# Contents

Acknowledgements xiii

1 Introduction 1

2 Formal Learning Theory 7
  2.1 History .................................................. 7
  2.2 Basic Definitions ........................................ 11
  2.3 Learning and Convergence ................................ 12
    2.3.1 The Power of Identification in the Limit ............... 14
    2.3.2 The Weakness of Identification in the Limit .......... 14
    2.3.3 Function Learning ..................................... 15
  2.4 Finite Elasticity ........................................ 16
  2.5 Elementary Formal Systems ................................ 19
    2.5.1 Inductive Inference of Monotonic Formal Systems ...... 20
    2.5.2 Context-Sensitive Grammars ............................ 21
    2.5.3 Linear Prolog Programs ................................. 21
  2.6 Constraints on Learning Functions ......................... 22
    2.6.1 Constraints on Environments ...................... 29
  2.7 Variations on Identification in the Limit ............... 30
  2.8 Algorithms for Learning ................................ 30
    2.8.1 Uniform Learning: Synthesizing Learners .......... 31
  2.9 Time Complexity of Learning Functions .................. 32

3 Classical Categorial Grammar 37
  3.1 Basic Definitions ........................................ 37
  3.2 Decidable and Undecidable Questions about Classical Categorial Grammars ........................................ 39
  3.3 Substitutions and Standardizations ......................... 41
    3.3.1 Substitutions ........................................ 41
    3.3.2 Grammars in Reduced Form and Grammars Without Useless Types ........................................ 42

4 Learning Classes of Categorial Grammars 45
  4.1 Learning Rigid Grammars: the Algorithm RG .............. 46
  4.2 Learning \( k \)-Valued Grammars .......................... 48
    4.2.1 Learning Functions Based on \( V_k \) .................. 50
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>$k_{VG}$ is not conservative</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>Least-Valued Grammars</td>
<td>55</td>
</tr>
<tr>
<td>4.5</td>
<td>Optimal Grammars</td>
<td>56</td>
</tr>
<tr>
<td>4.6</td>
<td>Least Cardinality Grammars</td>
<td>57</td>
</tr>
<tr>
<td>4.7</td>
<td>Minimal Grammars</td>
<td>58</td>
</tr>
<tr>
<td>4.8</td>
<td>Learning $k$-valued Grammars from Strings</td>
<td>59</td>
</tr>
<tr>
<td>4.8.1</td>
<td>Algorithms for Learning $k$-Valued Grammars from Strings</td>
<td>60</td>
</tr>
<tr>
<td>4.9</td>
<td>Classes that are not Learnable from Strings</td>
<td>62</td>
</tr>
<tr>
<td>4.10</td>
<td>Summary</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>Complexity Issues</td>
<td>65</td>
</tr>
<tr>
<td>5.1</td>
<td>An Avalanche of Hypotheses</td>
<td>65</td>
</tr>
<tr>
<td>5.2</td>
<td>The Tractability of Producing Consistent Hypotheses</td>
<td>69</td>
</tr>
<tr>
<td>5.2.1</td>
<td>The Node-Cover Problem</td>
<td>70</td>
</tr>
<tr>
<td>5.3</td>
<td>The Complexity of Learning $G_k$-valued</td>
<td>70</td>
</tr>
<tr>
<td>5.4</td>
<td>The Complexity of Learning $G_k$-valued Revisited</td>
<td>75</td>
</tr>
<tr>
<td>5.5</td>
<td>The Complexity of Learning $G_2$-valued</td>
<td>77</td>
</tr>
<tr>
<td>5.6</td>
<td>Consistent Identification in the Limit of Rigid Grammars from Strings is NP-hard</td>
<td>85</td>
</tr>
<tr>
<td>5.7</td>
<td>Conclusions and Further Research</td>
<td>89</td>
</tr>
<tr>
<td>5.8</td>
<td>Detailed Proofs</td>
<td>90</td>
</tr>
<tr>
<td>6</td>
<td>Miscellaneous</td>
<td>95</td>
</tr>
<tr>
<td>6.1</td>
<td>General Combinatory Grammars</td>
<td>95</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Previous Results</td>
<td>96</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Restricting Combinatory Rules for Finite Elasticity</td>
<td>99</td>
</tr>
<tr>
<td>6.2</td>
<td>Learning Generalized Quantifiers</td>
<td>102</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Beyond First Order Generalized Quantifiers</td>
<td>103</td>
</tr>
<tr>
<td>6.3</td>
<td>Tree Adjoining Grammars</td>
<td>106</td>
</tr>
<tr>
<td>6.3.1</td>
<td>TAGs with the Empty String</td>
<td>109</td>
</tr>
<tr>
<td>6.3.2</td>
<td>The Class of Rigid TAGs is Not Learnable</td>
<td>110</td>
</tr>
<tr>
<td>6.4</td>
<td>Minimalist Grammars</td>
<td>112</td>
</tr>
<tr>
<td>6.4.1</td>
<td>The Class of 2-Valued MGs with Empty Categories is Not Learnable</td>
<td>114</td>
</tr>
<tr>
<td>6.5</td>
<td>Conclusions and Future Work</td>
<td>117</td>
</tr>
<tr>
<td>7</td>
<td>Learning Regular Tree Languages</td>
<td>119</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>119</td>
</tr>
<tr>
<td>7.2</td>
<td>Regular Tree Languages</td>
<td>120</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Trees are Terms</td>
<td>120</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Tree Automata</td>
<td>120</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Reversible Regular Tree Languages</td>
<td>121</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Reversible Regular Tree Grammars</td>
<td>121</td>
</tr>
<tr>
<td>7.2.5</td>
<td>Identification of Reversible Tree Languages</td>
<td>122</td>
</tr>
<tr>
<td>7.3</td>
<td>Identification of Reversible Tree Languages</td>
<td>122</td>
</tr>
<tr>
<td>7.3.1</td>
<td>An Algebraic Characterization of Reversible Tree Languages</td>
<td>122</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Characteristic Samples</td>
<td>124</td>
</tr>
<tr>
<td>7.3.3</td>
<td>An Efficient Learning Algorithm</td>
<td>125</td>
</tr>
<tr>
<td>7.4</td>
<td>Learning with Structural Examples</td>
<td>127</td>
</tr>
</tbody>
</table>