Mining Structured Data
Nijssen, Siegfried Gerardus Remius

Citation

Version: Corrected Publisher’s Version
License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from: https://hdl.handle.net/1887/4395

Note: To cite this publication please use the final published version (if applicable).
Bibliography


http://aria.njit.edu/mediadb/fast/.


Index

- Subsumption, 69

AcGM, 202
ADI-Mine, 198
AGM, 201
Aho, Hopcroft and Ullman's tree isomorphism algorithm, 134
Alternating path, 143
Anti-monotonicity, 40
Apriori algorithm, 14
property, 14
AprioriSMP, 233
Association rules, 9
Augmenting path, 143
Automorphisms, 179

Backbone depth tuple order, 172
Backtracking sequence, 133
Bias, 43, 73
Bipartite involved matchings problem, 143
Bongard datasets, 92

C-ArmR, 98
CAP, 60
CBA, 237
CBAmS, 237
CHARM, 64
Chemistry, 4, 94, 165, 211
CHOPPER, 151
Chung's subtree algorithm, 141, 201
CLOSE, 63
Closed itemsets, 51
CLOSEGraph, 162, 197
CMAR, 237
CMTREEMiner, 152
CN2-SD, 238
Complexities
- graph isomorphism algorithms, 163
- subgraph isomorphism algorithms, 163
- subtree isomorphism algorithms, 110
- tree isomorphism algorithms, 110
Confidence, 10
Constraint
- anti-monotonic, 41
- convertible, 43
- monotonic, 40
- Succinct, 44
Contingency table, 229
CorClass, 235
Cover, 10
Covers, 36
Cycle, 105

D-Lab, 97
Data tree, 110
Depth sequence, 114
Depth tuple, 114
DiffSets, 19, 191
Dryade, 153
DualMiner, 60

ECLAT, 18
Edge sequence, 162
Enumeration, 39, 130
Equivalence classes, 35
ExAnte property, 61

Farmer, 83
FFSM, 198
Formal concept analysis (FCA), 247
Formal concepts analysis (FCA), 52
<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP-Bonsai, 61</td>
<td>275</td>
</tr>
<tr>
<td>FP-Growth, 23, 60</td>
<td></td>
</tr>
<tr>
<td>Free itemsets, 51</td>
<td></td>
</tr>
<tr>
<td>FreeTreeMiner, 200, 201</td>
<td></td>
</tr>
<tr>
<td>Freqt, 139</td>
<td></td>
</tr>
<tr>
<td>Frequent itemsets, 9</td>
<td></td>
</tr>
<tr>
<td>FSG, 203</td>
<td></td>
</tr>
<tr>
<td>FST-Forest, 151</td>
<td></td>
</tr>
<tr>
<td>GBI, 204</td>
<td></td>
</tr>
<tr>
<td>GraphML, 113, 166</td>
<td></td>
</tr>
<tr>
<td>Graphs, 105</td>
<td></td>
</tr>
<tr>
<td>Greatest lower bound, 33</td>
<td></td>
</tr>
<tr>
<td>GSP, 64</td>
<td></td>
</tr>
<tr>
<td>gSpan, 162, 196</td>
<td></td>
</tr>
<tr>
<td>GXL, 113, 166</td>
<td></td>
</tr>
<tr>
<td>HybridTreeMiner, 151, 199</td>
<td></td>
</tr>
<tr>
<td>Hypergraphs, 112, 165</td>
<td></td>
</tr>
<tr>
<td>Inductive Logic Programming (ILP), 68, 165</td>
<td></td>
</tr>
<tr>
<td>Itemset occurrences, 10</td>
<td></td>
</tr>
<tr>
<td>Kirchoff’s matrix-tree theorem, 167</td>
<td></td>
</tr>
<tr>
<td>k-Prefix, 13</td>
<td></td>
</tr>
<tr>
<td>Large itemsets, 11</td>
<td></td>
</tr>
<tr>
<td>Lattice, 35</td>
<td></td>
</tr>
<tr>
<td>Learning from entailment, 97</td>
<td></td>
</tr>
<tr>
<td>Learning from interpretations, 97</td>
<td></td>
</tr>
<tr>
<td>Least upper bound, 33</td>
<td></td>
</tr>
<tr>
<td>Leftmost path, 172</td>
<td></td>
</tr>
<tr>
<td>Lexicographical order, 13</td>
<td></td>
</tr>
<tr>
<td>Maximal frequent itemsets, 50</td>
<td></td>
</tr>
<tr>
<td>Merge operators, 53</td>
<td></td>
</tr>
<tr>
<td>Merging</td>
<td></td>
</tr>
<tr>
<td>...of cyclic graphs, 185</td>
<td></td>
</tr>
<tr>
<td>...of free trees, 176</td>
<td></td>
</tr>
<tr>
<td>...of ordered trees, 115</td>
<td></td>
</tr>
<tr>
<td>...of unordered trees, 124</td>
<td></td>
</tr>
<tr>
<td>basic definitions, 54</td>
<td></td>
</tr>
<tr>
<td>downward, 55</td>
<td></td>
</tr>
<tr>
<td>Modes, 79</td>
<td></td>
</tr>
<tr>
<td>MoFA, 198</td>
<td></td>
</tr>
<tr>
<td>MolFea, 61</td>
<td></td>
</tr>
<tr>
<td>Monotonicity, 40</td>
<td></td>
</tr>
<tr>
<td>Multi-relational data mining, 4, 113, 165</td>
<td></td>
</tr>
<tr>
<td>Multicast dataset, 112</td>
<td></td>
</tr>
<tr>
<td>Nauty, 204</td>
<td></td>
</tr>
<tr>
<td>Next prefix node, 120</td>
<td></td>
</tr>
<tr>
<td>Non-derivable itemsets, 51</td>
<td></td>
</tr>
<tr>
<td>Object exchange model, 151</td>
<td></td>
</tr>
<tr>
<td>Object Identity, 71</td>
<td></td>
</tr>
<tr>
<td>Occurrence sequence, 136, 189</td>
<td></td>
</tr>
<tr>
<td>Occurrence tree, 145</td>
<td></td>
</tr>
<tr>
<td>Orders, 12</td>
<td></td>
</tr>
<tr>
<td>Path</td>
<td></td>
</tr>
<tr>
<td>...rooted, 106</td>
<td></td>
</tr>
<tr>
<td>...simple, 105</td>
<td></td>
</tr>
<tr>
<td>PathJoin, 151</td>
<td></td>
</tr>
<tr>
<td>Pattern tree, 110</td>
<td></td>
</tr>
<tr>
<td>PolyFarm, 99</td>
<td></td>
</tr>
<tr>
<td>Prefix trie, 15</td>
<td></td>
</tr>
<tr>
<td>PREFIXSPAN, 64</td>
<td></td>
</tr>
<tr>
<td>Primary key, 74</td>
<td></td>
</tr>
<tr>
<td>Projected database, 20</td>
<td></td>
</tr>
<tr>
<td>Proteins, 165</td>
<td></td>
</tr>
<tr>
<td>Query packs, 99</td>
<td></td>
</tr>
<tr>
<td>RAADAR, 101</td>
<td></td>
</tr>
<tr>
<td>Receiver Operating Characteristic (ROC), 229</td>
<td></td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>...of cyclic graphs, 179</td>
<td></td>
</tr>
<tr>
<td>...of free trees, 173</td>
<td></td>
</tr>
<tr>
<td>...of ordered trees, 115</td>
<td></td>
</tr>
<tr>
<td>...of unordered trees, 117</td>
<td></td>
</tr>
<tr>
<td>basic definitions, 29</td>
<td></td>
</tr>
<tr>
<td>downward, 36</td>
<td></td>
</tr>
<tr>
<td>suboptimal, 31</td>
<td></td>
</tr>
<tr>
<td>upward, 36</td>
<td></td>
</tr>
<tr>
<td>Relations, 12</td>
<td></td>
</tr>
<tr>
<td>Rightmost path, 114</td>
<td></td>
</tr>
<tr>
<td>SD-APRIORI, 237</td>
<td></td>
</tr>
<tr>
<td>Sequences, 13</td>
<td></td>
</tr>
<tr>
<td>SiGRAM, 164, 203</td>
<td></td>
</tr>
<tr>
<td>Simple occurrence sequence, 136</td>
<td></td>
</tr>
</tbody>
</table>
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.