ROB311 Quiz 2

Hanhee Lee

March 13, 2025

Contents

Probabilistic Inference Problems

1 Bayesian Networks

Definition: Vertices represent random variables and edges represent dependencies between variables.

1.1 Junction

Definition: A **junction** $\mathcal{J}$ consists of three vertices, X_1, X_2, and X_3, connected by two edges, e_1 and e_2:

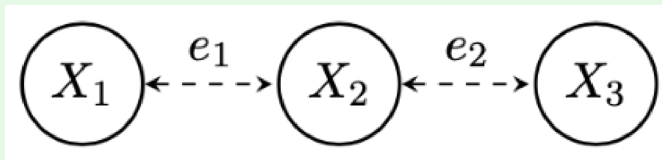


Figure 1

- X_1 and X_2 are not independent, X_2 and X_3 are not independent, but when is X_1 and X_3 independent?

1.1.1 Causal Chain

Definition: A causal chain is a junction $\mathcal{J}$ s.t.

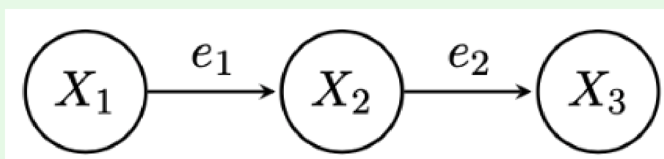


Figure 2

- X_1 and X_3 are not independent (unconditionally), but are independent given X_2.

Notes:
- **Analogy:** Given X_2, X_1 and X_3 are independent. Why? X_2's door closes when you know X_2, so X_1 and X_3 are independent.
- **Distinction b/w Causal and Dependence:** X_1 and X_2 are dependent. However, from a causal perspective, X_1 is influencing X_2 (i.e. $X_1 \rightarrow X_2$).

1.1.2 Common Cause

Definition: A common cause is a junction $\mathcal{J}$ s.t.

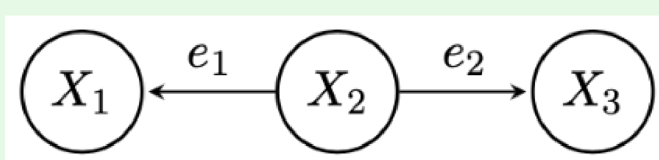


Figure 3

- X_1 and X_3 are not independent (unconditionally), but are independent given X_2.

Notes:
- **Analogy:** Given X_2, X_1 and X_3 are independent. Why? Consider the following example:
  - Let X_2 represent whether a person smokes or not, X_1 represent whether they have yellow teeth, X_3 represent whether they have lung cancer.
- Without knowing X_2, observing X_1 provides information about X_3 because yellow teeth are associated with smoking, which in turn increases the likelihood of lung cancer.
- If X_2 is known, then knowing whether a person has yellow teeth provides no additional information about whether they have lung cancer beyond what is already known from smoking status.

1.1.3 Common Effect

Definition: A common effect is a junction $\mathcal{J}$ s.t.

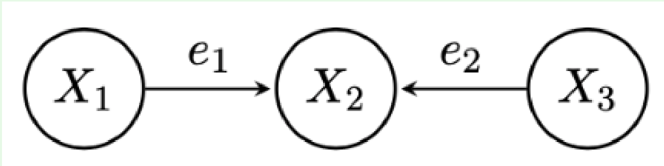


Figure 4

- X_1 and X_3 are independent (unconditionally), but are not independent given X_2 or any of X_2's descendents.

Notes:
- **Analogy:** Consider the following example:
  - Let X_2 represent whether the grass is wet, X_1 represent whether it rained, X_3 represent whether the sprinkler was on.
- Without knowing whether the grass is wet (X_2), the occurrence of rain (X_1) and the sprinkler being on (X_3) are independent events. The rain may occur regardless of the sprinkler, and vice versa.
- However, once we observe that the grass is wet (X_2), the two events become dependent:
  - If we learn that the sprinkler was not on, then the wet grass must have been caused by rain.
  - If we learn that it did not rain, then the wet grass must have been caused by the sprinkler.

2 Dependence Separation

2.1 Independence

Theorem: Any two variables, X_1 and X_2, in a Bayesian network, $\mathcal{B} = (\mathcal{V}, \mathcal{E})$, are independent given $\mathcal{K} \subseteq \mathcal{V}$ if every undirected path is blocked.

2.1.1 Blocked Undirected Path

Definition: An undirected path,

$$p = \langle (X_1, e_1, X_2), \ldots, (X_{|p|-1}, e_{|p|-1,|p|}, X_{|p|}) \rangle,$$

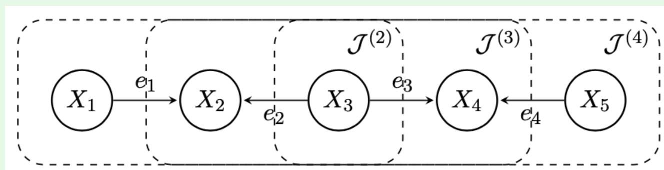


Figure 5

is **blocked** given $\mathcal{K} \subseteq \mathcal{V}$ if any of its junctions,

$$\mathcal{J}^{(n)} = \{(X_{n-1}, X_n, X_{n+1}), (e_{n-1}, e_n)\},$$

is blocked given $\mathcal{K}$.

2.1.2 Blocked Junction

Definition: $\mathcal{J} = (\{X_1, X_2, X_3\}, \{e_1, e_2\})$ is **blocked** given $\mathcal{K} \subseteq \mathcal{V}$ if X_1 and X_3 are independent given $\mathcal{K}$.

2.2 Consequence of Dependence Separation

Theorem: For any variable, $X \in \mathcal{V}$, it can be shown that X is independent of X's non-descendants, $\mathcal{V} \setminus \mathrm{des}(X)$, given X's parents, $\mathrm{pts}(X)$.

Notes:

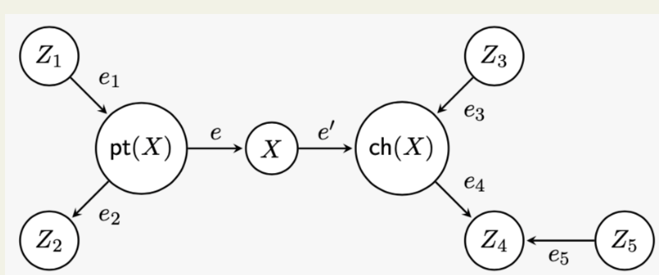


Figure 6

- Given X's parent, apply junction rules to determine that X is independent of its non-descendants.
- $\mathcal{J} = \{(Z_1, \mathrm{pt}(X), X), (e_1, e)\}$ shows that Z_1 and X are independent given $\mathrm{pt}(X)$ (causal chain).
- $\mathcal{J} = \{(Z_2, \mathrm{pt}(X), X), (e_2, e)\}$ shows that Z_2 and X are independent given $\mathrm{pt}(X)$ (common cause).
- Given $\mathrm{ch}(X)$'s parent, apply junction rules to determine that $\mathrm{ch}(X)$ is independent of its non-descendants.
- $\mathcal{J} = \{\mathrm{pt}(X), X, \mathrm{ch}(X)), (e, e')\}$ shows that $\mathrm{pt}(X)$ and $\mathrm{ch}(X)$ are independent given X (causal chain).

- Given Z_4's parent, apply junction rules to determine that Z_4 is independent of its non-descendants.
- $\mathcal{J} = \{X, \text{ch}(X), Z_4, (e', e_4)\}$ shows that X and Z_4 are independent given $\text{ch}(X)$ (causal chain).
- CHECK THIS OVER AGAIN WITH THE PROFESSOR.

2.3 Canonical Problems

2.3.1 Undirected Path Blocked?

Process:

1. **Given:** Undirected path p and $\mathcal{K}$
2. Check if any of the junctions on the undirected path are blocked given $\mathcal{K}$.
   - i.e. Check if X_1 and X_3 of the junction are independent given $\mathcal{K}$.

2.3.2 Independence

Process:

1. Given a Bayesian network w/ 2 variables to find independence.
2. Find all undirected paths between the 2 variables in the Bayesian network.
3. Identify a set of variables, $\mathcal{K}$, that block at least one junction in all undirected paths.
   - Test a junction by seeing junction given relationships.
4. If all undirected paths are blocked, then the 2 variables are independent given $\mathcal{K}$.

Warning:

- Be careful of common effect, in which it is blocked by default.

Example:

1. **Given:** Bayesian network.

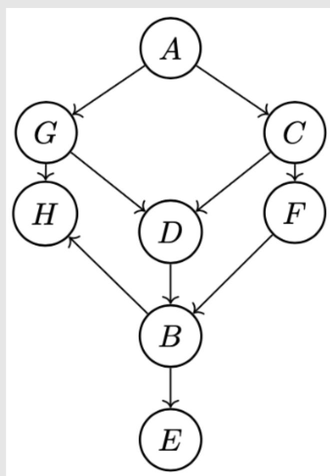


Figure 7

2. **Problem:** A and E are
   - independent if $\mathcal{K} =$
   - not necessarily independent for $\mathcal{K} =$
3. **Soln:**
   (a) **Undirected Paths:**
   - $A \to G \to H \to B \to E$
   - $A \to G \to D \to B \to E$
   - $A \to C \to F \to B \to E$
   - $A \to C \to D \to B \to E$

Example: Independent:

$\mathcal{K}$
$\{G, C\}$

- $A \iff G \iff H \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(A, G, H), (e_1, e_2)\}$ is blocked given G since A, H independent given G (causal chain)
- $A \iff G \iff D \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(A, G, D), (e_1, e_2)\}$ is blocked given G since A, D independent given G (causal chain)
- $A \iff C \iff F \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(A, C, F), (e_1, e_2)\}$ is blocked given C since A, F independent given C (causal chain)
- $A \iff C \iff D \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(A, C, D), (e_1, e_2)\}$ is blocked given C since A, D independent given C (causal chain)

$\{D, F\}$

- $A \iff G \iff H \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(G, H, B), (e_1, e_2)\}$ is blocked NOT given H since G, B independent NOT given H (common effect)
- $A \iff G \iff D \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(G, D, B), (e_1, e_2)\}$ is blocked given D since G, B independent given D (causal chain)
- $A \iff C \iff F \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(C, F, B), (e_1, e_2)\}$ is blocked given F since C, B independent given F (causal chain)
- $A \iff C \iff D \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(C, D, B), (e_1, e_2)\}$ is blocked given D since C, B independent given D (causal chain)

Not Necessarily Independent:

$\mathcal{K}$
$\{H, D, F\}$

- $A \iff G \iff H \iff B \iff E$ is unblocked given $\mathcal{K}$ since $\mathcal{J} = \{(G, H, B), (e_1, e_2)\}$ is unblocked given H since G, B not independent given H (common effect)
- $A \iff G \iff D \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(G, D, B), (e_1, e_2)\}$ is blocked given D (causal chain) since G, B independent given D (causal chain)
- $A \iff C \iff F \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(C, F, B), (e_1, e_2)\}$ is blocked given F since C, B independent given F (causal chain)
- $A \iff C \iff D \iff B \iff E$ is blocked given $\mathcal{K}$ since $\mathcal{J} = \{(C, D, B), (e_1, e_2)\}$ is blocked given D since C, B independent given D (causal chain)

Example: Determine all subsets of $\{B, C, D, F, G, H\}$ for which A and E are independent.

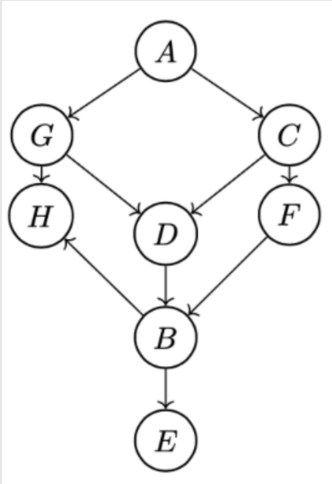


Figure 8

1. **Undirected Paths:**
    - $A \to G \to H \to B \to E$
    - $A \to G \to D \to B \to E$
    - $A \to C \to F \to B \to E$
    - $A \to C \to D \to B \to E$

$\mathcal{K}$

$\{B\}$ (Any subset that contains B will be independent)

- $AGHBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(H, B, E), (e_1, e_2)\}$ is b since H, E indep. given B (causal chain)
- $AGDBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(D, B, E), (e_1, e_2)\}$ is b since D, E indep. given B (causal chain)
- $ACFBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(F, B, E), (e_1, e_2)\}$ is b since F, E indep. given B (causal chain)
- $ACDBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(D, B, E), (e_1, e_2)\}$ is b since D, E indep. given B (causal chain)

$\{C\}$ (Not independent)

- $AGDBE$ is ub given $\mathcal{K}$ since $\forall \mathcal{J}$ on p, all are ub.

$\{D\}$ (Not indepedent)

- $ACFBE$ is ub given $\mathcal{K}$ since $\forall \mathcal{J}$ on p, all are ub.

$\{F\}$ (Not independent)

- $AGDBE$ is ub given $\mathcal{K}$ since $\forall \mathcal{J}$ on p, all are ub.

$\{G\}$ (Not independent)

- $ACFBE$ is ub given $\mathcal{K}$ since $\forall \mathcal{J}$ on p, all are ub.

$\{H\}$ (Not independent)

- $ACFBE$ is ub given $\mathcal{K}$ since $\forall \mathcal{J}$ on p, all are ub.

Example:

$\mathcal{K}$
$\{C, D\}$ (Any subset that contains C and D except H will be independent)
...

- $AGHBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(G, H, B), (e_1, e_2)\}$ is b since G, B indep. not given H (causal effect)
- $AGDBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(G, D, B), (e_1, e_2)\}$ is b since G, B indep. given D (causal chain)
- $ACFBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(A, C, F), (e_1, e_2)\}$ is b since A, F indep. given C (causal chain)
- $ACDBE$ is b given $\mathcal{K}$ since $\mathcal{J} = \{(A, C, D), (e_1, e_2)\}$ is b since A, D indep. given C (causal chain)

Example: Any set with

$$B \vee [(G \vee (D \wedge (\neg H))) \wedge (C \vee F)]$$

Example:

3 Probabilistic Inference

3.1 Problem Setup

Definition: Given a Bayesian network, $\mathcal{B} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{X_1, \dots, X_{|\mathcal{V}|}\}$, we want to find the value of:

$$\text{pr}(\mathbf{Q} \mid \mathbf{E}) := \text{pr}(Q_1, \dots, Q_{|\mathbf{Q}|} \mid E_1, \dots, E_{|\mathbf{E}|}) = \frac{\sum_{\mathcal{V}\setminus(\mathbf{Q}\cup\mathbf{E})} p(X_1, \dots, X_{|\mathcal{V}|})}{\sum_{\mathcal{V}\setminus\mathbf{E}} p(X_1, \dots, X_{|\mathcal{V}|})}$$

$$\text{pr}(\mathbf{Q} \mid \mathbf{E}) \propto \sum_{\mathcal{V}\setminus(\mathbf{Q}\cup\mathbf{E})} \left(p(X_1) \prod_{i\neq 1} p(X_i \mid \text{pts}(X_i)) \right)$$

- $\mathbf{Q} = \{Q_1, \dots, Q_{|\mathbf{Q}|}\}$: Query variables
- $\mathbf{E} = \{E_1, \dots, E_{|\mathbf{E}|}\} \subseteq \mathcal{V}$: Evidence variables
- $\mathbf{Q} \cap \mathbf{E} = \emptyset$.

Warning:
- Denominator: Normalization constant (assuming $\mathbf{E}$ is fixed)
- Therefore, only need to compute numerator (w/o specifying $\mathbf{Q}$), which we can then normalize w.r.t. $\mathbf{Q}$

3.1.1 Joint Distribution in a Bayesian Network

Derivation: For any joint distribution, the following factorization holds:

$$p(X_1, \dots, X_{|p|}) = p(X_1) \prod_{i\neq 1} p(X_i \mid X_1, \dots, X_{i-1})$$

Bayesian Network Conditions: If
- at least 1 variable will be an orphan (i.e. no parents)
- no variable is both ancestor and descendant of another.

then this allows us to order $X_1, \dots, X_{|\mathcal{V}|}$, so that if X_j is a descendent of X_i, then for any $j > i$,

$$\text{pts}(X_i) \subseteq \{X_1, \dots, X_{i-1}\} \text{ and } X_1, \dots, X_{i-1} \notin \text{des}(X_i)$$

Therefore, using the consequence of dependence separation, then

$$p(X_1, \dots, X_{|\mathcal{V}|}) = p(X_1) \prod_{i\neq 1} p(X_i \mid \text{pts}(X_i))$$

3.2 Method 1: Bayesian Network Inference

3.2.1 Markov Blanket

Definition: The **Markov blanket** of a variable X, denoted $\text{mbk}(X)$, consists of the following variables:
- X's children
- X's parents
- The other parents of X's children, excluding X itself.

which is when a variable, X, is "eliminated", the resulting factor's scope is the Markov blanket of X.

3.2.2 Graphical Interpretation

Notes: Pictorially, eliminating X is equivalent to replacing all hyper-edges that include X with their union minus X, and then removing X.

3.2.3 Elimination Ordering

Definition: The order that the variables are eliminated.
- This creates a sequence of hyper-graphs that depend on the elimination ordering.

3.2.4 Elimination Width

Definition: The **elimination width** of a sequence of hyper-graphs is the # of variables in the hyper-edge within the sequence with the most variables.

3.2.5 Heuristics for Elimination Ordering

Definition: Choose the elimination ordering to minimize the elimination width using the following heuristics:
1. Eliminate variable with the fewest parents.
2. Eliminate variable with the smallest domain for its parents, where

$$|\operatorname{dom}(\operatorname{pts}(X))| = \prod_{Z \in \operatorname{pnt}(X)} |\operatorname{dom}(Z)|.$$

3. Eliminate variable with the smallest Markov blanket.
4. Eliminate variable with the smallest domain for its Markov blanket, where

$$|\operatorname{dom}(\operatorname{mbk}(X))| = \prod_{Z \in \operatorname{embk}(X)} |\operatorname{dom}(Z)|.$$

Warning: Choosing the variable with the smallest domain for its Markov blanket is the most effective heuristic.

3.3 Method 2: Inference via Sampling

Definition: Generate a large # of samples and then approximate as:

$$p(\mathbf{Q} \mid \mathbf{E}) \approx \frac{\text{\# of samples w/ } \mathbf{Q} \text{ and } \mathbf{E}}{\text{\# of samples w/ } \mathbf{E}}.$$

- As # of samples $\to \infty$, the approximation becomes exact.

3.3.1 Inference via Sampling with Likelihood Weighting

Motivation: Most of the samples are wasted since they are not consistent with the evidence.

Definition: Generate a large # of samples and then approximate as:

$$p(\mathbf{Q} \mid \mathbf{E}) \approx \frac{\text{weight of samples w/ } \mathbf{Q} \text{ and } \mathbf{E}}{\text{weight of samples w/ } \mathbf{E}}.$$

- Weight for each sample: Probability of forcing the evidence, i.e. probability of the evidence given the sample.

3.4 Canonical Problems:

Example:

1. **Given:** Caveman is deciding whether to go hunt for meat. He must take into account several factors:
   - Weather
   - Possibility of over-exertion
   - Possibility encountering lion

   These factors can result in Cavemen's death. His decision will ultimately depend on the **chances** of his death.
2. **Binary Variables:**
   - $W = \{\text{Sun}, \text{Rainy}\}$: Weather
   - H: Whether the Cavemen goes hunting or not.
   - L: Whether the Cavemen encounters a lion or not.
   - T: Whether the Cavement is tired or not.
   - D: Whether the Cavemen dies or not
3. **Problem:** Cavemen must decide whether to go hunting or not.
   - He must consider the conditional probabilities (i.e. dependence) of each event.

Warning: Have to be discrete.

3.4.1 Bayesian Inference via Variable Elimination

Process:

1. Given Bayesian network w/ variables and their conditional probabilities.
2. Find the probability of the query variable given the evidence variable, $p(\mathbf{Q} \mid \mathbf{E})$.
3. Use $p(\mathbf{Q} \mid \mathbf{E}) = \dfrac{\sum_{\mathcal{V}\setminus(\mathbf{Q}\cup\mathbf{E})} p(X_1, \ldots, X_{|\mathcal{V}|})}{\sum_{\mathcal{V}\setminus\mathbf{E}} p(X_1, \ldots, X_{|\mathcal{V}|})}$.
4. Determine $p(X_1)\prod_{i\neq 1} p(X_i \mid \text{pts}(X_i))$ using the Bayesian network.
5. Write out the summation of the numerator in an order using heuristics to determine elimination ordering.
6. Start with inner summation and work outwards.
7. Calculate the probability of the query variable(s) given the evidence variable(s).

Example:

1. **Given:**

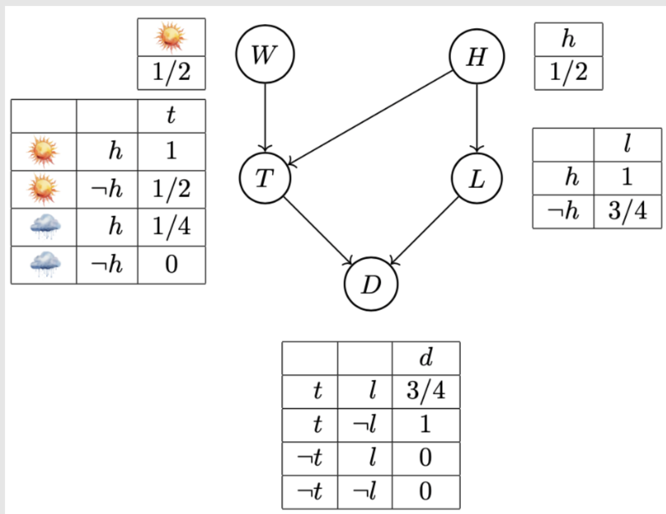


Figure 9

Variables	Values
W	$P(\text{Sunny}) = 0.5 \mid P(\text{Rainy}) = 0.5$
H	$P(h) = 0.5 \mid P(\neg h) = 0.5$
T	$P(t \mid \text{Sunny}, h) = 1 \mid P(t \mid \text{Sunny}, \neg h) = 0.5 \mid P(t \mid \text{Rainy}, h) = 0.25 \mid P(t \mid \text{Rainy}, \neg h) = 0$ $P(\neg t \mid \text{Sunny}, h) = 0 \mid P(\neg t \mid \text{Sunny}, \neg h) = 0.5 \mid P(\neg t \mid \text{Rainy}, h) = 0.75 \mid P(\neg t \mid \text{Rainy}, \neg h) = 1$
L	$P(l \mid h) = 1 \mid P(l \mid \neg h) = 0.75$ $P(\neg l \mid h) = 0 \mid P(\neg l \mid \neg h) = 0.25$
D	$P(d \mid t, l) = 0.75 \mid P(d \mid t, \neg l) = 1 \mid P(d \mid \neg t, l) = 0 \mid P(d \mid \neg t, \neg l) = 0$ $P(\neg d \mid t, l) = 0.25 \mid P(\neg d \mid t, \neg l) = 0 \mid P(\neg d \mid \neg t, l) = 1 \mid P(\neg d \mid \neg t, \neg l) = 1$

2. **Problem:** $p(d \mid h)$?
3. **Soln:**
   (a) $p(d \mid h) = \dfrac{p(d, h)}{p(h)} = \dfrac{\sum_{W,T,L} p(W, h, T, L, d)}{\sum_{W,T,L,D} p(W, h, T, L, d)}$ by definition of query and evidence equations.
   (b) $p(W, h, T, L, D) = p(h)p(W)p(L \mid h)p(t \mid W, h)p(D \mid T, L)$ by Bayesian network and $p(X_1, \ldots, X_{|\mathcal{V}|}) = p(X_1)\prod_{i\neq 1} p(X_i \mid \text{pts}(X_i))$.

Summation

$$\text{Numerator} : \underbrace{p(h)\sum_{L} p(L \mid h) \underbrace{\sum_{T} p(D \mid T, L) \underbrace{\sum_{W} p(W)p(T \mid W, h)}_{g_1(T)}}_{g_2(L,D)}}_{g_3(D)}$$

$$g_1(T) = p(\text{Sunny})p(T \mid \text{Sunny}, h) + p(\text{Rainy})p(T \mid \text{Rainy}, h)$$

$g_1(t) = p(\text{Sunny})p(t \mid \text{Sunny}, h) + p(\text{Rainy})p(t \mid \text{Rainy}, h) = 0.5 \cdot 1 + 0.5 \cdot 0.25 = 0.625$
$g_1(\neg t) = p(\text{Sunny})p(\neg t \mid \text{Sunny}, h) + p(\text{Rainy})p(\not t \mid \text{Rainy}, h) = 0.5 \cdot 0 + 0.5 \cdot 0.75 = 0.375$

$$g_2(L, D) = p(D \mid t, L)g_1(t) + p(D \mid \neg t, L)g_1(\neg t)$$

$g_2(l, d) = p(d \mid t, l)g_1(t) + p(d \mid \neg t, l)g_1(\neg t) = 0.75 \cdot 0.625 + 0 \cdot 0.375 = 0.46875$
$g_2(l, \neg d) = p(\neg d \mid t, l)g_1(t) + p(\neg d \mid \neg t, l)g_1(\neg t) = 0.25 \cdot 0.625 + 1 \cdot 0.375 = 0.53125$
$g_2(\neg l, d) = p(d \mid t, \neg l)g_1(t) + p(d \mid \neg t, \neg l)g_1(\neg t) = 1 \cdot 0.625 + 0 \cdot 0.375 = 0.625$
$g_2(\neg l, \neg d) = p(\neg d \mid t, \neg l)g_1(t) + p(\neg d \mid \neg t, \neg l)g_1(\neg t) = 0 \cdot 0.625 + 1 \cdot 0.375 = 0.375$

$$g_3(D) = p(h)p(l \mid h)g_2(l, D) + p(h)p(\neg l \mid h)g_2(\neg l, D)$$

$g_3(d) = p(h)p(l \mid h)g_2(l, d) + p(h)p(\neg l \mid h)g_2(\neg l, d) = (0.5)(1)(0.46875) + (0.5)(0)(0.625) = 0.234375$
$g_3(\neg d) = p(h)p(l \mid h)g_2(l, \neg d) + p(h)p(\neg l \mid h)g_2(\neg l, \neg d) = (0.5)(1)(0.53125) + (0.5)(0)(0.375) = 0.265625$

$$p(d \mid h) = \frac{g_3(d)}{g_3(d) + g_3(\neg d)} = \frac{0.234375}{0.234375 + 0.265625} = \frac{0.234375}{0.5} = 0.46875$$

Example:

Summation

$$\text{Numerator} : \underbrace{p(h)\sum_L p(L \mid h)\underbrace{\sum_W p(W)\underbrace{\sum_T p(T \mid W,h)p(D \mid T,L)}_{g_1(W,D,L)}}_{g_2(D,L)}}_{g_3(D)}$$

$$\text{Numerator} : \underbrace{p(h)\sum_W p(W)\underbrace{\sum_T p(T \mid W,h)\underbrace{\sum_L p(L \mid h)p(D \mid T,L)}_{g_1(D,T)}}_{g_2(D,W)}}_{g_3(D)}$$

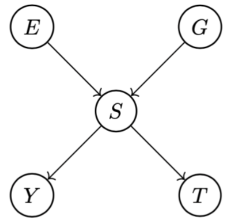
$$\text{Numerator} : \underbrace{p(h)\sum_W p(W)\underbrace{\sum_L p(L \mid h)\underbrace{\sum_T p(T \mid W,h)p(D \mid T,L)}_{g_1(W,D,L)}}_{g_2(W,D)}}_{g_3(D)}$$

$$\text{Numerator} : \underbrace{p(h)\sum_T p(T \mid W,h)\underbrace{\sum_W p(W)\underbrace{\sum_L p(L \mid h)p(D \mid T,L)}_{g_1(D,T)}}_{g_2(D,T)}}_{g_3(D)}$$

...

Example:

1. **Given:**



G	E	pr$[S=1 \mid G,E]$
1	1	0.7
1	0	0.5
0	1	0.4
0	0	0.1

S	pr$[T=1 \mid S]$
1	0.9
0	0.2

S	pr$[Y=1 \mid S]$
1	0.8
0	0.1

Figure 10

2. **Problem:** Compute $\Pr(s=1 \mid t=1)$ if $\Pr(G=1)=0.3$, $\Pr(E=1)=0.4$, and the conditional probability tables for S, Y, and T are given below.
3. **Solution:**
   (a) $p(s=1 \mid t=1) = \dfrac{p(s=1,t=1)}{p(t=1)} = \dfrac{\sum_{E,G,Y} p(E,G,Y,s=1,t=1)}{\sum_S p(t=1,S)}$
   (b) $p(E,G,Y,s=1,t=1) = p(G)p(E)p(s=1 \mid G,E)p(t=1 \mid s=1)p(Y \mid s=1)$
      - Conditional probability and individual probabilities come from Bayesian network, and set $t, s=1$ due to the query and evidence variables.

Summation

$$\text{Numerator}: \underbrace{p(t=1 \mid s=1)\sum_E p(E)\underbrace{\sum_G p(G)p(s=1 \mid G,E)\underbrace{\sum_Y p(Y \mid s=1)}_{g_1}}_{g_2}}_{g_3}$$

$$g_1 = p(Y=1 \mid S=1) + p(Y=0 \mid S=1) = 0.9 + 0.1 = 1$$

$$g_2(E) = (p(g=1)p(s=1 \mid g=1,E) + p(g=0)p(s=1 \mid g=0,E))g_1$$

$g_2(e=1,s=1) = 0.3(0.7) + 0.7(0.4) = 0.49$
$g_2(e=0,s=1) = 0.3(0.5) + 0.7(0.1) = 0.22$
$g_2(e=1,s=0) = 0.3(0.3) + 0.7(0.6) = 0.51$
$g_2(e=0,s=0) = 0.3(0.5) + 0.7(0.9) = 0.78$

- $g_3(t=1 \mid s=1) = 0.9p(e=1)g_2(e=1) + 0.9p(e=0)g_2(e=0) = 0.9(0.4)(0.49) + 0.9(0.6)(0.22) = 0.2952$
- $g_3(t=1 \mid s=0) = 0.2p(e=1)g_2(e=1) + 0.2p(e=0)g_2(e=0) = 0.2(0.4)(0.51) + 0.2(0.6)(0.78) = 0.1344$

$$p(s=1 \mid t=1) = \frac{g_3}{\sum_S p(t=1,S)} = \frac{0.2952}{0.2952+0.1344} = \frac{0.2952}{0.4296} = 0.6875$$

Example:

1. **Given:** Consider the following Bayesian network, where A, B, C, D are binary R.V. over $\{0, 1\}$

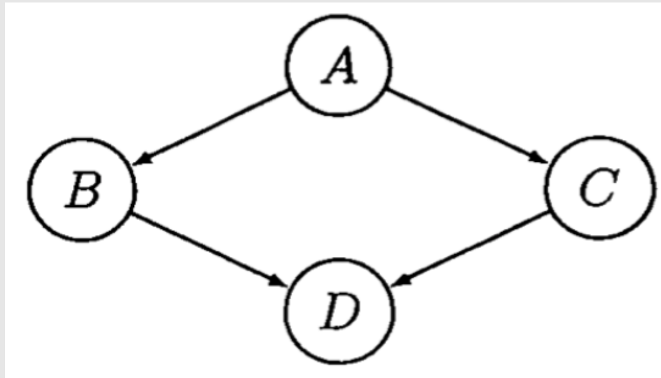


Figure 11

$P(A=1)$	P(A=0)
0.1	0.9

A	$P(B=1 \mid A)$	P(B=0|A)
0	0.4	0.6
1	0.5	0.5

A	$P(C=1 \mid A)$	P(C=0|A)
0	0.2	0.8
1	0.4	0.6

B	C	$P(D=1 \mid B, C)$	P(D=0|B,C)
0	0	0.5	0.5
0	1	0.2	0.8
1	0	0.5	0.5
1	1	0.4	0.6

Figure 12

2. **Problem:** Find $P(A = 0 \mid C = 0)$ and $P(D = 1 \mid C = 0)$.
3. **Solution:**
   (a) Derivation of $P(D = 1 \mid C = 0)$:

$$
\begin{aligned}
P(D = 1 \mid C = 0) &= \frac{P(D = 1, C = 0)}{P(C = 0)} \quad \text{by definition} \\
&= \frac{P(D = 1, C = 0)}{\sum_d P(D = d, C = 0)} \quad \text{marginalize over } D \\
&= \frac{\sum_{A,B} P(A, B, C = 0, D = 1)}{\sum_d \sum_{A,B} P(A, B, C = 0, D = d)} \quad \text{equation in problem setup}
\end{aligned}
$$

   - Summing over the variables that are not in the query and evidence variables.

   (b) Summation Term:

$$
\begin{aligned}
&\sum_{A,B} P(A)P(B \mid A)P(C = 0 \mid A)P(D = d \mid B, C = 0) \quad \text{Bayesian network} \\
&\sum_{A} P(A)P(C = 0 \mid A) \sum_{B} P(B \mid A)P(D = d \mid B, C = 0) \quad \text{(1st ordering)} \\
&\sum_{B} P(D = d \mid B, C = 0) \sum_{A} P(A)P(B \mid A)P(C = 0 \mid A) \quad \text{(2nd ordering)}
\end{aligned}
$$

(c) Choose:

$$\underbrace{\sum_B P(D = d \mid B, C = 0) \underbrace{\sum_A P(A)P(B \mid A)P(C = 0 \mid A)}_{g_1(B)}}_{g_2(d)}$$

(d) $g_1(B)$:

$$\begin{aligned} g_1(B) &= P(A = 0)P(B \mid A = 0)P(C = 0 \mid A = 0) + P(A = 1)P(B \mid A = 1)P(C = 0 \mid A = 1) \\ &= \begin{cases} 0.9(0.6)(0.8) + 0.1(0.5)(0.6) & \text{if } B = 0 \\ 0.9(0.4)(0.8) + 0.1(0.5)(0.6) & \text{if } B = 1 \end{cases} \\ &= \begin{cases} 0.462 & \text{if } B = 0 \\ 0.318 & \text{if } B = 1 \end{cases} \end{aligned}$$

(e) $g_2(d)$:

$$\begin{aligned} g_2(d) &= P(D = d \mid B = 0, C = 0)g_1(B = 0) + P(D = d \mid B = 1, C = 0)g_1(B = 1) \\ &= \begin{cases} 0.5(0.462) + 0.5(0.318) & \text{if } d = 0 \\ 0.5(0.462) + 0.5(0.318) & \text{if } d = 1 \end{cases} \\ &= \begin{cases} 0.39 & \text{if } d = 0 \\ 0.39 & \text{if } d = 1 \end{cases} \end{aligned}$$

(f) $P(D = 1 \mid C = 0) = \dfrac{g_2(1)}{g_2(0) + g_2(1)} = \dfrac{0.39}{0.39 + 0.39} = 0.5$

4. **Solution 2:**

(a) Derivation of $P(A = 0 \mid C = 0)$:

$$\begin{aligned} P(A = 0 \mid C = 0) &= \frac{P(A = 0, C = 0)}{P(C = 0)} \\ &= \frac{P(A = 0, C = 0)}{\sum_a P(A = a, C = 0)} \\ &= \frac{\sum_{B,D} P(A = 0, B, C = 0, D)}{\sum_a \sum_{B,D} P(A = a, B, C = 0, D)} \end{aligned}$$

(b) Summation Term:

$$\sum_{B,D} P(A = a)P(B \mid A = a)P(C = 0 \mid A = a)P(D \mid B, C = 0) \quad \text{Bayesian network}$$

$$P(C = 0 \mid A = a) \sum_B P(B \mid A = a)P(A = a \mid B, C = 0) \sum_D P(D \mid B, C = 0) \quad \text{(1st ordering)}$$

$$P(C = 0 \mid A = a) \sum_D P(D \mid B, C = 0) \sum_B P(B \mid A = a)P(A = a \mid B, C = 0) \quad \text{(2nd ordering)}$$

(c) Choose:

$$P(C = 0 \mid A = a) \underbrace{\sum_B P(B \mid A = a)P(A = a \mid B, C = 0) \underbrace{\sum_D P(D \mid B, C = 0)}_{g_1(B)}}_{g_2(A)}$$

(d) Same as before.

Warning:
- Write the complement probability to make life easier.
- To determine the conditional probability summation of a variable, look at its parents (inward arrows)
- Inner sum must have all probabilities with that variable in it that you are summing over.

3.4.2 Hypergraph

Process: Process of eliminating a variable.

1. Create a Hyper-graph by creating a node for each variable.
2. Create hyper-edges (factors) by circling the nodes based on of its parents (i.e. arrows pointing into a variable). If no parents, circle itself.
3. Select a variable v that we are summing over.
   (a) Circle all the variables that have v in their hyperedge into one big hyperedge (i.e. union of hyper-edges).
   (b) Eliminate v by removing the node.
   (c) Calculate the factor by multiplying the support of the variables in the union of hyperedges.
4. Repeat the process for all other v.
5. Select the smallest factor to eliminate first.
6. Repeat until all variables are eliminated to determine the best ordering of elimination.
   - The first eliminated variable will be the inner sum.

Example:

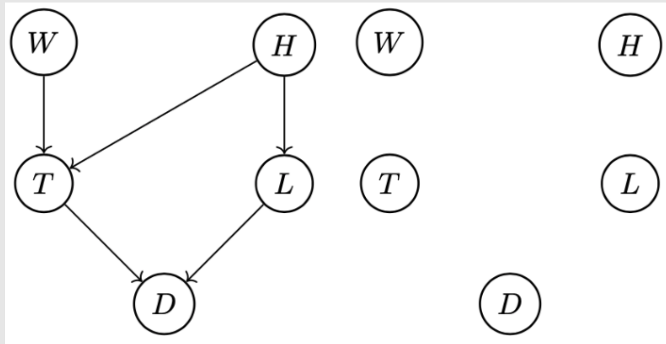


Figure 13

- Since these are all binary variables, we are selecting the factor with the least number of variables to eliminate first.

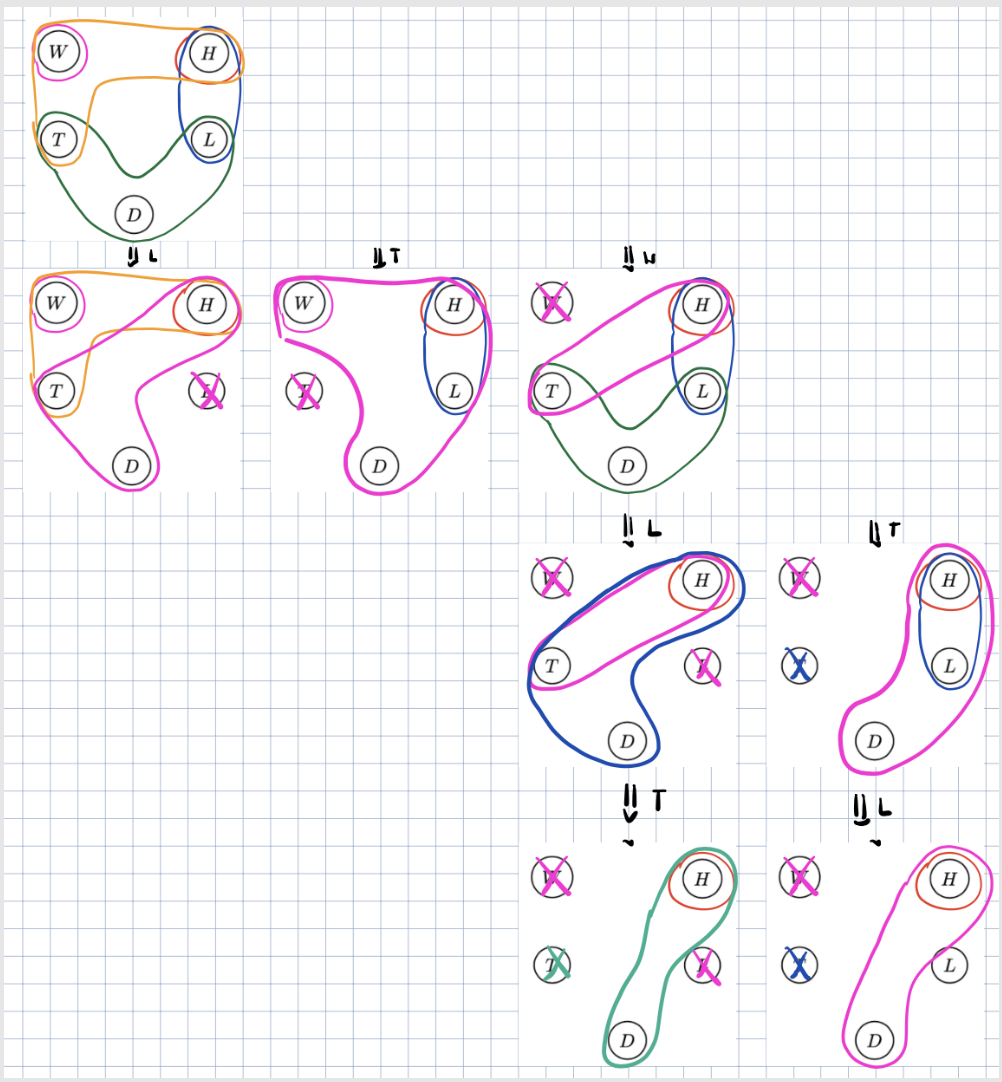


Figure 14

Example:

1. **Given:** Bayesian network

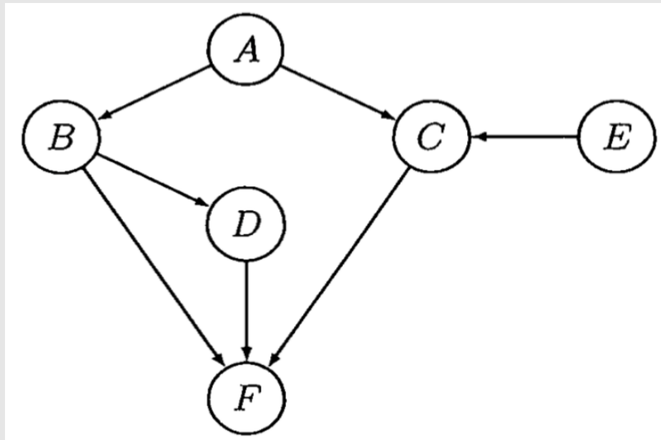


Figure 15

with cardinality of the support of each variable (i.e. number of values each variable can take on) as follows:
   - A: 2^4
   - B: 2^2
   - C: 2^{12}
   - D: 2^2
   - E: 2^3
   - F: 2^6

   Suppose elimination ordering is chosen so that the next variable eliminated is the one that results in the smallest factor (breaking ties alphabetically).
2. **Problem 1:** How many variables must be eliminated to compute $P(A, F \mid C)$?
3. **Solution 1:**
   (a) Since A, F are query, and C is evidence, we must eliminate B, D, and E, so 3 variables must be eliminated.
4. **Problem 2:** What is the first variable to be eliminated to compute $P(F \mid A)$?
5. **Solution 2:**
   (a) Try eliminating all variables that aren't query or evidence and count # of variables in union of hyperedges.

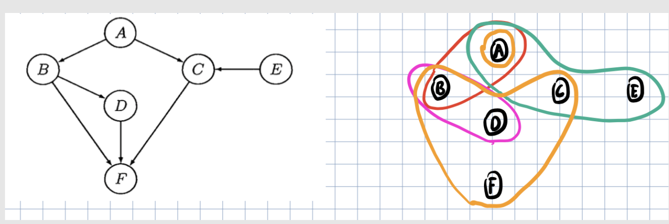


Figure 16

   i. Eliminate B: Hyperunion is ACDF $\rightarrow 2^4 \cdot 2^{12} \cdot 2^2 \cdot 2^6 = 2^{24}$
   ii. Eliminate C: Hyperunion is ABDEF $\rightarrow 2^4 \cdot 2^2 \cdot 2^2 \cdot 2^3 \cdot 2^6 = 2^{17}$
   iii. Eliminate D: Hyperunion is BCF $\rightarrow 2^2 \cdot 2^{12} \cdot 2^6 = 2^{20}$
   iv. Eliminate E: Hyperunion is AC $\rightarrow \boxed{2^4 \cdot 2^{12} = 2^{16}}$

   (b) Choose E as the first variable to be eliminated because it has the lowest support in its hyperunion.
6. **Problem 3:** What is the second variable to be eliminated to compute $P(F \mid A)$?
7. **Solution 3:**
   (a) Try eliminating all variable except F, A, E.

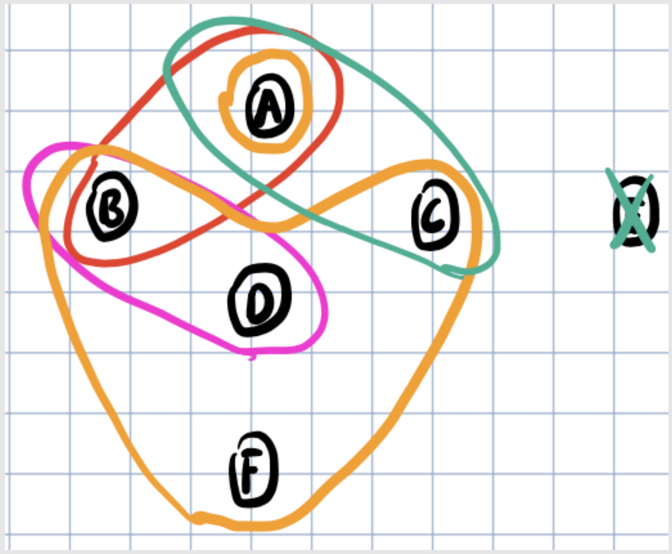


Figure 17

i. Eliminate B: Hyperunion is ACDF $\rightarrow 2^4 \cdot 2^{12} \cdot 2^2 \cdot 2^6 = 2^{24}$

ii. Eliminate C: Hyperunion is ABDF $\rightarrow \boxed{2^4 \cdot 2^2 \cdot 2^2 \cdot 2^6 = 2^{14}}$

iii. Eliminate D: Hyperunion is BCF $\rightarrow 2^2 \cdot 2^{12} \cdot 2^6 = 2^{20}$

(b) Choose C as the second variable to be eliminated because it has the lowest support in its hyperunion.

3.4.3 Inference via Sampling

Process:
1. Given samples
2. Calculate number of samples w/ the query and evidence variables.
3. Calculate number of samples w/ the evidence variables.
4. Approximate the probability of the query variable given the evidence variable by dividing the # of samples w/ the query and evidence variables by the # of samples w/ the evidence variables.

Example:
1. **Given:** Samples

W	H	T	L	D
	h	t	l	d
	h	t	l	d
	$\neg h$	$\neg t$	l	$\neg d$
	$\neg h$	t	l	d
	h	t	l	$\neg d$
	h	$\neg t$	l	d
	$\neg h$	$\neg t$	l	d
	$\neg h$	$\neg t$	$\neg l$	$\neg d$
	h	$\neg t$	$\neg l$	$\neg d$
	$\neg h$	$\neg t$	$\neg l$	d

Figure 18

2. **Problem:** Find the probability of $p(d \mid h)$.
3. **Soln:** $p(d \mid h) \approx \dfrac{\#\text{ of samples w/ } d \text{ and } h}{\#\text{ of samples w/ } h} = \dfrac{3}{5} = 0.6.$

4 Markov

4.1 General

4.1.1 Random Process

Definition: Time-varying random variables $S_0, S_1, S_2, \ldots$.

4.1.2 Markov Process

Definition: Random process + depends on previous time step only (memoryless)
- w.l.o.g. states can contain history of previous states.

4.2 Markov Chains (MCs)

Summary: In a **Markov Chain**, we assume that:
- there are no agents
- state transitions occur automatically
- S_t is the state *after* transition t
- the state transition process is stochastic and memoryless:

$$S_t \perp S_0, \ldots, S_{t-2} \mid S_{t-1}$$

  - S_t is independent of all previous states given S_{t-1}

Name	Function:
initial state distribution	$p_0(s) := \mathbb{P}[S_0 = s]$
transition distribution	$p(s'\|s) := \mathbb{P}[S_{t+1} = s'\|S_t = s]$
Prob. that state of the env. after T transitions is s	$p_T(s) := \mathbb{P}[S_T = s] = \sum_{s'} p_{T-1}(s')p(s\|s')$

- $p_{T-1}(s')$: Prob. s' at T-1 (given)
  - $p_0(s)$: Base case
- $p(s|s')$: Prob. s given s' (from graph)

4.2.1 Bayesian Network

Notes: $S_0, S_1, S_2, \ldots$ form a **Bayesian Network**:

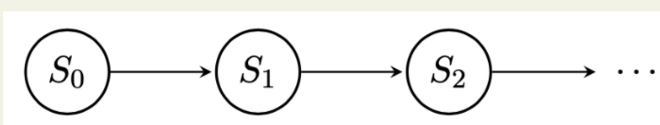


Figure 19

4.3 Markov Reward Processes (MRPs)

Summary: In a **Markov Reward Process**, we assume that:

- there is one agent
- state transitions occur automatically (i.e. agent has no control over actions)
- S_t is the state *after* transition t
- the state transition process is stochastic and memoryless:

$$S_t \perp S_0, \ldots, S_{t-2} \mid S_{t-1}$$

  - S_t is independent of all previous states given S_{t-1}
- R_t is the reward for transition t, i.e., $(S_{t-1}, \varnothing, S_t)$

Name	**Function:**
Initial state distribution	$p_0(s) := \mathbb{P}[S_0 = s]$
Transition distribution	$p(s'\|s) := \mathbb{P}[S_{t+1} = s'\|S_t = s]$
Reward function	$r(s, s') :=$ reward for transition $(s, \varnothing, s')$
Discount factor	$\gamma \in [0, 1]$
Return after T transitions	$\begin{aligned} U_T &= \sum_{t=1}^{T} \gamma^{t-1} R_t \\ &= U_{T-1} + \gamma^{T-1} R_T \end{aligned}$

- i.e. The (possibly discounted) sum of the rewards after T transitions (sequence of rewards)
- **Why?**
  - Future rewards are less valuable than immediate rewards.
  - Won't converge if sum goes to ∞ if $\gamma = 1$.

Expected return after T transitions	$\begin{aligned} \mathbb{E}[U_T] &= \mathbb{E}[U_{T-1}] + \gamma^{T-1}\mathbb{E}[R_t] \\ &= \mathbb{E}[U_{T-1}] + \gamma^{T-1} \sum_{s,s'} p_{T-1}(s)p(s'\|s)r(s, s') \end{aligned}$

- $p_{T-1}(s)p(s'|s)$: Prob. $s \to s'$
- $r(s, s')$: rwd $s \to s'$
- $\mathbb{E}[U_0] := 0$: Base case

4.3.1 Bayesian Network

Notes: $S_0, R_1, S_1, R_2, S_2, \ldots$ form a **Bayesian Network**:

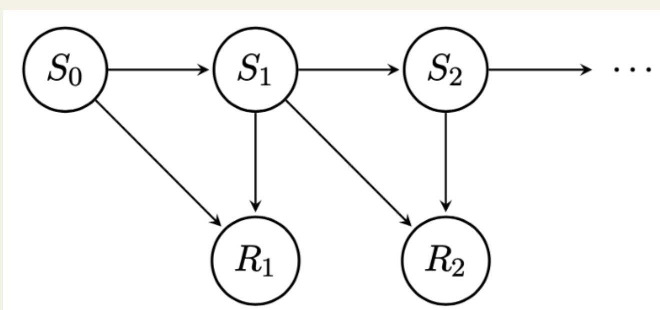


Figure 20

4.4 Markov Decision Processes (MDPs)

4.4.1 Setup

Summary: In a **Markov Decision Process (MDP)**, we assume that:

- there is one agent
- state transitions occur manually (after each action)
- S_t is the state *after* transition t
- A_t is the action inducing transition t
- the state transition process is stochastic and memoryless:

$$S_t \perp S_0, A_1, \ldots, S_{t-2}, A_{t-1} \mid S_{t-1}, A_t$$

  - S_t is independent of all previous states and actions given S_{t-1} and A_t
- R_t is the reward for transition t, i.e., (S_{t-1}, A_t, S_t)

Name	**Function:**
initial state distribution	$p_0(s) := \mathbb{P}[S_0 = s]$
transition distribution	$p(s'\|s, a) := \mathbb{P}[S_t = s'\|A_t = a, S_{t-1} = s]$
reward function	$r(s, a, s') :=$ reward for transition (s, a, s')
a time-invariant policy for choosing actions	$\pi(a\|s) := \mathbb{P}[A_t = a\|S_t = s]$
Maximum number of transitions • A Markov Decision Process can be either: – **Finite**: $T_{\max}$ is finite – **Infinite**: $T_{\max}$ is infinite ∗ For infinite MDPs, we must have $\gamma < 1$.	$T_{\max}$
Prob. that state of the env. after T transitions is s • $p_{T-1}(s)$: Prob. s' at T-1 • $\pi(a\|s')$: Action a from s' • $p(s\|s', a)$: Prob. s given s', a	$p_T(s) = \sum_{a,s'} p_{T-1}(s)\pi(a\|s')p(s\|s', a)$
Expected return after T transitions • $\mathbb{E}_\pi[R_t] = \sum_{s,a,s'} p_{T-1}(s)\pi(a \mid s)p(s' \mid s, a)r(s, a, s')$ • $\mathbb{E}_\pi[U_0] = 0$: Base case.	$\mathbb{E}_\pi[U_T]=\mathbb{E}_\pi[U_{T-1}] + \gamma^{T-1}\mathbb{E}_\pi[R_t]$
Future return after τ transitions • Starting at $\tau + 1$ for the future return.	$G_\tau = \sum_{t=\tau+1}^{T} \gamma^{t-(\tau+1)}R_t = R_{\tau+1} + \gamma G_{\tau+1}$
Expected future return after τ transitions given $S_\tau{=}s$ • $\mathbb{E}_\pi[G_{T_{\max}} \mid S_{T_{\max}} = s] = 0$: Base case.	$\mathbb{E}_\pi[G_\tau \mid S_\tau{=}s]=\mathbb{E}_\pi[R_{\tau+1} \mid S_\tau{=}s] + \gamma\mathbb{E}_\pi[G_{\tau+1} \mid S_\tau{=}s] = \sum_{a,s'} \pi(a \mid s)p(s' \mid s, a)\left(r(s, a, s') + \gamma\mathbb{E}_\pi[G_{\tau+1} \mid S_{\tau+1}{=}s']\right)$

Summary:

Name	**Function:**
Value function	$\begin{aligned} v_\pi(s,T) &:= \mathbb{E}_\pi[G_{T_{\max}-T} \mid S_{T_{\max}-T} = s] \\ &= \sum_{a,s'} \pi(a \mid s) p(s' \mid s,a) \left(r(s,a,s') + \gamma v_\pi(s', T-1)\right) \end{aligned}$ • Value of state s under the policy π with T transitions remaining. – i.e. How good the state is at time T (e.g. If $v(s,T) = 5$, then the expected future return at T is 5). • $v(s,0) = 0$ for all s: Base case
Optimal action	$\begin{aligned} a^*(s,T) &= \arg \max_{a \in \mathcal{A}(s)} \sum_{s'} p(s' \mid s,a) \left(r(s,a,s') + \gamma v_{\pi^*}(s', T-1)\right) \\ &= \arg \max_{a \in \mathcal{A}(s)} q^*(s,a,T) \end{aligned}$
Optimal policy	$\pi^*(a \mid s,T) = \arg \max_{\pi(a \mid s,T)} \mathbb{E}_\pi[G_\tau \mid S_\tau = s] = \begin{cases} 1 & \text{if } a = a^*(s,T) \\ 0 & \text{otherwise} \end{cases}$ • Choose $\pi(\cdot \mid s)$ to maximize the expected future return after T transitions given $S_\tau = s$. • **Note:** Policy always depends on transitions remaining so may omit.
Optimal value function	$v^*(s,T) = \max_a \sum_{s'} p(s' \mid a,s) \left(r(s,a,s') + \gamma v^*(s', \tau+1)\right)$ • Assume we use an optimal policy π^*. • $v^*(s,0) = 0$ for all s: Base case.
Q function (quality)	$\begin{aligned} q_\pi(s,a,T) &:= \mathbb{E}_\pi[G_{T_{\max}-T} \mid S_{T_{\max}-T} = s, A_{T_{\max}-(T-1)} = a] \\ &= \sum_{s'} p(s' \mid s,a) \left(r(s,a,s') + \gamma \sum_{a'} \pi(a' \mid s') q_\pi(s',a',T-1) \right) \end{aligned}$ • Quality of move (s,a) under policy π with T transitions remaining. • $q_\pi(s,a,0) = 0$ for all s,a: Base case.
Optimal Q function	$q^*(s,a,T) = \sum_{s'} p(s' \mid s,a) \left(r(s,a,s') + \gamma \max_{a'} q^*(s',a',T-1) \right)$ • $q^*(s,a,0) = 0$ for all s,a: Base case.

4.4.2 Bayesian Network

Notes: $S_0, A_1, R_1, S_1, A_2, R_2, S_2, \ldots$ form a **Bayesian Network**:

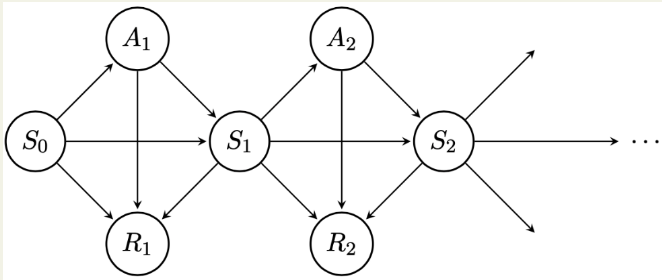


Figure 21

4.4.3 Intuition on Formulae

Notes:

$$\mathbb{E}_\pi[R_{\tau+1} \mid S_\tau = s] = \sum_{a,s'} \pi(a \mid s)p(s' \mid a, s)r(s, a, s')$$

- $\pi(a \mid s)p(s' \mid a, s)$: Prob. of getting to s' from s w/ action a
- $r(s, a, s')$: Reward of getting to s' from s w/ action a

$$\mathbb{E}_\pi[G_{\tau+1} \mid S_\tau = s] = \sum_{a,s'} \pi(a \mid s)p(s' \mid a, s)\mathbb{E}_\pi[G_{\tau+1} \mid S_{\tau+1} = s']$$

- $\pi(a \mid s)p(s' \mid a, s)$: Prob. of getting to s' from s w/ action a
- $\mathbb{E}_\pi[G_{\tau+1} \mid S_{\tau+1} = s']$: Expected future return at $\tau + 1$ from s' at $\tau + 1$.
- $\sum_{a,s'}$: Sum over all possible future states and current actions to get expected future return at $\tau + 1$ from s at τ.

4.5 Canonical Examples

4.5.1 Markov Chains

Example:

1. **Given:** Caveman needs to predict the weather, W, which is either sunny or rainy. Suppose the weather tomorrow depends on the weather today:

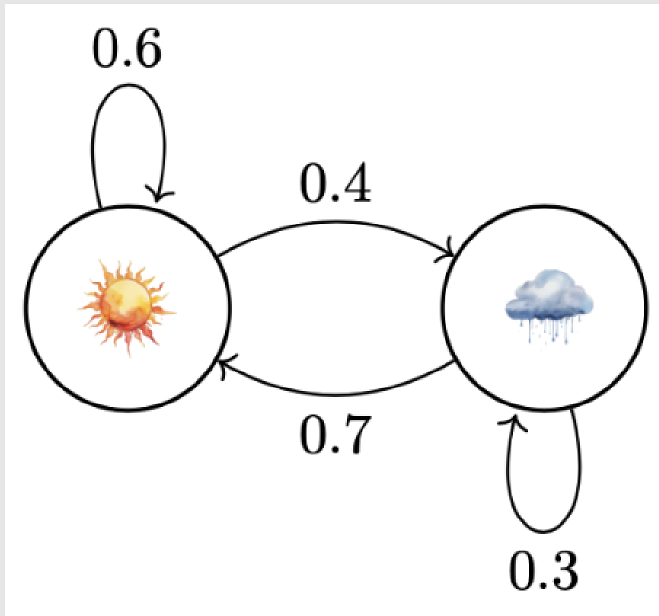


Figure 22

2. **Problem:** Caveman wants to predict the weather on a given day.

4.5.2 Markov Reward Processes

Example:

1. **Given:** Caveman needs to predict the weather, W, which is either sunny or rainy. Suppose the weather tomorrow depends on the weather today:

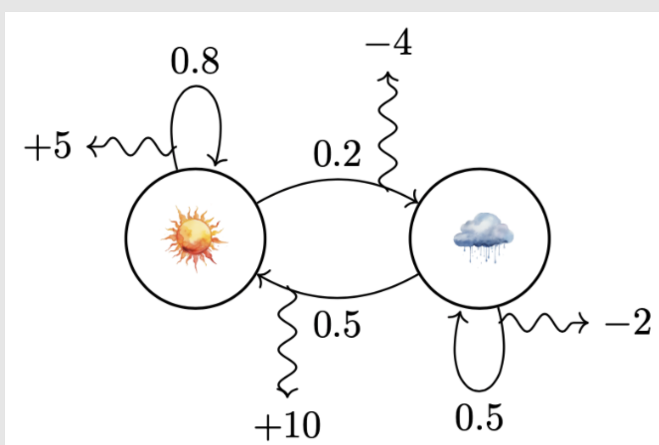


Figure 23

   - Depending on the transition, caveman may feel happier/sadder. This is quantified w/ the rewards.

2. **Problem:** Caveman wants to predict the weather on a given day that maximizes his happiness.

4.5.3 Markov Decision Processes

Example:

1. **Given:**

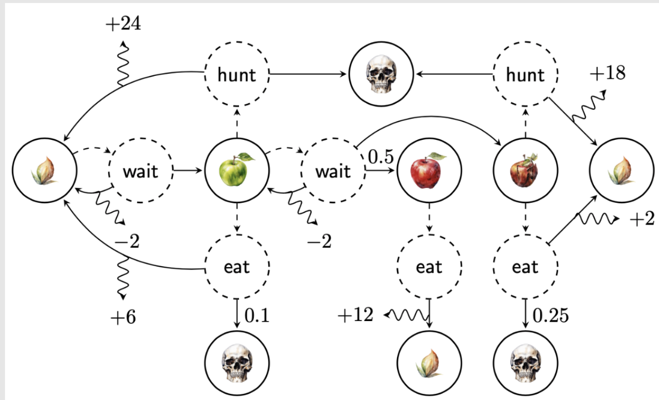


Figure 24

- Solid straight line: Outcome of action a from state s.
- Dotted straight line: Choice of action (policy) from state s.
  - If policy known, then reduced to MRP.
- Squiggly line: Reward for action a from state s to state s'.
- Assume uniform probability.
  - Since $\sum p = 1$, therefore count # of arrows going out of s and divide by 1 to get p.
- Same states have the same connections (i.e. all can use them just to hard to draw)

2. **Problem:** Find the optimal policy for $\gamma = 1$ and $T_{\max} = 5$.
3. **Soln:**

Warning:

- Be careful with the problems. Verify the answers. Go up to at least 2 steps since that tests everything.
- Be able to go through the formula quickly.
- 1st question on the quiz.

Example:

T	s	a	$q^*(s,a,T) = \sum_{s'} p(s' \mid s,a)\left(r(s,a,s') + \gamma \max_{a'} q^*(s',a',T-1)\right)$
0	-	-	0

- Best Action: $a^*(s,0) = \text{NA}$

1	seed	wait	$q^*(\text{seed},\text{wait},1) = \underbrace{0.5(-2+0)}_{s'=\text{seed}} + \underbrace{0.5(0+0)}_{s'=\text{ga}} = -1$

- Best Action: $a^*(\text{seed},1) = \text{wait}$

1	ga	wait	$q^*(\text{ga},\text{wait},1) = \underbrace{0.25(-2+0)}_{s'=\text{ga}} + \underbrace{0.5(0+0)}_{s'=\text{rea}} + \underbrace{0.25(0+0)}_{s'=\text{roa}} = -0.5$
1	ga	eat	$q^*(\text{ga},\text{eat},1) = \underbrace{0.1(0+0)}_{s'=\text{dead}} + \underbrace{0.9(6+0)}_{s'=\text{seed}} = 5.4$
1	ga	hunt	$q^*(\text{ga},\text{hunt},1) = \underbrace{0.5(24+0)}_{s'=\text{seed}} + \underbrace{0.5(0+0)}_{s'=\text{dead}} = 12$

- Best Action: $a^*(\text{ga},1) = \text{hunt}$

1	rea	eat	$q^*(\text{rea},\text{eat},1) = \underbrace{1(12+0)}_{s'=\text{seed}} = 12$

- Best Action: $a^*(\text{rea},1) = \text{eat}$

1	roa	eat	$q^*(\text{roa},\text{eat},1) = \underbrace{0.25(0+0)}_{s'=\text{dead}} + \underbrace{0.75(2+0)}_{s'=\text{seed}} = 1.5$
1	roa	hunt	$q^*(\text{roa},\text{hunt},1) = \underbrace{0.5(0+0)}_{s'=\text{dead}} + \underbrace{0.5(18+0)}_{s'=\text{seed}} = 9$

- Best Action: $a^*(\text{roa},1) = \text{hunt}$

1	dead	-	$q^*(\text{dead},-,1) = \underbrace{1(0+0)}_{s'=\text{end}} = 0$

- Best Action: $a^*(s,1) = -$

- Optimal Policy w/ 1 Transition Remaining: $\pi^*(a \mid s,1) = \begin{cases} 1 & \text{if } a = a^*(s,1) \\ 0 & \text{otherwise} \end{cases}$

Example:

T	s	a	$q^*(s,a,T) = \sum_{s'} p(s' \mid s,a) \Big(r(s,a,s') + \gamma \max_{a'} q^*(s',a',T-1) \Big)$
2	seed	wait	$q^*(\text{seed},\text{wait},2) = \underbrace{0.5(-2-1)}_{s'=\text{seed}} + \underbrace{0.5(0+12)}_{s'=\text{ga}} = 4.5$

- Best Action: $a^*(\text{seed},2) = \text{wait}$

2	ga	wait	$q^*(\text{ga},\text{wait},2) = \underbrace{0.25(-2+12)}_{s'=\text{ga}} + \underbrace{0.5(0+12)}_{s'=\text{rea}} + \underbrace{0.25(0+9)}_{s'=\text{roa}} = 10.75$
2	ga	eat	$q^*(\text{ga},\text{eat},2) = \underbrace{0.1(0+0)}_{s'=\text{dead}} + \underbrace{0.9(6-1)}_{s'=\text{seed}} = 4.5$
2	ga	hunt	$q^*(\text{ga},\text{hunt},2) = \underbrace{0.5(24-1)}_{s'=\text{seed}} + \underbrace{0.5(0+0)}_{s'=\text{dead}} = 11.5$

- Best Action: $a^*(\text{ga},2) = \text{hunt}$

2	rea	eat	$q^*(\text{rea},\text{eat},2) = \underbrace{1(12-1)}_{s'=\text{seed}} = 11$

- Best Action: $a^*(\text{rea},2) = \text{eat}$

2	roa	eat	$q^*(\text{roa},\text{eat},2) = \underbrace{0.25(0+0)}_{s'=\text{dead}} + \underbrace{0.75(2-1)}_{s'=\text{seed}} = 0.75$
2	roa	hunt	$q^*(\text{roa},\text{hunt},2) = \underbrace{0.5(0+0)}_{s'=\text{dead}} + \underbrace{0.5(18-1)}_{s'=\text{seed}} = 8.5$

- Best Action: $a^*(\text{roa},2) = \text{hunt}$

2	dead	-	$q^*(\text{dead},-,2) = \underbrace{1(0+0)}_{s'=\text{end}} = 0$

- Best Action: $a^*(s,2) = -$

- Optimal Policy w/ 2 Transitions Remaining: $\pi^*(a \mid s, 2) = \begin{cases} 1 & \text{if } a = a^*(s,2) \\ 0 & \text{otherwise} \end{cases}$

Example:

T	s	a	$q^*(s,a,T) = \sum_{s'} p(s' \mid s,a)\left(r(s,a,s') + \gamma \max_{a'} q^*(s',a',T-1)\right)$
3	seed	wait	$q^*(\text{seed},\text{wait},3) = \underbrace{0.5(-2+4.5)}_{s'=\text{seed}} + \underbrace{0.5(0+11.5)}_{s'=\text{ga}} = 7$

- Best Action: $a^*(\text{seed},3) = \text{wait}$

3	ga	wait	$q^*(\text{ga},\text{wait},3) = \underbrace{0.25(-2+11.5)}_{s'=\text{ga}} + \underbrace{0.5(0+11)}_{s'=\text{rea}} + \underbrace{0.25(0+8.5)}_{s'=\text{roa}} = 10$
3	ga	eat	$q^*(\text{ga},\text{eat},3) = \underbrace{0.1(0+0)}_{s'=\text{dead}} + \underbrace{0.9(6+4.5)}_{s'=\text{seed}} = 9.45$
3	ga	hunt	$q^*(\text{ga},\text{hunt},3) = \underbrace{0.5(24+4.5)}_{s'=\text{seed}} + \underbrace{0.5(0+0)}_{s'=\text{dead}} = 14.25$

- Best Action: $a^*(\text{ga},3) = \text{hunt}$

3	rea	eat	$q^*(\text{rea},\text{eat},3) = \underbrace{1(12+4.5)}_{s'=\text{seed}} = 16.5$

- Best Action: $a^*(\text{rea},3) = \text{eat}$

3	roa	eat	$q^*(\text{roa},\text{eat},3) = \underbrace{0.25(0+0)}_{s'=\text{dead}} + \underbrace{0.75(2+4.5)}_{s'=\text{seed}} = 4.875$
3	roa	hunt	$q^*(\text{roa},\text{hunt},3) = \underbrace{0.5(0+0)}_{s'=\text{dead}} + \underbrace{0.5(18+4.5)}_{s'=\text{seed}} = 11.25$

- Best Action: $a^*(\text{roa},3) = \text{hunt}$

3	dead	-	$q^*(\text{dead},-,3) = \underbrace{1(0+0)}_{s'=\text{end}} = 0$

- Best Action: $a^*(s,3) = -$

- Optimal Policy w/ 3 Transitions Remaining: $\pi^*(a \mid s,3) = \begin{cases} 1 & \text{if } a = a^*(s,3) \\ 0 & \text{otherwise} \end{cases}$

Example:

T	s	a	$q^*(s,a,T) = \sum_{s'} p(s' \mid s,a) \Big(r(s,a,s') + \gamma \max_{a'} q^*(s',a',T-1)\Big)$
4	seed	wait	$q^*(\text{seed}, \text{wait}, 4) = \underbrace{0.5(-2+7)}_{s' = \text{seed}} + \underbrace{0.5(0+14.25)}_{s' = \text{ga}} = 9.625$
			• Best Action: $a^*(\text{seed}, 4) = \text{wait}$
4	ga	wait	$q^*(\text{ga}, \text{wait}, 4) = \underbrace{0.25(-2+14.25)}_{s' = \text{ga}} + \underbrace{0.5(0+16.5)}_{s' = \text{rea}} + \underbrace{0.25(0+11.25)}_{s' = \text{roa}} = 14.125$
4	ga	eat	$q^*(\text{ga}, \text{eat}, 4) = \underbrace{0.1(0+0)}_{s' = \text{dead}} + \underbrace{0.9(6+7)}_{s' = \text{seed}} = 11.7$
4	ga	hunt	$q^*(\text{ga}, \text{hunt}, 4) = \underbrace{0.5(24+7)}_{s' = \text{seed}} + \underbrace{0.5(0+0)}_{s' = \text{dead}} = 15.5$
			• Best Action: $a^*(\text{ga}, 4) = \text{hunt}$
4	rea	eat	$q^*(\text{rea}, \text{eat}, 4) = \underbrace{1(12+7)}_{s' = \text{seed}} = 19$
			• Best Action: $a^*(\text{rea}, 4) = \text{eat}$
4	roa	eat	$q^*(\text{roa}, \text{eat}, 4) = \underbrace{0.25(0+0)}_{s' = \text{dead}} + \underbrace{0.75(2+7)}_{s' = \text{seed}} = 6.75$
4	roa	hunt	$q^*(\text{roa}, \text{hunt}, 4) = \underbrace{0.5(0+0)}_{s' = \text{dead}} + \underbrace{0.5(18+7)}_{s' = \text{seed}} = 12.5$
			• Best Action: $a^*(\text{roa}, 4) = \text{hunt}$
4	dead	-	$q^*(\text{dead}, -, 4) = \underbrace{1(0+0)}_{s' = \text{end}} = 0$
			• Best Action: $a^*(s, 4) = -$

- Optimal Policy w/ 4 Transitions Remaining: $\pi^*(a \mid s, 4) = \begin{cases} 1 & \text{if } a = a^*(s,4) \\ 0 & \text{otherwise} \end{cases}$

Example:

T	s	a	$q^*(s,a,T) = \sum_{s'} p(s' \mid s,a) \left(r(s,a,s') + \gamma \max_{a'} q^*(s',a',T-1) \right)$
5	seed	wait	$q^*(\text{seed}, \text{wait}, 5) = \underbrace{0.5(-2+9.625)}_{s'=\text{seed}} + \underbrace{0.5(0+15.5)}_{s'=\text{ga}} = 11.5625$
			• Best Action: $a^*(\text{seed}, 5) = \text{wait}$
5	ga	wait	$q^*(\text{ga}, \text{wait}, 5) = \underbrace{0.25(-2+15.5)}_{s'=\text{ga}} + \underbrace{0.5(0+19)}_{s'=\text{rea}} + \underbrace{0.25(0+12.5)}_{s'=\text{roa}} = 16$
5	ga	eat	$q^*(\text{ga}, \text{eat}, 5) = \underbrace{0.1(0+0)}_{s'=\text{dead}} + \underbrace{0.9(6+9.625)}_{s'=\text{seed}} = 14.0625$
5	ga	hunt	$q^*(\text{ga}, \text{hunt}, 5) = \underbrace{0.5(24+9.625)}_{s'=\text{seed}} + \underbrace{0.5(0+0)}_{s'=\text{dead}} = 16.8125$
			• Best Action: $a^*(\text{ga}, 5) = \text{hunt}$
5	rea	eat	$q^*(\text{rea}, \text{eat}, 5) = \underbrace{1(12+9.625)}_{s'=\text{seed}} = 21.625$
			• Best Action: $a^*(\text{rea}, 5) = \text{eat}$
5	roa	eat	$q^*(\text{roa}, \text{eat}, 5) = \underbrace{0.25(0+0)}_{s'=\text{dead}} + \underbrace{0.75(2+9.625)}_{s'=\text{seed}} = 8.71875$
5	roa	hunt	$q^*(\text{roa}, \text{hunt}, 5) = \underbrace{0.5(0+0)}_{s'=\text{dead}} + \underbrace{0.5(18+9.625)}_{s'=\text{seed}} = 13.8125$
			• Best Action: $a^*(\text{roa}, 5) = \text{hunt}$
5	dead	-	$q^*(\text{dead}, -, 5) = \underbrace{1(0+0)}_{s'=\text{end}} = 0$
			• Best Action: $a^*(s, 5) = -$

- Optimal Policy w/ 5 Transitions Remaining: $\pi^*(a \mid s, 5) = \begin{cases} 1 & \text{if } a = a^*(s,5) \\ 0 & \text{otherwise} \end{cases}$