis of returned products “, we can go back to those images and see if it’s a new problem or was always there.”

There is also a sound booth that checked that his speaker was working properly and did not exceed limits.

¹⁴ <https://forums.anki.com/t/any-one-know-what-error-code-50-is/40891>

16.2. OEM PARTITION FOR OWNER CERTIFICATES AND LOGS

The OEM partition is a read/writeable ext2 file-system. It is used to hold the SDK certificate folders:

Folder	Description
	The top level holds the log files.
<i>cloud</i>	Holds the SDK TLS certificate and signing keys. With newer firmware, the folder may also hold some other calibration information.
<i>nvStorage</i>	holds some binary “.nvdata” files

Table 16: OEM partition file hierarchy

17. SECURITY AND PRIVACY

Vector’s design includes a well thought-out system to protect against disclosing (i.e. providing to strangers) sensitive information, and allowing the operator to review and delete it at any time:

[Anki Security & Privacy Policy](#)

- Photographs taken by Vector are not sent to (nor stored in) a remote server. They are stored in encrypted file system, and only provided to authenticated applications on the local network. Each photograph can be individually deleted (via the mobile application).
- The image stream from Vector’s camera is not sent to a remote server. It is only provided to authenticated applications on the local network.
- The data used to recognize faces¹⁵ and the names that Vector knows are not sent to (nor stored in) a remote server. The information is stored in encrypted file system. The list of known faces (and their names) is only provided to authenticated applications on the local network. Any facial recognition data not associated with a name is deleted when Vector goes to sleep. Facial data associated with an individual name can be deleted (along with the name) via the mobile application.
- “[After] you say the wake words, “Hey Vector”, Vector streams your voice command to the cloud, where it is processed. Voice command audio is deleted after processing. Text translations of commands are saved for product improvement not associated with a user.”
- The audio stream from the microphone — if it had been finished being implemented – would have been provided to authenticated applications on the local network.
- Information about the owner
- Control of the robots movement, speech & sound, display, etc. is limited to authenticated applications on the local network.

Vector’s software is protected from being altered in a way that would impair its ability to secure the above.

Vector also indicates when it is doing something sensitive:

- When the microphone is actively listening, it is always indicated on the backpack lights (blue).

¹⁵ The Anki privacy and security documents logically imply that the face image is not sent to Anki servers to construct a recognition pattern. I would have guessed heavy crunching would have been performed on a server.

- The microphone is enabled by default, but only listening for the wake word, unless Vector's microphone has been disabled.
- When the camera is taking a picture (to be saved), Vector makes a sound
- When the camera is XYZ on?
- Unless the backpack lights are all orange, the WiFi is enabled. (All orange indicates it is disabled.)

17.1. ENCRYPTED COMMUNICATION

The personally identifying information and other data about the owner — photos, account information, WiFi passwords, and so one — is only sent on encrypted channels.

17.2. ENCRYPTED FILESYSTEM

The file system with the users data — photos, account information, WiFi passwords, and so one — is encrypted.

17.3. THE OPERATING SYSTEM

There is a chain of firmware signed by Anki. This is intended to protect Vector's software from being altered in a way that would impair its ability to secure the above information.

Android boot loaders typically include a few powerful (but unchecked) bits that disable the signature checking, and other security features. These bits typically are set either thru commands to the firmware during boot up, by applications, or possibly by hack/exploit. Sometimes this requires disassembling the device and shorting some pins on the circuit board.

Vector doesn't support those bits, nor those commands. Signature checking of the boot loader, kernel and RAM disk can't be turned off.

17.3.1 The possibility for future modifications to Vector's software

There may be a way to disable checking of the system file system and its software.

Anki created special Vectors for internal development. The software for these units has a special version of the kernel and RAM disk that does not check system room file system, and makes it writable. This file system has Vector's application soft, supports SSH. This software was tightly controlled, and “only .., available inside the Anki corporate network.” For purposes of customizing and updating Vector, this version is essential. (Note: the kernel and RAM disk can't be modified.)

Jane Fraser, 2019

Note: the OTA software has a “dev” (or development) set of OTA packages. Those packages are not the same; they are essential software release candidates being pushed out for test purposes.

17.4. AUTHENTICATION

The web services built into Vector require a token. This is used to prove that you have authenticated (with the more capable — and not physically accessible — servers). This authentication is to protect:

- Photos already on Vector
- The image stream from the camera

- The audio stream from the microphone — if it had finished being implemented
- The sensitive owner information
- Controlling the robot

(That is to say, to prevent disclosure)

18. CONFIGURATION AND ASSET FILES

The Anki vector software is configured by JSON files. Some of the JSON files were probably created by a person (for the trivial ones). Others were created by scripting / development tools; a few of these were edited by developers. These JSON files are clearly intended to be edited by people:

- The files are cleanly spaced, not in the most compact minimized size
- The JSON parser supports comments, which is not valid JSON. Many files have comments in them. Many have sections of the configuration that are commented out.

18.1. CONFIGURATION FILES

The top-level configuration file provides the paths to the network other configuration files. It is found at:

`/anki/etc/config/platform_config.json`

This path is hardcoded into the vic-dasmgr, and provided in the editable startup files for vic-anim and vic-engine. The configuration file contains a JSON structure with the following fields:

Field	Value	Description & Notes
<code>DataPlatformCachePath</code>	<code>"/data/data/com.anki.victor/cache"</code>	
<code>DataPlatformPersistentPath</code>	<code>"/data/data/com.anki.victor/persistent"</code>	
<code>DataPlatformResourcesPath</code>	<code>"/anki/data/assets/cozmo_resources"</code>	The path to most configuration files and assets

Table 17: The platform config JSON structure

When describing the configuration and asset files, a full path will be provided. When the path is constructed from different parts, the part that is specified in another configuration or binary file will be outlined. The path to a settings file might look like:

`/anki/assets/cozmo_resources/config/engine/settings_config.json`

The path leading up to the settings file (not outlined in red) is specified in an earlier configuration file, usually the platform configuration file described above.

19. SOFTWARE-HARDWARE LAYERS

- Body-board input/output software architecture
- The LCD display
- Camera

19.1. THE BODY BOARD INPUT/OUTPUT

The body-board input-output software has a structure like so:

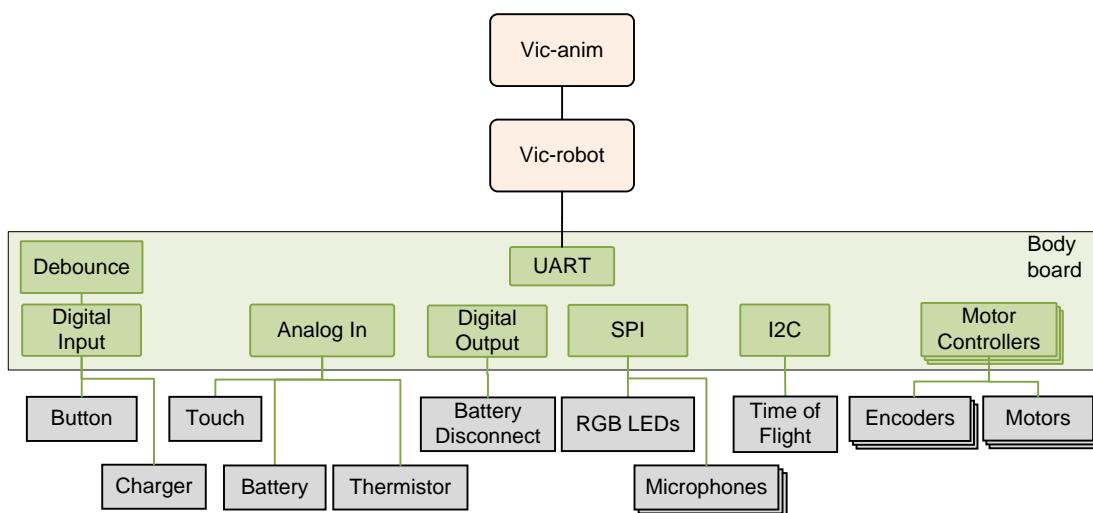


Figure 28: The body board-related architecture

19.2. THE LCD DISPLAY

Four different applications may access the display, albeit not at the same time:

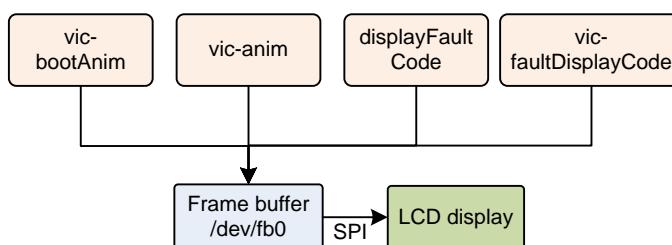


Figure 29: The LCD architecture

Note: `displayFaultCode` is present on Vector, but it is not called by any program.

The LCD is connected to the MPU via an SPI interface (`/dev/spidev1.0`). The frame buffer (`/dev/fb0`) is essentially a buffer with metadata about its width, height, pixel format, and orientation. Application modifies the frame buffer by `write()` or `memmap()` and modifies the bytes. Then the frame buffer has the bytes transfer (via SPI) to the display.

`vic-animate` employs a clever screen compositing system to create Vector's face (his eyes), animate text jumping and exploding, and small videos, such as rain or fireworks.

The `vic-faultDisplayCode` and Customer Care Information Screen of `vic-animate` have a visual aesthetic is unlike the rest of Vector. These modes employ a barebones system for the display.

The text appears to be rendered into the buffer using OpenCV's `putText()` procedure, and transferring it to the display without any further compositing.

Not sure if the transfer is in a driver, in the kernel, or in user space... or which process would have it in user space.

19.3. THE CAMERA

The camera subsystem has the following architecture:

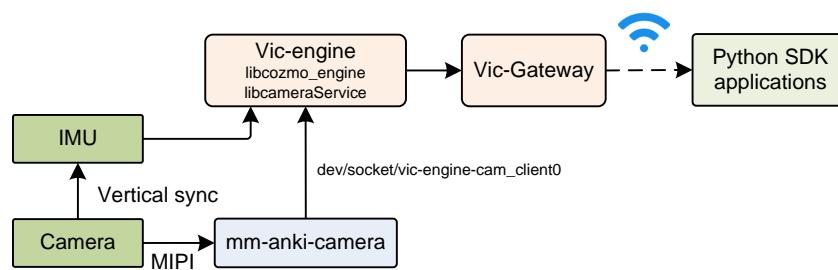


Figure 30: The camera architecture

Daniel Casner, 2019
Embedded Vision
Summit

The camera's vertical synchronization signal is connected to the interrupt line on IMU, triggering accelerometer and gyroscope sampling in sync with the camera frame. The vision is used as a navigation aid, along with the IMU data. The two sources of information are fused together in the navigation system (see chapter 18) to form a more accurate position and relative movement measure. The image must be closely matched in time with the IMU samples. However the transfer of the image from the camera to the processor, then thru several services to vic-engine introduces variable or unpredictable delays. The camera's vertical sync – an indication of when the image is started being sampled – is used to trigger the IMU to take a sample at the same time.

The camera is also used as an ambient light sensor when Vector is in low power mode (e.g. napping, or sleeping). In low power mode, the camera is suspended and not acquiring images. Although in a low power state, it is still powered. The software reads the camera's auto exposure/gain settings and uses these as an ambient light sensor. (This allows it to detect when there is activity and Vector should wake.)

20. REFERENCES & RESOURCES

Anki, *Elemental Platform*

<https://anki.com/en-us/company/elemental-platform.html>

Describes, as a marketing brochure, much of Anki's product network architecture.

Anki, *Vector Security & Privacy FAQs*, 2018

<https://support.anki.com/hc/en-us/articles/360007560234-Vector-Security-Privacy-FAQs>

Casner, Daniel, *Consumer Robots from Smartphone SoCs*, Embedded Systems Conference Boston, 2019 May 15

<https://schedule.esc-boston.com//session/consumer-robots-from-smartphone-socts/865645>

Stein, Andrew; Kevin Yoon, Richard Alison Chaussee, Bradford Neuman, Kevin M.Karol, US Patent US2019/01563A1, *Custom Motion Trajectories for Robot Animation*, Anki, filed 2018 Jul 13, published 2019 Apr 18,

Qualcomm, *Snapdragon™ 410E (APQ8016E) r1034.2.1 Linux Embedded Software Release Notes*, LM80-P0337-5, Rev. C, 2018 Apr 10
lm80-p0337-5_c_snapdragon_410e_apq8016e_r1034.2.1_linux_embedded_software_revc.pdf

Tariq, Talha *Securing Autonomous Robots at Scale*, 2018 Oct 3

https://www.infosecuritynorthamerica.com/RXUK/RXUK_InfosecurityNorthAmerica/16.05-16.30_Anki.pdf

Wiltz, Chris, *Lessons After the Failure of Anki Robotics*, Design News, 2019 May 21

<https://www.designnews.com/electronics-test/lessons-after-failure-anki-robotics/140103493460822>

CHAPTER 7

Startup

This chapter describes Vector's start up and shutdown processes:

- The startup process
- The shutdown steps

21. STARTUP

Vector's startup is based on the Android boot loader and Linux startup.¹⁶ These are otherwise not relevant to Vector, and their documentation is referred to. The boot process gets quite far before knowing why it booted up or being able to respond in complex fashion.

1. The backpack button is pressed, or Vector is placed into the charger. This powers the body board, and the head-board.
2. The body-board displays an animation of the backpack LEDs while turning on.
 - a. If turned on from a button press and the button is released before the LED segments are fully lit, the power will go off.
 - b. If held—about 5 seconds—the head-board will have reach a point in its boot process to direct the body-board to keep the battery switch closed.
 - c. If held for 15 seconds, the body-board will hold its TX line—the head-boards RX line—low during the boot process. This tells the system to boot into recovery mode.
3. While the head-board boots, the body-board performs several self tests. These include checking that the microcontroller can communicate with the 4 cliff (surface proximity) sensors, and the time of flight sensor.

21.1. QUALCOMM'S PRIMARY AND SECONDARY BOOTLOADER

Meanwhile, on the head-board:

1. “Qualcomm’s Primary Boot Loader is verified and loaded into [RAM] memory¹⁷ from BootROM, a non-writable storage on the SoC. [The primary boot loader] is then executed and brings up a nominal amount of hardware,” Nolen Johnson
2. The primary boot loader checks to see if a test point is shorted on the board, the unit will go into emergency download (EDL) mode. It is known that when F_USB pad on the head-board is pulled to Vcc, USB is enabled; this may be the relevant pin. Roee Hay
3. If the primary boot loader is not in EDL mode it “then verifies the signature of the next boot loader in the chain [the secondary bootloader], loads it, [and] then executes it.” The Nolen Johnson

¹⁶ An ideal embedded system has a fast (seemingly instant) turn on. Vector's startup *isn't* fast. The steps to check the integrity of the large flash storage – including checking the security signatures – and the complex processes that Linux provides each contribute to the noticeable slow turn on time. Checking the signatures is inherently slow, by design.

¹⁷ The boot loader is placed into RAM for execution to defeat emulators.

secondary boot loader is stored in the flash partition SBL.

4. If the secondary boot loader does not pass checks, the primary boot loader will go into emergency down load mode.
5. “The next boot loader(s) in the chain are SBL*/XBL (Qualcomm’s Secondary/eXtensible Boot Loader). These early boot loaders bring up core hardware like CPU cores, the MMU, etc. They are also responsible for bringing up core processes .. [for] TrustZone. The last purpose of SBL*/XBL is to verify the signature of, load, and execute aboot/ABL [Android boot loader].”

The Android boot loader (`aboot`) is stored on the “ABOOT” partition.

The secondary bootloader also supports the Sahara protocol; it is not known how to activate it.

Note: The keys for the boot loaders and TrustZone are generated by Qualcomm, with the public keys programmed into the hardware fuses before delivery to Anki or other customers. The signed key pair for the secondary boot loader is not necessarily the same signed key pair for the `aboot`. They are unique for each of Qualcomm’s customer. Being fuses they cannot be modified, even with physical access.

21.2. ANDROID BOOTLOADER (ABOOT)

1. “Aboot brings up most remaining core hardware then in turn normally verifies the signature of the boot image, reports the verity status to Android Verified boot through `dm-verity`... On many devices, Aboot/ABL can be configured to skip cryptographic signature checks and allow booting any kernel/boot image.”
 - a. On other devices, `aboot` reads the `DEVINFO` partition for a structure. It checks the header of the structure for a magic string (“ANDROID-BOOT!”) and then uses the values internally to indicate whether or not the device is unlocked, whether verity-mode enabled or disabled, as well as a few other settings. By writing a version of this structure to the partition, the device can be placed into unlock mode.

Vector does *not* support this method of unlocking.
 - b. “The build system calculates the SHA256 hash of the raw `boot.img` and signs the hash with the user’s private key... It then concatenates this signed hash value at the end of raw `boot.img` to generate signed `boot.img`.”
 - c. “During bootup, [Aboot]¹⁸ strips out the raw `boot.img` and signed hash attached at the end of the image. [Aboot] calculates the SHA256 hash of the complete raw `boot.img` and compares it with the hash provided in the `boot.img`. If both hashes match, kernel image is verified successfully.”
2. ABoot can either program the flash with software via boot loader mode, or load a kernel. The kernel can be flagged to use a recovery RAM disk or mount a regular system.
3. If recovery mode, it will load the kernel and file systems from the active RECOVERY partitions.
 - a. Recovery is entered if the active regular partition cannot be loaded, e.g. doesn’t exist or fails signature check, or

Roei Hay

Qualcomm LM80
P0436

¹⁸ The Qualcomm document speaks directly about Little Kernel; ABoot is based on Little Kernel.

- b. The RX signal from the body-board may be held low when `aboot` starts, indicating that the operator has held the button and wishes to initiate recovery mode.¹⁹ If this is the case, “`anki.unbrick=1`” is prepended to the command line passed to the kernel.
- 4. ABoot loads the kernel and RAM file system from the active “BOOT” partition and passes it command line to perform a check of the boot and RAM file system the signatures.²⁰ The command line is stored in the header of the boot partition; it is checked as part of the signature check of the boot partition and RAM file system. If the ABoot is compiled for a developer robot, it will add an “`anki.dev`” to the command line.

Many of these elements will be revisited in Chapter 30 where updating `aboot`, `boot`, and `system` partitions are discussed.

21.3. REGULAR SYSTEM BOOT

The boot partition holds the linux kernel, and a small RAM disk to initialize the system. It is passed parameters on the command line from `aboot` and from the `boot.img`. The purpose of the extra (Anki-specific) command line parameters are:

Table 18: linux command line parameters

Field	Description
<code>anki.unbrick</code>	This is used to trigger a boot into recovery.
<code>anki.dev</code>	This is set to confirm (to the linux system) that this robot is a development robot and can run development software systems.
<code>dm=</code>	The dm-verity command line used to verify the system file system

¹⁹ This seems to imply that the body-board a resets/restarts the head-board so that the bootloader to runs again.

²⁰ The check specifies the blocks on the storage to perform a SHA256 check over and provides expected signature result.

After the kernel has finished loading, it launches init. In Vector, it is a shell script with Anki-specific system checks:

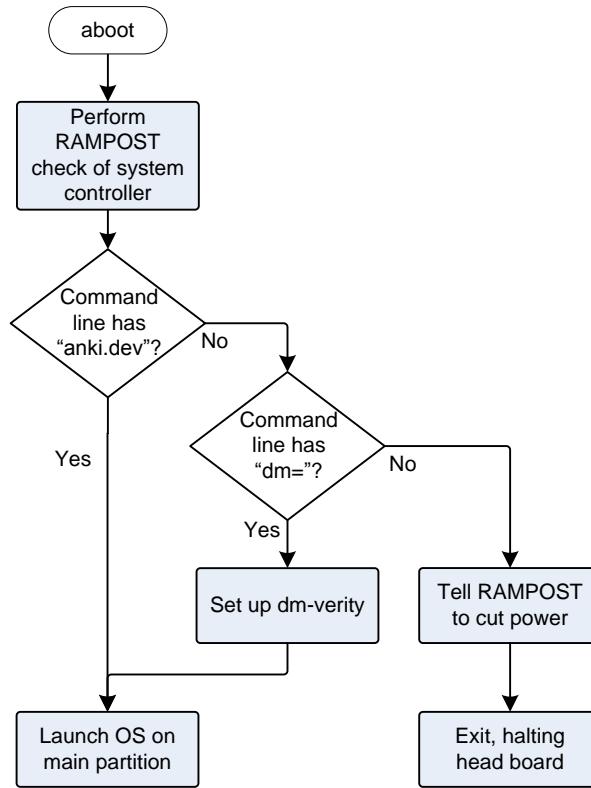


Figure 31: The linux boot-partition init-script flow

These Anki-specific system checks are:

1. The RAM file system contains primarily of two programs: init and /bin/rampost. init is a shell script and the first program launched by the kernel. This script turns on the LCD, its backlight and initiate communication with the body-board. (These occur ~6.7 seconds after power-on is initiated).
 - a. rampost will perform a firmware upgrade of the body-board if its version is out of date. It loads the firmware from syscon.dfu (Note: the firmware in the body-board is referred to as syscon.)
 - b. rampost checks the battery voltage, temperature and error flags. It posts any issues to /dev/rampost_error
 - c. All messages from rampost are prefixed with "@rampost."
2. Next, init performs a signature check of the system partition to ensure integrity. This is triggered by the command line which includes dm-verity options prefixed with "dm=". If the system does not pass checks, init fails and exits.
 - a. Note: none of the file systems in fstab marked for verity checking, so this is the only place where it is performed.
3. The main system file-system is mounted and launches the main system initialization.

The regular boot uses systemd to allow of the startup steps to be performed in parallel. The rough start up sequence is:

1. Starts the Qualcomm Secure Execution Environment Communicator (dev-qseecom.device) and ION memory allocator (dev-ion.device)
2. The encrypted user file system is checked and mounted (via the mount-data service). This file system is where all of the logs, people's faces, and other information specific to the individual Vector are stored. The keys to this file system are stored in the TrustZone in the MPU's SOC fuse area. This file system can only be read by the MPU that created it.
3. The MPU's clock rate is limited to 533Mhz, and the RAM is limited to 400MHz to prevent overheating.
4. The camera power is enabled
5. If Vector doesn't have a robot name, Vic-christen is called to give it one.
6. After that several mid-layer communication stacks are started:
 - a. usb-service any time after that
 - b. the WiFi connection manager (connman)
 - c. The time client (chronyd), to retrieve network time. (Vector does not have a clock that keeps time when turned off)
 - d. init-debuggerd
7. multi-user, sound, init_post_boot
8. The "Victor Boot Animator" is start (~8 seconds after power on). This is probably the sparks.
9. Victor Boot completes ~20.5 after power on, and the post boot services launches
10. vic-crashloader
11. vic-robot
12. Once the startup has sufficiently brought up enough the next set of animations the sound of boot
13. VicOS is running ~32 seconds after power on. The boot is complete, and Vector is ready to play

21.4. ABNORMAL SYSTEM BOOT

If there is a problem during startup – such as one of the services is unable to successfully start, a fault code associated with that service is stored in /run/fault_code and the fault code displayed. See chapter 31 for a description of the display of fault codes and diagnostics. See Appendix C for fault codes.

22. SHUTDOWN

- Turning Vector off manually
- Vector turning off spontaneously due to brown-out or significant loss of power
- Vector turning off (under low power) by direction of the head-board

Vector cannot be turned off via Bluetooth LE, or the local HTTPS SDK access. There are no exposed commands that do this. Using a verbal command, like "turn off" does not direct Vector to

shut off (disconnect the battery). Instead it goes into a quiet mode. Although it may be possible for a Cloud command to turn Vector off, this seems unlikely.

However, there is likely a command to automate the manufacture and preparation for ship process.

22.1. TURNING VECTOR OFF (INTENTIONALLY)

When the system decides it needs to shutdown, it internally posts one of the following codes at the reason for shutdown:

Element	Description & Notes
SHUTDOWN_BATTERY_CRITICAL_TEMP	Vector shut down automatically because the battery temperature was too high.
SHUTDOWN_BATTERY_CRITICAL_VOLT	Vector shut down automatically because the battery voltage was too low.
SHUTDOWN_BUTTON	Vector was shut down by a long button press.
SHUTDOWN_GYRO_NOT_CALIBRATING	Vector shut down automatically because of an IMU problem(?)
SHUTDOWN_UNKNOWN	Vector shut down unexpectedly; the reason is not known. Likely a brown-out or battery voltage dipped low faster than Vector could respond to.

Table 19: Vector shutdown codes

The shutdown code is logged but not otherwise stored.

22.2. UNINTENTIONALLY

The body-board is responsible for keeping the battery connected. However brownouts, self-protects when the voltage get to too low, and bugs can cause the battery to be disconnected.

22.3. GOING INTO AN OFF STATE

When Vector wants to intentionally turn off, it cleans up its state, to gracefully shutdown the linux system and tells the body-board to disconnect the battery.

23. REFERENCES & RESOURCES

Android, *Verified Boot*

<https://source.android.com/security/verifiedboot/>

Bhat, Akshay; *Secure boot on Snapdragon 410*, TimeSys, 2018 Aug 23

<https://www.timesys.com/security/secure-boot-snapdragon-410/>

Discusses how one can get the source to the secondary bootloader (SBL), the tools to sign it and aboot using sectools

<https://gitlab.com/cryptsetup/cryptsetup/wikis/DMVerity>

Hay, Roee. *fastboot oem vuln: Android Bootloader Vulnerabilities in Vendor Customizations*,

Aleph Research, HCL Technologies, 2017

<https://www.usenix.org/system/files/conference/woot17/woot17-paper-hay.pdf>

Hay, Roee; Noam Hadad. *Exploiting Qualcomm EDL Programmers*, 2018 Jan 22

Part 1: Gaining Access & PBL Internals

<https://alephsecurity.com/2018/01/22/qualcomm-edl-1/>

Part 2: Storage-based Attacks & Rooting

<https://alephsecurity.com/2018/01/22/qualcomm-edl-2/>

Part 3: Memory-based Attacks & PBL Extraction

<https://alephsecurity.com/2018/01/22/qualcomm-edl-3/>

Part 4: Runtime Debugger

<https://alephsecurity.com/2018/01/22/qualcomm-edl-4/>

Part 5: Breaking Nokia 6's Secure Boot

<https://alephsecurity.com/2018/01/22/qualcomm-edl-5/>

Johnson, Nolen; *Qualcomm's Chain of Trust*, Lineage OS, 2018 Sept 17

<https://lineageos.org/engineering/Qualcomm-Firmware/>

A good overview of Qualcomm's boot loader, boot process, and differences between versions of Qualcomm's process. Quotes are slightly edited for grammar.

Nakamoto, Ryan; *Secure Boot and Image Authentication*, Qualcomm , 2016 Oct

<https://www.qualcomm.com/media/documents/files/secure-boot-and-image-authentication-technical-overview-v1-0.pdf>

Qualcomm, *DragonBoard™ 410c based on Qualcomm® Snapdragon™ 410E processor Little*

Kernel Boot Loader Overview, LM80-P0436-1, Rev D, 2016 Jul

lm80-p0436-1_little_kernel_boot_loader_overview.pdf

<https://github.com/alephsecurity>

A set repositories researching tools to discover commands in Sahara and EDL protocols

<https://github.com/openpst>

A set of repositories researching and implementing an interface to the Sahara protocol.

CHAPTER 8

Power management

This chapter describes Vector's power management:

- The battery management
- Load shedding
- Charger info

24. POWER MANAGEMENT

24.1. BATTERY MANAGEMENT

There isn't a coulomb counter to track the remaining energy in the battery. At the broadest strokes, the battery voltage is used to predict the battery state of charge.

24.1.1 Battery levels

Vector maps the battery voltage into a battery level, taking into account whether or not the charger is active:

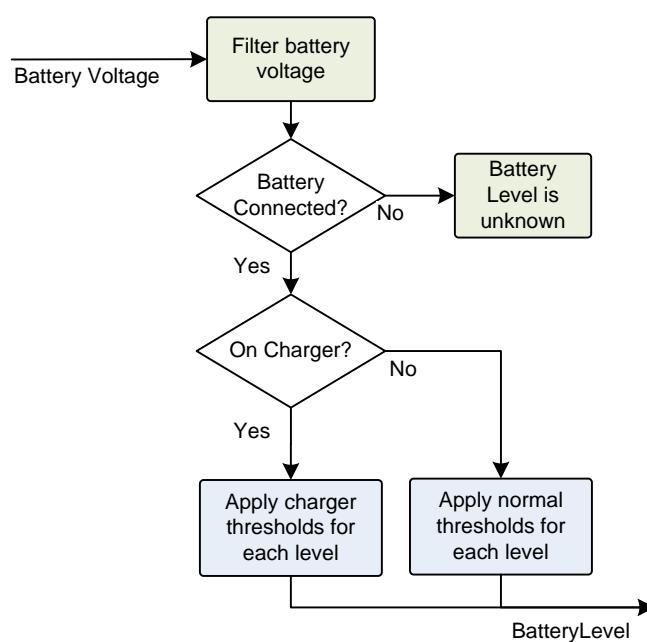


Figure 32: The battery level classification tree

Note: The battery voltage is filtered as the voltage will bounce around with activity by the motors, driving the speaker and processors.

The `BatteryLevel` enumeration is used to categorize the condition of the Vector battery:

Name	Value	Description
<code>BATTERY_LEVEL_FULL</code>	3	Vector's battery is at least 4.1V
<code>BATTERY_LEVEL_LOW</code>	1	Vector's battery is 3.6V or less; or if Vector is on the charger, the battery voltage is 4V or less.
<code>BATTERY_LEVEL_NOMINAL</code>	2	Vector's battery level is between low and full.
<code>BATTERY_LEVEL_UNKNOWN</code>	0	If the battery is not connected, Vector can't measure its battery.

Table 20:
`BatteryLevel` codes²¹
as they apply to
Vector

The battery levels are organized conventionally:

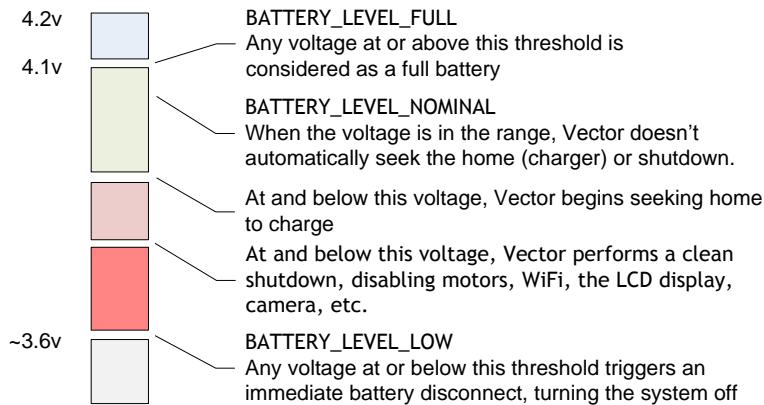


Figure 33: The battery thresholds

The current battery level and voltage can be requested with the Battery State command (see Chapter 13, section 47.2 *Battery State*). The response will provide the current battery voltage, and interpreted level.

24.1.2

A software “fuel gauge”

It is typical for larger battery packs to include a coulomb counter, often called a fuel gauge. They include it for a serious reason: it prevents fire and explosions that can result from overcharging a large multi-cell pack. The fancy “fuel gauge” and estimated useful life is a bonus.

For Vector, a fuel gauge would give him smarts about knowing he will need to plan to return home, or is getting low. His hardware doesn't have a coulomb counter, for a variety of reasons. Any effort, beyond simple battery voltage, to estimate the remaining play time would have to be based on software and tracking the battery performance.

24.2. RESPONSES, SHEDDING LOAD / POWER SAVING EFFORTS

Vector's main (power-related) activity modes are:

- active, interacting with others
- calm, where primarily sitting still, waiting for assistance or stimulation

²¹ The levels are from robot.py

- sleeping

Depending on the state of the battery – and charging – Vector may engage in behaviours that override others.

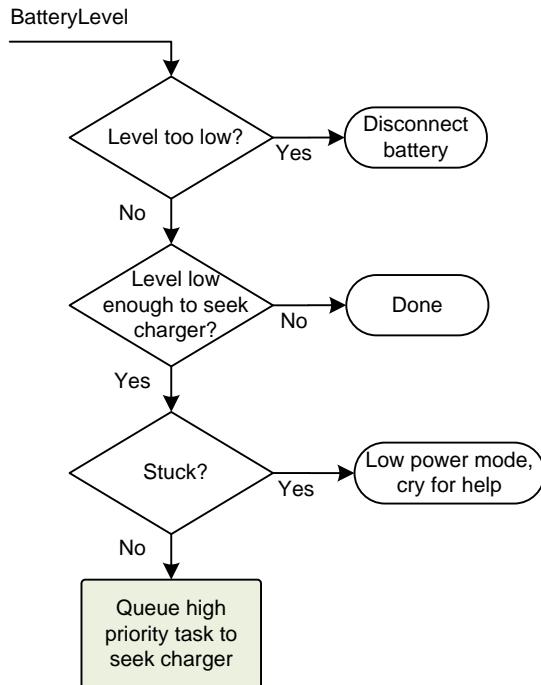


Figure 34: The response to battery level

If Vector is unable to dock (or even locate a dock) he sheds load as he goes into a lower state:

- He no longer responds to his trigger word or communicates with WiFi servers
- He turns off camera and LCD; presumably the time of flight sensor as well.
- He reduces processing on the processor

24.2.1

Temperature limits and related processing

The software tracks the temperature of the battery and head. As the temperature rises, more aggressive actions are taken to protect the battery and let the chips cool down.

- Around 90C, Vector displays the overheating icon.
- If the body board is overheated, a flag in the HTTPS API RobotStatus bit mask is set (see Chapter 13, section 40.1.2 *RobotStatus Note: this is speculated, not proven*.
- At some point past 90C, Vector starts a clean shut down (see earlier). The software in the head is idle, and turns off as many peripherals (e.g. WiFi, display, etc.) with “the goal to save enough power in the head to let the chip cool off, so we could continue driving home.”
- If the APQ8009 processor is hot, it will throttle its clocks. If the MP2617B charging chip is reaching the thermal limits related to charging, it will throttle the charging.
- If either the body or head board exceeds a maximum temperature, the system is completely shut down, and power is cut.

The battery overheated icon is displayed by `vic-faultCodeDisplay`, which has a hard coded path to the icon:

/anki/data/assets/cozmo_resources/config/devOnlySprites/independentSprites/battery_overheated.png

Version 1.6 uses very conservative thresholds (to protect the battery) with the intention of follow up releases fine tuning the thresholds.

24.2.2

Calm Power mode

Has a high-level power mode called “calm power mode.” This mode “is generally when Vector is sleeping or charging.” Vector [probably] turns off the sensors, lowers the CPU and camera clock rate – or may even suspend the camera. (See Chapter 17, section 76.5 *Illumination level sensing* for a description).

Vector being in calm power mode is reported in the `RobotStatus` message in the `status` field. (See Chapter 6 for details.) Vector is in a calm power model if the `ROBOT_STATUS_CALM_POWER_MODE` bit is set (in the `status` value).

24.2.3

When not moving

When Vector isn’t driving (or using his head and lift), he puts his motors and related sensors into a low power state:

- The encoders are mostly turned off; they “pulsed at 1% duty cycle and watched for changes” to detect someone moving Vector around;
- The time of flight sensor is turned to a lower sampling period

24.3. ACTIVITY LEVEL MANAGEMENT

Version 1.5 slowed down a lot of Vectors activities (by lowering his max clock rate), to reduce heat (prolonging the battery life) and allow him to play longer between charge cycles. Some of his behaviors were modified so that he doesn’t initiate exploring and playing as much, choosing instead to stay on the charger longer until there was more signs that people were around to play.

Version 1.6 may have gone further

25.

CHARGING

Vector tracks whether is charging is in process, and how long. The software has some initial estimates how long before charging is complete. This is similar to the software “fuel gauge.” It takes some model of the batteries capacity, and typical charging times given that.

The state of the charger is reported in the `RobotStatus` message in the `status` field. (See chapter 13 for details.) Vector is on the charger if the `ROBOT_STATUS_IS_ON_CHARGER` bit is set (in the `status` value), and charging if the `ROBOT_STATUS_IS_CHARGING` bit is set.

Version 1.5 slowed down the charging, to reduce heat, prolonging the battery life.

Additional information about the state of the charger can be requested with the `Battery State` command (see Chapter 13, section 47.2 *Battery State*). The response will provide flags indicating whether or not Vector is on the charger, and if it is charging. The response also provides a suggested amount of time to charge the batteries.

CHAPTER 9

Basic Inputs and Outputs

This chapter describes Vector's most basic input and output: his button, touch and LEDs:

- Touch and button input
- Backpack Lights control

Note: the audio sampling will be covered in a later chapter (Chapter 16)

26. BUTTON, TOUCH AND CLIFF SENSOR INPUT

Vector's backpack button used to wake (and silence) Vector, or place him into recovery mode. Touch is used to pet Vector and provide him stimulation. Four surface proximity IR sensors are used to detect cliffs and line edges. The responsibility for the button, touch, and proximity sensor input functions are divided across multiple processes and boards in Vector:

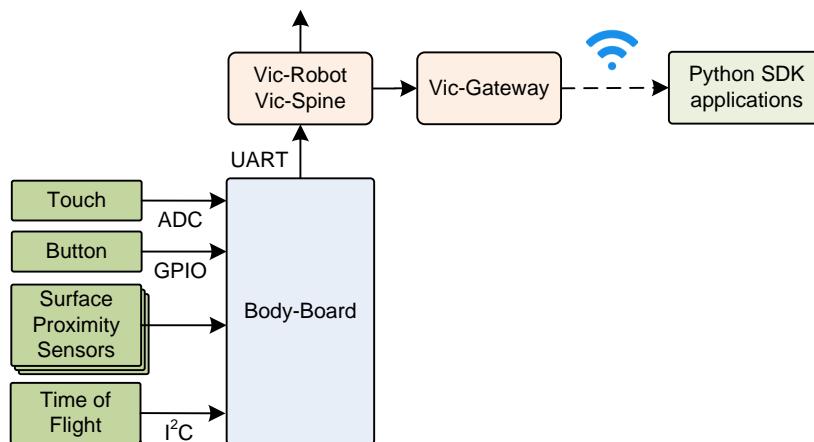


Figure 35: The touch and button input architecture

The state of the inputs (button, touch, surface proximity and time of flight sensors) are reported in the `RobotStatus` message. (See chapter 13 for details.) The button state can be found in the `status` field. The button is pressed if the `ROBOT_STATUS_IS_BUTTON_PRESSED` bit is set (in the `status` value).

The surface proximity sensors (aka “cliff sensors”) are used to determine if there is a cliff, or potentially in the air. The individual sensor values are not accessible. The cliff detection state can be found in the `status` field. A cliff is presently detected if the `ROBOT_STATUS_CLIFF_DETECTED` bit is set (in the `status` value).

26.1. TOUCH SENSING INFORMATION

The touch sensor is driven by the body-board, and the sample values are processed in the head-board. The sensors samples are filtered, to get a sense of the current “level” the sensor is at. A standard deviation is used as a measure of how solid the signal, to help distinguish between a real signal and ambient conditions like humidity and weather. These two measures – along with a timer to screen out transitory noise – can be used to decide that Vector is being touched or not.



Figure 36: The touch sensor and petting detector

These measures could potentially distinguish between light touch (e.g. tip of the finger), heavy touch (e.g. a full palm?), and perhaps even changing touch.

The touch sensor readings can be found in the `touch_data` field of the `RobotStatus` message. The values indicate whether Vector is being touched (e.g. petted).

The touch sensor module produces a JSON structure for internal use:

Field	Type	Units	Description
<code>max</code>	float		The maximum value seen
<code>min</code>	float		The minimum value seen
<code>stddev</code>	float		The standard deviation

Table 21: Touch sensor structure

26.2. TIME OF FLIGHT PROXIMITY SENSOR

The time of flight reading is given in the `prox_data` field. This indicates whether there is a valid measurement, the distance to the object, and a metric that indicates how good the distance measurement is. This will be processed by the mapping system. See *Chapter 18 section 86 Measuring the distance to objects*.

27. BACKPACK LIGHTS CONTROL

The backpack lights are used to show the state of the microphone, charging, WiFi and some other behaviours. (It is also used to show unusual error states.)

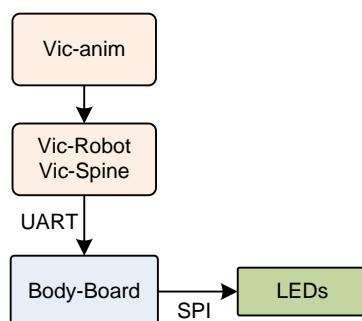


Figure 37: The backpack lights output architecture

The software can direct the body-board to illuminate the backpack lights with individually different colors and brightness's. The body-board “pulse width modulates” (PWM's) the LEDs to achieve different colors and intensities.

The body-board doesn't directly interface with the LED's (they're connected to a logic chip on a separate board), so it cannot delegate the work to an internal PWM peripheral. The body-board must implement its PWM in firmware, and send the GPIO states to the backpack every time there is a change. (See Chapter 4, section *10.3.1 The LED controls*)

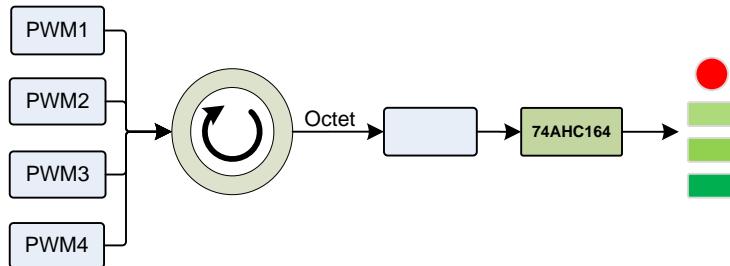


Figure 38: The firmware driving of the LEDs

The basic logic to drive the LEDs is:

1. Select LEDs for the time slice
2. Get the LED bit settings from the PWM(s)
3. Organize these into a format for the 74AHC164
4. Send the bits to the 74AHC164
5. Delay until the next time slice, and repeat

CHAPTER 10

Inertial Motion Sensing

This chapter describes Vector's motion sensing:

- Sensing motion and cliffs
- Detecting external events
- Measuring motion as feedback to motion control, and allow moving along paths in a smooth and controlled fashion

28. MOTION SENSING

Vector employs an IMU – an accelerometer and gyroscope in the same module – to detect motion, such as falling or being bumped, as well as measuring the results of motor-driven motions.

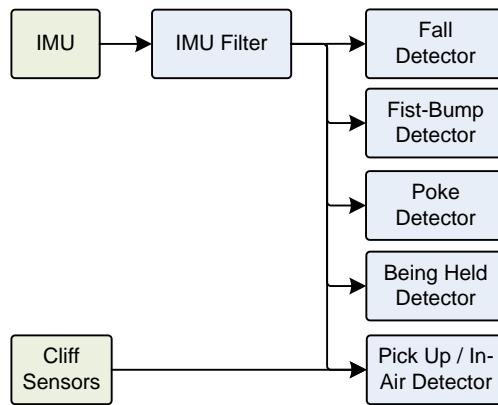


Figure 39: Sensing motion events

28.1. ACCELEROMETER AND GYROSCOPE

Neither the accelerometer nor a gyroscope by itself is sufficient to accurately measure change in position and orientation. Accelerometers measure force along 3 (XYZ) axes, including gravity. The accelerometer provides the orientation – if there is no other motion. The drawback is that accelerometers cannot correctly measure spins, and other rotations from other movements. Gyroscopes can measure rotations around the axes, but cannot measure linear motion along the axis. Gyroscopes also have a slight bias in the signal that they measure, giving the false signal that there is always some motion occurring.

By blending the accelerometer and gyroscope signals together, they can compensate and cancel each other's weaknesses out.

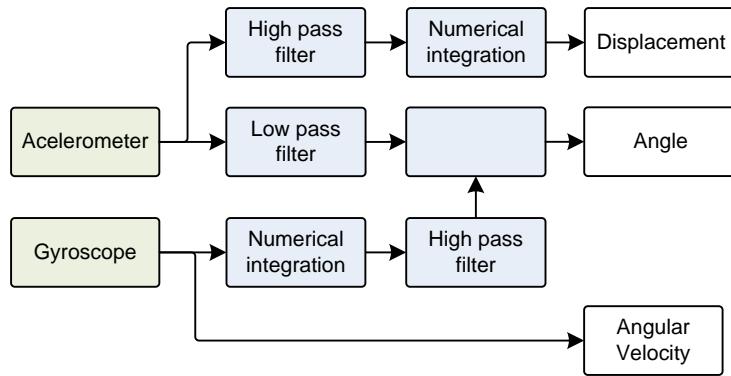


Figure 40:
Complementary filtering of the
accelerometer and
gyroscope

28.2. TILTED HEAD

The IMU can also measure how tilted Vector's head is. The IMU is located in Vector's head. This presents a small extra step of processing for the software to accommodate the impact of the head.

By combining the position & orientation of the IMU within the head, the current estimated angle of the head, the known joint that the head swivels on, and working backwards the IMU measures can be translated to body-centered measures.

28.3. SENSING MOTION

The IMU, located in Vector's head, can also measure how tilted Vector's head is. This presents a small extra step of processing for the software to accommodate the influence of the head tilt. By combining the position & orientation of the IMU within the head, the current estimated angle of the head, the known joint that the head swivels on, and working backwards the IMU measures can be translated to body-centered measures.

28.4. SENSING INTERACTIONS

The IMU (with some help from the cliff sensors) is also used to sense interactions and other environmental events – such as being picked up or held by a person, being poked or given a fist bump, or falling.

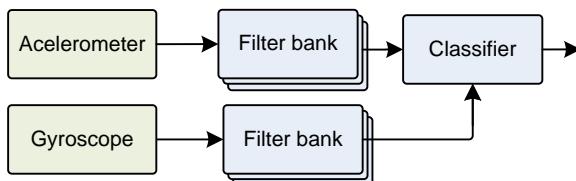


Figure 41: Classify
movement by filtering
the accelerometer and
gyroscope signals

By using combinations of high, low pass, and band filters, and looking for signature patterns, Vector identify the kinds of physical interactions that are occurring.

The taps and pokes may tilt Vector, but will also provide a “frequency” response to the signals that can be used to trigger on. The movement will change his position quickly and slight in small distance, but Vector will resume his prior position very quickly.

Fist-bumps are like pokes, except that the lift has already been raised, and most of the frequency response and motion will be predictable from receiving the bump on the lift.



Fall detector is similar. In free-fall, the force measured by the accelerometer gets very small. If Vector is tumbling, there is a lot of angular velocity that is taking Vector off of his driving surface.

Being picked up is distinct because of the direction of acceleration and previous orientation of Vector's body.

Being held is sensed, in part by first being picked up, and by motions that indicate it is not on a solid surface.

Patent filings (e.g. WO 2019/173321 indicates that Anki had ideas of how this could be extended to detect riding in a car, and even estimating how fast it is moving.



29. REFERENCES AND RESOURCES

AdaFruit, https://github.com/adafruit/Adafruit_9DOF/blob/master/Adafruit_9DOF.cpp
An example of how accelerometer and gyroscope measurements are fused.

Anderson, Ross *Robot Transportation Mode Classification*, Anki, WIPO WO 2019/173321 A1,
2019 Sept 12

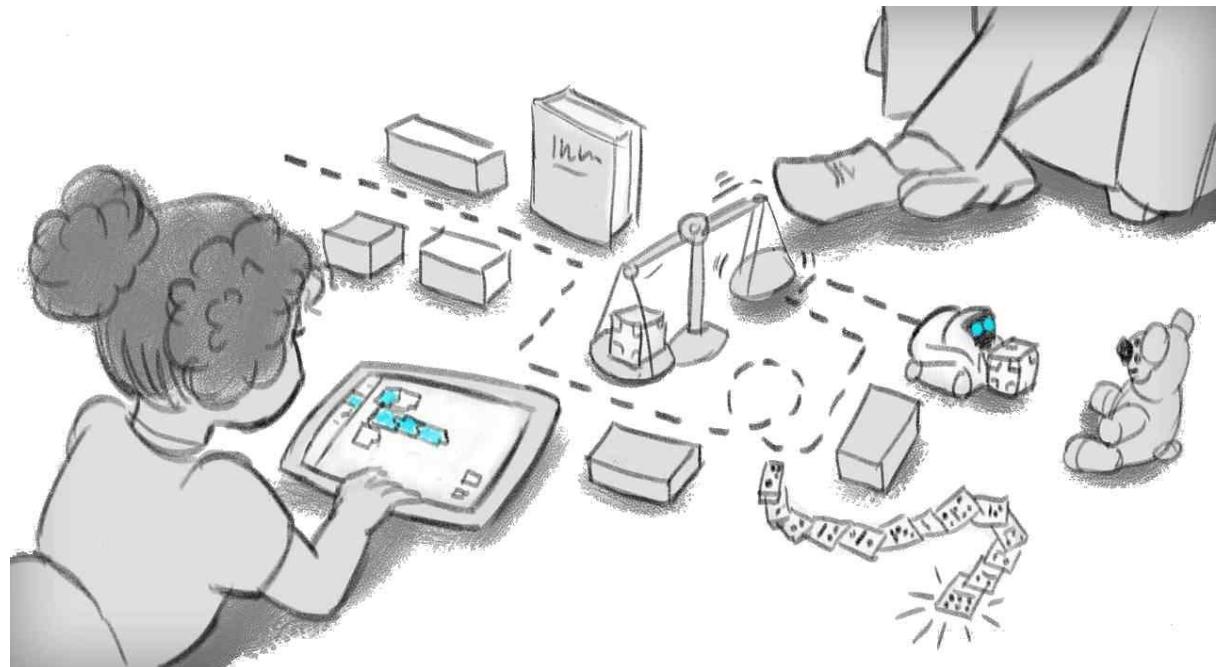
[This page is intentionally left blank for purposes of double-sided printing]

PART III

Communication

This part provides details of Vector's communication protocols. These chapters describe structure communication, the information that is exchanged, its encoding, and the sequences needed to accomplish tasks. Other chapters will delve into the functional design that the communication provides interface to.

- COMMUNICATION. A look at Vector's communication stack.
- BLUETOOTH LE. The Bluetooth LE protocol that Vector responds to.
- SDK PROTOCOL. The HTTPS protocol that Vector responds to.
- WEB-VISUALIZATION PROTOCOL. The web-sockets protocol(s) that Vector provides for debugging in development builds.
- CLOUD. A look at how Vector interacts with remote services



drawing by Jesse Easley

[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 11

Communication

This chapter describes the communication system:

- Internal communication with the body-board, and internal peripherals
- Bluetooth LE: with the Cube, and with the application
- WiFi: with the cloud, and with the application
- Internal support

30. OVERVIEW OF VECTOR'S COMMUNICATION INFRASTRUCTURE

A significant part of Vector's software is focused on communication.

- Internal IPC between processes
- Communication with local peripherals and the body-board processor
- Communication with external accessories and applications.

The communication stacks look something like:

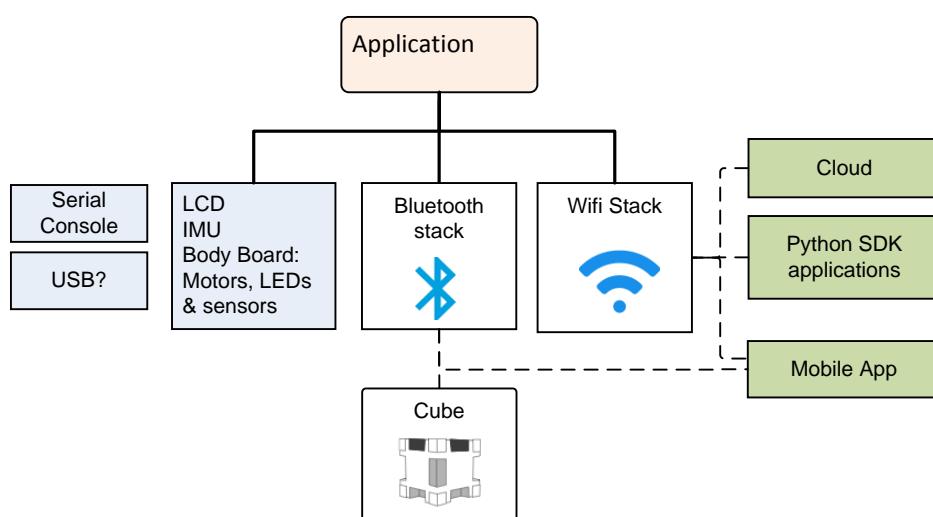


Figure 42: The overall communication infrastructure

31. INTERNAL COMMUNICATION WITH PERIPHERALS

Communication stack within the software. One part Linux, one part Qualcomm, and a big heaping dose of Anki's stuff.

31.1. COMMUNICATION WITH THE BODY-BOARD

The head board communicates with the body board using a serial interface. The device file is `/dev/ttyHS0`.

Data rate: 3 Mbits/sec²²

31.1.1 Messages from the body-board to the head board

The body-board sends packets at regular intervals to the head-board. The frame of the message in

[Unknown byte] AA₁₆ ‘B’ ‘2’ ‘H’ [16-bit packet type] [16-bit payload size] [payload bytes] [32-bit CRC]

(All multi-byte values are in little endian order.) The maximum packet size is 1280 bytes.

The packet type implies both the size of the payload, and the contents. If the packet type is not recognized, or the implied size does not match the passed payload size, the packet is considered in error.) The table below gives the different type codes:

Packet type	Payload Size	Description
6473 ₁₆	0	
6b61 ₁₆	4	
6466 ₁₆	768	The size of the message suggests that it holds 128 samples from one to three microphones (3 microphones × 2bytes/sample × 128 samples/microphone == 768 bytes) for the voice activity detection audio processing.
6662 ₁₆	4	
6675 ₁₆	1028	The size of the message suggests that it holds 128 samples from the microphones (4 microphones × 2bytes/sample × 128 samples/microphone == 1024 bytes) for the spatial audio processing.
736c ₁₆	16	
6d64 ₁₆	0	
7276 ₁₆	40	
7374 ₁₆	0	
7878 ₁₆	0	

The payload can contain (depending on the type of packet):

- The state of the backpack button
- The touch sensor voltage
- The microphone samples for all 4 microphones. (Most likely as 16 bits)
- The battery voltage

²² Value from analysis of the RAMPOST and vic-switchboard programs.

- The charging terminal voltages
- State of the charger (on dock/etc)
- The temperature of the charger
- The state of 4 motor encoders, possibly as encoder counters, possibly as IO state
- The time of flight reading, probably 16bits in mm
- The values from each of the 4 cliff proximity sensors
- Which peripherals are enabled and disabled (powered down)

The messages are sent fast enough to support microphone sample rate of 15625 samples/second for each of the 4 microphones.

31.1.2

Messages from the head-board to the body-board

The messages from the head board to the body-board have the content:

- The 4 LED RGB states
- Controls for the motors: possible direction and enable; direction and duty cycle; or a target position and speed.
- Power control information: disable power to the system, turn off distance, cliff sensors, etc.

The head-board can update the firmware in the body-board, by putting into DFU (device firmware upgrade) mode and downloading the replacement firmware image.

TIMEOUTS. The body-board maintains a timer to detect the loss of communication from the head-board – perhaps from a software crash. If the body-board does not receive communication within this timeout period, it will turn off power.

31.2. SERIAL BOOT CONSOLE

The head-board employs a serial port to display kernel boot up and log messages. The parameters are 115200 bits/sec, 8 data bits no parity, 1 stop bit; the device file is `/dev/ttyHSLO`. This serial port is not bi-directional, and can not be used to login.

Melanie T

31.3. USB

There are pins for USB on the head board. Asserting “F_USB” pad to VCC enables the port. During power-on, and initial boot it is a Qualcomm QDL port. The USB supports a Qualcomm debugging driver (QDL), but the readout is locked. It appears to be intended to inject software during manufacture.

Melanie T

The `/etc/initscriptsusb` file enables the USB and the usual functions adb. It lives in `/sbin/usr/composition/9091` (I think, if I understand the part number matching correctly). This launches ADB (DIAG + MODEM + QMI_RMNET + ADB)

Vectors log shows the USB being disabled 24 seconds after linux starts.

Melanie T

32. BLUETOOTH LE

Bluetooth LE is used for two purposes:

1. Bluetooth LE is used to initially configure Vector, to reconfigure him when the WiFi changes; and to pair him to with the companion cube accessory. Potentially allows some diagnostic and customization.

- Bluetooth LE is used to communicate with the companion Cube: to detect its movement, taps, and to set the state of its LEDs.

Vector's Bluetooth LE stack looks like:

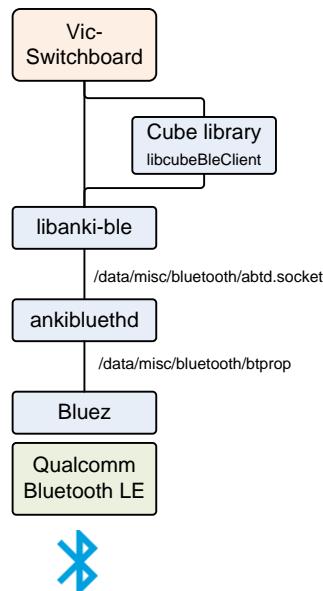


Figure 43: The Bluetooth LE stack

The elements of the Bluetooth LE stack include:

Element	Description & Notes
ankibluehd	A server daemon. The application layer communicates with it over a socket: /data/misc/bluetooth/abtd.socket
BlueZ	Linux's official Bluetooth stack, including Bluetooth LE support. The Anki Bluetooth daemon interacts with it over a socket: /data/misc/bluetooth/btprop
bccmd	A Bluetooth core command
btmon	A command-line Bluetooth tool
libanki-ble.so	Communicates with Anki Bluetooth daemon probably serves both the external mobile application interface and communication with the companion cube.
libcubeBleClient.so ²³	A library to communicate with the companion cube, play animations on its LEDs, detect taps and the orientation of the cube.
viccubetool	Probably used to update the firmware in the Cube.

Table 23: Elements of the Bluetooth LE stack

²³ The library includes a great deal of built in knowledge of the state of application (“game engine”), animations, and other elements

33. WIFI

WiFi networking is used by Vector for five purposes:

1. WiFi is used to provide the access to the remote servers for Vector's speech recognition, natural language processing
2. WiFi is used to provide the access to the remote servers for software updates, and providing diagnostic logging and troubleshooting information to Anki
3. To provide time services so that Vector knows the current time
4. To provide an interface, on the local network, that the mobile application can use to configure Vector, and change his settings.
5. To provide an interface, on the local network, that SDK applications can use to program Vector.
6. To provide interfaces, on the local network, that allow development Vectors (special internal versions) to be debugged and characterized

The WiFi network stack looks like:

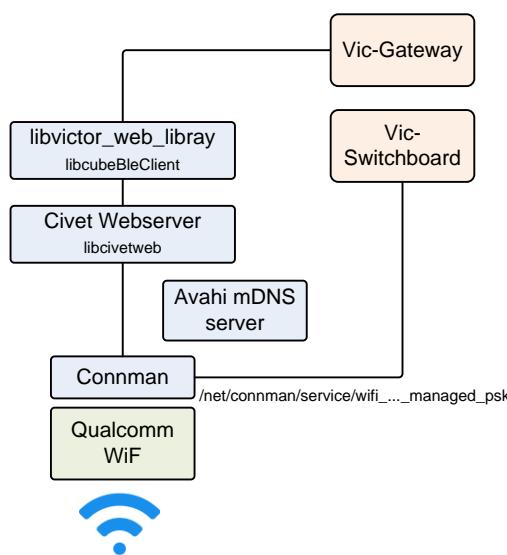


Figure 44: The WiFi stack

The elements of the stack include:²⁴

Element	Description & Notes
avahi 0.6.31	A mDNS server that registers Vector's robot name (with his network address) on the local network;
chronyd	Fetches the time from a network time server.
libcivetweb.so.1.9.1	Embedded web server
libvictor_web_library.so	Anki Vector Web Services.

Table 24: Elements of the Bluetooth LE stack

²⁴ All of the software versions include an Anki webserver service systemd configuration file whose executable is missing. The most likely explanation is that early architecture (and possibly early versions) included this separate server, and that the systemd configuration file is an unnoticed remnant.

33.1. FIREWALL

The network configuration includes a firewall set up with the usual configuration files:

`/etc/iptables/iptables.rules`
`/etc/iptables/iptables`
`/etc/iptables/ipv6tables.rules`
`/etc/iptables/ipv6tables`

Is set to block incoming traffic (but not internal traffic), except for:

1. Responses to outgoing traffic
2. DHCP
3. TCP port 443 for vic-gateway
4. UDP port 5353 for mDNS (Avahi)
5. And the ping ICMP

In developer builds the firewall also allows:

1. SSH access
2. Android Debugger (ADB) over TCP access
3. “Web-viz” access, which has web-server / websockets / webdav ports
4. Webots support
5. WWise profiler support

The firewall does not block outgoing traffic

33.2. WIFI CONFIGURATION

The WiFi is configured by the Vic-switchboard over Bluetooth LE. The WiFi settings cannot be changed by the remote servers or thru the WiFi-based API; nor is information about the WiFi settings is not stored remotely.

The WiFi is managed by connman thru the Vic-Switchbox:

- To provide a list of WiFi SSIDs to the mobile app
- To allow the mobile app to select an SSID and provide a password to
- Tell it forget an SSID
- To place the WiFi into Access Point mode

34. NETWORK COMMUNICATION

34.1. COMMUNICATING WITH MOBILE APP AND SDK

Vector’s *robot name* is something that looks like “Vector-E5S6”. This name is used consistently; it will be Vector’s:

- advertised Bluetooth LE peripheral name (although spaces are used instead of dashes)
- mDNS network name (dashes are used instead of spaces),
- the name used to sign certificates, and
- it will be the name of his WiFi Access Point, when placed into Access Point mode

34.1.1

Certificate based authentication

A *session token* is always provided by Anki servers.²⁵ It is passed to Vector to authenticate with him and create a client token. The session token is passed to Vector via the Bluetooth LE RTS protocol or the HTTPS-based SDK protocol; Vector will return a client token. The session token is single use only.

A *client token* is passed to Vector in each of the HTTPS-based SDK commands, and in the Bluetooth LE SDK Proxy commands. It is generated in one of two ways. One method is by the Bluetooth LE command (cloud session); the other is by send a User Authentication command (see Chapter 13 section 48.5 *User Authentication*). The client token should be saved indefinitely for future use. It is not clear if the client token can be shared between the two transport mechanisms.

A *certificate* is also generated by Vector in the case of the SDK request. The certificate is intended to be added to the trusted SSL certificates before an HTTPS communication session. The certificate issued by Vector is good for 100 years.

Note: the certificates are invalidated and new ones are created when recovery-mode is used. Vector is assigned a new robot name as well.

The typical information embedded in a Vector certificate:

Element	Value
Common Name	<i>Vector's robot name</i>
Subject Alternative Names	<i>Vector's robot name</i>
Organization	Anki
Locality	SF
State	California
Country	US
Valid From	<i>the date the certificate was created</i>
Valid To	<i>100 years after the date the certificate was created</i>
Issuer	<i>Vector's robot name, Anki</i>
Serial Number	

Table 25: Elements of a Vector certificate

The TLS certificates and signing keys are stored in the OEM partition, in the “cloud” folder:

File	Description
<i>AnkiRobotDeviceCert.pem</i>	The
<i>AnkiRobotDeviceKeys.pem</i>	The
<i>Info\${serialNum}.json</i>	A configuration file that
<i> \${serialNum}</i>	empty

Table 26: OEM cloud folder

²⁵ <https://groups.google.com/forum/#!msg/anki-vector-rooting/YIYQsX08OD4/fvkAOZ91CgAJ>
<https://groups.google.com/forum/#!msg/anki-vector-rooting/XAaBE6e94ek/OdES50PaBQAJ>

The Info\${serialNum}.json file has the following structure:

Field	Type	Description
<i>CertDigest</i>	base64 string	
<i>CertSignature</i>	base64 string	
<i>CertSignatureAlgorithm</i>	string	The name of openSSL signature algorithm to use, “sha256WithRSAEncryption”
<i>CommonName</i>	string	‘vic:’ followed by the serial number. (This is also called the “thing id” in other structures.
<i>KeysDigest</i>	base64 string	

Table 27: Cloud Info\${serialNum} structure

34.2. WEB-VIZ, A VISUAL CHARACTERIZATION TOOL

Development builds of Vectors software include an optional web-sockets API and web-visualization (webviz) tool. This feature is not present in the production releases, nor many of the development releases. This tool has some of the vic-server processes provide an HTTP web-server, and web socket over it:

Port	Description
8887	The webserver built into vic-webserver
8888	The webserver built into vic-engine
8889	The webserver built into vic-anim
8890	Not used

Table 28: Web-viz HTTP & web-socket server ports

The web-sockets provide access to internal variables and other software state. In some cases provide points of control. The web-server, esp thru the webdav support, allows files to be downloaded and uploaded into Vector. This includes the ability to add animation files that can be tested.

Note: the tool is rumoured to be consume a lot of resources, causing unusual faults to occur on Vector.

35. CLOUD SERVERS

The cloud servers are used for natural language processing, storing settings, tracking diagnostic information, and software updates.

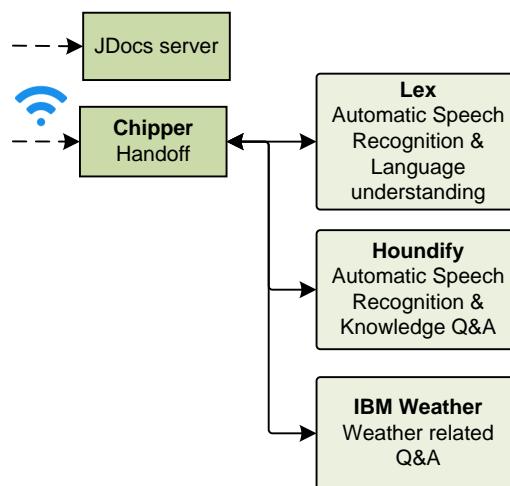


Figure 45: The cloud servers

For natural language processing, the audio stream (after the “Hey Vector”) sent to a group of remote servers for processing. The functions are divided up across several different servers which can provide specialized services:

Server	Description
<i>Chipper</i>	Chipper is a server that that hands off the audio processing.
<i>Houndify</i>	The “knowledge graph” Q&A server is handled by Sound Hound (Houndify). Note: the speech is sent to Houndify only if Lex is unable to handle the query.
<i>IBM Weather</i>	IBM handles the Weather related questions.
<i>Lex</i>	Lex handles most of Vectors speech recognition, natural language understanding, return an intent. (This is discussed a bit more in Chapter 16) The “I have a question” queries are handed off to Houndify. This server is hosted by AWS.

Table 29: Natural language processing servers

Chapter 15 describes the communication with these servers, including the responses that they send back.

Chapter 16 describes typical natural language processing, and the processing of intents.

36. REFERENCES & RESOURCES

PyCozmo.

<https://github.com/zayfod/pycozmo/blob/master/docs/protocol.md>
https://github.com/zayfod/pycozmo/blob/master/pycozmo/protocol_declaration.py

Vector has a couple UDP ports open internally; likely this is inherited from libcozmo_engine. The PyCozmo project has reverse engineered much of Cozmo’s UDP protocol.

CHAPTER 12

Bluetooth LE

Communication

Protocol

This chapter describes Vector's Bluetooth LE communication protocol.

- The kinds of activities that can be done thru communication channels
- The interaction sequences
- The communication protocol stack, including encryption, fragmentation and reassembly.

Note: communication with the Cube is simple reading and writing a characteristic, and covered in Appendix F.

37. COMMUNICATION PROTOCOL OVERVIEW

Vector advertises services on Bluetooth LE, with the Bluetooth LE peripheral name the same as his robot name (i.e. something that looks like "Vector-E5S6".)

Communication with Vector, once established, is structure as a request-response protocol. The request and responses are referred to as "C-Like Abstract Data structures" (CLAD) which are fields and values in a defined format, and interpretation. Several of these messages are used to maintain the link, setting up an encryption over the channel.

The application layer messages may be arbitrarily large. To support Bluetooth LE 4.1 (the version in Vector, and many mobile devices) the CLAD message must be broken up into small chunks to be sent, and then reassembled on receipt.

Combined with application-level encryption, the communication stack looks like:

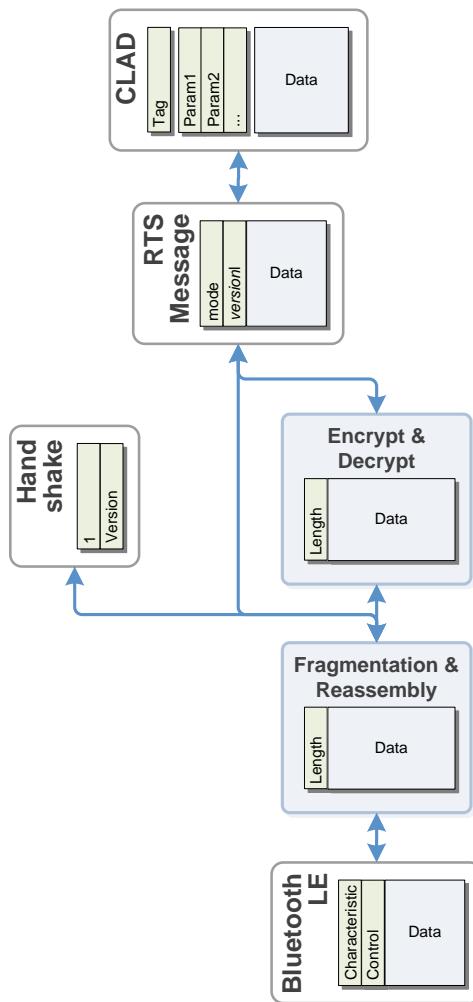


Figure 46: Overview of encryption and fragmentation stack

THE BLUETOOTH LE is the link/transport media. It handles the delivery, and low-level error detection of exchanging message frames. The frames are fragments of the overall message. The GUID's for the services and characteristics can be found in Appendix F.

THE FRAGMENTATION & REASSEMBLY is responsible for breaking up a message into multiple frames and reassembling them into a message.

THE ENCRYPTION & DECRYPTION LAYER is used to encrypt and decrypt the messages, after the communication channel has been set up.

THE RTS is extra framing information that identifies the kind of CLAD message, and the version of its format. The format changed with version, so this version code is embedded at this layer.

THE C-LIKE ABSTRACT DATA (CLAD) is the layer that decodes the messages into values for fields, and interprets them,

37.1. SETTING UP THE COMMUNICATION CHANNEL

It sometimes helps to start with the overall process. This section will walk thru the process, referring to later sections where detailed information resides.

If you use “first time” – or wish to re-pair with him – put him on the charger and press the backpack button twice quickly. He’ll display a screen indicating he is getting ready to pair.

If you have already paired the application with Vector, the encryption keys can be reused.

The process to set up a Bluetooth LE communication with Vector is complex. The sequence has many steps:

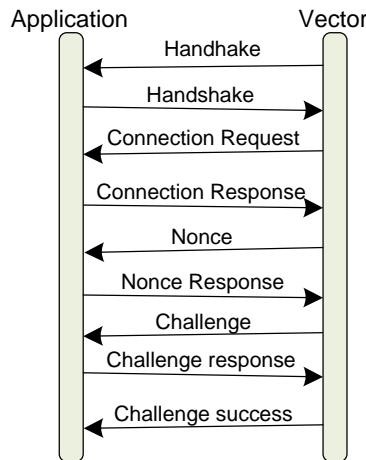


Figure 47: Sequence for initiating communication with Vector

1. The application opens Bluetooth LE connection (retrieving the service and characteristics handles) and subscribes to the “read” characteristic (see Appendix F for the UUID).
2. Vector sends *handshake* message; which the application receives. The handshake message structure is given below. The handshake message includes the version of the protocol supported.

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>type</i>	?
1	4	uint32_t	<i>version</i>	The version of the protocol/messages to employ

Table 30: Parameters for Handshake message

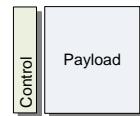
3. The application sends the handshake back
4. Then the Vector will send a *connection request*, consisting of the public key to use for the session. The application’s response depends on whether this is a first-time pairing, or a reuse.
 - a. First time pairing requires that Vector have already been placed into pairing mode prior to connecting to Vector. The application keys should be created (see section 37.3.1 *First time pairing* above).
 - b. Reconnection can reuse the public and secret keys, and the encryption and decryption keys from a prior pairing
5. The application should then send the *publicKey* in the response

6. If this is a first-time pairing, Vector will display a *pin code*. This is used to create the public and secret keys, and the encryption and decryption keys (see section 37.3.1 *First time pairing* above). These can be saved for use in future reconnection.
7. Vector will send a *nonce* message. After the application has sent its response, the channel will now be encrypted.
8. Vector will send a *challenge* message. The application should increment the passed value and send it back as a challenge message.
9. Vector will send a *challenge success* message.
10. The application can now send other commands

If the user puts Vector on the charger, and double clicks the backpack button, Vector will usually send a *disconnect* request.

37.2. FRAGMENTATION AND REASSEMBLY

An individual frame sent over Bluetooth LE is limited to 20 bytes. (This preserves compatibility with Bluetooth LE 4.1) A frame looks like:



The control byte is used to tell the receiver how to *reassemble* the message using this frame.

- If the MSB bit (bit 7) is set, this is the start of a new message. The previous message should be discarded.
- If the 2nd MSB (bit 6) is set, this is the end of the message; there are no more frames.
- The 6 LSB bits (bits 0..5) are the number of payload bytes in the frame to use.

The receiver would append the payload onto the end of the message buffer. If there are no more frames to be received it will pass the buffer (and size count) on to the next stage. If encryption has been set up, the message buffer will be decrypted and then passed to the RTS and CLAD. If encryption has not been set up, it is passed directly to the RTS & CLAD.

Fragmenting reverses the process:

1. Set the MSB bit of the control byte, since this is the start of a message.
2. Copy up to 19 bytes to the payload.
3. Set the number of bytes in the 6 LSB bits of the control byte
4. If there are no more bytes remaining, set the 2nd MSB bit of the control byte.
5. Send the frame to Vector
6. If there are bytes remaining, repeat from step 2.

37.3. ENCRYPTION SUPPORT

For the security layer, you will need the following:

```
uint8_t Vectors_publicKey[32];
uint8_t publicKey [crypto_kx_PUBLICKEYBYTES];
uint8_t secretKey [crypto_kx_SECRETKEYBYTES];
uint8_t encryptionKey[crypto_kx_SESSIONKEYBYTES];
uint8_t decryptionKey[crypto_kx_SESSIONKEYBYTES];
uint8_t encryptionNonce[24];
uint8_t decryptionNonce[24];
uint8_t pinCode[16];
```

Example 1: Bluetooth LE encryption structures

The variables mean:

Variable	Description	Table 31: The encryption variables
<i>decryptionKey</i>	The key used to decrypt each message from to Vector.	
<i>decryptionNonce</i>	An extra bit that is added to each message. The initial nonce's to use are provided by Vector.	
<i>encryptionKey</i>	The key used to encrypt each message sent to Vector.	
<i>encryptionNonce</i>	An extra bit that is added to each message as it is encrypted. The initial nonce's to use are provided by Vector.	
<i>pinCode</i>	6 digits that are displayed by Vector during an initial pairing.	
<i>Vectors_publicKey</i>	The public key provided by Vector, used to create the encryption and decryption keys.	

There are two different paths to setting up the encryption keys:

- First time pairing, and
- Reconnection

37.3.1 First time pairing

First time pairing requires that Vector be placed into pairing mode prior to the start of communication. This is done by placing Vector on the charger, and quickly double clicking the backpack button.

The application should generate its own internal *public* and *secret* keys at start.

```
crypto_kx_keypair(publicKey, secretKey);
```

Example 2: Bluetooth LE key pair

The application will send a *connection response* with first-time-pairing set, and the public key. After Vector receives the connection response, he will display the *pin code*. (See the steps in the next section for when this will occur.)

The session *encryption* and *decryption* keys can then created:

```
crypto_kx_client_session_keys(decryptionKey, encryptionKey, publicKey, secretKey,
    Vector_publicKey);
size_t pin_length = strlen(pin);

crypto_generichash(encryptionKey, sizeof(encryptionKey), encryptionKey,
    sizeof(encryptionKey), pin, pin_length);
crypto_generichash(decryptionKey, sizeof(decryptionKey), decryptionKey,
    sizeof(decryptionKey), pin, pin_length);
```

Example 3: Bluetooth LE encryption & decryption keys

37.3.2 Reconnecting

Reconnecting can reused the public and secret keys, and the encryption and decryption keys. It is not known how long these persist on Vector. {Next pairing? Next reboot? Indefinitely?}

37.3.3

Encrypting and decryption messages

Vector will send a *nonce* message with the *encryption* and *decryption nonces* to employ in encrypting and decrypting message.

Each received enciphered message can be decrypted from cipher text (cipher, and cipherLen) to the message buffer (message and messageLen) for further processing:

```
crypto_aead_xchacha20poly1305_ietf_decrypt(message, &messageLen, NULL, cipher,
                                              cipherLen, NULL, 0L, decryptionNonce, decryptionKey);
sodium_increment(decryptionNonce, sizeof decryptionNonce);
```

Example 4: Decrypting a Bluetooth LE message

Note: the decryptionNonce is incremented each time a message is decrypted.

Each message to be sent can be encrypted from message buffer (message and messageLen) into cipher text (cipher, and cipherLen) that can be fragmented and sent:

```
crypto_aead_xchacha20poly1305_ietf_encrypt(cipher, &cipherLen, message,
                                              messageLen, NULL, 0L, NULL, encryptionNonce, encryptionKey);
sodium_increment(encryptionNonce, sizeof encryptionNonce);
```

Example 5: Encrypting a Bluetooth LE message

Note: the encryptionNonce is incremented each time a message is encrypted.

37.4. THE RTS LAYER

There is an extra, pragmatic layer before the messages can be interpreted by the application. The message has two to three bytes at the header:

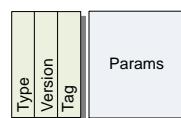


Figure 48: The format of an RTS frame

- The type byte is either 1 or 4. If it is 1 the version of the message format is 1.
- If type byte is 4, the version is held in the next byte. (If the type is 1, there is no version byte).
- The next byte is the tag – the value used to interpret the message.

The tag, parameter body, and version are passed to the CLAD layer for interpretation. This is described in the next section.

37.5. FETCHING A LOG

The process to set up a Bluetooth LE communication with Vector is complex. The sequence has many steps:

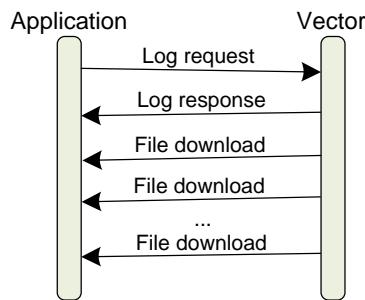


Figure 49: Sequence for initiating communication with Vector

The log request is sent to Vector. In principal this includes a list of the kinds of logs (called filter names) to be included. In practice, the “filter name” makes no difference.

Vector response, and if there will be a file sent, includes an affirmative and a 32-bit file identifier used for the file transfer.

Vector zips the log files up (as a tar.bz2 compressed archive) and sends the chunks to the application. Each chunk has this file identifier. (Conceptually there could be several files in transfer at a time.)

The file transfer is complete when the packet number matches the packet total.

38. MESSAGE FORMATS

This section describes the format and interpretation of the CLAD messages that go between the App and Vector. It describes the fields and how they are encoded, etc. Fields that do not have a fixed location, have no value for their offset. Some fields are only present in later versions of the protocol. They are marked with the version that they are present.

Except where otherwise stated:

- Requests are from the mobile application to Vector, and responses are Vector to the application
- All values in little endian order

	Request	Response	Min Version
Application connection id	1F ₁₆	20 ₁₆	4
Cancel pairing	10 ₁₆		0
Challenge	04 ₁₆	04 ₁₆	0
Challenge success	05 ₁₆		0
Connect	01 ₁₆	02 ₁₆	0
Cloud session	1D ₁₆	1E ₁₆	3
Disconnect	11 ₁₆		0
File download		1a ₁₆	2
Log	18 ₁₆	19 ₁₆	2
Nonce	03 ₁₆	12 ₁₆	
OTA cancel	17 ₁₆		2
OTA update	0E ₁₆	0F ₁₆	0
SDK proxy	22 ₁₆	23 ₁₆	5
Response		21 ₁₆	4
SSH	15 ₁₆	16 ₁₆	0
Status	0A ₁₆	0B ₁₆	0
WiFi access point	13 ₁₆	14 ₁₆	0
WiFi connect	06 ₁₆	07 ₁₆	0
WiFi forget	1B ₁₆	1C ₁₆	3
WiFi IP	08 ₁₆	09 ₁₆	0
WiFi scan	0C ₁₆	0D ₁₆	0

Table 32: Summary of the commands

38.1. APPLICATION CONNECTION ID

?

38.1.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	2	uint16_t	<i>name length</i>	The length of the application connection id; may be 0
2	varies	uint8_t*[name length]	<i>name</i>	The application connection id

Table 33: Parameters for Application Connection Id request

38.1.2 Response

There is no response.

38.2. CANCEL PAIRING

Speculation: this is sent by the application to cancel the pairing process

38.2.1 Request

The command has no parameters.

38.2.2 Response

There is no response.

38.3. CHALLENGE

This is sent by Vector if he liked the response to a nonce message.

38.3.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	4	uint8_t	value	The challenge value

Table 34: Parameters for challenge request

The application, when it receives this message, should increment the value and send the response (a challenge message).

38.3.2 Response

The parameters of the response body are:

Offset	Size	Type	Parameter	Description
0	4	uint8_t	value	The challenge value; this is 1 + the value that was received.

Table 35: Parameters for challenge response

If Vector accepts the response, he will send a *challenge success*.

38.4. CHALLENGE SUCCESS

This is sent by Vector if the challenge response was accepted.

38.4.1

Request

The command has no parameters.

38.4.2

Response

There is no response.

38.5. CLOUD SESSION

This command is used to request a cloud session.

38.5.1 Command

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	2	uint16_t	<i>session token length</i>	The number of bytes in the session token; may be 0
2	varies	uint8_t	<i>session token</i>	The session token, as received from the cloud server. ²⁶
	1	uint8_t	<i>client name length</i>	The number of bytes in the client name string; may be 0
	varies	uint8_t[]	<i>client name</i>	version >= 5 The client name string. Informational only. The mobile app uses the name of the mobile device. version >= 5
	1	uint8_t	<i>application id length</i>	The number of bytes in the application id string; may be 0; version >= 5
	varies	uint8_t[]	<i>application id</i>	The application id. Informational only. The mobile uses “companion-app”. version >= 5

Table 36: Parameters for Cloud Session request

38.5.2 Response result

The parameters for the connection response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>success</i>	0 if failed, otherwise successful
1	1	uint8_t	<i>status</i>	See Table 38: Cloud status enumeration
2	1	uint16_t	<i>client token GUID length</i>	The number of bytes in the client token GUID; may be 0
	varies	uint8_t[]	<i>client token GUID</i>	The client token GUID. The client token GUID should be saved for future use.

Table 37: Parameters for Cloud Session Response

The cloud status types are:

Index	Meaning
0	unknown error
1	connection error
2	wrong account
3	invalid session token
4	authorized as primary
5	authorized as secondary
6	reauthorization

Table 38: Cloud status enumeration

²⁶ <https://groups.google.com/forum/#!msg/anki-vector-rooting/YIYQsX08OD4/fvkAOZ91CgAJ>
<https://groups.google.com/forum/#!msg/anki-vector-rooting/XAaBE6e94ek/OdES50PaBQAJ>

38.6. CONNECT

The connect request *comes from Vector* at the start of a connection. The response is from the application.

38.6.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	32	uint8_t[32]	<i>publicKey</i>	The public key for the connection

Table 39: Parameters for Connection request

The application, when it receives this message, should use the public key for the session, and send a response back.

38.6.2 Response

The parameters for the connection response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>connectionType</i>	See Table 41: Connection types enumeration
1	32	uint8_t[32]	<i>publicKey</i>	The public key to use for the connection

Table 40: Parameters for Connection Response

The connection types are:

Index	Meaning
0	first time pairing (requests pin code to be displayed)
1	reconnection

Table 41: Connection types enumeration

The application sends the response, with its *publicKey* (see section 37.3 *Encryption support*). A “first time pairing” connection type will cause Vector to display a pin code on the screen

If a first time pairing response is sent:

- If Vector is not in pairing mode – was not put on his charger and the backpack button pressed twice, quickly – Vector will respond. Attempting to enter pairing mode now will cause Vector to send a *disconnect* request.
- If Vector is in pairing mode, Vector will display a pin code on the screen, and send a nonce message, triggering the next steps of the conversation.

If a reconnection is sent, the application would employ the public and secret keys, and the encryption and decryption keys from a prior pairing.

38.7. DISCONNECT

This may be sent by Vector if there is an error, and it is ending communication. For instance, if Vector enters pairing mode, it will send a disconnect.

The application may send this to request Vector to close the connection.

38.7.1

Request

The command has no parameters.

38.7.2

Response

There is no response.

38.8. FILE DOWNLOAD

This command is used to pass chunks of a file to Vector. Files are broken up into chunks and sent.

38.8.1 Request

There is no direct request.

38.8.2 Response

The parameters of the response body are:

Table 42: Parameters for File Download request

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>status</i>	
1	4	uint32_t	<i>file id</i>	
5	4	uint32_t	<i>packet number</i>	The chunk within the download
9	4	uint32_t	<i>packet total</i>	The total number of packets to be sent for this file download
13	2	uint16_t	<i>length</i>	The number of bytes to follow (can be 0)
varies	varies	uint8_t[length]	<i>bytes</i>	The bytes of this file chunk

38.9. LOG

This command is used to request the Vector send a compressed archive of the logs.

38.9.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>mode</i>	
1	2	uint16_t	<i>num filters</i>	The number of filters in the array
3	varies	filter[num filters]	<i>filters</i>	The filter names

Table 43: Parameters for Log request

Each filter entry has the following structure:

Offset	Size	Type	Parameter	Description
0	2	uint16_t	<i>filter length</i>	The length of the filter name; may be 0
2	varies	uint8_t[filter length]	<i>filter name</i>	The filter name

Table 44: Log filter

38.9.2 Response

It can take several seconds for Vector to prepare the log archive file and send a response. The response will be a “log response” (below) and a series of “file download” responses.

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>exit code</i>	
1	4	uint32_t	<i>file id</i>	A 32-bit identifier that will be used in the file download messages.

Table 45: Parameters for Log Response

38.10. NONCE

A nonce is sent by Vector after he has accepted your key, and the application sends a response

38.10.1 Request

The parameters for the nonce request message are:

Offset	Size	Type	Parameter	Description
0	24	uint8_t[24]	<i>toVectorNonce</i>	The nonce to use for sending stuff to Vector
24	24	uint8_t[24]	<i>toAppNonce</i>	The nonce for receiving stuff from Vector

Table 46: Parameters for Nonce request

38.10.2 Response

After receiving a nonce, if the application is in first-time pairing the application should send a response, with a value of 3.

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>connection tag</i>	This is always 3

Table 47: Parameters for Nonce response

After the response has been sent, the channel will now be encrypted. If vector likes the response, he will send a challenge message.

38.11. OTA UPDATE

This command is used to request the Vector download software from a given server

38.11.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>length</i>	The length of the URL; may be 0
1	varies	uint8_t[length]	<i>URL</i>	The URL string

Table 48: Parameters for OTA request

38.11.2 Response

The response will be one or more “OTA response” indicating the status of the update, or errors. Status codes ≥ 200 indicate that the update process has completed. The update has completed the download when the current number of bytes match the expected number of bytes.

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>status</i>	See Table 50: OTA status enumeration
1	8	uint64_t	<i>current</i>	The number of bytes downloaded
9	8	uint64_t	<i>expected</i>	The number of bytes expected to be downloaded

Table 49: Parameters for OTA Response

The OTA status codes are:

Status	Meaning
0	idle
1	unknown
2	in progress
3	complete
4	rebooting
5	error
200...	Status codes from the update-engine. See Appendix D, Table 512: OTA update-engine status codes.

Table 50: OTA status enumeration

Note: the status codes 200 and above are from the update-engine, and are given in Appendix D.

38.12. RESPONSE

This message will be sent on the event of an error. Primarily if the session is not cloud authorized and the command requires it.

Offset	Size	Type	Parameter	Description
0	1	uint16_t	<i>code</i>	0 if not cloud authorized, otherwise authorized
1	1	uint8_t	<i>length</i>	The number of bytes in the string that follows.
<i>varies</i>	uint8_t [<i>length</i>]		<i>text</i>	A text error message.

Table 51: Parameters for Response

38.13. SDK PROXY

This command is used to pass the gRPC/protobufs messages to Vector over Bluetooth LE. It effectively wraps a HTTP request/response. Note: the HTTPS TLS certificate is not employed with this command.

38.13.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>GUID length</i>	The number of bytes in the GUID string; may be 0
2	varies	uint8_t[<i>GUID length</i>]	<i>GUID</i>	The GUID string
	1	uint8_t	<i>msg length</i>	The number of bytes in the message id string
	varies	uint8_t[<i>msg id length</i>]	<i>msg id</i>	The message id string
	1	uint8_t	<i>path length</i>	The number of bytes in the URL path string
	varies	uint8_t[<i>path length</i>]	<i>path</i>	The URL path string
2	uint16_t		<i>JSON length</i>	The length of the JSON
varies	uint8_t[<i>JSON length</i>]		<i>JSON</i>	The JSON (string)

Table 52: Parameters for the SDK proxy request

38.13.2 Response

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>msg id length</i>	The number of bytes in the message id string; may be 0
2	varies	uint8_t[<i>msg id length</i>]	<i>msg id</i>	The message id string
	2	uint16_t	<i>status code</i>	The HTTP-style status code that the SDK may return.
	1	uint8_t	<i>type length</i>	The number of bytes in the response type string
	varies	uint8_t[<i>type length</i>]	<i>type</i>	The response type string
	2	uint16_t	<i>body length</i>	The length of the response body
varies	uint8_t[<i>body length</i>]		<i>body</i>	The response body (string)

Table 53: Parameters for the SDK proxy Response

38.14. SSH

This command is used to request the Vector allow SSH. SSH is supported only in developer releases (and not all). SSH is not supported in the production release software.

38.14.1

Request

The SSH key command passes the authorization key by dividing it up into substrings and passing the list of substrings. The substrings are appended together by the recipient to make for the overall authorization key.

The parameters for the request message are:

Offset	Size	Type	Parameter	Description
0	2	uint16_t	<i>num substrings</i>	The number of SSH authorization keys; may be 0
2	varies	substring[num substrings]	<i>substrings</i>	The array of authorization key strings (see below).

Table 54: Parameters for SSH request

Each authorization key substring has the following structure:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>substring length</i>	The length of the substring; may be 0
1	varies	uint8_t[substring length]	<i>substring</i>	UTF8 substring of the SSH authorization key

Table 55: SSH authorization key substring

38.14.2

Response

The response has no parameters.

38.15. STATUS

This command is used to request basic info from Vector.

38.15.1 Request

The request has no parameters.

38.15.2 Response

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>SSID length</i>	The number of bytes in the SSID string; may be 0
2	varies	uint8_t[SSID length]	<i>SSID</i>	The WiFi SSID (hex string).
1	uint8_t		<i>WiFi state</i>	See <i>Table 57: WiFi state enumeration</i>
1	uint8_t		<i>access point</i>	0 not acting as an access point, otherwise acting as an access point
1	uint8_t		<i>Bluetooth LE state</i>	0 if the Bluetooth
1	uint8_t		<i>Battery state</i>	
1	uint8_t		<i>version length</i>	The number of bytes in the version string; may be 0 version >= 2
varies	uint8_t [version length]		<i>version</i>	The version string; version >= 2
1	uint8_t		<i>ESN length</i>	The number of bytes in the ESN string; may be 0 version >= 4
varies	uint8_t[ESN length]		<i>ESN</i>	The <i>electronic serial number</i> string; version >= 4
1	uint8_t		<i>OTA in progress</i>	0 over the air update not in progress, otherwise in process of over the air update; version >= 2
1	uint8_t		<i>has owner</i>	0 does not have an owner, otherwise has an owner; version >= 3
1	uint8_t		<i>cloud authorized</i>	0 is not cloud authorized, otherwise is cloud authorized; version >= 5

Table 56: Parameters for Status Response

Note: a *hex string* is a series of bytes with values 0-15. Every pair of bytes must be converted to a single byte to get the characters. Even bytes are the high nibble, odd bytes are the low nibble.

The WiFi states are:

Index	Meaning
0	Unknown
1	Online
2	Connected
3	Disconnected

Table 57: WiFi state enumeration

38.16. WIFI ACCESS POINT

This command is used to request that the Vector act as a WiFi access point. This command requires that a “cloud session” have been successfully started first (see section 38.5 *Cloud session*).

If successful, Vector will provide a WiFi Access Point with an SSID that matches his robot name.

38.16.1 Request

The parameters of the request body are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>enable</i>	0 to disable the WiFi access point, 1 to enable it

Table 58: Parameters for WiFi Access Point request

38.16.2 Response

If the Bluetooth LE session is not cloud authorized a “response” message will be sent with this error. Otherwise the WiFi Access Point response message will be sent.

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>enabled</i>	0 if the WiFi access point is disabled, otherwise enabled
1	1	uint8_t	<i>SSID length</i>	The number of bytes in the SSID string; may be 0
2	varies	uint8_t[SSID length]	<i>SSID</i>	The WiFi SSID (hex string)
1	1	uint8_t	<i>password length</i>	The number of bytes in the password string; may be 0
varies	varies	uint8_t [password length]	<i>password</i>	The WiFi password

Table 59: Parameters for WiFi Access Point Response

38.17. WIFI CONNECT

This command is used to request Vector to connect to a given WiFi SSID. Vector will retain this WiFi for future use.

A pretty Wi-Fi for the little guy

38.17.1 Request

The parameters for the request message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>SSID length</i>	The number of bytes in the SSID string; may be 0
1	varies	uint8_t[SSID length]	<i>SSID</i>	The WiFi SSID (hex string)
1		uint8_t	<i>password length</i>	The number of bytes in the password string; may be 0
varies		uint8_t [password length]	<i>password</i>	The WiFi password
1		uint8_t	<i>timeout</i>	How long to given the connect attempt to succeed.
1		uint8_t	<i>auth type</i>	The type of authentication to employ; see <i>Table 61: WiFi authentication types enumeration</i>
1		uint8_t	<i>hidden</i>	0 the access point is not hidden; 1 it is hidden

Table 60: Parameters for WiFi Connect request

The WiFi authentication types are:

Index	Meaning
0	None, open
1	WEP
2	WEP shared
3	IEEE8021X
4	WPA PSK
5	WPA2 PSK
6	WPA2 EAP

Table 61: WiFi authentication types enumeration

38.17.2 Response

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>SSID length</i>	The length of the SSID that was deleted; may be 0
1	varies	uint8_t[SSID length]	<i>SSID</i>	The SSID (hex string) that was deleted
1		uint8_t	<i>WiFi state</i>	See <i>Table 57: WiFi state enumeration</i>
1		uint8_t	<i>connect result</i>	version >= 3

Table 62: Parameters for WiFi Connect command

38.18. WIFI FORGET

This command is used to request Vector to forget a WiFi SSID.

38.18.1 Request

The parameters for the request message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>delete all</i>	0 if Vector should delete only one SSID; otherwise Vector should delete all SSIDs
1	1	uint8_t	<i>SSID length</i>	The length of the SSID that to be deleted; may be 0
2	varies	uint8_t[SSID length]	<i>SSID</i>	The SSID (hex string) to be deleted

Table 63: Parameters for WiFi Forget request

38.18.2 Response

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>did delete all</i>	0 if only one; otherwise Vector deleted all SSIDs
1	1	uint8_t	<i>SSID length</i>	The length of the SSID that was deleted; may be 0
2	varies	uint8_t[SSID length]	<i>SSID</i>	The SSID (hex string) that was deleted

Table 64: Parameters for WiFi Forget response

38.19. WIFI IP ADDRESS

This command is used to request Vector's WiFi IP address.

38.19.1 Request

The request has no parameters

38.19.2 Response

The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>has IPv4</i>	0 if Vector doesn't have an IPv4 address; other it does
1	1	uint8_t	<i>has IPv6</i>	0 if Vector doesn't have an IPv6 address; other it does
2	4	uint8_t[4]	<i>IPv4 address</i>	Vector's IPv4 address
6	32	uint8_t[16]	<i>IPv6 address</i>	Vector's IPv6 address

Table 65: Parameters for WiFi IP Address response

38.20. WIFI SCAN

This command is used to request Vector to scan for WiFi access points.

*finding hot signals in
Vectors area*

38.20.1 Request

The command has no parameters.

38.20.2 Response

The response lists the Wi-Fi access points Vector can find. The parameters for the response message are:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>status code</i>	
1	1	uint8_t	<i>num entries</i>	The number of access points in the array below
2	varies	AP[num entries]	<i>access points</i>	The array of access points

Table 66: Parameters for WiFi scan response

Each access point has the following structure:

Offset	Size	Type	Parameter	Description
0	1	uint8_t	<i>auth type</i>	The type of authentication to employ; see <i>Table 61: WiFi authentication types enumeration</i>
1	1	uint8_t	<i>signal strength</i>	The number of bars, 0..4
2	1	uint8_t	<i>SSID length</i>	The length of the SSID string
3	varies	uint8_t[SSID length]	<i>SSID</i>	The SSID (hex string)
	1	uint8_t	<i>hidden</i>	0 not hidden, 1 hidden; version >= 2
	1	uint8_t	<i>provisioned</i>	0 not provisioned, 1 provisioned; version>= 3

Table 67: Parameters access point structure

CHAPTER 13

The HTTPS based API

This chapter describes the communication with Vector via the local HTTPS.

Note: the information in this chapter comes from the protobuf specification files in the python SDK, from the SDK itself, and some analysis of the mobile application. All quotes (unless otherwise indicated) are from the SDK.

39. OVERVIEW OF THE SDK HTTPS API

The descriptions below²⁷ give the JSON keys, and their value format. It is implemented as gRPC/protobufs interaction over HTTP. (Anki has frequently said that the SDK included code (as python) with the protobuf spec so that others could use their own preferred implementation language.) Each command is requested by POST-ing the request structure to the given relative URL (relative to Vector's address or local network name) and interpreting the returned body as the response structure.

The HTTPS header should include

- Bearer *BASE64KEY*
- Content-Type: application/json

(The JSON request is posted in the body)

39.1. SDK MESSAGE GROUPINGS

The major groups of messages here are:

- Accessories and custom objects
- Actions and behaviors – setting the current priority and cancelling actions
- Alexa configuration – configuring Vector to use Alexa's services
- Audio – playing sounds on Vector, and submitting text to speech
- Battery – the current state of charge
- Connection – authenticating with the remote servers to allow access to Vector, connection management, event stream, and end-point version info
- Cube – commands to manage and interact with the cube
- Diagnostics – checking the connection with the cloud, and uploading log information
- Display – display images on Vector's LCD

²⁷ The protocol was specified in Google Protobuf.

- Faces (of people, not Vector's face)
- Features and entitlements – the features that are enabled (or disabled)
- Image processing – Getting a video stream, and enabling (or disabling) video processing steps
- Interactions with objects (outside of the cube)
- JDocs, the JSON document storage interface
- Map and Navigation
- Motion Control
- Motion Sensing – how Vector senses that he is moving
- Onboarding
- Photos – commands to access (and delete) photographs and their thumbnails
- Settings and Preferences
- Software Updates, used to update Vector's software – operating system, applications, assets, etc.

40. COMMON ELEMENTS

The enumerations and structures in this section are common to many commands.

40.1. ENUMERATIONS

40.1.1 ResultCode

The ResultCode enumeration has the following named values:

Name	Value	Description
ERROR_UPDATE_IN_PROGRESS	1	The settings could not be applied; there is already another update to the settings in process.
SETTINGS_ACCEPTED	0	The settings were successfully saved.

Table 68: ResultCode Enumeration

40.1.2 RobotStatus

The RobotStatus is a bit mask used to indicate the Vector is doing, and the status of his controls. It is used in the RobotState message. The enumeration has the following named bits (any number may be set). Note that some bits have two names; the second name is one employed by Anki's python SDK.

Name	Value	Description
ROBOT_STATUS_NONE	00000 ₁₆	
ROBOT_STATUS_IS_MOVING	00001 ₁₆	This bit is set “if Vector is currently moving any of his motors (head, arm or wheels/treads).”
ROBOT_STATUS_ARE_MOTORS_MOVING	00001 ₁₆	
ROBOT_STATUS_IS_CARRYING_BLOCK	00002 ₁₆	This bit is set “if Vector is currently carrying a block.”
ROBOT_STATUS_IS_PICKING_OR_PLACING	00004 ₁₆	This bit is set “if Vector has seen a marker and is actively heading toward it (for example his charger or cube).”
ROBOT_STATUS_IS_DOCKING_TO_MARKER	00004 ₁₆	
ROBOT_STATUS_IS_PICKED_UP	00008 ₁₆	This bit is set “if Vector is currently picked up (in the air),” being held or is on his side. Vector “uses the IMU data to determine if the robot is not on a stable surface with his treads down.” If Vector is not on stable surface (with his treads down), this bit is set.
ROBOT_STATUS_IS_BUTTON_PRESSED	00010 ₁₆	This bit is set “if Vector's button is pressed.”
ROBOT_STATUS_IS_FALLING	00020 ₁₆	This bit is set “if Vector is currently falling.”
ROBOT_STATUS_IS_ANIMATING	00040 ₁₆	This bit is set “if Vector is currently playing an animation.”
ROBOT_STATUS_IS_PATHING	00080 ₁₆	This bit is set “if Vector is currently traversing a path.”
ROBOT_STATUS_LIFT_IN_POS	00100 ₁₆	This bit is set “if Vector's arm is in the desired position.” It is clear “if still trying to move it there.”
ROBOT_STATUS_HEAD_IN_POS	00200 ₁₆	This bit is set “if Vector's head is in the desired position.” It is clear “if still trying to move there.”
ROBOT_STATUS_CALM_POWER_MODE	00400 ₁₆	This bit is set “if Vector is in calm power mode.”

Table 69: RobotStatus Enumeration

		Calm power mode is generally when Vector is sleeping or charging.”
<i>reserved</i>	00800_{16}	<i>This bit is not defined</i>
<i>ROBOT_STATUS_IS_ON_CHARGER</i>	01000_{16}	This bit is set “if Vector is currently on the charger.” Note: Vector may be on the charger without charging.
<i>ROBOT_STATUS_IS_CHARGING</i>	02000_{16}	This bit is set “if Vector is currently charging.”
<i>ROBOT_STATUS_CLIFF_DETECTED</i>	04000_{16}	This bit is set “if Vector detected a cliff using any of his four cliff sensors.”
<i>ROBOT_STATUS_ARE_WHEELS_MOVING</i>	08000_{16}	This bit is set “if Vector's wheels/treads are currently moving.”
<i>ROBOT_STATUS_IS_BEING_HELD</i>	10000_{16}	This bit is set “if Vector is being held.” Note: <i>ROBOT_STATUS_IS_PICKED_UP</i> will also be set when this bit is set.
		Vector “uses the IMU to look for tiny motions that suggest the robot is actively being held in someone's hand.” This is used to distinguish from other cases, such as falling, on its side, etc.
<i>ROBOT_STATUS_IS_MOTION_DETECTED</i>	20000_{16}	This bit is set “if Vector is in motion. This includes any of his motors (head, arm, wheels/tracks) and if he is being lifted, carried, or falling.”
<i>ROBOT_STATUS_IS_ROBOT_MOVING</i>		
<i>ROBOT_STATUS_IS_BATTERY_OVERHEATED</i>	40000_{16}	<i>Speculated to exist in 1.6.</i> This bit is set if Vector's battery temperature is considered too hot.
<i>ROBOT_STATUS_ENCODERS_DISABLED</i>	80000_{16}	<i>Speculated to exist in 1.6.</i> This bit is set if Vector has turned off the motor encoders. This is done to save power when the motors are idle.

40.2. STRUCTURES

40.2.1 CladPoint

The CladPoint is used to represent a 2D rectilinear point on an image or in the 2D map. It has the following fields:

Field	Type	Units	Description
x	float	<i>pixels</i>	The x-coordinate of the point
y	float	<i>pixels</i>	The y-coordinate of the point

Table 70: CladPoint JSON structure

40.2.2

CladRect

The CladRect is used to represent a 2D rectilinear rectangle on an image. It has the following fields:

Field	Type	Units	Description
<i>height</i>	float	<i>pixels</i>	The height of the rectangle
<i>width</i>	float	<i>pixels</i>	The width of the rectangle
<i>x_top_left</i>	float	<i>pixels</i>	The x-coordinate of the top-left corner of the rectangle within the image.
<i>y_top_left</i>	float	<i>pixels</i>	The y-coordinate of the top-left corner of the rectangle within the image.

Table 71:
CladRectangle JSON structure

40.2.3

PoseStruct

The PoseStruct is used to represent a 3D rectilinear point and orientation on the map. It has the following fields:

Field	Type	Units	Description
<i>origin_id</i>	uint32		Which version of the map this pose is in (0 for none or unknown). See Chapter 18 for a description of the mapping origin id.
<i>q0</i>	float		Part of the rotation quaternion
<i>q1</i>	float		Part of the rotation quaternion
<i>q2</i>	float		Part of the rotation quaternion
<i>q3</i>	float		Part of the rotation quaternion
<i>x</i>	float	<i>mm</i>	The x coordinate
<i>y</i>	float	<i>mm</i>	The y coordinate
<i>z</i>	float	<i>mm</i>	The z coordinate

Table 72: PoseStruct JSON structure

40.2.4

ResponseStatus

The ResponseStatus is “a shared response message sent back as part of most requests. This will indicate the generic state of the request.” It has the following fields:

Field	Type	Units	Description
<i>code</i>	StatusCode		“The generic status code to give high-level insight into the progress of a given message.”

Table 73:
ResponseStatus JSON structure

The StatusCode is used to indicate state of the request.

Name	Value	Description
UNKNOWN	0	
RESPONSE_RECEIVED	1	“The message has completed as expected.”
REQUEST_PROCESSING	2	“The message has been sent to the robot.”
OK	3	“The message has been handled successfully at the interface level.”
FORBIDDEN	100	“The user was not authorized.”
NOT_FOUND	101	“The requested attribute was not found.”
ERROR_UPDATE_IN_PROGRESS	102	“Currently updating values from another call.”

Table 74: StatusCode Enumeration

41. ACCESSORIES AND CUSTOM OBJECTS

This section describes the objects that Vector can see and track in his map. Specialized accessories – the charger and cube – are broken out into their own sections.

See also section 49 *Cube* and section 55 *Interactions with Objects*

You too can create custom objects for Vector to... at least see and perceive. Maybe even love.

There are four kinds of custom objects that you can define:

- A fixed, unmarked cube-shaped object. The object is in a fixed position and orientation, and it can't be observed (since it is unmarked). So there won't be any events related to this object. “This could be used to make Vector aware of objects and know to plot a path around them.”
- A flat wall with only a front side,
- A cube, with the same marker on each side.
- A box with different markers on each side.

A note about object id's: The object id may change: “a cube disconnecting and reconnecting it's removed and then re-added to robot's internal world model which results in a new ID.”

The client should employ a timer for each potential visual object. If there isn't an “object observed” event received in the time period, it should be assumed “that Vector can no longer see an object.”

41.1. ENUMERATIONS

- The CustomObjectMarker enumerates the marker symbols
- The CustomType refers to the one of the 20 possible custom objects that can be defined
- The ObjectFamily is an older, now deprecated method, of enumerating the kind of object (as in, charger, light cube, wall, box, or custom cube).
- The ObjectType enumeration is the preferred method of enumerating the kinds of objects

41.1.1 CustomObjectMarker

The CustomObjectMarker is used represent the marker symbol used. The symbols are predefined, with the images that Vector recognizes included in the SDK. The enumeration has the following named values:

Name	Value	Description	Table 75: CustomObjectMarker Enumeration
CUSTOM_MARKER_UNKNOWN	0		
CUSTOM_MARKER_CIRCLES_2	1		
CUSTOM_MARKER_CIRCLES_3	2		
CUSTOM_MARKER_CIRCLES_4	3		
CUSTOM_MARKER_CIRCLES_5	4		
CUSTOM_MARKER_DIAMONDS_2	5		
CUSTOM_MARKER_DIAMONDS_3	6		
CUSTOM_MARKER_DIAMONDS_4	7		
CUSTOM_MARKER_DIAMONDS_5	8		
CUSTOM_MARKER_HEXAGONS_2	9		
CUSTOM_MARKER_HEXAGONS_3	10		
CUSTOM_MARKER_HEXAGONS_4	11		
CUSTOM_MARKER_HEXAGONS_5	12		
CUSTOM_MARKER_TRIANGLES_2	13		
CUSTOM_MARKER_TRIANGLES_3	14		
CUSTOM_MARKER_TRIANGLES_4	15		
CUSTOM_MARKER_TRIANGLES_5	16		
CUSTOM_MARKER_COUNT	16		

41.1.2

CustomType

The CustomType is used to represent the identifier of object that a symbol is attached to. The enumeration has the following named values:

Table 76: CustomType Enumeration

Name	Value	Description
<i>INVALID_CUSTOM_TYPE</i>	0	
<i>CUSTOM_TYPE_00</i>	1	
<i>CUSTOM_TYPE_01</i>	2	
<i>CUSTOM_TYPE_02</i>	3	
<i>CUSTOM_TYPE_03</i>	4	
<i>CUSTOM_TYPE_04</i>	5	
<i>CUSTOM_TYPE_05</i>	6	
<i>CUSTOM_TYPE_06</i>	7	
<i>CUSTOM_TYPE_07</i>	8	
<i>CUSTOM_TYPE_08</i>	9	
<i>CUSTOM_TYPE_09</i>	10	
<i>CUSTOM_TYPE_10</i>	11	
<i>CUSTOM_TYPE_11</i>	12	
<i>CUSTOM_TYPE_12</i>	13	
<i>CUSTOM_TYPE_13</i>	14	
<i>CUSTOM_TYPE_14</i>	15	
<i>CUSTOM_TYPE_15</i>	16	
<i>CUSTOM_TYPE_16</i>	17	
<i>CUSTOM_TYPE_17</i>	18	
<i>CUSTOM_TYPE_18</i>	19	
<i>CUSTOM_TYPE_19</i>	20	
<i>CUSTOM_TYPE_COUNT</i>	20	

41.1.3

ObjectFamily

The ObjectFamily is a deprecated method used to represent the type of object that a symbol is attached to. ObjectType should be used instead, where possible. The enumeration has the following named values:

Name	Value	Description
INVALID_FAMILY	0	This value represents a kind of object that is not properly set.
UNKNOWN_FAMILY	1	This value is used when there is an object, but its kind is not known.
BLOCK	2	This is the identifier used for blocks/cubs other than the companion-cube
LIGHT_CUBE	3	This is the identifier used for the companion-cube
CHARGER	4	This is the identifier used for the home charging station.
CUSTOM_OBJECT	7	This is the identifier used for as custom object definition.
OBJECT_FAMILY_COUNT	7	

Table 77: ObjectType Enumeration

41.1.4

ObjectType

The ObjectType is used represent the type of object that a symbol is attached to. The enumeration has the following named values:

Name	Value	Description
INVALID_OBJECT	0	This value represents an object id used when there isn't an object associated.
UNKNOWN_OBJECT	1	This value is used when there is an object, but it is not recognized.
BLOCK_LIGHTCUBE1	2	This is the identifier used for the companion-cube
CHARGER_BASIC	6	This is the identifier used for the home charging station.
FIRST_CUSTOM_OBJECT_TYPE	15	The custom objects all have types greater than or equal to this.

Table 78: ObjectType Enumeration

41.2. EVENTS

These are the events that are sent to inform the application of an objects state (and availability).

41.2.1

ObjectEvent

The ObjectEvent event is sent (see *Event* message) when the state of an object has changed. The structure has one (and only one) of the following fields:

Field	Type	Description
<i>cube_connection_lost</i>	CubeConnectionLost	This event is sent when cube no longer is connected via Bluetooth LE.
<i>robot_observed_object</i>	RobotObservedObject	This even is sent the object is visually seen by Vector.
<i>object_available</i>	ObjectAvailable	This event is sent when cube a Bluetooth LE connection to the cube is established.
<i>object_connection_state</i>	ObjectConnectionState	The information about the Bluetooth LE identity of the cube, and whether is connected (or not).
<i>object_moved</i>	ObjectMoved	The object has changed position.
<i>object_stopped_moving</i>	ObjectStoppedMoving	The object had change position previously, but has now come to reset.
<i>object_tapped</i>	ObjectTapped	The cube was tapped.
<i>object_up_axis_changed</i>	ObjectUpAxisChanged	The object was rotated and has a new upward face.

Table 79: ObjectEvent JSON structure

41.2.2

ObjectAvailable

The ObjectAvailable event is sent (see section 41.2.1 *ObjectEvent*) when Vector has received Bluetooth LE advertisements from the object (cube).

See also section 49.2.2 *CubeConnectionLost*

This event structure has the following fields:

Field	Type	Units	Description
<i>factory_id</i>	string		The identifier for the cube. This is built into the cube.

Table 80: ObjectAvailable JSON structure

41.2.3

ObjectConnectionState

The ObjectConnectedState event is to “indicate that a cube has connected or disconnected to the robot. This message will be sent for any connects or disconnects regardless of whether it originated from us or underlying robot behavior.”

See also section 49.2.2 *CubeConnectionLost*

This event structure has the following fields:

Field	Type	Units	Description
<i>connected</i>	bool		True if Vector has a Bluetooth LE connection with the Cube.
<i>factory_id</i>	string		The identifier for the cube. This is built into the cube.
<i>object_id</i>	uint32		The identifier of the object that Vector is (or was) connected to.
<i>object_type</i>	ObjectType		The type of object referred to.

Table 81:
ObjectConnectedState
JSON structure

41.2.4

ObjectMoved

The ObjectMoved event is sent (see section 41.2.1 *ObjectEvent*) when an object has changed its position. The structure has the following fields:

Field	Type	Units	Description
<i>object_id</i>	uint32		The identifier of the object that moved.
<i>timestamp</i>	uint32		The time that the event occurred on. The format is milliseconds since Vector’s epoch.

Table 82: *ObjectMoved*
JSON structure

41.2.5

ObjectStoppedMoving

The ObjectStoppedMoving event is sent (see section 41.2.1 *ObjectEvent*) when an object previously identified as moving has come to reset. The structure has the following fields:

Field	Type	Units	Description
<i>object_id</i>	uint32		The identifier of the object that was moving.
<i>timestamp</i>	uint32		The time that the event occurred on. The format is milliseconds since Vector’s epoch.

Table 83:
ObjectStoppedMoving
JSON structure

41.2.6 ObjectUpAxisChanged

The ObjectUpAxis event is sent (see section 41.2.1 *ObjectEvent*) if the orientation of the object has significantly changed, leaving it with a new face upward. The structure has the following fields:

Field	Type	Units	Description
<i>object_id</i>	uint32		The identifier of the object whose axis has changed.
<i>timestamp</i>	uint32		The time that the event occurred on. The format is milliseconds since Vector's epoch.
<i>up_axis</i>	UpAxis		The orientation of object, represented as which axis is pointing upwards

Table 84: ObjectUpAxis JSON structure

The UpAxis is used represent the orientation of an object. The enumeration has the following named values:

Name	Value	Description
<i>INVALID_AXIS</i>	0	The orientation of the object is not known.
<i>X_NEGATIVE</i>	1	The positive direction along the body's x-axis is upward.
<i>X_POSITIVE</i>	2	The negative direction along the body's x-axis is upward.
<i>Y_NEGATIVE</i>	3	The positive direction along the body's y-axis is upward.
<i>Y_POSITIVE</i>	4	The negative direction along the body's y-axis is upward.
<i>Z_NEGATIVE</i>	5	The positive direction along the body's z-axis is upward.
<i>Z_POSITIVE</i>	6	The negative direction along the body's z-axis is upward.
<i>NUM_AXES</i>	7	

Table 85: UpAxis Enumeration

41.2.7

RobotObservedObject

The RobotObservedObject event is sent when “an object with [the] specified ID/Type was seen at a particular location in the image and the world.” This event structure has the following fields:

Table 86:
RobotObservedObject
JSON structure

Field	Type	Units	Description
<i>img_rect</i>	CladRect		The position of the object within the vision image.
<i>is_active</i>	uint32		
<i>object_family</i>	ObjectFamily		<i>Deprecated.</i> “Use ObjectType instead to reason about groupings of objects.”
<i>object_id</i>	int32		The identifier of the object that has been seen. Note that this is signed (int32 instead of uint32) for internal compatibility reasons.
<i>object_type</i>	ObjectType		The type of object referred to.
<i>pose</i>	PoseStruct		The observed pose of this object. <i>Optional.</i>
<i>timestamp</i>	uint32		The time that the object was most recently observed. The format is milliseconds since Vector’s epoch.
<i>top_face_orientation_rad</i>	float	radians	“Angular distance from the current reported up axis. “ “absolute orientation of top face, iff isActive==true”

41.3. CREATE FIXED CUSTOM OBJECT

This command “creates a permanent custom [cube-shaped] object instance in the robot's world” except this object has “no markers associated with it.” The object “will remain in the specified pose as an obstacle forever (or until deleted).” The object can't be observed, and won't create any events related to being observed. The fixed, custom object can “be used to make Vector aware of objects and know to plot a path around them.”

Post: “/v1/create_fixed_custom_object”

41.3.1 Request

The CreateFixedCustomObjectRequest structure has the following fields:

Field	Type	Units	Description
<i>pose</i>	PoseStruct		The position and orientation of this object.
<i>x_size_mm</i>	float	mm	The size of the object that the marker symbol is on, along the x-axis.
<i>y_size_mm</i>	float	mm	The size of the object that the marker symbol is on, along the y-axis.
<i>z_size_mm</i>	float	mm	The size of the object that the marker symbol is on, along the z-axis.

Table 87:
CreateFixedCustomObjectRequest JSON structure

41.3.2 Response

The CreateFixedCustomObjectResponse structure has the following fields:

Field	Type	Description
<i>object_id</i>	uint32	The object identifier assigned to this object.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 88:
CreateFixedCustomObjectResponse JSON structure

41.4. DEFINE CUSTOM OBJECT

“Creates a custom object with distinct custom marker(s)” on one or more its faces. This can create a wall, a box, a cube (similar to a box, but each side is the same size as every other, and has the same marker). Once the object has been created, “the robot will now detect the markers associated with this object and send a RobotObservedObject message when they are seen. The markers must be placed in the center of their respective sides.”

Note: “No instances of this object are added to the world until they have been seen.”

See also *Create Fixed Custom Object*, *Delete Custom Objects*

Post: “/v1/define_custom_object”

41.4.1

Request

The DefineCustomObjectRequest structure has the following fields:

Table 89:
DefineCustomObjectRequest JSON structure

Field	Type	Units	Description
<i>custom_type</i>	CustomType		The object type to be assigned to this object.
<i>is_unique</i>	bool		If true, “there is guaranteed to be no more than one object of this type present in the world at a time.”
<i>custom_box</i>	CustomBoxDefinition		The definition of a box with different markers on each side.
<i>custom_cube</i>	CustomCubeDefinition		The definition of a cube, with the same marker on each side.
<i>custom_wall</i>	CustomWallDefinition		The definition of a flat wall with only a front side.

Note: only one of “*custom_box*,” “*custom_cube*,” or “*custom_wall*” can be used in the request.

The CustomBoxDefinition “defines a custom object of the given size with the given markers centered on each side.” The structure has the following fields:

Field	Type	Units	Description
<i>marker_back</i>	CustomObjectMarker		The marker symbol used on the back surface of the box. This marker must be unique (not used by any of the other side’s on this box or in any other shape).
<i>marker_bottom</i>	CustomObjectMarker		The marker symbol used on the bottom surface of the box. This marker must be unique (not used by any of the other side’s on this box or in any other shape).
<i>marker_front</i>	CustomObjectMarker		The marker symbol used on the front surface of the box. This marker must be unique (not used by any of the other side’s on this box or in any other shape).
<i>marker_left</i>	CustomObjectMarker		The marker symbol used on the left-hand side of the box. This marker must be unique (not used by any of the other side’s on this box or in any other shape).
<i>marker_right</i>	CustomObjectMarker		The marker symbol used on the right-hand side of the box. This marker must be unique (not used by any of the other side’s on this box or in any other shape).
<i>marker_top</i>	CustomObjectMarker		The marker symbol used on the top surface of the box. This marker must be unique (not used by any of the other side’s on this box or in any other shape).
<i>marker_height_mm</i>	float	mm	The height of the marker symbol.
<i>marker_width_mm</i>	float	mm	The width of the marker symbol.
<i>x_size_mm</i>	float	mm	The size of the object, along the x-axis, that the marker symbol is on.
<i>y_size_mm</i>	float	mm	The size of the object, along the y-axis, that the marker symbol is on.
<i>z_size_mm</i>	float	mm	The width of the object, along the z-axis, that the marker symbol is on.

The CustomCubeDefinition “defines a custom cube of the given size.” The structure has the following fields:

Field	Type	Units	Description
<i>marker</i>	CustomObjectMarker		The marker symbol used on all of the cube surfaces; “the same marker [must] be centered on all faces.”
<i>marker_height_mm</i>	float	mm	The height of the marker symbol
<i>marker_width_mm</i>	float	mm	The width of the marker symbol
<i>size_mm</i>	float	mm	The height, width, and depth of the object that the marker symbol is on.

Table 90:
CustomBoxDefinition
JSON structure

The `CustomWallDefinition` “defines a custom wall of the given height and width... The wall's thickness is assumed to be 1cm (and thus there are no markers on its left, right, top, or bottom).” The structure has the following fields:

Field	Type	Units	Description
<code>marker</code>	CustomObjectMarker		The marker symbol used on the wall surfaces; “the same marker centered on both sides (front and back)”
<code>marker_height_mm</code>	float	<i>mm</i>	The height of the marker symbol
<code>marker_width_mm</code>	float	<i>mm</i>	The width of the marker symbol
<code>height_mm</code>	float	<i>mm</i>	The height of the object that the marker symbol is on.
<code>width_mm</code>	float	<i>mm</i>	The width of the object that the marker symbol is on.

Table 92:
`CustomWallDefinition`
JSON structure

41.4.2 Response

The `DefineCustomObjectResponse` type has the following fields:

Field	Type	Description
<code>status</code>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<code>success</code>	bool	True if the thumbnail was successfully retrieved; otherwise there was an error.

Table 93:
`DefineCustomObjectResponse`
JSON structure

41.5. DELETE CUSTOM OBJECTS

This command “causes the robot to forget about custom objects it currently knows about.” All custom objects that match the given pattern are removed.

Post: “/v1/delete_custom_objects”

41.5.1

Request

The DeleteCustomObjectsRequest type has the following fields:

Field	Type	Description
mode	CustomObjectDeletionMode	The kind of custom objects to remove.

Table 94:
DeleteCustomObjectsRequest JSON structure

The CustomObjectDeletionMode is used to specify which kinds of custom objects should be deleted from the internal database. The enumeration has the following named values:

Name	Value	Description
DELETION_MASK_UNKNOWN	0	
DELETION_MASK_FIXED_CUSTOM_OBJECTS	1	Delete the custom objects that are “fixed” (don't have any marker symbols).
DELETION_MASK_CUSTOM_MARKER_OBJECTS	2	Delete the objects with marker symbols.
DELETION_MASK_ARCHETYPES	3	Deletes everything but the fixed objects and their marker symbols.

Table 95:
CustomObjectDeletionMode Enumeration

41.5.2

Response

The DeleteCustomObjectsResponse type has the following fields:

Field	Type	Description
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 96:
DeleteCustomObjectsResponse JSON structure

42. ACTIONS AND BEHAVIOUR

Actions and “behaviors represent a complex task which requires Vector’s internal logic to [carry out]. This may include combinations of animation, path planning or other functionality.”

See also section 49 *Cube*, and section 55 *Interactions with Objects*, which covers actions/behaviors that involve interacting with objects and faces.

Actions often have tags (an arbitrary value given to it by the SDK application), and have result code. And action can be cancelled using this tag. Behaviors do not have tags.

Behaviors are submitted at the priority level associated with the connection. If the connection has released control, requested behaviors and actions are ignored. When control is requested, a priority level is requested by the SDK application at the time. Behaviors requested by Vector’s internal AI with a lower priority will be ignored; behaviors with a high priority will take control (causing the SDK to lose control). By giving up control, or changing the control priority the SDK can effectively cancel the behavior it requested.

Request control at the *RESERVE_CONTROL*_priority level “can be used to suppress the ordinary idle behaviors of the Robot and keep Vector still between SDK control instances. Care must be taken when blocking background behaviors, as this may make Vector appear non-responsive.”

See chapter 26 *Behaviors* for a description of behaviors and priorities.

42.1. ENUMERATIONS

42.1.1 ActionTagConstants

This is the range of numbers in which we can assign an identifier for the action so that we can cancel it later.

Name	Value	Description
<i>INVALID_SDK_TAG</i>	0	
<i>FIRST_SDK_TAG</i>	2000001	An assigned action tag must be equal to or greater than this value.
<i>LAST_SDK_TAG</i>	3000000	An assigned action tag must be less than or equal to this value.

42.1.2 BehaviorResults

The BehaviorResults is used TBD. The enumeration has the following named values:

Name	Value	Description
<i>BEHAVIOR_INVALID_STATE</i>	0	
<i>BEHAVIOR_COMPLETE_STATE</i>	1	
<i>BEHAVIOR_WONT_ACTIVATE_STATE</i>	2	

42.2. EVENTS

42.2.1 FeatureStatus

The FeatureStatus status event is sent as Vector's behavior focus changes. The structure has the following fields:

Field	Type	Description
<i>feature_name</i>	string	The current active behaviour (feature). See Appendix H, table <i>Table 527: The AI behaviour features</i> for a list and description.
<i>source</i>	string	Where the direction to do this behavior came from: "Voice", "App", "AI", "Unknown". Voice is for responses to voice commands and intents; "App" is for application submitted intents; AI is behaviors initiated by the high-level AI.

Table 99: FeatureStatus JSON structure

Note: for Vector-OS feature flags, see section *53 Features & Entitlements*.

42.2.2 StimulationInfo

The StimulationInfo event is used report events that impact Vector's emotion state and overall stimulation level. The structure has the following fields:

Field	Type	Units	Description
<i>accel</i>	float	<i>mm/sec</i> ²	The acceleration at the time of the stimulation.
<i>emotion_events</i>	string[]		The list of event names related to the emotion. The names of emotion events and their description can be found in Appendix J <i>Table 531: The emotion event names. Optional</i> .
<i>max_value</i>	float		The minimum stimulation value. Typically 1
<i>min_value</i>	float		The maximum stimulation value. Typically 0
<i>value</i>	float		The stimulation value after applying the events.
<i>value_before_event</i>	float		The stimulation value before the event(s). "matches value if there were no emotion events"
<i>velocity</i>	float	<i>mm/sec</i>	The speed at the time of the stimulation.

Table 100: StimulationInfo JSON structure

42.3. STRUCTURES

42.3.1 ActionResults

“The possible results of running an action.” The structure has the following fields:

Field	Type	Description
code	ActionResultCode	The results

The ActionResultCode is used to provide “the possible results of running an action.”

Name	Value	Description
ACTION_RESULT_SUCCESS	0	“Action completed successfully.”
ACTION_RESULT_RUNNING	16777216	“Action is still running.”
ACTION_RESULT_CANCELLED_WHILE_RUNNING	33554432	“Action was cancelled by SDK request”
NOT_STARTED	33554433	“Initial state of an Action to indicate it has not yet started.”
ABORT	50331648	“Action aborted itself (e.g. had invalid attributes, or a runtime failure).”
ANIM_ABORTED	50331649	“Animation Action aborted itself (e.g. there was an error playing the animation).”
BAD_MARKER	50331650	“There was an error related to vision markers.”
BAD_MESSAGE_TAG	50331651	“There was a problem related to a subscribed or unsupported message tag”
BAD_OBJECT	50331652	“There was a problem with the Object ID provided (e.g. there is no Object with that ID).”
BAD_POSE	50331653	“There was a problem with the Pose provided.”
BAD_TAG	50331654	“The SDK-provided tag was bad.”
CHARGER_UNPLUGGED_ABORT	50331655	“Vector is on the charger but cannot sense the contacts. Charger may be unplugged.”
CLIFF_ALIGN_FAILED_TIMEOUT	50331656	
CLIFF_ALIGN_FAILED_NO_TURNING	50331657	
CLIFF_ALIGN_FAILED_OVER_TURNING	50331658	
CLIFF_ALIGN_FAILED_NO_WHITE	50331659	
CLIFF_ALIGN_FAILED_STOPPED	50331660	
FAILED_SETTING_CALIBRATION	50331661	“Shouldn’t occur outside of factory.”
FOLLOWING_PATH_BUT_NOT_TRAVERSING	50331662	“There was an error following the planned path.”
INTERRUPTED	50331663	“The action was interrupted by another Action or Behavior.”
INVALID_OFF_TREADS_STATE	50331664	“The robot ended up in an “off treads state” not valid for this action (e.g. the robot was placed on its back while executing a turn).”
MISMATCHED_UP_AXIS	50331665	“The Up Axis of a carried object doesn’t match the desired placement pose.”

Table 101:
ActionResults JSON structure

Table 102:
ActionResultCode Enumeration

<i>NO_ANIM_NAME</i>	50331666	“No valid Animation name was found.”
<i>NO_DISTANCE_SET</i>	50331667	“An invalid distance value was given.”
<i>NO_FACE</i>	50331668	“There was a problem with the Face ID (e.g. Vector doesn't know where it is).”
<i>NO_GOAL_SET</i>	50331669	“No goal pose was set.”
<i>NO_PREACTION_POSES</i>	50331670	“No pre-action poses were found (e.g. could not get into position).”
<i>NOT_CARRYING_OBJECT_ABORT</i>	50331671	“No object is being carried, but the action requires one.”
<i>NOT_ON_CHARGER_ABORT</i>	50331672	“Vector is expected to be on the charger, but is not.”
<i>NULL_SUBACTION</i>	50331673	“No sub-action was provided.”
<i>PATH_PLANNING_FAILED_ABORT</i>	50331674	“Vector was unable to plan a path.”
<i>PICKUP_OBJECT_UNEXPECTEDLY_MOVING</i>	50331675	“The object that Vector is attempting to pickup is unexpectedly moving (e.g it is being moved by someone else).”
<i>SEND_MESSAGE_TO_ROBOT_FAILED</i>	50331676	“Shouldn't occur in SDK usage.”
<i>STILL_CARRYING_OBJECT</i>	50331677	“Vector is unexpectedly still carrying an object.”
<i>TIMEOUT</i>	50331678	“The Action timed out before completing correctly.”
<i>TRACKS_LOCKED</i>	50331679	“One or more movement tracks (Head, Lift, Body, Face, Backpack Lights, Audio) are already being used by another Action.”
<i>UNEXPECTED_DOCK_ACTION</i>	50331680	“There was an internal error related to an unexpected type of dock action.”
<i>UNKNOWN_TOOL_CODE</i>	50331681	“Shouldn't occur outside of factory.”
<i>UPDATE_DERIVED_FAILED</i>	50331682	“There was a problem in the subclass's update on the robot.”
<i>VISUAL_OBSERVATION_FAILED</i>	50331683	“Vector did not see the expected result (e.g. unable to see cube in the expected position after a related action).”
<i>SHOULDNT_DRIVE_ON_CHARGER</i>	50331684	“Action is not permitted on the charger.”
<i>RETRY</i>	67108864	“The Action failed, but may succeed if retried.”
<i>DID_NOT_REACH_PREACTION_POSE</i>	67108865	“Failed to get into position.”
<i>FAILED_TRAVERSING_PATH</i>	67108866	“Failed to follow the planned path.”
<i>LAST_PICK_AND_PLACE_FAILED</i>	67108867	“The previous attempt to pick and place an object failed.”
<i>MOTOR_STOPPED_MAKING_PROGRESS</i>	67108868	“The required motor isn't moving so the action cannot complete.”
<i>NOT_CARRYING_OBJECT_RETRY</i>	67108869	“Not carrying an object when it was expected, but may succeed if the action is retried.”
<i>NOT_ON_CHARGER_RETRY</i>	67108870	“Driving onto the charger failed, but may succeed if the action is retried.”
<i>PATH_PLANNING_FAILED_RETRY</i>	67108871	“Vector was unable to plan a path, but may succeed if the action is retried.”
<i>PLACEMENT_GOAL_NOT_FREE</i>	67108872	“There is no room to place the object at the desired

		destination.”
<i>PICKUP_OBJECT_UNEXPECTEDLY_NOT_MOVING</i>	67108873	“The object that Vector thought he was lifting didn’t start moving, so he must have missed.”
<i>STILL_ON_CHARGER</i>	67108874	“Vector failed to drive off the charger.”
<i>UNEXPECTED_PITCH_ANGLE</i>	67108875	“Vector’s pitch is at an unexpected angle for the Action.”

42.4. BEHAVIOR CONTROL AND ASSUME BEHAVIOR CONTROL

These commands are used to setup the ability to submit actions and behaviors into Vector's AI system. This control is needed "to be able to directly control Vector's motors, override his screen, play an animation, etc."

The request specifies a priority level. After control is granted, Vector's AI will suppress internal behaviors with a lower priority. When a behavior is commanded by the SDK, it will be associated with the priority level selected here. Note: the priority level is represented by a number where lower values represent higher priorities, and higher values represent lower priorities. See Chapter 26 for a detailed description of behavior priorities.

There are two entry points: `AssumeBehaviorControl` and `BehaviorControl`. Both employ the same request and response message structures. The response is a stream that includes information when the control was acquired, and lost.

Post: "/v1/assume_behavior_control"

42.4.1 Request

The `BehaviorControlRequest` is used to request control of Vector's behavior stream, and to release it. This structure includes one (and only one) of the following fields:

Field	Type	Description	
<code>control_release</code>	{}	This is used to when the application is releasing control back to Vector; the value is an empty dictionary.	
<code>control_request</code>	<code>ControlRequest</code>	This is used when the application is requesting control of Vector; see below for a description.	

Table 103:
`BehaviorControlRequest`
JSON structure

The `ControlRequest` is used to request control of the behavior system at a given priority. This structure has the following fields:

Field	Type	Description	
<code>priority</code>	<code>Priority</code>	This is the priority level that should be employed for requested behaviors; internal behaviors with a priority lower than this will be suppressed.	

Table 104:
`ControlRequest` JSON
structure

The `Priority` enumeration has the following named priority level values:

Name	Value	Description	
<code>UNKNOWN</code>	0	"Unknown priority. Used for versions that doesn't understand old priority levels."	
<code>OVERRIDE_BEHAVIORS</code>	10	"Highest priority level. Suppresses most automatic physical reactions, use with caution."	
<code>DEFAULT</code>	20	"Normal priority level. Directly under mandatory physical reactions."	
<code>RESERVE_CONTROL</code>	30	This priority level is "used to disable idle behaviors." It is intended to "enable long-running SDK control between script executions. Not.. for regular behavior control."	

Table 105: Priority level Enumeration

42.4.2

Response

The response is a stream of BehaviorControlResponse structures that includes information when the control was acquired, and lost. This structure includes one (and only one) of the following fields:

Field	Type	Units	Description
<code>control_granted_response</code>	{}		The application is now in control of the behavior stream and is “free to run any actions and behaviors they like. Until a ControlLostResponse is received, they are directly in control of Vector’s behavior system.”
<code>control_lost_event</code>	{}		“This informs the user that they lost control of the behavior system... to a higher priority behavior.” “This control can be regained through another” BehaviorControlRequest.
<code>keep_alive</code>	KeepAlivePing		“Used by Vector to verify the connection is still alive.”
<code>reserved_control_lost_event</code>	{}		The “behavior system lock has been lost to another connection.” “This control can be regained through another” BehaviorControlRequest. This is sent when the SDK is at RESERVE_CONTROL priority level.

Table 106:
BehaviorControlResponse
e JSON structure

42.5. CANCEL ACTION BY ID TAG

Cancel “a previously-requested action.”

Post: “/v1/cancel_action_by_id_tag”

42.5.1 Request

The CancelActionByldTagRequest structure has the following fields:

Field	Type	Description
<i>id_tag</i>	uint32	“Use the id_tag provided to the action request”

Table 107:
CancelActionByldTagRequest JSON structure

42.5.2 Response

The CancelActionByldTagResponse type has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 108:
CancelActionByldTagResponse JSON structure

42.6. LOOK AROUND IN PLACE

This has Vector turn around (in place) and see what is around him. See also section 52.3.5 *RobotObservedFace*, section 41.2.7 *RobotObservedObject*

Post: “/v1/look_around_in_place”

42.6.1 Request

The LookAroundInPlaceRequest structure has no fields.

42.6.2 Response

The LookAroundInPlaceResponse structure has the following fields:

Field	Type	Description
<i>result</i>	BehaviorResults	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 109:
LookAroundInPlaceResponse JSON structure

43. ALEXA

43.1. ENUMERATIONS

43.1.1 AlexaAuthState

The AlexaAuthState is used represent how far in the Alexa Voice Services authorization process Vector is. The enumeration has the following named values:

Name	Value	Description
ALEXA_AUTH_INVALID	0	“Invalid/error/versioning issue”
ALEXA_AUTH_UNINITIALIZED	1	“Not opted in, or opt-in attempted but failed”
ALEXA_AUTH_REQUESTING_AUTH	2	“Opted in, and attempting to authorize”
ALEXA_AUTH_WAITING_FOR_CODE	3	“Opted in, and waiting on the user to enter a code”
ALEXA_AUTH_AUTHORIZED	4	“Opted in, and authorized / in use”

Table 110:
AlexaAuthState
Enumeration

43.2. EVENTS

43.2.1 AlexaAuthEvent

The AlexaAuthEvent is used to post updates to SDK application (via the *Event* message) when the authorization with Alexa Voice Services change. The structure has the following fields:

Field	Type	Description
auth_state	AlexaAuthState	
extra	string	

Table 111:
AlexaAuthEvent JSON
structure

43.3. ALEXA AUTH STATE

This is used to find out whether Vector has been authenticated and authorized to use Alexa Voice Services.

Post: “/v1/alexa_auth_state”

43.3.1 Request

The AlexaAuthStateRequest structure has no fields.

43.3.2 Response

The AlexaAuthStateResponse structure has the following fields:

Field	Type	Description
<i>auth_state</i>	AlexaAuthState	
<i>extra</i>	string	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

43.4. ALEXA OPT IN

This is used to enable Alexa Voice Services on Vector.

Post: “/v1/alexa_opt_in”

43.4.1 Request

The AlexaOptInRequest structure has the following fields:

Field	Type	Description
<i>opt_in</i>	bool	True, if Vector should employ Alexa Voice services; otherwise Vector should not.

43.4.2 Response

The AlexaOptInResponse structure has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

44. ANIMATION

Some things related to animation but we haven't figured it all out yet.

44.1. STRUCTURES

44.1.1 Animation

This structure is used to provide the name of an animation. The Animation structure has the following fields:

Field	Type	Description
<i>name</i>	string	“The name of a given animation”

Table 115: Animation JSON structure

44.1.2 AnimationTrigger

This structure is used to provide the name of an animation group (aka its trigger name). The AnimationTrigger structure has the following fields:

Field	Type	Description
<i>name</i>	string	“The name of a given animation trigger”

Table 116: AnimationTrigger JSON structure

44.2. LIST ANIMATIONS

“Constructs and returns a list of animations.”

Post: “/v1/list_animations”

44.2.1 Request

The ListAnimationsRequest has no fields.

44.2.2 Response

The ListAnimationsResponse structure has the following fields:

Field	Type	Description
<i>animation_names</i>	Animation[]	“The animations that Vector knows..”
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 117: ListAnimationsResponse JSON structure

44.3. LIST ANIMATION TRIGGERS

“Constructs and returns a list of animation triggers.”

Post: “/v1/list_animation_triggers”

44.3.1 Request

The ListAnimationTriggersRequest has no fields.

44.3.2 Response

The ListAnimationTriggersResponse structure has the following fields:

Field	Type	Description
<i>animation_trigger_names</i>	AnimationTrigger[]	“The animations triggers that Vector knows.”
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 118:
ListAnimationTriggersResponse JSON structure

44.4. PLAY ANIMATION

“Requests that Vector play an animation.”

44.4.1 Request

The PlayAnimationRequest structure has the following fields:

Field	Type	Units	Description
<i>animation</i>	Animation		“The animation to play.”
<i>ignore_body_track</i>	bool		“Ignore any movement of Vector's body when playing the animation.”
<i>ignore_head_track</i>	bool		“Ignore any movement of Vector's head when playing the animation.”
<i>ignore_lift_track</i>	bool		“Ignore any movement of Vector's lift when playing the animation.”
<i>loops</i>	uint32		“The number of times to play the animation in a row.”

Table 119:
PlayAnimationRequest JSON structure

44.4.2 Response

The PlayAnimationResponse structure has the following fields:

Field	Type	Description
<i>animation</i>	Animation	“The animation that the robot executed.”
<i>result</i>	BehaviorResults	“Information on whether the animation played successfully.”
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 120:
PlayAnimationResponse JSON structure

44.5. PLAY ANIMATION TRIGGER

“Requests that Vector play an animation trigger.”

44.5.1 Request

The PlayAnimationTriggerRequest structure has the following fields:

Field	Type	Units	Description
<i>animation_trigger</i>	AnimationTrigger		“The animation trigger to play.”
<i>ignore_body_track</i>	bool		“Ignore any movement of Vector's body when playing the animation.”
<i>ignore_head_track</i>	bool		“Ignore any movement of Vector's head when playing the animation.”
<i>ignore_lift_track</i>	bool		“Ignore any movement of Vector's lift when playing the animation.”
<i>loops</i>	uint32		“The number of times to play the animation in a row.”
<i>use_lift_safe</i>	bool		“Automatically ignore the lift track if Vector is currently carrying an object.”

Table 121:
PlayAnimationTriggerRequest JSON structure

44.5.2 Response

See the response for Play Animation.

45. ATTENTION TRANSFER

Note: this attention event is unlikely to be sent and the response to getting the latest attention transfer is likely to invalid or empty, as the “AttentionTransfer” feature is disabled in all software releases.

45.1. EVENTS

45.1.1 AttentionTransfer

This event is sent when TBD. The AttentionTransfer structure has the following fields:

Field	Type	Description	Table 122: AttentionTransfer JSON structure
<i>reason</i>	AttentionTransferReason	The reason that the attention was changed.	
<i>seconds_ago</i>	float	How long ago the attention was changed.	

The AttentionTransferReason is used to represent why the attention was transferred. The enumeration has the following named values:

Name	Value	Description	Table 123: AttentionTransferReason Enumeration
<i>Invalid</i>	0		
<i>NoCloudConnection</i>	1		
<i>NoWifi</i>	2		
<i>UnmatchedIntent</i>	3		

45.2. GET LATEST ATTENTION TRANSFER

Part of the behaviour component

Post: “/v1/get_latest_attention_transfer”

45.2.1 Request

The GetLatestAttentionTransferRequest has no fields.

45.2.2 Response

The GetLatestAttentionTransferResponse has the following fields:

Field	Type	Description
<i>latest_attention_transfer</i>	LatestAttentionTransfer	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 124:
GetLatestAttentionTransferResponse JSON structure

The LatestAttentionTransfer structure has the following fields:

Field	Type	Description
<i>attention_transfer</i>	AttentionTransfer	When and why the attention was changed.

Table 125:
LatestAttentionTransfer JSON structure

46. AUDIO

This section describes events and commands related to Vectors audio input and output.

46.1. ENUMERATIONS

46.1.1 AudioProcessingMode

The `AudioProcessingMode` is used to represent the different ways that Vector can process the microphone audio. The enumeration has the following named values:

Name	Value	Description
<code>AUDIO_UNKNOWN</code>	0	“error value”
<code>AUDIO_OFF</code>	1	The audio settings from the HTTPS API will not be used.
<code>AUDIO_FAST_MODE</code>	2	The spatial audio processing is disabled; the sound is used from a single microphone. This has the lowest processing overhead.
<code>AUDIO_DIRECTIONAL_MODE</code>	3	Use “beamforming support for focusing on specific direction – [this] sounds cleanest”
<code>AUDIO_VOICE_DETECT_MODE</code>	4	Use “multi-microphone non-beamforming. [This is] best for voice detection programs.”

Table 126:
`AudioProcessingMode`
Enumeration

46.1.2 MasterVolumeLevel

The `MasterVolumeLevel` is used to control the volume of audio played by Vector, including text to speech. It is used in the `MasterVolumeLevelRequest`. The enumeration has the following named values:

Name	Value	Description
<code>VOLUME_LOW</code>	0	
<code>VOLUME_MEDIUM_LOW</code>	1	
<code>VOLUME_MEDIUM</code>	2	
<code>VOLUME_MEDIUM_HIGH</code>	3	
<code>VOLUME_HIGH</code>	4	

Table 127:
`MasterVolumeLevel`
Enumeration

46.1.3 UtteranceState

The UtteranceState is used to represent the state of audio playback by Vector, including text to speech. It is used in the SayTextResponse. The enumeration has the following named values:

Name	Value	Description
INVALID	0	
GENERATING	1	Vector is generating the audio and other animation for the text to speech.
READY	2	Vector has completed generating the audio and animation.
PLAYING	3	Vector is playing the speech and related animation.
FINISH	4	Vector has finished playing the audio and animation.

Table 128:
UtteranceState
Enumeration

46.2. EVENTS

The following events are sent in the *Event* message. When a person speaks the wake word, the *WakeWordBegin* event will be sent, followed by the *WakeWordEnd* event and possibly a *UserIntent* event.

46.2.1 AudioSendModeChanged

Note: *this event is not available*; it was defined in the API protocol, but never implemented and removed. It is reproduced here for information purposes; it may be in future releases.

This event is “sent when the robot changes the mode it’s processing and sending audio” in.

See Chapter 16, section 73.2 *Spatial audio processing* for more information

The event structure has the following fields:

Field	Type	Description
mode	AudioProcessingMode	The requested audio processing mode.

Table 129:
AudioSendModeChanged
JSON structure

46.2.2 UserIntent

The UserIntent event is sent by Vector when an intent is received (from the cloud), after a person has said the wake word and spoken. The UserIntent structure has the following fields:

Field	Type	Description
intent_id ²⁸	uint32	The identifier for the intent. See Appendix I Table 530: <i>Mapping of different intent names</i> for an enumeration.
json_data	string	The parameters as a JSON formatted string. This may be empty if there is not additional information.

Table 130: UserIntent
JSON structure

²⁸ The use of an enumeration rather than a string is unusual here, and seems limiting.

46.2.3

WakeWord

This event is sent when the wake word is heard, and then when the cloud response is received. The WakeWord structure has the following fields, only one is present at any time:

Field	Type	Description	Table 131: WakeWord JSON structure
<i>wake_word_begin</i>	WakeWordBegin	This is sent when the wake word is heard. The structure has no contents.	
<i>wake_word_end</i>	WakeWordEnd	This is sent when the response (and potential intent) is received from the cloud. This is sent before the <i>UserIntent</i> event (if any).	

The WakeWordEnd structure has the following fields:

Field	Type	Description	Table 132: WakeWordEnd JSON structure
<i>intent_heard</i>	bool	True if a sentence was recognized with an associated intent; false otherwise.	
<i>intent_json</i>	string	The intent and parameters as a JSON formatted string. This is empty if an intent was not heard (<i>intent_heard</i> will be false), or if the client does not have control. In the later case, a <i>UserIntent</i> event with the intent JSON data will be sent.	

46.3. APP INTENT

This command allows the mobile application or SDK application to send an intent to Vector.

See also section 46.2.2 *UserIntent*, and section 46.2.3 *WakeWord*

Post: “/v1/app_intent”

46.3.1 Request

The AppIntentRequest structure has the following fields:

Field	Type	Description
<i>intent</i>	string	The name of the intent to request; Vector (probably) will only honor the intents listed in the “App Intent” column in Appendix I, <i>Table 530: Mapping of different intent names</i>
<i>param</i>	string	The parameters for the intent. This is usually a JSON formatted string. This can be empty if the intent does not require any additional information.

Table 133:
AppIntentRequest
JSON structure

This intent_meet_victor intent has the parameter following fields:

Field	Type	Units	Description
<i>param</i>			

Table 134:
intent_meet_victor
parameters

The intent_clock_settimer intent parameter isn’t used. Instead the length of the param is used as the number of seconds to set the timer for.

46.3.2 Response

The AppIntentResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 135:
AppIntentResponse
JSON structure

46.4. AUDIO FEED (FROM THE MICROPHONES)

Note: *this command is not available*; it was defined in the API protocol, but never implemented and removed. It is reproduced here for information purposes; it may be in future releases.

This command is used to request an audio feed from Vector.

See Chapter 16, section 73.2 *Spatial audio processing* for more information

Post: “/v1/audio_feed”

46.4.1 Request

This AudioFeedRequest has no fields.

46.4.2 Response

The response is a stream of the following AudioFeedResponse structure. This structure has the following fields:

Field	Type	Description	Table 136: AudioFeedResponse JSON structure
<i>direction_strengths</i>	bytes	“Histogram data of which directions this audio chunk came from.”	
<i>group_id</i>	uint32	“The index of this audio feed response”	
<i>noise_floor_power</i>	uint32	The background noise level, as a “power value, convert to db with $\log_{10}(\text{value})$ ”	
<i>robot_time_stamp</i>	uint32	The “robot time at the transmission of this audio sample group”	
<i>signal_power</i>	bytes	The stream of sound that Vector hears, as a “mono audio amplitude samples”. This is probably 1600 “16-bit little-endian PCM audio” samples.	
<i>source_confidence</i>	uint32	The “accuracy of the calculated source_direction”	
<i>source_direction</i>	uint32	0-11: The index of the direction that the voice or key sound is coming. 12: There is no identifiable sound or the direction cannot be determined.	

46.5. AUDIO PROCESSING MODE

Note: *this command is not available*; it was defined in the API protocol, but never implemented and removed. It is reproduced here for information purposes; it may be in future releases.

This command is used to “request how the robot should process and send audio.” Specifically it can turn off the audio processing, and enable or disable the spatial audio processing.

See Chapter 16, section 73.2 *Spatial audio processing* for more information

46.5.1 Request

This AudioSendModeRequest has the following fields:

Field	Type	Description	Table 137: AudioSendModeRequest JSON structure
<i>mode</i>	AudioProcessingMode	The requested audio processing mode.	

46.5.2 Response

There is no response.

46.6. EXTERNAL AUDIO STREAM PLAYBACK

This command is used to stream sound files to Vector play on his speaker. The audio is sent as single channel (mono) 16-bit little-endian PCM (i.e. without compression or other format) at a sample rate between 8000 samples/sec to 16205 samples/sec.

The audio is done by

1. Setting up the audio playback, by sending the “audio_stream_prepare” substructure with the audio rate and value
2. Sending the audio data in chunks (up to 1024 bytes, or 512 samples) using the “audio_stream_chunk” structure
3. repeating #2 until all of the sound data has been sent
4. Sending the “audio_stream_complete” or “audio_stream_cancel” to end the playback.

46.6.1 Request

The ExternalAudioStreamRequest is used to stream a chunk of audio to Vector. This structure has one (and only one) of the following fields:

Field	Type	Description	
<i>audio_stream_cancel</i>	{}	“Cancel a playing external robot audio stream”	
<i>audio_stream_chunk</i>	ExternalAudioStreamChunk	“Send chunk of audio data to stream on robot.”	
<i>audio_stream_complete</i>	{}	“Send notification of last chunk of audio sent to robot”	
<i>audio_stream_prepare</i>	ExternalAudioStreamPrepare	This is used to set up the audio channel, with the sample rate and playback volume.	

The ExternalAudioStreamPrepare structure has following fields:

Field	Type	Description	
<i>audio_frame_rate</i>	uint32	The sample rate for the audio. This must be in the range of 8000 to 16025 samples/sec.	
<i>audio_volume</i>	uint32	The volume to play the audio at. 0-100	

The ExternalAudioStreamChunk structure has following fields:

Field	Type	Description	
<i>audio_chunk_samples</i>	byte[]	The audio samples, encoded as 16-bit values in little-endian order. This must be 1024 or few bytes	
<i>audio_chunk_size_bytes</i> ²⁹	uint32	The number of bytes sent; the max is 1024 (i.e. a max of 512 samples).	

Table 138:
ExternalAudioStreamRequest JSON structure

Table 139:
ExternalAudioStreamPrepare JSON structure

Table 140:
ExternalAudioStreamChunk JSON structure

²⁹ I am curious. Why does this field exist? The array intrinsically knows its size...

46.6.2

Response

The ExternalAudioStreamResponse is provides the response to streamed a audio chunk. This structure has one (and only one) of the following fields:

Field	Type	Description	Table 141: ExternalAudioStreamResponse JSON structure
<code>audio_stream_playback_complete</code>	{}	“Audio has been played on the Robot”	
<code>audio_stream_playback_failyer</code> ³⁰	{}	There was an error playing the audio.	
<code>audio_stream_buffer_overrun</code>	ExternalAudioStreamBufferOverrun	“Audio has been sent to robot that would overrun the memory buffer”	

The ExternalAudioStreamBufferOverrun structure has following fields:

Field	Type	Description	Table 142: ExternalAudioStreamBufferOverrun JSON structure
<code>audio_samples_played</code>	uint32	The number of samples that were played.	
<code>audio_samples_sent</code>	uint32	The number of audio samples that were sent [To Vector? To the audio subsystem?]	

³⁰ Yes, that mis-spelling is correct

46.7. MASTER VOLUME

This command is used to set the volume of Vector's audio playback and sound effects.

46.7.1 Request

The *MasterVolumeRequest* has the following fields:

Field	Type	Description
<i>volume_level</i>	MasterVolumeLevel	This is used to set the volume of Vector's audio playback.

Table 143:
MasterVolumeRequest
JSON structure

46.7.2 Response

The *MasterVolumeResponse* has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 144:
MasterVolumeResponse
JSON structure

46.8. SAY TEXT

This command is used to request the state of Vector speak the given text.

Post: “/v1/say_text”

46.8.1

Request

The SayTextRequest structure has the following fields:

Field	Type	Units	Description
<i>duration_scalar</i>	float	<i>ratio</i>	This controls the speed at which Vector speaks. 1.0 is normal rate, less than 1 increases the speed (e.g. 0.8 causes Vector to speak in just 80% of the usual time), and a value larger than one slows the speed (e.g. 1.2 causes Vector to take 120% of the usual time to speak). Default: 1.0
<i>text</i>	string		The text (the words) that Vector should say.
<i>use_vector_voice</i>	bool		True if the text should be spoken in “Vector’s robot voice; otherwise, he uses a generic human male voice.”

Table 145:
SayTextRequest JSON structure

46.8.2

Response

The SayTextResponse structure has the following fields:

Field	Type	Description
<i>state</i>	UtteranceState	Where in the speaking process Vector is currently.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 146:
SayTextResponse JSON structure

TBD: are multiple responses sent as the task progresses?

47. BATTERY

See section 40.1.2 *RobotStatus* for a flag indicating that Vector is charging.

See section 55 *Interactions with Objects* for actions to drive onto and off of the charger.

47.1. ENUMERATIONS

The *BatteryLevel* enumeration is located in Chapter 8, Power Management, *Table 20: BatteryLevel codes as they apply to Vector*

47.2. BATTERY STATE

This command is used to request the state of Vector’s battery and the cube battery. The state includes its voltage, and whether Vector is charging.

Post: “/v1/battery_state”

47.2.1 Request

No parameters

47.2.2 Response

The *BatteryStateResponse* structure has the following fields:

Table 147:
BatteryStateResponse
JSON structure

Field	Type	Units	Description
<i>battery_level</i>	BatteryLevel		The interpretation of the battery level.
<i>battery_volts</i>	float	volts	The battery voltage.
<i>cube_battery</i>	CubeBatteryLevel		The status of the companion Cube’s battery.
<i>is_on_charger_platform</i>	bool		True if Vector is on his “home,” aka charger.
<i>is_charging</i>	bool		True if Vector is charging, false otherwise.
<i>status</i>	ResponseStatus		A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<i>suggested_charger_sec</i>	float	seconds	Suggested amount of time to charge.

48. CONNECTION

This section describes the events and commands used to establish and maintain a connection with Vector. This includes the ability to get the versions of the connection protocol, and the software used.

48.1. EVENTS

48.1.1 ConnectionResponse

The ConnectionResponse structure has the following fields:

Field	Type	Description	Table 148: ConnectionResponse JSON structure
<i>is_primary</i>	bool		
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

48.1.2

Event

The Event structure is to deliver messages that some event has occurred. It is received in periodic response to the part of the *Event Stream* command. All the events are carried in this one has one (and only) of the following fields:

Field	Type	Description
<code>alexa_auth_event</code>	AlexaAuthEvent	
<code>attention_transfer</code>	AttentionTransfer	<i>Not implemented?</i>
<code>check_update_status_response</code>	CheckUpdateStatusResponse	This event is sent when the update status has changed.
<code>connection_response</code>	ConnectionResponse	
<code>cube_battery</code>	CubeBattery	This event is sent when the cube's battery level has changed.
<code>jdocs_changed</code>	JdocsChanged	This event is sent when Vector's preference settings have changed.
<code>keep_alive</code>	KeepAlivePing	"Used by Vector to verify the connection is still alive."
<code>mirror_mode_disabled</code>	MirrorModeDisabled	This event is sent when the display system has disabled mirror mode.
<code>object_event</code>	ObjectEvent	This event is sent when an object is seen, tapped, lost, moved, a connection was established or lost.
<code>onboarding</code>	Onboarding	
<code>photo_taken</code>	PhotoTaken	This event is sent when a photograph has been taken.
<code>robot_state</code>	RobotState	This event is regularly sent to give the status of the robot.
<code>robot_changed_observed_face_id</code>	RobotChangedObservedFaceID	This event is sent when Vector recognizes a face.
<code>robot_observed_face</code>	RobotObservedFace	This event is sent when a face is seen.
<code>robot_observed_motion</code>	RobotObservedMotion	This event is sent when some visual motion is seen. <i>Note: this event is not available; it has been removed, but is included here for clarity</i>
<code>stimulation_info</code>	StimulationInfo	This event is sent when stimulation from internal or external events is received.
<code>time_stamped_status</code>	TimeStampedStatus	
<code>user_intent</code>	UserIntent	This event is sent when a user intent has been received and is being acted upon.
<code>vision_modes_auto_disabled</code>	VisionModesAutoDisabled	This event is sent when the vision system has disabled further updates.
<code>wake_word</code>	WakeWord	This event is sent when the wake word has been heard.

Table 149: Event JSON structure

48.1.3

KeepAlivePing

This is "a null message used by streams to verify that the client is still connected." This message has no fields.

48.1.4

TimeStampedStatus

The TimeStampedStatus structure has the following fields:

Field	Type	Description
<i>status</i>	Status	
<i>timestamp_utc</i>	uint32	The time that the status occurred on. The format is unix time: seconds since 1970, in UTC.

Table 150:
TimeStampedStatus
JSON structure

The Status structure has one (and only one) of the following fields:

Field	Type	Description
<i>face_enrollment_completed</i>	FaceEnrollmentComplete	
<i>feature_status</i>	FeatureStatus	This event is sent when the high-level AI changes Vector's behavior.
<i>meet_victor_face_scan_complete</i>	Meet Victor Face Scan Complete	
<i>meet_victor_face_scan_started</i>	Meet Victor Face Scan Started	

Table 151: Status JSON structure

48.2. EVENT STREAM

This command is used to request a stream of events from Vector.

Post: “/v1/event_stream”

Get: “/v1/event_stream”

48.2.1 Request

The EventRequest has the following fields:

Field	Type	Description
<i>black_list</i>	FilterList	The list of events to not include. ?
<i>connection_id</i>	string	
<i>white_list</i>	FilterList	The list of events to include.

Table 152:
EventRequest JSON
structure

The FilterList structure has the following fields:

Field	Type	Description
<i>list</i>	string[]	A list of events

Table 153: FilterList
JSON structure

48.2.2 Response

The response is a stream of EventResponse structures. These have the following fields:

Field	Type	Description
<i>event</i>	Event	The event that occurred. This structure is described above in the subsection <i>Events</i>
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 154:
EventResponse JSON
structure

48.3. PROTOCOL VERSION

“Checks the supported protocol version by passing in the client version and minimum host version and receiving a response to see whether the versions are supported.”

Post: “/v1/protocol_version”

“The valid versions of the protocol. Protocol versions are updated when messages change significantly: new ones are added and removed, fields deprecated, etc. The goal is to support as many old versions as possible, only bumping the minimum when there is no way to handle a prior version.”

48.3.1 Request

The ProtocolVersionRequest has the following fields:

Field	Type	Description	Table 155: ProtocolVersionRequest JSON structure
<i>client_version</i>	int64	The version of the protocol that the client is using.	
<i>min_host_version</i>	int64	The minimum version level of the protocol that robot should support.	

48.3.2 Response

The ProtocolVersionResponse has the following fields:

Field	Type	Description	Table 156: ProtocolVersionResponse JSON structure
<i>host_version</i>	int64	The version of the protocol that the robot supports.	
<i>result</i>	Result	Whether or not the protocol version supported by the robot is compatible with the client. See below.	

The Result is used to indicate whether the client version is supported. The enumeration has the following named values:

Name	Value	Description	Table 157: Result Enumeration
<i>SUPPORTED</i>	1	The protocol supports the client version and the minimum host (robot) version of the protocol.	
<i>UNSUPPORTED</i>	0	The protocol is unable to support the client; either the client version is not supported, or the host is unable to support a compatible version of the protocol.	

48.4. SDK INITIALIZATION

“SDK-only message to pass version info for device OS, Python version, etc.”

Post: “/v1/sdk_initialization”

48.4.1 Request

The `SDKInitializationRequest` has the following fields:

Field	Type	Description
<code>cpu_version</code>	string	The CPU model that the client (SDK) is using; <i>informational only</i> .
<code>os_version</code>	string	The version of operating system that the client (SDK) is using; <i>informational only</i> .
<code>python_implementation</code>	string	
<code>python_version</code>	string	The version of python that the client (SDK) is using. <i>Informational only</i> .
<code>sdk_module_version</code>	string	The version of the SDK software that the client is using.

Table 158:
`SDKInitializationRequest` JSON structure

48.4.2 Response

The `SDKInitializationResponse` type has the following fields:

Field	Type	Description
<code>status</code>	<code>ResponseStatus</code>	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 159:
`SDKInitializationResponse` JSON structure

48.5. USER AUTHENTICATION

This command is used to authenticate

Post: “/v1/user_authentication”

48.5.1 Request

The UserAuthenticationRequest has the following fields:

Field	Type	Description
<i>client_name</i>	bytes	
<i>user_session_id</i>	bytes	

Table 160:
UserAuthenticationRequest JSON structure

48.5.2 Response

The UserAuthenticationResponse has the following fields:

Field	Type	Description
<i>client_token_guid</i>	bytes	The token bytes to be included in subsequent HTTPS postings. This token should be saved for future use.
<i>code</i>	Code	The result of the authentication request
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 161:
UserAuthenticationResponse JSON structure

The Code enumeration is:

Name	Value	Description
<i>UNAUTHORIZED</i>	0	
<i>AUTHORIZED</i>	1	

Table 162: Code Enumeration

48.6. VERSION STATE

Retrieves Vector's version information.

Post: “/v1/version_state”

48.6.1 Request

The VersionStateRequest has no fields.

48.6.2 Response

The VersionStateResponse type has the following fields:

Field	Type	Description	Table 163: VersionStateResponse JSON structure
<i>engine_build_id</i>	string	The robot's software build identifier.	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	
<i>os_version</i>	string	The identifier of the robot's software version.	

49. CUBE

This section describes the structures and commands to interact with the cube.

Comment: Many of the commands are specific to interacting with a cube, but appear to have been intended to be generalized to work with a wider range of objects.

See also section [41.4 Define Custom Object](#) for a description how to create custom box and cube objects.

The cube's unique identifier is called “factory_id” in these messages.

49.1. ENUMERATIONS

49.1.1 AlignmentType

The AlignmentType is used to indicate how Vector should align with the object. The enumeration has the following named values:

Name	Value	Description
ALIGNMENT_TYPE_UNKNOWN	0	
ALIGNMENT_TYPE_LIFT_FINGER	1	“Align the tips of the lift fingers with the target object”
ALIGNMENT_TYPE_LIFT_PLATE	2	“Align the flat part of the lift with the object (useful for getting the fingers in the cube's grooves)”
ALIGNMENT_TYPE_BODY	3	“Align the front of Vector's body (useful for when the lift is up)”
ALIGNMENT_TYPE_CUSTOM	4	“For use with distanceFromMarker parameter”

49.1.2 CubeBatteryLevel

The CubeBatteryLevel enumeration is used to categorize the condition of the Cube battery:

Name	Value	Description
BATTERY_LEVEL_LOW	0	The Cube battery is 1.1V or less.
BATTERY_LEVEL_NORMAL	1	The Cube battery is at normal operating levels, i.e. >1.1v

³¹ The levels are from robot.py

49.2. EVENTS

49.2.1 CubeBattery

The CubeBattery structure has the following fields:

Field	Type	Units	Description
<i>battery_volts</i>	float	<i>volts</i>	The battery voltage.
<i>factory_id</i>	string		The text string reported by the cube via Bluetooth LE.
<i>level</i>		CubeBatteryLevel	The interpretation of the battery level.
<i>time_since_last_reading_sec</i>	float	<i>seconds</i>	The number of seconds that have elapsed since the last Bluetooth LE message from the cube with a battery level measure.

Table 166: CubeBattery JSON structure

49.2.2 CubeConnectionLost

“Indicates that the connection subscribed through ConnectCube has been lost.”

See also *ObjectConnectionState*

The ConnectCubeRequest has no fields.

49.2.3 ObjectTapped

The ObjectTapped event is sent (see *ObjectEvent*) when an object has received a finger-tap. This event is only sent by the cube. Note: this event can have false triggers; it may sent when Vector is picking up, carrying, or putting down the Cube.

The structure has the following fields:

Field	Type	Units	Description
<i>object_id</i>	uint32		The identifier of the object tapped.
<i>timestamp</i>	uint32		The time that the event occurred on. The format is milliseconds since Vector’s epoch.

Table 167:
ObjectTapped JSON structure

49.3. CONNECT CUBE

“Attempt to connect to a cube. If a cube is currently connected, this will do nothing.”

Post: “/v1/disconnect_cube”

49.3.1 Request

The ConnectCubeRequest has no fields.

49.3.2 Response

The ConnectCubeResponse type has the following fields:

Field	Type	Description
<i>factory_id</i>	string	The identifier for the cube. This is built into the cube.
<i>object_id</i>	uint32	The identifier of the cube that we connected with. This is Vector’s internal identifier, and only the preferred cube is assigned one.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<i>success</i>	bool	True if Vector was able to successfully connect, via Bluetooth LE, with the cube.

Table 168:
ConnectCubeResponse
JSON structure

49.4. CUBES AVAILABLE

Have Vector scan for cubes via Bluetooth LE and report the ones heard.

Post: “/v1/cubes_available”

49.4.1 Request

The CubesAvailableRequest has no fields.

49.4.2 Response

The CubesAvailableResponse is sent to indicate whether the action successfully completed or not.
This structure has the following fields:

Field	Type	Description
<i>factory_ids</i>	string[]	A list of the cubes that were seen via Bluetooth LE. The cubes internal identifier (it’s factor id) is sent.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 169:
CubesAvailableResponse
JSON structure

49.5. DISCONNECT CUBE

“Requests a disconnection from the currently connected cube.”

Post: “/v1/disconnect_cube”

49.5.1 Request

The DisconnectCubeRequest has no fields.

49.5.2 Response

The DisconnectCubeResponse is sent to indicate whether the action successfully completed or not.

This structure has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 170:
DisconnectCubeResponse JSON structure

49.6. DOCK WITH CUBE

“Tells Vector to dock with a light cube with [an optional] given approach angle and distance.”
“While docking with the cube, Vector will use path planning.”

This action requires the use of the wheels (tracks). “Actions that use the wheels cannot be performed at the same time; otherwise you may see a TRACKS_LOCKED error.”

Post: “/v1/dock_with_cube”

49.6.1

Request

The DockWithCubeRequest structure has the following fields:

Field	Type	Units	Description
<i>alignment_type</i>	AlignmentType		“Which part of the robot to align with the object.”
<i>approach_angle_rad</i>	float	radians	“The angle to approach the cube from. For example, 180 degrees will cause Vector to drive past the cube and approach it from behind.”
<i>distance_from_marker_mm</i>	float	mm	“The distance from the object to stop. This is the distance between the origins.” 0mm to dock.
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional</i> .
<i>motion_prof</i>	PathMotionProfile		Modifies how Vector should approach the cube. <i>Optional</i> .
<i>num_retries</i>	int32		Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.
<i>object_id</i>	int32		The identifier of the object to dock with.
<i>use_approach_angle</i>	bool		If true, Vector will approach the cube from the given approach angle; otherwise Vector will approach from the most convenient angle.
<i>use_pre_dock_pose</i>	bool		“Whether or not to try to immediately [dock with the] object or first position the robot next to the object.” Recommended to set this to the same as <i>use_approach_angle</i> .

Table 171:
DockWithCubeRequest
JSON structure

49.6.2

Response

The DockWithCubeResponse is sent to indicate whether the action successfully completed or not.

This structure has the following fields:

Field	Type	Description
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 172:
DockWithCubeResponse
JSON structure

49.7. FLASH CUBE LIGHTS

“Plays the default cube connection animation on the currently connected cube's lights.”

Note: “This [command] is intended for app level user surfacing of cube connectivity, not for SDK cube light control.”

Post: “/v1/flash_cube_lights”

49.7.1 Request

The FlashCubeLightsRequest has no fields.

49.7.2 Response

The FlashCubeLightsResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 173:
FlashCubeLightsResponse JSON structure

49.8. FORGET PREFERRED CUBE

“Forget the robot's preferred cube. This will cause the robot to connect to the cube with the highest RSSI (signal strength) next time a connection is requested. Saves this preference to disk. The next cube that the robot connects to will become its preferred cube.”

See also section [49.15 Set Preferred Cube](#)

Post: “/v1/forget_preferred_cube”

49.8.1 Request

The ForgetPreferredCubeRequest has no fields.

49.8.2 Response

The ForgetPreferredCubeResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 174:
ForgetPreferredCubeResponse JSON structure

49.9. PICKUP OBJECT

“Instruct the robot to pick up the supplied object.” “While picking up the cube, Vector will use path planning.”

“Note that actions that use the wheels cannot be performed at the same time, otherwise you may see a TRACKS_LOCKED error.”

49.9.1

Request

The PickupObjectRequest structure has the following fields:

Field	Type	Units	Description
<i>approach_angle_rad</i>	float	radians	“The angle to approach the cube from. For example, 180 degrees will cause Vector to drive past the cube and approach it from behind.”
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>
<i>motion_prof</i>	PathMotionProfile		<i>Optional.</i>
<i>num_retries</i>	int32		Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.
<i>object_id</i>	int32		The identifier of the object to pick up. ‘Negative value means currently selected object’
<i>use_approach_angle</i>	bool		If true, Vector will approach the cube from the given approach angle; otherwise Vector will approach from the most convenient angle.
<i>use_pre_dock_pose</i>	bool		“Whether or not to try to immediately pick up an object or first position the robot next to the object.”

Table 175:
PickupObjectRequest
JSON structure

49.9.2

Response

The PickupObjectResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 176:
PickupObjectResponse
JSON structure

49.10. PLACE OBJECT ON GROUND HERE

“Ask Vector to place the object he is carrying on the ground at the current location.”

49.10.1 Request

The PlaceObjectOnGroundRequest structure has the following fields:

Field	Type	Units	Description
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>
<i>num_retries</i>	int32		Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.

Table 177:
PlaceObjectOnGroundRequest JSON structure

49.10.2 Response

The PlaceObjectOnGroundResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 178:
PlaceObjectOnGroundResponse JSON structure

49.11. POP A WHEELIE

“Tell Vector to ‘pop a wheelie’ using his cube.” Vector will approach the cube, then “push down on [it] with [his] lift, to start the wheelie.”

49.11.1 Request

The PopAWheelieRequest structure has the following fields:

Field	Type	Units	Description
<i>approach_angle_rad</i>	float	radians	“The angle to approach the cube from. For example, 180 degrees will cause Vector to drive past the cube and approach it from behind.”
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>
<i>motion_prof</i>	PathMotionProfile		<i>zOptional.</i>
<i>num_retries</i>	int32		Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.
<i>object_id</i>	int32		The identifier of the object to used to pop a wheelie. Negative value means currently selected object’
<i>use_approach_angle</i>	bool		If true, Vector will approach the cube from the given approach angle; otherwise Vector will approach from the most convenient angle.
<i>use_pre_dock_pose</i>	bool		“Whether or not to try to immediately [use the] object or first position the robot next to the object.” Recommended to set this to the same as <i>use_approach_angle</i> .

Table 179:
PopAWheelieRequest
JSON structure

49.11.2 Response

The PopAWheelieResponse is sent to indicate whether the action successfully completed or not.

This structure has the following fields:

Field	Type	Description
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 180:
PopAWheelieResponse
JSON structure

49.12. ROLL BLOCK

“Make Vector roll his block, regardless of relative position and orientation.” This triggers a behaviour, where Vector will look for his block, then “move into position as necessary based on relative distance and orientation.”

See also section [49.13 Roll Object](#)

Post: “/v1/roll_block”

49.12.1

Request

The RollBlockRequest has no fields.

49.12.2

Response

The RollBlockResponse structure has the following fields:

Field	Type	Description
<i>result</i>	BehaviorResults	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 181:
RollBlockResponse
JSON structure

49.13. ROLL OBJECT

“Tell Vector to roll his cube.” This triggers an action.

49.13.1 Request

The RollObjectRequest structure has the following fields:

Field	Type	Units	Description	Table 182: RollObjectRequest JSON structure
<i>approach_angle_rad</i>	float	radians	“The angle to approach the cube from. For example, 180 degrees will cause Vector to drive past the cube and approach it from behind.”	
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>	
<i>motion_prof</i>	PathMotionProfile		<i>Optional.</i>	
<i>num_retries</i>	int32		Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.	
<i>object_id</i>	int32		The identifier of the object to roll. ‘Negative value means currently selected object’	
<i>use_approach_angle</i>	bool		If true, Vector will approach the cube from the given approach angle; otherwise Vector will approach from the most convenient angle.	
<i>use_pre_dock_pose</i>	bool		“Whether or not to try to immediately [roll the] object or first position the robot next to the object.” Recommended to set this to the same as <i>use_approach_angle</i> .	

49.13.2 Response

The RollObjectResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description	Table 183: RollObjectResponse JSON structure
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

49.14. SET CUBE LIGHTS

“Set each of the lights on the currently connected cube based on two RGB values each and timing data for how to transition between them.”

“Sets each LED on [Vector]’s cube. Two states are specified designated ‘on’ and ‘off’, each with a color, duration, and state transition time.”

See also the Chapter 21 section *100 Cube lights Animation*

49.14.1

Request

The SetCubeLightsRequest event is used to specify the light pattern on the cube. The structure has the following fields:

Table 184:
SetCubeLightsRequest
JSON structure

Field	Type	Units	Description
<i>object_id</i>	uint32		The internal id for the cube.
<i>make_relative</i>	MakeRelativeMode		<i>Should be off(1)</i>
<i>off_color</i>	array of uint32[]		Each color corresponds to each of the 4 cube lights. Each color is represented as four values (red, green, blue, and alpha), in the range of 0..255.
<i>off_period_ms</i>	uint32[] ms		The “off” duration for each of the 4 cube lights. This is the duration to show each cube light in its corresponding “off” color (in <i>off_color</i>).
<i>offset</i>	int32[4]		<i>recommended: set four 0’s.</i>
<i>on_color</i>	array of uint32[]		Each color corresponds to each of the 4 cube lights. Each color is represented as four values (red, green, blue, and alpha), in the range of 0..255.
<i>on_period_ms</i>	uint32[] ms		The “on” duration for each of the 4 cube lights. This is the duration to show each cube light in its corresponding “on” color (in <i>onColors</i>).
<i>relative_to_x</i>	float		Should be 0.0
<i>relative_to_y</i>	float		Should be 0.0
<i>rotate</i>	boolean		? Possibly to have the colors be assigned to the next clockwise (or counterclockwise) light periodically? Should be <i>false</i>
<i>transition_off_period_ms</i>	uint32[] ms		The time (in ms) to transition from the on color to the off color.
<i>transition_on_period_ms</i>	uint32[] ms		The time (in ms) to transition from the off color to the on color

The `MakeRelativeMode` is used to indicate how Vector should align with the object. The enumeration has the following named values:

Name	Value	Description
<code>UNKNOWN</code>	0	
<code>OFF</code>	1	
<code>BY_CORNER</code>	2	
<code>BY_SIDE</code>	3	

Table 185:
`MakeRelativeMode`
Enumeration

49.14.2 Response

The `SetCubeLightsResponse` is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<code>status</code>	<code>ResponseStatus</code>	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 186:
`SetCubeLightsResponse`
JSON structure

49.15. SET PREFERRED CUBE

“Set the robot’s preferred cube and save it to disk. The robot will always attempt to connect to this cube if it is available. This is only used in simulation ~~for now~~.”

Post: “/v1/set_preferred_cube”

49.15.1 Request

The `SetPreferredCubeRequest` structure has the following fields:

Field	Type	Units	Description
<code>factory_id</code>	string		The identifier of the cube to use. This is built into the cube.

Table 187:
`SetPreferredCubeRequest`
JSON structure

49.15.2 Response

The `SetPreferredCubeResponse` is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<code>status</code>	<code>ResponseStatus</code>	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 188:
`SetPreferredCubeResponse`
JSON structure

50. DIAGNOSTICS

This section include commands intended to help diagnose trouble: checking the connection with the cloud servers; and uploading logs from Vector to help diagnose his problems.

50.1. CHECK CLOUD CONNECTION

This command is used to check the connection with the remote servers.

Post: “/v1/check_cloud_connection”

50.1.1 Request

The CheckCloudRequest has no fields.

50.1.2 Response

The CheckCloudResponse has the following fields:

Table 189:
CheckCloudResponse
JSON structure

Field	Type	Description
code	ConnectionCode	Whether the cloud is available, or the relevant connection error
expected_packets	int32	The number of packets expected to have been exchanged with the cloud server.
num_packets	int32	The number of packets actually exchanged with the cloud server.
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
status_message	string	

The ConnectionCode is used to indicate whether the cloud is available. It is used in the response to the CheckCloudConnectionRequest command. The ConnectionCode enumeration has the following named values:

Table 190:
ConnectionCode
Enumeration

Name	Value	Description
AVAILABLE	1	The cloud is connected, and has authenticated successfully
BAD_CONNECTIVITY	2	The internet or servers are down
FAILED_AUTH	4	The cloud connection has failed due to an authentication issue
FAILED_TLS	3	The cloud connection has failed due to [TLS certificate?] issue
UNKNOWN	0	There is an error connecting to the cloud, but the reason is unknown

50.2. UPLOAD DEBUG LOGS

TBD: Request that the logs be uploaded to the server for analysis.

Post: “/v1/upload_debug_logs”

50.2.1 Request

The UploadDebugLogsRequest structure has no fields.

50.2.2 Response

The UploadDebugLogsResponse structure has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<i>url</i>	string	

Table 191:
UploadDebugLogsResponse JSON structure

51. DISPLAY

This section describes commands that are used to display imagery on Vector's LCD.

51.1. EVENTS

51.1.1 MirrorModeDisabled

The `MirrorModeDisabled` event is sent (see *Event*) “if `MirrorMode` (camera feed displayed on face) is currently enabled but is automatically being disabled.”

The `MirrorModeDisabled` structure has no fields.

51.2. DISPLAY IMAGE RGB

“Sets screen (Vector's face) to” display the passed image.

Post: “/v1/display_face_image_rgb”

51.2.1 Request

The `DisplayFacelImageRGBRequest` structure has the following fields:

Field	Type	Units	Description
<code>duration_ms</code>	<code>uint32</code>	<code>ms</code>	“How long to display the image on the face.”
<code>face_data</code>	<code>bytes</code>		The raw data for the image to display. The LCD is 184x56, with RGB565 pixels (16 bits/pixel).
<code>interrupt_running</code>	<code>bool</code>		“If this image should overwrite any current images on the face.”

Table 192:
`DisplayFacelImageRGBRequest` JSON structure

51.2.2 Response

The `DisplayFacelImageRGBResponse` structure has the following fields:

Field	Type	Description
<code>status</code>	<code>ResponseStatus</code>	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 193:
`DisplayFacelImageRGBResponse` JSON structure

51.3. ENABLE MIRROR MODE

“When enabled, camera feed will appear on the robot’s face, along with any detections” (if enabled).

Post: “/v1/enable_mirror_mode”

51.3.1 Request

The EnableMirrorModeRequest message has the following fields:

Field	Type	Description
enable	bool	If true, enables displaying the camera feed (and detections) on the LCD.

Table 194:
EnableMirrorModeRequest JSON structure

51.3.2 Response

The EnableMirrorModeResponse structure has the following fields:

Field	Type	Description
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 195:
EnableMirrorModeResponse JSON structure

51.4. SET EYE COLOR

This is used to set Vector’s current eye color. See also section *61 Settings and Preferences*

Post: “/v1/set_eye_color”

51.4.1 Request

The SetEyeColorRequest has the following fields:

Field	Type	Description
hue	float	The hue to set Vector’s eyes to.
saturation	float	The saturation of the color to set Vector’s eyes to.

Table 196:
SetEyeColorRequest JSON structure

51.4.2 Response

The SetEyeColorResponse structure has the following fields:

Field	Type	Description
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 197:
SetEyeColorResponse JSON structure

52. FACES

This section describes the commands and queries related to Vector's detection of faces, and managing what he knows about them. For a description of the facial detection and recognition process, see Chapter 17 section 79 *Face and Facial features recognition*.

Note: an int32 identifier is used to distinguish between faces that are seen. Each face will have a separate identifier. A positive identifier is used for a face that is known (recognized). This value will be the same when the face disappears and reappears later; the value likely persists across reboots. A negative identifier is used for face that is not recognized; as unknown faces appear and disappear they may be assigned different subsequent negative numbers. If a face becomes recognized, a RobotChangedObservedFaceID event will be sent, along with a change in identifier used.

see also section 59 *On boarding*

52.1. ENUMERATIONS

52.1.1 FaceEnrollmentResult

The FaceEnrollmentResult is used to represent the success of associating a face with a name, or an reason code if there was an error. The enumeration has the following named values:

Name	Value	Description	Table 198: FacialExpression Enumeration
SUCCESS	0	A face was seen, its facial signature and associated name were successfully saved.	
SAW_WRONG_FACE	1		
SAW_MULTIPLE_FACES	2	Too many faces were seen, and Vector did not know which one to associate with the name.	
TIMED_OUT	3		
SAVED_FAILED	4	There was an error saving the facial signature and associated name to non-volatile storage.	
INCOMPLETE	5		
CANCELLED	6	See Cancel Face Enrollment.	
NAME_IN_USE	7		
NAMED_STORAGE_FULL	8	There was no more room in the non-volatile storage to hold another facial signature and associated name.	
UNKNOWN_FAILURE	9		

52.1.2 FacialExpression

The FacialExpression is used to estimate the emotion expressed by each face that vector sees. The enumeration has the following named values:

Name	Value	Description
<i>EXPRESSION_UNKNOWN</i>	0	The facial expression could not be estimated. Note: this could be because the facial expression estimation is disabled.
<i>EXPRESSION_NEUTRAL</i>	1	The face does not appear to have any particular expression.
<i>EXPRESSION_HAPPINESS</i>	2	The face appears to be happy
<i>EXPRESSION_SURPRISE</i>	3	The face appears to be surprised.
<i>EXPRESSION_ANGER</i>	4	The face appears
<i>EXPRESSION_SADNESS</i>	5	The face appears to be sad.

Table 199:
FacialExpression
Enumeration

52.2. STRUCTURES

52.2.1 RobotRenamedEnrolledFace

The RobotRenamedEnrolledFace structure has the following fields:

Field	Type	Description
<i>face_id</i>	int32	The identifier code for the face.
<i>name</i>	string	The name now associated with the face.

Table 200:
RobotRenamedEnrolled
Face JSON structure

52.3. EVENTS

52.3.1 FaceEnrollmentComplete

The FaceEnrollmentComplete structure has the following fields:

Field	Type	Description
<i>face_id</i>	int32	The identifier code for the face.
<i>name</i>	string	The name associated with the face.
<i>result</i>	FaceEnrollmentResult	Whether or not the face enrollment was successful; an error code if not.

Table 201:
FaceEnrollmentComplete
JSON structure

52.3.2 Meet Victor Face Scan Complete

The `MeetVictorFaceScanComplete` structure has no fields.

52.3.3 Meet Victor Face Scan Started

The `MeetVictorFaceScanStarted` structure has no fields.

52.3.4 RobotChangedObservedFaceID

This event occurs when a tracked (but not yet recognized) face is recognized and receives a positive ID. This happens when Vector's view of the face improves. This event can also occur "when face records get merged" "(on realization that 2 faces are actually the same.)"

The `RobotChangeObservedFaceID` structure has the following fields:

Field	Type	Description
<code>new_id</code>	int32	The new identifier code for the face that has been recognized.
<code>old_id</code>	int32	The identifier code that was used for the face until now. Probably negative

Table 202:
`RobotChangedObservedFaceID` JSON structure

52.3.5 RobotObservedFace

The `RobotObservedFace` event is sent when faces are observed within the field of view. This event is only sent if face detection is enabled. This structure has the following fields:

Field	Type	Description
<code>face_id</code>	int32	The identifier code for the face; negative if the face is not recognized, positive if it has been recognized.
<code>expression</code>	<code>FacialExpression</code>	The estimated facial expression seen on the face.
<code>expression_values</code>	uint32[]	An array that represents the histogram of confidence scores in each individual expression. If the expression is not known (e.g. expression estimation is disabled), the array will be all zeros. Otherwise, will sum to 100.
<code>img_rect</code>	<code>CladRect</code>	The area within the camera view holding the face.
<code>name</code>	string	The name associated with the face (if recognized). Empty if a name is not known.
<code>pose</code>	<code>PoseStruct</code>	The position and orientation of the face.
<code>left_eye</code>	<code>CladPoint[]</code>	A polygon outlining the left eye, with respect to the image rectangle.
<code>mouth</code>	<code>CladPoint[]</code>	A polygon outlining the mouth; the coordinates are in the camera image.
<code>nose</code>	<code>CladPoint[]</code>	A polygon outlining the nose; the coordinates are in the camera image.
<code>right_eye</code>	<code>CladPoint[]</code>	A polygon outlining the right eye; the coordinates are in the camera image.
<code>timestamp</code>	uint32	The time that the most recent facial information was obtained. The format is milliseconds since Vector's epoch.

Table 203:
`RobotObservedFace` JSON structure

52.4. CANCEL FACE ENROLLMENT

Cancels the request to look for a face and associate the face with a name.

post: “/v1/cancel_face_enrollment”

52.4.1 Request

The CancelFaceEnrollmentRequest structure has no fields.

52.4.2 Response

The CancelFaceEnrollmentResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 204:
CancelFaceEnrollmentResponse JSON structure

52.5. ENABLE FACE DETECTION

This command enables (or disables) face detection, facial expression detection, blink and gaze detection. Disabling one or more of these features reduces the number of events sent by Vector, and reduces his processing overhead.

post: “/v1/enable_face_detection”

52.5.1 Request

The EnableFaceDetectionRequest structure has the following fields:

Field	Type	Description
<code>enable</code>	bool	If true, face detection (and recognition) is enabled; otherwise face detection processes are disabled.
<code>enable_blink_detection</code>	bool	If true, Vector will attempt “to detect how much detected faces are blinking.” Note: the blink amount is not reported.
<code>enable_expression_estimation</code>	bool	If true, Vector will attempt to estimate facial expressions.
<code>enable_gaze_detection</code>	bool	If true, Vector will attempt “to detect where detected faces are looking.” Note: the gaze direction is not reported.
<code>enable_smile_detection</code>	bool	If true, Vector will attempt “to detect smiles in detected faces.” Note: the smile is not reported.

Table 205:
EnableFaceDetectionRequest JSON structure

52.5.2 Response

The EnableFaceDetectionResponse has the following fields:

Field	Type	Description
<code>status</code>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 206:
EnableFaceDetectionResponse JSON structure

52.6. ERASE ALL ENROLLED FACES

This command is used to erase all of the known faces (and their identity).

post: “/v1/erase_all_enrolled_faces”

52.6.1 Request

The EraseAllEnrolledFacesRequest structure has no fields.

52.6.2 Response

The EraseAllEnrolledFacesResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 207:
EraseAllEnrolledFacesResponse JSON structure

52.7. ERASE ENROLLED FACE BY ID

This command is used to erase the identify feature (and identity) of a known face.

post: “/v1/erase_enrolled_face_by_id”

52.7.1 Request

The EraseEnrolledFaceByIDRequest structure has the following fields:

Field	Type	Description
<i>face_id</i>	int32	The identifier code for the face to erase.

Table 208:
EraseEnrolledFaceByIDRequest JSON structure

52.7.2 Response

The EraseEnrolledFaceByIDResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 209:
EraseEnrolledFaceByIDResponse JSON structure

52.8. FIND FACES

This causes Vector to look around for faces. He does this by turning in place and moving his head up and down. This is carried out by the TBD behaviour.

post: “/v1/find_faces”

52.8.1 Request

The FindFacesRequest structure has no fields.

52.8.2 Response

The FindFacesResponse structure has the following fields:

Table 210:
FindFacesResponse
JSON structure

Field	Type	Description
<i>result</i>	BehaviorResults	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

52.9. REQUEST ENROLLED NAMES

This command is used to list the faces known to Vector, their names, and some other useful information.

post: “/v1/request_enrolled_names”

52.9.1 Request

The RequestEnrolledNamesRequest structure has no fields.

52.9.2 Response

The RequestEnrolledNamesRequest structure has the following fields:

Field	Type	Description	
<i>faces</i>	LoadedKnownFace[]	An array of the faces that are associated with names.	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

The LoadedKnownFace structure has the following fields:

Table 211:
RequestEnrolledNames
Response JSON
structure

Field	Type	Units	Description	
<i>face_id</i>	int32		The identifier code for the face.	
<i>last_seen_seconds_since_epoch</i>	int64	seconds	The timestamp of the time the face was last seen. The format is unix time: seconds since 1970, in UTC?	
<i>name</i>	name		The name associated with the face.	
<i>seconds_since_first_enrolled</i>	int64	seconds	The number of seconds since the face was first associated with a name and entered into the known faces database.	
<i>seconds_since_last_seen</i>	int64	seconds	The number of seconds since the face was last seen	
<i>seconds_since_last_updated</i>	int64	seconds	The number of seconds since (?) the name associated with the face was last changed.(?)	

Table 212:
LoadedKnownFace
JSON structure

52.10. SET FACE TO ENROLL

This command is can used to assign a name to unrecognized face, or to update the recognition pattern (and name) for an already known face. This command initiates a behaviour that can be configured.

post: “/v1/set_face_to_enroll”

52.10.1 Request

The SetFaceToEnrollRequest structure has the following fields:

Field	Type	Description
<i>name</i>	string	The name to associate with the face.
<i>observed_id</i>	int32	If non-zero, the identifier code for a specific observed face to enroll. Note the identifier is negative if the face is not already recognized, positive if it has been recognized. If zero, Vector will use the next face he sees.
<i>save_id</i>	int32	If non-zero, Vector will use this ID as the ID for the face. (Note: this must be “the ID of an existing face”). If zero, Vector will use the <i>observedID</i> for the ID.
<i>save_to_robot</i>	bool	If true, “save to robot’s NVStorage when done (NOTE: will (re)save everyone enrolled!)”
<i>say_name</i>	bool	If true, “play say-name/celebration animations on success before completing.”
<i>use_music</i>	bool	If true, “starts special music during say-name animations (will leave music playing!)”

Table 213:
SetFaceToEnrollRequest
JSON structure

52.10.2 Response

The SetFaceToEnrollResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 214:
SetFaceToEnrollResponse
JSON structure

52.11. UPDATE ENROLLED FACE BY ID

This command is used to change the name associated with a face.

post: “/v1/update_enrolled_face_by_id”

52.11.1 Request

The UpdateEnrolledFaceByIDRequest structure has the following fields:

Field	Type	Description
<i>face_id</i>	int32	The identifier code for the face.
<i>new_name</i>	string	The new name to associate with the face.
<i>old_name</i>	string	The name associated (until now) with the face. This name must match the one Vector has for the <i>face_id</i> . If not the command will not be honored.

Table 215:
UpdateEnrolledFaceByIDRequest JSON structure

52.11.2 Response

The UpdateEnrolledFaceByIDResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 216:
UpdateEnrolledFaceByIDResponse JSON structure

53. FEATURES & ENTITLEMENTS

Vector has granular features that can be enabled and disabled thru the use of feature flags. This section describes the queries related to list Vector's features flags, and their state. For a description of feature flags, see Chapter 29 *Settings, Preferences, Features, and Statistics*. For a list of the features, and a description of each, see Appendix H, *Table 526: The features*.

Note: the API does not include the ability to enable a feature.

Note: For AI behaviour “features” see section *42.2.1 FeatureStatus*.

53.1. ENUMERATIONS

53.1.1 UserEntitlement

The UserEntitlement enumeration has the following named values:

Name	Value	Description
KICKSTARTER_EYES	0	<i>Note: This was an entitlement that was explored, but not used.</i>

Table 217:
UserEntitlement
Enumeration

53.2. GET FEATURE FLAG

Request the current setting of a feature flag.

post: “/v1/feature_flag”

53.2.1 Request

The FeatureFlagRequest message has the following fields:

Field	Type	Description
<i>feature_name</i>	string	The name of the feature; this feature name should be one of those listed in response to <i>Get Feature Flag List</i> (section 53.3). See Appendix H, <i>Table 526: The features</i>

Table 218:
FeatureFlagRequest
JSON structure

53.2.2 Response

The FeatureFlagResponse type has the following fields:

Field	Type	Description
<i>feature_enabled</i>	bool	True if the feature is enabled, false if not
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<i>valid_feature</i>	bool	True if the given feature name is a valid name of a feature; false if not.

Table 219:
FeatureFlagResponse
JSON structure

53.3. GET FEATURE FLAG LIST

Request the list of the current feature flags. Note: to see which flags are enabled, use the *Get Feature Flag* command (section 53.2).

post: “/v1/feature_flag_list”

53.3.1 Request

The following is streamed... to the robot?

Field	Type	Description
<i>request_list</i>	string	

Table 220:
FeatureFlagListRequest
JSON structure

53.3.2 Response

The FeatureFlagListResponse type has the following fields:

Field	Type	Description
<i>list</i>	string[]	An array of the feature flags; see Appendix H, <i>Table 526: The features</i> for a description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 221:
FeatureFlagListResponse
JSON structure

53.4. UPDATE USER ENTITLEMENTS

UpdateUserEntitlements

Post: “/v1/update_user_entitlements”

53.4.1 Request

The UpdateUserEntitlementsRequest has the following fields:

Field	Type	Description	Table 222: JSON Parameters for UpdateUserEntitlementsRequest
<i>user_entitlements</i>	UserEntitlementsConfig		

The UserEntitlementsConfig has the following fields:

Field	Type	Description	Table 223: JSON Parameters for UserEntitlementsConfig
<i>kickstarter_eyes</i>	bool		

53.4.2 Response

The UpdateUserEntitlementsResponse type has the following fields:

Field	Type	Description	Table 224: UpdateUserEntitlements Response JSON structure
<i>code</i>	resultCode		
<i>doc</i>	Jdoc		
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

54. IMAGE PROCESSING

See also section 52 *Faces*, for detecting and recognizing faces, and enabling the features

54.1. ENUMERATIONS

54.1.1 ImageEncoding

The `ImageEncoding` is used to describe the format of the image data contained in the chunk. The enumeration has the following named values:

Table 225:
ImageEncoding
Enumeration

Name	Value	Description
<code>NONE_IMAGE_ENCODING</code>	0	Image is not encoded. TBD: does this mean no image?
<code>RAW_GRAY</code>	1	“No compression”
<code>RAW_RGB</code>	2	“no compression, just [RGBRGBRG...]”
<code>YUYV</code>	3	
<code>YUV420SP</code>	4	
<code>BAYER</code>	5	
<code>JPEG_GRAY</code>	6	
<code>JPEG_COLOR</code>	7	
<code>JPEG_COLOR_HALF_WIDTH</code>	8	
<code>JPEG_MINIMIZED_GRAY</code>	9	“Minimized grayscale JPEG - no header, no footer, no byte stuffing”
<code>JPEG_MINIMIZED_COLOR</code>	10	“Minimized grayscale JPEG – no header, no footer, no byte stuffing, with added color data.”

54.2. EVENTS

54.2.1 RobotObservedMotion

Note: *this event is not available*; it was defined in an early version of the API protocol, and removed. It is reproduced here for information purposes; the event may be in future releases.

This RobotObservedMotion event structure has the following fields:

Field	Type	Units	Description	Table 226: RobotObservedMotion parameters
<i>bottom_img_area</i>	float	<i>area fraction</i>	“Area of the supporting region for the point, as a fraction of the bottom region”	
<i>bottom_img_x</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>bottom_img_y</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>ground_area</i>	float	<i>area fraction</i>	“Area of the supporting region for the point, as a fraction of the ground ROI. If unable to map to the ground, area=0.”	
<i>ground_x</i>	int32	<i>mm</i>	“Coordinates of the point on the ground, relative to robot, in mm.”	
<i>ground_y</i>	int32	<i>mm</i>	“Coordinates of the point on the ground, relative to robot, in mm.”	
<i>img_area</i>	float	<i>area fraction</i>	“Area of the supporting region for the point, as a fraction of the image”	
<i>img_x</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>img_y</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>left_img_area</i>	float	<i>area fraction</i>	“Area of the supporting region for the point, as a fraction of the left region.”	
<i>left_img_x</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>left_img_y</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>right_img_area</i>	float	<i>area fraction</i>	“Area of the supporting region for the point, as a fraction of the right region.”	
<i>right_img_x</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>right_img_y</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>timestamp</i>	uint	<i>ms</i>	“Timestamp of the corresponding image”	
<i>top_img_area</i>	float	<i>area fraction</i>	“Area of the supporting region for the point, as a fraction of the top region”	
<i>top_img_x</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	
<i>top_img_y</i>	int32	<i>pixel</i>	“Pixel coordinate of the point in the image, relative to top-left corner.”	

54.2.2

VisionModesAutoDisabled

This event is “sent when vision modes are automatically disabled due to the SDK no longer having control of the robot.”

The VisionModesAutoDisabled structure has no fields.

54.3. CAMERA FEED

“Request a camera feed from the robot.”

Post: “/v1/camera_feed”

54.3.1

Request

The CameraFeedRequest has no fields.

54.3.2

Response

The response is a stream of the following CameraFeedResponse structure. This structure has the following fields:

Field	Type	Description	<i>Table 227: CameraFeedResponse JSON structure</i>
<i>data</i>	bytes	The bytes of the image	
<i>frame_time_stamp</i>	uint32		
<i>image_encoding</i>	ImageEncoding		
<i>image_id</i>	uint32		
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

54.4. CAPTURE SINGLE IMAGE

“Request a single image to be captured and sent from the robot.”

Post: “/v1/capture_single_image”

54.4.1 Request

The CaptureSingleImageRequest has no fields.

54.4.2 Response

The CaptureSingleImageResponse structure has the following fields:

Field	Type	Description	
<i>data</i>	bytes	The bytes of the image	Table 228: <i>CaptureSingleImageResponse</i> JSON structure
<i>frame_time_stamp</i>	uint32		
<i>image_encoding</i>	ImageEncoding		
<i>image_id</i>	uint32		
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

54.5. ENABLE IMAGE STREAMING

“Toggle image streaming at the given resolution”

Post: “/v1/enable_image_streaming”

54.5.1 Request

The EnableImageStreamingRequest type has the following fields:

Field	Type	Description	
<i>enable</i>	bool	True if a stream of images from the camera should be sent.	Table 229: <i>EnableImageStreamingRequest</i> JSON structure

54.5.2 Response

The EnableImageStreamingResponse has the following fields:

Field	Type	Description	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	Table 230: <i>EnableImageStreamingResponse</i> JSON structure

54.6. ENABLE MARKER DETECTION

This enables and disables the processing of custom marker symbols and generating events when a marker symbol is seen. If enabled, when an marker symbol is seen, the RobotObservedObject event will be sent.

Note: The custom marker detection may remain internally enabled, even if disabled by the SDK, “if another subscriber (including one internal to the robot) requests this vision mode be active.”

Post: “/v1/enable_marker_detection”

54.6.1 Request

The EnableMarkerDetectionRequest has the following fields:

Field	Type	Description
enable	bool	If true, enable search for marker symbols, and generating events when they are detected,

Table 231: JSON Parameters for EnableMarkerDetection Request

54.6.2 Response

The EnableMarkerDetectionResponse has the following fields:

Field	Type	Description
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 232: EnableMarkerDetection Response JSON structure

54.7. ENABLE MOTION DETECTION

Enables detection visual motion, and sending RobotObservedMotion (not implemented) events.

Note: this command is has no useful effect since Vector does not support sending the events it enables.

Post: “/v1/enable_motion_detection”

54.7.1 Request

The EnableMotionDetectionRequest structure has the following fields:

Field	Type	Description
enable	bool	True if RobotObservedMotion events should be sent.

Table 233:
EnableMotionDetectionRequest JSON structure

54.7.2 Response

The EnableMotionDetectionResponse has the following fields:

Field	Type	Description
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 234:
EnableMotionDetectionResponse JSON structure

54.8. IS IMAGE STREAMING ENABLED

“Request whether or not image streaming is enabled on the robot”

Post: “/v1/is_image_streaming_enabled”

54.8.1 Request

The IsImageStreamingRequest has no fields.

54.8.2 Response

The IsImageStreamingResponse “indicates whether or not image streaming is enabled on the robot.”

The structure has the following fields:

Field	Type	Description
enable	bool	True if image streaming is enabled, false otherwise

Table 235:
IsImageStreamingResponse JSON structure

55. INTERACTIONS WITH OBJECTS

These commands are used to interact with faces and objects. These initiate behaviours.

- Some behaviours can be assigned a tag that can be used to cancel it later.
- Some behaviours accept a parameter to modify their motion profile.
- Behaviour results value

Actions

- Actions can be assigned a tag that can be used to cancel it later.
- Action results value

See also section 42 *Actions and Behaviour*, and section 49 *Cube*

55.1. STRUCTURES

55.1.1 PathMotionProfile

This structure contains “all the information relevant to how a path should be modified or traversed.”

Table 236:
PathMotionProfile
JSON structure

Field	Type	Units	Description
<i>accel_mmmps2</i>	float	<i>mm/sec²</i>	How fast Vector should accelerate to achieve the speed.
<i>decel_mmmps2</i>	float	<i>mm/sec²</i>	How fast Vector should decelerate to the speed.
<i>is_custom</i>	bool		
<i>dock_accel_mmmps2</i>	float	<i>mm/sec²</i>	How fast Vector should accelerate when performing the docking procedure.
<i>dock_decel_mmmps2</i>	float	<i>mm/sec²</i>	How fast Vector should decelerate when performing the docking procedure.
<i>dock_speed_mmmps</i>	float	<i>mm/sec</i>	The speed that Vector should perform the docking procedure at.
<i>point_turn_accel_mmmps2</i>	float	<i>mm/sec²</i>	How fast Vector should accelerate when turning (in place).
<i>point_turn_decel_mmmps2</i>	float	<i>mm/sec²</i>	How fast Vector should decelerate when turning (in place).
<i>point_turn_speed_mmmps</i>	float	<i>mm/sec</i>	The speed that Vector should perform a turn (in place) at.
<i>reverse_speed_mmmps</i>	float	<i>mm/sec</i>	How fast Vector should move when backing up
<i>speed_mmmps</i>	float	<i>mm/sec</i>	The speed that Vector should move along the path

55.2. DRIVE OFF CHARGER

This command directs Vector to drive off his charger – if he is on it. This will initiate a behavior.

Post: “/v1/drive_off_charger”

55.2.1 Request

The DriveOffChargerRequest structure has no fields.

55.2.2 Response

The DriveOffChargerResponse type has the following fields:

Field	Type	Description
<i>result</i>	BehaviorResults	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 237:
DriveOffChargerResponse JSON structure

55.3. DRIVE ON CHARGER

This command directs Vector to drive onto his charger – if he is not already on it. “Vector will attempt to find the charger and, if successful, he will back onto it and start charging. Vector’s charger has a visual marker so that the robot can locate it for self-docking.” This will initiate a behavior.

Post: “/v1/drive_on_charger”

55.3.1 Request

The DriveOnChargerRequest structure has no fields.

55.3.2 Response

The DriveOnChargerResponse type has the following fields:

Field	Type	Description
<i>result</i>	BehaviorResults	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 238:
DriveOnChargerResponse JSON structure

55.4. GO TO OBJECT

“Tell Vector to drive to the specified object” (i.e. his cube). Note: custom objects “are not supported.” This initiates an action.

55.4.1 Request

The GoToObjectRequest structure has the following fields:

Field	Type	Units	Description	Table 239: GoToObjectRequest JSON structure
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>	
<i>motion_prof</i>	PathMotionProfile		<i>Optional.</i>	
<i>num_retries</i>	int32		Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.	
<i>object_id</i>	int32		The identifier of the object to drive to. Note: custom objects “are not supported”	
<i>distance_from_object_origin_mm</i>	float	mm	“The distance from the object to stop. This is the distance between the origins. For instance, the distance from the robot’s origin (between Vector’s two front wheels) to the cube’s origin (at the center of the cube) is ~40mm.”	
<i>use_pre_dock_pose</i>	bool		Set this to false	

55.4.2 Response

The GoToObjectResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description	Table 240: GoToObjectResponse JSON structure
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

55.5. TURN TOWARDS FACE

“Tell Vector to turn towards this face.” This initiates an action.

55.5.1 Request

The TurnTowardsFaceRequest structure has the following fields:

Field	Type	Description
<i>face_id</i>	int32	The identifier of the face to look for
<i>id_tag</i>	int32	This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>
<i>max_turn_angle_rad</i>	float	Recommend value of 180°
<i>num_retries</i>	int32	Maximum of times to attempt to reach the object. A retry is attempted if Vector is unable to reach the target object.

Table 241:
TurnTowardsFaceRequest JSON structure

55.5.2 Response

The TurnTowardsFaceResponse is sent to indicate whether the action successfully completed or not.

This structure has the following fields:

Field	Type	Description
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 242:
TurnTowardsFaceResponse JSON structure

56. JDOCS

This section discussed the commands for “Jdocs” (short for “JSON Documents”), which are JSON objects that are passed to Vic-Engine and Vic-Cloud. See the next chapter for interactions with a remote Jdocs server, using a sibling protocol.

56.1. ENUMERATIONS

56.1.1 JdocType

The JdocType enumeration has the following named values:

Name	Value	Description
ACCOUNT_SETTINGS	2	Refers to the owner’s account settings
ROBOT_LIFETIME_STATS	1	Refers to the robot’s settings (owner preferences)
ROBOT_SETTINGS	0	Refers to the robot’s lifetime stats.
USER_ENTITLEMENTS	3	Refers to the owner’s entitlements.

Table 243: JdocType Enumeration

Items of these types are described in more detail in Chapter 29.

56.2. STRUCTURES

56.2.1 JDoc

The Jdoc type has the following fields:

Field	Type	Description
client_meta	string	
doc_version	uint64	
fmt_version	uint64	
json_doc	string	

Table 244: Jdoc JSON structure

56.2.2 NamedJDoc

The NamedJdoc type has the following fields:

Field	Type	Description
doc	Jdoc	The JSON structure and meta-data about the document
jdoc_type	JdocType	The type of document provided in “doc”

Table 245: JSON NamedJdoc structure

56.3. EVENTS

56.3.1 JdocsChanged

The JdocsChanged message has the following fields:

Field	Type	Description
<i>jdoc_types</i>	JdocType[]	

Table 246: JSON
JdocsChanged request
structure

56.4. PULL JDOCS

Post: “/v1/pull_jdocs”

56.4.1 Request

The PullJdocsRequest has the following fields:

Field	Type	Description
<i>jdoc_types</i>	JdocType[]	

Table 247: JSON
PullJdocsRequest
structure

56.4.2 Response

The PullJdocResponse has the following fields:

Field	Type	Description
<i>named_jdocs</i>	NamedJdoc[]	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 248: JSON
PullJdocsResponse
structure

57. MOTION CONTROL

This section describes commands to drive Vector, and to control his lift & head position. See also section 55 *Interactions with Objects*.

57.1. DRIVE STRAIGHT

This will initiate an action of Vector driving in a straight line.

Note: “Vector will drive for the specified distance (forwards or backwards) Vector must be off of the charger for this movement action. [A] that use the wheels cannot be performed at the same time; otherwise you may see a TRACKS_LOCKED error.”

57.1.1 Request

The DriveStraightRequest has the following fields:

Field	Type	Units	Description
<i>dist_mm</i>	float	mm	The distance to drive. (Negative is backwards)
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional</i> .
<i>is_absolute</i>	uint32		If 0, turn by <i>angle_rad</i> relative to the current orientation. If 1, turn to the absolute angle given by <i>angle_rad</i> .
<i>num_retries</i>	int32		Maximum of times to attempt to move the head to the height. A retry is attempted if Vector is unable to reach the target angle
<i>should_play_animation</i>	bool		If true, “play idle animations whilst driving (tilt head, hum, animated eyes, etc.)”
<i>speed_mmmps</i>	float	mm/sec	The speed to drive at. This should be positive.

57.1.2 Response

The DriveStraightResponse has the following fields:

Field	Type	Description
<i>response</i>	ActionResult	Whether the action is able to be run. If not, an error code indicating why not.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

57.2. DRIVE WHEELS

Sets the speed and acceleration for Vector's wheel motors.

57.2.1 Request

The DriveWheelsRequest has the following fields:

Field	Type	Units	Description	Table 251: DriveWheelsRequest JSON structure
<i>left_wheel_mmmps</i>	float	mm/sec	The initial speed to set the left wheel to.	
<i>left_wheel_mmmps2</i>	float	mm/sec ²	How fast to increase the speed of the left wheel.	
<i>right_wheel_mmmps</i>	float	mm/sec	The initial speed to set the right wheel to.	
<i>right_wheel_mmmps2</i>	float	mm/sec ²	How fast to increase the speed of the right wheel.	

To unlock the wheels, set all values to 0.

57.2.2 Response

The DriveWheelsResponse has the following fields:

Field	Type	Description	Table 252: DriveWheelsResponse JSON structure
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

57.3. GO TO POSE

“Tells Vector to drive to the specified pose and orientation.” This will initiate an action.

“In navigating to the requested pose, Vector will use path planning.

“Since the robot understands position by monitoring its tread movement, it does not understand movement in the z axis. This means that the only applicable elements of pose in this situation are position.x position.y and rotation.angle_z.

“Note that actions that use the wheels cannot be performed at the same time, otherwise you may see a TRACKS_LOCKED error.”

Post: “/v1/go_to_pose”

57.3.1 Request

The GoToPoseRequest structure has the following fields:

Field	Type	Units	Description
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>
<i>motion_prof</i>	PathMotionProfile		
<i>num_retries</i>	int32		Maximum of times to attempt to reach the pose. A retry is attempted if Vector is unable to reach the target pose.
<i>rad</i>	float	radians	The angle to change orientation to.
<i>x_mm</i>	float	mm	The x-coordinate of the position to move to.
<i>y_mm</i>	float	mm	The y-coordinate of the position to move to.

Table 253:
GoToPoseRequest JSON structure

57.3.2 Response

The GoToPoseResponse is sent to indicate whether the action successfully completed or not. This structure has the following fields:

Field	Type	Description
<i>result</i>	ActionResult	An error code indicating the success of the action, the detailed reason why it failed, or that the action is still being carried out.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 254:
GoToPoseResponse JSON structure

57.4. MOVE HEAD

Move Vector's head

57.4.1 Request

The *MoveHeadRequest* has the following fields:

Field	Type	Units	Description
<i>speed_rad_per_sec</i>	float	radian/sec	The speed to drive the head motor at. Positive values are up, negative move down. A value of 0 will unlock the head track.

Table 255:
MoveHeadRequest
JSON structure

57.4.2 Response

The *MoveHeadResponse* has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 256:
MoveHeadResponse
JSON structure

57.5. MOVE LIFT

Move Vector's lift

57.5.1 Request

The *MoveLiftRequest* has the following fields:

Field	Type	Units	Description
<i>speed_rad_per_sec</i>	float	radian/sec	The speed to drive the lift at. Positive values are up, negative move down. A value of 0 will unlock the lift track.

Table 257:
MoveLiftRequest JSON structure

57.5.2 Response

The *MoveLiftResponse* has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 258:
MoveLiftResponse JSON structure

57.6. SET HEAD ANGLE

This will initiate an action to move Vector's head to a given angle.

57.6.1 Request

The SetHeadAngleRequest has the following fields:

Field	Type	Units	Description	Table 259: SetHeadAngleRequest JSON structure
<i>accel_rad_per_sec2</i>	float	radian/sec ²	How fast to increase the speed the head is moving at. Recommended value: 10 radians/sec ²	
<i>angle_rad</i>	float	radians	The target angle to move Vector's head to. This should be in the range -22.0° to 45.0°.	
<i>duration_sec</i>	float	sec	"Time for Vector's head to move in seconds. A value of zero will make Vector try to do it as quickly as possible."	
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>	
<i>max_speed_rad_per_sec</i>	float	radian/sec	The maximum speed to move the head. (This clamps the speed from further acceleration.) Recommended value: 10 radians/sec	
<i>num_retries</i>	int32	count	Maximum of times to attempt to move the head to the height. A retry is attempted if Vector is unable to reach the target angle	

57.6.2 Response

The SetHeadAngleResponse has the following fields:

Field	Type	Description	Table 260: SetHeadAngleResponse JSON structure
<i>response</i>	ActionResult	Whether the action is able to be run. If not, an error code indicating why not.	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

57.7. SET LIFT HEIGHT

This will initiate an action to move Vector's lift to a given height.

57.7.1 Request

The SetLiftRequest has the following fields:

Field	Type	Units	Description
<i>accel_rad_per_sec2</i>	float	radian/sec ²	How fast to increase the speed the lift is moving at/ Recommended value: 10 radians/sec ²
<i>duration_sec</i>	float	sec	“Time for Vector's lift to move in seconds. A value of zero will make Vector try to do it as quickly as possible.”
<i>height_mm</i>	float	mm	The target height to raise the lift to. Note: the python API employs a different range for this parameter
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional.</i>
<i>max_speed_rad_per_sec</i>	float	radian/sec	The maximum speed to move the lift at. (This clamps the speed from further acceleration.) Recommended value: 10 radians/sec
<i>num_retries</i>	int32	count	Maximum of times to attempt to move the lift to the height. A retry is attempted if Vector is unable to reach the target height.

Table 261:
SetLiftRequest JSON
structure

57.7.2 Response

The SetLiftResponse has the following fields:

Field	Type	Description
<i>response</i>	ActionResult	Whether the action is able to be run. If not, an error code indicating why not.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 262:
SetLiftResponse JSON
structure

57.8. STOP ALL MOTORS

Stop all motor commands for the head, lift and wheels

57.8.1 Request

The StopAllMotorsRequest structure has no fields.

57.8.2 Response

The StopAllMotorsResponse has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 263:
StopAllMotorsResponse
JSON structure

57.9. TURN IN PLACE

This command initiates an action to turn Vector around its current position.

Note: “Vector must be off of the charger for this movement action. Note that actions that use the wheels cannot be performed at the same time, otherwise you may see a TRACKS_LOCKED error.”

57.9.1

Request

The TurnInPlaceRequest has the following fields:

Field	Type	Units	Description	Table 264: TurnInPlaceRequest JSON structure
<i>accel_rad_per_sec2</i>	float	radian/sec ²	How fast to increase the speed the body is moving at	
<i>angle_rad</i>	float	radians	If <i>is_absolute</i> is 0, turn relative to the current heading by this number of radians; positive means turn left, negative is turn right. Otherwise, turn to the absolute orientation given by this angle.	
<i>id_tag</i>	int32		This is an action tag that can be assigned to this request and used later to cancel the action. <i>Optional</i> .	
<i>is_absolute</i>	uint32		If 0, turn by <i>angle_rad</i> relative to the current orientation. If 1, turn to the absolute angle given by <i>angle_rad</i> .	
<i>num_retries</i>	int32	count	Maximum of times to attempt to move the head to the height. A retry is attempted if Vector is unable to reach the target angle	
<i>speed_rad_per_sec</i>	float	radian/sec	The speed to move around the arc.	
<i>tol_rad</i>	float	count	“The angular tolerance to consider the action complete (this is clamped to a minimum of 2 degrees internally).”	

57.9.2

Response

The TurnInPlaceResponse has the following fields:

Field	Type	Description	Table 265: TurnInPlaceResponse JSON structure
<i>response</i>	ActionResult	Whether the action is able to be run. If not, an error code indicating why not.	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

58. MOTION SENSING AND ROBOT STATE

This section describes the structures and events that describe the sensed motions.

The values are given with respect to a “coordinate space is relative to Vector, where Vector's origin is the point on the ground between Vector's two front wheels. The X axis is Vector's forward direction, the Y axis is to Vector's left, and the Z axis is up.”

58.1. STRUCTURES

58.1.1 AccelData

This structure is used to report the accelerometer readings, as part of the RobotState structure. The accelerometer is located in Vector's head, so its XYZ axes are not the same as Vector's body axes. When motionless, the accelerometer can be used to calculate the angle of Vector's head. The AccelData has the following fields:

Field	Type	Units	Description
x	float	mm/s ²	The acceleration along the accelerometers X axis.
y	float	mm/s ²	The acceleration along the accelerometers Y axis.
z	float	mm/s ²	The acceleration along the accelerometers Z axis.

Table 266: AccelData JSON structure

When at rest, there will be a constant 9810 mm/s² downward acceleration from gravity. This most likely will be distributed across multiple axes.

58.1.2 GyroData

This structure is used to report the gyroscope readings, as part of the RobotState structure. The gyroscope is located in Vector's head, so its XYZ axes are not the same as Vector's body axes. The GyroData has the following fields:

Field	Type	Units	Description
x	float	radian/s	The angular velocity around the X axis.
y	float	radian/s	The angular velocity around the Y axis.
z	float	radian/s	The angular velocity around the Z axis.

Table 267: GyroData JSON structure

58.1.3

ProxData

This structure is used to report the “time of flight” proximity sensor readings, as part of the RobotState structure.

“The proximity sensor is located near the bottom of Vector between the two front wheels, facing forward. The reported distance describes how far in front of this sensor the robot feels an obstacle is. The sensor estimates based on time-of-flight information within a field of view which the engine resolves to a certain quality value.”

The distance measurement may not be valid; “four additional flags are supplied by the engine to indicate whether this proximity data is considered valid for the robot’s internal pathfinding.” It is recommended that an application track the most recent proximity data from the robot, and the most recent proximity data which did not have the lift blocking.

The ProxData has the following fields:

Field	Type	Units	Description
<i>distance_mm</i>	uint32	mm	The distance to the object
<i>found_object</i>	bool		True if “the sensor detected an object in the valid operating range.”
<i>is_lift_in_fov</i>	bool		True if “Vector’s lift (or an object in the lift) is blocking the time-of-flight sensor. While the lift will send clear proximity signals, it’s not useful for object detection.”
<i>signal_quality</i>	float		An estimate of the “reliability of the measurement.” “The proximity sensor detects obstacles within a given field of view; this value represents the likelihood of the reported distance being a solid surface.”
<i>unobstructed</i>	bool		True if “the sensor has confirmed it has not detected anything up to its max range.” (The opposite of <i>found_object</i>)

Table 268: ProxData JSON structure

58.1.4

TouchData

This structure is used to report the touch sensor readings, as part of the RobotState structure. The TouchData has the following fields:

Field	Type	Units	Description
<i>is_being_touched</i>	bool		True if Vector is currently being touched.
<i>raw_touch_value</i>	uint32		“Raw input from the touch sensor.”

Table 269: TouchData JSON structure

58.2. EVENTS

58.2.1 RobotState

The RobotState structure is periodically posted in an *Event* message. The structure has the following fields:

Table 270: RobotState JSON structure

Field	Type	Units	Description
<i>accel</i>	AccelData		The accelerometer readings
<i>carrying_object_id</i>	int32		-1 if no object is being carried. Otherwise the identifier of the cube (or other object) being carried.
<i>carrying_object_on_top_id</i>	int32		<i>Not supported</i>
<i>gyro</i>	GyroData		The gyroscope readings
<i>head_angle_rad</i>	float	radian	The angle of Vector's head (how much it is tilted up or down).
<i>head_tracking_object_id</i>	int32		The identifier “of the object the head is tracking to.” If no object is being tracked, this will be -1.
<i>last_image_time_stamp</i>	uint32		“The robot's timestamp for the last image seen.” The format is milliseconds since Vector's epoch.
<i>left_wheel_speed_mmmps</i>	float	mm/s	The speed of Vector's left wheel.
<i>lift_height_mm</i>	float	mm	“Height of Vector's lift from the ground.”
<i>localized_to_object_id</i>	int32		The identifier “of the object that the robot is localized to.” If no object, this will be -1.
<i>pose</i>	PoseStruct		“The current pose (position and orientation) of Vector.”
<i>pose_angle_rad</i>	float	radian	“Vector's pose angle (heading in X-Y plane).”
<i>pose_pitch_rad</i>	float	radian	“Vector's pose pitch (angle up/down).”
<i>right_wheel_speed_mmmps</i>	float	mm/s	The speed of Vector's right wheel.
<i>prox_data</i>	ProxData		The time-of-flight proximity sensor readings.
<i>status</i>	uint32		A bit map of active states of Vector; the bits are described in the RobotStatus enumeration.
<i>touch_data</i>	TouchData		The touch sensor readings.

59. ON BOARDING

The section describes the command and events used to introduce Vector to his new human, his human to Vector's features.

59.1. ENUMERATIONS

59.1.1 OnboardingPhase

The OnboardingPhase enumeration has the following named values:

Name	Value	Description	Table 271: OnboardingPhase Enumeration
<i>InvalidPhase</i>	0		
<i>Default</i>	1		
<i>LookAtPhone</i>	2		
<i>WakeUp</i>	3		
<i>LookAtUser</i>	4		
<i>TeachWakeWord</i>	5		
<i>TeachComeHere</i>	6		
<i>TeachMeetVictor</i>	7		

59.1.2 OnboardingPhaseState

The OnboardingPhaseState enumeration has the following named values:

Name	Value	Description	Table 272: OnboardingPhaseState Enumeration
<i>PhaseInvalid</i>	0		
<i>PhasePending</i>	1		
<i>PhaseInProgress</i>	2		
<i>PhaseComplete</i>	3		

59.2. EVENTS

59.2.1 Onboarding

The Onboarding event is sent as different stages of the onboarding process have been completed. This structure has the following fields:

Field	Type	Description
<code>onboarding_1p0_charging_info</code>	<code>Onboarding1p0ChargingInfo</code>	
<code>onboarding_state</code>	<code>OnboardingState</code>	
<code>onboarding_wakeup_finished</code>	{}	This structure contains no fields.

Table 273: Onboarding JSON structure

59.2.2 Onboarding1p0ChargingInfo

This structure is used to report whether Vector needs to charge, and an estimated duration. It is part of the Onboarding event structure. This structure has the following fields:

Field	Type	Units	Description
<code>needs_to_charge</code>	bool		If true, Vector needs to charge before onboarding can continue.
<code>on_charger</code>	bool		If true, Vector is on his charger, and there is power supplied to the charger.
<code>suggested_charger_time</code>	float		The estimated amount of time to charge Vector before completing the onboarding process.

Table 274:
Onboarding1p0ChargingInfo JSON structure

59.2.3 OnboardingState

The OnboardingState type has the following fields:

Field	Type	Description
<code>stage</code>	<code>OnboardingStages</code>	Where in the onboarding process we are

Table 275:
OnboardingState JSON structure

The OnboardingStages enumeration has the following named values:

Name	Value	Description
<code>NotStarted</code>	0	The onboarding process has not started yet.
<code>TimedOut</code>	1	The onboarding process has halted because an operation timed out.
<code>Complete</code>	3	The onboarding has completed.
<code>DevDoNothing</code>	4	

Table 276:
OnboardingStages Enumeration

59.3. ONBOARDING INPUT

Post: “/v1/send_onboarding_input”

59.3.1 Request

The OnboardingInputRequest has one (and only one) of the following fields:

Field	Type	Description	Table 277: OnboardingInputRequest JSON structure
<code>onboarding_charge_info_request</code>	{}	This is a request for charging information; it contains no fields.	
<code>onboarding_complete_request</code>	{}	This is a request to complete the onboarding; it contains no fields.	
<code>onboarding_mark_complete_and_exit</code>	{}	This contains no fields.	
<code>onboarding_restart</code>	{}	This is a request to restart the onboarding process; it contains no fields.	
<code>onboarding_skip_onboarding</code>	{}	This is a request to skip the onboarding; it contains no fields.	
<code>onboarding_phase_progress_request</code>	{}	This contains no fields.	
<code>onboarding_set_phase_request</code>	OnboardingSetPhaseRequest	See below.	
<code>onboarding_wake_up_request</code>	{}	This is a request to wake up Vector; it contains no fields.	
<code>onboarding_wake_up_started_request</code>	{}	This contains no fields.	

The OnboardingSetPhaseRequest type has the following fields:

Field	Type	Description	Table 278: OnboardingSetPhaseRequest JSON structure
<code>phase</code>	OnboardingPhase	The desired phase to be in	

59.3.2

Response

The OnboardingInputResponse has a status field one (and only one) of the remaining following fields (which will correspond to the one sent in the request):

Field	Type	Description
onboarding_charge_info_response	OnboardingChargingInfoResponse	<i>See below</i>
onboarding_complete_response	OnboardingCompleteResponse	<i>See below</i>
onboarding_phase_progress_response	{}	<i>See below</i>
onboarding_set_phase_response	OnboardingSetPhaseResponse	<i>See below.</i>
onboarding_wake_up_response	OnboardingWakeUpResponse	<i>See below</i>
onboarding_wake_up_started_response	OnboardingWakeUpStartedResponse	
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 279:
OnboardingInputResponse JSON structure

ONBOARDINGCHARGINGINFORESPONSE

This structure is used to report whether Vector needs to charge, and an estimated duration. It is part of the OnboardingInputResponse event structure. This structure has the following fields:

Field	Type	Units	Description
needs_to_charge	bool		If true, Vector needs to charge before onboarding can continue.
on_charger	bool		If true, Vector is on his charger, and there is power supplied to the charger.
required_charge_time	float		The estimated amount of time to charge Vector before completing the onboarding process.
status	ResponseStatus		A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 280:
OnboardingInputResponse structure

Note: this structure is similar to the Onboarding1p0ChargingInfo structure. That structure is older, but retained as software had already been developed against it.

ONBOARDINGCOMPLETERESPONSE

The OnboardingCompleteResponse type has the following fields:

Field	Type	Description
completed	bool	True if the onboarding has completed

Table 281:
OnboardingCompleteResponse JSON structure

ONBOARDINGPHASEPROGRESSRESPONSE

This structure is used to report how far we are in a particular phase of onboarding. It is part of the OnboardingInputResponse event structure. This structure has the following fields:

Field	Type	Units	Description	Table 282: OnboardingPhaseProgre ssResponse structure
<i>last_set_phase</i>	OnboardingPhase			
<i>last_set_phase_state</i>	OnboardingPhaseState			
<i>percent_completed</i>	int32	%	How far we are in the phase.	
<i>status</i>	ResponseStatus		A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

ONBOARDINGSETPHASERESPONSE

The OnboardingSetPhaseResponse type has the following fields:

Field	Type	Units	Description	Table 283: OnboardingSetPhaseRes ponse JSON structure
<i>last_set_phase</i>	OnboardingPhase			
<i>last_set_phase_state</i>	OnboardingPhaseState			
<i>status</i>	ResponseStatus		A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	

ONBOARDINGWAKEUPRESPONSE

The OnboardingWakeupResponse type has the following fields:

Field	Type	Description	Table 284: OnboardingWakeupRes ponse JSON structure
<i>charging_info</i>	Onboarding1p0ChargingI nfo	Whether or not Vector needs to charge after waking up.	
<i>waking_up</i>	bool	True if Vector is waking up.	

ONBOARDINGWAKEUPSTARTEDRESPONSE

The OnboardingWakeupStartedResponse type has the following fields:

Field	Type	Description	Table 285: OnboardingWakeupStart edResponse JSON structure
<i>already_started</i>	bool	True if the onboarding has completed	

59.4. ONBOARDING STATE

Post: “/v1/get_onboarding_state”

59.4.1 Request

The OnboardingStateRequest structure has no fields.

59.4.2 Response

The OnboardingStateResponse type has the following fields:

Field	Type	Description
<i>onboarding_state</i>	OnboardingState	Where in the onboarding process we are.
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 286:
OnboardingStateResponse JSON structure

59.5. ONBOARDING COMPLETE REQUEST

59.5.1 Request

The OnboardingCompleteRequest structure has no fields.

59.5.2 Response

The OnboardingCompleteResponse type has the following fields:

Field	Type	Description
<i>completed</i>	bool	True if the onboarding process has completed.

Table 287:
OnboardingCompleteResponse JSON structure

59.6. ONBOARDING WAKE UP REQUEST

59.6.1 Request

The OnboardingWakeUpRequest structure has no fields.

59.6.2 Response

The OnboardingWakeUpResponse type has the following fields:

Field	Type	Description
<i>already_started</i>	bool	True if the process of waking Vector up for onboarding has already been started.

Table 288:
OnboardingWakeUpResponse JSON structure

59.7. ONBOARDING WAKE UP STARTED REQUEST

59.7.1 Request

The OnboardingWakeUpStartedRequest structure has no fields.

59.7.2 Response

The OnboardingWakeUpStartedResponse type has the following fields:

Field	Type	Description	Table 289: OnboardingWakeUpStar tedResponse JSON structure
<i>charging_info</i>	Onboarding1p0ChargingInfo	The state of Vectors initial charging	
<i>waking_up</i>	bool	True if TBD	

60. PHOTOS

This section describes the commands and queries related to accessing and managing photographs (and their thumbnail images) on Vector. For a description of the photos, see Chapter 17 *Image Processing*.

60.1. STRUCTURES

60.1.1 Photo Path

The PhotoPathMessage type has the following fields:

Field	Type	Description	Table 290: PhotoPathMessage JSON structure
<i>full_path</i>	string		
<i>success</i>	bool	True if the photograph exists; otherwise, the photograph does not exist.	

60.1.2 Thumbnail Path

The ThumbnailPathMessage type has the following fields:

Field	Type	Description	Table 291: ThumbnailPathMessage JSON structure
<i>full_path</i>	string		
<i>success</i>	bool	True if the thumbnail image exists; otherwise, the thumbnail does not exist.	

60.2. EVENTS

60.2.1 PhotoTaken

The PhotoTaken event is sent after Vector has taken a photograph and stored it. This structure has the following fields:

Field	Type	Description	Table 292: PhotoTaken JSON structure
<i>photo_id</i>	uint32	The identifier of the photograph taken.	

60.3. DELETE PHOTO

This command is used to delete a photograph and its thumbnail
Post: “/v1/delete_photo”

60.3.1 Request

The DeletePhotoRequest has the following fields:

Field	Type	Description
photo_id	uint32	The identifier of the photograph to delete.

Table 293:
DeletePhotoRequest
JSON structure

60.3.2 Response

The DeletePhotoResponse type has the following fields:

Field	Type	Description
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
success	bool	True if the photograph was successfully removed; otherwise there was an error.

Table 294:
DeletePhotoResponse
JSON structure

60.4. PHOTO

This command is used to retrieve the photograph’s image.
Post: “/v1/photo”

60.4.1 Request

The PhotoRequest structure has the following fields:

Field	Type	Description
photo_id	uint32	The identifier of the photograph to request.

Table 295:
PhotoRequest JSON structure

60.4.2 Response

The PhotoResponse type has the following fields:

Field	Type	Description
image	bytes	The data that make up the photograph’s image
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
success	bool	True if the photograph was successfully retrieved; otherwise there was an error.

Table 296:
PhotoResponse JSON structure

60.5. PHOTOS INFO

This command is used to get the list of photographs available on Vector.

Post: “/v1/photos_info”

60.5.1 Request

The PhotosInfoRequest structure has no fields.

60.5.2 Response

The PhotosInfoResponse type has the following fields:

Field	Type	Description
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<i>photo_infos</i>	PhotoInfo[]	An array of information about the photographs available on Vector.

The PhotoInfo type has the following fields:

Table 297:
PhotosInfoResponse
JSON structure

Table 298: PhotoInfo
JSON structure

Field	Type	Description
<i>photo_copied_to_app</i>	bool	True if the photograph has been downloaded to the application.
<i>photo_id</i>	uint32	The identifier of this photograph. This can be used to retrieve the thumbnail, photograph or to delete it.
<i>thumb_copied_to_app</i>	bool	True if the thumbnail image has been downloaded to the application.
<i>timestamp_utc</i>	uint32	The time that the photograph was taken. The format is unix time: seconds since 1970, in UTC.

60.6. THUMBNAIL

This command is used to retrieve the thumbnail image of a photograph.

Post: “/v1/thumbnail”

60.6.1 Request

The ThumbnailRequest structure has the following fields:

Field	Type	Description
<i>photo_id</i>	uint32	The identifier of the photograph to request a thumbnail for.

Table 299:
ThumbnailRequest
JSON structure

60.6.2 Response

The ThumbnailResponse type has the following fields:

Field	Type	Description
<i>image</i>	bytes	The data that make up the thumbnail’s image
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.
<i>success</i>	bool	True if the thumbnail was successfully retrieved; otherwise there was an error.

Table 300:
ThumbnailResponse
JSON structure

61. SETTINGS AND PREFERENCES

This section describes the commands and queries related to settings and preferences on Vector.

For a description of the settings and what they mean, see Chapter 29 *Settings, Preferences, Features, and Statistics*. That chapter includes definitions for the following types:

- RobotSettingsConfig

61.1. STRUCTURES

61.1.1 AccountSettingsConfig

The AccountSettingsConfig type has the following fields:

Field	Type	Description
app_locale	string	
data_collection	boolean	

Table 301:
AccountSetting JSON
structure

61.2. UPDATE SETTINGS

Post: “/v1/update_settings”

61.2.1 Request

The UpdateSettingsRequest has the following fields:

Field	Type	Description
settings	RobotSettingsConfig	

Table 302:
UpdateSettingsRequest
JSON structure

61.2.2 Response

The UpdateSettingsResponse type has the following fields:

Field	Type	Description
code	ResultCode	
doc	JDoc	
status	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 303:
UpdateSettingsRespons
e JSON structure

61.3. UPDATE ACCOUNT SETTINGS

Post: “/v1/update_account_settings”

61.3.1 Request

The UpdateAccountsSettingsRequest has the following fields:

Field	Type	Description
<i>account_settings</i>	AccountSettingsConfig	

Table 304: JSON Parameters for *UpdateAccountSettings Request*

61.3.2 Response

The UpdateAccountsSettingsResponse type has the following fields:

Field	Type	Description
<i>code</i>	ResultCode	
<i>doc</i>	JDoc	
<i>status</i>	ResponseStatus	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 305: *UpdateAccountSettings Response* JSON structure

62. SOFTWARE UPDATES

These commands are siblings to the OTA Update and related commands in Chapter 12 Bluetooth LE protocol. However, they differ: in some cases, less information, in others present the same information in different ways.

62.1. ENUMERATIONS

62.1.1 **UpdateStatus**

The `UpdateStatus` enumeration has the following named values:

Name	Value	Description	Table 306: <i>UpdateStatus Enumeration</i>
<code>IN_PROGRESS_DOWNLOAD</code>	2	The software update is currently being downloaded.	
<code>NO_UPDATE</code>	0	There are no software updates available.	
<code>READY_TO_INSTALL</code>	1	The software update has been downloaded and is ready to install.	

62.2. START UPDATE ENGINE

“StartUpdateEngine cycles the update-engine service (to start a new check for an update) and sets up a stream of `UpdateStatusResponse` events.”

Post: “/v1/start_update_engine”

This command uses the same request and response structures as `CheckUpdateStatus`

62.3. CHECK UPDATE STATUS

“CheckUpdateStatus tells if the robot is ready to reboot and update.”

Post: “/v1/check_update_status”

62.3.1 **Request**

The `CheckUpdateStatusRequest` structure has no fields.

62.3.2 **Response**

This is streamed set of update status. The `CheckUpdateStatusResponse` type has the following fields:

Field	Type	Description	Table 307: <i>CheckUpdateStatusResponse JSON structure</i>
<code>expected</code>	int64	The number of bytes expected to be downloaded	
<code>progress</code>	int64	The number of bytes downloaded	
<code>status</code>	<code>ResponseStatus</code>	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.	
<code>update_status</code>	<code>UpdateStatus</code>		
<code>update_version</code>	string		

62.4. UPDATE AND RESTART

Post: “/v1/update_and_restart”

62.4.1 Request

The UpdateAndRestartRequest structure has no fields.

62.4.2 Response

The UpdateAndRestartResponse has the following fields:

Field	Type	Description
<code>status</code>	<code>ResponseStatus</code>	A generic status of whether the request was able to be carried out, or an error code indicating why it was unable to be carried out.

Table 308:
`UpdateAndRestartResponse` JSON structure

63. HISTORICAL ODDITIES

The Github repository is interesting for the SDK commands that were *removed*. The last version of the repository for many of these commands is:

<https://github.com/anki/vector-python-sdk/tree/c14082af5a947c23016111c1f73a445d8356dbf8>

Some commands removed from this repository (possibly later) because they were not implemented on Vector include:

- Ability to cancel remote
- Network statistics
- Ability to set and fetch the camera settings
- Motion observed events
- An audio stream from Vector’s microphone to the SDK application

CHAPTER 14

The Web Visualization Protocol

This chapter describes the internal web visualization (“web-viz”) and web-sockets interface.

64. WEBSOCKET OVERVIEW

Development builds of Vectors software include an optional web-visualization (webviz) tool, built on a web-socket interface. This protocol predates the HTTPS based protocol, which has since superseded, at least in some areas. Studying the protocol does provide some insight into the internal structures of Vectors modules.

Cavaet: This feature is not present in the production releases, nor many of the development releases. As this is a debugging tool, the schema for the data provided over the web-socket probably changed with each software version.

From the application level the protocol is organized into a stream of module-specific JSON messages that share a common web-socket, plus some management of that stream:

- Subscribe, and unsubscribe to module-specific events
- Receiving module-specific messages from the robot
- Sending module-specific message to the robot

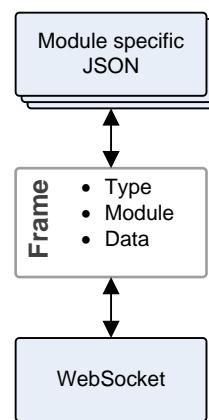


Figure 50: Overview of the websocket wrapper

Before we go further we'll need to know the module identifiers. The modules differ by ports that vend their events:

Module Name	Port(s)	Description
<i>Alexa</i>	8889	<i>Alexa related events</i>
<i>AnimationEngine</i>	8888	
<i>Animations</i>	8889	
<i>AudioEvents</i>	8889	
<i>BeatDetector</i>	8888	
<i>Behaviors</i>	8888	
<i>BehaviorConds</i>	8888	
<i>CloudIntents</i>	8888	
<i>Cpu</i>	8888	
<i>CpuProfile</i>	8888 8889	
<i>Cubes</i>	8888	
<i>Features</i>	8888	
<i>Habitat</i>	8888	
<i>HeldInPalm</i>	8888	
<i>IMU</i>	8888	
<i>Intents</i>	8888	
<i>MicData</i>	8889	
<i>Mood</i>	8888	8888
<i>NavMap</i>	8888	
<i>ObservedObjects</i>	8888	
<i>Power</i>	8888	
<i>Sleeping</i>	8888	
<i>SoundReactions</i>	8888	
<i>SpeechRecognizerSys</i>	8889	
<i>Touch</i>	8888	
<i>VisionScheduleMediator</i>	8888	

Table 309: Module names and their ports

64.1. SETTING UP THE COMMUNICATION CHANNEL

The URL to connect to is:

`ws://address:port/socket`

Where the address is the address of the Vector of interest, and the port is the shared port for the modules of interest.

To subscribe to events from a module post the following JSON structure to the web socket:

Field	Type	Description
<i>module</i>	string	The name of the module to subscribe to events from. See <i>Table 309: Module names and their ports</i> for the module names
<i>type</i>	string	“subscribe”

Table 310: Subscribing to a module's events

To unsubscribe from events from a module post the following JSON structure to the web socket:

Field	Type	Description
<i>module</i>	string	The name of the module to unsubscribe to events from. See <i>Table 309: Module names and their ports</i> for the module names
<i>type</i>	string	“unsubscribe”

Table 311: Unsubscribing from a module's events

Events related to the module will come with the following structure:

Field	Type	Description
<i>data</i>	JSON	The event structure for the given module
<i>module</i>	string	The name of the module to unsubscribe to events from. See <i>Table 309: Module names and their ports</i> for the module names

Table 312: Audio event parameters

To send a request to the module, it is wrapped with:

Field	Type	Description
<i>data</i>	JSON	The event structure for the given module
<i>module</i>	string	The name of the module to send the data to. See <i>Table 309: Module names and their ports</i> for the module names
<i>type</i>	string	“data”

Table 313: Audio event parameters

65. WEBSOCKETS JSON MESSAGES

65.1. RECEIVED EVENTS

65.1.1 Alexa State

These Alexa state events have the following fields:

Field	Type	Description
<i>authState</i>	uint	Whether or not Alexa is authenticated
<i>uxState</i>	uint	

Table 314: Alexa state event parameters

65.1.2 Animation Engine

These animation engine events have the following fields:

Field	Type	Description
<i>clip</i>	string	
<i>group</i>	string	
<i>headAngle_deg</i>	float	
<i>mood</i>	string	
<i>trigger</i>	string	

Table 315: Animation engine event parameters

65.1.3 Animation State

These animation state events includes a *type* field that describes how to interpret the rest of the structure. The events have the following fields:

Field	Type	Description
<i>animation</i>	string	the name of the animation
<i>type</i>	string	“start” if the animation is starting. “stop” if the animation has completed.

Table 316: Animation state event parameters

65.1.4 Audio events

See Chapter 23 for details of the audio events. The event includes a *type* field that describes how to interpret the rest of the structure.

When an audio event is sent, the event structure has the following fields:

Field	Type	Description
<i>eventName</i>	string	The name of the event (or its identifier) that was sent to the audio engine.
<i>gameObjectId</i>	string	The identifier that the audio engine will use for the object; this identifier is expected to link to items within Vector’s “game world” – his mental model.
<i>hasCallback</i>		
<i>time</i>		The time that this occurred at
<i>type</i>	string	“PostAudioEvent”

Table 317: Audio post event parameters

When all audio events are stopped, the event structure has the following fields:

Field	Type	Description
<i>eventName</i>	string	The name of the event (or its identifier) that was sent to the audio engine.
<i>gameObjectId</i>	string	The identifier that the audio engine will use for the object; this identifier is expected to link to items within Vector’s “game world” – his mental model.
<i>time</i>		The time that this occurred at

Table 318: Audio stop all event parameters

<code>type</code>	string	“StopAllAudioEvents”
-------------------	--------	----------------------

When an audio group is set to a new state, the event structure has the following fields:

Field	Type	Description
<code>stateGroupId</code>	string	The state group to modify
<code>stateId</code>		The new state that the group is being put into
<code>time</code>		The time that this occurred at
<code>type</code>	string	“SetState”

Table 319: Audio state group event parameters

When an audio switch group is set to a new state, the event structure has the following fields:

Field	Type	Description
<code>gameObjectId</code>	string	The new state that the switch is being changed into.
<code>switchGroupId</code>	string	The switch being modified.
<code>switchStateId</code>		The new state that the switch is being put into
<code>time</code>		The time that this occurred at
<code>type</code>	string	“SetSwitchState”

Table 320: Audio switch group event parameters

65.1.5 Beat Detector

See Chapter 16 section 73.5.3 *Beat Detector Outputs* for the event structures.

65.1.6 Behaviors

When a behavior event is sent, the event structure has the following fields:

Field	Type	Description
<code>activeFeature</code>	string	The name of the active AI feature. <i>optional</i>
<code>debugState</code>	string	<i>optional</i>
<code>stack</code>	string[]	An array of behavior identifiers. The last item is the current active behavior. <i>optional</i>
<code>time</code>		
<code>tree</code>	TreeNode[]	The behavior’s and their parents

Table 321: Behavior event parameters

The behavior TreeNode structure has the following fields:

Field	Type	Description
<code>behaviorID</code>	string	The identifier for a given behavior node
<code>parent</code>	string	The identifier of the parent node for this behavior

Table 322: TreeNode parameters

65.1.7

Behavior Conditions (BehaviorConds)

Note: “when the stack changes, we’ll receive new conditions before receiving the stack.”

When a behavior condition event is sent, the event structure has the following fields:

Field	Type	Description
<i>factors</i>	Factor[]	The current set of factors being examined. <i>optional</i>
<i>inactive</i>	string	The name of condition that is now inactive. <i>optional</i>
<i>owner</i>	string	The name of the owning behavior. <i>Optional</i> . Note: if factors array is present, this is not.
<i>stack</i>	string[]	An array of behavior identifiers. The last item is the current active behavior. <i>optional</i>
<i>time</i>		
<i>tree</i>	TreeNode[]	The behavior’s and their parents

Table 323: Behavior condition events parameters

The Factor structure has the following fields:

Field	Type	Description
<i>areConditionsMet</i>	boolean	True of the conditions for the behavior were met: they have all been true
<i>conditionLabel</i>	string	The name of the condition
<i>ownerDebugLabel</i>	string	The name of the owning behavior.

Table 324: Factor parameters

65.1.8

CPU Profile

When a CPU profile event is sent, the event structure has the following fields:

Field	Type	Description
<i>sample</i>	Sample[]	An array of CPU usage samples for the thread
<i>threadName</i>	string	The name of the thread that was measured for this CPU profile
<i>time</i>	float	Time stamp

Table 325: CPU profile parameters

The Sample structure has the following fields:

Field	Type	Description
<i>max</i>	float	The maximum CPU usage during the time interval
<i>mean</i>	float	The average CPU usage during the time interval
<i>min</i>	float	The minimum CPU usage during the time interval
<i>name</i>	string	The name of the thread that this belongs to

Table 326: Sample parameters

65.1.9 Habitat

The habitat event is a structure with the following fields:

Field	Type	Units	Description
<i>habitatState</i>	string		Whether or not Vector has detected his habitat (tray)
<i>reason</i>	string		
<i>stopOnWhiteEnabled</i>	string	"yes", "no"	
<i>whiteThresholds</i>			

Table 327: Habitat event parameters

65.1.10 IMU State

When a IMU state event is sent, the event structure has the following fields:

Field	Type	Description
<i>fall_impact_count</i>	uint	Fall impact count

Table 328: IMU state visualization structure

65.1.11 Microphone data (MicData)

See also section [65.1.15 SoundReactions](#)

The microphone data event is a structure with the following fields:

Field	Type	Units	Description
<i>activeState</i>	boolean		True if a voice has been detected.
<i>beatDetector</i>	BeatDetector		Information about the beat(s) heard
<i>delayTime</i>	float	ms	
<i>direction</i>	uint		The direction to the origin of the strongest sound.
<i>directions</i>	uint[]		The confidence in each of the directions
<i>latestNoiseFloor</i>	float		A measure of the ambient noise level.
<i>latestPowerValue</i>	float		A measure of the most recent sound loudness
<i>maxConfidence</i>	int		The maximum confidence value in the directions array
<i>selectedDirection</i>	uint		The direction that is picked for the origin of the sound of interest.
<i>time</i>	float		
<i>triggerDetected</i>	boolean		True if the trigger word has been detected

Table 329: Microphone data event parameters

The BeatDetector structure has the following fields:

Field	Type	Units	Description
<i>confidence</i>	float		How confident the analysis is in the tempo measurement.
<i>tempo_bpm</i>	float	beats / minute	The measured number of beats per minute.

Table 330: BeatDetector parameters

65.1.12 Mood

These structures are similar to, but differ from, those in Chapter 27.

When mood event is sent, the event structure has the following fields:

Field	Type	Units	Description
<i>emotionEvent</i>	string		The name of the event (see appendix J <i>Table 531: The emotion event names</i>). <i>Optional</i>
<i>info</i>	Info[]		
<i>moods</i>	Mood[]		An array emotion event structures (see below).
<i>simpleMood</i>	string		<i>Optional</i>
<i>time</i>	float	ms	

Table 331: Mood event parameters

The Emotion structure has following fields:

Field	Type	Description
<i>emotionName</i>	string	
<i>emotionType</i>	string	The dimension or type of emotion (“Happy”, “Confident”, “Stimulated”, “Social”, or “Trust”)
<i>max</i>	float	The maximum value that the dimension can take on
<i>min</i>	float	The minimum value that the dimension can take on

Table 332: Emotion structure parameters

The Info structure has following fields:

Field	Type	Description
<i>emotions</i>	Emotion[]	An array of emotions?
<i>simpleMoods</i>	dictionary	This dictionary maps a mood name to data to send to enter the mood

Table 333: Emotion Info structure parameters

The Mood structure has following fields:

Field	Type	Description
<i>emotion</i>	string	The dimension or type of emotion (“Happy”, “Confident”, “Stimulated”, “Social”, or “Trust”)
<i>value</i>	float	The value to add to the emotional state. The range is -1 to 1

Table 334: Emotion affecter parameters

65.1.13 Power

The power event is structure with the following fields (or an array of these structures):

Field	Type	Description
<i>powerSaveEnabled</i>	boolean	True if in power saving mode
<i>powerSaveRequesters</i>	string	

Table 335: Power event parameters

65.1.14 Sleeping

The sleeping event is a structure with the following fields:

Field	Type	Units	Description
<i>last_sleep_reason</i>	string		
<i>last_wake_reason</i>	string		
<i>reaction_state</i>	string		
<i>sleep_cycle</i>	string		
<i>sleep_debt_hours</i>	float	hours	

Table 336: Sleeping event parameters

65.1.15 SoundReactions

See also section 65.1.11 Microphone data (*MicData*)

The sound reaction event is a structure with the following fields:

Field	Type	Units	Description
<i>activeState</i>	boolean		True if a voice has been detected.
<i>confidence</i>	int		
<i>direction</i>	uint		The direction to the origin of the strongest sound.
<i>isTriggered</i>	boolean		True if the trigger word has been heard
<i>latestNoiseFloor</i>	float		A measure of the ambient noise level.
<i>latestPowerValue</i>	float		A measure of the most recent sound loudness
<i>powerScore</i>	float		
<i>powerScoreAvg</i>	float		
<i>powerScoreThreshold</i>	float		
<i>powerScoreMinThreshold</i>	float		
<i>selectedDirection</i>	uint		The direction that is picked for the origin of the sound of interest.
<i>time</i>	float	ms	
<i>triggerDirection</i>	uint		The reaction direction
<i>triggerConfidence</i>	int		The reaction confidence
<i>triggerScore</i>	float		The reaction score

Table 337:
SoundReactions event parameters

65.1.16 Speech Recognizer

The speech recognizer event is a structure with the following fields (or an array of these structures):

Field	Type	Units	Description
<i>endSampleIndex</i>	int	<i>index</i>	
<i>endTime_ms</i>	int	<i>ms</i>	
<i>result</i>			
<i>notch</i>			
<i>playback</i>			
<i>score</i>			
<i>startSampleIndex</i>	int	<i>index</i>	
<i>startTime_ms</i>	int	<i>ms</i>	

Table 338: Speech recognizer event parameters

65.1.17 Touch Sensor

The touch sensor event message has the following fields:

Field	Type	Description
<i>calibrated</i>	string	“yes” or “no”
<i>count</i>		
<i>enabled</i>	string	“true” “false”

Table 339: Touch sensor event parameters

65.1.18 Vision Schedule Mediator

The vision schedule event message has the following fields:

Field	Type	Units	Description
<i>fullSchedule</i>			
<i>numActiveModes</i>	uint		The number of vision modes that are currently enabled.
<i>patternWidth</i>			

Table 340: Vision schedule mediator event parameters

The Schedule structure has the following fields:

Field	Type	Units	Description
<i>offset</i>	uint		
<i>updatePeriod</i>	uint	<i>frames?</i>	
<i>visionMode</i>	string		The name of the vision processing step.

Table 341: Schedule parameters

65.2. POSTED EVENTS

65.2.1 Feature

The feature settings can be enabled, disabled, or reset. The posted structure includes a *type* field that describes how to interpret the rest of the structure.

To enable or disable a feature:

Field	Type	Units	Description
<i>name</i>	string		The name of the AI feature
<i>override</i>	string	"none", "enabled", "disabled"	Whether or not the feature should be enabled or disabled
<i>type</i>	string		"override"

Table 342:
Enable/disable Feature parameters

To reset all of the features to their default state:

Field	Type	Units	Description
<i>type</i>	string		"reset"

Table 343: Reset all features parameters

65.2.2 Touch Sensor

Can send a command to enable and disable the touch sensor:

Field	Type	Description
<i>enabled</i>	string	"true" to enable the touch sensor "false" to disable the touch sensor <i>Optional</i>
<i>resetCount</i>	string	"true" to reset the touch count. <i>Optional</i> .

Table 344: Touch sensor control parameters

CHAPTER 15

The Cloud Services

This chapter describes the remote servers that provide functionality for Vector.

- JSON document storage server
- The crash uploader
- The diagnostic logger
- The token/certificate system
- The natural language processing

66. CONFIGURATION

The server URLs are specified in

`/anki/data/assets/cozmo_resources/config/server_config.json`

The path to this JSON file is hardcoded in vic-cloud

Element	Description & Notes
<code>appkey</code>	A base64 token used to communicate with servers. “oDoa0quieSeir6goowai7f”
<code>check</code>	The server to use for connection checks
<code>chipper</code>	The natural language processing server
<code>jdocs</code>	The remote JSON storage server
<code>logfiles</code>	The server to upload log files to
<code>tms</code>	The token server where Vector gets authentication items like certificates and tokens

Table 345: The cloud services configuration file

The crash upload URL is given in

`/anki/etc/vic-crashuploader.env`

The OTA download URL is given in

`/anki/etc/update-engine.env`

The DAS server to contact is given in

`/anki/data/assets/cozmo_resources/config/DASConfig.json`

(This path is hardcoded in vic-DASMgr)

67. JDOCS SERVER

The Vic-Cloud services stores information on a “JDocs” server. This unusual name appears to be short for “JSON Documents.” Vic-Cloud uses the “jdocs” tag in the cloud services configuration file to know which server to contact. It uses the file

/anki/data/assets/cozmo_resources/**config/engine/jdocs_config.json**

to adjust how often it contacts the server. (The path to this JSON file is hardcoded in `libcozmo_engine.`)

The interactions are basic: store, read, and delete a JSON blob by an identifier. The description below³² gives the JSON keys, value format. It is implemented as gRPC/protobuf interaction over HTTP.

67.1. JDOCS INTERACTION

The JDdoc message has the following fields

Field	Type	Description
<code>client_meta</code>	string	
<code>doc_version</code>	uint64	
<code>fmt_version</code>	uint64	
<code>json_doc</code>	string	

Table 346: JSON Parameters for JDdoc request

67.2. DELETE DOCUMENT

67.2.1 Request

The DeleteDocReq request message has the following fields

Field	Type	Description
<code>account</code>	string	The account to read from
<code>doc_name</code>	string	The name of the document to delete.
<code>thing</code>	string	The thing id is a ‘vic:’ followed by the serial number

Table 347: JSON Parameters for JDdoc request

67.2.2 Response

The DeleteDocResp response message has the following fields:

Field	Type	Description
<code>latest_version</code>	uint64	The current version of the document in the repository.
<code>status</code>	string	

Table 348: JSON Parameters for JDdoc request

³² The protocol was specified in Google Protobuf. Vic-Cloud and Vic-Gateway were both written in Go. There is enough information in those binaries to reconstruct significant portions of the Protobuf specification in the future.

67.3. ECHO TEST

67.3.1 Request

EchoReq

- data

67.3.2 Response

EchoResp

- data

67.4. READ DOCUMENTS

67.4.1 Request

The ReadDocsReq request message has the following fields

Field	Type	Description
<i>account</i>	string	The account to read from
<i>items</i>	[]	Array of names document names(?)
<i>thing</i>	string	The thing id is a ‘vic:’ followed by the serial number

Table 349: JSON Parameters for JDoc read request

67.4.2 Response

ReadDocsResp

- items

The ReadDocsResp response message has the following fields:

Field	Type	Description
<i>items</i>	ReadDocsReq_i tem[]	The documents (?)

Table 350: JSON Parameters for JDoc read response

67.5. READ DOCUMENT ITEM

67.5.1 Request

The ReadDocsReq_Item request message has the following fields

Field	Type	Description	Table 351: JSON Parameters for JDoc read item request
<i>doc_name</i>	string	The name of the document to request	
<i>my_doc_version</i>	UInt64	The version to retrieve(?)	

67.5.2 Response

The ReadDocsResp_Item response message has the following fields:

Field	Type	Description	Table 352: JSON Parameters for JDoc read item response
<i>doc</i>	JDoc	The document structure.	
<i>status</i>	string		

67.6. WRITE DOCUMENT

67.6.1 Request

The WriteDocReq request message has the following fields

Field	Type	Description	Table 353: JSON Parameters for JDoc write request
<i>account</i>	string	The account to write to	
<i>doc</i>	JDoc	The document structure.	
<i>doc_name</i>	string	The name of the document to write.	
<i>thing</i>	string	The thing id is a ‘vic:’ followed by the serial number	

67.6.2 Response

The WriteDocResp response message has the following fields:

Field	Type	Description	Table 354: JSON Parameters for JDoc write response
<i>latest_doc_version</i>	uint64	The current version of the document in the repository.	
<i>status</i>	string		

68. NATURAL LANGUAGE PROCESSING

The audio after a “Hey Vector” is sent to servers for processing. They send a response back, in the form of an *intent*. This is a code and a structure that represents an action to carry out in response to the spoken request, query, or statement; it may represent the action requested, an answer to a query, or an action that emotionally responds to what was said. The intents received are listed in the “Cloud Intent” column in Appendix I, *Table 530: Mapping of different intent names*.

68.1. PARAMETERS FOR THE CLOUD INTENTS

The following are the parameters for each of the cloud intents.

This *intent_clock_settimer_extend* intent has the parameter following fields:

Field	Type	Units	Description
<i>timer_duration</i>	int	seconds	The number of seconds to set the timer to.

Table 355:
intent_clock_settimer_extend parameters

This *intent_global_delete_extend* intent has the parameter following fields:

Field	Type	Units	Description
<i>entity_behavior_deletable</i>	bool		

Table 356:
intent_global_stop_delete parameters

This *intent_global_stop_extend* intent has the parameter following fields:

Field	Type	Units	Description
<i>entity_behavior_stoppable</i>	bool		

Table 357:
intent_global_stop_extend parameters

This *intent_imperative_eyecolor_extend* intent has the parameter following fields:

Field	Type	Units	Description
<i>eye_color</i>	string		The name of the color to set the eye color to

Table 358:
intent_imperative_eyecolor_extend parameters

This *intent_imperative_volumellevel_extend* intent has the parameter following fields:

Field	Type	Units	Description
<i>volume_level</i>	string		

Table 359:
intent_imperative_volumellevel_extend parameters

This *intent_knowledge_response_extend* intent has the parameter following fields:

Field	Type	Units	Description
<i>answer</i>	string		The text to be spoken(?)
<i>query_text</i>	string		The text of the question asked(?)

Table 360:
intent_knowledge_response_extend parameters

This intent_message_playmessage_extend intent has the parameter following fields:

Field	Type	Units	Description
<i>given_name</i>	string		The name of the person to send the message to

Table 361:
intent_message_playmessage_extend parameters

This intent_names_username_extend intent has the parameter following fields:

Field	Type	Units	Description
<i>username</i>	string		The name of the user

Table 362:
intent_names_username_extend parameters

This intent_photo_take_extend intent has the parameter following fields:

Field	Type	Units	Description
<i>entity_photo_selfie</i>	string		Empty string if taking a photo, “photo_selfie” if taking a selfie.

Table 363:
intent_photo_take_extend parameters

This intent_weather_extend intent has the parameter following fields:

Field	Type	Units	Description
<i>condition</i>	string		The current weather conditions. One of “Clear”, “Cloudy”, “Cold”, “Rain”, “Snow”, “Stars”, “Sunny”, “Thunderstorms”, or “Windy”
<i>is_forecast</i>	string	“false” or “true”	“false” if it is the current weather conditions; “true” if forecasted weather conditions.
<i>local_datetime</i>	string		The local time (where the weather conditions apply) in UTC ISO 8601 format.
<i>speakable_location_string</i>	string		The location name that Vector could employ in his verbal description of the temperature.
<i>temperature</i>	string	degrees	The current or forecasted temperature, in the given units.
<i>temperature_unit</i>	string		F or C, for the units

Table 364:
intent_weather_extend parameters

69. LOG UPLOADER

The logs are uploaded by performing a HTTP PUT to the server. The URL is the “logfiles” URL in the server configuration file, with the compressed log file “/victor-\${esn}-\${timestamp}- \${pid}.log.gz” appended. The HTTP headers are:

HTTP header	Description
<i>Anki-App-Key:</i>	The appKey from the server configuration file.
<i>Usr-RobotESN:</i>	Vector’s serial number
<i>Usr-RobotOSRevision:</i>	The OS revision string from /etc/os-version-rev
<i>Usr-RobotOSVersion:</i>	The OS version string from /etc/os-version
<i>Usr-RobotRevision:</i>	The Anki revision string from /anki/etc/revision
<i>Usr-RobotTimestamp:</i>	The time of Vector’s internal clock.
<i>Usr-RobotVersion:</i>	The Anki version string from /anki/etc/version
<i>Usr-Username:</i>	

Table 365: Log upload
HTTP header fields

Note: the log uploader does not appear to be enabled in the production software.

70. CRASH UPLOADER

Minidumps are uploaded to a backtrace.io server. The URL (including the key) is hard coded in anki-crashuploader. This is done using a POST. The HTTP headers are:

Form fields	Description
<i>attachment_messages.log</i>	The “.log” file associated with the minidump. This is optional; only included if /run/das_allow_upload exists
<i>hostname</i>	<code> \${hostname}</code>
<i>robot.esn</i>	Vector’s serial number
<i>robot.os_version</i>	The OS version string from /etc/os-version
<i>robot.anki_version</i>	The Anki version string from /anki/etc/version
<i>upload_file</i>	The minidump “.dmp” file

Table 366: Crash
upload form fields

71. DAS MANAGER

DAS Manager uploads TBD to an Amazon “Simple Queue Service” (SQS) server. Amazon’s API uses the following key/value pairs in a URL encoded form:

Keys	Value	Table 367: DAS Manager SQS key-value pairs
Action	SendMessage	
MessageAttribute.1.Name	DAS-Transport-Version	
MessageAttribute.1.Value.DataType	Number	
MessageAttribute.1.Value.StringValue	2	
MessageAttribute.2.Name	Content-Encoding	
MessageAttribute.2.Value.DataType	String	
MessageAttribute.2.Value.StringValue	gzip, base64	
MessageAttribute.3.Name	Content-Type	
MessageAttribute.3.Value.DataType	String	
MessageAttribute.3.Value.StringValue	application/vnd.anki.json; format=normal; product=vic	
MessageBody		
Version	2012-11-05 ³³	

Note: there may be a body of compressed JSON data. These values are hardcoded in vic-dasmgr and libcozmo_engine

Unless you create an account and log in, Analytics Data is stored under a unique ID and not connected to you.

72. REFERENCES AND RESOURCES

Davis, Jason; *File Attachments in Backtrace*, Backtrace.io

<https://help.backtrace.io/en/articles/1852523-file-attachments-in-backtrace>

³³ This date is very far in the past, before Vector or Cozmo were developed. This was the time frame of the Overdrive product development.

[This page is intentionally left blank for purposes of double-sided printing]

PART IV

Advanced Functions

This part describes items that are Vector's primary function.

- **AUDIO INPUT.** A look at Vector's ability to hear spoken commands, and ambient sounds.
- **IMAGE PROCESSING.** Vector vision system is sophisticated, with the ability to recognize marker, faces, and objects; to take photographs, and acts as a key part of the navigation system.
- **MAPPING, NAVIGATION.** A look at Vector's mapping and navigation systems
- **ACCESSORIES.** A look at Vector's home (charging station), companion cube and custom objects.



Steph Dere

drawing by Steph Dere

[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 16

Audio Input

This chapter describes the sound input system:

- The audio input
- The audio filtering, and triggering of the speech recognition

73. AUDIO INPUT

The audio input is used to both give Vector verbal interaction, and to give him environmental stimulation:

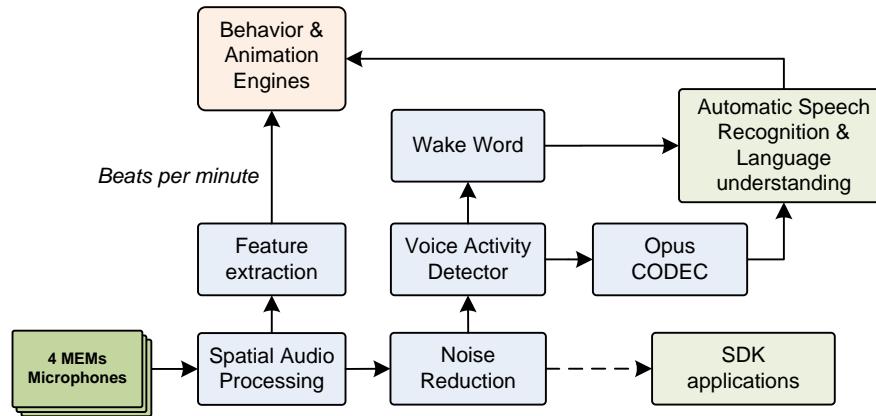


Figure 51: The audio input functional block diagram

- Spatial audio processing localizes the sound of someone talking from the background music.
- The feature extraction detects the ambient activity, and the tempo of the music. If the tempo is right, Vector will dance to it. This also provides basic stimulation to Vector.
- Noise reduction makes for the best sound.
- Voice activity detector usually triggered off of the signal before the beam-forming.
- A wake word is used to engage the automatic speech recognition system. *Note: the wake word is also referred to as the trigger word.*
- A CODEC is used to compress the audio before sending it to the remote server; Alexa Voice Services use the Opus audio CODEC.
- The speech recognition system is on a remote server. The audio sent to the automatic speech recognition system is compressed to reduce data usage.

The responsibility for these functions is divided across multiple processes and boards in Vector:

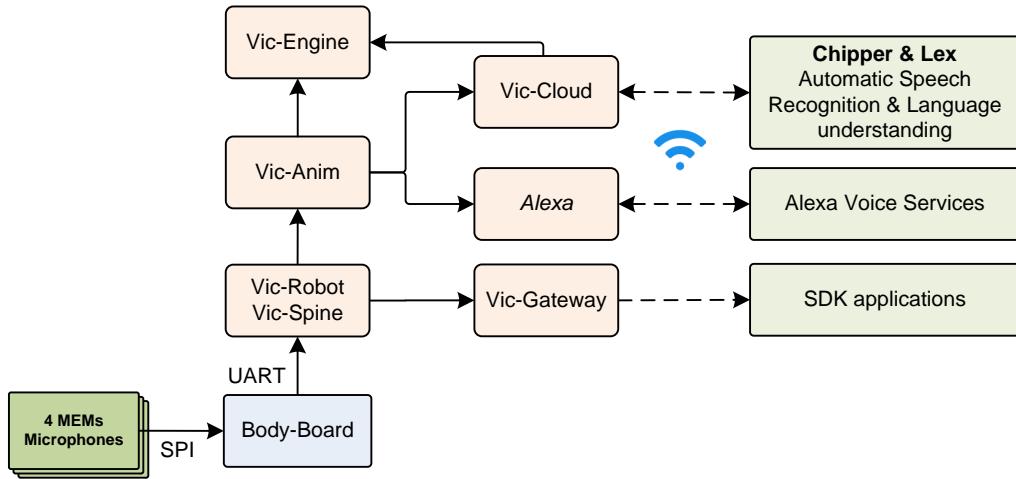


Figure 52: The audio input architecture

Note: providing the audio input to the SDK (via Vic-gateway) was never completed. It will be discussed based on what was laid out in the protobuf specification files.

The audio processing blocks, except where otherwise discussed, are part of Vic-Anim. These blocks were implemented by Signal Essence, LLC. They probably consulted on the MEMs microphones and their configuration. The Qualcomm family includes software support for these tasks, as part of the Hexagon DSP SDK; it is not known how much of this Signal Essence took advantage off.

73.1. THE MICROPHONES AND CONVERSION TO AUDIO SAMPLES

The microphone array is 4 far-field MEMs PDM microphones that sample the incoming sound and transfer the samples to body-board. (See Chapter 4 section 11.4.4 PDM Microphones for a description of the low-level bit-moving.)

microphone array architecture

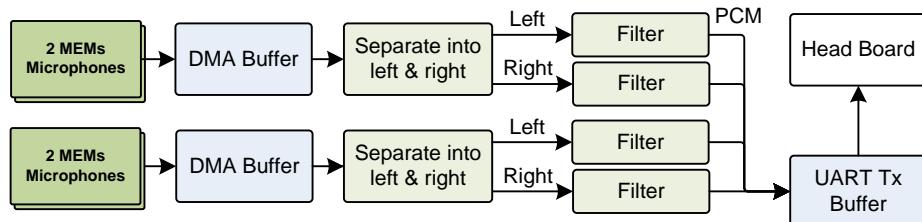


Figure 53: Sampling the microphones and converting to PCM format

The body-board samples each microphone at 1.5 M samples/sec – but at only 1 bit/sample! It passes the stream of samples thru a filter, produces audio at 15,625 samples/sec, with 16 bits/sample (effectively it may have anything in the range 10 to 16 bits, and padding out the rest). The filter also acts as a low pass filter, removing high frequency sampling artifacts. The most important part is that it preserves “phase information” so that the beam forming and direction finding steps work well. (More on this in a later section).

The audio samples are transferred to the Vic-spine module (part of Vic-robot) in regular communication with the head-board. The message from the body-board to the head board for sending 4 channels of audio samples has a payload size of 1028 bytes. This works out to 128 samples per channel (512 samples total).

The samples are extracted from the received message and forward to the Vic-Anim process. The software treats the audio as if its sample rate was 16,000 samples/sec. (“As a result, the pitch is altered by 2.4%.”) The signal processing is done in chunks of 160 samples.

73.1.1

The likely operation

Each SPI is configured to run at 3Mb/s (the slowest it can drive the microphones), using a DMA to transfer data into a buffer (~1536 bytes in size). The DMA’s are configured to use *circular mode*, where they never stop; instead they automatically wrap around to the start of the buffer after filling it. Because both SPI’s are tied together, only one DMA is configured to generate interrupts.

*transferring 4Mb/s
from 4 microphones
and filtering it into
PCM audio*

The input system triggers the SPIs to start gathering the data into their respective buffers. After that:

1. When the DMA has filled half of the buffer, it generates an interrupt. The filtering on all four channels is initiated for this half of the buffer, puts the result into the outgoing message buffer.
2. When the DMA has filled the second half of the buffer, it generates an end of transfer interrupt. The filtering on all four channels is initiated for this second half of the buffer, again putting the result into the outgoing message buffer. (In the mean time, the DMA has automatically looped back to the start of buffer and kept the SPI transferring the bits.)
3. If the outgoing buffer is full (i.e., after the DMA buffers have been filled twice), the UART transmit is initiated.

It is possible that the firmware uses two buffers, one that is filled by the filtering, and another that is sending data on the UART, and swapping every time it’s filled. It is more likely that only that the body-board *fills the same output buffer as data is being sent from it to the head-board*, to save on memory usage. Although the SPI is 2-3x faster than the UART, the filter stage takes 6 bits for every for every data bit that is sent to the head board. The UART can effectively send data at least 2x faster than the SPIs receive.

- Each microphone is driven at 1.5 M samples/sec (half the SPI clock frequency). The ratio between this input sample rate and the output sample rate (15,625) – called the *decimation* – is 96:1.
- Since it takes 96 input samples (bits) to get one output 16-bit sample³⁴, the bit-rate reduction is 6:1.

Altogether the audio sampling, filtering/decimation, and sending to the head-board uses at least 4KB of the MCU’s 8KB of RAM.

³⁴ The filtering may give the audio samples an effective range ~11 or 12 bits. The Customer Care Information Screen (CCIS) shows the microphones to be about 1024 when quiet.

73.2. SPATIAL AUDIO PROCESSING

The spatial audio processing uses multiple microphones to pick-out the wanted sound signal and cancel out the unwanted. Note: The spatial audio processing is bypassed until voice activity has been detected. The direction finding can isolate the audio from 12 different directions, each 30° wide:

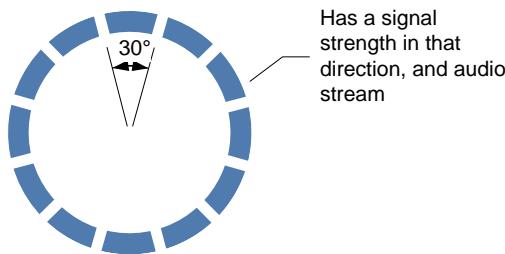


Figure 54: The audio can be isolated into one of 12 different directions

Note: forward is 0° .

The direction finding and isolation is typically a two-stage process:

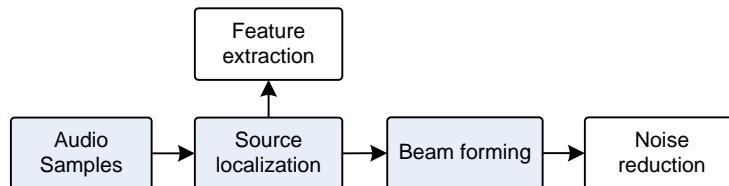


Figure 55: Typical spatial audio processing flow

THE SOURCE LOCALIZATION estimates direction of arrival of the person talking.

BEAM-FORMING combines the multiple microphone inputs to cancels audio coming from other directions.

The output of this stage includes:

- A histogram of the directions that the sound(s) in this chunk of audio came from. There are 12 bins, each representing a 30° direction.
- A measure of the background noise
- The direction that is picked for the origin of the sound of interest
- A confidence value for that direction
- The sound stream isolated for the picked direction, in the form of 160 16-bit PCM audio samples.

See also:

- Chapter 13, section *46.5 Audio Processing Mode* for a potential method to enable and disable the spatial sound processing;
- Chapter 13, section *46.4 Audio Feed (from the Microphones)* for potential access to the audio stream via the HTTPS API.

73.3. NOISE REDUCTION

Noise reduction identifies and eliminates noise and echo in the audio input:

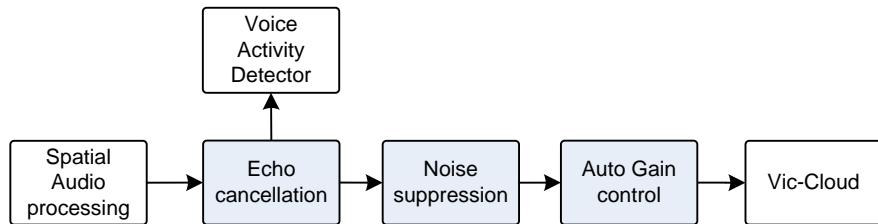


Figure 56: Typical audio noise reduction flow

ACOUSTIC ECHO CANCELLATION cancels slightly delayed repetitions of a signal.

NOISE SUPPRESSION is used to eliminate noise.

The combination of spatial processing and noise reduction gives the cleanest sound (as compared with no noise reduction and/or no spatial processing).

73.4. DETECTING ACTIVITY

Vector includes a module to detect sound activity (as distinguished from noise). The sound reaction behavior uses this to stimulate Vector from his sleep, get his attention, and encourage him to be more active. One way this could be done is through a set of filters to measure power levels:

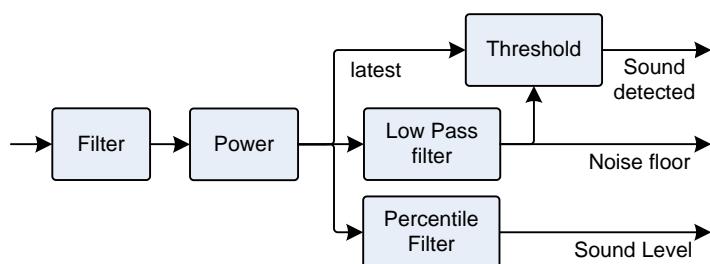


Figure 57: Sound and noise level estimator, and activity detector

The TBD {loudness estimator} might use an algorithm similar to the following steps:

1. First, the sound filter is to make the sound better reflect how our ears hear it, and/or remove elements that would cause false triggers. Two popular approaches are “equal loudness” by David Robinson and “a-weighting.” Both take into account how people perceive sounds loudness by giving less weight to some frequency regions (the very low and high), and more weight to others (the very middle).
2. Every few tens or hundreds of milliseconds the “power” level of the sound is computed. This is the logarithm of the root mean square (RMS) of the filtered values –squaring each value, averaging that, taking the square root and then computing its logarithm. Often this calculation is rearranged to be a bit faster, by skipping the square root and adjusting the logarithm scaling factor.
3. The computed power can then be compared against an estimate of the noise floor (the generic ambient sound level), to see if there is some activity, even the beat of a music.
4. The power levels are also tracked for a second (or a few seconds). The values could be averaged. Or the values could be sorted, from smallest to largest. The first value ~95% of the way into the sorted before could be used as the current overall sound level. (This avoids taking the loudest transient sound.).

The noise floor could be taken from the lowest value in the sorted array (step 4) – or the value that is, say, 5% into the array can be treated as the noise floor. Or it could be estimated by taking a low pass filter on the lowest values. The key is that so even though the sound level is increasing, the noise level is slow move up. A low pass filter has the advantage of not taking a large amount of memory – using a large percentile filter window (and using the lowest value in it) would take much more memory to prevent confusing several minutes of music with silence.

73.5. BEAT DETECTION

The details of how Vector's beat detection is implemented are not known, but beat detection is a common signal processing task. This section describes a typical implementation.

The beat detection is made of two related sub-functions. The first is a fast detector that can be used for quick dance responses in time to the music. The second finds the tempo – the beats per minute – of the music, which is also good indication that there is music playing (and not other activity).

Note: See chapter 23, section *107.1.1 Pitch tracker* for how to find the pitch.

73.5.1

A quick beat-detector

A simple responsive beat detector works by filtering the sound thru a band pass filter (say with a range of 100 Hz to 350 Hz) and then look for the magnitude to above a threshold:



Figure 58: Typical beat detection

Once a beat is detected, it holds off sending another event until the signal has dropped below a threshold for at least half a second or more. Another timer may be used to tell when the music has stopped: the timer is reset whenever a new beat is detected, and expires if a beat has not been detected for a few seconds.

Although simple to implement, loud noises can trick it, and it is not very good at measuring the tempo (the speed of the music in beats per second).

73.5.2

Tempo

A more accurate approach is to use a spectrogram to measure the tempo. The beats are very low frequency signals in the spectrograph. Music might be in the range of 50-110bpm (0.7Hz to 2Hz). The approach is to search the spectrogram in this frequency range for signals above a minimum threshold (to screen out generic sounds), and pick the strongest.

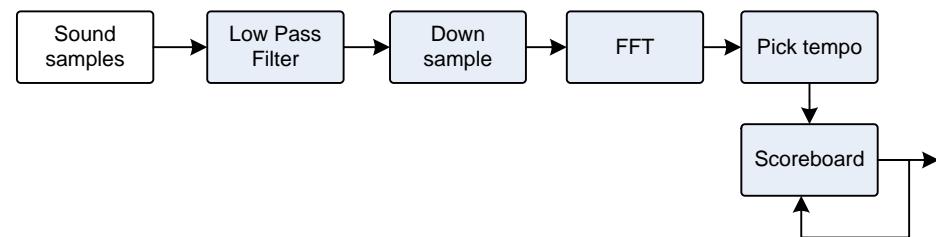


Figure 59: Typical tempo measurement

This will use an FFT to compute the spectrogram. It will need a wide window – a few seconds wide – to detect the beats, since they are at such a low frequency. This is down-sampled to a lower sampling rate – this reduces the memory required, and the amount of computation required to fit the task at hand. The basic algorithm is:

1. Take the sound input, and perform a low pass filter in it; this eliminates aliasing noises that can come from down-sampling
2. Next is to down sample the audio to only a few samples per second, and hold the results in a window a few seconds wide.
3. Periodically – every few seconds – an FFT is performed to create a new spectrograph. Note: the window can be “rolling” (instead of being thrown out and repopulated each time) to allow faster updates to the measured beats.
4. The FFT results are examined to find frequencies with a power above a threshold. These are the potential beats (in Hz)
5. The beats are then tracked in a scoreboard. The scoreboard tracks which beats are consistent and which are transitory. The beat-rates that haven’t been heard in a while are discounted or cleared out with time.
6. A beat rate, perhaps the highest persistent beats/minute, is then reported as the most likely rate.

The drawback of this approach is that it is “slow” and can’t be used to dance in time to the music with. The time window to find slower beats (the ones about every second) is very long, it can take a few seconds before it will have anything about the music beats.

73.5.3

Beat Detector Outputs

The beat detection modules produce several JSON structures for developer websocket and internal use. The main structure has the following fields:

Field	Type	Units	Description
<i>beatInfo</i>	BeatInfo[]		Information on the beat (or different possible tempos present in the music)
<i>detectorInfo</i>	DetectorInfo		Information from the detector on the beats

The *DetectorInfo* structure has the following fields:

Field	Type	Units	Description
<i>beatDetected</i>	boolean		True if a beat has just been detected.
<i>latestConf</i>	float		How confident the analysis is in the tempo measurement.
<i>latestTempo_bpm</i>	float	beats / minute	The measured number of beats per minute.
<i>possibleBeatDetected</i>	string	“yes” or “no”	Whether or not a potential music beat was detected.

Table 368: Beat detection event parameters

Table 369: *DetectorInfo* parameters

The BeatInfo structure has the following fields:

Field	Type	Units	Description
<i>aboveThresh</i>	boolean		Are the beats per minute above a threshold??
<i>conf</i>	float		How confident the analysis is in the tempo measurement.
<i>tempo_bpm</i>	float	<i>beats / minute</i>	The measured number of beats per minute.
<i>timeSinceBeat</i>	float	<i>seconds</i>	The number of seconds since the last beat; this can be useful for deciding that the music has stopped.

Table 370: BeatInfo parameters

73.6. RECORDING TO A FILE

The microphone module can store sound – either raw or processed – to a wave file. This may be for diagnostic purposes, left over as part of testing different microphone settings.

73.7. VOICE ACTIVITY DETECTOR AND WAKE WORD

The voice activity detector is given cleaned up sound from multiple microphones without beam-forming. When it detects voice activity, then the spatial audio processing is fully enabled.³⁵

Detecting that speaking is going on is more refined and specific than simply detecting that there is some interesting sound.

The voice activity detector and the wake word are used so that downstream processing – the wake word detection, and the automatic speech recognition system – are not engaged all the time. They are both expensive (in terms of power and CPU load), and the speech recognition is prone to misunderstanding.

When the voice activity detector triggers – indicating that a person may be talking – the spatial audio processing is engaged (to improve the audio quality) and the audio signals are passed to the Wake Word Detector.

The detector for the “Hey, Vector” is provided by Sensory, Inc. Pryon, Inc provided the detector for “Alexa.”³⁶ The recognition is locale dependent, detecting different wake words for German, etc. It may be possible to create other recognition files for other wake words.

When the “Hey, Vector” wake word is heard,

1. A connection (via Vic-Cloud) is made to the remote speech processing server for automatic speech recognition.
2. If there was an intent found (and control is not reserved), the intent is mapped to a local behaviour to be carried out. This is described in a later section.

73.7.1

Wake work configuration file

The configuration file for the wake word is located at:

/anki/data/assets/cozmo_resources/config/micData/micTriggerConfig.json

³⁵ Vector’s wake word detection, and speech recognition is pretty hit and miss. Signal Essence’s demonstration videos show much better performance. The differences are they used more microphones and the spatial audio filtering in their demos.

³⁶ This appears to be standard for Alexa device SDKs.

This file has dictionary structure with the following fields:³⁷

Field	Type	Description & Notes	Table 371: The micTriggerConfig JSON structure
<i>alexa_pryon</i>	WakeWordLocale[]	The wake word speech recognition models for Alexa in each of the supported locales, using models for Pryon’s toolkit (the default for Alexa Voice Services).	
<i>alexa_thf</i>	WakeWordLocale[]	The wake word speech recognition models for Alexa in each of the supported locales, using models for Sensory’s Truly HandsFree toolkit.	
<i>hey_vector_thf</i>	WakeWordLocale[]	The wake word speech recognition models for “Hey Vector” in each of the supported locales, using models for Sensory’s Truly HandsFree toolkit	

A WakeWordLocale is used to map a language locale to the wake word recognition models to use. This structure has the following fields:

Field	Type	Description & Notes	Table 372: The WakeWordLocale JSON structure
<i>defaultModelType</i>	string	e.g. “size_500kb” or “size_1mb”	
<i>locale</i>	string	The IETF language tag of the human companion’s language preference – American English, UK English, Australian English, German, French, Japanese, etc. default: “en-US”	
<i>modelList</i>	WakeWordModel[]	The wake word speech recognition models, in a variety of sizes	

³⁷ The names of the structures here were created for clarity; they are not actually used in the files.

Each WakeWordModel provides a set of word recognition models that can be used. The structure has the following fields:

Field	Type	Description & Notes
<i>dataDirectory</i>	string	The path (relative to the TBD) holding the recognition models.
<i>defaultSearchFileIndex</i>	uint	The index of the model (in <i>searchFileList</i>) to use by default.
<i>modelType</i>	string	e.g. “size_500kb” or “size_1mb”
<i>netFileName</i>	string	Name of a file.
<i>searchFileList</i>	WakeWordFile[]	The wake word speech recognition models, in a variety of sizes

Table 373: The WakeWordLocale JSON structure

Each WakeWordFile structure has the following fields:

Field	Type	Description & Notes
<i>searchFileIndex</i>	uint	The index of the model (in <i>searchFileList</i>) to use by default.
<i>searchFileList</i>	string	The name of the file...? (relative to the data directory). “NA” if a file name is not applicable.

Table 374: The WakeWordLocale JSON structure

73.8. CONNECTIONS WITH VIC-GATEWAY AND SDK ACCESS

An application has access to the wake-word events and the received user intent events as they occur. When the “Hey, Vector” wake word is heard,

1. A WakeWordBegin (see Chapter 13 section 46.2.3 *WakeWord*) event message is posted to Vic-Engine and Vic-Gateway. Vic-Gateway may forward the message on to a connected application.
2. A StimulationInfo (see Chapter 13, section 42.2.2 *StimulationInfo*) event message, an emotion event “ReactToTriggerWord,” is posted to Vic-Gateway for possible forwarding to a connected application.
3. A WakeWordEnd (see Chapter 13 section 46.2.3 *WakeWord*) event message is sent (to Vic-Gateway for possible forwarding to a connected application) when the Vic-cloud has received a response back. If control has not been reserved, and intent was received, the intent JSON data structure is included.
4. If there was no intent found, a StimulationInfo (see Chapter 13, section 42.2.2 *StimulationInfo*) event message is post (to Vic-Gateway), with an emotion event such as NoValidVoiceIntent
5. If there was an intent found (and control is reserved), a UserIntent (see Chapter 13, section 46.2.2 *UserIntent*) event is posted to Vic-Gateway for possible forwarding to a connected application. In this case, the intent will not be carried out.

An external application can send an intent to Vector using the AppIntent command (see Chapter 13, section 46.3 *App Intent*).

73.8.1

Audio Stream

It is clear that Anki made provisions to connect the audio stream to Vic-Gateway but were unable to complete the features before they ceased operation. The SDK would have been able to:

- Enable and disable listening to the microphone(s)
- Select whether the audio would have the spatial audio filter and noise reduction processing done on it.
- Include the direction of sound information from the spatial audio processing (see section 73.2 *Spatial audio processing*)
- 1600 audio samples; Note: this is 10x the chunk size of the internal processing size

74. CLOUD SPEECH RECOGNITION

The audio stream (after the “Hey Vector”) sent to a group of remote servers for processing. The servers perform automatic speech recognition (ASR), language understanding steps, and craft a response.

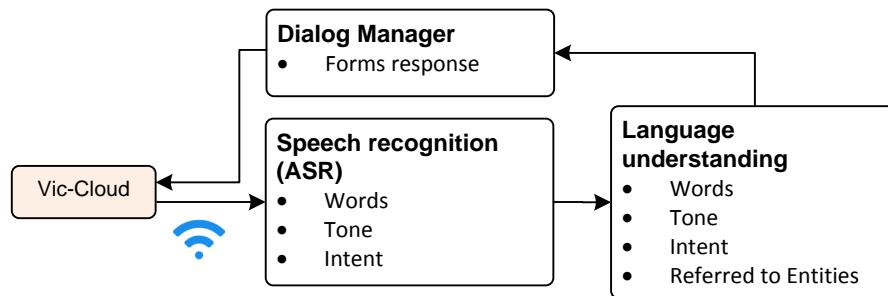


Figure 60: Typical speech recognition processes



What the user said is mapped to a *user intent*. This is a code and structure that represents an action to carry out in response to the spoken request, query, or statement; it may represent the action requested, an answer to a query, or an action that emotionally responds to what was said. The intent includes some supporting information – the colour to set the eyes to, for instance. Many of the phrase patterns and the intent they map to can be found in Appendix I. The intent may be further handled by Anki servers; the intent is eventually sent back to Vector.

The intent system does some replacement on the intent names and parameters³⁸ from the cloud and SDK application to names used internally within Vector’s engine.

³⁸ The complexity suggests that the development of the server, mobile application and Vector were not fully coordinated and needed this to bridge a gap.

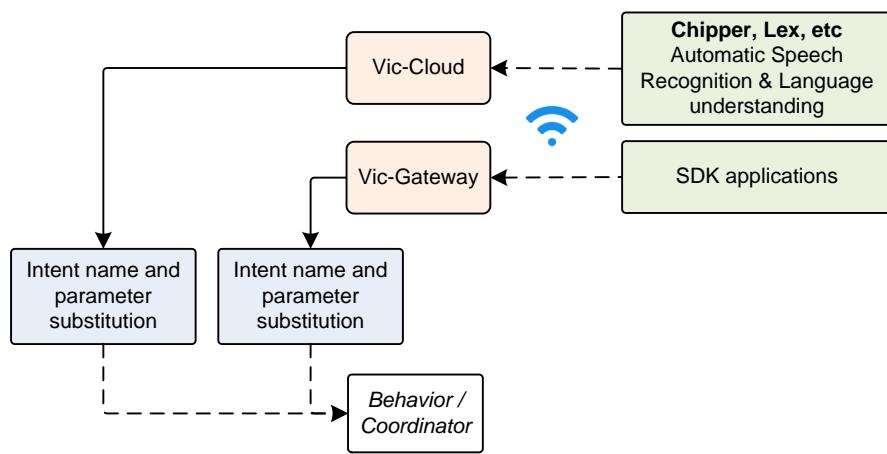


Figure 61: The filtering of intents

It uses separate tables for the intents passed by the cloud and those passed from an SDK application. With the cloud based intent,

1. Looks up to see if there is a rule matching the name of the passed intent. If there is no match, the intent (may be) is passed to the next stage. If the internal intent name associated with the rule will be used, and
2. Each of the passed intent parameter names is checked to see if the name should be changed to an internal name. If so it is changed to the internal name; otherwise the parameter's passed name is (probably) used.

The intents passed by the SDK application also go thru a filtering phase:

1. Looks up to see if there is a rule matching the name of the passed intent. If there is no match, the intent is *discarded*. If the internal intent name associated with the rule will be used, and
2. Each of the passed intent parameter names is checked to see if the name should be changed to an internal name. If so it is changed to the internal name; otherwise the parameter is *discarded*.

The intent is also checked to see if it is enabled. Each intent can be associated with a feature flag; if it is, the flag is looked up to see if the corresponding feature is enabled. (see also Chapter 29 section *130 Feature Flags*).

An intent may initiate a behavior, or a coordinator. A coordinator is receptive to further intents in addition to physical stimulation.

74.1. INTENT PARAMETERS

The `global_delete` intent has the parameter following fields:

Field	Type	Units	Description
<code>what_to_stop</code>	string		

Table 375:
`global_delete`
parameters

The global_stop intent has the parameter following fields:

Field	Type	Units	Description
<i>what_to_stop</i>	string		

Table 376: *global_stop* parameters

This imperative_eyecolor_specific intent has the parameter following fields:

Field	Type	Units	Description
<i>eye_color</i>	string		The name of the color to set the eye color to

Table 377: *imperative_eyecolor_specific* parameters

This imperative_volumelvel intent has the parameter following fields:

Field	Type	Units	Description
<i>volume_level</i>	string		

Table 378: *imperative_volumelvel* parameters

This knowledge_response intent has the parameter following fields:

Field	Type	Units	Description
<i>answer</i>	string		The text to be spoken(?)
<i>query_text</i>	string		The text of the question asked(?)

Table 379: *knowledge_response* parameters

This meet_victor intent has the parameter following fields:

Field	Type	Units	Description
<i>username</i>			

Table 380: *meet_victor* parameters

This message_playback intent has the parameter following fields:

Field	Type	Units	Description
<i>given_name</i>	string		The name of the person to send the message to

Table 381: *message_playback* parameters

The set_timer intent has the parameter following fields:

Field	Type	Units	Description
<i>time_s</i>	int	seconds	The number of seconds to set the timer to

Table 382: *set_timer* parameters

The take_a_photo intent has the parameter following fields:

Field	Type	Units	Description
<i>empty_or_selfie</i>	string		Empty string if taking a photo, “photo_selfie” if taking a selfie.

Table 383: *take_a_photo* parameters

This test_name intent has the parameter following fields:

Field	Type	Units	Description
<i>name</i>	string		

Table 384: *test_name* parameters

This test_timeWithUnits intent has the parameter following fields:

Field	Type	Units	Description
<i>time</i>	uint		
<i>units</i>	string		

Table 385: *test_timeWithUnits* parameters

This weather_response intent has the parameter following fields:

Field	Type	Units	Description
<i>condition</i>	string		The current weather conditions. One of “Clear”, “Cloudy”, “Cold”, “Rain”, “Snow”, “Stars”, “Sunny”, “Thunderstorms”, or “Windy”
<i>isForecast</i>	string	“false” or “true”	“false” if it is the current weather conditions; “true” if forecasted weather conditions.
<i>localDateTime</i>	string		The local time (where the weather conditions apply) in UTC ISO 8601 format.
<i>speakableLocationString</i>	string		The location name that Vector could employ in his verbal description of the temperature.
<i>temperature</i>	string	degrees	The current or forecasted temperature, in the given units.
<i>temperatureUnit</i>	string		F or C, for the units

Table 386: *weather_response* parameters

74.2. INTENT MAPPING CONFIGURATION FILE

The configuration file holding the mapping of the clouds external intent names and parameters to those used internally within Vector’s engine is located at:

/anki/data/assets/cozmo_resources/**config/engine/behaviorComponent/user_intent_map.json**

The path is hard coded into libcozmo_engine.so. The file has the following structure:

Field	Type	Description
<i>user_intent_map</i>	array of X	A table that maps the intent received by the cloud or application to the intent name used internally
<i>unmatched_intent</i>	string	The intent to employ if cloud’s intent cannot be found in the table above. Default: “unmatched_intent”

Table 387: The *user_intent_map* JSON structure

Each of the mapping entries has the following structure:

Table 388: The intent mapping JSON structure

Field	Type	Description
<i>app_intent</i>	string	The intent name sent by the SDK application. See the “App Intent” column in Appendix I <i>Table 530: Mapping of different intent names</i> for a list of intent names. <i>Optional</i> .
<i>app_substitutions</i>	dictionary	A dictionary whose keys are the keys provided by the application’s intent structure, and maps to the keys used internally. <i>Optional</i> .
<i>cloud_intent</i>	string	The intent name returned by the cloud. See the “Cloud Intent” column in Appendix I <i>Table 530: Mapping of different intent names</i> for a list of intent names.
<i>cloud_numerics</i>	array of strings	Names of keys that used as parameter values by the behaviour..?? <i>Optional</i> .
<i>cloud_substitutions</i>	dictionary	A dictionary whose keys are the keys provided by the cloud’s intent structure, and maps to the keys used internally. <i>Optional</i> .
<i>feature_gate</i>	string	The name of the feature that must be enabled before this intent can be processed. <i>Optional</i> .
<i>test_parsing</i>	bool	Default: true. <i>Optional</i> .
<i>user_intent</i>	string	The name of the intent used internally within Vector’s engine.

75. REFERENCES AND RESOURCES

<https://github.com/ARM-software/ML-KWS-for-MCU>

A reference keyword listener for ARM microcontrollers.

<https://github.com/MTG/essentia/tree/master/test/src/descriptortests/equalloudness>

A reference implementation of an equal-loudness measure

Hydrogen Audio, *ReplayGain 1.0 specification*, 2018 Nov 19

http://wiki.hydrogenaud.io/index.php?title=ReplayGain_1.0_specification

A detailed description of how the sound loudness can be measured and used to adjust the volume of music playback.

Note: the filter implementation for audio effects can be very complex; for sound detection it is very simple.

ST Microelectronics, *Reference manual, STM32F030x4/x6/x8/xC and STM32F070x6/xB advanced ARM®-based 32-bit MCUs*, 2017 Apr, Rev 4

https://www.st.com/resource/en/reference_manual/dm00091010-stm32f030x4-x6-x8-xc-and-stm32f070x6-xb-advanced-arm-based-32-bit-mcus-stmicroelectronics.pdf

Wikipedia, *A-weighting*

<https://en.wikipedia.org/wiki/A-weighting>

Wikipedia, *Robinson–Dadson curves*

<https://en.wikipedia.org/wiki/Robinson%20%93Dadson%20curves>

CHAPTER 17

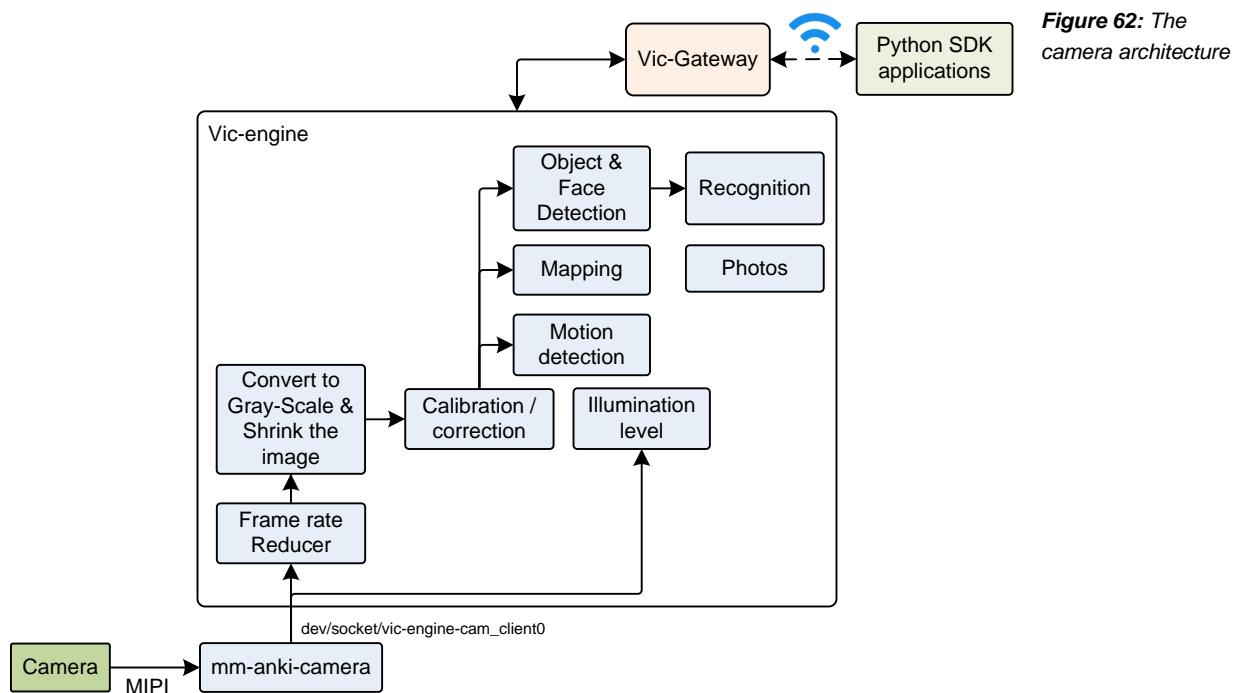
Image Processing

This chapter describes Vector's image processing system:

- Camera operation including calibration
- Visual stimulation
- Recognizing symbols and specially marked objects
- Detecting faces, recognizing people, and estimating emotion
- Hand, pet and other object detection
- Taking photos
- Sending video stream to the SDK
- Vision is primarily used for navigation purposes: Recognizing the floor (or ground), odometry and “simultaneous localization and mapping”

76. CAMERA OPERATION

Vector has a 1280x720 camera with a wide field of view to see around it without moving its head, similar to how an animal can see a wide area around it by moving its eyes. The camera is connected to the processor through a MIPI interface. The data from the camera passes to device drivers, then to a separate daemon service and eventually passes to Vic-engine for the processing:



Vector visually recognizes some elements in its environment:

- Special visual markers; Vector treats all marked objects as moveable... and all other objects in its driving are as fixed & unmovable.
- Faces
- Hands
- Pets (feature not completed)
- Other objects, like fruit, etc. (feature not completed)
- LASER pointers (feature not completed)

76.1. CAMERA OPERATION

To reduce computing load the camera frame rate is reduced and the image size is scaled down. This pattern is common throughout the image processing:

- More pixels require much more memory at each stage of image the image processing.
- It takes much, much longer (and more power) to process larger frames. There is the added time to process each of the added pixels. Second, the neural-net models (used for human, pet and object recognition) are much larger as well, taking much longer to process with the many stages involved in these models.
- That extra processing is among the most power expensive items in Vector, and rapidly depleting his battery, shortening the time between charges,
- The extra processing also generates heat in the head board, and
- Image processing tasks don't need more pixels. There is rarely any improvement in visual detection from using more pixels or higher frame rates

The software in reduces the frame rate by skipping frames (no fancy interpolation needed). Then the image is converted to gray scale and scaled down to quarter size (640x360). (This was also the case with Cozmo.)

76.2. CAMERA CALIBRATION

The camera is calibrated at manufacturing time. This is necessary so that the Vector can accurately dock with a cube, getting the small lift fingers into the cube's holes. The calibration primarily compensates for the image being slightly offset, and unit to unit variation of focal length.

Vector's camera has ~120° diagonal *field of view*.³⁹ For comparison the iPhone's camera has a 73° field of view, and the human eye is approximately 95°. The cropped sensor image has a 90° horizontal field of view and a 50° vertical field of view.

³⁹ The press release for Vector reported a 120° field of view, but should be discounted as this number does not match the frame field of view numbers given in the SDK documentation.

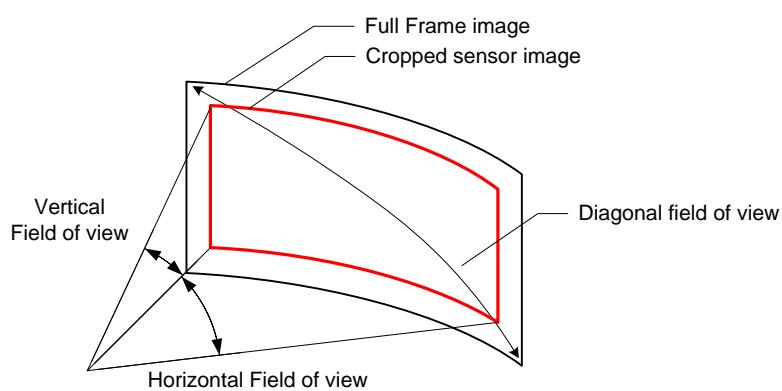


Figure 63: The camera field of view

Vector's calibration uses focal length instead of field of view. The two values are related:

$$\text{fieldOfView} = 2 \arctan \frac{\text{sensor size}}{2 \text{focalLength}}$$

The following structure is reported in the robot logs for the camera calibration:

Field	Type	Description
<i>cx</i>	float	"The position of the optical center of projection within the image. It will be close to the center of the image, but adjusted based on the calibration of the lens at the factory."
<i>cy</i>	float	
<i>distortionCoeffs</i>	float[]	
<i>fx</i>	float	The "focal length combined with pixel skew (as the pixels aren't perfectly square), so there are subtly different values for x and y."
<ify< i=""></ify<>	float	
<i>ncols</i>	int	The width of the image in pixels. The value given is 640.
<i>nrows</i>	int	The height of the image in pixels. The value given is 360
<i>skew</i>	float	

Equation 1:
Relationship between field of view and focal length

Table 389: The camera calibration JSON structure

Quotes are from Anki Cozmo SDK.

"A full 3x3 calibration matrix for doing 3D reasoning based on the camera images would look like:"

$$\begin{pmatrix} \text{focalLength}_x & 0 & \text{center}_x \\ 0 & \text{focalLength}_y & \text{center}_y \\ 0 & 0 & 1 \end{pmatrix}$$

Equation 2: Camera calibration matrix

76.3. CORRECTION

With each image frame to be processed, the software applies some processing to improve the image contrast. This helps with the low-light that is common in rooms and at night. (The software also monitors the illumination levels and tweaks the exposure settings so that image is as good as possible before it gets to the software stage.)

One technique to improve the contrast is *contrast-limited adaptive histogram equalization* (CLAHE). This looks at a histogram of the pixel gray-scale values in a wide window around each pixel, and then maps the used gray-scale to a different set of gray values. This spreads the grays out a bit further. Being adaptive, it helps with different lighting levels across the scene – some areas being well lit, others in shadow.

76.4. VISION MODES

The vision process happen at different rates, many execute together in a shared group.

The vision processing system has many detectors, and functions. Some have their software run at different rates. While most are independent of each other, they are often grouped together.

Table 390: The Vision processes

Vision Mode	Executes with	Description and notes
<i>AutoExp</i>		This mode is used to control the auto-exposure control level.
<i>AutoExp_Cycling</i>	<i>AutoExp</i>	This mode is used to detect
<i>AutoExp_MinGain</i>	<i>AutoExp</i>	
<i>BrightColors</i>		This mode is used to detect colors that may be interesting to explore.
<i>Faces</i>		Used for face detection, and to trigger facial identification.
<i>Faces_Blink</i>	<i>Faces</i>	This mode is used to detect and count eye blinks.
<i>Faces_Crop</i>	<i>Faces</i>	This mode is used to detect faces that are obscured or partly out of view.
<i>Faces_Expression</i>	<i>Faces</i>	This mode is used to estimate the facial expression.
<i>Faces_Gaze</i>	<i>Faces</i>	Detects the gaze and looks deep into their eyes with wonder and the hope of biscuits.
<i>Faces_Smile</i>	<i>Faces</i>	This mode is used to detect smiles.
<i>Illumination</i>		This mode is used to estimate the level of illumination in the scene.
<i>Lasers</i>		This mode is used to detect laser pointer activity.
<i>Markers</i>		Detects Vector's special square marker symbols.
<i>Markers_Off</i>	<i>Markers</i>	
<i>Markers_ChargerOnly</i>	<i>Markers</i>	This part of the process of detecting Vector's special square marker symbols.
<i>Markers_Composite</i>	<i>Markers</i>	This part of the process of detecting Vector's special square marker symbols.
<i>Markers_FastRotation</i>	<i>Markers</i>	This part of the process of detecting Vector's special square marker symbols.
<i>Markers_FullFrame</i>	<i>Markers</i>	This part of the process of detecting Vector's special square marker symbols.
<i>Markers_FullHeight</i>	<i>Markers</i>	This part of the process of detecting Vector's special square marker symbols.

<i>Markers_FullWidth</i>	Markers	This part of the process of detecting Vector's special square marker symbols.
<i>MirrorMode</i>		Displays the camera image on the LCD
<i>Motion</i>		The mode is used to detect visual motion
<i>OverheadEdges</i>		
<i>OverheadMap</i>		disabled
<i>People</i>		This mode is used to detect people, rather than faces
<i>Pets</i>		This mode is used to detect pets, such as cats and dogs.
<i>Hands</i>		Used to detect hands (for purposes of pouncing on them).
<i>SaveImages</i>		This mode is used to save the camera image as a photograph.
<i>Stats</i>		This is probably used to compute statistics about the images or image processing
<i>Viz</i>		This module creates a marked up image showing where Vector sees the charger, cubes, faces, and other interesting things.
<i>WhiteBalance</i>		This mode is used to estimate the

76.5. ILLUMINATION LEVEL SENSING

Vector estimations the amount of illumination in the room. Dark rooms would encourage him to go to sleep, while bright or changing illumination would encourage him to be active. The illumination is pretty each to compute: sum up the brightness of each pixel in the image, or the number of pixels above a threshold of brightness.

The camera is also used as an ambient light sensor when Vector is in low power mode (e.g. napping, or sleeping). In low power mode, the camera is suspended and not acquiring images. Although in a low power state, it is still powered. The software reads the camera's auto exposure/gain settings and uses these as an ambient light sensor. (This allows it to detect when there is activity and Vector should wake.)

76.6. VISUAL MOTION DETECTION

Note: this is not the same as chapter 10, which sensed Vector's motion.

Vector can detect visual movement in its field of view. This motion detector looks in two regions of the camera view (the low left and the top right) for movement, and it looks at its projected view of the ground for movement.

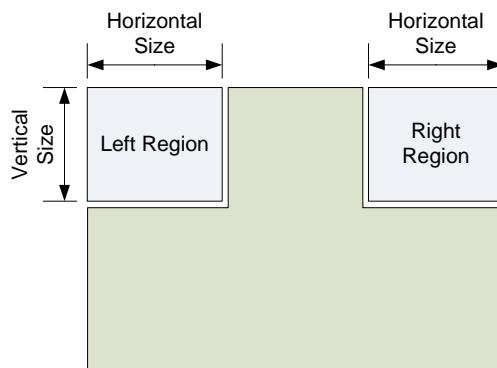


Figure 64: The movement detection regions

The detector likely does pixel subtraction in these regions between frames, computing a score for the number of pixels that changed and how wide of an area it was (the centroid). Then it adds this with the past value (using the inverse of DecreaseFactor as a weight). If score is above a threshold (MaxValue?) it concludes that there is motion in that region.

See Section 82.1.7 *MotionDetector* for a description of the motion detectors configuration.

The motion detector is used by the pouncing behaviors – see Chapter 28, section 121.2 *Pouncing*

The RobotObservedMotion event (see Chapter 13 section 54.2.1 *RobotObservedMotion*) is intended to indicate when visual motion is detected. {Note: this event is not supported in current software}

77. THE CAMERA POSE: WHAT DIRECTION IS CAMERA POINTING IN?

The camera is located in Vector's head. The pose of Vector's camera – its position and orientation, including its tilt up or down, can be estimated from Vector's pose, the angle of his head, the known position of the camera within the head and the position of the joint around which the head swivels. Note: the values are for Cozmo, but are assumed to be representative of Vector:

```
# Neck joint relative to robot origin
NECK_JOINT_POSITION = [-13, 0, 49]

# camera relative to neck joint
HEAD_CAM_POSITION = [17.52, 0, -8]
DEFAULT_HEAD_CAM_POS = list(HEAD_CAM_POSITION)

DEFAULT_HEAD_CAM_ROTATION = [
    0,      -0.0698,   0.9976,
    -1,      0,          0,
    0,      -0.9976,  -0.0698 ]

# Compute pose from robot body to camera
# Start with a pose defined by the DEFAULT_HEAD_CAM_ROTATION (rotation matrix)
# and the initial position DEFAULT_HEAD_CAM_POS
default_head_pose = Matrix3d(DEFAULT_HEAD_CAM_ROTATION, DEFAULT_HEAD_CAM_POS)

# Rotate that by the head angle
rotation_vector = RotationVectorAroundYAxis(-robot.head_angle.radians);
current_head_pose = default_head_pose.rotate_by(rotation_vector)

# Get the neck pose (transform the initial offset by the robot's pose)
neck_pose = TransformPose(NECK_JOINT_POSITION, robot.pose)

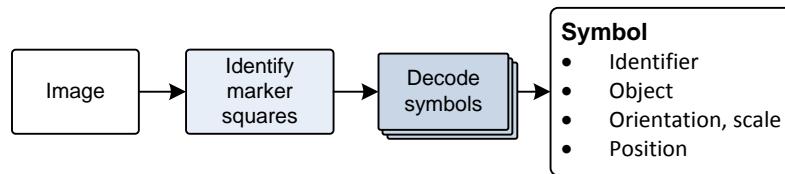
# Precompose with robot-to-neck-pose
camera_pose = current_head_pose.pre_compose_with(neck_pose);
```

Example 6: Computing the camera pose
source: Anki⁴⁰

⁴⁰ <https://forums.anki.com/t/camera-matrix-for-3d-positionning/13254/5>

78. MARKERS

Anki considered QR codes to mark accessories and special items... but they were universally rejected in the feedback received during development. So Anki created their own visual labeling system, starting with Cozmo. Vector has a newer set of visual labels that is not compatible with Cozmos. (There isn't a clear reason for the incompatibility.) The algorithm used is among the most documented of Anki's internally developed modules for Vector.



A key characteristic of the markers is a big, bold square line around it:

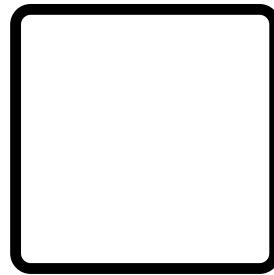


Figure 65: The processing of the image for symbols and objects

Figure 66: A typical rectangle around the visual markers

The square is used to estimate the distance and relative orientation (pose) of the marker and the object is on. Vector, internally, knows the physical size of marker. The size of the square in the view — and being told how big the shape really is —lets Vector know enough to compute the likely physical distance to the marked item. And since the “true” mark has parallel lines, Vector can infer the pose (relative angles) of the surface the mark is on.

The process of finding and decoding the marker symbols is very straightforward, since there is quite a lot known about the structure of the marker image ahead of time. This allows the use of computation friendly algorithms.

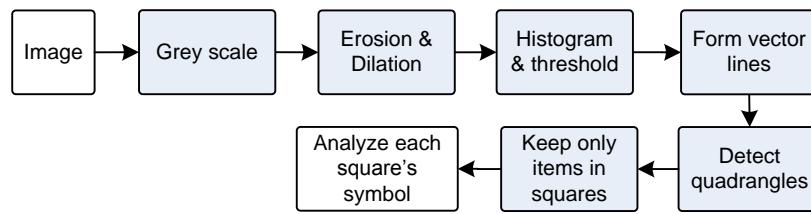


Figure 67: Preparing image for scanning for symbols

The steps in processing are:

1. Acquire a gray scale image,
2. Apply classic erosion-dilation and Sobel transforms to build a vector representation (no pun intended) of the image; this is most familiar as “vector drawing” vs bitmap images
3. Detect the squares – the parallel and perpendicular lines – in the vector drawing. This will be the potential area that a symbol is in.
4. Analyze square to determine its size, and affine transform – how it is tilted up-and-down, and tilted away from the camera.

5. Screen the squares, tossing out those that are horribly distorted,
6. Analyze the pixels in the square to identify the code

78.1. THE INITIAL PREPARATION STEPS

The image is initially prepared for analysis by:

1. The image is converted to grey scale, since color is of no value.
2. Performs (erosion, dilation) that strip out noise, fill in minor pixel gaps. There are no small features, so fine detail is not important.
3. The image is then converted to high-contrast black and white (there is no signal in grey scale). This is done by performing a histogram of the grey scale colors, finding a median value. This value is used as a threshold value: greys darker than this are considered black (a 1 bit), and all others are white (0 bit).

78.2. DETECT AND ANALYZE SQUARES

The detection of squares then:

1. Typically a pair of Sobel filters is applied to identify edges of the black areas, and the gradients (the x-y derivative) of the edges.
2. The adjacent (or nearby) pixels with similar gradients are connected together into a list. Straight line segments will have very consistent gradients along them. In other words, the bitmap is converted into a vector drawing. In jargon, this is called the *morphology*.
3. The lists of lines are organized into a containment tree. A bounding box (min and max positions of the points in the list) can be used to find which shapes are around others. The outermost shape is the boundary.
4. “Corners of the boundaries are identified... by filtering the (x,y) coordinates of the boundaries and looking for peaks in curvature. This yields a set of quadrilaterals (by removing those shapes that do not have four corners).” Stein, 2017
5. A perspective transformation is computed for the square (based on the corners), using homography (“which is a mathematical specification of the perspective transformation”). This tells how tilted the square is.
6. The list of squares is filtered, to keep those that are big enough to analyze, and not distorted with a high skew or other asymmetries.

78.3. DECODING THE SQUARES

The next step is to decode the symbol. Vector has a set of probe locations within the marker square that it probes for black or white reading. These are usually centered in the cells of a grid.

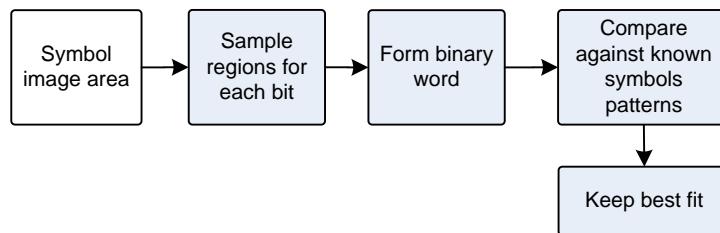


Figure 68: Decoding the symbol

The steps in decoding the symbol are:

1. The software uses the perspective transform to map the first point location to one in the image;
2. The pixels at that point in the image are sampled and used to assign a 0 or 1 bit for the sample point.
3. The bit is stored, in a small binary word
4. The above steps are repeated for the rest of the probe locations

This process allows Vector to decode images warped by the camera, its lens, and the relative tilt of the area.

Next, the bit patterns are compared against a table of known symbol patterns. The table includes multiple possible bit patterns for any single symbol, to accommodate the marker being rotated. There is always a good chance of a mistake in decoding a bit. To find the right symbol, Vector:

1. XOR's the decoded bit pattern with each in its symbol table,
2. Counts the number of bits in the result that are set. (A perfect match will have no bits set, a pattern that is off by one bit will have a single bit set in the result, and so on.)
3. Vector keeps the symbol with the *fewest* bits set in the XOR result.

78.4. REVAMPING SIZE AND ORIENTATION

The different rotations of the symbol would change the order that it sees the bits. Each bit pattern in the table might also include a note on how much the symbol is rotated (i.e. 0, +90°, -90°, or 180°). When matching a bit pattern, Vector can know the major rotation of the symbol.

Combined with the angle of the symbol square, the full rotation of the symbol can be computed.

78.5. INFERRING KNOWLEDGE ABOUT OBJECTS

Vector associates an object with symbol. Some objects can have many symbols associated with them. Cubes have different symbols used for sides of cubes. This allows Vector to know what object it is looking at, and what side of the object. And, with some inference, the orientation of the object.

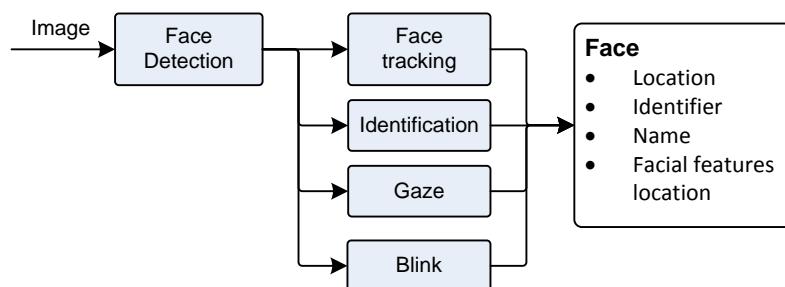
Vector knows (or is told) the physical size of the symbol, and the object holding the symbol.

Combining this with the visual size of the object, time of flight distance measurement (if any), and Vector's known position, this allows Vector infer the objects place in the map.

79. FACE AND FACIAL FEATURES RECOGNITION

Vector “is capable of recognizing human faces, tracking their position and rotation (“pose”) and assigning names to them via an enrollment process.” Vector’s facial detection and recognition is based on the OKAO vision library. This lets Vector know when one (or more) people are looking at it. This library is primarily used by Vector for facial recognition tasks:

- Face detection ability – the ability to sense that there is a face in the field of view, and locate it within the image.
- Face recognition, the ability to identify whose face it is, looking up the identify for a set of known faces
- Recognize parts of the face, such as eyes, nose and mouth, and where they are located within the image.



Anki Cozmo SDK

Figure 69: The face detection and recognition processes

There are a couple of areas that Vector includes access to in the SDK API, but did not incorporate fully into Vector’s AI:

- The ability to recognize the facial expression: happiness, surprise, anger, sadness and neutral. This is likely to be unreliable; that is the consensus of research on facial expression software.
- Ability to estimate the direction of gaze

And there are several features in OKAO that are not used

- The ability to estimate the gender and age of the person
- Human upper body detection
- Hand detection and the ability to detect an open palm. The hand detection used in Vector is done in a different way (which we will discuss in a section below.)

79.1. FACE DETECTION

OpenCV also has facial detection, but not recognition. OpenCV’s classic face detector is an implementation of an algorithm developed by Viola-Jones. Since we know how that works, we can discuss it as representative of how OKAO may work. Viola-Jones applies a series of fast filters (called a “cascade” in the jargon) to detect low-level facial features (called Haar feature selection) and then applies a series of classifiers (also called a cascade). This divides up interesting areas of the image, identify facial parts, and makes conclusions about where a face is.

Vector’s face detector (and facial recognition) can’t tell that it is looking at an image of face – such as a picture, or on a computer screen – rather than an actual face. One thing that Anki was considering for future products was to move the time of flight sensor next to the camera. This

Daniel Casner, 2019

would allow Vector to estimate the size of the face (and its depth variability) but measuring the distance.

Side note: Anki was exploring ideas (akin to the idea of object permanence) to keep track of a known person or object in the field of view even when it was too small to be recognized (or detected).

79.2. FACE IDENTIFICATION AND TRAINING

When it sees a face, it forms a description of the facial features using twelve points:

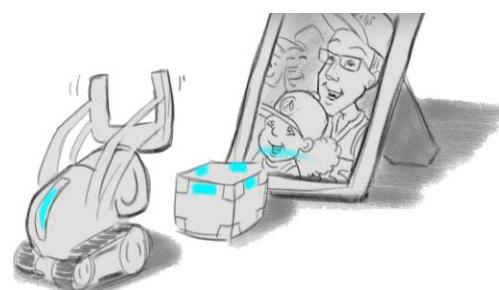
- Each eye has three points,
- The nose has two,
- The mouth has four points

If you introduce yourself to Vector by voice, you are permitting the robot to associate the name you provide with Facial Features Data for you. Facial Features Data is stored with the name you provide, and the robot uses this data to enhance and personalize your experience and do things like greet you by that name. This data is stored locally on the robot and in the robot's app. It is not uploaded to Anki nor shared, and you can delete it anytime.

79.3. COMMUNICATION INTERFACE

There are several commands to manage the faces that Vector recognizes, and to keep informed of the faces that Vector sees. See Chapter 13 section *52 Faces* for more details.

- The Enable Face Detection (see Chapter 13 section *52.5 Enable Face Detection*) command enables and disables face detection and analysis stages.
- The RobotChangedObservedFaceID and RobotObservedFace (see Chapter 13 section *52.3.4 RobotChangedObservedFaceID* and *52.3.5 RobotObservedFace*) events are used to indicate when a face is detected, and tracking it: the identity of the face (if known), where it is in the field of view, the facial expression, where key parts of the face are (in the view), etc
- The Set Face to Enroll (see Chapter 13 section *52.10 Set Face to Enroll*) command is used to ability assign a name to face, and the Update Enrolled Face By ID (see Chapter 13 section *52.10 Set Face to Enroll*) command is used to change the name of a known face
- The Request Enrolled Names (see Chapter 13 section *52.9 Request Enrolled Names*) command is used to retrieve a list the known faces
- The ability to remove a facial identity (see Chapter 13 section *52.7 Erase Enrolled Face By Id*), or all facial entities (see Chapter 13 section *52.6 Erase All Enrolled Faces*)
- The Find Faces (see Chapter 13 section *52.8 Find Faces*) command initiates the search for faces



drawing by Jesse Easley

80. TENSORFLOW LITE, DETECTING HANDS, PETS... AND THINGS?

Vector includes support to detect hands, and has preliminary support for detecting pets and a wide variety of objects. These are done using TensorFlow Lite⁴¹ (aka TFLite), an inference only neural-net discriminator.

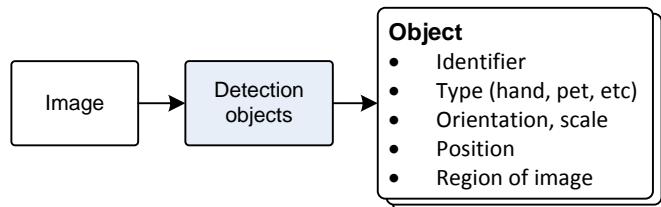


Figure 70: The processing of the image for symbols and objects

Vector's hand detection is done with a custom TensorFlow Lite DNN model.⁴² Vector also has a custom person detector; this may be used to quickly identify whether there is a face in view before engaging the potentially more expensive OKAO framework.

80.1. DETAILS ON TENSORFLOW LITE

From a distance, the TensorFlow Lite framework acts much the same as a classification trees, taking inputs, examining properties and producing a result, such as “this is a hand.” Internally the framework is designed as a modular virtual machine for signal-processing-like computation. A “model” is the program for this virtual machine, with information describing its memory structures, inputs, outputs and the instructions. The analog of a software procedure in the model are called a *graph*. The instructions are called *operations*. Full TensorFlow supports 800+ different operations out of the box, and custom ones can be added. TensorFlow Lite supports 122+ different operations, and custom ones can be added as well. TensorFlow Lite supports one graph in a model.

Warden & Situnayake, 2019

The host application has to do preprocessing such as feature extraction, prepare the input for the system. For instance, the image must be converted to grey scale and scaled down to 128 pixels by 128 pixels. (More pixels require much more memory and processing steps often with no improvement in detection; some higher quality models do use slightly larger image sizes.)

Then each of the operations in the model is carried out. An operation might perform a simple calculation like summing values, keeping the smallest or largest, etc or an operation might be a complex calculation such as a convolution. Once all of the operations have completed, the results are not a “this is a hand” or other conventional software result. Instead, the results are a big list of values on how confident it is for each possible item. An application typically chooses the top item or two as the output – if their confidence is high enough.

⁴¹ Since Tensorflow Lite was both introduced at the end of 2017, there has been a steady trickle of improvements to TensorFlow Lite. There is a lower power version that targets microcontrollers.

⁴² There are four different hand detector models – only one is used – which suggests that the hand detector was actively being tweaked and improved.

TensorFlow Lite includes build time support to replace the key operation implementations with fast, processor-specific ones:

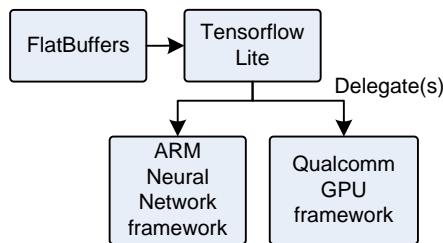


Figure 71:
TensorFlow lite with
hardware specific
accelerators

In addition, applications using TensorFlow Lite can provide their own, faster or more efficient implementations of operations.

Each TensorFlow Lite model is probably run in its own thread. The benchmarks posted by TensorFlow⁴³ using smartphones to run model tens to hundreds of milliseconds. Putting each model on its own thread and waiting for posted results allows the rest of the processing to execute in a consistent fashion.

80.1.1 SalientPoint data structure

The SalientPoint JSON data structure is produced from the neural networks analysis of the image. These points are used by the behavior system as something interesting to react to as well. The structure has the following fields:

Field	Type	Units	Description
<i>area_fraction</i>	float		Area of the region that was identified as a salient point.
<i>color_rgba</i>	uint	RGBA	How to color this region, for web visualization
<i>description</i>	string		
<i>salientType</i>	string		An enumerated type describing the kind of salient point found.
<i>score</i>	float		A metric relating how interesting the point/region is
<i>shape</i>	array of points?		An array of points outlining the interesting area?
<i>timestamp</i>	uint	ms	The timestamp that this point was identified on
<i>x_img</i>	float	pixel	Pixel coordinate of the upper left hand corner of the region.
<i>y_img</i>	float	pixel	Pixel coordinate of the upper left hand corner of the region.

Table 391: SalientPoint parameters

⁴³ <https://www.tensorflow.org/lite/performance/benchmarks>

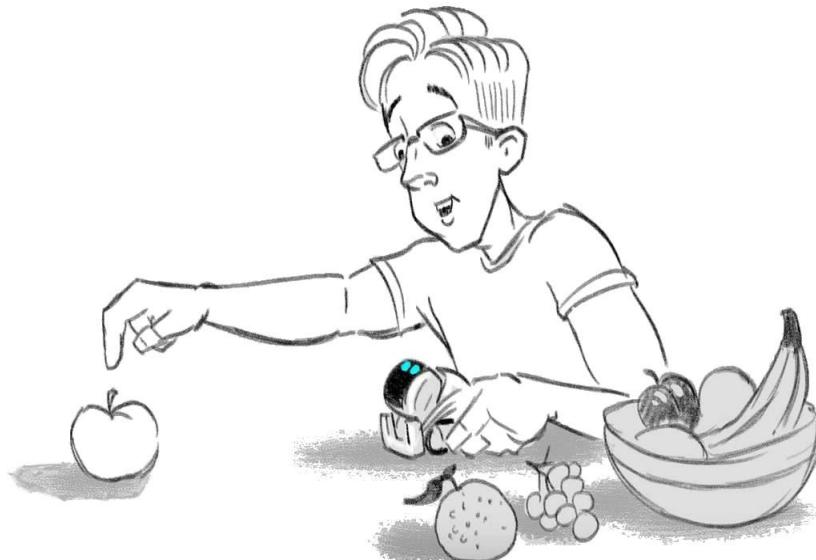
80.2. OTHER IDEAS THAT WEREN'T FULLY REALIZED AND FUTURE POTENTIAL

Vector also includes the stock MobileNet V1 (0.5, 128) model to classify images, although it does not appear to have been used yet. This model was likely intended to give Vector the ability to identify a wide variety of things, and pets.⁴⁴

MobileNet V1 includes higher quality models than the one employed that may be explored. Since this model was released, a version 2 and version 3 of MobileNet have been developed and released. Version 2 is reported to be faster, higher quality, and/or require fewer processor resources. (Version 3 is slower and takes more processor resources, but is much more accurate.)

The configuration file shows experimentation with MobileNet V2 (using 192x192 input images), but it was disabled.

drawing by Jesse Easley



⁴⁴ Or a special model for recognizing pets may have been under development

81. PHOTOS/PICTURES

Vector has the ability to take pictures. The photographs were taken with less than the full camera resolution. (It isn't known if Anki intended to eventually take photographs at a higher resolution.) These pictures are stored on Vector, not in the cloud. The mobile application and SDK applications can view, delete or share pictures taken by Vector.

The camera/image processing pipeline in Vector is entirely focused on his AI features with as low as practical battery impact. The images available for taking a picture are not filtered, or cleaned up, so the pictures that Vector takes are noisy and smaller.

Commentary: The quality of photos seen on a mobile phone is achieved using a camera processing pipeline to enhance the images, removing noise and applying special filters to reconstruct textures. It is conceivable that the camera processing framework(s) from Qualcomm and Android could be added to an open-source Vector. That would come at the cost of battery performance, heat, and potentially overwhelm the memory resources (there are still bugs in Vector where the memory use becomes too high, and the system thrashes, slowing noticeably down and eventually crashes.)

It is more practical, in a future open-source Vector, to export the raw camera images (in its RAW format and at different illumination levels) and process the images on a PC or mobile device. The availability of sophisticated image processing frameworks are much wider for those devices. See Chapter 13, section *54 Image Processing* for the camera access API.

81.1. COMMUNICATION INTERFACE

There are several commands to manage the photographs that Vector has taken. See Chapter 13 section *60 Photos* for more details.

- The PhotoTaken event (see Chapter 13 section *60.2.1 PhotoTaken*) is used to receive a notification when Vector has taken a photograph.
- The Photos Info (see Chapter 13 section *60.5 Photos Info*) command is used to retrieve a list of the photographs that Vector currently has
- The Photo (see Chapter 13 section *60.4 Photo*) command is used to retrieve a photo
- The Delete Photo command (see Chapter 13 section *60.3 Delete Photo*) removes a photo from the system
- The Thumbnail (see Chapter 13 section *60.6 Thumbnail*) command retrieves a small version of the image, suitable for displaying as a thumbnail

82. CONFIGURATION FILES

82.1. VISION CONFIG

The vision system' main configuration file is located at:

/anki/data/assets/cozmo_resources/config/engine/vision_config.json

This path is hardcoded into libcozmo_engine.so. It configures each of the image processing module, and the schedule defaults. The file is a structure with the following fields:

Field	Type	Description	Table 392: The vision configuration JSON structure
<i>ColorImages</i>	boolean	“whether color images are enabled on startup (can still be toggled later)”	
<i>FaceAlbum</i>	string		
<i>FaceRecognition</i>	struct	Configures when the face recognition runs	
<i>GroundPlaneClassifier</i>	struct	Configuration of the ground plane classifier	
<i>IlluminationDetector</i>	struct	Configuration of the illumination detector, and a link to the configuration file for the classifier	
<i>ImageCompositing</i>	struct	Configuration of the image compositing module	
<i>ImageQuality</i>	struct	Controls the cameras auto-exposure settings.	
<i>InitialModeSchedule</i>	Struct	“VisionModes that need to be scheduled by default go here. Basically things that are always running.”	
<i>MotionDetector</i>	struct	Configuration of the motion detection – the size of image area to look for peripheral motion	
<i>NeuralNets</i>	NeuralNets struct	Configures the use of the TensorFlow Lite detection modules	
<i>NumOpenCvThreads</i>	int	“Number of threads to use with OpenCV. 0 means no threading according to docs. Only affects calls from VisionSystem thread.”	
<i>OverheadMap</i>	struct	The size of the overhead map	
<i>PerformanceLogging</i>	struct	Configures how often to log information about the image processing stats	
<i>PetTracker</i>	struct	Configures the pet tracker – the number of pets, face size, thresholds, etc.	

82.1.1 FaceRecognition

The FaceRecognition structure includes the following fields:

Field	Type	Description	Table 393: The FaceRecognition structure
<i>RunMode</i>	string	Can be either “asynchronous” or “synchronous”	

82.1.2

GroundPlaneClassifier

The GroundPlaneClassifier (also called “desk classifier”) is used to visually identify where the driving surface is. The structure includes the following fields:

Field	Type	Description	
<i>FileOrDirName</i>	string	Path to the ground plane classifier file.	
<i>MaxDepth</i>	uint		
<i>MinSampleCount</i>	uint		
<i>OnTheFlyTrain</i>	booleab		
<i>PositiveWeight</i>	float		
<i>TruncatePrunedTree</i>	boolean		
<i>Use1SERule</i>	boolean		

Table 394: The *GroundPlaneClassifier* structure

Note: The Ground Plane classifier is a bit unusual. It is one of only two YAML files. The YAML file is an openCV based classifier tree, instead of TensorFlow Lite. This suggests it may have been older (i.e. from Cozmo), and/or it may have been more efficient to implement in openCV.

82.1.3

IlluminationDetector

The IlluminationDetector structure includes the following fields:

Field	Type	Description	
<i>AllowMovement</i>	boolean	If true, “continue to run even if robot is moving”	
<i>ClassifierConfigPath</i>	string	Path to the illumination classifier configuration file	
<i>DarkenedMaxProbability</i>	float	The “max probability to result in ‘Darkened’ class”	
<i>FeaturePercentileSubsample</i>	uint	“Stride for building histogram to compute percentiles”	
<i>IlluminatedMinProbability</i>	float	The “min probability to result in ‘Illuminated’ class”	

Table 395: The *IlluminationDetector* structure

82.1.4

InitialModeSchedules

The InitialModeSchedules provides the default frequency that each vision processing step is run. (And step not listed here, the default is that it is not scheduled to run). The structure includes the following fields:

Field	Type	Description	
<i>AutoExp</i>	uint	Run the auto exposure step every n frames.	
<i>Markers</i>	uint	Run the markers step every n frames.	
<i>WhiteBalance</i>	uint	Run the white balance step every n frames	

Table 396: The *InitialModeSchedules* structure

82.1.5 ImageCompositing

The ImageCompositing The structure includes the following fields:

Field	Type	Description
<i>imageReadyPeriod</i>	uint	
<i>numImageReadyCyclesBeforeReset</i>	uint	
<i>percentileForMaxIntensity</i>	uint	

Table 397: The ImageCompositing structure

82.1.6 ImageQuality

The ImageQuality structure includes the following fields:

Field	Type	Description
<i>CyclingTargetValues</i>	float[]	“When ‘cycling’ exposure is enabled, the target values to use.” Each value must be in the range 0 to 255
<i>HighPercentile</i>	float	Range 0.0 to 1.0
<i>InitialExposureTime_ms</i>	uint	“Sent each time we request camera calibration”
<i>LowPercentile</i>	float	Range 0.0 to 1.0
<i>MaxChangeFraction</i>	float	“Relative amount we can change current exposure/wb each update, zero disables.” Range: 0.0 to inf
<i>MeterFromDetections</i>	boolean	“Base auto-exposure on detected markers, faces, etc, if any”
<i>RepeatedErrorMessageInterval_ms</i>	uint	The “time between error messages once triggered”
<i>SubSample</i>	uint	
<i>TargetPercentile</i>	float	Range 0.0 to 1.0
<i>TargetValue</i>	uint	“Try to make targetPercentile have this value.” Range 0 to 254
<i>TimeBeforeErrorMessage_ms</i>	uint	“How long [the] Vision System must detect ‘bad’ quality before notifying game”
<i>TooBrightValue</i>	uint	“‘Too Bright’ if LowPercentile is above this”
<i>TooDarkValue</i>	uint	“‘Too Dark’ if HighPercentile is below this”

Table 398: The ImageQuality structure

82.1.7

MotionDetector

The MotionDetector structure includes the following fields:

Field	Type	Description
<i>CentroidStability</i>	float	“How quickly should peripheral motion detection track the source of motion.”
<i>DecreaseFactor</i>	float	“The higher this number, the more quickly motion detection forgets about motion.”
<i>HorizontalSize</i>	float	“Fraction of the width of the image to be used for peripheral motion detection (right and left).”
<i>IncreaseFactor</i>	float	“The higher this number, the more sensitive is motion detection to motion.”
<i>MaxValue</i>	float	“The higher this value, the sooner peripheral motion detection will be triggered.”
<i>VerticalSize</i>	float	“Fraction of the height of the image to be used for peripheral motion detection (top)”

Table 399: The MotionDetector structure

82.1.8

NeuralNets

The NeuralNets structure includes the following fields:

Field	Type	Description
<i>Models</i>	Model[]	An array of TensorFlow Lite model configuration structures (see below).
<i>ProfilingEventLogFrequency_ms</i>	uint	How often to log information about the model execution timing.
<i>ProfilingPrintFrequency_ms</i>	uint	How often to print (to TBD) information about the model execution timing.

Table 400: The NeuralNets structure

The `Models` is an array of structures. Each structure has the following fields:

Field	Type	Description
<code>architecture</code>	string	
<code>benchmarkRuns</code>	uint	If non-zero, gather duration and resource usage information for each run.
<code>graphFile</code>	string	The name of the TensorFlow Lite file (.tflite) to load to implement this model
<code>inputHeight</code>	uint	The height, in pixels, of the input image. Typically 128
<code>inputLayerName</code>	string	The name of the input layer in the TensorFlow Lite file.
<code>inputWidth</code>	uint	The width, in pixels, of the input image. Typically 128
<code>inputScale</code>	float	When the input data type is “float”, the data is first scaled by this number, then the shift value is added: $\text{float_input} = \text{data} / \text{inputScale} + \text{inputShift}$
<code>inputShift</code>	int	
<code>labelsFile</code>	string	The name of the text file (.txt) that gives text strings for the classification output of the model.
<code>memoryMapGraph</code>	uint	?If non-zero, memory-map the TensorFlow Lite file in, rather than loading it with file reads.
<code>minScore</code>	float	If the highest “score” for a label is below this value, none of the items was recognized in the image.
<code>modelType</code>	string	“TFLite” for TensorFlow Lite files.
<code>networkName</code>	string	The name of the vision processing step.
<code>numGridCols</code>	uint	<i>Optional.</i>
<code>numGridRows</code>	uint	<i>Optional.</i>
<code>outputLayerNames</code>	string	The name of the output layer in the TensorFlow Lite file.
<code>outputType</code>	string	“classification” vs “binary_localization”
<code>pollPeriod_ms</code>	uint	
<code>timeoutDuration_sec</code>	float	?The time to allow the model to run in a background thread without any results before it is considered timed out, and must be restarted?
<code>useFloatInput</code>	uint	If non-zero, use float data type within the model
<code>useGrayscale</code>	uint	
<code>verbose</code>	uint	If non-zero, provide verbose output during the interpretation of the model.

Table 401: The TensorFlow Lite model configuration structure

TENSORFLOW MODEL FILES

The TensorFlow Lite models are stored in:

`/anki/data/assets/cozmo_resources/config/engine/`**dnn_models**

The last part of the path is hardcoded into `libcozmo_engine.so`.

82.1.9 OverheadMap

The OverheadMap is a floor map being constructed by Vector. This structure is used to configure the size. “Bigger sizes do not impact computation, only memory.” This structure includes the following fields:

Field	Type	Description
<i>NumCols</i>	uint	The number of columns in the map.
<i>NumRows</i>	uint	The number of rows in the map.

Table 402: The OverheadMap structure

82.1.10 PerformanceLogging

The PerformanceLogging provides the frequency to log stats about the vision processing. The structure includes the following fields:

Field	Type	Description
<i>DropStatsWindowLength_sec</i>	uint	“How long to average dropped image stats”
<i>TimeBetweenProfilerDasLogs_sec</i>	uint	“How often to print Profiler info messages to the logs”
<i>TimeBetweenProfilerInfoPrints_sec</i>	uint	“How often to log Profiler DAS events”

Table 403: The PerformanceLogging structure

82.1.11 PetTracker

The PetTracker structure includes the following fields:

Field	Type	Description
<i>DetectionThreshold</i>	uint	Range is 0 to 1000
<i>InitialSearchCycle</i>	uint	Range is 1 to 45
<i>NewSearchCycle</i>	uint	Range is 5 to 45
<i>MaxFaceSize</i>	uint	
<i>MaxPets</i>	uint	The maximum number animals that are detectable & trackable at the same time.
<i>MinFaceSize</i>	uint	
<i>TrackSteadiness</i>	uint	“10 to 30 in Percentage change required to actually update size/position”

Table 404: The PetTracker structure

Comment: The ability to search for a lost pet would have been *really cool*.

82.2. SCHEDULE MEDIATOR CONFIGURATION FILES

The scheduling for the different vision system modules – how often each processing step is run – is configuration in file located at:

`/anki/data/assets/cozmo_resources/config/engine/visionScheduleMediator_config.json`

This path is also hardcoded into `libcozmo_engine.so`.

This is an array of structures. Each structure gives the frequency to run a given image processing step, for each of the vision processing subsystems modes. 1 means “runs every frame,” 4 every fourth frame, and so on. The structure has the following fields:

Field	Type	Description
<i>high</i>	uint	When in high “mode” run the image processing step every n frames. This value must be a power of two.
<i>low</i>	uint	When in low “mode” run the image processing step every n frames. This value must be a power of two.
<i>med</i>	uint	When in medium “mode” run the image processing step every n frames. This value must be a power of two.
<i>mode</i>	string	The name of the image processing step
<i>relativeCost</i>	uint	A “heuristic weighting to drive separation of heavy-weight tasks between frames where 1 should indicate our lowest cost process e.g. “Markers” is ~16x as resource intensive as “CheckingQuality”
<i>standard</i>	uint	When in medium “mode” run the image processing step every n frames. This value must be a power of two.

Table 405: The vision schedule configuration JSON structure

82.3. PHOTOGRAPHY CONFIGURATION FILES

The photography subsystem configuration in file located at:

`/anki/data/assets/cozmo_resources/config/engine/photography_config.json`

This path is also hardcoded into `libcozmo_engine.so`.

This structure has the following fields:

Field	Type	Description
<i>MaxSlots</i>	uint	
<i>MedianFilterSize</i>	uint	“If > 0, enables a median filter before saving. Must be odd. 3 or 5 are reasonable values.”
<i>SharpeningAmount</i>	float	0.0 disables sharpening
<i>RemoveDistortion</i>	boolean	
<i>SaveQuality</i>	uint	
<i>ThumbnailScale</i>	float	

Table 406: The photography configuration JSON file

83. RESOURCES & RESOURCES

ARM, Neural-network Machine learning software repo
<https://github.com/ARM-software/armnn>

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. Emotional *expressions reconsidered: Challenges to inferring emotion from human facial movements*. Psychological Science in the Public Interest, 20, 1–68. (2019). doi:10.1177/1529100619832930
<https://journals.sagepub.com/stoken/default+domain/10.1177%2F1529100619832930-FREE/pdf>

This paper describes the limitations and high error rate of facial expression software.

FloydHub, *Teaching My Robot With TensorFlow*, 2018 Jan 24,
<https://blog.floydhub.com/teaching-my-robot-with-tensorflow/>

Google, *MobileNets: Open-Source Models for Efficient On-Device Vision*, 2017, Jun 14,
<https://ai.googleblog.com/2017/06/mobilenets-open-source-models-for.html>

Hollemans, Matthijs. *Google's MobileNets on the iPhone*, 2017 Jun 14
<https://machinethink.net/blog/googles-mobile-net-architecture-on-iphone/>

MobileNet version 2, 2018 Apr 22
<https://machinethink.net/blog/mobilenet-v2/>

These two blog posts give an excellent overview of the mechanics of the MobileNet architecture.

Omron, *Human Vision Component (HVC-P2) B5T-007001 Evaluation Software Manual*, 2016
<http://www.farnell.com/datasheets/255338.pdf>

Omron, *OKAO Vision Software Library*
<https://www.components.omron.com/sensors/image-sensing/solution/software-library>

Qualcomm, *How can Snapdragon 845's new AI boost your smartphone's IQ?*, 2018 Feb 1
<https://www.qualcomm.com/news/onq/2018/02/01/how-can-snapdragon-845s-new-ai-boost-your-smartphones-iq>

Qualcomm, *Snapdragon Neural Processing Engine SDK Reference Guide*,
<https://developer.qualcomm.com/docs/snpe/overview.html>

Qualcomm Neural Processing software development kit (SDK) for advanced on-device AI, the Qualcomm Computer Vision Suite

Situnayake, Daniel; Pete Warden, *TinyML*, O'Reilly Media, Inc. 2019 Dec,
<https://www.oreilly.com/library/view/tinyml/9781492052036/>

Stein, Andrew; *Decoding Machine-Readable Optical codes with Aesthetic Component*, Anki, Patent US 9,607,199 B2, 2017 Mar. 28

TensorFlow, Mobile Net v1
https://github.com/tensorflow/models/blob/master/research/slim/nets/mobilenet_v1.md

“small, low-latency, low-power models” that can recognize a variety of objects (including animals) in images, while running on a microcontroller

TensorFlow, *TensorFlow Lite GPU delegate*
<https://www.tensorflow.org/lite/performance/gpu>

TensorFlow, *TensorFlow Lite inference*
<https://www.tensorflow.org/lite/guide/inference>

This Week in Machine Learning (TWIMLAI), episode 102, *Computer Vision for Cozmo, the Cutest Toy Robot Everrrrr!* with Andrew Stein
<https://twimlai.com/twiml-talk-102-computer-vision-cozmo-cutest-toy-robot-everrrrr-andrew-stein/>

Viola, Paul; Michael Jones, *Rapid Object Detection using a Boosted Cascade of Simple Features*, Accepted Conference on Computer Vision and Pattern Recognition, 2001
http://wearables.cc.gatech.edu/paper_of_week/viola01rapid.pdf

Wikipedia, *Adaptive histogram equalization*
https://en.wikipedia.org/wiki/Adaptive_histogram_equalization

Wikipedia, *Viola-Jones object detection framework*
https://en.wikipedia.org/wiki/Viola%20%26%20Jones_object_detection_framework

Viola, Paul; Michael Jones, *Robust Real-time Object Detection*, International Journal of Computer Vision (2001)
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.110.4868>

Example code for running a TensorFlow Lite model on a PC
<https://github.com/ctuning/ck-tensorflow/blob/master/program/object-detection-tflite/detect.cpp>

CHAPTER 18

Mapping & Navigation

Vector builds an internal map to track where he can drive, where objects and faces are.

- Mapping Overview
- Navigation and Path Planning

84. MAPPING OVERVIEW

Vector tracks objects in two domains:

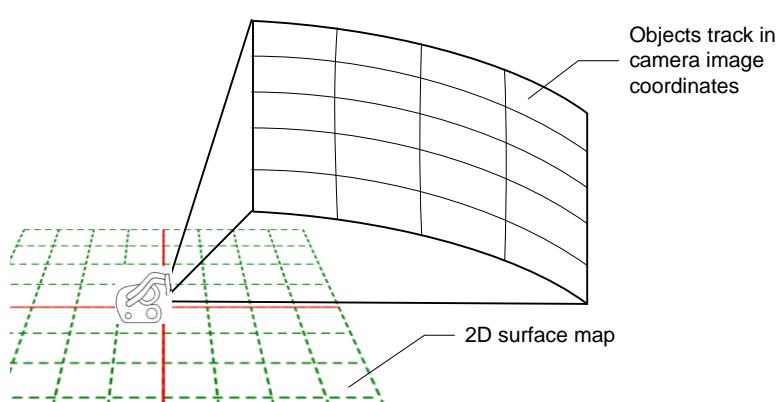


Figure 72: Mapping contexts

- A 2D map that is used to track where objects (especially objects whose marker symbols he recognizes), cliffs, and other things are on the surfaces that he can drive on. Vector uses this map to navigate. This map has an arbitrary origin and orientation.
- Vector also tracks where faces, pets and some kinds of recognized objects are in his camera image area; these objects are tracked in the image pixels. (Never mind that the camera pose can change!)

Vector's 2-D surface map system works with the localization and navigation subsystem. It uses several sensors to know

- Cliff sensors to detect edge, and lines
- Time of flight sensor to measure distances
- Vision to detect the edges, and the location of a hand
- Vision to identify accessories by recognizing markers

85. MAP REPRESENTATION

Vector keeps track of the surface that he can drive on with a navigable 2D map. The map's orientation and position of its origin are arbitrary – Vector just picks a spot and goes with it. The surface map is represented in a compressed format called a quad-tree.

Vector tracks accessory objects, immovable obstacles, and cliffs in terms of this map. The map's units are in mm.

85.1. QUAD-TREE MAP REPRESENTATION BASICS

A *quad-tree* is a structured way of “compressing” the information in the map, to reduce the amount of memory required. A quad-tree that recursively breaks down a grid into areas that are interesting, and those that are not.

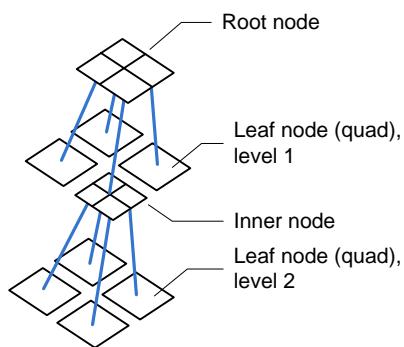


Figure 73: Structure of a quad-tree

The tree has two kinds of nodes: inner nodes and leaf nodes:

- The inner nodes do not hold any information about the region (except its size). Instead they point to 4 child nodes at the next lower layer. The top most node is called the *root node*.
- The *leaf nodes* of the tree are square cells (called *quads*) that hold information about what is there (or that the area is unexplored).

Each node represents a square area. The size of the square depends on how many levels it is from the root node. The root node covers the whole map. The nodes in the next layer down are half the width and height of the root node. (In general, a node is half the width and height of a node the next layer up.) *Nodes (including quads) at the same level – the same distance from the root node – are the same size.* Each node’s coordinates can be figured in a similar way by knowing the coordinates of the root node.

When Vector reaches the edge of his map area and needs to expand it, he has to add a new node at the top (this becomes the new root node) and adds nodes down until it can contain the info at the edge of the map.

85.2. THE MAP'S STARTING POINT

Vector doesn't have a “north pole” or other global reference point to center his map on. When he powers on, or “whenever Vector is delocalized (i.e. whenever Vector no longer knows where he is – e.g. when he's picked up), Vector creates a new pose starting at (0,0,0) with no rotation. As Vector drives around, his pose (and the pose of other objects he observes – e.g. faces, his [cube], charger, etc.) is relative to this initial position and orientation.”

Client applications (the ones that talk via the HTTPS API) may also wish to know that the map was thrown out and a new one created – and thus know they should toss out their map and location of objects. Vector associates a unique identifier with each generation of the map called *origin_id*. Whenever a new map is created the “*origin_id* [is] incremented to show that [the] poses [of the

Anki SDK

map, Vector, and objects in the world] cannot be compared with earlier ones.” “Only poses of the same `origin_id` can safely be compared or operated on.”

85.3. HOW THE MAP IS SENT FROM VECTOR TO SDK APPLICATIONS

The full tree is not sent from Vector to an SDK, only the leaf nodes (quads). Quads are the only part of the tree that hold information about what is in an area. They also have sufficient information to reconstruct a quad-tree, which is a useful access structure.

86. MEASURING THE DISTANCE TO OBJECTS

Vector has a time of flight sensor, pointing straight ahead. He can use the sensor to measure the distance to the objects, barriers, open spots on the map, and to estimate his position. The sensor can be blocked by the arms, if they are in just the right lowered position – such as approaching an object and docking with it.

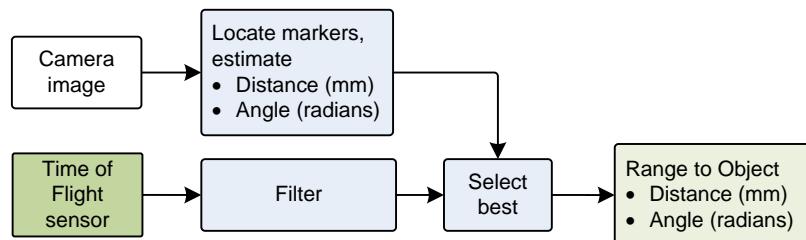


Figure 74: A typical localization and mapping functional block diagram

If the sensor is not blocked:

- The samples of distances reported by the sensor are gathered,
- A filter is applied to them (probably a median filter), throwing out values that are too near or too far.
- Combining this with Vectors current position and orientation, and the distance to the object, he can estimate the objects position; and
- Vector can infer that the space between him and the object are free of other objects and obstacles. (This means splitting up the map quads into a fine-grained resolution along the narrow beam path.)

In addition to this, if the object has a known marker, the vision system estimates the angle of the object, and a distance to it. This is based on the known visual size of the marker, and the observed size. If the time of flight sensor is not blocked, only the angle need be used. If the sensor is blocked, the visually estimated distance to the object can be used instead.

86.1. FILTERING

The time of flight sensor emits a stream of pulses that are detected by a grid of *single photon avalanche diode* (SPAD) detectors. The detectors measure two things:

1. The duration from the time that the pulse was emitted; this is a direct measure of the distance to the object.
2. A count of the number of photons received back from the object. This is a measure of how reflective the object is. This can potentially be used to distinguish between two different objects.

The sensor gathers the distances into a histogram counting the number of times each distance was measured. At regular intervals the body-board firmware gets the data from the sensor, resetting the counts and the histogram. The data is then sent to the head-board.

The software has to clean up the histogram since it is very noisy, with lots of spikes:

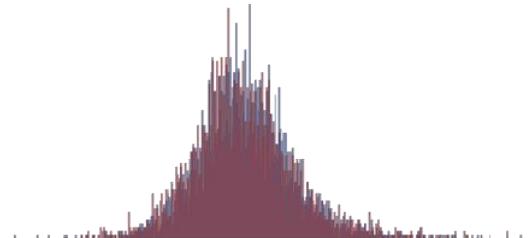


Figure 75: A typical histogram from a time of flight sensor.
ST Microelectronics

Comment: The histogram is actually part of how the sensor cleverly rejects noise. The detectors will pick up light from other sources, such as bright sunlight. By using pulses and controlling when they are sent, the sensor can measure the background (or ambient) light level, and better discriminate its own light pulses from the rest. The noise can come from the light be reflected back by dirt on the sensor lens, dust in the atmosphere, light bouncing around and coming back a little later than the directly reflected light. Gathering the measurements into a histogram spreads the noise out, mostly randomly, making it easier to pick out the useful measurement.

The easiest way to eliminate the histogram spikes is to do a pass over, setting each points value to be the weighted average of the values to the left and right. Values below a noise floor can be tossed out.

Then a good distance measurement can be found by looking for the peak or by finding the median.

86.1.1

VL53L1X next generation sensor

Vector is built with a VL53L0X sensor. But Vector's software is structured to support processing data from its much more powerful sibling VL53L1X sensor. Anki was investigating this sensor for use in future Cozmo generations, and performing engineering evaluations using a modified Vector.

The VL53L1X's detector is a 16x16 grid of SPAD detectors. The sensor can be configured to use rectangular areas of the detector grid, called the *region of interest* (ROI), instead of the whole grid:

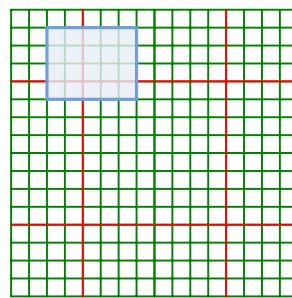


Figure 76: Sensing regions of interest

With the sensors field of view, different regions look in different directions. By creatively choosing regions to get measurements from – and using the reflectivity measurement to distinguish between objects – the software could look around, track multiple objects and scan the driving surface. In other words, it could work like a low-resolution depth sensing camera, with a very good measurement of depth and surface reflectivity. It can even detect swiping motions.

Since Anki had placed the time of flight sensor in the robot's *head*, near the camera, there was more potential for smarter interaction. Obviously, the head could scan up and down, giving a

much wider range of looking for cliffs, distance to objects higher than a few centimeters and such. The recognition would have had another signal to help it improve hand, face, and pet detection. The face recognition software would have better ability to tell if it was looking at an image or a real face – by knowing the distance, it could tell if the face was a plausible size.

86.2. INTERNAL DATA STRUCTURES

The ranging modules produce several JSON structures for internal use. The first main output structure, `DistanceSensorData`, has the following fields:

Field	Type	Units	Description
<code>proxDistanceToTarget_mm</code>	float	mm	The distance to object, as measured by the time of flight sensor.
<code>visualAngleAwayFromTarget_rad</code>	float	radians	The targets relative orientation angle, as estimated by the vision system.
<code>visualDistanceToTarget_mm</code>	float	mm	The distance to object, as estimated by the vision system.

Table 407:
`DistanceSensorData`
parameters

86.2.1 Raw Range Data

The second main output data structure, `RangeSensorData`, is very similar, but links to source data. It has the following fields:

Field	Type	Units	Description
<code>headAngle_rad</code>		radians	The angle (tilt) of the robots head.
<code>rangeData</code>		RangeDataRaw	The data from the time of flight sensor.
<code>visualAngleAwayFromTarget_rad</code>	float	radians	The targets relative orientation angle, as estimated by the vision system.
<code>visualDistanceToTarget_mm</code>	float	mm	The distance to object, as estimated by the vision system.

Table 408:
`RangeSensorData`
parameters

The sensor-related data structures involve a complex nesting of structures. To help clarify:

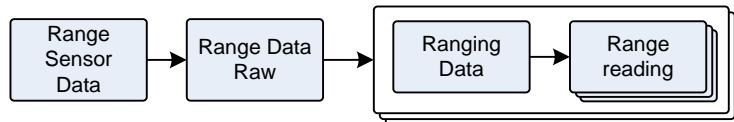


Figure 77: The time of flight data structures.

The `RangeDataRaw` structure is just a link to an array of arrays of measurements. It has the following fields:

Field	Type	Units	Description
<code>data</code>	<code>RangingData[]</code>		An array of the sensor data to process, and the results.

Table 409:
`RangeDataRaw`
parameters

The RangingData structure has the following fields:

Field	Type	Units	Description	Table 410: RangingData parameters
<i>numObjects</i>	uint	<i>count</i>	A count of the number of objects seen in the regions.	
<i>processedRange_mm</i>	int	<i>mm</i>	The range to the object after processing & filtering of the data	
<i>readings</i>	RangeReading[]		The range readings reported from the time of flight sensor	
<i>roi</i>	uint		The region of interest that was measured.	
<i>roiStatus</i>	uint		A code indicating whether there is a valid measurement for this region.	
<i>spadCount</i>	float		“the time difference (shift) between the reference and return [detector] arrays.” This translates to distance to the target.	

The RangeReading structure is basically identical to the structure in ST’s software to interface with the time of flight sensor. It has the following fields:

Field	Type	Units	Description	Table 411: RangeReading parameters
<i>ambientRate_mcps</i>	float	<i>mcps</i>	The ambient number of counts; this is the noise floor	
<i>rawRange_mm</i>	int	<i>mm</i>		
<i>sigma_mm</i>	float	<i>mm</i>	The distance to the target	
<i>signalRate_mcps</i>	float	<i>mcps</i>	The “return signal rate measurement... represents the amplitude of the signal reflected from the target and detected by the device.”	
<i>status</i>	uint		A code with 0 indicating a valid measurement, otherwise indicating an error during measurement or processing.	

Note: *mcps* is mega counts per second.

86.2.2 Display-Ready Range Data

The RangeDataDisplay structure is just a link to an array of arrays of measurements. It has the following fields:

Field	Type	Units	Description	Table 412: RangeDataDisplay parameters
<i>data</i>	RangingDataDisplay[]		The ranging data, for potential display.	

The RangingDataDisplay structure is basically identical to the structure in ST's software to interface with the time of flight sensor. It has the following fields:

Field	Type	Units	Description	Table 413: RangingDataDisplay parameters
<i>padding</i>	uint		Likely a CLAD structure field that is reserved for future use that was automatically converted to JSON.	
<i>processedRange_mm</i>	float	mm	The range to the object after processing & filtering of the data	
<i>roi</i>	uint		The region of interest that was measured.	
<i>roiStatus</i>	uint		A code indicating whether there is a valid measurement for this region.	
<i>spadCount</i>	float	count	“the time difference (shift) between the reference and return [detector] arrays.” This translates to distance to the target.	
<i>signalRate_mcps</i>	float	mcps	The “return signal rate measurement... represents the amplitude of the signal reflected from the target and detected by the device.”	
<i>status</i>	uint		A code with 0 indicating a valid measurement, otherwise indicating an error during measurement or processing.	

87. BUILDING THE MAP

The map is made as Vector drives around – when he is on a mission, or just exploring. Each of the leaf quads (in the map) is associated with information about that space and what is contained there:

- What Vector knows is in the quad –a cliff, the edge of a line, an object with a marker symbol on it, or an object without a symbol (aka an obstacle),
- A list of what Vector *doesn't* know about quad – i.e. that he doesn't know whether or not there is a cliff or interesting line edge there,
- Whether Vector has visited the quad or not.

Vector subdivides quads to better represent the space. The quad probably is only slight bigger than the object in it. But the quad (probably) can be smaller than the object, to accommodate the object not oriented and aligned to fit quite perfectly in the quad. More than quad can refer to a contained object.

87.1. MAPPING CLIFFS AND EDGES

If a cliff (surface proximity) sensor has a large, significant change in value, Vector will make a note that there is a cliff sensor there. If the value has a smaller, but still noticeable change, he might make a note that there is a line edge there – an edge between a dark area and a light area.

87.2. TRACKING OBJECTS

Vector tracks objects (especially objects with markers) using the map, and other cross-referencing structures. Vector associates the following information with each object it tracks:

- The object's kind (dock, cube, etc)
- A pose. The image skew of the marker symbol gives some partial attitude (relative orientation) information about the object and Vector can compute an estimated orientation (relative to the coordinate system) of the object from this and Vector's own pose. Vector can estimate the objects position from his own position, orientation, and the distance measured by the time of flight sensor.
- A size of the object. Vector is told the size of objects with the given symbol.
- A link to a control structure for the kind of object. For instance, accessory cubes can be blinked and sensed.

If he sees a symbol, he uses the objects known size, the image scale, its pose (if known) and any time-of-flight information to (a) refine his estimated location on the map, (b) update the location and orientation of that object.

87.3. BUILDING A MAP WITH SLAM

Vector employs a mapping technique known as *simultaneous location and mapping* (SLAM) to integrate these (and other) steps. SLAM is a method to identify Vector's current position and orientation (relative to the map), and to construct that map.

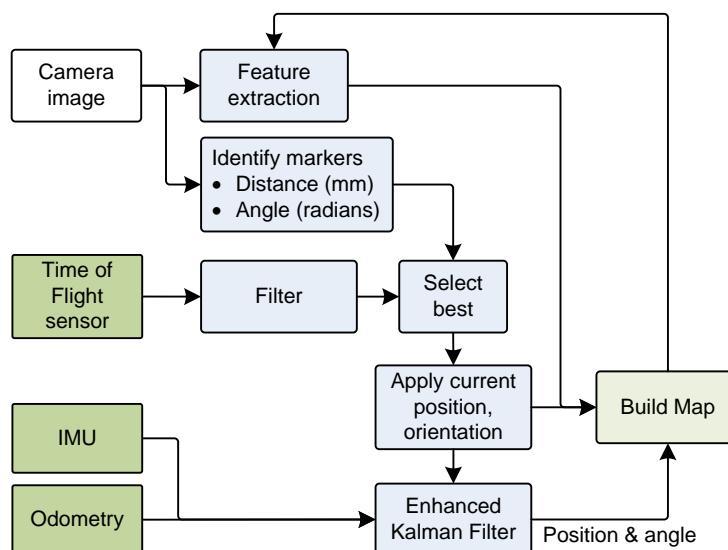


Figure 78: A typical localization and mapping functional block diagram

SLAM consists of multiple parts. It integrates the sensor for distance and movement. It also uses image processing to figure it out where it is at. It identifies landmarks, and information about them. In a sophisticated integration process, it can estimate Vector's orientation *and* if an object has moved. The estimate of Vectors orientation is based on turn information from the IMU, and refined by what it can see.

88. NAVIGATION AND PLANNING

Path planning is devising a path around obstacles without collision, to accomplish some goal, such as docking with the “home” (charger) or accessory cube. Intuitively, all you need to do with a rectilinear grid is to figure out the x-y points to go from point A to B. Vector (and Cozmo) is longer than they are wide – especially when carrying a cube. If this isn’t taken into account by the planner, Vector could get stuck going down some path he can’t fit in or turn around in. Cozmo had an XY-theta planner to construct paths that he could traverse.

Vector’s path planning approach is unknown.

89. RESOURCES & RESOURCES

Riisgaard, Søren; Morten Rufus Blas; *SLAM for Dummies: A Tutorial Approach to Simultaneous Localization and Mapping*
<http://www-inst.eecs.berkeley.edu/~ee290t/fa18/readings/Slam-for-dummies-mit-tutorial.pdf>

ST Microelectronics, *UM2039 World smallest Time-of-Flight ranging and gesture detection sensor Application Programming Interface*, Rev, 2016 Jun
https://www.st.com/resource/en/user_manual/dm00279088-world-smallest-timeofflight-ranging-and-gesture-detection-sensor-application-programming-interface-stmicroelectronics.pdf

ST Microelectronics, *UM2600 Counting people with the VL53L1X long-distance ranging Time-of-Flight sensor*, Rev 1, 2019 Jun
https://www.st.com/resource/en/user_manual/dm00626942-counting-people-with-the-vl53l1x-longdistance-ranging-timeofflight-sensor-stmicroelectronics.pdf

Wikipedia, *Occupancy grid mapping*,
https://en.wikipedia.org/wiki/Occupancy_grid_mapping

Vector’s map is based on occupancy grids, except it does not use probabilities.

CHAPTER 19

Accessories

Vector's accessories include his charging station, companion cube, and custom items that can be defined thru the SDK.

- Accessories in general: symbols, docking
- Home & Charging Station
- Companion cube
- Custom items

90. ACCESSORIES IN GENERAL

Accessories have at least one maker symbol that Vector can recognize. Vector tracks the location and orientation based on this

90.1. DOCKING

Docking is a behaviour/action that is used for both approaching the cube, charging station (home), and other marked items.

It has specialized steps depending on whether it is a cube, the home, etc.

91. HOME & CHARGING STATION

Vector has a rich set of behaviours associated with its Home / Charger. In retrospect, this makes sense, as it is Vectors home, his nest, his comfy chair.

91.1. DOCKING

Vector's step in docking with the charging station are:

1. Approach and line up with the charger
2. Turn around (rotate 180°)
3. Reverse and back up the ramp. Vector uses a line follower, with his cliff sensors, to drive straight backwards. (Since he is going backwards, he can't use vision.) He uses the tilt of the ramp to confirm that he is on the charger
4. He also checks that he is in the right spot by looking for power to his charging pads, as reported by body-board charging circuit. If he is unable to find the spot, he grumbles about it, drives off and retries.

Vector has a cute low light mode that turns on most of the pixels on his display to see a bit more, and locate his home.

92. COMPANION CUBE

Vector has a companion cube that he can pickup, illuminate the lights on, and detect taps. The cubes design is described in chapter 5.

Vector can roll his cube, shove it around, use it to “pop a wheelie,” and pick it up. To do these, he must line up squarely with cube. Vision was found to be needed in the Cozmo to align precisely enough to get the lift hooks into the cube.

92.1. COMMUNICATION

Vector connects with the cube via Bluetooth LE. This communication link provides the ability for Vector to:

- Discover cubes
- Pair with a cube (note that Vector can pair with only one cube, and if he is not already paired, he will automatically pair with the first cube he receives Bluetooth LE advertising for.)
- Check the firmware version
- Update the cube firmware
- Check the cube’s battery level
- Detect the cube orientation
- Detect taps on the cube
- Turn the cubes lights on and off.

The HTTPS API provides the following Cube-related commands:

- List the available cubes, see Chapter 13, section *49.4 Cubes Available*
- Forget (or unpair) from his preferred cube, see Chapter 13, section *49.8 Forget Preferred Cube*
- Pair to the first cube detected, see Chapter 13, section *49.15 Set Preferred Cube*
- Connect to his cube , see Chapter 13, section *49.3 Connect Cube*
- Disconnect from the cube, see Chapter 13, section *49.5 Disconnect Cube*
- Dock with his cube, see Chapter 13, section *49.6 Dock With Cube*
- Flash cube lights, see Chapter 13, section *49.7 Flash Cube Lights* and *49.14 Set Cube Lights*. The later allows using a complex pattern
- Pick up an object (his cube), see Chapter 13, section *49.9 Pickup Object*
- Place his object (his cube) on the ground, see Chapter 13, section *49.10 Place Object on Ground Here*
- Pop a wheelie, see Chapter 13, section *49.11 Pop A Wheelie*
- Roll his cube, see Chapter 13, section *49.12 Roll Block* and *49.13 Roll Object*

The state of the cube is reported to the HTTPS API.

As the state of the cube changes, the following events are posted to the API:

- The cubes battery level, see Chapter 13, section 49.2.1 *CubeBattery*
- A loss of the connection with the cube, see Chapter 13, section 49.2.2 *CubeConnectionLost*
- The robot state event (see Chapter 13, section 58.2.1 *RobotState*) provides other info about Vector's attempt to interact with the cube. This includes what object he is carrying. There are bits to indicate when
 - Vector is carrying his cube
 - His picking up or moving to dock with his cube
- The object event (see Chapter 13, section 41.2.1 *ObjectEvent*) provides other info about the state of the cube as it happens: taps, loss of connection, state of connection, being moved, etc.

92.2. TAP DETECTION

The software detects taps on the cube by streaming the accelerometer data, filtering and looking for a tap pattern:

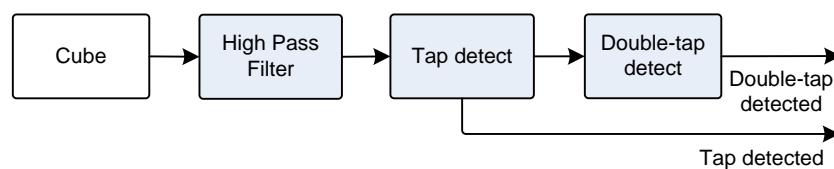


Figure 79: The tap detector

93. CUSTOM ITEMS

Vector can be told about custom objects. Once Vector knows about these, he can identify the object and track it on his map, or navigate around them. Objects with markers are reported to SDK based applications for custom processing.

Defining a custom object takes three kinds of information. First, the shape of the item – whether it is a “wall,” box or cube. Second, assign some of the handful of predefined symbols to the item; this is optional. And third, measure the size of the marker symbols and object.

There are four kinds of custom objects that can be defined:

- A fixed, unmarked cube-shaped object.
- A flat wall with only a front side,
- A cube, with the same marker on each side.
- A box with different markers on each side.

93.1. A FIXED, UNMARKED OBJECT (CUBE-SHAPED)

The object is in a fixed position and orientation. This cube can't be observed since it is unmarked. So there won't be any events related to this object. “This could be used to make Vector aware of objects and know to plot a path around them.”

93.2. CUSTOM WALL DEFINITION

The second type of custom object is a *wall*. It has a single marker on the front face.

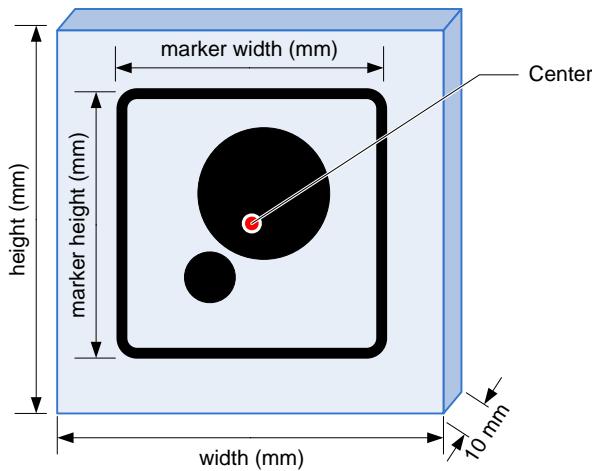


Figure 80: The custom wall parameters

The marker must be horizontally and vertically centered. The width of the marker doesn't have to be the same as the height... but probably should be.

The body origin is the 5mm behind the center of the face. When Vector is tracking the position and orientation of this object, the position it gives for the point in the wall 5mm behind the face, at half the height and width – the center of the wall.

93.3. CUSTOM CUBE DEFINITION

The third type of custom object is a *cube*. A cube's width, height and depth are all the same size. A cube has the same marker on all 6 faces (not shown below):

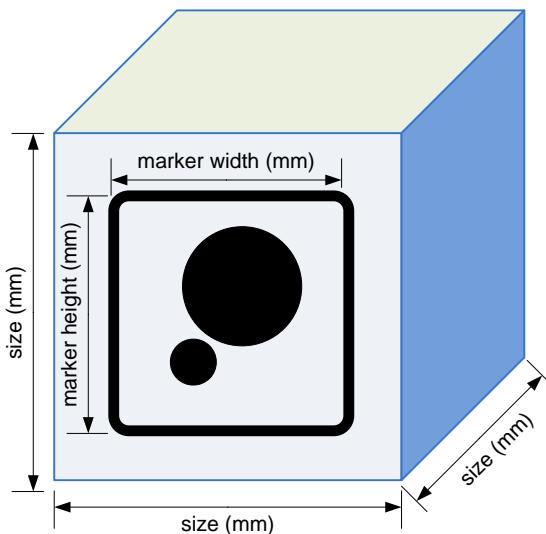


Figure 81: The custom cube parameters

The marker must be horizontally and vertically centered on each face. The width of the marker doesn't have to be the same as the height... but probably should be.

The body origin is the very center of the cube. When Vector is tracking the position and orientation of this object, the position it gives is for the very center of the cube, not for a visible face.

93.4. CUSTOM BOX DEFINITION

The fourth (and final) type of custom object is a *box*. Although they have similar names, a custom box differs from a custom cube in two ways. With a box, the height, width and depth can all be different sizes. Second, each face has a different marker symbol associated with it, so that Vector can match it up with the size of that side.

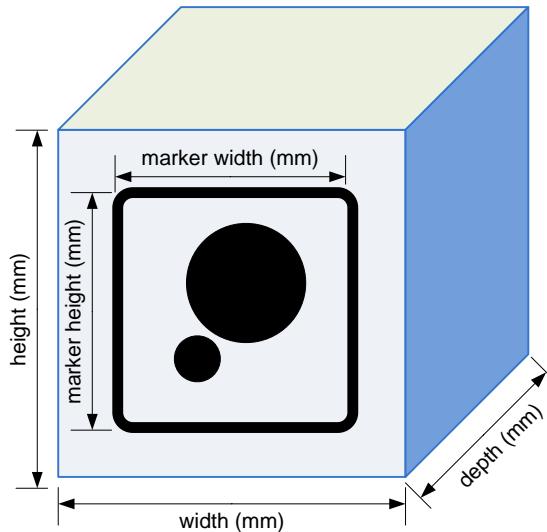


Figure 82: The custom box parameters

The marker must be horizontally and vertically centered on each face. The width of the marker doesn't have to be the same as the height... but probably should be.

The body origin is the very center of the box. When Vector is tracking the position and orientation of this object, the position it gives is for the very center of the box, not for a face.

93.5. COMMUNICATION

The Chapter 13 HTTPS API provides the following custom-object related commands:

- Create a custom unmarked object (see Chapter 13 section *41.3 Create Fixed Custom Object*) or one with markers that can be tracked (see Chapter 13 section *41.4 Define Custom Object*)
- Drive to the object, see Chapter 13 section *55.4 Go To Object*. Note Vector thinks in terms of the center of the object, not the face; for larger objects add the distance from the center to the face for Vector's position.

As the state of the cube changes, the following events are posted to the API:

- The object event (see Chapter 13, section *41.2.1 ObjectEvent*) provides other info about the state of the object as it happens: that is observed or lost, being moved, that it's orientation has changed etc.

[This page is intentionally left blank for purposes of double-sided printing]

PART V

Animation

Vector uses animations – “sequence[s] of highly coordinated movements, faces, lights, and sounds” – “to demonstrate an emotion or reaction.” This part describes how the animation system works.

- ANIMATION. An overview how Vector’s scripted animations represents the “movements, faces, lights and sounds,” and how they are coordinated
- LIGHT ANIMATION. An overview of the backpack and cube light animation.
- DISPLAY & PROCEDURAL FACE. Vector displays a face to convey his mood and helps form an emotional connection with his human.
- AUDIO PRODUCTION. A look at Vector’s sound effects and how he speaks
- MOTOR CONTROL. A look at how Vector’s moves.
- ANIMATION FILE FORMAT. The format of Vector’s binary animation file.



drawing by Steph Dere

[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 20

Animation

This chapter describes Vector's animation engine:

- Animation Engine, animation groups, triggers, and events
- Animation file formats

94. ANIMATION TRIGGERS AND ANIMATION GROUPS

An *animation* is a scripted “sequence of highly coordinated movements, faces, lights, and sounds.” Vector uses animations “to demonstrate an emotion or reaction” as well as many other physical moments. Vector’s small, light frame allows him to make quick physical motions to represent his little tantrums and other emotions.

animation

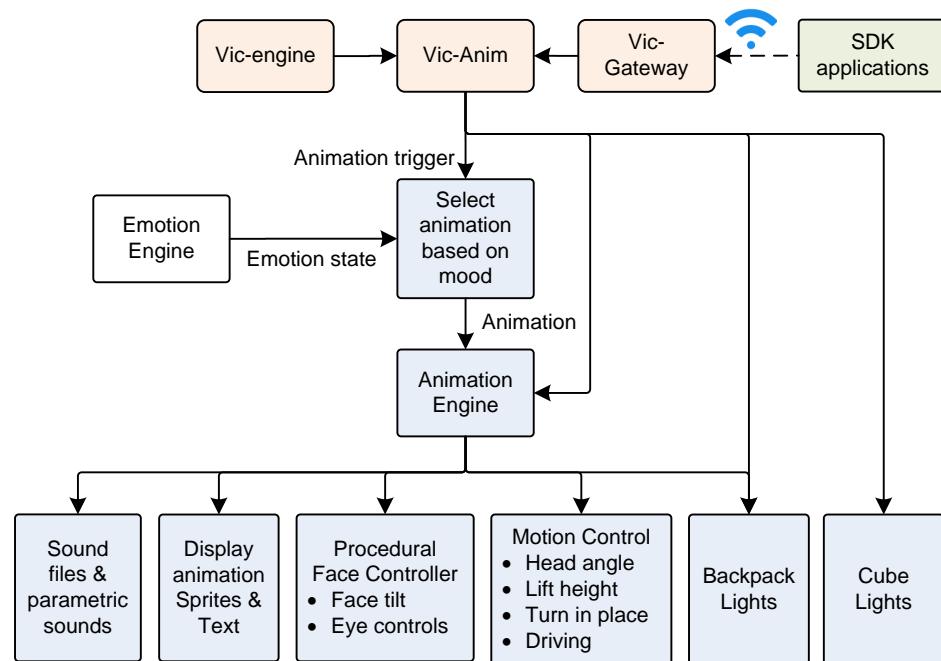


Figure 83: The behaviour-animation flow

Not surprising, much of the animation is carried out by vic-anim. The motor controls, including driving along a path, are performed in vic-robot.

Vector employs two levels of referring to an animation. Individual animations have an *animation name*. Animations are also grouped together by type, with an identified for the group called an *animation trigger name*. Vector “pick[s] one of a number of actual animations to play based on Vector’s mood or emotion, or with random weighting. Thus playing the same trigger twice may not result in the exact same underlying animation playing twice.”

animation name
animation trigger name

94.1. FILES

The animation system employs many files, working in concert to effect the animation. The top level files map trigger names to the next level, which may be groups of animations, display compositing tables, or (in the case of the lights) the patterns to illuminate the lights with. In the case of animation groups, these further map to sprite sequences to display (which then maps to image files), and sounds play. The compositing image maps also map to these sprite sequences.

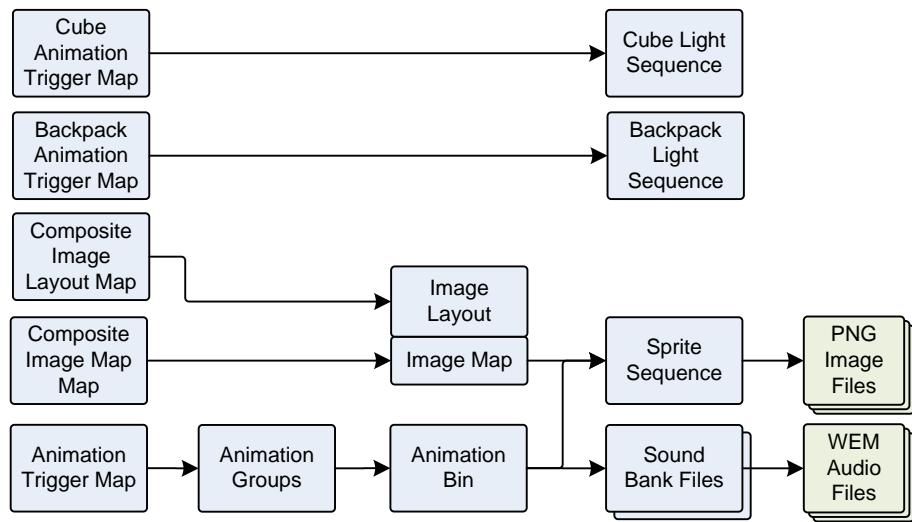


Figure 84: The behaviour-animation flow

There are seven types of animation files and other animation sources:

- JSON files that describe how the backpack lights should behave (see Chapter 21)
- JSON files that describe how the cube lights should behave (see Chapter 21)
- Binary animation files holding one or more of related animations that coordinate sophisticated sounds, eye animations, linking together sprite sequences, and coordinate head & lift movements with driving (see Chapter 24 for details of this file)
- Sprite sequences (see Chapter 22), which are folders of PNG image files to display in sequence
- Composed screens (see Chapter 22) showing icons and text information driven by the behaviors and cloud server intents.
- Sound files (see Chapter 23) holding pre-recorded sound effects
- A procedural animation that is generated with vic-anim. These perform text to speech, driving around obstacles, animating Vector's eyes, and other tasks that are not practical to script in a file.

And there are four kinds of files gluing these together:

- JSON files that map the trigger names to the *animation groups*, and to the backpack and cube light animations. (These will be described below.)
- The animation group files with rules on how to select one or more individual animations, including with the weighted randomization and emotion filters. (These will be described below.)
- The animation binary file may direct a sprite sequence and/or audio file to play. These will be described in Chapter 25

- JSON files to layout the display; these may call out animation sequences and places to composite icons and text. These will be described in Chapter 22

94.2. TRIGGER MAP CONFIGURATION FILES

The list of animation triggers provided to the SDK is built into `libcozmo_engine.so`. The internal configuration files support a much wider range of animation triggers; it is not known if passing one these will work, or will be filtered out.

mapping trigger names to animation groups

The animation trigger name is mapped to an animation file (and group of animations). The table that defines this mapping is found in the following file:

`/anki/data/assets/cozmo_resources/assets/cladToFileMaps/AnimationTriggerMap.json`

(This path is hardcoded into `libcozmo_engine.so`.)

The format of the files is the same. The file is an array of structures. Each structure has the following fields:

Field	Type	Description
<code>AnimName</code>	string	The name of the JSON file (without the “.json” suffix) for the animation.
<code>CladEvent</code>	string	This is the animation trigger name to match when looking up the animation.

The cube’s and backpack light animation file name is usually the same as the trigger, except that the first letter is lower case.

Table 414: JSON structure mapping trigger name to animation group

94.3. ANIMATION GROUP FILES

In the “`AnimationTriggerMap.json`” file (describe above), the “`AnimName`” field value maps (when the suffix “.json” is appended) to animation groups are located in the following folder tree:

`/anki/data/assets/cozmo_resources/assets/animationGroups`

This path is hardcoded into `libcozmo_engine`. Inside are folders (grouping the animation groups), each of which holds the JSON files. By convention, the animation group file names are all lower case, starts with “`ag_`” and often look similar to the trigger name (but not always).

Each animation group JSON file is a structure with the following fields:

Field	Type	Description
<code>animations</code>	<code>AnimationGroupItem[]</code>	An array of animations to select from to play.

Table 415: Animation group file JSON structure

The AnimationGroupItem structure describes the specific animation clip to use. It may also specify some head movement, with some variability; this is optional. The structure has the following fields:

Field	Type	Units	Description
<i>CooldownTime_Sec</i>	float	<i>seconds</i>	The minimum duration, after this animation has completed, before it can be used again. Typically 0.0
<i>HeadAngleMin_Deg</i>	float	<i>degree</i>	The head is to move to random angle greater (or equal) to this. Only used if <i>UseHeadAngle</i> is true.
<i>HeadAngleMax_Deg</i>	float	<i>degree</i>	The head is to move to random angle less than (or equal) to this. Only used if <i>UseHeadAngle</i> is true.
<i>Mood</i>	string	<i>emotion name</i>	The emotion name or “Default”
<i>Name</i>	string		The name of the animation clip within one of the animation binary files.
<i>UseHeadAngle</i>	bool		If true, this enables the head to be moved to some random within the specified range. <i>Optional, default is false</i>
<i>Weight</i>	float		How much “weight” to give to this entry. Typically 1.0

Table 416:
AnimationGroupItem
structure

The possible animations are screened for being applicable to the current emotional state (with *Mood*). The result set is randomly selected from: The weights (of the select items) are summed up and normalized, giving the probability that that entry would be selected.

95. ANIMATIONS

95.1. ANIMATION TRACKS

Each animation may use one or more tracks. The tracks can control:

animation tracks

- The display graphics – compositing a movie , text, and eyes layers
- Animating the backpack and cube lights
- Audio effects
- Moving the head
- Moving the lift
- Moving the treads, or navigating

When an animation is played, it locks the tracks that it is using, for the duration of the animation. If one of the tracks that it needs to use is already locked, the animation can't be played (and generates an internal error).

When an animation is submitted to be played, a several of tracks (the lift, head and body) can be flagged to be ignored if they already used elsewhere by animation system.

95.2. ANIMATION FILES

The animation binary files are held in the following folder:

/anki/data/assets/cozmo_resources/assets/animations

This path is hardcoded into vic-anim. Each of these files may contain several animations (called clips). By convention, the name of the animation starts with the name of the file. See Chapter 25 for a detailed description of these files.

95.3. ANIMATION NAMES MANIFEST

A list of animation names and their duration is located in manifest file at:

*animation duration
manifest*

/anki/data/assets/cozmo_resources/**assets/sprites/anim_manifest.json**

This path is hardcoded into libcozmo_engine. The file is an array of structures. Each structure has the following fields:

Field	Type	Units	Description
<i>length_ms</i>	int	<i>ms</i>	The duration of the animation (when played)
<i>name</i>	string		The name of the animation clip within one of the animation binary files.

Table 417: JSON structure mapping animation name to duration

96. SDK COMMANDS TO PLAY ANIMATIONS

The HTTPS SDK includes commands to list and play animations.

- A list of animations triggers can be retrieved with the List Animation Triggers command (see Chapter 13 section *44.3 List Animation Triggers*).
- A list of animations can be retrieved with the List Animation command (see Chapter 13 section *44.2 List Animations*).
- An animation can be play by selecting the animation trigger (see Chapter 13 section *44.5 Play Animation Trigger* command). Vector will select the specific animation from the group. Or,
- An animation can be play by the giving the specific animation name (see Chapter 13 section *44.4 Play Animation*).

As the individual animations are low-level they are the most likely to change, be renamed or removed altogether in software updates. Anki strongly recommends using the trigger names instead. “Specific animations may be renamed or removed in future updates of the app.”

CHAPTER 21

Lights Animation

This chapter describes the light animations:

- The Cube Spinner game, which is one user of this type of lights animation
- Backpack Lights animation
- Cube Lights Animation

97. LIGHTS ANIMATION OVERVIEW

The backpack lights are used to show the state of the microphone, charging, WiFi and some other behaviours. The companion cube lights are show setting up the cube, entertainment while Vector is interacting with it, and for games. There are three interrelated sources of light animation.

- Animation binary file to animate the backpack lights. This drives most of the light animation.
- JSON files for the Cube and backpack light animation. There are four kinds of JSON files: files for the Cube's light sequence, files for the backpack light sequence, two files to map animation trigger names to each of those light sequences.
- The Cube Spinner game, which is a notable client of the JSON-driven light animations..

The light animations may be triggered by the cube spinner game configuration, or by behaviors (within `libcozmo_engine`), such as those related to exploring, interaction, pouncing etc.

The companion cube and backpack light animations are very similar, so they have been grouped here for discussion.

98. CUBE SPINNER GAME

The cube spinner game was “like a little roulette wheel on the cube. The lights would spin around and you and Vector competed to make them stop at the right combination.” Although developed early, it was not enabled in any of the Anki software releases. It is thought that version 1.7 would have enabled it.

cube spinner game

The Cube Spinner game’s configuration file is located with the behavior folders:

/anki/data/assets/cozmo_resources/**config/engine/behaviorComponent/cubeSpinnerLight**
Maps.json

This is path hardcoded into `libcozmo_engine`.

This configuration file is unlike the behavior files in that it doesn't have a behavior class or ID. It is used to map game events to animation triggers for backpack and cube lights. The events names are defined in `libcozmo_engine.so`. The configuration file structure has the following fields:

Field	Type	Description
<code>lightMap</code>	<code>LightMap[]</code>	This is an array of alternatives mappings from events to the animation triggers.
<code>playerErrorCubeLights</code>	<code>string</code>	The animation trigger name for the <code>playerErrorCubeLights</code> event.
<code>startGameCubeLights</code>	<code>string</code>	The animation trigger name for the <code>startGameCubeLights</code> event.

Table 418: Cube spinner light map JSON structure

Each of the `LightMap` structures has the following fields:

Field	Type	Description
<code>backpackLights</code>	<code>BackpackLightMap</code>	This structure maps event to animation trigger names appropriate for the backpack light animation.
<code>cubeLights</code>	<code>CubeLightMap</code>	This structure maps event to animation trigger names appropriate for the cube light animation.
<code>debugColorName</code>	<code>string</code>	A name like "blue". This likely is used to provide the active mapping to a tool during development.

Table 419: `LightMap` JSON structure

`BackpackLightMap` is a structures used to map an event to an animation trigger name. The animation trigger name is mapped to backpack light animation. This structure has the following fields:

Field	Type	Description
<code>celebration</code>	<code>string</code>	The animation trigger name for the <code>celebration</code> event.
<code>holdTarget</code>	<code>string</code>	The animation trigger name for the <code>holdTarget</code> event.
<code>selectTarget</code>	<code>string</code>	The animation trigger name for the <code>selectTarget</code> event.

Table 420: `BackpackLightMap` JSON structure

`CubeLightMap` is a structures used to map an event to an animation trigger name. The animation trigger name is mapped to cube light animation. This structure has the following fields:

Field	Type	Description
<code>celebration</code>	<code>string</code>	The animation trigger name for the <code>celebration</code> event.
<code>cycle</code>	<code>string</code>	The animation trigger name for the <code>cycle</code> event.
<code>locked</code>	<code>string</code>	The animation trigger name for the <code>locked</code> event.
<code>lockedPulse</code>	<code>string</code>	The animation trigger name for the <code>lockedPulse</code> event.
<code>lockIn</code>	<code>string</code>	The animation trigger name for the <code>lockIn</code> event.

Table 421: `CubeLightMap` JSON structure

99. BACKPACK LIGHTS ANIMATION

The sequence to illuminate the backpack lights is found by

1. The Cube Spinner name produces an animation trigger name
2. The animation trigger name is mapped to an animation file
3. The animation file provides the sequence to illuminate the backpack lights.

99.1. TRIGGER MAP CONFIGURATION FILES

The table mapping the animation trigger name to the backpack lights animation file is found in the following file:

/anki/data/assets/cozmo_resources/**assets/cladToFileMaps/BackpackAnimationTriggerM
ap.json**

*mapping trigger
names to light
sequences*

This path is hard coded into vic-anim. This file maps the trigger name to the name of the animation file. The file's schema is the same as in Chapter 20, section 94.2 *Trigger Map Configuration files*

99.2. THE BACKPACK LIGHTS PATTERN

Vic-anim controls the backpack lights based on specifications in JSON files in

/anki/data/assets/cozmo_resources/**config/engine/lights/backpackLights/**

The path is hard coded into vic-anim. All of the JSON files have the same structure with the following fields:

Table 422: The Backpack LEDs JSON structure

Field	Type	Units	Description
<i>offColors</i>	array of 3 colors	RGBA	Each color corresponds to each of the 3 lower back pack lights. Each color is represented as an array of 4 floats (red, green, blue, and alpha), in the range 0..1. Alpha is always 1.
<i>offPeriod_ms</i>	int[3]	ms	The “off” duration for each of the 3 back pack lights. This is the duration to show each light in its corresponding “off” color (in <i>offColors</i>).
<i>offset</i>	int [3]	ms	This holds how many milliseconds each light’s clock is advanced from the clock driving the animation. This is used to stagger each lights progression through the animation sequence.
<i>onColors</i>	array of 3 colors	RGBA	Each color corresponds to each of the 3 lower back pack lights. Each color is represented as an array of 4 floats (red, green, blue, and alpha), in the range 0..1. Alpha is always 1 (the value is ignored).
<i>onPeriod_ms</i>	int[3]	ms	The “on” duration for each of the 3 lights. This is the duration to show each light in its corresponding “on” color (in <i>onColors</i>).
<i>transitionOffPeriod_ms</i>	int [3]	ms	The time to transition from the on color to the off color.
<i>transitionOnPeriod_ms</i>	int [3]	ms	The time to transition from the off color to the on color.

Note: These sequences do not have the parametric variation based on emotion or random weighting.

100. CUBE LIGHTS ANIMATION

The sequence to illuminate the backpack lights is found by

1. The Cube Spinner name produces an animation trigger name
2. The animation trigger name is mapped to an animation file
3. The animation file provides the sequence to illuminate the cube lights.

100.1. TRIGGER MAP CONFIGURATION FILES

The table mapping the animation trigger name to the cube lights animation file is found in the following file:

/anki/data/assets/cozmo_resources/**assets/cladToFileMaps/CubeAnimationTriggerMap.json**

This path is hardcoded into libcozmo_engine.so. This file maps the trigger name to the name of the animation file.

The file's schema is the same as in Chapter 20, section 94.2 *Trigger Map Configuration files*,

100.2. CUBE ANIMATIONS

The cube light animation files are located in:

/anki/data/assets/cozmo_resources/**config/engine/lights/cubeLights**

and within folders (and sub-folders) therein. This path is hard-coded into libcozmo-engine.

All of the cube light animation JSON files have the same structure. They are an array of structures. (There is usually one item, but there may be more.) Each structure may contain the following fields:

Field	Type	Units	Description
<i>canBeOverridden</i>			Default is true. <i>Optional</i> .
<i>duration_ms</i>	float	ms	If zero, do this until told to stop, otherwise perform the animation for this period and proceed to next structure or stop.
<i>pattern</i>			A structure describing the light patterns. Described below.
<i>patternDebugName</i>	string		A text name that is associated with this structure. <i>Optional</i> .

The pattern structure contains the following fields:

Field	Type	Units	Description
<i>offColors</i>	array of 4 colors	<i>RGBA</i>	Each color corresponds to each of the 4 cube lights. Each color is represented as an array of 4 integers (red, green, blue, and alpha), in the range 0..255. Alpha is always 255.
<i>offPeriod_ms</i>	int[4]	<i>ms</i>	The “off” duration for each of the 4 cube lights. This is the duration to show each cube light in its corresponding “off” color (in <i>offColors</i>).
<i>offset</i>	int[4]		This holds how many milliseconds each light’s clock is advanced from the clock driving the animation. This is used to stagger each lights progression through the animation sequence.
<i>onColors</i>	array of 4 colors	<i>RGBA</i>	Each color corresponds to each of the 4 cube lights. Each color is represented as an array of 4 integers (red, green, blue, and alpha), in the range 0..255. Alpha is always 255.
<i>onPeriod_ms</i>	int[4]	<i>ms</i>	The “on” duration for each of the 4 cube lights. This is the duration to show each cube light in its corresponding “on” color (in <i>onColors</i>).
<i>rotate</i>	boolean		? Possibly to have the colors be assigned to the next clockwise (or counterclockwise) light periodically?
<i>transitionOffPeriod_ms</i>	int[4]	<i>ms</i>	The time to transition from the on color to the off color.
<i>transitionOnPeriod_ms</i>	int[4]	<i>ms</i>	The time to transition from the off color to the on color.

Table 424: The Cube LEDs pattern structure

This structure is very similar to that used by the backpack lights. The obvious differences are:

- There are 4 lights (instead) of three,
- The RGBA value range is 0..255; and
- There is a boolean “rotate” field.

CHAPTER 22

Video Display & Face

Vector's LCD is used to display his face, which conveys his mood and forms an emotional connection with his human, and to display the results of behaviour interactions:



- The image compositor
- The sprite manager
- Animated compositions
- Procedural face

101. OVERVIEW OF THE DISPLAY

Vector displays imagery – often moving imagery – on his display. Drawn on screen includes

- Full screen sprites — each frame is a PNG image that covers the whole display. A sequence of frames (PNGs) is drawn regularly to create the animated effect.
- Composed images and text
- Procedural face to draw the face in a complex way (more on this later)

The first two are used as part of behaviors and intents. A visual “movie” is shown when the behavior starts and another is to provide the response. The compositor map allows mixing in iconography, digits and text to show information in the response.

Vector's eyes are drawn in one of two ways:

- Using the full-screen sprites above, with the eyes pre-drawn in the PNG's
- Using procedural face which synthesizes the eyes

Note: the sprite and procedural face can be drawn at the same time, with sprites drawn over the eyes. This is done to create weather effects over Vector's face.

101.1. ORIGIN

The display system – especially the procedural face module – was pioneered in Cozmo. To prevent burn in and discoloration of the OLED display, Cozmo was given two features. First, Cozmo was given regular eye motion, looking around and blinking. Second, the illuminated rows were regularly alternated to give a retro-technology interlaced row effect, like old CRTs.

US Patent 20372659

Vector's eyes are more refined, but kept the regular eye motion. The interlacing was made optional, and disabled by default.

101.2. RENDERING SYSTEM

The display is layered, placing sprites on top, followed by the compositional layout and, finally, the procedural face.

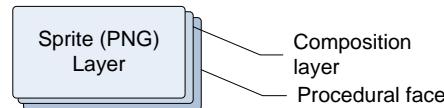


Figure 85: The display layers

The rendering and compositing these different layers look like:

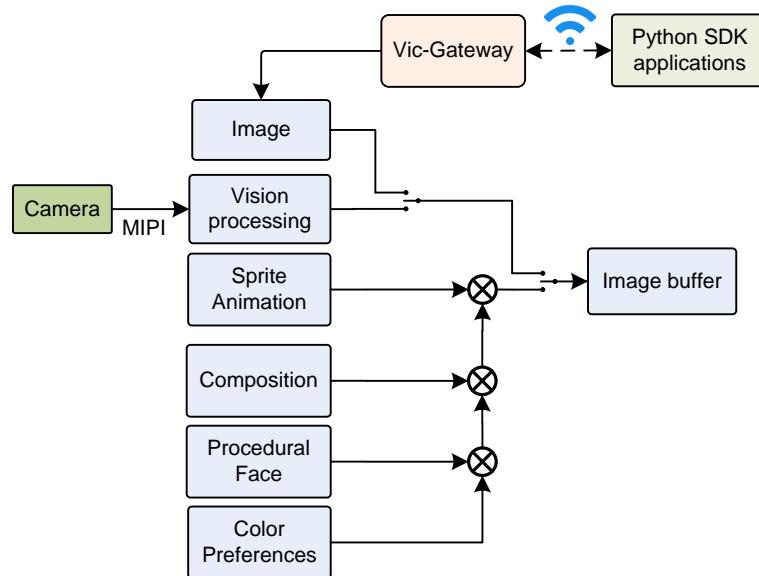


Figure 86: The display animations composition

As a functional flow, the top level is:

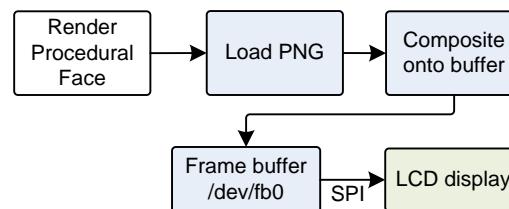


Figure 87: The display animations functional flow

102. IMAGE LAYOUT, COMPOSITION, AND SPRITE SEQUENCES

The system maps some animation trigger names to:

- A screen layout defining rectangular areas on the display (called *sprite boxes*) where images and sprite sequences will be drawn.
- Sprite sequence to display in the layout areas. Not all screen layouts have an associated sprite sequence.

102.1. MAPPING ANIMATION TRIGGER NAMES TO LAYOUT AND SPRITE SEQUENCES

There are two related files used to map animation trigger names to the layout and possible sprite sequence to display.

102.1.1

Maps to layout

The table mapping the trigger name to the layout file is found in the following file:

/anki/data/assets/cozmo_resources/assets/cladToFileMaps/CompositelImageLayoutMap.js
on

This path is hardcoded into libcozmo_engine.so. The format of the file is an array of structures.

Each structure has the following fields:

Field	Type	Description
<i>CladEvent</i>	string	This is the animation trigger name to match when looking up the animation. (This name is also defined in libcozmo_engine)
<i>LayoutName</i>	string	The name of the JSON file (without the “.json” suffix) for the animation.

Table 425: JSON structure mapping trigger name to composite image layout map

102.1.2

Maps to Image maps

The following table is used to translate an animation trigger name to an image map. This translation table is:

/anki/data/assets/cozmo_resources/assets/cladToFileMaps/CompositelImageMapMap.json

This path is hardcoded into libcozmo_engine.so. The format of the file is an array of structures.

Each structure has the following fields:

Field	Type	Description
<i>CladEvent</i>	string	This is the animation trigger name to match when looking up the animation. (This name is also defined in libcozmo_engine)
<i>MapName</i>	string	The name of the JSON file (without the “.json” suffix) for the animation.

Table 426: JSON structure mapping an animation trigger name to an image map

102.2. LAYOUT FILE

A screen layout defines rectangular areas on the display where images and sprite sequences will be drawn. The layouts are held in folders within:

/anki/data/assets/cozmo_resources/assets/compositelImageResources/imageLayouts

This path is hardcoded into libcozmo_engine.

Each layout is formatted as the array of zero or more structures, although most have a single structure. Each structure has the following fields:

Field	Type	Description
<i>images</i>	SpriteBox[]	An array of sprite boxes for showing icons and other images.
<i>layerName</i>	string	The name of the layer. (This name is also defined in vic-anim and libcozmo_engine) The animation engine may use this to select the procedure(s) in charge managing the layer and sprite boxes.

Table 427: Display layout JSON structure

A sprite box defines a rectangular region on the display to draw an image from a file. Each SpriteBox structure has the following fields:

Field	Type	Units	Description
<i>height</i>	int	<i>pixels</i>	The height of the sprite box.
<i>spriteBoxName</i>	string		The name of the sprite box. (This name is also defined in vic-anim and libcozmo_engine.) The animation engine may use this to select the procedure(s) in charge managing the layer and sprite boxes. If an image map is available for this animation, the sprite sequence it describes will be displayed in this rectangle.
<i>spriteRenderMethod</i>	string		“CustomHue” if the PNG images should be converted from gray scale to the colour using the current eye colour setting. “RGBA” if ??
<i>width</i>	int	<i>pixels</i>	The width of the sprite box.
<i>x</i>	int	<i>pixel</i>	The x coordinate of the upper left hand corner of the sprite box. The coordinate must be within the display area (0..183).
<i>y</i>	int	<i>pixel</i>	The y coordinate of the upper left hand corner of the sprite box. The coordinate must be within the display area (0..95).

Table 428: Sprite box JSON structure placing an image on the display

The position plus the size of the area must fit within the display area.

$$\begin{aligned}x + \text{width} &< 184 \\y + \text{height} &< 96\end{aligned}$$

See also Chapter 25 section *113.16 SpriteBox* for an alternate method to define a sprite box.

102.3. IMAGE MAP FILE

An image map describes which sprite sequence to display (it just has a lot of extra steps). The image map files are held in folders within:

/anki/data/assets/cozmo_resources/assets/compositeImageResources/imageMaps

This path is hardcoded into libcozmo_engine.

Each image map file is formatted as the array of zero or more structures, although most have a single structure. Each structure has the following fields:

Field	Type	Description
<i>images</i>	SpriteMapBox[]	An array of sprite boxes for showing sprite sequences.
<i>layerName</i>	string	The name of the layer. This name is also defined in <code>vic-anim</code> and <code>libcozmo_engine</code> . The animation engine may use this to select the procedure(s) in charge managing the layer and sprite boxes.

Table 429: Image map JSON structure

Each SpriteMapBox structure has the following fields:

Field	Type	Units	Description
<i>spriteBoxName</i>	string		The name of the sprite box. (This name is also defined in <code>vic-anim</code> and <code>libcozmo_engine</code> .)
<i>spriteName</i>	string		The name of a sprite sequence.

Table 430: Sprite map box JSON structure

102.4. SPRITE SEQUENCES

Sprite sequences are PNG files that are displayed sequentially— each PNG makes up each frame of the animation. These image files are held in folder with the same name as the sprite sequence name. The sprite sequence folders are held in:

/anki/data/assets/cozmo_resources/assets/sprites/spriteSequences

and

/anki/data/assets/cozmo_resources/config/devOnlySprites/spriteSequences

sprite sequences

These paths are hardcoded into `vic-anim`.

The sprite sequence PNG files are sized to fill the display. The images *must* match the width and height of the display (184x96). The images may be colored, or in gray scale with an alpha channel. If the sprite is grey-scale, it will be colourized with the current eye colour setting, using the gray scale for the pixel brightness level.

These sprites are displayed as a sequence. The frame number is appended to the file name, starting with 0.

The images are composited on top of the eye layer. The eyes may haven be turned off, or they may be present.

102.5. DISPLAYING TEXT ON THE SCREEN

When Vector is operating, almost all of the text displayed is composed from image files (sprites). There are two additional procedures that Vector can use to put text on the display:

- `drawTextOnScreen()` (part of `libcozmo_engine`)
- `OpenCV's putText()` (part of OpenCV)

These are procedures are only used in exceptional circumstances. (The typeface is inelegant, if they were something Vector used more frequently; undoubtedly they'd have improved typeface

designs.) They are used to display the fault codes (via Vic-faultCodeDisplay), when the system is unable to operate the software; and to display information on the customer care information screen (CCIS) in vic-anim.

103. PROCEDURAL FACE

Vector's dynamic, moving eyes are brilliant, forming *the gateway for an emotion connection*. They allow Vector to give eye contact, facial expressions, and his current sentiment. These eyes are drawn by the procedural face manager.

The parameters of the face controls are divided into the overall view of the face and the individual characteristics of each eye:

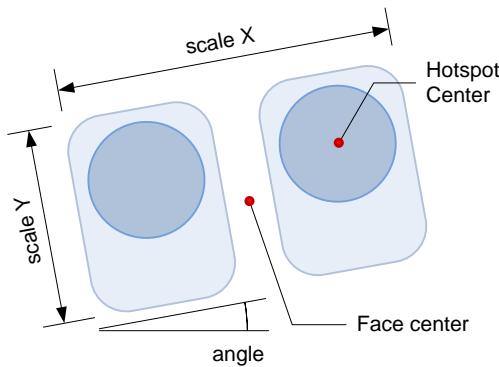


Figure 88: Face control points

The high level face animation parameters include:

- The color to draw the eyes in. Vector's eye color is a preference setting, but can be temporarily overridden by the SDK.
- The position of the center of the face
- The angle of the face; tilt (or rotation) of the face gives the impression of tilting the head
- The scaling of the height and width of the face
- The illusion of gaze – the intuition that Vector is looking at something – is achieved by giving each eye a soft spherical rounding effect. The center of the shading, the equivalent of a pupil, may be moved around the eye area. This gives a sense of where Vector is looking – and by moving the center, Vector can appear to be looking around. Coordinated with the face detector, Vector can make (and maintain) direct eye contact.
- The outer shape of the eyes, which gives a sense of the emotions – smiling, frustration, sleep etc.
- There is a *scan line opacity* factor. This controls how much alternating lines are illuminated and darkened. A value of 1.0 has odd and even lines with the same coloring.

Where the eyes are looking is controlled within the procedural face manager, rather than in the animation files. It controls the blink rate, the focus of the eyes and how much the eyes dart around. The manager contributes looking at a face and making eye contact.

103.1. THE RENDERING OF INDIVIDUAL EYES

The eyes are rendered with a gradient and a shape that is controllable by the animations. These create a soft-face feel, combining the soft glow of the eye, along with rounded eyelids and cheeks.

103.1.1

The Hot spot

The interior of the eye is rendered as a radial gradient from the eyes pupil, with the shape of the eyes forming the clip path. The location of the center of each eye's 'pupil' is called the *hotspot center*:

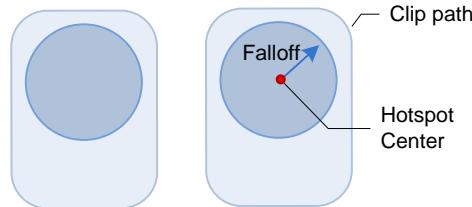


Figure 89: The hotspot and center

The shape of the eyes is parametrically controlled by the animation engine. An internal configuration variable controls how fast the shading falls off from the center toward the edge. A bit of random noise is added to the gradient to give the eyes shading a little texture. This too has an internal configuration variable to control the noise factor.

103.1.2

The eye-shape clip path

Each eye has individual animation parameters that control its shape. These create the rounded eyelids and cheeks by masking off some of the eye pixels. A line is drawn along the outer path to complete the effect.

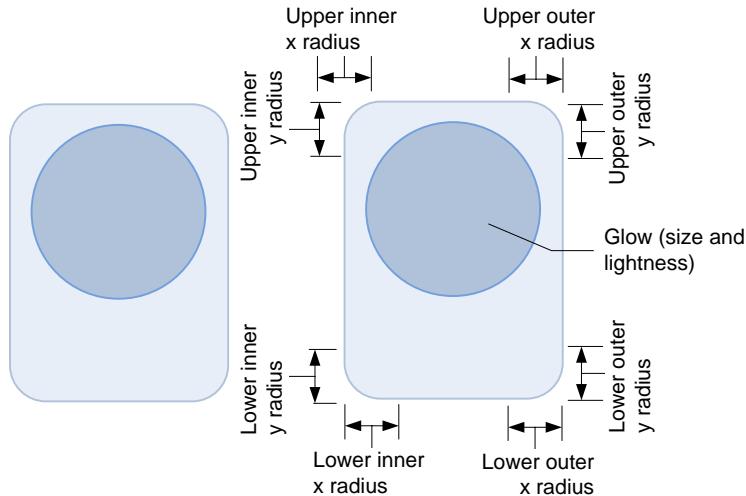


Figure 90: Basic parameters of an individual eye control

- The basic shape of the eye is controlled by the roundedness of the corners.
- The position of each eye is controlled by the center of the face;
- The size and width of the eye is created by the face's scaling factors

Each eye has controls for its eyelids (or cheek, depending on your perspective):

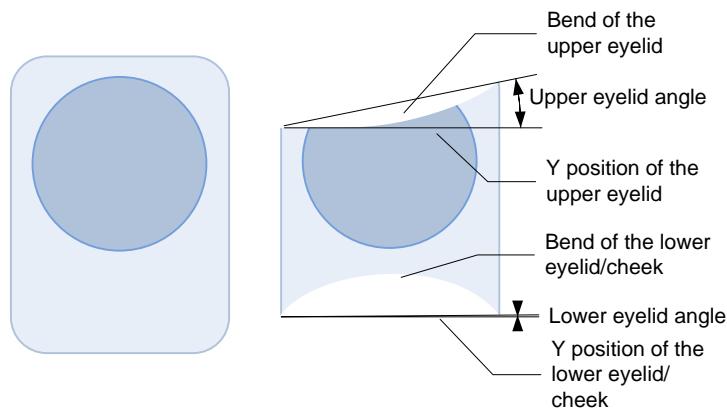


Figure 91:
Parameters of an eyes
eyelids (or cheeks)

- An arc represents the upper eye-lid and erases (or occludes) the upper portion of the eyes; these help create the sleepy, frustrated/angry emotions.
- An arc represents the lower eyelid and cheek, and erases (or occludes) the lower portion of the eyes; these help create the happy emotions

An eye can be made smaller – or to squint – by having no bend to the eyelids, but moving the eyelids position closer to the center.

103.2. THE PROCESS OF DRAWING THE PROCEDURAL FACE

Each eye of the procedural face can be drawn with a process like:

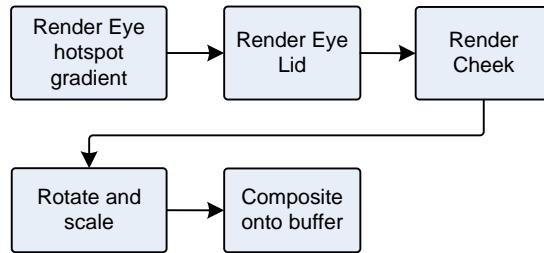


Figure 92: The
functional flow of
drawing an eye

Assuming that a complex clipping path is less efficient, the eye could be render as

1. The eye is rendered as a gradient pattern into a buffer, with the scale
2. The eye lid is drawn, forcing the pixels (of the eye lid area) to become transparent
3. The cheek is drawn, forcing the pixels (of the cheek area) to become transparent
4. The rectangle area where the eye will go is scaled, rotated, and offset for where the eye will go
5. Each pixel in the translated rounded rectangle region is map to one in the eye pixel buffer, and copied to the display buffer

104. COMMANDS

The HTTPS SDK API (Chapter 13) includes commands that affect the display

- Display RGB image (see Chapter 13 section 51.2 *Display Image RGB*)

- Mirror display (see Chapter 13 section *51.3 Enable Mirror Mode*)

CHAPTER 23

Audio Production

This chapter describes how Vector produces sounds and the audio output system:

- An overview of the audio output
- Text to speech
- Audio Effects

105. SPEAKER

Vector uses sound to convey emotion and activities, to speak, and to play sounds streamed from SDK applications and Alexa's remote servers. There are five sources of sound:

- Sound effects from playing pre-recorded audio files
- Sound effects from parametrically generating audio
- Sound from an audio stream sent by the SDK application to Vector
- Text to speech
- A sound stream from Alexa Voice Services

To support this, Vector includes a sophisticated audio architecture:

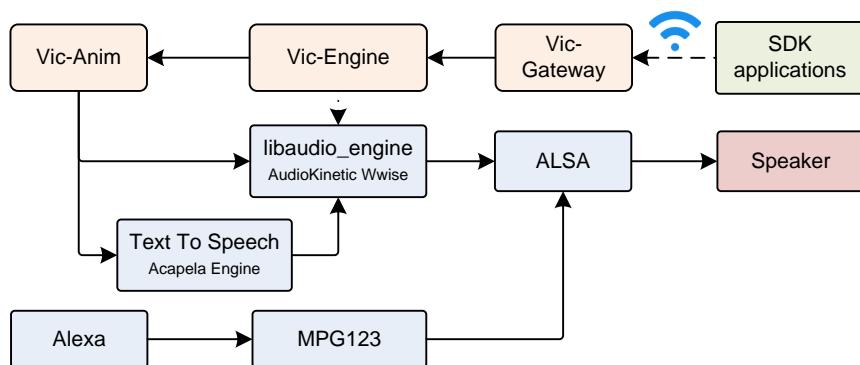


Figure 93: The audio output architecture

Compression is not used to send audio from SDK applications to Vector. The vic-engine passes the received samples to audio engine to mix in its playback.

Some key elements of the audio production are:

Element	Description
ALSA	This is Linux's sound player, which provides the device drivers to allow software access to the speaker.
Alexa Voice Services	These are a set of remote servers that provide an audio stream of Alexa's voice responses, music, and other sounds.
libaudio_engine	This is Vector's sound framework. It is built on AudioKinetic's Wwise framework. It handles audio file loading and playback, sound effects, and others.
MPG123	The MPG123 is used to decode the sound stream sent from Alexa Voice Services servers.
Text to Speech	The text to speech facility is used to convert text (written sentences) to spoken words. This is built on Acapela's speech engine.

Table 431: The audio systems functional elements

106. SOUND EFFECTS FROM AUDIO FILES AND PROCEDURES

Vector uses AudioKinetic's Wwise (*WaveWorks Interactive Sound Engine*) toolkit for sound playback. Wwise is one of the most popular high-end game sound frameworks. It has sophisticated composition tools, extensive documentation, and a redistributable player for many platforms. Such a powerful tool seems overkill on device with only one output channel and a tiny speaker. In context, it makes sense.

AudioKinetic Wwise

Wwise was used in Cozmo's mobile application. The application which was designed as a kind of video game, and employs a lot video game design approaches. So it makes sense that an audio tool targeting video games would be used there. In turn, Vector is draws on Cozmo's frameworks—both the mobile application and what ran on the hardware—and creation tools it isn't surprising that the same framework would be employed by Vector.

The key features of Wwise (at least for Vector) are:

- Triggering sound effects, muting sounds, and changing parameters of sound playback (by sending the framework audio events)
- Playing pre-recorded sounds, including looping
- Playing procedural sounds
- Playing music (in the case of Cozmo)
- Sound effects, including fading
- Change the sample rates from different sources to the one played
- Mixing different sound sources together
- Managing a library of files that specify how to respond to audio events, how to create music and sound effects, and can hold pre-recorded sounds.

106.1. SOUND PLUGINS

The Wwise framework provides hooks that allow it to be integrated into the rest of the software system, and given extra functionality. This is accomplished by *plug-in* modules:

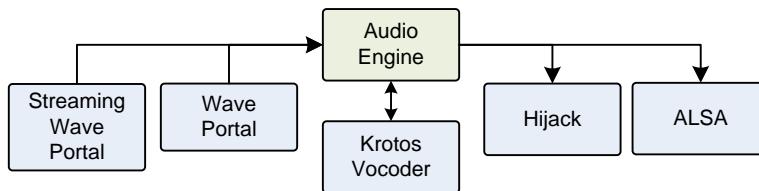


Figure 94: Plug-ins in the Wwise audio pipeline

The most import plug-ins are the ones receiving the audio output (aka “sink” plug-ins). This is how the audio sounds are taken from the audio engine and sent to Vector’s speaker.

- The ALSA plug-in gathers the audio output and passes it to the “Advanced Linux Sound System” (ALSA) sound handler, which in turn passes it thru to Qualcomm’s audio driver.
- The Hijack plug-in is probably unused on Vector, but is used on desktop computers to allow recording of Vector’s sounds...? (It may also have been intended to be used as part of the message-recording, with the microphone audio piped thru the audio engine to be filtered/cleaned, and then saved.)

There are two plug-ins allowing audio from external sources to be processed by the audio engine and delivered to Vector’s speaker:

- Wave Portal
- Streaming Wave portal. This receives the audio sounds from vic-engine for playback

Finally, there are the sound effects plug-ins:

- The Krotos “Dehumaniser” vocoder is used to give Vector his unique vocal qualities.

106.2. AUDIO PIPELINE

In a sense, Wwise can be thought of having multiple, configurable pipelines to produce sound. Each “game object” –perhaps a character, tool or machine, etc. – have its own pipelines for each kind of sound/sound-effect it would make:

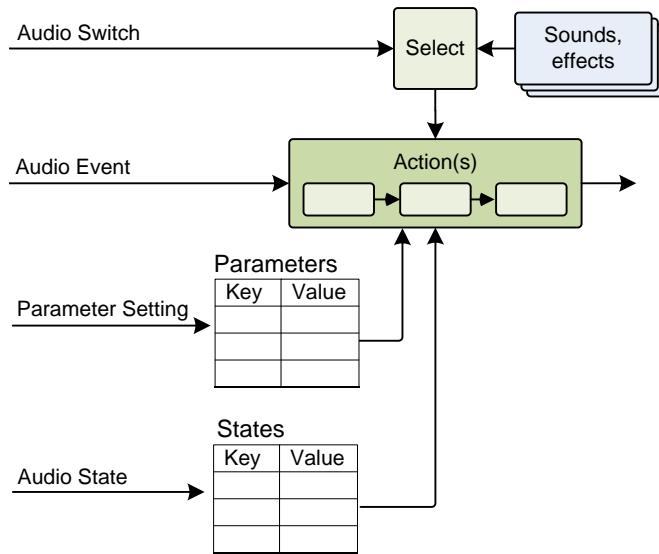


Figure 95: Overview of the Wwise audio pipeline

The audio pipeline is driven by *audio events* it receives from the main application. These events are like the animation triggers. The metaphor is that when something occurs in the application (usually a video game), it represents this as an event distributed to a variety of subsystems to respond to, including the audio engine. Typically an event will cause the audio engine to play a sound, but the event system is much more powerful than that. Events trigger an *action*, which is the heart of the pipeline, and it is the action that plays the sound. The actions, in turn, can be configured to change sound parameters, stop playback, and so on:

“[Audio] events apply actions to the different sound objects or object groups in your project hierarchy. The actions you select specify whether the Wwise objects will play, stop, pause, .. mute, set volume, enable effect bypass, and so on.”

The Wwise framework employs event ID *numbers* to refer to events. Events can also be referred to using lexical names – as strings. A later section will describe how to translate a string to an event ID.

Audio parameters are settable values used by the actions that control how they sound. Vector mainly uses these to adjust the sounds based on his current mood. Like the action, these parameters are on a per object (within the audio engine) basis.

An *audio state* is used to set the context for sound system overall, so that the right sounds and effects are used in responding to events, and actions general across all of the game objects.

An *audio switch* is similar, but it sets up the context so that the right sound (or sound effect, etc) is used for a particular object or event. AudioKinetic gives an example of foot-step events triggering a footstep sound.. but the audio switch is set to the kind of surface that the is being walked on – selecting walking on grass, gravel, pavement, etc.

audio event

audio action

audio state

audio state

audio switch

106.3. PARAMETRIC OR PROCEDURAL SOUND GENERATION

Vector has the ability to modify sound effects parametrically. The behaviour-animation systems uses this convey Vector's emotional state – sadness, approval, etc. These are like *intersectional tones* that people make – grumbles, and grunting. When an action is being performed, the animation may trigger the sound effect, perhaps to “simulate the physical movement” he is making. The sound effect parameters are modified by the current emotional state, or anticipation – to convey whether he is struggling to do the task, is excited, is frightened, etc.

The mood manager sets audio parameters based on the current mood. This gives “feedback cues about the robot's emotion state.” See Chapter 27, section 117.2 *The emotion model*.

The binary animation file can trigger sending audio events (by ID), setting parameters, states and switches. See Chapter 25, section 113.3 *AudioEventGroup* for more information.

The behaviors may emit audio events.

106.4. EQUALIZER

The Wwise sound equalizer is used to compensate for some of the distortion of Vectors small speaker. It is also used to “prevent the higher pitches from ever getting very loud” – something that physically is possible despite the speakers small dimensions. The standards for toy sound levels vary by country, but typically are limited to 75-80dB at the ear.

106.5. THE CONFIGURATION

The audio configuration and sound files are located in:

/anki/data/assets/cozmo_resources/sound

There are three kinds of files located there: configuration files, sound bank files (which may include sounds), and sound files.

106.5.1 The configuration files

These configuration file is

/anki/data/assets/cozmo_resources/sound/SoundbankBundleInfo.json

This file name is hard coded into libaudio_engine. The file is n array of structures, each with the following fields:

Field	Type	Description & Notes	
<i>bundle_name</i>	string	A name?	
<i>language</i>	string	“SFX” or “English(US)” (It isn't clear how to interpret these.	
<i>path</i>	string	The path of the sound bank file, relative to the location of the configuration file.	
<i>soundbank_name</i>	string	The name of the sound bank file, without the “.bnk” extension.	

Table 432: The sound bank bundle info JSON structure

106.6. THE SOUND FILES

106.6.1

Sound Bank Files (BNK)

A sound bank is a binary file, composed of series tagged sections. Some sections are

- The sound bank file has setups for how the sounds flow from files (and other inputs), thru mixers, and other filters to the output. It calls this the *audio bus hierarchy*.
- Sound effects,
- Music compositions to play, (these probably are used heavily in Cozmo, but appear unused in Vector)
- State transition management, how to altering the settings of effects during play.
- A map of *audio events* to the actions to carry out when that event occurs, such as playing a sound, stopping other sounds, changing mixer settings, and so on
- The set of other sound bank files to load.
- WEM sound files, either embedded, or a reference to an external file.

Wwise always has an “Init.bnk” sound bank. It is loaded first, since it holds sections that are shared across all of the sound bank files. It does not contain any sounds.

106.6.2

“Wwise Encoded Media” WEM Files

WEM files are sound files. They are considered *containers*, since it is possible for the sound (in the file) to be encoded in different formats, some standard and some custom. The file format is custom to AudioKinetic, and automatically produced by AudioKinetic’s WWise tool. Like the BNK file format, a WEM file is organized as tagged sections. The file names are ID numbers with a “.wem” extension. (More on the generation of the ID numbers in the section.)

WEM files also including optional looping parameters. A sound file can be configured to loop indefinitely or a fixed number of times. looping

In practice, all of Vector’s WEM sound files are single channel (mono), 16-bit little-endian IMA ADPCM files. The only thing that varies that there sample rate varies. “IMA ADPCM” is an ugly, confusing acronym, but it is rather simple in practice:

IMA ADPCM
adaptive-differential
audio encoding

- The decoder produces 16-bit values to feed to a digital to analog converter (DAC) and the amplifier and speaker; that is the *pulse code modulation* (PCM) portion.
- The sound file only has 4 bit values for each sample. Each 4 bit value is used as the index into a pair of look up tables for how much to add to or subtract from the previous 16-bit value for the new output; this is the *adaptive differential* (AD) portion.
- The tables and their interpretation are standardized by a committee, which is the IMA portion.

This approach is easy for the processor, and takes little working memory. It does make for larger files than, say, MP3 or other more sophisticated compression. That is acceptable since the sound segments are all short, and Vector has a large storage area to hold the files.

This approach has one drawback. It uses only 4 bits in each sample to represent the change in the analog waveform. Often this isn’t enough; it takes several samples to add or subtract enough for the output to catch up to the desired values (from the source material). At “low” sample rates, this can create audible distortion. The fix is to use a higher sample rate for encoding. First, there is less change between two points closer together in time. Second, the higher rate lets the decoder

catch up faster; this effectively makes the distortion at high, inaudible frequencies. The play back can be done with this higher sample rate, or it can be down-sampled again after decoding.

Vector (probably) down-samples many of the sound files (after decoding) during playback. The audio files have sample rates much higher than is supported through the SDK audio channels: many at 30,000 samples/sec, some as high as 44,100 sample/sec.

106.6.3

Wave files

The audio engine includes a facility to load wave files for sound (so long as they are RIFF with PCM data). It likely it was added in preparation for the “messaging” feature. A message from a friend could be recorded, and distributed to a particular Vector. vic-engine could trigger the playback of the audio by vic-anim.

106.7. MAPPING AUDIO EVENT AND SOUND NAMES TO ID NUMBERS

The ID number used to identify the events and audio files is not random or entirely arbitrary. It is formed from the text names used by sound engineers and software developers. The number is automatically made from the text names of the events and sound by the AudioKinetic software⁴⁵ using a Fowler-Noll-Vo hash function. (The number is used instead of strings to reduce runtime memory and processing overhead during game play.)

*IDs from names
Fowler-Noll-Vo hash*

The algorithm to compute the 32-bit hash number:

1. “Start with an initial hash value of FNV offset basis.” (Use 16777619 for this offset)
2. “For each [character] in the input,
 - a. “multiply [the] hash by the FNV prime,” (use 2166136261 for the prime.)
 - b. convert the next character to lower case
 - c. XOR the hash with the lower case character

107. TEXT TO SPEECH

Vector includes a text to speech (TTS) facility. The engine is based on Cozmo’s text-to-speech subsystem, with the text-to-speech engine from Acapela. The text to speech engine is part of vic-anim, with some components in libcozmo_engine. Vector treats the process of speaking as a type of animation, one with just an audio track. The audio for the speech is generated, and then handled by the animation manager.⁴⁶

text to speech

The text-to-speech voices are stored in

`/anki/data/assets/cozmo_resources/tts`

The voice files include:

- co-French-Bruno-22khz
- co-German-Klaus-22khz
- co-Japanese-Sakura-22khz
- co-USEnglish-Bendnn-22khz

⁴⁵ https://www.audiokinetic.com/library/edge/?source=SDK&id=_ak_f_n_v_hash_8h_source.html

⁴⁶ <https://forums.anki.com/t/multiple-actions-possibility-for-sdk/104>

107.1. THAT DISTINCT ROBOTIC VOICE QUALITY

The text to speech engine produces human-sounding speech... so how does it get to be Vector's *robotic* voice? The audio is pumped through a vocoder (the Krotos Dehumaniser) in the sound engine to give it a distinct sound.



Figure 96: High-level vocoder functional flow

The exact implementation isn't known, but there are common techniques. A typical vocoder works by estimating the pitch of the text-to-speech voice every 30ms, and then adjusting the gains settings on a multi-band equalizer.

107.1.1 Pitch tracker

Krotos' pitch detection – or pitch tracker – can be configured to use one of five different algorithms:

1. Autocorrelation
2. Cepstrum
3. McLeod pitch detection method (MPM)
4. Frequency spectrum-based
5. YIN

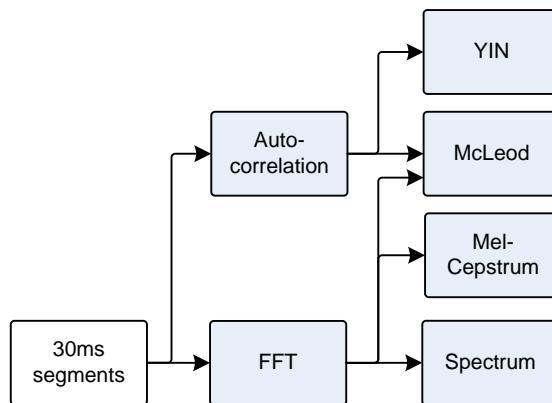


Figure 97: Pitch tracking methods

All these methods have distinctions between them, but there are far more similarities. They all build on basic techniques like autocorrelation, and *fast-fourier transform* (FFTs). An FFT computes a spectrograph, giving the strength of each frequency. The simplest (or naive) approach is to find the strongest frequency and call it a pitch. This approach is easily fooled. A variety of other techniques have been developed to work around this.

107.1.2 Autocorrelation

Autocorrelation is a slow, brute-force algorithm for finding the pitch. It works by shifting the signal in slight amounts until the shifted signal best matches the original one. The core algorithm looks something like:

```
for each sample offset (1... to big number)
    sum = 0
    for each sample index 0...num samples
        diff = samples[sampleIdx+sampleOfs] - samples[sampleIdx];
        sum += diff*diff;
```

Example 7:
Autocorrelation pseudo-code

keep track of the sample offsets with the smallest sums.

It does this compute the difference between each of the samples in the shifted waveform and the waveform, square that difference, and then summing it up. The shift offset with the smallest offset is the best match, and used to compute the pitch. Note: It only needs to try the offsets that cover the first few kHz at the given sample rate.

This approach is an expensive way to find the pitch: Every audio sample must be scanned a huge number of times.

107.1.3

YIN detection

YIN is uses auto-correlation and adds a few more “polishing” steps to improve its estimate: it weights the different offsets, and blends together a few of the offsets that are close to the best one.

107.1.4

McLeod Pitch Method (MPM)

McLeod’s pitch tracker (named for its author) improves on YIN in two ways. First is that it can be implemented using a FFT to perform the autocorrelation step, and to estimate the pitch. (It can still use the brute force method, if an FFT isn’t available.) This FFT autocorrelation is done by

1. Performing an FFT on the sample window.
2. Compute the square magnitude of each complex value – that is, multiply each complex number by its complex conjugate. (Or, for those not steeped in the jargon, $real*real*+imag*imag$, ignore the part where multiplying two imaginary numbers becomes negative)
3. Compute the inverse FFT of this to compute the power vs frequency
4. Identify the frequency with the highest power associated with it

McLeod’s method, adds a few more polishing steps as well to clean this up and mix together a few of the best results to get a better one.

107.1.5

Mel Cepstrum

The Mel Cepstrum method is similar to the FFT-base auto correlation, but tweaked to take into account the psychology of pitch:

1. Performing an FFT on the sample window.
2. Compute the square magnitude of each complex value
3. Change from frequency (Hz) to “mel scale,” which is a perceptual scale for pitch
4. Compute the logs of all of those numbers on this scale
5. Compute a discrete cosine transform of this to compute the amplitude vs frequency
6. Identify the frequency with the highest amplitude associated with it

107.2. THE CONFIGURATION AND LOCALIZATION FILES

The configuration file for the text to speech engine is located at:

/anki/data/assets/cozmo_resources/config/engine/tts_config.json

(This path is hardcoded into vic-anim.) This file is organized as dictionary whose key is the operating system. The “vicos” key is the one relevant for Vector.⁴⁷ This dereferences to a dictionary whose key is the language base: “de”, “en”, “fr”, or “ja”. The language dereferences to a structure with the following fields:

Field	Type	Description & Notes
<i>pitch</i>	float	This is a pitch setting field. This is not supported by all voices / platforms. (The comment says that this is Acapela TTS SDK.) “Pitch... adjustment is actually performed by audio layer.” <i>Optional</i> .
<i>shaping</i>	optional	
<i>speed</i>	float	
<i>speedTraits</i>	speedTraits[]	Array of speed traits structures (see below). <i>Optional</i>
<i>voice</i>	string	a path to the ini file within the [assets/cozmo_resources/tts] folder

Table 433: The JSON structure

Each speedTraits structure has the following fields:

Field	Type	Description & Notes
<i>rangeMax</i>	uint	
<i>rangeMin</i>	uint	
<i>textLengthMax</i>	uint	
<i>textLengthMin</i>	uint	

Table 434: The speedTraits JSON structure

107.2.1 Localization

Vector internally has support for German, French, and Japanese, but the application-level language settings only really support US, UK, and Australian dialects of English. The files for non-English localization were not completed.

The localization files for feature stores its text strings (to be spoken) in

/anki/data/assets/cozmo_resources/LocalizedStrings

This path is not present in versions before v1.6. The folder holds sub-folders based on the language:

de-DE en-US fr-FR

Note: there is no ja-JA, but it may be possible to create.

Inside of each are three files intended to provide the strings, for a behaviour, in the locale:

- BehaviorStrings.json
- BlackJackStrings.json
- FaceEnrollmentStrings.json

⁴⁷ The other OS key is “osx” which suggests that Vector’s software was development on an OS X platform.

Each JSON file is a dictionary with the following fields:

Field	Type	Description & Notes
<i>smartling</i> ⁴⁸	structure	see below to the structure below. Note all smartling structures examined are the same.

Table 435: The JSON structure

The dictionary also includes keys, such as “BehaviorDisplayWeather.Rain” that map to a locale specific string. These have the following fields:

Field	Type	Description & Notes
<i>translation</i>	string	The text in the given locale. The string may have placeholders, such as {0}, where text is substituted in.

Table 436: The JSON structure

Each smartling structure has the following fields:

Field	Type	Description & Notes
<i>placeholder_format_custom</i>	array of strings	An array of patterns that represent possible placeholder patterns.
<i>source_key_paths</i>	array of strings	“/{*}” Strings are path of a JSON key?
<i>translate_paths</i>	array of strings	“*/translation” Strings are path of a JSON key?
<i>translate_mode</i>	string	“custom”
<i>variants_enabled</i>	boolean	

Table 437: The JSON structure

This is handled by libcozmo_engine, including the key strings.

WEATHER FILES

The weather behaviour stores its text strings (to be spoken) in

/anki/data/assets/cozmo_resources/config/engine/behaviorComponent/weather/condition_to_tts.json

This path is hardcoded into libcozmo_engine. The JSON file is an array of structures. Each structure has the following fields:⁴⁹

Field	Type	Description & Notes
<i>Condition</i>	string	e.g. “Cloudy”, “Cold”
<i>Say</i>	string	The key used in the BehaviourStrings.json file to look up the localized test. (In previous versions, this was the text to say, in English.)

Table 438: The JSON structure

107.3. CUSTOMIZATION

Vector’s voice files are from Acapela. Acapela sells language packs for book readers, but the format appears different and likely very difficult to modify or create.

⁴⁸ it really has this key.

⁴⁹ That this file (and many others) is a simple 1:1 transform lends the suspicion that the localization process is needlessly complex.

Cozmo's employs a different English voice (in the Cozmo APK). This likely could be extracted and used on Vector. (In turn, Vectors voice could probably be used with Cozmo.)

Customization of the Localization TTS would give Vector a bit more personality.

108. COMMANDS

The HTTPS SDK API (Chapter 13) includes commands that affect the sounds

- Audio stream commands (see Chapter 13 section *46.6 External Audio Stream Playback*)
- Text to speech (see Chapter 13 section *46.8 Say Text*) An external application can direct Vector to speak using the Say Text command. The response(s) provide status of where Vector is in the speaking process.
- Vector's volume can be set as a setting using the UpdateSettings command (see Chapter 13 section *61.2 Update Settings*) and the RobotSettingsConfig structure (see chapter 29), or using the Master Volume command (see Chapter 13 section *46.7 Master Volume*). Note: the volume levels using settings doesn't fully match those in the master volume command.

Note: can also trigger animations which play sounds effects as well.

109. REFERENCES AND RESOURCES

AudioKinetic , *Wwise Fundamentals* (2015)

https://wwwaudiokineticcom/download/documents/Wwise_Fundamentals.pdf

AudioKinetic, *Wwise User's Guide*

https://wwwaudiokineticcom/files/?get=2015.1.9_5624/Wwise_UserGuide_en.pdf

AudioKinetic, *The Wwise Project Adventure*

https://wwwaudiokineticcom/download/documents/WwiseProjectAdventure_en.pdf

AudioKinetic, *Get Started Using Wwise*

https://wwwaudiokineticcom/download/documents/Wwise_GetStartedGuide.pdf

AudioKinetic, *Wwise 101*

https://wwwaudiokineticcom/download/lessons/wwise101_en.pdf

AudioKinetic, *Wwise Fundamentals, Understanding Events*

https://wwwaudiokineticcom/library/edge/?source=WwiseFundamentalApproach&id=understanding_events_understanding_events

Cheveigne, Alain de. (2002). "YIN, a fundamental frequency estimator for speech and music," *Journal of the Acoustical Society of America*. Vol 111(4), pp 1917-30.

http://audition.ens.fr/adc/pdf/2002_JASA_YIN.pdf

Krotos, *Dehumaniser Live*

<https://wwwkrotosaudiocom/dehumaniser-live/>

<https://wwwkrotosaudiocom/products/dehumaniser2/>

<https://s3-us-west-2.amazonaws.com/dehumaniser/Manuals/Dehumaniser+2+Manual.pdf>

Krotos, *Dehumaniser Live, Integration With Wwise*, 2017

<https://s3-us-west-2.amazonaws.com/dehumaniser/Manuals/Dehumaniser+Live+for+Wwise+Manual.pdf>

McLeod, Philip and Wyvill, G., "A Smarter Way to Find Pitch" (2003), *Proc. International Computer Music Conference*, Barcelona, Spain, September 5-9, 2005, pp 138-141.
<http://miracle.otago.ac.nz/tartini/papers.html>

Sweet, Michael, *Creating Music for Robots: Interview with the Composing Team for Cozmo*, Design Music Now, 2016 Dec 12

<https://wwwdesigningmusicnowcom/2016/12/12/creating-music-robots-interview-composing-team-cozmo/>

Wikipedia, *Cepstrum*
<https://en.m.wikipedia.org/wiki/Cepstrum>

Wikipedia, *Fowler-Noll-Vo hash function*
https://en.wikipedia.org/wiki/Fowler%20%93Noll%20%93Vo_hash_function

Wikipedia, *Mel-Frequency Cepstrum*
https://en.m.wikipedia.org/wiki/Mel-frequency_cepstrum

Wikipedia, *Mel scale*
https://en.wikipedia.org/wiki/Mel_scale

Wikipedia, *Pitch detection algorithms*
https://en.wikipedia.org/wiki/Pitch_detection_algorithm

Xentax, *Wwise SoundBank (*.bnk)*, 2012 Dec 6
[http://wiki.xentax.com/index.php/Wwise_SoundBank_\(*.bnk\)](http://wiki.xentax.com/index.php/Wwise_SoundBank_(*.bnk))

This site provides a wealth of information on the format of the Sound Bank files.
Unfortunately not all sections of the file have been documented, and there are sections in
Vector's Sound Bank files that were not known when this page was written

Some example code for YIN and McLeod pitch tracking
<https://github.com/ashokfernandez/Yin-Pitch-Tracking>
https://github.com/adamski/pitch_detector/blob/master/source/PitchMPM.h
<https://github.com/sevagh/pitch-detection/blob/master/src/mpm.cpp>

CHAPTER 24

Motion Control

This chapter describes the motion control subsystem:

- The control of the motors
- Performing head and lift movements
- Moving along paths in a smooth and controlled fashion

110. MOTION CONTROL

The motion control is designed to take a path of movements from the path planner or the animation systems. The path consists of arc, line, and turn (in place) movement commands. These can be coordinated with the head and lift, by the animation system.

Note: the animation system is described in chapter 20

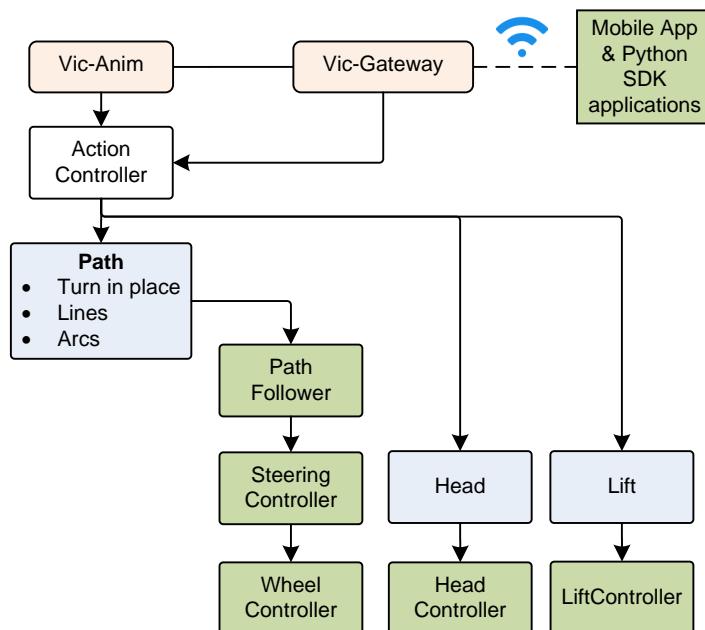


Figure 98: Motion controller

The path planner thinks of the word and robot coordinates within it in terms of x,y and θ (theta) coordinates. The θ being the direction angle that Vector is facing at the time. It builds a list of straight line segments, arcs, and point turns. The PathFollower carries these out. Each of the motors is independently driven and controlled, with the steering controller coordinating the driving actions. Sets gains, executes turns, does docking,

The individual motors have controllers to calibrate, move, prevent motor burnout, and perform any special movements.

110.1. FEEDBACK

The motion controller may take position and orientation feedback from

- The linear speed can be estimated from the motors shaft rotation speed (and some estimated tread slip), merged with IMU information
- The speed that the robot is rotating can be measured by the IMU and the vision system.
- The navigation and localization subsystem, which employs a sophisticated Kalman filter on all of the above position.

110.2. MOTOR CONTROL

Vector's small, light frame and powerful (for their size) motors allow him to make quick actions representing his emotions and little tantrums, as well as drive about smoothly. A typical speed and position controller for the brushed DC motors is a set of PID control loops. (Although the “d” – derivative – term is often small or not needed.)

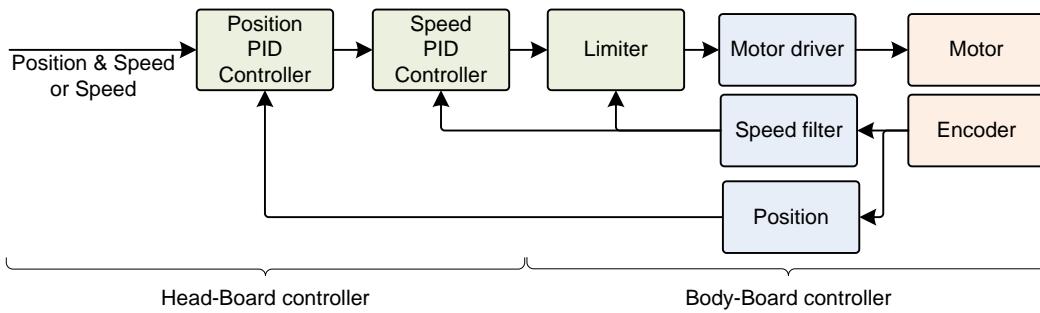


Figure 99: A typical motor control loops, as they might appear in Vector

The motor control loops are implemented in the head-board. They are implemented using floating point (rather than the fixed point⁵⁰ that the body-board's M0 microcontroller would require), and are updated 200 times per second. The body-board is responsible for driving the motors and sampling the encoder. It is also responsible for protecting the motors in case of a stall.

The lift and head motors are position-controlled. The motors can be commanded to travel to an encoder position at a speed (given in radians/sec). The position – the cumulative number of radians that the shaft has turned – can be computed by counting the encoder events, with the expected direction that the motor has turned.

The speed of rotation is also computed, indirectly, from the encoder count. The typical approach is to regularly take a derivative of the position (say once every millisecond), and filter it. Since the encoder is discrete, at slow speeds its update rate will produce false measures of shaft speed.

110.3. BURN OUT PROTECTION

The body-board is responsible for protecting the motors. It monitors the speed of the motors. If a stall is detected, it limits the current to the motor (by limiting the duty-cycle). The motor is still driven, but at such a low power that burn out is not a risk. This lets the head and lift hold position. In case the hand or other obstruction is removed, the motor can sense the change (the encoder will show that they are able to turn again) and resume.

⁵⁰ Although both approaches work, fixed point (using integers and scaling factors rather than floating point) takes a bit more effort to tune, as small but important parts of the feedback signals are dropped... this can introduce effects like jerkiness, stutter or motor noise from.

Luckily, those motors can't overheat instantaneously – it takes at least 15 seconds of being stalled at full power before you risk permanent damage. The firmware in the body board watches the encoders on all 4 motors, and turns down the power on stalled channels. It never turns the power down to 0, since it doesn't have to. All 4 motors can push continuously (gently) without stalling.

So if you drive a motor toward the limit but someone is pulling on it the other way, it might push hard at first, then quickly “relax” to a voltage that's safe for continuous use, but never stop pushing just in case you let go.

110.4. NO PINCHING FINGERS!

The motors can also be “unlocked” – allowed to be spun by external forces. This allows a person to raise and lower the lift, as well as raise and lower the head. Both of these are used as inputs to enter diagnostic modes.

In addition, the software control loops can detect when a person is playing with Vectors lift (or head or tracks), and then unlock the motors.

the PID controller violently fights your attempt to pull the lift, smacking your fingers and oscillating and otherwise causing trouble. The PID controller is pretty feisty, because it has to operate across a huge range of forces – between flipping or lifting the robot's entire weight and delicately setting down or lifting cubes without flinging them.



110.5. GETTING THE LIFT AND HEAD POSITIONS JUST RIGHT

The head and lift motors need to have their positions calibrated.

Both head and lift angle must be known exactly, since we need to know exactly where the tongs (on the lift) are relative to what the camera sees. Otherwise we couldn't engage (lift) and disengage (pull out) the block.

At startup Vector performs a calibration procedure, “which is just an animation that pushes the head/lift to [their] hard stop.” Both the lift and head have hard stops at their most downward position, which serves a well-defined starting point. When these motors reach the end of travel, the measured speed will fall below a threshold, and the software knows to zero estimated position.

Vector's software has two backups in case the position is wrong. This can happen if the calibration was wrong – something, perhaps a block or impatient human companion – prevented the head or lift from moving further. Or if someone moved his lifts or head (since the position encoder is single step, Vector won't be able to tell which direction they were moved).

1. The body-board firmware has motor burnout prevent features. This quickly drops the power applied to the motor if there is a stall.
2. If a motor is stalled unexpectedly or the motor isn't stopped (by the hard stop) within 5% of where it should, Vector schedules another calibration procedure. (This is handled by the `ReactToUncalibratedHeadAndLift` behavior.)

110.6. DIFFERENTIAL DRIVE KINEMATICS

Under ideal circumstances, these motions are straight-forward to accomplish:

- To turn in place, the treads turn at the same rate, but in opposite directions. The speed of the turn is proportional to the speed of treads

- To drive straight, both treads turn at the same speed. The speed of motion is proportional to the speed of the treads.
- To drive in an arc is done by driving the treads at two different speeds.

To drive in an arc, the left and right treads are driven speeds:

$$v_{left} = \omega \left(radius + \frac{1}{2} width \right)$$

$$v_{right} = \omega \left(radius - \frac{1}{2} width \right)$$

Equation 3: Tread speeds based on arc radius

Where

- *width* is Vector's body width
- ω is the angular velocity, i.e. speed to drive around the arc
- *radius* is the distance from the center of the arc to the center of Vector's body:

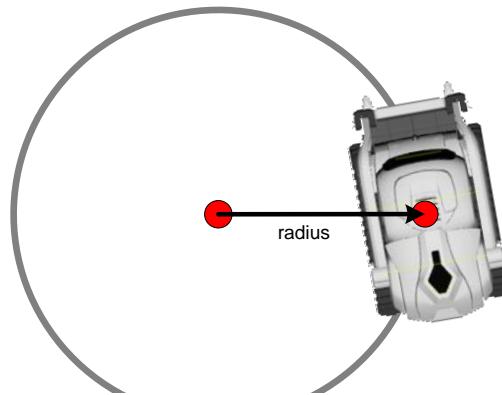


Figure 100: Radius of arc measurement

110.6.1 Slip

In practice, Vector's actual movement won't quite match what he attempted to do. Mainly this will come from how the treads slip a bit (especially while trying to push an object), and some variation in how driving the motors maps to actual motion.

111. MOTION CONTROL COMMANDS

The HTTPS SDK API (Chapter 13) includes commands to control the motors, and to initiate driving actions. The lower level commands, below the action processor are:

- *Drive Wheels*
- *Move Head*
- *Move Lift*
- *Stop All Motors*

The higher level commands, part of the action system are:

- *Drive Straight*
- *Stop All Motors*
- *Turn In Place*
- *Set Head Angle*
- *Set Lift Height*

- *Go to Pose*
- *Turn Towards face*
- *Go To Object*

CHAPTER 25

Animation File Format

The binary animation files are the most significant of Vector's animation files. The file format provides for coordinating the display, motion, and sound responses.

- Animation file format overview
- Structures in the file

112. ANIMATION BINARY FILE FORMAT

The schema and format of the animation binary file is given as a flatbuffer specification. This specification is located at:

`/anki/data/assets/cozmo_resources/config/cozmo_anim.fbs`

Note: this file is not read by any program in Vector. A compatible parser is compiled in.

112.1. OVERVIEW OF THE FILE FORMAT

The animation file contains one or more named animation "clips." Each clip has one or more tracks that represent the scripted motions (and lights & sounds) that Vector should perform. There are tracks for moving Vector's head, lift, driving, modifying his facial expressions, displaying images on the LCD, audio effects, and controlling the backpack lights.

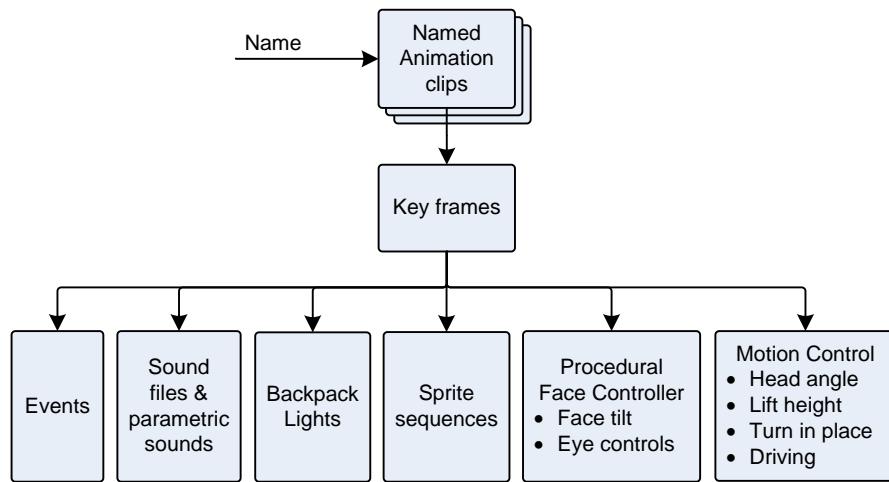


Figure 101: The animation file structure

Each of the tracks within the clip is composed of key frames (with settings for each of the relevant tracks) that are triggered at different points in time.

112.2. RELATIONSHIP WITH COZMO

Vector's file format for animations is derived from the file format used with Cozmo. This presents the possibility of converting Cozmo's animation files to Vector, and vice-versa. The key

differences between Cozmo and Vector's formats are that Vector includes audio tracks, plus some minor extra fields, and fewer backlight LEDs.

The PyCozmo project has the (experimental) ability extract Cozmo's animations, and may be useful for this transcoding.

113. STRUCTURES

The animation file starts with an `AnimClips` structure. Unless specified otherwise, these are the same as in Cozmo.

By default, all fields are optional unless specified otherwise.

113.1. ANIMCLIPS

The `AnimClips` structure is the “root” type for the file. It provides one or more animation “clips” in the file. Each clip has one or more tracks. The structure following fields:

Field	Type	Description
<code>clips</code>	<code>AnimClip[]</code>	An array of animation clips

Table 439: `AnimClips` structure

113.2. ANIMCLIP

The `AnimClip` is a named animation that can be played. This structure has the following fields:

Field	Type	Description
<code>Name</code>	<code>string</code>	The name of the animation clip; this is also called the <i>animation trigger name</i> .
<code>keyFrames</code>	<code>KeyFrames</code>	The key frames for each of the tracks for this animation clip

Table 440: `AnimClip` structure

113.3. AUDIOEVENTGROUP

The `AudioEventGroup` structure is used to randomly select an audio event (and volume), and send it to the audio subsystem. See Chapter 23, section 106.2 *Audio Pipeline* for a description of audio events. This structure has the following fields:

Field	Type	Units	Description
<code>eventIds</code>	<code>uint[]</code>		The audio event IDs, weighted by a probability.
<code>volumes</code>	<code>float[]</code>		The volume to play the selected audio at.
<code>probabilities</code>	<code>float[]</code>	<code>0..1</code>	The probability weight that a given event will be selected.

Table 441: `AudioEventGroup` structure

This structure is new to Vector.

113.4. AUDIOPARAMETER

The AudioParameter structure is used to set one of the sound parameters in the AudioKinetic Wwise subsystem. See Chapter 23, section *106.2 Audio Pipeline* for a description of audio parameters. This structure has the following fields:

Field	Type	Units	Description	
<i>parameterId</i>	uint		The identifier of the parameter to set. Default: 0	Table 442: <i>AudioParameter</i> structure
<i>value</i>	float		The value to set the parameter to. Default: 0	
<i>time_ms</i>	uint	ms	The time at which the parameter should be set. Default: 0	
<i>curveType</i>	ubyte		default: 0	

This structure is new to Vector.

113.5. AUDIOSTATE

The AudioState structure is used to put the audio system into a particular state. See Chapter 23, section *106.2 Audio Pipeline* for a description of audio state. This structure has the following fields:

Field	Type	Description	
<i>stateGroupId</i>	uint	The state group to modify. Default: 0	Table 443: <i>AudioState</i> structure
<i>stateId</i>	uint	The new state to put the group into. Default: 0	

This structure is new to Vector.

113.6. AUDIOSWITCH

The AudioSwitch structure is used to put an audio switch into a particular setting. See Chapter 23, section *106.2 Audio Pipeline* for a description of audio switches. This structure has the following fields:

Field	Type	Description	
<i>switchGroupId</i>	uint	The switch to modify. Default: 0	Table 444: <i>AudioSwitch</i> structure
<i>stateId</i>	uint	The new state to put the switch into. Default: 0	

This structure is new to Vector.

113.7. BACKPACKLIGHTS

The BackpackLights structure is used to animate the LEDs on Vector's back. This structure has the following fields:

Field	Type	Units	Description	Table 445: BackpackLights structure
<i>triggerTime_ms</i>	uint	ms	The time at which the backlights animation should begin.	
<i>durationTime_ms</i>	uint	ms	The duration before a transition to the next backlight setting may begin. During this time the lights should be illuminated with these colors; after this the colors may transition from these to the next colors.	
<i>Front</i>	float[4]	RGBA	Each color is represented as 4 floats (red, green, blue, and alpha), in the range 0..1. Alpha is always 0 (the value is ignored).	
<i>Middle</i>	float[4]	RGBA	Each color is represented as 4 floats (red, green, blue, and alpha), in the range 0..1. Alpha is always 0 (the value is ignored).	
<i>Back</i>	float[4]	RGBA	Each color is represented as 4 floats (red, green, blue, and alpha), in the range 0..1. Alpha is always 0 (the value is ignored).	

see also: Chapter 21 section 27 *Backpack lights control* for a similar JSON structure.

Note: Cozmo's animation structure includes a left and right LED animation.

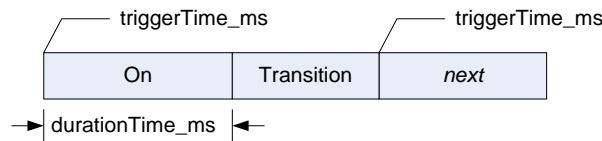


Figure 102: Time course of the backlight colors

The best interpretation is that, once a frame is triggered, the LED is set to the given color. The LED won't be changed for at least *durationTime_ms*. Once that time has expired, the LED color is ramped linearly to the color of the next frame.

113.8. BODYMOTION

The BodyMotion structure is used to specify driving motions for Vector. This structure has the following fields:

Field	Type	Units	Description	Table 446: BodyMotion structure
<i>triggerTime_ms</i>	uint	ms	The time at which the motion should begin	
<i>durationTime_ms</i>	uint	ms	The duration that the robot should drive.	
<i>radius_mm</i>	string	mm		
<i>speed</i>	short		The speed at which the robot should move.	

Note: it is possible that the driving should ramp to the speed in the given duration. This is a TBD.

113.9. EVENT

The Event structure is used to pause the animation at the given time code until the event is received or cancelled. When the event is received, the animation resumes the given time code. This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	When the event occurs it triggers animation to begin at this time.(?) Or, at this time, emit the event?
<i>event_id</i>	string		The name of the event.

Table 447: Event structure

The event names include

- CHANGE_EYE_COLOR
- CUBE_LIGHT_TOGGLE
- DANCE_BEAT_SYNC
- DEAL_CARDS_BEGIN
- FLIP_DOWN_BEGIN
- LISTENING_BEGIN
- STRAIGHT
- SWIPE_CARDS_BEGIN
- TAPPED_BLOCK
- TOGGLE_NUMBERS_DISPLAY
- TURN_IN_PLACE

Note: unless otherwise specified the animations are not allowed to have event key frames – the behavior wouldn't expect to send the events to them.

113.10. FACEANIMATION

The FaceAnimation structure is used to specify the JSON file to animation Vector's display. This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time at which the motion is triggered.
<i>animName</i>	string		The name of the sprite-sequence animation to play. See Chapter 22 section <i>102.4 Sprite Sequences</i> . <i>Required</i>
<i>scanlineOpacity</i>	float		This is new for Vector. Default: 1.0

Table 448:
FaceAnimation structure

The *scanlineOpacity* is new to support Vector's display. With Cozmo “the screen is displayed interlaced, with only every other line displayed. This alternates every time the image is changed (no longer than 30 seconds) to prevent screen burn-in. Therefore to ensure the image looks correct on either scan-line offset we use half the vertical resolution”

Cozmo SDK (Anki)

113.11. HEADANGLE

The HeadAngle structure is used to specify how to move Vector's head. The head should reach the target angle in the duration given, ramping up the movement speed smoothly (with some variability) until it reaches that point. This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time at which the head motion should begin.
<i>durationTime_ms</i>	uint	ms	How long the head motion should last.
<i>angle_deg</i>	ubyte	deg	The angle to move the head to.
<i>angleVariability_deg</i>	ubyte	deg	The amount of randomness allowed for the target head angle. Default: 0

Table 449: HeadAngle structure

113.12. LIFTHEIGHT

The LiftHeight structure is used to specify how to move Vector's lift. The lift should reach the target height in the duration given, ramping up the movement speed smoothly (with some variability) until it reaches that. This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time at which the lift should begin motion.
<i>durationTime_ms</i>	uint	ms	How long the lift motion should last.
<i>height_mm</i>	ubyte	mm	The height to lift the arms to.
<i>heightVariability_mm</i>	ubyte	mm	The amount of randomness allowed for the target height. default: 0

Table 450: LiftHeight structure

113.13. KEYFRAMES

The KeyFrames structure the following fields:

Field	Type	Description	Table 451: KeyFrames structure
<i>LiftHeightKeyFrame</i>	LiftHeight[]	A series of key frames describing when and how the lift should move.	
<i>ProceduralFaceKeyFrame</i>	ProceduralFace[]	A series of key frames describing when and how the eyes should move.	
<i>HeadAngleKeyFrame</i>	HeadAngle[]	A series of key frames describing when and how the head should move.	
<i>RobotAudioKeyFrame</i>	RobotAudio[]		
<i>BackpackLightsKeyFrame</i>	BackpackLights[]	A series of key frames describing when and how the backpack lights should be illuminated.	
<i>FaceAnimationKeyFrame</i>	FaceAnimation[]		
<i>EventKeyFrame</i>	Event[]	Note: many behaviors do not support event key frames; those that do expect a specific event, and number of event keyframes.	
<i>BodyMotionKeyFrame</i>	BodyMotion[]	A series of keyframes to drive and turn the body.	
<i>RecordHeadingKeyFrame</i>	RecordHeading[]	A series of key frames to record the current heading of the robot so that the animation can return to them later.	
<i>TurnToRecordedHeadingKeyFrame</i>	TurnToRecordedHeading[]	A series of key frames use to return the robot to a previously saved heading after a movement.	
<i>SpriteBoxKeyFrame</i>	SpriteBox[]	A series of key frames for the visual sprite box animation. <i>New in version 1.7.</i>	

Note: Each of the structures has a time code. Within each array, the time code(s) must be in ascending order; no two entries in the same array can share the same time code.

113.14. PROCEDURALFACE

The ProceduralFace structure is used squash, stretch and shake Vectors face in cartoonish ways. It does not affect where his eyes are focused. See Chapter 22 section *103 Procedural face* for a description of Vectors face, and how these parameters influence it. The structure has the following fields:

Table 452:
ProceduralFace
structure

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time at which the motion is triggered.
<i>faceAngle</i>	float		default: 0
<i>faceCenterX</i>	float		default: 0
<i>faceCenterY</i>	float		default: 0
<i>faceScaleX</i>	float		default: 1.0
<i>faceScaleY</i>	float		default: 1.0
<i>leftEye</i>	float[]		If present, these describe modifications to the eye – lid, cheeks, and shape of the eye. They have the structure given below.
<i>rightEye</i>	float[]		If present, these describe modifications to the eye – lid, cheeks, and shape of the eye. They have the structure given below.
<i>scanlineOpacity</i>	float		This is new for Vector. default: 1.0

The arrays of floats for each eye in animations for *Cozmo* have been deciphered, and are presumed to be the same for Vector. They are presumed to be the same for Vector:

PyCozmo

Field	Default	Description
<i>lower_inner_radius_x</i>	0.5	
<i>lower_inner_radius_y</i>	0.5	
<i>lower_outer_radius_x</i>	0.5	
<i>lower_outer_radius_y</i>	0.5	
<i>upper_inner_radius_x</i>	0.5	
<i>upper_inner_radius_y</i>	0.5	
<i>upper_outer_radius_x</i>	0.5	
<i>upper_outer_radius_y</i>	0.5	
<i>upper_lid_y</i>	0.0	The vertical position of the upper eye lid (which occludes the eye).
<i>upper_lid_angle</i>	0.0	The angle of the upper eye lid.
<i>upper_lid_bend</i>	0.0	The bend to the upper eye lid.
<i>lower_lid_y</i>	0.0	The vertical position of the lower eye lid / cheek (which occludes the eye).
<i>lower_lid_angle</i>	0.0	The angle of the lower eye lid / cheek.
<i>lower_lid_bend</i>	0.0	The bend to the lower eye lid / cheek.

113.15. RECORDHEADING

The RecordHeading structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time when the robot should record his heading?

Table 453:
RecordHeading structure

113.16. SPRITEBOX

The SpriteBox structure is new to Vector, introduced in version 1.7 of the software. A sprite box defines a rectangular region on the display to draw an image from a file. This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time when Vector should begin to use this sprite box.
<i>spriteBoxName</i>	string		The name of the sprite box. (This name is also defined in <i>vic-anim</i> and <i>libcozmo_engine</i> .) The animation engine may use this to select the procedure(s) in charge managing the layer and sprite boxes. If an image map is available for this animation, the sprite sequence it describes will be displayed in this rectangle. <i>Required</i>
<i>layer</i>	string		The name of the layer. (This name is also defined in <i>vic-anim</i> and <i>libcozmo_engine</i>) The animation engine may use this to select the procedure(s) in charge managing the layer and sprite boxes. <i>Required</i>
<i>assetName</i>	string		<i>Required</i>
<i>renderMethod</i>	string		“CustomHue” if the PNG images should be converted from gray scale to the colour using the current eye colour setting. “RGBA” if ???. <i>Required</i>
<i>spriteSeqEndType</i>	string		<i>Required</i>
<i>alpha</i>	float	%	The opacity of the image pixels. Default is 100.0
<i>x</i>	int	<i>pixels</i>	The x coordinate of the upper left hand corner of the sprite box. The coordinate must be within the display area (0..183). Default: 0
<i>y</i>	int	<i>pixels</i>	The y coordinate of the upper left hand corner of the sprite box. The coordinate must be within the display area (0..95). Default: 0
<i>width</i>	uint	<i>pixels</i>	The width of the sprite box.
<i>height</i>	uint	<i>pixels</i>	The height of the sprite box.

Table 454: SpriteBox structure

The position plus the size of the area must fit within the display area.

$$x+width < 184$$

$$y+height < 96$$

The box coordinates and area should smoothly move and change size to the reach the target position and size by the given trigger time.

See also Chapter 22 section *102.2 Layout file* for another method of defining a sprite box.

113.17. ROBOTAUDIO

The RobotAudio structure is new to Vector; a very different structure with a similar name was used with Cozmo. This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time the audio events should be sent, and the parameters should be set.
<i>eventGroups</i>	AudioEventGroup[]		The set of possible audio events to send.
<i>state</i>	AudioState[]		
<i>switches</i>	AudioSwitch[]		The configuration of the audio context (aka, setting the audio “switches” to use the right sounds and effects for the circumstances).
<i>parameters</i>	AudioParameter[]		A set of changes to the audio playback parameters.

Table 455: RobotAudio structure

113.18. TURNTORECORDEDHEADING

The TurnToRecordedHeading is used to specify how Vector should turn to the previously recorded heading. The robot reach the target heading in the duration given, ramping up the movement speed smoothly until it reaches that position (within some tolerance). This structure has the following fields:

Field	Type	Units	Description
<i>triggerTime_ms</i>	uint	ms	The time when Vector should begin to turn to the recorded heading.
<i>durationTime_ms</i>	uint	ms	The amount of time to move to the recorded heading.
<i>offset_deg</i>	short	deg	default: 0
<i>speed_degPerSec</i>	short	deg/sec	
<i>accel_degPerSec2</i>	short	deg/sec ²	default: 1000
<i>decel_degPerSec2</i>	short	deg/sec ²	default: 1000
<i>tolerance_deg</i>	ushort	deg	This specifies how close the actual heading is to the target before considering the movement complete. Default: 2
<i>numHalfRevs</i>	ushort		default: 0
<i>useShortestDir</i>	bool		default: false

Table 456:
TurnToRecordedHeading structure

PART VI

High Level AI

This part describes items that are Vector's behaviour function.

- BEHAVIOR. A look at Vectors behaviors, and emotions
- EMOTION MODEL. A look at how Vector emulates emotions
- BEHAVIOR TREES. A look at how the behaviors are selected and their settings



Steph Dere
drawing by Steph Dere

[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 26

Behavior

This chapter describes Vector’s action, behaviour, and emotion system:

- Actions and behaviour queues
- The emotion-behaviour system, and stimulation

114. OVERVIEW

How does Vector get excited from praise, and then decide to go exploring and play? How does he decide it’s time to take a nap?

Vector’s high-level AI – his emotions, sense of the environment and himself, and behaviors – are a key part of how he creates a compelling character. He has an emotional state that is seen in his affect – his facial expression, head and arm posture – how he behaves and responds, as well as the actions he initiates.

115. ACTIONS AND BEHAVIORS

Actions and “behaviors represent a complex task which requires Vector’s internal logic to determine how long it will take. This may include combinations of animation, path planning or other functionality.”

Anki SDK

115.1. ACTIONS AND THE ACTION QUEUES

Animations can be submitted with which tracks of the animation to disable. This allows multiple actions can be run at the same time. If the action requires a track that is already in use, the action isn’t run, and returns an error. Actions can automatically retry if a problem was encountered.

Actions can have associated “tag” used to refer to that running instance. The client can cancel the action.

115.2. BEHAVIORS

Unlike actions, only one behavior can be active at a time. The others are waiting in a stack. A behavior is submitted (to be run) with a priority; if its priority is higher priority the current one, it is run instead. The old behavior is pushed down in the stack. When a behavior completes, the next high priority one is resumed.

115.2.1

Priority Levels

The behaviors requested by Vector’s internal AI are submitted to the stack with a priority based on that behavior. If the SDK has requested control, the behaviors it requests are submitted with the priority level set when control was requested. As long as the SDK is connected and has control, behaviors submitted at a lower priority will not activate, even if the SDK is not currently running a behavior. The SDK can lose control if a higher priority is submitted (e.g. “like returning to the charger due to low battery”), gives up control or closes the connection.

The priority levels are organized with lower numbers being higher priority (and larger numbers being lower priority). The built in behaviors have different associate priority levels:

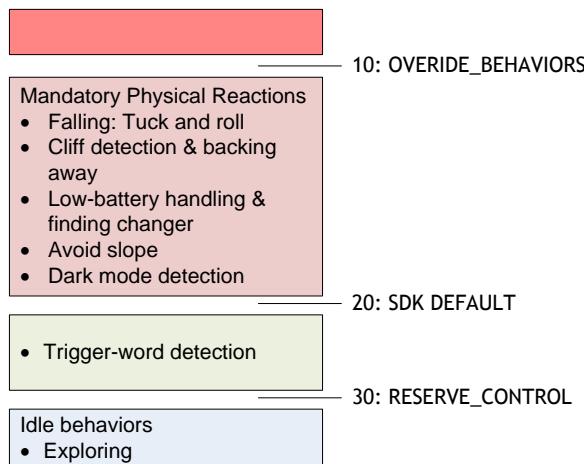


Figure 103: The behaviour priorities

The behaviors are grouped, from the highest priority to the least, into the following categories:

- MandatoryPhysicalReactions
- TriggerWordDetected
- SDKDefault (the behaviors submitted via the SDK if the default priority was used)
- SingletonWallTimeCoordinator
- TimerUtilityCoordinator
- WeatherResponses
- TakeAPhotoCoordinator
- ReactToRobotShaken
- ReactToTouchPetting
- BasicVoiceCommands (“simple voice commands that we want to ignore obstacles”)
- ReactToObstacle
- InterruptingVoiceReactions
- ChangeEyeColor
- ReactToUnclaimedIntent
- HeldInPalmDispatcher
- WhileInAirDispatcher
- ReactToPutDown
- ReactToDarkness
- GreetAfterLongTime
- ReactToUncalibratedHeadAndLift
- DanceToTheBeatCoordinator
- StayOnChargerUntilCharged
- ReactToSoundAwake
- ConfirmHabitat

- HighLevelAI

115.2.2

Other properties of a behavior

Besides a priority, a behavior has other properties:

- They have a string identifier
- A given behavior is an instance of a class
- It can have conditions (usually on the current executing environment) that must be met before the behavior can be activated, and other conditions that they must be met to keep running.
- A behavior can have a cool down period associated with – a period of time after the end of its last use before it can be run again.
- A behavior can trigger animations or actions when it is activated (referred to as the “get in” animations), and when it stops running (the “get out” animations)
- A behavior can submit other behaviors to be run

115.3. PATH PLANNING AND OTHER SMART THINGS TO SUPPORT US

For some commands, “Vector uses path planning, which refers to the problem of navigating the robot from point A to B without collisions. Vector loads known obstacles from his map, creates a path to navigate around those objects, [and] then starts following the path. If a new obstacle is found while following the path, a new plan may be created.”



Anki SDK

“For commands such as go_to_pose, drive_on_charger and dock_with_cube, Vector uses path planning, which refers to the problem of navigating the robot from point A to B without collisions. Vector loads known obstacles from his map, creates a path to navigate around those objects, and then starts following the path. If a new obstacle is found while following the path, a new plan may be created.”

115.4. DECIDING ON THE BEHAVIOR TO USE

A behavior can be initiated in two different ways. The libcozmo engine can on startup or based on internal state or events, choose a behavior to submit to run. The other is that the *behavior tree* can decide which behavior should be submitted to run:

Vector’s behavior follows a hierarchy. “The highest level is what kind of things should the robot be doing right now – Should he be quiet? Should he be engaging? Should he be sleeping? Is his battery super-low, and he needs to recharge?” Different behaviors flow from these high-level states, in response to events and the states of his Emotion Engine.

Captain 2018 quoting Brad Neuman

The behavior tree works by allowing the currently executing behavior to submit other behaviors behaviors to run; but those behaviors can have sophisticated rules (and priorities) that govern whether can run, or should stop running. The details of the behavior tree will be examined in the next chapter.

115.5. INITIATING THE BEHAVIOR

In both cases, the identifier (a text string) of the behavior is passed to the behavior engine, along with a priority to run at. The engine checks to see if this is a lower priority (higher number) than the current priority level. If so, the behavior is rejected. . The engine also checks that the

behavior is not already on the stack; if so, the behavior is rejected. Otherwise, the behavior id is used to look up (in a table) the relevant behavior node:

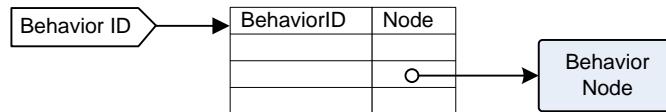


Figure 104: Mapping a behavior identifier to the behavior tree node

An instance of the behavior is created from the behavior node – the node specifies the class, and its configuration, but the state is not preserved between uses. Then:

1. The behavior is given a preliminary check: can the behavior run?
 - a. Is the behavior still in cool down period? If so, the behavior is rejected
 - b. A behavior node can have optional conditions attached to it that say whether or not it can run. Have these conditions been met? If not, the behavior is rejected
2. Next, the active behavior is suspended
3. Update the stats for behavior activation
4. The new behavior is pushed on to the stack
5. The behaviors associated AI Feature is tracked, to aid in debugging and statistics of usage
6. If the behavior has an emotion event (emotionEventOnActivated) associated, it is posted to the Mood Manager to update Vector's emotional state.
7. The behaviors "get in" activities are carried out – those that are done when it is activated. This often includes animations.

115.6. MANAGING THE ACTIVE AND PAUSED BEHAVIORS

The behavior engine regularly checks the behaviors in the stack. It checks that the top most behavior can still run; if not, the behavior is cancelled: its "get out" animation is started, and the behavior is removed from the stack. Perhaps all behaviors on the stack are checked to see if they can still run, and (if not) they and their children are removed from the stack; with the suspended behaviors not running their "get out" animation.

Then the top most behavior carries out any updates to its activities and state. The behavior may also choose to cancel itself, or to initiate another behavior.

If a behavior exits – or is cancelled – the next one on the stack is resumed.

CHAPTER 27

Emotion Model

This chapter describes Vector's action, behaviour, and emotion system:

- Actions and behaviour queues
- The emotion-behaviour system, and stimulation

116. OVERVIEW

How does Vector get excited from praise? Vector has an emotional state that is seen in his affect – his facial expression, head and arm posture – how he behaves and responds, as well as the actions he initiates.

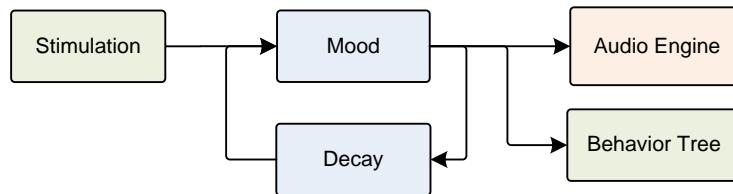


Figure 105: The functional flow of the mood

Vectors mood is affected by external stimulation, and his feedback on his successes (and failures) in his activities. His current mode affects the choices he makes and the behaviors he takes, including those in response to events and stimulation. His emotional state is also reflected in how the audio engine modulates its effects, even potentially choosing other effects or sounds. Vectors' emotions are transitory though: heightened emotions decay, based on the stimulus and behavior that drove them.

This emotion model and coupling their effects with other systems is managed by the “MoodManager.”

117. EMOTIONS, AND STIMULATION

It's thru stimulation of these emotions that Vector responds to praise. The label “emotion” shouldn't be taken too seriously, as it doesn't model psychological moods, and other concepts. It does just enough to convey character.



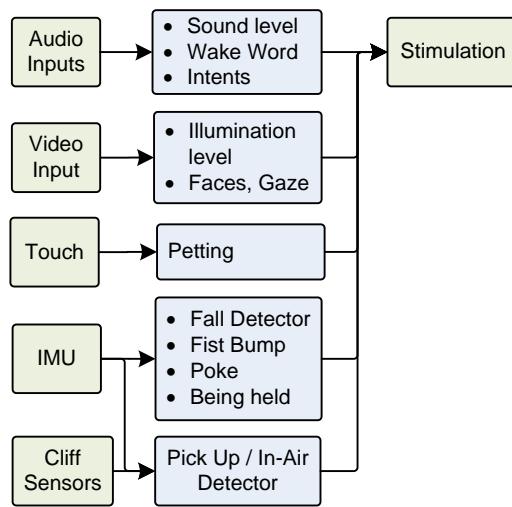
117.1. STIMULATION

Vector uses a concept of a stimulation level to guide how much he should initiate

“When stimulation is low, the robot is chill,”.. Vector is studiously observing but not acting out. “Then if you start making noise, or make eye contact with the robot, and certainly if you say ‘Hey Vector,’ that spikes [stimulation] way up...” But Vector also picks up subtler actions–peripheral movement and noises, for instance, or the room lights turning on and off. “If he gets stimulated enough, he’ll drive off his charger and start to socialize with you, ... say your name, greet you, give you a fist-bump, potentially.”

Captain 2018 quoting
Brad Neuman

Figure 106: The stimulation from sensations



117.2. THE EMOTION MODEL

Stimulation is just one of the dimensions in Vector's emotional model. Some dimensions are influenced by the kind stimulation he is receiving, but others are from internal feedback Vectors behaviors. Altogether he has five dimensions to his emotional state.

- Stimulated (or the stimulation level) is from those sensory experiences described earlier;
- Social, or “how eager [he] is to interact with users generally.” “Hearing his name stimulates Vector, for instance, but it also makes him more social.”
- Confident: “Vector’s confidence is affected by his success in the real world. The hooks on his arms sometimes don’t line up with those on his cube, for instance, and he can’t pick it up. Sometime he gets stuck while driving around. These failures make him feel less confident, while successes make him more confident and more happy.”
- Happy. This is Vectors sense that, overall, things are going well.
- Trust⁵¹

Wolford et al, 2018
Captain 2018

Overall, Vector possesses just enough dimensions/aspects to his emotion model to drive responses and his goal-driven behaviour, giving him a personality. When more dimensions are used, it is harder to get them right, and the less convincing the character is.

117.3. INTERACTION WITH THE BEHAVIOR ENGINE

The behavior engine receives the stimulation events, and using a behavior tree, posts *emotion event* to the mood manager. (The engine may base its decision what to post on the current mood.) An emotion event is just an identifier string. The mood manager maps the emotion event to an *emotion affector*. This is the emotion dimensions impacted by the event, values describing *how much* the event impacts those emotion dimensions, and a decay graph that describes how the heightened emotion fades away toward a neutral state.

emotion event

⁵¹ Trust was added in version 1.6. Vector initially only had the first four. Cozmo had nine, so it seems plausible that Vector would have developed more dimensions over time.

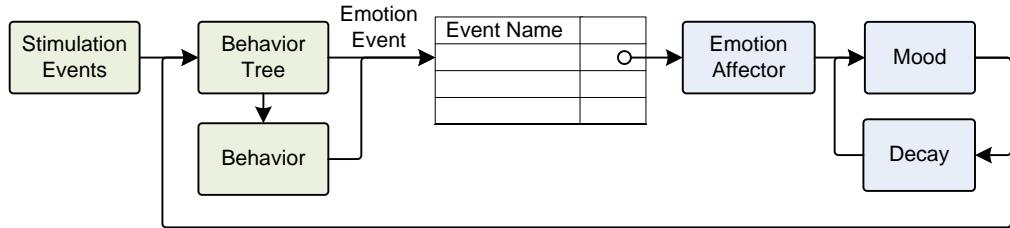


Figure 107: Mapping an emotion event to how it impacts the emotion model

The active behaviors (which are also selected by the behavior tree) may post emotion events to the mood manager as well.

117.4. MOOD MANAGER CONFIGURATION

At start up, the mood manager scans the configuration files building a table mapping the emotion event names to the emotion affector.

The configuration files for the mood manager are located in a folder at:

/anki/data/assets/cozmo_resources/**config/engine/emotionevents**

This is path hardcoded into libcozmo_engine. It is a folder that contains a set of JSON files, all with the same structure. Each of these files is loaded. The files is a structure containing the following fields:

Field	Type	Description
<i>emotionEvents</i>	EmotionEvent[]	An array emotion event structures (see below).

Table 457:
CheckUpdateStatusResponse JSON structure

Field	Type	Description
<i>name</i>	string	The name of the event (see appendix J, <i>Table 531: The emotion event names</i>)
<i>emotionAffectors</i>	EmotionAffectuator[]	The impact on the emotion state
<i>repetitionPenalty</i>	RepetitionPenalty	This is a “time ratio” describing how the value decays. <i>Optional</i> .

Table 458:
EmotionEvent JSON structure

Field	Type	Description
<i>emotionType</i>	string	The dimension or type of emotion (“Happy”, “Confident”, “Stimulated”, “Social”, or “Trust”)
<i>value</i>	float	The value to add to the emotional state. The range is -1 to 1

Table 459:
EmotionAffectuator JSON structure

Altogether, the files respond to the following “emotion event” names. Some are external stimuli, some are events in general, some are events regarding whether or not a behaviour succeed, or failed (failed with retry, failed with abort)

117.4.1

The RepetitionPenalty

The RepetitionPenalty structure contains the following fields:

Field	Type	Description
nodes	XY[]	This is a “time ratio” describing how the value decays.

Table 460:
RepetitionPenalty
structure

117.4.2

The XY decay graph

The XY structure is used to define how a value (often the value associated with an emotion dimension) should decay with time. This structure contains the following fields:

Field	Type	Description
X	float	With time graphs, this is “the time in seconds since the most recent event (which changed the emotion by more than some delta)”. With value slopes, this is “the emotion value”
Y	float	With time graphs, this is “the ratio of the original value that should be reached by the given time” With value slopes, this is “the amount it decays (towards zero) per <i>minute</i> as a fixed amount (not a ratio).” The value never goes below zero.

Table 461: XY structure

117.5. MOOD CONFIGURATION

A mood configuration files is located at:

/anki/data/assets/cozmo_resources/[config/engine/mood_config.json](#)

This is path hardcoded into libcozmo_engine.

The file is a structure containing the following fields:

Field	Type	Description
actionResultEmotionEvents	TBD struct	
audioParameterMap	struct	This is a structure that maps an emotion type to an audio parameter’s string name. The audio parameter is set to current emotion type’s value. <i>Optional</i>
decayGraphs	DecayGraph[]	This describes how an emotion decays.
defaultRepetitionPenalty	RepetitionPenalty	This is a “time ratio” describing how the value decays. <i>Optional</i> .
eventMapper	struct	
simpleMoodAudioParameters	struct	<i>Optional</i>
valueRanges	ValueRange[]	The value ranges for given emotion types.

Table 462: Mood config JSON structure

The DecayGraph structure contains the following fields:

Field	Type	Description
<i>emotionType</i>	string	The dimension or type of emotion (“Happy”, “Confident”, “Stimulated”, “Social”, or “Trust”). “default” Matches
<i>graphType</i>	string	“TimeRatio” or “ValueSlope”
<i>nodes</i>	XY[]	Array of structures describing how the value decays with time.

Table 463: DecayGraph structure

The ValueRange structure contains the following fields:

Field	Type	Description
<i>emotionType</i>	string	The dimension or type of emotion (“Happy”, “Confident”, “Stimulated”, “Social”, or “Trust”)
<i>max</i>	float	The maximum value for the emotion type
<i>min</i>	float	The minimum value for the emotion type

Table 464: ValueRange structure

118. REFERENCES & RESOURCES

Captain, Sean; *Can emotional AI make Anki’s new robot into a lovable companion?* , Fast Company, 2018-8-8
<https://www.fastcompany.com/90179055/can-emotional-ai-make-ankis-new-robot-into-a-lovable-companion>

CHAPTER 28

Behavior Tree

This chapter describes Vector's behaviour tree and how behaviors are configured:

- Behavior trees, parameters for behaviors, conditions that allow a behavior or stop a behavior
- Cube spinner event mapping.

119. OVERVIEW

Behaviors are why Vector wants to shove stuff off of the desk.

Vector employs a *behavior tree* that decides if a behavior can run or can no longer run. It doesn't take it to the extreme a detailed decision tree scripting every action and response. Most of the behavior tree is focused on ensuring that transition between behaviors isn't too abrupt, and provides the settings (or preferences) for the behavior.

The fields and structures of the behavior tree are pretty ad hoc though. This seems to be the norm in the video game industry

The “design principles” listed in this paper are a rather transparent attempt to impose a structure on what might otherwise appear to be a random grab-bag of ideas – interesting, perhaps, in and of themselves but not terribly cohesive as a whole. Isla 2005

120. BEHAVIOR TREE

The behavior tree is composed of nodes in JSON files. Each of the behavior nodes has a unique identifier, called `behaviorID`. This is a way to make the records unique so that they can be looked up. It is unlike that it links to any special code or modules within `libcozmo_engine`.

The nodes also have a field – `behaviorClass` – that says how to interpret the node parameters, if the behavior is activated. This class name links to code/modules within `libcozmo_engine`. There are 86 different behaviour classes.

Behavior nodes can initiate other behaviors. The identity of the behavior they launch may be called out in the configuration of the node, or be hardcoded internally. To prevent loops, the chain of the nodes must be *acyclic*. The concern is that a behavior node kicks off another (and so on), eventually to a child node initiate another copy of the first node, leading to an infinite loop of behaviors being started on pushed onto the stack. Not only doesn't it give expected results, eventually the software will run out of memory, and crash.

`libcozmo_engine` kicks off the initial behavior that forms the root of the tree. Vector, at the top level, has 7 broad states:

- PR demo
- Factory test (e.g. the playpen tests)
- Acoustic testing
- On-boarding

- Normal
- Developer
- Post on-boarding

These states are mapped to initial behavior identifier. Some have the mapping built-in to the software (hardcode), the others this mapping is in the above JSON configuration file (in the `victor_behavior_config.json` file; more on this below). In normal operation, this is the “InitNormalOperation” behavior.

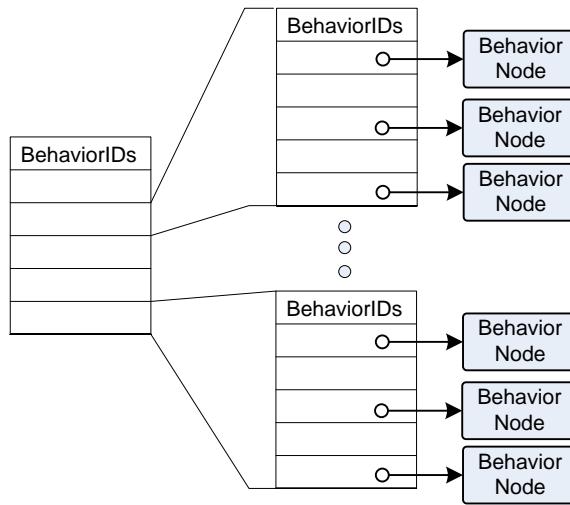


Figure 108: The behavior tree fan out

It is built on the `DispatcherStrictPriority` behavior class, listing behaviors to sequentially check to see if they can be run. The top node only refers to behaviors that in a list of behaviors to invoke sequentially. The behaviors listed at that second level in turn reference behaviors that carry out actual AI features.

The decision tree logic is called out with the nodes. There is a portion of the logic that is used to check to see if the behavior *can be run*. This logic can be used to delay running the behavior until some clean or stabilization of other stuff has occurred. And there is a portion of the logic that is used to check to see if the behavior should be cancelled.

120.1. TIMERS

Behaviors can have an associated timer, similar to animations can have a cool down timer. This prevents the behavior from re-engaging too quickly. These timers can be used as part of the conditional rules that enable or disable a behavior.

Timer	Description
<i>FistBump</i>	
<i>ObservingOffChager</i>	
<i>ObservingOnCharger</i>	
<i>ReactToIllumination</i>	
<i>ReactToJoltInPalm</i>	

Table 465: Behavior timers

The timers are handled by `BehaviorTimerManager()`.

120.2. CONFIGURATION

The configuration files for the behavior tree are located:

```
/anki/data/assets/cozmo_resources/config/engine/behaviorComponent/victor_behavior_
config.json
```

This is path hardcoded into libcozmo_engine.

Note: most of the names of the structures in this chapter are arbitrary. They were made up to ease readability and documentation. The files do not reference any such structure names.

120.3. BEHAVIOR NODE

The following fields are common to all behavior nodes:

Table 466: Behavior JSON structure

Field	Type	Description
<i>anonymousBehaviors</i>	Behavior[]	A list of behaviors that are executed. <i>Optional</i>
<i>associatedActiveFeature</i>	string	Note: this is the high level AI feature, not the feature gate. <i>Optional</i>
<i>behaviors</i>	string[]	Array of behavior names, in order; these behaviorName are in the anonymousBehaviors array. Can also be in...? <i>Optional</i>
<i>behaviorClass</i>	string	Often these are the same
<i>behaviorID</i>	string	
<i>behaviorName</i>		
<i>behaviors</i>	string	If it is a string, this is the behaviorID of the behavior to run.
	string[]	If it is an array of strings, this is one or more behaviors to run in sequence. That is, the first behavior is initiated, and the current behavior node is paused until the new completes. Then the next one in the sequence is initiated, and so on in order.
	BehaviorConfig[]	If it is an array of structures, this is a set of possible behaviors to run; they are picked randomly based on weight.
<i>behaviorStatToIncrement</i>	string	<i>Optional</i>
<i>delegateID</i>	string	
<i>driveOffChargerBehavior</i>		
<i>emotionEventOnActivated</i>	string	e.g. RespondToShortVoiceCommand or DanceToTheBeat. <i>Optional</i> .
<i>faceSelectionPenalties</i>		
<i>getInAnimation</i>	string	The animation trigger name of the animation to play when starting the behavior. <i>Optional</i>
<i>getOutAnimation</i>	string	The animation trigger name of the animation to play when exiting the behavior. <i>Optional</i>
<i>headAngle_deg</i>	float	<i>Optional</i>

<i>postBehaviorSuggestion</i>	string	<i>Optional</i>
<i>resetTimers</i>		
<i>respondToUserIntents</i>	TBD[]	<i>Optional</i> . Only key is “type” (with a value such greeting_goodbye) The intent string is the <i>User Intent</i> (see table in appendix)
<i>wantsToBeActivatedCondition</i>	Condition	If the condition is false, the rest of the behavior is skipped. Some of the conditions are used to wait until the robot is in a state that he can carry out the behavior (and past stuff that could cause false triggers are behind us) <i>Optional</i>
<i>wantsToCancelSelfCondition</i>	Condition	If the condition is true, the rest of the behavior is skipped. <i>Optional</i>

BEHAVIORCONFIG

The BehaviorConfig structure has the following fields:

Field	Type	Units	Description
<i>behavior</i>	string		The name of a behaviorID. The anonymous behaviors in the current behavior node are checked first to find a behavior node with this id. Then the global table.
<i>cooldown_s</i>	float	seconds	The amount of time after this behaviour completes before it can be run again
<i>weight</i>	float		<i>Optional</i>

Note: the weights do not have to sum to 1.0

Given an array of BehaviorConfig structures, the list is prescreened to eliminate behaviors that already running or still in cooldown. A behavior is randomly selected from this list based on its weighting, and launched.

Table 467:
BehaviorConfig
parameters

120.4. CONDITION NODES

A condition node is used as part of the behavior tree to determine if a behavior is eligible, or if a running behavior should be cancelled. The interpretation of the condition is (mostly) controlled by the *conditionType* field. This type defines what other fields will be looked at.

The following are the kinds of condition nodes:

conditionType	Description
<i>AlexaInteractionActive</i>	This condition is true if Vector’s Alexa modules are currently interacting with a person.
<i>BatteryLevel</i>	This type of condition compares the current battery state with a specified state. Although any of the states can be used, the current behavior tree nodes only check for low battery.
<i>BeatDetected</i>	This condition is true if a music beat has been detected
<i>BeingHeld</i>	This condition is true if Vector is being held.
<i>BehaviorTimer</i>	This condition is true if

Table 468: Types of
condition nodes

<i>CliffDetected</i>	This type of condition is true if a cliff sensor has detected an edge.
<i>Compound</i>	This type of condition is used for boolean logic. It is the only kind that recurses to other compound structures.
<i>Emotion</i>	This condition is true if the value for the given emotion dimension is above a threshold.
<i>FeatureGate</i>	This type of condition is true if a specified AI feature is active or inactive.
<i>HighTemperature</i>	This type of condition is true if Vector's temperature is above an acceptable limit.
<i>InCalmMode</i>	This type of condition is true if Vector is in a calm emotional state.
<i>IsMaintenanceReboot</i>	This type of condition is true if Vector has rebooted for a software update or other maintenance reasons.
<i>IsNightTime</i>	This type of condition is true if Vector [thinks it is night? the illumination level is dark?]
<i>MotionDetected</i>	This type of condition is true if Vector sees some motion in his peripheral vision.
<i>ObjectInitialDetection</i>	This condition is true if
<i>ObjectKnown</i>	This condition is true if
<i>OffTreadsState</i>	This type of condition is true if Vector's current state matches a conditions such as on his tread, or has been picked up, is being held, is being put back down, has fallen (on his side). There is a time component to ensure that state is stable, to prevent overreacting.
<i>ProxInRange</i>	
<i>RobotPickedUp</i>	This type of condition is true if Vector is picked up – being held.. (does “in the air” count?)
<i>RobotPitchInRange</i>	This type of condition is true if Vector's pitch is within a specified range.
<i>RobotRollInRange</i>	This type of condition is true if Vector's roll is within a specified range.
<i>RobotShaken</i>	This type of condition is true if Vector has been shaken.
<i>RobotTouched</i>	This type of condition is true if Vector has been touched/petted (for at least a minimum duration).
<i>StuckOnEdge</i>	This type of condition is true if at least one tread has gone over the edge. Perhaps the cliff sensors on one side show in open space, perhaps Vector can tell that the tread isn't moving the robot?
<i>TimedDedup</i>	This condition is true if
<i>TimerInRange</i>	This condition is true if a specified timer is within a specified value range.

The Condition structure has the following possible fields:

Field	Type	Condition Type	Description	Table 469: Condition JSON structure
<i>and</i>	Condition[]	<i>Compound</i>	<i>Optional</i>	
<i>allowPotentialBeat</i>	boolean	<i>BeatDetected</i>		
<i>begin_s</i>	float	<i>TimerInRange</i>	Wait for the timer to have been going for at least this number of seconds.	
<i>conditionType</i>	string		One of the conditions listed in <i>Table 468: Types of condition nodes</i>	
<i>cooldown_s</i>	float	<i>BehaviorTimer</i>	The minimum duration between behaviors.	
<i>dedupInterval_ms</i>	int	<i>TimedDedup</i>		
<i>emotion</i>	string	<i>Emotion</i>	The name of the emotion dimension to fetch the value for.	
<i>expected</i>	boolean	<i>FeatureGate</i>		
<i>feature</i>	string	<i>FeatureGate</i>	“HowOldAreYou”	
<i>firstTimeOnly</i>		<i>ObjectInitialDetection</i>		
<i>invalidSensorReturn</i>		<i>ProxInRange</i>	If true, matches when the proximal sensor is unable to measure the distance to the object. Default is false. <i>Optional</i>	
<i>maxAge_ms</i>	int	<i>ObjectKnown</i>		
<i>maxPitchThreshold_deg</i>		<i>RobotPitchInRange</i>	The maximum acceptable pitch reported by the IMU.	
<i>maxProxDist_mm</i>		<i>ProxInRange</i>	The maximum acceptable distance reported by the time of flight sensor.	
<i>maxRollThreshold_deg</i>		<i>RobotRollInRange</i>	The maximum acceptable roll reported by the IMU.	
<i>maxTimeSinceChange_ms</i>	int	<i>OffTreadsState</i>	<i>Optional</i>	
<i>min</i>		<i>Emotion</i>	The minimum acceptable value associated with the given emotion dimension.	
<i>minAccelMagnitudeThreshold</i>		<i>RobotShaken</i>		
<i>minDuration_ms</i>		<i>CliffDetected</i>		
<i>minPitchThreshold_deg</i>		<i>RobotPitchInRange</i>	The minimum acceptable pitch reported by the IMU.	
<i>minProxDist_mm</i>		<i>ProxInRange</i>	The minimum acceptable distance reported by the time of flight sensor.	
<i>minRollThreshold_deg</i>		<i>RobotRollInRange</i>	The minimum acceptable roll reported by the IMU.	
<i>minTimeSinceChange_ms</i>	int	<i>OffTreadsState</i>	<i>Optional</i>	
<i>minTouchTime</i>	int	<i>RobotTouched</i>		
<i>motionArea</i>	string	<i>MotionDetected</i>	“Left”, “Right”, (potentially “Ground”)	
<i>motionLevel</i>	string	<i>MotionDetected</i>	“High”	
<i>not</i>	Condition	<i>Compound</i>	<i>Optional</i>	
<i>numCliffDetectionsToTrigger</i>		<i>CliffDetected</i>		

<i>or</i>	Condition[]	<i>Compound</i>	<i>Optional</i>
<i>shouldBeHeld</i>	boolean	<i>BeingHeld</i>	
<i>shouldDetectNoCliffs</i>	boolean	<i>CliffDetected</i>	
<i>subCondition</i>	Condition	<i>TimedDedup</i>	
<i>targetBatteryLevel</i>	string	<i>BatteryLevel</i>	e.g. “Low”
<i>targetState</i>	string	<i>OffTreadsState</i>	“OnTreads”
<i>timerName</i>	string	<i>BehaviorTimer</i>	“ReactToIllumination”

121. A LOOK AT SOME INTERESTING BEHAVIORS

There are too many behavior classes to dig into. But a few are particularly fun to look at.

121.1. SHOVING STUFF OFF OF THE TABLE

The BumpObject class is likely the behavior that drives Vector to shove stuff off of desk. There is only one behavior tree node with this class. It has the id ExploringBumpObject and is held in the file

`/anki/data/assets/cozmo_resources/config/engine/behaviorComponent/behaviors/victorBehaviorTree/highLevelDelegates/exploring/exploringBumpObject.json`

The BumpObject class takes the following extra parameters:

Field	Type	Units	Description	Table 470: BumpObject behavior structure
<i>maxDist_mm</i>	uint	mm	The maximum distance to the potential object to bump.	
<i>minDist_mm</i>	uint	mm	The minimum distance to the potential object to bump	
<i>pBumpWhenEvil</i>	float		Probability of bumping things when being evil?	
<i>pBumpWhenNotEvil</i>	float		Probability of bumping things when not being evil?	
<i>pEvil</i>	float		Probability of being evil?	

121.2. POUNCING

Pouncing is where Vector springs forward to leap on an object, such as a finger.



- Vector detects visual motion, and turns toward that (see motion detection). In this he turns left (or right) to where he detected the motion (the Turn behavior class)
- When he has a distance measurement (from the proximity sensor) (The PounceWithProx behavior class)
- When he is close enough, the animation takes over; he'll make his facial expressions, moves his arms, and tries to pin the object with his arms ("mousetrap"). Note: the animations can't be used to drive toward the target earlier; they aren't linked into the proximity sensors for driving.
- If nothing else is happening, he'll wait for up to 30 seconds before losing interest.

A behavior tree node, using the DispatcherStrictPriority behavior class, coordinates these. The DispatcherStrictPriority class takes the following extra parameters:

Field	Type	Units	Description	Table 471: DispatcherStrictPriority behavior structure
<i>interruptActiveBehavior</i>	boolean			

The Turn class takes the following extra parameters:

Field	Type	Units	Description	Table 472: Turn behavior structure
<i>shouldTurnClockwise</i>	boolean		True if Vector should turn clockwise; false if he should turn counter clockwise	
<i>turnDegrees</i>	int		The number of degrees to turn.	

121.3. REACTING TO SOUND

Vector has two related behaviors for reacting to sound – deciding that there is some activity and that he should react, or even play.

The ReactToSound behavior class is used to rouse Vector and respond if there are any sudden noises, or there sounds like activity in the room:

Table 473:
ReactToSound
parameters

Field	Type	Units	Description
<i>micAbsolutePowerThreshold</i>	float		“a mic power above this will always be considered a valid reaction sound” 0...4?
<i>micConfidenceThresholdAtMinPower</i>	float		Used in conjunction with <i>micMinPower?</i> 0...5000
<i>micDirectionReactionBehavior</i>	string		The behavior ID to use for reactions
<i>micMinPowerThreshold</i>	float		“a mic power above this will require a confidence of at least <i>kRTS_ConfidenceThresholdAtMinPower</i> to be considered a valid reaction sound” 0..3 ? 999.9 is considered impossibly high”

The *ReactToMicDirection* behavior class is used to allow Vector to respond to direction that the sound is coming from. It maps the sound direction to the terms “TwelveOClock” “OneOClock”, and has conditions like “OnSurface” and “OnCharger”

See Chapter16,section 73.2 *Spatial audio processing* for where it the microphone sound is coming from.

121.4. DANCING

The dancing can be initiated two different ways. The first step is if a beat is detected. The second is if Vector is verbally told to dance.

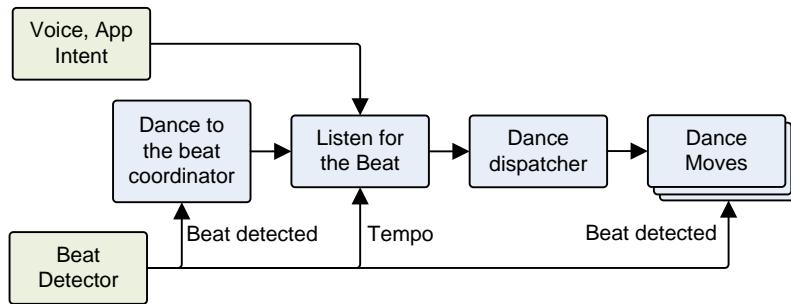


Figure 109: Flow of the detecting and dancing to music

The details of the beat detector and tempo measurement are in Chapter 16 section *73.5 Beat Detection*.

121.4.1

Dancing if a beat is detected

A behavior node (of behavior class “`DanceToTheBeatCoordinator`”) is regularly invoked as part of the behavior tree (see section *120 Behavior Tree*). This node has the pre-condition that a beat has been detected. This isn’t quite the same as reacting sounds, but it is similar. The `BeatDetected` condition structure has the following parameters:

Field	Type	Units	Description
<code>allowPotentialBeat</code>	boolean		Default: false. <i>Optional</i>

Table 474:
`BeatDetected`
parameters

If a beat has been heard, the `DanceToTheBeatCoordinator` proceeds in two phases. The first kicks off a helper behavior to listen for music. If it detects music (beats), it then fires off a dance behavior: there are two such behaviors, depending on whether or not it was on the charger. If there is no music detected – or Vector is no longer on his treads – this behavior exits.

The behavior’s configuration structure has the following parameter fields:

Field	Type	Units	Description
<code>listeningBehavior</code>	string	<code>behaviorID</code>	The name of a behavior node to invoke.
<code>longListeningBehavior</code>	string	<code>behaviorID</code>	The name of a behavior node to invoke.
<code>offChargerDancingBehavior</code>	string	<code>behaviorID</code>	The name of a behavior node to invoke.
<code>onChargerDancingBehavior</code>	string	<code>behaviorID</code>	The name of a behavior node to invoke.

Table 475:
`DanceToTheBeatCoordinator`
parameters

121.4.2

Dancing by voice command

A behavior node (of behavior class “`DanceToTheBeatVoiceCommand`”) is regularly invoked as part of the behavior tree (see section *120 Behavior Tree*). Part of pre-conditions for being able to execute this node is that someone has given Vector a command to dance. This is done by the “`respondToUserIntents`” condition including the “`imperative_dance`” intent. If there is no such user intent pending, the node is skipped.

Otherwise, the intent is dequeued, Vector drives off of the charger, listen for the music beats, and (if there are any) then begins dancing.

121.4.3

Listening for the beat

The ListenForBeats behavior class is used to listen for multiple beats and to get the tempo. The behavior exits once music has been heard, or if there is a timeout. Then the behavior node that invoked it is then resumed to initiate the next step.

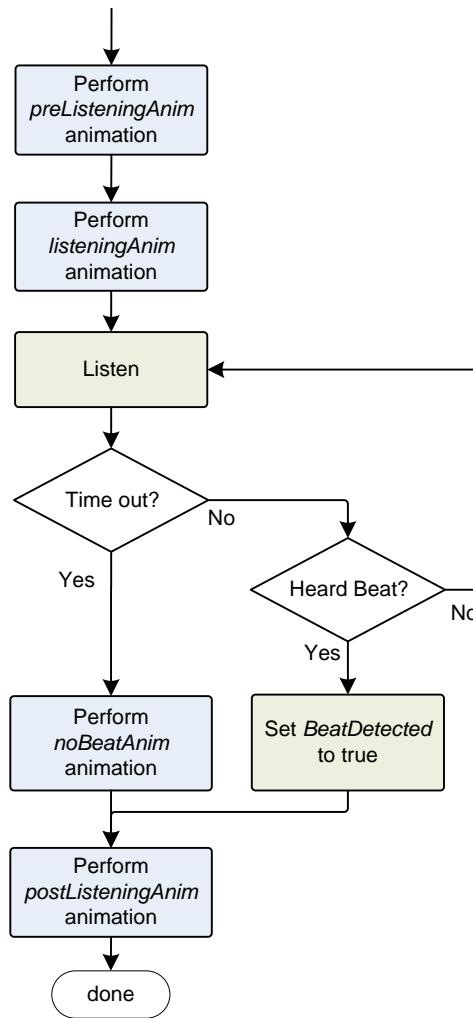


Figure 110:
ListenForBeats behavior
function

The behavior plays animations when it begins, ends, and if it doesn't hear any beats. (If it does hear beats, the dancing behaviors will play their own animations.) It sets the behavior tree variable "BeatDetected" to true if it heard beats; otherwise it is set to false.

The behavior's configuration structure has the following parameter fields:

Field	Type	Units	Description	Table 476: ListenForBeats parameters
<i>cancelSelfIfBeatLost</i>	boolean		If true, exits the behavior when the beat has been lost,	
<i>listeningAnim</i>	string		The name of an animation trigger that is played while listening for music and getting the tempo.	
<i>maxListeningTime_sec</i>	float	seconds	The maximum amount of time to listen for music	
<i>minListeningTime_sec</i>	float	seconds	Listen for at least this amount before concluding that there is no music and exiting.	
<i>noBeatAnim</i>	string		The name of an animation trigger that is played when the behavior exits because there is no music playing.	
<i>postListeningAnim</i>	string		The name of an animation trigger that is played after music has been detected, and is transitioning to the good stuff	
<i>preListeningAnim</i>	string		The name of an animation trigger that is played when this behavior is started, before listening for music has fully started.	

The dance feature employs multiple instances of these to detect beats.

121.4.4

The Robo-boogie

Once Vector has decided to his groove on, he chooses dance from the many kinds of dances of that he knows about. The dances themselves are partly-randomized sequences of dance move animations that are coordinated the beats of the music.



The selection of the dance is performed using a node with the DispatcherRandom behavior class. The different dances (as behaviors nodes) are listed in the “behaviors” array (along with some weighting to help randomize which dance is selected). A new behavior – one that performs the dancing – is randomly selected from this.

A particular dance is an instance of the DanceToTheBeat behavior class. A dance includes the dance moves, whether the back pack lights can play along, and the facial expression.

Field	Type	Description	Table 477: DanceToTheBeat parameters
<i>backpackAnim</i>	string	The backpack animation trigger name to	
<i>danceSessions</i>	DanceSession[]	A dance moves that make up the dance.	
<i>eyeHoldAnim</i>	string	The animation trigger name?	
<i>getOutAnim</i>	string	The animation trigger name to play when	
<i>useBackpackLights</i>	boolean	If true, play the backpack lights animation. Default is false(?)	

The dance moves – called *dance sessions* – are made up of smaller animated movements called “dance phrases.” They can (optionally) be coordinated with the beat of the music. These are captured in the DanceSession structure:

Field	Type	Description
<i>canListenForBeats</i>	boolean	If true, then the animation (in the dance phrases) will be synchronized with beats with the beat. If false, then the beat events from the beat-detector will be ignored.
<i>dancePhrases</i>	DancePhraseConfig []	The sequence of (randomized) animations. These are played in order.
<i>playGetoutIfInterrupted</i>	boolean	If true, and is interrupted by another animation, it plays the animation specified by <i>getOutAnim</i> (in the containing structure).

Table 478:
DanceSession
parameters

The dancing motions– dance phrases – are “made up of one or more possible dance animations … strung together and played on sequential musical beats.” The DancePhraseConfig structure “specifies the rules by which dance phrases are generated when the behavior is run.” These differ from animation groups: here a random *list* of animations to play is created, rather than selecting just one. This structure has the following fields:

Field	Type	Description
<i>anim</i> s	string[]	The list of animation names (rather than trigger names) to randomly draw from. There must be at least one animation given.
<i>maxBeats</i>	uint	The number of times the animation is played no more than this number.
<i>minBeats</i>	uint	The number of times the animation is played at least this number.
<i>multipleOf</i>	uint	The number of times the animation is played is a multiple of this number.

Table 479:
DancePhraseConfig
parameters

“The number of animations that make up the phrase is random, but is always between ‘minBeats’ and ‘maxBeats’, and is always a multiple of ‘multipleOf’.” The “animations are randomly drawn from the [anim] list in accordance with the min/max beats.”

If *canListenForBeats* (in the containing structure) is true, the animation (may) have an event key frame that pauses the animation until a musical beat is heard and a DANCE_BEAT_SYNC event is sent to the animation engine. In this case, the animations must have one event key frame, and the *event_id* must be “DANCE_BEAT_SYNC”.

122. USER CONDITIONS

There is an unusual configuration file that looks like it was intended to allow some user defined behaviors, tuning of behaviours or responses:

/anki/data/assets/cozmo_resources/config/engine/userDefinedBehaviorTree/conditionTo
BehaviorMaps.json

This is path hardcoded into libcozmo_engine. It is a configuration that links to a bunch of backpack and cube lights patterns.

123. REFERENCES & RESOURCES

Isla, Damian; *Handling Complexity in the Halo 2 AI*, GDC 2005 Proceeding, 2005 March 11
https://www.gamasutra.com/view/feature/130663/gdc_2005_proceeding_handling_.php

It is said that this presentation kicked off the widespread use of behavior trees in video games.

[This page is intentionally left blank for purposes of double-sided printing]

PART VII

Maintenance

This part describes practical items to support Vector's operation.

- SETTINGS, PREFERENCES, FEATURES AND STATISTICS. A look at how Vector syncs with remote servers
- SOFTWARE UPDATES. How Vector's software updates are applied.
- DIAGNOSTICS & STATS. The diagnostic support built into Vector, including logging and usage statistics



drawing by Steph Dere

[This page is intentionally left blank for purposes of double-sided printing]

CHAPTER 29

Settings, Preferences, Features, and Statistics

This chapter describes:

- The owner's account settings and entitlements
- The robot's settings (owner preferences)
- The robot's lifetime stats

124. THE ARCHITECTURE

The architecture for setting and storing settings, statistics, account information is:

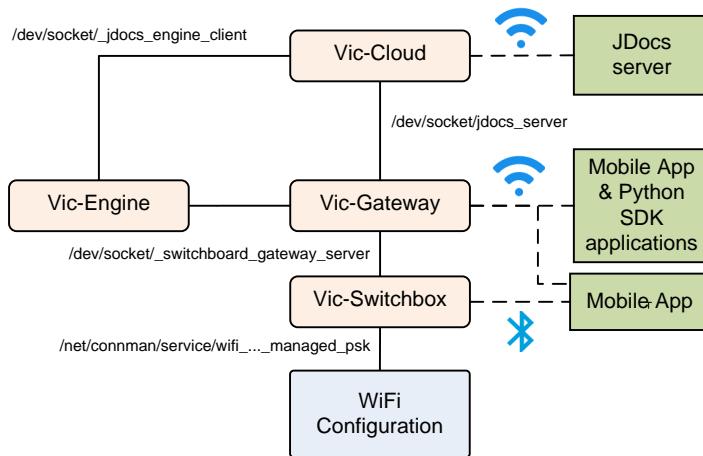


Figure 111: The architecture for storing preferences, account info, entitlements, and tracking stats

The Vic-Cloud service accesses information on a remote server.

The Vic-Switchbox interacts with the WiFi subsystem (connman) to allow the mobile App to set the preferred WiFi network to use. The mobile app must use Bluetooth LE to do this.

Vic-Gateway interacts with the mobile App and SDK programs to changes the robot settings.

Vic-Engine receives the preferences from the Vic-Cloud and Vic-Gateway, to carry out an changes in behaviour of Vector.

125. WIFI CONFIGURATION

The WiFi configuration (aka settings or preferences) is entirely local to the Vector robot. The information about the WiFi settings is not stored remotely.

The mobile application can configuration the WiFi settings via Vic-Switchbox commands. The WiFi is managed by connman thru the Vic-Switchbox:

- To provide a list of WiFi SSIDs to the mobile app
- To allow the mobile app to select an SSID and provide a password to
- Tell it forget an SSID
- To place the WiFi into Access Point mode

126. THE OWNER ACCOUNT INFORMATION

The owner account information is sent from the mobile application to Anki servers at time of registration and setting up a Vector. The owner account information includes:⁵²

Table 480: The owners account information

JSON Key	units	Description & Notes
<i>user_id</i>	base64	A base64 token to identify the user
<i>created_by_app_name</i>	string	The name of the mobile application that register the owner. Example: “chewie”
<i>created_by_app_platform</i>	string	The mobile OS version string when the mobile application created the owners account. Example “ios 12.1.2; iPhone8,1
<i>created_by_app_version</i>	string	The version of the mobile application that register the owner. Example: “1.3.1”
<i>deactivation_reason</i>		
<i>dob</i>	YYYY-MM-DD	The owner’s date of birth (the one given at time of registration)
<i>drive_guest_id</i>	GUID	A GUID to identify the owner. This is the same as the “player_id”
<i>email</i>	string	The email address used to register the account; the same as the user name.
<i>email_failure_code</i>		The reason that the email was unable to be verified
<i>email_is_blocked</i>	boolean	
<i>email_is_verified</i>	boolean	True if the email verification has successfully completed. False otherwise.
<i>email_lang</i>	IETF language tag	The IETF language tag of the owner’s language preference. example: “en-US”
<i>family_name</i>	string	The surname of the owner; null if not set
<i>gender</i>	string	The gender of the owner; null if not set
<i>given_name</i>	string	The given of the owner; null if not set
<i>is_email_account</i>	boolean	
<i>no_autodelete</i>	boolean	
<i>password_is_complex</i>	boolean	
<i>player_id</i>	GUID	A GUID to identify the owner. This is the same as the

⁵² It is not clear why there is so much information, and why this is sent from the Jdocs server in so many cases.

		“drive_guest_id”
<i>purge_reason</i>		
<i>status</i>	string	Example “active”
<i>time_created</i>	string	The time, in ISO8601 format, that the account was created
<i>user_id</i>	base64	A base64 token to identify the owner
<i>username</i>	string	Same as the email address

127. PREFERENCES & ROBOT SETTINGS

The following settings & preferences are stored in (and retrieved from) the JDoc server. They are set by the mobile app or python SDK program using the HTTPS protocol described in chapter 13. They may also be set (in some cases) by the cloud in response to verbal interaction with the owner, via vic-cloud (e.g. “Hey Vector, set your eye color to teal.”).

127.1. ENUMERATIONS

127.1.1 ButtonWakeWord

When Vector’s backpack button is pressed once for attention, he acts as if someone has said his wake word. The ButtonWakeWord enumeration describes which wake word is treated as having been said:

Name	Value	Description
BUTTON_WAKEWORD_ALEXA	1	When the button is pressed, act as if “Alexa” was said.
BUTTON_WAKEWORDHEY_VECTOR	0	When the button is pressed, act as if “Hey, Vector” was said.

Table 481:
ButtonWakeWord
Enumeration

127.1.2 EyeColor

This is the selectable colour to set Vector’s eyes to. The JdocType enumeration maps the playful name to the following value used in the RobotSettingsConfig (and vice-versa) and the colour specification:

Name	Value	Hue	Saturation	Description
CONFUSION_MATRIX_GREEN	6	0.30	1.00	
FALSE_POSITIVE_PURPLE	5	0.83	0.76	
NON_LINEAR_LIME	3	0.21	1.00	
OVERFIT_ORANGE	1	0.05	0.95	
SINGULARITY_SAPPHIRE	4	0.57	1.00	
TIP_OVER_TEAL	0	0.42	1.00	
UNCANNY_YELLOW	2	0.11	1.00	

Table 482: EyeColor
Enumeration

The mapping from to enumeration to color values is held in

/anki/assets/cozmo_resources/config/engine/eye_color_config.json

(This path is hardcoded into libcozmo_engine.so.) This JSON configuration file is a hash that maps the EyeColor *name* (not the numeric value) to a structure with the “Hue” and “Saturation” values suitable for the SetEyeColor API command. The structure has the following fields:

Field	Type	Description & Notes
<i>Hue</i>	float	The hue to use for the color
<i>Saturation</i>	float	The saturation to use for the color.

Table 483: The eye colour JSON structure

This structure has the same interpretation as the SetEyeColor request, except the first letter of the keys are capitalized here.

The mapping of the number to the JSON key for the eye colours configuration file is embedded in Vic-Gateway. Adding more named colours would likely require successful complete decompilation and modification. Patching the binary is unlikely to be practical. The colours for the existing names can be modified to give custom, permanent eye colours.

127.1.3

Volume

This is the volume to employ when speaking and for sound effects. Note: the MasterVolume API enumeration is slightly different enumeration.

Name	Value	Description
<i>MUTE</i>	0	
<i>LOW</i>	1	
<i>MEDIUM_LOW</i>	2	
<i>MEDIUM</i>	3	
<i>MEDIUM_HIGH</i>	4	
<i>HIGH</i>	5	

Table 484: Volume Enumeration

127.2. ROBOTSETTINGSCONFIG

The RobotSettingsConfig structure has the following fields:

Field	Type	Description & Notes	Table 485: The RobotSettingsConfig JSON structure
<i>button_wakeword</i>	ButtonWakeWord	When the button is pressed, act as if this wake word (“Hey Vector” vs “Alexa”) was spoken. default: 0 (“Hey Vector”)	
<i>clock_24_hour</i>	boolean	If false, use a clock with AM and PM and hours that run from 1 to 12. If true, use a clock with hours that run from 1 to 24. default: false	
<i>default_location</i>	string	default: “San Francisco, California, United States”	
<i>dist_is_metric</i>	boolean	If true, use metric units for distance measures; if false, use imperial units. default: false	
<i>eye_color</i>	EyeColor	The colour used for the eyes. The colour is referred to by one of an enumerated set. (Within the SDK, the eyes can be set to a colour by hue and saturation, but this is not permanent.) default: 0 (TIP_OVER_TEAL)	
<i>locale</i>	strong	The IETF language tag of the owner’s language preference – American English, UK English, Australian English, German, French, Japanese, etc. default: “en-US”	
<i>master_volume</i>	Volume	default: 4 (MEDIUM_HIGH)	
<i>temp_is_fahrenheit</i>	boolean	If true, use Fahrenheit for temperature units; otherwise use Celsius. ⁵³ default: true	
<i>time_zone</i>	string	The “tz database name” for time zone to use for the time and alarms. default: “America/Los_Angeles”	

The default settings are held in

`/anki/assets/cozmo_resources/config/engine/settings_config.json`

(This path is hardcoded into libcozmo_engine.so.) The file is a JSON structure that maps each of the fields of RobotSettingsConfig to a control structure. The control structure has the following fields:

Field	Type	Description & Notes	Table 486: The setting control structure
<i>defaultValue</i>		The value to employ unless one has been given by the operator or other precedent.	
<i>updateCloudOnChange</i>	boolean	true if the value is pushed to the colour when it is changed by the operator. False if not. Won’t be restored..?	

⁵³ Anyone else notice that metric requires a true for distance, but a false for temperature? Parity.

It is implied that the setting value is to be pulled from the Cloud when the robot is restored after clearing.

128. OWNER ENTITLEMENTS

An entitlement is a family of features or resources that the program or owner is allowed to use. It is represented as set of key-value pairs. This is a concept that Anki provided provision for but was not used in practice.

The only entitlement defined in Vector's API (and internal configuration files) is "kickstarter eyes" (JSON key "KICKSTARTER_EYES"). Anki decided not to pursue this, and its feature(s) remain unimplemented.

The default entitlement settings are held in

/anki/assets/cozmo_resources/config/engine/userEntitlements_config.json

(This path is hardcoded into libcozmo_engine.so.) The file is a JSON structure that maps each of the entitlement to a control structure. The control structure is the same as *Table 486: The setting control structure*, used in settings in the previous section.

129. VESTIGAL COZMO SETTINGS

There is an "account settings" file held in

/anki/etc/config/engine/accountSettings_config.json

This path is hardcoded into libcozmo_engine.so and these settings are only read (possibly) by vic-gateway. The file is a JSON structure that maps each of the settings to a control structure. The control structure is the same as *Table 486: The setting control structure*, used in settings in an earlier section.

The settings include:

Field	Type	Description & Notes
APP_LOCALE	string	The IETF language tag of the owner's language preference – American English, UK English, Australian English, German, French, Japanese, etc. default: "en-US"
DATA_COLLECTION	boolean	default: false

Table 487: The Cozmo account settings

130. FEATURE FLAGS

Vector has granular features that can be enabled and disabled thru the use of feature flags. Feature flags allow the code to be deployed, and selectively enabled. As a software engineering practice, a feature is usually not enabled because the feature is:

- not yet fully developed, or
- specific to a customer, or
- mostly developed and being tested in some groups, or
- only enabled when there is some error occurs or other functionality is not working intended, or

- a special/premium function sold at a cost or reward (like entitlement).

Many of these possibilities do not apply to Anki. But some do. Many of the disabled features are probably disabled because they are incomplete, do not work, and likely not to work for without further development.

130.1. CONFIGURATION FILE

The features flag configuration file is located at:

/anki/data/assets/cozmo_resources/config/features.json

(This path is hardcoded into libcozmo_engine.so.) This file is organized as an array of structures with the following fields:

Field	Type	Description & Notes
<i>enabled</i>	boolean	True if the feature is enabled, false if not
<i>feature</i>	string	The name of the feature

Table 488: The feature flag structure

The set of feature flags and their enabled/disabled state can be found in Appendix H. The features are often used as linking mechanisms of the modules. It is likely modules of behavior / functionality.

130.2. COMMUNICATION INTERFACE TO THE FEATURES

The list of features can be queried with the `GetFeatureFlagList` command. The status of each individual feature (whether it is enabled or not) can be found with the `GetFeatureFlag` query. See Chapter 13 section 53 *Features & Entitlements* for more details.

131. ROBOT LIFETIME STATISTICS & EVENTS

Vector summarizes his experiences and activities into a set of fun measures. The measures can be found in Appendix K *Table 544: The robot lifetime stats schema*

The intent is that they can be shared as attaboy and novel dashboard. The lifetime statics are held in the server, updated by the robot (I don't know if the robot has a local copy), and retrievable by the application. The statistics may also be sent the DAS sever.

132. REFERENCES & RESOURCES

Wikipedia, *List of tz database time zones*,
https://en.wikipedia.org/wiki/List_of_tz_database_time_zones

CHAPTER 30

The Software Update process

This chapter describes the software update process

- The software architecture
- The software update process
- How to extract official program files



133. THE ARCHITECTURE

The architecture for updating Vector's software is:

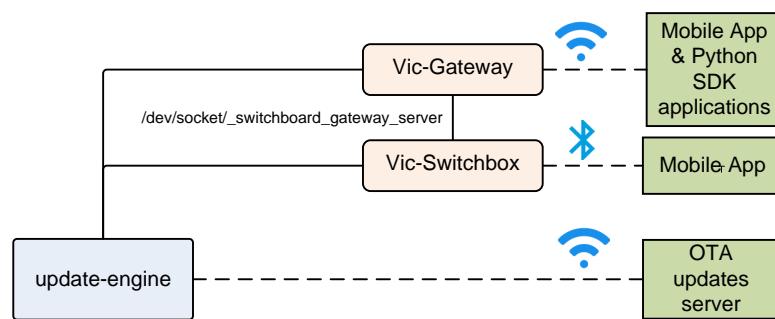


Figure 112: The architecture for updating Vector's software

The Vic-Gateway and Vic-Switchbox both may interact with the mobile App and SDK programs to receive software update commands, and to provide update status information.

The update-engine is responsible for downloading the update, validating it, applying it, and providing status information to Vic-Gateway and Vic-switchbox. [It isn't known yet how they kick off the update]. The update-engine provides status information in a set of files with the "/run/update-engine" folder

134. THE UPDATE FILE

The update files are TAR files with a suffix "OTA". The TAR file has a fixed structure, with some of the files encrypted. There are 3 kinds of update files

- Factory updates. These modify the RECOVERY and RECOVERYFS partitions.
- Production updates. These modify the ABOOT, BOOT, and SYSTEM partitions
- Delta updates. These modify the file system partitions; by sending only the changes to the underlying partitions, the updates can be very compact.

The archive contains 3 to 5 files, and they must be in a specific order:

file name	Description
<i>manifest.ini</i>	Provides a descriptor which Vector units this update can be applied to, a list of the update files, including their compression & encryption schemes, and their signature.
<i>manifest.sha256</i>	This is sha256 produced when the manifest is signed with secret OTA key. This is used to ensure that the manifest is valid.
<i>apq8009-robot-delta.bin.gz</i>	This holds the changes to the disk. This is present only in delta updates. It is optionally encrypted.
<i>apq8009-robot-emmc_appsboot.img.gz</i>	This provides a new ABOOT partition. This is present only in factory updates. To unlock Vectors, making them developer units, the modified (and signed) aboot is provided using this type of updater. It is optionally encrypted.
<i>apq8009-robot-boot.img.gz</i>	In factory updates it will be applied to the RECOVERY partition; otherwise it will be applied to the BOOT partition. This is not present in delta updates. It is optionally encrypted.
<i>apq8009-robot-sysfs.img.gz</i>	In factory updates it will be applied to the RECOVERYFS partition; otherwise it will be applied to the SYSTEM partition. This is not present in delta updates. It is optionally encrypted.

Table 489: The contents of an over-the-air update archive file

134.1. MANIFEST.INI

The *manifest.ini* is checked by verifying its signature⁵⁴ against *manifest.sha256* using a secret key (/anki/etc/ota.pub):

```
openssl dgst \
    -sha256 \
    -verify /anki/etc/ota.pub \
    -signature /run/update-engine/manifest.sha256 \
    /run/update-engine/manifest.ini
```

Example 8: Checking the *manifest.ini* signature

Note: the signature check that prevents turning off encryption checks in the manifest below. At this time the signing key is not known.

All forms of update have a [META] section. This section has the following structure:

Key	Description
<i>ankidev</i>	0 if production release, 1 if development
<i>manifest_version</i>	Acceptable versions include 0.9.2, 0.9.3, 0.9.4, 0.9.5, or 1.0.0
<i>num_images</i>	The number of img.gz files in the archive. The number must match that of the type of update file it is. 1, 2, or 3
<i>qsn</i>	The Qualcomm Serial Number; if there are three images (ABOOT, RECOVERY, RECOVERYFS) present, the software is treated as a factory update. The QSN must match the robot's serial number. Optional.
<i>reboot_after_install</i>	0 or 1. 1 to reboot after installing.
<i>update_version</i>	The version that the system is being upgrade to, e.g. 1.6.0.3331

Table 490: *manifest.ini* META section

⁵⁴ I'm using the information originally at: <https://github.com/GooeyChickenman/victor/tree/master/firmware>

After the [META] section, there are 1 to 3 sections, depending on the type of update:

- A delta update has a [DELTA] section
- A regular update has a [BOOT], [SYSTEM] sections; both must be present/
- A factory update has [ABOOT], [RECOVERY], and [RECOVERYFS] sections; all 3 must be present.

Each of these sections has the same structure:

Key	Description
<i>base_version</i>	The version that Vector must be at in order to accept this update. Honored only in delta updates.
<i>bytes</i>	The number of bytes in the uncompressed archive
<i>compression</i>	gz (for gzipped). This is the only supported compression type.
<i>delta</i>	1 if this is a delta update; 0 otherwise
<i>encryption</i>	1 if the archive file is encrypted; 0 if the archive file is not encrypted.
<i>sha256</i>	The digest of the decompressed file must match this
<i>wbits</i>	31. Not used by update-engine

Table 491:
manifest.ini image stream sections

134.1.1 Version numbers

When performing version checks on the update file, looks at the number in *update_version*, the suffix in the *update_version* and the *ankidev* field. The update-engine ensures that production Vectors will not install software with the *ankidev* field set – and that developer Vector’s will not install production software. (This is probably to prevent work at allowing further development updates to be installed on the unit.)

There are also subtle different kinds of development software. This is indicated in the suffix at the end of the version string – blank, “d” or “ud”. The update-engine ensures that a Vector cannot be changed from running software with one kind of suffix to another kind.

Wire/Kerigan
Creighton

Type	anki.dev	suffix	Description
<i>developer</i>	1	<i>d</i>	This can include developer tools, such as SCP, SSH, AWK, a web server, and a web-viz browser-based visualization application. It has logging, simulation visualization, microphone processing information, beat detection visualization, cloud intents, CPU usage statistics, etc.
<i>production</i>	0		This is the end-consumer released software. (The boot partition is signed, and dmverity is enabled for the system partition.)
<i>release candidate</i>	1		This is almost identical to the production software releases, and is likely used to test the units.
<i>userdev</i>	1	<i>ud</i>	Some have SSH installed, but often do not include web-viz & web-server.

Table 492: Different kinds of Vector software updates⁵⁵

⁵⁵ <https://docs.google.com/document/d/1KZ93SW7geM0gA-LBXHdt55a9NR1jfKp7UZyqlRuokno/edit>

134.2. HOW TO DECRYPT THE OTA UPDATE ARCHIVE FILES⁵⁶

How to decrypt the OTA update archive files:

```
openssl enc -d -aes-256-ctr -pass file:ota.pas -in apq8009-robot-boot.img.gz -out apq8009-robot-boot.img.dec.gz  
openssl enc -d -aes-256-ctr -pass file:ota.pas -in apq8009-robot-sysfs.img.gz -out apq8009-robot-sysfs.img.dec.gz
```

Example 9: Decrypting the OTA update archives

To use OpenSSL 1.1.0 or later, add “-md md5” to the command:

```
openssl enc -d -aes-256-ctr -pass file:ota.pas -md md5 -in apq8009-robot-boot.img.gz -out apq8009-robot-boot.img.dec.gz  
openssl enc -d -aes-256-ctr -pass file:ota.pas -md md5 -in apq8009-robot-sysfs.img.gz -out apq8009-robot-sysfs.img.dec.gz
```

Example 10: Decrypting the OTA update archives with Open SSL 1.1.0 and later

Note: the password on this file is insecure (ota.pas has only a few bytes⁵⁷) and likely intended only to prevent seeing the assets inside of the update file. The security comes from (a) the individual image files are signed (this is checked by the updater), and (b) the file systems that they contain are also signed, and are checked by aboot and the initial kernel load. See Chapter 7 *Startup* for the gory details.

Signing the files is a whole other kettle of fish.

135. THE UPDATE PROCESS

135.1. STATUS DIRECTORY

The update-engine provides its status thru a set of files in the /run/update-engine folder.

File	Description
<i>done</i>	If this file exists, the update has completed
<i>error</i>	The error code representing why the update failed. See Appendix D, <i>Table 512: OTA update-engine status codes</i>
<i>expected-download-size</i>	The expected file size (the given total size of the OTA file) to download.
<i>expected-size</i>	In non-delta updates, the total number of bytes of the unencrypted image files. This is the sum of the “bytes” field in the sections.
<i>progress</i>	Indicates how many of the bytes to download have been completed, or how much of the partitions have been written.

Table 493: update-engine status file

This folder also holds the unencrypted, uncompressed files from the OTA file:

- manifest.ini
- manifest.sha256
- delta.bin
- aboot.img
- boot.img

⁵⁶ <https://groups.google.com/forum/#!searchin/anki-vector-rooting/ota.pas%7Csort:date/anki-vector-rooting/YIYQsX08OD4/fvkAOZ91CgAJ>

⁵⁷ Opening up the file in a UTF text editor will show Chinese glyphs; google translate reveals that they say “This is a password”. This password is a bit of humour to comply with a security consultant.

135.2. PROCESS

The update process if there is an error at any step, skips the rest, deletes the bin and img files.

1. Remove everything in the status folder
2. Being downloading the OTA file. It does *not* download the TAR then unpack it. The file is unpacked as it is received.
3. Copies the manifest.ini to a file in the status folder
4. Copies the manifest.sha256 to a file
5. Verifies the signature of the manifest file
6. Validates that the update to the OTA version is allowed. .

Note: the software can be downgraded by going into the recovery mode (running factory software). The recovery mode does not check the version number or suffix. These checks were likely included (in the main software) to prevent any server bugs from accidentally downgrading Vectors – wiping out bug fixes –at home.

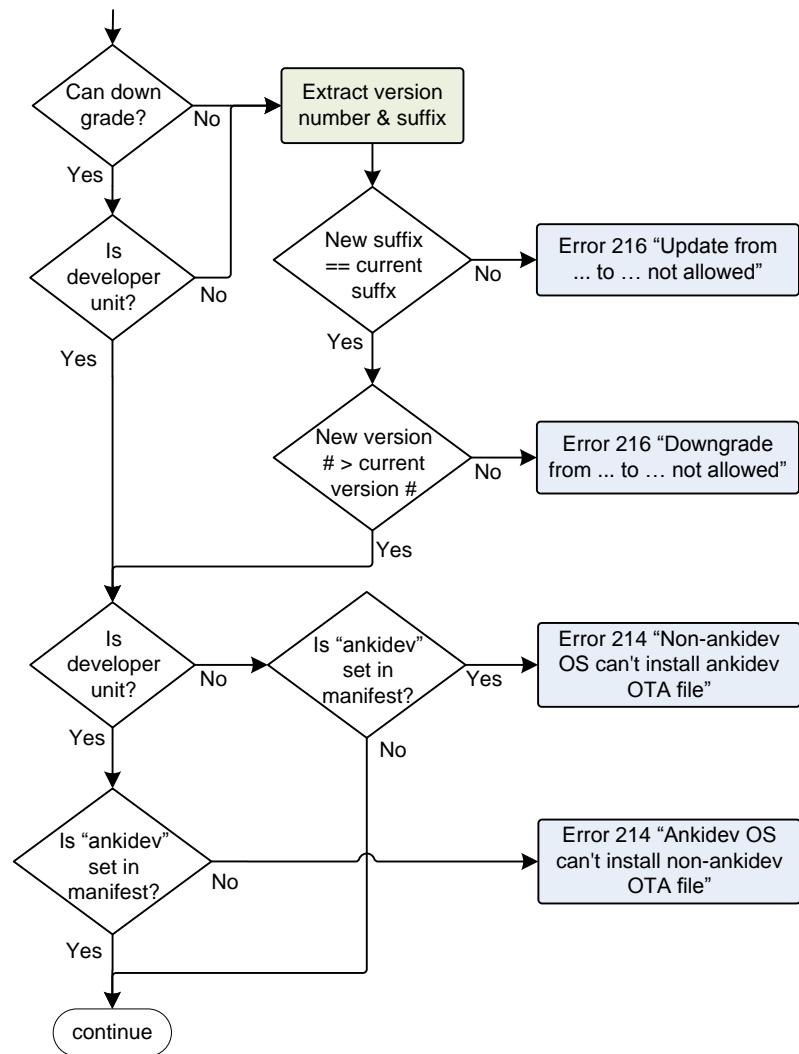


Figure 113: The update-engine's version check process

- a) If this is a development Vector (i.e. anki.dev is set on the linux boot command line), and the current software has UPDATE_ENGINE_ALLOW_DOWNGRADE internally set (to true), the next two checks are skipped (until step d). Otherwise,
 - b) Does the suffix at the end of the version number in the new manifest match the suffix in the currently running software? If not, a 216 error code is produced.
 - c) Is the new version number in the new manifest greater than the one in the currently running software? If not, a 216 error code is produced.
 - d) The ankidev variable in the manifest must be set on developer units, and must not be set on production units; otherwise a 214 error code is produced.
7. If this is factory update, it checks that the QSN in the manifest matches Vector's QSN.
 8. It marks the target partition slots as unbootable
 9. Checks the img and bin contents
 - a) delta
 - b) boot & system
 - c) If this is a factory update, aboot, recovery, and recoveryfs
 10. If this is a factory update:
 - a) Sets /run/wipe-data. This will trigger erasing all of the user data on the next startup
 - b) Makes both a and b slots for BOOT and SYSTEM partitions as unbootable
 11. If this is not factory update
 - a) Sets the new target slot as active
 12. Deletes any error file
 13. Sets the done file
 14. Posts a DAS event robot.ota_download_end to success + next version
 15. If the reboot_after_install option was set, reboots the system

136. RESOURCES & RESOURCES

<https://source.android.com/devices/bootloader/flashing-updating>

Describes the a/b process as it applies to android

CHAPTER 31

Diagnostics

This chapter describes the diagnostic support built into Vector

- The customer care information screen
- The logging of regular use
- Crash logs
- Gathering usage, and performance data

137. OVERVIEW

Anki gathers “analytics data to enable and improve the services and enhance your gameplay... Analytics Data enables us to analyze crashes, fix bugs, and personalize or develop new features and services.” There are lot different services that accomplish the analytics services. This data is roughly: logs, crash dumps and “DAS manager”

Logging and diagnostic messages are typically not presented to the owner, neither in use with Vector or thru the mobile application... nor even in the SDK.

The exception is gross failures that display a 3-digit error code. This is intended to be very exceptional.

Diagnostic and logging information is available thru undocumented interfaces.⁵⁸

137.1. THE SOFTWARE INVOLVED

There are many different programs and libraries used in the diagnostic and logging area. The table below summarizes of them:

Program / Library	Description	
<i>animfail</i>	This is started by the <i>animfail</i> service.	
<i>anki-crash-log</i>	Copies the last 500 system messages and the crash dump passed to the command line to a given log file. This is called by <i>vic-cloud</i> , <i>vic-dasmgr</i> , <i>vic-engine</i> , <i>vic-gateway</i> , <i>vic-log-kernel-panic</i> , <i>vic-log-upload</i> , <i>vic-robot</i> , <i>vic-switchboard</i> , and the <i>anki-crash-log</i> service.	
<i>ankitrace</i>	This program wraps the Linux tract toolkit (LTTng). This program is not present in Vector’s file system. This is called by <i>fault-code-handler</i> .	
<i>cti_vision_debayer</i>	This is not called.	
<i>diagnostics-logger</i>	Bundles together several log and configuration states into a compressed tar file. This is called by <i>vic-switchboard</i> , in a response to a Bluetooth LE log command.	
<i>displayFaultCode</i>	Displays error fault codes on the LCD. This is not called; see <i>vic-faultCodeDisplay</i> .	

Table 494: Vector diagnostic & logging software

⁵⁸ The lack of documentation indicates that this was not intended to be supported and employed by the public... at least not until other areas had been resolved.

<i>fault-code-clear</i>	This clears and pending or displayed faults (by deleting the relevant files). This allows new fault code to be displayed. This is called by <i>vic-init.sh</i> .
<i>fault-code-handler</i>	This is called by the fault-code service.
<i>librobotLogUploader.so</i>	Sends logs to cloud. This library is employed by <i>libcozmo_engine</i> , <i>vic-gateway</i> and <i>vic-log-upload</i> .
<i>libosState</i>	Used to profile the CPU temperature, frequency, load; the WiFi statistics, and etc. This is used by <i>libvictor_web_library</i> , <i>vic-anim</i> , and <i>vic-dasmgr</i> .
<i>libwhiskeyToF</i>	This unusually named library ⁵⁹ has lots of time of flight sensor diagnostics. This is present only in version 1.6. This library is employed by <i>libcozmo_engine</i> .
<i>rampost</i>	This is to perform initial communication and version check of the firmware on the body-board (syscon). This exists within the initial RAM disk, and is called by init.
<i>vic-anim</i>	Includes the support for the Customer Care Information Screen. This is started by the <i>vic-anim</i> service.
<i>vic-crashuploader-init</i>	Removes empty crash files, renames the files ending in “.dmp~” to “.dmp”. This is called by the <i>vic-crashuploader</i> service.
<i>vic-crashuploader</i>	A script that sends crash mini-dump files to backtrace.io. This is called by the <i>vic-crashuploader</i> service.
<i>vic-dasmgr</i>	This is started by the <i>vic-dasmgr</i> service.
<i>vic-faultCodeDisplay</i>	Displays error fault codes on the LCD. This is called by <i>fault-code-handler</i> .
<i>vic-init.sh</i>	Takes the log messages from <i>rampost</i> and places them into the system log, forwards any kernel panics. This is started by the <i>vic-init</i> service.
<i>vic-log-event</i>	A program that is passed an event code in the command line. This is called by TBD.
<i>vic-log-forward</i>	This is called by <i>vic-init.sh</i>
<i>vic-log-kernel-panic</i>	This is called by <i>vic-init.sh</i>
<i>vic-log-upload</i>	This is called by <i>vic-log-uploader</i>
<i>vic-log-uploader</i>	“This script runs as a background to periodically check for outgoing files and attempt to upload them by calling ‘ <i>vic-log-upload</i> ’.” This is started by the <i>vic-log-uploader</i> service.
<i>vic-logmgr-upload</i>	“This script collects a snapshot of recent log data” into a compressed (gzip) file, then uploads the file” and software revision “to an Anki Blobstore bucket.” This is not called.
<i>vic-on-exit</i>	Called by <i>systemd</i> after any service stops. This script places the fault code associated with the service (if another fault code is not pending) into <i>/run/fault_code</i> for display.
<i>vic-powerstatus.sh</i>	Record every 10 seconds the CPU frequency, temperature and the CPU & memory usage of the “ <i>vic-</i> ” processes. This is not called.

(Quotes from Anki scripts.) Support programs are located in */bin*, */anki/bin*, and */usr/bin*

⁵⁹ Anki has taken great care for squeaky-clean image, even throughout the internal files, so it was a surprise to see one that might appear named after a rude acronym (WTF). The name is a result of the internal product codes: *Whiskey* was the code name for a new generation of Cozmo in development. This was its time of flight (ToF) sensor library, using a modified Vector (called “Spiderface”) as a development prototype. On Whiskey, the time of flight sensor would connect directly to the main processor.

138. SPECIAL SCREENS AND MODES

Vector has 3 special screens and two special modes. The screens are

- Customer Care Info Screen (CCIS) that can display sensor values and other internal measures,
- Debug screen used to display Vector's serial number (ESN) and IP address, and
- The fault code display which is used to display a 3-digit fault code when there is an internal failure (this screen is only displayed if there is a fault, and can't be initiated by an operator.)

Vector has two special modes

- Entering recovery mode, to force Vector use factory software and download replacement firmware. (This mode doesn't delete any user data.)
- "Factory reset" which erases all user data, and Vector's robot name

138.1. CUSTOMER CARE INFORMATION SCREEN

Customer Care Info Screen (CCIS). It has a series of screens that display sensor values and other readings.

See <https://www.kinvert.com/anki-vector-customer-care-info-screen-ccis/> for a walk thru

138.2. VECTORS' DEBUG SCREEN (TO GET INFO FOR USE WITH THE SDK)

Steps to enter the debug screen

1. Place Vector on the charger,
2. Double-click his backpack button,
3. Move the arms up and down

This will display his ESN (serial number) and IP address. The font is much smaller than normal, and may be hard to read.

138.3. DISPLAY FAULT CODES FOR ABNORMAL SYSTEM SERVICE EXIT / HANG

If there is a problem while the system is starting or running— such as one of the services exits (e.g. crashes) – a fault code associated with that service is stored in `/run/fault_code` and the fault code displayed. See Appendix D for fault codes.

138.4. RECOVERY MODE

Vector includes a *recovery mode* that is used to force Vector to boot using factory software. The recovery mode will not delete any user data or software that had previously been installed via Over-The-Air (OTA) update.

The recovery mode is intended to help with certain issues such as Vector failing to boot up using the regular firmware. He may have been unable to charge (indicated by teal Back Lights), or encountered other software bugs⁶⁰.

The application in the recovery mode attempts to download and reinstall the latest software. This is likely done under the assumption that the firmware may be corrupted, or not the latest, and that a check for corruption would take so long as to not be useful.

138.5. “FACTORY RESET”

Erases all user data, include pictures, faces, and API certificates. It clears out the robot name. The Vector will be given a new robot name when he is set up again.

The name “factory reset” is controversial, as this does not truly place Vector into an identical software state as robot in the factory.

139. BACKPACK LIGHTS

The lights on the backpack are primarily set by Vic-robot, but driven by the body-board. If the body-board firmware (syscon) is unable to communicate with Vic-robot, it will set the lights on its own.

140. DIAGNOSTIC COMMANDS

There are several HTTPS commands that are useful for diagnosing errors:

The connectivity with the cloud can be checked to see if the servers can be reached, if the authentication (i.e. username and password) is valid, if the server certificate is valid. See Chapter 13, section *50.1 Check Cloud Connection*

The debug logs can be requested to be sent to the server for analysis. See the Upload Debug Logs command in Chapter 13, section *50.2 Upload Debug Logs*

⁶⁰ The web page says that are “indicated by a blank screen. If you get a status code between 200-219, recovery mode will also help.”

141. LOGS

- Logs can be downloaded to a PC or mobile application using the Bluetooth LE API
- The Logs can be used to the server using the SDK command X

141.1. GATHERING LOGS, ON DEMAND

The logs can be requested by issuing a log fetch command via Bluetooth LE. Vic-switchboard handles the request, delegating the preparation of the log files to diagnostics-logger. This utility gathers the following tars and compresses them:

File	Description
<i>connman-services.txt</i>	connmanctl services
<i>dmesg.txt</i>	dmesg
<i>ifconfig.txt</i>	ifconfig wlan0
<i>iwconfig.txt</i>	iwconfig wlan0
<i>log.txt</i>	Concatenates /var/log/messages.1.gz (uncompressed) and /var/log/messages
<i>netstat-ptlnu.txt</i>	netstat -ptlnu
<i>ping-anki.txt</i>	Ping's anki.com for connectivity and latency.
<i>ping-gateway.txt</i>	Looks up the IP address (using netstat) of the gateway that Vector is using and pings it for connectivity and latency.
<i>ps.txt</i>	Process stats (ps) of Anki's "Vic" processes
<i>top.txt</i>	top -n 1

Table 495: Files in the log archive

This utility is triggered by:

- vic-switchboard when issued a log fetch command (via Bluetooth LE).
- Vic-gateway when the upload log command is issued
- Other

142. CONSOLE FILTER

The logging by functional blocks (primarily in Vic-engine) can be configured. The logging configuration file is located at:

/anki/data/assets/cozmo_resources/config/engine/console_filter_config.json

This file is organized as dictionary whose key is the operating system. The “vicos” key is the one relevant for Vector.⁶¹ It dereferences to a structure with the following fields:

Field	Type	Description & Notes
channels	array	An array of the channel logging enable structures
levels	array	An array of logging level enable structures

This “channels” is as an array of structures with the following fields:

Field	Type	Description & Notes
channel	string	The name of the channel
enabled	boolean	True if should log information from the channel, false if not.

This “levels” is an array of structures with the following fields:

Field	Type	Description & Notes
enabled	boolean	True if should log information at that level, false if not.
level	string	“event” or “debug”

The features are used as linking mechanisms of the modules. It is likely modules of behavior / functionality. It is not clear how it all ties together.

Channel	enabled	Description & Notes
Actions	false	
AIWhiteboard	false	
Alexa	false	
Audio	false	
Behaviors	false	
BlockPool	false	
BlockWorld	false	
CpuProfiler	true	
FaceRecognizer	false	
FaceWorld	false	
JdocsManager	true	
MessageProfiler	true	

Table 496: The console filter channel structure

Table 497: The channel logging enable structure

Table 498: The logging level enable structure

Table 499: The channels

⁶¹ The other OS key is “osx” which suggests that Vector’s software was development on an OS X platform.

<i>Microphones</i>	false
<i>NeuralNets</i>	false
<i>PerfMetric</i>	true
<i>SpeechRecognizer</i>	false
<i>VisionComponent</i>	false
<i>VisionSystem</i>	false
*	false

143. USAGE STUDIES AND PROFILING DATA

Anki had ambitious to perform engagement studies and experiments with device settings:

“The Services collect gameplay data such as scores, achievements, and feature usage. The Services also automatically keep track of information such as events or failures within them. In addition, we may collect your device make and model, an Anki-generated randomized device ID for the mobile device on which you run our apps, robot/vehicle ID of your Anki device, ZIP-code level data about your location (obtained from your IP address), operating system version, and other device-related information like battery level (collectively, “Analytics Data”).”

The DAS manager protocol’s version identifier dates to development of Overdrive. One patent on their “Adaptive Data Analytics Service” is quite an ambitious plan to tune and improve systems.

“A closed-loop service, referred to as an Adaptive Data Analytics Service (ADAS), characterizes the performance of a system or systems by providing information describing how users or agents are operating the system, how the system components interact, and how these respond to external influences and factors. The ADAS then builds models and/or defines relationships that can be used to optimize performance and/or to predict the results of changes made to the system(s). Subsequently, this learning provides the basis for administering, maintaining, and/or adjusting the system(s) under study. Measurement can be ongoing, even after the operating parameters or controls of a system under the administration or monitoring of the ADAS have been adjusted, so that the impact of such adjustments can be determined. This recursive process of observation, analysis, and adjustment provides a closed-loop system that affords adaptability to changing operating conditions and facilitates self-regulation and self-adjustment of systems.”

There is no information on whether this was actually accomplished, or that these techniques were used in Cozmo or Vector. Anki developed “both batch and real-time dashboards to gain insights over device and user behavior,” according to their Elemental toolkit literature.

143.1. EVENT TRACING

The DAS manager on Vector and the mobile application posts event such as when an activity begins, key milestones along the way, and when the activity ends. The events can include extra parameters such as text and values. In the case of the mobile application, this is the name of each button pressed, screen displayed, error encountered, and so forth.

Speculated purpose:

- To identify how far people got in a process, or what their flow thru an interaction is

- To estimate durations of activities, such as onboarding, how long Vector can play between charge cycles, and how long a charge cycle is.
- To identify unusual events (such as errors).
- May allow detailed reconstruction of the setup, configuration and interaction

The event naming pattern is *[module name].[some arbitrary name]*

When these are logged in Vector’s text log files they are prefixed with an ‘@’ symbol.⁶²

For examples of DAS events, see Appendix K.

143.2. PROFIILING AND LIBOSSTATE

The tools in Vector gather a variety of diagnostic information about

- Basic information about the robot – the version of software it is running, and what the robot’s identifier/serial number is.
- Whether Vector is booted into recovery mode when it is sending the information.
- The uptime – how long Vector has been running since the last reboot or power on.
- The WiFi performance, to understand the connectivity at home since Vector depends so heavily on cloud connectivity for his voice interactions.
- The CPU temperature profile, to find the balance between overheating and AI performance. Some versions and features of Vector can cause faults due to the processor overheating. Anki probably wanted to identify unusual temperatures and whether their revised settings addressed it.
- The CPU and memory usage statistics for the “vic-” application services. Anki probably sought to identify typical and on unusual processing loads and heavy use cases.
- The condition of the storage system – information about the flash size & partitions, whether the user space is “secure”, and whether the EMR is valid.

Speculated purpose: To identify typical and on unusual processing loads and temperatures. The heavy uses cases are likely undesired and would be something to identify.

The data gather in Vector for these is primarily based in a library called libosState.

143.2.1 WiFi Stats

libosState gathers the following information about the WiFi network:

- The WiFi MAC address
- The WiFi SSID (and flagged if it isn’t valid)
- The assigned IP Address (and flagged if it isn’t valid)
- The number of bytes received and sent
- The number of transmission and receive errors

⁶² This is a very helpful feature

The key files employed to access this information:

File	Description
/sys/class/net/wlan0/address	The IP address assigned to Vector
/sys/class/net/wlan0/statistics/rx_bytes	The number of bytes received
/sys/class/net/wlan0/statistics/rx_errors	The number of receive errors
/sys/class/net/wlan0/statistics/tx_bytes	The number of transmit errors
/sys/class/net/wlan0/statistics/tx_errors	The number of bytes sent

Table 500: The WiFi related stats /proc files

How this is used: to get a sense of WiFi connectivity in the home, and rooms where Vector is used. Anki's internal research showed that rooms in a home can have a wide range of connectivity characteristics.

Jane Fraser, 2019

143.2.2 CPU stats

libosState gathers the following information about the CPU temperature:

- The CPU temperature
- The CPU target and actual frequency
- Whether the CPU is being throttled
- The limits set on the CPU frequency

The key files employed to access this information:

File	Description
/sys/devices/system/cpu/cpu0/cpufreq/cpuinfo_cur_freq	
/sys/devices/system/cpu/cpu0/cpufreq/scaling_max_freq	
/sys/devices/system/cpu/cpu0/cpufreq/scaling_governor	
/sys/devices/system/cpu/cpu0/cpufreq/scaling_setspeed	
/sys/devices/virtual/thermal/thermal_zone3/temp	

Table 501: Named device and control files

How this is used: This information was probably intended to find the balance between overheating and AI performance.

143.3. EXPERIMENTS

There is an experiments file. This is in libcozmo. Cozmo's APK has a file with the same structure. The file has the following high-level structure:

Field	Type	Description & Notes
meta	meta structure	A structure that describes what project the experiment applies to and the versioning info of the structure.
experiments	array of experiment structures	An array of experiments, each with their own conditions and parameters.

Table 502: The experiments TBD structure

The meta structure has the following fields:

Field	Type	Description & Notes
<i>project_id</i>	string	“cozmo” ⁶³
<i>revision</i>	int	1
<i>version</i>	int	2

Table 503: The meta JSON structure

An experiment structure has the following fields:

Field	Type	Description & Notes
<i>activation_mode</i>	string	“automatic”
<i>audience_tags</i>	array of TBD	
<i>forced_variations</i>	array of TBD	
<i>key</i>	string	“report_test_auto”
<i>pop_frac_pct</i>	int	Portion of the population, as a percentage, that will take part in this experiment.
<i>pause_time_utc_iso8601</i>	string	
<i>resume_time_utc_iso8601</i>	string	
<i>start_time_utc_iso8601</i>	string	The date and time that the experiment will commence.
<i>stop_time_utc_iso8601</i>	string	The date and time that the experiment will end.
<i>variations</i>	array of variation	
<i>version</i>	int	0

Table 504: The experiment JSON structure

A variation structure has the following fields:

Field	Type	Description & Notes
<i>key</i>	string	One of at least two populations subject to the test: “control” or “treatment”
<i>pop fract pct</i>	int	Portion of the population, as a percentage, that will be in this subject group.

Table 505: The variation JSON structure

144. REFERENCES & RESOURCES

Anki, Privacy policy, 2018 Oct 5

<https://anki.com/en-us/company/privacy.html>

DeNeale, Patrick; Tom Eliaz; *Adaptive data analytics service*, Anki, USPTO
US9996369B2, 2015-Jan-05

os-release — Operating system identification

<https://www.freedesktop.org/software/systemd/man/os-release.html>

Describes the /etc/os-version and /etc/os-version-rev files

⁶³ I suspect that this would have changed once experiments were initiated with Vector

References & Resources

Note: most references appear in the margins, significant references will appear at the end of their respective chapter.

145. CREDITS

Credit and thanks to Anki who made Vector possible; CORE, Melanie T for access to the flash partitions, file-systems, decode keys, board shots, unusual LED codes, information on the electronics, and OTA URLs. Wire/Kerigan Creighton for board shots, the Project Victor website & public relations, finding the web-visualization tool, OTA URLs, and identifying the valuable OTA versions. Fictiv for board shots. (The board shots helped identify parts on the board and inter-connection on the board.) GooeyChickenman for the github repository. Cyril Peponet for aboot analysis, finding OTA v1.7, and pointing me valuable past discord postings. Alexander Entinger for connector signal information. HSReina for Bluetooth LE protocol information. Wayne Venables for crafting a C# version of the SDK. Silvarius/Silvarius613 & nammo for info on the other Anki products that were under development. nammo for information on error codes, shaft encoders, battery life, signal processing, and much more. Several drawings were adapted from Steph Dere, and Jesse Easley's twitter & instagram.

Thank-you to Digital Dream Labs (DDL) for continuing support for Vector; DDL and Drew Zhrodague for providing error tables and cloud information.

146. REFERENCE DOCUMENTATION AND RESOURCES

146.1. ANKI

Anki, *Vector Quick Start Guide*, 293-00036 Rev: B, 2018

Anki, *Vector Pillars*, 2018

Casner, Daniel, *Sensor Fusion in Consumer Robots*, Embedded Vision Summit, 2019 May

<https://www.embedded-vision.com/platinum-members/embedded-vision-alliance/embedded-vision-training/videos/pages/may-2019-embedded-vision-summit-casner>

<https://www.youtube.com/watch?v=NTU1egF3Z3g>

Casner, Daniel; Lee Crippen, Hanns Tappeiner, Anthony Armenta, Kevin Yoon; *Map Related Acoustic Filtering by a Mobile Robot*, Anki, US Patent 0212441 A1, 2019 Jul 11

Fraser, Jane, *IoT: How it Changes the Way We Test*, Spring 2019 Software Test Professionals Conference, 2019 Apr 3

<https://spring2019.stpcon.com/wp-content/uploads/2019/03/Fraser-IoT-How-it-changes-the-way-we-test-updated.pdf>

Jameson, Molly; Daria Jerjomina; *Cozmo: Animation pipeline for a physical robot*, Anki, 2017 Game Developers conference

Kaiser (Anki), *(Cozmo) Code Lab Block Glossary*, 2017 Dec
<https://forums.anki.com/t/code-lab-block-glossary/10958>

Monson, Nathaniel; Andrew Stein, Daniel Casner, *Reducing Burn-in of Displayed Images*, Anki, US Patent 20372659 A1; 2017 Dec 28

Stein, Andrew; *Making Cozmo See*, Embedded Vision Alliance, 2017 May 25
<https://www.slideshare.net/embeddedvision/making-cozmo-see-a-presentation-from-anki>
<https://youtu.be/Ypz7sNgSzyI>

Wolford, Jason; Ben Gabaldon, Jordan Rivas, Brian Min *Condition-Based Robots Audio Techniques*, Anki , USPTO Pub.No: US2019/0308327A1, 2018 Apr 6

146.2. OTHER

Coull, Ashley, *Sound for Robots: An Interview with Sr. Sound Designer Ben Gabaldon*, 2016 Nov 15, Designing Sound
<http://designingsound.org/2016/11/15/sound-for-robots-an-interview-with-sr-sound-designer-ben-gabaldon/>

cozmopedia.org

Crowe, Steven, *Anki was developing security robots before shutdown*, The Robot Report, 2020 Feb 25
<https://www.therobotreport.com/anki-developing-security-robots-before-shutdown/>

Easley, Jesse
<https://fatralla.tumblr.com/>

FCC ID 2AAIC00010 internal photos
<https://fccid.io/2AAIC00010>

FCC ID 2AAIC00011 internal photos
<https://fccid.io/2AAIC00011>

FPL, *FlatBuffers*
<https://google.github.io/flatbuffers/>

Kinvert, *Anki Vector Customer Care Info Screen (CCIS)*
<https://www.kinvert.com/anki-vector-customer-care-info-screen-ccis/>

Sriram, Swetha, *Anki Vector Robot Teardown*, Fictiv, 2019 Aug 6
<https://www.fictiv.com/blog/anki-vector-robot-teardown>

Tenchov, Kaloyan; *PyCozmo*
<https://github.com/zayfod/pycozmo/tree/master/pycozmo>

Venable, Wayne; *Anki.Vector.SDK*
<https://github.com/codaris/Anki.Vector.SDK>
<https://github.com/codaris/Anki.Vector.Samples>
<https://weekendrobot.com/>

Zaks, Mazim *FlatBuffers Explained*, 2016-Jan-30
<https://github.com/mzaks/FlatBuffersSwift/wiki/FlatBuffers-Explained>

146.3. QUALCOMM

Although detailed documentation isn't available for the Qualcomm APQ8009, there is documentation available for the sibling APQ8016 processor.

Qualcomm, *APQ8016 Application Processor Tools & Resources*,
<https://developer.qualcomm.com/hardware/apq-8016/tools>

Qualcomm, *DragonBoard™ 410c based on Qualcomm® Snapdragon™ 410E processor ADB Debugging Commands Guide*, LM80-P0436-11, Rev C, 2016 Sep
lm80-p0436-11_adb_commands.pdf

Qualcomm, *DragonBoard™ 410c based on Qualcomm® Snapdragon™ 410E processor Software Build and Installation Guide, Linux Android*, LM80-P0436-2, Rev J, 2016 Dec
lm80-p0436-2_sw-build-and-install_gd_linux_android_dec2016.pdf

[This page is intentionally left blank for purposes of double-sided printing]

Appendices

These appendices provide extra material supplemental to the main narrative. These include tables of information, numbers and keys.

- **ABBREVIATIONS, ACRONYMS, & GLOSSARY.** This appendix provides a gloss of terms, abbreviations, and acronyms.
- **TOOL CHAIN.** This appendix lists the tools known or suspected to have been used by Anki to create, and customize the Vector, and for the servers. Tools that can be used to analyze Vector.
- **ALEXA MODULES.** This appendix describes the modules used by the Alexa client
- **FAULT AND STATUS CODES.** This appendix provides describes the system fault codes, and update status codes.
- **FILE SYSTEM.** This appendix lists the key files that are baked into the system.
- **BLUETOOTH LE PROTOCOLS.** This appendix provides information on the Bluetooth LE interfaces to the companion Cube, and to Anki Vector
- **SERVERS.** This appendix provides the servers that the Anki Vector and App contacts
- **FEATURES.** This appendix enumerates the Vector OS “features” that can be enabled and disabled; and the AI behavior’s called “features.”
- **PHRASES.** This appendix reproduces the phrases that the Vector keys off of.
- **EMOTION EVENTS.** This appendix provides a list of the emotion events that Vector internally responds to.
- **DAS EVENTS.** This appendix describes the identified DAS events
- **PLEO.** This appendix gives a brief overview of the Pleo animatronic dinosaur, an antecedent with many similarities.



[This page is intentionally left blank for purposes of double-sided printing]

APPENDIX A

Abbreviations, Acronyms, Glossary

Abbreviation / Acronym	Phrase
ADC	analog to digital converter
AG	animation group
ALSA	advanced Linux sound architecture
APQ	application processor Qualcomm (used when there is no modem in the processor module)
ASR	automatic speech recognition
AVS	Alexa Voice Service
BIN	binary file
BMS	battery management system
BNK	AudioKinetic sound bank file
CCIS	customer care information screen
CLAD	C-like abstract data structures
CLAHE	contrast-limited adaptive histogram equalization
CNN	convolution neural network
CRC	cyclic redundancy check
CSI	Camera serial interface
DAS	<i>unknown (diagnostic/data analytics service?)</i>
DFU	device firmware upgrade
DTTB	Dance to the beat
DVT	design validation test
EEPROM	electrical-erasable programmable read-only memory
EMR	electronic medical record
ESD	electro-static discharge
ESN	electronic serial number
EVT	engineering validation test
FBS	flat buffers
FDE	full disc encryption
FFT	fast Fourier transform

Table 506: Common acronyms and abbreviations

GPIO	general purpose IO
gRPC	Google remote procedure call
GUID	globally unique identifier (effectively same as UUID)
HLAI	high-level AI
I2C	inter-IC communication
IMA ADPCM	interactive multimedia association adaptive pulse-code modulation
IMU	inertial measurement unit
IR	infrared
JDocs	JSON Documents
JSON	JavaScript Object Notation
JTAG	Joint Test Action Group
LCD	liquid crystal display
LED	light emitting diode
LUKS	Linux unified key setup
MCU	microcontroller
mDNS	multicast domain name service (DNS)
MEMS	micro-electromechanical systems
MIPI	mobile industry processor interface
MISO	master-in, slave-out
MOSI	master-out, slave-in
MPU	microprocessor
MSM	mobile station modem, the APC processor and a modem.
MSRP	manufacturer's suggest retail price
OLED	organic light-emitting diode display
OTA	over the air updates
PCB	printed circuit board
PCBA	printed circuit board assembly (PCB with the components attached)
PCM	pulse-code modulation
PDM	pulse-density modulation
PMIC	power management IC
PNG	portable network graphics; an image file format
PWM	pulse width modulation
PVT	production validation test
QSN	Qualcomm serial number
ROI	region of interest
RPM	resource power management
RRT	rapidly-expanding random tree
SCLK	(I2C) serial clock

SDA	(I2C) serial data
SDK	software development kit
SLAM	simultaneous localization and mapping
SOC	system on a chip
SPAD	single photon avalanche diode
SPI	serial-peripheral interface
SSH	secure shell
SSID	service set identifier (the name of the Wifi network)
STM32	A microcontroller family from ST Microelectronics
SWD	single wire debug
SYSCON	system controller
TAR	tape archive file
TTS	text to speech
UART	universal asynchronous receiver/transmitter
USB	universal serial bus
UUID	universally unique identifier (effectively same as GUID)
vic	short for Victor (Vector's working name)
WEM	AudioKinetic Wwise Encoded Media. (a type of sound file)

Phrase	Description
<i>A*</i>	A path finding algorithm
<i>aboot</i>	The Android boot-loader used to launch Vector's linux system.
<i>accelerometer</i>	A sensor used to measure the angle of Vector's head, and acceleration (change in velocity).
<i>animation</i>	A scripted "sequence of highly coordinated movements, faces, lights, and sounds ... to demonstrate an emotion or reaction."
<i>animation trigger name</i>	An identifier of a group of related animations; Vector "pick[s] .. [an] actual animations to play based on Vector's mood or emotion, or with random weighting. Thus playing the same trigger twice may not result in the exact same underlying animation playing twice."
<i>attitude</i>	Vector's orientation, esp relative to the direction of travel
<i>autocorrelation</i>	A technique to find how repetitive a signal is; it works by finding how much one has to shift version of the signal before it (mostly) matches the original signal again.
<i>backpack board</i>	The circuit board in Vectors head with lights, push button, microphones and touch sensor
<i>beam forming</i>	A technique using multiple microphones to listen to a single speaker by selectively paying attention to sound only coming from that direction.
<i>behavior</i>	Behaviors represent a complex task [that] may include combinations of animation, path planning or other functionality. Examples include" driving to the charger, set the lift height, etc.

Table 507: Glossary
of common terms and
phrases

<i>behavior tree</i>	A decision tree that decides if a behavior can run or can no longer run, which related behaviors to start, and the parameter settings to run the behavior with.
<i>body board</i>	The circuit board in Vector's belly that manages the battery and drives the motors
<i>boot loader</i>	A piece of software used to load and launch the application software.
<i>C-like abstract data structure (CLAD)</i>	Anki's phrase for how they pack information into fields and values with a defined binary format. "Any data [passed] over the wire, [is] define[d with] enums, structures and messages in ".clad" files.. [with a] syntax [that] looks a lot like C structs. [A tool] auto-generate[s] Python, C++ and C# code for each of these structures, along with code to serialize and deserialize to efficiently packed byte streams of data." ⁶⁴ (FlatBuffers are used for the same purpose, but were not available when CLAD was developed.)
<i>capacitive touch</i>	A type of sensing where light contact, such as touch, is detected without requiring pressing a mechanism.
<i>cascade</i>	Applies a series of fast to compute filters and classifiers to detect low-level features and identify things like faces.
<i>cepstrum</i>	A way of using the frequency spectrum to analyze a voice.
<i>certificate</i>	Vector generates an SSL certificate that can be used for the secure communications.
<i>characteristic (Bluetooth LE)</i>	A key (or slot) that holds a value in the services key-value table. A characteristic is uniquely identified by its UUID.
<i>client token</i>	A string token provided by Vector that is passed with each SDK command.
<i>control</i>	Responsible for motors and forces to move where and how it is told to. (smooth arcs)
<i>cooldown</i>	A period of time after an action, animation, or behavior has run before it can be run again. see also <i>hold-off timer</i>
<i>D*-lite</i>	A path-finding algorithm
<i>decimation</i>	The amount (or ratio) that something is down sampled by.
<i>delocalization</i>	"Whenever Vector no longer knows where he is – e.g. when he's picked up," or falls.
<i>device mapper verity (dm-verity)</i>	A feature of the Linux kernel that checks the boot and RAM file systems for alteration, using signed keys
<i>electronic medical record</i>	A software record of Vector's serial number, model, lot code, manufacturing & test dates, and other information. This is stored in flash.
<i>electronic serial number</i>	Vector's serial number, but the copy that is in his <i>electronic medical record</i> .
<i>entitlement</i>	An entitlement is a family of features or resources that the program or owner is allowed to use.
<i>face detection</i>	The ability to realize that there is a face in the image, and where it is
<i>face recognition</i>	The ability to know the identity of a face seen.
<i>feature flags aka feature toggle</i>	A setting that enables and disables features, especially those still in development. This allows developing the code and integrating its structure before the module or function is completely ready. Otherwise it is very difficult to keep the different branches of development in sync and merge them when the feature is ready.
<i>field of view</i>	How wide of an area in the world that the camera can see
<i>firmware</i>	A type of software held (and usually executed from) in ROM or flash. It may have a (minimal) operating system, but often does not.

⁶⁴ <https://forums.anki.com/t/what-is-the-clad-tool/102/3>

<i>flash</i>	A type of persistent (non-volatile) storage media.
<i>guidance</i>	Builds the desired path
<i>gyroscope</i>	A sensor that is used to measure how fast Vector is turning (the angular velocity) along its x, y, and z axes.
<i>Haar feature</i>	Facial features picked out using Haar wavelets
<i>Haar wavelet</i>	A fast, low-cost that can be used to pick out (or recognize) simple features in an image.
<i>head board</i>	The circuit board in Vector's head with the main processor, WiFi (and Bluetooth LE), camera, speaker, etc.
<i>hold-off timer</i>	a timer that prevents another trigger or event for a period of time (after the previous one). <i>see also cool down.</i>
<i>hotword</i>	aka <i>wake word</i>
<i>inertial measurement unit</i>	The combination of an accelerometer and gyroscope to measure linear acceleration and rotational velocity.
<i>inner node</i>	A node in a tree data structure that does links to other nodes below it. Often it does not hold any other information.
<i>intent</i>	An intent is an internal code (and accompanying structure) that is used to represent the how to respond to the question or other phrases spoken by a person. It may represent the action requested, an answer to a query, or an action that emotionally responds to what was said.
<i>Kalman filter</i>	Used to merge two or more noisy signals together to estimate a proper signal.
<i>leaf node</i>	A node in a tree data structure that does not link to any other nodes below it. It holds the information that was being looked up.
<i>navigation</i>	Knowing where it is in the map
<i>nonce</i>	An initially random number, incremented after each use.
<i>odometry</i>	Estimate motion – displacement and rotation – from inertial measurement units and wheel & track rotation.
<i>path planning</i>	Forms smooth arcs and line segments to move in around an environment to avoid collisions, blocked paths, and cliffs. This is often used to navigate from point A to point B.
<i>playpen</i>	The playpen is a testing area with ramps, barriers, camera targets at a variety of angles, cube and a charging station.
<i>playpen test</i>	The playpen test is a check of the robot's sensors, camera calibration, motor function, microphone and a check over his overall functions. During testing, Vector's is checked for correct function: that he can correctly navigate, detect cliffs, see markers (getting their type and size correct), dock, and charge.
<i>pose</i>	The position and orientation of an object relative to a coordinate system
<i>power source</i>	Where the electric energy used to power Vector comes from.
<i>quad-tree</i>	A way of compressing a 2D map down into regions.
<i>rapidly-expanding random tree</i>	A path-finding algorithm
<i>recovery mode</i>	A separate, independent operating system that Vector can boot into for purposes of downloading software to replace a damaged partition.
<i>robot name</i>	Vector's robot name looks like "Vector-E5S6". It is "Vector-" followed by a 4 letters and numbers.
<i>serial number</i>	A unique number that is stamped on Vector's hardware (on the bottom) and in his <i>electronic medical record</i> .

<i>session token</i>	A string token provided by the Anki servers that is passed to Vector to authenticate with him and create a <i>client token</i> .
<i>simultaneous localization and mapping</i>	A vision-based technique for building a map of the immediate world for purposes of positioning oneself within it and detecting relative movements.
<i>service (Bluetooth LE)</i>	A key-value table grouped together for a common purpose. A service is uniquely identified by its UUID.
<i>software</i>	Software is distinct from firmware in that is often loaded from external storage to be run in RAM, and is based on dynamic linking, allowing the use of other (replaceable) software elements. It does not access hardware directly; instead it employs sophisticated features of the operating system.
<i>system controller (syscon)</i>	The name of the body-board microcontroller, and the firmware program running on it.
<i>tempo</i>	The pace of music, in beats per second
<i>text to speech</i>	A process of reading aloud a word, phrase, sentence, etc.
<i>trigger word</i>	aka <i>wake word</i>
<i>Trust Zone</i>	A security mode on ARM processor where privileged/special code is run. This includes access to encryption/decryption keys.
<i>universally unique identifier (UUID)</i>	A 128bit number that is unique. (effectively same as GUID)
<i>vocoder</i>	A sound effect that analyzes and transforms a voice; in this case to give Vector his unique vocal sound.
<i>wake word</i>	The phrase (“Hey, Vector”) used to activate Vector so that he will respond to spoken interaction.

Quotes are from Anki SDK.

APPENDIX B

Tool chain

This appendix tries to capture the tools that Anki is known or suspected to have used for the Anki Vector and its cloud server.

Table 508: Tools used by Anki

Tool	Description
<i>Acapela</i>	Vector uses Acapela's text to speech synthesizer, and the Ben voice. https://www.acapela-group.com/
<i>Advanced Linux Sound Architecture (alsa)</i>	The audio system https://www.alsa-project.org
<i>Amazon Alexa</i>	A set of software tools that allows Vector to integrate Alexa voice commands, probably in the AMAZONLITE distribution https://github.com/anki/avs-device-sdk https://developer.amazon.com/alexa-voice-service/sdk
<i>Amazon Simple Queue Service (SQS)</i>	Vector employs Amazon's SQS for its DAS functions.
<i>Amazon Simple Storage Service (S3)</i>	Vector's cloud interface uses Amazon's AWS go module to interact with Amazon's service: https://docs.aws.amazon.com/sdk-for-go/api/service/s3/ https://docs.aws.amazon.com/AmazonS3/latest/API/API_Operations_Amazon_Simple_Storage_Service.html
<i>Amazon Web services</i>	used on the server https://aws.amazon.com/
<i>android boot-loader</i>	Vector uses the Android Boot-loader; the code can be found in the earlier archive.
<i>ARM NN</i>	ARM's neural network support https://github.com/ARM-software/armnn
<i>AudioKinetic Wwise</i> ⁶⁵	Used to craft the parametric sound effects, and play pre-recorded effects. https://wwwaudiokinetic.com/products/wwise/
<i>Backtrace.io</i>	A service that receives uploaded minidumps from applications in the field and provides tools to analyze them. https://backtrace.io
<i>clang</i>	A C/C++ compiler, part of the LLVM family https://clang.llvm.org
<i>bluez v5</i>	Bluetooth LE support http://www.bluez.org/
<i>busybox</i>	The shell on the Anki Vector linux https://busybox.net

⁶⁵ <https://blog.audiokinetic.com/interactive-audio-brings-cozmo-to-life/>

<i>chromium update</i>	?
<i>civetweb</i>	The embedded webserver that allows Mobile apps and the python SDK to communicate with Vector. https://github.com/civetweb/civetweb
<i>connman</i>	Connection manager for WiFi https://01.org/connman
<i>Eigen</i>	A linear algebra library http://eigen.tuxfamily.org/
<i>gemmlowp</i>	A low-precision general matrix multiplication library https://github.com/google/gemmlowp
<i>GNU C Compiler (gcc)</i>	GCC version 4.9.3 was used to compile the kernel
<i>golang</i>	Go is used on the server applications, and (reported) some of Vector's internal software.
<i>Google Breakpad</i>	Google Breakpad is used to generate tracebacks and mini-dump files of programs that crash. Results are sent to http://backtrace.io https://chromium.googlesource.com/breakpad/breakpad
<i>Google FlatBuffers</i>	Google FlatBuffers is used to encode the animation data structures. "It is similar to protocol buffers, but the primary difference is that FlatBuffers does not need a parsing/unpacking step to a secondary representation before you can access data, often coupled with per-object memory allocation. Also, the code footprint of FlatBuffers is an order of magnitude smaller than protocol buffers" ⁶⁶ https://github.com/google/flatbuffers
<i>Google Protobuf</i>	Google's Protobuf interface-description language is used to describe the format/encoding of data sent over gRPC to and from Vector. This is used by mobile and python SDK, as well as on the server. https://developers.google.com/protocol-buffers
<i>Google RPC (gRPC)</i>	A "remote procedure call" standard, that allows mobile apps and the python SDK to communicate with Vector. https://grpc.io/docs/quickstart/cpp/
<i>hdr-histogram</i>	Unknown use https://github.com/HdrHistogram/HdrHistogram
<i>libsodium</i>	Cryptography library suitable for the small packet size in Bluetooth LE connections. Used to encrypt the mobile applications Bluetooth LE connection with Vector. https://github.com/jedisct1/libsodium
<i>linux, yocto</i> ⁶⁷	The family of linux distribution used for the Anki Vector (v3.18.66)
<i>linux</i>	on the server
<i>linux unified key storage (LUKS)</i>	
<i>Maya</i>	A character animation tool set, used to design the look and movements of Cozmo and Vector. The tool emitted the animation scripts.
<i>mpg123</i>	A MPEG audio decoder and player. This is needed by Alexa; other uses are unknown. https://www.mpg123.de/index.shtml
<i>ogg vorbis</i>	Audio codec https://xiph.org/vorbis
<i>Omron OKAO Vision</i>	Vector uses the Omron Okao Vision library for face recognition and tracking. https://plus-sensing.omron.com/technology/position/index.html
<i>open CV</i>	Used for the first-level image processing – to locate faces, hands, and possibly accessory symbols. https://opencv.org/
<i>openssl</i>	used to validate the software update signature

⁶⁶ <https://nlp.gitbook.io/book/tensorflow/tensorflow-lite>

⁶⁷ <https://www.designnews.com/electronics-test/lessons-after-failure-anki-robotics/140103493460822>

	https://www.openssl.org
<i>opkg</i>	Package manager, from yocto https://git.yoctoproject.org/cgit.cgi/opkg/
<i>Opus codec</i>	Audio codec; to encode speech sent to servers http://opus-codec.org/
<i>perl</i>	A programming language, on Victor https://www.perl.org
<i>Pretty Fast FFT pfffft</i>	Julien Pommier's FFT implementation for single precision, 1D signals https://bitbucket.org/jpommier/pfffft
<i>Pryon, Inc</i>	The recognition for the Alexa keyword at least the file system includes the same model as distributed in AMAZONLITE https://www.pryon.com/company/
<i>python</i>	A programming language and framework used with desktop tools to communicate with Vector. Vector has python installed. Probably used on the server as well. https://www.python.org
<i>Qualcomm</i>	Qualcomm's device drivers, camera support and other kit are used.
<i>Segger ICD</i>	A high-end ARM compatible in-circuit debugging probe. Rumoured to have been used by Anki engineers, probably with the STM32F030 https://www.segger.com/products/debug-probes/j-link/
<i>Sensory TrulyHandsFree</i>	Vectors recognition for "Hey Vector" and Alexa wake word is done by Sensory, Inc's TrulyHandsfree SDK 4.4.23 (c 2008) https://www.sensory.com/products/technologies/trulyhandsfree/ https://en.wikipedia.org/wiki/Sensory,_Inc.
<i>Signal Essence</i>	Designed the microphone array, and the low-level signal processing of audio input. https://signalessence.com/
<i>Sound Hound, inc Houndify</i>	Vector's Q&A "knowledge graph" is done by Sound Hound, using their Houndify product https://blog.soundhound.com/hey-vector-i-have-a-question-3c174ef226fb https://www.houndify.com/
<i>SQLite</i>	This is needed by Alexa; other uses are unknown https://www.sqlite.org/index.html
<i>systemd</i>	Used by Vector to launch the internal services https://www.freedesktop.org/software/systemd/
<i>tensor flow lite (TFLite)</i>	TensorFlow lite is used to recognize hands, the desk surface, and was intended to support recognizing pets and common objects. https://www.tensorflow.org/lite/microcontrollers/get_started

147. REFERENCES & RESOURCES

Several of the tools have licenses requiring Anki to post that the tools was listed and/or to post their versions of the tools, and their modification. The following archives of some of the open source tools are listed in the "acknowledgements" section of the mobile application:⁶⁸

<https://anki-vic-pubfiles.anki.com/license/prod/1.0.0/licences/OStarball.v160.tgz>
<https://anki-vic-pubfiles.anki.com/license/prod/1.0.0/licences/engineTarball.v160.tgz>

Note: Other open source tools may have been used by Anki were used without Anki posting their version (or modifications), and the licenses may not require them to.

⁶⁸ You can only read the acknowledgements in the mobile application if you are connected to a robot.

APPENDIX C

Alexa modules

This Appendix outlines the modules used by the Alexa client built into Vector (using the Alexa Client SDK). Alexa's modules connect together like so:

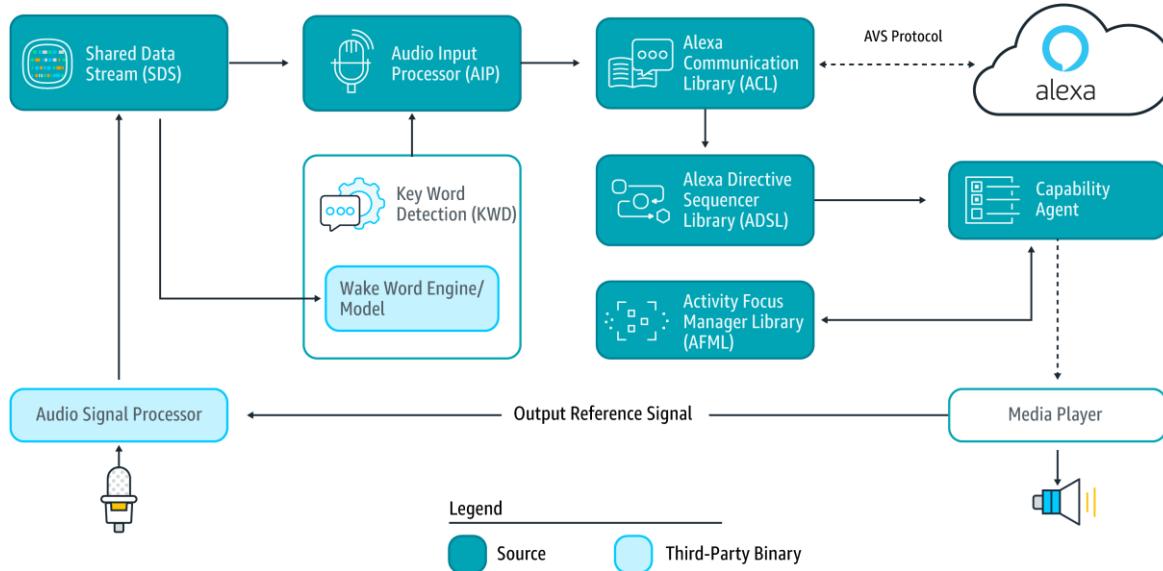


Figure 114: Alexa's function blocks
(image courtesy Amazon)

Alexa's modules include:

Library	Description & Notes
libACL.so	Alexa Communication Library. “Serves as the main communications channel between the device and the Alexa Voice Service.”
libAIP.so	Audio Input Processor. “Handles the audio input to Alexa Voice Service from on-device microphones, remote microphones and other audio input sources.”
libADSL.so	Alexa Directive Sequencer Library (Directive Router, Processor, Sequencer; Message Interpreter).
libAFML.so	Activity Focus Manager Library, including Audio Activity Tracker, Visual Activity tracker. “Prioritizes the channel inputs and outputs as specified by the AVS Interaction Model”
libAlerts.so	Alexa alert scheduler; “The interface for setting, stopping, and deleting timers and alarms.”
libAudioPlayer.so	Alexa’s audio player. “The interface for managing and controlling audio playback.”

Table 509: Alexa files

<code>libAudioResources.so</code>	Alexa's audio resources, including calls
<code>libAVSCommon.so</code>	Alexa's voice service support
<code>libAVSSystem.so</code>	Alexa's voice service support
<code>libCapabilitiesDelegate.so</code>	Alexa capabilities. "Handles Alexa-driven interactions; specifically, directives and events. Each capability agent corresponds to a specific interface exposed by the AVS API."
<code>libCBLAuthDelegate.so</code>	Alexa Authorization
<code>libCertifiedSender.so</code>	Alexa certified sender
<code>libContextManager.so</code>	Alexa's context manager
<code>libESP.so</code>	Alexa ESP, Dummy ESP
<code>libInteractionModel.so</code>	"This interface allows a client to support complex interactions initiated by Alexa, such as Alexa Routines."
<code>libNotifications.so</code>	Alexa Notifications. "The interface for displaying notifications indicators." Uses SQLite
<code>libPlaybackController.so</code>	"The interface for navigating a playback queue via GUI or buttons."
<code>libPlaylistParser.so</code>	Alexa playlist
<code>libRegistrationManager.so</code>	Alexa's registration manager
<code>libSettings.so</code>	Alexa's settings & preferences module
<code>libSpeakerManager.so</code>	
<code>libSpeechSynthesizer.so</code>	"The interface for Alexa speech output."

Note: quotes from Amazon Alexa Voice Services SDK documentation

APPENDIX D

Fault and status codes

The following are system status codes that may be produced during startup (*Quotes from “Anki Vector Error Codes”*):

Code	Meaning
1..10	Systemd failed...?
1	“Switchboard: unknown status”
2	“Switchboard: [Over the Air Update is] in progress”
3	“Switchboard: [Over the Air Update has] completed”
4	“Switchboard: rebooting”
5	“Switchboard: other [Over the Air Update] error”
10	“OS: Unknown system error”
001-099	Playpen
100-199	Error related to the body-board (syscon)
200-219	Software update status codes, see table below
220-299	Error codes in the range of 220-299 refer to problems from the software processes within Vector’s OS
300-799	Error codes in the range of 300-799 refer to problems expected during factory tests.
700	The robot was shutdown because the button was pressed.
701	The gyroscope sensor is out of range or failed (it wasn’t able to calibrate), so the robot shutdown.
702	The robot was shutdown because the battery voltage was too low.
703-704	Internal sensor out of range or failed.
705	The robot was shutdown because the battery was too hot to safely operate.
800-999	Error codes in the range of 800-999 refer to “power on self check” failures.
800	Vic-anim was unable to start or crashed.
801	The process to update the cube firmware failed.
840	The camera calibration is missing.
850	There is a problem with the cloud certificate
851	There is a problem with the cloud token store
852	There is a problem reading the cloud electronic serial number (ESN).
870	The front right MEMS microphone failed during power-on self test.
871	The front left MEMS microphone failed during power-on self test.
872	The back right MEMS microphone failed during power-on self test.
873	The back left MEMS microphone failed during power-on self test.
890	The front right cliff sensor failed during power-on self test.
891	The front left cliff sensor failed during power-on self test.
892	The back right cliff sensor failed during power-on self test.

Table 510: The system fault codes

893	The back left cliff sensor failed during power-on self test.
894	The time-of-flight distance sensor failed during power-on self test.
895	The touch sensor failed (during power-on self test?)
898	The main board is unable to communicate with the body-board. The cable between boards may be disconnected. This also appears as a software bug in version 1.6.
899	“No Body” This may mean that the main board is unable to communicate with the body-board. The cable between boards may be disconnected. This also appears as a software bug in version 1.6.
911	The audio system (hardware or software?) is not working properly.
913	Vic-switchboard was unable to start or crashed
914	Vic-engine was unable to start or crashed. “This was what zombie 915 was (915 on the screen, robot still drove around)”
915	Vic-engine stopped responding.
916	Vic-robot was unable to start or crashed
917	Vic-anim stopped responding (Or Vic-robot stopped responding.)
919	systemd is not working properly
920	Vic-gateway-cert was unable to generate a x509 certificate for vic-gateway
921	Vic-gateway was unable to start or crashed
923	Vic-cloud was unable to start or crashed
960	The IMU (accelerometer and gyroscope) has failed, or is not communicating properly.
919	
970	The WiFi hardware has failed.
980	“These codes indicate issues with the camera. These issues are typically caused by mm-anki-camera hanging when we try to stop the camera stream on vic-engine stop. We have to manually kill it and start it again.”
981	The camera stopped responding. “These codes indicate issues with the camera. These issues are typically caused by mm-anki-camera hanging when we try to stop the camera stream on vic-engine stop. We have to manually kill it and start it again.”
990	The LCD display is not communicating properly with the processor.

The following are the Playpen Failure codes. These overlap some of the error number ranges for other status codes:

Status	Meaning
0	Unknown: “Not possible”
1	Success: “Passed playpen”
2	There was a problem with one of the CLAD data structures. This error should not happen.
3	The lift or head failed. This error should not happen.
4	The “robot not detecting charger/ not charging. Make sure charger is plugged in; Check charge, contacts/charge circuit”
5	The charger is unavailable. This error should not happen.
6	The charger is not connected. This error should not happen.
7	The IMU is faulty or not communicating. The IMU is faulty. “Check/replace IMU. Maybe robot is shaking/moved while on charger.”
8	Still on charger. This error should not happen.
9	Failed to go to the calibration pose. This error should not happen.
10	The “robot saw a cliff and then stopped seeing it. Check cliff sensors; Check cliff slot in playpen; Check playpen surface for dirt”
11	The “robot detected a cliff” where there is none. “Check cliff sensors; Check playpen surface for dirt.”
12	The robot is not in the calibration pose. This error should not happen.
13	The calibration has failed. The “robot is seeing the calibration target but calibration [is] taking too long.” “Check calibration target; Check camera position and lens”
14	The “camera calibrated but the calibration is outside normal range. Check camera position and lens”
15	“Failed to write [camera calibration] data to [the] robot [non-volatile storage]. Should never happen”
16	“Failed to write [camera calibration image] data to [the] robot [non-volatile storage]. Should never happen”
17	“Failed to write [calibration pose] data to [the] robot [non-volatile storage]. Should never happen”
18	“Too many calibration images. Not possible.”
19	“Calibration pose failed. Not possible.”
20	“Read tool code failed. Not possible.”
21	“Tool code positions [out of range]. Not possible”
22	“Tool code write failed. Not possible.”
23	“Goto pre pickup pose action failed. Not possible.”
24	“Not in pre pickup pose. Not possible.”
25	“Not seeing cube to pickup. Check [that the] cube is in correct spot; Check [the] camera; Check [the] wheels if robot is not facing the cube”
26	“Seeing cube to pickup but robot thinks the cube is somewhere else. Check [that the] cube is in correct spot; Check [the] camera position.”
27	“Failed to pickup cube. Check [the] cube position; check [the] lift motor/gearbox.”
28	“Failed to place cube. Check [the] lift motor/gearbox.”

Table 511: Playpen Failure codes

29	“Unexpected observed object. Not possible.”
30	“Goto pre mount charger pose action failed. Not possible”
31	The charger was not found. “Not possible”
32	The charger dock failed. “Not possible”
33	Queue action failed. “Not possible.”
34	“A motor randomly calibrated. Check for issue with all motors: Sticky gearbox? Encoder problem?”
35	“Failed to write [test result] data to [the] robot [non-volatile storage]. Should never happen”
36	A test timed out. “Some step of playpen took too long and never completed. Try rerunning”
37	The test was cancelled. “Not possible”
38	The “robot detected being picked up. Was the robot picked up during the test or lifted off the ground in some way? Check cliff sensors; Check IMU”
39	“Tool code images write failed. Not possible.”
40	The “touch sensor readings [are] not stable, [too noisy]. Check touch sensor”
41	“Failed to write [cube pose] data to [the] robot [non-volatile storage]. Should never happen”
42	The “lift motor randomly calibrated. Check lift motor/gearbox”
43	The “head motor randomly calibrated. Check head motor/gearbox”
44	“Touch sensor readings [are] too small or [too] large. Check touch sensor”
45	The robot “did not pass all previous fixtures. Run robot through previous fixtures”
46	The robot has not been tested. “Not possible.”
47	The “head/lift motor failed to calibrate. Check head/lift motor/gearbox”
48	“Failed to write [birth certificate] data to [the] robot [non-volatile storage]. Should never happen”
49	“Failed to write data to [the] robot [non-volatile storage]. Should never happen”
50	“Failed to write data to [the] robot [non-volatile storage]. Should never happen”
51	There are “too many tool code images. Not possible.”
52	“Failed to write [calibration meta information] data to robot. [This] should never happen.”
53	“Failed to write [IMU] data to robot. [This] should never happen.”
54	“No [Bluetooth LE] advertising packet [was received] from a cube. Check battery of cube; Check [Bluetooth LE] radio on robot.”
55	“Touch sensor readings [standard deviation is] too large. Check [the] touch sensor.”
56	Failed to compute the camera pose. “Not possible.”
57	The camera pose is out of range. “Not possible.”
58	Failed to play sound. “Not possible.”
59	Was unable to read the results of the previous test. “Not possible.”
60	The wrong firmware version. “Not possible.”
61	A cliff sensor value is too high. “Not possible.”
62	A cliff sensor value is too low. “Not possible.”
63	The “robot did not backup straight while picking up the cube. Check wheel treads/motors/gearboxes”

-
- 64 The “battery is too low. Check battery; Leave robot on a charger for a couple of minutes to charge”
- 65 No IMU data. “Not possible.”
- 66 The “robot did not backup straight after picking up the cube. Check [the] wheel treads/motors/gearboxes”
- 67 The wrong body hardware version. This error should not happen.
- 68 There is no body hardware version information. This error should not happen.
- 69 The non-volatile storage erase operation failed. “Not possible.”
- 70 Was unable to parse the header of [??]. “Not possible.”
- 71 “No WiFi [access points were] found. Check [the] radio on [the] robot”
- 72 “Unknown body color. Not possible.”
- 73 “Failed to write data to robot. [This] should never happen”
- 74 The behavior is not runnable. “Software bug. Try rerunning.”
- 75 The “robot [is] not seeing [the] camera calibration target. Check camera position and lens”
- 76 The “robot detected being on a charger randomly. Check charge contacts/circuit”
- 77 “Head motor randomly disabled. Check head motor/gearbox”
- 78 “Lift motor randomly disabled. Check lift motor/gearbox”
- 79 “Some motor randomly disabled. Check all motors/gearboxes”
- 80 The “robot detected unexpected movement. [It] probably ran into something in playpen. Check wheels; Check IMU”
- 81 “Some action [that] the robot was trying to do failed.
- 82 The “robot failed to detect front cliffs. Check [the] front two cliff sensors.”
- 83 The “robot failed to detect back cliffs. Check [the] back two cliff sensors.”
- 84 The “robot failed to undetect front cliffs after detecting them. Check [the] front two cliff sensors.”
- 85 The “robot failed to undetect back cliffs after detecting them. Check [the] back two cliff sensors.”
- 86 The “robot thinks [that the] cube [is too] low in the ground. Check [the] cube position; Check camera position/rotation”
- 87 The “robot thinks [that the] cube [is too] high above the ground. Check [the] cube position; Check camera position/rotation”
- 88 The “camera calibrated but the calibration is outside normal range. Check camera position and lens.”
- 89 The “robot [is] not seeing [the] distance sensor marker. Check [the] robot[s] position at [the] time of failure, why wasn’t it seeing the marker. Check [the] distance sensor marker.”
- 90 The “robot [is] seeing [the] distance sensor marker [too] close or far away. [The] robot is too far or close to [the] distance sensor marker; check [the] wheels.”
- 91 The “front right [microphone is] not working. Check [the] front right [microphone].”
- 92 The “front left [microphone is] not working. Check [the] front left [microphone].”
- 93 The “back right [microphone is] not working. Check [the] back right [microphone].”
- 94 The “back left [microphone is] not working. Check [the] back left [microphone].”
- 95 “Either [the] speaker [is] not working or all [of the microphones are] not
-

-
- working. If [the] robot played [a] sound, check all [of the microphones, otherwise] check [the] speaker.”
- 96 “Never received FFT result from [the microphone] check.”
- 97 The “touch sensor reporting unexpected [out of range] values. Check [the] touch sensor.”
- 98 The time-of-flight “distance sensor reporting incorrect values. Check [the] distance sensor.”
- 99 The certificates were checked but have been found to be invalid. “Invalid [certificates] written by previous fixture. Run robot through previous fixtures”
-

The following are the update-engine status codes that may be produced during the update process:

Status	Meaning	Table 512: OTA update-engine status codes
200	The TAR contents did not follow the expected order.	
201	Unhandled section format for expansion, or The manifest version is not supported, or The OTA has the wrong number of images for the type, or The OTA is missing a BOOT or SYSTEM image, or The manifest configuration is not understood	
202	Could not mark target, a, or b slot unbootable, or Could not set target slot as active	
203	Unable to construct automatic update URL, or The URL for the update could not be opened	
204	The file (from the update URL) wasn't a valid TAR file, or is corrupt	
205	The compression scheme is not supported, or Decompression failed, the file may be corrupt	
206	"Block error" (<i>Note: this error code is not present in Vector's update software, and may be reserved.</i>)	
207	Delta payload error	
208	Couldn't sync OS images to disk, or Disk error while transferring OTA file.	
209	The manifest failed signature validation; or the aboot, boot image, system image, or delta.bin hash doesn't match signed manifest	
210	The encryption scheme is not supported.	
211	Vector's current version doesn't match the baseline for a delta update.	
212	The decompression engine had an unexpected, undefined error.	
213	The processor serial number (QSN) doesn't match the one in the manifest	
214	There is a mismatch: development Vectors can't install release OTA software, and release Vectors can't install development OTA software.	
215	OTA transfer failed, due to timeout. (There may be poor network connectivity)	
216	OS version name in the update file doesn't follow an acceptable pattern (the version suffixes – for production, release candidate, userdev, and development – must match the already installed software), or it is not allowed to upgrade or downgrade from the current version to the new version.	
219	Other unexpected, undefined error while transferring OTA file.	

148.

REFERENCES AND RESOURCES

Anki Vector Error Codes, 2020-2-26

<https://documents.project-victor.org/Release-Anki-Vector-error-codes-20200226.pdf>
<https://github.com/GooeyChickenman/victor/blob/master/documentation/Anki-Vector-error-codes%20-%2020200226.pdf>

APPENDIX E

File system

This Appendix describes the file systems on Vector's flash. As the Vector uses the Android bootloader, it reuses – or at least reserves – many of the Android partitions⁶⁹ and file systems. Many are probably not used. Quotes are from Android documentation.

The file system table tells us where they are stored in the partitions, and if they are non-volatile.

Mount point	Partition name	Description & Notes
/	BOOT_A	The primary linux kernel and initramfs
/data ⁷⁰	USERDATA	The data created for the specific robot (and user) that customizes it. A factory reset wipes out this user data. This portion of the file system is encrypted using “Linux Unified Key Setup” (LUKS).
/firmware	MODEM	The firmware for the WiFi/Bluetooth radio
/factory	OEM	Customizations, such as bootloader property values. ... Or the factory recovery?
/persist	PERSIST	Device specific “data which shouldn't be changed after the device is shipped, e.g. DRM related files, sensor reg file (sns.reg) and calibration data of chips; wifi, bluetooth, camera etc.”
/media/ram /run /var/volatile /dev/sm		Internal temporary file systems; holds temporary files, interprocess communication

Table 513: The file system mount table

The partition table⁷¹ found on the Vector:

Partition name	Size	Description & Notes
ABOOT	1 MB	The primary and backup Android boot loader, which may load the kernel,
ABOOTBAK	1 MB	recovery, or fastboot. This is in the format of a signed, statically linked ELF binary.
BOOT_A	32 MB	These are the primary and backup linux kernel and initramfs. Updates modify the
BOOT_B	32 MB	non-active partition, and then swap which one is active.
CONFIG	512 KB	This partition is not employed by Vector. It is zero'd out.
DDR	32 KB	Configuration of the DDR RAM.
DEVINFO	1 MB	This partition is not read by Vector. It is zero'd out. In typical aboot implementations this partition is used to hold “device information including: is_unlocked (aboot), is_tampered, is_verified, charger_screen_enabled, display_panel, bootloader_version, radio_version etc. Contents of this partition are displayed by “fastboot oem device-info” command in human readable format. Before loading boot.img or recovery.img, [the] boot

Table 514: The partition table
adapted from Melanie T

⁶⁹ <https://forum.xda-developers.com/android/general/info-android-device-partitions-basic-t3586565>

⁷⁰ This is mounted by “mount-data.service”. The file has a lot of information on how it unbricks

⁷¹ Much information from: <https://source.android.com/devices/bootloader/partitions-images>

		loader verifies the locked state from this partition.”
		Vector’s aboot will write to this partition to indicate tampering when it finds that the boot image does not pass integrity checks.
EMR	16 MB	This is Vectors “Electronic Medical Record.” It holds Vector’s Model, Serial Number, and such. It is a binary data structure, rather than a file system.
FSC	1KB	“Modem FileSystem Cookies”
FSG	1.5 MB	Golden backup copy of MODEMST1, used to restore it in the event of error
KEYSTORE	512 KB	“Related to [USERDATA] Full Disk Encryption (FDE)”
MISC	1MB	This is “a tiny partition used by recovery to communicate with bootloader store away some information about what it’s doing in case the device is restarted while the OTA package is being applied. It is a boot mode selector used to pass data among various stages of the boot chain (boot into recovery mode, fastboot etc.). e.g. if it is empty (all zero), system boots normally. If it contains recovery mode selector, system boots into recovery mode.”
MODEM	64 MB	Binary “blob” for the WiFi/Bluetooth radio firmware
MODEMST1	1.5MB	A FAT file-system holding executables and binary “blobs” for the WiFi/Bluetooth radio firmware.
MODEMST2	1.5MB	Several are signed by Anki. Includes a lot of test code, probably for emissions testing.
OEM	16MB	A modifiable ext2/4 file system that holds the logs, robot name, some calibration info, and SDK TLS certificates.
PAD	1MB	“related to OEM”
PERSIST	64MB	This partition is not employed by Vector. It is zero’d out.
RECOVERY	32 MB	An alternate partition holding kernel and initial RAM filesystem that allows the system boot into a mode that can download a new system. Often used to wipe out the updates.
RECOVERYFS	640 MB	An alternate partition holding systems applications and libraries that let the application boot into a mode that can download a new system. Often used to wipe out the updates. This partition holds v0.90 of the Anki software.
RPM	512KB	The primary and backup partitions for resource and power management. This is in the format of a signed, statically linked ELF binary.
RPMBAK	512KB	
SBL1	512KB	The primary and back up partitions for the secondary boot-loader. Responsible for loading aboot; has an “Emergency” download (EDL) mode using Qualcomm’s Sahara protocol.
SBL1BAK	512KB	This is in the format of a signed, statically linked ELF binary.
SEC	16KB	The secure boot fuse settings, OEM settings, signed-bootloader stuff
SSD	8KB	“Secure software download” for secure storage, encrypted RSA keys, etc
SYSTEM_A	896MB	The primary and backup system applications and libraries with application specific code.
SYSTEM_B	896MB	Updates modify the non-active partition, and then swap which one is active.
SWITCHBOARD	16 MB	This is a modifiable data area used by Vic-switchboard to hold persistent communication tokens. This appears to be a binary data structure, rather than a file system.
TZ	768KB	The primary and backup TrustZone. This is in the format of a signed, statically linked ELF binary.
TZBAK	768KB	This code is executed with special privileges to allow encrypting and decrypting key-value pairs without any other modules (or debuggers) having access to the secrets.
USERDATA	768MB	The data created for the specific robot (and user) that customizes it. A factory reset wipes out this user data. This partition is encrypted using “Linux Unified Key Setup” (LUKS).

The following files are employed in the Vector binaries and scripts:

File	Description	Table 515: Files
/anki/etc/revision	Contains the robot revision number	
/anki/etc/version	Contains the robot version number	
/data/data/com.anki.victor		
/data/data/com.anki.victor/cache/crashDumps	Holds the crash dump files	
/data/data/com.anki.victor/cache/outgoing		
/data/data/com.anki.victor/cache/vic-logmgr	A folder used to hold the log files while constructing the compressed archive file that will be uploaded.	
/data/diagnostics/	Used to hold the diagnostic logs as the archive is constructed and compressed.	
/data/etc		
/data/etc/localtime	The time zone	
/data/fault-reports		
/data/lib/connman/	The contents of /var/lib/connman are copied here	
/data/maintenance_reboot	This is set when the system has rebooted for maintenance reasons (e.g. updates)	
/data/misc/bluetooth	A folder to hold communication structures for the Bluetooth LE stack.	
/data/misc/bluetooth/abtd.socket	The IPC socket interface to Anki's Bluetooth LE service	
/data/misc/bluetooth/btprop	The IPC socket interface to BlueZ Bluetooth LE service.	
/data/misc/camera		
/data/panics		
/data/run/connman		
/data/data/com.anki.victor/persistent/switchboard_sessions	Used by Vic-switchboard to hold persistent session information, e.g. tokens	
/data/unbrick		
/data/usb		
/data/vic-gateway		
/dev/block/bootdevice/by-name/emr	File system access to the manufacturing records, including serial number	
/dev/block/bootdevice/by-name/switchboard	File system access to switchboards persistent data.	
/dev/rampost_error	The status of the rampost checks of the body board.	
/dev/socket/_anim_robot_server_	The IPC socket with Vector's animation controller	
/dev/socket/_engine_gateway_server_	The IPC socket interface to Vector's Gateway [TBD] server	
/dev/socket/_engine_gateway_proto_server_	The IPC socket interface to Vector's Gateway [TBD] server	
/dev/socket/_engine_switch_server_	The IPC socket interface to Vector's Switchbox [TBD] server	
/etc/os-version	Contains the OS (linux) version string.	
/etc/os-version-rev	Contains the OS (linux) revision string.	
/factory/cloud/something.pem		
/proc/sys/kernel/random/boot_id	A random identifier, created each boot	

<code>/sys/devices/system/cpu/possible</code> ⁷²	The number of CPUs and whether they can be used.
<code>/sys/devices/system/cpu/present</code>	
<code>/run/after_maintenance_reboot</code>	This is set to indicate to Vectors services that the system was rebooted for maintenance reasons, and they should take appropriate action. This will be set, on boot, if <code>/data/maintenance_reboot</code> had been set.
<code>/run/anki-crash-log</code>	
<code>/run/das_allow_upload</code>	If this exists, the crash log files can be uploaded to the backtrace.io servers; if it does not exist, the files are not uploaded. This file probably always exists, but was intended to be a user settable feature.
<code>/run/fake-hwclock-cmd</code> ⁷³	Sets the fake time to the time file (Vector doesn't have a clock)
<code>/run/fault_code</code>	This is set to the fault code (see Appendix C) if a program is unable to carry out a significant task, or crashes. The fault display program may present this code on the LCD display.
<code>/run/fault_code.pending</code>	The next fault code in queue to be handled
<code>/run/fault_code.showing</code>	The fault code being displayed
<code>/run/fault_restart_count</code>	This is incremented with each restart, and cleared by a reboot.
<code>/run/fault_restart_uptime</code>	
<code>/tmp/data_cleared</code>	
<code>/tmp/vision/neural_nets</code>	

Key named device files employed in Vector binaries:

File	Description
<code>/dev/fb0</code>	The display framebuffer
<code>/dev/spidev0.0</code>	The SPI channel to communicate with the IMU
<code>/dev/spidev1.0</code>	The SPI channel to communicate with the LCD
<code>/dev/ttyHS0</code>	Serial connection with the body-board
<code>/dev/ttyHSL0</code>	Console log
<code>/sys/class/android_usb/android0/iSerial</code>	Set to Vector's serial number
<code>/sys/class/gpio/gpio83</code>	Used to control the camera power
<code>/sys/class/leds/face-backlight-left/brightness</code>	LCD left backlight control
<code>/sys/class/leds/face-backlight-right/brightness</code>	LCD right backlight control
<code>/sys/devices/platform/soc/1000000.pinctrl/gpio/gpiochip0/base</code>	LCD backlight enable (left or right?) GPIO config
<code>/sys/devices/system/cpu/cpu0/cpufreq/scaling_max_freq</code>	The maximum frequency that the CPU can run at. Initially set to 533MHz
<code>/sys/kernel/debug/msm_otg/bus_voting</code>	Disabled to prevent the USB from pinning RAM to 400MHz.
<code>/sys/kernel/debug/rpm_send_msg/message</code>	Used to control the RAM controller. The RAM is set to a maximum of 400MHz.
<code>/sys/devices/soc/1000000.pinctrl/gpio/gpiochip0/base</code>	LCD backlight enable (left or right?) GPIO config

Table 516: Named device and control files

⁷² <https://www.kernel.org/doc/Documentation/ABI/testing/sysfs-devices-system-cpu>

⁷³ <https://manpages.debian.org/jessie/fake-hwclock/fake-hwclock.8.en.html>

`/sys/devices/soc.0/1000000.pinctrl/gpio/gpiochip911/base` LCD backlight enable (left or right?) GPIO config

`/sys/module/spidev/parameters/bufsiz` The buffer size for SPI transfers. This is set to the size of the LCD frame (184 pixels × 96 pixels × 2 bytes/pixel).

APPENDIX F

Bluetooth LE Services & Characteristics

This Appendix describes the configuration of the Bluetooth LE services – and the data access they provide – for the accessory cube and for Vector.

149. CUBE SERVICES

The basic Bluetooth LE services:

Service	UUID ⁷⁴	Description & Notes
Device Info Service ⁷⁵	180A ₁₆	Provides device and unit specific info –it's manufacturer, model number, hardware and firmware versions
Generic Access Profile ⁷⁶	1800 ₁₆	The device name, and preferred connection parameters
Generic Attribute Transport ⁷⁷	1801 ₁₆	Provides access to the services.
Cube's Service	C6F6C70F-D219-598B-FB4C-308E1F22F830 ₁₆	Service custom to the cube, reporting battery, accelerometer and date of manufacture

Table 517: The Bluetooth LE services

Note: It appears that there isn't a battery service on the Cube. When in over-the-air update mode, there may be other services present (i.e. by a bootloader)

Element	Value
Device Name (Default)	"Vector Cube"
Firmware Revision	"v_5.0.4"
Manufacturer Name	"Anki"
Model Number	"Production"
Software Revision	"2.0.0"

Table 518: The Cube's Device info settings

⁷⁴ All values are a little endian, per the Bluetooth 4.0 GATT specification

⁷⁵ http://developer.bluetooth.org/gatt/services/Pages/ServiceViewer.aspx?u=org.bluetooth.service.device_information.xml

⁷⁶ http://developer.bluetooth.org/gatt/services/Pages/ServiceViewer.aspx?u=org.bluetooth.service.generic_access.xml

⁷⁷ http://developer.bluetooth.org/gatt/services/Pages/ServiceViewer.aspx?u=org.bluetooth.service.generic_attribute.xml

149.1. CUBE'S ACCELEROMETER SERVICE

Values are little-endian, except where otherwise stated.

UUID	Access	Size	Notes	
0EA75290-6759-A58D-7948-598C4E02D94A ₁₆	Write	unknown		Table 519: Cube's accelerometer service characteristics
450AA175-8D85-16A6-9148-D50E2EB7B79E ₁₆	Read		The date and time of manufacture (?) char[] <i>A date and time string</i>	
43EF14AF-5FB1-7B81-3647-2A9477824CAB ₁₆	Read, Notify, Indicate		Reads the battery and accelerometer uint16_t <i>battery ADC value</i> uint16_t <i>accelerometer X ADC value #1</i> uint16_t <i>accelerometer Y ADC value #1</i> uint16_t <i>accelerometer Z ADC value #1</i> uint16_t <i>accelerometer X ADC value #2</i> uint16_t <i>accelerometer Y ADC value #2</i> uint16_t <i>accelerometer Z ADC value #2</i> uint16_t <i>accelerometer X ADC value #3</i> uint16_t <i>accelerometer Y ADC value #3</i> uint16_t <i>accelerometer Z ADC value #3</i>	
9590BA9C-5140-92B5-1844-5F9D681557A4 ₁₆	Write		Unknown	

Presumably some of these will cause the Cube to go into over the air update (OTAU) mode, allowing its firmware to be updated.

Others turn the RGB on to an RGB color, possibly duty cycle and pulsing duty cycle

150. VECTOR SERVICES SERVICE

times and other feature parameters:

Service	UUID ⁷⁸	Description & Notes	
Generic Access Profile	1800 ₁₆	The device name, and preferred connection parameters	
Generic Attribute Transport	1801 ₁₆	Provides access to the services.	
Vector's Serial Service	FEE3 ₁₆	The service with which we can talk to Vector.	

It appears that there isn't a battery service on the Vector.

Element	Value	
Device Name (Default)	"Vector" followed by his serial number	Table 521: The Vector's Device info settings

⁷⁸ All values are a little endian, per the Bluetooth 4.0 GATT specification

150.1. VECTOR'S SERIAL SERVICE

UUID	Access	Format	Notes
30619F2D-0F54-41BD-A65A-7588D8C85B45 ₁₆	Read, Notify,Indicate		
7D2A4BDA-D29B-4152-B725-2491478C5CD7 ₁₆	write		

Table 522: Vector's serial service characteristics

APPENDIX G

Servers & Data Schema

This Appendix describes the servers that Vector contacts⁷⁹

Server	Description & Notes
chipper.api.anki.com:443	The speech recognition engine lives here
conncheck.global.anki-services.com/ok	Used to check to see if it can connect to Anki
jdocs.api.anki.com:443	Storage of some sort of data. Name, faces, prefs?
token.api.anki.com:443	Used to get the API certificate. ⁸⁰
https://anki.sp.backtrace.io:6098/post?format=minidump&token=6fd2bd053e8dd542ee97c05903b1ea068f090d37c7f6bbfa873c5f3b9c40b1d9	Vector posts crashes (linux minidumps) to this server. This is hard coded in anki-crashuploader
https://sns.us-west-2.amazonaws.com/792379844846/DasProd-dasprodSqs-1845FTIME3RHN	This is used to synchronize with data analytics services.
https://ota.global.anki-services.com/vic/prod/	Where Vector checks for updates
https://ota.global.anki-dev-services.com/vic/rc/lo8awreh23498sf/	For the Developer branch
amazon.com/code	

The mobile application contacts the following servers:

Server	Description & Notes
https://locations.api.anki.com/1/locations	This is used to provide a list of locations to the mobile application that the Chipper servers will recognize. Without this, you cannot change Vector's location in the mobile application

Table 523: The servers that Vector contacts.

Table 524: The servers that the mobile application contacts.

⁷⁹ Todo: sync up with info at: <https://github.com/anki-community/vector-archive>

⁸⁰ Project Victor had a write up, reference that.

The Alexa modules contact the following servers:

Server	Description & Notes	Table 525: The Amazon Alexa Voice Service servers that Vector contacts.
https://api.amazon.com/auth/O2/	Used to authenticate the account for the Alexa device.	
https://avs-alexa-na.amazon.com	The Alexa Voice Service that accepts the spoken audio and returns a rich intent. Amazon changed preferred URLs on 2019 May 22, and this is considered legacy. ⁸¹	

⁸¹ <https://developer.amazon.com/docs/alexa-voice-service/api-overview.html>

APPENDIX H

Features

The following is the set of application-level feature flags and whether they are enabled (i.e. sufficiently developed to be used) in Vector:

Feature	enabled	Description & Notes
<i>ActiveIntentFeedback</i>	true	
<i>Alexa</i>	true	The ability to use Alexa
<i>Alexa_AU</i>	true	The ability to use Alexa, localized for Australia
<i>Alexa_UK</i>	true	The ability to use Alexa, localized for the UK
<i>AttentionTransfer</i>	false	
<i>CubeSpinner</i>	false	
<i>Dancing</i>	true	The ability for Vector to dance to music.
<i>Exploring</i>	true	The ability for Vector to explore his area
<i>EyeColorVC</i>	true	The ability to set Vector's eye color through a voice command
<i>FetchCube</i>	true	The ability for Vector to fetch his cube
<i>FindCube</i>	true	The ability for Vector to find his cube
<i>GazeDirection</i>	false	
<i>GreetAfterLongTime</i>	true	
<i>HandDetection</i>	true	The ability for Vector to spot hands
<i>HeldInPalm</i>	true	
<i>HowOldAreYou</i>	true	The ability for Vector to track how long it has been since he was activated (his age) and use that info to respond to the question "How old are you?"
<i>Invalid</i>	false	
<i>Keepaway</i>	true	
<i>KnowledgeGraph</i>	true	The ability for Vector to answer a question when asked "Hey Vector, I have a question..."
<i>Laser</i>	false	
<i>Messaging</i>	false	
<i>MoveCube</i>	true	
<i>PopAWheelie</i>	true	The ability for Vector pop a wheelie using his cube
<i>PRDemo</i>	false	
<i>ReactToHeldCube</i>	true	
<i>ReactToIllumination</i>	true	
<i>RollCube</i>	true	The ability for Vector to drive up and roll his cube

Table 526: The features

<i>StayOnChargerUntilCharged</i>	true	
<i>TestFeature</i>	false	
<i>Volume</i>	true	The ability to set Vector's volume by voice command.

The following is the set of AI features (related to, but the same as the feature flags), which identify an active behavior:

Table 527: The AI behaviour features

Feature	Description & Notes
<i>Alexa</i>	This behavior is used to perform Alexa-based interaction.
<i>AskForHelp</i>	? Is Vector asking for help?
<i>BasicVoiceCommand</i>	Vector is responding to a wake word and intent.
<i>BeQuiet</i>	Vector is responding to the intent TBD to be quiet (make no sounds and not move).
<i>Blackjack</i>	Vector is playing a game of Blackjack.
<i>CantDoThat</i>	
<i>ComeHere</i>	Vector is going to the speaker.
<i>CubeSpinner</i>	
<i>DanceToTheBeat</i>	Vector is dancing to music.
<i>Exploring</i>	Vector is driving and exploring his area.
<i>FetchCube</i>	Vector is fetching his cube.
<i>FindCube</i>	Vector is looking for his cube. This is done prior to fetching a cube, if Vector doesn't know where it is.
<i>FindHome</i>	Vector is looking for his home (the charger). This is done if Vector needs to charger, and doesn't know where the charger is.
<i>FistBump</i>	If Vector has received a fist bump. Note this can be a result of the shaking of lifting the cube, driving with the cube, or putting it down.
<i>Frustrated</i>	Vector is frustrated and throwing a little tantrum.
<i>GoHome</i>	Vector is driving home to his charger, often in response to a low battery.
<i>HeldInPalm</i>	Vector is held in the palm of a hand; he may coo or throw a little tantrum.
<i>HowOldAreYou</i>	Vector has been asked how long it has been since he was activated (his age) and telling his human.
<i>InteractWithFaces</i>	
<i>InTheAir</i>	Vector has detected that his in the air. If he thinks he is falling, he may engage in "tuck and roll" where lowers his lift, and tilts his head down.
<i>KeepAway</i>	
<i>KnowledgeGraph</i>	Vector has been ask to answer a question ("Hey Vector, I have a question...") and this behaviour is used to perform the rest of the interaction.
<i>ListeningForBeats</i>	Vector thinks that music may be playing and is listening for the beat of the music to dance to. (He may follow this with the <i>DanceToTheBeat</i> feature).
<i>LookAtMe</i>	Vector is looking for a person face, to look into their gaze.

<i>LowBattery</i>	Vector's battery level is low and he needs to begin looking for the charger.
<i>MeetVictor</i>	Vector is performing the on-boarding steps.
<i>MoveCube</i>	
<i>MovementBackward</i>	
<i>MovementForward</i>	
<i>MovementLeft</i>	
<i>MovementRight</i>	
<i>MovementTurnAround</i>	
<i>NoFeature</i>	When Vector's mind isn't doing anything and his mind is blank... he'll probably pick exploring, observing, or sleeping as his next activity.
<i>Observing</i>	Vector is looking around.
<i>ObservingOnCharger</i>	Vector is looking around while on his charger.
<i>Petting</i>	Vector is being petting.
<i>PlayingMessage</i>	<i>The messaging features are not yet support.</i>
<i>PopAWheelie</i>	Vector is attempting to pop a wheelie using his cube.
<i>ReactToAbuse</i>	Vector is responding to verbally abusive statements (represented as an intent).
<i>ReactToAffirmative</i>	Vector is responding to verbal complements (represented as an intent).
<i>ReactToApology</i>	Vector is responding to an apology (represented as an intent).
<i>ReactToCliff</i>	Vector has detected a cliff while driving, and is reacting to it.
<i>ReactToGazeDirection</i>	Vector has detected a face looking at him (the gaze) and is reacting to it.
<i>ReactToGoodBye</i>	Vector is responding to a verbal goodbye (represented as an intent).
<i>ReactToGoodMorning</i>	Vector is responding to a verbal good morning (represented as an intent).
<i>ReactToHand</i>	Vector has seen a hand and is reacting to it.
<i>ReactToHello</i>	Vector is responding to a verbal hello (represented as an intent).
<i>ReactToLove</i>	Vector is responding to a verbal statement of affection (represented as an intent).
<i>ReactToNegative</i>	Vector is responding to verbal abuse.
<i>ReactToRobotOnSide</i>	Vector has fallen (possibly from driving off the edge of his area) and is on his side.
<i>RecordingMessage</i>	<i>The messaging features are not yet support.</i>
<i>RequestCharger</i>	Vector is asking his human to help him by putting him on his charger. This happens if Vector can't get to his charger – he is stuck or doesn't know where it is.
<i>RobotShaken</i>	Vector has detected being shaken, like a snow globe
<i>RollBlock</i>	The ability for Vector to drive up and roll his cube
<i>SDK</i>	
<i>SeasonalHappyHoliday</i>	Vector is animating a little celebration video suitable for Christmas and other holidays.

<i>SeasonalHappyNewYear</i>	Vector is animating a little celebration video suitable for New Years.
<i>ShutUp</i>	
<i>Sleeping</i>	Vector is sleeping, usually on his charging, and waiting for stimulation.
<i>StuckOnEdge</i>	Vector has driven at least one of
<i>TakeAPhoto</i>	Vector is taking a photo.
<i>TimerCanceled</i>	Vector is cancelling the timer, as part of the timer behavior
<i>TimerChecked</i>	Vector is answering “how long is left on the timer”, as part of the timer behavior.
<i>TimerReminder</i>	
<i>TimerRinging</i>	Vector is playing the timer ring (i.e. the timer has expired) animation as part of the timer behavior.
<i>TimerSet</i>	
<i>UnmatchedVoiceIntent</i>	The cloud wasn’t able to identify an intent based on what was said (if anything) after the <i>Hey Vector</i> wake word.
<i>VolumeAdjustment</i>	Vector’s volume was adjusted by a voice command.
<i>Weather</i>	Vector is looking up the weather (from the cloud) and animating the results.
<i>WhatsMyName</i>	Vector is looking for a face and identifying it.

APPENDIX I

Phrases and their Intent

This Appendix maps the published phrases that Vector responds to and their intent:

Intent	Enumeration	Phrase
<i>movement_backward</i>	23	Back up
<i>imperative_scold</i>	18	Bad robot
<i>imperative_quiet</i>		Be quiet
<i>global_stop</i>	3	Cancel the timer
		Change/set your eye color to [blue, green, lime, orange, purple, sapphire, teal, yellow].
<i>check_timer</i>	1	Check the timer
<i>imperative_come</i>	10	Come here
<i>imperative_dance</i>	11	Dance.
<i>play_popawheelie</i>	34	Do a wheelstand
<i>imperative_fetchcube</i>	12	Fetch your cube
<i>imperative_findcube</i>	13	Find your cube
<i>play_fistbump</i>	32	Fist Bump
<i>play_fistbump</i>	32	Give me a Fist Bump
<i>movement_backward</i>	23	Go backward
<i>explore_start</i>	2	Go explore
<i>movement_forward</i>	22	Go forward.
<i>movement_turnleft</i>	24	Go left
<i>movement_turnright</i>	25	Go right
<i>system_sleep</i>		Go to sleep
<i>system_charger</i>		Go to your charger
		Good afternoon
<i>greeting_goodbye</i>	4	Goodbye
		Good evening
<i>greeting_goodnight</i>		Good night
<i>greeting_goodmorning</i>	5	Good morning
<i>imperative_praise</i>	16	Good robot

Table 528: The “Hey Vector” phrases

<i>seasonal_happyholidays</i>	36	Happy Holidays
<i>seasonal_happynewyear</i>	37	Happy New Year
<i>greeting_hello</i>	6	Hello He's behind you
<i>character_age</i>	0	How old are you
<i>imperative_abuse</i>	7	I hate you.
<i>knowledge_question</i>	27	I have a question ...
<i>imperative_love</i>	15	I love you.
<i>imperative_apology</i>	9	I'm sorry.
<i>play_blackjack</i>	31	Let's play Blackjack Listen to music
<i>imperative_lookatme</i>	14	Look at me Look behind you My name is [Your Name]
<i>imperative_negative</i>	17	No
<i>play_pickupcube</i>	33	Pick up your cube.
<i>play_anygame</i>	29	Play a game
<i>play_anymove</i>	30	Play a move
<i>play_blackjack</i>	31	Play Blackjack
<i>play_popawheelie</i>	34	Pop a wheelie.
<i>play_rollcube</i>	35	Roll your Cube
<i>imperative_quiet</i>		Quiet down Run
<i>set_timer</i>	38	Set a timer for [length of time]
<i>imperative_shutup</i>		Shut up
<i>explore_start</i>	2	Start Exploring Stop Exploring
<i>global_stop</i>	3	Stop the timer
<i>take_a_photo</i>	40	Take a picture of [me/us]
<i>take_a_picture</i>	40	Take a picture
<i>take_a_selfie</i>	40	Take a selfie
<i>movement_turnaround</i>	26	Turn around
<i>movement_turnleft</i>	24	Turn left Turn off
{same as be quiet }		Turn off
<i>movement_turnright</i>	25	Turn right
<i>imperative_volumelvel</i>	19	Volume [number].
<i>imperative_volumedown</i>	21	Volume down
<i>imperative_volum eup</i>	20	Volume up. Volume maximum

<i>names_ask</i>	28	What's my name?
<i>weather_response</i>	41	What's the weather in [City Name]?
<i>weather_response</i>	41	What's the weather report?
<i>show_clock</i>	39	What time is it?
<i>blackjack_hit</i>		
<i>blackjack_playagain</i>		
<i>blackjack_stand</i>		
<i>global_delete</i>		
<i>imperative_lookoverthere</i>		
<i>knowledge_response</i>		
<i>knowledge_unknown</i>		
<i>meet_victor</i>		
<i>message_playback</i>		
<i>message_record</i>		
<i>silence</i>		
<i>status_feeling</i>		
<i>imperative_affirmative</i>	8	Yes

Note: Vector's NLP server doesn't recognize "home" ..

Questions

Subject	Example Phrase
<i>Current conversion</i>	What's 1000 Yen in US Dollars?
<i>Flight status</i>	What is the status of American Airlines Flight 100?
<i>Equation solver</i>	What is the square root of 144?
<i>General knowledge</i>	What is the tallest building?
<i>places</i>	What is the distance between London and New York?
<i>People</i>	Who is Jarvis?
<i>Nutrition</i>	How many calories are in an avocado?
<i>Sports</i>	Who won the World Series?
<i>Stock market</i>	How is the stock market?
<i>Time zone</i>	What time is it in Hong Kong?
<i>Unit conversion</i>	How fast is a knot?
<i>Word definition</i>	What is the definition of Artificial Intelligence?

Table 529: The Vector questions phrases

User Intent	Cloud Intent	App Intent	Feature Flag
<i>amazon_signin</i>	intent_amazon_signin		
<i>amazon_signout</i>	intent_amazon_signout		
<i>blackjack_hit</i>	intent_blackjack_hit		

Table 530: Mapping of different intent names

<i>blackjack_playagain</i>	intent_blackjack_playagain		
<i>blackjack_stand</i>	intent_blackjack_stand		
<i>character_age</i>	intent_character_age		HowOldAreYou
<i>check_timer</i>	intent_clock_checktimer		
<i>explore_start</i>	intent_explore_start	explore_start	Exploring
<i>global_delete</i>	intent_global_delete_extend		
<i>global_stop</i>	intent_global_stop_extend		
<i>greeting_goodbye</i>	intent_greeting_goodbye		
<i>greeting_hello</i>	intent_greeting_hello		
<i>greeting_goodmorning</i>	intent_greeting_goodmorning		
<i>greeting_goodnight</i>	intent_greeting_goodnight		
<i>imperative_abuse</i>	intent_imperative_abuse		
<i>imperative_affirmative</i>	intent_imperative_affirmative		
<i>imperative_apology</i>	intent_imperative_apologize		
<i>imperative_come</i>	intent_imperative_come	intent_imperative_come	
<i>imperative_dance</i>	intent_imperative_dance	intent_imperative_dance	
<i>imperative_eyecolor</i>	intent_imperative_eyecolor		EyeColorVC
<i>imperative_eyecolor_specific</i>	intent_imperative_eyecolor_extend		EyeColorVC
<i>imperative_fetchcube</i>	intent_imperative_fetchcube	intent_imperative_fetchcube	FetchCube
<i>imperative_findcube</i>	intent_imperative_findcube	intent_imperative_findcube	FindCube
<i>imperative_lookatme</i>	intent_imperative_lookatme	intent_imperative_lookatme	
<i>imperative_lookingoverthere</i>	intent_imperative_lookingoverthere	intent_imperative_lookingoverthere	GazeDirection
<i>imperative_love</i>	intent_imperative_love		
<i>imperative_negative</i>	intent_imperative_negative		
<i>imperative_praise</i>	intent_imperative_praise		
<i>imperative_scold</i>	intent_imperative_scold		
<i>imperative_quiet</i>	intent_imperative_quiet	intent_imperative_quiet	
<i>imperative_shutup</i>	intent_imperative_shutup	intent_imperative_shutup	
<i>imperative_volumedown</i>	intent_imperative_volumedown		Volume
<i>imperative_volumolevel</i>	intent_imperative_volumolevel_extend		Volume
<i>imperative_volumeup</i>	intent_imperative_volumeup		Volume
<i>knowledge_question</i>	intent_knowledge_promptquestion	knowledge_question	KnowledgeGraph
<i>knowledge_response</i>	intent_knowledge_response_extend	knowledge_response	KnowledgeGraph
<i>knowledge_unknown</i>	intent_knowledge_no_response	knowledge_unknown	KnowledgeGraph
<i>meet_victor</i>	intent_names_username_extend	intent_meet_victor	
<i>message_playback</i>	intent_message_playmessage_extend	intent_message_playmessage	Messaging
<i>message_record</i>	intent_message_recordmessage_extend	intent_message_recordmessage	Messaging
<i>movement_backward</i>	intent_imperative_backup		
<i>movement_forward</i>	intent_imperative_forward		
<i>movement_turnaround</i>	intent_imperative_turnaround		
<i>movement_turnleft</i>	intent_imperative_turnleft		
<i>movement_turnright</i>	intent_imperative_turnright		
<i>names_ask</i>	intent_names_ask	intent_names_ask	

<i>play_anygame</i>	intent_play_anygame	
<i>play_antrick</i>	intent_play_antrick	
<i>play_blackjack</i>	?play_blackjack	
<i>play_fistbump</i>	?play_fistbump	
<i>play_pickupcube</i>	?play_pickupcube	
<i>play_popawheelie</i>	?play_popawheelie	
<i>play_rollcube</i>	?play_rollcube	
<i>play_specific</i>	intent_play_specific_extend	intent_play_specific
<i>seasonal_happyholidays</i>	intent_seasonal_happyholidays	
<i>seasonal_happynewyear</i>	intent_seasonal_happynewyear	
<i>set_timer</i>	intent_clock_settimer_extend	intent_clock_settimer
<i>show_clock</i>	intent_clock_time	
<i>silence</i>	intent_system_noaudio	
<i>status_feeling</i>	intent_status_feeling	
<i>system_charger</i>	intent_system_charger	intent_system_charger
<i>system_sleep</i>	intent_system_sleep	intent_system_sleep
<i>take_a_photo</i>	intent_photo_take_extend	
<i>unmatched_intent</i>		
<i>weather_response</i>	intent_weather_extend	

APPENDIX J

Emotion Events

The following is the set of emotion names used by Vector's mood manager. Some are from external events. Many whether or not a behavior or action succeeded, or failed (failed with retry, failed with abort).

	Emotion Name	Description and notes	<i>Table 531: The emotion event names</i>
<i>Ambient light</i>	ReactToDark		
<i>Charger</i>	DriveOffCharger	Vector was able to drive off of the charger.	
	MountChargerSuccess	Vector was able to drive onto of the charger successfully.	
	PlacedOnCharger	Vector was place on the charger.	
<i>Cube</i>	CubeSpinner		
	KeepawayPounce		
	KeepawayStarted		
	PickingOrPlacingActionFailedWithAbort	Vector was unable to pick up his cube or to place it successfully, and the action was aborted.	
	PickingOrPlacingActionFailedWithRetry	Vector was unable to pick up his cube or to place it successfully, even with retries.	
	PickupSucceeded	Vector picked up his cube successfully.	
	RollSucceeded	Vector rolled his cube successfully.	
<i>Driving</i>	CliffReact	Vector encountered a cliff and reacted.	
	DrivingActionFailedWithAbort	Vector was unable to drive to his target successfully, and the action was aborted.	
	DrivingActionFailedWithRetry	Vector was unable to drive to his target successfully, even with retries.	
	DrivingActionSucceeded	Vector was able to drive to his target successfully.	
	ExploringExamineObstacle	Vector found an obstacle (possibly a marked object) to look at while exploring.	
	FoundObservedObject	Vector found an object he knows about.	
	ReactToObstacle	Vector is reacting to the obstacle he encountered while driving,	
<i>Faces</i>	DrivingToFace	Vector is driving to a face.	
	DriveToFaceSuccess	Vector was able to drive to a face successfully.	
	EnrolledNewFace	Vector associated a new face with a name.	
	EyeContactReaction	Vector detected the gaze of someone looking at him.	
	GreetingSayName	Vector said the name of someone he recognized.	
	InteractWithFaceRetry		

	InteractWithNamedFace	Vector is looking at the face of someone he knows.
	InteractWithUnnamedFace	Vector is looking at the face of someone he doesn't know.
	LookAtFaceVerified	
<i>Intents</i>	BeQuietVoiceCommand	Vector was told to be quiet.
	FistBumpLeftHanging	Vector attempted a fist bump, but no one gave him one.
	FistBumpSuccess	Vector received a fist bump.
	NoValidVoiceIntent	The wake word was sent, but an intent was either unable to be identified, or the returned intent is not recognized.
	OnboardingStarted	Vector has started the process of on boarding his human companion.
	ReactToTriggerWord	The wake word was said, and Vector has reacted to it.
	RespondToGoodNight	Vector responded to a verbal goodnight (represented as an intent)
	RespondToShortVoiceCommand	Vector responded to an intent that didn't trigger a more complex behavior interaction.
<i>Motion sensing</i>	ReactToMotion	
	ReactToPickedUp	Vector has been picked up and has reacted to it.
	ReactToUnexpectedMovement	
	RobotShaken	Vector was shaken.
<i>Power State</i>	Asleep	Vector is asleep
	Sleeping	Vector is sleeping
<i>Petting</i>	PettingBlissLevelIncrease	Vector is being petted.
	PettingReachedMaxBliss	Vector is being petted.
	PettingStarted	Vector is being petted.
<i>Sound</i>	DanceToTheBeat	Vector is dancing to music that he hears (he may be the only one to hear it)
	ReactToSoundAsleep	Vector heard some noise while asleep or sleeping.
	ReactToSoundAwake	Vector heard some noise (but wasn't asleep)
<i>Timer</i>	TimerRinging	Vector's timer has expired and is ringing.

APPENDIX K

DAS Tracked Events and Statistics

This Appendix captures the events and statistics that are posted to Anki's the diagnostics / analytics services (see Chapter 31)

151. DAS TRACKED EVENTS AND STATISTICS

151.1. BASIC INFORMATION

151.1.1 Version Information

The following are version-information events that are posted to the diagnostic logger:

Event	Description & Notes
hal.body_version	
robot.boot_info	
robot.cpu_info	

Table 532: Version info, posted to DAS

151.1.2 Settings and Preferences Information

The following are settings and preference related events that are posted to the diagnostic logger:

Event	Description & Notes
das.allow_upload	
engine.language_locale	
robot.locale	
robot.timezone	

Table 533: Start up information, posted to DAS

151.1.3

Start-up Information not described elsewhere

The following are start-up events that are posted to the diagnostic logger:

Event	Description & Notes
ntp.timesync	
profile_id.start	
rampost.lcd_check	
rampost.rampost.exit	The rampost has completed and exited.
random_generator.seed	
robot.engine_ready	
switchboard.hello	
vic.cloud.hello.world	

Table 534: Settings and preferences, posted to DAS

Note: other startup events are covered elsewhere with their functional groups.

151.2. POWER MANAGEMENT EVENTS AND STATISTICS

The power management posts the following set of related events:

Event	Description & Notes
hal.active_power_mode	
PowerModeManager.DisableActiveMode	
PowerModeManger.EnableActiveMode	
robot.power_on	

Table 535: Power management events, posted to DAS

151.2.1

Battery Statistics and Events

The battery management posts the following battery related events and state information:

Event	Description & Notes
battery.battery_level_changed	This is set when the battery level has changed from event posting.
battery.encoder_power_stats	?? Strange name. Is this the voltage seen on the charger input and charging duration stats?
battery.fully_charged_voltage	The battery voltage seen when the charger reported the battery to be fully charged.
battery.voltage_reset	
battery.voltage_stats	Information about the range of battery voltages that have been observed; e.g. min/max, average, etc.
rampost.battery_flags	
rampost.battery_level	

Table 536: Battery level events and statistics

151.2.2

Charger Statistics and Events

The charging function of the battery management system posts the following events and state information:

Event	Description & Notes
battery.cooldown	Indicates that Vector is or needed to pause charging and activity to let the battery cool down.
battery.is_charging_changed	This is set when the state of charging has changed from event posting.
battery.on_charger_changed	
battery.saturation_charging	
battery.temp_crossed_threshold	
battery.temperature_stats	Information about the range of battery temperatures that have been observed; e.g. min/max, average, etc.
rampost.battery_temperature	

Table 537: Charger statistics and events, posted to DAS

151.3. MOTOR AND IMU STATISTICS AND EVENTS

The motor controllers post the following events and statistics:

Event	Description & Notes
calibrate_motors	
head_motor_calibrated	
head_motor_uncalibrated	
imu_filter.gyro_calibrated	
lift_motor_calibrated	
lift_motor_uncalibrated	

Table 538: Motor events

151.4. COMMUNICATION RELATED EVENTS POSTED TO DAS

151.4.1 Body-Board / Spine Related Events Posted to DAS

The communication with the based-board controller posts the following events:

Event	Description & Notes
rampost.spine.configure_serial_port	The rampost program was successful in configuring the serial port to communicate with the body-board.
rampost.spine.open_serial	The rampost was able to open the serial port to communicate with the body-board.
rampost.spine.select_timeout	There was a timeout in communicating with rampost the body-board.
touch_sensor.activate_charger_mode_check	
touch_sensor.charger_mode_check.no_baseline_change	
ttyHSL0.service	

Table 539: Body-board / spine related DAS events

Note: see the updates section for events related to updating the body-board firmware

151.4.2 Accessory-Related Events Posted to DAS

The communication with the mobile application and SDK posts the following events:

Event	Description & Notes
cube.battery_voltage	The cube's battery voltage
cube.connected	Vector was able to connect to his accessory companion cube.
cube.connection_failed	Vector was unable to connect to his accessory companion cube.
cube.disconnected	Vector lost the connection with his accessory companion cube.
cube.firmware_mismatch	Vector is unable to use his accessory companion cube as is, since the firmware version is not compatible.
cube.unexpected_connect_disconnect	

Table 540: Accessory cube related DAS events

Note: see the updates section for events related to updating the cube firmware

151.4.3

Mobile-App / SDK Related Events Posted to DAS

The communication with the mobile application and SDK posts the following events:

Event	Description & Notes
ble.connection	
ble_conn_id.stop	
ble.disconnection	
wifi.initial_state	

Table 541: Mobile application / SDK related DAS events

151.5. UPDATE-RELATED EVENTS POSTED TO DAS

The following are events posted by the update subsystem:

Event	Description & Notes
cube.firmware_flash_success	
rampost.dfu.desired_version	
rampost.dfu.installed_version	
rampost.dfu.open_file	
rampost.dfu.request_version	
robot.ota_download_end	On success the parameters include the new version; on failure the parameters include the version identifier, error code, and some explanatory text.
robot.ota_download_stalled	
robot.ota_download_start	

Table 542: Update events

151.6. BEHAVIOUR, FEATURE, MOOD, AND ENGINE RELATED EVENTS POSTED TO DAS

The engine/animation controller posts the following behavior-related events:

Event	Description & Notes
action.play_animation	The specified animation will be played
behavior.feature.end	
behavior.feature.pre_start	The behaviour for specified feature will begin.
behavior.feature.start	The behaviour for specified feature are begun.
behavior.hlai.change	There was a change in the high-level AI state. Some possible supplemental parameters include “ObservingOnCharger”
dttb.coord_activated	The “dance to the beat” feature has been activated.
dttb.coord_no_beat	
engine.state	
mood.event	
mood.simple_mood_transition	
robot.object_located	
robot.reacted_to_sound	

Table 543: Behaviour, feature, mood and engine related DAS events

151.7. ROBOT LIFETIME STATISTICS & EVENTS

Vector tracks the following behavior and events

Key	units	Description & Notes
Alive.seconds	seconds	Vector’s age, since he was given preferences (a factory reset restarts this)
Stim.CumlPosDelta		Cumulative stimulation of some kind
BStat.AnimationPlayed	count	The number of animations played
BStat.BehaviorActivated	count	
BStat.AttemptedFistBump	count	The number of fist bumps (attempted)
BStat.FistBumpSuccess	count	
BStat.PettingBlissIncrease		
BStat.PettingReachedMaxBliss		
BStat.ReactedToCliff	count	
BStat.ReactedToEyeContact	count	
BStat.ReactedToMotion	count	
BStat.ReactedToSound	count	
BStat.ReactedToTriggerWord	count	
Feature.AI.DanceToTheBeat		
Feature.AI.Exploring		

Table 544: The robot lifetime stats schema

Feature.AI.FistBump		
Feature.AI.GoHome		
Feature.AI.InTheAir		
Feature.AI.InteractWithFaces	count	The number of times recognized / interacted with faces
Feature.AI.Keepaway		
Feature.AI.ListeningForBeats		
Feature.AI.LowBattery		
Feature.AI.Observing		
Feature.AI.ObservingOnCharger		
Feature.AI.Onboarding		
Feature.AI.Sleeping		
Feature.AI.Petting	ms	The amount of time petted
Feature.AI.ReactToCliff		
Feature.AI.StuckOnEdge		
Feature.AI.UnmatchedVoiceIntent		
Feature.Voice.VC_Greeting		
FeatureType.Autonomous		
FeatureType.Failure		
FeatureType.Sleep		
FeatureType.Social		
FeatureType.Play		
FeatureType.Utility		
Odom.LWheel		The left wheel odometer – how far it has driven
Odom.Rwheel		The right wheel odometer – how far it has driven
Odom.Body		
Pet.ms	ms	The cumulative time petted

APPENDIX L

Pleo

The Pleo, sold in 2007 –a decade prior to Vector – has many similarities. The Pleo was a software skinned animatronic baby dinosaur created by Caleb Chung, John Sosuka and their team at Ugobe. Ugobe went bankrupt in 2009, and the rights were bought by Innvo Labs which introduced a second generation in 2010. This appendix is mostly adapted from the Wikipedia article and reference manual.

Sensing for interacting with a person

- Two microphones, could do beat detection allowing Pleo to dance to music. The second generation (2010) could localize the sound and turn towards the source.
- 12 touch sensors (head, chin, shoulders, back, feet) to detect when petted,

Environmental sensors

- Camera-based vision system (for light detection and navigation). The first generation treated the image as gray-scale, the second generation could recognize colors and patterns.
- Four ground foot sensors to detect the ground. The second generation could prevent falling by detecting drop-offs
- Fourteen force-feedback sensors, one per joint
- Orientation tilt sensor for body position
- Infrared mouth sensor for object detection into mouth, in the first generation. The second generation could sense accessories with an RFID system.
- Infrared detection of objects
- Two-way infrared communication with other Pleos
- The second generation include a temperature sensor

Annuciators and Actuators

- 2 speakers, to give it sounds
- 14 motors
- Steel wires to move the neck and tail (these tended to break in the first generation)

The processing

- Atmel ARM7 microprocessor was the main processor.
- An NXP ARM7 processor handle the camera system, audio input
- Low-level motor control was handled by four 8-bit processors

A developers kit – originally intended to be released at the same time as the first Pleo – was released ~2010. The design included a virtual machine intended to allow “for user programming of new behaviors.”⁸²

151.8. SALES

Pleo’s original MSRP was \$350, “the wholesale cost of Pleo was \$195, and the cost to manufacture each one was \$140” sold ~100,000 units, ~\$20 million in sales⁸³

The second generation (Pleo Reborn) had an MSRP of \$469

151.9. RESOURCES

Wikipedia article. <https://en.wikipedia.org/wiki/Pleo>

iFixit’s teardown. <https://www.ifixit.com/Teardown/Pleo+Teardown/597>

Ugobe, *Pleo Monitor*, Rev 1.1, 2008 Aug 18

Ugobe, *Pleo Programming Guide*, Rev 2, 2008 Aug 15

⁸² <https://news.ycombinator.com/item?id=17755596>

⁸³ <https://www.idahostatesman.com/news/business/article59599691.html>

<https://www.robotshop.com/community/blog/show/the-rise-and-fall-of-pleo-a-farewell-lecture-by-john-sosoka-former-cto-of-ugobeJohnSosoka>