Decision making in the presence of uncertainty

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

Many real-world problems require to choose future actions in the presence of uncertainty

Examples: patient management, investment decisions

Main issues:
• How to model the decision process in the computer?
• How to make decisions about actions in the presence of uncertainty?
(Stochastic) Decision tree

- Decision tree:
  - decision node
  - chance node
  - outcome (value) node

Sequential (multi-step) problems

The decision tree can be built to capture multi-step decision problems:
- Choose an action
- Observe a stochastic outcome
- And repeat

How to make decisions for multi-step problems?
- Start from the leaves of the decision tree (outcome nodes)
- Compute expectations at chance nodes
- Maximize at the decision nodes
Algorithm is sometimes called expectimax
Multi-step problem example

Assume:

- Two investment periods
- Two actions: stock and bank

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
150   125
   0.5  1.0
Bank
95   90
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

```
Stock  Bank
117   110
   0.4  0.6
Bank
95   110
   (up) (down)
```

- Notice that the probability of stock going up and down in the 2\textsuperscript{nd} step is independent of the 1\textsuperscript{st} step (=0.5)

Conditioning in the decision tree

- But this may not hold in general. In decision trees:
  - Later outcomes can be conditioned on the earlier stochastic outcomes and actions

Example: stock movement probabilities. Assume:

- \( P(1\textsuperscript{st}=\text{up})=0.4 \)
- \( P(2\textsuperscript{nd}=\text{up}|1\textsuperscript{st}=\text{up})=0.4 \)
- \( P(2\textsuperscript{nd}=\text{up}|1\textsuperscript{st}=\text{down})=0.5 \)

Tree Structure: every observed stochastic outcome = 1 branch
- \( P(1^{st}=\text{up})=0.4 \)
- \( P(2^{nd}=\text{up}|1^{st}=\text{up})=0.4 \)
- \( P(2^{nd}=\text{up}|1^{st}=\text{down})=0.5 \)

Trajectory payoffs
- Outcome values at leaf nodes (e.g. monetary values)
  - Rewards and costs for the path trajectory

Example: stock fees and gains. Assume:
Fee per period: $5 paid at the beginning
Gain for up: 15%, loss for down 10%

\[
\begin{align*}
(1000-5) \times 1.15 & = 1310.14 \\
(1000-5) \times 0.9 & = 1025.33
\end{align*}
\]
Constructing a decision tree

- The decision tree is rarely given to you directly.
  - Part of the problem is to construct the tree.

Example: stocks, bonds, bank for k periods

**Stock:**
- Probability of stocks going up in the first period: 0.3
- Probability of stocks going up in subsequent periods:
  - \( P(\text{kth step}=\text{Up}|\ (k-1)\text{th step }=\text{Up})=0.4 \)
  - \( P(\text{kth step }=\text{Up}|\ (k-1)\text{th step }=\text{Down})=0.5 \)
- Return if stock goes up: 15% if down: 10%
- Fixed fee per investment period: $5

**Bonds:**
- Probability of value up: 0.5, down: 0.5
- Return if bond value is going up: 7%, if down: 3%
- Fee per investment period: $2

**Bank:**
- Guaranteed return of 3% per period, no fee

Information-gathering actions

- Many actions and their outcomes irreversibly change the world
- **Information-gathering (exploratory) actions:**
  - make an inquiry about the world
  - Key benefit: reduction in the uncertainty
- **Example: medicine**
  - Assume a patient is admitted to the hospital with some set of initial complaints
  - We are uncertain about the underlying problem and consider a surgery, or a medication to treat them
  - But there are often lab tests or observations that can help us to determine more closely the disease the patient suffers from
  - **Goal of lab tests:** Reduce the uncertainty of outcomes of treatments so that better treatment option can be chosen
Decision-making with exploratory actions

In decision trees:
- **Exploratory actions** can be represented and reasoned about the same way as other actions.

How do we capture the effect of exploratory actions in the decision tree model?
- Information obtained through exploratory actions may affect the probabilities of later outcomes
  - Recall that the probabilities on later outcomes can be conditioned on past observed outcomes and past actions
  - Sequence of past actions and outcomes is “remembered” within the decision tree branch

Oil wildcatter problem.

An oil wildcatter has to make a decision of whether to drill or not to drill on a specific site

- **Chance of hitting an oil deposit:**
  - Oil: 40% \( P(\text{Oil} = T) = 0.4 \)
  - No-oil: 60% \( P(\text{Oil} = F) = 0.6 \)
- **Cost of drilling:** 70K
- **Payoffs:**
  - Oil: 220K
  - No-oil: 0 K

![Decision Tree Diagram](attachment://decision_tree.png)
Oil wildcatter problem.

An oil wildcatter has to make a decision of whether to drill or not to drill on a specific site

- **Chance of hitting an oil deposit:**
  - Oil: 40% \( P(Oil = T) = 0.4 \)
  - No-oil: 60% \( P(Oil = F) = 0.6 \)

- **Cost of drilling:** 70K

- **Payoffs:**
  - Oil: 220K
  - No-oil: 0 K

```
Drill

0.4

0.6

220-70=150

-70

No-drill

1.0

0
```

Oil wildcatter problem

- Assume that in addition to the drill/no-drill choices we have an option to run the **seismic resonance test**

- **Seismic resonance test results:**
  - **Closed pattern** (more likely when the hole holds the oil)
  - **Diffuse pattern** (more likely when it is empty)

\[
P(\text{Seismic resonance test} | \text{Oil})
\]

\[
\begin{array}{c|cc}
\text{Seismic resonance test pattern} & \text{closed} & \text{diffuse} \\
\hline
\text{Oil} & \text{True} & \text{False} \\
\hline
\text{True} & 0.8 & 0.2 \\
\text{False} & 0.3 & 0.7
\end{array}
\]

- **Test cost:** 10K
Oil wildcatter problem.

- Decision tree

- Compute outcomes

Oil: + 220
Drill: - 70
Test: - 10
Oil wildcatter problem.

- Compute outcomes

Oil: + 220
Drill: - 70
Test: - 10

Oil: + 220
Drill: - 70
Test: - 10
Oil wildcatter problem.

- Compute outcomes

**Oil:** +220  
**Drill:** -70  
**Test:** -10

- Compute probabilities

**Oil:** +220  
**Drill:** -70  
**Test:** -10
Oil wildcatter problem.

• Decision tree probabilities

Test 0

NoTest

P(?)

Drill

No-drill

P(Oil = T | Test = closed)

P(Oil = T | Test = closed) = ?
Oil wildcatter problem.

- Decision tree probabilities

\[ P(\text{Oil} = T \mid \text{Test} = \text{closed}) = \frac{P(\text{Test} = \text{closed} \mid \text{Oil} = T)P(\text{Oil} = T)}{P(\text{Test} = \text{closed})} = \frac{0.8 \times 0.4}{0.8 \times 0.4 + 0.6 \times 0.2} = 0.64 \]

\[ P(\text{Oil} = F \mid \text{Test} = \text{closed}) = \frac{P(\text{Test} = \text{closed} \mid \text{Oil} = F)P(\text{Oil} = F)}{P(\text{Test} = \text{closed})} = \frac{0.8 \times 0.4}{0.8 \times 0.4 + 0.6 \times 0.2} = 0.64 \]
Oil wildcatter problem.

- Decision tree probabilities

\[
P(\text{Oil} = F \mid \text{Test} = \text{closed}) = \frac{P(\text{Test} = \text{closed} \mid \text{Oil} = F)P(\text{Oil} = F)}{P(\text{Test} = \text{closed})} = \frac{0.8 \times 0.4}{0.8 \times 0.4 + 0.6 \times 0.2} = 0.36
\]

\[
P(\text{Oil} = T \mid \text{Test} = \text{closed}) = \frac{P(\text{Test} = \text{closed} \mid \text{Oil} = T)P(\text{Oil} = T)}{P(\text{Test} = \text{closed})} = \frac{0.8 \times 0.4}{0.8 \times 0.4 + 0.6 \times 0.2} = 0.64
\]

\[
P(\text{Test} = \text{closed}) = P(\text{Test} = \text{closed} \mid \text{Oil} = F)P(\text{Oil} = F) + P(\text{Test} = \text{closed} \mid \text{Oil} = T)P(\text{Oil} = T) = 0.5
\]
Oil wildcatter problem.

- Decision tree probabilities

\[ P(\text{Test} = \text{closed}) = P(\text{Test} = \text{closed} \mid \text{Oil} = F)P(\text{Oil} = F) + P(\text{Test} = \text{closed} \mid \text{Oil} = T)P(\text{Oil} = T) \]
\[ P(\text{Test} = \text{diff}) = P(\text{Test} = \text{diff} \mid \text{Oil} = F)P(\text{Oil} = F) + P(\text{Test} = \text{diff} \mid \text{Oil} = T)P(\text{Oil} = T) \]
Oil wildcatter problem.

- Decision tree

- Alternative model
Oil wildcatter problem.

- Decision tree

The presence of the test and its result affected our decision:

- if test = closed then drill
- if test = diffuse then do not drill
Value of information

• When the test makes sense?
• Only when its result makes the decision maker to change his mind, that is he decides not to drill.

• Value of information:
  – Measure of the goodness of the information from the test
  – Difference between the expected value with and without the test information

• Oil wildcatter example:
  – Expected value without the test = 18
  – Expected value with the test = 25.4
  – Value of information for the seismic test = 7.4

Using utility to measure the outcomes
Selection based on expected values

- **Until now:** The optimal action choice was the option that maximized the expected monetary value.
- **But is the expected monetary value always the quantity we want to optimize?**

![Diagram with decision nodes]

- **Answer:** Yes, but only if we are risk-neutral.

- **But what if we do not like the risk (we are risk-averse)?**
- In that case we may want to get the premium for undertaking the risk (of loosing the money)
- **Example:**
  - we may prefer to get $101 for sure against $102 in expectation but with the risk of loosing the money
- **Problem:** How to model decisions and account for the risk?
- **Solution:** use utility function, and utility theory
Utility function

- **Utility function (denoted U)**
  - Quantifies how we “value” outcomes, i.e., it reflects our preferences
  - Can be also applied to “value” outcomes other than money and gains (e.g. utility of a patient being healthy, or ill)

- **Decision making:**
  - uses expected utilities (denoted EU)
    
    $$EU(X) = \sum_{x \in \Omega_X} P(X = x)U(X = x)$$

  $$U(X = x)$$ the utility of outcome $x$

  **Important!!!**
  - Under some conditions on preferences we can always design the utility function that fits our preferences

Utility theory

- **Defines axioms on preferences** that involve uncertainty and ways to manipulate them.
- Uncertainty is modeled through **lotteries**
  - **Lottery:**
    
    $$[p : A; (1 – p) : C]$$
    
    - Outcome A with probability $p$
    - Outcome C with probability $(1-p)$
  - The following six constraints are known as the axioms of utility theory. The axioms are the most obvious semantic constraints on preferences with lotteries.

- **Notation:**
  - $\succ$ - preferable
  - $\sim$ - indifferent (equally preferable)
Axioms of the utility theory

- **Orderability**: Given any two states, a rational agent prefers one of them, else the two as equally preferable.
  \[(A \succ B) \lor (B \succ A) \lor (A \sim B)\]
- **Transitivity**: Given any three states, if an agent prefers \(A\) to \(B\) and prefers \(B\) to \(C\), the agent must prefer \(A\) to \(C\).
  \[(A \succ B) \land (B \succ C) \Rightarrow (A \succ C)\]
- **Continuity**: If some state \(B\) is between \(A\) and \(C\) in preference, then there is a \(p\) for which the rational agent will be indifferent between state \(B\) and the lottery in which \(A\) comes with probability \(p\), \(C\) with probability \((1-p)\).
  \[(A \succ B \succ C) \Rightarrow \exists p \ [p : A; (1-p) : C] \sim B\]

- **Substitutability**: If an agent is indifferent between two lotteries, \(A\) and \(B\), then there is a more complex lottery in which \(A\) can be substituted with \(B\).
  \[(A \sim B) \Rightarrow [p : A; (1-p) : C] \sim [p : B; (1-p) : C]\]
- **Monotonicity**: If an agent prefers \(A\) to \(B\), then the agent must prefer the lottery in which \(A\) occurs with a higher probability
  \[(A \succ B) \Rightarrow (p > q \iff [p : A; (1-p) : B] \succ [q : A; (1-q) : B])\]
- **Decomposability**: Compound lotteries can be reduced to simpler lotteries using the laws of probability.
  \[\ [p : A; (1-p) : [q : B; (1-q) : C]] \Rightarrow \ [p : A; (1-p)q : B; (1-p)(1-q) : C]\]
Utility theory

If the agent obeys the axioms of the utility theory, then
1. there exists a real valued function \( U \) such that:
   \[ U(A) > U(B) \iff A \succ B \]
   \[ U(A) = U(B) \iff A \sim B \]

2. The utility of the lottery is the expected utility, that is the sum of utilities of outcomes weighted by their probability
   \[ U[p : A; (1 - p) : B] = pU(A) + (1 - p)U(B) \]

3. Rational agent makes the decisions in the presence of uncertainty by maximizing its expected utility

Utility functions

We can design a utility function that fits our preferences if they satisfy the axioms of utility theory.

- But how to design the utility function for monetary values so that they incorporate the risk?
- What is the relation between utility function and monetary values?

- Assume we loose or gain $1000.
  - Typically this difference is more significant for lower values (around $100 - 1000) than for higher values (~ $1,000,000)
- What is the relation between utilities and monetary value for a typical person?
Utility functions

- What is the relation between utilities and monetary value for a typical person?
- **Concave function** that flattens at higher monetary values

Utility functions

- Expected utility of a sure outcome of 750
Utility functions

Assume a lottery \( L \) \([0.5: 500, 0.5:1000]\)

- Expected value of the lottery = 750
- Expected utility of the lottery \( EU(L) \) is different:
  - \( EU(L) = 0.5U(500) + 0.5*U(1000) \)

- Risk aversion – a bonus is required for undertaking the risk

Expected utility of the lottery \( EU(\text{lottery } L) < EU(\text{sure } 750) \)
Decision making with utility function

- Original problem with monetary outcomes

```
Stock 1
\[ \begin{array}{c}
102 \\
0.6 \\
\text{(up)} \\
110 \\
0.4 \\
\text{(down)} \\
90 \\
\end{array} \]
```

```
Stock 2
\[ \begin{array}{c}
104 \\
0.4 \\
\text{(up)} \\
140 \\
0.6 \\
\text{(down)} \\
80 \\
\end{array} \]
```

```
Bank
\[ \begin{array}{c}
101 \\
1.0 \\
\text{(up)} \\
101 \\
1.0 \\
\text{(down)} \\
\end{array} \]
```

```
Home
\[ \begin{array}{c}
100 \\
1.0 \\
\text{(up)} \\
100 \\
1.0 \\
\text{(down)} \\
\end{array} \]
```

---

Decision making with the utility function

- Utility function \( \log(x) \)

```
Stock 1
\[ \begin{array}{c}
2.00653 \\
0.6 \\
\text{(up)} \\
2.0413 \\
0.4 \\
\text{(down)} \\
1.9542 \\
\end{array} \]
```

```
Stock 2
\[ \begin{array}{c}
2.0003 \\
0.4 \\
\text{(up)} \\
2.1461 \\
0.6 \\
\text{(down)} \\
1.9030 \\
\end{array} \]
```

```
Bank
\[ \begin{array}{c}
2.004 \\
1.0 \\
\text{(up)} \\
2.0043 \\
1.0 \\
\text{(down)} \\
2.0000 \\
\end{array} \]
```

---