GRAPHICAL MARKOV MODELS

Overview

Probabilistic Markov models on graphs represent a model class widely applied in many areas of computer science, such as computer networks, data security, robotics and pattern recognition. Hidden Markov Models on chains are for instance the most widespread and successful model class for automated speech recognition. Considering the view of artificial intelligence and machine learning, this model class is especially well suited for demonstrating how careful theoretical considerations lead to sometimes unexpected but highly practicable algorithms.

The course presents and discusses essential concepts, problems and algorithms for this model class being relevant for artificial intelligence and machine learning. The lectures are divided into two main chapters. The first chapter considers Markov models on chains and trees. In this situation, practically all inference and learning tasks including structure learning can be solved by efficient algorithms. The situation is more complicated for Markov models on general graphs. Here on the contrary, practically all inference and learning tasks are NP-complete. The focus of the second chapter is therefore on efficient approximative algorithms for inference and learning.

The course objectives are

(1) To gain in depth understanding of the chain "model–problem–algorithm" in the context of probabilistic Markov models especially for inference and learning problems.
(2) To enable students to adapt and apply the learned problem formulations and algorithms for different application contexts. To provide a basic ability to recognise the applicability of the learned concepts for AI/ML specific applications.

Annotation: Markov models on graphs represent a model class widely applied in many areas of computer science, such as computer networks, data security, robotics and pattern recognition. The first part of the course covers inference and learning algorithms for Markov models on chains and trees. The second part addresses graphical models on general graphs. Practically all inference and learning tasks are NP-complete in this situation. The focus is therefore on efficient approximative algorithms.

Prerequisites: Basics of probability theory, graphs and graph algorithms

Form of Teaching: 2p, 1s, 1c, winter semester

Credits: 6 CP

Lectures

Chapter I (Hidden) Markov Models on chains and trees.

L01: Markov chains, equivalent representations, ergodicity, convergence theorem for homogeneous Markov chains.
L02: Hidden Markov Models on chains for speech recognition: pre-processing, dynamic time warping, HMM-s
L03: Recognising the generating model – calculating the emission probability for a measured signal sequence.
L04: Recognising the most probable sequence of hidden states and the sequence of most probable states.
L05: Possible formulations for supervised and unsupervised learning tasks (parameter estimation).
L06: Algorithms for supervised and unsupervised learning according to the Maximum-Likelihood principle, the Expectation Maximisation algorithm for HMM (aka Baum-Welch algorithm).
L07: Hidden Markov models on acyclic graphs (trees). Estimating the graph structure.
L08: Hidden Markov models with continuous state spaces. Kalman filter and particle filters.

Chapter II Markov Models on general graphs.

L09: Markov Random Fields - Markov models on general graphs. Equivalence to Gibbs models, Examples from Computer Vision.
L10: Relations to Constraint Satisfaction Problems and Energy Minimisation tasks, unified formulation, semi-rings.
L11: Searching the most probable state configuration: transforming the task into a MinCut-problem for the submodular case.
L12: Searching the most probable state configuration: approximative algorithms for the general case.
L13: The partition function and marginal probabilities: Approximative algorithms for their estimation.
L14: Duality between marginal probabilities and Gibbs potentials. The Expectation Maximisation algorithm for parameter learning - a "difference of convex functions" algorithm.

Exercises

The lectures will be accompanied by exercises given in a "mixed mode": advanced exercises will be discussed in seminars. Their aim is to deepen the learned knowledge and to develop skills in recognising the applicability of the learned concepts. On the other hand, medium level programming assignments will be discussed and implemented in practical exercises. The focus is on efficient implementation and scalability. Groups of 2-3 students choose an application and implement the inference and learning for this specific application. Possible applications and data sets will be provided.

Annotation: Seminars (every two weeks): discussion of homework and advanced exercises. Project work (every two weeks): groups of 2-3 students choose an application and implement the inference and learning for this application. Possible applications and data sets will be provided.

P01: Decision on the HMM application project, data set, pre-processing
S02: Markov chains, application examples
P03: Implementation of the application environment, synthetic data generation
S04: Inference algorithms for HMM (I)
P05: Implementation of the inference
S06: Inference algorithms for HMM (II)
P07: Implementation of supervised learning
S08: Empirical risk minimisation, semi-supervised learning
P09: Implementation of the chosen learning approach
S10: Graphical image models
P11: Choosing the Computer Vision application, implementing the inference (using available Open Source software)
S12: Inference for Markov Random Field models
P13: Implementing the Gibbs sampler and supervised learning
S14: Presentation of the application projects

References

[1] Stan Z. Li. *Markov Random Field Modeling in Image Analysis*. Computer Science Workbench. Springer Verlag, 2. edition, 2001.

[2] Michail I. Schlesinger and Vaclav Hlavač. *Ten Lectures on Statistical and Structural Pattern Recognition*, volume 24 of *Computational Imaging and Vision*. Kluwer Academic Press, 2002.

[3] Gerhard Winkler. *Image analysis, random fields and Markov chain Monte Carlo methods: A Mathematical Introduction*. Springer, 2. edition, 2003. (For additional reading).