Hypertrees: a study in language specification

William Allen Baldwin
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd

Part of the Computer Sciences Commons

Recommended Citation
http://lib.dr.iastate.edu/rtd/8449

This Dissertation is brought to you for free and open access by Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106
Baldwin, William Allen

HYPERTREES -- A STUDY IN LANGUAGE SPECIFICATION

Iowa State University  Ph.D.  1983

University Microfilms International  300 N. Zeeb Road, Ann Arbor, MI 48106
Hypertrees -- A study in language specification

by

William Allen Baldwin

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Computer Science

Approved:
Signature was redacted for privacy.

In Charge of Major Work
Signature was redacted for privacy.

For the Major Department
Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1983
# Table of Contents

| Section                                                        | Page |
|                                                               |      |
| Abstract                                                      | 1    |
| Introduction                                                  | 2    |
| Basic Definitions                                             | 15   |
| Representation of Level N Hypertrees                          | 77   |
| Regular Hypertree Grammars                                    | 105  |
| Finite Automata                                               | 132  |
| Regular Expressions                                           | 152  |
| Characterization of the Algebraic Language Hierarchy          | 183  |
| Containment in Context Sensitive                              | 236  |
| Conclusion                                                    | 285  |
| Bibliography                                                  | 288  |
| Acknowledgements                                              | 292  |
| Appendix: Linearly Bounded Grammars                           | 294  |
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>Some Examples of Hypertrees over {a,b,c}</td>
<td>20</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>Equation c of Definition 9</td>
<td>46</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>Some Examples of n-ary Encoded Hypertrees</td>
<td>101</td>
</tr>
<tr>
<td>TABLE 4</td>
<td>Regular Expressions over {a}</td>
<td>160</td>
</tr>
<tr>
<td>TABLE 5</td>
<td>Miscellaneous Routines</td>
<td>239</td>
</tr>
<tr>
<td>TABLE 6</td>
<td>The Over Routine</td>
<td>240</td>
</tr>
<tr>
<td>TABLE 7</td>
<td>The Search Routine</td>
<td>241</td>
</tr>
<tr>
<td>TABLE 8</td>
<td>The Apply Routine</td>
<td>243</td>
</tr>
<tr>
<td>TABLE 9</td>
<td>The Frontier Routine</td>
<td>245</td>
</tr>
<tr>
<td>TABLE 10</td>
<td>The String Routine</td>
<td>246</td>
</tr>
<tr>
<td>TABLE 11</td>
<td>Over Derivation for ( \gamma \in F_{n,0}^{k'} )</td>
<td>249</td>
</tr>
<tr>
<td>TABLE 12</td>
<td>Over Derivation for ( \gamma \in I )</td>
<td>249</td>
</tr>
<tr>
<td>TABLE 13</td>
<td>Over Derivation for ( \gamma \in F_{n,m-1}^{k'+1} )</td>
<td>250</td>
</tr>
<tr>
<td>TABLE 14</td>
<td>Over for ( \gamma \in F_{n,m-1}^{k'+1} )</td>
<td>251</td>
</tr>
<tr>
<td>TABLE 15</td>
<td>Over for ( \gamma \in &lt;\text{frontier}<em>n^{k'} F</em>{n+1,m-1}&lt;^\gamma )</td>
<td>252</td>
</tr>
<tr>
<td>TABLE 16</td>
<td>Over Derivation for ( \gamma \in F_{n,m}^0 )</td>
<td>252</td>
</tr>
<tr>
<td>TABLE 17</td>
<td>( S_{n+1}^k ) Derivation for ( \gamma \in H_{n+1,0}^k )</td>
<td>254</td>
</tr>
<tr>
<td>TABLE 18</td>
<td>( S_{n+1}^k ) Derivation for ( \gamma \in I )</td>
<td>254</td>
</tr>
<tr>
<td>TABLE 19</td>
<td>( S_{n+1}^k ) for ( \gamma \in H_{n+1,m-1}^{n+2,k} )</td>
<td>255</td>
</tr>
<tr>
<td>TABLE 20</td>
<td>( S_{n+1}^k ) Derivation for ( \gamma \in H_{n,m-1}^{k+1} )</td>
<td>256</td>
</tr>
<tr>
<td>TABLE 21</td>
<td>( S_{n+1}^k ) for ( \gamma \in H_{n+1,m-1}^{k+1,k} )</td>
<td>257</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>22</td>
<td>$S_n^0$ Derivation for $\gamma \in H_{n,m}^0$</td>
<td>258</td>
</tr>
<tr>
<td>23</td>
<td>$&lt;S_1^2_n&gt;$ Derivation for $\gamma_2 \in F_{n,m}^n$</td>
<td>259</td>
</tr>
<tr>
<td>24</td>
<td>$&lt;S_1^k_n&gt;$ for $\gamma_2 \in F_{n,m}^{k+1}[k F_{n,n}^{k-1}n]$ and $j = k_i$</td>
<td>260</td>
</tr>
<tr>
<td>25</td>
<td>$&lt;S_1^k_n&gt;$ Derivation for $\gamma_2 \in F_{n,m}^{k+1}$ and $j \in F_{n,n}^{k+1}$</td>
<td>261</td>
</tr>
<tr>
<td>26</td>
<td>$&lt;S_1^k_n&gt;$ for $\gamma_2 \in F_{n,m}^{k+1}[k F_{n,n}^{k-1}n]$ and $j \in F_{n,n}^{k+1}$</td>
<td>262</td>
</tr>
<tr>
<td>27</td>
<td>$&lt;S_1^0_n&gt;$ Derivation</td>
<td>263</td>
</tr>
<tr>
<td>28</td>
<td>$&lt;A_1^k_n&gt;$ Derivation for $\gamma_1 \in X_{k,n}$ and $k &lt; n$</td>
<td>263</td>
</tr>
<tr>
<td>29</td>
<td>$&lt;A_1^2_n&gt;$ Derivation for $\gamma_1 = X_{j,n}^j$</td>
<td>264</td>
</tr>
<tr>
<td>30</td>
<td>$&lt;A_1^{n+1}_n&gt;$ Derivation</td>
<td>265</td>
</tr>
<tr>
<td>31</td>
<td>$&lt;A_1^k_n&gt;$ Derivation for $\gamma_1 \in H_{n,m}^{k+1}$</td>
<td>265</td>
</tr>
<tr>
<td>32</td>
<td>$&lt;A_1^k_n&gt;$ for $\gamma_1 \in H_{n,m-1}^{k+1}[k H_{n,n}^{k-1}n]$</td>
<td>266</td>
</tr>
<tr>
<td>33</td>
<td>$&lt;A_1^0_n&gt;$ Derivation</td>
<td>267</td>
</tr>
<tr>
<td>34</td>
<td>Frontier Derivation for Basis</td>
<td>267</td>
</tr>
<tr>
<td>35</td>
<td>Frontier Derivation for $\gamma = \alpha_{n+1}^{\beta n+1}$</td>
<td>268</td>
</tr>
<tr>
<td>36</td>
<td>Frontier Derivation for $k = n$, $\gamma \in H_{n,n}^{n+1}$</td>
<td>269</td>
</tr>
<tr>
<td>37</td>
<td>Frontier for $k = n$, $\gamma = \alpha_{n}^{\beta n}$ and $\alpha \notin \Sigma$</td>
<td>270</td>
</tr>
<tr>
<td>38</td>
<td>Frontier for $k = n$, $\gamma = \alpha_{n}^{\beta n}$ and $\alpha \notin \Sigma$</td>
<td>273</td>
</tr>
<tr>
<td>39</td>
<td>Frontier Derivation for $\gamma \in H_{n,n}^{k+1}$</td>
<td>275</td>
</tr>
<tr>
<td>40</td>
<td>Frontier Derivation for $\gamma = \alpha_{k}^{\beta k}$</td>
<td>276</td>
</tr>
<tr>
<td>41</td>
<td>Frontier Derivation for $k = 0$</td>
<td>277</td>
</tr>
<tr>
<td>42</td>
<td>Construction for Lemma 103</td>
<td>325</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

FIGURE 1. Sample Tree Grammar ................. 10
FIGURE 2. Language Defined by the Sample Grammar .... 11
FIGURE 3. Illustration of a Context Free Tree Language . 12
FIGURE 4. Known Language Types .................. 13
FIGURE 5. A Projection of a Level 3 Hypertree ....... 21
FIGURE 6. A Tree .................................. 36
FIGURE 7. A Hypertree of level 4 and degree 2 ........ 43
<table>
<thead>
<tr>
<th>Definition</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td>2</td>
</tr>
<tr>
<td>Semantics</td>
<td>2</td>
</tr>
<tr>
<td>Regular Grammars</td>
<td>4</td>
</tr>
<tr>
<td>Context Free Grammars</td>
<td>6</td>
</tr>
<tr>
<td>Context Sensitive Grammar</td>
<td>7</td>
</tr>
<tr>
<td>Unrestricted Grammars</td>
<td>8</td>
</tr>
<tr>
<td>Classical Strings</td>
<td>16</td>
</tr>
<tr>
<td>Classical Trees</td>
<td>16</td>
</tr>
<tr>
<td>Modified Classical Trees</td>
<td>16</td>
</tr>
<tr>
<td>Hypertrees</td>
<td>20</td>
</tr>
<tr>
<td>Hyperforests</td>
<td>23</td>
</tr>
<tr>
<td>Hyperforests of Depth m</td>
<td>26</td>
</tr>
<tr>
<td>n-paths of Degree k</td>
<td>37</td>
</tr>
<tr>
<td>Frontierable Hypertrees</td>
<td>38</td>
</tr>
<tr>
<td>Frontierable Hyperforests</td>
<td>39</td>
</tr>
<tr>
<td>Frontierable Hyperforests of Depth m</td>
<td>39</td>
</tr>
<tr>
<td>Search</td>
<td>45</td>
</tr>
<tr>
<td>Apply</td>
<td>52</td>
</tr>
<tr>
<td>Frontier</td>
<td>57</td>
</tr>
<tr>
<td>Strings</td>
<td>65</td>
</tr>
<tr>
<td>Trees</td>
<td>66</td>
</tr>
<tr>
<td>Definition</td>
<td>Title</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------</td>
</tr>
<tr>
<td>22</td>
<td>Ranked Set</td>
</tr>
<tr>
<td>23</td>
<td>Rank</td>
</tr>
<tr>
<td>24</td>
<td>Hypertrees with Arguments</td>
</tr>
<tr>
<td>25</td>
<td>Hyperforest with Arguments</td>
</tr>
<tr>
<td>26</td>
<td>Hyperforests with Arguments of Depth m</td>
</tr>
<tr>
<td>27</td>
<td>n-ary Trees</td>
</tr>
<tr>
<td>28</td>
<td>n-ary Encoded Hyperforests</td>
</tr>
<tr>
<td>29</td>
<td>n-ary Encoded Hypertrees</td>
</tr>
<tr>
<td>30</td>
<td>n-ary Encoded Hyperforests of Depth m</td>
</tr>
<tr>
<td>31</td>
<td>n-ary Encoding</td>
</tr>
<tr>
<td>32</td>
<td>Recursive Encoding</td>
</tr>
<tr>
<td>33</td>
<td>Regular Grammar</td>
</tr>
<tr>
<td>34</td>
<td>IO Derivation Relation</td>
</tr>
<tr>
<td>35</td>
<td>IO Derivation Closure</td>
</tr>
<tr>
<td>36</td>
<td>IO Language of a Grammar</td>
</tr>
<tr>
<td>37</td>
<td>OI Derivation Relation</td>
</tr>
<tr>
<td>38</td>
<td>Extract Function</td>
</tr>
<tr>
<td>39</td>
<td>String Function</td>
</tr>
<tr>
<td>40</td>
<td>IO String Language</td>
</tr>
<tr>
<td>41</td>
<td>m-normal Form</td>
</tr>
<tr>
<td>42</td>
<td>Finite Automata</td>
</tr>
<tr>
<td>43</td>
<td>Complete Transition Function</td>
</tr>
<tr>
<td>44</td>
<td>Language of a Finite Automaton</td>
</tr>
<tr>
<td>45</td>
<td>Deterministic Finite Automata</td>
</tr>
<tr>
<td>46</td>
<td>Concatination using $v$</td>
</tr>
</tbody>
</table>
DEFINITION 47. Closure of a Set 157
DEFINITION 48. Regular Expression of Level n and Degree k 159
DEFINITION 49. Regular Expressions of Level n, Degree k and Depth m 161
DEFINITION 50. Language Function of Level n and Degree k 162
DEFINITION 51. Normal Regular Expression 163
DEFINITION 52. Partial Finite Automata 169
DEFINITION 53. Extracted Language of Level k 184
DEFINITION 54. String Language 184
DEFINITION 55. Algebraic Language Hierarchy of Level n and Degree k 185
DEFINITION 56. Completed Grammars 197
DEFINITION 57. Linear Paths 205
DEFINITION 58. Linear X 205
DEFINITION 59. Linear Grammar to Level i 205
DEFINITION 60. Linear Grammar 206
DEFINITION 61. The Count Function 221
DEFINITION 62. E Function 227
DEFINITION 63. The F Function 247
DEFINITION 64. The S Function 253
DEFINITION 65. Alternative Hypertree Definition 1 287
DEFINITION 66. Derivation over Grammar G 295
DEFINITION 67. Width of Derivation 295
LIST OF THEOREMS

THEOREM 1. \( h_n^k \) and \( h_n \) Relationship ............ 23

COROLLARY 2. \( h_n = h_n^k \) .................. 25

LEMMA 3. \( h_n^j \subset h_n^k \) .................. 26

LEMMA 4. Alternative Line 3 in Hyperforest Definition ............ 28

LEMMA 5. \( h_n^k, m \subset h_n^k \) .................. 29

LEMMA 6. \( h_n^k, m \) is Nested in \( m \) .................. 32

THEOREM 7. \( h_n^k, m \) and \( h_n^k \) Relationship ............ 33

THEOREM 8. \( H_n^k \) and \( H_n \) Relationship ................ 40

COROLLARY 9. \( H_n = H_n^k \) .................. 40

LEMMA 10. \( H_n^j \subset H_n^k \) .................. 40

LEMMA 11. \( H_n^k, m \subset H_n^k \) .................. 41

LEMMA 12. \( H_n^k, m \) is Nested in \( m \) .................. 41

LEMMA 13. Alternative Line 3 In Frontierable Hyperforest Definition ............ 41

THEOREM 14. \( H_n^k, m \) and \( H_n^k \) Relationship ............ 41

LEMMA 15. Codomain for search Using \( H_n^k, m \) .................. 47

THEOREM 16. Search Function Codomain ................ 51

LEMMA 17. Codomain for apply Using \( H_n^k, m \) .................. 53

THEOREM 18. Apply Function Codomain ................ 56

LEMMA 19. Codomain for frontier Using \( H_n^k, m \) .................. 58

THEOREM 20. Frontier Function Codomain ................ 62
LEMMA 21. \( h^k_{n,m} \subseteq H^k_{n,m} \) ........................................ 66
THEOREM 22. \( h_n \subseteq H_n \) ........................................... 68
THEOREM 23. \( H^k_n(I,V) \) and \( H_n(I,V) \) Relationship ............ 73
COROLLARY 24. \( H_n(I,V) = H^k_n(I,V) \) ............................. 73
LEMMA 25. \( H^j_n(I,V) \subseteq H^k_n(I,V) \) ............................. 73
LEMMA 26. \( H^k_{n,m}(I,V) \subseteq H^k_n(I,V) \) ......................... 74
LEMMA 27. \( H^k_{n,m}(I,V) \) is Nested in \( m \) ......................... 74
LEMMA 28. Alternative Line 3 in Frontierable Hyperforests with Arguments Definition ........ 74
THEOREM 29. \( H^k_{n,m}(I,V) \) and \( H^k_n(I,V) \) Relationship ........ 74
LEMMA 30. Nesting of Hypertree Types ................................ 76
THEOREM 31. \( h_n \subseteq H_n(I,V) \) ..................................... 76
LEMMA 32. \( T^j_n \subseteq T^k_n \) ........................................ 80
LEMMA 33. \( T^k_{n,m} \subseteq T^k_n \) ..................................... 82
LEMMA 34. \( T^k_{n,m} \) is Nested in \( m \) .............................. 85
THEOREM 35. \( T^k_{n,m} \) and \( T^k_n \) Relationship ...................... 87
LEMMA 36. \( T^k_{n,m} \) is a Set of \( n \)-ary Trees ....................... 89
LEMMA 37. \( T^k_n \) is a Set of \( n \)-ary Trees ......................... 94
THEOREM 38. \( T_n \) is an \( n \)-ary Tree ................................ 94
THEOREM 39. Encode Function Domain -- \( n \)-ary Encoded Hypertrees .................. 97
LEMMA 40. \( G^k_n \rightarrow \) Relation Codomain Lemma .................. 107
THEOREM 41. \( G^k_n \rightarrow \) Relation Codomain ....................... 111
THEOREM 42. \( G^k_n \rightarrow \) Relation Codomain ....................... 112
THEOREM 43. \( m \)-normal Grammar Exists ............................. 119
LEMMA 44. \( m-l \)-normal Grammar Exists ............................ 119
THEOREM 45. Normal Grammar Exists .......... 131
THEOREM 46. Existence of a Deterministic Finite
Automaton ................................. 138
LEMMA 47. Grammar -- Automata Construction .... 146
LEMMA 48. Normal Grammar -- Automata Equivalence . 150
THEOREM 49. Finite Automata -- Regular Grammar
Equivalence ............................... 150
LEMMA 50. Concatination Codomain for Hypertrees of
Depth m ................................. 155
THEOREM 51. Concatination Codomain .............. 157
THEOREM 52. Closure of a Set Codomain .......... 158
THEOREM 53. $R_{n,m}^k$ and $R_n^k$ Relationship .... 161
THEOREM 54. Regular Set Language Codomain ....... 163
THEOREM 55. Equivalence of Partial and Full Finite
Automata ................................. 170
THEOREM 56. Regular Expression and Automata
Equivalence ............................... 172
THEOREM 57. Grammar and Regular Expression
Equivalence ............................... 174
COROLLARY 58. $ALH_n^k$ Using Automata and Regular
Expressions .............................. 185
THEOREM 59. Hypertree Closure Under Union ....... 186
THEOREM 60. Hypertree Closure Under Intersection ... 186
COROLLARY 61. $ALH$ Closed Under Union ............. 190
THEOREM 62. Closure Under Intersection with Regular
Sets ................................. 191
THEOREM 63. Completed Grammar Exists .......... 197
COROLLARY 64. Existance of Non-dropping Grammar ... 201
THEOREM 65. Only $V_i \ k > i > 0$ are needed ........ 202
COROLLARY 66. Only $V_n$ needed for $ALH^1$ n n 203
COROLLARY 67. $G_n^L$ exists for $L \in ALH^k$ n n 204
THEOREM 68. $G_n^L$ for $L \in ALH^k$ modified n n 204
THEOREM 69. Linear Grammar for $L \in ALH^k$ n n 205
COROLLARY 70. Linear Grammars for Strings n n 206
COROLLARY 71. Linear Grammars for trees n n 207
COROLLARY 72. Linear Grammars for Hypertrees of Level 3 208
THEOREM 73. ALH Level 1 Corresponds to Regular Languages n n 209
THEOREM 74. ALH Level 2 Corresponds to Context Free Languages n n 210
THEOREM 75. ALH Level 3 Corresponds to Macro Languages n n 213
THEOREM 76. Pumping Lemma n n 222
LEMMA 77. Limit on Count of the Frontier n n 223
LEMMA 78. Language to Establish Hierarchy Introduced n n 227
LEMMA 79. Language in Lemma 78 is in the Hierarchy 230
THEOREM 80. Proof of Hierarchy n n 232
COROLLARY 81. Proof of Hierarchy at Level k n n 233
LEMMA 82. Language not in the Hierarchy Introduced 234
LEMMA 83. Language in Lemma 82 is Context Sensitive 234
THEOREM 84. Non-Equivalence with Context Sensitive n n 235
LEMMA 85. Over Derivation Width n n 248
COROLLARY 86. Completed Over Derivation n n 251
LEMMA 87. Frontier Derivation to $frontier^L_n$ n n 253
LEMMA 88. Search Derivation Width n n 256
LEMA 89. Apply Derivation Width ... 261

LEMA 90. Frontier Derivation Width ... 267

LEMA 91. Frontier Function Works ... 277

LEMA 92. String Function Works ... 277

LEMA 93. Linearly Bounded Sequence of Frontiers is Sufficient ... 278

LEMA 94. K Bounded Grammar Exists ... 279

THEOREM 95. Containment in Context Sensitive ... 283

LEMA 96. $P \subset (V^+ \rightarrow V^*) \ U (V \rightarrow \lambda) \ U (V \rightarrow \Sigma)$ Exists ... 296

LEMA 97. Reduce Width of Productions by One ... 302

LEMA 98. Width of Productions Reduced to 2 ... 306

LEMA 99. $P \subset (V \rightarrow V) \ U (V \rightarrow VV) \ U (V \rightarrow \lambda) \ U (V \rightarrow \Sigma) \ U (V \rightarrow V)$ ... 307

LEMA 100. $P \subset (V \rightarrow V) \ U (V \rightarrow VV) \ U (V \rightarrow \Sigma) \ U (V \rightarrow \lambda)$ ... 309

THEOREM 101. Normal Form Grammar Exists ... 312

LEMA 102. Length Bounded Grammar Corresponds to Context Sensitive ... 317

LEMA 103. Width can be Halved ... 323

LEMA 104. Arbitrary $I$ is/is not in a Bounded Language ... 327

COROLLARY 105. $\lambda$ is/is not in a Bounded Language ... 329

LEMA 106. Context Sensitive Grammar for K Bounded Grammars ... 329

THEOREM 107. Linearly Bounded Languages are Context Sensitive ... 331
A hierarchy of data structures, called the hypertree hierarchy, is presented which has strings and trees as its smallest two elements. A generalized frontiering operation is presented. Grammars, automata and regular expressions are extended to this hierarchy. These are shown to be equivalent and the resulting languages are called regular. This leads to a hierarchy of regular languages on the hypertree hierarchy. These are projected onto the set of strings by the frontier operation resulting in a true hierarchy of string languages. This hierarchy is called the (10) algebraic language hierarchy. It has regular languages, context free languages and macro languages as its first three levels. It is contained in the set of context sensitive languages but is not equal to it. Other characterizations are presented.
INTRODUCTION

In the area of programming languages, two words are used quite extensively and with some variation as to the meaning. They are "syntax" and "semantics". In order to be certain that there is no confusion in the use of these terms, a definition will be given here.

Definition 1:
The syntax of a language is the set of all legal programs in the language.

Definition 2:
The semantics of a language is function which maps the syntax of the language to a set of meanings.

The term "meanings" will remain undefined in this context. This paper concerns itself with the specification of a syntax a formal language.

Typically, the syntax of a formal language is specified by giving a context free grammar for the language. However, this has its drawbacks because some of the aspects of common programming languages cannot be specified using a context free grammar. For example, in Pascal a variable must be declared before it is used, yet this simple fact cannot be
expressed via a context free grammar. Therefore, in addition to the grammar of most programming languages there has to be a piece of English prose which describes these sets of facts. English, being a natural language, has a problem in that it is subject to different interpretations by different people. Even more important, the English prose is too complex and ill-defined to allow automatic generation of a parser given these specifications. Therefore, typically a parser is written which recognizes the context free language specified, and the errors that are defined by the English prose are postponed and allowed to show up in the code generation (semantic analysis) phase of a compiler. For this reason, they are typically called semantic errors, although, in fact, they are not errors in the semantics of the program at all but rather are syntax errors. All these consideration have led some people to explore other means of specifying syntax.

There are five methods of specifying syntax that are of interest at this point. The first four were collected into a hierarchy by Chomski in the 1950s, and are therefore known as the Chomski hierarchy. They are: regular grammars, context free grammars, context sensitive grammars and unrestricted grammars. The fifth method, macro grammars, is due to Fischer (Fischer 1968). In 1974, it was shown by
Wand (Wand 1975) and Maibaum (Maibaum 1974) that of these, regular grammars, context free grammars and macro grammars form a natural hierarchy. This hierarchy will be called the algebraic language hierarchy in this paper since Wand and Maibaum showed that this hierarchy was algebraic in nature.

Regular grammars are defined as follows:

Definition 3:

A regular grammar is a four-tuple, \((\Sigma, N, P, S)\), where

1. \(\Sigma\) is a set called the terminals.
2. \(N\) is a set called the nonterminals.
3. \(P\) is a finite set called the productions.
4. \(S\) is a symbol called the start symbol.

In addition, \(S \in N\) and \(P \subseteq N \times (\Sigma^*N \cup \Sigma^*)\).

An example of a regular grammar is

\[\{(a,b),(S,B),(S,aB),(B,bS),(S,a)\},S\]

which is typically written as

\[S \rightarrow aB\]
\[B \rightarrow bS\]
\[S \rightarrow a\]

where each of the lines is called a production. Note that \(A \rightarrow \alpha\) is a production is exactly equivalent to \((A,\alpha) \in P\).
Also, typically, the start symbol is written as the first symbol, that is, the left hand side of the first production. This notation will be used here. These may be used to generate a language as follows: If A → β is a production, then,

\[ αAγ \Rightarrow αβγ \]

for all α and γ. Let this (informally) define \( \Rightarrow \). Then, let \( *\Rightarrow \) be the reflexive and transitive closure of \( \Rightarrow \). The language defined by the grammar above is, therefore,

\[ \{ α \mid S \Rightarrow α \text{ and } α \in Σ^* \} \]

Therefore, the grammar above defines the language

\[ \{(ab)^nα\}_{n≥0} \]

It is a well-known fact that regular grammars are very limited in the languages they can define. They are used for defining only very simple languages such as are used in lexical analyses. Therefore, more powerful grammars are typically used to define real programming languages. The most common is context free grammars. Context free grammars are defined as follows:
Definition 4:
A context free grammar is a four-tuple \((\Sigma, N, P, S)\)
where
1. \(\Sigma\) is a set called the terminals.
2. \(N\) is a set called the nonterminals.
3. \(P\) is a finite set called the productions.
4. \(S\) is a symbol called the start symbol.

Further, \(S \in N\) and \(P \subseteq N \times (\Sigma \cup N)^*\).

The definition of \(\Rightarrow\) and the language generated is exactly analogous to the corresponding notation used in regular grammars above. An example of a context free grammar is

\[
\begin{align*}
S & \rightarrow E \\
E & \rightarrow (E) \\
E & \rightarrow a
\end{align*}
\]

Note, this grammar defines the language

\[
\{(a^n)^n\}_{n \geq 0}
\]

It is easy to show this language cannot be represented by a regular grammar; therefore, it is known that context free grammars can define a larger set of languages than regular grammars can define. For this reason, context free grammars are used to define a number of programming languages as was
The third type of grammar is context sensitive grammars. It is defined as follows:

Definition 5:

A context sensitive grammar is a four-tuple

$(\Sigma, N, P, S)$ where

1. $\Sigma$ is a set of terminals.
2. $N$ is a set of nonterminals.
3. $P$ is a finite set of productions.
4. $S$ is a symbol called the start symbol.

Further, $S \in N$ and $P \subseteq (\Sigma \cup N)^* (\Sigma \cup N)^* X (\Sigma \cup N)^*$ such that if $(A, B) \in P$ then the length of $A$ is less than or equal to the length of $B$.

The definition of $\Rightarrow$ requires now that the $A$ may be a string of symbols. An example of a context sensitive grammar is

\[
\begin{align*}
S & \rightarrow aSBc \\
S & \rightarrow abc \\
cB & \rightarrow Bc \\
bB & \rightarrow bb
\end{align*}
\]

It can be shown that this grammar defines the language

\[
\{a^n b^n c^n \}_{n>0}
\]
which cannot be expressed with a context free grammar.

The fourth type of grammar in the Chomski hierarchy is called unrestricted.

Definition 6:

An unrestricted grammar is a four-tuples $(\Sigma, \mathcal{N}, P, S)$ where

1. $\Sigma$ is a set of terminals.
2. $\mathcal{N}$ is a set of nonterminals.
3. $P$ is a finite set of productions.
4. $S$ is a symbol called the start symbol.

In addition, $S \in \mathcal{N}$ and $P \subseteq (\Sigma \cup \mathcal{N})^* \mathcal{N} (\Sigma \cup \mathcal{N})^* \Sigma (\Sigma \cup \mathcal{N})^*$.

The language defined by this grammar is analogous to the previous definitions.

The fifth type of grammar was first formally defined by Fischer (Fischer 1968). It is called macro grammars and will not be formally defined here. In general, macro grammars allow productions in which each nonterminal is allowed zero or more arguments. For example,

\[
\begin{align*}
S & \rightarrow T(a,b,c) \\
T(X_1,X_2,X_3) & \rightarrow T(X_1a,X_2b,X_3c) \\
T(X_1,X_2,X_3) & \rightarrow X_1X_2X_3
\end{align*}
\]
Here, the "(a,b,c)" following the first nonterminal \( T \) is a listing of the three arguments, "a", "b" and "c".

Similarly, \( \Rightarrow \) must be defined differently. If \( A(X_1 \ldots X_n) \Rightarrow \beta \) is a production, then

\[
aA(\beta_1 \ldots \beta_n)\gamma \Rightarrow a\beta'\gamma
\]

where \( \beta' \) is \( \beta \) with all occurrences of \( X_1 \) replaced by \( \beta_1 \), \( X_2 \) replaced by \( \beta_2 \) etc. \( \ast \Rightarrow \) is the reflexive and transitive closure as before. Using these rules, it can easily be seen that the grammar given produces the language

\[
\{a^n b^n c^n\}_{n \geq 0}
\]

so that macro grammars are more powerful than context free grammars. Fischer also showed that macro grammars are not as powerful as context sensitive grammars. The same paper also shows that there are two types of macro grammars, IO and OI. IO macro grammars insist that in order to apply the production given above, \( \beta_1 \ldots \beta_n \) must be elements of \( I^* \). OI, on the other hand, insists that the \( A \) not be an argument to any other nonterminal at the point where the production is applied. This gives the same language as is obtained if there is no restrictions at all placed on the order of evaluation.
At the same time, some of the grammars that have just been described were extended to operate on trees by Rounds (Rounds 1969). These included regular grammars and context-free grammars. An example of a regular grammar over trees is given in figure 1 where each nonterminal is required to be on the frontier of the tree. This will define the language given in figure 2.

```
S → 1
  | E
E → 2
    [ E ]
E → 3
    a
```

FIGURE 1. Sample Tree Grammar

Note that an alternative method of representing trees is to put parentheses around the subtrees at each node. Therefore, the above grammar can be written
such that $n > 0$.

FIGURE 2. Language Defined by the Sample Grammar

$$S \rightarrow 1(E)$$
$$E \rightarrow 2([E])$$
$$E \rightarrow 3(a)$$

and the language can be represented as

$$\{1(2([])^n3(a){})^n\}_{n \geq 0}$$

Note that the frontier of the elements of this language form the language

$$\{[n \ a ]^n\}_{n \geq 0}$$

which is a context free language. In general, this is the case. The frontier of any regular tree language is a context free string language.
FIGURE 3. Illustration of a Context Free Tree Language

Rounds also talks about context free grammars over trees (Rounds 1969). For example,

\begin{align*}
S & \rightarrow T(abc) \\
T(X_1X_2X_3) & \rightarrow T(*(X_1a)*(X_2b)*(X_3c)) \\
T(X_1X_2X_3) & \rightarrow *(X_1X_2X_3)
\end{align*}

Some examples of the trees resulting from this grammar are given in figure 3. Note that the frontier of this language is the language
which is a macro language. Rounds showed that in general this is the case. Also, for each context free tree grammar, there is a macro grammar which describes the frontier language. Note that just as there were two distinct classes of macro grammars, 10 and 01, there are two classes of context free tree grammars on trees -- top-down and bottom-up.

<table>
<thead>
<tr>
<th></th>
<th>Regular</th>
<th>Context Free</th>
<th>Indexed</th>
</tr>
</thead>
<tbody>
<tr>
<td>String</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Tree</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 4. Known Language Types

These last few remarks lead to the figure given in figure 4. Here each of the known grammars are represented by an asterisk, with an arrow representing the frontier function. This is similar to the table given by Maibaum (Maibaum 1974), with his hierarchy of algebras replaced by a hierarchy (of two) data structures. As Maibaun noted, this
suggests a hierarchy of data structures, and it is this hierarchy of data structures that is the subject of this paper. This hierarchy is called the hypertree hierarchy.
BASIC DEFINITIONS

The set of hypertrees over $I$ has two well-known special cases. These are the set of strings over $I$ and the set of trees over $I$, which are designated as level 1 hypertrees and level 2 hypertrees respectively.

Historically, there have been a few discrepancies in the treatment of strings and trees. For instance, a null string is provided, but a null tree is not. Also, any element of $I$ can appear at any location in a string, but only a subset can appear at any particular location on a tree, such as at the frontier, due to the ranking relation (Rounds 1970). In each case, a decision had to be made as to which structure to follow; a similar development could be made using the other choices. In the above examples, it was decided that eliminating the null string and the ranking relation would make the smoothest development. Therefore, $I^+$ is used as the set of strings instead of $I^*$. Also, each of the ranks $I_1$'s in the ranked set $I$ is assumed to be equal to $I$ itself. Normally, only a finite number of the ranks may be nonempty so that tree automata may be developed, however, it will be shown in chapter 5 that this is not needed.

A recursive definition of the set of strings over $I$, $I^+$, is given by
Definition 7:

The set of classical strings over $\Sigma$, denoted $\Sigma^+$, is the smallest set such that

$$\Sigma^+ = \Sigma \cup \Sigma \Sigma^+$$

Similarly, a recursive definition of the set of trees over $\Sigma$, $T_\Sigma$, is given by

Definition 8:

The set of classical trees over the ranked set $\Sigma$, denoted $T^\Sigma$, is the smallest set such that

$$T^\Sigma = \Sigma_0 \cup \bigcup_{i \geq 1} \Sigma_i [(T^\Sigma)^i]$$

where $\Sigma_i$ is the $i^{th}$ rank of $\Sigma$.

Keeping in mind the above discussion, this may be rewritten as

Definition 9:

The set of modified classical trees over $\Sigma$, denoted $MT^\Sigma$, is the smallest set such that

$$MT^\Sigma = \Sigma \cup \bigcup_{i \geq 1} \Sigma [(MT^\Sigma)^i]$$

since each of the $\Sigma_i$'s is equal to $\Sigma$.

In developing the definition of hypertrees, first a common notation is needed for the set of strings over $\Sigma$ and
the set of trees over Σ. Since a level number is also needed, an entirely new notation was developed rather than expand either the $T_\Sigma$ notation or the $\Sigma^+$ notation. Therefore, the set of strings over Σ, which is the set of hypertrees of level 1 over Σ, is denoted by $H_1(\Sigma)$, and the hypertrees of level 2 over Σ, the set of trees over Σ, is denoted as $H_2(\Sigma)$. In either case, if the hypertrees are over Σ for the remainder of this paper, the "$(\Sigma)$" will be elided unless confusion would result. Therefore, the set of strings over Σ is denoted as $H_1$, the set of trees over Σ is denoted $H_2$, and in general, the set of hypertrees of level n over Σ will be denoted by $H_n$.

In order to develop the general definition for hypertrees, some additional changes need to be made to the definition of strings given in definition 7 and the definition of trees given in definition 9. Using the new notation, the equation in definition 7 may be written as

$$H_1 = \Sigma U \Sigma H_1$$

At this level, there is no ambiguity about what follows the initial element of Σ in the string, but in general this is not the case. Therefore, the part of the string which follows Σ is indicated by the use of brackets.

$$H_1 = \Sigma U \Sigma [ H_1 ]$$
In order that these brackets not be confused with brackets at other levels of the hierarchy, a subscript is introduced. Therefore, the equation becomes

\[(1) \quad H_1 = \Sigma \cup \Sigma[1 H_1 1]\]

Similarly, the equation in definition 9 may be written as

\[H_2 = \Sigma \cup \Sigma[1 \Sigma[1 H_2 i 2]][2 (H_2)^i ]\]

As above, subscripts need to be added to avoid confusion. The equation now becomes

\[H_2 = \Sigma \cup \Sigma[1 \Sigma[1 \Sigma[1 \Sigma[1 \Sigma[1 H_2 i 2]][2 (H_2)^i ] 2]][2 (H_2)^i ] 2][2 (H_2)^i ] 2]\]

A reasonable interpretation for the \(U_{i \geq 1}(H_2)^i\) is the set of strings over \(H_2\), that is the set of strings of trees, which is \(H_1(H_2)\). Therefore, the above equation becomes

\[(2) \quad H_2 = \Sigma \cup \Sigma[1 \Sigma[1 H_1(H_2) 2]]\]

Note that since \(H_2\) is an infinite set, the set of strings must be defined over infinite sets. This is also extended to all hypertrees.
A logical extension of the hypertree hierarchy in the downward direction is to the set of elements of $\mathcal{E}$, that is, to $\mathcal{E}$ itself. This is reasonable if it is noted that trees have two dimensions, height and width, whereas strings have only one dimension, length. Continuing on down, the elements of $\mathcal{E}$ have neither height, width nor length and are, therefore, zero dimensional objects. Therefore, it is reasonable to expect that the set of level 0 hypertrees over $\mathcal{E}$ is simply $\mathcal{E}$ itself. That is

\begin{equation}
H_0 = \mathcal{E}
\end{equation}

Using this fact, equation 1 may be rewritten as

\begin{equation}
H_1 = \mathcal{E} \cup \mathcal{E} \{ \mathcal{E} \cup \mathcal{E} \}
\end{equation}

since $H_0$ is the identity function.

From equations 2 and 4, it should be evident that in general

\begin{equation}
H_n = \mathcal{E} \cup \mathcal{E} \{ \mathcal{E} \cup \mathcal{E} \}
\end{equation}

and indeed this is the case. The one exception has been noted above in equation 3. It will be seen later that this definition is somewhat restrictive, therefore, in the formal definition of hypertrees a lower case $h$ is used in place of the upper case $H$. 

Definition 10:

The set of hypertrees over $\Sigma$ of level $n$, denoted $h_n$, is the smallest set such that

(a) $h_0 = \Sigma$

(b) $h_n = \Sigma \cup \Sigma[h_{n-1}(h_n)_n]$ if $n > 0$.

TABLE 1. Some Examples of Hypertrees over {a,b,c}

<table>
<thead>
<tr>
<th>line</th>
<th>level</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>a</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>a[_b[1c_1]_1]</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>a</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>a[2a[1b[1c_1]_1]_1]_2</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>a[2a[2a[1b[1c_1]_1]_1]_2][1b[1c[2a[2a[1b[1c_1]_1]_2]_1]_1]_2</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>a</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>a[3a[2a[1b[1c_1]_1]_2]_3]</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>a[3a[2b[1c[3a[2b[1c[1c_1]_2]_3]_2a[1a[1a_1]_1]_2]_3]_3]</td>
</tr>
</tbody>
</table>

If $\Sigma = \{a, b, c\}$, then some examples of hypertrees are given in table 1. In each case, at level 2 and below, in view of the above discussion, the usual parenthesized
notation for the string or tree can be recovered by deleting the one level brackets and deleting the subscripts on the two level brackets. Therefore, line 6 of table 1 is usually written as a[a[ab]bc[a[ab]]].

FIGURE 5. A Projection of a Level 3 Hypertree

Throughout this paper, hypertrees are defined and used as sets of strings rather than graphs. The reason for this is that strings and trees, being one and two dimensional objects, have simple two dimensional projections. However, this is not the case in general. A level n hypertree is an n dimensional object. For example, the hypertree given in line 9 of table 1 is of level 3, and, therefore, is a three dimensional object. This object may be represented by a 3 dimensional graph. A projection of this graph into 2 dimensions is given in figure 5. Hypertrees of levels higher than 3 are not uncommon, so even models cannot be built of these. For this reason, no further attempts will
be made at using graphs to illustrate hypertrees in this paper.

In definition 10, line b, the expression "h_{n-1}(h_n)" occurs. This expression occurs so often in the following pages that a special notation is given for it

\[ h^k_n = h_k(h_{k+1}( \ldots (h_{n-1}(h_n)) \ldots ) \]

In the case of trees, this reduces to

\[ h^1_2 = h_1(h_2) \]

which is the set of strings of trees, or in other words the set of forests over \( I \). For this reason, \( h^k_n \) is referred to as the set of hyperforests of degree \( k \) and level \( n \) over \( I \).

In order to reduce the number of special cases in the formal definition, the hyperforest \( h^{n+1}_n \) is taken as being equal to \( I \). Note also that

\[ h^0_n = h^1_0(h^1_n) = h^1_n \]

In general,

\[ h^k_n = h_k(h^k_{n+1}) = h^k_n \cup h^k_n[h_{k-1}(h^k_n) k] = h^k_{n+1} \cup h^k_n[h_{k-1} h^k_n k] \]
Since it is often useful to prove things about hyperforests and then note that a hypertree is a special case of a hyperforest, a formal definition of hyperforests is in order.

Definition 11:

The set of hyperforests of degree k and level n over I, denoted \( h_n^k \), is the smallest set such that

(a) \( h_n^{n+1} = \emptyset \)

if \( n \geq 0 \).

(b) \( h_n^k = h_n^{k+1} \cup h_n^{k+1}[h_n^{k-1}] \)

if \( n \geq k > 0 \).

(c) \( h_n^0 = h_n^1 \)

if \( n \geq 0 \).

This definition of hyperforest does not obviously have the connection to hypertrees that is mentioned in the above discussion. This connection is proved in the following theorem.

Theorem 1:

For all \( n \geq k \geq 0 \),

\[ h_n^k = h_k(h_{k+1}(\ldots (h_n)\ldots )) \]

Proof:

Define

\[ a_n^k(\xi) = h_k(h_{k+1}(\ldots (h_n)\ldots )) \]
for all \( n \geq k \geq 0 \). Also, define

\[
\alpha_{n+1}^n(\xi) = \xi
\]

for all \( n \geq 0 \). If \( n > k \geq 0 \), then

\[
\alpha_k^n(\xi) = h_k(\alpha_k^{n+1}(\xi))
\]

If \( n = k \), then

\[
\alpha_k^n(\xi) = h_n
\]

\[
= h_n(\alpha_n^{n+1}(\xi))
\]

\[
= h_k(\alpha_k^{n+1}(\xi))
\]

By definition 10, therefore, \( \alpha_k^n(\xi) \) is the smallest set such that

\[
\alpha_0^n(\xi) = h_0(\alpha_0^{n}(\xi))
\]

\[
= \alpha_n^{1}(\xi)
\]

if \( n \geq 0 \).

\[
\alpha_k^n(\xi) = h_k(\alpha_k^{n+1}(\xi))
\]

\[
= \alpha_n^{k+1}(\xi) \cup \alpha_n^{k+1}(\xi)[_k h_{k-1}(h_k(\alpha_n^{k+1}(\xi)))]
\]

\[
= \alpha_n^{k+1}(\xi) \cup \alpha_n^{k+1}(\xi)[_k \alpha_n^{k-1}(\xi)]
\]
if \( n \geq k > 0 \). But equations 5, 6 and 7 are precisely the
same as equations a, c and b respectively in definition 11 except that \( a_n^k(\xi) \) replaces \( h_n^k \). Therefore, since \( h_n^k \) is the
smallest set that meets these conditions, and \( a_n^k(\xi) \) is also
the smallest set, it must be true that

\[
h_n^k = a_n^k(\xi)
= h_k(h_{k+1}( \ldots (h_n) \ldots ))
\]

and the theorem is proved.

- - -

Corollary 2:

For all \( n \geq 0 \),

\[ h_n = h_n^n \]

Proof:

Substitute \( n \) for \( k \) in theorem 1.

- - -

A lot of the proofs in this paper depend on the use of
induction. However, definition 11 does not lend itself readily to the use of induction because of the occurrence of
the "\( k+1 \)" as well as the "\( k-1 \)" in equation b. Therefore, a
new variable is introduced to act as the induction variable.
Since it corresponds loosely with the height of trees, this
variable is called the depth of the hypertree. The
definition corresponds rather closely to the definition of
hyperforests, except that "m" is used to keep track of how many times the definition has recurred.

Definition 12:

The set of hyperforests over $\mathcal{I}$ of level $n$, degree $k$ and depth $m$, denoted $h_{n,m}^k$, is the smallest set such that

(a) $h_{n,m}^{n+1} = \sum$
if $n > 0$ and $m > 0$.

(b) $h_{n,0}^k = \emptyset$
if $n > k > 0$.

(c) $h_{n,m+1}^k = h_{n,m}^k \cup h_{n,m}^k U h_{n,m}^{k+1} U h_{n,m}^{k-1}$
if $n > k > 0$ and $m > 0$.

(d) $h_{n,m}^0 = h_{n,m}^1$
if $n > 0$ and $m > 0$.

The task now is to prove that this definition does in fact allow proofs about hyperforests as has been claimed.

This will be accomplished by showing that

$$h_n^k = \bigcup_{m \geq 0} h_{n,m}^k$$

This, in turn, requires some lemmas.

Lemma 3:

For all $n \geq 0$ and $n+1 \geq j \geq k \geq 0$,

$$h_n^j \subset h_n^k$$
Proof by induction on $k$:

Basis: if $j = k$ then

$$h_n^j = h_n^k$$

Therefore,

$$h_n^j < h_n^k$$

Induction step: Assume that if $k' > k$ then

$$h_n^j < h_n^{k'}$$

Since $k \geq 0$, there are two cases to be considered:

1. If $k > 0$, then

$$h_n^k = h_n^{k+1} \cup h_n^{k+1}[k \quad h_n^{k-1} \quad k]$$

and by induction hypothesis,

$$h_n^j < h_n^{k+1}$$

$$< h_n^k$$

2. If $k = 0$, then

$$h_n^j < h_n^1$$

by induction hypothesis. Also,

$$h_n^1 = h_n^0$$

= $h_n^k$

Therefore,

$$h_n^j < h_n^k$$

(9)
In both 8 and 9,
\[ h_n^j \subseteq h_n^k \]
so the theorem is proved by induction.

---

**Lemma 4:**

If \( n \geq k > 0 \) and \( m \geq 0 \), then
\[ h_{n,m+1}^k = h_{n,m}^k U h_{n,0}^k U h_{n,m+1}^{k+1}[k h_{n,m}^{k-1}] \]

Proof by induction on \( m \):

**Basis:** If \( m = 0 \), then
\[ h_{n,1}^k = h_{n,0}^k U h_{n,0}^k U h_{n,0}^{k+1}[k h_{n,0}^{k-1}] \]
\[ = h_{n,m}^k U h_{n,0}^k U h_{n,m+1}^{k+1}[k h_{n,m}^{k-1}] \]
by definition 12.

**Induction step:** Assume that if \( m \geq 0 \) then
\[ h_{n,m+1}^k = h_{n,m}^k U h_{n,0}^k U h_{n,m+1}^{k+1}[k h_{n,m}^{k-1}] \]
Substituting this equation into the appropriate line in definition 12 will allow the expansion

\[ h_{n,m+2}^k = h_{n,m+1}^k U h_{n,m+1}^k U h_{n,0}^k \]
\[ U h_{n,m+1}^{k+1}[k h_{n,m}^{k-1}] U h_{n,m+1}^{k+1}[k h_{n,m+1}^{k-1}] \]
\[ = h_{n,m+1}^k U h_{n,0}^k U h_{n,m+1}^{k+1}[k h_{n,m+1}^{k-1}] \]
since
\[ h_{n,m}^{k+1} < h_{n,m+1}^{k+1} \]
and
\[ h_{n,m}^{k-1} < h_{n,m+1}^{k-1} \]
Notice that equation 10 is the same as the induction hypothesis except each occurrence of \( m \) has been incremented by one. Therefore,
\[ h_{n,m+1}^k = h_{n,m}^{k+1} U h_{n,0}^k U h_{n,m}^{k+1} h_{n,m+k}^{k-1} \]
for all \( m \geq 0 \) by induction.

---

Lemma 5:
For all \( m \geq 0, n \geq 0 \) and \( n+1 \geq k \geq 0 \),
\[ h_{n,m}^k \subseteq h_n^k \]
Proof by induction on \( m \):
Basis: if \( m = 0 \) then there are three cases that need to be considered, \( k = n+1, n \geq k > 0 \) and \( k = 0 \).

1. If \( k = n+1 \), then
\[ h_{n,m}^k = \emptyset \]
\[ = h_n^k \]
2. If \( n \geq k > 0 \), then
\[ h_{n,m}^k = \emptyset \]
\[ \subseteq h_n^k \]
3. If \( k = 0 \), then
$h^0_{n,m} = h^1_{n,m}$

By part 1 if $n = 0$ or part 2 above if $n > 0$

$h^1_{n,m} < h^1_n$

But since $h^1_n = h^0_n$ it must be that

$h^0_{n,m} < h^0_n$

Therefore, in all three cases if $m = 0$ then

$h^k_{n,m} < h^k_n$

Induction step: Assume that if $m' < m$ then

$h^k_{n,m'} < h^k_n$

There are three cases that need to be considered.

1. If $k = n+1$, then

$h^{n+1}_{n,m} = \emptyset$

$= h^{n+1}_n$

Therefore,

$\frac{11}{h^k_{n,m} < h^k_n}$

2. If $n \geq k > 0$, then

$h^k_{n,m} = h^{k+1}_{n,m-1} U h^k_{n,m-1} U h^{k+1}_{n,m-1} [k h^k_{n,m-1} k]$

by lemma 4. Therefore, since $h^k_{n,0} = \emptyset$ it must be that
\[
\begin{align*}
\text{k}^n, \text{k+1} & = \text{k+1} U \text{k}^n, \text{k-1} \\
\text{By induction hypothesis} & \\
\text{k+1} \text{h}^n, \text{n-1} & = \text{h}^n \\
\text{k-1} \text{h}^n, \text{n-1} & = \text{h}^n \\
\text{therefore,} & \\
\text{(12)} & \\
\text{h}^n \text{h}^n, \text{n} = \text{h}^n, \text{n-1} U \text{h}^n, \text{n-1} \text{h}^n, \text{n-1} \\
\text{c} \text{h}^n \\
3. \text{ If k = 0, then} & \\
\text{h}^0 \text{h}^0, \text{n} = \text{h}^1 \text{h}^0, \text{n} \\
\text{But} & \\
\text{h}^1 \text{h}^0, \text{n} = \text{h}^1 \\
\text{as a special case of equation 11 or 12 depending on} & \\
\text{the value of n. Therefore,} & \\
\text{h}^0 \text{h}^0, \text{n} = \text{h}^1 \text{h}^0, \text{n} \\
\text{c} \text{h}^0 \\
\text{So, in this case also} & \\
\text{(13)} & \\
\text{h}^k \text{h}^k, \text{n} = \text{h}^k \text{h}^k, \text{n} \\
\text{Therefore, by equations 11, 12 and 13} & \\
\text{h}^k \text{h}^k, \text{n} = \text{h}^k
if $m > 0$ and the induction hypothesis is assumed. Therefore, the lemma is proved by induction.

---

Lemma 6:

If $n \geq 0$, $n+1 \geq k \geq 0$ and $m \geq 0$, then

$$h_{n,m}^k = U_{m \geq m' \geq 0} h_{n,m'}^k$$

Proof by induction on $m$:

Define

$$a_{n,m}(\xi) = U_{m \geq m' \geq 0} h_{n,m'}^k$$

Basis: if $m = 0$ then

$$a_{n,0}^k(\xi) = U_{m'=0} h_{n,m'}^k = h_{n,0}^k$$

Induction step: Assume that if $m > 0$ then

$$a_{n,m}^k(\xi) = h_{n,m}^k$$

Then by the definition of union,

$$a_{n,m+1}^k(\xi) = h_{n,m+1}^k U a_{n,m}^k(\xi)$$

Therefore, by induction hypothesis
Since $h_{n,m}^k < h_{n,m+1}^k$, it must be true that

$$a_{n,m+1}^k(\xi) = h_{n,m+1}^k$$

if $m \geq 0$. Therefore, by induction

$$a_{n,m}^k(\xi) = h_{n,m}^k$$

Which is the same as saying

$$h_{n,m}^k = U_{m \geq m' \geq 0} h_{n,m'}^k$$

Therefore, the lemma is proved.

- • -

Theorem 7:

For all $n \geq 0$ and $n+1 \geq k \geq 0$,

$$h_{n}^k = U_{m \geq 0} h_{n,m}^k$$

Proof:

To prove that

$$h_{n}^k = U_{m \geq 0} h_{n,m}^k$$

it will be shown that $U_{m \geq 0} h_{n,m}^k$ has all the properties that define $h_{n}^k$. Define

$$a_{n,m}^k(\xi) = U_{m \geq m' \geq 0} h_{n,m'}^k$$
By lemma 6
\[ a_{n,m}^k(\xi) = b_{n,m}^k \]

Therefore, \( a_{n,m}^k(\xi) \) is the smallest set such that the following three equations are true:

1. If \( n \geq 0 \), \( k = n+1 \) and \( m \geq 0 \), then

\[ a_{n,m}^{n+1}(\xi) = \xi \]

2. If \( n \geq k > 0 \) and \( m \geq 0 \), then by lemma 4

\[ a_{n,m}^k(\xi) = a_{n,m}^{k+1}(\xi) \cup b_{n,0}^k(\xi) \cup a_{n,m}^{k+1}(\xi) \cup [k a_{n,m}^{k-1}(\xi)] \]

3. If \( n \geq 0 \) and \( m \geq 0 \), then

\[ a_{n,m}^0(\xi) = a_{n,m}^1(\xi) \]

By the definition of infinite union

\[ \lim_{m \to \infty} a_{n,m}^k(\xi) = \bigcup_{m \geq 0} b_{n,m}^k \]

Therefore, if \( a_{n}^k(\xi) \) is defined as

\[ a_{n}^k(\xi) = \bigcup_{m \geq 0} b_{n,m}^k \]

then \( a_{n}^k(\xi) \) is the smallest set such that
1. If $n \geq 0$ and $k = n+1$, then

$$a_{n}^{n+1}(\xi) = \xi$$

2. If $n \geq k > 0$, then

$$a_{n}^{k}(\xi) = a_{n}^{k+1}(\xi) U h_{n,0}^{k} U a_{n}^{k+1}(\xi)[h_{n}^{k-1}(\xi) U]$$

3. If $n \geq 0$ and $k = 0$, then

$$a_{n}^{0}(\xi) = a_{n}^{1}(\xi)$$

But since $h_{n,0}^{k} = \phi$ the second equation becomes

$$a_{n}^{k}(\xi) = a_{n}^{k+1}(\xi) U a_{n}^{k+1}(\xi)[h_{n}^{k-1}(\xi) U]$$

Now since $h_{k}^{n}$ is the smallest set with properties 14, 15 and 16 it must also be true that

$$h_{n}^{k} = a_{n}^{k}(\xi) = U_{m \geq 0} h_{n,m}^{k}$$

and the theorem is proved.
As was stated earlier, the definition of hypertrees given in definition 10 is somewhat restrictive. Consider the case of the frontier. For a tree, this is simple. All that needs to be done is that the leaves of the tree -- those nodes which don't have any subtrees -- are listed in the order that they occur. For example, the tree in figure 5 has the frontier "abbab". However, consider the hypertree in figure 5. Since this is a level 3 hypertree, the frontier should be a level 2 hypertree, a tree. It is easy to see that the frontier starts out as a[ba[bc]a], but the bottommost "a" doesn't have any subhypertrees, and, therefore, by the logic given above, should be on the frontier. But where does the "a" get attached? This particular problem leads to the redefinition of hypertree.
In this new definition, it has to be explicitly stated where the subtrees are to be attached during the frontiering operation, as will be seen.

First, some way of referring to a particular subhypertree in a hypertree must be found. This can be done by finding a way to refer to the head of the hypertree only since once the location of the head is known it is quite easy to find a complete hypertree of any given level. In the case of level 2 hypertrees (trees), this is trivial, and even at level 3 it is not a difficult problem but at higher levels it becomes very much more difficult. Therefore, a set of "paths" are defined for locating a particular subhypertree in a hypertree. A path describes how to follow the pointers in a hypertree to get to the head of the desired subhypertree.

Definition 13:

The set of legal $n$-paths of degree $k$, denoted $P^k_n$, is the smallest set such that

(a) $P^0_n = \{\lambda\}$
if $n \geq 0$.

(b) $P^k_n = P^{k+1}_n \cup kP^{k-1}_n$
if $n > k > 0$.

(c) $P^0_n = P^1_n$
if $n > 0$. 

Having defined n-paths, a new definition for hypertrees can be given. Since this definition allows for frontiering, these are called frontierable hypertrees. In frontierable hypertrees, a new type of symbol is introduced. It is called "X" and is used to hold one of the n-paths just defined. Since in the definition of frontiering, it is important to know a level also, there is a level encoded in the X. The use of the X will be obvious when frontiering is introduced.

Definition 14:

The set of frontierable hypertrees over $\Sigma$ of level $n$, denoted $H^+_n$, is the smallest set such that

(a) $H^+_0 = \Sigma$

(b) $H^+_n = \Sigma \cup X_n \cup \Sigma[H^+_n(\Sigma)]$

if $n > 0$ and $X_n = \{X^j_n | j \in \mathbb{P}^{n-1} \}$.

The same discussion which lead to the definition of hyperforests of degree $k$, definition 11, leads to the definition of frontierable hyperforests of degree $k$ which correspond directly to hyperforests. Similarly, hyperforests of degree $k$ and depth $m$ have a counterpart called frontierable hyperforests of degree $k$ and depth $m$. These are defined as follows:
Definition 15:

The set of frontierable hyperforests over $I$ of level $n$ and degree $k$, denoted $H_n^{k}$, is the smallest set such that

(a) $H_n^{n+1} = I$

if $n \geq 0$.

(b) $H_n^{0} = H_n^{1}$

if $n \geq 0$.

(c) $H_n^{k} = H_n^{k+1} \cup X_k \cup H_n^{k+1}[k H_n^{k-1} k]$

if $n \geq k > 0$ and $X_k = \{X_k^j | j \in P^{k-1}_k\}$.

Definition 16:

The set of frontierable hyperforests over $I$ of level $n$, degree $k$, and depth $m$, denoted $H_{n,m}^{k}$, is the smallest set such that

(a) $H_{n,m}^{n+1} = I$

if $n \geq 0$ and $m \geq 0$.

(b) $H_{n,0}^{k} = X_k$

if $n \geq k > 0$ and $X_k = \{X_k^j | j \in P^{k-1}_k\}$.

(c) $H_{n,m}^{k} = H_{n,m}^{k+1} \cup H_{n,m}^{k} \cup H_{n,m}^{k+1}[k H_{n,m}^{k-1} k]$

if $n \geq k > 0$ and $m \geq 0$.

(d) $H_{n,0}^{0} = H_{n,0}^{1}$

if $n \geq 0$ and $m \geq 0$. 
There is also a set of theorems corresponding to theorems 1 to 7. In each case, the proof follows the corresponding proof in the original theorem so closely that it need not be given here.

Theorem 8:
For all $n > k \geq 0$,

$$H_n^k = H_k(H_{k+1}(\ldots (H_n)\ldots ))$$

Proof:
This proof is the same as the proof of theorem 1.

---

Corollary 9:
For all $n \geq 0$,

$$H_n = H_n^n$$

Proof:
This proof is the same as the proof of corollary 2.

---

Lemma 10:
For all $n \geq 0$ and $n+1 \geq j \geq k \geq 0$,

$$H_n^j \subseteq H_n^k$$

Proof:
This proof is the same as the proof for lemma 3.

---
Lemma 11:
For all \( n \geq 0, n+1 \geq k \geq 0 \) and \( m \geq 0 \),
\[
H_{n,m}^k \subseteq H_n^k
\]
Proof:
This proof is the same as the proof of lemma 5.

Lemma 12:
If \( n \geq 0, n+1 \geq k \geq 0 \) and \( m \geq 0 \), then
\[
H_{n,m}^k = \bigcup_{m' \geq m \geq 0} H_{n,m'}^k
\]
Proof:
This proof is the same as lemma 6.

Lemma 13:
For all \( n \geq k > 0 \) and \( m \geq 0 \),
\[
H_{n,m+1}^k = H_{n,m}^{k+1} \cup H_{n,0}^k \cup H_{n,m}^{k+1} \setminus \left( H_{n,m}^{k-1} \right)
\]
Proof:
This proof is the same as the proof for lemma 4.

Theorem 14:
For all \( n \geq 0 \) and \( n+1 \geq k \geq 0 \),
\[
H_n^k = \bigcup_{m \geq 0} H_{n,m}^k
\]
Proof:
This proof is the same as the proof for theorem 7.

- - -

Having defined a path as a means of selecting a subhypertree, and having defined frontierable hypertrees, the selection function can now be defined. This function, called the search function, takes as its arguments a path and a hyperforest and returns the selected hypertree. The method of doing this is to consider the top level of the hyperforest -- that is the root nodes of all the sub-hypertrees -- and proceed down the path in the following manner. If there is a subhyperforest in the top level hyperforest described above with a degree corresponding to the first number in the path, then that hyperforest is selected and the top number is removed from the path. If there is no hyperforest corresponding to that number, then the function is not defined. This continues recursively until the end of the path is reached at which time the subhypertree corresponding to the first node is selected.

For example, if the hyperforest of level 4 and degree 2 is the hyperforest given in figure 7 and the path is "21321" then, after the first recursion the path becomes "1321" and the hyperforest is

\[ b[1c[3a[2b[1c_1]_2]_3][2a_2][1a_1]_1] \]
n-path: 21321

Selected Hypertree

\[ a_{2b[1c_{12}]}[3a_{2b[1c_{12}]}3][2a_{2}[1a_{11}2]] \]

FIGURE 7. A Hypertree of level 4 and degree 2

After recursion 2, the path is "321" and the hyperforest is

\[ c_{3a_{2b[1c_{12}]}3}[2a_{2}[1a_{11}]] \]

After the third recursion, the path is "21" and the hyperforest is

\[ a_{2b[1c_{12}]} \]
Then, after this pass, the path becomes "1" and the hyperforest is

\[ b[1c_1] \]

And lastly, the path becomes null and the hyperforest is

\[ c \]

so the selected hypertree is "c". Note that although this is a level 4 hypertree there are no level four brackets so that any selected subhypertrees will be single nodes. If the path had been "221", which is illegal, then recursion 2 would have had path "21" and hyperforest

\[ b[1c[3a[2b[1c_1]_2]_3][2a_2][1a_1]_1] \]

which has no subhypertree corresponding to number 2, so the select function fails and is not defined for these arguments.

Notice that the level of the hypertree has to be known at the time that the hypertree is finally selected so that the end of the selected hypertree can be determined. Also, it is convenient to know the level of the hyperforest that
is being input. Therefore, instead of a single selection function there is a family of selection functions, each corresponding to hyperforests of a given level and degree.

Definition 17:

Search is a family of functions, denoted \( \text{search}_{n}^{k} \), such that

\[
\text{search}_{n}^{k}: P_{n}^{k} \times H_{n}^{k} \rightarrow H_{n}
\]

if \( n \geq k \geq 0 \). The value of the functions are defined as follows:

(a) \( \text{search}_{n}^{n}(j, \gamma) = \gamma \) if \( n > 0 \).

(b) \( \text{search}_{n}^{0}(j, \gamma) = \text{search}_{n}^{1}(j, \gamma) \) if \( n > 0 \).

(c)

\[
\text{search}_{n}^{k}(j, \gamma) = \text{search}_{n}^{k+1}(j, \gamma)
\]

if \( \gamma \in H_{n}^{k+1} \) and \( j = ki \).

= undefined

if \( \gamma \in X_{n} \).

= undefined

if \( \gamma \in X_{n} \).

\[
\text{search}_{n}^{k+1}(j, \alpha)
\]

if \( \gamma = \alpha[k \beta k] \) and \( j \in P_{n}^{k+1} \).

= \text{search}_{n}^{k+1}(i, \beta)

if \( \gamma = \alpha[k \beta k] \) and \( j = ki \).

if \( n \geq k > 0 \).
Notice in equation a of definition 17, \( j \) must equal \( \lambda \) due to domain constraints.

**TABLE 2. Equation c of Definition 9**

<table>
<thead>
<tr>
<th>( j )</th>
<th>other</th>
<th>( k_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>other</td>
<td>( \text{search}^{k+1}_n(j, \gamma) )</td>
<td>error</td>
</tr>
<tr>
<td>( \alpha[k, \beta_k] )</td>
<td>( \text{search}^{k+1}_n(j, \alpha) )</td>
<td>( \text{search}^{k-1}_n(i, j) )</td>
</tr>
</tbody>
</table>

if \( n > k > 0 \).

Table 2 gives a diagram to aid in the understanding of definition 17. Notice that since \( n > k > 0 \)

\[ j \in p^{k+1}_n \cup k_p^{k-1}_n \]

by line c of definition 13. Therefore,

\[ j \in p^{k+1}_n \]

or else \( j \) is a string that begins with \( k \). Similarly,

\[ \gamma \in h^{k+1}_n \cup x_k \cup h^{k+1}_n \cup k \cdot h^{k-1}_n \cdot k. \]
In general, if the \( n \)-path \( j \) does not start with a number corresponding to the level of this search function then the next higher level search function is called on that part of the hypertree that forms the next higher level. This will continue to search the hyperforest until either the level which corresponds with the first number in \( j \) is called or \( \text{search}^n \) is called. In the latter case,

\[
  j \in P^n = \{1\}
\]

so that a hypertree is selected via line \( a \) of the definition. In the former case, the part of the hypertree in the brackets that correspond to the level selected is searched using the remainder of the path. If nothing is bracketed at that level, then the search is undefined.

It has been stated that the codomain of the \( \text{search}^n \) function is \( H^n \). However, there is nothing in the definition which guarantees that this is the case. Therefore, a theorem is presented which proves that the elements defined by the search function are in fact hypertrees, or, in other words, that the search function is well-formed.

Lemma 15:

If \( n \geq k \geq 0, m \geq 0, \gamma \in H_{n,m}^k, j \in P_n^k \) and \( \text{search}_n^k(j, \gamma) \) is defined, then

\[
  \text{search}_n^k(\gamma) \in H^n
\]
Proof by induction on m:

Basis: if \( m = 0 \) then either \( n = k \), \( n > k > 0 \) or \( k = 0 \).

1. If \( n = k \), then

\[
\gamma \in H_{n,0}^n
\]

\[
\in H_n^n
\]

\[
\in H_n
\]

by theorem 14 and corollary 9. Therefore,

\[
\text{search}_{n}(j, \gamma) = \gamma
\]

\[
\in H_n
\]

by definition 17.

2. If \( n > k > 0 \), then

\[
\gamma \in H_{n,0}^k
\]

\[
\in X_k
\]

by definition 16. Therefore, \( \text{search}_{n}(j, \gamma) \) is undefined.

3. If \( k = 0 \), then

\[
\text{search}_{n}^{0}(j, \gamma) = \text{search}_{n}^{1}(j, \gamma)
\]

and is undefined if \( n > 0 \). If \( n = 0 \), then

\[
\text{search}_{n}^{0}(j, \gamma) \in H_n
\]

by part 1 of this proof.
In any event, if \( \text{search}^k_n(j, \gamma) \) is defined then
\[
\text{search}^k_n(j, \gamma) \in H_n
\]

Induction step: Assume that if \( m > m' \geq 0 \) and \( \gamma \in H^k_{n,m'} \), then
\[
\text{search}^k_n(j, \gamma) \in H_n
\]
if it is defined. If \( \gamma \in H^k_{n,m} \), then one of the following cases must be true:

1. If \( k = n \), then

\[
\text{search}^k_n(j, \gamma) = \gamma
\]

by theorem 14 and corollary 9.

2. If \( n > k > 0 \), \( \gamma \in H^k_{n,m-1} \) and \( j \in P^{k+1}_n \), then

\[
\text{search}^k_n(j, \gamma) = \text{search}^{k+1}_n(j, \gamma)
\]

\( \in H_n \)

or is not defined by induction hypothesis.

3. If \( n > k > 0 \), \( \gamma \in H^k_{n,m-1} \) and \( j \in kP^{k-1}_n \), then \( \text{search}^k_n(j, \gamma) \) is not defined.

4. If \( n > k > 0 \), \( \gamma \in H^k_{n,m-1} \) and \( \text{search}^k_n(j, \gamma) \) is defined, then

\[
\text{search}^k_n(j, \gamma) \in H_n
\]
by induction hypothesis.

5. If $n > k > 0$, $\gamma = \alpha[\beta \beta \beta \ldots \beta]$ and $j \in \mathbb{P}^{k+1}_{n}$, then

$$\alpha \in H^{k+1}_{n,m=1}$$

and

$$\text{search}^{k}_{n}(j,\gamma) = \text{search}^{k+1}_{n}(j,\alpha) \in H_{n}$$

6. If $n > k > 0$, $\gamma = \alpha[\beta \beta \beta \ldots \beta]$ and $j = ki$, then

$$\beta \in H^{k-1}_{n,m=1}$$

and

$$\text{search}^{k}_{n}(j,\gamma) = \text{search}^{k-1}_{n}(i,\beta) \in H_{n}$$

by induction hypothesis.

7. If $k = 0$, then

$$\text{search}^{k}_{n}(j,\gamma) = \text{search}^{1}_{n}(j,\gamma) \in H_{n}$$

as a special case of one of the previous cases of this proof.

In all cases, assuming the induction hypothesis allows a proof that

$$\text{search}^{k}_{n}(j,\gamma) \in H_{n}$$
or else it is not defined, so that the lemma is proved by induction.

---

Theorem 15: If \( n \geq k \geq 0, j \in F_n^k \) and \( \text{search}_n^k(j, y) \) is defined, then

\[ \text{search}_n^k(j, y) \in H_n \]

Proof:
If \( y \in H_n^k \), then there must be an \( m \) such that

\[ y \in H_n^k \]

by theorem 14. Therefore,

\[ \text{search}_n^k(j, y) \in H_n \]

if it is defined and the theorem is proved.

---

Another function which is needed for the definition of the frontier function is the apply function. The apply function will take a hyperforest and replace all of the \( X_n^j \)'s with corresponding hypertrees. This is done by searching each of the nodes of the input hyperforest for nodes which are of the form \( X_n^j \). All other nodes are ignored. At these nodes, the appropriate search function is called to locate the subhypertree in the hyperforest that is to replace the
Like search, apply is a family of functions, denoted \( \text{apply}^k \), each returning a different level or degree of hyperforest.

**Definition 18:**

Apply is a family of partial functions, denoted \( \text{apply}^k \), such that

\[
\text{apply}^k : H^k_n \times H^{n-1}_n \rightarrow H^k_n
\]

for all \( n > 0 \) and \( n+1 \geq k \geq 0 \). The value of the functions are defined as follows:

(a) \( \text{apply}^{n+1}_n(\gamma, \psi) = \gamma \)

if \( n > 0 \).

(b) \( \text{apply}^0_n(\gamma, \psi) = \text{apply}^1_n(\gamma, \psi) \)

if \( n > 0 \).

(c) \( \gamma = \gamma \)

if \( \gamma \in X^j_k \) and \( n > k \).

\( \text{apply}^k_n(\gamma, \psi) = \text{search}^{n-1}_n(j, \psi) \)

if \( \gamma = X^j_n \) and \( n = k \).

\( = \text{apply}^{k+1}_n(\gamma, \psi) \)

if \( \gamma \in H^{k+1}_n \).

\( = \text{apply}^{k+1}_n(\alpha, \psi)[_k \text{apply}^{k-1}_n(\beta, \psi) _k] \)

if \( \gamma = \alpha[_k \beta _k] \).

if \( n \geq k > 0 \).
Just as was the case with the search function, the statement of the codomain of the apply function was made with no proof that the definition in fact maps the domain into the codomain. This proof will now be supplied in the following theorem.

Lemma 17:

If \( n > 0, n+1 \geq k \geq 0, m \geq 0, \gamma \in H^k_n, \psi \in H^k_{n,m} \) and \( \text{apply}^k_n(\gamma, \psi) \) is defined, then

\[ \text{apply}^k_n(\gamma, \psi) \in H^k_n \]

Proof by induction on \( m \):

Basis: If \( m = 0 \), then there are four possible cases:

1. If \( k = n+1 \), then by the definition of apply

\[ \text{apply}^k_n(\gamma, \psi) = \gamma \]

\[ \in H^{n+1}_n \]

2. If \( k = n \) and \( \text{search}^{n-1}_n(j, \psi) \) is defined, then

\[ \text{apply}^k_n(\gamma, \psi) = \text{search}^{n-1}_n(j, \psi) \]

\[ \in H^n_n \]

\[ \in H^n_n \]

by theorem 16 and corollary 9 where \( \gamma = x^j_n \).

3. If \( n > k > 0 \), then
\[(19) \quad \text{apply}^k_n(\xi, \psi) = \xi \quad \in H^k_n \]

since \( \xi \in X_k \) by definition 16.

4. If \( k = 0 \), then

\[
\text{apply}^k_n(\xi, \psi) = \text{apply}^1_n(\xi, \psi) \quad \in H^1_n
\]

as a special case of equation 19. Notice that \( n \neq 0 \).

Therefore, since \( H^0_n = H^1_n \),

\[
\text{apply}^k_n(\xi, \psi) \in H^0_n
\]

Since \( \text{apply}^k_n(\xi, \psi) \in H^k_n \) in each case, the basis is proved.

Induction step: Assume that if \( m > m' \geq 0 \), \( \xi \in H^k_{n,m'} \) and \( \psi \in H^{n-1}_n \) then

\[
\text{apply}^k_n(\xi, \psi) \in H^k_n
\]

if it is defined. If \( \xi \in H^k_{n,m'} \) then one of the cases given below must be true:

1. If \( k = n+1 \), then

\[
\text{apply}^k_n(\xi, \psi) = \xi \quad \in H^{n+1}_n
\]

due to the domain constraint.
2. If \( n \geq k > 0 \), \( \gamma \in H^{k+1}_{n,m-1} \) and \( \text{apply}^k_n(\gamma, \psi) \) is defined, then

\[
\text{apply}_n^k(\gamma, \psi) = \text{apply}^{k+1}_n(\gamma, \psi) \\
\in H^{k+1}_{n,m-1}
\]

by induction hypothesis. Also, since \( H_{n}^{k+1} \subseteq H_{n}^{k} \),

\[
\text{apply}_n^k(\gamma, \psi) \in H_{n}^{k}
\]

3. If \( n \geq k > 0 \), \( \gamma \in H^{k}_{n,m-1} \) and \( \text{apply}_n^k(\gamma, \psi) \) is defined, then

\[
\text{apply}_n^k(\gamma, \psi) \in H_{n}^{k}
\]

by induction hypothesis.

4. If \( n \geq k > 0 \) and \( \gamma = \alpha[k \beta]_k \), then \( \alpha \in H_{n,m-1}^{k+1} \) and \( \beta \in H_{n,m-1}^{k-1} \). Therefore, if \( \text{apply}_n^{k+1}(\alpha, \psi) \) and \( \text{apply}_n^{k-1}(\beta, \psi) \) are both defined then

\[
\text{apply}_n^k(\gamma, \psi) = \text{apply}_n^{k+1}(\alpha, \psi)[_k \text{apply}_n^{k-1}(\beta, \psi)_k] \\
\in H_{n}^{k+1}[k H_{n}^{k-1} k]
\]

by induction hypothesis. But since

\[
H_{n}^{k+1}[k H_{n}^{k-1} k] \subseteq H_{n}^{k}
\]

(20) \( \text{apply}_n^k(\gamma, \psi) \in H_{n}^{k} \)

5. If \( k = 0 \), then
apply^k_n(\gamma, \psi) = apply^1_n(\gamma, \psi)
\in H^1_n

if it is defined. This is a special case of equation 20. But since H^0_n = H^1_n',
apply^k_n(\gamma, \psi) \in H^0_n
\in H^k_n

Therefore, since
apply^k_n(\gamma, \psi) \in H^k_n

in each case, the lemma is proved by induction.

- • -

Theorem 18:
If n > 0, n+1 \geq k \geq 0, \gamma \in H^k_{n'}, \psi \in H^{n-1}_n and
apply^k_n(\gamma, \psi) is defined, then
apply^k_n(\gamma, \psi) \in H^k_n

Proof:
By theorem 14 there exists an m such that
\gamma \in H^k_{n,m}

Therefore, by lemma 17,
apply^k_n(\gamma, \psi) \in H^k_n

if it is defined at that point.

- • -

Using the apply function, the frontier may now be defined
Definition 19:

Frontier is a family of partial functions, denoted \( \text{frontier}_n^k \), such that

\[
\text{frontier}_n^{n+1} : H_{n+1}^n \to H_n^n
\]

if \( n \geq 0 \).

\[
\text{frontier}_n^k : H_k^n \to H_k^n
\]

if \( n \geq k \geq 0 \). The values are defined as follows:

(a) \[
\text{frontier}_n^k(\gamma) = \gamma \quad \text{if} \quad \gamma \in H_{n+1}^{k+1}.
\]

(b) \[
\text{frontier}_n^0(\gamma) = \text{frontier}_n^1(\gamma)
\]

if \( n \geq 0 \).

(c) \[
\begin{align*}
\text{frontier}_n^k(\gamma) &= \gamma & \text{if} \quad \gamma \in H_k^n \text{ and } n \geq k > 0. \\
&= \text{frontier}_n^{k+1}(\gamma) & \text{if} \quad \gamma \in H_{n+1}^{k+1} \text{ and } n \geq k > 0. \\
&= \text{frontier}_n^{k+1}(a)[_k \text{frontier}_n^{k-1}(\beta)]_k & \text{if} \quad \gamma = a[_k \beta]_k \text{ and } n > k. \\
&= \text{apply}_n(\text{frontier}_n^{n+1}(a), \text{frontier}_n^{n-1}(\beta)) & \text{if} \quad \gamma = a[_k \beta]_k, \ a \notin \Sigma \text{ and } n = k. \\
&= a[_k \text{frontier}_n^{n-1}(\beta)]_k & \text{if} \quad \gamma = a[_k \beta]_k, \ a \in \Sigma \text{ and } n = k.
\end{align*}
\]
As has been done previously, a theorem will now be proved establishing that the codomain stated in the definition of the frontier is correct.

**Lemma 19:**

If \( m \geq 0 \), \( y \in H_{n+1,m}^k \) and \( \text{frontier}_n^k(y) \) is defined,

then

\[
\text{frontier}_n^k(y) \in H_n^k
\]

if \( n \geq k \geq 0 \).

\[
\text{frontier}_n^k(y) \in H_n^k
\]

if \( n \geq 0 \) and \( k = n+1 \).

**Proof by induction on** \( m \):

**Basis:** if \( m = 0 \) then one of the following cases is true:

1. If \( k = n+1 \), then

\[
y \in X_{n+1}
\]

and \( \text{frontier}_n^k(y) \) is not defined.

2. If \( n \geq k > 0 \), then

\[
y \in X_k
\]

and

\[
\text{frontier}_n^k(y) = y
\]

But since \( H_{n,0}^k = X_k \) it must be true that

\[
y \in H_{n,0}^k
\]

and
3. If \( k = 0 \) and \( n = 0 \), then

\[
\text{frontier}^k_n(\gamma) = \text{frontier}^1_0(\gamma)
\]

But by case 1 above, \( \text{frontier}^1_0(\gamma) \) is not defined. Therefore, \( \text{frontier}^k_n(\gamma) \) is not defined.

4. If \( k = 0 \) and \( n \geq 0 \), then

\[
\text{frontier}^k_n(\gamma) = \text{frontier}^1_n(\gamma)
\]

By case 2 above,

\[
\text{frontier}^k_n \subseteq H^1_n
\]

Since \( H^0_n = H^1_n \), it follows that

\[
\text{frontier}^k_n(\gamma) \subseteq H^0_n
\]

Since in each case, either the frontier was not defined or was an element of the appropriate codomain, the basis is established.

Induction step: Assume that if \( m > m' \geq 0 \) and \( \gamma \subseteq H^k_{n+1,m'} \) then \( \text{frontier}^k_n(\gamma) \) is an element of the appropriate codomain or else is undefined. If \( \gamma \subseteq H^k_{n+1,m'} \) then one of the following conditions is true:

1. If \( k = n+1 \) and \( \gamma \subseteq H^{n+2}_{n+1,m-1} \), then \( \gamma \subseteq \Sigma \) and

\[
\text{frontier}^k_n(\gamma) = \gamma
\]

Since \( \Sigma \subseteq H^k_{n,m'} \).
2. If \( k = n+1 \) and \( \gamma \in H_{n+1,m-1}^{n+1} \), then
\[
\text{frontier}_{n}^{k}(\gamma) \in H_{n}^{n}
\]
or is not defined by the induction hypothesis.

3. If \( k = n+1 \) and \( \gamma = a[k \beta k] \), then
\[
\text{frontier}_{n}^{k}(\gamma) = \text{frontier}_{n}^{n}(\beta)
\]
But \( \beta \in H_{n+1,m-1}^{n} \) and by induction hypothesis
\[
\text{frontier}_{n}^{n}(\beta) \in H_{n}^{n}
\]
if \( \text{frontier}_{n}^{n}(\beta) \) is defined. Therefore,
\[
\text{frontier}_{n}^{k}(\gamma) \in H_{n}^{n}
\]
if \( \text{frontier}_{n}^{k}(\gamma) \) is defined.

4. If \( n \geq k > 0 \) and \( \gamma \in H_{n+1,m-1}^{k+1} \), then
\[
\text{frontier}_{n}^{k}(\gamma) = \text{frontier}_{n}^{k+1}(\gamma)
\]
But since \( \text{frontier}_{n}^{k+1}(\gamma) \in H_{n}^{k+1} \) by induction hypothesis, and \( H_{n}^{k+1} \subseteq H_{n}^{k} \)
\[
\text{frontier}_{n}^{k}(\gamma) \in H_{n}^{k}
\]
if it is defined.

5. If \( n \geq k > 0 \) and \( \gamma \in H_{n+1,m-1}^{k} \), then
\[
\text{frontier}_{n}^{k}(\gamma) \in H_{n}^{k}
\]
by induction hypothesis provided that \( \text{frontier}^k_n(\gamma) \) is defined.

6. If \( k = n \) and \( \gamma = \alpha[\beta_1 \ldots \beta_k] \), then

\[
\text{frontier}^k_n(\gamma) = \text{apply}_n^n(\text{frontier}^{n+1}_n(\gamma), \text{frontier}^{n-1}_n(\beta))
\]

But by theorem 18, if the frontier is defined

\[
\text{frontier}^k_n(\gamma) \in H^n
\]

(Note that the arguments to apply are in domain for apply by induction hypothesis.)

7. If \( n > k > 0 \) and \( \gamma = \alpha[\beta_1 \ldots \beta_k] \), then

\[
\text{frontier}^k_n(\gamma) = \text{frontier}^{k+1}_n(\alpha)[\beta_1 \ldots \text{frontier}^{k-1}_n(\beta)]
\]

Notice that since \( n > k \), neither \( k+1 \) nor \( k-1 \) can equal \( n+1 \). Therefore,

\[
\text{frontier}^{k+1}_n(\alpha) \in H^{k+1}_n
\]

and

\[
\text{frontier}^{k-1}_n(\beta) \in H^{k-1}_n
\]

if they are defined, and

\[
\text{frontier}^k_n(\gamma) \in H^k_n[\beta_1 \ldots \text{frontier}^{k-1}_n(\beta)]
\]

if it is defined.

8. If \( k = 0 \), then
Therefore, by a previous step of this proof,

\[ \text{frontier}^1_n(\tau) \subseteq \mathcal{H}_n \]

and finally,

\[ \text{frontier}^0_n(\tau) \subseteq \mathcal{H}_n \]

In any event, if \( \text{frontier}^k_n(\tau) \) is defined it is in the appropriate codomain, so the lemma is proved by induction.

---

**Theorem 20:**

If \( \tau \in \mathcal{H}^k_{n+1} \) and \( \text{frontier}^k_n(\tau) \) is defined, then if \( n \geq k \geq 0 \) then

\[ \text{frontier}^k_n(\tau) \subseteq \mathcal{H}^k_n \]

If \( n \geq 0 \) and \( k = n+1 \), then

\[ \text{frontier}^k_n(\tau) \subseteq \mathcal{H}^n_n \]

**Proof:**

This follows directly from lemma 19.

---

For example, the frontier of the tree given on line 5 of table 1 is given by

\[ \text{frontier}^2_1(a_2a_1b[1c_1]_1)_2) = \text{frontier}^1_1(a_1b[1c_1]_1) \]

\[ = a[1 \text{frontier}^0_1(b[1c_1]_1)_1] \]
which is the string "abc". The tree given may also be written $a(abc)$, so the frontier function obtained the traditional frontier. If the tree given in line 6 of table 1 is chosen, then the frontier will be

$$\text{frontier}^2_1(a_2a_2[a_1b_1]_2)[_1b_1c_2[a_2a_2[a_1b_1]_2]_2]_2]_2)$$

$$= \text{frontier}^1_1(a_2a_2[a_1b_1]_2)[_1b_1c_2[a_2a_2[a_1b_1]_2]_2]_2]_2)$$

$$= \text{apply}^1_1(\text{frontier}^2_1(a_2a_2[a_1b_1]_2),$$

$$\text{frontier}^0_1(b_1c_2[a_2a_2[a_1b_1]_2]_2)_2)$$

$$= \text{apply}^1_1(\text{frontier}^1_1(a_1b_1),$$

$$\text{frontier}^1_1(b_1c_2[a_2a_2[a_1b_1]_2]_2)_2)$$

$$= \text{apply}^1_1(\text{a_1frontier}^0_1(b),$$

$$b_1\text{frontier}^0_1(c_2[a_2a_2[a_1b_1]_2]_2)_2)$$

$$= \text{apply}^1_1(\text{a_1frontier}^1_1(b),$$

$$b_1\text{frontier}^1_1(c_2[a_2a_2[a_1b_1]_2]_2)_2)$$

$$= \text{apply}^1_1(\text{a_1frontier}^2_1(b),$$

$$b_1\text{frontier}^2_1(c_2[a_2a_2[a_1b_1]_2]_2)_2)$$

$$= a_1\text{frontier}^1_1(b_1c_1)_1]$$
Note that the frontier corresponds to the string "ab". If the tree a(a(ab)bc(a(ab))) is drawn out as a graph, it can be seen that the frontier should be "abbab". This discrepancy is caused by the fact that with level 2 hypertrees it must be indicated how the frontier is hooked up using the "X"'s. If a portion of the nodes with no subtrees is not referenced by an "X", then those nodes are dropped. This is exactly analogous to dropping an argument in a macrogrammar. Therefore, the hypertree corresponding to a(a(ab)bc(a(ab))) is actually

\[ a[2a[2a[a[1b[1X_1]_1]_2][1b[1c[2a[2a[a[1b]_1]_2]_1]_2]]_2] \]

The frontier of this hypertree is
frontier\(^2\)(a\([2a\[2a[1b[1\la\sub{1}l_1]_2]_2]_2c[2a[2a[1b[2]_1]_2]_1]_2])

= frontier\(^1\)(a\([2a[1b[1\la\sub{1}l_1]_2]_2]_2c[2a[2a[1b[2]_1]_2]_1]_2))

= apply\(^1\)(frontier\(^2\)(a\([2a[1b[1\la\sub{1}l_1]_2]_2]))

= apply\(^0\)(apply\(^1\)(a\([1a[1b[1\la\sub{1}l_1]_2]_2]))

= apply\(^0\)(a, b\([1a[1b[1\la\sub{1}l_1]_2]_2]))

which is the "abbbab" that was expected.

At this point, two definitions can be made which will formalize the premise of this paper.

Definition 20:

A string is a frontierable hypertree of level 1.
Definition 21:

A tree is a frontierable hypertree of level 2.

Previously in this paper it has been stated that hypertrees over \( \Xi, h_n^\Xi \), is a restricted case of frontierable hypertrees over \( \Xi, H_n \). This can now be proven. First, a lemma must be proved to establish that hyperforests are a restricted case of frontierable hyperforests. This is done by induction on the depth of the hyperforest.

Lemma 21:

For all \( n \geq 0, m \geq 0 \) and \( n+1 \geq k \geq 0 \),

\[
\begin{align*}
\text{Lemma } 21: \\
\forall n \geq 0, m \geq 0 \text{ and } n+1 \geq k \geq 0,
\end{align*}
\]

\[
h_n^k, m \subseteq h_n^k, m
\]

Proof by induction on \( m \):

Basis: If \( m = 0 \), then one of the following is the case

1. If \( k = n+1 \), then

\[
h_n^{n+1} = \Xi
\]

\[
= h_n^{n+1}
\]

therefore,

\[
h_n^{n+1} \subseteq h_n^{n+1}
\]

2. If \( n \geq k > 0 \), then

\[
h_n^k, 0 = \emptyset
\]

\[
\subseteq h_n^k, 0
\]

3. If \( k = 0 \), then
Since in each of these cases,
\[ h_{n,0}^k = h_{n,0}^k \]
the basis is proved.

Induction step: Assume that if \( m > m' \geq 0 \) then
\[ h_{n,m}^k \leq h_{n,m}^k \]
If \( m > 0 \), then there are three cases to consider
1. If \( k = n+1 \) then
\[ h_{n,m}^{n+1} = h_{n,m}^{n+1} = h_{n,m}^{n+1} \]
Therefore,
\[ h_{n,m}^{n+1} \leq h_{n,m}^{n+1} \]
(21)
\[ h_{n,m}^{n+1} \leq h_{n,m}^{n+1} \]

2. If \( n \geq k > 0 \), then
\[ h_{n,m}^k = h_{n,m-1}^k U h_{n,m-1}^k U h_{n,m-1}^k \]
but since by induction hypothesis
\[ h_{n,m-1}^{k+1} \leq h_{n,m-1}^{k+1} \]
\[ h_{n,m-1}^k \leq h_{n,m-1}^k \]
\[ h_{n,m-1}^{k-1} \leq h_{n,m-1}^{k-1} \]
Therefore, the previous equation becomes

\[(22) \quad h_{n,m}^k \leq H_{n,m-1}^{k+1} U H_{n,m-1}^{k} U H_{n,m-1}^{k+1} [k H_{n,m-1}^{k-1} k] \leq H_{n,m}^k \]

so that the theorem is true in this case also.

3. Lastly, if \( k = 0 \) then

\[h_{n,m}^0 = h_{n,m}^1\]

But by equation 21 or 22 depending on the value of \( n \)

\[h_{n,m}^1 \leq H_{n,m}^1 \leq H_{n,m}^0 \]

Therefore, in this case also

\[h_{n,m}^k \leq H_{n,m}^k\]

Since in each of the above cases, \( h_{n,m}^k \leq H_{n,m}^k \), the lemma is proved by induction.

---

Theorem 22:

For all \( n \geq 0 \),

\[h_n \leq H_n\]

Proof:
$$h_n = h_n^m$$
$$= \bigcup_{m \geq 0} h_{n,m}^m$$
$$\subseteq \bigcup_{m \geq 0} H_{n,m}^m$$

by lemmas 2, 7 and 21. Therefore,

$$h_n \subseteq H_{n}^m$$
$$\subseteq H_n$$

Therefore,

$$h_n \subseteq H_n$$

for all $n \geq 0$ and the theorem is proved.

---

The original justification for studying hypertrees was the characterization of the algebraic language hierarchy. This requires that each level of the hypertree hierarchy have at least a regular grammar defined over it. The frontier function will be used then to extract the elements of the algebraic language hierarchy from the regular hypertree languages. Since a regular language requires that nonterminals appear only on the "frontier" of each of the right hand sides of the productions in the hypertree grammar, a new definition must be given for the hierarchy which will allow nonterminals to appear on the "frontier". With the definitions given above, requiring the nonterminals to be on the frontier can only be done informally since the
phrase "on the frontier" has not been defined. This new definition of hypertrees, hypertrees with arguments, arguments are provided which are only allowed on the frontier of the hypertrees.

Before a formal definition for hypertrees with arguments can be given, a formal definition for ranked sets must be given since these will be used in the definition of hypertrees with arguments.

Definition 22:

A ranked set \( V \) is an ordered pair \((V',r)\) where \( V' \) is a set and \( r \) is a function such that

\[
r: V' \to N
\]

where \( N \) is the set of positive integers.

Definition 23:

A rank of the ranked set \( V \), denoted \( V_k \), is defined by

\[
V_k = r^{-1}(k) = \{ a \mid r(a) = k \}
\]

Note that since a function is used in defining ranked sets the ranks are disjoint. The definition of hypertrees
and hyperforests with arguments can now be given. As should
be expected, these definitions correspond rather closely to
the definition of hypertrees, hyperforests, frontierable
hypertrees and frontierable hyperforests already given.

Definition 24:

The set of hypertrees over $\mathcal{Z}$ of degree $n$ with
arguments from the ranked set $\mathcal{V}$, denoted $H_n(\mathcal{Z}, \mathcal{V})$,
is the smallest set such that

(a) $H_0(\mathcal{Z}, \mathcal{V}) = \mathcal{Z}$
(b) $H_n(\mathcal{Z}, \mathcal{V}) = \mathcal{Z} \cup V_n \cup X_n \cup \bigcup_{m \leq n} H_{n-1}(H_{n}(\mathcal{Z}, \mathcal{V}), \mathcal{V})$

if $n > 0$.

Definition 25:

The set of hyperforests over $\mathcal{Z}$ of level $n$ and
degree $k$ with arguments from the ranked set $\mathcal{V}$,
denoted $H_n^k(\mathcal{Z}, \mathcal{V})$, is the smallest set such that

(a) $H_n^{n+1}(\mathcal{Z}, \mathcal{V}) = \mathcal{Z}$

if $n \geq 0$.
(b) $H_n^0(\mathcal{Z}, \mathcal{V}) = H_n^{-1}(\mathcal{Z}, \mathcal{V})$

if $n \geq 0$.
(c) $H_n^k(\mathcal{Z}, \mathcal{V}) =$

$H_n^{k+1}(\mathcal{Z}, \mathcal{V}) \cup V_k \cup X_k \cup H_n^{k+1}(\mathcal{Z}, \mathcal{V})[k H_n^{k-1}(\mathcal{Z}, \mathcal{V}) k]$

if $n \geq k > 0$. 
Definition 26:

The set of hyperforests over $\Sigma$ of level $n$, degree $k$ and depth $m$ with arguments from the ranked set $V$, denoted $H_{n,m}^k(\Sigma, V)$, is the smallest set such that

(a) $H_{n,m}^{n+1}(\Sigma, V) = \Sigma$
if $n \geq 0$ and $m \geq 0$.

(b) $H_{n,0}^k(\Sigma, V) = V_k \cup X_k$
if $n \geq k > 0$.

(c) $H_{n,m+1}^k(\Sigma, V) = H_{n,m}^{k+1}(\Sigma, V) \cup H_{n,m}^k(\Sigma, V) \cup H_{n,m}^{k+1}(\Sigma, V)[H_{n,m}^{k-1}(\Sigma, V) \cup X_k]$
if $n \geq k > 0$ and $m \geq 0$.

(d) $H_{n,m}^0(\Sigma, V) = H_{n,m}^1(\Sigma, V)$
if $n \geq 0$ and $m \geq 0$.

Notice that throughout these definitions $V_k$ always appears beside the $X_k$. From this fact, it is apparent that there is a relationship between the $V_k$ and the $X_k$. In each case, the argument or the "$X$" may be replaced with an element of $H_{n}^k(\Sigma, V)$ and a well-formed hypertree will result. In the case of the "$X$"'s, this fact was used to substitute the appropriate subhypertree into the frontier. In the case of the argument, this will be used to develop a grammar.

Just as in the case of a hypertree over $\Sigma$, the $V$ will be elided throughout the rest of this paper if no confusion
would result. This will result in $H_n^k(\Sigma)$ and $H_n^k(\Sigma, V)$ both being represented by $H_n^k$, but in most instances this will not cause a confusion because it will be obvious which notation is being referenced.

Theorems 1 to 7 can also be extended to hypertrees with arguments.

**Theorem 23:**

For all $n \geq k \geq 0$,

$$H_n^k = H_k(H_{k+1}( \ldots (H_n(\Sigma, V), V) \ldots ), V)$$

**Proof:**

This proof is the same as theorem 1.

---

**Corollary 24:**

For all $n \geq 0$,

$$H_n(\Sigma, V) = H_n^0(\Sigma, V)$$

**Proof:**

This proof is the same as corollary 2.

---

**Lemma 25:**

For all $n \geq 0$ and $n+1 \geq j \geq k \geq 0$,

$$H_n^j(\Sigma, V) \subset H_n^k(\Sigma, V)$$

**Proof:**

This proof is the same as the proof for lemma 3.

---
Lemma 26:

For all $m \geq 0$, $n \geq 0$ and $n+1 \geq k \geq 0$,

$$H^k_{n,m}(\Sigma, V) \subseteq H^k_n(\Sigma, V)$$

Proof:

This lemma is proved along the same lines as lemma 5.

Lemma 27:

If $n \geq 0$, $n+1 \geq k \geq 0$ and $m \geq 0$, then

$$H^k_{n,m}(\Sigma, V) = \bigcup_{m \geq m' \geq 0} H^k_{n,m}(\Sigma, V)$$

Proof:

This proof is the same as the proof for lemma 6.

Lemma 28:

For all $n \geq k \geq 0$ and $m \geq 0$,

$$H^k_{n,m+1}(\Sigma, V) = H^k_{n,m} \cup H^k_{n,0} \cup H^k_{n,m}[k \mid H^k_{n,m} \]$$

Proof:

This proof follows the lines of lemma 4.

Theorem 29:

For all $n \geq 0$ and $n+1 \geq k \geq 0$,

$$H^k_n = \bigcup_{m \geq 0} H^k_{n,m}$$

Proof:

This proof is the same as theorem 7.
Without formally stating them, it will be assumed that the search, apply and frontier functions may be extended to hypertrees with arguments. For the search function, this extension will only require the appropriate change in the domain and codomain. For the apply function, besides the domain change, the line
\[ \text{apply}_n^k(\tau, \psi) = \tau \]
needs to be added to handle the case where \( \tau \in V_k \). Besides the domain changes, the frontier function needs the line
\[ \text{frontier}_n^k(\tau, \psi) = \tau \]
added to handle the case where \( \tau \in V_k \) and \( n > k \geq 0 \), and the line
\[ \text{frontier}_n^k(\tau, \psi) = \text{undefined} \]
needs to be added to handle the \( n = k > 0 \) case. Similarly, a string with arguments and a tree with arguments will be taken as the obvious extension of the definition of trees and strings.

Theorems 21 and 22 will now be extended.
Lemma 30:

For all \( n \geq 0, m \geq 0 \) and \( n+1 \geq k \geq 0 \),

\[
    h_{n,m}^k(\Xi) \subseteq H_{n,m}^k(\Xi) \subseteq H_{n,m}^k(\Xi,V)
\]

Proof:

This proof is the same as the proof for lemma 21.

---

Theorem 31:

For all \( n \geq 0 \),

\[
    h_n(\Xi) \subseteq H_n(\Xi) \subseteq H_n(\Xi,V)
\]

Proof:

This is proved in the same fashion as theorem 22.

---

This concludes the basic definitions that are needed for this development of hypertrees. Subsequent chapters will be involved in showing additional relationships between trees, strings and hypertrees. Also, grammars, automata and regular expressions will be developed to describe regular languages on hypertrees, and these languages will be characterized. This characterization then will be used to characterize the algebraic language hierarchy.
In order to be useful, there must be some reasonable way of representing hypertrees in a computer. Perhaps the most reasonable representation is obtained by extending the concept of binary encoded trees to hypertrees. It will be shown that, in fact, linked list representation of strings is related to the binary representation of trees. In general, it will be shown that a level n hypertree may be conveniently represented as an n-ary tree. The two special cases presented are that of trees, which may be represented as binary trees, and strings, which may be represented as unary trees (linked lists).

In order to establish this relationship, it is first necessary to give a formal definition of what is meant by n-ary trees.

**Definition 27:**

The set of n-ary trees over \( \Sigma \), denoted \( T'_n(\Sigma) \), is the smallest set such that

\[
T'_n(\Sigma) = \{\lambda\} \cup \Sigma [ (T'_n(\Sigma))^n ]
\]

for all \( n \geq 0 \).

As in the previous sections, the linear representation of n-ary trees will be used instead of the graph theoretic
representation. Also, the set \( \mathcal{I} \) will be elided for the remainder of this paper unless it is not clear what set the trees are over.

In order to define a simple mapping from hypertrees to n-ary trees, a subset of the n-ary trees will be defined which exactly matches the image of the hypertrees of level \( n \). Then it will be proved that, in fact, this image is a subset of the n-ary trees over \( \mathcal{I} \). Also, it will be shown that a corresponding image can be defined for \( H_{n}^{k}(\mathcal{I}) \) and \( H_{n,m}^{k}(\mathcal{I}) \).

**Definition 28:**

The set of n-ary encoded hyperforests of degree \( k \), denoted \( T_{n}^{k} \), is defined as

\[
T_{n}^{k} = t_{n}^{k}(\mathcal{I})^{\lambda \max(k-1,0)}
\]

if \( n \geq 0 \) and \( n+1 \geq k \geq 0 \), where \( t_{n}^{k}(\mathcal{I}) \) is the smallest set such that

(a) \( t_{n}^{n+1}(\mathcal{I}) = \mathcal{I}[ \]
if \( n \geq 0 \).

(b) \( t_{n}^{0}(\mathcal{I}) = t_{n}^{1}(\mathcal{I}) \)
if \( n \geq 0 \).

(c) \( t_{n}^{k}(\mathcal{I}) = t_{n}^{k+1}(\mathcal{I})^{\lambda} \cup t_{n}^{k+1}(\mathcal{I})T_{n}^{k-1} \)
if \( n \geq k > 0 \).
Definition 29:

The set of $n$-ary encoded hypertrees over $ξ$, denoted $T_n$, is defined as

$$T_n = T_n^n$$

if $n \geq 0$.

Definition 30:

The set of $n$-ary encoded hypertrees over $ξ$ of degree $k$ and depth $m$, denoted $T_{n,m}^k$, is defined as

$$T_{n,m}^k = T_{n,m}^k(ξ)^{\lambda \max(k-1,0)}$$

if $n \geq 0$, $n+1 \geq k \geq 0$ and $m \geq 0$ where $T_{n,m}^k(ξ)$ is the smallest set such that

(a) $T_{n,m}^{n+1}(ξ) = ξ$ if $n \geq 0$ and $m \geq 0$.

(b) $T_{n,0}^k(ξ) = Φ$ if $n \geq k > 0$.

(c) $T_{n,m+1}^k(ξ) = T_{n,m}^k(ξ) \cup T_{n,m}^{k+1}(ξ) T_{n,m}^{k-1}$ if $n \geq k > 0$ and $m \geq 0$.

(d) $T_{n,m}^0(ξ) = T_{n,m}^1(ξ)$ if $n \geq 0$ and $m \geq 0$.

These definitions are so similar to the definitions of hypertrees given in the previously in chapter 2 that a set
of theorems may be given which correspond to lemma 3 to theorem 7 in that chapter.

Lemma 32:

For all \( n \geq j \geq k \geq 0 \)

\[
T_n^j \subseteq T_n^k
\]

Proof (patterned after the proof of lemma 3 in chapter 2):
Define \( T_n^k(\Sigma) \) as it is defined in definition 28. Prove the lemma by induction on \( k \).

Basis: If \( k = j \), then

\[
T_n^j = T_n^k
\]

Therefore,

\[
T_n^j \subseteq T_n^k
\]

Induction step: Assume that if \( j \geq k' > k \geq 0 \) then

\[
T_n^j \subseteq T_n^{k'}
\]

Since \( k \geq 0 \) it must be true that either \( k > 0 \) or \( k = 0 \).

1. If \( k > 0 \), then

\[
(1) \quad T_n^k = t_n^k(\Sigma)^{\max(k-1,0)}] = t_n^{k+1}(\Sigma)^{\lambda \max(k-1,0)}] = t_n^{k+1}(\Sigma)T_n^{k-1,\max(k-1,0)}]
\]

Now, since \( k > 0 \),

\[
\max(k-1,0) = k-1
\]
equation 1 becomes

\[ T_n^k = t_n^{k+1}(z)\lambda^{\max((k+1)-1,0)} \quad \cup \quad t_n^{k+1}(z)T_n^{k-1}\lambda^{\max(k-1,0)} \]

But by induction hypothesis

\[ T_n^j \subset T_n^{k+1} \]

Now,

\[ T_n^{k+1} = t_n^{k+1}(z)\lambda^{\max((k+1)-1,0)} \]

\[ \subset T_n^k \]

therefore,

\[ T_n^j \subset T_n^k \]

2. If \( n \geq j > k = 0 \), then

\[ T_n^0 = t_n^0(z)\lambda^{\max(0-1,0)} \]

\[ = t_n^1(z)\lambda^{\max(0-1,0)} \]

\[ = t_n^1(z)\lambda^{\max(1-1,0)} \]

\[ = T_n^1 \]

By induction hypothesis

\[ T_n^j \subset T_n^1 \]

therefore,

\[ T_n^j \subset T_n^0 \]

and the lemma is proved by induction.
Lemma 33:

For all m \geq 0, n \geq 0 and n+1 \geq k \geq 0 it is true that

\[ g^{k, n, m} = T^{k, n} \]

Proof (patterned after the proof for lemma 5 in chapter 2):

Define \( t^k_n(\xi) \) as it is defined in definition 28 and \( t^k_{n,m}(\xi) \) as it is defined in definition 30. Prove the lemma by induction on m.

Basis: if m = 0 then

\[ T^{k, n, 0} = t^{k, n, 0}(\xi) = t^{k, n}(\xi) \]

and one of the following is true:

1. If k = n+1, then

\[ t^{k, n, 0}(\xi) = \xi \]

therefore,

\[ t^{k, n, 0}(\xi) = t^{k, n}(\xi) \]

and

\[ T^{k, n, 0} = t^{k, n}(\xi) = T^{k, n}(\xi) \]

2. If n \geq k > 0, then

\[ t^{k, n, 0}(\xi) = \emptyset \]

therefore,
\[ T_{n, 0}^k = \emptyset \]

and

\[ T_{n, 0}^k \subset T_n^k \]

since the null set is a subset of every set.

3. If \( k = 0 \), then

\[ T_{n, 0}^k = T_{n, 1}^k \]

\[ \subset T_n^1 \]

by either step 1 or step 2 above. Since \( T_n^0 = T_n^1 \),

\[ T_{n, 0}^k \subset T_n^0 \]

and the basis is proven.

Induction step: Assume that if \( m > m' \geq 0 \) then

\[ T_{n, m'}^k \subset T_n^k \]

If \( m > 0 \), then

\[ T_{n, m}^k = t_{n, m}^k (\Sigma \lambda^\max(k-1, 0)) \]

and one of the following conditions must be true:

1. If \( k = n+1 \), then

\[ T_{n, m}^{n+1} = \Sigma [\lambda^\max((n+1)-1, 0)] \]

\[ = t_n^{n+1} (\Sigma \lambda^\max((n+1)-1, 0)) \]

\[ = T_n^{n+1} \]
2. If \( n \geq k > 0 \), then

\[
T_{n,m}^k = t_{n,m-1}^{k+1}(\xi)\lambda^{\max(k-1,0)}
\]

\[
U t_{n,m-1}^{k+1}(\xi)T_{n,m-1}^{k-1}\lambda^{\max(k-1,0)}
\]

Note that

\[
T_{n,m}^{k+1} = T_{n,m-1}^{k+1}
\]

by induction hypothesis, so that

\[
t_{n,m-1}^{k+1}(\xi)\lambda^{\max(k-1,0)} \leq t_{n}^{k+1}(\xi)\lambda^{\max(k-1,0)}
\]

and therefore,

\[
t_{n,m-1}^{k+1}(\xi) \leq t_{n}^{k+1}(\xi)
\]

Therefore, by this equation and induction hypothesis

\[
T_{n,m}^k \leq t_{n}^{k+1}(\xi)\lambda^{\max(k-1,0)}
\]

\[
U t_{n}^{k+1}(\xi)T_{n,m-1}^{k-1}\lambda^{\max(k-1,0)}
\]

\[
= (t_{n}^{k+1}(\xi) \leq t_{n}^{k+1}(\xi)T_{n,m-1}^{k-1})\lambda^{\max(k-1,0)}
\]

\[
\leq t_{n}^{k}(\xi)\lambda^{\max(k-1,0)}
\]

\[
= t_{n}^{k}
\]

and the lemma is true in this case also.

3. If \( k = 0 \), then

\[
T_{n,m}^0 = t_{n,m}^0(\xi)\lambda^{\max(0-1,0)}
\]

\[
= t_{n,m}^1(\xi)\lambda^{\max(0-1,0)}
\]
85

= t_{n,m}^1 (\lambda)_{\max(1-1,0)}

= T_{n,m}

and by either part 1 or part 2 of this proof depending on the value of n

\[ T_{n,m}^1 \leq T_{n}^1 \leq T_{n}^0 \]

Therefore,

\[ T_{n,m}^0 \leq T_{n}^0 \]

in this case also.

Since the lemma is true for each case, the lemma is proved by induction.

---

Lemma 34:

If \( n \geq 0, n+1 \geq k \geq 0 \) and \( m \geq 0 \), then

\[ T_{n,m}^k = U_{m \geq m' \geq 0} T_{n,m'}^k \]

Proof (patterned after the proof of lemma 6):

Define \( t_{n,m}^k (\xi) \) as it is defined in definition 30. Prove the theorem by induction on m:

Define

\[ a_{n,m}^k (\xi) = U_{m \geq m' \geq 0} T_{n,m}^k \]
Basis: if \( m = 0 \) then

\[
\alpha_{n,0}^k(\xi) = U_{m'=0}^k t_{n,m}^k = t_{n,0}^k
\]

Induction step: Assume that if \( m \geq 0 \) then

\[
\alpha_{n,m}^k(\xi) = t_{n,m}^k
\]

then by the definition of union,

\[
\alpha_{n,m+1}^k(\xi) = t_{n,m+1}^k U t_{n,m}^k(\xi)_{\lambda}^{\max(k-1,0)}
\]

by induction hypothesis. Therefore,

\[
\alpha_{n,m+1}^k(\xi) = t_{n,m+1}^k(\xi)_{\lambda}^{\max(k-1,0)}
\]

since \( t_{n,m}^k(\xi) \subseteq t_{n,m+1}^k(\xi) \). Therefore,

\[
\alpha_{n,m}^k(\xi) = t_{n,m}^k
\]

if \( m \geq 0 \) by induction. Therefore, the lemma is proved.

\[\cdots\]
Theorem 35:

For all \( n \geq 0 \) and \( n+1 \geq k \geq 0 \)

\[
T_n^k = \bigcup_{m \geq 0} T_{n,m}^k
\]

Proof:

To prove that

\[
T_n^k = \bigcup_{m \geq 0} T_{n,m}^k
\]

a similar result will be shown for \( t_n^k(\Sigma) \) and \( t_{n,m}^k(\Sigma) \). If \( t_n^k(\Sigma) \) is defined as in definition 28 and \( t_{n,m}^k(\Sigma) \) is defined as in definition 30, then it is sufficient to show that

\[
t_n^k(\Sigma) = \bigcup_{m \geq 0} t_{n,m}^k(\Sigma)
\]

since

\[
T_n^k = t_n^k(\Sigma) \lambda^{\max(k-1,0)}
\]

and

\[
U_{m \geq 0} T_{n,m}^k = \bigcup_{m \geq 0} U_{m \geq m' \geq 0} t_{n,m'}^k(\Sigma) \lambda^{\max(k-1,0)}
\]

Therefore, it will be shown that \( U_{m \geq 0} t_n^k(\Sigma) \) has all the properties that define \( t_n^k(\Sigma) \). Define

\[
a_{n,m}(\Sigma) = \bigcup_{m \geq m' \geq 0} t_{n,m'}^k(\Sigma)
\]

By lemma 34

\[
a_{n,m}(\Sigma) = t_{n,m}^k(\Sigma)
\]

due to equations 2 and 3 given above. Therefore, \( a_{n,m}(\Sigma) \) is the smallest set such that

1. If \( n \geq 0 \) and \( m \geq 0 \), then
2. If $n \geq 0$ and $m \geq 0$, then
\[
\alpha_{n,m}^{0}(\Sigma) = \alpha_{n,m}^{1}(\Sigma)
\]

3. If $n \geq k > 0$ and $m \geq 0$, then
\[
\alpha_{n,m}^{k}(\Sigma) = \alpha_{n,m}^{k+1}(\Sigma)_{\Sigma}^{\lambda} U \alpha_{n,m}^{k+1}(\Sigma)_{\Sigma}^{\lambda} t_{n,m}^{k-1}
\]
\[
= \alpha_{n,m}^{k+1}(\Sigma)_{\Sigma}^{\lambda}
\]
\[
U \alpha_{n,m}^{k+1}(\Sigma)_{\Sigma}^{\lambda} \alpha_{n,m}^{k-1}(\Sigma)_{\Sigma}^{\lambda} \max((k-1)-1,0)\]

By the definition of infinite union
\[
\lim_{m \to \infty} \alpha_{n,m}^{k}(\Sigma) = U_{m \geq 0} t_{n,m}^{k}(\Sigma)
\]

Therefore, if $\alpha_{n}^{k}(\Sigma)$ is defined as
\[
\alpha_{n}^{k}(\Sigma) = U_{m \geq 0} t_{n,m}^{k}(\Sigma)
\]

then it must be true that

1. If $n \geq 0$, then
\[
(4) \quad \alpha_{n}^{n+1}(\Sigma) = \Sigma
\]

2. If $n \geq 0$, then
\[
(5) \quad \alpha_{n}^{0}(\Sigma) = \alpha_{n}^{1}(\Sigma)
\]

3. If $n \geq k > 0$, then
Now, since $t_k^n(\Sigma)$ is the smallest set with properties 4, 5 and 6 it must be true that

$$t_k^n(\Sigma) = a_k^n(\Sigma)$$

and the theorem is proved.

---

A theorem can now be proved which will show that n-ary encoded hypertrees are, in fact, n-ary trees. After this theorem is proved, it will be shown that these definitions allow a quite natural encoding of level n hypertrees, and hyperforests as n-ary trees.

Lemma 36:

If $n \geq 0$, $n+1 \geq k \geq 0$ and $m \geq 0$, then

$$t_k^n, m(\Sigma) = \Sigma$$

Proof by induction on $m$:

Basis: if $m = 0$ then one of the following three cases is true:

1. If $k = n+1$, then

$$t_k^n(\Sigma) = \Sigma$$
\[
\begin{align*}
\tau_{n,m}^k &= \xi[\lambda^{\max((n+1)-1,0)}] \\
&= \xi[\lambda^n]
\end{align*}
\]

Since \( \lambda \in T'_n \) it must be also true that 
\[\xi[\lambda^n] \subset T'_n\]

by definition 27. Therefore, 
\[T_{n,m}^k \subset T'_n\]

2. If \( n \geq k > 0 \), then 
\[T_{n,0}^k = T_{n,0}^k \xi[\lambda^{\max(0-1,0)}] = \emptyset\]

Therefore, 
\[T_{n,0}^k \subset T'_n\]

3. If \( k = 0 \), then 
\[T_{n,m}^k = T_{n,m}^1\]

But by either part 1 or part 2 of this proof, 
\[T_{n,m}^1 \subset T'_n\]

Therefore, 
\[T_{n,m}^0 \subset T'_n\]
Induction step: Assume that if $m > m' \geq 0$ then

$$T_{n,m'}^k \subseteq T_n^k$$

If $m > 0$, then

$$T_{n,m}^k = t_{n,m}^k(\xi)\lambda^{\max(k-1,0)}$$

Therefore, one of the following cases must be true:

1. If $k = n+1$, then

$$T_{n+1,m}^n = \Sigma[\lambda^{\max((n+1)-1,0)}]$$

$$= \Sigma[\lambda^{\max(n,0)}]$$

$$= \Sigma[\lambda^n]$$

Notice that $\lambda \in T'_n$, therefore,

$$\Sigma[\lambda^n] \subseteq T'_n$$

and

$$T_{n+1,m}^n \subseteq T'_n$$

2. If $n \geq k > 0$, then

$$T_{n,m}^k = t_{n,m}^k(\xi)\lambda^{\max(k-1,0)}$$

$$= t_{n,m}^k(\xi)\lambda^{k-1}$$

by definition 30. Also by the definition of $t_{n,m}^k(\xi)$
\[ T_{n,m}^k = T_{n,m-1}^{k+1}(\xi)^{\lambda_{k-1}} U T_{n,m-1}^{k+1}(\xi)T_{n,m-1}^{k-1} \lambda_{k-1} \]

\[ = T_{n,m-1}^{k+1}(\xi)^{\lambda_{(k+1)-1}} U T_{n,m-1}^{k+1}(\xi)T_{n,m-1}^{k-1} \lambda_{k-1} \]

and by the definition of \( T_{n,m-1}^{k+1} \)

\[ (7) \quad T_{n,m}^k = T_{n,m-1}^{k+1} U T_{n,m-1}^{k+1}(\xi)T_{n,m-1}^{k-1} \lambda_{k-1} \]

Notice that since

\[ (8) \quad T_{n,m-1}^{k+1} \subseteq T_{n}^{k} \]

by induction hypothesis, and since

\[ \lambda \notin T_{n,m-1}^{k+1} \]

it must be that

\[ T_{n,m-1}^{k+1} \subseteq \Sigma[(T_{n}^{k})^n] \]

Also note that

\[ T_{n,m-1}^{k+1} = t_{n,m-1}^{k+1}(\xi)\lambda^k \]

since \((k+1)-1 \geq 0\). Therefore,

\[ (9) \quad t_{n,m-1}^{k+1}(\xi)\lambda^k \subseteq \Sigma[(T_{n}^{k})^n] \]
By noting that \( \lambda \in T'_n \) and matching subtrees it can be seen that

\[
t^k_{n,m-1}(\xi) \subseteq \Sigma [(T'_n)^{n-k}]
\]

Therefore,

\[
t^k_{n,m-1}(\xi)T^{k-1}_{n,m-1} \subseteq \Sigma [(T'_n)^{n-k+1}]
\]

by induction hypothesis and

\[
t^k_{n,m-1}(\xi)T^{k-1}_{n,m-1}\lambda^{k-1} \subseteq \Sigma [(T'_n)^{n-k+1}\lambda^{k-1}]
\]

Since \( \lambda \in T'_n \) it must be that

\[
t^k_{n,m-1}(\xi)T^{k-1}_{n,m-1}\lambda^{k-1} \subseteq \Sigma [(T'_n)^{n-k+1}(T'_n)^{k-1}]
\]

\[
\subseteq \Sigma [(T'_n)^n]
\]

and that

\[
(10) \quad t^{k+1}_{n,m-1}(\xi)T^{k-1}_{n,m-1}\lambda^{k-1} \subseteq T'_n
\]

by equations 7, 8 and 10 therefore,

\[
t^k_{n,m} \subseteq T'_n
\]

3. Finally, if \( n \geq k = 0 \) then

\[
T^{0}_{n,m} = t^0_{n,m}(\xi)^{\max(0-1,0)}
\]
Since
\[ T_{n,m}^1 \subset T'_n \]
as a special case of step 1 or step 2 of this proof
\[ T_{n,m}^0 \subset T'_n \]

Therefore, the lemma is true in this case also.

Since the lemma is true in each of the above cases, the lemma is proved by induction.

---

Lemma 37:
If \( n \geq 0 \) and \( n+1 \geq k \geq 0 \), then
\[ T_n^k \subset T'_n \]

Proof:
This follows immediately from lemma 36 and theorem 35.

---

Theorem 38:
If \( n \geq 0 \), then
\[ T_n \subset T'_n \]

Proof:
By lemma 37

\[ T_n^n < T'_n \]

Therefore,

\[ T_n^n < T'_n \]

by definition 29 and the theorem is proved.

Note that this theorem does not imply that

\[ T_n = T'_n \]

In fact, this is not true, as can be demonstrated by noting that

\[ a[b[\lambda^2]c[\lambda^2]] \]

is a legal n-ary tree, but not a legal n-ary encoded hypertree. This is true since this example is a 2-ary tree, corresponding to 2-ary encoded hypertrees, but 2-ary encoded hypertrees must all have the postfix "\(\lambda\)". This example does not.

The encoding that may be used will now be demonstrated. It will be shown recursively that \(H_{n,m}^k(\lambda)\) can be easily encoded in \(T_{n,m}^k\), and once this is shown, it will be assumed that this encoding can be extended to \(H_n^k(\lambda)\) and \(H_n(\lambda)\).
Definition 31:

Encode is a family of functions, denoted \( \text{encode}^k_{n,m} \), such that

\[
\text{encode}^k_{n,m}: h_{n,m}^k(X) \to t_{n,m}^k
\]

where

(a) \( \text{encode}^{n+1}_{n,m}(\gamma) = \gamma[\lambda^n] \) if \( n \geq 0 \) and \( m \geq 0 \).

(b) \( \text{encode}^k_{n,0} \) is a null function if \( n \geq k > 0 \).

(c) \( \text{encode}^k_{n,m+1}(\gamma) = \text{encode}^k_{n,m}(\gamma) \) if \( \gamma \in h_{n,m}^k \).

where \( n \geq k > 0, m \geq 0 \) and if \( \text{encode}^{k+1}_{n,m}(\alpha) = \nu[\lambda^k] \) then \( \nu = \nu_1 \text{encode}^{k-1}_{n,m}(\beta) \lambda^{k-1} \).

(d) \( \text{encode}^0_{n,m}(\gamma) = \text{encode}^1_{n,m}(\gamma) \) if \( n \geq 0 \) and \( m \geq 0 \).
Since this is the definition of a function, a proof of the codomain is in order.

Theorem 39:

If \( n \geq 0, \ n+1 \geq k \geq 0, \ m \geq 0 \) and \( \gamma \in h^k_{n,m} \), then

\[
\text{encode}^k_{n,m}(\gamma) \in T^k_{n,m}
\]

Proof by induction on \( m \):

Define \( t^k_{n,m}(\xi) \) as it is defined in definition 30.

Basis: If \( m = 0 \), then one of the following three cases is true:

1. If \( k = n+1 \), then

\[
\text{encode}^{n+1}_{n,m}(\gamma) = \gamma[\lambda^n]
\]

Now, \( \gamma \in \Sigma \) therefore,

\[
\text{encode}^{n+1}_{n,m}(\gamma) \in \Sigma[\lambda^n]
\]

\[
\epsilon t^{k+1}_{n,0^n}
\]

\[
\epsilon T^{n+1}_{n,m}
\]

2. If \( n \geq k > 0 \), then \( \gamma \) does not exist since

\[
\gamma \in h^k_{n,0}
\]

\[
\epsilon \emptyset
\]

3. If \( k = 0 \), then

\[
\text{encode}^k_{n,m}(\gamma) = \text{encode}^1_{n,m}(\gamma)
\]
Now, if \( n > 0 \) then
\[
y \in h^{1}_{n,0}
\]
\( \notin \phi \)
so that \( y \) does not exist and this case cannot happen.
If \( n = 0 \), then
\[
y \in h^{1}_{0,0}
\]
\( \notin \Sigma \)
and by case 1 above
\[
\text{encode}^{1}_{0,0}(y) \in T^{1}_{0,0}
\]
But \( T^{1}_{0,0} = T^{0}_{0,0} \) so that
\[
\text{encode}^{0}_{0,0}(y) \in T^{0}_{0,0}
\]
Induction step: Assume that if \( m > m' \geq 0 \) and \( y \in h^{k}_{n,m} \) then
\[
\text{encode}^{k}_{n,m}(y) \in T^{k}_{n,m'}
\]
If \( y \in h^{k}_{n,m} \), then one of the following cases is true:
1. If \( k = n+1 \), then
\[
y \in h^{n+1}_{n,m}
\]
\( \notin \Sigma \)
and
\[
\text{encode}^{k}_{n,m}(y) = \Sigma[\lambda^{n}]
\]
Note that \( t^{k}_{n,m}(\Sigma) = \Sigma[ \) so that this becomes
\[ \text{encode}_{n,m}(\gamma) \in T_{n,m}^{k}(\gamma) \]

Since \( \max((n+1)-1,0) = n \), it must be true that

\[ \text{encode}_{n,m}(\gamma) \in T_{n,m}^{k}(\lambda)^{\max((n+1)-1,0)} \]

\[ \in T_{n,m}^{k} \]

2. If \( n \geq k > 0 \) and \( \gamma \in h_{n,m}^{k+1} \) then

\[ \text{encode}_{n,m}(\gamma) = \text{encode}_{n,m-1}^{k+1}(\gamma) \]

But by induction hypothesis

\[ \text{encode}_{n,m-1}^{k+1}(\gamma) \in T_{n,m-1}^{k+1} \]

Therefore,

\[ \text{encode}_{n,m}(\gamma) \in T_{n,m}^{k+1} \]

\[ \in T_{n,m}^{k} \]

since \( T_{n,m-1}^{k+1} \subset T_{n,m}^{k} \)

3. If \( n \geq k > 0 \) and \( \gamma \in h_{n,m-1}^{k} \) then

\[ \text{encode}_{n,m}(\gamma) = \text{encode}_{n,m-1}^{k}(\gamma) \]

\[ \in T_{n,m-1}^{k} \]

by induction hypothesis. Since \( T_{n,m-1}^{k} \subset T_{n,m}^{k} \)

\[ \text{encode}_{n,m}(\gamma) \in T_{n,m}^{k} \]

4. If \( n \geq k > 0 \) and \( \gamma = \alpha[k \beta_k] \), then
where $V$ is defined in definition 30. Now,

$$encode_{n,m}^{k+1}(\alpha) \in T_{n,m}^{k+1}$$

by induction hypothesis. Therefore,

$$encode_{n,m-1}^{k+1}(\alpha) = v_{1} \lambda k$$

where $v_{1} \in T_{n,m-1}^{k+1}(\xi)$. Therefore,

$$encode_{n,m}^{k}(\alpha) = v_{1} encode_{n,m-1}^{k-1}(\beta) \lambda k-1$$

Since $encode_{n,m-1}^{k-1}(\beta) \in T_{n,m-1}^{k-1}$ by induction hypothesis,

$$encode_{n,m}^{k}(\alpha) \in T_{n,m}^{k+1}(\xi) T_{n,m-1}^{k-1}(\beta) \lambda k-1$$

$$\in T_{n,m}^{k}$$

Since if $k > 0$ then $\max(k-1,0) = k-1$.

5. If $k = 0$, then

$$encode_{n,m}^{0} = encode_{n,m}^{1}$$

and by a previous part of this proof

$$encode_{n,m}^{0}(\alpha) \in T_{n,m}^{1}$$

Since $T_{n,m}^{1} = T_{n,m}^{0}$ it must be true that

$$encode_{n,m}^{0}(\alpha) \in T_{n,m}^{0}$$
Therefore,

\[ \text{encode}_{n,m}^k (\bar{x}) \in \mathbb{T}_{n,m}^k \]

in all cases and the theorem is proved by induction.

---

**TABLE 3. Some Examples of n-ary Encoded Hypertrees**

<table>
<thead>
<tr>
<th>line level</th>
<th>hypertree</th>
<th>encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>a[ ]</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>a[λ]</td>
</tr>
</tbody>
</table>

line encodings with n-ary trees represented with pointers.

<table>
<thead>
<tr>
<th>line</th>
<th>encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a[ ]</td>
</tr>
<tr>
<td>2</td>
<td>a[λ]</td>
</tr>
<tr>
<td>3</td>
<td>a[λ] b[λ] c[λ]</td>
</tr>
<tr>
<td>4</td>
<td>a[λλ]</td>
</tr>
<tr>
<td>5</td>
<td>a[λ] a[λ] b[λ] c[λλ]</td>
</tr>
<tr>
<td>6</td>
<td>a[λλ] a[λ] a[λλ] c[λλλ]</td>
</tr>
</tbody>
</table>
Some examples of hypertrees and their n-ary representations are given in table 3. Note that, as had been expected earlier, line 3 corresponds to the linked list representation of the string "abc". A detailed look at this example will show that any encoded string will result in its linked list representation. Similarly, line 5 is the binary encoded tree for "a(abc)", and it should be obvious that any tree will be encoded as the corresponding binary tree.

While the n-ary tree encoding of hypertrees is the only encoding which will be discussed at length in this paper, it might be wise to note a few other encoding schemes. Of course, one other scheme has been used rather extensively in this paper. This is the "normal" encoding using brackets, subscripts and superscripts. It is hoped that this is the easiest encoding for the reader to understand. For example

(11) \[ a[2b[1c[1d]_1]_2] \]

Another significant encoding scheme is the recursive encoding. This encoding follows directly from the definition of hypertree given in definition 12 of chapter 2.
Definition 32:

Recursive is a family of functions, denoted $\text{recursive}_n^k$, such that

(a) $\text{recursive}_n^{n+1}(\gamma) = [1,\gamma]$ if $n \geq 0$ and $\gamma \in h_n^{n+1}(\xi)$.

(b) $\text{recursive}_n^0(\gamma) = [2,\text{recursive}_n^1(\gamma)]$ if $n \geq 0$ and $\gamma \in h_n^0(\xi)$.

(c) $\text{recursive}_n^k(\gamma) = [3,\text{recursive}_n^{k+1}(\gamma)]$ if $n \geq k > 0$ and $\gamma \in h_n^{k+1}(\xi)$.

$= [4,\text{recursive}_n^{k+1}(\alpha),\text{recursive}_n^{k-1}(\beta)]$ if $n \geq k > 0$ and $\gamma = \alpha[\varnothing_k^k \beta_k^k]$.

Notice that although $k$ is not mentioned, its value can always be inferred from the structure. $n$, on the other hand, must be stored somewhere independent of this encoding as its value cannot be inferred. Since each of the alternatives above have at most two possible recursions, this can be stored in a binary tree. It could be demonstrated that this encoding is directly related to the binary encoding of the n-ary tree encoding of hypertrees. The example given in equation 11 can be encoded as

$$2[4,[1,a],[4,[3,[1,b]],[2,[4,[3,[1,c], [2,[3,[1,d]]]]]]]]$$
At first glance, this encoding scheme seems worse than the n-ary tree encoding of hypertrees given earlier, but if the level of the hypertrees that are encoded could vary over a large range of values, this encoding could result in less storage being required, since every node is at most binary.

In this chapter, three different encodings for hypertrees are mentioned. In the next chapter, the concept of a hypertree grammar is discussed so that work can begin toward characterizing the algebraic language hierarchy, which was the original justification for the existence of hypertrees.
REGULAR HYPERTREE GRAMMARS

The original justification for hypertrees was that they could model the algebraic language hierarchy. This requires that at least regular grammars be defined on hypertrees at all levels. This chapter attempts to do this. Context free grammars are not defined as a separate entity, but rather that term is reserved for the frontier of a next higher level hypertree grammar. Similarly, macrogrammars are not explicitly defined.

In the case of macrogrammars, there are two different methods of deriving a language from a given grammar. These are referred to as the inside-out grammars and the outside-in grammar by Fischer. This leads to two different methods of deriving a language for a given hypertree grammar. This paper will concern itself with the inside-out (IO) approach due to its pleasing properties. The outside in (OI) approach will be defined but characterization will be left for further study.

Regular grammars are defined in a method that is analogous to the definition of regular grammars over strings.
Definition 33:
A regular grammar of level $n$ and degree $k$, denoted $G_n^k$, is a four-tuple $(\mathcal{L}, \mathcal{V}, P, S)$ where $\mathcal{L}$ is the set of terminals, $\mathcal{V}$ is a ranked set of nonterminals, $P$ is a finite set of productions and $S$ is the start symbol. In addition,
(a) $S \in \mathcal{V}_k$
(b) $P \subseteq \bigcup_{i \geq 0} (\mathcal{V}_i \rightarrow H_n^k(\mathcal{L}, \mathcal{V}))$

By using a ranked set of nonterminals it will be possible to generate all of the set $H_n^k(\mathcal{L})$. As is commonly done, the grammar $G_n^k$ will commonly be designated by specifying the elements of $P$. These will be written vertically with the first production specifying the start symbol. For example

$$G_2^2 = ((a), ((S,T),(S \rightarrow 2, T \rightarrow 1)),$$

$$\{ S \rightarrow a[2 T_2], T \rightarrow a[1 T_1], T \rightarrow a, S \}^2$$

will be written as

$$S \rightarrow a[2 T_2]$$
$$T \rightarrow a[1 T_1]$$
$$T \rightarrow a$$

An additional shorthand will be to combine productions with the same left hand side by separating the right hand sides with a "|". Thus, the above grammar may be written as
S → a[2 T 2]
T → a[1 T 1] | a

In addition to defining what is meant by regular grammar, it is also necessary to define the derivation relation so that eventually the language described may be defined.

Definition 34:

The 10 derivation relation over grammar \( G^k_n \), denoted \( G^k_n \Rightarrow \), is defined as

\[
G^k_n \Rightarrow : H^k_n(\Sigma, V) \rightarrow H^k_n(\Sigma, V)
\]

if \( n \geq k > 0 \) such that, if \( G^k_n = (\Sigma, V, P, S)^k_n \) and

\( \alpha V \beta \in H^k_n(\Sigma, V) \) where \( V \in V_j \) and \( V \gamma \beta \in P \) then

\( \alpha V \beta \in G^k_n \Rightarrow (\alpha V \beta) \)

As has been done previously, a proof will be offered to show that the codomain stated is correct. First a lemma needs to be established.

Lemma 40:

If \( n \geq 0, n+1 \geq k \geq 0, m \geq 0, \gamma \in H^k_{n,m}, \gamma = \alpha V \beta, \n V \in V_j \) and \( \psi \in H^i_{n}, \) then

\( \alpha \psi \beta \in H^k_n \)

Proof by induction on \( m \):
Basis: if \( m = 0 \) then

1. If \( k = n+1 \), then

\[ \gamma \in \mathcal{I} \]

so that \( \gamma \) cannot equal \( \alpha \nu \beta \), so this cannot be the case.

2. If \( n \geq k > 0 \), then

\[ \gamma \in V_k \cup X_k \]

But since \( \gamma = \alpha \nu \beta \), it must be that \( \gamma \in V_k \), \( \alpha = \lambda \) and \( \beta = \lambda \). Also, \( k = i \) since \( \nu \in V_k \), \( \nu \in V_i \) and the ranks must be disjoint. Therefore,

\[ \alpha \psi \beta = \psi \]

But \( \psi \in H_{n}^{k} \) so that

\[ \alpha \psi \beta \in H_{n}^{k} \]

3. If \( k = 0 \), then

\[ \gamma \in H_{n,0}^{1} \]

Therefore, by part 2 of this proof,

\[ \alpha \psi \beta \in H_{n}^{k} \]

if \( \gamma = \alpha \nu \beta \).

Therefore, the basis is proven.
Induction step: Assume that if $m > m' \geq 0$ and $\gamma = \alpha \psi \beta$ then

$$\alpha \psi \beta \in H_n^k$$

If $\gamma \in H_n^k$ and $\gamma = \alpha \psi \beta$, then

1. If $k = n+1$, then

$$\gamma \in \Sigma$$

so that this cannot be the case.

2. If $n \geq k > 0$ and $\gamma \in H_{n,m-1}^{k+1}$, then

$$\alpha \psi \beta \in H_n^{k+1}$$

by induction hypothesis. But $H_n^{k+1} \subset H_n^k$ so

$$\alpha \psi \beta \in H_n^k$$

3. If $n \geq k > 0$ and $\gamma \in H_{n,m-1}^k$, then

$$\alpha \psi \beta \in H_n^k$$

by induction hypothesis.

4. If $n \geq k > 0$, $\gamma = \gamma_1[k \gamma_2 k]$ and $\gamma_1 = \alpha_1 \psi_1 \beta_1$, then

$$\alpha_1 \psi_1 \beta_1 \in H_n^{k+1}$$

by induction hypothesis. Also, $\alpha = \alpha_1$ and

$$\beta = \beta_1[k \gamma_2 k]$$

so that

$$\alpha \psi \beta = \alpha_1 \psi_1 \beta_1[k \gamma_2 k]$$

$$\in H_n^{k+1}[k \beta_1^{-1} H_n^{k-1} k]$$
5. If $n \geq k > 0$, \( \gamma = \gamma_1[k \gamma_2 k] \) and \( \gamma_2 = \alpha_2 \psi \beta_2 \), then

\[ \alpha_2 \psi \beta_2 \in H_{n}^{k-1} \]

by induction hypothesis. Also, \( \alpha = \gamma_1[k \alpha_2 \gamma_2 k] \) so that

\[ \alpha \psi \beta = \gamma_1[k \alpha_2 \psi \beta_2 k] \]

\[ \in H_{n}^{k+1}[k H_{n}^{k-1} k] \]

\[ \in H_{n}^{k} \]

6. If $k = 0$, then

\[ \gamma \in H_{n,m}^{1} \]

and by the previous parts of this induction step

\[ \alpha \psi \beta \in H_{n}^{1} \]

But since $H_{n}^{1} = H_{n}^{0}$,

\[ \alpha \psi \beta \in H_{n}^{k} \]

and the lemma is proved by induction.
Theorem 41:

If \( n \geq k > 0, \ \gamma_1 \in H_n^k \) and \( \gamma_1 \Rightarrow \gamma_2 \), then

\[ \gamma_2 \in H_n^k \]

Proof:

If \( \gamma_1 \Rightarrow \gamma_2 \), then

\[ \gamma_1 = \alpha \psi \beta \]

and

\[ \gamma_2 = \alpha \psi \beta \]

for some \( \alpha, \beta \) and \( \psi \), where \( \nu \rightarrow \psi \in P \). But, if \( \nu \rightarrow \psi \in P \) then

\[ \nu \in V_i \]

and

\[ \psi \in H_i \]

for some \( i \). Therefore,

\[ \gamma_2 \in H_n^k \]

by lemma 40 and the theorem is proved.

---

As is commonly done in the literature on languages,

\( \beta \in G_n^k(\alpha) \) will be written in infix notation as \( \alpha G_n^k \Rightarrow \beta \).

Also, if the grammar \( G_n^k \) is obvious, this will be elided so that \( \alpha G_n^k \Rightarrow \beta \) will be written as \( \alpha \Rightarrow \beta \). The transitive and reflexive closure of \( G_n^k \Rightarrow \) will be written as \( G_n^{k*} \Rightarrow \), as
defined in the following definition. Also, the $G^k_n$ will be elided whenever the grammar is obvious in this case also.

Definition 35:

The IO derivation relation closure over grammar $G^k_n$, denoted $G^k_n*$, is defined as follows:

$$G^k_n*: H^k_n(\Sigma, V) \rightarrow H^k_n(\Sigma, V)$$

such that if $\alpha \in H^k_n(\Sigma, V)$ then

$$\alpha G^k_n* \alpha$$

Also, if $\alpha \in H^k_n(\Sigma, V)$ and $\alpha G^k_n* \beta$ and $\beta G^k_n* \gamma$ then

$$\alpha G^k_n* \gamma$$

As stated above, the $G^k_n$ will be elided, and the *$\Rightarrow$ will be written in infix notation if no confusion would result.

Theorem 42:

If $n \geq k > 0$, $\gamma_1 \in H^k_n$ and $\gamma_1 G^k_n* \gamma_2$, then

$$\gamma_2 \in H^k_n$$

Proof:

This follows directly from theorem 41 by induction.

---

The IO language of $G^k_n$ may now be defined
Definition 35:

The language defined by grammar $G^k_n = (\Sigma, V, P, S)^k_n$, denoted $language(G^k_n)$, is defined as

$language(G^k_n) = \{ \gamma \mid S \xrightarrow{G^k_n} \gamma \text{ and } \gamma \in H^k_n(\Sigma) \}$

For example, given the hypertree grammar above, then

\[
S \xrightarrow{*} a[2 \ T \ 2]
\]
\[
\xrightarrow{*} a[2a[1 \ T \ 1]_2]
\]
\[
\xrightarrow{*} a[2a[1a[1]_2]
\]

Also, since the second production could have been substituted for the T instead of the third production in step 2 above

\[
S \xrightarrow{*} a[2a[1 \ T \ 1]_2]
\]
\[
\xrightarrow{*} a[2a[1a[1 \ T \ 1]_1]_2]
\]
\[
\xrightarrow{*} a[2a[1a[1a[1]_1]_2]
\]

From this short example it should be evident that this grammar will generate all trees of height one over the one element alphabet \{a\}.

The above example illustrates that hypertree grammars over trees are more powerful than the classical approach defined by Brainard (Brainard 1967). In the classical approach, it would be impossible to write a grammar giving all trees of height one since the root node could only be in
a finite number of the ranks of the ranked alphabet that the language is over.

As was stated earlier, there is another method of deriving a language for a grammar $G^k_n$ called the OI language. It will be defined here, but analysis of this approach will be left for further research.

Definition 37:

The OI derivation relation over grammar $G^k_n$ denoted $G^k_n \Rightarrow^*$, is defined as follows:

If $n > k > 0$.

$$G^k_n \Rightarrow^* : H^k_{n-1}(\Sigma, V) \rightarrow H^k_{n-1}(\Sigma, V)$$

if $n > 0$. If $a$ is an element of the domain and $a \Rightarrow B$, then

$$G^k_n \Rightarrow : H^k_{n-1}(\Sigma, V) \rightarrow H^k_{n-1}(\Sigma, V)$$

As in the IO case, $\beta \in G^k_n \Rightarrow^*(a)$ may be written as $a \Rightarrow^* \beta$. Also, the transitive closure may be written as $a \Rightarrow^* \beta$.

Notice that this actually defines a language $\text{language}'(G^k_n)$ which is in $H^k_{n-1}(\Sigma)$ rather than $H^k_n(\Sigma)$ as is the case with the IO derivation. For proper comparison, therefore, the frontier language of $\text{language}(G^k_n)$ should be
compared with language'($G_n^k$). Call this language, the
frontier language of language($G_n^k$), Frlanguage($G_n^k$). The
basic difference between Frlanguage($G_n^k$) and language'($G_n^k$) is
that Frlanguage($G_n^k$) first builds up a large level n
hyperforest then takes the frontier, whereas language'($G_n^k$)
takes the frontier after each derivation. The former
approach guarantees that each given nonterminal always
develops into the same hypertree, even if more than one copy
of the result of the derivation of that nonterminal is made
in the resulting sentence. The latter approach, on the
other hand, allows each copy of the nonterminal to develop
independently of the other copies so that different copies
can have different values. As expected, this is the same as
the difference that exists between IO macro languages and OI
macro languages.

If the example given above is expanded using the OI
derivation method, then an example of the resulting language
is

\[
S \Rightarrow 'T \\
\Rightarrow 'a[1 \ T \ 1] \\
\Rightarrow 'a[a[1 \ T \ 1]_1] \\
\Rightarrow 'a[a[a[1 \ a_1]_1]_1]
\]

Notice that since this grammar is of level 2 the derivation
yields a string which is equal to the frontier of one of the
trees generated by the IO derivation method. The grammar generated all trees of height one using the IO derivation method, however, it should be obvious that the set of all strings over the alphabet \{a\} is generated by the OI derivations.

The languages of ultimate interest are string languages, therefore, a new function is defined for extracting the string language associated with a given hypertree language.

Definition 38:

Extract is a family of functions, denoted extract\_n, such that

\[
\text{extract}_n : U_{n' \geq n} H_{n'}^0 \rightarrow H_n^0
\]

if \( n \geq 0 \) where

(a) \[ \text{extract}_n(y) = y \]
if \( y \in H_n^0 \).

(b) \[ \text{extract}_n(y) = \alpha \]
if \( y \in H_n^0, \beta = \text{frontier}_{n'}^0(y) \) and
\[ \text{extract}_{n-1}(\beta) = \alpha. \]

This function uses the frontier function to repeatedly extract the frontier of a hypertree until the hypertree level is reduced to the desired level. Since the only requirement for the degree of the domain is that it be zero,
any hyperforest of any higher level can be operated on since $H_n^k \subseteq H_n^0$ for any value of $k$. (See lemma 8 chapter 2.)

As mentioned above, the special case of the extract function which returns strings is important enough to be given a special name.

Definition 39:

String is another name for the function $\text{extract}_1$.

Also, since the string language associated with a given grammar is important, the following definition is made.

Definition 40:

The IO string language defined by grammar $G_n^k$, denoted $\text{string}(G_n^k)$, is defined as

$$\text{string}(G_n^k) = \{ \alpha \mid \alpha = \text{string}(\beta) \text{ and } \beta \in \text{language}(G_n^k) \}$$

Again using the example above, it should be evident that the IO string language defined by the grammar is the set of all strings over \{a\}.

In order to simplify the proofs involving hypertree grammars, a special normal form of the grammar has been developed.
Definition 41:
A grammar \( G_n^k = (\Sigma, V, P, S)^k_n \) is said to be in \( m \)-normal form if
\[
P \subseteq \bigcup_{n \geq k > 0} (V_k \rightarrow H_{n,m}^k)
\]

If a grammar is simply said to be in normal form, then it is in 1-normal form.

Given the grammar above
\[
S \rightarrow a[2T_2]
\]
\[
T \rightarrow a[1T_1] | a
\]
notice that this is a 2-normal grammar. \( a[2T_2] \in H_{2,2}^2 \)
since \( a \in \Sigma = H_{2,1}^3 \) and \( T \in H_{2,1}^1 \). The latter is true since
\( T \in H_{2,0}^1 \). Similarly, \( a[1T_1] \in H_{2,2}^1 \) and \( a \in H_{2,2}^1 \).

Therefore, the grammar is 2-normal. Notice that the grammar, by definition, is also 3-normal since, for instance, \( a \in H_{2,3}^1 \) in addition to \( H_{2,2}^1 \). By similar reasoning, the grammar is 125-normal. In general, any grammar which is \( m \)-normal is also \( m' \)-normal where \( m' \geq m \).

It will be shown that to an extent the opposite of this is also true, the difference being that some changes to the grammar may be needed to decrease the normalcy \( m \). But first, a theorem is presented which demonstrates that any grammar in \( m \)-normal for some \( m \).
Theorem 43:

For every grammar $G_n^k = (\Sigma, V, P, S)^n_k$ there is an $m$ such that $G_n^k$ is $m$-normal.

Proof:

By definition 33,

$$P \subseteq \bigcup_{n \geq k > 0} (V_k \rightarrow H_n^k)$$

and $P$ is finite. If $v \rightarrow a \in P$, then $a$ must be an element of some $H_{n,m}^k$ since

$$H_{n,m}^k = \bigcup_{n \geq 0} H_{n,m}^k$$

Since $P$ is finite, each element of $P$ may determine a value in this fashion. Let the largest of these numbers be the desired $m$.

Lemma 44:

If $G_n^k = (\Sigma, V, P, S)^n_k$ is $m$-normal and $m > 1$, then there is a grammar $G_n^{k'} = (\Sigma', V', P', S')^n_k$ which is $m-1$-normal such that

$$\text{language}(G_n^k) = \text{language}(G_n^{k'})$$

Proof:

If grammar $G_n^k$ is also $m-1$-normal, then the original grammar is the desired grammar. If $G_n^k$ is not $m-1$-normal, then there
is a finite number of productions, $v_1 \rightarrow \gamma_1$, $v_2 \rightarrow \gamma_2$, ..., $v_j \rightarrow \gamma_j$ such that

$$\gamma_r \in \mathbb{H}_{n,m}^i$$

for some $i$ and

$$\gamma_r \notin \mathbb{H}_{n,m-1}^i$$

Therefore, construct $G_n^{k'} = (\mathcal{E}', V', P', S')_n^k$ as follows:

1. $\mathcal{E}' = \mathcal{E}$.
2. $V_i'$ consists of $V_i$ and elements required for the construction of $P'$.
3. The construction for $P'$ is given below.
4. $S' = S$.

$P'$ is constructed as follows:

1. $P' - \{v_1 \rightarrow \gamma_1 \ldots v_j \rightarrow \gamma_j\} \subset P'$.
2. If $v_r \rightarrow \gamma_r$ is one of the productions given above, $v_r \in V_i$ and $\gamma_r \in \mathbb{H}_{n,m-1}^{i+1}$ where $i < n$ then

$$\{v_r \rightarrow X_r, X_r \rightarrow \gamma_r\} \in P'$$

where $X_r \in V_{i+1}'$. Note that it is not possible for $v_r \rightarrow \gamma_r$ to be one of the productions given above where $\gamma_r \in \mathbb{H}_{n,m-1}^{i+1}$ and $i = n$ since in this case $\gamma_r \in \mathcal{E}$ and, therefore, is in $\mathbb{H}_{n,0}^{i+1}$.
3. If $v \rightarrow \gamma_r$ is one of the equations given above, $\gamma_r = a[i \beta i]$ and $n > i > 0$ then

$$\{v_r \rightarrow Y_{r[i}, Z_{r[i]}, Y_r \rightarrow a_r, Z_r \rightarrow \beta r\} \subset P'$$
where \( Y_r \in V_{i+1}' \) and \( Z_r \in V_{i-1}' \).

4. If \( v_r \rightarrow \gamma_r \) is one of the productions given and
   \( \gamma = a[n \beta n] \), then
   \[
   \{v_r \rightarrow a[n Z_r i], Z_r \rightarrow \beta\} \subset P'
   \]
   where \( Z_r \in V_{n-1}' \).

5. If \( v_r \rightarrow \gamma_r \) is one of the productions given above
   and \( \gamma = a[l \beta l] \), then
   \[
   \{v_r \rightarrow Y_r[l Z_r l], Y_r \rightarrow a_r, Z_r \rightarrow \beta_r\} \subset P'
   \]
   where \( Y_r \in V_{2}' \) and \( Z_r \in V_{1}' \).

6. These are all of the elements of \( P' \) and \( V' \).

\( k \) is \( m-1 \)-normal since each of the elements of
   \[
   P-(v_1 \rightarrow \gamma_1 \ldots v_j \rightarrow \gamma_j)
   \]
   is in \( U_{n \geq k \geq 0} V_k \rightarrow H_{n,m-1}^k \) by definition of the set
   \( \{v_1 \rightarrow \gamma_1 \ldots v_j \rightarrow \gamma_j\} \). Also, the production \( v_r \rightarrow X_r \) above must
   be in the required set since \( X_r \in V_{k'+1}' \) which is in
   \[
   H_{n,0}^{k'+1} \subset H_{n,1}^k \subset H_{n,m}^{k'}
   \]
   Similarly, \( X_r \rightarrow \gamma_r \) is in the set since \( \gamma_r \in H_{n,m-1}^{k'+1} \) by
   definition. Similar considerations will show that
   \( v_r \rightarrow Y_r[l_k Z_r k'] \), \( Y_r \rightarrow \gamma_r \) and \( Z_r \rightarrow \beta_r \) are all in the required
   set.
If $\gamma \in \text{language}(G^k_n)$, then

$$S \Rightarrow \gamma$$

Form a sequence $\gamma_0, \gamma_1, \ldots, \gamma_r$ such that $\gamma_0 = S$, $\gamma_r = \gamma$ and for all $r \geq i \geq 0$

$$\gamma_i \Rightarrow \gamma_{i+1}$$

By theorem 42

$$\gamma_i \in H^k_n$$

To prove that for all $i$

$$S \Rightarrow \gamma_i$$

an induction proof will be presented on $i$

Basis: if $i = 0$ then

$$S \Rightarrow \gamma_0$$

and since $\gamma_0 = S$

$$S \Rightarrow \gamma_0$$

Induction step: Assume that if $i < r$ then

$$S \Rightarrow \gamma_i$$

Then,

$$\gamma_i \Rightarrow \gamma_{i+1}$$

Therefore, there is a $v_i \in V_j$ such that
\[ \gamma_i = \alpha_i v_i \beta_i \]

and

\[ \gamma_{i+1} = \alpha_i \delta_i \beta_i \]

where

\[ v_i \rightarrow \delta_i \in P \]

Therefore, one of the following cases must be true:

1. If \( \delta_i \in H_{n,m-1}^j \), then

\[ v_i \rightarrow \delta_i \in P' \]

since these productions were copied from \( P \).

Therefore,

\[ \gamma_i \Rightarrow' \gamma_{i+1} \]

and by definition

\[ S \Rightarrow' \gamma_{i+1} \]

2. If \( \delta_i \in H_{n,m-1}^{j+1} \) and \( \delta \notin H_{n,m-1}^j \), then

\[ v_i \rightarrow X_i \in P' \]

where \( X_i \) occurs only in two specified productions in \( P' \). Therefore,

\[ \gamma_i = \alpha_i v_i \beta_i \]

\[ \Rightarrow' \alpha_i X_i \beta_i \]

also, \( X_i \rightarrow \delta_i \in P' \) so that
Therefore, since $S \Rightarrow^* \gamma_i$

$S \Rightarrow^* \gamma_{i+1}$

3. If $\delta_i' = \delta_{1,k}^i$ and $\delta \notin H_{n,m-1}^j$, then there are $Y_i$ and $Z_i$ such that

$$v_i \Rightarrow Y_i[k Z_i k] \in P'$$

Therefore,

$$\gamma_i = a_i v_i \beta_i \Rightarrow^* a_i Y_i[k Z_i k] \beta_i$$

Also, the productions to expand the $Y_i$ and $Z_i$ are included in $P'$. They are

$$Y_i \Rightarrow \delta_1 \in P'$$
$$Z_i \Rightarrow \delta_2 \in P'$$

Therefore,

$$a_i Y_i[k Z_i k] \beta_i \Rightarrow^* a_i \delta_1[k \delta_2 k] \beta_i \Rightarrow^* a_i \delta_i \beta_i \Rightarrow^* \gamma_{i+1}$$

Therefore, by the definition of $\Rightarrow^*$

$$S \Rightarrow^* \gamma_{i+1}$$
4. If $\delta = \delta_1[n \delta_2 n]$ and $\delta_i \notin H_{n,m-1}^n$, then there is a production

$$V_i \rightarrow \delta[n Z_i n]$$

therefore,

$$\gamma_i = a_i V_i \beta_i$$

$$\Rightarrow a_i \delta_1[n Z_i n] \beta_i$$

Also,

$$Z_i \rightarrow \delta_2 \in P'$$

so that

$$\gamma_i \Rightarrow^* a_i \delta_1[n \delta_2 n] \beta_i$$

$$\Rightarrow^* a_i \delta \beta_i$$

$$\Rightarrow^* \gamma_{i+1}$$

and by induction hypothesis

$$S \Rightarrow^* \gamma_{i+1}$$

Therefore, by induction

$$S \Rightarrow^* \gamma_i$$

for all $i$. In particular,

$$S \Rightarrow^* \gamma_r$$

Therefore, by the definition of $\text{language}(G_n^k)$

$$\text{language}(G_n^k) \subseteq \text{language}(G_n^{k'})$$
If \( \gamma \in \text{language}(G^k) \), then there is a sequence
\[
\gamma_0', \gamma_1', \ldots, \gamma_r'
\]
such that
\[
\gamma_0 = S \\
\gamma_r = \gamma
\]
and for all \( r > i \geq 0 \)
\[
\gamma_i \Rightarrow \gamma_{i+1}
\]
Construct a new sequence
\[
\gamma_0', \gamma_1', \ldots, \gamma_r'
\]
such that \( \gamma_i' = \gamma_i \) with all of the new nonterminals introduced in step 2 of the construction of \( G^k \) replaced by the right-hand-sides of the corresponding productions.
To prove: For all \( r \geq i \geq 0 \)
\[
S \Rightarrow \gamma_i'
\]
This will be proved by induction on \( i \).
Basis: if \( i = 0 \) then
\[
\gamma_0' = \gamma_0 \\
= S
\]
and by definition of \( \Rightarrow \)
\[
S \Rightarrow S
\]
Therefore,

\[ S \Rightarrow \gamma_0 \]

Induction step: Assume that if \( i < r \) then

\[ S \Rightarrow \gamma_i' \]

And there are two possible cases for the relationship between \( \gamma_i' \) and \( \gamma_{i+1}' \)

1. If \( \gamma_i' = \gamma_{i+1}' \), then

\[ S \Rightarrow \gamma_{i+1}' \]

by induction hypothesis.

2. If \( \gamma_i' = \alpha_i \nu_i \beta_i' \), \( \gamma_{i+1}' = \alpha_i \delta_i \beta_i \) and \( \nu_i \rightarrow \delta_i \in \mathcal{P} \), then

\[ \gamma_i' = \alpha_i \nu_i \beta_i \]

\[ \Rightarrow \alpha_i \delta_i \beta_i \]

\[ \Rightarrow \gamma_{i+1} \]

and by the definition of \( \Rightarrow \)

\[ S \Rightarrow \gamma_{i+1}' \]

These are the only possible cases because suppose

\[ \gamma_i = \alpha_i \nu_i \beta_i \]

and

\[ \gamma_{i+1} = \alpha_i \delta_i \beta_i \]
where

\[ v_i \rightarrow \delta_i \in P' \]

and

\[ v_i \rightarrow \delta_i \notin P \]

then either \( v_i \) is one of the new nonterminals added due to the construction, in which case

\[ \gamma_i' = \gamma_{i+1}' \]

by construction or \( \delta_i \) contains one or more of the new nonterminals. In this latter case, if \( \delta_i = \delta \) with the new nonterminals replaced by the corresponding right hand sides, then

\[ v_i \rightarrow \delta_i' \in P \]

by construction. Therefore,

\[ \gamma_i' = a_i'v_i\beta_i' \]

\[ * \Rightarrow a_i'\delta_i'\beta_i' \]

\[ * \Rightarrow \gamma_{i+1}' \]

where \( a_i' \) is \( a_i \) with all the occurrences of the new nonterminals replaced by the corresponding right hand sides, etc. Therefore, for all \( i \)

\[ S * \Rightarrow \gamma_i \]

In particular,
Therefore, since $\gamma_r = \gamma$, 

$$S \Rightarrow^* \gamma$$

and it must be that 

$$\gamma \in \text{language}(G_n^k)$$

Therefore, 

$$\text{language}(G_n^k) = \text{language}(G_n^{k'})$$

---

The example that is being dealt with above is 2-normal. By this theorem, there should be a 1-normal grammar which has the same language. The productions given are

(a) $S \rightarrow a[2 T_2]$  
(b) $T \rightarrow a[1 T_1]$  
(c) $T \rightarrow a$

with $S \in V_2$ and $T \in V_1$. In this example, production a is also already in 1-normal form and is, therefore, carried into the new set of productions unchanged. Production b will generate three new productions,

$$T \rightarrow X[1 Y_1]$$  
$$X \rightarrow a$$  
$$Y \rightarrow T$$
each of which is 1-normal. Similarly, production c generates two new productions.

\[ T \rightarrow Z \]
\[ Z \rightarrow a \]

Therefore, a new grammar which is 1-normal and has the same language as the 2-normal grammar presented is

\[ S \rightarrow a[T_2] \]
\[ T \rightarrow X[Y_1] \]
\[ X \rightarrow a \]
\[ Y \rightarrow T \]
\[ T \rightarrow Z \]
\[ Z \rightarrow a \]

where the ranked nonterminals are

\[ V_2 = \{S,X,Z\} \]
\[ V_1 = \{T,Y\} \]

Since any grammar has a corresponding grammar at the next lower level, and that grammar has another grammar at the level below that, and so forth until the grammar has a 1-normal grammar that has the same language, it seems reasonable that any grammar has to have a normal grammar with the same language. It could not be continued to a
0-normal grammar since lemma 44 requires that \( m \) be greater than one in order to reduce the level. This fact is stated in the following theorem

Theorem 45:

For every grammar \( G_n^k \) there is a corresponding normal grammar \( G_n^{k'} \).

Proof:

This follows directly from theorem 43 and lemma 44.

---

In this chapter, grammars have been discussed, and a normal grammar has been presented. In the following chapter, it will be shown that not only is the generative approach of grammars available for hypertrees, but the recognizing approach of automata is also available.
An alternative method of describing regular languages that is used in the literature is called finite automata. This is a method of describing a function which takes as an argument an element of $H_n^k$. This function will, then, return a value of a state. If this state is one of a predefined set of states called the final states, then the original element is said to be in the language described by the finite automaton. In place of the nonterminals from the previous chapter, an automaton has states, and in place of productions, an automaton has a finite transition function.

Definition 42:

A finite automata, denoted $A_n^k$, is a four-tuple $(\Sigma, V, \delta, F)_n^k$ where $n \geq k > 0$ and

1. $\Sigma$ is a set of terminal symbols.
2. $V$ is a ranked set of states.
3. $\delta$ is a set of functions called the transition functions

$$\delta_n^i: H_n^i, l \rightarrow P(V_i)$$

where $n \geq i > 0$ and $\delta_n^i(\alpha) = \phi$ except for a finite number of $\alpha$'s.

4. $F$ is a finite subset of $V_k$ called the set of final states.
The similarity between the definition of finite automata and regular grammars should be noted. As in the case of grammars, a finite automaton can be specified by giving only the transition functions and the final states. A result that will be assumed but not proven is that any automaton has a corresponding automaton -- one describing the same language -- which has finite sets for terminals and states.

An example of a finite automaton is as follows:

\[
\begin{align*}
\delta_2^2(a & [T]) = \{S\} \\
\delta_2^1(X & [Y]) = \{Y,T\} \\
\delta_2^2(a) &= \{X,Z\} \\
\delta_2^1(Z) &= \{Y,T\}
\end{align*}
\]

final states = \{S\}

It will be possible to show that this finite automaton describes the same language as the grammar given in the last chapter.

As in the case of grammars in the last chapter, a function needs to be defined which will allow the application of this automaton to the selection of the elements of a language.
Definition 43:

Given a finite automaton $A_n^k = (\Sigma, \Gamma, \delta, F)^k_n$ the complete transition function of degree $i$, denoted $A_n^{k,i}$, is defined such that

$$A_n^{k,i}: H_n^i(\Sigma, \Gamma) \rightarrow P(\Gamma) \text{ if } i = 0.$$

where $A_n^{k,i}(\emptyset)$ is the smallest set such that

(a) $A_n^{k,n+1}(\emptyset) = \{\emptyset\}$

(b) $A_n^{k,0}(\emptyset) = A_n^{k,1}(\emptyset)$

(c) $A_n^{k,i}(\emptyset) = \delta_n^i(\emptyset)$

if $\emptyset \in X_1$.

$$A_n^{k,i}(\emptyset) = U_n^i(A_n^{k,i+1}(\emptyset))$$

if $\emptyset \in H_n^{i+1}$. 

$$A_n^{k,i}(\emptyset) = U_n^i(A_n^{k,i+1}(a)[A_n^{k,i-1}(b) \cup \emptyset])$$

if $\emptyset = a[\cup \emptyset]$. 

where $n \geq k > 0$ and $n \geq i > 0$.

Definition 44:

The language identified by finite automaton $A_n^k$, denoted $\text{language}(A_n^k)$, is defined as

$$\text{language}(A_n^k) = \{\emptyset | \emptyset \in H_n^k(\Sigma) \text{ and } F_n A_n^{k,\emptyset}(\emptyset) \neq \emptyset\}$$
An example of the complete transition function for the automaton given earlier is

\[ A_{2,2}^2(a_{[1a[1a_1]_1)_2}) = \mathcal{U}_{2}^2(A_{2,3}^2(a)[_2^2A_{2,1}^2(a[1a[1a_1]_1))_2]) \]

But

\[ A_{2,2}^2(a) = a \]

and

\[ A_{2,1}^1(a_{[1a[1a_1]_1}) = \mathcal{U}_{2}^1(A_{2,2}^2(a)[_2^2A_{2,0}^2(a[1a_1]_1))_1) \]

and

\[ A_{2,2}^2(a) = \mathcal{U}_{2}^2(A_{2,3}^2(a)) \]
\[ = \mathcal{U}_{2}^2((a)) \]
\[ = \mathcal{U}(\{X,Z\}) \]
\[ = \{X,Z\} \]

also

\[ A_{2,0}^0(a_{[1a_1]} = A_{2,1}^1(a_{[1a_1]}) \]
\[ A_{2,1}^1(a_{[1a_1]} = \mathcal{U}_{2}^1(A_{2,2}^2(a)[_2^2A_{2,0}^2(a[1a_1]_1))) \]

and

\[ A_{2,0}^0(a) = A_{2,1}^2(a) \]
\[ = \mathcal{U}_{2}^1(A_{2,2}^2(a)) \]
\[ = \mathcal{U}_{2}^1(\{X,Z\}) \]
\[ = \{Y,T\} \]

Therefore,
\[ A_2^{0,1}(a[_{1a1}_1]) = U^1_2(({X,Z}_1}{Y,T}_1) \]
\[ = \{Y,T\} \]

and

\[ A_2^{0,0}(a[_{1a1}_1]) = \{Y,T\} \]

\[ A_2^{2,1}(a[_{1a1}_1]) = U^1_2(({X,Z}_1}{Y,T}_1) \]
\[ = \{Y,T\} \]

and

\[ A_2^{2,2}(a[_{2a[_{1a1}_1]_1)_2}] = U^2_2(a[_{2}\{Y,T\}_2]) \]
\[ = \{S\} \]

Since S is a final state, \( a[_{2a[_{1a1}_1]_1)_2} \) is in the language specified by the finite automaton. It was stated above that this automaton specified the same language as the grammar given in the last chapter, therefore, this result is expected.

It might also be instructive to consider what happens when the complete transition function is applied to an invalid hypertree. For example, the hypertree \( a[_{2a[_{2a2}_2]}_2] \) is of depth 2 and is, therefore, not expected to be in the language.

\[ A_2^{2,2}(a[_{2a[_{2a2}_2]}_2]) = U^2_2(A_2^{2,3}(a)[_{2A_2^{2,1}(a[_{2a2}_2])}_2]) \]
\[ = U^2_2(a[_{2U^1_2(a[_{2a2}_2])}_2]) \]
\[ = U^2_2(a[_{2U^1_2(U^2_2(\delta^2_2(A_2^{2,3}(a)[_{2A_2^{2,1}(a)_2}])_2))}])_2) \]
Notice that here there are no states corresponding to the input tree, therefore, there are no final states.

Another concept that is helpful in using automatons is that of deterministic finite automata. In a deterministic finite automaton, each element of the domain of the transition function has only one state to go to. This is equivalent to saying that the codomain of $\delta^k_n$ is further restricted so that each of the sets in the codomain have at most one element of $V_k$ in it. That is, the codomain is $V_k$ instead of $P(V_k)$.

Definition 45:
A finite automaton $A^k_n = (\Sigma, V, \delta, F)^k_n$ is deterministic if for all $n \geq k > 0$

$$\delta^k_n: B^k_n \rightarrow V_k$$

An example of a deterministic finite automaton is as follows:
\[ \delta_2^2(a[2T_2]) = S \]
\[ \delta_2^1(X[1T_1]) = T \]
\[ \delta_2^2(a) = X \]
\[ \delta_2^1(X) = T \]

final state = \( S \)

It should be evident that this automaton recognizes the same language as the previous nondeterministic automaton. In general, this is the case.

Theorem 46:

Given a finite automaton \( A_n^k = (\Sigma, V, \delta, F) \) there exists a deterministic finite automaton \( A_n^{k'} = (\Sigma', V', \delta', F') \) such that

\[ \text{language}(A_n^k) = \text{language}(A_n^{k'}) \]

Proof:

Construct the new finite automaton such that

1. \( \Sigma' = \Sigma \).
2. \( V_i' = P(V_i) \) where \( n \geq i > 0 \).
3. The construction for the \( \delta_{i'}^n \) functions is given below.
4. \( F' = \{ V \mid V \subset V_n' \text{ and } F \cap V \neq \emptyset \} \)

The construction of the \( \delta_{i'}^n \) functions is

1. If \( \gamma \in \Sigma \cup X_{i+1} \cup X_i \cup X_{i+1}[i \times_{i-1} i] \), then

\[ \delta_{i'}^n(\gamma) = (\delta_{i'}^n(\gamma)) \]
2. If
\[ y \in V_i \cup V_{i+1} \cup (P(\Sigma)UV_{i+1}UX_{i+1})[V_{i+1}UX_{i+1}] \],

then

\[ \delta_n^i(y) = \{U\delta_n^i(y)\} \]

A_n^k, is deterministic, since the value of \( \delta_n^k(y) \) is a single element of \( V_k' = P(V_k) \) in each case. If \( y \in H_{n,m}^i(\Sigma) \), then

\[ A_n^{k,i}(y) = \{A_n^{k,i}(y)\} \]

if \( m \geq 0 \) and \( n \geq i \geq 0 \), since by induction on \( i \)

Basis: if \( m = 0 \) then if \( i > 0 \) then

\[ y = X_1 \]

Therefore,

\[ A_n^{k,i}(y) = \delta_n^i(y) \]

\[ = \{\delta_n^i(y)\} \]

\[ = \{A_n^{k,i}(y)\} \]

If \( i = 0 \), then

\[ y \in X_1 \]

and

\[ A_n^{k,0}(y) = A_n^{k,1}(y) \]

\[ = \{A_n^{k,1}(y)\} \]

\[ = \{A_n^{k,0}(y)\} \]
Induction step: Assume that if \( m > m' \geq 0 \) and \( \gamma \in H^i_{n,m} \), then

\[
A^{k,i}_n(\gamma) = \{A^{k,i}_n(\gamma)\}
\]

If \( \gamma \in H^i_{n,m} \), then

1. If \( i = n \) and \( \gamma \in H^{n+1}_{n,m-1} \), then

\[
A^{k,n'}_n(\gamma) = U_{n'}^{n'}(A^{k,n+1}_n(\gamma)) = U_{n'}^{n'}(\{\gamma\}) = \delta^{n'}_n(\gamma) = \{\delta^{n'}_n(\gamma)\} = \{U_{n'}^{n'}(\{\gamma\})\} = \{A^{k,n}_n(\gamma)\}
\]

2. If \( i = n \) and \( \gamma \in H^n_{n,m-1} \), then

\[
A^{k,n'}_n(\gamma) = \{A^{k,n}_n(\gamma)\}
\]

by induction hypothesis.

3. If \( i = n \) and \( \gamma = \alpha[A_n \beta_n] \), then

\[
A^{k,n'}_n(\gamma) = U_{n'}^{n'}(A^{k,n+1}_n(\alpha[A_n A^{k,n-1}_n(\beta)_n])) = U_{n'}^{n'}(\{\alpha[A_n A^{k,n-1}_n(\beta)_n]\}) = U_{n'}^{n'}(\{\alpha[A_n A^{k,n-1}_n(\beta)_n]\}) = \delta^{n'}_n(\alpha[A_n A^{k,n-1}_n(\beta)_n]) = \{U_{n'}^{n'}(\alpha[A_n A^{k,n-1}_n(\beta)_n])\}
\]
4. If $n > i > 0$ and $\gamma \in H_{n,m-1}^{i+1}$, then

$$A_n^{k,i}(\gamma) = \delta_n^{i'}(A_n^{k,i+1}(\gamma))$$

$$= \delta_n^{i'}(\{A_n^{k,i+1}(\gamma)\})$$

$$= \delta_n^{i'}(A_n^{k,i+1}(\gamma))$$

$$= \{\delta_n^{i'}(A_n^{k,i+1}(\gamma))\}$$

$$= \{A_n^{k,i}(\gamma)\}$$

5. If $n > i > 0$ and $\gamma \in H_{n,m-1}^{i}$, then

$$A_n^{k,i}(\gamma) = (A_n^{k,i}(\gamma))$$

by induction hypothesis.

6. If $n > i > 0$ and $\gamma = a[i \beta i]$, then

$$A_n^{k,i}(\gamma) = U\delta_n^{i'}(A_n^{k,i+1}(a)[i A_n^{k,i-1}(\beta) i])$$

$$= U\delta_n^{i'}(\{A_n^{k,i+1}(a)[i A_n^{k,i-1}(\beta) i]\})$$

$$= U\delta_n^{i'}(\{A_n^{k,i+1}(a)[i A_n^{k,i-1}(\beta) i]\})$$

$$= \delta_n^{i'}(\{A_n^{k,i+1}(a)[i A_n^{k,i-1}(\beta) i]\})$$

$$= \{U\delta_n^{i'}(A_n^{k,i+1}(a)[i A_n^{k,i-1}(\beta) i])\}$$

$$= \{A_n^{k,i}(a)\}$$
7. If \( i = 0 \), then

\[
A_n^{k,0}(\gamma) = A_n^{k,1}(\gamma) = (A_n^{k,1}(\gamma))
\]

by one of the previous sections of this proof (depending on the value of \( n \)).

Therefore, since \( i = n \), \( n > i > 0 \) or \( i = 0 \),

\[
A_n^{k,i}(\gamma) = (A_n^{k,i}(\gamma))
\]

if \( \gamma \in H_{n,m}^k \). Therefore, the conjecture is proved by induction.

If \( \gamma \in \text{language}(A_n^k) \), then

\[
A_n^{k,k}(\gamma) \cap F \neq \emptyset
\]

Therefore, \( A_n^{k,k}(\gamma) \in F' \) by point 4 of the construction and

\[
\{A_n^{k,k}(\gamma)\} \cap F' = \{A_n^{k,k}(\gamma)\}
\]

\[
\neq \emptyset
\]

Therefore, by the conjecture,

\[
A_n^{k,k}(\gamma) \cap F' \neq \emptyset
\]

Therefore, \( \gamma \in \text{language}(A_n^k) \) and

\[
\text{language}(A_n^k) \subset \text{language}(A_n^k)
\]
If \( \gamma \in \text{language}(A_n^k) \), then

\[
A_n^{k,k'}(\gamma) \cap F' \neq \emptyset
\]

Therefore, by the conjecture,

\[
(A_n^{k,k'}(\gamma)) \cap F' \neq \emptyset
\]

Since \( (A_n^{k,k'}(\gamma)) \) has only one element, it must be true that

\[
A_n^{k,k'}(\gamma) \in F'
\]

But by the construction this implies that

\[
F \cap A_n^{k,k'}(\gamma) \neq \emptyset
\]

Since this is the same condition that is required of \( \gamma \) for it to be an element of \( \text{language}(A_n^k) \). Therefore, \( \gamma \in \text{language}(A_n^k) \) and

\[
\text{language}(A_n^{k'}) \subset \text{language}(A_n^k)
\]

and \( A_n^{k'} \) is a deterministic finite automaton which describes the same language as the original automaton \( A_n^k \), and the theorem is proved.

---

At this point, it might be well to see how this definition of finite automata compares to the automata given in the literature. Considering the case of string automata...
first, it will be noted that the transition function $\delta_{1}^{1}$ (the only one) makes the following mapping

$$\delta_{1}^{1}: H_{1,1}^{1} \rightarrow P(V_{1})$$

But notice that $H_{1,1}^{1}$ is a very limited case. Therefore,

$$H_{1,1}^{1} = \Sigma \cup V_{1} \cup X_{1} \cup \Sigma[_{1} (V_{1} \cup X_{1})_{1}]$$

It can be seen that this automata differs from the automata given in the literature, since null transitions can be specified. In particular, this happens when a value of $V_{1}$ is mapped by $\delta_{1}^{1}$. In the definition of the complete transition function, $\delta_{1}^{1}$ is arbitrarily restricted so that this doesn't happen so that

$$H_{1,1}^{1} = \Sigma \cup \Sigma[_{1} (V_{1} \cup X_{1})_{1}]$$

Since $X_{1}$ cannot appear in classical strings they can be similarly deleted.

$$H_{1,1}^{1} = \Sigma \cup \Sigma[_{1} V_{1} \cup 1]$$
If a new state is chosen, call it $q_0$, then a new function, $\delta$, can be defined

$$\delta: \Sigma \times V \rightarrow \mathcal{P}(V)$$

where $V = V_1 \cup \{q_0\}$ and

$$\delta(a,v) = \delta_1(a[1_1 v_1]) \text{ if } v \neq q_0.$$ 
$$= \delta_1(a) \text{ if } v = q_0.$$ 

The function $\delta$ is precisely the transition function needed to generate a classical automaton $(V, \Sigma, \delta, q_0, F)$.

A similar development would show that classical tree automata as defined by Thatcher and Wright (Thatcher 1968) is a special limited case of hypertree automata also. The method used is similar to that used in chapter 4 for grammars over trees. As was the case with grammars, hypertree finite automata can describe a larger class of tree languages than classical tree automata as the example given earlier in this chapter indicates. It has been shown that the set of languages accepted by finite automata over trees is the same as the set of languages described by regular grammars over trees in the classical case, so this could be expected. This same property is true in the case of hypertrees also, as the following theorem proves.
Lemma 47:

For every normal regular grammar $G_n = (L, V, P, S)_n^k$, there exists a finite automaton $A_n^k = (\Sigma', V', \delta', F')_n^k$ such that if

\[ v \Rightarrow^* \gamma \]

then

\[ v \in A_n^{k,i}(\gamma) \]

for all $v \in V_i$, $n \geq k > 0$ and $n \geq i > 0$.

Proof:

Since any production of the form $A \rightarrow B$ where $A$ and $B$ are both in $V_i$ for some $i$ can be easily eliminated by forward substitution assume no productions of this type occur in $G_n^k$. In a similar fashion, it is possible to alter $G_n^k$ so that $X$'s only appear in productions of the form

\[ v \rightarrow X_i \]

Therefore, assume this is true of $G_n^k$ also. Construct the finite automaton such that

1. $\Sigma' = \Sigma$.
2. $V' = V$.
3. If $v \Rightarrow \gamma \in P$ and $v \in V_i$, then $v \in \delta_n^{i}(\gamma)$ and these are the only elements is $\delta_n^{i}(\gamma)$.
4. $F' = \{S\}$.
Claim: $A^k_n$ is the required automaton. This will be proved by induction on the number of applications of $=>$ used to produce $\gamma$.

Basis: if only one step is required to produce $\gamma$ then there must be a production

$$v \rightarrow \gamma$$

in $P$. Therefore, by construction

$$v \in \delta^i_n(\gamma)$$

Therefore,

1. If $\gamma \in X_i$, then

$$v \in \delta^i_n(\gamma)$$

$$\epsilon A^k_n$$

2. If $\gamma \in \Sigma$ and $i = n$, then

$$v \in U\delta^i_n(\gamma)$$

$$\epsilon U\delta^i_n(A^k_n)$$

$$\epsilon A^i_{n}(\gamma)$$

Note that $\gamma \neq \alpha[i \beta_i]$, since $\beta$ must be a state. Therefore,

$$\gamma \in H^i_n(\Sigma) \text{ and } \gamma \notin V_i U V_{i+1}.$$
Induction step: Assume that if $v \Rightarrow^* \gamma$ in less than $m$ steps then

$$v \in A_n^{k,i}(\gamma)$$

Suppose that $v \Rightarrow^* \gamma$ in $m > 1$ steps where $v \in V_1$. Then it must be true that $v \Rightarrow \gamma_1 \Rightarrow^* \gamma$ where $v \Rightarrow \gamma_1 \in P$. Therefore, $v \in \delta_n^i(\gamma_1)$ by construction. $\gamma_1 \notin I \cup X_1$, since this would have to be the last production if that were the case and it was assumed that there were more than one production.

1. If $\gamma_1 \in V_{i+1}$, then $\gamma_i \Rightarrow^* \gamma$ in fewer than $m$ steps, so that

$$\gamma_i \in A_n^{k,i+1}(\gamma)$$

by induction hypothesis. Therefore,

$$\delta_n^i(\gamma_i) \subseteq U\delta_n^i(A_n^{k,i+1}(\gamma))$$

$$\subseteq A_n^{k,i}(\gamma)$$

Therefore,

$$v \in A_n^{k,i}(\gamma)$$

2. If $\gamma_1 = \alpha[n \beta_1 n]$, then

$$\gamma = \alpha[n \beta n]$$

where $\alpha = \alpha_1$ and $\beta_1 \Rightarrow^* \beta$ in fewer than $m$ steps. Therefore,

$$\beta_1 \in A_n^{k,n-1}(\beta)$$
also

\[ \delta_n^n(y_1) = \delta_n^n(\alpha[n, \beta_n]) \]

\[ = \bigcup \delta_n^n((\alpha)[n A_{k, n-1}(\beta) n]) \]

\[ = \bigcup \delta_n^n(A_{k, n+1}(\alpha)[n A_{n, k, n-1}(\beta) n]) \]

\[ = A_{k, n}^n(\alpha[n, \beta_n]) \]

\[ = A_{k, n}^n(y) \]

and \( v \in A_{k, n}^n(y) \).

3. If \( y_1 = \alpha_1[i, \beta_1 i] \) and \( n > i \), then

\[ y = \alpha[i, \beta_i] \]

where \( \alpha_1 \Rightarrow_1 \alpha \) and \( \beta_1 \Rightarrow_1 \beta \) with fewer than \( m \) steps each. Therefore, \( \alpha_1 \in A_{n, k, i+1}(\alpha) \), \( \beta_1 \in A_{n, k, i-1}(\beta) \) and

\[ \delta_n^i(y_1) = \delta_n^i(\alpha_1[i, \beta_1 i]) \]

\[ = \bigcup \delta_n^i(A_{n, k, i+1}(\alpha)[i A_{n, k, i-1}(\beta) i]) \]

\[ = A_{n, k, i}^i(\alpha[i, \beta_i]) \]

\[ = A_{n, k, i}^i(y) \]

and \( v \in A_{n, k, i}^i(y) \).

Since \( v \in A_{n, k, i}^i(y) \) in each case the lemma is proved by induction.

---
Lemma 48:

For all normal grammars $G^k_n = (\Sigma, V, P, S)^k_n$ there is a finite automaton $A^k_n$ such that

$$\text{language}(A^k_n) = \text{language}(G^k_n)$$

for all $n \geq k > 0$.

Proof:

Use the automaton constructed for lemma 47.

$$\text{language}(G^k_n) = \{\gamma | S^* \Rightarrow \gamma \text{ and } \gamma \in H_n(\Sigma)\}$$

$$= \{\gamma | S \in A^k_n, \gamma \in (\gamma) \text{ and } \gamma \in H_n(\Sigma)\}$$

$$= \{\gamma | A^k_n, \gamma \in (\gamma) \cup \{S\} \neq \emptyset \text{ and } \gamma \in H_n(\Sigma)\}$$

$$= \text{language}(A^k_n)$$

Q.E.D.

---

Theorem 49:

For every regular grammar $G^k_n$ there exists a finite automaton $A^k_n$ such that

$$\text{language}(G^k_n) = \text{language}(A^k_n)$$

for all $n \geq k > 0$.

Proof:

This follows directly from lemma 48.

---
Therefore, any regular grammar has a finite automaton that accepts the same language as the grammar generates. It is also true that any finite automaton has a grammar which produces the same language as the automaton accepts, however, before this is proved a third method of describing regular languages will be discussed, namely regular expressions.
Yet a third method of describing regular sets is called regular expressions. In this chapter, it will be shown that regular expressions, which have been defined over strings and trees, may be extended to hypertrees of any level. Again, in the case of trees the hypertree regular expressions will be more powerful than those described by Thatcher and Wright (Thatcher and Wright 1968).

Regular expressions describe regular sets by using operations on sets. These are put together into an expression, called a regular expression, which then describes the regular set. Using the ideas presented by Thatcher and Wright, rather than concatenating sets as is done in regular string expressions, elements or the frontier of hypertrees will be replaced to effect concatenation. However, in this case the symbols to be replaced will be collected together into a ranked set called the set of auxiliary symbols. These will correspond roughly to the nonterminals used in regular grammars and the states used in finite automata.

Before regular expressions can be defined, two new operations need to be defined for hypertrees.
Definition 45:
Concatenation using $v$ over $H_n^k$ is an operation, denoted $*_v^k$, such that

$$*_v^k: H_n^k \times P(H_n) \rightarrow P(H_n^k)$$

where $v \in V_1$, $n \geq 0$, $n+1 \geq k \geq 0$, $n \geq i > 0$ and the value of $*_v^k$ is the smallest set such that

(a) $*_v^{n+1}(A,B) = \{A\}$ if $n \geq 0$.

(b) $*_v^0(A,B) = *_v^1(A,B)$ if $n \geq 0$.

(c) $*_v^k(A,B) = \{A\}$ if $A \in H_n^k$.

$*_v^n(A,B) = \{A\}$ if $A \in X_k \cup V_k$ and $A \neq V$.

$= B$ if $A = V$.

$= *_v^{k+1}(\alpha, \beta) \left[ k^* v^{k-1}(\beta, B) \right]^k$ if $A = \alpha \left[k^* \beta \right]^k$.

if $n \geq k > 0$.

Even though concatenation has only been defined for individual elements of $H_n^k$ in the first argument, it will be extended to subsets of $H_n^k$ with no formal definition. Also, if $A$ is a subset of $H_n^k$, $B$ is a subset if $H_n^l$ and $v$ is an
element of $V_i$ then $\ast v^k_n(A, B)$ will be written using infix notation as

$$A \ast v B = U \ast v^k_n(A, B)$$

Even though the $k, n, \Sigma$ and $V$ do not appear in this notation they will be assumed.

Notice that concatenation replaces all occurrences of the argument symbol $v$ with a string from the set $B$. For example, if

$$A = a[2, a[2, v[2][1, v[1]]2]$$

and

$$B = \{c, d\}$$

where $v \in V_1$ then

$$A \ast v B = a[2, a[2, v[2][1, v[1]]2] \ast v B$$

$$= (a \ast v B)[2, a[2, v[2][1, v[1]]2] \ast v B[2]]$$

$$= a[2, a[2, v[2]] \ast v B[2], (v \ast v B[2]), 1, 2]$$

$$= a[2, a[2, v B[2], v B[2]], 1, 2]$$

$$= a[2, a[2, c, d[2][1, c, d[1]]2]]$$

$$= \{a[2, a[2, c[2]], 1, c[1][2]], a[2, a[2, c[2]], 1, d[1][2]], a[2, a[2, d[2]], 1, c[1][2]], a[2, a[2, d[2]], 1, d[1][2]]\}$$
As has been previously the case, the codomain of the concatenation operation has been assumed. This codomain will now be proved to be correct.

Lemma 50:

If \( n + 1 \geq k \geq 0, \ n \geq i > 0, \ m \geq 0, \ v \in V_i, \ A \in H^k_n, m \) and \( B \in P(H^i_n) \), then

\[
A \ast_v B \subseteq H^k_n
\]

Proof by induction on \( m \):

Basis: If \( m = 0 \), then one of the following cases is true:

1. If \( k = n + 1 \), then

\[
A \ast_v B = \{A\}
\]

but \( A \in H^{n+1}_n \), therefore, \( \{A\} \subseteq H^{n+1}_n \) and

\[
A \ast_v b \subseteq H^{n+1}_n
\]

2. If \( n \geq k > 0 \) and \( A \neq v \), then \( A \in X_k \cup V_k \).

\[
A \ast_v B = A \\
\subseteq H^k_n
\]

3. If \( A = v \), then \( i = k \) so that \( B \subseteq H^k_n \). Therefore,

\[
A \ast_v B = B \\
\subseteq H^k_n
\]
4. If $k = 0$, then $A \ast_v B$ is a special case of one of the previous sections of this proof since $A \in H_n^1$.

Induction step: Assume that if $m > m' \geq 0$ and $A \in H_{n,m}'$ then $A \ast_v B \subseteq H^k_n$.

If, therefore, $A \in H_{n,m}'$, then one of the following is true.

1. If $k = n+1$, then

\[ A \ast_v B = \{A\} \subseteq H^{n+1}_n \]

2. If $n \geq k > 0$ and $A \in H^{k+1}_{n,m-1}$, then $A \ast_v B \subseteq H^{k+1}_n$ by induction hypothesis. But $H^{k+1}_n \subseteq H^k_n$. Therefore,

\[ A \ast_v B \subseteq H^k_n \]

3. If $n \geq k > 0$ and $A \in H^k_{n,m-1}$, then

\[ A \ast_v B \subseteq H^k_n \]

by induction hypothesis.

4. If $n \geq k > 0$ and $A = a[k \beta_k]$, then $a \ast_v B \subseteq H^{k+1}_n$ and $\beta \ast_v B \subseteq H^{k-1}_n$ by induction hypothesis. Since

\[ A \ast_v B = a \ast_v B[k \beta \ast_v B k] \]

it must be true that

\[ A \ast_v B \subseteq H^k_n \]
5. If $K = 0$, then $A \in \mathbb{H}_{n,m}^1$ and $A \ast_{\mathbf{v}} B \subset \mathbb{H}_{n}^1$ by a previous section of this proof. But since $\mathbb{H}_{n}^1 = \mathbb{H}_{n}^0$,

$$A \ast_{\mathbf{v}} B \subset \mathbb{H}_{n}^0$$

Since in each of these cases $A \ast_{\mathbf{v}} B \subset \mathbb{H}_{n}^k$ the lemma is proved by induction.

---

Theorem 51:

If $n+1 > k \geq 0$, $n \geq i > 0$, $\forall \in \mathbb{V}_{i}$, $A \in \mathbb{H}_{n}^k$ and $B \in \mathbb{H}_{n}^i$, then

$$A \ast_{\mathbf{v}} B \subset \mathbb{H}_{n}^k$$

Proof:

This follows directly from the lemma.

---

Just as is the case with strings, concatenation will be extended to form the closure of a set.

Definition 47:

The closure of a set using $\mathbf{v}$ over $\mathbb{H}_{n}^k$ is an operation, denoted $^{+\mathbf{v}}$, such that

$$^{+\mathbf{v}} : \mathcal{P}(\mathbb{H}_{n}^k) \rightarrow \mathcal{P}(\mathbb{H}_{n}^k)$$

where $\forall \in \mathbb{V}_{i}$, $n \geq 0$, $n \geq k > 0$ and $^{+\mathbf{v}}(A)$ is the smallest set such that

$$^{+\mathbf{v}}(A) = A \cup ^{+\mathbf{v}}(A) \ast_{\mathbf{v}} A$$
In order to simplify the notation, $^+v(A)$ is written $A^+v$ in this paper. The following theorem establishes the domain of the operation.

Theorem 52:

If $n \geq k > 0$, $V \in V_k$ and $A \subseteq H_k^n$, then

$$A^+v \subseteq H_k^n$$

Proof:

Note that the smallest solution to the equation

$$A^+v = A \cup A^+v$$

is the set defined by the following inductive definition

1. $A_0^+v = A$.

2. $A_j^+v = A \cup A^+v \cap A_{j-1}^+v$.

where

$$A^+v = \bigcup_{j \geq 0} A_j^+v$$

Therefore, proof by induction on $j$:

Basis: If $j = 0$, then

$$A_0^+v = A \subseteq H_k^n$$

Induction step: Assume that if $j \geq 0$ then

$$A_j^+v \subseteq H_k^n$$

Note that
\[ A_{j+1}^+ = A \cup A \uparrow \backslash A_j^+ \]

but \( A \uparrow \backslash A_j^+ \subseteq H_n^k \) by theorem 51. Therefore,

\[ A_{j+1}^+ \subseteq H_n^k \]

and

\[ A_j^+ \subseteq H_n^k \]

for all \( j \geq 0 \) by induction. Since \( V^+ = U_{j \geq 0} A_j^+ \) and all of the \( A_j^+ \) are subsets of \( H_n^k \) the theorem is proved.

---

At this point, the definition for regular expressions may be given.

**Definition 48:**

The set of regular expressions over \( \Sigma \) of level \( n \) and degree \( k \) with auxiliary symbols \( V \), denoted \( R_n^k(\Sigma, V) \), is the smallest set such that

(a) \( R_n^{n+1} = \Sigma \)

if \( n \geq 0 \).

(b) \( R_n^0 = R_n^1 \)

if \( n \geq 0 \).

(c) \( R_n^k = R_n^{k+1} \cup \Sigma \cup U_{k \geq 1} R_n^{k-1} \cup [k \cdot R_n^k \cup k] \cup R_n^{k+1} \)

\( \cup \Sigma \cup U_{k \geq 0} (R_n^k \cup V_i R_n^i) \)

if \( n \geq k > 0 \).
Table 4 gives some examples of regular expressions.

As was the case with hyperforests, proofs using regular expressions are difficult because of the expression $R_{n}^{k+1} \left[ k R_{k-1}^{\frac{k-1}{k}} \right]$. For this reason, a second definition is added with an induction variable similar to the depth in hypertrees.
Definition 49:

The set of regular expressions over \( \Sigma \) of level \( n \), degree \( k \) and depth \( m \) with auxiliary symbol \( V \), denoted \( R^k_n(\Sigma, V) \), is the smallest set such that

(a) \( R^{n+1}_{n,m} = \Sigma \)

if \( n \geq 0 \) and \( m \geq 0 \).

(b) \( R^k_{n,0} = X_k \cup V_k \)

if \( n \geq k > 0 \).

(c) \( R^k_{n,m+1} = R^k_{n,m} \cup R^k_{n,m} \cup R^k_{n,m} [ R^{k-1}_{n,m,k} \]

\[ \cup R^k_{n,m} \circ R^k_{n,m} \cup R^k_{n,m} \circ V_k \cup U_{n \geq i > 0}(R^k_{n,m} \circ V_i \circ R^{i}_{n,m}) \]

if \( n \geq k > 0 \) and \( m \geq 0 \).

(d) \( R^0_{n,m} = R^n_{n,m} \)

if \( n \geq 0 \) and \( m \geq 0 \).

Theorem 53:

If \( n \geq 0 \) and \( n+1 \geq k \geq 0 \), then

\( R^k_n(\Sigma, V) = \cup_{m \geq 0} R^k_{n,m}(\Sigma, V) \)

Proof:

This proof follows the corresponding proof of theorem 7 in chapter 2, and will not be given here.

- - -
Definition 50:

The language function of level $n$ and degree $k$, denoted $\text{language}_n^k$, is defined such that

$$\text{language}_n^k : \mathcal{R}_n^k \rightarrow \mathcal{P}(\mathcal{H}_n^k)$$

where

(a) $\text{language}_n^{n+1}(A) = \{A\}$ if $A \in \mathcal{I}$ and $n \geq 0$.

(b) $\text{language}_n^0(A) = \text{language}_n^1(A)$ if $n \geq 0$.

(c) $\text{language}_n^k(A) = \text{language}_n^{k+1}(A)$

if $A \in \mathcal{R}_n^{k+1}$.

$= \{A\}$

if $A \in X_k \cup Y_k$.

$\text{language}_n^k(A) = \text{language}_n^{k+1}(A)[k \text{language}_n^{k-1}(\beta) \in k]$

if $A = [k \beta]$.

$= \text{language}_n^k(\alpha) \cup \text{language}_n^k(\beta)$

if $A = \alpha + \beta$.

$= (\text{language}_n^k(\alpha))^+v$

if $A = \alpha^v$.

$= \text{language}_n^k(\alpha) \cdot_v \text{language}_n^i(\beta)$

if $A = \alpha \cdot_v \beta$ and $v \in V_i$.

if $n \geq k > 0$.

*
Definition 51:
A regular expression, $A \in R_n^k$, is a normal regular expression if

$$\text{language}_n^k(A) \subseteq H_n^k(\Sigma)$$

If no confusion will result, the $k$ and $n$ in $\text{language}_n^k$ will be elided.

As has been the case, a proof is now in order to show that the codomain described in Definition 50 is in fact correct.

Theorem 54:
If $n \geq 0$, $n+1 \geq k \geq 0$, $m \geq 0$ and $A \in R_n^k(\Sigma, \mathcal{V})$, then

$$\text{language}(A) \subseteq H_n^k$$

Proof by induction on the induction variable $m$:

Basis: If $m = 0$, then

1. If $k = n+1$, then $A \in \Sigma$ and

$$\text{language}(A) = \{A\}$$

$$\subseteq H_{n+1}^k$$

2. If $n \geq k > 0$, then $A \in X_k^k U V_k$ so that

$$\text{language}(A) = \{A\}$$

$$\subseteq X_k^k U V_k$$

$$\subseteq H_n^k$$

3. If $k = 0$, then
language(A) = language_n^{1}(A)
but by either case 1 or case 2,

language_n^{1}(A) \subseteq H_n^{1}

since H_n^{1} = H_n^{0},

language(A) \subseteq H_n^{k}

Therefore, the basis is proved.

Induction step: Assume that if m > m' \geq 0 and A \in R^{k}_{n,m}', then

language(A) \subseteq H_n^{k}

If A \in R^{k}_{n,m}', then one of the following cases holds:

1. If k = n+1, then A \in \Sigma and

language(A) = \{A\}
\subseteq H_n^{k}

2. If n \geq k > 0 and A \in R^{k+1}_{n,m-1}', then

language(A) = language_n^{k+1}(A)

but

language_n^{k+1}(A) \subseteq H_n^{k+1}

by induction hypothesis. Since H_n^{k+1} \subseteq H_n^{k}

language(A) \subseteq H_n^{k}

3. If n \geq k > 0 and A \in R^{k}_{n,m-1}', then
by induction hypothesis.

4. If \( n \geq k > 0 \) and \( A = \alpha[k \beta k] \), then

\[
\text{language}(A) = \text{language}^k_n(\alpha)[k \text{ language}^{k-1}_n(\beta) k]
\]

but both \( \text{language}^{k+1}_n(\alpha) \subseteq H^{k+1}_n \) and

\( \text{language}^{k-1}_n(\beta) \subseteq H^{k-1}_n \) by induction hypothesis so

\[
\text{language}(A) \subseteq H^k_n
\]

5. If \( n \geq k > 0 \) and \( A = \alpha + \beta \), then

\[
\text{language}(A) = \text{language}^k_n(\alpha) + \text{language}^k_n(\beta)
\]

but, by induction hypothesis \( \text{language}^k_n(\alpha) \subseteq H^k_n \) and

\( \text{language}^k_n(\beta) \subseteq H^k_n \), therefore,

\[
\text{language}(A) \subseteq H^k_n
\]

6. If \( n \geq k > 0 \) and \( A = \alpha^* \), then

\[
\text{language}(A) = (\text{language}^k_n(\alpha))^+\nu
\]

but since \( \text{language}^k_n(\alpha) \subseteq H^k_n \) by induction hypothesis,

\[
\text{language}(A) \subseteq H^k_n
\]

by theorem 52.

7. If \( n \geq k > 0 \) and \( A = \alpha \nu \beta \), then
language(\(A\)) = \(\text{language}_{n}^{k}(\alpha) \ast_{v} \text{language}_{n}^{i}(\beta)\)

But \(\text{language}_{n}^{k}(\alpha) \subset H_{n}^{k}\) and \(\text{language}_{n}^{i}(\beta) \subset H_{n}^{i}\) by induction hypothesis. Therefore,

\[\text{language}(A) \subset H_{n}^{k}\]

by theorem 51.

8. If \(k = 0\), then

\[\text{language}(A) = \text{language}_{n}^{1}(A)\]

by case 1 if \(n = 0\) or cases 2 to 7 if \(n > 0\),

\[\text{language}_{n}^{1}(A) \subset H_{n}^{1}\]

but

\[H_{n}^{1} = H_{n}^{0}\]

therefore,

\[\text{language}(A) \subset H_{n}^{k}\]

and the theorem is proved by induction.

---

If \(n = 1\) and \(k = 1\), then the definition of regular expressions becomes

\[R_{1}^{1} = R_{1}^{2} U X_{1} U V_{1} U R_{1}[1 \, R_{1}^{0}] U R_{1}^{1}+R_{1}^{1} U R_{1}^{1*} V U R_{1}^{1*} R_{1}\]

notice that \(R_{1}^{0} = R_{1}^{1}\) so that
\[ R_1^1 = R_1^2 \cup X_1 \cup V_1 \cup R_1^2[R_1^1 R_1^1] \cup R_1^1 R_1^1 \cup R_1^1 v R_1^1 \]

also, \( R_1^1 = \Sigma \) so

\[ R_1^1 = \Sigma \cup X_1 \cup V_1 \cup \Sigma[R_1^1 R_1^1] \cup R_1^1 R_1^1 \cup R_1^1 v R_1^1 \]

If \( V_1 = \{ \lambda \} \), then

\[ R_1^1 = \lambda \cup X_1 \cup \{ \lambda \} \cup \Sigma[R_1^1 R_1^1] \cup R_1^1 R_1^1 \cup R_1^1 v R_1^1 \]

If it is decided arbitrarily that all elements of \( R_1^1 \) are to end in \( \lambda \), then (call this set "\( R_1^1 \)"

\[ R_1^1 = \{ \lambda \} \cup \Sigma[R_1^1 R_1^1] \cup R_1^1 R_1^1 \cup R_1^1 v R_1^1 \]

Notice that \( \Sigma[R_1^1] \) describes the same set as \( \Sigma[\lambda] \) so the later may be substituted for the former in a new set of regular expressions \( R_2^1 \)

\[ R_2^1 = \lambda \cup \Sigma[\lambda \lambda] \cup R_2^1 R_2^1 \cup R_2^1 v R_2^1 \]

Lastly, if a new set of regular expressions is defined, denoted \( R \), with the extraneous \( \lambda \)'s elided, and the \( [1 \) and \( 1 \) elided, then the equation becomes

\[ R = \lambda \cup \Sigma \cup R \cup R v R \]

which is exactly the set of regular expression over \( \Sigma \) in the classical string case.
If \( n = 2 \) and \( k = 2 \) (the case of trees over \( \Sigma \)), then the definition of regular expressions becomes

\[
R_2^2 = R_2^3 \cup X_2 \cup V_2 \cup R_2^3 \left[ \{ R_2 \} \right] \cup R_2^2 + R_2^2 \cup R_2^2 * v \cup R_2^2 * V_2
\]

But \( R_2^2 = \Sigma \) so this becomes

\[
R_2^2 = \Sigma \cup X_2 \cup V_2 \cup \{ R_2 \} \cup R_2^2 + R_2^2 \cup R_2^2 * v \cup R_2^2 * V_2
\]

If a new set of regular expressions are defined with only elements of \( \Sigma \) on the frontier, and \( V_2 \) is set equal to \( \Sigma \), then the resulting regular set, \( R_1 \), is

\[
R_1 = \Sigma \cup \{ R_2 \} \cup R_1 + R_1 \cup R_1 * v \cup R_1 * V_1
\]

where \( v \in \Sigma \). If it is assumed that \( R_2^1 \) can only take a limited form, in particular, any element of \( \Sigma \) can only have a string of a fixed number of sub-expressions after it, and if this number is designated by ranking \( \Sigma \), then the equation becomes

\[
R_2 = \Sigma_0 \cup \Sigma_1 \cup \{ R_2 \} \cup R_1 + R_1 \cup R_1 * v \cup R_1 * V_1
\]

where \( v \in \Sigma_0 \). The first two terms can be used to build up any tree. Both of these first two terms can be described by a finite set of trees and concatenation, these may be replaced with the set of trees over \( \Sigma \), \( T_\Sigma \).

\[
R = T_\Sigma \cup R + R \cup R^* \Sigma \cup R^* R
\]
is the resulting regular expressions, which are the regular expressions described by the Thatcher and Wright.

In order to prove that a normal regular expression exists that corresponds to any given automaton, a new type of automaton will be defined.

Definition 52:

A partial finite automata, denoted $P_n^k$, is a fivetuple $(\Sigma, V, \nu, \delta, F)_n^k$ where $n \geq k > 0$ and

1. $\Sigma$ is a set of terminal symbols.
2. $V$ is a ranked set of argument symbols.
3. $\nu$ is a ranked set of states.
4. $\delta$ is a set of incomplete finite functions called the transition functions.

$$\delta_n^i: \nu_n^i(\Sigma, V U V) \rightarrow P(\nu_1)$$

where $n \geq i > 0$.

5. $F$ is a finite subset of $V_k^*$ called the set of final states.

Notice that if $V = \emptyset$ this corresponds directly to the automata defined in chapter 5. The language of $P_n^k$, denoted $\text{language}(P_n^k)$, is defined as in the standard automata given except that elements of $V$ are taken as terminal symbols in addition to $\Sigma$. An interesting theorem is
Theorem 55:

If \( n \geq k > 0 \) and \( P^k_n = (\Sigma, \mathcal{V}, \mathcal{V}, \delta, F)^n \) is a partial finite automata, then there exists a finite automata \( A^k_n = (\Sigma', \mathcal{V}', \delta', F') \) such that

\[
\text{language}(A^k_n) = \text{language}(P^k_n)
\]

Proof:

Construct \( A^k_n \) as follows

1. \( \Sigma' = \Sigma \cup \mathcal{V} \).
2. \( V_i' = V_i \cup V_i^i \) where \( V_i^i = \{ a_i | \ a \in \mathcal{V} \} \).
3. The construction of \( \delta^i_n \) is given below.
4. \( F' = F \).

The construction of the \( \delta^i_n \) functions is as follows

1. If \( a \in H_{i+1}(\Sigma) \) and \( \delta^i_n(a) = P \), then

\[
\delta^i_n'(a) = P
\]

2. If \( \delta^i_n(a[i \ \beta \ i]) = P \) and \( a \in V_{i+1} \), then

\[
\delta^i_n'(a[i+1[i \ \beta \ i]) = P
\]

3. If \( \delta^i_n(a[i \ \beta \ i]) = P \) and \( \beta \in V_{i-1} \), then

\[
\delta^i_n'(a[i \ \beta_i-1 \ i]) = P
\]

4. If \( \delta^i_n(a[i \ \beta \ i]) = P \), \( a \in V_{i+1} \) and \( \beta \in V_{i-1} \), then

\[
\delta^i_n'(a[i+1[i \ \beta_i-1 \ i]) = P
\]
5. If $\delta_n^i(\alpha) = \mathcal{P}$ and $\alpha \in \mathcal{V}_{i+1}$, then

$$\delta_n^i(\alpha_{i+1}) = \mathcal{P}$$

6. If $\delta_n^1(\alpha) = \mathcal{P}$ and $\alpha \in \mathcal{V}_i$, then

$$\delta_n^i(\alpha_i) = \mathcal{P}$$

7. If $\alpha \in \mathcal{V}_i$, then

$$\delta_n^i(\alpha) = \{\alpha_n\}$$

and

$$\delta_n^j(\alpha_{j+1}) = \{\alpha_j\}$$

for all $n > j > 0$.

8. These are all the elements of $\delta$.

A simple inductive proof will show that

$$\text{language}(A_n^k) = \text{language}(F_n^k)$$

so that the theorem is proved.

---

It is now time to close the loop of equivalences. That is, to prove that regular sets, finite automata, and regular grammars all describe the same set of languages. The first theorem will establish that there is a normal regular expression for the language described by any finite automata.
Theorem 55:
If \( n > k > 0 \) and \( A_n^k = (\Sigma, V, \delta, F)_n^k \) is a finite automaton, then there exists a normal regular expression \( B \in R_n^k \) such that
\[
\text{language}(B) = \text{language}(A_n^k)
\]

Proof:
Note that if \( A_n^{'k} = (\Sigma, V', \delta, F)_n^k \) then
\[
\text{language}(A_n^{'k}) = \text{language}(A_n^k)
\]
if \( V' \) is the set of those elements of \( V \) which occur in \( \delta \).
Therefore, there is no loss in generality by assuming the \( V \)
is finite. Since \( V \) is finite, the elements of \( V \) can be
ordered. Therefore, \( V \) may be represented as the set
\( \{V_1, \ldots, V_m\} \). Define the partial finite automata
\( Q_j, T = (\Sigma', V', V', \delta', F')_n^k \) where \( m > i \geq 1, m > j \geq 1 \) and
\( T = \{v_{i+1}, \ldots, v_m\} \) such that
1. \( \Sigma' = \Sigma \).
2. \( V' = T \).
3. \( V' = V - T \).
4. \( \delta^i_n(a) \cap \{v_1, \ldots, v_i\} \).
5. \( F' = \{v_j\} \).
Define \( S_j^i, T = \text{language}(Q_j, T) \) if \( i > 0 \). If \( i = 0 \), then
\[
S_j^0, T = \{a \mid v_j \in \delta^i_n(a) \text{ for some } i\}
\]
It can be easily seen that
language(\(A_n^k\)) = \(U_{j} F S_j^m\)

To prove this theorem a slightly stronger proof will be presented.

Conjecture: For all \(m \geq i \geq 0\), \(m \geq j \geq 1\) and \(T \subseteq \{v_{i+1} \ldots v_m\}\), there is a regular expression for \(S_j^i, T\).

Proof by induction on \(i\):

Basis: If \(i = 0\), then \(S_j^0, T\) is a finite set for all \(j\) and \(T\) so it has a regular expression.

Induction step: Assume that if \(i' < i\) then there is a regular expression \(R_j^{i'}, T\) such that

\[ \text{language}(R_j^{i'}, T) = S_j^{i'}, T \]

Let

\[ A = R_j^{i-1} \_
\]

\[ B = R_i^{i-1} \_
\]

and

\[ C = R_i^{i-1} \_
\]

A, B and C exist by induction hypothesis. It is claimed that

\[ S_j^i, T = A \_ V_i B^+ V_i \_ V_i C \]
This may be established by a simple but lengthy induction on the induction variable for each element of $S_{j,T}^i$ and again for each element of the expression on the right of the equal sign. Therefore, in particular, $R_{j,\phi}^m$ exists were $v_j \in F$. A regular expression for the entire language is

$$R_{j_1,\phi}^m + \ldots + R_{j_p,\phi}^m$$

where

$$F = \{v_{j_1}, \ldots, v_{j_p}\}$$

---

Now the last link in the loop of equivalences will be proved.

**Theorem 57:**

If $n \geq k > 0$, $m \geq 0$ and $B \in R_{n,m}^k$, then there is a grammar $G_n^k = (\Sigma', \mathcal{V}', \mathcal{P}', S')_n^k$ such that

$$\text{language}(G_n^k) = \text{language}(B)$$

**Proof by induction on $m$:**

**Basis:** If $m = 0$, then $B \in \mathcal{X}_k^* \cup \mathcal{V}_k^*$. Construct $G_n^k$ as follows:

1. $\Sigma' = \{B\}$.
2. $\mathcal{V}_k = \{S'\}$.
3. $\mathcal{P}' = \{S' \rightarrow B\}$.
4. $S'$ does not occur elsewhere.
Induction step: Assume that if \( m > m' > 0 \) and \( B \in R_{n,m}^k \), then there is a grammar \( G_n^k \) such that

\[
\text{language}_n^k(B) = \text{language}(G_n^k)
\]

If \( B \in R_{n,m'}^k \), then one of the following cases is true:

1. If \( k = n \) and \( B \in R_{n,m-1}^{n+1} \), then \( B \in \Sigma \). Construct \( G_n^k \) as follows:
   1. \( R = \{B\} \).
   2. \( V' = \{S'\} \).
   3. \( P' = \{S' \rightarrow B\} \).
   4. \( S' \) does not occur elsewhere.

2. If \( n > k > 0 \) and \( B \in R_{n,m-1}^{k+1} \), then there is a grammar such that

\[
\text{language}((\Sigma, V, P, S)_n^k) = \text{language}^k(B) = \text{language}_n^k(B)
\]

Construct the new grammar \( G_n^k \) such that

1. \( \Sigma' = \Sigma \).
2. \( V_i' = V_{i\alpha} \) if \( i \neq k \) and \( V_k' = V_{k\alpha} \cup \{S\} \).
3. \( P' = P_{\alpha} \cup \{S' \rightarrow S\} \).
4. \( S' \) does not occur elsewhere.

The language is given by

\[
\text{language}(G_n^k) = \{y | S^* \rightarrow y\}
\]
\[
= \{y | S \Rightarrow S_{\alpha}^* \Rightarrow y\}
\]
3. If \( n \geq k > 0 \) and \( B \in \mathcal{R}_{n,m-1}^k \), then there is a grammar \( G_n^k \) such that

\[
\text{language}_n(B) = \text{language}(G_n^k)
\]

by induction hypothesis.

4. If \( k = n \) and \( B = \alpha[\kappa \beta \kappa] \), then there is a grammar \( (\Sigma_\beta', V_\beta', P_\beta', S_\beta') \) such that

\[
\text{language}((\Sigma_\beta', V_\beta', P_\beta', S_\beta')^k) = \text{language}_{n-1}^k(\beta)
\]

by induction hypothesis. Construct the new grammar \( G_n^k \) such that

1. \( \Sigma' = \Sigma_\beta \cup \{\alpha\} \).
2. \( V_i' = V_i \beta \) if \( i \neq n \) and \( V_n' = V_n \beta \cup \{S\} \).
3. \( P' = P_\beta \cup \{S' \rightarrow \alpha[\kappa \beta \kappa]\} \).
4. \( S' \) does not occur elsewhere.

In this case, the language of \( G_n^k \) is given as

\[
\text{language}(G_n^k) = \{\gamma \mid S \Rightarrow \gamma\}
\]

\[
= \{\gamma \mid S \Rightarrow \alpha[\kappa \beta \kappa]\}
\]

\[
= \text{language}_n^{k+1}(\alpha)[\kappa(\gamma_1 \mid S_\beta \Rightarrow \gamma_1)_{\kappa}]
\]

\[
= \text{language}_n^{k+1}(\alpha)[\kappa \text{language}_{n-1}^{k-1}(\beta)_{\kappa}]
\]

\[
= \text{language}_n^{k}(\alpha[\kappa \beta \kappa])
\]
Therefore, \( G_n^k \) is the required grammar.

5. If \( n > k > 1 \) and \( B = \alpha [k \beta k] \), then there are grammars such that

\[
\text{language}\left(\left(\Sigma_\alpha, V_\alpha, P_\alpha, S_\alpha\right)_n^{k+1}\right) = \text{language}(\alpha)
\]

and

\[
\text{language}\left(\left(\Sigma_\beta, V_\beta, P_\beta, S_\beta\right)_n^{k-1}\right) = \text{language}(\beta)
\]

where \( V_\alpha \cap V_\beta = \emptyset \). Construct the new grammar \( G_n^k \) such that

1. \( \Sigma' = \Sigma_\alpha U \Sigma_\beta \).
2. \( V_i' = V_i\alpha U V_i\beta \) if \( i \neq k \) and
   \( V_k' = V_k\alpha U V_k\beta U \{S'\} \).
3. \( P' = P_\alpha U P_\beta U \{S' \rightarrow S_\alpha[k S_\beta k]\} \).
4. \( S' \) does not occur elsewhere.

\( \text{language}(G_n^k) \) is the required language since

\[
\text{language}(G_n^k) = \{\gamma \mid S^* \Rightarrow \gamma\}
\]

\[
= \{\gamma \mid S \Rightarrow S_\alpha [k S_\beta k]^* \Rightarrow \gamma\}
\]

Since \( V_\alpha \) and \( V_\beta \) have no nonterminals in common, the derivation of \( S_\alpha \) must be independent of the derivation of \( S_\beta \). Therefore,

\[
\text{language}(G_n^k) = \{\gamma_1 \mid S_\alpha^* \Rightarrow \gamma_1\}[k \{\gamma_2 \mid S_\beta^* \Rightarrow \gamma_2\}]_k
\]

\[
= \text{language}_n^{k+1}(\alpha)[k \text{language}_n^{k-1}(\beta) \_k]
\]
6. If $B = \alpha[1 \beta 1]$, then there are grammars such that

$$\text{language}((Z^\alpha,V^\alpha,P^\alpha,S^\alpha)_n^2) = \text{language}((\alpha)$$

and

$$\text{language}((Z^\beta,V^\beta,P^\beta,S^\beta)_n^2) = \text{language}((\beta)$$

where $V^\alpha \cap V^\beta = \emptyset$. Construct the new grammar $G_n^1$ such that

1. $\Sigma' = \Sigma^\alpha \cup \Sigma^\beta$.
2. $V_i' = V_i^\alpha \cup V_i^\beta$ if $i \neq 1$ and
   $$V_1' = V_1^\alpha \cup V_1^\beta \cup \{S'\}.$$
3. $P' = P^\alpha \cup P^\beta \cup \{S' \rightarrow S^\alpha[1 S^\beta 1]\}$.
4. $S'$ does not occur elsewhere.

$\text{language}(G_n^k)$ is the required language since

$$\text{language}(G_n^1) = \{y \mid S^* \Rightarrow y\}$$

$$= \{y \mid S \Rightarrow S^\alpha[1 S^\beta k]^k \Rightarrow y\}$$

Since $V^\alpha$ and $V^\beta$ have no nonterminals in common, the derivation of $S^\alpha$ must be independent of the derivation of $S^\beta$. Therefore,

$$\text{language}(G_n^1) = \{y_1 \mid S^\alpha \Rightarrow y_1\}[k \{y_2 \mid S^\beta \Rightarrow y_2\}_k]$$
7. If \( n > k > 0 \) and \( B = \alpha + \beta \), then there are two grammars such that

\[
\text{language}(\Sigma, V, P, S) = \text{language}(B)
\]

Choose these two grammars such that

\[ V_\alpha \cap V_\beta = \emptyset \]

Construct the new grammar \( G_n^k \) such that

1. \( \Sigma' = \Sigma_\alpha \cup \Sigma_\beta \).
2. \( V_i' = V_{i\alpha} \cup V_{i\beta} \) if \( i \neq k \) and
   \( V_k' = V_{k\alpha} \cup V_{k\beta} \cup S \).
3. \( P' = P_\alpha \cup P_\beta \cup \{ S' \rightarrow S_\alpha, S' \rightarrow S_\beta \} \)
4. \( S' \) does not occur elsewhere.

The language of \( G_n^k \) is given by

\[
\text{language}(G_n^k) = \{ \gamma \mid S \Rightarrow^* \gamma \}
\]

\[ = \{ \gamma_1 \mid S \Rightarrow S_\alpha \Rightarrow^* \gamma_1 \} \cup \{ \gamma_2 \mid S \Rightarrow S_\beta \Rightarrow^* \gamma_2 \} \]

\[ = \text{language}_n^k(\alpha) \cup \text{language}_n^k(\beta) \]
8. If \( n \geq k > 0 \) and \( B = \alpha^*v \), then there is a grammar such that

\[
\text{language}\left( (\Sigma_\alpha^k, V_\alpha^k, P_\alpha^k, S_\alpha^k) \right) = \text{language}^k(\alpha)
\]

Construct the new grammar \( G^k_n \) such that

1. \( I' = I_\alpha \).
2. \( V_i' = V_i \alpha \) if \( i \neq k \) and \( V_k' = V_k \alpha \cup \{v\} \).
3. \( P' = P_\alpha \cup \{ v \to S_\alpha, S' \to v \} \).
4. \( S' \) does not occur elsewhere.

Notice that \( v \in V_k' \), therefore, it must occur in the right spots to be legitimately a member of \( V_k' \).

\[
\text{language}(G^k_n) = \{ \gamma \mid S' \Rightarrow \gamma \}
\]

\[
= \{ \gamma \mid S' \Rightarrow v \Rightarrow \gamma_1 \Rightarrow \gamma \text{ where } \gamma_1 \in \text{language}^k_n(\alpha) \}
\]

\[
= \{ \gamma_1 | v \Rightarrow \gamma_1 \} \ast_v (\{ \gamma_2 | v \Rightarrow \gamma_2 \} \cup \{ \gamma_1 | v^* \Rightarrow \gamma_1 \})
\]

\[
= \text{language}^k_n(\alpha) \ast_v \text{language}(G^k_n) \cup \text{language}^k_n(\alpha)
\]

\[
= \text{language}^k_n(\alpha) + v
\]

\[
= \text{language}^k_n(\alpha^*v)
\]

\[
= \text{language}(B)
\]
9. If \( n \geq k > 0 \) and \( B = \alpha \ast \nu \beta \) where \( B \in R_{n,m}^j \), then there are two grammars such that

\[
\text{language}(\Sigma_{\alpha}, V_{\alpha}, P_{\alpha}, S_{\alpha})^k = \text{language}^k_n(\alpha),
\]

and

\[
\text{language}(\Sigma_{\beta}, V_{\beta}, P_{\beta}, S_{\beta})^j = \text{language}^j_n(\beta)
\]

by induction hypothesis. Construct these such that

\[
V_{\alpha} \cap V_{\beta} = \emptyset
\]

Construct the new grammar \( G^k_n \) such that

1. \( \Sigma' = (\Sigma_{\alpha} - \{v\}) \cup \Sigma_{\beta} \).
2. \( V_i' = V_{i\alpha} \cup V_{i\beta} \) for all \( i \).
3. \( P' = \{ v' \mid v' \in P_{\alpha} \text{ and } v' = \gamma \} \) with all \( v' \)'s
   replaced by \( S_{\beta} \} \cup P_{\beta} \).
4. \( S' = S_{\alpha} \).

Now

\[
\text{language}(G^k_n) = \{ \gamma \mid S_{\alpha} \ast \Rightarrow \gamma \}
\]

\[
= \{ \gamma \mid S_{\alpha} \ast \Rightarrow \gamma_1 \ast \Rightarrow \gamma \}
\]

where \( \gamma_1 \in \text{language}(\alpha) \) with all \( v' \)'s replaced by \( S_{\beta} \).

\[
\text{language}(G^k_n) = \{ \gamma_1 \mid S_{\alpha} \Rightarrow \gamma_1 \} \ast_{S_{\beta}} \{ \gamma_2 \mid S_{\beta} \ast \Rightarrow \gamma_2 \}
\]

\[
= X \ast_{S_{\beta}} \text{language}^j_n(\beta)
\]
where \( X \) is the set language\(_n^k(\alpha) \) with all occurrences of \( v \) on the frontier replaced by \( S_\beta \). But in the concatenation given above, the resulting language would be the same if the \( S_\beta \) were replaced by \( v \) so

\[
\text{language}(G_n^k) = \text{language}_n^k(\alpha) \ast_v \text{language}_n^j(\beta)
\]

\[
= \text{language}_n^k(\alpha \ast_v \beta)
\]

\[
= \text{language}(B)
\]

Since assuming the induction hypothesis in each case allowed the proof that there is a \( G_n^k \) such that

\[
\text{language}(G_n^k) = \text{language}(B)
\]

when \( B \in R_{n,m}^k \), the lemma is proved by induction.

---

From the preceding chapters, it can be seen that there is a type of language called a regular language, and that these language may be described in any one of three ways, using grammars, automata or regular expressions. In the following chapters, these languages will be characterized, and also the related languages of the algebraic language hierarchy.
CHARACTERIZATION OF THE ALGEBRAIC LANGUAGE HIERARCHY

In this chapter, various characteristics of the hypertree hierarchy and their influence on the algebraic language hierarchy will be explored. These will include some closure properties for the hypertree hierarchy -- closure under union and intersection -- as well as a pumping lemma. These will imply closure under union for the algebraic language hierarchy, however it will be shown that this does not imply closure under intersection. Closure under intersection with regular sets is also shown. It will also be shown that not all context sensitive languages are in the algebraic language hierarchy. Lastly, it will be shown that the algebraic language hierarchy does, in fact, start out as designed: level 1 is the set of regular language, level 2 is the set of context free languages and level 3 is the set of macro languages.

Before this discussion can begin it is necessary to give some rather obvious formal definitions. These definitions will allow a formal definition of the algebraic language hierarchy just as was done for the hypertree hierarchy in chapter 2.
Definition 53:
If $G^k_n$ is a hypertree grammar, then the extracted language of level $k'$, denoted $\text{extract}_{k'}(G^k_n)$, is the set

$$\text{extract}_{k'}(G^k_n) = \{ \gamma \mid \gamma \in \text{language}(G^k_n) \text{ and } \gamma = \text{extract}_{k'}(\gamma') \}$$

also, if $A^k_n$ is a hypertree automaton the extracted language of level $k'$, denoted $\text{extract}_{k'}(A^k_n)$, is the set

$$\text{extract}_{k'}(A^k_n) = \{ \gamma \mid \gamma \in \text{language}(A^k_n) \text{ and } \gamma = \text{extract}_{k'}(\gamma') \}$$

and finally, if $R^k_n$ is a normal regular hypertree expression then the extracted language of level $k'$, denoted $\text{extract}_{k'}(R^k_n)$, is the set

$$\text{extract}_{k'}(R^k_n) = \{ \gamma \mid \gamma \in \text{language}(R^k_n) \text{ and } \gamma = \text{extract}_{k'}(\gamma') \}$$

Definition 54:
The string language for a hypertree grammar, automata or regular expression, denoted $\text{string}(A^k_n)$ or $\text{string}(R^k_n)$ respectively, is another name for the $\text{extract}_1$ function.
Definition 55:

If there exists a regular hypertree grammar $G_n^{k'}$ such that

$$\gamma = \text{extract}_k(G_n^{k'})$$

then $\gamma$ is a member of level $n$ and degree $k$ of the algebraic language hierarchy, denoted $ALH_n^k$. These are all the elements of $ALH_n^k$.

Corollary 58:

$\gamma \in ALH_n^k$

If and only if there exists a hypertree automata $A_n^{k'}$, a normal hypertree regular expressions $R_n^{k'}$, as well as a hypertree grammar $G_n^{k'}$ such that

$$\gamma = \text{extract}_k(A_n^{k'})$$
$$\gamma = \text{extract}_k(R_n^{k'})$$

and

$$\gamma = \text{extract}_k(G_n^{k'})$$

Proof:

This follows directly from the definition.

---

The first set of theorems that will be presented are the closure theorems. A pair of closure theorems will be offered for level $n$ of the hypertree hierarchy.
Theorem 59:
If \( A_n^k \) and \( B_n^k \) are regular hypertree languages, then there is a regular hypertree language \( C_n^k \) such that
\[
C_n^k = A_n^k \cup B_n^k
\]
Proof:
This follows directly from the definition of regular sets.

Theorem 60:
If \( A_n^k \) is a regular hypertree language and \( B_n^k \) is another regular hypertree language, then there is a third regular hypertree language \( C_n^k \) such that
\[
C_n^k = A_n^k \cap B_n^k
\]
Proof:
Construct a normal grammar for \( A_n^k \) and \( B_n^k \). Then, in each case, eliminate productions of the form \( A \rightarrow B \) where \( A \in V_i \) by forward substitution. Lastly, for each \( X_i^j \), add a production of the form \( A \rightarrow X_i^j \) where \( A \) is a new nonterminal, and substitute \( A \) in for \( X_i^j \) wherever it occurs. Note that \( A \in V_i \) in this case. Call the resulting grammars \( G_n^k = (\Sigma, V, P, S_n^k) \) and \( G_n^k = (\Sigma', V', P', S') \) respectively.
Then, construct a new grammar \( G_n^{k''} = (\Sigma'', V'', P'', S'') \) such that
1. \( \Sigma'' = \Sigma \cap \Sigma' \)
2. \( V_k'' = V_k \times V_k' \)
3. The construction for \( P'' \) is given below.
4. \( S'' = (S, S') \)

\( P'' \) is constructed as follows:

1. If \( A \xrightarrow{a} P \) and \( B \xrightarrow{a} P \) where \( a \in \Sigma \cap \Sigma' \), then
   
   \((A, B) \rightarrow a \in P''\)

2. If \( A \rightarrow B \in P \) and \( C \rightarrow D \in P' \) where \( A \in V_i, B \in V_{i+1}, C \in V_i' \) and \( D \in V_{i+1}' \), then
   
   \((A, C) \rightarrow (B, D) \in P''\)

3. If \( A \xrightarrow{X_i^j} P \) and \( B \xrightarrow{X_i^j} P' \), then
   
   \((A, B) \rightarrow X_i^j \in P''\)

4. If \( A \rightarrow a[n_B, n_c] \in P \) and \( C \rightarrow a[n_D, n_c] \in P' \) where \( a \in \Sigma \cap \Sigma' \), then
   
   \((A, C) \rightarrow a[n(B, D), n_c] \in P''\)

5. If \( A \rightarrow B[i, C_i] \in P \) and \( D \rightarrow E[i, F_i] \in P' \), then
   
   \((A, D) \rightarrow (B, E)[i, (C, F), i] \in P''\)

for all \( n > i > 0 \).

6. These are all the elements of \( P'' \).

If \( \gamma \in \text{language}(G_n^{kn}) \), then pick any derivation for \( \gamma \) in \( G_n^{kn} \).

It is easy to show by induction that a derivation for \( \gamma \) in \( G_n^{kn} \) is obtained if each nonterminal in this derivation is
replaced by the corresponding nonterminal from $V$, that is the left hand projection of the nonterminal pair. A similar inductive proof using the right hand projection of the pair yields a derivation for $\gamma$ in $G_n^{k'}$. Therefore,

$$\text{language}(G_n^{k})^* \subseteq \text{language}(G_n^k) \cap \text{language}(G_n^{k'})$$

In order to prove containment the other way, a slightly stronger proof will be presented. Namely, if $A \in V_i$, $B \in V_i'$, $A \Rightarrow^* \gamma$ and $B \Rightarrow^* \gamma$ then

$$(A,B) \Rightarrow^* \gamma$$

This will be proved by induction on the length of the derivation.

Basis: If the length of the derivations is one, then

$$(A,B) \Rightarrow \gamma$$

by construction.

Induction step: Assume that this is true for all derivations of length less than $m$. If either of the derivations is of length $m$, where $m > 1$, then

1. If $\gamma \in H_n^{i+1}$, then

$$A \Rightarrow C \Rightarrow^* \gamma$$

and

$$B \Rightarrow D \Rightarrow^* \gamma$$
Since $C \Rightarrow^* \gamma$ and $D \Rightarrow^* \gamma$ are both of length less than $m$,

$$(C,D) \Rightarrow^* \gamma$$

But by construction $(A,B) \Rightarrow (C,D)$ so that

$$(A,B) \Rightarrow^* \gamma$$

2. $\gamma$ cannot be an $X_j^k$ since the derivation is longer than one.

3. If $\gamma = a[i \beta i]$ and $n > i > 0$, then

$$A \Rightarrow C[i D i] \Rightarrow^* a[i \beta i]$$

and

$$B \Rightarrow E[k F i] \Rightarrow^* a[i \beta i]$$

where $C \Rightarrow^* a$, $D \Rightarrow^* \beta$, $E \Rightarrow^* a$ and $F \Rightarrow^* \beta$.

Therefore, by induction hypothesis $(C,E) \Rightarrow^* a$ and $(D,F) \Rightarrow^* \beta$ and by construction

$$(A,B) \Rightarrow (C,E)[i(D,F)i]$$

Therefore,

$$(A,B) \Rightarrow^* \gamma$$

and the hypothesis is true in this case also.

4. If $\gamma = a[n \beta n]$, then

$$A \Rightarrow a[n D n] \Rightarrow^* a[n \beta n]$$
and

\[ B \Rightarrow a_k F_i \Rightarrow a_i \beta_i \]

where \( D \Rightarrow \beta \) and \( F \Rightarrow \beta \). Therefore, by induction hypothesis \((D,F) \Rightarrow \beta \) and by construction

\[ (A,B) \Rightarrow a_i (D,F)_i \]

Therefore,

\[ (A,B) \Rightarrow \gamma \]

and the hypothesis is true in this case also.

Therefore, since these are the only possible cases, the hypothesis is proved by induction.

If \( \gamma \in \text{language}(G^k_n) \) and \( \gamma \in \text{language}(G^k_n)' \), then \( S \Rightarrow \gamma \) and \( S' \Rightarrow \gamma \), therefore, by the conjecture,

\[ (S,S') \Rightarrow \gamma \]

and \( \gamma \in \text{language}(G^k_n)'' \), so the theorem is proved.

\[ - \cdot - \]

A corollary to theorem 59 is

Corollary 61:

If \( L_1 \) and \( L_2 \) are two languages from \( ALH^k_n \) then there is a third language \( L_3 \) such that \( L_3 \in ALH^k_n \) and

\[ L_3 = L_1 \cup L_2 \]
Proof:
Construct the union of the corresponding hypertree grammar. This gives \( L_3 \) as its extracted language.

A similar corollary to theorem 60 is not given because it is not true. Consider the languages

\[ \{a^n b^n c^m\}_{n \geq 1, m \geq 1} \]

and

\[ \{a^m b^n c^n\}_{n \geq 1, m \geq 1} \]

Both of these are context free and, therefore, in \( \text{ALH}^1 \). However, the intersection of these two languages is the language

\[ \{a^n b^n c^n\}_{n \geq 1} \]

which is not context free. A proof of closure under intersection with regular sets will now be presented.

Theorem 62:
If \( L \in \text{ALH}^k_n \) and \( M \in \text{ALH}^k_{n'} \), then
\[ L \cap M \in \text{ALH}^k_n \]

Proof:
Construct a normal grammar $G_n^L = (\Sigma, V, P, S)_n^L$ for $L$ and another grammar $G_k^M = (\Sigma', V', P', S')_k^M$ for $M$. In each case, construct the grammar such that Xs only appear in productions of the form $A \rightarrow X^j_k$ and there are no productions of the form $A \rightarrow B$ where $A$ and $B$ are in $V_i$ for some $i$. Construct $G_n^{L''} = (\Sigma'', V'', P'', S'')_n^{L''}$ as follows:

1. $\Sigma'' = \Sigma \cap \Sigma'$.
2. $V_i'' = V_i \times V_i' \times X \times N \times P(V \times V' \times X \times N)$ for all $k \geq i > 0$ where $V$ is the set of all the Xs used in $G_n^L$ and $N$ is the set $\{1\ldots n\}$.
3. $V_i'' = V_i \times V_k \times X \times N \times P(V \times V' \times X \times N)$ for all $n \geq i > k$.
4. The construction of $P''$ is given below.
5. $S''$ does not occur elsewhere.

Construct $P''$ as follows:

1. If $A \rightarrow a \in P$ and $A' \rightarrow a \in P'$ where $a \in \Sigma \cap \Sigma'$, then

$$(A, A', n, a) \rightarrow a \in P''$$

for all $a \in V \times V' \times X \times N$.

2. If $A \rightarrow X^j_k \in P$ and $A' \rightarrow X^j_{k'} \in P'$ (Note that this implies that $k' \leq k$), then

$$(A, A', k', a) \rightarrow X^j_{k'} \in P''$$

for all $a \in V \times V' \times X \times N$. 

3. If $A \rightarrow X^j_k \in P$ and $k' \geq k$, then

$$(A, X, I, \alpha) \rightarrow X^j_{k'} \in P''$$

for all $X \in V^{k'}$ and $\alpha \in V \times V' \times N$ such that

$$(X^j_{k'}, X, I) \in \alpha.$$  

4. If $A \rightarrow B \in P$ and $A' \rightarrow B' \in P'$ where $A \in V_i$, $B \in V_{i+1}$,

$A' \in V_{i'}$, and $B' \in V_{i'+1}$ for some $k \geq i > 0$, then

$$(A, A', I, \alpha) \rightarrow (B, B', I, \alpha) \in P''$$

for all $\alpha \in V \times V \times N$ and $I \in N$.

5. If $A \rightarrow B \in P$ where $A \in V_i$ and $B \in V_{i+1}$ for some $n \geq i > k$, then

$$(A, X, I, \alpha) \rightarrow (B, X, I, \alpha) \in P''$$

for all $X \in V^{k'}$, $\alpha \in V \times V' \times N$ and $I \in N$.

6. If $A \rightarrow a[n, B^n] \in P$ and $A' \rightarrow a[n, B'^n]$ where

$a \in \Sigma^n I'$, $B \in V_i$, and $B' \in V_{i-1}$', then

$$(A, A', I, \alpha) \rightarrow a[n (B, B', I, \alpha) n] \in P''$$

for all $n > I > 0$. Also,

$$(A, A', n-1, \alpha) \rightarrow a[n (B, B', n, \alpha) n] \in P''$$

for all $\alpha \in V \times V' \times N$. Note that in this case $n = k$. 

7. If $A \rightarrow a_n B_n \in P$ and $n > k$, then

$$(A, X, I, \alpha) \rightarrow a_n (B, X, I, \alpha) \in P^n$$

for all $n > I > 1$. Also,

$$(A, X, n-1, \alpha) \rightarrow a_n (B, X, n, \alpha) \in P^n$$

for all $X \in V_k'$ and $\alpha \in V \times V' \times N$.

8. If $A \rightarrow B[k', C_k'] \in P$ and $A' \rightarrow B'[k', C', k'] \in P'$, then

$$(A, A', I, \alpha) \rightarrow (B, B', J, \alpha)[k', (C, C', K, \alpha)_{k'}] \in P^n$$

for all $J \in N, K \in N$ and $\alpha \in V \times V' \times N$ such that

$I = \min(k', K, J)$.

9. If $A \rightarrow B[k', C_k'] \in P$ and $n > k' > k$, then

$$(A, X, I, \alpha_A) \rightarrow (B, X, I, \alpha_B)[k', E_{k'-1}(C, P_C, \alpha_C)_{k'}] \in P^n$$

for all $k' > I > 0, X \in V_k'$ and $\alpha_A, \alpha_B$ and $\alpha_C$ are subsets of $V \times V' \times N$ such that

$$\alpha_B = (\alpha_A \times X_k', X V' \times N) \cup \{(X^j_k', C, I) \mid (j, C, I) \in P_C\}$$

$$\alpha_C = \alpha_A$$

$$P_C = \{j \mid X^j_1 \in V\} \times V' \times N$$

The function $E$ which returns a set is defined below.

$$(A, X, k', \alpha) \rightarrow (B, X, I, \alpha)[k', (C, X, J, \alpha)_{k'}] \in P^n$$
for all $\alpha \in V \times V' \times N$, $n \geq I > k'$ and $n \geq J > k'$.

$$(A,X,J,\alpha)\rightarrow(B,X,I,\alpha)[k'] (C,X,J,\alpha) [k'] \in P''$$

for all $\alpha \in V \times V' \times N$, $n \geq I \geq k'$ and $k' \geq J > 0$.

10. $S''\rightarrow(S,S',I,\emptyset) \in P''$ for all $I \in N$.

11. These are all the elements of $P''$.

$E$ is a set of finite functions $\{E'_{k''}\}_{n \geq k'' > 0}$ such that $E'_{k''}(A,p,\alpha)$ is the smallest set such that

1. If $A\rightarrow B[k'' C k''] \in P$, then

$$E''_{k''+1}(B,p_B,\alpha)[k'' E'_{k''-1}(C,p_C,\alpha) [k''] \subset E''_{k''}(A,p,\alpha)$$

for all $\alpha \in V \times V' \times N$, $p_B$ and $p_C$ such that

$$p = p_B \cup \{(k''j,C',I) | (j,C,I) \in P_C\}$$

$$p_B \subset P_{k''+1} \times V \times N$$

and

$$p_C \subset P_{k''-1} \times V \times N$$

such that if $(j,C,I) \in p_B \cup p_C$ then $j$ is a suffix of a path in $V \cap X_1$.

2. If $A \rightarrow B \in P$, then

$$E''_{k''+1}(B,p,\alpha) \subset E''_{k''}(A,p,\alpha)$$

where $A \in V_{k''}$ and $B \in V_{k''+1}$ for all $p \subset P_{k''+1} \times V \times N$ such that if $(j,C,I)$ is in $p$ then $j$ is the suffix of a path in $V \cap X_1$. 
3. If $C \in V_{k''}$, then

$$E_{k''}(C, \phi, \alpha) = \{(C, X, I, \alpha) : X \in V' \text{ and } I \in N\}$$

and

$$E_{k''}(C, \{(\lambda, C', I')\}, \alpha) = \{(C, C', I', \alpha)\}$$

for all $C' \in V_{k''}$ if $k'' \leq k$ or $C' \in V_k'$ if $n \geq k'' > k$,

all $I' \in N$ and $\alpha \subseteq V \times V' \times X \times N$.

Note that if $(A, A', I, \alpha) \in V''$ and $(A, A', I, \alpha) \Rightarrow \beta$ then

$A \Rightarrow \beta$ in $G^1_n$ and $A' \Rightarrow \text{extract}_k(\beta')$ where $\beta'$ is the same as $\beta$ except that each $X^j_i$ that references a subhypertree that is not contained in $\beta$ is replaced by $C'$ where $(X^j_i, C', I) \in \alpha$.

Also, due to the construction if $(X^j_i, C', I) \in \alpha$ then during the frontiering operation $X^j_i$ will be replaced with a hypertree that was generated from a nonterminal $(C, C', I, \alpha)$ for some $C$ and $\alpha$. Also, the $I$ represents the level of the first frontiering operation for $\beta'$ that will return a single element of $I$. Each of these facts can be established by simple induction and, therefore, $G^R_n$ is a grammar for $L \cap M$ and the theorem is proved.

---

This last theorem suggests another normal form which may be used to simplify some proofs. This normal form is called "completed" since no $X$ will ever copy just a piece of the subhypertree generated by a given nonterminal. That is, each subhypertree remains "complete".
Definition 56:
A grammar $G_n^k$ is said to be a completed grammar if during any extract operation on any element of $\text{language}(G_n^k)$ a piece of the structure generated by any nonterminal is copied then all of the structure generated from that nonterminal is copied.

The following theorem proves this to be a useful concept.

Theorem 63:
If $G_n^k = (\mathcal{L}, V, P, S)_n^k$ is a regular grammar, then there is a completed grammar $G_n^{k'} = (\mathcal{L}', V', P', S')_n^k$ such that $\text{language}(G_n^k) = \text{language}(G_n^{k'})$

Proof:
Construct $G_n^k$ such that it is a normal grammar and there are no productions of the form $A \rightarrow B$ where $A$ and $B$ are in $V_i$ for some $i$. Also, construct $G_n^k$ such that all $X$s are in productions of the form $A \rightarrow X_j^i$. Construct $G_n^{k'}$ as follows:
(Note that this construction is quite similar to that used in theorem 62.)

1. $\mathcal{L}' = \mathcal{L}$.
2. $V_i' = V_i \times N \times P(\mathcal{V} \times N)$ where $\mathcal{V}$ is the set of all $X$s that are used in $G_n^k$ and $N$ is the set $\{1...n\}$.

Also, $S' \in V_k$.
3. The construction of $P'$ is given below.
4. $S'$ does not occur elsewhere.
Construct $P'$ as follows:

1. If $A \rightarrow a \in P$ where $a \in \Sigma$, then

$$(A, n, a) \rightarrow a \in P'$$

for all $a \in \mathcal{V} \times \mathbb{N}$.

2. If $A \rightarrow X^j_k \in P$, then

$$(A, I, a) \rightarrow X^j_k \in P'$$

for all $a \in \mathcal{V}$ such that $(X^j_k, I) \in a$.

3. If $A \rightarrow B \in P$, then

$$(A, I, a) \rightarrow (B, I, a) \in P'$$

for all $a \in \mathcal{V} \times \mathbb{N}$.

4. If $A \rightarrow a[B \in P$, then

$$(A, I, a) \rightarrow a[B \in P'$$

for all $n > I > 0$ and $a \in \mathcal{V} \times \mathbb{N}$. Also,

$$(A, n-1, a) \rightarrow a[B \in P'$$

5. If $A \rightarrow B[k' \in C \in P$ and $n > k' > 0$, then

$$(A, I, a) \rightarrow (B, I, a)[k' \in E_{k'-1}(C, P_C, a_C) \in P'$$

for all $k' > I > 0$, $a_A$, $a_B$, and $a_C \in \mathcal{V} \times \mathbb{N}$ such that
\( a_B = (a_A - X_{k'} \times N) \cup \{(X^j_{k'}, I) \mid (j, I) \in P_C\} \)

\( a_C = a_A \)

\( P_C = \{j \mid X^j_{k'} \in V\} \times N \)

The function \( E \) is defined below.

\[(A, k', \alpha) \rightarrow (B, I, \alpha)[k', (C, J, \alpha) \ k] \in P' \]

for all \( \alpha \in V \times N \), \( n \geq I \geq k' \) and \( n \geq J > k' \).

\[(A, J, \alpha) \rightarrow (B, I, \alpha)[k', (C, J, \alpha) \ k] \in P' \]

for all \( \alpha \in V \times N \), \( n \geq I \geq k' \) and \( k' \geq J > 0 \).

6. \( S' \rightarrow (S, I, V) \in P' \) for all \( I \in N \).

7. These are all the elements of \( P' \).

\( E \) is a set of finite functions \( \{E_{k''}\}_{n \geq k'' > 0} \) such that \( E_{k''}(A, p, \alpha) \) is the smallest set such that

1. If \( A \rightarrow B[k'' C] \in P \), then

\[ E_{k'' + 1}(B, P_B, \alpha)[k'' E_{k'' - 1}(C, P_C, \alpha) \ k''] \subset E_{k''}(A, p, \alpha) \]

for all \( \alpha \), \( P_B \) and \( P_C \) such that

\[ p = P_B \cup \{(k'' j, I) \mid (j, I) \in P_C\} \]

\[ P_B \subset P_i^{k'' + 1} \times N \]

and

\[ P_C \subset P_i^{k'' - 1} \times N \]
such that if \((j, I) \in P_B \cup P_C\) then \(j\) is the suffix of a path in \(V \cap X_i\).

2. If \(A \rightarrow B \in P\), then

\[ E_{k''+1}(B, p, \alpha) \subseteq E_{k''}(A, p, \alpha) \]

where \(A \in V_{k''}\) and \(B \in V_{k''+1}\) for all \(p \in P_{i}^{k''+1} \times N\) such that if \((j, I) \in p\) then \(j\) is the suffix of a path in \(V \cap X_i\) and \(\alpha \in V \times N\).

3. If \(C \in V_{k''}\), then

\[ E_{k''}(C, \beta, \alpha) = \{(C, I, \alpha)\} \]

for all \(\alpha \in V \times N\) and \(\beta \subseteq \{\lambda\} \cup \{j \mid j \in P_{i}^{k''}, j \notin P_{i}^{k''+1}\} \times \{I\}\).

Notice that if \(\beta = \emptyset\) then the corresponding structure is dropped when the \(i^{th}\) level frontier is taken. If \(\beta = \{\lambda\}\), then the structure is copied when the \(i^{th}\) level frontier is taken. If \(\beta \neq \emptyset\) and \(\beta \neq \{\lambda\}\), then the \(i^{th}\) frontier is possibly not defined since this corresponds to possibly having an \(X\) which doesn't reference anything. As in the proof for theorem 62, note that \((A, I, \alpha) \Rightarrow \beta\) if and only if \(A \Rightarrow \beta\) and the set of \(X\)s which reference anything outside of \(\beta\) is a subset of \(\alpha\). The number associated with the \(X\) is the level of frontier which first produces a singleton.
Also, the number associated with the nonterminal is the level of frontier that first transforms the resulting structure into a singleton. These facts can be established by induction on the length of the production. Therefore, $S \Rightarrow^* \beta$ if and only if $(S, I, V) \Rightarrow^* \beta$ for some $I$ so that

$$\text{language}(G_n^k) = \text{language}(G_n^{k'})$$

It could also be established that $G_n^{k'}$ is complete since the $E$ functions force nonterminals to be expanded until all the paths that are possible are included in one production. Therefore, the theorem is proved.

---

Corollary 64:

If $L \in ALH_n^k$, then there is a grammar

$$G_n^{k'} = (\Sigma', V', P', S')$$

for $L$ such that if $a \in \text{language}(G_n^k)$ then $\text{extract}_k(a)$ exists.

Proof:

Construct the completed grammar for $L$ as in theorem 63 except replace step 3 of the construction of the $E$ function with

3. If $C \in V_n^k$ and $k > k'' > 0$, then

$$E_{k''}(C, \beta, a) = \emptyset$$
for all $\alpha \in V \times N$ and $\beta \in \{(\lambda) \cup \{j \mid j \in P_i^{k''+1}\}$
and $j$ is the suffix of a path in $V \cap X_i)$ $\times \{I\}$ where $\beta \neq \emptyset$ and $\beta \not\in \{(\lambda) \times N$.

4. If $C \in V_{k''}$ and $n \geq k'' \geq k$, then

$$E_{k''}(C, \emptyset, \alpha) = \{(C, I, \alpha) \mid I \in N\}$$

and

$$E_{k''}(C, \{(\lambda), I\}, \alpha) = \{(C, I, \alpha)\}$$

The meaning of the ordered triples is exactly the same as in the theorem. If an $X$ were generated in the original grammar which would force the desired extract operation to be undefined, then the element could not be generated by $G_{n}^{k'}$, since ultimately a call to $E_{k''}(C, \beta, \alpha)$ would be made in the construction of $G_{n}^{k'}$ to generate the required productions, but this is equal to $\emptyset$ so the needed productions would not be generated.

---

Theorem 65:

For any language, $L \in ALH_{n}^{k'}$, there is a grammar $G_{n}^{k'} = (V', P', S')$ such that $V_{i} = \emptyset$ for all $n > i \geq k$ and

$$L = \text{extract}_k(G_{n}^{k'})$$

Proof:
Construct a completed grammar $G_n^k = (\Sigma, V, P, S)_n^k$ for $L$.

Construct $G_n^{k'}$ as follows:

1. $\Sigma' = \Sigma \cup \{\ast\}$.

2. $V_n' = \bigcup_{n \geq i \geq k} V_i$, $V_i' = \emptyset$ if $n > i \geq k$ and $V_i' = V_i$ if $k > i > 0$.

3. The construction for $P'$ is given below.

4. $S' = S$.

Construct $P'$ as follows:

1. If $\alpha \to \alpha \in P$ and $\alpha \in V_i$ where $k > i > 0$ or $i = n$, then

   $\alpha \to \alpha \in P'$

2. If $\alpha \to \alpha \in P$ and $\alpha \in V_i$ where $n > i \geq k$, then

   $\alpha \to \ast[n \ast[n-1 \ast \cdots \ast[i+1 \alpha \ast i+1] \ast \cdots \ast n-1 \ast] \ast] \in P'$

A simple inductive proof would show that

$$L = \text{extract}_k(G_n^{k'})$$

---

Corollary 66:

If $L \in \text{ALH}_n^1$, then there is a grammar

$$G_n^{k'} = (\Sigma, V, P, S)_n^k$$

such that $V_i = \emptyset$ if $i \neq n$ and

$$L = \text{extract}_1(G_n^{k'})$$

Proof:

Substitute 1 into theorem 65.
Corollary 57:
If $L \in \text{ALH}_n^k$ and $G_n^{k'} = (\Sigma, V, P, S)_n^{k'}$ is a grammar for $L$ where $k' > k$, then there is a grammar $G_n^{k'} = (\Sigma', V', P', S')_n^{k'}$ such that

$L = \text{extract}_{k'}(G_n^{k'})$

Proof:
This follows directly from the theorem.

---

Theorem 68:
If $L \in \text{ALH}_n^k$ and $G_n^{k'} = (\Sigma, V, P, S)_n^{k'}$ is a grammar for $L$, then there is a grammar $G_n^n = (\Sigma', V', P', S')_n^n$ such that

$* [k' * [k-1 ... *[k'' L_{k''}]...k-1] k] = \text{extract}_k(G_n^n)$
where $k' + 1 \geq k'' \geq 1$.

Proof:
Construct $G_n^n$ as follows:
1. $\Sigma' = \Sigma \cup \{*\}$.
2. $V_i' = V_i$ if $n > i > 0$. $V_n' = V_n \cup \{S\}$.
3. $P' = P \cup \{S' \rightarrow *[n ... *[k'' S_{k''}]...n]\}.$
4. $S'$ does not occur elsewhere.

A simple induction over $k$ will prove the theorem.

---
Note that these theorem and corollaries allow $L$ to be treated as though there were a grammar $G_n^k$ for it if $L \in \text{ALH}_n^k$ in most cases.

The following theorem will show that in fact it is not necessary to consider all the possible $X$s which might occur in a grammar. But first it is necessary to define a subset of the $X$s that will be proved to be sufficient in most cases.

Definition 57:

The set of linear paths of level $n$, denoted $LP_n$, is defined as

$$LP_n = \{(n-1)\}^*$$

Definition 58:

The set of linear $X$s of level $n$, denoted $LX_n$, is defined as

$$LX_n = \{X_n^j | j \in LP_n\}$$

Definition 59:

A grammar $G_n^k$ is said to be linear to level $i$ if $X_i^j$, appears in $G_n^k$ and $i' > i$ implies that $X_i^j \in LX_i$. 


Definition 60:
A grammar $G_n^k$ is said to be linear if it is linear to level 1.

Theorem 69:
If $L \in \text{ALH}_n^k$, then there is a linear grammar to level $k+1$, $G_n^{k'} = (\Sigma, V, P, S)^{k'}_n$, such that

$L = \text{extract}_{k}(G_n^{k'})$

Proof:
Construct a completed grammar $G_n^{k'} = (\Sigma', V', P', S')^{k'}_n$ for $L$ as was done in theorem 63. Construct $G_n^{k'}$ as follows:

1. $\Sigma = \Sigma' \cup \{\ast\}$.
2. $V = V'$.
3. The construction for $P$ is given below.
4. $S = S'$.

Order the $X$s at each level greater than $k$. Call the first $X_{k''}(0)$, the second $X_{k''}(1)$, and so forth for each $n \geq k'' > k$.

Construct $P$ as follows:

1. If $A \rightarrow B \in P'$ where $B \in V' \cup \Sigma$, then
   
   $A \rightarrow B \in P$

2. If $A \rightarrow a[n \alpha_n] \in P'$ where $a \in \Sigma$, then
   
   $A \rightarrow a[n \alpha_n] \in P$
3. If $A \rightarrow B[i \alpha i] \in P'$ and $n \geq i > k$ where $B \in \mathcal{V}_{i+1}'$, $B = (X,I,\beta)$ and $i \geq I > 0$, then

$$A \rightarrow B[i C_0 C_{i-1} C_{i-1} \cdots C_{m-1} i-1] i \in P$$

where $C_j = *$ if $(X_i(j),I)$ is not in $\beta$ for any $I$, and if $(X_i(j),I)$ is in $\beta$ then $C_i = \alpha$ is the nonterminal from $\alpha$ that is lead to by the path in $X_i(j)$. If $n \geq I > i$, then

$$A \rightarrow* [i \alpha i] \in P$$

4. If $A \rightarrow B[i \alpha i] \in P'$ and $k \geq i > 0$, then

$$A \rightarrow B[i \alpha i] \in P$$

5. If $A \rightarrow X^j_i \in P'$, then

$$A \rightarrow X^{(i-1)m}_i \in P$$

where $X^j_i = X_i(m)$.

Simple induction would show that these two grammars define the same language.

---

**Corollary 70:**

There is a linear grammar for all string languages in the algebraic language hierarchy.

**Proof:**
Substitute 1 for \( k' \) in theorem 69. Note that \( X_1^\lambda \) is linear and the only \( X \) that can appear in strings.

---

**Corollary 71:**

There is a linear grammar for all tree languages in the algebraic language hierarchy.

**Proof:**

Substitute 2 for \( k \) in theorem 69. Note that only \( X_1^\lambda \) and elements of \( X_2 \) can appear on trees. Both of these types of \( X \)s must be linear.

---

**Corollary 72:**

If \( L \in \text{ALH}^3 \) and \( \text{frontier}_2^0(L) \) exists for each element of the language \( L \), then there is a linear grammar for \( L \).

**Proof:**

Substitute 3 for \( k' \) in theorem 69. Note that if \( X_3^j \) is in \( \alpha \in L \) then \( \text{frontier}_2^0(L) \) is not defined since ultimately this will require that \( \text{frontier}_2^3(X_3^j) \) be evaluated and this frontier is not defined.

---

Previously it was mentioned that context free languages where in \( \text{ALH}^1_2 \). So far, this has not been established
formally, so this will be shown in one of the next three theorems. This correspondence is not direct, though, in that the hypertree languages have some extraneous symbols -- the Xs and the brackets. It will be assumed, therefore, that the brackets and level 1 Xs will be deleted from the sentences to arrive at the resulting language. Note that no level 2 or higher Xs can occur.

Theorem 73:

Subject to the above constraints, \( L \in ALH^1 \) if and only if there is a regular grammar for \( L \).

Proof:
Construct a normal regular hypertree grammar \( G^1 = (\Sigma, V, P, S)_1 \) for \( L \). Then, construct the regular grammar
\[ G' = (\Sigma', V', P', S') \] such that
1. \( \Sigma' = \Sigma \).
2. \( V' = V_1 \).
3. \( P' \) is given below.
4. \( S' = S \).

Construct \( P' \) as follows:
1. If \( A \rightarrow a \in P \) where \( A \in V_1 \) and \( a \in \Sigma \), then
   \[ A \rightarrow a \in P' \]
2. If \( A \rightarrow B \in P \), then
   \[ A \rightarrow B \in P' \]
3. If $A \rightarrow a_1 B_1 \in P$, then
   
   $A \rightarrow aB \in P'$

4. If $A \rightarrow a_1 X_1 \in P$, then
   
   $A \rightarrow a \in P'$

The correspondence is easy to show via induction on the length of the derivation. The proof of containment in the other direction is similar.

---

Theorem 74:

Subject to the above constraints, $L \in \text{ALH}_2^1$ if and only if there is a context free grammar for $L$.

Proof:

If $L$ is a context free language, construct a grammar

$G = (I,V,P,S)$ for language with productions of the form

- $A \rightarrow BC$
- $A \rightarrow a$

where $A$, $B$ and $C$ are nonterminals and "a" is a terminal.

Construct $G_2' = (\Sigma',V',P',S')_2$ such that

1. $\Sigma' = \Sigma U \{\ast\}$.
2. $V_2' = V$, $V_1' = \emptyset$.
3. $P'$ is given below.
4. $S' = S$. 

P' is constructed as follows:

1. If \( A \rightarrow BC \in P \), then

\[
A \rightarrow^* [2B_1C_1X_1^11_11_2] \in P'
\]

2. If \( A \rightarrow a \in P \), then

\[
A \rightarrow a \in P'
\]

A simple, though involved, induction will show these two grammars generate the same language.

If \( L \in \text{ALH}^1 \), then construct a normal grammar

\[
G_2 = (\Sigma, V, P, S)_2
\]

for language subject to the constraints that

- \( X \)s only appear in production of the form \( A \rightarrow X^j \)
- there are no productions of the form \( A \rightarrow B \) where \( A \) and \( B \) are either both from \( V_1 \) or both from \( V_2 \).

Construct

\[
G' = (\Sigma', V', P', S')
\]

as follows:

1. \( \Sigma' = \Sigma \).
2. \( V' = (V_1 \cup V_2) \times \{1,2,3\} \cup S' \).
3. \( P' \) is given below.
4. \( S' \) does not occur elsewhere.

Construct \( P' \) as follows:

1. If \( A \rightarrow a \in P \), then

\[
(A, 1) \rightarrow a \in P'
\]

2. If \( A \rightarrow a[2B_2] \in P \), then
(A,3)→(B,1) ∈ P'
(A,2)→(B,2) ∈ P'
(A,3)→(B,3) ∈ P'

3. If A→B ∈ P, then
   (A,1)→(B,1) ∈ P'
   (A,2)→(B,2) ∈ P'
   (A,3)→(B,3) ∈ P'

4. If A→X^1 ∈ P, then
   (A,2)→X ∈ P'

5. If A→B[1C1] ∈ P, then
   (A,3)→(B,1)(C,1) ∈ P'
   (A,2)→(B,1)(C,2) ∈ P'
   (A,3)→(B,1)(C,3) ∈ P'
   (A,3)→(B,2)(C,1) ∈ P'
   (A,2)→(B,2)(C,2) ∈ P'
   (A,3)→(B,2)(C,3) ∈ P'
   (A,3)→(B,3) ∈ P'

6. \{S'→(S,1), S'→(S,2), S'→(S,3)\} ∈ P'.

7. These are all the elements of P'.

Note that productions of the form A → X^1_j are dropped since if an X^1_j appears in the resulting hypertree the frontier is
not defined. It can be easily seen that the two grammars
define the same language if it is noted that (A, 1) *=> 'y if
and only if A *=> y and y is one symbol only. (A, 2) *=> y
if and only if A *=> y_1 and frontier^0(y_1) = yX_1^λ. Also,
(A, 3) *=> y if and only if A *=> y_1 and frontier^0(y_1) = y
(with no trailing X_1^λ). A simple inductive proof shows the
two languages are the same.

---

Theorem 75:

Subject to the above constraints, L ∈ ALH^1 if and only
if there is an IO macrogrammar for L.

Proof:
If L is an IO macrolanguage, then construct an IO
macrogrammar G = (I, F, V, ρ, S, P) where language(G) = L, I is
the set of terminals, F is the set of nonterminals, V is the
set of arguments, ρ is the ranking function for F, S is the
start symbol and P is the set of productions. Construct a
hypertree grammar G^3_3 = (E', V', P', S')^3 as follows:
1. I' = Ε U {*}
2. V_1' = φ, V_2' = φ and V_3' = F.
3. P' is given below.
4. S' = S.

For each production, α → β in P, there is one production in
P' constructed as follows:
1. The nonterminal in \( \alpha \) is the nonterminal on the left hand side of the production.

2. For each occurrence of an argument on the right side of the production, substitute an \( X \) such that, if it were the first argument, then substitute \( X_2^1 \), the second argument, \( X_2^1 \), the third argument \( X_2^3 \), etc.

3. Bracket each right hand side with an "*\[3" on the left and an "3"] on the right.

4. Change each ( to an [2 and each ) to an ]2.

5. Bracket each argument on the right hand side with an *[2 on the left and an ]2 on the right.

6. Add [1 and 1] as needed to make the strings in the production legal. Add \( X_1^\lambda \) wherever it can legally be added.

For example, the grammar

\[
S \to A(1)
\]

\[
A(a) \to A(aa)
\]

\[
A(a) \to a
\]

yields the grammar

\[
S \to *[3A[2*[2^1 X_1^\lambda 1_2[1 X_1^\lambda 1_2]3]]
\]

\[
A \to *[3A[2*[2 X_2^\lambda [1 X_2^\lambda 1_2[1 X_1^\lambda 1_2]3]]
\]

\[
A \to *[3 X_2^\lambda 3]
\]
It can be seen by induction that these two grammars give the same language. Notice in the latter case that if the same argument is copied in twice that it must be copied in as the same thing since the hypertree is completely built before the frontier is taken.

If \( L \in ALH^1_3 \), then construct a completed grammar \( G_3^3 = (\Sigma, V, P, S)_3^3 \) for \( L \) as was done in theorem 63. Construct an IO macrogrammar \( G' = (L', F', V', P, S', P') \) as follows:

1. \( \Sigma' = \Sigma \).
2. \( F' = \{S', D\} \cup \{1, 2\} \times \{1, 2, 3\} \times \{0, 1, 2, 3\}^m \).
3. \( V' = \{a_1, a_2, \ldots, a_m\} \) where each \( a_i \) does not appear elsewhere, and \( m \) is the maximum number of ones in a single path in the grammar.
4. \( \rho' \) maps all of \( V \) to \( m \), and \( S' \) to zero.
5. \( S' \) does not occur elsewhere.
6. \( P' \) is constructed as shown below.

Notice that for each nonterminal in this grammar there are two projections associated with the nonterminal and one projection for each possible \( X_2 \) that might exist in the grammar. The meaning of these projections are

1. The first projection:
   1. This nonterminal produces a singleton.
   2. This nonterminal does not produce a singleton.
2. The second projection:
1. This nonterminal produces a singleton after the first frontier operation.

2. This nonterminal produces a string which ends in an $X_1^\lambda$ after the first frontier operation.

3. This nonterminal produces a string which does not end in $X_1^\lambda$.

3. Each projection associated with a path: For each argument, 1, 2 and 3 mean the same as above except that the associated argument produces the string, etc. A zero means the associated argument is not to be used.

Therefore,

1. If $A \rightarrow a \in P$ where $A \in V_3$ and $a \in \Sigma$, then

$$(A,1,1,X_1\ldots X_m)(a_1,a_2\ldots a_m) \rightarrow a \in P'$$

for all $X_1\ldots X_m$.

2. If $A \rightarrow a_1[a_2B_3] \in P$, then

$$(A,2,2,X_1\ldots X_m)(a_1\ldots a_m) \rightarrow (B,2,2,X_1\ldots X_m)(a_1\ldots a_m) \in P'$$

$$(A,2,1,X_1\ldots X_m)(a_1\ldots a_m) \rightarrow (B,1,1,X_1\ldots X_m)(a_1\ldots a_m) \in P'$$

$$(A,2,1,X_1\ldots X_m)(a_1\ldots a_m) \rightarrow (B,2,1,X_1\ldots X_m)(a_1\ldots a_m) \in P'$$

$$(A,2,3,X_1\ldots X_m)(a_1\ldots a_m) \rightarrow (B,2,3,X_1\ldots X_m)(a_1\ldots a_m) \in P'$$
217

for all $X_1 \ldots X_m$.

3. If $A \rightarrow B \in P$ where $A \in V_2$ and $B \in V_3$, then

$$(A,1,1,X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B,1,1,X_1 \ldots X_m)(a_1 \ldots a_m) \in P'$$

$$(A,2,1,X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B,2,1,X_1 \ldots X_m)(a_1 \ldots a_m) \in P'$$

$$(A,2,2,X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B,2,2,X_1 \ldots X_m)(a_1 \ldots a_m) \in P'$$

$$(A,2,3,X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B,2,3,X_1 \ldots X_m)(a_1 \ldots a_m) \in P'$$

for all $X_1 \ldots X_m$.

4. If $A \rightarrow X_2^j \in P$ and $|j| = i$, then

$$(A,2,X_i,X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow a_i \in P'$$

for all $X_1 \ldots X_m$ except $X_i = 0$. If $X_i = 0$, no production is generated. Note that the second projection must equal $X_i$.

5. If $A \rightarrow B[2C_2] \in P$ and

$$C = C_1[1C_2[1C_3 \ldots [1C_i 1] \ldots 1]_1]$$

where each $C_j \in V_2$ except $C_{m+1}$ which is an element of $V_1$ and $m+1 \geq i > 0$, then

$$(A,2,3,X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (C_1Z_1Y_1X_1 \ldots X_m)(a_1 \ldots a_m) \ldots$$
\[(C_j, Z_j, 3, X_1 \ldots X_m) \in P' \]

if \(i \geq j > 0\).

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (C_1, Z_1, Y_1, X_1 \ldots X_m)(a_1 \ldots a_m) \ldots \]

\[(C_m, Z_1, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \in P' \]

if \(B \Rightarrow a\) where \(a \in \Sigma\).

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B, 2, 1, Y_1 \ldots Y'_1, 0 \ldots 0)(\]

\[(C_1, Z_1, Y_1, X_1 \ldots X_m)(a_1 \ldots a_m) \ldots \]

\[(C_m, Z_1, Y_1, X_1 \ldots X_m)(a_1 \ldots a_m, D, \ldots D) \in P' \]

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B, 2, 2, Y_1 \ldots Y'_1, 0 \ldots 0)(\]

\[(C_1, Z_1, Y_1, X_1 \ldots X_m)(a_1 \ldots a_m) \ldots \in P' \]

\[(A, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B, 2, 3, Y_1 \ldots Y'_1, 0 \ldots 0)(\]

\[(C_1, Z_1, Y_1, X_1 \ldots X_m)(a_1 \ldots a_m) \in P' \]

for all \(X_1 \ldots X_m, Y_1 \ldots Y_m\) and \(Z_1 \ldots Z_m\). Note that if \(i = m\) or \(m+1\) the sequence of zeros and \(D\)'s above is null.

6. If \(A \Rightarrow B \in P\) where \(A \in V_1\) and \(B \in V_2\), then

\[(A, 1, 1, X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B, 1, 1, X_1 \ldots X_m)(a_1 \ldots a_m) \in P' \]

\[(A, 2, 1, X_1 \ldots X_m)(a_1 \ldots a_m) \rightarrow (B, 2, 1, X_1 \ldots X_m)(a_1 \ldots a_m) \in P' \]
219

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

\[(A, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

for all possible combinations of \(X_1 \ldots X_m\).

7. If \(A \Rightarrow X_1 \in P\), then

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow \lambda \in P'\]

for all \(X_1 \ldots X_m\).

8. If \(A \Rightarrow B_1 \cdot C_1 \in P\), then

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 1, 1, X_1 \ldots X_m)(a_1 \ldots a_m)\]

\[(C, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 2, 1, X_1 \ldots X_m)(a_1 \ldots a_m)\]

\[(C, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

\[(A, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m)\]

\[(C, 2, 2, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

\[(A, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 1, 1, X_1 \ldots X_m)(a_1 \ldots a_m)\]

\[(C, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

\[(A, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 1, 1, X_1 \ldots X_m)(a_1 \ldots a_m)\]

\[(C, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \in P'\]

\[(A, 2, 3, X_1 \ldots X_m)(a_1 \ldots a_m) \Rightarrow (B, 2, 1, X_1 \ldots X_m)(a_1 \ldots a_m)\]
for all $X_1 \ldots X_m$.

9. These are all the elements of $P'$.

   It can be shown by induction that these two grammars define the same language.

   - - -

   At this point, a pumping lemma will be introduced. This will be used eventually to prove that the algebraic language hierarchy is a true hierarchy, that is, it has an infinite number of distinct levels.

   Before this is proven, the count function will be formally introduced as a means of determining the "length" of a hypertree.
Definition 61:

Count is a family of functions, denoted $\text{Count}_n^k$, such that

$$\text{Count}_n^k : H_n^k(I, V) \to \mathbb{N}$$

where $\mathbb{N}$ is the set of positive integers. The values for the count function are given by

(a) $\text{Count}_n^{n+1}(\gamma) = 1$
if $n \geq 0$.

(b) $\text{Count}_n^0(\gamma) = \text{Count}_n^1(\gamma)$
if $n \geq 0$.

(c) $\text{Count}_n^k(\gamma) = \text{Count}_{n+1}^k(\gamma)$
if $\gamma \in H_n^{k+1}$.

$\text{Count}_n^k(\gamma) = 1$
if $\gamma \in X_k \cup V_k$.

$\text{Count}_n^k(\gamma) = \text{Count}_{n+1}^k(\alpha) + \text{Count}_{n+1}^{k-1}(\beta)$
if $\gamma = \alpha[k^k\beta]$. 

if $n \geq k > 0$.

Notice that count function returns the sum of the number of nodes in a hypertree, the number of Xs in the hypertree plus the number of variables in the given hypertree.

Notationally, $\text{Count}_n$ will refer to $\text{Count}_n^0$, count will refer to $\text{Count}_n^0$, where $n'$ is larger than any level under consideration. In each case, a simple proof will show that the domain and codomain given is correct.
Theorem 76:

If $L$ is a regular hypertree language of level $n$ and degree $k$ and $j$ is a positive integer, then there exists an $N$ such that if $\gamma \in L$ and $\text{Count}(\gamma) > N$ then there exists $P_1 P_2 \ldots P_j$ such that $P_1 \in H_n^k$ and $P_2 \ldots P_j \in H_n^{k'}$ such that nonterminal $T$ occurs exactly once in $P_1 \ldots P_{j-1}$ and not all in $P_j$. Also, $T$ is the only nonterminal appearing, $\sum_{j \geq i > 1} \text{Count}(P_i) < N$ and if $L'$ is the language with productions

$$S' \rightarrow P_1$$
$$T' \rightarrow P_i$$

for all $j \geq i > 1$ then $L' \subset L$ and

$$S' \Rightarrow \gamma_1 \Rightarrow \gamma_2 \ldots \Rightarrow \gamma_j$$

where $\gamma_j = \gamma$ and $\gamma_{i-1} \Rightarrow \gamma_i$ uses the $i^{th}$ production above.

Proof:
Assume $G_n^k = (L, V, P, S)_n^k$ is a normal grammar for $P$. Proceed as in the standard proof for a context sensitive except choose $N$ large enough that one of the nonterminals have to occur $j-1$ times.
Note that if \( j \) is taken as one or two the result is trivial (in fact, \( N \) could be taken as zero). If three is chosen for \( j \), then in the string case this becomes the standard pumping lemma for regular sets. Also, in the tree case, this suggests the pumping lemma for context free sets.

To prove that the algebraic language hierarchy is a true hierarchy a language will be presented at each level which cannot belong to the next lower level. A few lemmas must be proven first.

Lemma 77:

If \( A \in H_n^k \), then

\[
\text{Count}(\text{frontier}_{n-1}^k(A)) < 2^{\text{Count}(A)}
\]

Proof by induction on the induction variable:

Basis: If \( m = 0 \), then \( \text{Count}(A) = 1 \), and in any event

\[
A = \text{frontier}_{n-1}^k(A)
\]

The lemma then becomes \( 1 < 2^1 \) which is obviously true.

Induction step: Assume that the lemma is true for all \( m' < m \). Let \( A \in H_n^k \). Then, one of the following is true:

1. If \( k = n+1 \), then the lemma is true as above.

2. If \( k = n \) and \( A = a[n\beta] \), then

\[
\text{frontier}_{n-1}^n(A) = \text{frontier}_{n-1}^{n-1}(\beta)
\]

and by the induction hypothesis.
Count(frontier^{n}_{n-1}(A)) = Count(frontier^{n-1}_{n-1}(\beta)) < 2\text{Count(\beta)}

Also, notice that \text{Count(A)} = \text{Count(\beta)} + 1, so that

\text{Count(frontier}^{n}_{n-1}(A)) < 2\text{Count(A)}

3. If \( k = 0 \), then

\text{frontier}^{0}_{n-1}(\gamma) = \text{frontier}^{1}_{n-1}(\gamma)

Therefore,

\text{count(frontier}^{0}_{n-1}(\gamma)) < 2\text{count(\gamma)}

by induction hypothesis.

4. If \( n > k > 0 \) and \( \gamma \in X_{k} \), then

\text{frontier}^{k}_{n-1}(\gamma) = \gamma

so

\text{count(frontier}^{k}_{n-1}(\gamma)) < 2\text{count(\gamma)}

since both \text{count(frontier}^{k}_{n-1}(\gamma)) \text{ and count(\gamma) are one.}

5. If \( n > k > 0 \) and \( A \in H_{n,m-1}^{n} \), then

\text{frontier}^{k}_{n-1}(A) = \text{frontier}^{n-1}_{n-1}(A)

and by the induction hypothesis

\text{Count(frontier}^{k}_{n-1}(A)) < 2\text{Count(A)}

6. If \( n-1 > k > 0 \) and \( A = \alpha[k^{\beta}, k] \), then
frontier_{n-1}^k(A) = frontier_{n-1}^{k+1}(a)[ frontier_{n-1}^{k-1}(\beta) ]

and by the induction hypothesis

\[
\text{Count}(\text{frontier}_{n-1}^k(A)) < 2^{\text{Count}(a)} + 2^{\text{Count}(\beta)} < 2^{\text{Count}(a)} \times 2^{\text{Count}(\beta)} < 2^{\text{Count}(A)}
\]

7. If \( k = n-1 \), \( A = a[n-1 \beta n-1] \) and \( a \notin \Sigma \), then

\[
\text{frontier}_{n-1}^k(A) = \text{apply}_{n-1}^{n-1}(\text{frontier}_{n-1}^n(a), \text{frontier}_{n-1}^{n-2}(\beta))
\]

But by induction hypothesis

\[
\text{Count}(\text{frontier}_{n-1}^n(a)) < 2^{\text{Count}(a)} \quad \text{and} \quad \text{Count}(\text{frontier}_{n-1}^{n-2}(\beta)) < 2^{\text{Count}(\beta)}
\]

The worst possible case is if every symbol in \( a \) copies all of \( \beta \). In this case, the count of the result is the product of the counts of the frontiers. That is

\[
\text{Count}(\text{frontier}_{n-1}^k(A)) \leq 2^{\text{Count}(a)} \times 2^{\text{Count}(\beta)} < 2^{\text{Count}(a)} + \text{Count}(\beta)
\]

But \( \text{Count}(A) = \text{Count}(a) + \text{Count}(\beta) \) by definition, so
Count(frontier[^k][n-1](A)) < 2^Count(A)

8. If \( k = n-1 \), \( A = \alpha[n-1 \beta n-1] \) and \( \alpha \in \Sigma \), then

\[
\text{frontier}[^k][n-1](A) = \alpha[n-1 \text{ frontier}[^{n-2}][n-1](\beta) n-1]\]

and by induction hypothesis

\[
\text{Count(frontier}[^{n-2}][n-1](\beta)) < 2^\text{Count}(\beta)
\]

which means that

\[
\text{Count(frontier}[^k][n-1](A)) = \text{Count(frontier}[^{n-2}][n-1](\beta)) + 1
\]

\[
< 2^\text{Count}(\beta) + 1
\]

But since it is known that \( \text{Count}(\beta) \geq 1 \) it must be true that

\[
\text{Count(frontier}[^k][n-1](A)) < 2^\text{Count}(\beta) + 2^\text{Count}(\beta)
\]

\[
< 2^\text{Count}(\beta) + 1
\]

\[
< 2^\text{Count}(A)
\]

Therefore, since in every case above

\[
\text{Count(frontier}[^k][n-1](A)) < 2^\text{Count}(A)
\]

the lemma is true by induction.

- - -
This last lemma in connection with the pumping lemma will be used in a proof that there exists at least one language at each level of the hierarchy that is not in that hierarchy. In order to describe this language, a special function needs to be defined.

Definition 62:

$E$ is a function such that

$E : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$

where $\mathbb{N}$ is the set of positive integers. The values are given by the recursive definition

(a) $E(0,m) = m$
(b) $E(n+1,m) = 2E(n,m)$

Note that $E(0,m) = m$, $E(1,m) = 2^m$ etc. This will be used to define the language described above.

Lemma 78:

The language $L = \{a^\gamma \mid \gamma = E(n,m) \text{ for all } m\}$ where $n$ is fixed cannot be in $\text{ALH}^1_k$ where $k \leq n+1$.

Proof:
Assume that $L \in \text{ALH}^1_k$ where $k \leq n+1$. Then, there must be a grammar $G_k^k = (L,V,P,S)_k^k$ such that

$L = \text{extract}_1(G_k^k)$
Let $L' = \text{language}(G_k^k)$. Without loss of generality, assume that $\text{extract}_i(y)$ exists for all $y \in L'$. Let $N$ be chosen as described in the pumping lemma for $j = 4$. Let $\bar{y}$ be the smallest hypertree such that $\text{extract}_1(\bar{y}) = a^{E(n,m)}$ and $a^{E(n,m)}$ is not the frontier of any hypertree with fewer than $N$ nodes. Note that $\text{Count}(\bar{y}) > N$. Therefore, there exist $P_1', P_2', P_3', \text{and} P_4'$ such that there is a sublanguage $L'$ with productions

\[
S' \rightarrow P_1' \\
T \rightarrow P_2' \\
T \rightarrow P_3' \\
T \rightarrow P_4'
\]

and $S' \Rightarrow \bar{y}_1 \Rightarrow \bar{y}_2 \Rightarrow \bar{y}_3 \Rightarrow \bar{y}$ by using the productions in order. Note that neither $P_2'$ nor $P_3'$ can be dropped by the frontier operation since if they were $S' \Rightarrow P_1' \Rightarrow \bar{y}_2' \Rightarrow \bar{y}'$ where the dropped substructure is not used would result in a shorter element of $L'$ for $a^{E(n,m)}$. Therefore, consider the sequence of elements of $L'$ generated by

\[
S' \Rightarrow P_1 \Rightarrow \bar{y}_2'' \Rightarrow \bar{y}_3'' \Rightarrow \bar{y}''
\]

such that the first production is used the first time, followed by $i$ applications of the second rule, followed by the third and fourth rule. Clearly this is a subset of $L'$. 
Since $P_2$ does not drop $P_3$, this will yield an increasing sequence from $L$ such that the elements of the second sequence are arrived at by taking $\text{extract}_2$'s of the first sequence. Let the first sequence be represented by

$$\gamma_0', \gamma_1', \gamma_2'...$$

where $\gamma_1 = \gamma$ and the second sequence be represented by

$$\gamma_0'', \gamma_1'', \gamma_2''...$$

where $\gamma_i'' = \text{extract}_2(\gamma_i')$ for all $i$. Then, form a third sequence

$$\gamma_0''', \gamma_1''', \gamma_2'''...$$

such that $\gamma_i''' = \text{frontier}_1(\gamma_i'')$. Also, clearly

$$\text{Count}(\gamma_i''') \geq E(n, m+i-1)$$

if $i > 1$ since this must be an increasing sequence in $L$. Also,

$$\text{Count}(\gamma_i') = \text{Count}(P_1) + i*\text{Count}(P_2) + \text{Count}(P_3) + \text{Count}(P_4)$$

or in other words $\text{Count}(\gamma_i') = K_1 + K_2*i$ where $K_1$ and $K_2$ are constants. To perform the $\text{extract}_2$ requires $k-2$ frontier operations, so an obvious induction using lemma 77 would show that

$$\text{Count}(\gamma_i') < E(k-1, K_1 + K_2*i)$$
The final tree frontier can only drop nodes -- that is, it cannot duplicate nodes -- so

\[ \text{Count}(\gamma_i^\prime) \leq \text{Count}(\gamma_i') \]

So it must be true that

\[ E(n,m+i-l) < E(k-1,K_1+K_2^*i) \]

or, by another short induction

\[ E(n-K+1,m+i-l) < E(0,K_1+K_2^*i) \]
\[ < K_1+K_2^*i \]

But this last statement is clearly false since the left side of this inequality is exponential whereas the right side is linear. Therefore, the lemma must be true since assuming it is not true leads to a contradiction.

---

At this point, it will be shown that the language given in lemma 78 is in \( \mathsf{ALH}_{n+2}^1 \). This will be done by inductively showing a grammar for the language.

Lemma 79:

The language given in lemma 78 is in \( \mathsf{ALH}_{n+2}^1 \).

Proof by induction on \( n \):

Basis: If \( n = 0 \), then
S → A
A → *[2^a[1 A 1]2] | a

is a grammar for the language.

Induction step: Assume that a grammar $G_{n}^{n}$ exists for the language at level $n-1$. Then, rotate this grammar 90 degrees. That is, increment every number in the grammar by 1 to form a new grammar for a "linear" tree. For example, the above grammar would give

S → A
A → *[3^a[2 A 2]3] | a

Next, for each "a" substitute

*[\_n\_{\_n-1}\ldots[3*[\_3^x_2^x\_x_2^x[x_1^x_1^x]]]]_3\ldots n-1\_n]

The "*[\_n\ldots\_3]" is used to hold the structure "*[3*[2^x_2^x\_x_2^x[x_1^x_1^x]]]3\ldots n-1\_n]" together until the frontier$_3^0$ phase of the extract. During the frontier$_2^0$ phase, each $X_2^\lambda$ will pull in the entire structure following it which will create a full binary tree. All of this copying will result in a frontier of length $2^q$ where $q$ is the length of the corresponding element of the original grammar. Each element of the resulting string frontier would be $X_2^\lambda$ if it were
legal. Therefore, another alternation must be made to substitute an "a" for the $x_2^\lambda$ in each case. This is to add a new start symbol, $S'$, and the production

$$S' \to \ast[n \ast[n-1\ldots\ast[3S_3][2a_2]_3]\ldots_n]$$

Frontiering will keep this structure together until the frontier$_3^0$, when the "a" will be substituted for the $x_2^\lambda$ mentioned above. Each of the facts stated above could be established by an induction proof, and therefore, this altered grammar is the required grammar.

- • -

The previous two lemmas then make the proof of hierarchy easy.

Theorem 80:

$$ALH_n^1 \subset ALH_{n+1}^1$$

and

$$ALH_n^1 \not= ALH_{n+1}^1$$

for all $n > 0$.

Proof:

Clearly this is true for $n = 1$ (regular and context free language). At all other levels, containment is obvious, and the previous two lemmas give an example of language in each level that is not in the previous level.

- • -
Corollary 81:

\[ \text{ALH}_n^k \subset \text{ALH}_n^{k+1} \]

and

\[ \text{ALH}_n^k \neq \text{ALH}_n^{k+1} \]

for all \( n \geq k \geq 1 \).

Proof:

Containment is obvious. Suppose that

\[ \text{ALH}_n^k = \text{ALH}_n^{k+1} \]

for some \( n \) and \( k \). Then, it would be easy to show by induction that

\[ \text{ALH}_n^1 = \text{ALH}_n^{1+1} \]

which contradicts the theorem.

-- * --

The proof that all members of \( \text{ALH}_n^1 \) are context sensitive will be deferred until chapter 8 due to the length of the proof. It will be shown here, however, that \( U_{n>0} \text{ALH}_n^1 \) is not equal to the set of context sensitive language. This will be done by diagonalization of the above hierarchy of language. That is,

\[ \{ a_y \mid y = E(n,1) \}_{n>0} \]

is context sensitive, but not any level of the hierarchy.
Lemma 82:

The language

\[ L = \{ ay \mid r = E(m,1) \text{ for } m > 0 \} \]

is not in \( \text{ALH}_n^1 \) for any \( n \).

Proof:
This proof parallels exactly the proof of lemma 78.

---

Lemma 83:

The language given above is context sensitive.

Proof:
This proof consists of demonstrating a linearly bounded grammar for the above language. Let \( G \) be the grammar

\[
\begin{align*}
S & \to <(A)> \\
A> & \to RD< > \\
AR & \to RD \\
<(R & \to <(A> \\
AD & \to DAA \\
<(D & \to <(> \\
A< > & \to a \\
Aa & \to aa \\
<(>a & \to a
\end{align*}
\]

It is easy to show by induction that the above grammar generates the desired language. It can also easily be seen
that if \( \gamma \in \text{language}(G) \) then there is a derivation bounded by \( |\gamma| + 3 \). Note that the grammar works by generating each successive element of the language as a string of As with \(<(> and \langle <\rangle around it. At each point, \( A<(> \) goes to \( RD<(> \) will go to the next element whereas "\( A<(>" goes to "a" will force the production to stop with a string of "a"'s. Note, the three extra symbols are the \(<(>, <\rangle) and the D.

---

Theorem 84:

The set of languages \( U_{n>0} \text{ALH}_{n}^{1} \) is not equal to the set of context sensitive languages.

Proof:

This follows directly from the last two lemmas.

---

This concludes the chapter on miscellaneous theorems about the hierarchies. Chapter 8 will be dedicated to showing that the algebraic language hierarchy is contained in the set of context sensitive languages.
This chapter will show that any language in the algebraic language hierarchy is, in fact, contained in the set of context sensitive languages. This will be done by "programming" a frontier taking algorithm using unrestricted grammars. It will be shown that this is a linearly bounded operation -- that is, it is linearly bounded in both its input and its output. After that, it will be shown that for any grammar, an equivalent grammar can be found which yields the same string language such that the intermediate frontiers are linearly bounded. This, then, will be a linearly bounded grammar for the resulting language, and, therefore, the language must be context sensitive.

Rather than treat the general case of hypertree grammars, only grammars which meet the following three restrictions will be handled:

1. The set of nonterminals contains only those nonterminals which appear in the productions.

2. The set of terminals contains only those terminals which appear in the productions.

3. If B is a nonterminal from \( V_k \), then there is a \( \psi \) such that \( B \Rightarrow^* \psi \) and \( \psi \in H_n^k \). That is, all nonterminals derive something.
It should be rather obvious that these "restrictions" are not restrictions at all, since a simple algorithm would generate a grammar following these restrictions from any grammar.

Throughout this chapter a special set, which is denoted $V$, is used. This is a finite set containing all the symbols from the initial hypertree grammar which are viewed as terminals in resulting grammar. This will include elements of $I$, brackets and $X$'s that occur in the initial hypertree grammar. It is finite since each production is of finite length and there are only a finite number of production. Also, throughout this development the symbols $n$, $k$ and $k'$ will be used. It assumed that these are integers such that $n' \geq n \geq 0$, $n \geq k > 0$ and $n+1 \geq k' \geq 0$ unless otherwise stated, where $n'$ is the level of the original grammars.

That is, if the production

$$<A_{1n}^k> \alpha <) > \rightarrow \alpha$$

where $\alpha \in \{A,B\}$ is given, and the grammar given is of level 2, this corresponds to saying that

$$<A_2^2> A <) > \rightarrow A$$

$$<A_2^2> B <) > \rightarrow B$$

$$<A_2^1> A <) > \rightarrow A$$
\[ <\text{Al}_2^1> \ B \ (<)> \rightarrow B \]
\[ <\text{Al}_1^1> \ A \ (<)> \rightarrow A \]
\[ <\text{Al}_1^1> \ B \ (<)> \rightarrow B \]

are all productions in the grammar.

Similarly in the "search" function, the symbol "j" appears. "j" is used to stand for any postfix of a path which occurs in \( V \). It is assumed that any nonterminal or terminal which is presented in this grammar does not occur in the original grammar unless specifically stated. Therefore, given a grammar \((\mathcal{Z}, V, P, S)^k_n\), the additional productions needed to make this an unrestricted grammar are given according to the function which is to be performed by the productions. Each of these "functions" will be discussed in the following paragraphs. Note that this corresponds to writing a program in a particular applicative language.

The miscellaneous functions given in table 5 are self-evident and will not be discussed separately.

Table 6 gives an implementation of the "over" function. This function is called by placing the nonterminal \(<\text{over}_n^k>\) on the left side of a string of characters with the \(<\text{over}_n^k>\) preceded by a nonterminal from the set \( V' \). The effect of the routine is to place that nonterminal on the right hand
TABLE 5. Miscellaneous Routines

Dup:
\[ <\text{Dup}>\rangle \rightarrow \langle\rangle\langle\rangle\]

Blitz:
\[ <\text{Blitz}>\alpha \rightarrow <\text{Blitz}> \]
\[ \text{if } \alpha \notin \{\langle\rangle,\langle,\rangle\}-\langle\text{end}\rangle\} \]
\[ <\text{Blitz}><\text{end}>\rangle \rightarrow \lambda \]

Switch:
\[ <\text{Switch}_{k}\rangle \langle\rangle \rightarrow \langle\rangle_{k}\rangle \]
\[ \text{where } n' \geq k > 0 \]

Die:
\[ <\text{Die}>\rangle \rightarrow \langle\rangle \]

where \( <\text{Dup}>, <\text{End}>, <\text{Switch}_{k}\rangle \) and \( <\text{Die}> \) are elements of \( V' \) (see the over routine)

side of a hypertree of level \( n \) and degree \( k \). This is done by use of the terminal \( <\text{post}_{n}^{k'}> \) where the \( k' \) and \( n \) are always equal to the level and degree of hypertrees that the nonterminal and \( <\text{over}_{n}^{k}> \) have passed over. If at any point,
TABLE 5. The Over Routine

a. $\beta_{n}^{k} \alpha \rightarrow \alpha \beta_{n}^{k} \langle \text{past}_{n}^{n+1} \rangle$

   if $\alpha \in \Sigma$

b. $\beta_{n}^{k} \langle \text{past}_{n}^{1} \rangle \rightarrow \beta_{n}^{k} \langle \text{past}_{n}^{0} \rangle$

c. $\beta_{n}^{k} \langle \text{past}_{n}^{k+1} \rangle \rightarrow \beta_{n}^{k} \langle \text{past}_{n}^{k'} \rangle$

   $\beta_{n}^{k} \langle \text{past}_{n}^{k'} \rangle \rightarrow \alpha \beta_{n}^{k} \langle \text{past}_{n}^{k'} \rangle$

   if $\alpha \in X_{k'} \cap V$

   $\beta_{n}^{k} \langle \text{past}_{n}^{k'+1} \rangle \rightarrow \beta_{n}^{k} \langle P_{1}^{k'} \rangle$

   $\beta_{n}^{k} \langle \text{past}_{n}^{k'-1} \rangle \langle P_{1}^{k'} \rangle \rightarrow k' \beta_{n}^{k} \langle \text{past}_{n}^{k'} \rangle$

   $\langle P_{1}^{k'} \rangle \alpha \rightarrow \alpha \langle P_{1}^{k'} \rangle$

   if $\alpha \in V U V U \{<),<\text{frontier}_{n}^{k''}\}$

(frontier)

   $\beta_{n}^{k} \langle \text{frontier}_{n}^{k'} \rangle \rightarrow \langle \text{frontier}_{n}^{k'} \rangle \beta_{n+1}^{k} \langle P_{2}^{k'} \rangle$

   $\beta_{n+1}^{k} \langle \text{past}_{n+1}^{k'} \rangle \langle P_{2}^{k'} \rangle \langle<\rangle \rightarrow <\rangle \beta_{n}^{k} \langle \text{past}_{n}^{k'} \rangle$

   $\beta_{n+1}^{k} \langle \text{past}_{n+1}^{k'+1} \rangle \langle P_{2}^{n+1} \rangle \langle<\rangle \rightarrow <\rangle \beta_{n}^{k} \langle \text{past}_{n}^{k'} \rangle$

   $\langle P_{2}^{k'} \rangle \alpha \rightarrow \alpha \langle P_{2}^{k'} \rangle$

   if $\alpha \in V U \{<\text{frontier}_{n}^{k''},<\rangle\}$

(finish)

   $\beta_{n}^{k} \langle \text{past}_{n}^{k} \rangle \rightarrow \beta$
TABLE 7. The Search Routine

\[
\begin{align*}
\text{<search}\_n^j & \rightarrow \text{<mark}>\text{<setup}\_n^j} \\
\text{<setup}\_n^j & \alpha \rightarrow \alpha\text{<setup}\_n^j \\
& \quad \text{if } \alpha \in \mathcal{V} \cup \{\text{<frontier}\_n^k, \text{<}\}\} \\
\text{<setup}\_n^j & \text{<>, } \rightarrow \text{<>, }\text{<}S1\_n^k\text{>}\text{<Dup}>\text{<over}\_n^{n-1} \\
\text{a. } \text{<}S1\_n^k\text{>} & \rightarrow \text{<move}>\text{<S5}>\text{<over}\_n^n \\
& \quad \text{<move}>\alpha \rightarrow \alpha\text{<return}>\alpha\text{<move>} \\
& \quad \quad \text{where } \alpha \in \mathcal{V} \cup \{\text{<frontier}\_n^k, \text{<}\}\} \\
\beta\text{<return> } & \rightarrow \alpha\text{<return}>\beta \\
& \quad \quad \text{where } \beta \in \mathcal{V} \cup \{\text{<frontier>, <\}, <, <}\} \\
\text{<mark>\alpha\text{<return> } } & \rightarrow \alpha\text{<mark>} \\
\text{<mark> } & \rightarrow \lambda \\
\text{<move>\text{<S5>}} & \rightarrow \lambda \\
\text{b. } \text{<}S1\_n^0\text{>} & \rightarrow \text{<}S1\_n^1\text{>} \\
\text{c. } \text{<}S1\_n^k\text{>} & \rightarrow \text{<}S2\_n^k\text{>}\text{<over}\_n^{k+1} \\
& \quad \text{if } j = k_i \\
\text{<}S2\_n^k\text{>} & \rightarrow \text{[<}S1\_n^{k-1}j\text{>}\text{<switch}>\text{<over}\_n^{k-1} \\
\text{<}S1\_n^k\text{>} & \rightarrow \text{<}S1\_n^{k+1}j\text{>}\text{<}S3\_n^k\text{>}\text{<over}\_n^{k+1} \\
& \quad \text{if } j \in p\_n^{k+1} \\
\text{<}S3\_n^k\text{>} & \rightarrow < > \\
\text{<}S3\_n^k\text{>} & \rightarrow < > \text{[<}S4\_n^k\text{>}\text{<over}\_n^{k-1} \\
\text{<}S4\_n^k\text{>} & \rightarrow \text{[<} \\
\text{where } \text{<}S2\_n^k\text{>, } \text{<}S3\_n^k\text{>, } \text{<}S4\_n^k\text{>, and } \text{<S5> are elements of } \mathcal{V}'.
\end{align*}
\]
the terminal \(<\text{past}^k_n\rangle\), indicating that a hypertree of level \(n\) and degree \(k\) has been passed over, an alternative production \(\beta<\text{over}^k_n><\text{past}^k_n\rangle \rightarrow \beta\) allows the routine to end with the \(\beta\) deposited on the right end of a hypertree. Notice that the over function will work even if a frontier function, indicated by \(<\text{frontier}^k_n\rangle\), is included in the string without requiring the function to be evaluated. It is the responsibility of the calling routine to insure that the proper hypertree is chosen. For example, "B<over^N_n>a[n a_n]\)" could place the B after either the "a" to yield "aB[n a_n]\)" or after the "a[n]\)" to yield "a[n a_n]B" since both "a" and "a[n a_n]\)" are legal hypertrees of level \(n\). Quite often if the wrong hypertree is chosen the B gets stranded and can never be eliminated from the resulting structure so that no output element is generated.

The search routine is given in table 7. It works on a string structure of the form

\(<\text{search}^k_n\rangle \gamma_1 \langle\rangle \gamma_2 \langle\rangle\)

where \(\gamma_1\) is an unknown string of elements of \(V\) and nonterminals, and \(\gamma_2\) is the argument to the search. The reason for this arrangement is that the argument may have to be used by other search routine calls and, therefore, must be held independently of the calls. If the argument were
TABLE 8. The Apply Routine

\[
<\text{apply}^k_n> \rightarrow <A_{2n}> <A_{1n}^k> <A_{4n}> <\text{over}^k_n>
\]

\[
<A_{2n}^k> \alpha \rightarrow \alpha <A_{2n}^k>
\]
if \( \alpha \in \Sigma \).

\[
<A_{2n}^k> <\_> \rightarrow <\text{Blitz}> <\text{end}> <\text{over}^{n-1}_n>
\]

\[
<A_{4n}^k> <\_> \rightarrow <\_><\_>
\]

a. \( <A_{1n}^{n+1}> \alpha <\_> \rightarrow \alpha \)
if \( \alpha \in \Sigma \).

b. \( <A_{1n}^0> \rightarrow <A_{1n}^1> \)

c. \( <A_{1n}^k> \alpha <\_> \rightarrow \alpha \)
if \( \alpha \in X^k_n \cap \Sigma \).

\[
<A_{1n}^k> \alpha <\_> \rightarrow <\text{search}^{n-1}_n j>
\]
if \( \gamma = X^j_n \) where \( \gamma \in \Sigma \)

\[
<A_{1n}^k> \rightarrow <A_{1n}^{k+1}> <A_{3n}^k> <\text{over}^{k+1}_n>
\]

\[
<A_{3n}^k> <\_> \rightarrow <\_>
\]

\[
<A_{3n}^k> [k \rightarrow <\_>] [k <A_{1n}^{k-1}> <\text{switch}_k> <\text{over}^{k-1}_n>
\]

where \( <A_{4n}> \) and \( <A_{3n}^k> \) are in \( \Sigma' \).
copied to be held adjacent to the \(<\text{search}_{n}^{kj}\rangle\), then the resulting grammar would not be linearly bounded. Beyond this the implementation is straightforward. The nonterminal \(<\text{setup}_{n}^{kj}\rangle\) is used to change the string structure to
\[
<\text{mark}> \gamma_1 <,> <\text{S1}_{n}^{kj}\gamma_2 <)> <\rangle>
\]
Then, the \(<\text{S1}_{n}^{kj}\rangle\) nonterminal is used to do the search that is requested. When the "j" is exhausted and \(<\text{S1}_{n}^{kj}\rangle\) is encountered, the search ends and the nonterminal \(<\text{move}\rangle\) is used to send the characters selected back to the \(<\text{mark}\rangle\).
\(<\text{mark}\rangle\) can disappear, but if not all the characters are moved yet a nonterminal \(<\text{return}\rangle\) will get stranded. Note that \(<\text{frontier}_{n}^{k}\rangle\) and \(<\rangle\) are included in the characters that can be moved so that the frontier functions do not have to be evaluated before the search function is used.

The apply function is given in table 8. It works on the structure
\[
<\text{apply}_{n}^{k}\gamma_1 <,> \gamma_2 <j>
\]
The initial productions transforms this into
\[
<\text{A2}_{n}> <\text{A1}_{n}^{k}\gamma_1 <>)<<,> \gamma_2 <rangle>
\]
which should look something like the string required for the search routine. The nonterminal \(<\text{A1}_{n}^{k}\rangle\) is used to do
TABLE 9. The Frontier Routine

<table>
<thead>
<tr>
<th>Case</th>
<th>Description</th>
</tr>
</thead>
</table>
| a.  | $\langle \text{frontier}^{n+1}_n \alpha \rangle \rightarrow \alpha$  
   |   if $\alpha \epsilon \Sigma$.  
   |   $\langle \text{frontier}^{n+1}_n \alpha \rangle \rightarrow \langle \text{frontier}^n_n \text{F1}_n \rangle \langle \text{over}^n_{n+1} \rangle$  
   |   if $\alpha \epsilon \Sigma$.  
   |   $\langle \text{F1}_n \rangle_{n+1} \rangle \rightarrow \langle \rangle$  
| b.  | $\langle \text{frontier}^0_n \rangle \rightarrow \langle \text{frontier}^1_n \rangle$  
| c.  | $\langle \text{frontier}^k_n \rangle \alpha \rangle \rightarrow \alpha$  
   |   if $\alpha \epsilon X_k \cap V$.  
   |   $\langle \text{frontier}^k_n \rangle \rightarrow \langle \text{frontier}^{k+1}_n \text{F2}_n \rangle \langle \text{over}^{k+1}_{n+1} \rangle$  
   |   if $k < n$.  
   |   $\langle \text{F2}_n \rangle^k \rangle \rightarrow \langle \rangle$  
   |   $\langle \text{F2}_n \rangle^k \rangle \rightarrow \langle \text{frontier}^{k-1}_n \text{switch}_k \rangle \langle \text{over}^{k-1}_{n+1} \rangle$  
   |   $\langle \text{frontier}^n_n \rangle \rightarrow \langle \text{frontier}^n_n \text{Die} \rangle \langle \text{over}^n_{n+1} \rangle$  
   |   $\langle \text{frontier}^n_n \rangle \alpha \rangle \rightarrow \langle \text{apply}^n \rangle \langle \text{frontier}^n_n \text{F3}_n \rangle \langle \text{over}^n_{n+1} \rangle \alpha \rangle \langle \text{F4}_n \rangle \langle \text{over}^n_{n+1} \rangle$  
   |   if $\alpha \epsilon \Sigma$.  
   |   $\langle \text{F4}_n \rangle_{n+1} \rangle \rightarrow \langle \rangle$  
|   | $\langle \text{F3}_n \rangle \rightarrow \langle \rangle$  
|   | $\langle \text{F3}_n \rangle \rightarrow \langle \text{frontier}^{n+1}_n \text{F4}_n \rangle \langle \text{over}^n_{n+1} \rangle$  
|   | $\langle \text{F4}_n \rangle \rightarrow \langle \rangle$  
| where $\langle \text{F1}_n \rangle$, $\langle \text{F2}_n \rangle$, $\langle \text{F3}_n \rangle$ and $\langle \text{F4}_n \rangle$ are in $V'$.  


actual work of the apply function, where the second argument is always separated as per the discussion of the search routine. The actual implementation is very simple.

Table 9 gives the productions for the frontier routine. This is an extremely straightforward implementation of this function, and need not be discussed here.

TABLE 10. The String Routine

\[
\begin{align*}
S' & \rightarrow \langle \text{string}_n, \text{S} \rangle \\
\langle \text{string}_n \rangle & \rightarrow \langle \text{string}_{n-1} \rangle \langle \text{frontier}^0_{n-1} \rangle \langle \text{Dup} \rangle \langle \text{over}^0_n \rangle \\
\langle \text{string}_1 \rangle & \rightarrow \langle \text{S} \rangle \langle \text{over}_1 \rangle \\
\langle \text{S} \rangle \langle \rangle & \rightarrow \lambda \\
\end{align*}
\]

where the original grammar is \((L, V, P, S)^k\) and

\[S'\] is the new start symbol.

String is a routine which is used to call the frontier function the required number of times to get the string frontier. The implementation of the string routine is given in Table 10. At this point, the start symbol of the initial grammar is referenced so that the initial hypertree can be generated.
The proofs involving these routines require the use of the following data structure which is based on hypertrees.

Definition 63:

\( F_{n,m}^k \) is the smallest set such that

(a) \( F_{n,m}^{n+1} = \emptyset \)

if \( n \geq 0 \) and \( m \geq 0 \).

(b) \( F_{n,0}^k = X_k \)

if \( n \geq k > 0 \).

(c) \( F_{n,m+1}^k = F_{n,m}^{k+1} \cup F_{n,m}^k \cup F_{n,m}^{k+1}[F_{n,m}^{k-1}] \)

\( \cup \langle \text{frontier}_n^k \rangle_{F_{n+1,m}^k} \rangle \)

if \( n > k > 0 \) and \( m \geq 0 \).

(d) \( F_{n,m}^0 = F_{n,m}^1 \)

if \( n > 0 \) and \( m \geq 0 \).
Notice that if the frontier function works properly, and \( \gamma \in F_{n,m}^k \) then \( \gamma \Rightarrow \gamma' \) where \( \gamma' \in H_{n}^k \). The purpose of this data structure is to allow the definition of hypertrees where not all of the frontier functions have been evaluated.

In view of this definition, the following lemmas summarize the use of these functions.

Lemma 85:

If \( \gamma \in F_{n,m}^{k'} \cap (V U \{<\text{frontier}^k_n,\langle\rangle^k_n>\})^* \) and \( \beta \in V' \), then

\[
\beta^{\langle\text{over}^k_n\rangle}\gamma \Rightarrow \gamma^{\langle\text{over}^k_n\rangle}\beta^{\langle\text{past}^k_n\rangle}
\]

and there is a derivation of width less than or equal to \( 2|\gamma|+3 \).

Proof by induction on \( m \):

Basis: If \( m = 0 \), then

\[
F_{n,0}^{k'} = X_{k'}
\]

and the derivation is given in table 11.

Induction step: Assume that if \( m' < m \) then

\[
\beta^{\langle\text{over}^k_n\rangle}\gamma \Rightarrow \gamma^{\langle\text{over}^k_n\rangle}\beta^{\langle\text{past}^k_n\rangle}
\]

where \( \gamma \in F_{n,m}' \). If \( \gamma \in F_{n,m}^k \), then
TABLE 11. Over Derivation for $\gamma \in F_{n,0}^{k'}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 $\beta &lt;_{n}^{k}\gamma$</td>
<td>4</td>
</tr>
<tr>
<td>2 $\gamma \beta &lt;<em>{n}^{k}&lt;\text{past}</em>{n}^{k'}$</td>
<td></td>
</tr>
</tbody>
</table>

1. If $k' = n+1$, then $\gamma \in \Sigma$ and the derivation is given in table 12.

TABLE 12. Over Derivation for $\gamma \in \Sigma$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 $\beta &lt;_{n}^{k}\gamma$</td>
<td>4</td>
</tr>
<tr>
<td>2 $\gamma \beta &lt;<em>{n}^{k}&lt;\text{past}</em>{n}^{n+1}$</td>
<td></td>
</tr>
</tbody>
</table>

2. If $n \geq k' > 0$ and $\gamma \in F_{n,m-1}^{k+1}$, then the derivation is given in table 13. Notice that in going from line 1 to line 2 the induction hypothesis had to be used.

3. If $n \geq k > 0$ and $\gamma \in F_{n,m-1}^{k}$, then the lemma is true by induction hypothesis.
TABLE 13. Over Derivation for $\gamma \in F_{n,m-1}^{k+1}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\beta^{k}_{n} \gamma$</td>
<td>$2</td>
</tr>
<tr>
<td>2. $\gamma \beta^{k}<em>{n} \gamma^{k+1}</em>{n}$</td>
<td>$</td>
</tr>
<tr>
<td>3. $\gamma \beta^{k}<em>{n} \gamma^{k}</em>{n}$</td>
<td></td>
</tr>
</tbody>
</table>

4. If $n \geq k > 0$ and $\gamma \in F_{n,m-1}^{k+1}$, then the derivation is given in table 14. Notice in this table $|\gamma| = |\gamma_{1}| + |\gamma_{2}| + 2$ therefore $2|\gamma|+3 = 2|\gamma_{1}|+2|\gamma_{2}|+7$.

5. If $n \geq k > 0$ and $\gamma \in <\text{frontier}_{n}^{k}> F_{n+1,m-1}^{k}$, then the derivation is given in table 15. Here, $2|\gamma|+3 = 2|\gamma_{1}|+7$.

6. If $k = n$ and $\gamma \in <\text{frontier}_{n}^{k+1}> F_{n+1,m-1}^{n+1}$, then the derivation would be similar to that given in table 15.

7. If $k = 0$, then the derivation is given in table 16. Notice that between lines 1 and 2 is a use of the derivation given for one of the previous parts of this induction step.

Note that in each case the maximum width is less than $2|\gamma|+3$ so that the lemma is proved.
### TABLE 14. Over for $\gamma \in F_{n,m-1}^{k'+1} \cdot F_{n,m-1}^{k'-1}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\beta^{\overline{k}} \gamma_{1[k} \gamma_{2,k]}$</td>
<td>$2</td>
</tr>
<tr>
<td>2. $\gamma_{1[k'} \beta^{\overline{k}} &lt;\text{past}<em>{n</em>{k'}}^{k'+1}&gt;\gamma_{2,k'}]$</td>
<td>$</td>
</tr>
<tr>
<td>3. $\gamma_{1[k'} \beta^{\overline{k}} &lt;\text{P}<em>{n</em>{k'}}^{k'}\gamma_2,k']$</td>
<td>$</td>
</tr>
<tr>
<td>4. $\gamma_{1[k'} \beta^{\overline{k}} &lt;\text{P}<em>{n</em>{k'}}^{k'}\gamma_2,k']$</td>
<td>$</td>
</tr>
<tr>
<td>5. $\gamma_{1[k'} \gamma_{2[k} \beta^{\overline{k}} &lt;\text{past}<em>{n</em>{k'}}^{k'-1}&gt;\gamma_{2,k'}]$</td>
<td>$</td>
</tr>
<tr>
<td>6. $\gamma_{1[k'} \gamma_{2,k']} \beta^{\overline{k}} &lt;\text{past}<em>{n</em>{k'}}^{k'}$</td>
<td></td>
</tr>
</tbody>
</table>

---

**Corollary 86:**

If $\gamma \in F_{n,m}^{k} \cup (\forall U (\langle \text{frontier}_{n}^{k'}, \langle \rangle \rangle)^{*} \text{ and } \beta \in \forall'$, then

$$\beta^{\overline{k}} \gamma \rightarrow \gamma \beta$$

and there is a derivation of width less than or equal to $2|\gamma|+3$.

**Proof:**
### TABLE 15. Over for \( \gamma \in \langle \text{frontier}^{k}_{n} F^{k}_{n+1,m-1} \rangle \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( \bar{k}\langle \text{frontier}^{k}<em>{n} \rangle \langle \gamma</em>{1} \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>2. ( \langle \text{frontier}^{k_{1}}<em>{n} \rangle \bar{k}\langle \text{past}^{k</em>{1}}<em>{n+1} \rangle \langle \gamma</em>{1} \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>3. ( \langle \text{frontier}^{k_{1}}<em>{n} \rangle \bar{k}\langle \text{past}^{k</em>{1}}<em>{n+1} \rangle \langle \gamma</em>{1} \rangle \langle \text{past}^{k_{1}}_{n+1} \rangle )</td>
<td>(2</td>
</tr>
<tr>
<td>4. ( \langle \text{frontier}^{k_{1}}<em>{n} \rangle \langle \gamma</em>{1} \rangle \bar{k}\langle \text{past}^{k_{1}}<em>{n+1} \rangle \langle \text{past}^{k</em>{1}}_{n+1} \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>5. ( \langle \text{frontier}^{k_{1}}<em>{n} \rangle \langle \gamma</em>{1} \rangle \bar{k}\langle \text{past}^{k_{1}}_{n} \rangle )</td>
<td>(</td>
</tr>
</tbody>
</table>

### TABLE 16. Over Derivation for \( \gamma \in F^{0}_{n,m} \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( \bar{k}\langle \gamma \rangle )</td>
<td>(2</td>
</tr>
<tr>
<td>2. ( \gamma\bar{k}\langle \text{past}^{l}_{n} \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>3. ( \gamma\bar{k}\langle \text{past}^{0}_{n} \rangle )</td>
<td>(</td>
</tr>
</tbody>
</table>

This follows directly from lemma 85.
Definition 64:

If \( n+1 \geq k \geq 0 \) then \( S_n^k \) is a function,

\[
S_n^k : H_n^k \to U_m \geq 0 F_n^k
\]

such that if \( S_n^k(y) = \psi \), then

\[
<\text{frontier}_n^k y> *\Rightarrow \psi
\]

and \( <\text{frontier}_n^k> \) is the only frontier nonterminal to occur in \( \psi \) and no \( <\text{frontier}_n^k> \) nonterminal is expanded.

Lemma 87:

If \( y \in H_n^k \) and \( S_n^k(y) = \psi \), then

\[
<\text{frontier}_n^k y> *\Rightarrow \psi
\]

has a derivation with a width of less than \( 2|\psi|+5 \).

Also, the length of \( \psi \) may be less than or equal to \( 2|\psi|+1 \).

Proof by induction on \( m \):

Basis: If \( m = 0 \), the derivation is given in table 17. In this case, \( 2|\psi|+5 = 7 \).

Induction step: Assume that the lemma is true for \( m' < m \).

If \( y \in H_n^k, m' \), then

\[
<\text{frontier}_n^k y> *\Rightarrow \psi
\]
TABLE 17. $s_n^k$ Derivation for $\gamma \in H_n^{k+1,0}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle \text{frontier}_n^k \rangle \gamma \langle \rangle$</td>
<td>3</td>
</tr>
<tr>
<td>2. $\gamma$</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 18. $s_n^{k+1}$ Derivation for $\gamma \in \Sigma$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle \text{frontier}_n^{k+1} \rangle \gamma \langle \rangle$</td>
<td>3</td>
</tr>
<tr>
<td>2. $\gamma$</td>
<td></td>
</tr>
</tbody>
</table>

1. If $k = n+1$, then the derivation given in either table 18 or 19. In table 18, $2|\gamma| + 5 = 7$ and, in table 19, $2|\gamma| + 5 = 2|\gamma_2| + 7$.

2. If $k = n$, then the string is already in the required form.

3. If $n > k > 0$, then either the derivation given in table 20 holds, or the derivation given in table 21 is true or $\gamma \in H_n^{k, m-1}$ in which case the lemma is true by
induction hypothesis. In table 21, notice that
\[ 2|\gamma|+5 = 2|\gamma_1|+2|\gamma_2|+9. \]
4. If \( k = 0 \), then the derivation is given in table 22.
Since in each of these cases the width is less than \( 2|\gamma|+5 \), that part of the lemma is proved by induction. Inspection will also show that the output \( \psi \) has a length less than \( 2|\gamma|+1 \) in all cases, so this part of the lemma is proved by induction.

---

TABLE 19. \( S_n^{n+1} \) for \( \gamma \in \mathbb{H}_n^{n+2} \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (&lt;\text{frontier}_n^{n+1}\gamma_1[n+1 \gamma_2 n+1]&lt;&lt;&gt;)&gt;)</td>
<td>(</td>
</tr>
<tr>
<td>2. (&lt;\text{frontier}_n^n\gamma_2&lt;\text{Fl}_n&gt;\gamma_2 n+1]&lt;&lt;&gt;)&gt;)</td>
<td>(2</td>
</tr>
<tr>
<td>3. (&lt;\text{frontier}_n^n\gamma_2&lt;\text{Fl}_n&gt;\gamma_2 n+1]&gt;&lt;\gamma_2&lt;&lt;&gt;)&gt;)</td>
<td>(</td>
</tr>
<tr>
<td>4. (&lt;\text{frontier}_n^n\gamma_2&lt;&lt;&gt;)&gt;)</td>
<td></td>
</tr>
</tbody>
</table>

---
TABLE 20. $S_n^k$ Derivation for $\gamma \in H_n^{k+1, m-1}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $&lt;\text{frontier}_n^k&gt;\gamma&lt;\rangle&gt;$</td>
<td>$</td>
</tr>
<tr>
<td>2. $&lt;\text{frontier}_n^{k+1}&gt;F2_n^k&lt;\text{over}^{k+1}_n&gt;\gamma&lt;\rangle&gt;$</td>
<td>$2</td>
</tr>
<tr>
<td>3. $&lt;\text{frontier}_n^{k+1}&gt;\gamma&lt;F2_n^k&lt;\rangle&gt;$</td>
<td>$</td>
</tr>
<tr>
<td>4. $&lt;\text{frontier}_n^{k+1}&gt;\gamma&lt;\rangle&gt;$</td>
<td>$2</td>
</tr>
</tbody>
</table>

5. $\psi$

Lemma 88:

If $\gamma_1 \in (\forall U \{<\text{frontier}_n^k>, \langle>, \langle>, \rangle\})^*$ and

$\gamma_2 \in F_{n,m}^k \cap (\forall U \{<\text{frontier}_n^k>, \langle>, \rangle\})^*$, then

$<\text{mark}>\gamma_1 <S_{n,j}^k>\gamma_2<\rangle> \Rightarrow \psi_1 \gamma_2$

where $\psi \in F_{n,m}^n \cap (\forall U \{<\text{frontier}_n^k>, \langle>, \rangle\})^*$ has a
derivation with a width of less than $|\gamma_1|+2|\gamma_2|+6$.

Proof by induction on $m$:
TABLE 21. \( S^k_n \) for \( \gamma \in H^{k+1}_{n+1,m-1} \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( \text{&lt;frontier}_n^k \gamma_1[k \gamma_2 k] )</td>
<td>(</td>
</tr>
<tr>
<td>2. ( \text{&lt;frontier}_n^k \gamma_1[k \gamma_2 k] )</td>
<td>( 2</td>
</tr>
<tr>
<td>3. ( \text{&lt;frontier}_n^k \gamma_1[k \gamma_2 k] )</td>
<td>(</td>
</tr>
<tr>
<td>4. ( \text{&lt;frontier}_n^k \gamma_1[k \text{&lt;switch}_n^k \gamma_2 k] )</td>
<td>(</td>
</tr>
<tr>
<td>5. ( \text{&lt;frontier}_n^k \gamma_1[k \text{&lt;switch}_n^k \gamma_2 k] )</td>
<td>(</td>
</tr>
<tr>
<td>6. ( \text{&lt;frontier}_n^k \gamma_1[k \text{&lt;switch}_n^k \gamma_2 k] )</td>
<td>( 2</td>
</tr>
<tr>
<td>7. ( \psi_1[k \text{&lt;switch}_n^k \gamma_2 k] )</td>
<td>( 2</td>
</tr>
<tr>
<td>8. ( \psi_1[k \psi_2 k] )</td>
<td></td>
</tr>
</tbody>
</table>

Basis: If \( m = 0 \), then \( \gamma_2 \in f_n^{k} = X_k \). But there are no
TABLE 22. $S^0_n$ Derivation for $\gamma \in H^0_{n,m}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle \text{frontier}^0_n \rangle \langle \gamma \rangle$</td>
<td>$</td>
</tr>
<tr>
<td>2. $\langle \text{frontier}^1_n \rangle \langle \gamma \rangle$</td>
<td>$2</td>
</tr>
<tr>
<td>3. $\psi$</td>
<td></td>
</tr>
</tbody>
</table>

derivations for this case, so the hypothesis must be false and the lemma is true by default in this case.

Induction step: Assume that the lemma is true for $m' < m$.

If $\gamma \in F^k_{n,m'}$, then

1. If $k = n$, then the derivation is given in table 23. Note that between steps 3 and 4 the nonterminal $\langle \text{return} \rangle$ occurs. It is assumed that this derivation only allows one $\langle \text{return} \rangle$ at a time.

2. If $n > k > 0$ and $j = ki$, then the derivation is given in table 24. Note that if $\gamma_2 \in F^k_{n,m-1}$ then the $\langle S^0^{ki}_n \rangle$ in step 3 will be stranded so that case is not possible. Also $|\gamma_1| + 2|\gamma_2| + 6 = |\gamma_1| + 2|\gamma| + 2|\beta| + 10$.

3. If $n > k > 0$, $\gamma_2 \in F^k_{n,m-1}$ and $j \in F^k_{n}$, then the derivation is given in table 25.
4. If $n > k > 0$, $\gamma_2 \in F^{n+1}_{n,k+1}$ and $j \in F^{k+1}_{n,n}$, then the derivation is given in table 26. In this case, $|\gamma_1| + 2 |\gamma_2| + 6 = |\gamma_1| + 2 |\alpha| + 2 |\beta| + 10$.

5. If $n > k > 0$ and $\gamma_2 \in F^{n+1}_{n,m-1}$, then the lemma is true by induction hypothesis.

6. If $k = 0$, then the derivation is given in table 27. Note that as before the previous cases are used here.
Note that since $<\text{frontier}_n^n>$ is the only frontier nonterminal to be in $\gamma_2$, $\gamma_2$ cannot equal $<\text{frontier}_n^n\gamma>$ if $n > k > 0$. The case where $k = n$ is handled in step 1. Therefore, the lemma is proved.
TABLE 25. $<S1_n^k>$ Derivation for $\gamma_2 \in F_{n+1}^k$ and $j \in F_{n+1}^k$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <code>&lt;mark&gt;</code>$\gamma_1 &lt;S1_n^k&gt;&lt;\gamma_2&gt;)`</td>
<td>$</td>
</tr>
<tr>
<td>2. <code>&lt;mark&gt;</code>$\gamma_1 &lt;S1_n^{k+1}j&gt;&lt;S3_n^k&gt;&lt;over_{n}^{k+1}&gt;&lt;\gamma_2&gt;)`</td>
<td>$</td>
</tr>
<tr>
<td>3. <code>&lt;mark&gt;</code>$\gamma_1 &lt;S1_n^{k+1}j&gt;&lt;\gamma_2 &lt;S3_n^k&gt;&lt;)`</td>
<td>$</td>
</tr>
<tr>
<td>4. <code>&lt;mark&gt;</code>$\gamma_1 &lt;S1_n^{k+1}j&gt;&lt;\gamma_2&gt;)`</td>
<td>$</td>
</tr>
<tr>
<td>5. $\psi \gamma_1 \gamma_2$</td>
<td></td>
</tr>
</tbody>
</table>

Lemma 89:

If $\gamma_1 \in F_{n,m}^k \cap V^*$, $\gamma_2 \in (V \cup \{<\text{frontier}^k_n, <\})^*$ and $\gamma_3 \in U_{n \geq 0} F_{n,m}^{n-1} \cap (V \cup \{<\text{frontier}^n_n, <\})^*$, then

$<A1_n^k><\gamma_1><\gamma_2><<\gamma_3>\> *\Rightarrow \psi \gamma_2<><\gamma_3<>$

where $\psi \in (V+\{<\text{frontier}^n_n, <\})^*$ has a derivation of width

$2|\gamma_1| + |\gamma_2| + 2|\gamma_3| + |\psi| + 6$

Proof by induction on $m$:

Basis: If $m = 0$ and $k < n$, then the derivation is given in table 28. If $k = n$, then the derivation is given in table
### TABLE 26. \( <S_{1n}^{k}> \) for \( \gamma_2 \in P_{n}^{k+1}[k \ F_{n}^{k-1} \ k] \) and \( j \in P_{n}^{k+1} \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( \langle \gamma_1 &lt; S_{1n}^{k} j &gt; a k \beta k \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>2. ( \langle \gamma_1 &lt; S_{1n}^{k+1} j &gt; S_{3n}^{k} &lt; \overline{\gamma_{n}^{k+1}} a k \beta k \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>3. ( \langle \gamma_1 &lt; S_{1n}^{k+1} j &gt; a &lt; S_{3n}^{k} &gt; k \beta k \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>4. ( \langle \gamma_1 &lt; S_{1n}^{k+1} j &gt; a \rangle [k &lt; S_{4n}^{k} &lt; \overline{\gamma_{n}^{k-1}} \beta k \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>5. ( \langle \gamma_1 &lt; S_{1n}^{k+1} j &gt; a \rangle [k \beta &lt; S_{4n}^{k} &gt; k \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>6. ( \langle \gamma_1 &lt; S_{1n}^{k+1} j &gt; a \rangle [k \beta k )</td>
<td>(</td>
</tr>
<tr>
<td>7. ( \psi \gamma_1 a k \beta k )</td>
<td></td>
</tr>
</tbody>
</table>

29. In either case, \( 2|\gamma_1| + |\gamma_2| + |\psi| + 6 = |\gamma_2| + 2|\gamma_3| + |\psi| + 8 \). Note that in table 29 between line 7 and 8 the width is \( |\gamma_2| + 2|\gamma_3| + 8 \). This is less than \( 2|\gamma_1| + |\gamma_2| + 2|\gamma_3| + |\psi| + 6 \) since \( |\gamma_1| \geq 1 \) and \( |\psi| \geq 1 \).

Induction step: Assume that the lemma is true for all \( m' < m \). If \( \gamma_1 \in H_{n,m}^k \), then one of the following 5 cases is true:
TABLE 27. $<S_{1}^{0}_{n}>$ Derivation

| $<\text{mark}>\gamma_{1}<S_{1}^{0}_{n}>\gamma_{2}<\text{mark}>\gamma_{1}<S_{1}^{1}_{n}>\gamma_{2}<\text{mark}>\gamma_{1}<S_{1}^{2}_{n}>\gamma_{2}$ | $|\gamma_{1}|+|\gamma_{2}|+3$  |
|---|---|
| $\psi\gamma_{1}\gamma_{2}$ | $|\gamma_{1}|+2|\gamma_{2}|+6$ |

TABLE 28. $<A_{1}^{k}_{n}>$ Derivation for $\gamma_{1} \in X_{k}$ and $k < n$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $&lt;A_{1}^{k}<em>{n}\gamma</em>{1}&lt;&gt;\gamma_{2},&lt;&gt;\gamma_{3}$&gt;</td>
<td>$</td>
</tr>
<tr>
<td>2. $\gamma_{1}\gamma_{2},&lt;&gt;\gamma_{3}$&gt;</td>
<td></td>
</tr>
</tbody>
</table>

1. If $k = n+1$, then the derivation is given in table 30. Here, $2|\gamma_{1}|+|\gamma_{2}|+2|\gamma_{3}|+|\psi|+6 = |\gamma_{2}|+2|\gamma_{3}|+|\psi|+8$.
2. If $n \geq k > 0$ and $\gamma_{1} \in H_{n,m-1}^{k+1}$, then the derivation is given in table 31.
3. If $n \geq k > 0$ and $\gamma_{1} \in H_{n,m-1}^{k}$, then the lemma is true by induction hypothesis.
TABLE 29. \(<Al_n^N>\) Derivation for \(\tau_1 = X^n_j\)

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (&lt;Al_n^N&gt;\gamma_1&gt;&lt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>2. (&lt;search_{n-j}^n&gt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>3. (&lt;mark&gt;&lt;setup_{n-j}^n&gt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>4. (&lt;mark&gt;&lt;setup_{n-j}^n&gt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>5. (&lt;mark&gt;&lt;setup_{n-j}^n&gt;\gamma_2&gt;&lt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>6. (&lt;mark&gt;&lt;setup_{n-j}^n&gt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>7. (&lt;mark&gt;&lt;setup_{n-j}^n&gt;\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
<tr>
<td>8. (\psi\gamma_2&gt;&lt;\gamma_3&gt;&lt;\gamma_2</td>
<td>\gamma_3</td>
</tr>
</tbody>
</table>

4. If \(n \geq k > 0\) and \(\gamma \in H_{n,m-1}^{k+1} [H_{n,m-1}^{k-1} k]\), then the derivation is given in table 32. Here,
\[2|\gamma_1|+|\gamma_2|+|\gamma_3|+|\psi|+6 = 2|\alpha|+2|\beta|+|\gamma_2|+2|\gamma_3|+|\psi_1|+|\psi_2|+12.\]

5. If \(k = 0\), then the derivation is given in table 33.
TABLE 30. $\langle A_1^{n+1} \rangle$ Derivation

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle A_1^n \rangle \langle 1 \rangle \langle 2 \rangle \langle 3 \rangle$</td>
<td>$</td>
</tr>
<tr>
<td>2. $\langle 1 \rangle \langle 2 \rangle \langle 3 \rangle$</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 31. $\langle A_1^k \rangle$ Derivation for $\gamma_1 \in H_n^{k+1}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle A_1^n \rangle \langle 1 \rangle \langle 2 \rangle \langle 3 \rangle$</td>
<td>$</td>
</tr>
<tr>
<td>2. $\langle A_1^n \rangle \langle A_2^k \rangle \langle \overline{2} \rangle \langle 1 \rangle \langle 2 \rangle \langle 3 \rangle$</td>
<td>$2</td>
</tr>
<tr>
<td>3. $\langle A_1^n \rangle \langle 1 \rangle \langle A_2^k \rangle \langle 2 \rangle \langle 3 \rangle$</td>
<td>$</td>
</tr>
<tr>
<td>4. $\langle A_1^n \rangle \langle 1 \rangle \langle 2 \rangle \langle 3 \rangle$</td>
<td>$2</td>
</tr>
<tr>
<td>5. $\gamma_2 \langle 3 \rangle$</td>
<td></td>
</tr>
</tbody>
</table>

In each case, the required bound is correct. Therefore, the lemma is proved.
TABLE 32. $\langle A^k_n \rangle$ for $\gamma_1 \in H^{k+1}_{n,m-1}[k, H^{k-1}_{n,m-1} k]$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle A^k_n \rangle \alpha (k, \beta, k) \gamma_2, \gamma_3$</td>
<td>$</td>
</tr>
<tr>
<td>2. $\langle A^k_n \rangle \alpha A^3_n \langle \overline{\gamma} \rangle \alpha (k, \beta, k) \gamma_2, \gamma_3$</td>
<td>$2</td>
</tr>
<tr>
<td>3. $\langle A^k_n \rangle \alpha A^3_n \langle \overline{\gamma} \rangle \alpha (k, \beta, k) \gamma_2, \gamma_3$</td>
<td>$</td>
</tr>
<tr>
<td>4. $\langle A^k_n \rangle \alpha \langle \gamma \rangle \langle \overline{\gamma} \rangle \alpha (k, \beta, k) \langle \overline{\gamma} \rangle \gamma_2, \gamma_3$</td>
<td>$</td>
</tr>
<tr>
<td>5. $\langle A^k_n \rangle \alpha \langle \gamma \rangle \langle \overline{\gamma} \rangle \alpha (k, \beta, k) \langle \overline{\gamma} \rangle \gamma_2, \gamma_3$</td>
<td>$</td>
</tr>
<tr>
<td>6. $\langle A^k_n \rangle \alpha \langle \gamma \rangle \langle \overline{\gamma} \rangle \alpha (k, \beta, k) \langle \overline{\gamma} \rangle \gamma_2, \gamma_3$</td>
<td>$2</td>
</tr>
<tr>
<td>7. $\phi_1 (k, \beta) \gamma_2, \gamma_3$</td>
<td>$2</td>
</tr>
<tr>
<td>8. $\phi_1 (k, \phi_2, k) \gamma_2, \gamma_3$</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 33. \( \langle A_1^n \rangle \) Derivation

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( \langle A_1^n \rangle y_1 \rangle y_2 \rangle y_3 \rangle )</td>
<td>(</td>
</tr>
<tr>
<td>2. ( \langle A_1^n \rangle y_1 \rangle y_2 \rangle y_3 \rangle )</td>
<td>( 2</td>
</tr>
<tr>
<td>3. ( \psi y_2 \rangle y_3 \rangle )</td>
<td></td>
</tr>
</tbody>
</table>

Lemma 90:
If \( y \in H_{n+1,m'} \), then
\( \langle \text{frontier}_n^k y \rangle \) \( \Rightarrow \) \( \psi \)

where \( \psi \in \mathcal{V}^* \) then there is a derivation of width
less than
\( 4|y| + 5|\psi| + 3 \)

Proof by induction on \( m \):

Basis: If \( m = 0 \), then the derivation is given in table 34.

TABLE 34. Frontier Derivation for Basis

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( \langle \text{frontier}_n^k y \rangle )</td>
<td>3</td>
</tr>
<tr>
<td>2. ( y )</td>
<td></td>
</tr>
</tbody>
</table>
Induction step: Assume that the lemma is true for all \( m' < m \). If \( \gamma \in H_{n+1,m}^k \) then

1. If \( k = n+1 \) and \( \gamma \in I \), then the derivation is given in table 34.

2. If \( k = n+1 \) and \( \gamma = a[n+1 \beta n+1] \), then the derivation is given in table 35. Here,
\[
4|\gamma| + 5|\psi| + 3 = 4|\beta| + 5|\psi| + 15.
\]

**TABLE 35. Frontier Derivation for \( \gamma = a[n+1 \beta n+1] \)**

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (&lt;\text{frontier}_n^{n+1}a[n+1 \beta n+1]&lt;&gt;))</td>
<td>(</td>
</tr>
<tr>
<td>2. (&lt;\text{frontier}_n^n&lt;\text{Fl}<em>n&gt;&lt;\text{over}^n</em>{n+1}\beta n+1&lt;&gt;))</td>
<td>(2</td>
</tr>
<tr>
<td>3. (&lt;\text{frontier}_n^n\beta&lt;\text{Fl}_n&gt;n+1&lt;&gt;))</td>
<td>(</td>
</tr>
<tr>
<td>4. (&lt;\text{frontier}_n^n\beta&lt;&gt;))</td>
<td>(4</td>
</tr>
<tr>
<td>5. (\psi)</td>
<td></td>
</tr>
</tbody>
</table>

3. If \( k = n \) and \( \gamma \in H_{n+1,m-1}^{n+1} \), then the derivation is given in table 36.
TABLE 36. Frontier Derivation for $k = n, \gamma \in H_{n+1}^n$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $&lt;\text{frontier}^n_n\gamma&gt;$</td>
<td>$</td>
</tr>
<tr>
<td>2. $&lt;\text{frontier}^{n+1}<em>n\gamma&gt;&lt;\text{die}&gt;&lt;\text{over}^{n+1}</em>{n+1}\gamma&gt;$</td>
<td>$2</td>
</tr>
<tr>
<td>3. $&lt;\text{frontier}^{n+1}_n\gamma&gt;&lt;\text{die}&gt;$</td>
<td>$</td>
</tr>
<tr>
<td>4. $&lt;\text{frontier}^{n+1}_n\gamma&gt;$</td>
<td>$4</td>
</tr>
<tr>
<td>5. $\psi$</td>
<td></td>
</tr>
</tbody>
</table>

4. If $k = n$ and $C \in H_{n+1,m-1}^n$, then the lemma is true by induction hypothesis.

5. If $k = n$, $\gamma = \alpha[n\beta_n]$ and $\alpha \in \Sigma$, then the derivation is given in table 37. Here,
$$4|\gamma|+5|\psi|+3 = 4|\beta|+5|\psi|+30.$$ 

6. If $k = n$, $\gamma = \alpha[n\beta_n]$ and $\alpha \notin \Sigma$, then the derivation is given in table 38. Here,
$$4|\gamma|+5|\psi|+3 = 4|\alpha|+4|\beta|+5|\psi|+11.$$ Notice that $\psi_0$ is the frontier of $\alpha$ and must be shorter than or equal to $|\psi|$ since each "X" that is replaced in $\psi_0$ is replaced with at least one character. $\psi_1 = S_{n-1}^n(\beta)$ and by lemma 87.
TABLE 37. Frontier for $k = n$, $\gamma = \alpha[n \beta n]$ and $\alpha \in \Sigma$.

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\langle \text{frontier}_n^n \alpha[n \beta n] \rangle$</td>
<td>$</td>
</tr>
<tr>
<td>2. $\alpha[n \langle \text{frontier}_n^{n-1} \rangle \langle \text{switch}_n \rangle \langle \text{over}_n^{n-1} \beta n \rangle]$</td>
<td>$2</td>
</tr>
<tr>
<td>3. $\alpha[n \langle \text{frontier}_n^{n-1} \beta \rangle \langle \text{switch}_n n \rangle]$</td>
<td>$</td>
</tr>
<tr>
<td>4. $\alpha[n \langle \text{frontier}_n^{n-1} \beta \rangle_n]$</td>
<td>$4</td>
</tr>
<tr>
<td>5. $\alpha[n \psi_2 n]$</td>
<td></td>
</tr>
</tbody>
</table>

$|\psi_1| \leq 2|\beta| + 1$

Between step 11 and step 12 $\psi_2$ is the same as $\psi_1$ except that the frontier has been taken of everything that is to be copied into $\psi_0$. The width is calculated as follows. Let

$$|\psi_1| = \psi_{11} + 2\psi_{12} + \psi_{13}$$

where $\psi_{11}$ is the length of the portion of $\psi_1$ which is not copied, $\psi_{12}$ is the number of frontier operations in that portion of $\psi_1$ that is copied, and $\psi_{13}$ is the...
total of the length of the arguments to the frontier operations in that portion of $\psi_1$ that is not copied. Let

$$|\psi_2| = \psi_{11} + \psi_{22}$$

where $\psi_{22}$ is the sum of the lengths of all of the frontiers in $\psi_2$ that are copied. Let $X$ be the number of "X"'s in $\psi_0$. Then, if the frontiers are taken one at a time the width must be less than

$$|\psi_0| + \psi_{11} + 4\psi_{13} + 5\psi_{22} + 8$$

It would be easy to show that $\psi_{11} \leq 2(|\beta| - \psi_{13}) + 1$ using lemma 87 as a guide. Therefore, a bound is

$$|\psi_0| + 2|\beta| + 2\psi_{13} + 5\psi_{22} + 9$$

Since $\psi_{13} \leq |\beta|$ another bound is

$$|\psi_0| + 4|\beta| + 5\psi_{22} + 9$$

Now, consider the two possible cases:

1. If $\psi_0 = X_n^j$, then a bound is $4|\beta| + 5\psi_{22} + 10$
   since $|\psi_0| = 1$. Also, $\psi_{22} \leq |\psi|$ so another bound is $4|\beta| + 5|\psi| + 10$.

2. If $\psi_0 \neq X_n^j$, then it can be shown by induction that $|\psi_0| \geq 2X$. It is also true that
$|\psi| \geq |\psi_0| + \psi_{22} - X$ since all of $\psi_0$ except the X's appears in $\psi$, and by definition all of $\psi_{22}$ appears in $\psi$. Therefore,

\[ 2|\psi| \geq 2|\psi_0| + 2\psi_{22} - 2X \]
\[ \geq |\psi_0| + 2\psi_{22} + (|\psi_0| - 2X) \]
\[ \geq |\psi_0| + 2\psi_{22} \]

Therefore,

\[ 5|\psi| \geq |\psi_0| + 5\psi_{22} \]

and a possible bound is $4|\beta| + 5|\psi| + 9$.

In either case, a bound is $4|\beta| + 5|\psi| + 10$.

Between 12 and 13 the bound is $2|\psi_0| + 2|\psi_2| + |\psi| + 7$ by lemma 89. But $|\psi| \geq |\psi_0|$ and

\[ |\psi_2| = \psi_{11} + \psi_{22} \]
\[ \leq \psi_{11} + |\psi| \]
\[ \leq |\psi| + 2|\beta| + 1 \]

by lemma 87 so a bound is $4|\beta| + 5|\psi| + 9$. Similarly for the rest of the table.
TABLE 38. Frontier for $k = n, \gamma = \alpha[n \beta n]$ and $\alpha \notin \Sigma$.

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. &lt;frontier$_n^n\alpha[n \beta n]\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>2. &lt;apply$_n^n\alpha[F3_n]n_1^n_1\alpha[n \beta n]\rangle$</td>
<td>$2</td>
</tr>
<tr>
<td>3. &lt;apply$_n^n\alpha[F3_n]n_1^n_1\alpha[n \beta n]\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>4. &lt;apply$_n^n\alpha[F3_n]n_1^n_1\alpha[\overline{n_1^n_1}\alpha[n \beta n]\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>5. &lt;apply$_n^n\alpha[F3_n]n_1^n_1\alpha[\overline{n_1^n_1}\beta[F4_n]n\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>6. &lt;apply$_n^n\alpha[F3_n]n_1^n_1\alpha[\overline{n_1^n_1}\beta[\psi_0]n\rangle$</td>
<td>$4</td>
</tr>
<tr>
<td>7. &lt;apply$_n^n\alpha[\psi_0]n_1^n_1\beta[\psi_0]n\rangle$</td>
<td>$</td>
</tr>
<tr>
<td>8. &lt;apply$_n^n\alpha[\psi_0]n_1^n_1\beta[\psi_0]n\rangle$</td>
<td>$</td>
</tr>
</tbody>
</table>
TABLE 38. Continued

9. \( \langle A_2 \rangle \langle A_1^n \psi_0 \rangle \langle A_4 \rangle \langle, \rangle \langle \text{frontier}_n^{n-1} \beta \rangle \langle \rangle \rangle \)
   \(|\beta| + |\psi_0| + 7\)

10. \( \langle A_2 \rangle \langle A_1^n \psi_0 \rangle \langle, \rangle \langle \text{frontier}_n^{n-1} \beta \rangle \langle \rangle \rangle \)
    \(2|\beta| + |\psi_0| + 10\)

11. \( \langle A_2 \rangle \langle A_1^n \psi_0 \rangle \langle, \rangle \langle \psi_1 \rangle \)
    \(4|\beta| + 5|\psi| + 10\)

12. \( \langle A_2 \rangle \langle A_1^n \psi_0 \rangle \langle, \rangle \langle \psi_2 \rangle \)
    \(4|\beta| + 5|\psi| + 9\)

13. \( \langle A_2 \rangle \langle \psi_1, \psi_2 \rangle \)
    \(2|\beta| + 2|\psi| + 4\)

14. \( \psi \langle A_2 \rangle \langle, \rangle \langle \psi_2 \rangle \)
    \(2|\beta| + |\psi| + 5\)

15. \( \psi \langle \text{blitz} \rangle \langle \text{end} \rangle \langle \text{over}_n^{n-1} \psi_2 \rangle \)
    \(4|\beta| + 3|\psi| + 7\)

16. \( \psi \langle \text{blitz} \rangle \langle \psi_2 \rangle \langle \text{end} \rangle \langle \rangle \)
    \(2|\beta| + 2|\psi| + 4\)

17. \( \psi \)

7. If \( n > k > 0 \) and \( \bar{\varepsilon} \in H_n^{k+1, m-1} \), then the derivation is given in the table 39.

8. If \( n > k > 0 \) and \( \bar{\varepsilon} \in H_n^{k, m-1} \), then the lemma is true by induction hypothesis.
TABLE 39. Frontier Derivation for $\gamma \in H_{n+1}^{k+1}$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $&lt;\text{frontier}_{n}^{k}&gt;\gamma&lt;$&gt;</td>
<td>$</td>
</tr>
<tr>
<td>2. $&lt;\text{frontier}<em>{n}^{k+1}&gt;\gamma&lt;\text{F2}</em>{n}^{k}&lt;\text{over}_{n+1}^{k+1}&lt;$&gt;</td>
<td>$2</td>
</tr>
<tr>
<td>3. $&lt;\text{frontier}<em>{n}^{k+1}&gt;\gamma&lt;\text{F2}</em>{n}^{k}&lt;$&gt;</td>
<td>$</td>
</tr>
<tr>
<td>4. $&lt;\text{frontier}_{n}^{k+1}&gt;\gamma&lt;$&gt;</td>
<td>$4</td>
</tr>
<tr>
<td>5. $\psi$</td>
<td></td>
</tr>
</tbody>
</table>

9. If $n > k > 0$ and $\gamma = a[\beta^k]$, then the derivation is given in table 40. Note that $4|\gamma|+5|\psi|+3 = 4|a|+4|\beta|+5|\psi_1|+5|\psi_2|+21$.

10. If $k = 0$, then the derivation is given in table 41. In this case, note that the derivations given in the previous steps must be used to complete the proof.

Now, since each of these are bound by $4|\gamma|+5|\psi|+3$, the lemma is proved by induction.
TABLE 40. Frontier Derivation for γ = α[k β k]

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( &lt;\text{frontier}_{k}^{n}\alpha[k \beta k]) &gt;</td>
<td>(</td>
</tr>
<tr>
<td>2. ( &lt;\text{frontier}<em>{k}^{n+1}\alpha[F2</em>{n}\overline{\text{over}}_{n}^{k+1}\alpha[k \beta k]) &gt;</td>
<td>(2</td>
</tr>
<tr>
<td>3. ( &lt;\text{frontier}<em>{k}^{n+1}\alpha[F2</em>{n}^{k}\beta k]) &gt;</td>
<td>(</td>
</tr>
<tr>
<td>4. ( &lt;\text{frontier}<em>{k}^{n+1}\alpha[k\text{switch}</em>{n}\overline{\text{over}}_{n}^{k-1}\beta_k]) &gt;</td>
<td>(</td>
</tr>
<tr>
<td>5. ( &lt;\text{frontier}<em>{k}^{n+1}\alpha[k\text{switch}</em>{n}^{k-1}\beta_k]) &gt;</td>
<td>(</td>
</tr>
<tr>
<td>6. ( &lt;\text{frontier}<em>{k}^{n+1}\alpha[k\beta</em>{k}]) &gt;</td>
<td>(</td>
</tr>
<tr>
<td>7. ( &lt;\text{frontier}_{k}^{n+1}\alpha[k ψ_2 k]) &gt;</td>
<td>(4</td>
</tr>
<tr>
<td>8. ( ψ_1[k ψ_2 k])</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 41. Frontier Derivation for $k = 0$

<table>
<thead>
<tr>
<th>Line</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $&lt;\text{frontier}^0_n\gamma&gt;\gamma&lt;\gamma&gt;$</td>
<td>$</td>
</tr>
<tr>
<td>2. $&lt;\text{frontier}^1_n\gamma&gt;\gamma&lt;\gamma&gt;$</td>
<td>$4</td>
</tr>
<tr>
<td>3. $\phi$</td>
<td></td>
</tr>
</tbody>
</table>

Lemma 91:
If $\gamma \in \mathcal{H}^k_{n+1,m'}$ then

$$<\text{frontier}^k_n\gamma>\gamma<\gamma> \Rightarrow \Rightarrow \phi$$

where $\phi \in \mathcal{V}^*$ then

$$\psi = \text{frontier}^k_n(\gamma)$$

Proof:
This can be established by a simple but lengthy induction on $m$. It would be necessary to prove a similar result as a lemma for search and apply, but these will not be written down here.

---

Lemma 92:
If $\gamma \in \mathcal{H}^k_n$, then

$$<\text{string}^k_n\gamma>\gamma<\gamma> \Rightarrow \Rightarrow \phi$$

where $\phi \in \mathcal{V}^*$ then $\psi = \text{string}(\gamma)$. 
Proof:
This is a simple induction on n using lemma 91.

Lemma 93:
Given the grammar presented, the language is the same as the string language of the original grammar. Also, if a linearly bounded sequence of frontiers can be found for each element of the string language, then the string language has a context sensitive grammar.

Proof:
The first part of this lemma follows directly from lemma 92. The second part requires inductive type proof using lemma 90.

This last lemma reduces the problem to one of showing that for every hypertree grammar there is a (possible the same) hypertree grammar which has the same string language and each element of the string language has a linearly bounded set of frontiers. The approach that will be used here is to generate a normalized grammar for the language, then some alternations are made on the language which do not effect the resulting string language. Finally, this grammar is presented as a linearly bounded grammar, proving the language to be context sensitive.
Lemma 94:

If $G_n^k = (\Sigma, V, P, S)_n^k$ is a normal hypertree grammar, then there is a grammar $G_n^{k'} = (\Sigma', V', P', S')_n^k$ and a $K$ such that

$$\text{string}(G_n^k) = \text{string}(G_n^{k'})$$

and is $a \in \text{string}(G_n^{k'})$ then there is a $\psi$ such that

$$\text{extract}_n(\psi) = a$$

and

$$\text{extract}_j(\psi) < K|a|$$

for all $n \geq j \geq 1$.

Proof:

Construct a completed linear grammar for $\text{string}(G_n^k)$ by constructing a linear grammar as was done in corollary 70 in chapter 7. Then, construct a completed grammar as was done in theorem 63 in chapter 7. Call this new grammar $G_n^{k'}$. Let

$V = \{X_n^{i,j} \mid X_n^{i,j} \text{ appears in } G_n^{k'}\}$.

Let $K_1$ be the number of characters in the righthand side of the longest production in $G_n^{k'}$.

If $A \Rightarrow \psi$ and $a = \text{extract}_i(\psi)$ where $a \in \Sigma U U_{n>0}X_n^i$, then define $G(A, \psi, a)$ as the number of characters in the longest $\psi_1$ such that

$$\psi_1 = \text{extract}_j(\psi)$$

for some $j$. Let

$$K_2 = \max_{A, a} (\min_{\psi} (G(A, \psi, a)))$$
If $A \Rightarrow \psi$, $\alpha = \text{string}(\psi)$ and $\psi$ is the smallest hypertree for which this is true, then a limit on the length of $\text{extract}_j(\psi)$ is

$$((\frac{|V|}{K_2 + K_1 + 4})|V||V||V|)^{j-1}$$

Proof by induction on $j$:

Let $K_3 = (\frac{|V|}{K_2 + K_1 + 4})|V||V||V|$.

Basis: If $j = 1$, then

$$\text{extract}_j(\psi) = \alpha$$

so that

$$|\text{extract}_j(\psi)| = |\alpha|$$

$$\leq K_3^0|\alpha|$$

$$\leq K_3^{j-1}|\alpha|$$

Induction step: Assume that if $j' < j$ then

$$|\text{extract}_{j'}(\psi)| \leq K_3^{j'-1}|\alpha|$$

Notice that in the definition of frontier only two lines result in a frontier that is shorter than the frontiered hypertree. These are

1. $\text{frontier}_{n+1}^n(\alpha[\beta]) = \text{frontier}_{n}^n(\beta)$.
2. $\text{frontier}_{n}^n(\alpha[\beta]) = \text{apply}_{n}(\text{frontier}_{n+1}^n(\alpha), \text{frontier}_{n}^n(\beta))$. 
Define a function $F$ which maps a hypertree to a set of nonterminals. In particular, $F(\alpha) = \{A \mid \text{extract}_{k'}(\alpha) = \text{extract}_{k'}(\psi') \text{ where } \psi' \text{ is that portion of } \alpha \text{ generated by nonterminal } A\}$. It should be obvious that

$$|F(\alpha)| \leq |V|$$

for all $\alpha$. It should also be obvious that if the nonterminal that generates $\alpha$ is in $F(\alpha)$ then there is a shorter hypertree which could have been used to generate the same frontier as $\alpha$. This limits the maximum length of any recursion which does not change the frontier to $|V|$ in $\psi$.

It is possible that the hypertree may be changed by just changing the $X$'s and not changing the length, but there are less than $|V||V|$ possible ways that the $X$'s can be arranged. Note that once two $X$'s become the same during the frontiering process they must remain the same throughout any further recursions. The total number of recursions that need to be allowed, therefore, are fewer than $|V||V||V|$. Note that the number of characters dropped by one recursion is at most $|V|K_2K_1+3$ since

1. If the recursion is of type one, then only 3 characters are dropped.
2. If the recursion is of type two, then less than $K_1$ characters can be deleted from the $\beta$ in the
equation since $\beta$ has a maximum length of $K_1$. Also, a maximum of $|V|K_2 + K_1 + 3$ characters can be deleted from the $a$ since

1. If $a = a_{n+1 \beta_{n+1}}$, $\text{frontier}(\beta_1) \neq X^n_i$ and all of $\beta$ is dropped, then $\text{frontier}(\beta_1)$ is the hypertree that is copied. Therefore, 6 characters ($a$, $[n+1, n+1]$, $[n, * \mid n]$) are dropped in this case.

2. If $a = a_{n+1 \beta_{n+1}}$, $\text{frontier}(\beta_1) \neq X^n_i$ and not all of $\beta$ is deleted, and since by design $\text{frontier}(\beta_1) = \text{frontier}(a_{n \beta n})$, therefore, it must be true that the only copying that is allowed is for one $X$ to replace itself with another $X$. Each of these $X$'s must have been generated by the shortest derivation possible since $\psi$ is the shortest hypertree. Therefore, each of these hypertrees must have a length of less than $K_2$ characters. In addition, there are at most $|V|$ of these, plus at most $K_1$ characters which are dropped from the $\beta$ directly plus 3 characters ($a$, $[n+1, n+1]$) that are dropped from the $a$. That is

$$|V|K_2 + K_1 + 3$$

characters are dropped.
3. If $\alpha = a_{n+1} \beta_1 a_{n+1}$ where $\text{frontier}(\beta_1) = x_n$, then at most $K_2$ characters are dropped from $\beta_1$ by definition of $K_2$ and less than $K_1$ characters are dropped from $\beta$. Therefore, at most $K_1 + K_2 + 3$ characters are dropped in this case.

In each of these cases, fewer than $|v|K_2 + K_1 + 3$ characters are dropped so

$$|\text{extract}_{j+1}(\alpha)| < K_3 |\text{extract}_j(\alpha)| < (K_3)^{j-1} |\alpha|$$

and the conjecture is proven by induction. Also, if $K = (K_3)^{n-1}$ the lemma is proved.

---

Theorem 95:

If $G_n^k$ is a regular hypertree grammar, then the string language for $G_n^k$ is context sensitive.

Proof:

This follows directly from lemma 94 and lemma 93.

---
Note that the linear bound on this language potentially quite large, so that it may not always be readily apparent what the context sensitive grammar is, however, given patience, this paper will eventually allow its construction.
CONCLUSION

The results shown, therefore, illustrate that there is an infinite hierarchy of languages called the algebraic language hierarchy. This hierarchy can be expressed via hypertrees -- using either grammars, automata or regular expressions. It has also been shown that this entire hierarchy is not as powerful as the set of context sensitive languages. It is hoped that further research will be done to determine exactly what can and cannot be expressed using this hierarchy. Also, research is currently being done to determine a feasible algorithm for parsing languages in this hierarchy.

Some of the potential usefulness of this hierarchy can be seen by comparing the set of three languages given by Aho and Ullman (Aho and Ullman 1977) modeling various real programming language constructs which are not context free. In each case, a macro grammar can be easily given which generates the language. Although this is greatly oversimplified, perhaps the additional power of the higher levels of the hierarchy can model the more complex real life situation.

Quite probably the concept of hypertrees may find uses which do not directly relate to the specification of programming language. Already, Dr. George Strawn of Iowa
State has found a very practical use for hypertree modeling in the area of idiom matching (Strawn 1982). Trees may be used to effectively encode the parse of a string language in most programming languages. However, recently some languages have been developed which are written in a tree format. It may be possible to parse these languages using a hypertree of level three in place of the parse tree. This last suggestion might be extended to the parse tree itself so that in addition to the parse tree there would be the parse of the parse tree. Lastly, it might be possible to use the concept of hypertrees as a starting point to find other types of grammars, etc. for specifying languages.

Research is already being done to answer the question, "What happens if the paths are redefined to eliminate the restriction that an \( n^{th} \) level path cannot have any \( n \)'s in it?"

Throughout this entire paper it has been assumed that the set of hypertrees did not include the null hypertree. This was just a matter of taste as was indicated in the introduction -- it would have been equally as easy to have a characterization which included a null hypertree. Therefore, an alternative definition of hypertree might be
Definition 65:

The set of hypertrees over $\Sigma$ of level $n$, denoted $H_n(\Sigma)$, is the smallest set such that

(a) $H_{-1}(\Sigma) = \{\lambda\}$

(b) $H_n(\Sigma) = \Sigma \cup X_n \cup \Sigma[H_{n-1}(H_n(\Sigma))_k]$

In this definition, $\lambda$ is the null hypertree. Note that definition of a path would have to be altered to allow a $P_n^{-1} = \emptyset$. Although this addition could possibly make some major changes in the proofs, it is felt that this will not significantly effect the final results.

And finally, in this paper an exploration has been made of the IO hierarchy because the author felt it was the most natural. What happens with the OI hierarchy? It can easily be seen that it is a true hierarchy since the language hierarchy given for the IO case will, with minor modifications, show the hierarchy for the OI case. The proof of containment in context sensitive language, although potentially simplified by the proofs in this paper, is by no means a trivial proof. It has been shown that OI macro grammars can model a language which requires variables to be declared. It might, therefore, be possible that the OI hierarchy may model some aspects of real languages even better than the IO algebraic language hierarchy that has been presented here.


Baker, B. S. "Composition of Top-down and Bottom-up Tree Transducers". Information and Control 41 (1979): 186-213.


ACKNOWLEDGEMENTS

I would like to acknowledge the contribution of Dr. George O. Strawn to the completion of this paper. His constant advice and insights were greatly appreciated. I would like to also express appreciation to those graduate students who were involved in the various aspects of this research. I also appreciate the contribution of the other members of my graduate committee for being an inspiration to me in various ways. I appreciate the contribution of the other members of the Iowa State faculty who have inspired me through the years.

Also, appreciation is extended to the Iowa State Computer Science Department and Computer Center for the computer time and resources used in the editing of this paper.

Any set of acknowledgements must also include some of the people who contributed by supplying the background that was needed for this project. In particular, Mr. Jack Freeze, my junior high mathematics teacher, who recognized the talent that I had and was willing to inspire me to pursue it. Also, Mr. Leonard Johnsen, my high school mathematics teacher who taught me to think logically.

I would like to also acknowledge the contribution of my parents, whose loving home was a constant source of strength
to me. I particularly appreciated the contribution of my father who often put my educational needs above his own interests. For this I am very grateful.

Last, but far from least, I appreciate the contribution of my wife, Feng-chuan Hsiao. In a practical sense, she entered most of this dissertation onto the computer. It is to her that I dedicate this paper, as I have already dedicated my life.
APPENDIX: LINEARLY BOUNDED GRAMMARS

This appendix is presented to establish a convenient method of proving that a given grammar describes a context sensitive language. This method, called linearly bounded grammar, is used extensively in this paper. Note that this is somewhat similar to the approach taken by Kuroda (Kuroda 1964) and Landweber (Landweber 1963) except that the major thrust in both papers was the development of linearly bounded automata whereas this proof emphasizes the grammars. Also, the idea of "linearly bounded" is much more restrictive in these papers.

In order to simplify the proofs, a normal form is established which is an extension of Chomski Normal form to unrestricted grammars. In this normal form, productions are allowed to be in the form

\[ A \rightarrow a \]

\[ A \rightarrow BC \]

\[ AB \rightarrow C \]

\[ A \rightarrow \lambda \]
where capital letters represent nonterminals and small letters represent terminals. It will be seen later that an important concept involved in linearly bounded grammars is that of the width of a derivations. To define this, derivation will be formally defined.

Definition 66:
A derivation for $\gamma$ over grammar $G=(\Sigma, V, P, S)$ where $\gamma \in \text{language}(G)$ is a sequence of sentential forms

$$\gamma_0, \gamma_1, \ldots, \gamma_n$$

such that $S = \gamma_0$, $\gamma_i \Rightarrow \gamma_{i+1}$ for all $n > i \geq 0$ and $\gamma_n = \gamma$.

Definition 67:
The width $w$ of derivation $\gamma_0, \gamma_1 \ldots \gamma_n$ is defined as

$$w = \max_{n \geq i \geq 0}(\vert \gamma_i \vert)$$

It will be shown that the normal form grammar described preserves the width of derivation. This fact simplifies proofs involving linearly bounded grammars.
Lemma 96:

For every grammar $G = (\Sigma, V, P, S)$, there is a corresponding grammar $G' = (\Sigma', V', P', S')$ such that

$$\text{language}(G) = \text{language}(G')$$

and

$$P' \subseteq (V'^+ \rightarrow V'^*) \cup (V' \rightarrow \Sigma')$$

and if there is a derivation for $\gamma$ over $G$ of width $w$, then there is also a derivation for $\gamma$ over $G'$ of width $w$.

Proof:

Construct $G'$ as follows:

1. $\Sigma' = \Sigma$.
2. $V' = V \cup \Sigma''$ where $\Sigma''$ has elements of the form $(a)$ where $a \in \Sigma$.
3. $P'$ is described below.
4. $S'$ is the same as $S$.

$P'$ is constructed as follows:

1. If $\alpha \rightarrow \beta$ is in $P$, then $\alpha' \rightarrow \beta'$ is in $P'$ where $\alpha'$ is $\alpha$ with all terminals replaced by the corresponding element from $\Sigma''$, and $\beta'$ is $\beta$ with its terminals replaced by the corresponding nonterminal.
2. If $a \in \Sigma$ and appears in $P$, then add $(a) \rightarrow a$.
3. These are all the elements of $P'$.

Claim: $G'$ is the required grammar.
Proof:
If $\gamma \in \text{language}(G')$ and $\gamma_0', \gamma_1', \ldots, \gamma_n'$ is a derivation for $\gamma$, then construct a new sequence $\gamma_0'', \gamma_1'', \ldots, \gamma_n''$ such that for all $n' \geq i \geq 0$, $\gamma_i''$ is the same as $\gamma_i'$ except that each element of $\Sigma''$ in $\gamma_i'$ is replaced by the corresponding element from $\Sigma$. Therefore, it must be true that if $n' \geq i \geq 0$ then

$$S \Rightarrow \gamma_i''$$

since, by induction.

Basis: If $i = 0$, then $\gamma_i'' = S$ but

$$S \Rightarrow S$$

by definition.

Induction step: Assume that if $i' \leq i$ then

$$S \Rightarrow \gamma_i'''$$

Since $\gamma_i' \Rightarrow \gamma_{i+1}'$, it must be true that the production used in the construction of $\gamma_{i+1}'$ either come from the first half of the construction for $P'$ or the second half. If it is from the second half, that is, of the form

$$(a) \rightarrow a$$

then $\gamma_i'' = \gamma_{i+1}''$ since $\gamma_i''$ and $\gamma_{i+1}''$ differ only in that the chosen element in $\Sigma''$ is replaced by its corresponding element in $\Sigma$. If it is from the first half, then
\[ \gamma_i'' \Rightarrow \gamma_{i+1}'' \]
since by definition \( \gamma_i = \gamma_{i1} \alpha' \gamma_{i3} \) and
\[ \gamma_{i+1} = \gamma_{i1} \beta' \gamma_{i3} \]
where \( \gamma' \rightarrow \beta' \in P' \). But, by construction this implies that
\[ \alpha \rightarrow \beta \in P \]
where \( \alpha \) is the same as \( \alpha' \) except elements of \( \Sigma'' \) are replaced by the corresponding elements of \( \Sigma \) etc. Also,
\[ \gamma_i'' = \gamma_{i1}'' \alpha \gamma_{i3}'' \]
and
\[ \gamma_{i+1}'' = \gamma_{i1}'' \beta \gamma_{i3}'' \]
where \( \gamma_{i1}'' \) and \( \gamma_{i3}'' \) are the same as \( \gamma_{i1} \) and \( \gamma_{i3} \) respectively except the elements of \( \Sigma'' \) have been replaced by the corresponding elements of \( \Sigma \). But this implies that
\[ \gamma_i'' \Rightarrow \gamma_{i+1}'' \]
But since \( S \Rightarrow \gamma_i'' \) it must be true that
\[ S \Rightarrow \gamma_{i+1}'' \]
and by induction the statement that
\[ S \Rightarrow \gamma_i'' \]
is established. Notice that \( \gamma_n' \) does not contain any elements of \( \Sigma'' \) since it consists only of terminal symbols. Therefore, \( \gamma_n' = \gamma_n'' \) and since \( \gamma_n' = \gamma \),

\[ \gamma \in \text{language}(G) \]

Therefore,

\[ \text{language}(G') \subset \text{language}(G) \]

If \( \gamma \in \text{language}(G) \) and \( \gamma_0, \gamma_1, \ldots, \gamma_n \) is a derivation for \( \gamma \), then construct a new sequence

\[ \gamma_0', \gamma_1', \ldots, \gamma_n', \gamma_{n+1}', \ldots, \gamma_n'' \]

such that if \( n \geq i \geq 0 \) then \( \gamma_i' \) is the same as \( \gamma_i \) except that each element of \( \Sigma \) in \( \gamma_i \) is replaced by the corresponding element of \( \Sigma'' \). If \( n' \geq i \geq n \), then \( \gamma_{i+1}' \) is the same as \( \gamma_i \) except one element of \( \Sigma'' \) is replaced by the corresponding element of \( \Sigma \). Also,

\[ \gamma_n', \gamma_i' \in \Sigma^* \]

Notice that the construction of \( \gamma_{n+1}' \), to \( \gamma_n' \) is not unique. This new sequence is a derivation over \( G' \) for \( \gamma \) and, further, the width of this derivation is the same as the original derivation, since by induction

Basis: if \( i = 0 \) then \( \gamma_0 = S \) and since \( S \in V \), \( \gamma_0' = S \).

Therefore,

\[ S \implies \gamma_0' \]
Induction step: Assume that if i' ≤ i then

\[ S \Rightarrow \gamma_i' \]

By definition,

\[ \gamma_i = \gamma_{i1} \alpha \gamma_{i2} \]

and

\[ \gamma_{i+1} = \gamma_{i1} \beta \gamma_{i2} \]

where \( \alpha \rightarrow \beta \in P \). Therefore,

\[ \gamma_i' = \gamma_{i1}' \alpha' \gamma_{i3}' \]

and

\[ \gamma_{i+1}' = \gamma_{i1}' \beta' \gamma_{i3}' \]

where \( \gamma_{i1}' = \gamma_{i1} \) with all the elements of \( \Sigma \) replaced by the corresponding element of \( \Sigma' \) etc. But if that is the case, then \( \alpha' \rightarrow \beta' \in P' \) by construction, so that

\[ \gamma_i' \Rightarrow \gamma_{i+1}' \]

Since \( S \Rightarrow \gamma_i' \) by induction hypothesis,

\[ S \Rightarrow \gamma_{i+1}' \]

If \( n' > i \geq n \), then

\[ \gamma_i' \Rightarrow \gamma_{i+1}' \]

since the production \( \alpha \rightarrow a \in P' \) by construction. Therefore,

\[ S \Rightarrow \gamma_i' \]
in this case also. Therefore, by induction, \( S \Rightarrow \gamma'_i \) for all \( n' \geq i \geq 0 \). Since \( \gamma'_n \in \Sigma^* \), this new sequence is a derivation. Notice that \( \gamma'_n \) is the same as \( \gamma_n \) except that all elements of \( \Sigma \) are replaced by elements of \( \Sigma'' \). Also, due to construction, \( \gamma'_n \) is the same as \( \gamma'_n \) except all the elements of \( \Sigma'' \) are replaced by the elements of \( \Sigma \).

Therefore, \( \gamma_n = \gamma'_n \) and

\[
\gamma_n' = \gamma
\]

so that

\[ \gamma \in \text{language}(G') \]

and

\[ \text{language}(G) \subset \text{language}(G') \]

Notice that in the above derivation,

\[ |\gamma_i| = |\gamma_i'| \]

for all \( n \geq i \geq 0 \) so that

\[ \max_{n \geq i \geq 0}(|\gamma_i|) = \max_{n \geq i \geq 0}(|\gamma_i'|) \]

also when \( n' \geq i \geq n \)

\[ |\gamma_i| = |\gamma_{i+1}| \]

it must be true that

\[ \max_{n \geq i \geq 0}(|\gamma_i|) = \max_{n' \geq i \geq 0}(|\gamma_i'|) \]
Therefore, the width of the two derivations is the same, and
the final statement in the lemma is proved.

---

Lemma 97:

Given a grammar $G = (I, V, P, S)$ where

$$P \subseteq (U_{n \geq i > 0} V^i \rightarrow U_{n \geq i \geq 0} V^i) \cup (V \rightarrow I)$$

with $n > 2$ then there is a grammar $G' = (I', V', P', S')$ such that

$$P' \subseteq (U_{n \geq i > 0} V'^i \rightarrow U_{n \geq i \geq 0} V'^i) \cup (V' \rightarrow I)$$

such that

$$language(G) = language(G')$$

and if $\gamma \in language(G)$ has a derivation of width $w$
then there is a derivation for $\gamma$ in $G'$ with a width of $w$.

Proof:

Construct $G'$ as follows:

1. $I' = I$.
2. $V' = V \cup V \times V$.
3. The construction for $P'$ is given below.
4. $S' = S$.

$P'$ is constructed as follows:

1. If $i < n$, $j < n$ and

$$\alpha_1 \alpha_2 \ldots \alpha_i \beta_1 \beta_2 \ldots \beta_j \in P$$

then this production is also in $P'$. 
2. If \( i = n, j < n \) and

\[
a_1a_2 \ldots a_i \rightarrow^* \beta_1\beta_2 \ldots \beta_j \in P
\]

then the productions \( a_1a_2 \rightarrow (a_1a_2) \) and

\[
(a_1a_2)a_3 \ldots a_i \rightarrow \beta_1\beta_2 \ldots \beta_j
\]

are in \( P' \).

3. If \( i < n, j = n \) and the production given above is in \( P \), then the productions \( (\beta_1\beta_2) \rightarrow \beta_1\beta_2 \) and

\[
a_1a_2 \ldots a_i \rightarrow (\beta_1\beta_2)\beta_3 \ldots \beta_j
\]

are in \( P' \).

4. If both \( i = n, j = n \) and the given production is in \( P \), then the productions \( a_1a_2 \rightarrow (a_1a_2) \),

\( (\beta_1\beta_2) \rightarrow \beta_1\beta_2 \) and

\[
(a_1a_2)a_3 \ldots a_i \rightarrow (\beta_1\beta_2)\beta_3 \ldots \beta_j
\]

are in \( P' \).

5. These are all the elements of \( P' \).

The proof that this is the required grammar follows the proof of lemma 96 quite closely so only an outline will be given here. If \( \gamma \in \text{language}(G') \) and

\[
\gamma_0, \gamma_1 \ldots \gamma_j
\]

is a derivation for \( \gamma \), then construct a new sequence
such that each nonterminal of the form \((\alpha_1\alpha_2)\) is replaced by the corresponding \(\alpha_1\alpha_2\). A simple inductive proof will then establish that

\[ S \Rightarrow^* \gamma_i' \]

using grammar \(G\) for all \(i\). Since \(\gamma = \gamma_j = \gamma_j'\),

\[ S \Rightarrow^* \gamma \]

and

\[
\text{language}(G') \subseteq \text{language}(G)
\]

If \(\gamma \in \text{language}(G)\) and

\[ \gamma_0', \gamma_1', \ldots, \gamma_j' \]

is a derivation for \(\gamma\), then construct a derivation for \(\gamma\) over \(G'\) by inserting the following sentential forms into the sequence:

1. If \(\gamma_i \Rightarrow \gamma_{i+1}\) and

\[
\begin{align*}
\gamma_i &= \gamma_{i1} \alpha_1\alpha_2 \ldots \alpha_k \gamma_{i3} \\
\gamma_{i+1} &= \gamma_{i1} \beta_1\beta_2 \ldots \beta_j \gamma_{i3}
\end{align*}
\]

where

\[
\alpha_1\alpha_2 \ldots \alpha_k \Rightarrow \beta_1\beta_2 \ldots \beta_j \in \mathcal{P}
\]
k < n and j < n then do not insert anything between $a_i$ and $a_{i+1}$.

2. If $k = n$ and $j < n$ in the above equations, then insert

$$\gamma_i \, (\alpha_1 \alpha_2) \alpha_3 \ldots \alpha_k \, \gamma_{i+1}$$

between $\gamma_i$ and $\gamma_{i+1}$.

3. If $k < n$ and $j = n$ in the equations given above, then insert

$$\gamma_i \, (\beta_1 \beta_2) \beta_3 \ldots \beta_j \, \gamma_{i+1}$$

between $\gamma_i$ and $\gamma_{i+1}$.

4. If $k = n$ and $j = n$ in the above equations, then insert

$$\gamma_i \, (\alpha_1 \alpha_2) \alpha_3 \ldots \alpha_k \, \gamma_{i+1}$$

and

$$\gamma_i \, (\beta_1 \beta_2) \beta_3 \ldots \beta_j \, \gamma_{i+1}$$

between $\gamma_i$ and $\gamma_{i+1}$ in order.

If this new sequence is relabeled as

$$\gamma_0', \gamma_1' \ldots \gamma_j'$$
then a simple induction would establish that $\gamma_i' \Rightarrow^* \gamma_{i+1}'$
so that this is a derivation. Since all of the newly added
sentential forms are shorter than one or the other of the
adjacent sentential forms, the width of this production must
be the same as the width of the original production.

---

Lemma 98:

Given a grammar $G = (\Sigma, V, P, S)$ such that

$$P \subseteq (V^+ \rightarrow V^*) \cup (V \rightarrow \Sigma)$$

then there is a grammar $G' = (\Sigma', V', P', S')$ such

that

$$P' \subseteq \left( \bigcup_{i \geq 0} V_i^+ \rightarrow V_i' \right) \cup (V' \rightarrow \Sigma)$$

such that

$$\text{language}(G) = \text{language}(G')$$

and if $\gamma \in \text{language}(G)$ has a derivation $D$ of width $w$, then there is a derivation for $\gamma$ in $G'$ such that
the second derivation has a width of $w$ also.

Proof:

This is easily established by a simple induction using lemma 97.

---
Lemma 99:

Given a grammar \( G = (\Sigma, V, P, S) \) such that

\[
P \subseteq (\bigcup_{i \geq 0} V^i) \cup (V \rightarrow \Sigma)
\]

then there is a grammar \( G' = (\Sigma', V', P', S') \) where

\[
P' \subseteq (VV \rightarrow V) \cup (V \rightarrow VV) \cup (V \rightarrow \lambda) \cup (VV \rightarrow VV)
\]

such that

\[
\text{language}(G) = \text{language}(G')
\]

and width is preserved.

Proof:

Note that

\[
P \subseteq (V \rightarrow VV) \cup (V \rightarrow V) \cup (V \rightarrow \lambda) \cup (VV \rightarrow VV)
\]

\[
\cup (VV \rightarrow V) \cup (VV \rightarrow \lambda) \cup (V \rightarrow \Sigma)
\]

therefore, only productions of the form \( AB \rightarrow CD \) and \( AB \rightarrow \lambda \) need to be eliminated. Construct \( G' \) as follows:

1. \( \Sigma' = \Sigma \).

2. \( V' = V \cup (VV \rightarrow VV) \cup (VV \rightarrow \lambda) \).

3. The construction of \( P' \) is given below.

4. \( S' = S \).

\( P' \) is constructed as follows:
1. If $a \rightarrow \beta \in P$ and $|a| = 1$ or $|\beta| = 1$, then $a \rightarrow \beta$ is an element of $P'$.

2. If $a \rightarrow \beta \in P$, $|a| \neq 1$ and $|\beta| \neq 1$, then $a \rightarrow (a \rightarrow \beta)$ and $(a \rightarrow \beta) \rightarrow \beta$ are elements of $P'$.

Note that the offending productions are eliminated since these are the only productions where $|a| \neq 1$ and $|\beta| \neq 1$.

To prove that this is the required grammar, use the same approach as was used in lemma 97. If $\gamma \in \text{language}(G')$ with derivation

$$\gamma_0, \gamma_1, \ldots, \gamma_n$$

then replace any nonterminal $(a \rightarrow \beta)$ added in step two of the construction with $a$ (note $\beta$ could have been used) to form a new sequence

$$\gamma'_0, \gamma'_1, \ldots, \gamma'_n$$

An inductive proof establishes that $S \Rightarrow \gamma_i'$ using grammar $G$ so

$$\text{language}(G') \subset \text{language}(G)$$

If $\gamma \in \text{language}(G)$ and

$$\gamma_0, \gamma_1, \ldots, \gamma_n$$

is a derivation for $\gamma$, then construct a new sequence
by inserting a new sentential form between any two
sentential forms which use one of the offending production.
The new sentential form is

$$\gamma_{i1} (\alpha \rightarrow \beta) \gamma_{i3}$$

where $\alpha \rightarrow \beta$ is the offending production, $\gamma_i = \gamma_{i1} \alpha \gamma_{i3}$ and $\gamma_{i+1} = \gamma_{i1} \beta \gamma_{i3}$ are the sentential forms. A simple proof would show that this is the required derivation. Proving the lemma.

---

**Lemma 100:**

Given a grammar $G = (\Sigma, V, P, S)$ such that

$$P \subseteq (VV \rightarrow V) \cup (V \rightarrow VV) \cup (V' \rightarrow \Sigma) \cup (V \rightarrow V) \cup (V \rightarrow \lambda)$$

there is a grammar $G' = (\Sigma', V', P', S')$ where

$$P' \subseteq (V'V' \rightarrow V') \cup (V' \rightarrow V'V') \cup (V' \rightarrow \Sigma) \cup (V \rightarrow \lambda)$$

such that

$$\text{language}(G) = \text{language}(G')$$

and width is preserved.

**Proof:**

Let $\text{Right}(G) = \{X \mid X \in V \text{ and } X \text{ is on the right hand side of a production of the form } A \rightarrow B \text{ in } G\}$. Proof by induction on the number of elements in $\text{Right}(G)$.
Basis: If right(G) is null, then G has no production of the form A → B where A and B are nonterminals, so G itself is the required grammar.

Induction step: Assume that if Right(G) has fewer than n elements there is a grammar with the desired properties. If Right(G) has n elements, then choose one element from Right(G), call it X. Form the set

\[ L_X = \{ a | a \in V \text{ and } a \rightarrow X \in P \} - \{ X \} \]

Construct a new grammar G' = (\Sigma', V', P', S') as follows:

1. \( \Sigma' = \Sigma \).
2. \( V' = V \).
3. The construction of P' is given below.
4. S' = S.

P' is constructed as follows:

1. If \( \alpha \rightarrow \beta \in P \text{ and } \alpha \rightarrow \beta \text{ is not of the form } \gamma \rightarrow X \) where \( \gamma \) is an element of V, then \( \alpha \rightarrow \beta \) is in \( P' \).
2. If \( X \rightarrow \beta \in P \text{ and } \alpha \in L_X \), then the production \( \alpha \rightarrow \beta \) is in \( P' \).
3. If \( X \beta \rightarrow \gamma \in P \text{ and } \alpha \epsilon L_X \), then the production \( \alpha \beta \rightarrow \gamma \) is in \( P' \).
4. If \( \beta X \rightarrow \gamma \in P \text{ and } \alpha \epsilon L_X \), then the production \( \beta \alpha \rightarrow \gamma \) is in \( P' \).
5. If $XX \rightarrow \gamma \in P$ and also $\alpha$ and $\beta \in L_X$, then the production $\alpha \beta \rightarrow \gamma$ is in $P'$.

Observe that there are no production of the form $\alpha \rightarrow X$ where $\alpha \in V$ in $G'$, and, therefore, $\text{Right}(G') = \text{Right}(G') - \{X\}$.

If $\gamma \in \text{language}(G)$ with derivation

$$\gamma_0, \gamma_1, \ldots, \gamma_n$$

then form a new sequence

$$\gamma_0', \gamma_1', \ldots, \gamma_n'$$

by deleting all members of the sequence $\gamma_{i+1}$ where

$$\gamma_i = \gamma_{i1} \alpha \gamma_{i3}$$

$$\gamma_{i+1} = \gamma_{i1} \gamma_{i3}$$

$\alpha \in V$ and $\alpha \rightarrow X \in P$. It can be shown through a simple inductive proof that the second sequence is a derivation over $G'$ for $\gamma$. Therefore,

$$\text{language}(G) \subseteq \text{language}(G')$$

Also, notice that each of the deleted sentential forms is the same length as the preceding sentential form in the first sequence. The result is that the width of the new derivation is the same as the width of the original derivation.

If $\gamma \in \text{language}(G')$ with derivation
construct a derivation over $G$ for $\gamma$ by inserting

$$\gamma_{i1} X \gamma_{i3}$$

into the sequence between $\gamma_i$ and $\gamma_{i+1}$ where

$$\gamma_i = \gamma_{i1} \alpha \gamma_{i3}$$
$$\gamma_{i+1} = \gamma_{i1} \beta \gamma_{i3}$$

and $\alpha \to \beta$ is one of the production in $P'$ that is not in $P$.

A simple inductive proof would then establish that this is a legal derivation for $\gamma$ over $G$ so that

$$\text{language}(G') \subseteq \text{language}(G)$$

and the lemma is proved by induction.

---

**Theorem 101:**

Given a grammar $G = (\Sigma, \Gamma, P, S)$ then there is a grammar $G' = (\Sigma', \Gamma', P', S')$ such that

$$P' \subseteq (\Gamma' \Gamma' \rightarrow \Gamma') \cup (\Gamma' \rightarrow \Gamma' \Gamma') \cup (\Gamma' \rightarrow \Sigma') \cup (\Gamma' \rightarrow \lambda)$$

where

$$\text{language}(G) = \text{language}(G')$$

and if $\gamma \in \text{language}(G)$ has a derivation of width $w$ then there is a derivation over $G'$ for $\gamma$ of width $w$.

**Proof:**
By lemma 96 there is a grammar $G_1 = (\Sigma_1, V_1, P_1, S_1)$ such that

$$P_1 \subset (V_1^+ \to V_1^*) \cup (V_1' \to \Sigma_1')$$

width is preserved and $\text{language}(G) = \text{language}(G)$. By lemma 98 there is a grammar $G_2 = (\Sigma_2, V_2, P_2, S_2)$ such that

$$P_2 \subset (U_{2i \geq 0}V_2 \to U_{2i \geq 0}V_2^i) \cup (V_2 \to \Sigma_2)$$

width is preserved and $\text{language}(G_2) = \text{language}(G)$. By lemma 99 there is a grammar $G_3 = (\Sigma_3, V_3, P_3, S_3)$ such that

$$P_3 \subset (V_3V_3 \to V_3) \cup (V_3 \to V_3V_3) \cup (V_3 \to \Sigma_3) \cup (V_3 \to V_3)$$

width is preserved and $\text{language}(G_3) = \text{language}(G)$. Finally, by lemma 100 there is a grammar $G_4 = (\Sigma_4, V_4, P_4, S_4)$ such that

$$P_4 \subset (V_4V_4 \to V_4) \cup (V_4 \to V_4V_4) \cup (V_4 \to \Sigma_4) \cup (V_4 \to \lambda)$$

width is preserved and $\text{language}(G_4) = \text{language}(G_3)$. This last grammar, $G_4$, is the desired grammar.

- - -

Note that this last theorem has the required normal form.

As an example of the normal form, consider the grammar

$$S \to aTbc$$

$$T \to aTbU$$

$$T \to \lambda$$

$$Ub \to bU$$

$$Uc \to cc$$
This grammar defines the language \( \{a^n b^n c^n \}_{n>0} \). If the first transformation given in lemma 96 is applied to this grammar, then the grammar becomes

\[
\begin{align*}
S & \rightarrow ATBC \\
T & \rightarrow ATBU \\
T & \rightarrow \lambda \\
UB & \rightarrow BU \\
UC & \rightarrow CC \\
A & \rightarrow a \\
B & \rightarrow b \\
C & \rightarrow c
\end{align*}
\]

If \( n \) is set to 4 and the transformation given in lemma 97 is applied, then the grammar becomes

\[
\begin{align*}
S & \rightarrow VBC \\
V & \rightarrow AT \\
T & \rightarrow WBU \\
W & \rightarrow AT \\
T & \rightarrow \lambda \\
UB & \rightarrow BU \\
UC & \rightarrow CC \\
A & \rightarrow a \\
B & \rightarrow b \\
C & \rightarrow c
\end{align*}
\]
If \( n \) is set to 3 and the transformation is applied again, the grammar becomes

\[
\begin{align*}
S & \rightarrow XC \\
X & \rightarrow VB \\
V & \rightarrow AT \\
T & \rightarrow YU \\
Y & \rightarrow WB \\
W & \rightarrow AT \\
T & \rightarrow \lambda \\
UB & \rightarrow BU \\
UC & \rightarrow CC \\
A & \rightarrow a \\
B & \rightarrow b \\
C & \rightarrow c
\end{align*}
\]

Again, the transformation given in lemma 99 produces the grammar

\[
\begin{align*}
S & \rightarrow XC \\
X & \rightarrow VB \\
V & \rightarrow AT \\
T & \rightarrow YU \\
Y & \rightarrow WB \\
W & \rightarrow AT \\
T & \rightarrow \lambda
\end{align*}
\]
Since there are no productions of the form \( A \rightarrow B \) where \( A \) and \( B \) are in \( V \), this is the required grammar. If the element \( aabbcc \) of the language is selected and the derivation

\[
S, aTbc, aaTbUbc, aabUbC, aabbUc, aabbcc
\]

is selected, note that width of the derivation is 7. A derivation in the transformed grammar for the same element of width 7 is

\[
S, XC, VBC, ATBC, AYUBC, AWBUBC, AATBUBC, AABUBC, AABZC, AABBUC, AABBD, AABBCC, aABBCC, aaBBCC, aabBCc, aabbcC, aabbcc
\]

Notice that the first, fourth, seventh, eighth, tenth and twelfth sentential form in the sequence are the same as in the original sequence except all of the terminals are replaced by the respective nonterminals.
Now that a normal form for a grammar has been developed, this form will be used to simplify the proof of the theorem that this appendix is written to provide. This theorem states that if the elements of a language all have a linearly bounded derivation, then the language is context sensitive. It should be noted that only one derivation for each element need be linearly bounded. Other derivations may exist which may or may not be linearly bounded, but they are irrelevant. This fact is used extensively in the containment proof given in this paper.

Lemma 102:

Given a grammar $G = (\Sigma, V, P, S)$ which is in normal form such that if $\gamma \in \text{language}(G)$ there exists at least one derivation for $\gamma$ of width $|\gamma|$ then there is a grammar $G' = (\Sigma', V', P', S')$ such that

$$\text{language}(G) = \text{language}(G')$$

and $G'$ is context sensitive.

Proof:

Construct grammar $G' = (\Sigma', V', P', S')$ as follows:

1. $\Sigma' = \Sigma$.
2. $V' = V \cup \{S', Z\}$.
3. The construction for $P'$ follows.
4. $S'$ does not appear elsewhere.

$P'$ is constructed such that

1. $S' \rightarrow S'Z$ and $S' \rightarrow S$ are in $P'$.
2. If $AB \rightarrow C \in P$, then $AB \rightarrow CZ$ is in $P'$.

3. If $A \rightarrow BC \in P$, then $AZ \rightarrow BC$ is in $P'$.

4. If $A \rightarrow \alpha \in P$, then $A \rightarrow \alpha$ is in $P'$.

5. If $A \rightarrow \lambda \in P$, then $A \rightarrow Z$ is in $P'$.

6. If $A \in V' \cup \Sigma'$, then $AZ \rightarrow ZA$ and $ZA \rightarrow AZ$ are in $P'$.

Claim: $G'$ is the required context sensitive language. Note that $G'$ is context sensitive since each production added to $P'$ has the same number of tokens on each side of the arrow with the exception of the very first production.

If $\gamma \in \text{language}(G')$ with derivation

$$\gamma_0, \gamma_1, \ldots, \gamma_n$$

then construct a new sequence

$$\gamma_0', \gamma_1', \ldots, \gamma_n'$$

by deleting all occurrences of the nonterminal $Z$ and replacing the nonterminal $S'$ by $S$. A simple induction proof will show that $S \Rightarrow \gamma_i'$ over $G$ for all $n \geq i \geq 0$.

Therefore, in particular,

$$S \Rightarrow \gamma_n'$$

and

$$\text{language}(G') \subseteq \text{language}(G)$$
Note that the fact that \( G \) is linearly bounded is not used in this part of the proof, so that if this construction is performed on a grammar which is not bounded by \( |y| \) then the grammar describes a context sensitive subset of the original grammar \( G \).

If \( y \in \text{language}(G) \) and

\[
\gamma_0, \gamma_1, \ldots, \gamma_n
\]

is a derivation for \( y \) of width \( |y| \), then construct a new sequence as follows:

1. Before \( \gamma_0 \) insert the sequence

\[
S', S'Z, S'ZZ, \ldots, S'|y| - 1
\]

2. Expand each sentential form in the original sequence to length \(|\gamma|\) by putting as many occurrences of \( Z \) on the end of the sentential form as are needed. Note that this step is not possible if the original derivation has a width of more than \(|\gamma|\).

3. If \( \gamma_i = \gamma_{i1} \alpha \gamma_{i3} \) and \( \gamma_{i+1} = \gamma_{i1} \beta \gamma_{i3} \) where \( \alpha \rightarrow \beta \in P \) and is of the form \( AB \rightarrow C \), then insert the sequence of sentential forms

\[
\gamma_{i1} \beta Z \gamma_{i3} Z^w
\]
320

\[ \gamma_{i1} \beta \gamma_{i3}^1 Z \gamma_{i3}^2 \cdots \gamma_{i3}^{n'w} \]

\[ \gamma_{i1} \beta \gamma_{i3}^1 Z \gamma_{i3}^2 \cdots \gamma_{i3}^{n'-1} Z \gamma_{i3}^{n'w} \]

each of length \(|\gamma|\) between \(\gamma_i\)' and \(\gamma_{i+1}'\), where

\[ w = |\gamma| - |\gamma_i| \quad \text{and} \quad \gamma_{i3} = \gamma_{i3}^1 \gamma_{i3}^2 \cdots \gamma_{i3}^{n'} \]

4. If the production is of the form \(A \rightarrow BC\), then insert the sequence of sentential forms

\[ \gamma_{i1} A \gamma_{i3}^1 \gamma_{i3}^2 \cdots \gamma_{i3}^{n'-1} Z \gamma_{i3}^{n'w} \]

\[ \gamma_{i1} A \gamma_{i3}^1 \gamma_{i3}^2 \cdots \gamma_{i3}^{n'-1} Z \gamma_{i3}^{n'w} \]

\[ \gamma_{i1} A \gamma_{i3}^1 \gamma_{i3}^2 \cdots \gamma_{i3}^{n'-1} Z \gamma_{i3}^{n'w} \]

\[ \gamma_{i1} A \gamma_{i3}^1 \gamma_{i3}^2 \cdots \gamma_{i3}^{n'-1} Z \gamma_{i3}^{n'w} \]

each of length \(|\gamma|\) between \(\gamma_i\)' and \(\gamma_{i+1}'\), where

\[ w = |\gamma| - |\gamma_{i+1}| \quad \text{and} \quad \gamma_{i3} = \gamma_{i3}^1 \gamma_{i3}^2 \cdots \gamma_{i3}^{n'} \]

5. If the production is of the form \(A \rightarrow a\), do not insert anything between \(\gamma_i\) and \(\gamma_{i+1}'\).
6. If the production is of the form $A \rightarrow \lambda$, then insert the sequence of sentential forms

$$
\gamma_i^1 \gamma_i^2 \cdots \gamma_i^n \gamma_i^w \\
\gamma_i^1 \gamma_i^2 \cdots \gamma_i^n \gamma_i^w \\
\gamma_i^1 \gamma_i^2 \cdots \gamma_i^n \gamma_i^w \\
\vdots \\
\gamma_i^1 \gamma_i^2 \cdots \gamma_i^{n-1} \gamma_i^n \gamma_i^w
$$

between $\gamma_i$ and $\gamma_{i+1}'$. Label this new sequence as

$$\gamma_0', \gamma_1', \ldots, \gamma_n'$$

This new sequence is a derivation for $\gamma$ over the grammar $G'$, since

1. The first symbol is $S'$, the start symbol for $G'$.
2. If $\gamma_i'$ was added in step one, then the first production in the construction of $P'$ is used to generate $\gamma_i'$.
3. If $\gamma_i'$ is the last sentential form in this set, then the second production is used to generate $\gamma_{i+1}'$. 
4. The sentential forms added in steps 3 to 6 above are applications of the productions added in step 6 of the construction. In general, they move the Zs to where they are needed in order for the productions added in steps 2 through 5 of the construction to be applied.

Since \( E' = \Sigma \) and the last item in the second sequence is equal to \( \lambda_n = \lambda \), an element of \( \Sigma^* \), \( \lambda \) must be an element of \( \text{language}(G) \) and

\[
\text{language}(G) = \text{language}(G')
\]

so the lemma is proved.

---

Note that \( \lambda \) cannot be an element of these languages, unless an ad hoc statement is made that the production

\[
S \rightarrow \lambda
\]

is allowed, and is simply carried along as an exception in each case. This is also true of context sensitive (and even context free) languages since the general characterization (production of the form \( a \rightarrow \beta \) where \(|a| \leq |eta|\)) does not allow the null string.

An example of a grammar which is not context sensitive but is bounded as defined in lemma 102 is as follows:
Note that this grammar defined the language 

\[ \{ a^n b^n c^n \}_{n>0} \]

Since \( Ub \rightarrow X \) is a production, this is not a context sensitive grammar. However, the \( X \) must eventually be expanded by \( X \rightarrow bU \) so that the intermediate sentential forms must all be shorter than the final generated string.

Lemma 103:

Given a Grammar \( G = (\Sigma, \Gamma, P, S) \) in normal form then there is a grammar \( G' = (\Sigma', \Gamma', P', S') \) such that if \( \gamma \in \text{language}(G) \) with a derivation of width \( w \) then there is a derivation over \( G' \) for \( \gamma \) of width \( w' \) such that

\[ w' \leq \max(\frac{(w+1)}{2}, |\gamma|, |\gamma|) \]

and \( \text{language}(G) = \text{language}(G') \).

Proof:
Construct $G' = (\Sigma', V', P', S')$ as follows:

1. $\Sigma' = \Sigma$.
2. $V' = V \cup (V \times V) \cup (V \times V) \times (V \times V)$.
3. The construction of $P'$ is given below.
4. $S' = S$.

$P'$ is constructed as follows:

1. If $AB \rightarrow C \in P$ and $D$ and $E$ are elements of $V \cup \Sigma$, then $(AB) \rightarrow C$, $(DA)(BE) \rightarrow D(CE)$ and $A(BE) \rightarrow (CE)$ are in $P'$.
2. If $A \rightarrow BC \in P$ and $D \in V \cup \Sigma$, then $(AD) \rightarrow (BCD)$, $(DA) \rightarrow (DCB)$ and $A \rightarrow (BC)$ are in $P'$.
3. If $A \rightarrow \lambda \in P$ and $D \in V \cup \Sigma$, then $(AD) \rightarrow D$, $(DA) \rightarrow D$ and $A \rightarrow \lambda$ are in $P'$.
4. If $A \rightarrow a \in P$ and $D \in V \cup \Sigma$, then $(AD) \rightarrow (aD)$, $(DA) \rightarrow (Da)$ and $A \rightarrow a$ are in $P'$.
5. If $A$, $B$, $C$, $D$ and $E$ are elements of $V \cup \Sigma$, then

\[
(AB)C \rightarrow A(BC) \\
(AB)(CDE) \rightarrow (ABC)(DE) \\
A(BCD) \rightarrow (AB)(CD) \\
(ABC) \rightarrow A(BC) \\
A(BC)D \rightarrow (AB)(CD)
\]

are in $P'$.

6. If $A$ and $B \in \Sigma$, then $(AB) \rightarrow AB$ are in $P'$.

These are arranged in table 42 for easy understanding.
TABLE 42. Construction for Lemma 103

Start of Production

(., )

\begin{align*}
AB \rightarrow C & \quad (AB) \rightarrow C \\
A \rightarrow BC & \quad (AD) \rightarrow (BCD) \\
A \rightarrow \lambda & \quad (AD) \rightarrow D \\
A \rightarrow a & \quad (AD) \rightarrow (aD)
\end{align*}

(., )

\begin{align*}
(AB)^C & \quad (DA)(BE) \rightarrow D(CE) \\
(AB)^C & \quad A(BD) \rightarrow (CE) \\
(AD)^D & \quad (DA) \rightarrow D \\
A \rightarrow X & \quad (AD) \rightarrow (Da) \\
A \rightarrow a & \quad (AD) \rightarrow (aD)
\end{align*}

(., )

\begin{align*}
A \rightarrow X & \quad (AD) \rightarrow D \\
A \rightarrow \lambda & \quad (DA) \rightarrow D \\
A \rightarrow a & \quad (AD) \rightarrow (Da)
\end{align*}

\begin{align*}
A \rightarrow \lambda & \quad (DA) \rightarrow D \\
A \rightarrow a & \quad (AD) \rightarrow (Da)
\end{align*}

\begin{align*}
A \rightarrow \lambda & \quad (DA) \rightarrow D \\
A \rightarrow a & \quad (AD) \rightarrow (Da)
\end{align*}

If \( \gamma \in \text{language}(G') \) with derivation

\[
\gamma_0, \gamma_1, \ldots, \gamma_n
\]

then construct the sequence

\[
\gamma_0', \gamma_1', \ldots, \gamma_n'
\]

where \( \gamma_i' \) is the same as \( \gamma_i \) except every occurrence of \((AB)\) has been replaced by \(A B\) and every occurrence of \((ABC)\) has been replaced by \(A B C\). A simple inductive proof shows that \( S \Rightarrow \gamma_i \) so

\[
\text{language}(G') \subset \text{language}(G)
\]

If \( \gamma \in \text{language}(G) \) with derivation

\[
\gamma_0, \gamma_1, \ldots, \gamma_n
\]
then construct a new derivation by

1. Grouping the characters from each element \( \gamma_i \) into groups of two and putting parentheses around them. If there are an odd number of characters, leave the first by itself.

2. If \( \gamma_i = \gamma_{i1} \alpha \gamma_{i3} \) and \( \gamma_{i+1} = \gamma_{i1} \beta \gamma_{i3} \) where \( \gamma \rightarrow \beta \in P \), construct a new element \( \gamma_{i\alpha} \) which results from using the appropriate production from table 42 depending on the form of \( \alpha \rightarrow \beta \) and the starting position of the production in \( \gamma_i \). If \( \gamma_{i\alpha} \) is the same as the element corresponding to \( \gamma_{i+1} \), then do not insert anything into the new derivation.

3. If \( \gamma_{i\alpha} \) is not the same as the element corresponding to \( \gamma_{i+1} \), then insert \( \gamma_{i\alpha} \) and as many other sentential forms as are needed to bring the extra nonterminal or ordered triple to the front of the production by using the productions of the form \((AB)C \rightarrow A(BC)\) and \((AB)(CDE) \rightarrow (ABC)(DE)\). Then use productions of the form

\[
(ABC) \rightarrow A(BC)
\]
\[
A(BC)D \rightarrow (AB)(CD)
\]

or

\[
A(BCD) \rightarrow (AB)(CD)
\]
to combine the elements which are not ordered pairs at
the end. These will generate the element
corresponding to \( Y_{i+1} \).

4. Lastly, use productions of the form \((AB) \rightarrow AB\) to
expand the element corresponding to \( Y_n \) back to \( Y \).

This can be proved to be the required derivation by a simple
inductive proof. Note that the width of this derivation is
\((w+1)/2\) if the maximum width is reached at or before the
element corresponding to \( Y_n \), \(|Y|\) if the maximum is reached
at the end or 1 if \( w = 1 \). Therefore, the lemma is proved.

---

Lemma 104:

Let \( G = (E,V,P,S) \) be a grammar and \( f \) be a function
such that if \( Y \in \text{language}(G) \) there exists at least
one derivation for \( Y \) of width \( w \) such that
\[ w \leq f(Y) \]
then if \( \alpha \) is an arbitrary element of \( E^* \) it is
possible to determine whether or not
\[ \alpha \in \text{language}(G) \]

Proof:

Given a string \( \alpha \), construct a sequence as follows

1. \( Y_1 = \alpha \).

2. Given \( Y_i \) then the construction for \( Y_{i+1} \) is given
below.

\( Y_{i+1} \) is constructed as follows:
1. Identify all substrings of $\gamma_i$ which are the right hand sides of productions in $P$.
2. Try to replace the substrings one at a time with the corresponding left hand side of the production to generate $\gamma_{i+1}$ until either $\gamma_{i+1}$ does not fail or all the substrings have been tried.

$\gamma_{i+1}$ can fail for any of the following reasons
1. $|\gamma_{i+1}| > f(\alpha)$.
2. $\gamma_{i+1} = \gamma_k$ for any $1 \leq k \leq i$.
3. All of the possible constructions for $\gamma_{i+1}$ fail and $\gamma_{i+1} \neq S$.

Note that if $G$ is a context free grammar, the substrings are tried in left to right order and strings are allowed to fail for reason 3 above only then this algorithm is a bottom-up backtracking parsing algorithm. If $\gamma_1$ succeeds, then $\alpha$ is in language($G$). It is easy to prove in this case that

$$\gamma_n', \gamma_{n-1}', \ldots, \gamma_1$$

where $\gamma_n = S$ is a parse for $\alpha$. Note that this algorithm terminates since points 1 and 2 above limit the length of the derivation to $f(\alpha)!$. Therefore, it can be determined that an arbitrary element of $\Sigma^*$ is in the bounded grammar so the lemma is proved.

- - -
Corollary 105:

Given a grammar $G = (\Sigma, V, P, S)$ such that if $\gamma \in \text{language}(G)$ then there is a derivation for $\gamma$ of width $w$ such that

$$w \leq K_1|\gamma| + K_2$$

then it can be determined if $\lambda$ is an element of the language.

Proof:

This follows directly from lemma 104.

---

Lemma 106:

Given a grammar $G = (\Sigma, V, P, S)$ such that if $\gamma \in \text{language}(G)$ then there is a derivation for $\gamma$ of width $w$ such that

$$w \leq K|\gamma|$$

where $K \geq 1$ then there exists a context sensitive grammar $G' = (\Sigma', V', P', S')$ such that

$$\text{language}(G) = \text{language}(G')$$

Proof by induction on $K$:

Basis: If $K = 1$, then by Theorem 101 there is a normal form grammar that is equivalent to $G$ and by lemma 102 there is a corresponding context sensitive sensitive language.
Induction step: Assume that if $K' < K$ then the lemma is true for a grammar bounded by $K'$. If $G = (I, V, P, S)$ is a grammar that is bounded by $K$, then by Theorem 101 there is a normal form grammar for $\text{language}(G)$ with the same bound. By lemma 103 there is a corresponding grammar where the derivations for $y$ are bounded by

$$w = \max((K|y|+1)/2, |y|, 1).$$

Since $K \geq 2$, it must be true that

$$(K|y| + 1)/2 > |y|$$

$|y| \geq 1$ because $\lambda$ is not an element of $G$ (otherwise the width of a derivation for $\lambda$ must be zero, but that is not possible). Therefore,

$$w = (K|y|+1)/2$$

Since the bound must be an integer, $w = K|y|/2$ if $K|y|$ is even. In particular, if $K = 2$, then the new bound is $K|y|/2$. But $K > K/2$ so by induction hypothesis there is a corresponding context sensitive grammar. If $K > 2$, then if a production is bounded by $(K|y|+1)/2$, it is also bounded by $(K+1)|y|/2$. Since $K$ is an integer, if $(K+1)/2$ is not an integer then $(K+2)/2$ is an integer and $(K+2)|y|/2$ is a bound for production. Now

$$K = K/2 + K/2 > K/2 + 1$$
since \( K > 2 \) and

\[
K > \frac{K + 2}{2} > \frac{K + 1}{2}
\]

In any event, there is an integer bound for the grammar given by lemma 103 which is less than \( K \), so by induction hypothesis there is a context sensitive grammar for the original language and lemma is proved by induction.

\[
\ldots
\]

Theorem 107:

Given a grammar \( G = (\Sigma, V, P, S) \) such that if \( C \in \text{language}(G) \) then there is a derivation for \( \gamma \) of width \( w \) such that

\[
w \leq K_1 |\gamma| + K_2
\]

then there exists a context sensitive grammar \( G' = (\Sigma', V', P', S') \) such that

\[
\text{language}(G) = \text{language}(G')
\]

Proof:

\( K_1 \geq 1 \) since if that were not the case, when \( |\gamma| \) becomes greater than \( K_2 / (1 - K_1) \) then the length of a derivation must be less than \( |\gamma| \) which is impossible. If \( K_2 < 0 \), then \( K_1 |\gamma| \) is also a bound. If \( \lambda \) is not an element of \( G \) (this can be determined by corollary 105), then in all cases,

\[
K_1 |\gamma| + K_2 \leq (K_1 + K_2) |\gamma|
\]
and by lemma 106 the required context sensitive grammar exists. If \( \lambda \in G \), then the constructions used in the proof of lemma 106 will result in equivalent grammar describing the set language \( G - \{\lambda\} \). Therefore, add a new start symbol \( S' \) and two productions

\[
S' \rightarrow S \\
S' \rightarrow \lambda
\]

and this is the required grammar.

- - -

For examples of linearly bounded grammars, see chapter 8 of this paper.