Structural Pattern Recognition

Prof. em. Horst Bunke
Institute of Informatics and Applied Mathematics
University of Bern, Switzerland

bunke@iam.unibe.ch

Contents

1 General Introduction **2**

I General Methodology **6**

2 Introduction to Pattern Recognition with Symbolic Data Structures **7**

3 Pattern Recognition with Strings **9**
3.1 Basic Definitions 9
3.2 Edit Distance of Strings 12
3.3 Application Example: 2-D Shape Recognition 18

4 Pattern Recognition with Graphs **23**
4.1 Basic Definitions 23
4.2 Basic Algorithms 27
4.2.1 Graph and Subgraph Isomorphism 27
4.2.2 Maximum Common Subgraph 32
4.3 Graph Edit Distance 36
4.4 Other Graph Distance Measures 39
4.5 Applications 42
4.5.1 Character Recognition 42
4.5.2 Diatom Classification 45
4.5.3 Classification of Web Documents 49
4.6 Other Algorithms 52
4.6.1 Network based Approach 53
4.6.2 Decision Tree based Method 55
4.7 Spectral Graph Theory 60
4.8 The Median Graph Problem 66
4.9 Suboptimal Graph Edit Distance Computation 71
4.10 Graphs with Unique Node Labels 75
4.11 Hypergraphs 77
4.12 Information Retrieval with Graphs 81
4.13 Graph Databases for Experimental Work 84

II Kernel Functions for Symbolic Data Structures 91

5 String Kernel Functions 93
5.1 P-spectrum Kernel 93
5.2 All Subsequence Kernel 95
5.3 Gap-weighted Subsequence Kernel 96

6 Graph Kernel Functions 99
6.1 Graph Kernels based on Common Substructures 99
6.2 Walk Kernels 102
6.3 Dissimilarity Based Graph Embedding 106

7 General Kernel Theory 113
7.1 General Introduction 113
7.2 Nearest-Neighbor Classifier 118
7.3 Minimum Distance Classifier 118
7.4 Perceptron Algorithm 120
7.5 Support Vector Machine 123
7.6 K-means Clustering 125
7.7 Properties of Kernel Functions 127