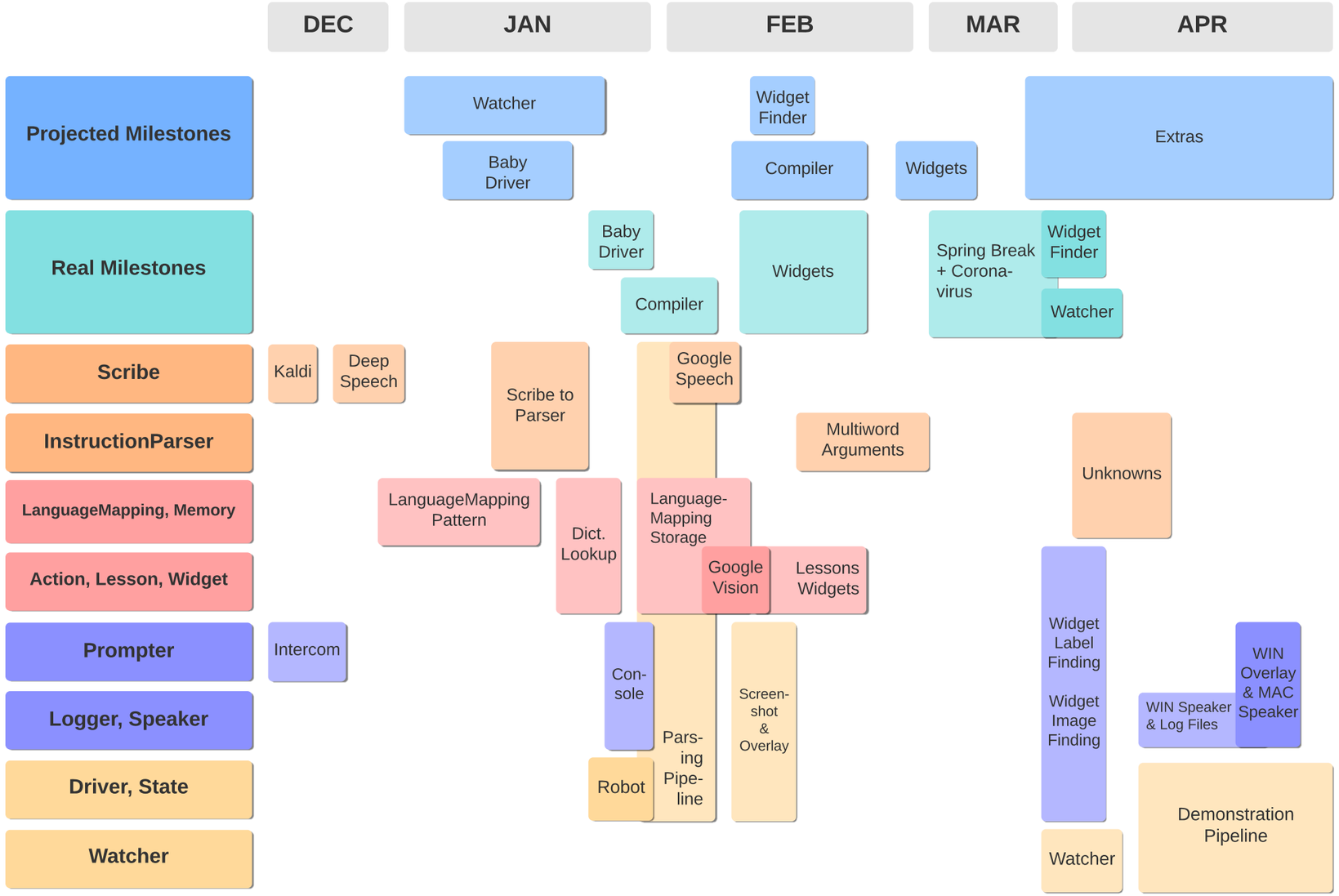
Terry Timeline (Retrospective)

Owen Gallagher | April 12, 2020



Kaldi - created first scribe transcript with Kaldi Aspire
Intercom - created the prompter intercom window with the glasses icon, that registers mouse clicks
DeepSpeech - created scribe transcript with DeepSpeech
LanguageMapping.Pattern - the pattern class stores a graph representation of ways an instruction can be said, and can convert from expression representation (ex: |hi,[whats_up],) ?world))
Scribe to Parser - handles the connection between two asynchronous classes to pass the transcript to the instruction parser
Dict. Lookup - created a dictionary for quick reference to known language mappings in memory, accounting for edit distance
Robot - first primitive actions with driver: point, type, click, shut-down; first driver demonstration with primitive action sequence
Console - added console window to prompter, connected to logger
Parsing Pipeline - the pipeline fully connects the intercom, to the scribe, to the instruction parser, to memory, to the driver for execution
Google Speech - improved transcription accuracy by switching to Google Speech API
LanguageMapping Storage - actions and dictionary can now be read and written to files in src/res/memory/
Google Vision - Google Vision API added for optical character recognition
Screenshot & Overlay - driver can now take a screen capture hiding prompter windows; prompter can create an invisible overlay over GUI
Lessons - first lesson for widget definition created
Widgets - widgets with labels can be created via lesson and referenced in instructions
Multiword Arguments - struggles with multi-word arguments in instruction expressions, with eventual success
Spring Break + Coronavirus - about a 1-month pause for Spring break, the break extension, and later cancellation of campus classes at JC
Widget Label Finding - uses Google Vision to find the widget label in a screenshot, then can direct the driver to move the mouse to that location
Widget Image Finding - defines widget appearance and can search for a widget by appearance to move the mouse to that location
Watcher - created watcher to read system keyboard and mouse events with JNativeHook. I have been struggling to get it to work consistently even now
Unknowns - Terry partially recognizes unknown actions and widgets to be learned
WIN Speaker & Log Files - added speaker to windows using voice.exe and enabled log file creation
WIN Overlay & MAC Speaker - fixed windows overlay and added mac speaker
Demonstration Pipeline - drafted the workflow for demonstration learning, but without a functioning watcher

Baby Driver - functioning driver demonstration with primitive actions
Widget Finder - theoretically find widgets on screen by label and appearance
Compiler - instruction parser to driver pipeline complete
Widgets (milestone) - widgets can be learned, referenced, found, saved, and loaded
Watcher (milestone) - create watcher connected ot peripheral input events
Extras - everything I hoped to do but did not reach: functioning demonstration learning, instruction learning, conditional and argument actions, multipattern actions