

Contents

List of Contributors xv

About the Authors xxi

Part I Fundamentals 1

1 What Is Neurosymbolic AI? An Overview and Frontier Problems 3

Alvaro Velasquez, Lucas White, Patrick Cooper, Antony Zhao, and Lekai Chen

1.1 Introduction 3

1.2 Neurosymbolic Artificial Intelligence 4

1.2.1 Explicit to Implicit: From Symbolic Representations to Neural Networks 5

1.2.2 Implicit to Explicit: From Neural Networks to Symbolic Representations 6

1.3 Frontiers Problems 7

1.3.1 Neurosymbolic AI for Synthetic Biology 7

1.3.2 Neurosymbolic AI for Robust Autonomy 9

1.3.3 Neurosymbolic AI for Creative Scientific Discovery 11

1.4 Conclusion 11

References 12

2 Reasoning in Neurosymbolic AI 15

Son Tran, Edjard Mota, and Artur d'Avila Garcez

2.1 What Is Reasoning in Neural Networks? 15

2.1.1 Reasoning in LLMs 16

2.1.2 AI from a Neurosymbolic Perspective 19

2.2 Background: Logic and RBMs 21

2.2.1 Illustrating Logical Reasoning with the Sudoku Puzzle 23

2.2.2	Sudoku with Strategies of Sampling	26
2.2.3	Restricted Boltzmann Machines	27
2.3	Symbolic Reasoning with Energy-based Neural Networks	28
2.3.1	Related Work	28
2.3.2	Knowledge Representation in RBMs	30
2.3.3	Reasoning in RBMs	33
2.3.4	Logical Boltzmann Machines	36
2.3.5	Experimental Results	39
2.3.6	Extensions of LBMs	43
2.4	LBMs for MaxSAT	49
2.4.1	LBM with Dual Annealing	52
2.4.2	Experimental Results of LBM for MaxSAT	52
2.5	Integrating Learning and Reasoning in LBMs	54
2.6	Challenges for Neurosymbolic AI	57
2.6.1	Nonmonotonic Logic	58
2.6.2	Planning	58
2.6.3	Learning from Its Mistakes	59
2.7	Conclusion	60
	References	62
3	Neurosymbolic Assurance Using Concept Probes in Foundation Models	69
	<i>Ramneet Kaur, Anirban Roy, and Susmit Jha</i>	
3.1	Introduction	69
3.2	Neural Features and Concept Probes	71
3.3	Foundation Models as Specification Lens	72
3.4	Symbolic Specification of ML Models Using Concept Probes	75
3.5	Implementation and Evaluation	78
3.6	Conclusion and Open Challenges	86
	References	87
4	Toward Assured Autonomy Using Neurosymbolic Components and Systems	89
	<i>Abhishek Dubey, Taylor T. Johnson, Xenofon Koutsoukos, Baiting Luo, Diego Manzanas Lopez, Miklos Maroti, Ayan Mukhopadhyay, Nicholas Potteiger, Serena Serbinowska, Daniel Stojcsics, Yunuo Zhang, and Gabor Karsai</i>	
4.1	Introduction	89
4.2	Problem Formulation and Challenges: Maneuver Control for Autonomous Vehicles	90
4.3	Software Architecture: Components and Interactions	91

4.4	Probabilistic World Model	93
4.4.1	Obstacle Map Calculation	94
4.4.2	Reward Map Calculation	96
4.5	Planner	97
4.5.1	Formalization	98
4.5.2	Online Planning Through Monte Carlo Search	98
4.5.3	Scalability Through Hierarchical Planning	100
4.5.4	Evaluation and Analysis	101
4.5.5	Neurosymbolic Extensions for Planning Under Partial Observability	101
4.6	Trajectory Control with Evolving Behavior Trees (EBTs)	103
4.6.1	Safe Autonomous UAV Navigation	103
4.6.2	Safe EBTs for Navigation	104
4.6.3	Evaluation	106
4.7	Assurance for Neurosymbolic Systems	108
4.7.1	Neurosymbolic Verification with BehaVerify	109
4.7.2	Assurance on Grid Abstractions	111
4.7.3	Timing Results and Conclusions	112
4.7.4	Future Work	113
4.8	Conclusions	114
	References	115
5	Safe Neurosymbolic Learning and Control	119
	<i>Somil Bansal and Jaime F. Fisac</i>	
5.1	Problem Setup	119
5.1.1	Dynamical Safety Problem	120
5.1.2	Running Example: Air Collision Avoidance	122
5.2	Hamilton–Jacobi (HJ) Reachability	123
5.2.1	Methods to Solve HJI-VI and Compute Unsafe Set	126
5.2.2	Running Example: Air Collision Avoidance	127
5.3	A Neurosymbolic Perspective on Learning Safe Controllers	12
5.3.1	Self-supervised Neurosymbolic Learning for Synthesizing Safe Controllers	129
5.3.2	Neurosymbolic Reinforcement Learning for Synthesizing Safe Controllers	135
5.3.3	Connections Between Reinforcement and Self-supervised Neurosymbolic Learning	143
5.4	Safety Assurances for Learned Controllers	144
5.4.1	Probabilistic Safety Assurances Through Conformal Prediction	145

5.4.2	Robust Safety Assurances Through Forward Reachability	148
5.5	Frontiers, Open Questions, and Promising Directions	150
	References	151
6	Controllable Generation via Locally Constrained Resampling	159
	<i>Kareem Ahmed, Kai-Wei Chang, and Guy Van den Broeck</i>	
6.1	Introduction	159
6.2	Background	160
6.2.1	Notation and Preliminaries	160
6.2.2	A Probability Distribution over Sentences	161
6.2.3	The State of Conditional Sampling	162
6.3	Locally Constrained Resampling: A Tale of Two Distributions	163
6.3.1	Inducing a Local Tractable Distribution	164
6.3.2	Tractable Operations via Compilation	165
6.3.3	Intermezzo: Constraint Circuits and DFAs	168
6.3.4	Correcting Sample Bias: Importance Sampling... and Resampling	168
6.4	Related Work	170
6.5	Experimental Evaluation	171
6.6	Conclusion and Future Work	175
	Appendix A Controllable Generation via Locally Constrained Resampling	175
A.1	Language Detoxification	175
A.2	Sudoku	176
A.3	Warcraft Shortest Path	176
A.4	Broader Impact	177
	References	177
7	Tractable and Expressive Generative Modeling with Probabilistic Flow Circuits	183
	<i>Sahil Sidheekh and Sriraam Natarajan</i>	
7.1	Introduction	183
7.2	Tractable Probabilistic Modeling	188
7.2.1	Inference Queries	189
7.2.2	The Expressivity-tractability Trade-off	190
7.3	Probabilistic Circuits	191
7.3.1	Defining a Probabilistic Circuit	192
7.3.2	Structural Properties	193
7.3.3	Tractable Inference with PCs	194
7.3.4	Parameter Learning for PCs	195
7.3.5	Structure Learning for PCs	195

7.4	Normalizing Flows: A Primer	197
7.4.1	Sampling and Inference in Flows	199
7.5	Integrating Normalizing Flows and PC	200
7.5.1	The Challenge	200
7.5.2	τ -Decomposability	201
7.6	Probabilistic Flow Circuits	205
7.7	Experiments and Results	210
7.7.1	Modeling Complex 3D Manifolds	211
7.7.2	Scaling to High-dimensional Data	212
7.7.3	Sample Generation and Inference	215
7.7.4	Ablation: Influence of PC Complexity	215
7.8	Conclusion and Discussion	216
7.8.1	Key Takeaways	217
7.8.2	Limitations and Future Directions	217
	Acknowledgements	218
	References	219
8	Toward Verifiable and Scalable In-context Fine-tuning in Neurosymbolic AI	223
	<i>Neel P. Bhatt, Alvaro Velasquez, Zhangyang Wang, and Ufuk Topcu</i>	
8.1	Introduction	223
8.2	Neurosymbolic Fine-tuning Using Automated Feedback from Formal Verification	225
8.2.1	Introduction	225
8.2.2	Preliminaries	226
8.2.3	Methodology	227
8.2.4	Experimental Results	232
8.2.5	Conclusion	240
8.3	Uncertainty-aware Fine-tuning and Inference for Multimodal Foundation Models	242
8.3.1	Introduction	242
8.3.2	Conformal Prediction	242
8.3.3	Perception Uncertainty	244
8.3.4	Decision Uncertainty	245
8.3.5	Estimating Decision Uncertainty Score	248
8.3.6	Targeted Interventions	248
8.3.7	Experiments	251
8.3.8	Automated Refinement	253
8.3.9	Conclusion	257
8.4	Toward a Hybrid Architecture: Dynamic Interleaving of Neural and Symbolic Reasoning	257

8.5	Conclusion and Future Directions	260
8.5.1	Extending the Scope: Symbolic Tool Use for Mathematical Reasoning	261
	References	262
	Part II Advanced Topics	267
9	Physics-informed Deep Learning	269
	<i>Nithin Chalapathi, Yiheng Du, Sanjeev Raja, and Aditi S. Krishnapriyan</i>	
9.1	Introduction	269
9.1.1	Data Generation in Physics-informed Machine Learning	271
9.1.2	Architectures	274
9.1.3	Training Objectives	282
9.1.4	Open Challenges	288
9.1.5	Connections to Atomistic Modeling	289
	References	291
10	Causal Representation Learning	307
	<i>Burak Varici, Chandler Squires, and Pradeep Ravikumar</i>	
10.1	Introduction	307
10.2	Background	310
10.2.1	Model Classes and Identifiability	311
10.2.2	Causal Graphical Models and Interventions	312
10.2.3	Causal Representation Models	314
10.2.4	CRL Identifiability and Equivalence Classes	315
10.3	Interventional CRL	317
10.4	CRL with Linear SCMs	320
10.4.1	Linear Mixing on Linear Latent SCMs	321
10.4.2	General Mixing on Linear Latent SCMs	323
10.5	CRL with General SCMs	324
10.5.1	Linear Mixing on General Latent SCMs	326
10.5.2	Multi-node Interventions	330
10.5.3	General Mixing on General Latent SCMs	332
10.6	Experiments	335
10.6.1	Linear Mixing with Synthetic Data	336
10.6.2	Experiments on Image Data	337
10.7	Other Approaches	339
10.8	Summary	340
	References	341

11	Neurosymbolic Computing: Hardware–Software Co-design	347
	<i>Xiaoxuan Yang, Zhangyang Wang, Miroslav Pajic, Hai “Helen” Li, Yiran Chen, X. Sharon Hu, Chris H. Kim, Shimeng Yu, and Rajit Manohar</i>	
11.1	Introduction	347
11.2	Background	348
11.2.1	Neurosymbolic Artificial Intelligence	348
11.2.2	Emerging Hardware Computing Platforms	350
11.3	Trends and Challenges	351
11.3.1	Enhance Reasoning and Generalization	351
11.3.2	Enable Compositionality	352
11.3.3	Handle Uncertainty	353
11.3.4	Improve System Efficiency	354
11.3.5	Demonstrate Full-stack NeSy Systems	354
11.4	Applications and Future Topics	355
11.5	Conclusions	356
	References	356
12	Programmatic Reinforcement Learning	365
	<i>Swarat Chaudhuri</i>	
12.1	Introduction	365
12.2	Programmatic RL	367
12.3	Imitation-projected Policy Gradients	369
12.4	Related Work	373
12.5	Conclusion	374
	References	376
	Part III Applications	381
13	From Symbolic to Neurosymbolic Information Extraction	383
	<i>Mihai Surdeanu, Marco A. Valenzuela-Escárcega, Gus Hahn-Powell, Robert Vacareanu, Gwendolen Herongrove, Enrique Noriega-Atala, Özgün Babur, Emek Demir, and Clayton T. Morrison</i>	
13.1	Motivation and Overview	383
13.2	An Example of Symbolic IE	386
13.2.1	Introduction	386
13.2.2	Approach	387
13.2.3	Intrinsic Evaluation: Machine Reading Performance	394
13.2.4	Extrinsic Evaluation: Discovery of Biological Hypotheses	396
13.2.5	Conclusion	401

13.3	Problems of Symbolic IE Systems	401
13.4	Generating Rules	402
13.4.1	Introduction	402
13.4.2	Approach	403
13.4.3	Evaluation	405
13.4.4	Conclusion	409
13.5	Matching Rules	409
13.5.1	Introduction	409
13.5.2	Approach	411
13.5.3	Evaluation	415
13.5.4	Conclusion	421
13.6	Take Away	421
	References	422
14	Neurosymbolic AI for Legal AI-TRISM: Trustworthy, Reliable, Interpretable, Safe Models	429
	<i>Deepa Tilwani, Yash Saxena, Ankur Padia, Srinivasan Parthasarathy, and Manas Gaur</i>	
14.1	Introduction	429
14.1.1	Neurosymbolic RAG	431
14.1.2	Advantages of Using Neurosymbolic RAG	432
14.2	Limitation of Using LLM as Legal Assistant	433
14.3	Neurosymbolic AI for Legal Domain	434
14.4	AI-TRISM with Neurosymbolic AI	436
14.4.1	KG Construction	436
14.4.2	Graph Construction Methodology	437
14.5	Symbiosis of LLM and KG for Neurosymbolic RAG in Legal Domain	439
14.6	Related Work	442
14.6.1	KG Construction	442
14.6.2	Legal Classification	444
14.6.3	Legal Question Answering	444
14.6.4	Legal Article and Case Retrieval	445
14.6.5	Citation Recommendation and Interoperability	445
14.6.6	Other Related Work	446
	Acknowledgement	447
	References	447
	Index	455