COMS W4701: Artificial Intelligence

Lecture 22: Decision Networks

Tony Dear, Ph.D.
Department of Computer Science
School of Engineering and Applied Sciences

Today

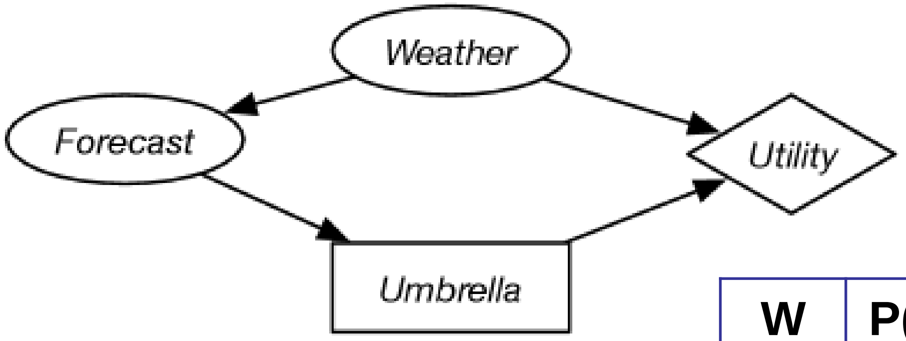
- Decision networks
- Variable elimination for decision networks
- Value of perfect information
- DDNs and POMDPs

Decision Networks

- **Decision network**: Bayes net combined with decision and utility nodes

- *Chance* nodes still represent random variables with an associated CPT
- *Decision* nodes may be explicitly assigned specific values by the agent

- A single *utility* node enumerates possible numerical utility values depending on the combination of parent node values



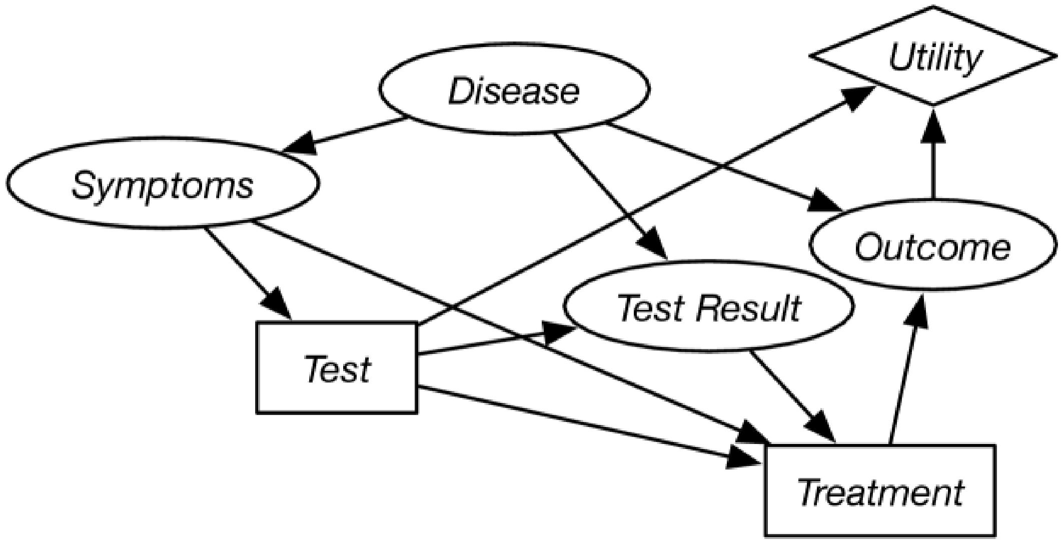
W	P(W)
Dry	0.7
Wet	0.3

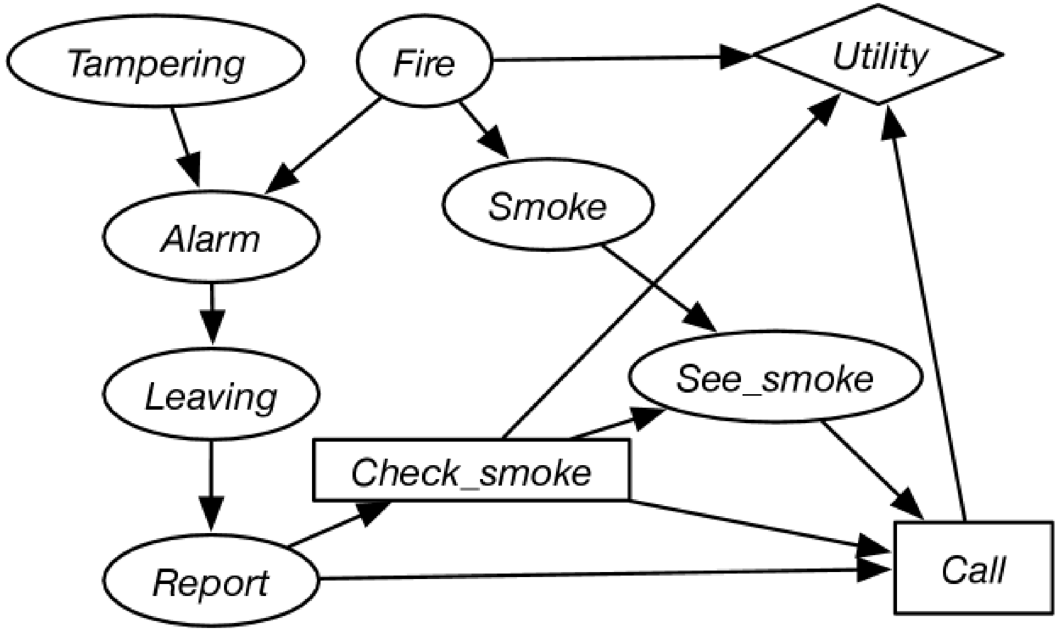
W	F	P(F|W)
Dry	Sunny	0.7
Dry	Cloudy	0.2
Dry	Rainy	0.1
Wet	Sunny	0.15
Wet	Cloudy	0.25
Wet	Rainy	0.6

W	U	Utility
Dry	Take	20
Dry	Leave	100
Wet	Take	70
Wet	Leave	0

Sequential Decisions

- Can model *sequential decisions* with many decision nodes following topological order
- **No-forgetting** network: For all $i < j$, all parents of D_i are also parents of D_j
- Such an agent remembers all previous decisions and relevant information





Decision Functions and Policies

- A **decision function** for a decision node maps each combination of its parents' values to a decision value ("action")

- Number of possible functions is exponential in number of parents
- Given n binary parent nodes, there are 2^n parent value combinations, and thus 2^{2^n} possible decision functions for a binary decision node

- A **policy** is a set of decision functions, one for each decision variable
- An **optimal policy** maximizes the expected utility (**value**) of the network

Expected Utility

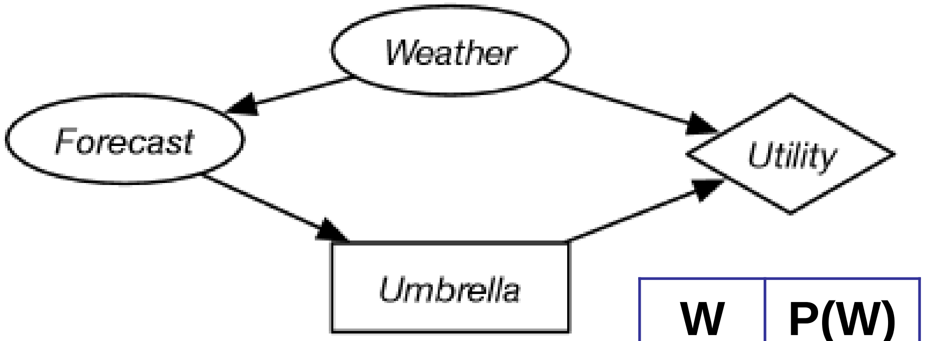
- Consider the policy π: Leave umbrella if F is sunny, take umbrella otherwise

- The **expected utility** is the sum of utilities in all possible scenarios

$$EU = \sum_{w,f,u=\pi(f)} P(w)P(f|w)Utility(w,u)$$

$$= 0.7(0.7(100) + 0.2(20) + 0.1(20)) + 0.3(0.15(0) + 0.25(70) + 0.6(70))$$

$$= 71.05$$



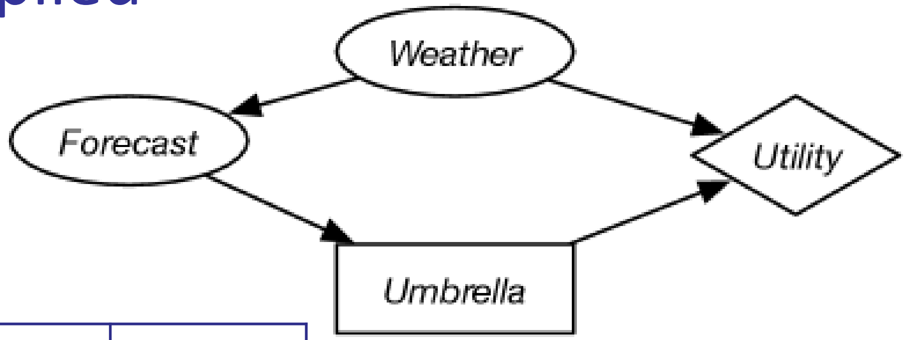
W	P(W)
Dry	0.7
Wet	0.3

W	F	P(F|W)
Dry	Sunny	0.7
Dry	Cloudy	0.2
Dry	Rainy	0.1
Wet	Sunny	0.15
Wet	Cloudy	0.25
Wet	Rainy	0.6

W	U	Utility
Dry	Take	20
Dry	Leave	100
Wet	Take	70
Wet	Leave	0

Factor Operations

- We can also compute utility using *factor* operations
- Both chance utility node factors can be multiplied and marginalized in the same way



$$EU = \sum_{w,f,u=\pi(f)} f_W(w) f_F(w,f) Utility(w,u)$$

F	U	Utility
Sunny	Leave	49
Cloudy	Take	8.05
Rainy	Take	14

$= \sum$

W	f_W
Dry	0.7
Wet	0.3

$\times$

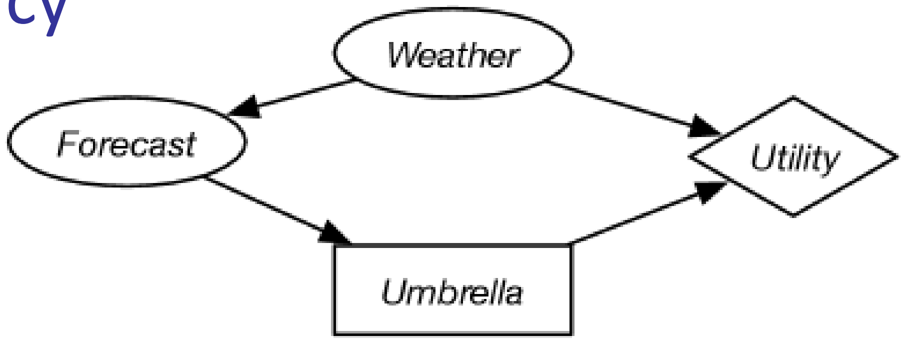
W	F	f_F
Dry	Sunny	0.7
Dry	Cloudy	0.2
Dry	Rainy	0.1
Wet	Sunny	0.15
Wet	Cloudy	0.25
Wet	Rainy	0.6

$\times$

W	U	Utility
Dry	Take	20
Dry	Leave	100
Wet	Take	70
Wet	Leave	0

Finding Optimal Policies

- Suppose that we want to *find* an optimal policy
- We first obtain a factor over all possible decision and parent values and the utility



$$f_{joint}(F, U) = \sum_{w} f_W(w) f_F(w, F) Utility(w, U)$$

- For each “group” of parent values, the *optimal* decision is the one that *maximizes* the utility in f_{joint}

$$\operatorname*{argmax}_{u} f_{joint}(F, u)$$

- Analogous to chance variable elimination, but with max instead of sum!

Example: Optimal Policy

$$\operatorname*{argmax}_{u} \sum_{w} f_W(w) f_F(w, F) Utility(w, u)$$

W	F	U	Utility
Dry	Sunny	Take	9.8
Dry	Cloudy	Take	2.8
Dry	Rainy	Take	1.4
Wet	Sunny	Take	3.15
Wet	Cloudy	Take	5.25
Wet	Rainy	Take	12.6
Dry	Sunny	Leave	49
Dry	Cloudy	Leave	14
Dry	Rainy	Leave	7
Wet	Sunny	Leave	0
Wet	Cloudy	Leave	0
Wet	Rainy	Leave	0

$\sum_{w}$

F	U	Utility
Sunny	Take	12.95
Sunny	Leave	49
Cloudy	Take	8.05
Cloudy	Leave	14
Rainy	Take	14
Rainy	Leave	7

argmax_u

F	U	Utility
Sunny	Leave	49
Cloudy	Leave	14
Rainy	Take	14

Optimal decision function for Umbrella

Variable Elimination

- When we eliminate (argmax) a decision variable, we must do so over a factor that contains itself, its parents, and the utility

- If we have multiple decision variables, we must eliminate decision variables from *last* to *first* following topological order

- Repeat while there are unassigned decision nodes:
  - Join and eliminate chance nodes that are *not* parents of unassigned decision nodes
  - Eliminate the topologically *last* unassigned decision node, by assigning decision function that *maximizes* expected utility for each parent value combination

Example: Fire Alarm

1. Eliminate chance nodes: Tampering, Fire, Alarm, Smoke, Leaving

$$f_1(R, CS, SS, C) = \sum_{t,f,a,s,l} f_T(t) f_F(f) f_S(f,s) f_A(t,f,a) f_L(a,l) f_R(l,R) f_{SS}(s,CS,SS) Utility(f,CS,C)$$

2. Eliminate Call:

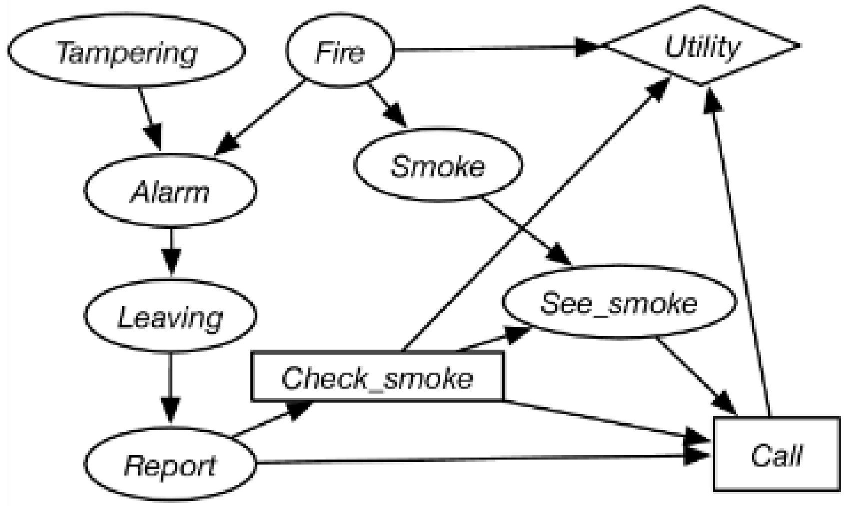
$$f_2(R, CS, SS) = \underset{c}{\operatorname{argmax}}\, f_1(R, CS, SS, c)$$

3. Eliminate See_smoke:

$$f_3(R, CS) = \sum_{ss} f_2(R, CS, ss)$$

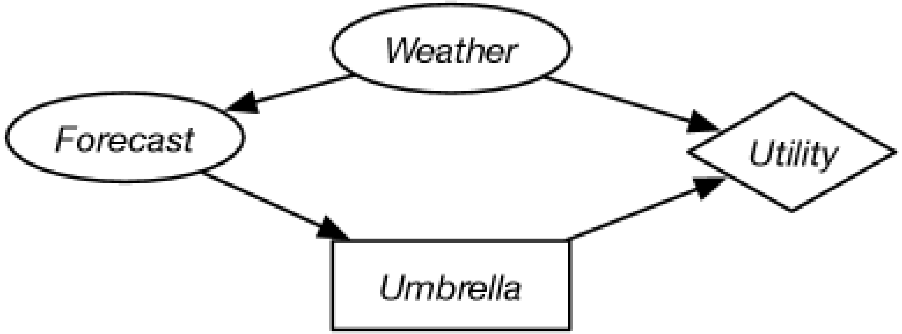
4. Eliminate Check_smoke:

$$f_4(R) = \underset{cs}{\operatorname{argmax}}\, f_3(R, cs)$$



Variable Elimination Considerations

- In a no-forgetting network, a decision variable depends on all past decisions and information



- The initial joint factor will be exponential in most if not all nodes in the network
- Finding exactly optimal policies is still NP-hard!

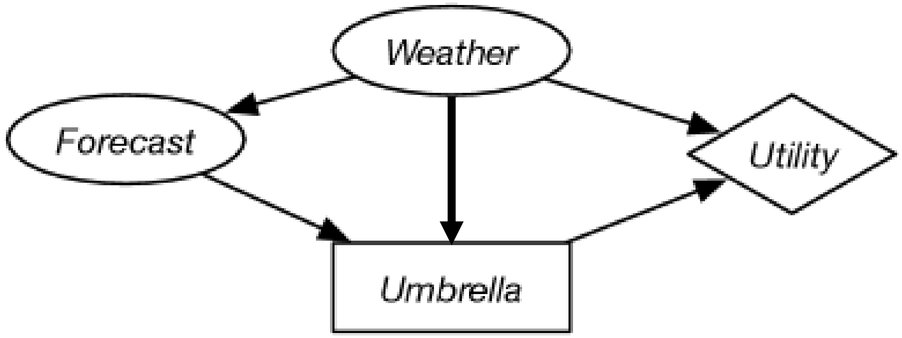
F	U	Utility
Sunny	Leave	49
Cloudy	Leave	14
Rainy	Take	14

- After finding a complete policy, the **maximum expected utility (MEU)** is the utility sum over all possible situations

$$MEU = 49 + 14 + 14 = 77$$

Additional Information

- What if Umbrella depends on both Forecast *and* Weather?
- The decision function should be computed directly from the joint factor without eliminating Weather
- Decision function is "larger" than in previous case

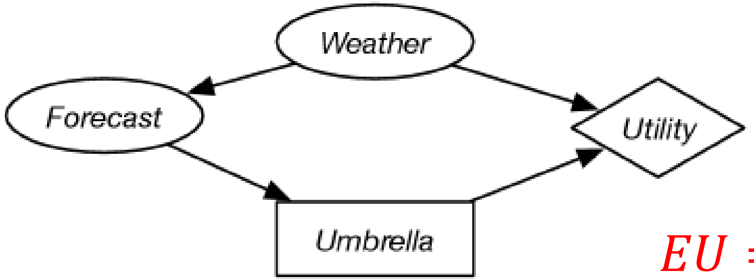


$$MEU = 91$$

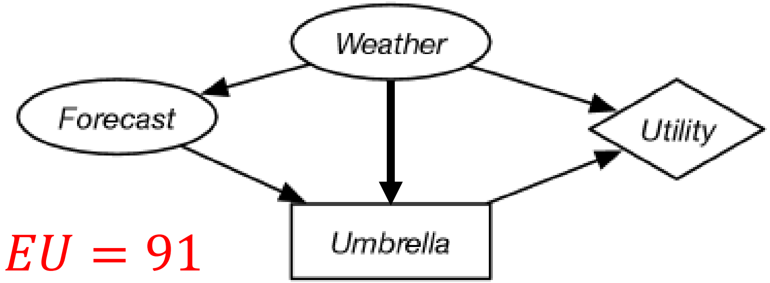
W	F	U	Utility
Dry	Sunny	Leave	49
Dry	Cloudy	Leave	14
Dry	Rainy	Leave	7
Wet	Sunny	Take	3.15
Wet	Cloudy	Take	5.25
Wet	Rainy	Take	12.6

W	F	U	Utility
Dry	Sunny	Take	9.8
Dry	Cloudy	Take	2.8
Dry	Rainy	Take	1.4
Wet	Sunny	Take	3.15
Wet	Cloudy	Take	5.25
Wet	Rainy	Take	12.6
Dry	Sunny	Leave	49
Dry	Cloudy	Leave	14
Dry	Rainy	Leave	7
Wet	Sunny	Leave	0
Wet	Cloudy	Leave	0
Wet	Rainy	Leave	0

Value of Information



$EU = 77$



$EU = 91$

- The MEU of the latter network is higher since we have more information
- Since the utility function only depends on Weather and Umbrella, the Forecast is irrelevant to making the decision!

- The **value of (perfect) information** of Weather is the *difference*: $VPI = 91 - 77 = 14$
- In general, VPI of X for decision D can be computed as the difference in maximum expected utility after adding dependency for D on X

Properties of VPI

- VPI can be used to quantify “information-seeking” decisions
- In contrast to “physical” actions in which agent interacts with environment
- Example: Diagnostic tests for a patient, as opposed to treatment plan

- VPI is an upper bound on the amount of utility that an agent would *pay* for information on a random variable

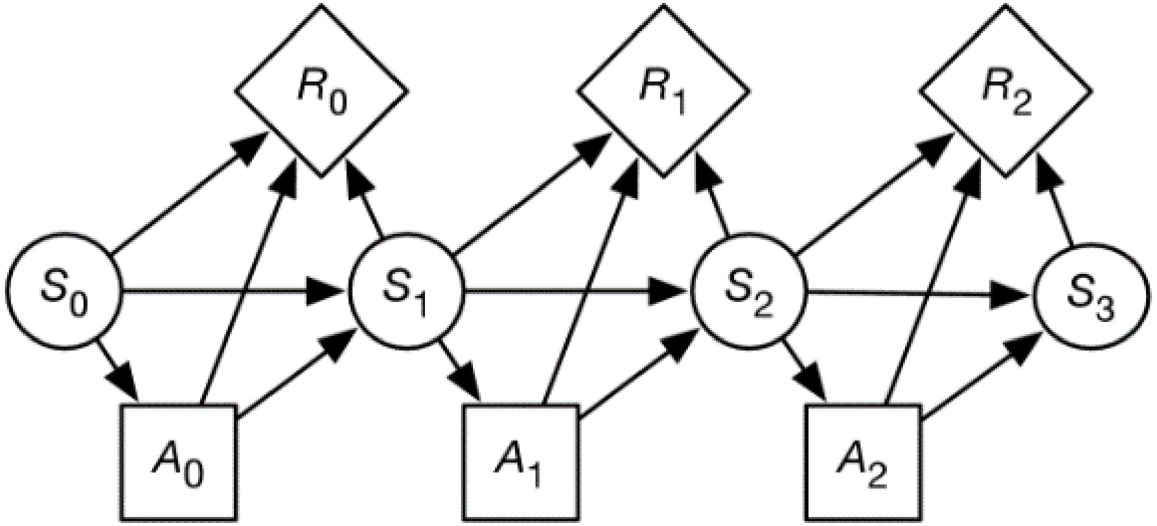
- VPI ≥ 0; expected utility can never decrease with more information
- If VPI $= 0$, then optimal policy does not depend on the random variable

Dynamic Decision Networks

- Just as HMMs represent Bayes nets repeated sequentially, **dynamic decision networks** allow for (possibly infinite) repeated decision network structure
- MDPs can be modeled using DDNs!

- State features: Chance nodes with CPTs from *transition models* $P(X_{t+1}|X_t, A_t)$
- Same set of actions available in each step, represented as decision nodes

- Rewards or utilities, possibly in each timestep and/or at the very end



Partially Observable MDPs

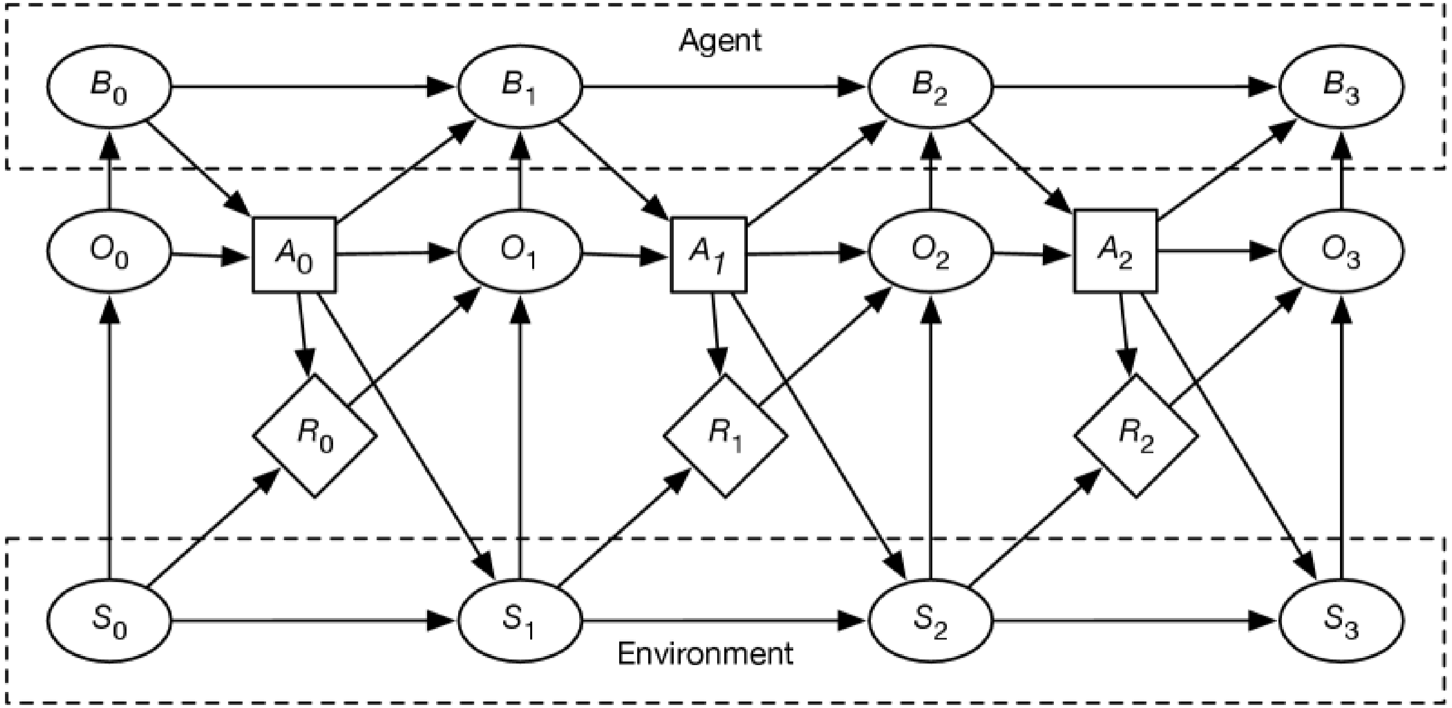
- A MDP requires that an agent knows its state in order to follow its policy
- A HMM separately considers state and observation variables to model partially observable environments

- A **partially observable MDP (POMDP)** is a decision process in which the agent sees only partial or noisy observations about its true state

- Observations can be used to maintain and update the agent's *belief state*, e.g. a probability distribution over its possible true states

Partially Observable MDPs

- POMDPs can also be represented as dynamic decision processes
- Solving for optimal policies via variable elimination or value iteration can become intractable, often have to turn to real-time filtering methods



Summary

- Decision networks model relationships between unobserved variables, decision variables, and utilities
- A policy assigns a decision function to each decision variable

- The expected utility of a policy is the average of all possible scenarios
- An optimal policy can be found by iterating between summing out irrelevant hidden variables and optimizing decisions

- Dynamic decision networks can be used to model hidden Markov models and partially observable MDPs