Reasoning and Decision-Making under Uncertainty

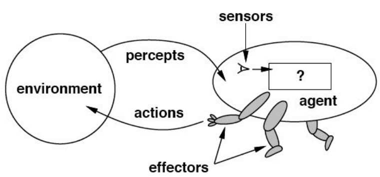
8. Session:

Probabilistic Decision-Making (1)

Prof. Dr.-Ing. Stefan Kopp
Center of Excellence „Cognitive Interaction Technology“
AG Sociable Agents

CITEC Sociable Agents

Making decisions



Classical A.I. approach (Newell 1982):

Principle of rationality:

> *"If an agent has knowledge that one of its actions will lead to one of its goals, then the rational agent will select that action."*

Making Decisions

Question: When to leave to the airport to catch a flight?

- Goal G = *catch flight, don't be too late*
- Action A_t = *leave for airport t minutes before flight*

Requires to check if an A_t gets me there on time

- purely logical reasoning won't work
  - A_{90} will get me there on time *if* there's no accident on the bridge *and* it doesn't rain *and* my tires remain intact *and* .....
  - plan success not inferable (qualification problem)
- probabilistic reasoning to deal with inherent uncertainty of assumptions about actions and events:
  - degree of belief: $Pr(A_{25} \mid \text{no reported accidents}) = 0.06$
  - can be updated as new (soft or hard) evidence comes in

Decision-making

But: Degree of belief cannot account for decision-making alone

- suppose the agent believes the following:
  - $Pr(A_{25}$ gets me there on time | …) = 0.04
  - $Pr(A_{90}$ gets me there on time | …) = 0.70
  - $Pr(A_{120}$ gets me there on time | …) = 0.95
  - $Pr(A_{1440}$ gets me there on time | …) = 0.999

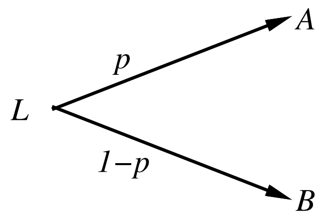
Rational decision-making must depend on two things:

1. likelihood that a goal can be achieved to a necessary degree
2. subjective and relative assessment of the goals
   - can be modeled as preferences for outcomes (risks, costs, rewards, etc.),

decision theory = probability theory + utility theory

Basics of utility theory - preferences

An agent chooses among prizes (A, B, etc.) and lotteries, i.e., situations with uncertain prizes



Lottery $L = [p, A;\ (1-p), B]$

L=[pi,Ci] ↔ outcome C_i can occur with probability p_i

Notation:

$A \succ B$	A preferred to B
$A \sim B$	indifference between A and B
$A \succsim B$	B not preferred to A

<u>Key question:</u> how to use such coarse and discrete preferences when making decisions?

From preferences to utility functions

Idea: preferences of a rational agent must obey constraints.
Rational preferences $\Rightarrow$
behavior describable as maximization of expected utility

Constraints:

<u>Orderability</u>

$(A \succ B) \vee (B \succ A) \vee (A \sim B)$

Agent cannot avoid deciding

<u>Transitivity</u>

$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$

<u>Continuity</u>

$A \succ B \succ C \Rightarrow \exists p\ [p, A;\ 1-p, C] \sim B$

Indifferent between lottery A vs. C, and getting B for sure

<u>Substitutability</u>

$A \sim B \Rightarrow [p, A;\ 1-p, C] \sim [p, B; 1-p, C]$

Lotteries with comparable prizes are comparable

<u>Monotonicity</u>

$A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A;\ 1-p, B] \succsim [q, A;\ 1-q, B])$

From preferences to utility functions

Rationally constrained preferences are a basic property of rational agents. Then, the existence of a utility function follows from two principles:

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944):
Given preferences satisfying the constraints
there exists a real-valued function U such that

$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$ — Rational Utility

$U([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i p_i U(S_i)$ — Expected Utility

That is, a continuous utility function can be formulated in accord with the preferences. This allows for modeling decision-making as utility maximization.

➔ Acting rationally when trying to maximize the expected utility

Problem often: The preferences/utilities need to be defined carefully

Example

You are asked if you wish to take a bet on the outcome of tossing a fair coin. If you bet and win, you gain £100. If you bet and lose, you lose £200. If you don't bet, the cost to you is zero.

$$U(\text{win, bet}) = 100 \qquad U(\text{lose, bet}) = -200$$
$$U(\text{win, no bet}) = 0 \qquad U(\text{lose, no bet}) = 0$$

Our expected winnings/losses are:

$$U(\text{bet}) = p(\text{win}) \times U(\text{win, bet}) + p(\text{lose}) \times U(\text{lose, bet})$$
$$= 0.5 \times 100 - 0.5 \times 200 = -50$$

$$U(\text{no bet}) = 0$$

Based on taking the decision which maximises expected utility, we would therefore be advised not to bet.

Example

You have $\pounds 1,000,000$ in your bank account. You are asked if you would like to participate in a fair coin tossing bet in which, if you win, your bank account will become $\pounds 1,000,000,000$. However, if you lose, your bank account will contain only $\pounds 1000$. Should you take the bet?

$$U(\text{bet}) = 0.5 \times 1,000,000,000 + 0.5 \times 1000 = 500,000,500.00$$

$$U(\text{no bet}) = 1,000,000$$

Based on expected utility, we are therefore advised to take the bet.

- In reality few people who are millionaires are likely to be willing to risk losing almost everything in order to become a billionaire.
- This means that the subjective utility of money is not simply the quantity of money.

Probabilistic decision-making

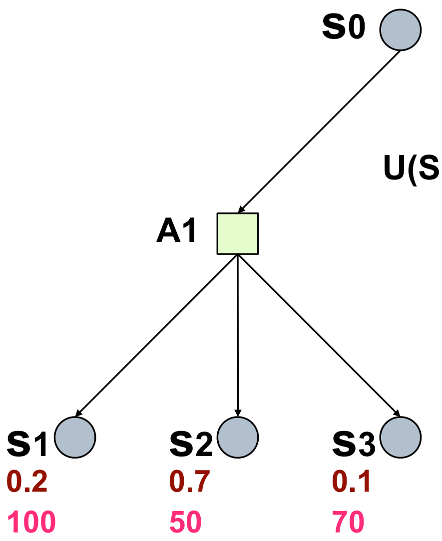
Principle of maximum expected utility (MEU)
An agent is rational *iff* it chooses the action that yields the highest expected utility, averaged over all possible outcomes of the action

Note: a non-deterministic action can have several outcomes $Result_i(A)$

Prior to executing A, the agent needs to...

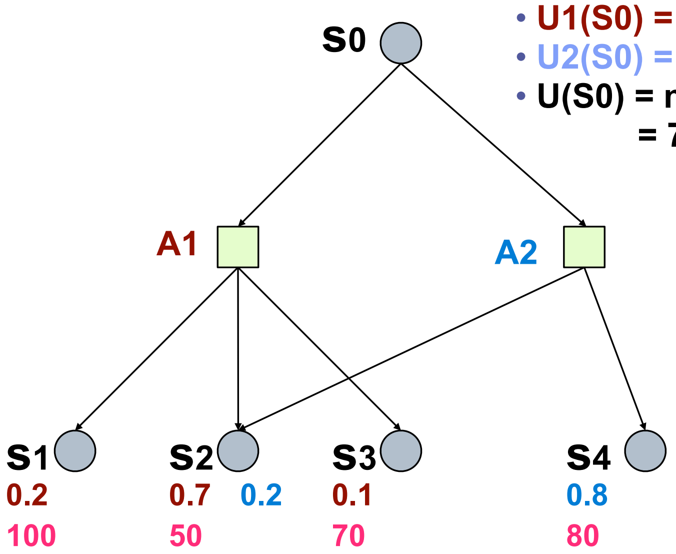
1. determine the probabilities $P(Result_i(A)|Do(A),E)$
2. calculate the expected utility of A, given evidence E:
   $EU(A|E)=\sum_i P(Result_i(A)|Do(A),E)\ U(Result_i(A))$
   with $U(S)$ utility function of state S
3. pick action A that maximizes the EU

One State/One Action Example



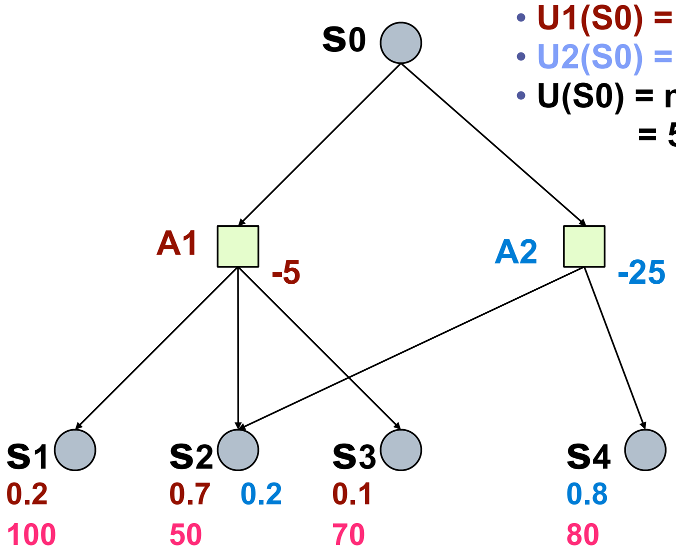
U(S0) = 100 x 0.2 + 50 x 0.7 + 70 x 0.1
= 20 + 35 + 7
= 62

One State/Two Actions Example



- U1(S0) = 62
- U2(S0) = 74
- U(S0) = max{U1(S0),U2(S0)}
  = 74

Introducing Action Costs



- U1(S0) = 62 – 5 = 57
- U2(S0) = 74 – 25 = 49
- U(S0) = max{U1(S0),U2(S0)} = 57

Using causal models for decision-making

Possible consequences of actions (when executed under given information about the world) and their probabilities can be inferred using causal models.

- specify how observations change degrees of belief about causally linked state variables: $P(x_j \mid x_i)$
- thus represent evidential causal structure

Caution: Be careful when using causal models for decision-making!

- need to infer belief states that result from (possible) actions
- but actions are interventions, i.e. external changes to the world, not brought about by causal links
- knowledge about world state due to actions differs from knowledge due to some sort of evidence

Interventions

Decision-making includes assuming to take actions. This changes the causal model of the world -- it doesn't behave as normally observed!

➔ Interventions: formalized using the *do(x)* operator

- simplest („atomic") intervention: variable forced to take a value
  - *do(Xi = xi)* or simply *do(xi)*
  - sets variable *Xi*, leaving all other mechanisms untouched,
  - PLUS new network structure with all links from *Pa(Xi)* to *Xi* pruned
- new model yields causal effect of *Xi* on *Xj* when solved for distribution of *Xj*
  - we write: *P(xj | ^xi)* or *P(xj | do(xi))*

Decision theory

Can no longer use „evidential decision theory"

- treats actions as ordinary events and encourages decision-making based on evidence an action would provide
- maximizes EU based on $U(x) = \sum_y P(y|X_i = x, E)U(y)$

Instead, we need to use a „causal decision theory"

- instructs agents to choose *x* that maximizes EU based on

$$U(x) = \sum_y P(y|do(X_i = x), E)U(y)$$

- action-specific conditionalization for any further evidence
- derive degrees of beliefs from new causal model

How to determine causal effects?

Assume: Causal diagram G, wish to estimate effect of *do*(X=x) on a set of variables **Y**, where **X** and **Y** are both subsets of a variable set **V**

- P(y|^x)=? from a sample estimate of P(v)

Back-door criterion (BDC): Variables **Z** are „back-doors" for the causal influence of Xi on Xj, if

(I) not descendants of Xi in **Z**

(II) **Z** blocks (d-seperates) every path between Xi, Xj that contains arrow into Xi

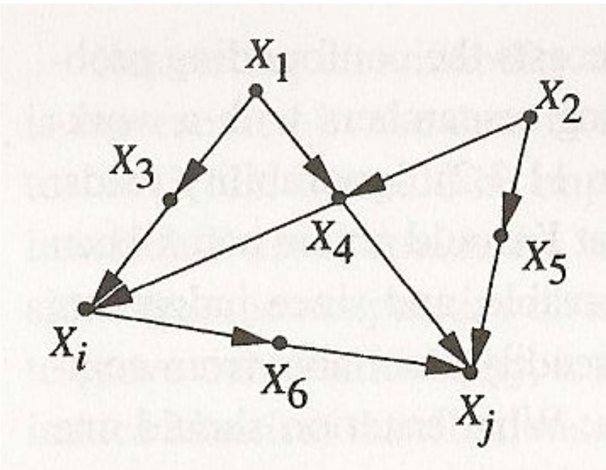
That is, **Z** captures the spurious (back-door) paths. If **Z** satisfies BDC relative to (**X**,**Y**) then

$$P(y|\hat{x}) = \sum_z P(y|x,z)P(z)$$

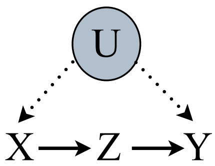
Causal effects

Example: Back-Door Criterion

- **Z1**={X3,X4}, **Z2**={X4,X5} meet the BDC
- **Z3**={X4} not, because the path (Xi,X3,X1,X4,X2,X5,Xj) is not blocked, i.e. even if X6 and X4 given, Xi and Xj are still dependent.



Causal effects



Front-Door Criterion

- Z does not satisfy BDC
- measuring Z can enable consistent estimate of P(y|^x)
  - P(x,y,z,u)=P(u)P(x|u)P(z|x)P(y|z,u) ➔ intervention do(x)
  - P(y,z,u|^x)=P(y|z,u)P(z|x)P(u) ➔ summing over z and u
  - $P(y|\hat{x}) = \sum_z P(z|x) \sum_u P(y|z,u)P(u)$
- Can eliminate all unobserved variables from r.h.s using

$$\sum_u P(y|z,u)P(u) = \sum_x \sum_u P(y|z,u)P(u|x)P(x)$$

$$= \sum_x \sum_u P(y|x,z,u)P(u|x,z)P(x) = \sum_x P(y|x,z)P(x)$$

- So, can estimate effect of X on Y via mediating („front-door") variable Z

$$P(y|\hat{x}) = \sum_z P(z|x) \sum_{x'} P(y|x',z)P(x')$$

Causal effects

Front-door criterion (FDC): Variables **Z** are „front-doors" for the causal influence of X on Y, if

(I) **Z** intercepts *all* direct paths from X to Y

(II) all back-door paths from X to **Z** are blocked

(III) all back-door paths from **Z** to Y are blocked by X

- captures the direct (front-door) paths via two-step application of back-door paths between X and **Z**, and between **Z** and Y

If **Z** satisfies FDC relative to (X,Y) *and* P(x,z)>0, then causal effect of X on Y is given by

$$P(y|\hat{x}) = \sum_z P(z|x) \sum_{x'} P(y|x',z)P(x')$$

Example: smoking & genotype theory

What about smoking (X) and lung cancer (Y)?

- smoking industry: „cancer solely due to some genotype (U)“
- does amount of tar in lung (Z) satisfy the FDC ?
  - condition (i): assume that smoking has no effect on lung cancer except through tar deposit
  - condition (ii)+(iii): assume that genotype has no effect on tar deposit (except indirectly through smoking), and no other factor affecting tar deposit has influence on smoking
  - condition P(x,z)>0: high levels of tar are not only the result of smoking, but also of other possible factors, and tar maybe absent in some smokers

➔ How to assess whether smoking increases risk of cancer?

Example

Hypothetical & unrealistic data set

	Group Type	$P(x, z)$ Group Size (% of Population)	$P(Y = 1 \mid x, z)$ % of Cancer Cases in Group
$X = 0, Z = 0$	Nonsmokers, No tar	47.5	10
$X = 1, Z = 0$	Smokers, No tar	2.5	90
$X = 0, Z = 1$	Nonsmokers, Tar	2.5	5
$X = 1, Z = 1$	Smokers, Tar	47.5	85

Probability that a random person will get cancer when non-/smoking?

➔ may use FDC along with „do-calculus“

$$P(Y = 1 \mid do(X = 1)) = .05(.10 \times .50 + .90 \times .50) + .95(.05 \times .50 + .85 \times .50) = .05 \times .50 + .95 \times .45 = .4525,$$

$$P(Y = 1 \mid do(X = 0)) = .95(.10 \times .50 + .90 \times .50) + .05(.05 \times .50 + .85 \times .50) = .95 \times .50 + .05 \times .45 = .4975.$$

Probabilistic decision-making

Problem: How to determine the optimally achievable expected utility together with the corresponding action (action sequence)?

Complex Problem -- even if you had a complete model of actions, states & utilities, quickly becomes computationally intractable. Truly rational agents take into account the *utility/costs of reasoning* as well.

- Bounded rationality (Simon, 1957)

Nevertheless, great progress has been made and we are able to solve much more complex decision-theoretic problems than ever before

- Decision Trees -- ordered action sequences
- Bayesian Decision Networks -- partially ordered action sequences
- Markov Decision Processes -- general orderless action policies

Decision Trees

Consider the decision problem as to whether or not to go ahead with a fund-raising garden party. If we go ahead with the party and it subsequently rains, then we will lose money (since very few people will show up). If we don't go ahead with the party and it doesn't rain we're free to go and do something else fun.

$$p(Rain = \text{rain}) = 0.6, \; p(Rain = \text{no rain}) = 0.4$$

$$U(\text{party}, \text{rain}) = -100, \quad U(\text{party}, \text{no rain}) = 500$$

$$U(\text{no party}, \text{rain}) = 0, \quad U(\text{no party}, \text{no rain}) = 50$$

Should we go ahead with the party? Since we don't know what will actually happen to the weather, we compute the expected utility of each decision:

$$U(\text{party}) = \sum_{Rain} U(\text{party}, Rain)p(Rain) = -100 \times 0.6 + 500 \times 0.4 = 140$$

$$U(\text{no party}) = \sum_{Rain} U(\text{no party}, Rain)p(Rain) = 0 \times 0.6 + 50 \times 0.4 = 20$$

Based on expected utility, we are therefore advised to go ahead with the party.

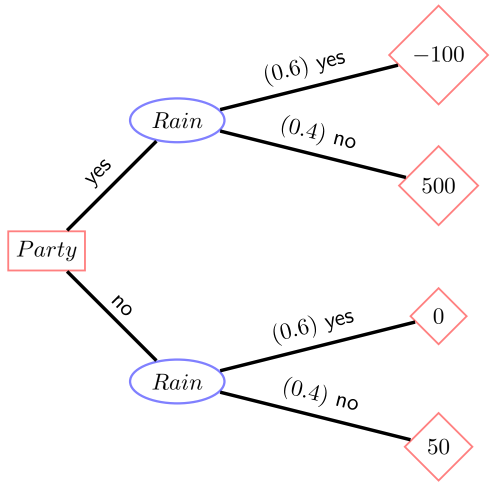
Decision Trees

Build a tree containing

- chance nodes
- decision nodes
- utility nodes

Explicit enumeration of possible choices and their order (action sequences):

- begin with leftmost choice
- probabilities on links out of chance nodes

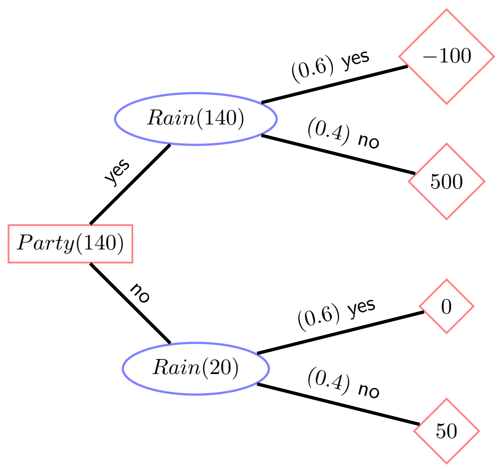


Solving a Decision Tree

Work backwards from leaves:

- chance parent is expectation of its children
- decision parent is max of its children

$$-100 \times 0.6 + 500 \times 0.4 = 140$$
$$0 \times 0.6 + 50 \times 0.4 = 20$$
$$\max(140, 20) = 140$$



Party-Friend Problem

If we decide not to go ahead with the party, we will consider going to visit a friend. In making the decision not to go ahead with the party we have utilities

$$U_{party}(\text{no party}, \text{rain}) = 0, \quad U_{party}(\text{no party}, \text{no rain}) = 50$$

$$U_{party}(\text{party}, \text{rain}) = -100, \quad U_{party}(\text{party}, \text{no rain}) = 500$$

$$p(Rain = \text{rain}) = 0.6, \quad p(Rain = \text{no rain}) = 0.4$$

The probability that the friend is in depends on the weather according to

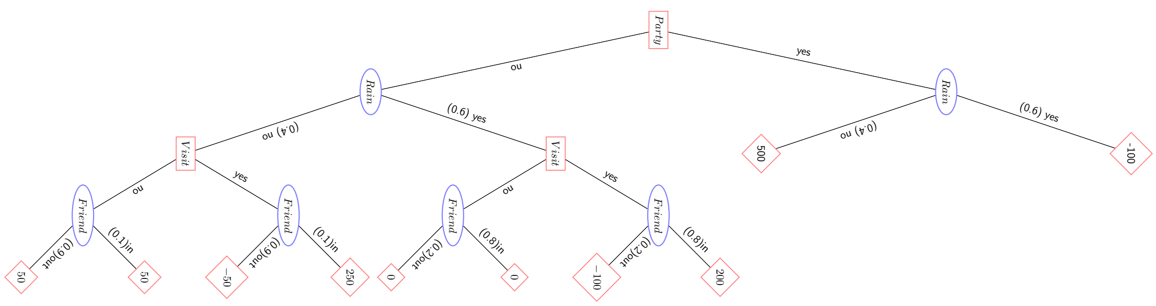
$$p(Friend = \text{in}|\text{rain}) = 0.8, \quad p(Friend = \text{in}|\text{no rain}) = 0.1,$$

The other probabilities are determined by normalisation. We additionally have

$$U_{visit}(\text{friend in}, \text{visit}) = 200, \quad U_{visit}(\text{friend out}, \text{visit}) = -100$$

with the remaining utilities zero. The two sets of utilities add up so that the overall utility of any decision sequence is $U_{party} + U_{visit}$.

Party-Friend Problem



Decision Trees can get unwieldily even for simple problems.

Need a more compact graphical representation...