

Blockchain: Fundamentals, Data Structures and Algorithms for Data Science

Akcora, Kantarcioglu and Gel

The latest version can be downloaded from:
<https://cakcora.github.io/book>

Contents

	<i>Preface</i>	<i>page 1</i>
1	Money and Cryptocurrencies	5
1.1	A Brief History of Digital Money	6
1.2	Attributes of Money	7
1.2.1	Financial Attributes	7
1.2.2	Digital Attributes	10
1.3	Road to Bitcoin	11
1.4	Reading List	15
	Exercises	15
2	Bitcoin: the First Blockchain	16
2.1	Tenets of Bitcoin	17
2.2	Hash Functions	20
2.3	Address and Output	22
2.4	Transaction	25
2.5	Improvements in Address and Transaction	27
2.5.1	Hierarchical Deterministic wallets	28
2.5.2	Segregated Witness	29
2.5.3	Mast, Schnorr and Taproot	30
2.6	Block	31
2.7	Block Mining	34
2.8	Proof-of-Work	37
2.9	Mining Reward and Fees	44
2.10	Bitcoin Script Language	46
2.11	Peer-to-Peer Network	47
2.12	Forks, Splits and Shared Origins	48
2.13	Log Messages	49
	Exercises	52

3	Ethereum: the World Computer	56
3.1	Beyond Cryptocurrencies	57
3.2	Address	60
3.3	Transactions in Terms of Their Purpose	61
3.4	Gas Cost, Gas Price, and Gas Limit	62
3.5	Smart Contract Data Storage	65
3.5.1	Stack	65
3.5.2	Memory	66
3.5.3	Storage	66
3.6	Transaction and Message Call	67
3.7	Transaction Types	72
3.8	Transaction Costs	74
3.9	Merkle Patricia Trie	77
3.10	Block	78
3.11	Consensus Mechanisms on Ethereum	81
3.12	Validator Rewards and Fees	82
3.13	Ethereum Nodes and Clients	84
3.14	Event and Log	85
3.15	Application Binary Interface	86
3.16	Call Data	87
3.17	Upgradeable Contracts	88
3.18	dApps and Web3	90
	Exercises	92
4	Solidity Coding for Ethereum	95
4.1	Primitive Data Types	96
4.2	Reference Data Types	99
4.3	Functions	103
4.4	Variables in Solidity	103
4.5	Visibility Specifiers and Function Types	105
4.6	Loops and Conditions	107
4.7	Security and Error Handling	109
4.8	Function Modifiers	110
4.9	Making and Receiving Payments	113
4.10	Events	116
4.11	Gas Considerations in Solidity	118
4.12	Advanced Solidity Features	119
4.12.1	Oracle Usage in Solidity	119
4.12.2	Inline Assembly	120
	Exercises	122

5	Ripple: the Currency Exchange	124
5.1	Ripple: History and Origins	125
5.2	Consensus and Trust-based Validation	126
5.3	The Software Ecosystem and Rippled	128
5.4	XRP Properties and Sales History	129
5.5	User-issued Currencies	130
5.6	Trust Lines	131
5.7	Accounts in the XRP Ledger: Special Addresses and Transaction History	133
5.7.1	Reserves	135
5.7.2	Open, Closed, and Validated Ledgers	136
5.8	Fee Structures in the XRP Ledger	137
5.8.1	Mandatory Fees	137
5.8.2	Optional Fees	138
5.9	Transaction Timeline and Processes	138
5.10	Source and Destination Tags	139
5.11	Transaction Costs	140
5.12	Finality of Results in the XRP Ledger	141
5.13	Ledger Payment Types	142
5.13.1	Direct Payments	142
5.13.2	Cross-currency Payments	143
5.13.3	Partial Payments	144
5.14	Financial Concepts	144
5.14.1	Offer	145
5.14.2	Check	146
5.14.3	Escrow	146
5.14.4	Payment Channels	147
5.15	Tokens	148
5.16	Paths	149
5.17	Decentralized Exchanges in the XRP Ledger	150
5.18	Infrastructure Developments	150
	Exercises	152
6	Privacy Coins	158
6.1	Zcash	159
6.1.1	Zcash Fundamentals	159
6.1.2	Zcash Transaction Types	161
6.1.3	Mining and Consensus on Zcash	162
6.1.4	Currency Supply in Zcash	163
6.2	Dash	163

6.2.1	Privacy-Enhancing Techniques in Dash	164
6.2.2	Consensus Protocol and Masternodes in Dash	165
6.2.3	Currency Supply in Dash	165
6.3	Monero	165
6.3.1	Consensus Protocol, and Mining in Monero	168
6.3.2	Evolution of Decoy Input Selection (Mixin Strategy)	169
6.4	Comparative Analysis of Privacy Coins	172
	Exercises	173
7	Blockchain - Next Generation	174
7.1	Alternative Consensus Mechanisms: Proof-of-X	175
7.1.1	Energy Efficiency and Sustainability	177
7.2	Scalability Solutions for Blockchain	179
7.2.1	Layer-One Protocols	180
7.2.2	Layer-Two Protocols	185
7.2.3	A Comparison of Layers	194
7.3	Blockchain Interoperability	195
7.3.1	Bridges	196
7.3.2	Protocol-level Interoperability	198
7.4	Namecoin: Pioneering Blockchain-Based Domain Services	199
	Exercises	201
8	Decentralized Finance	205
8.1	Financial Primitives	206
8.2	Building Blocks of DeFi	210
8.2.1	DeFi Asset Types, Roles and usage	210
8.2.2	Tokens	212
8.2.3	Stablecoins	221
8.2.4	Wrapped Tokens	228
8.3	DeFi Protocols and Mechanisms	229
8.3.1	Oracles	229
8.3.2	Decentralized Exchanges	233
8.3.3	Lending Protocols	241
8.3.4	Derivative Platforms	247
8.3.5	Yield Farming	252
8.4	Management and Governance in DeFi	253
8.4.1	Keepers	254
8.4.2	Governance	255
8.5	Decentralized Autonomous Organizations	255

8.6	Future Directions in DeFi	256
	Exercises	258
9	Blockchain Transaction Networks	260
9.1	Bitcoin: A Canonical UTXO Network, and Extensions in Monero and Zcash	261
9.1.1	Graph Rules for UTXO Blockchains	264
9.1.2	UTXO Transaction Graph	266
9.1.3	UTXO Address Graph	267
9.1.4	Monero and Zcash Transaction Networks	269
9.1.5	Chainlets for UTXO Transaction Networks	274
9.1.6	Occurrence and Amount Information in Chainlets	276
9.2	Ethereum: Account Networks	279
9.2.1	Account Transaction Network	281
9.2.2	Token Transaction Networks	283
9.2.3	Trace Network	285
9.3	Ripple: Credit Networks	288
9.3.1	Ripple trust graph	291
9.3.2	Ripple payment graph	294
	Exercises	297
10	Analyzing Blockchain Entities: Clustering, Mixing, and Cen- trality	300
10.1	Clustering Blockchain Networks	300
10.1.1	Address Clustering	301
10.2	Coin Mixing	305
10.2.1	Centralized Coin Mixing	308
10.2.2	Decentralized Coin Mixing	310
10.3	Centrality and Influence in Blockchain Networks	313
10.3.1	Centrality in UTXO-Based Blockchains	313
10.3.2	Centrality in Account-Based Blockchains	318
10.3.3	Applications: Influence, Propagation, and Illicit Activity	322
	Exercises	325
11	Privacy and Security on Blockchain	327
11.1	Requirements for Privacy and Security	328
11.2	Attacks on the P2P Network	328
11.2.1	Sybil and Eclipse Attacks	329
11.2.2	Partition Attacks	331
11.2.3	Timejacking Attacks	332
11.3	Attacks on Transaction and Block Mining	334

11.3.1	Consensus Delay	334
11.3.2	Empty Block Attack	334
11.3.3	Fee Sniping and Undercutting Attacks	335
11.3.4	Confirmation and Race Attacks	335
11.3.5	51% Attack	335
11.3.6	Selfish Mining	337
11.3.7	Block Withholding Attacks	338
11.3.8	Stake Attacks	339
11.3.9	MEV, Ordering and Bribe Attacks	340
11.4	Attacks on Data Privacy	342
11.4.1	IP Linking Attacks	342
11.4.2	Dust Attacks	343
11.4.3	Malicious Value Fingerprinting	344
11.4.4	Privacy Coin Attacks	345
11.5	Attacks on Smart Contracts and DeFi Protocols	349
11.5.1	Smart Contract Vulnerabilities	349
11.5.2	Attacks on Protocols	353
11.5.3	Attacks Relying on the Execution Order of Transactions in a Block	355
11.5.4	Attacks Executed Within a Single Transaction	359
	Exercises	365
12	E-crime on Blockchains	368
12.1	Thefts and Hacks	369
12.2	Ponzi Schemes and Fraudulent Investment Offers	370
12.3	Cryptoasset Price Pump and Dump Schemes	373
12.4	Darknet Markets	375
12.5	Sextortion	377
12.6	Ransomware	379
12.7	Money Laundering on Blockchains	382
12.7.1	Tracing Coins in UTXO-based Blockchains	382
12.7.2	Tracing Coins in Account-based Blockchains	384
12.7.3	Evasion Tactics in Money Laundering	386
12.8	A Brief History of E-Crime on Blockchains	388
	Exercises	392
13	Temporal Analysis	395
13.1	Background on Time Series and Forecasting	395
13.2	Temporal Data Analysis on Blockchains	398
13.2.1	Crypto-Asset Time Series Tools	399

	13.2.2 From Time Series to Dynamic Transaction Graphs	400
	Exercises	404
14	Conclusion	406
	References	409
	<i>References</i>	409
	14.1 Answers to Questions in Chapter Money and Cryptocurrencies	435
	14.2 Answers to Questions in Chapter Bitcoin: the First Blockchain	436
	14.3 Answers to Questions in Chapter Ethereum: the World Computer	441
	14.4 Answers to Questions in Chapter Solidity Coding for Ethereum	445
	14.5 Answers to Questions in Chapter Ripple: the Currency Exchange	447
	14.6 Answers to Questions in Chapter Privacy Coins	453
	14.7 Answers to Questions in Chapter Blockchain - Next Generation	454
	14.8 Answers to Questions in Chapter Decentralized Finance	458
	14.9 Answers to Questions in Chapter Blockchain Transaction Networks	460
	14.10 Answers to Questions in Chapter Analyzing Blockchain Entities: Clustering, Mixing, and Centrality	463
	14.11 Answers to Questions in Chapter Privacy and Security on Blockchain	465
	14.12 Answers to Questions in Chapter E-crime on Blockchains	469
	<i>Index</i>	474

Preface

The famous author Douglas Adams remarked that technologies invented between the ages of fifteen and thirty-five are exciting and revolutionary, and that you can probably build a career in them. He also added that anything in the world when you are born is normal and ordinary, and is just a natural part of the way the world works. We first started studying blockchain in our thirties, and we found it new and exciting. Meanwhile, Bitcoin and its underlying technology, blockchain, have continued to gain popularity across a range of applications, from finance and insurance to clinical trials. A decade later, we now teach it to the youth with the solemn conviction that it's part of the natural order of things, and everyone should know the basics of blockchain.

Originally designed to facilitate a secure distributed platform without central authorities, blockchain was heralded as a paradigm that would be as powerful as Big Data, Cloud Computing, and Machine Learning. Expectations around blockchain have ebbed and flowed, but its importance can not be belittled. Daily blockchain trading volumes have surged past 120 billion USD [113], making blockchain a major financial avenue.

Blockchain brought together ideas from various fields, such as public key encryption and distributed systems. As such, early readers often studied resources that explained blockchain technology from only a narrow perspective, leaving them with more questions than clarity.

In this book, we offer a holistic view of blockchain. Table 0.1 gives a compact tour of the blockchain landscape, where coins and platforms differ in structure, purpose, and ideology. Starting with a brief historical overview, we discuss the building blocks of blockchain and systematically explain their interactions. We also devote a section to the next generation of blockchain and explain such extensions as Decentralized Autonomous Organizations. As graph mining plays an important role in blockchain, we elaborate on how graph theoretical aspects can be used in blockchain technology and data analytics.

	Data Structure			Functionality		System Properties		
	UTXO	Account	DAG	Platform	Cryptocurrency	Privacy-enabled	Permissioned	Public
Bitcoin	✓				✓			✓
Litecoin	✓				✓			✓
ZCash	✓				✓	✓		✓
Dash	✓				✓	✓		✓
Monero	✓				✓	✓		✓
Ripple		✓			✓		✓	
Ethereum		✓		✓	✓			✓
IOTA			✓	✓	✓		semi	✓

Table 0.1 *Blockchain systems grouped by data structure, functionality, and system properties. Privacy-enabled indicates support for optional or default privacy features. Permissioned means validator participation is restricted, as in Ripple. Ethereum is permissionless, as any node can stake 32 ETH and become a validator. “semi” denotes transitional or partially centralized designs.*

We adopt a chronological approach to teaching technology, where chapters follow the historical progression of developments rather than their technical complexity. For example, in the context of scalability, we first introduce layer one advancements, such as Proof of Stake, before moving to earlier solutions like segregated witness (a layer zero solution). This mirrors the actual sequence of events in the field, as people did not first develop layer zero solutions and then progress to higher layers. By teaching in this order, we offer students a clearer understanding of how the field has evolved and where it is headed, allowing them to contextualize current and future innovations.

Blockchains involve multiple disciplines, such as graph theory, finance, and game theory. To fully grasp blockchain, it must be considered as a concept that interacts with the world. In this sense, we follow the Annales School methodology [81], examining not only the technology itself but also the broader ecosystem in which it operates. By taking this comprehensive, world-context view, we aim to provide students with a deeper understanding of blockchain’s development and its wide-reaching impact.

Without assuming any prior reader’s expertise, we aim to provide a concise but complete description of blockchain technology and its capabilities.

Roadmap for This Book

Figure 0.1 shows an example reading order for the book. Chapters 1 to 2 introduce the origins of digital money and Bitcoin. Readers seeking a foundational understanding of money, financial systems, and the motivations behind Bitcoin

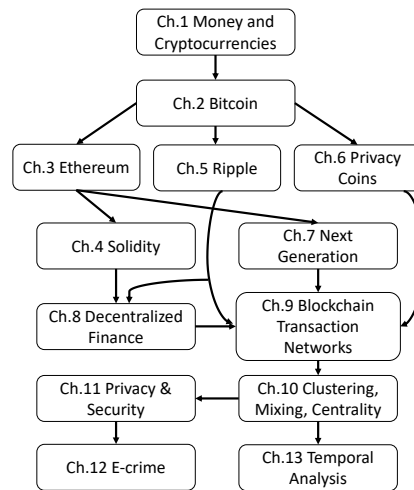


Figure 0.1 Suggested reading flow through the chapters. Arrows indicate conceptual or technical dependencies.

should begin here. Chapter 3 teaches Ethereum as a programmable blockchain, detailing accounts, gas mechanics, and data structures. Chapter 4 focuses on Solidity and smart contract development. It assumes familiarity with Ethereum concepts, so readers without coding experience might wish to skim rather than study it deeply at first. Chapter 5 explains Ripple and its architecture as a bridge currency and trust-based ledger. Chapter 6 teaches Dash, Monero, and Zcash privacy coins. Chapter 7 expands into next-generation technologies such as alternative consensus mechanisms, scalability, and blockchain interoperability. It provides valuable foresight for students focused on research or system-level design. Chapter 8 introduces Decentralized Finance, including tokens, exchanges, lending protocols, and DAOs. This section is best read after mastering Ethereum. Chapter 9 takes a graph-theoretical approach, describing the blockchain as transaction networks. Chapter 10 shows how addresses can be clustered, how important nodes can be discovered and how address influence can be measured. Readers interested in analytics, data science, or network theory will find Chapters 9 and 10 particularly useful. Chapters 11 and 12 focus on privacy and security and the use of blockchains in e-crime. These are especially relevant for students in cybersecurity or digital forensics. Chapter 13 introduces temporal analysis, useful for advanced blockchain research and time-aware systems.

Each chapter concludes with questions to prompt reflection and enable self-assessment.

The book is designed to stand as both a course textbook and a reference work. Whether used in a university classroom or for independent study, its structure encourages a layered, cumulative understanding of blockchain technologies.

1

Money and Cryptocurrencies

This chapter traces the intellectual and technological lineage of Bitcoin and digital money. While Nakamoto's white paper launched Bitcoin, its roots extend through decades of economic theory, cryptographic innovation, and activist movements. We examine how the Austrian and Chicago Schools of Economics provided a framework for stateless and non-inflationary money, and how Cypherpunk ideals shaped the push for privacy and decentralization. The chapter reviews early experiments with digital currencies such as DigiCash, b-money, and e-gold, highlighting the technical shortcomings, regulatory battles, and user adoption barriers that prevented their success but furnished essential building blocks for Bitcoin. We then contrast the classical financial attributes of money—medium of exchange, unit of account, and store of value—with additional digital requirements such as offline spendability, identity-less spendability, and fungibility. Finally, we show how Bitcoin resolved the long-standing double-spending problem without a central authority through proof-of-work, situating it as both a culmination of earlier efforts and the starting point for a new era of cryptocurrencies.

Although Nakamoto's paper marked the launch of Bitcoin, the system's intellectual roots lie in a wider lineage. Earlier work on blinded signatures by David Chaum, the structure and vulnerabilities of centralized digital gold projects like e-gold, and the ideological foundations laid by the Cypherpunks all contributed essential ingredients. Moreover, regulatory responses after 9/11 and ongoing battles over anonymity and control further shaped the environment in which Bitcoin emerged. It is more accurate to view Bitcoin as a culmination of technical lessons and philosophical motivations stretching across decades, rather than a sudden and solitary invention.