CSCI E-82a
Probabilistic Programming and AI
Lecture 5
Decision Making Under Uncertainty

Steve Elston

HARVARD
Extension School

Introduction to Decisions Under Uncertainty

In a **stochastic world** how can an agent make **optimal decisions?**

- Use **planning methods** where an Agent makes optimal decisions
  - **Optimize utility**
  - Requires a **probability model of the environment**
- **Single decision** models
  - Combine CPD table nodes with decision and utility nodes
- **Sequential decision** models
  - Use **dynamic programming**
  - Covered in another lecture

Introduction to Decisions Under Uncertainty

In a **stochastic world** how can an agent make **optimal decisions?**

- Two types of planning algorithms
- **Single decision** models
  - Combine CPD table nodes with decision and utility nodes
  - Topic for this lecture
- **Sequential decision** models
  - Use **dynamic programming**
  - Covered in another lecture

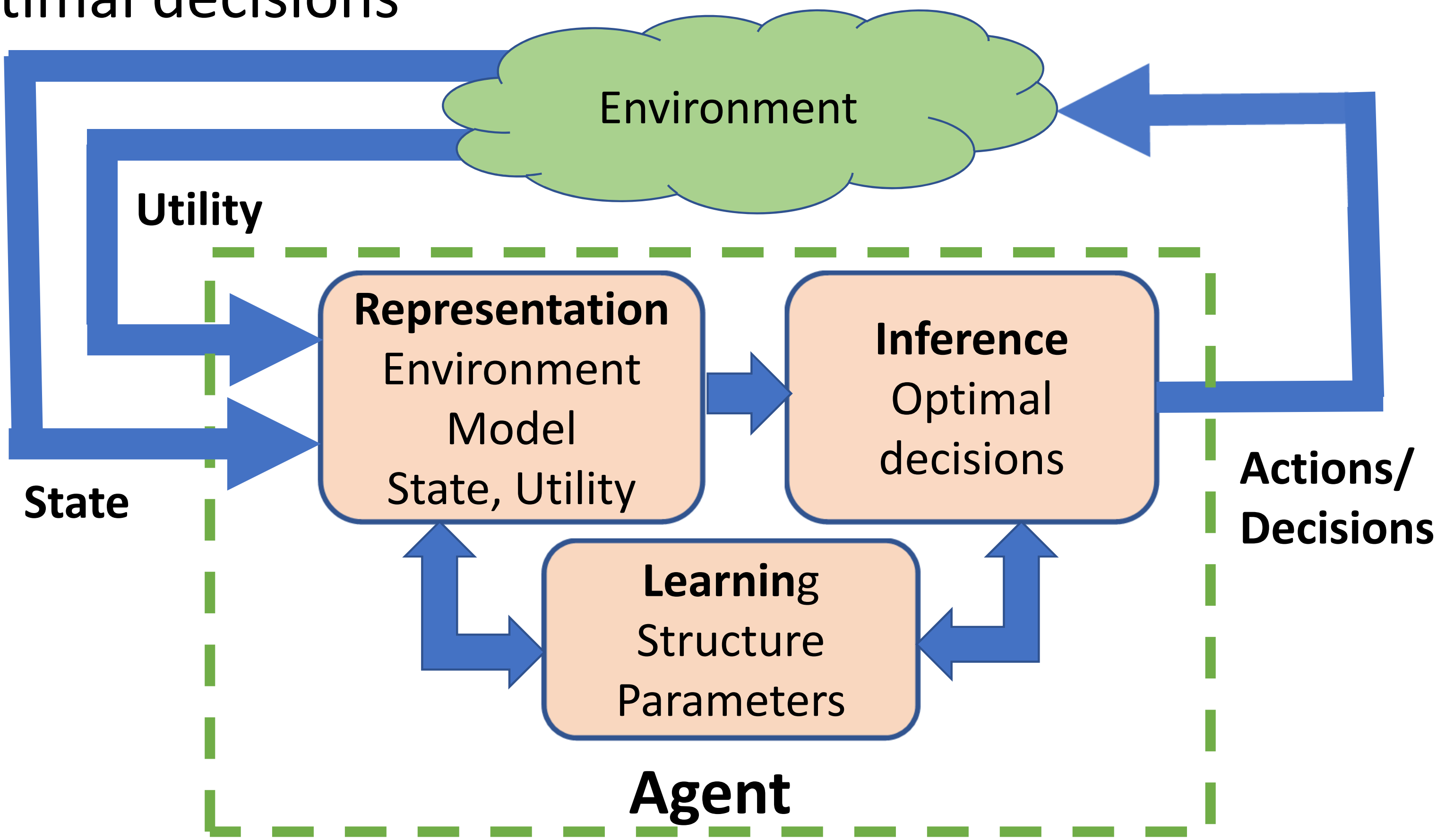
Intelligent Agents for Decisions Under Uncertainty

Intelligent agents use state and utility information to make optimal decisions

- Agent interacts with the **environment**
- Agent is not part of the environment
- Agent outputs **commands for actions** to **actuators** in the environment
- Agent receives **rewards** and **state updates** from **sensors** in the environment
- Agent maintains a **model** of the environment including **representations** of **state** and **rewards**

Intelligent Agents for Decisions Under Uncertainty

Intelligent agents use state and utility information to make optimal decisions



Intelligent Agents for Decisions Under Uncertainty

Intelligent agents use three components for a single decision model

- **Utility theory:** Allows us to **quantify the value of a system state**
- **Influence diagrams**: Extension of the **representation** we use for Bayesian graphical models
- **Inference**: to compute **optimal decisions**

Introduction to Utility Theory

A simple example of **utility**

- You go to a charity event
- There is a raffle with 100 tickets
- Each ticket is $100
- The winner will receive a $1,000 prize
- Your joy from supporting the charity is 200
- What is your utility for buying a ticket?

Introduction to Utility Theory

A simple example of **utility**

- The utility of not buying a ticket is 0
- How do we calculate the utility of buying a ticket?
- Need to **multiply the utility by the probability**

$$
\begin{aligned}
U(1) &= -p(buy) * cost + p(feeling) * value + p(win) * win\ amount \\
&= -1.0 * 100 + 1.0 * 200 + 0.01 * 1000 \ = 110
\end{aligned}
$$

- Is the **rational decision** to buy a ticket?
- **Yes!** The utility is positive

Introduction to Utility Theory

A simple example of **utility**

- Should you buy a second ticket?
- A second ticket may not double your joy
- Perhaps your joy increases to 300
- The utility for two tickets is now:

$$U(2) = -1.0 * 2 * 100 + 1.0 * 300 + 0.01 * 2 * 1000 = 120$$

- The utility increases only slightly
- Utility does **not** need to be **linear with money!**

Introduction to Utility Theory

Preferences are important in utility theory

- We can state the following relationships between preferences
- If $A \succ B$ then *A* is preferred to *B*
- If $A \sim B$ then there is indifference between *A* and *B*
- If $A \succeq B$ then *A* is preferred to *B* or there is indifference between *A* and *B*

Introduction to Utility Theory

von Neumann-Morgenstern axioms

John von Neumann and Oskar Morgenstern proposed axioms as **constraints on preferences** in the 1940s

- **Completeness:** Only one of the following can hold$A \succ B$, $B \succ A$, or $A \sim B$.
- **Transitivity:** If $A \succeq B$ and $B \geq C$, *then* $A \geq C$.
- **Continuity:** If $A \geq C \succeq B$, then for a probability *p*, $[A : p; C : 1 - p] \sim C$.
- **Independence:** If $A \succeq B$, the for any *C* and probability *p*, $[A : p; C : 1 - p] \geq [B : p; C : 1 - p]$.
- Where the notation $X : p$ means *X* with probability *p.*

Introduction to Utility Theory

How do we compute utility given probability of state?

- Given a set of states, *S*, the utility can be computed:

$$U(S) = \sum_{s} p(s)\, u(s)$$

Where,

$$p(s) = probability\ of\ state\ s$$

$$u(s) = utility\ of\ state\ s$$

Introduction to Utility Theory

How can you compute expectation for a probabilistic process?

- The expectation is the sum of the probability weighted values of the states

$$\mathbb{E}(states) = \sum_{states} p(state)\ value(state)$$

- The normalization of the expectation is correct since:

$$\sum_{states} p(state) = 1.0$$

Introduction to Utility Theory

An agent takes action to maximize utility though actions

- Given actions, *a*, and observations, *o*, the **expected utility,** for states *s'*:

$$E\big[U(a \mid o)\big] = \sum_{s'} p(s' \mid a, o)\, U(s')$$

- The agent must find the **optimal action** to maximize expected utility:

$$argmax_a E\big[U(a \mid o)\big] = argmax_a \sum_{s'} p(s' \mid a, o)\, U(s')$$

- The solution to the above equation may look straight forward
- But direct solution is not computationally feasible, except for simple decision trees

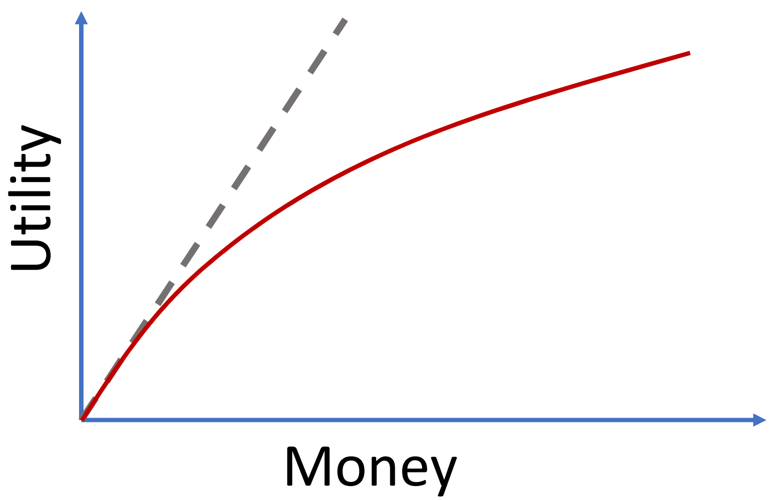
Introduction to Utility Theory

Utility leads to **rational decisions**

- People do not make rational decisions
- Say, offers you the 50% chance of winning $1M on a bet of $100,
  - The expected gain is $499,950 = -0.5 * $100 + 0.5 * $1M
  - Almost anyone will take that bet
- On the other hand, having won the $1M an offer is made to for a 50% chance win an additional $5M by betting the $1M.
  - The expected gain is $2M = -0.5 * $1M + 0.5 * $5M
  - But, many people will not take this second bet
- People exhibit **risk aversion**
  - Risk aversion is specific to an individual
- Utility is a function of risk aversion, reward, and state

Introduction to Utility Theory

Utility is **not equal to money**



- Most people will be happy to work for one hour for $1,000
- But, for a billionaire this offer might not be interesting
- Consider the utility of an incremental amount of money
- Utility could increase linearly with money
- But, the more money someone has, the less the utility of an incremental increase
- Here, amount of money available is a state variable

Influence Networks

Single decision models expressed as **influence diagrams**

- Influence diagrams have 3 components
- **CDP nodes** for modeling conditional probability distributions
  - Same as used in DAGs
- **Decision nodes** are like switches and initiate actions in the environment
  - Have no distribution
  - Can optimize utility
- **Utility nodes** measure the value of the states of the environment
  - Evaluate the utility of decisions

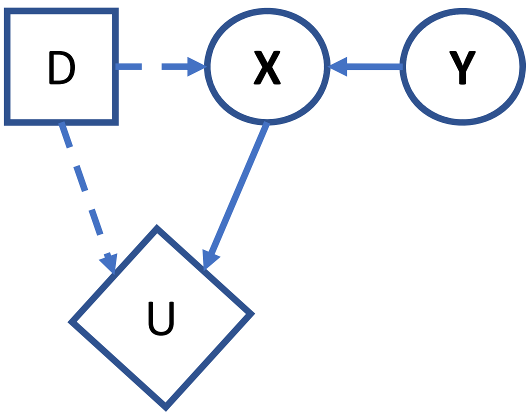
Influence Networks

Single decision models expressed as **influence diagrams**

- There are three types of **edges** in influence diagrams
- **Conditional information edges,** propagate belief, as used in DAGs
- **Informational edges** which propagate information that is not related to a distribution or belief
- **Functional edges** which end in utility nodes which propagate the information needed for the utility calculation

Influence Networks

Use a **graphical language** to express an influence diagram



- Conditional probability distribution (CPD) nodes shown as circles
- **Conditional information edges** connect CPDs
- **Decision nodes** shown as rectangles
- **Information edges** propagate non-probabilistic information – Shown as dotted lines
- **Utility nodes** shown as diamonds
- **Functional edges** propagate information required for utility calculations

Consistency and Partial Ordering

For **multiple decision problems** order matters!

- Multiple decision problems require **causal consistency**
- Some decisions must be made before there is information to make subsequent decisions
- This sequence is call **partial ordering**
- Consider a sequence of probabilistic variables, χ_i, and decisions, D_i
  - The variable χ_{i+1} will not be known until decision D_i is made
  - Using the precedence symbol, $\prec$, we can write a general expression for partial ordering as:

$$\chi_1 \prec D_1 \prec \chi_2 \prec D_2 \ldots \chi_n \prec D_n$$

- Partial ordering leads to a factorization

Influence Network Example

Planning for pizza delivery robot

- A pizza delivery robot must arrive on time at a customer site with a hot pizza
- Excessive delay leads to cold pizza and an unhappy customer
- We express the utility of on time pizza delivery as:

On Time	Slightly Late	Very Late
25.0	0.5	-25.0

Influence Network Example

Planning for pizza delivery robot

- There are two obstacles to on time pizza delivery
  - Heavy traffic on the route to the customer
  - The shortest route has a draw bridge which may be open and therefore cannot be crossed
- The Robot must make two decisions
  - Leave early which has negative utility, since the robot is unavailable for other delivery:

Leave On Time	Leave Early
0.0	-5.0

  - Take a longer route to avoid the bridge, increasing the probability of late delivery

Influence Network Example

Planning for pizza delivery robot

- Given the decision to leave early with utility, *U*1, the decision to take the shorter route over the bridge, and the utility of on time arrival, *U*2, we can express the partial ordering for this problem:

$$leave\ early \prec U1 \prec P(traffic\ delay \mid traffic, D1)$$

$$\prec take\ bridge \prec p(bridge\ delay \mid bridge\ open, D1, D2)$$

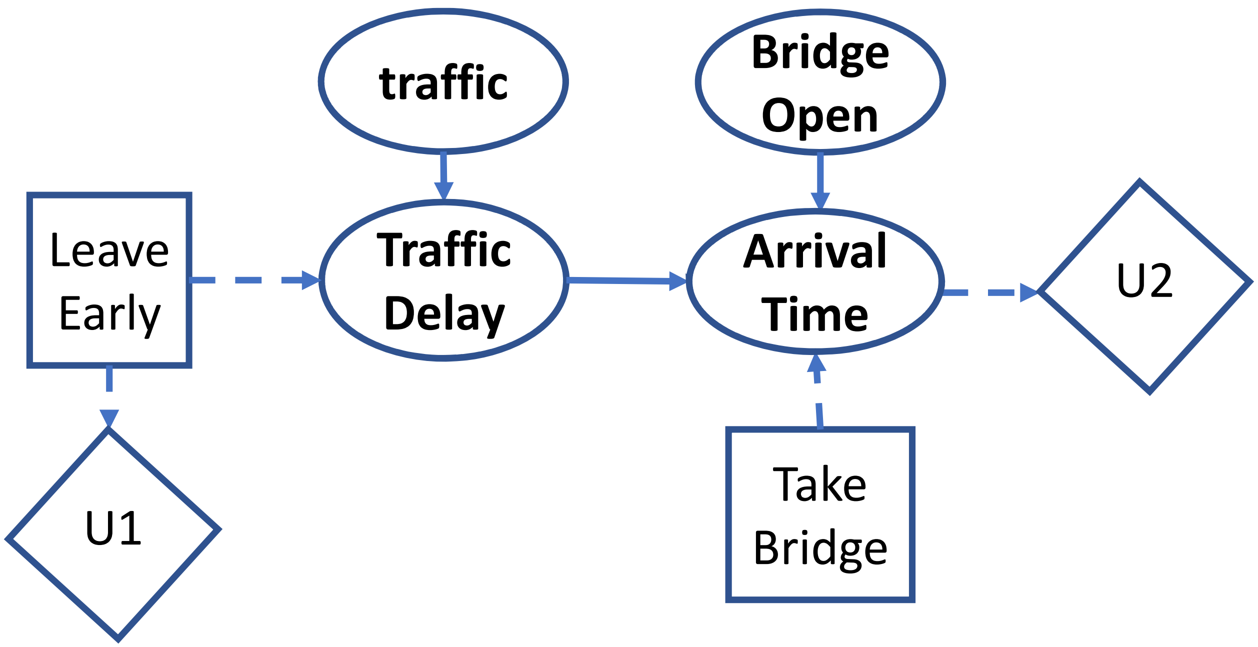
$$\prec p(arrival\ time \mid D1, D2, traffic, bridge) \prec U2$$

- Notice that the decision to take the short route over the bridge depends on how early or late the robot is at the decision point

Influence Network Example

Planning for pizza delivery robot

What does the influence diagram look like?



We can find partial ordering from influence diagram or vice versa

Inference on Influence Network

How can we perform inference on an influence Networks?

- We have already explored several inference methods which can be applied to influence networks
  - Variable elimination
  - Belief Propagation
  - Junction Tree Algorithm
- These same methods apply to influence networks

Inference on Influence Network

Variable elimination for influence network

- Given the partial ordering of a set of random variables x_i and decision variables d_i, the probability of the *Tth* state:

$$p(x_{1:T}, d_{1:T}) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, d_{1:t})$$

- Multiply by the utilities of the states, $u(x_{1:T}, d_{1:T})$, and then maximize the sum over the variables for each decision:

$$max_{d_1} \sum_{x_1} \cdots max_{d_T} \sum_{x_T} \prod_{t=1}^{T} p(x_t \mid x_{t-1}, d_{1:t})\, u(x_{1:T}, d_{1:T})$$

Inference on Influence Network

Variable elimination for influence network

- First step in variable elimination, rearrange terms:

$$max_{d_1} \sum_{x_1} \cdots max_{d_{T-1}} \sum_{x_{T-1}} \prod_{t=1}^{T-1} p(x_t \mid x_{t-1}, d_{1:t}) \; max_{d_T} \sum_{x_T} p(x_T \mid x_{1:T-1}, d_{1:T}) u(x_{1:T}, d_{1:T})$$

- Eliminating a variable produces a marginal utility $\widetilde{u}(x_{1:T-1}, d_{1:T-1})$

$$= max_{d_1} \sum_{x_1} \cdots max_{d_{T-1}} \sum_{x_{T-1}} \prod_{t=1}^{T-1} p(x_t \mid x_{t-1}, d_{1:t}) \; \widetilde{u}(x_{1:T-1}, d_{1:T-1})$$

- Factorization works since each decision is independent of subsequent decisions

Inference on Influence Network

How can we perform inference on an influence Networks?

- There is no decision variable in most software; e.g. pgmpy
- What to do?
  - Use a CPD as decision variable
  - Then use evidence to represent decision
- Can find maximum utility by setting equal probability for all options