Bibliography


# Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta)-subsumption</td>
<td>69</td>
</tr>
<tr>
<td>AcGM, 202</td>
<td></td>
</tr>
<tr>
<td>ADI-Mine, 198</td>
<td></td>
</tr>
<tr>
<td>AGM, 201</td>
<td></td>
</tr>
<tr>
<td>Aho, Hopcroft and Ullman’s tree isomorphism algorithm, 134</td>
<td></td>
</tr>
<tr>
<td>Alternating path, 143</td>
<td></td>
</tr>
<tr>
<td>Anti-monotonicity, 40</td>
<td></td>
</tr>
<tr>
<td>Apriori algorithm, 14</td>
<td></td>
</tr>
<tr>
<td>Apriori property, 14</td>
<td></td>
</tr>
<tr>
<td>AprioriSMP, 233</td>
<td></td>
</tr>
<tr>
<td>Association rules, 9</td>
<td></td>
</tr>
<tr>
<td>Augmenting path, 143</td>
<td></td>
</tr>
<tr>
<td>Automorphisms, 179</td>
<td></td>
</tr>
<tr>
<td>Backbone depth tuple order, 172</td>
<td></td>
</tr>
<tr>
<td>Backtracking sequence, 133</td>
<td></td>
</tr>
<tr>
<td>Bias, 43, 73</td>
<td></td>
</tr>
<tr>
<td>Bipartite involved matchings problem, 143</td>
<td></td>
</tr>
<tr>
<td>Bongard datasets, 92</td>
<td></td>
</tr>
<tr>
<td>C-ARMR, 98</td>
<td></td>
</tr>
<tr>
<td>CAP, 60</td>
<td></td>
</tr>
<tr>
<td>CBA, 237</td>
<td></td>
</tr>
<tr>
<td>CBAMs, 237</td>
<td></td>
</tr>
<tr>
<td>CHARM, 64</td>
<td></td>
</tr>
<tr>
<td>Chemistry, 4, 94, 165, 211</td>
<td></td>
</tr>
<tr>
<td>CHOPPER, 151</td>
<td></td>
</tr>
<tr>
<td>Chung’s subtree algorithm, 141, 201</td>
<td></td>
</tr>
<tr>
<td>Close, 63</td>
<td></td>
</tr>
<tr>
<td>Closed itemsets, 51</td>
<td></td>
</tr>
<tr>
<td>CLOSEGRAPH, 162, 197</td>
<td></td>
</tr>
<tr>
<td>CMAR, 237</td>
<td></td>
</tr>
<tr>
<td>CMTREEMINER, 152</td>
<td></td>
</tr>
<tr>
<td>CN2-SD, 238</td>
<td></td>
</tr>
<tr>
<td>Complexities</td>
<td></td>
</tr>
<tr>
<td>graph isomorphism algorithms, 163</td>
<td></td>
</tr>
<tr>
<td>subgraph isomorphism algorithms, 163</td>
<td></td>
</tr>
<tr>
<td>subtree isomorphism algorithms, 110</td>
<td></td>
</tr>
<tr>
<td>tree isomorphism algorithms, 110</td>
<td></td>
</tr>
<tr>
<td>Confidence, 10</td>
<td></td>
</tr>
<tr>
<td>Constraint</td>
<td></td>
</tr>
<tr>
<td>anti-monotonic, 41</td>
<td></td>
</tr>
<tr>
<td>convertible, 43</td>
<td></td>
</tr>
<tr>
<td>monotonic, 40</td>
<td></td>
</tr>
<tr>
<td>Succinct, 44</td>
<td></td>
</tr>
<tr>
<td>Contingency table, 229</td>
<td></td>
</tr>
<tr>
<td>CorCLASS, 235</td>
<td></td>
</tr>
<tr>
<td>Cover, 10</td>
<td></td>
</tr>
<tr>
<td>Covers, 36</td>
<td></td>
</tr>
<tr>
<td>Cycle, 105</td>
<td></td>
</tr>
<tr>
<td>D-LAB, 97</td>
<td></td>
</tr>
<tr>
<td>Data tree, 110</td>
<td></td>
</tr>
<tr>
<td>Depth sequence, 114</td>
<td></td>
</tr>
<tr>
<td>Depth tuple, 114</td>
<td></td>
</tr>
<tr>
<td>Diffsets, 19, 191</td>
<td></td>
</tr>
<tr>
<td>Dryade, 153</td>
<td></td>
</tr>
<tr>
<td>DualMiner, 60</td>
<td></td>
</tr>
<tr>
<td>ECLAT, 18</td>
<td></td>
</tr>
<tr>
<td>Edge sequence, 162</td>
<td></td>
</tr>
<tr>
<td>Enumeration, 39, 130</td>
<td></td>
</tr>
<tr>
<td>Equivalence classes, 35</td>
<td></td>
</tr>
<tr>
<td>ExAnte property, 61</td>
<td></td>
</tr>
<tr>
<td>FARMER, 83</td>
<td></td>
</tr>
<tr>
<td>FFSM, 198</td>
<td></td>
</tr>
<tr>
<td>Formal concept analysis (FCA), 247</td>
<td></td>
</tr>
<tr>
<td>Formal concepts analysis (FCA), 52</td>
<td></td>
</tr>
</tbody>
</table>
INDEX

FP-Bonsai, 61
FP-Growth, 23, 60
Free itemsets, 51
FreeTreeMiner, 200, 201
FREQ, 139
Frequent itemsets, 9
FSG, 203
FST-Forest, 151

GBI, 204
GraphML, 113, 166
Graphs, 105
Greatest lower bound, 33
GSP, 64
gSpan, 162, 196
GXL, 113, 166

HybridTreeMiner, 151, 199
Hypergraphs, 112, 165

Inductive Logic Programming (ILP), 68, 165
Itemset occurrences, 10

Kirchoff’s matrix-tree theorem, 167
k-Prefix, 13

Large itemsets, 11
Lattice, 35
Learning from entailment, 97
Learning from interpretations, 97
Least upper bound, 33
Leftmost path, 172
Lexicographical order, 13

Maximal frequent itemsets, 50
Merge operators, 53
Merging
...of cyclic graphs, 185
...of free trees, 176
...of ordered trees, 115
...of unordered trees, 124
...basic definitions, 54
...downward, 55
Modes, 79
MoFA, 198
MolFea, 61

Monotonicity, 40
Multi-relational data mining, 4, 113, 165
Multicast dataset, 112

Nauty, 204
Next prefix node, 120
Non-derivable itemsets, 51

Object exchange model, 151
Object Identity, 71
Occurrence sequence, 136, 189
Occurrence tree, 145
Orders, 12

Path
...rooted, 106
...simple, 105
PathJoin, 151
Pattern tree, 110
PolyFarm, 99
Prefix trie, 15
PREFIXSPAN, 64
Primary key, 74
Projected database, 20
Proteins, 165

Query packs, 99

RADAR, 101
Receiver Operating Characteristic (ROC), 229

Refinement
...of cyclic graphs, 179
...of free trees, 173
...of ordered trees, 115
...of unordered trees, 117
...basic definitions, 29
downward, 36
suboptimal, 31
upward, 36

Relations, 12
Rightmost path, 114

SD-Apriori, 237

Sequences, 13
SiGram, 164, 203
Simple occurrence sequence, 136
SMILES, 5
Spin, 199
Stamp point, 229
STUCCO, 235
SUBDUE, 204
Subgraphs, 161
Subpaths, 36
Subsequences
...with \((\alpha,\beta)\) gaps, 34
...with unlimited gaps, 34
...without gaps, 34
Subtrees
bottom-up, 110
embedded ordered, 107
embedded unordered, 107
induced ordered, 107
induced unordered, 107
ordered leaf, 109
prefix ordered, 109
Support, 10
Symmetry, 170

Transaction, 9
Transaction based support, 47
TreeFINDER, 152
TreeMinerV, 135
Trees, 105

uFREQT, 141, 148
Ullman’s subgraph isomorphism algorithm, 204
UML, 100
uNot, 140

Version space, 50
VF Algorithm, 204

WARMR, 83

XML, 113, 166
Acknowledgements

I would like to thank Eric-Wubbo Lameijer for building the molecular model of Cuneane of which a photo is included in this thesis. I enjoyed the discussions that I had with Eric-Wubbo and Jeroen Kazius about mining molecular databases. These discussions have motivated me very much, and I would like thank them for that. Of course I would also like to thank all colleagues that I used to have lunch and ‘coffee’ breaks with for making the period in Leiden an enjoyable one.