

Contents

1 Gold-Style Learning Theory	1
John Case	
1.1 Language Learnability: Gold 1967 and Beyond	2
1.1.1 Gold's 1967 Model of Child Language Learning	2
1.1.2 Some Criteria of Successful Learning	3
1.1.3 Constraints on Learnability	4
1.1.4 Characterizations and Pattern Languages	5
1.1.5 Insensitivities to Presentation Order	6
1.1.6 Empirical U-Shaped Learning	8
1.1.7 Formal U-Shaped Learning	8
1.1.8 Results/Question Regarding U-Shaped Learning.	9
1.2 Complexity Considerations	10
1.2.1 Hurling Databases and Memory-Limited Learners	10
1.2.2 Deficiencies of Programs Learned	12
1.2.3 Complexity of Learner Updates	16
1.3 Summary Including Open Problems	19
References	20
2 Efficiency in the Identification in the Limit Learning Paradigm.	25
Rémi Eyraud, Jeffrey Heinz and Ryo Yoshinaka	
2.1 Introduction.	25
2.1.1 The Importance of Efficiency in Learning	25
2.1.2 Preliminary Definitions	26
2.2 PAC Learning and Other Learning Paradigms	28
2.2.1 PAC Paradigm	28
2.2.2 Other Learning Paradigms	29
2.3 The Limits of Gold's Paradigm	30
2.3.1 Identification in the Limit	30

2.3.2	Polynomial Time	32
2.3.3	Identification of a Language and the Size of a Target Representation	33
2.4	First Refinements	34
2.4.1	Mind Changes and Implicit Errors of Prediction	34
2.4.2	Characteristic Sample	35
2.5	Recent Refinements	39
2.5.1	Structurally Complete Set	39
2.5.2	Thickness	40
2.5.3	Comparison of the Two Refinements	41
2.6	Conclusion	42
	Appendix	43
	References	44
3	Learning Grammars and Automata with Queries	47
	Colin de la Higuera	
3.1	Introduction.	47
3.2	Definitions and Notations	49
3.3	Motivations and Applications	50
3.4	Query Learning Models	51
3.4.1	The Oracle	52
3.4.2	The Different Types of Queries	52
3.4.3	Query Learnability	56
3.4.4	PAC Query Learning	58
3.5	Some Important Hardness Results	59
3.5.1	Learning with Membership Queries and Little Else	59
3.5.2	Learning with Equivalence Queries	59
3.6	Learning DFA from an MAT	61
3.6.1	LSTAR.	61
3.6.2	Trading Off Equivalence Queries	64
3.6.3	Adapting to Correction Queries	65
3.6.4	Some Implementation Issues	65
3.7	Extensions.	65
3.7.1	Post Zulu Extensions	65
3.7.2	Co-learning.	66
3.7.3	Learning Other Types of Grammars	66
3.7.4	More About Querying	67
3.8	Conclusions.	68
	References	68
4	On the Inference of Finite State Automata from Positive and Negative Data	73
	Damián López and Pedro García	
4.1	Introduction.	73
4.2	Definitions and Notation	76

4.3	Inference of Deterministic Automata.	77
4.3.1	Non-merging Algorithms: Trakhtenbrot–Barzdin and Gold	77
4.3.2	Merging Algorithms: <i>RPNI</i> and Lang	80
4.3.3	Algorithms Guided by Heuristics: <i>EDSM</i> and <i>Blue-Fringe</i>	83
4.3.4	Inference of Teams of Automata	88
4.3.5	Identification in the Limit	89
4.4	Inference of Non-deterministic Automata	92
4.4.1	Inference of Canonical NFAs	93
4.4.2	Inference by Juxtaposition of Automata	98
4.4.3	Order-Independent Merging Inference.	100
4.5	(Experimental) Comparison of Different Approaches	105
4.6	Conclusions.	109
	References	110
5	Learning Probability Distributions Generated by Finite-State Machines	113
	Jorge Castro and Ricard Gavalda	
5.1	Introduction.	113
5.2	Preliminaries	115
5.2.1	Learning Distributions in the PAC Framework	116
5.2.2	Identification in the Limit Paradigm	117
5.2.3	Probabilistic Automata	117
5.2.4	Probabilistic Deterministic Automata and Distinguishability.	118
5.2.5	Hidden Markov Models	119
5.2.6	Weighted Automata	119
5.3	A Panoramic View of Known Results.	120
5.4	The Hankel Matrix Approach	122
5.5	Learning PDFA	125
5.5.1	An Oracle Algorithm for Learning DWA	125
5.5.2	State-Merging Learning Algorithms	127
5.6	Learning PFA	132
5.6.1	An Oracle Algorithm for Learning WA	132
5.6.2	The Spectral Method	134
5.6.3	Variations and Implementation of the Spectral Method	136
5.7	Future Work	137
	Appendix: The Baum–Welch Method.	137
	References	139
6	Distributional Learning of Context-Free and Multiple Context-Free Grammars.	143
	Alexander Clark and Ryo Yoshinaka	
6.1	Introduction.	143

6.2	Distributional Learning: A Historical Note.	144
6.3	Languages and Grammars	145
6.3.1	Learning Models	146
6.3.2	Contexts and Distributions	147
6.3.3	Observation Tables	148
6.4	Context-Free Algorithms.	149
6.4.1	Primal-Dual Distinction	149
6.5	Primal Algorithms	150
6.5.1	Congruential Languages	150
6.5.2	Substitutable Languages	154
6.5.3	Finite Kernel Property	155
6.6	Dual Algorithms	155
6.6.1	Context Deterministic Grammars	156
6.6.2	Finite Context Property.	156
6.7	Combined Primal-Dual Methods	157
6.8	Multiple Context-Free Grammars	157
6.8.1	Definition	159
6.8.2	Generalisation of Contexts and Substrings and Observation Tables	162
6.8.3	Substitutability.	163
6.8.4	MAT Result	165
6.8.5	Finite Kernel Property and Finite Context Property	168
6.9	Discussion.	168
6.9.1	Distributional Learning Beyond MCFGs and CFGs	168
6.9.2	Conclusion	169
	References	169
7	Learning Tree Languages.	173
	Johanna Björklund and Henning Fernau	
7.1	Introduction.	173
7.2	Definitions for Tree Languages	175
7.2.1	Trees	175
7.2.2	Operations on Trees	177
7.2.3	Finite Tree Automata	177
7.2.4	Weighted Tree Automata	179
7.2.5	Constructions for Tree Automata	180
7.3	Mathematical Results	182
7.3.1	Gold-Style Learning.	182
7.3.2	Active Learning.	187
7.3.3	Error-Correcting Grammatical Inference (ECGI).	191
7.4	Applications	192
7.4.1	XML Learning	192
7.4.2	Bioinformatics.	195
7.4.3	Computational Linguistics.	197

7.5	Conclusions and Perspectives	199
7.5.1	PAC Learning	199
7.5.2	Support Vector Machines	200
7.5.3	Unranked Trees	200
7.5.4	Infinite Trees	200
7.5.5	Alternative Representations	200
7.5.6	Learning Within or Beyond Regularity?	201
7.5.7	Shock Trees: A Potential Application	201
7.5.8	Beyond Trees	202
	References	202
8	Learning the Language of Biological Sequences	215
	François Coste	
8.1	Linguistic Metaphor	215
8.2	Discovering and Modeling Biological Words	217
8.2.1	Short DNA Words	217
8.2.2	Longer Words	223
8.3	Learning Syntax	234
8.4	Conclusion	241
	References	242