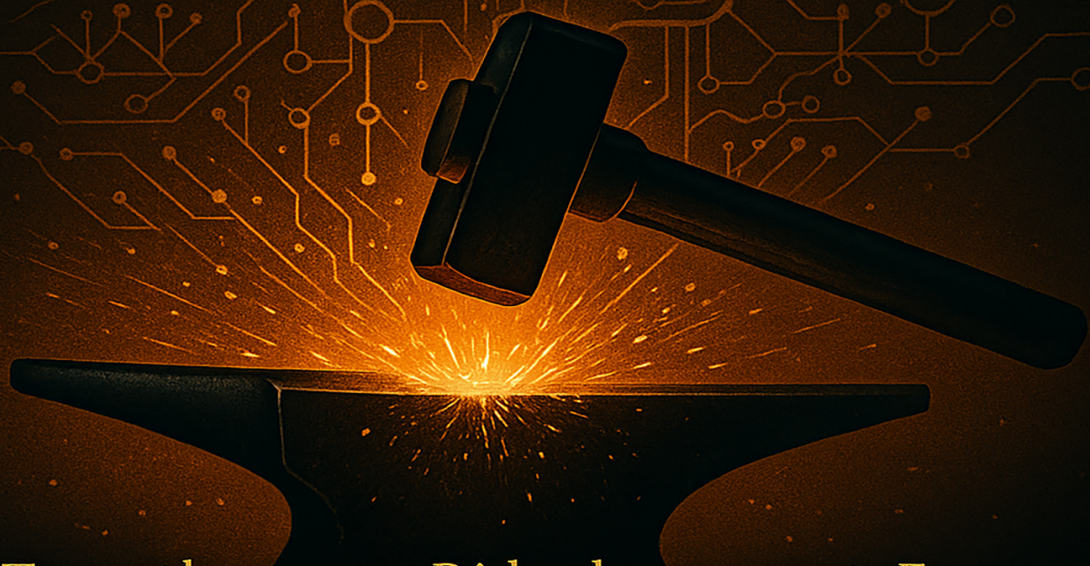


# THE AI FORGE

## VOLUME I

From Foundations to Frontiers



Trace the roots. Ride the current. Forge  
the future of AI.



# DR. JIANXIAN WU

# The AI Forge

**Volume I: From Foundations to Frontiers**

*Trace the roots. Ride the current. Forge the future of AI.*

# **The AI Forge, Volume I: From Foundations to Frontiers**

© 2025 by Jianxian Wu

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means —electronic, mechanical, photocopying, recording, or otherwise —without the prior written permission of the publisher, except in the case of brief quotations embodied in critical articles or reviews.

Published by PCA<sup>3</sup> Press

First Edition

ISBN: 979-8-9989817-1-5

Kindle Edition

# FOREWORD BY CHRIS ROWEN

If you’ve been living under a rock for the past five years, you might have missed the explosion of interest in artificial intelligence (AI). Otherwise, you’ve probably felt the buzz in the media, seen the flood of new products, and noticed the astonishing pace of technical progress in AI. But the complexity of these methods, the breadth of their applications, and the breathless hype may leave you yearning for thoughtful yet accessible explanations. The AI Forge is here to provide exactly that kind of grounding and to help you explore many of the most important questions:

- Where did AI come from?
- What kinds of problems can AI methods solve?
- What does AI progress reveal about the nature of intelligence?
- How can you internalize—and truly understand—the remarkable power (and occasional weaknesses) of AI?

Dr. Jianxian Wu’s book makes remarkably lucid progress on these questions, while offering a systematic journey through puzzles, games, and algorithms that enable you to learn AI by doing. Dr. Wu brings the ideal background to this task. His long career in high performance computing - spanning Teradyne, HP Labs, AMD, and Baidu, plus his own brave startup - led him naturally into the heart of advanced AI research, especially at the intersection of novel algorithms and scalable, efficient implementation. His rare combination of deep technical skill and educational clarity makes this book stand out as both authoritative and approachable. That experience shines through in this book, which combines practical insight with clear and structured explanations of core AI principles.

You’ll find The AI Forge weaving together four perspectives. First, it frames the evolution of AI through the stories of the pioneers who made breakthroughs in theory and practice. Second, it deepens that story with clear discussions of the foundational ideas—from the conceptual work of Alan Turing and Claude Shannon to the cutting-edge reasoning models of OpenAI and DeepSeek. Third, it grounds your learning in real examples, especially a deep exploration of the game Connect 4. This deceptively simple game serves as the backdrop for building up increasingly sophisticated AI strategies, from tree search, linear regression, and convolutional neural networks, to generative pretrained transformers (GPTs), reinforcement learning, and chat-oriented large language models. And finally, the book brings computing efficiency—often overlooked but critical—into sharp focus. By engaging with implementation challenges,

you'll learn how to build AI that is not only powerful, but also fast, efficient, scalable, and responsible. Along the way, you'll gain not only technical understanding but also a deeper appreciation of AI's potential to augment human insight and creativity.

Whether you're skimming through chapters of interest or diving deeply into the hands-on coding exercises and gameplay strategies, you'll find this book both accessible and rewarding. The AI Forge is a rare kind of resource: one that gives you a meaningful return on the time you invest, whether you spend two hours or two months with it.

# PREFACE: THE FORGE AND THE FLAME

In every medieval village, the blacksmith's workshop stands at the heart of the community. Apprentices start young, sweeping floors and watching silently. Over time, they earn trust to stoke the fire, shape small pieces of metal, and learn the rhythm of the hammer. As their understanding deepens, they earn their place at the anvil, crafting blades of remarkable strength and elegance.

The forge is a place where skill, patience, and insight come together, where mastery is shaped by hand, by mind, and by fire.

This book follows the same spirit. In the tradition of the old guild masters, it unfolds in two parts. First, Volume I: From Foundations to Frontiers, where we travel back through the early landscape of artificial intelligence. We meet the pioneers who faced the unknown with boldness and imagination, Alan Turing, Arthur Samuel, Ken Thompson, and explore the ideas they forged: ideas that changed the course of history. We seek not only to learn these foundations, but to understand their meaning, to appreciate their beauty, and to use them as stepping stones for our own creations.

Later, we enter Volume II: Beyond Frontiers, where the craft deepens. Here, we embrace performance-conscious, architecture-aware, AI-aligned design, using C not as a burden, but as a forge to shape faster, sharper, and more elegant AI systems. We learn to harness the full potential of the machine, building intelligence that is not only powerful, but beautifully efficient and enduring.

Whether you are a student preparing for an AI-driven world, a professional seeking deeper mastery, or a lifelong learner drawn to the mysteries of intelligence, this book invites you to walk the full path, from beginner to builder, from builder to craftsman.

Let us begin, with curiosity, with fire, and with the courage to forge something new.

And so, with the spirit of the forge alive in our hands, we take our first steps onto The Apprentice's Path. Here, we return to the beginnings of artificial intelligence, to the simple games, the bold ideas, and the brilliant minds who lit the first sparks. Let us walk with them, learn from them, and begin to shape our own craft.

# INTRODUCTION: TRACING THE PATH AHEAD

Artificial intelligence has risen from a quiet branch of computer science into a force that touches nearly every part of our lives. Yet even as AI becomes a defining technology of the modern world, confusion and uncertainty often surround it. What is AI, truly? How does it work beneath the surface? What role will it play in our future, and what role will we play alongside it?

This book was born from a desire to answer those questions, not with shortcuts or slogans, but by tracing the roots of intelligence itself, to understand where AI came from, how it grew, and what it means to build it with care and purpose. It is a journey through history, theory, and craft, told through the stories of the pioneers who first imagined machines that could think, and through the ideas that continue to shape how machines learn, decide, and create today.

We wrote *The AI Forge* for three kinds of readers. For the **AI enthusiasts**, those who have heard the echoes of AI everywhere and seek to understand its foundations, technologies, and meanings. For the **young generation**, high school and college students standing at the threshold of an AI-shaped future, wondering how to prepare themselves not just to survive it, but to ride its wave with skill and insight. And for the **professionals**, those already working with technology who seek a deeper, principled understanding of AI's roots and future, beyond headlines and buzzwords.

Our approach is different from many AI books. This is not a collection of recipes or checklists to memorize. It is a forge: a place where curiosity is tempered into understanding, where historical context, physical intuition, and hands-on building come together. In each chapter, we first meet the heroes, the human minds who dreamed, struggled, and invented, and then explore their ideas, not just mechanically, but philosophically: why they matter, how they connect to the physical realities of computation, and how they reshaped what we mean by intelligence.

You do not need to be a programmer to travel this path. Even if you never write a line of Python or C, the ideas and the journeys will speak to you. But for those who wish to build, we offer a steady path: simple, accessible projects in Python to bring AI concepts to life, and later, the transition into C, to forge those ideas into systems that are not only functional, but efficient, elegant, and scalable.

This structure reflects an ancient ideal, one that has long guided learning and mastery. In Chinese tradition, there is a saying: 登堂入室. First, you enter the

hall, gaining access to knowledge and beginning your training. Then, you step into the inner chamber, where true mastery is forged, through discipline, reflection, and deeper understanding.

**Volume I: From Foundations to Frontiers** is your first step into the hall. Here, you will trace the roots of artificial intelligence, learn through games like Connect 4, and explore the classical pillars of search, learning, and reasoning. You'll not only build your first agents, but also develop the physical and philosophical intuition that underpins AI's evolution.

**Volume II: Beyond Frontiers**, opens the door to the inner chamber. Here, the craft deepens. We turn to performance-conscious, architecture-aware, AI-aligned engineering, crafting AI systems that not only think, but think efficiently, sustainably, and harmoniously with the machines they run on. In a world where AI's energy demands are growing exponentially, this discipline is no longer optional. It is essential. The future of AI belongs not to those who can simply use models, but to those who can build them wisely, respecting every cycle, every watt, every design.

**The AI Forge** is thus more than a technical guide. It is a philosophy of learning, building, and shaping AI with care, clarity, and consciousness, preparing you not only to navigate the AI era, but to help lead it, crafting systems that are sustainable, elegant, and enduring.

Welcome to the forge. Let us begin.



# CONTENTS

<b>Foreword by Chris Rowen</b>	<b>iv</b>
<b>Preface: The Forge and the Flame</b>	<b>vi</b>
<b>Introduction: Tracing the Path Ahead</b>	<b>vii</b>
<b>1 The Rise of AI: The Evolution of Machine Intelligence</b>	<b>1</b>
Prologue: The First Fracture - Kasparov vs. Deep Blue . . . . .	1
The Last Spark: AlphaGo vs. Lee Sedol . . . . .	2
The Turning Point: ChatGPT and the Arrival of Everyday AI . . . . .	5
Three Sparks, One Flame . . . . .	6
History of Artificial Intelligence . . . . .	7
Before “Artificial,” There Was Intelligence . . . . .	7
Lights in the Dark Age . . . . .	10
AI from Learning . . . . .	14
The Cyclic Nature of AI . . . . .	19
A Cycle That Mirrors Our Minds . . . . .	20
Foundations That Echo Forward . . . . .	20
<b>2 Connect 4: Our AI Testbed</b>	<b>22</b>
Prologue: A Game of Surprises . . . . .	22
Why Connect 4? The Drosophila of AI . . . . .	23
The Hidden Depths . . . . .	25
From Brute Force to Learning . . . . .	26
A Solved Game, A Lasting Legacy . . . . .	27
What Connect 4 Teaches Us . . . . .	28
A Preview of Volume I . . . . .	28
What’s Next: Let’s Begin the Build . . . . .	29

<b>3</b>	<b>Python &amp; C for AI Development</b>	<b>31</b>
	Prologue: The Swordsmith – The Story of C . . . . .	31
	The Sculptor’s Tool – The Story of Python . . . . .	32
	Why Programming Still Matters in the Age of AI . . . . .	33
	What Makes Python Stand Out as the Language of AI . . . . .	33
	What Makes C Indispensable for Building AI at Scale . . . . .	35
	Enter PCA <sup>3</sup> : A Discipline for the GenAI Era . . . . .	36
	Looking Ahead: From Philosophy to Practice . . . . .	39
	From Vision to Action: Time to Set Up . . . . .	40
	Setting Up Our Development Environment . . . . .	40
	Python: The Language of Exploration . . . . .	41
	C: The Language of Execution . . . . .	41
	Version Control & Experiment Tracking . . . . .	42
	Git: Our Time Machine . . . . .	42
	Benchmarking AI Quality & Performance . . . . .	43
	What to Measure (and Why) . . . . .	43
	A Simple Python Benchmark Script . . . . .	44
	Benchmark Early, Benchmark Often . . . . .	44
	Key Takeaway: Simplicity Meets Scale . . . . .	45
<b>4</b>	<b>First AI: A Python-Based Connect 4 Player</b>	<b>46</b>
	Prologue: The Seeds of Foresight . . . . .	46
	Prologue: Simulating Chaos . . . . .	48
	The Bridge: Planning Amid Uncertainty . . . . .	49
	The First Move . . . . .	50
	The First Real Step . . . . .	50
	Designing for Growth: Why Structure Matters . . . . .	51
	What We Are Building (Quietly, Carefully) . . . . .	51
	The Journey Ahead . . . . .	52
	Understanding the Code: Building the System, Piece by Piece . . . . .	52
	Initializing the Game Board . . . . .	52
	Displaying the Board . . . . .	53
	Checking if a Move is Valid . . . . .	53

Placing a Piece on the Board . . . . .	53
Switching Players . . . . .	54
Understanding Win Detection Directions . . . . .	54
Detecting a Win . . . . .	54
Checking for a Draw . . . . .	54
Making a Random AI Move . . . . .	55
Game Loop – Running the Game . . . . .	56
Our First AI: ai_random . . . . .	56
The Full Program: Our First Playable AI . . . . .	57
What Comes Next: From Random to Reason . . . . .	59
<b>5 Baseline, Benchmark, and Best Practices</b>	<b>61</b>
Prologue: The Dream of Rules - Allen Newell and Herbert Simon . . . . .	61
Building Our First Rule-Based AI . . . . .	63
From Rules to Reality: The Need for Fair Testing . . . . .	69
Introducing BBB: Baseline, Benchmark, and Best Practices . . . . .	71
Setting Up the Arena: Our First True Tests . . . . .	72
First Benchmark . . . . .	75
Building the Simulation AI: Toward Probabilistic Reasoning . . . . .	76
Monte Carlo: From Solitaire to Strategy . . . . .	76
From Randomness to Intelligence . . . . .	77
Simulation AI . . . . .	78
Benchmarking All Players . . . . .	79
Key Takeaways: What We Have Built, What We Have Learned . . . . .	81
<b>6 Classical AI: Alpha-Beta Search &amp; Heuristics</b>	<b>83</b>
Prologue: The Men Who Thought Ahead . . . . .	83
What Is Classical AI? . . . . .	85
Minimax: Brute-Force Brilliance . . . . .	86
A Visual Example of Minimax Decision-Making . . . . .	88
The Human Version of the Minimax Algorithm . . . . .	89
Alpha-Beta Pruning: Think Smarter, Not Harder . . . . .	89
Knuth’s Revelation: Depth Without Exhaustion . . . . .	91

AlphaBeta Algorithm . . . . .	91
The Power of Alpha-Beta: Speed Without Sacrificing Strength . . . . .	92
A Note on Move Ordering . . . . .	94
Negamax: A Cleaner Way to Think . . . . .	94
Evaluation Functions: How AI “Feels” About a Game . . . . .	96
How the AI Sees the Board: Building Blocks of Strategy . . . . .	97
Alpha-Beta with Proper Evaluation: Strategic Intelligence Unleashed . . . . .	100
Benchmarking and BBB: Measuring, Not Guessing . . . . .	102
Benchmarking Depth vs Speed: The Cost of Thinking . . . . .	103
Depth Duel: Deeper Thinkers vs Faster Movers . . . . .	105
The Hardware Race: From Belle to Deep Blue . . . . .	107
Dr. Feng-hsiung Hsu and the Birth of Deep Thought . . . . .	108
Type A vs Type B: The Dreams of Shannon, the Reality of Hsu and Berliner	110
Man vs Machine: Play Against Our Creation . . . . .	111
Sample Session . . . . .	112
Closing Reflection: The Strengths and Limits of Classical AI . . . . .	113
<b>7 From Classical AI to Machine Learning</b>	<b>115</b>
Prologue: Arthur Samuel and the Birth of Learning Machines . . . . .	115
What Machine Learning Means to AI . . . . .	117
From Rigid Rules to Flexible Learning . . . . .	118
The Feedback Loop: Learn $\rightarrow$ Play $\rightarrow$ Learn . . . . .	118
Adaptation = Survival . . . . .	119
From Programming to Training . . . . .	119
Learning: The Missing Key to True Intelligence . . . . .	120
Transition: The Bridge to Learning AI . . . . .	121
Key Takeaway: The Shift from Rules to Learning . . . . .	121
<b>8 First ML Model: Linear Regression for Evaluation</b>	<b>123</b>
Prologue: Linear Regression – A Story Rooted in the Stars . . . . .	123
Types of Machine Learning: Understanding the Landscape . . . . .	125
Building a Training Dataset – Teaching Through Experience . . . . .	127
The Supervised Learning Pipeline – From Data to Understanding . . . . .	129

Writing the Feature Extractor and Data Generator . . . . .	130
The FeatureLogger Class . . . . .	131
Integrating with Connect4Arena . . . . .	132
Running the Generator . . . . .	133
Training a Linear Regression Model . . . . .	133
Using the Learned Model for Game Evaluation . . . . .	135
Experiment: Comparing Against Simulation-Based AIs . . . . .	137
Rebuilding Linear Regression from Scratch . . . . .	139
What Are We Really Doing? . . . . .	139
The Math Behind It (Gently) . . . . .	140
The Code (Using Only NumPy) . . . . .	141
Why This Matters . . . . .	141
Where We Go From Here . . . . .	142
<b>9 Deep Learning: A New Era</b>	<b>143</b>
Prologue: The Man Who Never Gave Up, Geoffrey Hinton’s Long Road . .	143
The Vision Behind Hinton’s Conviction . . . . .	144
From Outsider to Prophet . . . . .	144
Honors, Recognition, and a Legacy Forged in Patience . . . . .	145
The Shock of Depth – The AlexNet Moment . . . . .	145
The Secret Sauce? Depth, Data, and the Right Hardware . . . . .	146
A Shockwave Felt Around the World . . . . .	147
A Shockwave Felt Around the World . . . . .	147
Why Machine Learning Stalled Before Deep Learning . . . . .	147
The Blow That Changed a Generation . . . . .	148
The 3-Layer Myth: A Cautionary Tale of Mathematical Misuse . . .	149
Why Deep Learning Took So Long . . . . .	150
Why Deep Learning Works So Well . . . . .	151
The Training Revolution: Backpropagation + Gradient Descent . . . . .	153
Gradient Descent: The Mathematics of Learning . . . . .	153
Backpropagation: The Algorithm That Should Have Changed Every- thing . . . . .	154
Why It Took So Long . . . . .	154

ReLU: The Silent Hero of Deep Learning . . . . .	155
The Final Push: Data, Compute, and the Spark That Lit the Fire . . . . .	156
The Rise of Big Data: The Internet Becomes a Dataset . . . . .	156
The Accidental Savior: GPUs and the Return of Parallelism . . . . .	157
Open-Source Tools: From Obscurity to Accessibility . . . . .	157
A Perfect Storm . . . . .	158
Chapter Reflection: From Theory to Transformation . . . . .	159
Deep Learning in Action . . . . .	160
The Deep Learning Workflow: From Data to Decisions . . . . .	160
Step 1: Data Collection and Labeling . . . . .	161
Step 2: Preprocessing and Representation . . . . .	162
Step 3: Model Architecture . . . . .	164
Step 4: Training the Network . . . . .	165
Step 5: Evaluation and Integration . . . . .	166
Recap: The Deep Learning Pipeline in Five Steps . . . . .	168
Our First Deep Net: A Simple MLP That Learns to Play . . . . .	168
Data Collection & Labeling . . . . .	169
Preprocessing & Representation . . . . .	171
Model Architecture: Choosing the Right Network . . . . .	171
Training the Network . . . . .	172
Evaluation & Integration . . . . .	175
Code: A Full Deep Learning-Backed AI . . . . .	176
Running a Match . . . . .	177
From First Steps to Real Practice: Elevating Our Training Pipeline . . . . .	179
Training 2.0: Smarter Logs, Sharper Learning . . . . .	180
Smarter Self-Play with Fused Labels . . . . .	180
Cleaning the Data: One Position, One Voice . . . . .	183
Train-Validation Split: Learning with Feedback . . . . .	185
Smarter Training: Early Stops, Best Models, and Multiple Brains . . . . .	187
Results: More Brains, Smarter Play . . . . .	191
Same Brain, New Experience: mlp_126_64_1 vs mlp_64.pth . . . . .	191
Performance of Other Models: How Big Is too Big? . . . . .	192
Sweet Spot Found: 2-3 Hidden Layers, Moderate Width . . . . .	192

Diminishing Returns: And Why That's OK . . . . .	193
Takeaway: The Sweet Spot, for Now . . . . .	194
<b>10 Seeing the Board: CNNs and Spatial Intelligence</b>	<b>195</b>
Prologue – Yann LeCun and the Neural Network That Learned to Perceive	195
A New Kind of Perception: What CNNs Really Do . . . . .	197
AlexNet: The Breakthrough That Lit the Sky . . . . .	198
VGG: Depth with Elegance . . . . .	199
ResNet: Cracking the Barrier of Depth . . . . .	200
Why MLPs Flatten the Board, and Why That's a Problem . . . . .	202
Seeing Like a Player: A CNN's View of the Board . . . . .	204
Building a CNN for Connect 4 Evaluation . . . . .	205
Translating the Board into Sight . . . . .	206
A Concrete Example . . . . .	207
A Canvas Built for Spatial Learning . . . . .	208
Building the CNN Architecture . . . . .	209
VGG3: The Classic Stack . . . . .	209
ResNet6 – Deeper with Skip Connections . . . . .	210
Two Designs, One Goal . . . . .	211
Complete Code for VGG3 and ResNet6 . . . . .	211
Training CNNs . . . . .	213
Reusing the Self-Play Dataset . . . . .	213
Reusing the Training Loop . . . . .	213
Full Training Script . . . . .	214
Training Results . . . . .	216
Saved Models . . . . .	216
Testing CNNs . . . . .	217
Match Setup . . . . .	217
How CNNs Are Integrated . . . . .	217
Match Results: CNNs vs. Everyone Else . . . . .	218
What We Learned from the Matches . . . . .	218
Deep MLP vs. CNN – A Deeper Dive . . . . .	219
Parameter Count . . . . .	219

Inference Cost: Multiply-Accumulate Operations . . . . .	219
Real-World Performance . . . . .	220
Summary Table . . . . .	220
The Big Shift . . . . .	222
How CNNs Changed the Game . . . . .	223
SSmaller Models, Deeper Insight . . . . .	223
From Pattern Recognition to Real Intelligence . . . . .	223
What Comes After Seeing? . . . . .	224
<b>11 Mastering Reinforcement Learning</b>	<b>225</b>
Prologue: Richard Sutton & the Birth of Learning from Experience . . . . .	225
The Backgammon Moment . . . . .	226
The Sutton–Silver Lineage . . . . .	227
A Turing-Worthy Vision . . . . .	229
A Legacy That Learns . . . . .	230
Games: The Perfect Playground . . . . .	231
Breaking It Down: The Building Blocks of Reinforcement Learning . . . . .	231
Two Paths to Mastery: How Agents Learn to Act . . . . .	233
Learning to Judge: The Value-Based Mindset . . . . .	234
Learning to Act: The Policy-Based Mindset . . . . .	235
The Best of Both Worlds: Value Meets Policy . . . . .	235
Learning from the Next Step: TD(0) . . . . .	236
A Smarter Way to Learn: Combining TD(0) with Lookahead . . . . .	237
Closing the Loop: From Self-Play to Self-Improvement . . . . .	238
The Training Loop: How Our Agent Learns from Itself . . . . .	239
Turning Experience into Training: Building the Dataset and Updating the Model . . . . .	241
Full Listing of the Training Script . . . . .	244
A Note on Terminology: Our Setup vs. Standard RL . . . . .	248
Results: TD(0) vs. the Rest . . . . .	249
TD(0) from Scratch: Learning from Pure Experience . . . . .	249
TD(0) with a Head Start: Fine-Tuning the Supervised ResNet6 . . . . .	250
A Note on Search Depth, Data Limits, and the Road Ahead . . . . .	251



A New Direction: Learning to Act, Not Just Evaluate . . . . .	252
The Dual-Head ResNet6: A Network That Judges and Acts . . . . .	253
Key Changes to the Training Loop . . . . .	255
Cloning Knowledge: Loading from the Value-Only Model . . . . .	255
Labeling Both Value and Policy Targets . . . . .	256
Dual Loss: Value + Policy . . . . .	256
Replay Buffer Support . . . . .	257
Full Listing: chapter11_rl_dual.py . . . . .	258
Dual-Head Model Results . . . . .	264
What These Results Reveal . . . . .	264
Interpreting the Results: Value vs. Policy . . . . .	265
<b>12 Monte Carlo Tree Search (MCTS) &amp; RL Synergy</b>	<b>267</b>
Prologue: The Professor Who Wanted Machines to Act . . . . .	267
The Birth of REINFORCE: Learning Through Interaction . . . . .	268
Bridging the Gap: From Theory to Practice . . . . .	269
Reflecting on a Legacy . . . . .	269
From Fragile Feedback to Strong Signals . . . . .	270
Rémi Coulom: The Quiet Revolutionary . . . . .	270
A Detour into Confidence: The Origins of UCB1 . . . . .	273
The Casino Thought Experiment . . . . .	273
Enter UCB1: Optimism in the Face of Uncertainty . . . . .	273
From Bandits to Trees . . . . .	274
Why This Matters . . . . .	275
How MCTS Works: The Four Pillars of Smart Search . . . . .	275
Phase 1: Selection, Walk the Tree . . . . .	275
Phase 2: Expansion, Grow a New Leaf . . . . .	276
Phase 3: Simulation, Roll the Dice . . . . .	276
Phase 4: Backpropagation, Remember What Worked . . . . .	276
What Emerges: A Policy with Insight . . . . .	277
MCTS as a Policy Generator . . . . .	277
From Search to Supervision . . . . .	278
Why This Works So Well . . . . .	278

Replacing Gradients with Games . . . . .	279
Implementing MCTS for Connect 4 . . . . .	279
Step 1: The MCTS Node . . . . .	279
Step 2: Selection Using UCB1 . . . . .	280
Step 3: Expansion . . . . .	282
Step 4: Simulation (Rollout) . . . . .	283
Step 5: Backpropagation . . . . .	284
Putting It All Together . . . . .	285
MCTS random player . . . . .	286
Evaluation Results: MCTS vs Previous AIs . . . . .	287
MCTS (Random Rollout) vs Simulation AI . . . . .	287
Comparing Simulation and MCTS Against Other Players . . . . .	288
Final Thoughts . . . . .	288
MCTS with ResNet6 as a Value Function . . . . .	289
Replacing Simulation with ResNet6 . . . . .	289
Evaluating MCTS with ResNet6 as the Value Function . . . . .	290
MCTS with both Value and Policy Network– Toward AlphaZero . . . . .	292
Integrating Policy into MCTS with PUCT . . . . .	293
Integrating the Dual-Head ResNet6 into MCTS . . . . .	293
1. Nodes with Memory and Belief . . . . .	294
2. Smarter Selection with PUCT . . . . .	294
3. Expansion with Neural Guidance . . . . .	294
4. Backpropagation: No Change Needed . . . . .	295
5. The New MCTS Loop . . . . .	295
Results from MCTS-DualNet – AlphaZero Without the Learning . . . . .	296
Self-Play Training: Teaching MCTS to Learn . . . . .	298
Code Changes for Self-Play Training . . . . .	298
Results – After Self-Play Training . . . . .	304
Chapter Review – From Search to Self-Improvement . . . . .	306
Looking Ahead – From Decision to Generation . . . . .	307
<b>13 The Generative Leap</b>	<b>308</b>
Prologue: The First Tokens . . . . .	308

But Behind Every Leap, There’s a Hidden Engine . . . . . 310

Two Ideas, One Revolution . . . . . 311

When Size Became Power: The GPT-3 Shockwave . . . . . 311

The ChatGPT Moment: When the World Met Generative AI . . . . . 312

The Three Eras of Generative AI . . . . . 313

    1. The Imitation Era (2018–2019) . . . . . 313

    2. The Instruction Era (2020–2023) . . . . . 313

    3. The Reasoning Era (2024– ) . . . . . 314

Hall of Innovation: The Rise of Generative AI . . . . . 315

The Cyclical Nature of AI Development . . . . . 316

    What’s Old Is New Again . . . . . 316

    The Arc Echoes Again . . . . . 317

    History Doesn’t Repeat, It Refines . . . . . 317

    Why This Matters for Us . . . . . 318

Letting ChatGPT Play Connect 4 . . . . . 318

    Game 1: ChatGPT (X) vs AlphaBeta (O, depth=1) . . . . . 319

    Game 2: ChatGPT (O) vs AlphaBeta (X, depth=1) . . . . . 321

    Reflections on GPT as a Player . . . . . 322

    Where GPT Struggles (and Why) . . . . . 323

Building BabyGPT from Scratch . . . . . 324

    Revisiting GPT, One Word at a Time . . . . . 325

    How Self-Attention Works, And Why It Matters . . . . . 326

    What We’re About to Build: A Map of BabyGPT . . . . . 328

BabyGPT . . . . . 328

    Translating Moves into Meaning: Writing the Tokenizer . . . . . 329

    Building a Thinking Machine: Defining the BabyGPT Model . . . . . 331

    Feeding the Mind: Preparing BabyGPT’s Training Data . . . . . 333

    Learning to Play: Training BabyGPT Step by Step . . . . . 336

    Can It Play? Evaluating BabyGPT’s Understanding . . . . . 339

BabyV: BabyGPT with Vision . . . . . 341

    Reading the Board: Designing BabyV’s Input Encoding . . . . . 342

    Seeing and Deciding: Defining the BabyV Model . . . . . 343

    What the Board Tells Us: Preparing Training Data for BabyV . . . . . 345

Learning from Sight and Memory: Training BabyV . . . . .	348
Smarter by Sight? Evaluating BabyV’s Gameplay . . . . .	350
What BabyV Learned . . . . .	351
Afterword: The Road to Volume II . . . . .	352
<b>Volume II Foreword – The New Discipline of AI Engineering</b>	<b>354</b>
<b>Volume II Preface – The Master’s Workshop</b>	<b>359</b>
<b>Bonus Chapter —Optimizing for the Future</b>	<b>360</b>
<b>14 Optimizing for the Future: From Python to Peak C</b>	<b>361</b>
Prologue: John Hennessy & David Patterson . . . . .	361
PCA <sup>3</sup> : Performance-Conscious, Architecture-Aware, AI-Aligned Program- ming. . . . .	363
How Modern CPUs Think . . . . .	364
Memory Hierarchy: The Hidden Latency Monster . . . . .	365
Bonus Complexity: NUMA, GPUs, and Distributed Memory . . . . .	366
Registers & Cache Reuse: The Golden Zone . . . . .	367
Bad Layout (Scattered, Cache-Hostile) . . . . .	368
Good Layout (Flat, Cache-Friendly) . . . . .	368
Instruction-Level Parallelism & Pipelining . . . . .	369
Branches & Misprediction: Control Flow’s Hidden Tax . . . . .	372
Mini Benchmark: Branch vs Branchless . . . . .	374
Libraries & Dynamic Memory: The Hidden Costs . . . . .	375
Why Tiny Allocations Are Dangerous . . . . .	376
The Better Model: Bulk Allocation + Reuse . . . . .	376
A Note on Modern Languages: What We Don’t See Can Hurt Us . . .	377
Why C Still Wins . . . . .	378
Architecture-Aware Programming: Key Principles to Remember . . .	378
PCA <sup>3</sup> in Practice: From Python to Peak C . . . . .	379
Python Baseline: Easy to Write, Painful to Scale . . . . .	380
It’s Not Python vs. C, It’s What Each Language Encourages . . . . .	382
C Version 0: A Straightforward Rewrite . . . . .	383

Full Code: <i>chapter14_sim_v0.c</i> . . . . .	383
Performance: A 300× Speedup, Without Even Trying . . . . .	385
Version 1: Struct Layout & Cache Awareness . . . . .	386
The New Shape of the Game . . . . .	386
Summary: What PCA <sup>3</sup> Gave Us . . . . .	388
Performance Gains: ~3× from Better Layout . . . . .	388
Version 2 – Loop & Logic Cleanup + Smarter Board Layout . . . . .	391
Introducing the 8×8 Extended Board . . . . .	391
Cleaner Logic, Hotter Loops . . . . .	392
What We Learn . . . . .	392
Summary: What Changed in Version 2 . . . . .	393
Performance Gains: From 0.817 → 0.497 sec . . . . .	393
Version 3 – Branchless Win Detection . . . . .	395
Branchless Win Checking . . . . .	396
Why This Matters . . . . .	397
What We Learn . . . . .	397
Summary: What Changed in Version 3 . . . . .	397
Performance: Slight Regression (but an Investment) . . . . .	397
What We Learn . . . . .	398
Version 4 – Fast Random Number Generation . . . . .	400
Our Custom RNG: xorshift128+ . . . . .	401
A word About xorshift128+ . . . . .	401
Why Use Only 3 Bits at a Time? . . . . .	402
What We Learn . . . . .	403
Version 5 – OpenMP Parallelization . . . . .	405
A Brief History of OpenMP . . . . .	406
OpenMP in Action: Parallelism Made Simple . . . . .	406
What This Means . . . . .	407
Performance: From 0.349 → 0.115 sec (8-core MacBook Air M1) . . . . .	407
Measuring the Right Time: Wall Clock vs. CPU Time . . . . .	408
What We Learn . . . . .	408
Version 6 – Bitboard Encoding: Compact, Fast, and Cache-Aligned . . . . .	411
How Do Bitboards Represent a Connect 4 Game? . . . . .	412

Win Detection with Bitboards: Bitwise Brilliance . . . . .	413
A Simpler, Smarter RNG Buffer . . . . .	414
Performance: From 0.115 $\rightarrow$ 0.076 sec (single-threaded) . . . . .	415
PCA <sup>3</sup> in Action . . . . .	416
Full Listing . . . . .	416
Version 7 – Pthreads and Complete Control . . . . .	419
A Short History of Pthreads . . . . .	419
From Fork-Join to Persistent Threads . . . . .	419
Thread Safety and TLS . . . . .	420
Why pthreads Was Worth It . . . . .	420
Full Listing . . . . .	420
Final Performance: Peak Speed Achieved . . . . .	424
What We Learn . . . . .	424
Chapter 14 Wrap-Up – From 101 Years to 1 Day . . . . .	425
From 101 Years to 1 Day: Simulation Performance Evolution . . . . .	426
Final Takeaway . . . . .	426
<b>Bibliography – Volume I: From Foundations to Frontiers</b>	<b>428</b>
Chapter 1 . . . . .	428
Chapter 2 . . . . .	428
Chapter 3 . . . . .	429
Chapter 4 . . . . .	429
Chapter 5 . . . . .	429
Chapter 6 . . . . .	430
Chapter 7 . . . . .	430
Chapter 8 . . . . .	430
Chapter 9 . . . . .	431
Chapter 10 . . . . .	431
Chapter 11 . . . . .	431
Chapter 12 . . . . .	432
Chapter 13 . . . . .	432
Chapter 14 (Bonus Chapter) . . . . .	433

---

About the Author	437
Back Cover Description	438

# JOIN THE AI FORGE COMMUNITY

Scan to Join the Community



Visit [pca3press.com/ai-forge](https://pca3press.com/ai-forge) to:

- Register your copy of *The AI Forge*
- Access source code and bonus content
- Join the private reader community
- Leave a review or share feedback
- Get early access to future chapters and tools

*This book is just the beginning.  
Let's forge the future of AI together.*



## ABOUT THE AUTHOR

Dr. Jianxian Wu is a researcher, engineer, and lifelong practitioner of intelligence in all its forms. Trained in artificial intelligence during its coldest winter, he has spent decades exploring the boundaries between logic, search, learning, and computation, pushing the frontlines of what machines can understand and become.

His work bridges high-performance, high-efficiency systems and conceptual clarity, blending history, theory, and code into a deeper understanding of AI's foundations and future.

The AI Forge is both his tribute to the pioneers who came before, and a toolkit for the builders who come next.

# BACK COVER DESCRIPTION

What does it take to build true artificial intelligence, to understand its roots, its meaning, and its craft?

Welcome to **The AI Forge**, where curiosity leads to creation, and the spirit of craftsmanship breathes life into code.

This book invites you on a journey through the living history of AI. You will walk alongside pioneers like Alan Turing, Arthur Samuel, and Ken Thompson, discovering the ideas that reshaped how machines learn, reason, and evolve. At every step, you will explore the physical interpretations behind each breakthrough, seeing not only how key technologies work, but understanding the forces that make them possible.

By blending historical context, hands-on building, and deep conceptual insight, The AI Forge offers more than knowledge. It forges skill and understanding together, preparing you to think, design, and build with clarity and purpose.

From rule-based thinking to deep learning, from simple games to scalable systems, you will develop real fluency, beginning with Python to explore key ideas, and advancing to C to bring them to life with performance-conscious, architecture-aware, AI-aligned craftsmanship. Each chapter builds not just technical ability, but a craftsman's grasp of why AI works the way it does, and what it can become.

Whether you are a student preparing for an AI-driven world, a professional seeking deeper mastery, or a lifelong learner captivated by intelligence itself, The AI Forge invites you to trace the roots, ride the current, and forge the future.