PROBABILISTIC MODELS – PART 1: INTRODUCTION TO THIS CLASS

Gerhard Widmer

Institute of Computational Perception
Johannes Kepler University
Linz, Austria

gerhard.widmer@jku.at
www.cp.jku.at/people/widmer

October 2, 2025

Presentation partly based on and inspired by [Koller & Friedman, 2009] and [Russell & Norvig, 2021], including the use of some figures from their books and/or lecture slides.

Many thanks to Daphne Koller, Nir Friedman, Stuart Russell, and Peter Norvig for making these available (`pgm.stanford.edu`; `aima.cs.berkeley.edu`).

Goals of this Lecture

- Explain what this class will be about
- Place the topic in its appropriate context
- Motivate you to stay with us for the rest of the class.

Outline

A Special Message ...

... from the Institute “Integriert Studieren”, JKU:

- Service & Support-Center for ALL students with
  - **Disability** (sensory, mobility, manipulation),
  - **Chronic & mental health issues** and
  - **Neurodivergence** (e.g. Dyslexia, ASD / Asperger, AD[H]S)
- Service & Support with:
  - **Exam adaptation** (written and oral exams) (e.g. additional time, quiet room, laptop, assistive technologies, writing assistance)
  - **Missed compulsory lectures** due to health issues,
  - **Accessible adapted learning materials**,
  - **Problems** (e.g. minimum ECTS in first semesters)

A Special Message ...

... from the Institute “Integriert Studieren”, JKU:

Dyslexie-Lab-Selbsteinschätzung:

- Für Personen mit Lese-Rechtschreib-Schwierigkeiten, Legasthenie, Dyslexie (LRS)
- Kostenlose Diagnostik, Gutachten, Beratung, Unterstützung

Example 1 – Automatic Diagnosis of Complex Systems

Given:

- A structured model of a spacecraft (or some sub-system)
- Sensor signals indicating certain errors and other conditions (*"Houston, we have a problem ..."*)

Task:

- Identify most likely fault that could be the cause of this
- Evaluate what would happen if certain repair actions were taken
- Decide on measures to fix the problem.

Example 2 – Autonomous Driving

Given:

- A road map
- Streams of sensor signals (GPS, cameras, radar, infrared range sensors, ...)

Task:

- Autonomously drive a car through difficult terrain
- in the presence of other cars, obstacles, ...

Example 3 – Automatic Aircraft Tracking

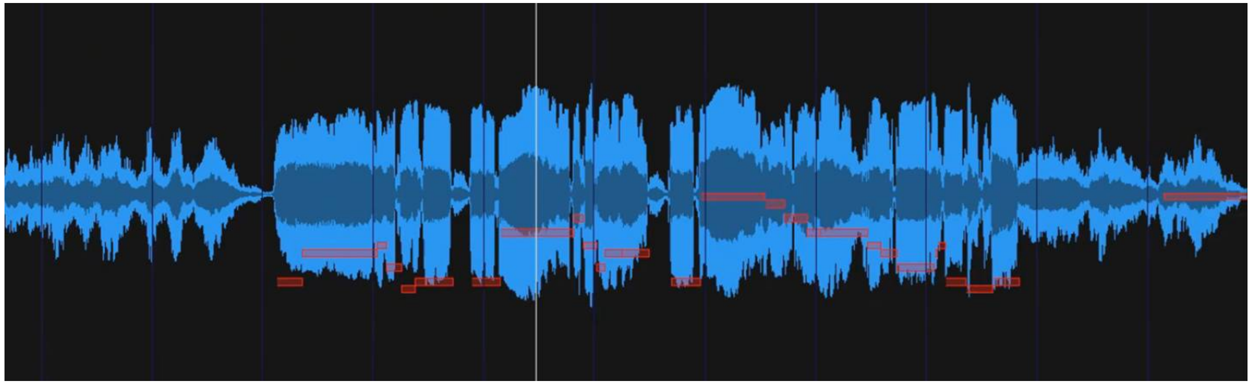
Given:

- Radar or video signals (intermittent, sometimes occluded)

Task:

- Track aircraft over time
- Predict where aircraft will be in 0.1 / 1.0 / 10 sec
- Determine how many aircraft there are
- Keep track of which is which

Example 4 – Speech Recognition and Voice Control



Given:

- Person speaking into a microphone (audio)

Task:

- Segment audio stream into separate words
- Understand / transcribe text that is spoken
- Use speech commands to control complex devices.

Example 5 – Real-time Music Following[1]

Given:

- Live music performance on stage (audio stream through microphone)
- Printed sheet music (score)

Task:

- Listen to performance and automatically track performers' position in score
- Synchronise live music with other events (e.g., page turning)

[1] https://www.youtube.com/watch?v=SUBtND_MJZs

What These Scenarios Have in Common

An autonomous system (robot, car, diagnosis system, tracker, ...)

- Receives information about its environment, at regular or irregular intervals
- Must identify objects, classify situations, make predictions, take decisions
- **Information may be incomplete, uncertain, noisy, partly wrong and contradictory**

☞ **Logic-based models or deterministic approaches not possible.**

What These Scenarios Have in Common

They can solved with PROBABILISTIC (GRAPHICAL) MODELS.

Probabilistic Models

- represent knowledge about the world, and about its uncertainty
- support logical and probabilistic inference (decision-making)
- can be learned from example observations
- technically: are graphical representations of **factorised probability distributions.**

Research Context

- Research field in the intersection of Artificial Intelligence, statistics, probability theory, machine learning, and algorithms.

The AI Philosophy: Declarative Modelling

General approach:

- Construct a **model** of the problem/world to be reasoned about
- Devise general **inference algorithms** that operate on this model and draw conclusions
- This is called the **declarative approach to problem solving**

  (as opposed to the *procedural approach* of developing problem-specific algorithms)

Advantages:

- Clear separation of knowledge and reasoning
- Representation (model) has a clear, well-defined semantics (interpretation), independent of the algorithms that operate on it
- Inference algorithms can be entirely generic and independent of the specific model they are applied to
- Can change / adapt a system simply by changing the knowledge base (model).

What are Probabilistic Graphical Models?

Probabilistic Graphical Models are

- A formalism for modelling the **structure** of some world or domain,
- and for quantifying the **degree of uncertainty** of information.
- Called "probabilistic" because they represent probability distributions
- Called "graphical" because they make dependencies explicit in a graph structure.

Central Issues:

- **Modelling**: representing the relevant aspects of a given world, and their connections
- **Inference**: taking (incomplete, uncertain, noisy) observations about the world, and using these and the model to reach conclusions about unobservable or future facts
- **Learning**: automatically building models of complex worlds from (large numbers of) observations.

What You Will Learn in this Class

SEMANTICS: Different types of models, and what they represent

- Static Models: Bayesian Networks
- Temporal Models: Dynamic Bayes Nets, HMMs, Kalman Filters
- Not covered in this class: Undirected and Semi-directed Models

INFERENCE: How to use models to answer questions

- Deterministic algorithms
- Probabilistic algorithms (stochastic sampling)

LEARNING: How to learn models from data

- Parameter learning
- Structure learning

Administrative Stuff

344.036 (VO, 2 hrs, 3 ECTS):

- Time and Place: Tue, 09:15 - 10:45, HS 7
  Live streaming via Zoom (for link, see Moodle page)
- Availability of instructor: (almost) always, via e-mail to `gerhard.widmer@jku.at`

344.037/... (UE, 1 hr, 1.5 ECTS):

- Taught be Paul Primus & Florian Schmid
- Opportunity to practically experiment with methods and algorithms presented in the VO, and to deepen your understanding

  (VO will be rather theoretical)
- Is a mandatory course in the AI curriculum; counts as an ‘elective course’ in the CS curriculum.

Class Materials

Lecture Slides and Videos:

- Lecture slides (pdf) and videos (recordings of live lectures) will be provided via Moodle
- Important: Videos and slides are for your own use only.
  I do not want to see them appear on the Web!

Literature:

- Is listed at the end of each set of slides
- Is not needed if you read the lecture slides.

Planned Schedule (WS 2025/26)

VO 1	7.10.2025:	Introduction; Basic Concepts of Probability Theory
VO 2	14.10.2025:	Probability: Inference and Independencies
VO 3	21.10.2025:	Bayesian Networks: Semantics
VO 4	28.10.2025:	Bayes Nets: Exact & Approximate Inference (1)
VO 5	4.11.2025:	Bayes Nets: Approximate Inference and MCMC
VO 6	11.11.2025:	Learning (1): General Issues & ML Parameter Learning
VO 7	18.11.2025:	Learning (2): Bayesian Parameter Learning
VO 8	25.11.2025:	Learning (3): Overfitting, Regularisation, Structure Learning
VO 9	2.12.2025:	Dynamic Bayes Nets and State-Observation Models
VO 10	9.12.2025:	Hidden Markov Models (1): Inference
VO 11	16.12.2025:	Hidden Markov Models (2): Learning and Applications
—		— Christmas Break —
VO 12	13.1.2026:	Modelling Continuous Linear Systems: The Kalman Filter
	20.1.2026:	**Exam (written; on-site)**

The Institute of Computational Perception

`www.jku.at/en/institute-of-computational-perception`

We develop and study computational models and algorithms that permit computers to perceive and ‘understand’ aspects of the external world, where we interpret ‘perception’ in the widest sense of the word, as the extraction of useful high-level information and knowledge from complex, possibly low-level data (audio, video, images, sensor data, texts, or even the Internet). This requires fundamental and applied research on **Artificial Intelligence, Machine Learning, Signal Analysis, and Statistical and Probabilistic Modelling**. Our current research has a particular focus on intelligent **Audio and Music Processing**. Further research topics include biometrics, cryptography, personalization, and recommender systems.
Our ambition is to be among the best research teams in the world in this field (and we are), and to provide an inspiring teaching environment that permits students to get involved in real research projects as early as possible.

If you are interested in these topics, work with us in

- courses and seminars
- practical projects, masters & PhD theses
- industrial and scientific research projects

Probabilistic Models: A First Intuitive Introduction

Consider a Simple Medical Diagnosis Setting:

- Two possible diseases: flu and hayfever
- Two possible symptoms: (sinus) congestion and muscle pain
- Also possibly relevant: the season (spring, summer, fall, winter)

Task: Build a system that automatically answers questions about a patient, given certain pieces of evidence

- How likely is it that a patient has the flu, given that it is fall, and that she has a sinus congestion but no muscle pain?
- How likely is it that a patient has *no* hayfever, given that she has a muscle pain and the flu?
- What is the most likely season for people to have a sinus congestion, but no hayfever?
- etc.

A Probabilistic Approach to Modelling

Basic Idea: Take a Probabilistic Approach

- Model all relevant entities in this world as a set $\mathcal{X}$ of *random variables*
- Model relationships/influences between these as (conditional) probabilities
- View question answering as the task of computing conditional probabilities, given some observations or things you know are true.

Will See:

- The answer to any desired query can be computed from the joint probability distribution over the variables in $\mathcal{X}$
- This joint distribution can often be represented in a very structured form (graph), as a *Graphical Model*
- Each Graphical Model corresponds to some joint distribution
- The answer to any query can be directly computed from such a model.

A Model of the Medical Diagnosis World

Modelling the diagnosis world with 5 discrete random variables:

Flu	$\in \{\mathit{true}, \mathit{false}\}$
$\mathit{Hayfever}$	$\in \{\mathit{true}, \mathit{false}\}$
$\mathit{Congestion}$	$\in \{\mathit{true}, \mathit{false}\}$
$\mathit{MusclePain}$	$\in \{\mathit{true}, \mathit{false}\}$
Season	$\in \{\mathit{spring}, \mathit{summer}, \mathit{fall}, \mathit{winter}\}$

$$\Rightarrow \mathcal{X} = \{\mathit{Flu}, \mathit{Hayfever}, \mathit{Congestion}, \mathit{MusclePain}, \mathit{Season}\}.$$

Assumptions:

- The variables describe the state of the world (a specific patient) at a given time point
- The current state of the world is characterised exclusively by the values of the random variables
- A variable always takes exactly one of its possible values
- Some of the values may be known to us, some may be hidden

A Model of the Medical Diagnosis World

Second Part of the Model: Joint Probabilities

- Assume that we know the probability for **each specific combination of values** of the 5 variables to occur in our world
- The set of these probabilities (covering all possible cases) is called the **(Full) Joint Distribution** over $\mathcal{X}$

$$
\begin{array}{lcl}
P(\neg flu, \neg hayfever, \neg congestion, \neg musclepain, spring) & = & 0.014 \\
P(\ \ flu, \neg hayfever, \neg congestion, \neg musclepain, spring) & = & 0.003 \\
P(\neg flu, \ \ hayfever, \neg congestion, \neg musclepain, spring) & = & 0.008 \\
P(\ \ flu, \ \ hayfever, \neg congestion, \neg musclepain, spring) & = & 0.001 \\
P(\neg flu, \neg hayfever, \ \ congestion, \neg musclepain, spring) & = & 0.009 \\
\ldots\ldots\ldots & & \ldots
\end{array}
$$

The Full Joint Distribution

- is a table of probabilities, one for each possible value combination
- sums to 1.0
- in our case, has $2 \times 2 \times 2 \times 2 \times 4 = 64$ entries

Challenges

In this Class:

- Will see that the questions we want to ask of the model are **conditional probability queries**

  Example: $P(Flu = true \mid Season = spring, MusclePain = false)$?
- Will show that the answer to **any query of this type** can be computed from the full joint distribution

☞ Have a **universal query answering algorithm** for probabilistic queries!

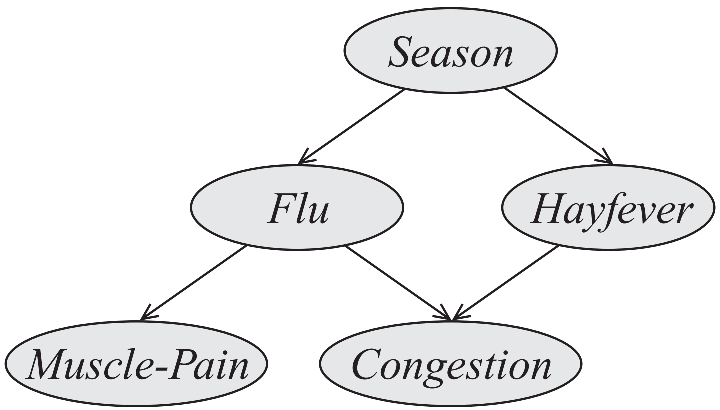
Problems:

- **Representational Complexity:**
  Joint distribution over large number of variables is too large to store (exponential in number of variables!)
- **Computational Complexity:**
  Computing the exact answer requires exponential time
- **Modelling Complexity:**
  Difficult to construct a valid model for a complex world.

How the Problems Will Be Addressed

Re: Representational Complexity

- We will study the concept of **(conditional) independence**
- use independencies to decompose (“factorise”) the full joint distribution
- and represent the factorised distribution as a compact **graphical model**,
- which reflects (in)dependencies (and sometimes also causal structure)



Effect: (Often) drastic reduction of complexity!

How the Problems Will Be Addressed

Re: Computational Complexity

- We will design algorithms that can compute the answer to any probabilistic query directly from a graphical model
- Look at ways to optimise these algorithms
- Develop algorithms that compute **approximate solutions**, **quickly**.

Re: Modelling Complexity

- We will design algorithms that can **automatically construct ("learn")** models from (large numbers of) concrete observations

☞ **Machine Learning**

For all that: Will need basic knowledge of Probability Theory.

☞ See next set of slides.

Literature

The following two books are excellent textbooks on Probabilistic Models and Artificial Intelligence, respectively. I recommend them to anyone seriously interested in these topics. However, you will not need to read the books in order to follow and understand this class. The slides should be self-explanatory.

Koller, Daphne and Friedman, Nir (2009).
Probabilistic Graphical Models: Principles and Techniques.
Cambridge, MA: MIT Press.

Russell, Stuart J. and Norvig, Peter (2021).
Artificial Intelligence: A Modern Approach. Fourth Edition.
Pearson Publishers.