

Probabilistic Reasoning and Decision Making

INF581 Advanced Machine Learning and Autonomous Agents

Jesse Read



Last Updated: January 8, 2025

Outline

1 Introduction

2 Bayesian Networks

3 Decision Theory

4 Summary

Introduction

1 Introduction

2 Bayesian Networks

3 Decision Theory

4 Summary

What is AI? What is an Autonomous Agent?

WIKIPEDIA: *AI is the study of “intelligent agents”: any device that **perceives** its environment and **takes actions** that maximize its chance of success at **some goal***

We wish to build a function

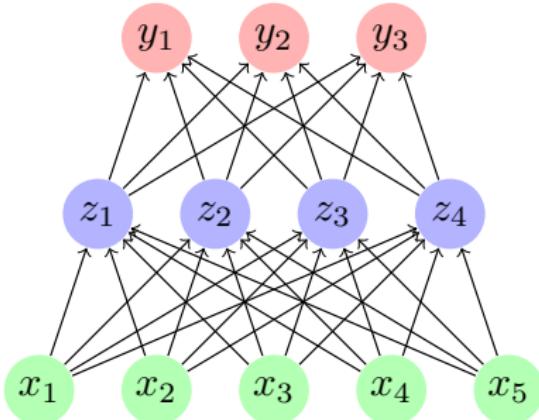
$$f : \text{Observation} \mapsto \text{Action}$$

to map **observation** (*input*; e.g., feature vector, images, text, sensor-readings, ...) to **action** (*output*; e.g., label, decisions, predictions, actuators) to attain some **goal** (*reward*; i.e., optimise some performance metric).

The question is: **How to build this function?**

The Existence of f

- Turing Machines (1936): Capable of simulating *any* computer algorithm
- Universal Approximation (1989): A simple neural network can approximate *any* continuous function f



So our function f can exist – we just need to obtain it. In this course: via data-driven machine learning.

Application Domains

- Autonomous vehicles, robotics and navigation
- Banking / finance
- Large Language Models (translation, chatbots, ...)
- Smart grids, dynamic resource management
- Healthcare (diagnostics, drug discovery, ...)
- Marketing; Recommendation systems
- Fraud prevention / anomaly detection / cybersecurity
- Games
- Art and design
- Agriculture, social challenges, politics



Descriptive vs Predictive vs Prescriptive Models

- **Descriptive models** (data mining, pattern mining, data analytics, ...)⇒ describe what we already have
- **Predictive models** (machine learning, statistical methods, extrapolation, forecasting, ...)⇒ use the past and present to predict the future
- **Prescriptive models** (optimisation and autonomous agents, reinforcement learning, solvers, planning, ...) ← In this course, we are here⇒ take actions that will change the future

Impressive Progress

2004:

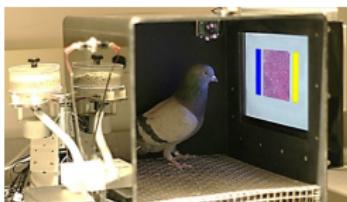
- Human can learn how to play Go in a few minutes and go on to beat the state-of-the-art AI
- DARPA Grand Challenge: None of the autonomous vehicles finished; the best team completed 11.78 of the 240 km
- Computer vision still largely unsolved. Focus on mining text associated with any given image, or 'bag of pixels'.
- Speech recognition had plateaued, needed heavy personalized training; Machine translation provided 'amusing' examples

2025, Already seems like old news:

- AI has beaten world champion in Go
- Autonomous vehicles drive hundreds of kilometres through urban environments
- Computer vision and speech recognition penetrating the mass market (Siri, Google Translate, etc.);
Impressive results from of language models, text-to-image models (ChatGPT, Midjourney, . . .).

But still a long way to go. Consider **AI vs Pigeon**:

- Getting from point A to B without crashing into anything, over long distances, in 3d coordinates, and **orders of magnitude less power** than modern AI
- Impressive visual skills: *Pigeons identify breast cancer 'as well as humans'*: *Pigeons, with training, did just as well as humans in a study testing their ability to distinguish cancerous from healthy breast tissue samples. [...] Likely no bigger than the tip of your index finger, the pigeon's brain nonetheless has impressive capabilities . . . Pigeons can distinguish identities and emotional expressions on human faces, letters of the alphabet, missshapen pharmaceutical capsules, and even paintings by Monet vs Picasso.*



– *BBC* (2015); Scientific article: Levenson et al., *Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images*, *PLOS ONE* 2015; Turner and Edward *The pigeon as a machine: Complex category structures can be acquired by a simple associative model*. *Iscience* 26.10 (2023).



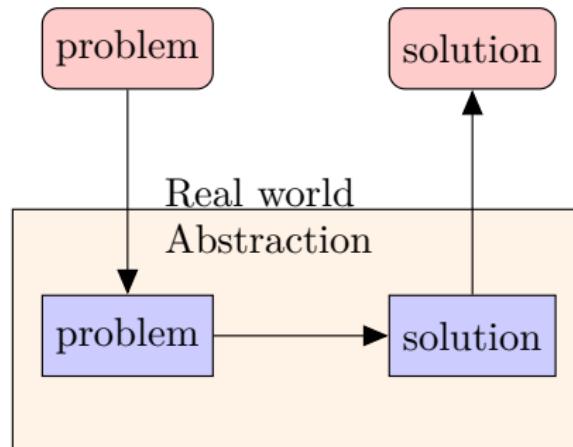
[1] ('Robot fails', YouTube) – Robots failing at basic tasks

Goals of this Course

If an agent perceives, and takes actions, to maximise its performance, we want to formalise this; map observations to action to maximize performance.

Where “good” performance, is such that the agent is useful, impressive, or provides us with greater understanding/knowledge.

In short, we are interested in **autonomous agents**, and we get them via **advanced machine learning**.



Theoretical abstraction, and practical intuitions.

Goals of Today

- What is an autonomous/intelligent agent
- A look at the pipeline: from perception (observation), to inferring state, to taking action.
- Understand Bayesian Networks as knowledge representation of knowledge and tool for probabilistic reasoning

$$P(S | x)$$

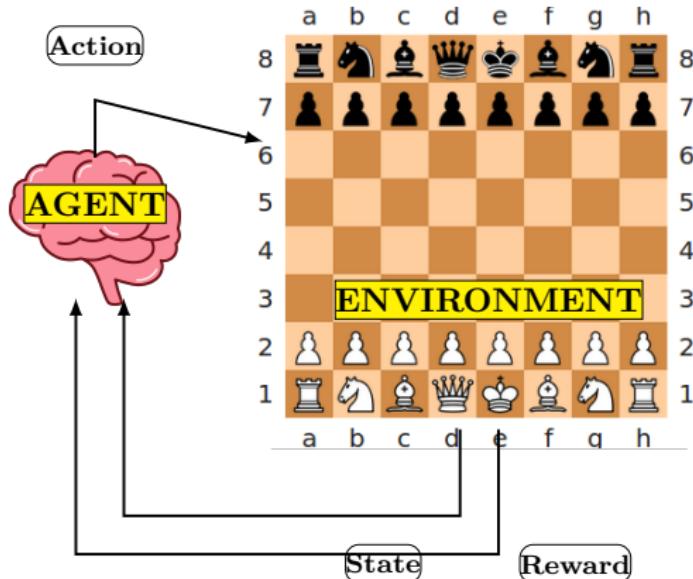
- Formalisation of decision making under uncertainty

$$\underset{a}{\operatorname{argmax}} \underbrace{\mathbb{E}_{S \sim P(S|x)}[r(S, a)]}_{Q(a)}$$

Key takeaways:

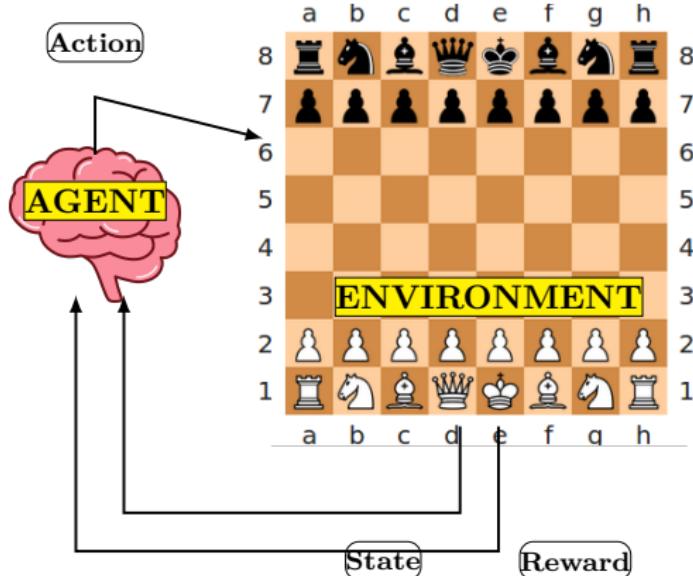
- ① understanding why this is challenging, and
- ② a probabilistic framework and set of tools with which to develop solutions.

Agents



An agent maps **perception** (state, reward) to **action** in an **environment**. Which action?

Agents



An agent maps perception (state, reward) to action in an environment. Which action? The one that maximises its performance metric (reward).

Components of an Agent

To build

$$f : \text{Observation} \rightarrow \text{Action}$$

we need:

- **Observation** ← Assume access to some perception
- **Knowledge and Representation** ← Today (and later)
- **Reasoning and Decision Making** ← Today
- **Planning/Sequential decision making** ← Later
- **Learning** ← From Week 2 onwards
- **Acting** ← Assume our choice can be actuated

Motivation: Reasoning, Knowledge and Planning

- We cannot plan/decide if we cannot reason.
- Without reasoning, we would have to store infinite information (a lookup table is not viable!)
- We reason from knowledge
- Thus we need to obtain knowledge (i.e., via perception, learning from experience),
- And we need to represent that knowledge for access and reasoning.

This is a difficult problem!

An example: you ask your AI agent:

Can you buy some beer for me for when I get home.

It needs to know ...

What is beer (what does it look like, sound like, . . .)? What is it used for? How much does beer weigh? How much is ‘some’ beer? Where to obtain beer? When to obtain it? How to get there? How to transport beer? Does the beer need transporting? . . . What is home? Where is home? How to buy something? What is money? How much does beer cost? What happens if payment doesn’t work? . . . Is this even a good idea to buy beer? Should I carry out this request? What are the consequences if I don’t? Should I even be taking requests?

Unblurring the Lines

Is a 'classifier' (via supervised learning) an autonomous agents?
Can an agent's actions (resulting from its decisions) affect its future observations; i.e., **interaction** with its environment?

Decision making vs Planning? A **decision is a single-step plan**.

More generally, planning involves pre-visualising a sequence of steps, involves having an internal model of the world; to obtain some non-immediate goal.

The tools we develop today for probabilistic reasoning and decision making can (and will) be used throughout!

Bayesian Networks

- 1 Introduction
- 2 Bayesian Networks
- 3 Decision Theory
- 4 Summary

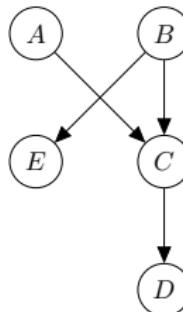
Probabilistic Graphical Models

A Probabilistic graphical model (PGM):

- A marriage between graphs and probabilistic models.
- Provide representation for knowledge, and mechanisms to reason from it
- Used throughout machine learning and beyond, especially when structured outputs are involved
- There are different types of PGM. In this course (and today) we look at Bayesian Networks (also known as Belief Networks)

Bayesian Networks: Main Concepts

- A directed acyclic graph (DAG), e.g.,



- Nodes are **random variables**, e.g., A, B, C, \dots
i.e., each has a distribution, $P(C | A, B)$, etc.
- Edges represent **conditional dependence** between variables,
e.g., $P(C | A, B)$ because A and B are the parents of C
- The graph specifies (**factorizes**) a **joint probability distribution**:

$$P(A, B, C, D, E) = P(A)P(B)P(C|A, B)P(D|C)P(E|B)$$

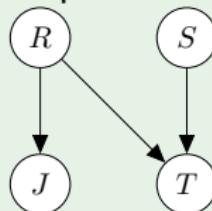
- N.B. Edge-relations are *not necessarily causal* – but can be

Bayesian Networks: Representing Knowledge

A Bayesian Network can be used as a representation of knowledge

Tracey's Grass (example from *D. Barber*)

$T = 1 \Leftrightarrow$ Tracey's grass is wet. $J = 1 \Leftrightarrow$ Jack's grass is wet.
 $S = 1 \Leftrightarrow$ The sprinkler is on. $R = 1 \Leftrightarrow$ It's been raining.



Knowledge:

$P(R = 1) = 0.2$	(sometimes it Rains)
$P(S = 1) = 0.1$	(sometimes Sprinkler is on)
$P(T = 1 R = 1, S = 0) = 1$	(Rain always wets grass)
$P(T = 1 R = 0, S = 0) = 0$	(only Rain/Sprinkler cause wet)
$P(T = 1 R = 0, S = 1) = 0.9$	(Sprinkler usually wets grass)
$P(T = 1 R = 1, S = 1) = 1$	(... and with rain, always)
...	...

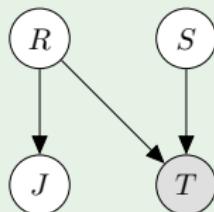
(each Node has a distribution).

Bayesian Networks for Probabilistic Reasoning (Inference)

A Bayesian Network can be used for reasoning (i.e., inference) from our knowledge.

Inference on Tracey's Grass

One morning Tracey leaves her house and realises that her grass is wet ($T = 1$). Was the Sprinkler left on overnight ($S = 1$)?



but first, ...

A Generalisation and Modern Context

In modern representations, we may have millions of parameters.

We typically organise into layers of variables, e.g., X_1, \dots, X_d

As before, each node X_j has a **conditional probability distribution**

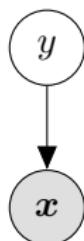
$$P(X_j | \text{pa}(X_j))$$

and **joint distribution**, over the m variables/nodes:

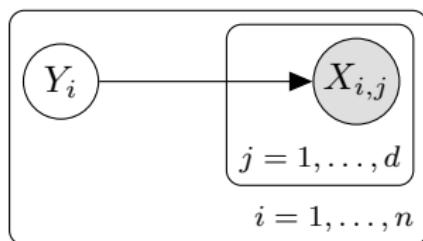
$$P(\mathbf{X}) = \prod_{j=1}^d P(X_j | \text{pa}(X_j))$$

For example,

$$P(\mathbf{X}, Y) = P(Y) \prod_{j=1}^d P(X_j | Y)$$



(or, in plate notation)



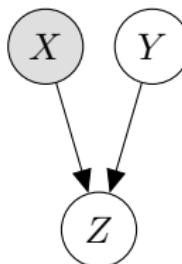
Queries and Observations/Evidence (Inference)

We formulate a **query**, given some **observation/evidence**, e.g.,

$$P(\mathbf{Y} = \mathbf{y} \mid \mathbf{X} = \mathbf{x})$$

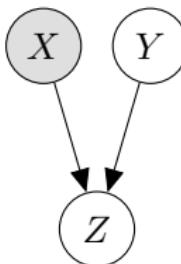
i.e., with

- ① **observation**, often \mathbf{x} N.B. these nodes are **shaded**
- ② **query**, often \mathbf{y}
- ③ **hidden/latent/'nuisance'** variable, often \mathbf{Z}
- ④ **knowledge** in the form of the graph



Marginalizing Out

We **marginalize out** (sum out, via sum rule/law of total probability) the hidden variables.



For example, with **knowledge** P (as defined by the graph), and observation x , we want to **reason** about y :

$$\begin{aligned} P(Y = \textcolor{magenta}{y}, X = \textcolor{brown}{x}) &= P(\textcolor{magenta}{y}, \textcolor{brown}{x}, Z) \quad \triangleright \text{Plug in query and observation} \\ &= P(Z|x, y)P(x)P(y) \quad \triangleright \text{Factorize (using graph)} \\ &= \sum_z P(\textcolor{brown}{z}|x, y)P(x)P(y) \quad \triangleright \text{Marginalize out } Z \\ &= P(x)P(y) \sum_z P(z|x, y) \quad \triangleright \text{Push sum right} \\ &= P(x)P(y) \quad \triangleright \text{Cancel/clean up} \end{aligned}$$

Joint vs Marginal vs Conditional (vs Query)

$$P(X, Y) = P(Y|X)P(X) = P(X|Y)P(Y):$$

	$x = 0$	$x = 1$
$y = 0$	0.3	0.2
$y = 1$	0.1	0.4

$$P(Y) = \sum_{x \in \{0,1\}} P(Y, x):$$

$y = 0$	0.5
$y = 1$	0.5

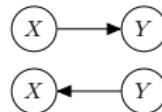
$$P(Y|X) = P(Y, X)/P(X):$$

	$x = 0$	$x = 1$
$y = 0$	0.75	0.33
$y = 1$	0.25	0.67

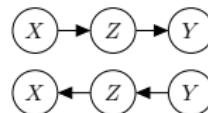
Plugging in query y and evidence x : $P(Y = 1|X = 0) =$

Dependence: When are X and Y dependent?

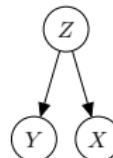
- Direct connection



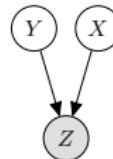
- Indirect connection (cascade)



- Common parent ('cause')



- Common child ('explaining away'); Z is *observed!*



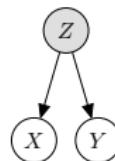
N.B. It does not matter if X or Y is observed/shaded or not

Independence: When are X and Y independent?

- Not connected (**marginal independence**)



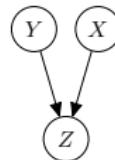
- Z is common evidence (**conditional independence**)



- Z blocks path between X and Y (**conditional independence**)



- Z is a **collider** node (N.B. Z is *not* observed in this case)



N.B. It does not matter if X or Y are observed/shaded or not

Marginal vs Conditional Independence

When X is (marginally) **independent** of Y we write:

$$X \perp\!\!\!\perp Y$$

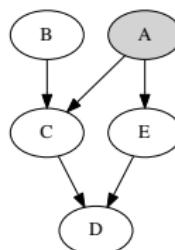
When X is **conditionally independent** of Y **given** Z we write:

$$X \perp\!\!\!\perp Y|Z$$

Dependence is **symmetric**:

$$X \perp\!\!\!\perp Y \Leftrightarrow Y \perp\!\!\!\perp X, \quad \text{and} \quad X \perp\!\!\!\perp Y|Z \Leftrightarrow Y \perp\!\!\!\perp X|Z$$

Example: Is $C \perp\!\!\!\perp E|A$?



Marginal vs Conditional Independence

When X is (marginally) **independent** of Y we write:

$$X \perp\!\!\!\perp Y$$

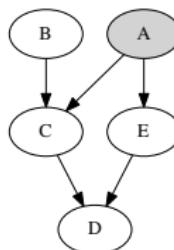
When X is **conditionally independent** of Y **given** Z we write:

$$X \perp\!\!\!\perp Y|Z$$

Dependence is **symmetric**:

$$X \perp\!\!\!\perp Y \Leftrightarrow Y \perp\!\!\!\perp X, \quad \text{and} \quad X \perp\!\!\!\perp Y|Z \Leftrightarrow Y \perp\!\!\!\perp X|Z$$

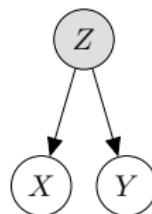
Example: Is $C \perp\!\!\!\perp E|A$?



Answer: Yes, influence is blocked above by 'common evidence', and blocked below by 'collider'.

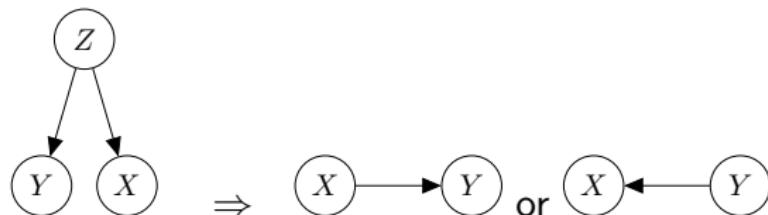
Common Evidence (Conditional Independence)

Where Z is observed as evidence, $X \perp\!\!\!\perp Y|Z$:



$$P(X, Y|Z = z) = P(X|Z = z) \cdot P(Y|Z = z) \quad \triangleright X \perp\!\!\!\perp Y|Z$$

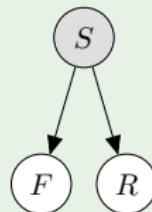
But this does not imply that $X \perp\!\!\!\perp Y$ *without* evidence Z ! In that case, we have to **marginalize out Z** :



$$P(X, Y) = \sum_{z \in \mathcal{Z}} P(X|Z = z) \cdot P(Y|Z = z) \cdot P(Z = z) \quad \triangleright X \not\perp\!\!\!\perp Y$$

Speeding

The probability of receiving a Fine $\perp\!\!\!\perp$ Red Car | Speed

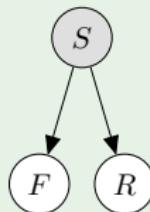


But, $F \not\perp\!\!\!\perp R$ when there is no observation on driving behaviour.

Car-insurance companies know this!

Speeding

The probability of receiving a Fine $\perp\!\!\!\perp$ Red Car | Speed



But, $F \not\perp\!\!\!\perp R$ when there is no observation on driving behaviour.

Car-insurance companies know this!

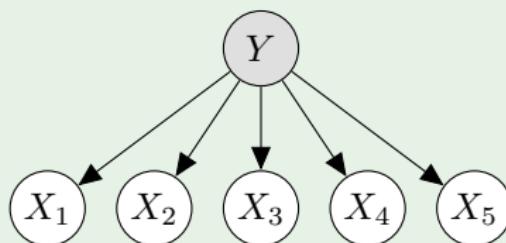
Also note ([3] AI fail from 2018):

Amazon has apparently abandoned an AI system aimed at automating its recruitment process . . . [It] penalised résumés that included the word ‘women’s’, as in ‘women’s chess club captain’ and marked down applicants who had attended women-only colleges . . .

they did *not* use gender field as evidence; marginalized out gender, ‘created’ dependence!

Naive Bayes (A Generative Model) as a Bayesian Network

An independence assumption in action (**likelihood**):

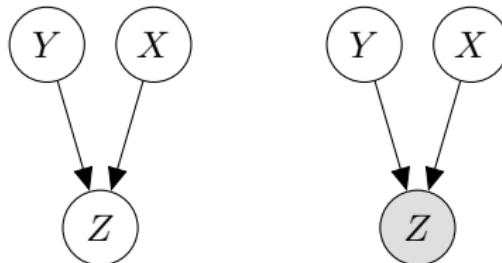


$$\begin{aligned} P(Y = y, \mathbf{X} = \mathbf{x}) &= P(Y = y) \prod_{j=1}^d P(X_j = x_j | Y = y) \\ &= P(Y = y) P(\mathbf{X} = \mathbf{x} | Y = y) \end{aligned}$$

$$\hat{y} = \operatorname{argmax}_y P(y | \mathbf{x}) = \operatorname{argmax}_y P(y, \mathbf{x})$$

Explaining Away

Recall: Two variables become dependent if a **common child** is observed

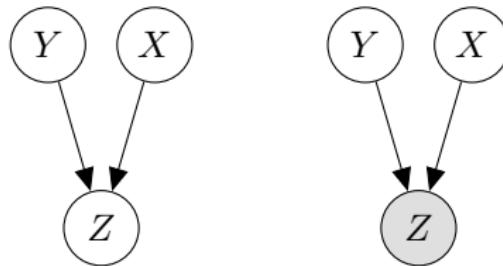


If Z is *not* observed (left),

$$\begin{aligned} P(X, Y) &= \sum_z P(X, Y, Z = z) \\ &= P(X)P(Y) \underbrace{\sum_z P(Z = z | X, Y)}_1 \\ &= P(X)P(Y) \quad \triangleright \text{independence} \end{aligned}$$

Explaining Away

Recall: Two variables become dependent if a **common child** is observed



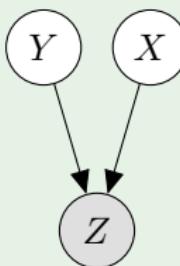
but if the common child is observed (right),

$$\begin{aligned} P(X, Y | Z = z) &= \frac{P(X, Y, Z = z)}{P(Z = z)} \quad \triangleright \text{cond. prob.} \\ &= \frac{P(Z = z | X, Y)P(X)P(Y)}{P(Z = z)} \quad \triangleright \text{explaining away} \end{aligned}$$

X and Y become dependent via the common child.

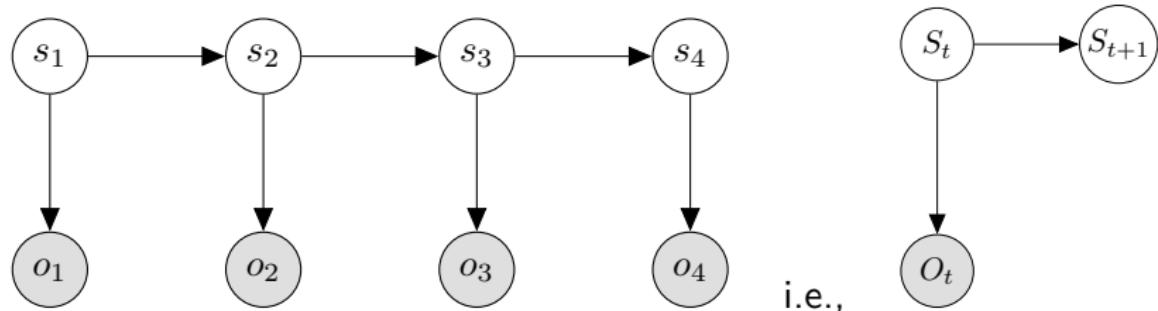
Missing Bike

My friend borrowing my bike ($X = 1$; suppose he has a key to the lock) and a thief stealing it ($Y = 1$) occur independently of each other. But if my bike is missing ($Z = 1$), then **X and Y are now dependent**: if $P(X = 1)$ is high (he suggested he would), then $P(Y = 1)$ becomes low, i.e., my friend ‘explains away’ much possibility of theft.



Can you detect the ‘explaining away’ in the Tracey’s Grass example?

Sequential Bayesian Networks (Markov Processes)



Important to recall: each variable has a distribution:

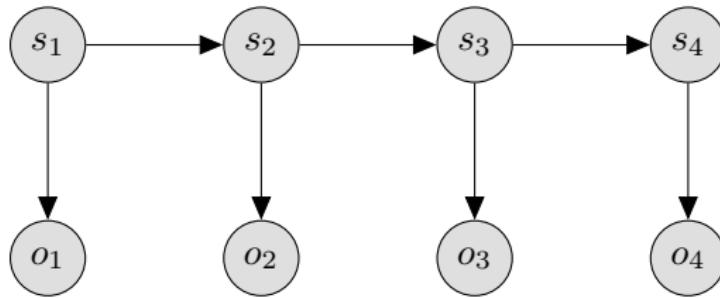
$$S_t \sim P(\cdot | S_{t-1})$$

$$O_t \sim P(\cdot | S_t)$$

We can generate a **trajectory**:

$$\underbrace{\{(S_t, O_t)\}_{t=1}^T}_{\tau} \sim P(\cdot)$$

Learning (Easy Case)



Supervised learning problem:

$$s' = h(s)$$

$$o = f(s)$$

Often unrealistic, as we would not usually have s_t !

Modern Bayesian Networks (Example: VAEs)

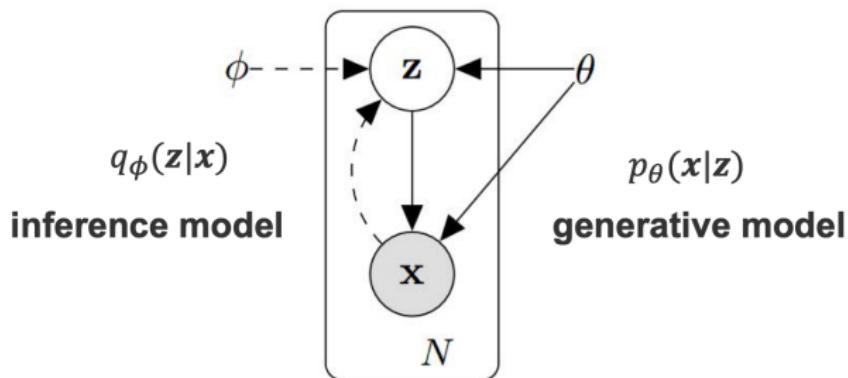


Figure courtesy: Kingma & Welling, 2014

Modern Bayesian Networks (Reinforcement Learning)

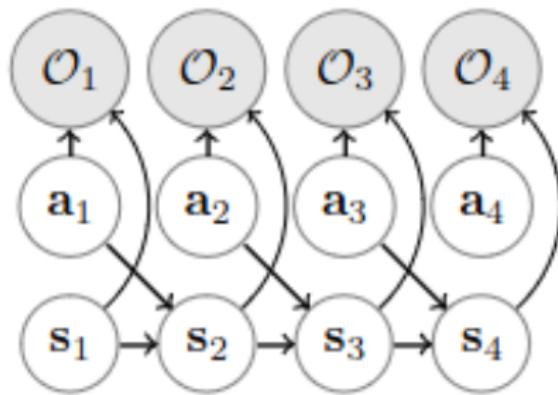


Image from: *Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review*, 2018 by Sergey Levine

Bayesian Networks: Summary

- A representation of knowledge
- A Directed Acyclic Graph, variables are nodes
- We can do reasoning (inference) across multiple variables by marginalizing out the variables we don't need.
- Based on conditional distributions, (parametrized by, e.g., θ); we should get θ from an expert or [mainly in this course] by learning from data and/or experience.

Decision Theory

- 1 Introduction
- 2 Bayesian Networks
- 3 Decision Theory
- 4 Summary

Decision Theory: From Beliefs to Actions

We have quantified our knowledge (beliefs) probabilistically via a Bayesian network (aka belief network).

Based on these beliefs, and observed evidence, we make **decisions**:

- What labels should I apply to an image?
- Should I speed?
- Should I accept a job offer?
- ...



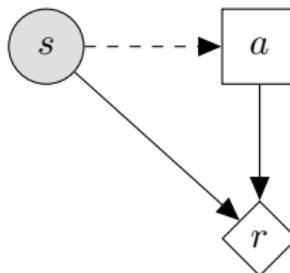
Key requirement: **loss function** (**error**, **cost**, ...; lower is better); or equivalently **reward function** (**utility**, **payoff**, ...; higher is better).

Umbrella Example: State, Reward, Decision (Action)?

- **State** space: It will rain ($s = \text{Rain}$) or not rain ($s = \neg\text{Rain}$)
- **Action** space: Take umbrella ($a = \text{Take Umbrella}$), or leave it ($a = \text{Leave Umbrella}$).
- **Reward** function $r(s, a)$:

	Rain	$\neg\text{Rain}$
Take Umbrella	10	-20
Leave Umbrella	-40	60

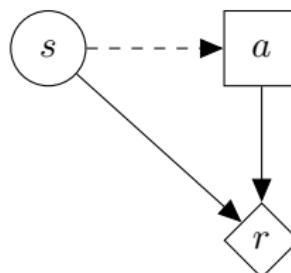
What is the **optimal decision** (*action* of the rational agent)?



When s is observed (e.g., $s = \text{Rain}$), then: $\operatorname{argmax}_a r(s, a)$.

Umbrella Example: Beliefs (Acting under Uncertainty)

What if we are **uncertain** of the state (don't know s)?



We rely on knowledge/beliefs, e.g., $P(S = \text{Rain}) = 0.4$

Again: what should a **rational agent** do (optimal decision)?

Answer: **Maximize expected reward** (i.e., **value**)

$$a^* = \operatorname{argmax}_{a \in \mathcal{A}} \underbrace{\mathbb{E}_{S \sim P(S)}[r(S, a)]}_{Q(a)}$$

Have you changed your decision (estimate of the optimal action)?

Value

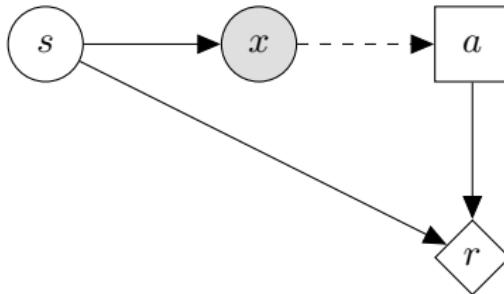
Value is an important concept. It's what drives agent learning and behaviour (choosing actions/making decisions, to maximise value).

$$Q(a) = \mathbb{E}_{S \sim P(S)}[r(S, a)] = \sum_{s \in S} r(s, a)P(s | x)$$

is the **value** of taking action a .

The agent should seek out valuable states!

Umbrella Example: Observations



In **state** s we obtain **observation** (evidence) x . We take **decision** $a(x)$, and obtain **reward** $r(s, a)$. The dashed line is the decision.

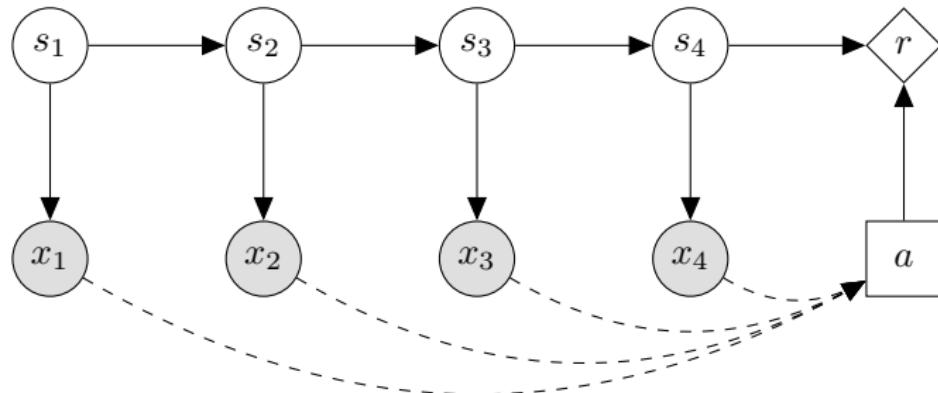
$$\mathbb{E}_{S \sim P(S|x)}[r(a, S)] = \sum_s r(a, s)p(x|s)p(s)$$

'Knowledge' based on "*Red sky at night, agent's delight. Red sky in the morning, agent's warning*":

$x = \text{red sky in the ...}$	morning	night	not at all
$p(X = x S = \text{Rain})$	0.6	0.1	0.3

You observe $x = \text{red sky in the morning}$, do you change your decision a ? Even if you don't, is the observation still useful? Hint: uncertainty.

Umbrella Example: Sequence of Observations



After 4 days inside (you have access to a window), you are going out ($t = 4$). Do you take your umbrella? What is the observation, state, here? Answered in the lab class!

Rational Agents: Summary

A **rational agent** (acts in a way to maximize its reward) considers

- ① **Actions** $\mathcal{A} = \{a_1, \dots, a_J\}$ (i.e., decisions)
- ② **States** $\mathcal{S} = \{s_1, \dots, s_K\}$
- ③ **Reward** $r(s, a)$ of taking action $a \in \mathcal{A}$ from state $s \in \mathcal{S}$.

and will model **uncertainty** (e.g., which state am I in?) via \mathbb{E} , built from P , the knowledge of the agent, under some **evidence** x .

The **value** of taking action a under observation x , is

$$Q(a) = \mathbb{E}_{S \sim P(S|x)}[r(S, a)] = \sum_{s \in \mathcal{S}} r(s, a) \cdot P(s | x) \quad (1)$$

and thus optimal action (optimizing Eq. (1)) is

$$a^* = \operatorname{argmax}_{a \in \mathcal{A}} Q(a) \quad (2)$$

What is a Reward?

A rational agent will optimize its expected reward. Thus, reward is a major consideration.



Machine learning from streaming data aims at extracting knowledge from a continuous, potentially never-ending source of data, and constructing a model that makes predictions. Examples of data streams include environmental sensing, [transactional](#), mobility or web traffic data. Learning from data streams has been an active research area over the last decades. [\[citepaper\]](#) with its fundamental concepts reaching back at least to 1990s.

This note overviews the literature



Complications arise: a good r is hard to design/obtain; **different agents can have different reward functions**, even for same state-action pair (and same environment dynamics), i.e., model reward and loss differently wrt, e.g., life, accepting a position, making a wrong classification, etc. Sometimes there are multiple agents.

Without proper knowledge, it is possible to assume that some agents do not appear to optimize their own expected reward.

Loss of Life

Program A: If adopted, exactly 400 out of 600 patients will die. Program B: If adopted, the probability that nobody will die is $1/3$, while the probability that all 600 will die is $2/3$.

Loss of Life

Program A: If adopted, exactly 400 out of 600 patients will die. Program B: If adopted, the probability that nobody will die is 1/3, while the probability that all 600 will die is 2/3.

The expected loss of life is equivalent, $\mathbb{E}[\ell(A)] = \mathbb{E}[\ell(B)]$. Yet *only 13 percent of the doctors^a* in this group chose to administer program A!

^aAccording to [5]

Doctors are not rational agents???

Discussion Points

Decision theory describes how a rational agent should act.

- Are humans rational agents?
 - “Human irrationality”
 - Phrasing questions in different ways
 - Ideology
 - Herd mentality (vs individualism)
 - Multiple agents (adversaries, collaborators)
 - Bias
 - Ego (> 50% of people believe they are more intelligent than average, better-than-average drivers, etc.)
 - People are in different states
 - People have different perception
 - People are different; act differently under uncertainty
- Is linear regression an ‘agent’? A rational agent?
- What is human $r(s, a)$ reward function?
- Is ChatGPT a rational agent?

In the context of the study of rational agents, a rational agent is typically defined as one that acts to achieve the best possible outcome or the best-expected outcome based on its knowledge and capabilities. While I am designed to process information and provide responses that are logical and aligned with the knowledge I have, my actions are determined by algorithms and are not driven by goals or desires like a traditional rational agent.

My responses are based on patterns, logic, and probabilistic reasoning, which could be seen as rational in a computational sense. However, since I don't have autonomous decision-making or goals, I don't fully fit the definition of a rational agent in the traditional sense of the study. My purpose is to assist and provide information based on the input I receive, rather than independently seeking optimal outcomes.

Alignment Issues



Conveying human objectives is a major challenge.

Summary

- 1 Introduction
- 2 Bayesian Networks
- 3 Decision Theory
- 4 Summary

Summary

- States s (and observations of states, x)
- Actions (decisions) a
- Reward function r
- The design (or choice) of these is important!
- Knowledge representation P (e.g., Bayesian Networks)
- Modeling uncertainty via expectations \mathbb{E}
- Optimal action (as taken by rational agents): the one which maximises value, $\text{argmax}_a Q(a)$

i.e., we have completed a pipeline for an intelligent agent, from **observation** (of states) to **actions** (decisions).

Focus of this lecture: Introduction **probabilistic graphical models** and **probabilistic reasoning**/inference. Should be able to provide an answer to: *What is an intelligent agent?*

What's missing (upcoming lectures): **Learning** (multi-output, structured output, reinforcement learning), and **Planning** (sequential decision making).

Probabilistic Reasoning and Decision Making

INF581 Advanced Machine Learning and Autonomous Agents

Jesse Read

