# Finish VE for Sequential Decisions & Value of Information and Control

Computer Science cpsc322, Lecture 35

*(Textbook Chpt 9.4)*

April, 3, 2009

# Lecture Overview

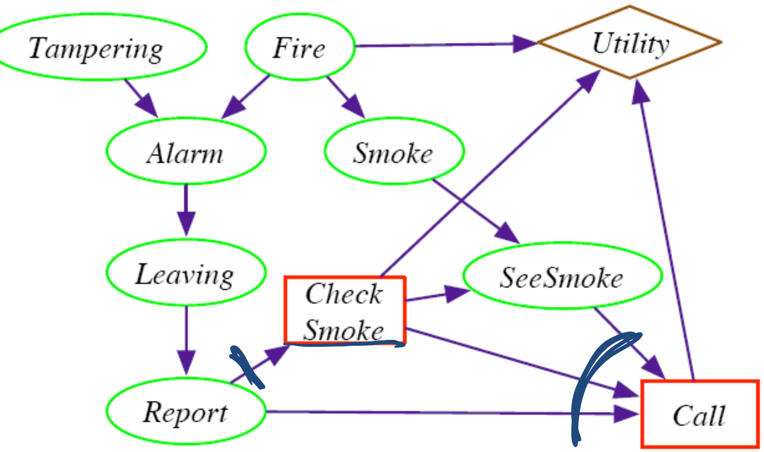
- Sequential Decisions
  - Optimal Policy
  - Variable Elimination
- Value of Information
- Value of Control

# Sequential decisions (Planning) under uncertainty

- What has been the output of planning so far? *A sequence of actions*
- Why this is not adequate?
  *The agent does not know how the environment will be when it will have to decide (i.e., to select an action)*
- What should the output be?
  *For each decision (action selection), a specification of what action will be the best in each possible configuration of the environment*

set of optimal decision functions: THE OPTIMAL POLICY

# Complexity of finding the optimal policy: how many policies?




- How many assignments to parents?
- How many decision functions? (binary decisions)
- How many policies?

- If a decision *D* has *k* binary parents, how many **assignments** of values to the parents are there?
- If there are *b* possible actions, **how many different decision functions** are there?
- If there are d decisions, each with *k* binary parents and *b* possible actions, **how many policies are there**?

# Finding the optimal policy more efficiently: VE

1. Remove all variables that are not ancestors of the utility node
2. Create a factor for each conditional probability table and a factor for the utility.
3. **Sum out** random variables that are not parents of a decision node.
4. **Eliminate** (aka sum out) the decision variables
5. **Sum out** the remaining random variables.
6. **Multiply the remaining factors**: this is the expected utility of the optimal policy.

# Eliminate the decision Variables: details

- Select a variable *D* that corresponds to the latest decision to be made
  - this variable will appear in only one factor with its parents

for each configuration of the parents

- Eliminate *D* by maximizing. This returns:
  - The optimal decision function for *D*, $\arg\max_D$ *Value*
  - A new factor to use in VE, $\max_D$ *Value*
- Repeat till there are no more decision nodes.

***Example: Eliminate CheckSmoke***

| Report | CheckSmoke | Value |
|---|---|---|
| true | true | -5.0 |
| true | false | -5.6 |
| false | true | -23.7 |
| false | false | -17.5 |

*New factor*

| Report | Value |
|---|---|
| true | -5 |
| false | -17.5 |

*Decision Function*

| Report | CheckSmoke |
|---|---|
| true | true |
| false | false |

# VE elimination reduces complexity of finding the optimal policy

- We have seen that, if a decision $D$ has $k$ binary parents, there are $b$ possible actions, If there are d decisions,
- Then there are: $(b^{2^k})^d$ *policies*
- Doing variable elimination lets us find the optimal policy after considering only $d \, b^{2^k}$ policies (we eliminate one decision at a time)

  linear in d

  - VE **is much more efficient** than searching through policy space.
  - However, this complexity is **still doubly-exponential** we'll only be able to handle relatively small problems.

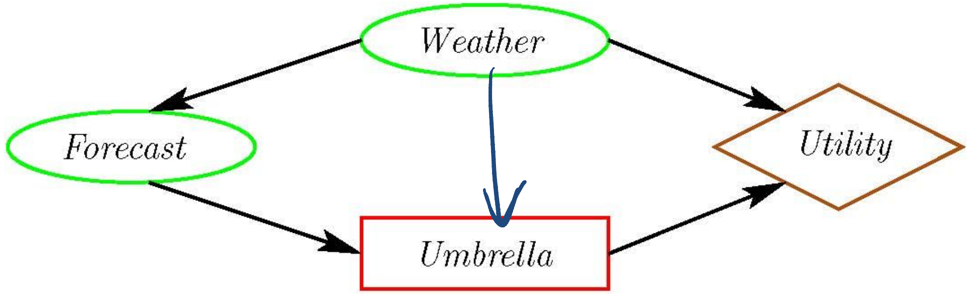
To solve Bigger problems – give up non forgetting

– approx. algorithms

# Lecture Overview

- Sequential Decisions
  - Optimal Policy
  - Variable Elimination
- Value of Information
- Value of Control

# Value of Information




- What would help the agent make a better *Umbrella* decision?

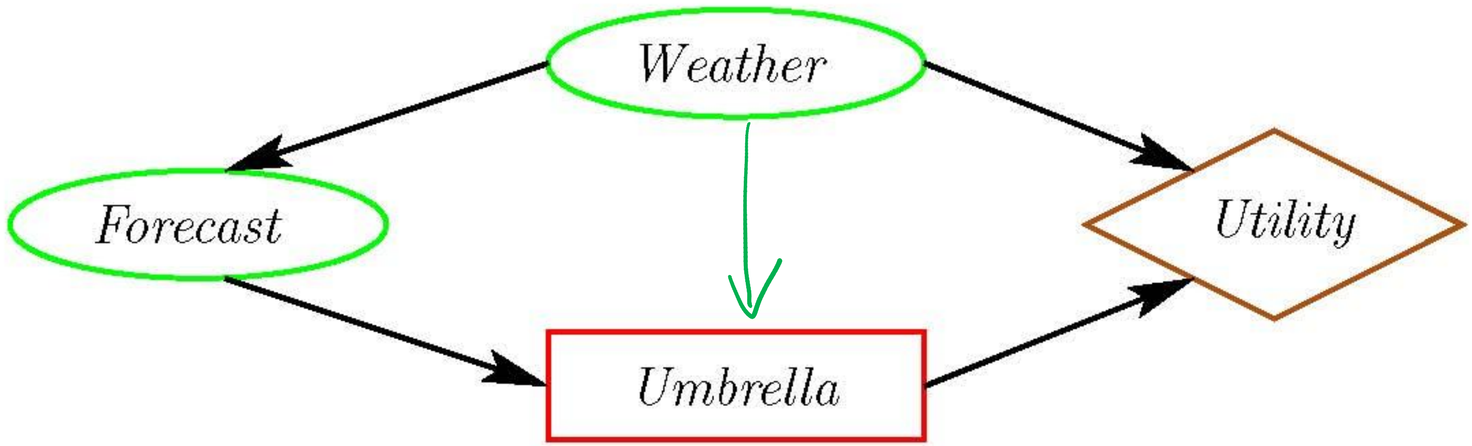
- The value of information of a random variable *X* for decision *D* is: EU(knowing X) − EU(not knowing X)

the utility of the network with an arc from *X* to *D* **minus** the utility of the network without the arc.

- Intuitively:
  - The value of information is always ≥ 0
  - It is positive only if the agent changes its policy

# Value of Information (cont.)

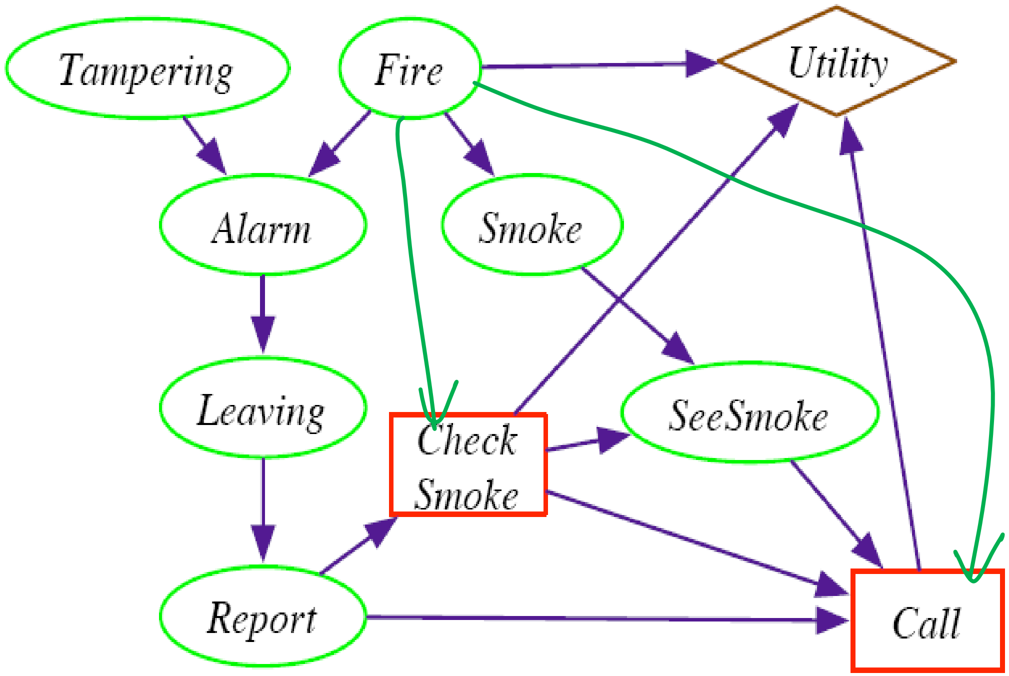
- The value of information provides a bound on how much you should be prepared to pay for a sensor. How much is a **perfect** weather forecast worth?




- Original maximum expected utility: 77
- Maximum expected utility when we know Weather: 91
- Better forecast is worth at most: 14

# Value of Information

- The value of information provides a bound on how much you should be prepared to pay for a sensor. How much is a **perfect** fire sensor worth?

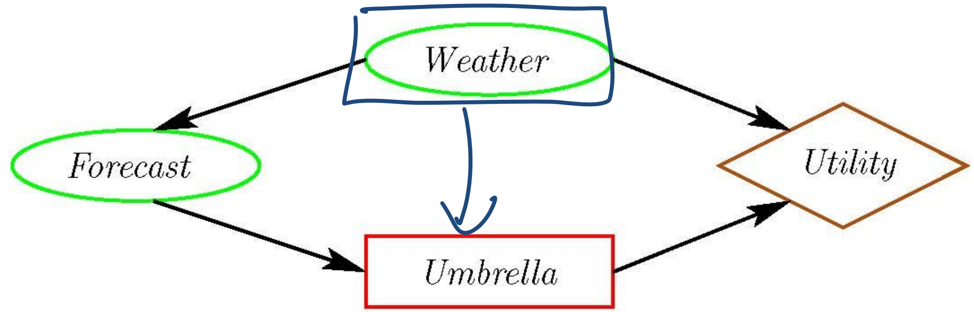



- Original maximum expected utility: -22.6
- Maximum expected utility when we know Fire: -2
- Perfect fire sensor is worth: 20.6

# Lecture Overview

- Sequential Decisions
  - Optimal Policy
  - Variable Elimination
- Value of Information
- Value of Control

# Value of Control




- What would help the agent to make an even better *Umbrella* decision? To maximize its utility.

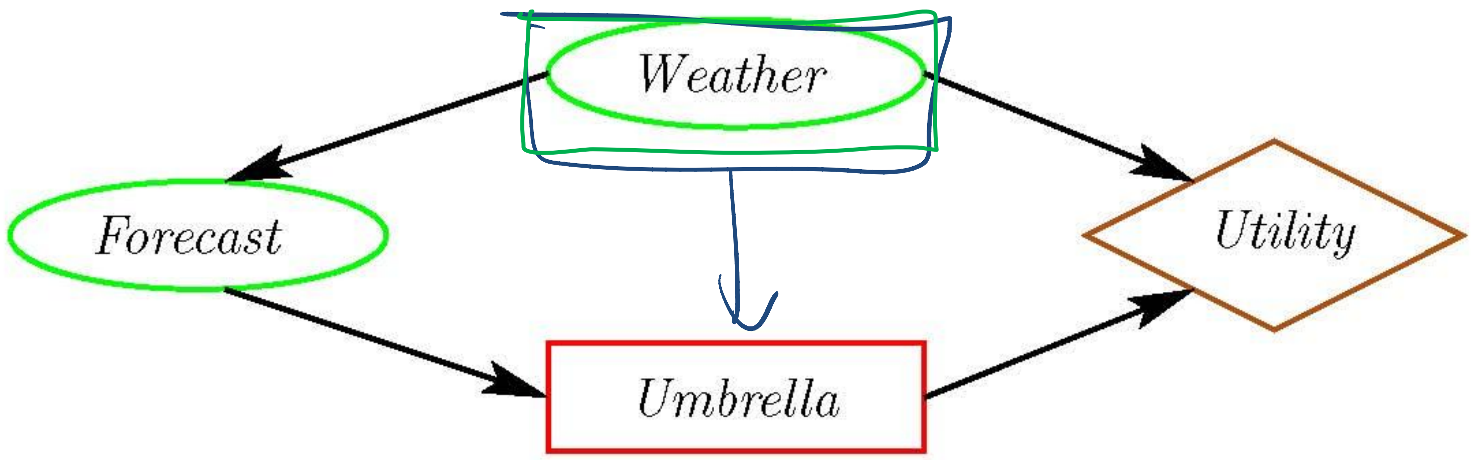
| Weather | Umbrella | Value |
|---|---|---|
| Rain | true | 70 |
| Rain | false | 0 |
| noRain | true | 20 |
| noRain | false | 100 |

- The value of control of a variable *X* is :

the utility of the network when you make X a decision variable **minus** the utility of the network when X is a random variable.

# Value of Control

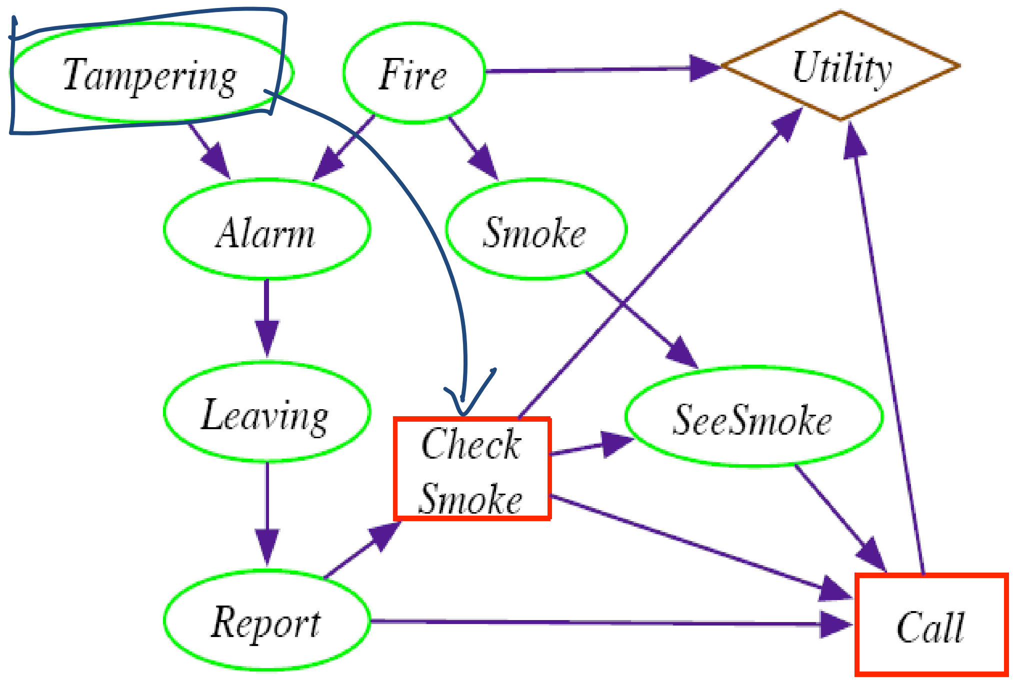
- What if we could control the weather?




- Original maximum expected utility: 77
- Maximum expected utility when we control the weather: 100
- Value of control of the weather: 23

# Value of Control

- What if we control Tampering?




- Original maximum expected utility: *-22.6*
- Maximum expected utility when we control the Tampering: *-20.7*
- Value of control of Tampering: 1.9
- Let's take a look at the policy
- Conclusion: do not tamper with fire alarms!

# Learning Goals for today's (and Wed) classes

**You can:**

- Represent **sequential decision problems** as decision networks. And explain the **non forgetting property**
- Verify whether a **possible world satisfies a policy** and define the **expected value of a policy**
- Compute the **number of policies** for a decision problem
- **Compute the optimal policy** by Variable Elimination
- Compute value of **information** and **control**

# Next class

- Markov Decision Processes (MDPs) (textbook 9.5)

## Assignment 4 due on Wed Apr 8

- Need to talk to student 521320