

Science of the Computer

Oscar E. Sanford

November 16, 2022

In an effort to retain the best knowledge and experiences of my years studying computer science at the University of Victoria and on my own, this compendium serves as an organized, concentrated volume of the most important concepts and practices that shaped my understanding of computer science and software engineering.

I am in no way able to claim ownership of the great research that many legendary and unsung computer scientists have undergone in order to arrive at our modern understanding of computing theory and their application in practical systems. The theories and ideas herein are neither mine nor complete - the internet is a fountain of free information for those that can access it. I've tried to take what I've read and write it in a way that I (and maybe we) can understand.

I believe education should be accessible to all who search for it. Hi, welcome to my classroom. I find myself aspiring to be a teacher, amongst many things. Unlike a classroom that can feel like a prison at times, you're free to close this document and rid yourself of my attempts at teaching. No hard feelings.

If you're totally new to this stuff, I recommend following the structure of the contents and referring to previous sections when necessary. My approach starts with foundational knowledge that is doable on pen and paper, then how modern computers work, and finally some practical applications. Have fun and keep learning!

- Oscar

Contents

1	Logic	9
1.1	Notation	9
1.2	Logic Laws	10
1.3	Set Theory	10
1.4	Proof by Induction	11
1.5	Time Complexity	11
2	Data Structures	13
2.1	Array	13
2.2	Linked List	14
2.3	Stack	16
2.4	Queue	17
2.5	Trees	18
2.5.1	Proper Binary Tree	18
2.5.2	Binary Search Tree	19
2.5.3	2-3 Tree	20
2.5.4	Heap	20
2.6	Graphs	20
2.6.1	Directed Graphs	22
2.6.2	Weighted Graphs	22
2.7	Hashing	23
2.8	Disjoint Sets	23
3	Algorithms	24
3.1	Algorithm Techniques	24
3.1.1	Divide and Conquer	24
3.1.2	Greedy	25
3.1.3	Dynamic Programming	25
3.2	Sorting	26
3.2.1	Selection Sort	26
3.2.2	Insertion Sort	27

3.2.3	Merge Sort	27
3.2.4	Quick Sort	28
3.2.5	Bucket Sort	30
3.3	Spanning Trees	30
3.3.1	Prim's	30
3.3.2	Kruskal's	30
3.4	Shortest Path	31
3.4.1	Dijkstra's	31
3.4.2	Bellman-Ford	31
3.4.3	Floyd-Warshall	32
3.5	Network Flow	33
4	Theory	34
4.1	Languages	34
4.1.1	Finite Automata	36
4.1.2	Nondeterministic Finite Automata	38
4.1.3	Regular Expressions	39
4.1.4	Pumping Lemma	40
4.2	Context-free Grammars	41
4.2.1	Parse Trees	42
4.2.2	Chomsky Normal Form (CNF)	43
4.2.3	Pushdown Automata	44
4.3	Turing Machines	46
4.3.1	Configurations	47
4.3.2	Recognizability and Decidability	48
4.3.3	Multiple Tapes	49
4.3.4	Nondeterminism	49
4.3.5	Simulation	49
4.4	Undecidable Problems	52
4.4.1	Cantor's Theorem	52
4.4.2	Common Problems	53
4.4.3	Reductions	53
4.4.4	P and NP	55
5	Computer Architecture	57
5.1	Number Bases	57
5.2	Data Units	58
5.3	Hardware	59
5.4	Memory	60
5.4.1	Call Stack	61
5.4.2	Dynamic Heap Memory	62

5.5	Cache	63
5.5.1	Mapping Methods	63
6	Operating Systems	65
6.1	I/O	65
6.1.1	Interfaces	66
6.2	System Management	67
6.2.1	Kernel	67
6.2.2	Protection Rings	68
6.2.3	Startup	69
6.3	Processes	69
6.3.1	Context Switching	70
6.3.2	Scheduling	71
6.3.3	Creation	73
6.3.4	Intercommunication	73
6.4	Threads	74
6.4.1	Management	75
6.4.2	Pthread	75
6.4.3	The Critical Section Problem	77
6.4.4	Mutex	77
6.4.5	Semaphore	81
6.5	File Systems	85
6.5.1	Links	85
6.5.2	FAT12	86
6.6	Memory Management	89
6.6.1	Address Binding	89
6.6.2	Allocation	90
6.6.3	Paging	90
6.6.4	Segmentation	91
7	Linux	93
7.1	File Hierarchy Standard	93
7.2	Tools & Customization	94
7.3	Networking	96
7.3.1	SSH	96
7.4	Encryption	97
7.5	Disk	99
7.6	Processes	99

8	Programming	100
8.1	Overview	100
8.1.1	Paradigms	101
8.1.2	Scheduling	101
8.1.3	Typing	102
8.2	Program Analysis	102
9	Security	104
9.1	Intro	104
9.1.1	Controls	105
9.1.2	Terminology	105
9.1.3	The Security Triad	106
9.2	Social Engineering	107
10	Databases	108
10.1	Functional Dependencies	108
10.2	PostgreSQL	108
11	Artificial Intelligence	109
11.1	Agents	109
11.1.1	Reflex Agents	111
11.1.2	Model-based Reflex Agents	111
11.1.3	Goal-based Agents	111
11.1.4	Utility-based Agents	111
11.1.5	Learning Agents	111
11.1.6	General Structure	112
11.2	Search	112
11.2.1	Uninformed Search	114
11.2.2	Informed Search	115
11.3	Constraint Satisfaction	116
11.3.1	Constraints	117
11.3.2	Backtracking Solver	118
11.3.3	Optimizing CSP	118
11.4	Knowledge Bases	119
11.4.1	Propositional Logic	119
11.4.2	First-Order Logic	120
12	Machine Learning	122
12.1	Probability	122
12.2	Inference	123
12.2.1	Bayesian Networks	124

12.2.2	Exact Inference	125
12.2.3	Approximate Inference	125
12.3	Temporal Models	126
12.3.1	Temporal Inference	127
12.3.2	Hidden Markov Models	127
12.3.3	Kalman Filters	128
12.4	Learning	129
12.4.1	Learning Methods	130
12.4.2	Hypotheses	130
12.4.3	Model Validation	131
12.5	Supervised Learning	133
12.5.1	Maximum Likelihood	133
12.5.2	Bayesian Learning	134
12.5.3	Support Vector Machines	136
12.6	Unsupervised Learning	136
12.6.1	Expectation-Maximization	136
12.6.2	K-Means	137
12.7	Deep Learning	137
13	Reinforcement Learning	138
13.1	Multi-Armed Bandits	138
13.1.1	ϵ -greedy	139
13.1.2	Upper-confidence bound (UCB)	139
13.2	Markov Decision Process	139
13.3	Dynamic Programming for RL	141
13.3.1	Policy Iteration	141
13.3.2	Value Iteration	142
13.4	Monte Carlo Methods	143
13.4.1	Off-Policy MC	145
13.5	Temporal Difference	147
13.5.1	Q-Learning	149
13.5.2	Expected Sarsa	150
13.5.3	Double Q-Learning	150
13.5.4	N-step & A Unified View	151
13.6	Planning and Learning	153
13.6.1	Dyna	153
13.6.2	Prioritized Sweeping	154
13.6.3	Expected vs. Sample Updates	155
13.6.4	Trajectory Sampling	155
13.7	Approximation	155
13.7.1	On-Policy Prediction	155

13.7.2	Stochastic Gradient Descent (SGD)	156
13.7.3	On-Policy Control	157
13.8	Eligibility Traces	159
13.8.1	λ -return	160
13.8.2	Offline λ -return	160
13.8.3	TD(λ)	161
13.8.4	Truncated n -step λ -return	161
13.8.5	Sarsa(λ)	162
13.9	Policy Gradient Methods	163
13.9.1	REINFORCE	164
13.9.2	Actor-Critic Methods	165
13.9.3	Deep Deterministic Policy Gradient (DDPG)	166

14 Computer Networks 168

14.1	Service Types	169
14.2	TCP/IP Model	169
14.3	The Application Layer	170
14.3.1	HTTP	170
14.3.2	Domain Name System	171
14.3.3	DNS Structure	171
14.4	The Transport Layer	173
14.4.1	TCP	173
14.4.2	TCP Flow Control	177
14.4.3	TCP Congestion Control	178
14.4.4	UDP	179
14.5	The Network Layer	179
14.5.1	IP Protocol	180
14.5.2	IP Addresses	181
14.5.3	Classless InterDomain Routing (CIDR)	182
14.5.4	Dynamic Host Configuration Protocol (DHCP)	183
14.5.5	Network Address Translation (NAT)	183
14.6	Routing Algorithms	184
14.6.1	Distance Vector Routing	184
14.6.2	Link State Routing	184
14.6.3	Hierarchical Routing	185
14.7	The Data Link Layer	187
14.7.1	The Channel Allocation Problem	187
14.7.2	ALOHA	187
14.7.3	Carrier Sense Multiple Access (CSMA)	188
14.7.4	Switched Ethernet	189
14.8	Wi-Fi	190

14.8.1	Wireless Problems	191
14.8.2	DCF	192
14.8.3	Handshakes	194
15	Software Development	196
15.1	Cloud Native	196
15.1.1	Docker Containers	197
15.1.2	Contexts & Registries	198
15.1.3	Kubernetes	199
16	Software Evolution	201
16.1	Open Source Software	201
16.2	Git & Source Control	201
16.3	GitHub	201
16.3.1	GitHub Actions	201