Syllabus
Graphical Models, Bayesian Learning, and Statistical Relational Learning, 6hp

Issued by the WASP graduate school management group 2021-06-11.

Main field of study

AI/math, AI/mlx

Course level

PhD Student course

Course offered for

PhD Students in the WASP graduate school

Entry requirements

The participants are assumed to have a background in mathematics corresponding to the contents of the WASP-course "Mathematics and Machine Learning".

The course requires basic understanding of probability theory and graph theory. In particular, the participants need to be comfortable with reasoning about (in-)dependence of random variables and factorizing joint probability distributions/density functions using the chain rule of probability, as well as reasoning about basic properties of graphs.

Module 2 requires familiarity with Bayesian statistics (as well as graphical models and basic message-passing algorithms, covered in module 1).

For module 3 working knowledge of propositional logic and first-order logic (including first-order structures) is required.

Intended learning outcomes

After completing the course, students should be able to

- Describe how undirected and directed acyclic graphs encode (probabilistic)independence information.
- Describe and use methods for learning directed acyclic graph representa-tions of conditional independence relations satisfied by a data-generatingdistribution.1
- Describe the Bayesian inference problem for graphical models, and explainwhen it can be solved exactly and when approximate inference algorithmsare needed.
- Describe and use methods for exact and approximate probabilistic infer-ence using graph representations of a data-generating distribution.
- Describe the key differences between deterministic and stochastic (i.e.,Monte–Carlo-based) approximate inference algorithms.
- Use basic methods of Inductive Logic Programming to learn a logic pro-gram and use if for inference.

- Use the concept of "probabilities on domains" to learn a parametrizedGraphical Model from a relational structure.
- Use the concepts of "probabilities on possible worlds" and parametrizedgraphical model to make inferences on arbitrary domains.

Course content

In module 1 we will introduce undirected graphical and directed acyclic graphical models as parametric statistical models whose parameterizations are defined according to a given undirected or directed acyclic graph. We then see how such models are equivalently defined as those distributions which satisfy sets of conditional independence relations encoded in the edge structure of their associated graph. We will apply these two different interpretations of our model to address observe the following: The parametric interpretation of graphical models will provide effective algorithms for exact probabilistic inference with computational complexity bounds related to the graph structure, and the conditional independence interpretation will provide some first algorithms for learning representations of cause-effect systems from data. Topics to be discussed include, undirected and directed acyclic graphical models, pairwise, local and global Markov properties for graphs, the Hammersley-Clifford Theorem, the variable elimination and clique-tree algorithms for exact probabilistic inference, Markov equivalence, and basic algorithms for causal discovery.

In Module 2 we consider Bayesian inference in graphical models where the involved distributions and/or the structure of the graph prevents exact solutions. We introduce several approximate inference algorithms to tackle this issue, specifically: expectation propagation, variational inference, and Markovchain Monte Carlo. We highlight some key differences, but also similarities, between these types of algorithms.

Module 3 introduces the basic ideas in Inductive Logic Programming. It explains the two main approaches to unifying logic and probability which are used in Statistical Relational Learning, the "probabilities on domains approach" and the "probabilities on possible worlds approach". It describes how the first approach can be used for learning a graphical model from a set of data consisting of objects and relations between them, and how the second approach, together with a parametrized graphical model, can be used for making inferences on arbitrary finite domains. The issue of dealing with large domains without exploding computational costs is discussed.

Teaching and working methods

Module 1 will consist of lectures, exercise sessions, and homework problem sets. The exercise sessions will follow the lectures, giving the students an opportunity to work with the newly introduced theory. The homework assignments will be used as examination. Module 2 will consist of lectures and practical sessions. The latter will be centered around implementing and evaluating approximate inference algorithms using numerical software (preferred language is Python). The knowledge and skills related to Module 3 are acquired from literature studies, problem solving and a few lectures.

Examination

Participants in the course will be assessed on all the modules using hand in assignments. For each module there will be a set of regular hand in assignments that are common to all. In addition to this, each participant is required to do an additional hand in assignment going deeper into the concepts covered by one of the three modules. This specialization is chosen at the start of the course.

Grades

Fail or Pass