Bayesian Networks 2: Bayesian Decision Making under Uncertainties

Annett Anders

IBK, ETH Zurich

Contents

Introduction

- Establishing a consistent decision basis such that the overall life cycle benefit of the facilities are maximized and the given requirements to the safety of personnel and environment specified by legislation or society are fulfilled.

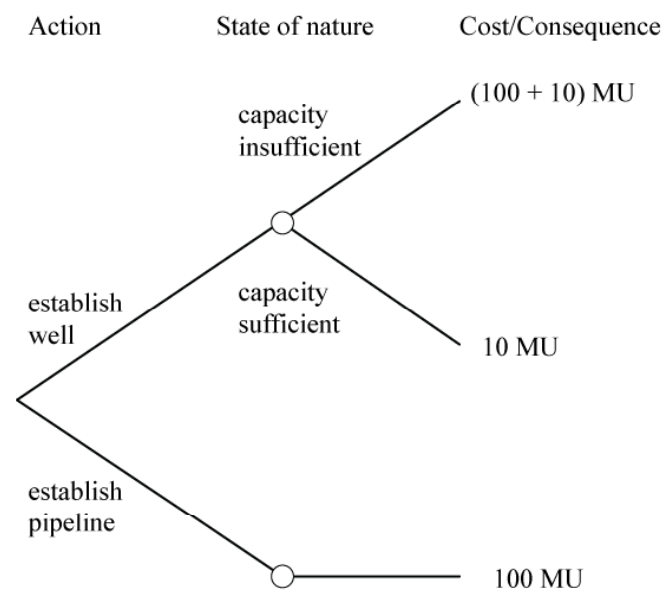
- As the available information is incomplete or uncertain $\Longrightarrow$ decision problem (DP) is a decision problem subject to uncertain information.

- The presentation introduces some fundamental issues of decision making subject to uncertain information and considers general aspects of decision theory illustrating these by using a simple example. Finally the risk analysis decision problem is defined in general terms within the context of decision theory.

Decision / Event Trees (DT)

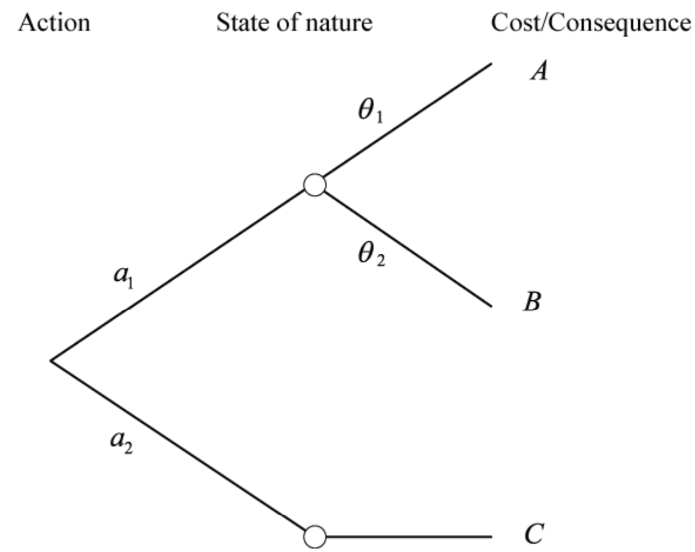
- If the number of alternative actions are extremely large, a framework for the systematic analysis of the corresponding consequences is needed.

- *Example:*
  - A production facility needs about 100 units of water per day
  - Known is that the underground contains a water reservoir but its capacity is uncertain.
  - Another option is to get the water from another location where a suitable well exists.



- Analyzing such DP in a way making consistent use of all information available, including her *degree of belief* in the possible states, her subsequent observed data and her *preferences* among the various possible action/state pairs.

Decision based on expected values



- Choose between actions a_1 and a_2. The consequence of a_2 is C with certainty whereas it is uncertain for a_1. The state of nature may be θ_1, in which case the consequence is A and the state of nature may be θ_1 in which case the consequence is B.

- Preferences expressed by any function u such that:
$$u(B) > u(C) > u(A)$$

- Find a particular function u (the *utility function)* such that it is logically consistent to decide between a_1 and a_2 by comparing $u(C)$ with the expected utility of the action a_1, namely:
$$pu(A) + (1 - p)u(B) \qquad \text{where } p = P[\theta].$$

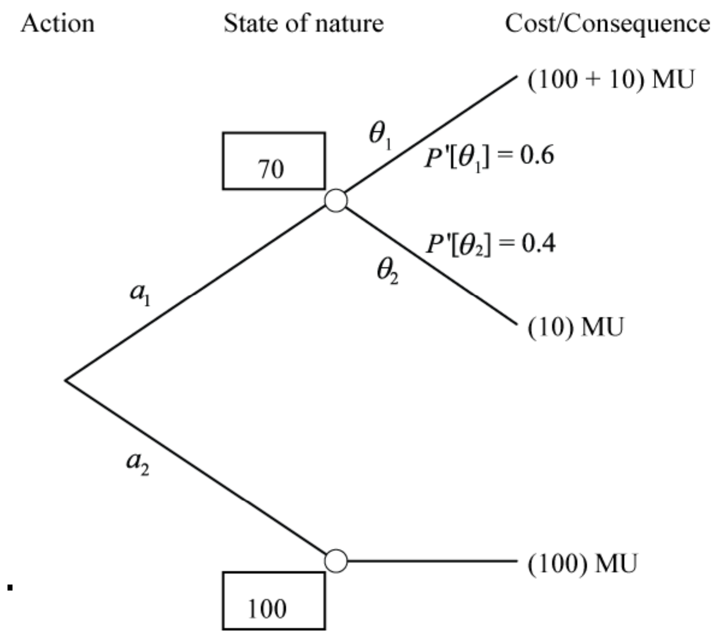
- Assuming that $u(A)$ and $u(B)$ have been given appropriate ⟹ choose $u(C)$ such that the expected utility is a valid decision criterion, i.e.
$$u(C) = p^* u(A) + (1 - p^*)u(B) \qquad \text{where } p^* \text{ is the indifference probability.}$$

Decision Making Subject to Uncertainty

- Having formulated the decision problem in terms of a decision/event tree, with proper assignment of utility and probability structure, the numerical evaluation of decision alternatives may be performed.

- Depending on the state of information at the time of the decision analysis, three different analysis types are distinguished, namely *prior analysis, posterior analysis* and *pre-posterior analysis*. Each of these are important in practical applications of decision analysis and are therefore discussed briefly in the following.

1) Decision Analysis with Given information – Prior Analysis



- The utility represented through the costs ➔ optimal decisions identified as the decisions minimizing expected costs.

- At this stage the probabilistic description $P[\theta]$ of the state of nature θ
  ➔ *prior probability* $P'[\theta]$.

- Choosing between two actions:
  a_1: Establish a new well.
  a_2: Establish a pipeline from an existing well.

- The possible states of nature:
  θ_1: Capacity insufficient , θ_2: Capacity sufficient

- Based on the prior information the expected cost $E'[C]$ amounts to:

$$E'[C] = \min\{(100+10)P'[\theta_1] + 10P'[\theta_2];100\} = \min\{70;100\} = 70$$

- It is seen that action alternative a_1 yields the smallest expense (largest expected utility) so this action alternative is the optimal decision.

2) Decision Analysis with Additional Information – Posterior Analysis

- Additional information becomes available ➔ update probability structure in the DP ➔ the decision analysis is unchanged in comparison to the situation with given - prior information.

- Given the result of an experiment z_k the *posterior probability* is denoted $P''[\theta]$ and may be evaluated by use of the Bayes' rule:

Likelihood of the result z_k given θ_i

Prior probability of θ_i

$$P''[\theta_i \mid z_k] = \frac{P\left[z_k \middle| \theta_i\right] P'[\theta_i]}{\sum_j P\left[z_k \middle| \theta_j\right] P'\left[\theta_j\right]}$$

Normalising constant

Back to the example:

- Assuming that information about the capacity of the local reservoir can be estimated using a less expensive test (cost of test = 1MU).
- Assuming the pump test provides indicators regarding the capacity of the local reservoir
- The capacity (units per day) of the reservoir is:
  - I_1 : larger than the given production requirements by 5%
  - I_2 : less than 95% of the required water production
  - I_3 : between 95 and 105 water units.
- The likelihood of the true capacity of the reservoir given the trial pump test results is given by the following table

	True capacity of the reservoir	
Indicators	θ_1 : less than 100	θ_2 : more than 100
I_1 : capacity >105	0.1	0.8
I_2 : capacity <95	0.7	0.1
I_3 : $95 \leq$ capacity ≤ 105	0.2	0.1

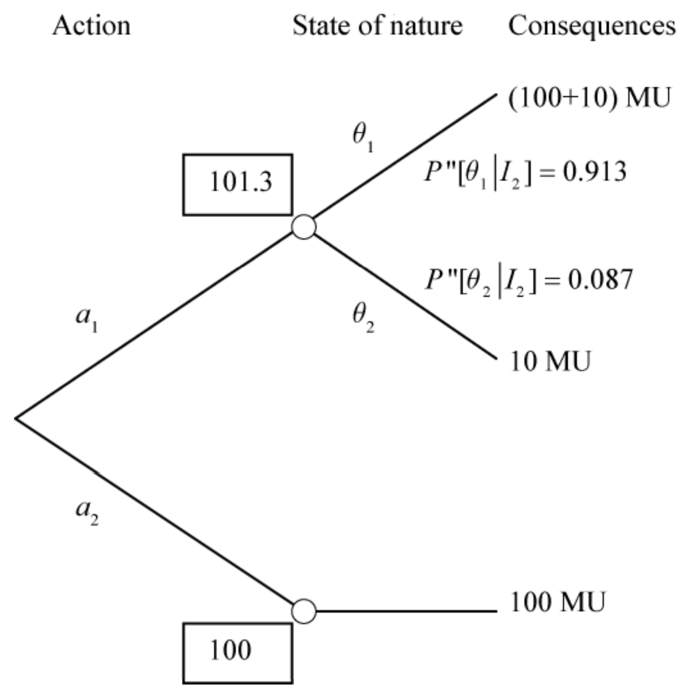
- The posterior probabilities given the new information:

$$P''[\theta_1 \mid I_2] = \frac{P[I_2 \mid \theta_1]P'[\theta_1]}{P[I_2 \mid \theta_1]P'[\theta_1] + P[I_2 \mid \theta_2]P'[\theta_2]}$$

$$= \frac{0.7 \cdot 0.6}{0.7 \cdot 0.6 + 0.1 \cdot 0.4} = \frac{0.42}{0.46} = 0.913$$

$$P''[\theta_2 \mid I_2] = \frac{P[I_2 \mid \theta_2]P'[\theta_2]}{P[I_2 \mid \theta_1]P'[\theta_1] + P[I_2 \mid \theta_2]P'[\theta_2]}$$

$$= \frac{0.1 \cdot 0.4}{0.7 \cdot 0.6 + 0.1 \cdot 0.4} = \frac{0.04}{0.46} = 0.087$$



⟹ the posterior expected values of the utility corresponding to the optimal action alternative:

$$E''[C \mid I_2] = \min\{P''[\theta_1 \mid I_2] \cdot (100+10) + P''[\theta_2 \mid I_2] \cdot 10;\ 100\}$$

$$= \min\{101.3; 100\} = 100 \text{ MU}$$

⟹ Considering the additional information the optimal decision has been switched to a_2 .

3) Decision Analysis with Additional Information – Pre-posterior Analysis

- 'Buy' additional information through an experiment before making her choice of action ➔ Comparing the cost of information to the potential value of the information (if different types of experiments are possible choose the one yielding the overall largest utility)

Back to the example:

- Decide whether or not to perform the trial pump tests
- Expected costs: $E\left[C\right]=\sum_{i=1}^{n}P'\left[I_i\right]E''\left[C\left|I_i\right.\right]=\sum_{i=1}^{n}P'\left[I_i\right]\min_{j=1,\dots,m}\{E''\left[C(a_j)\left|I_i\right.\right]\}$

  n... Number of possible experiments findings

  m... Number of different decision alternatives

  And $$P'\left[I_i\right]=P\left[I_i\left|\theta_1\right.\right]P'\left[\theta_1\right]+P\left[I_i\left|\theta_2\right.\right]P'\left[\theta_2\right]$$

⇨ $P'\left[I_1\right]=0.38$ ⇨ $E''\left[C\,|\,I_1\right]=25.8$ MU

$P'\left[I_2\right]=0.46$ $E''\left[C\,|\,I_2\right]=100$ MU

$P'\left[I_3\right]=0.16$ $E''\left[C\,|\,I_3\right]=85$ MU

- The expected cost corresponding to the situation where the experiment with the experiment costs C_P:

$$E\left[C\right]=E''\left[C\mid I_1\right]P'[I_1]+E''\left[C\mid I_2\right]P'[I_2]+E''\left[C\mid I_3\right]P'[I_3]$$
$$=(25.8+C_P)\cdot 0.38+(100+C_P)\cdot 0.46+(85+C_P)\cdot 0.16$$
$$=(69.4+C_P)\,\mathrm{MU}$$

- Comparing the results

$$E'\left[C\right]-E\left[C\right]=70-(69.4+C_P)=0.6-C_P$$

➔ The experiment should be performed if $C_P \leq 0.6$

The Risk Treatment Decision Problem

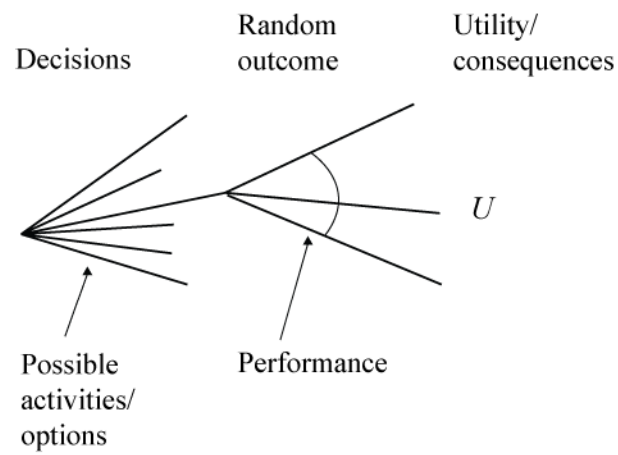
Prior decision analysis

- Simplest form of the risk analysis
- Risk is evaluated on the basis of statistical information and probabilistic modelling available prior to any decision and/or activity

$$R = E\left[U\right] = \sum_{i=1}^{n} P_i C_i$$

P_i i^{th} branching probability

- Basis for comparison of risks between different activities



Posterior decision analysis

- Same form as *prior analysis*, but changes in the branching probabilities and/or the consequences in the decision/event tree
  $\Longrightarrow$ effect of risk reducing measures, risk mitigating measures and/or collection of additional information

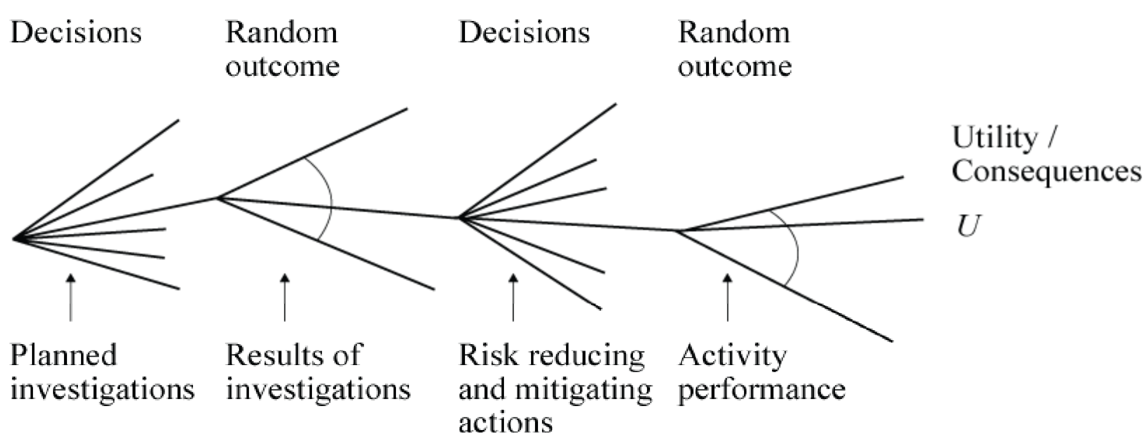
- Used to evaluate the effect of activities, which have been performed

Pre-posterior decision analysis

- Used for optimal decisions in regard to activities that may be performed in the future
- Decision rules formulated for specifying the future actions on the basis of the results of the planned activities
- optimal investigation a^* is identified through:

$$\min_a E'_Z\left[E''_Z[C(a(z),z)]\right] = \min_a E'_Z\left[\sum_{i=1}^{n} P_i''(a(z),z)C_i(a(z))\right]$$

$a(z)$ different possible actions considering the result of investigation z



Influence Diagrams (ID)

= directed acyclic graph (DAG) over chance nodes, decision nodes, and utility nodes, with

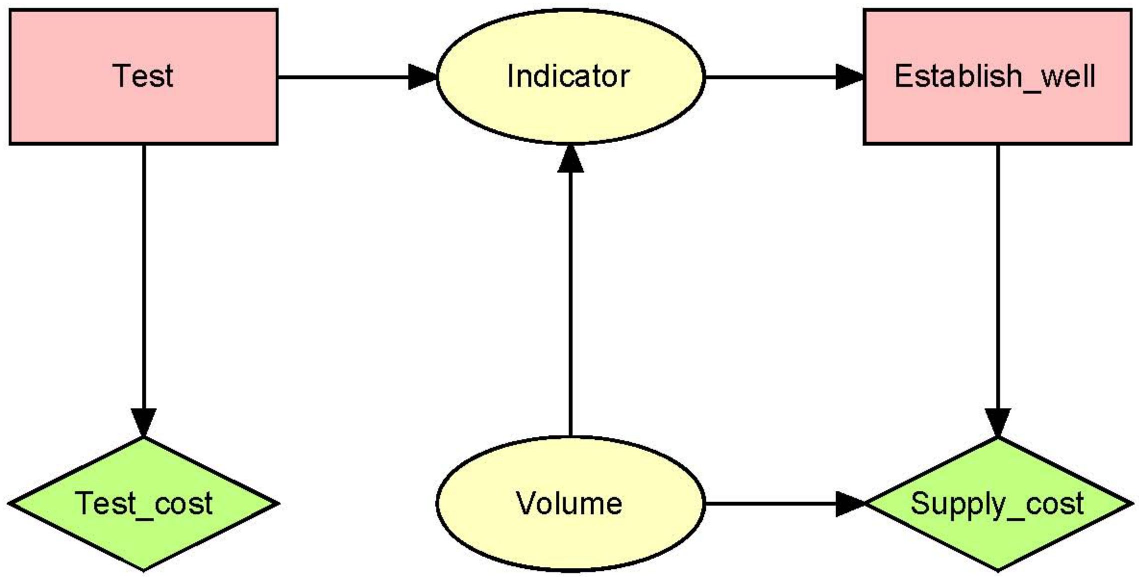
- There is a directed graph comprising all decision nodes;
- Utility nodes have no children;
- Decision nodes and chance nodes have finite set of mutually exclusive states;
- Utility nodes have no state;

An ID is *realized* when

- To each chance node A there be attached a conditional probability table $P[A \mid pa(A)]$;
- To each utility node V there be attached a real-valued function over $pa(V)$.

Jensen, F.V. and Nielsen, D.N. (2007) Bayesian Networks and Decision Graphs second edition, Springer

Example with HUGIN



Solving an ID

Definition: A *policy* for decision D_i is a mapping δ_i that for any configuration of the past of D_i yields a decision for D_i. That is,

$$\delta_i(I_0, D_1, \dots, D_{i-1}, I_{i-1}) \in sp(D_i)$$

A *strategy* for a ID is a set of policies, one for each decision. A *solution* to a ID is a strategy maximizing the expected utility.

- IDs well suited for symmetric decision problems
- Transforming the ID into a DT in order to solve it ➔ complexity problem
- Solution methods, e.g. (see Jensen and Nielsen (2007), Chapter 10)
  - Variable Elimination (strong junction trees)
  - Node Removal (removal of barren nodes, of chance nodes, of decision nodes and *arc reversal*)

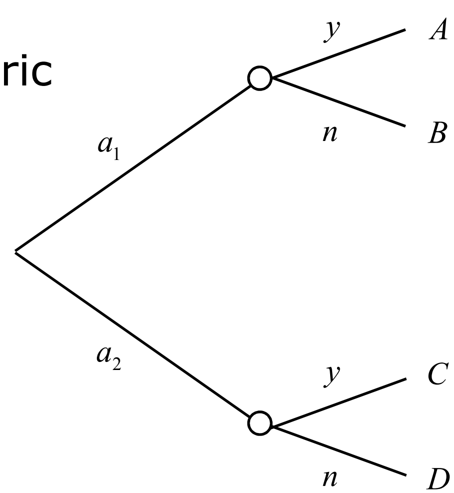
Asymmetric Decision Problems (ADP)

Definition: A decision problem is said to be *symmetric* if:

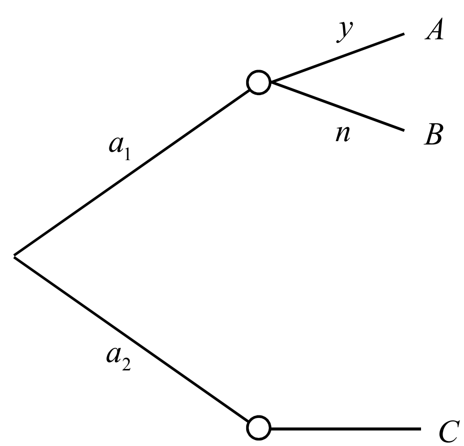
- in all of its DT representations, the number of scenarios is the same as the cardinality of the Cartesian product of the state space of all chance and decision variables, and
- in at least one decision tree representation, the sequence of chance and decision variables is the same in all scenarios.

⇒ Possible outcomes and decision options for a variable do not depend on previous observations and decisions

symmetric



asymmetric



- Using test decisions is a frequent cause of asymmetry
- No special treatment for test decisions ⟹ ordinary decision variables
- 3 types of asymmetry:
  - *Functional:* possible outcomes/decision depend on the past
  - *Structural:* the vary occurrence of an observation/decision depend on the past
  - *Order:* ordering of observation/decision not settled when specified

⟹ Unconstrained influence diagrams (UID)

Definition: An *UID* is a DAG over decision, chance and utility variables. Utility variables have no children. There are 2 types of chance variables, observables and nonobservables. A nonobservable cannot have a decision as a child.
Let U be a UID. The set of decision variables is denoted by D_U, and than the set of observables is denoted by O_U. The partial temporal order induced by U is denoted by $\prec_U$.

- Used to model order asymmetry ➔ no total ordering of the decisions
- An observable chance variable is released (for observation) when all its antecedent decision variables have been decided on
- Solving a UID: find the next action and find an optimal policy if the next action is a decision
- Solution is specified in term of an S-DAG:

Definition: Let *U* be a UID. An *S-DAG* is a DAG *G*. The nodes are labeled with variables from $D_U \cup O_U$ such that each maximal directed graph in *G* represents an admissible ordering of $D_U \cup O_U$.

Notation: - *Source node* = only node with no parents
- *Sink node* = only node with no child

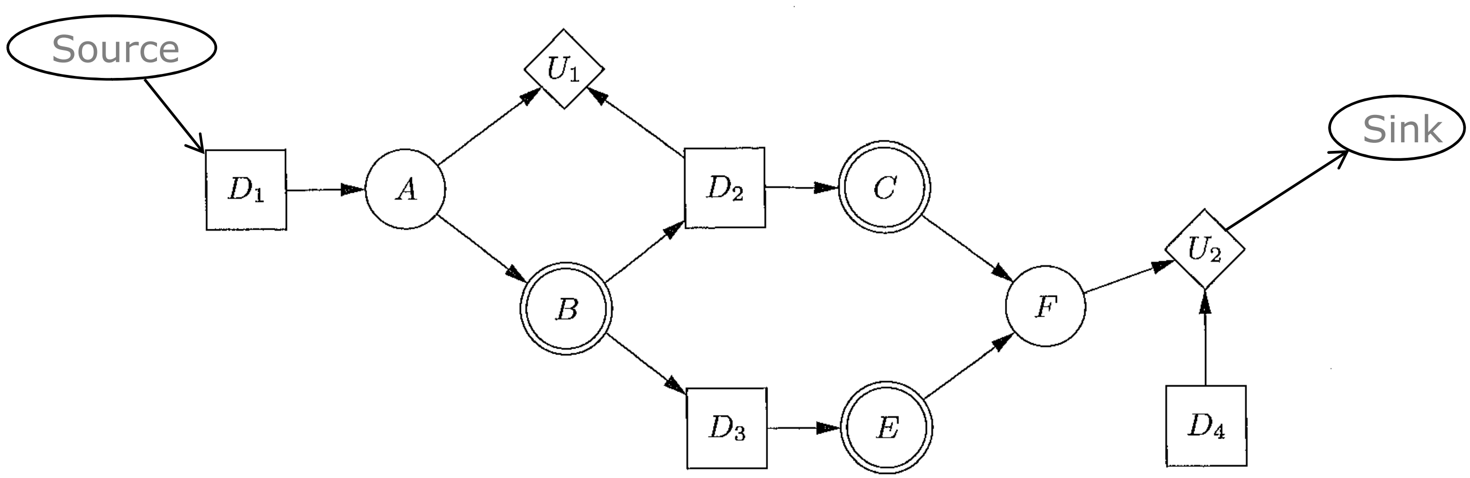


Fig. 9.35. An example UID.

The structure of a strategy for the UID

Definitions:

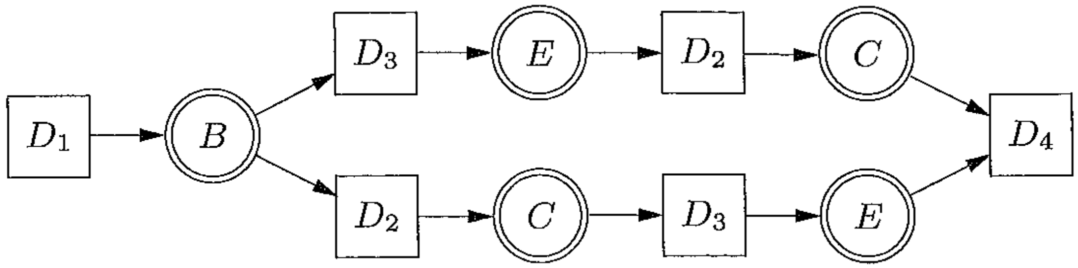
- A *step policy* for a node *N* in an S-DAG *G* is a function

$$\sigma: \ \mathrm{sp}(\mathrm{hst}(N)) \rightarrow \mathrm{ch}(N)$$

- A *step strategy* for *U* is a pair *(G,S),* where *G* is an S-DAG and *S* is a set of step policies, one for each node in *G* (except for the Sink). A policy for *N* is function

$$\delta: \ \mathrm{sp}(past(N)) \rightarrow \mathrm{ch}(N)$$

- A *strategy* for *U* is a step strategy together with a policy for each node.



$$\varnothing : choose\ option\ d_1^1$$

$$D_1, B : choose \begin{cases} d_2^3 & if\ B = b_1 \\ d_1^2 & if\ B = b_2 \end{cases}$$

$$D_1, B, D_3, E : choose \begin{cases} d_2^2 & if\ E = e_1 \\ d_1^2 & if\ E = e_2 \end{cases}$$

$$D_1, B, D_2, C : choose \begin{cases} d_1^3 & if\ C = c_1 \\ d_2^3 & if\ C = c_2 \end{cases}$$

This strategy can be unfolded to the following strategy tree

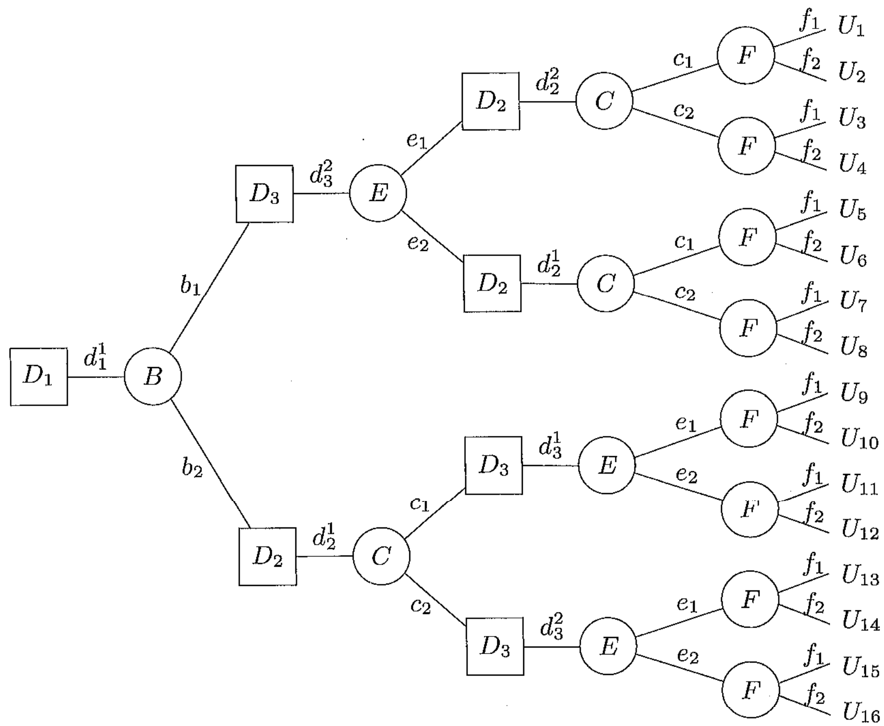


Fig. 9.37. The strategy from Figure 9.36 unfolded to a strategy tree.

➔ Rather than trying out all possible strategy trees in looking for an optimal strategy, there are efficient solution algorithms that exploit dynamic programming and work on a (single) S-DAG representation of the UID.

Summery

- Introduction to decion making under uncertainties with prior, posterior and pre-posterior analysis in the decision tree model
- Generalization of Bayesian Network ➔ Influence Diagrams, which are well-suited for so-called symmetric decision problems
- For asymmetric decision problems ➔ Introducion of Unconstrained Influence Diagrams

Thank you for your attention ☺