Two applications of Bayesian networks

Jiří Vomlel

Laboratory for Intelligent Systems
University of Economics, Prague

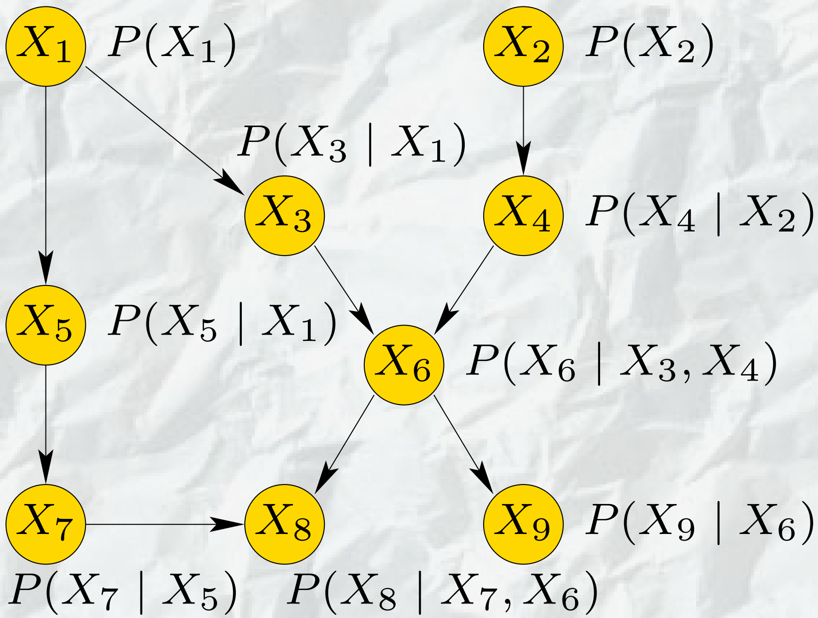
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

This presentation is available at:
`http://www.utia.cas.cz/vomlel/`

Contents:

- Bayesian networks as a model for reasoning with uncertainty
- Building probabilistic models
- Building “good” strategies using the models
- Application 1: Adaptive testing
- Application 2: Decision-theoretic troubleshooting

An example of a Bayesian network:

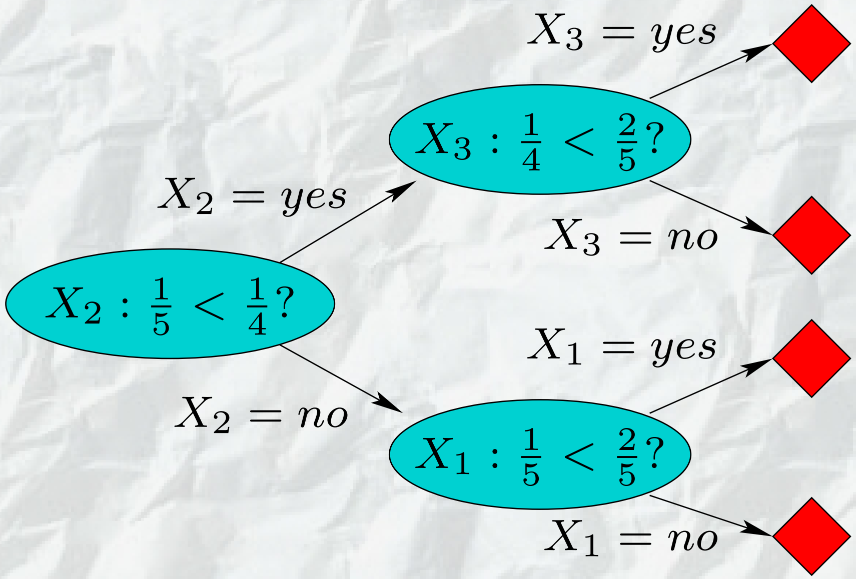


Building Bayesian network models

three basic approaches

- Discussions with domain experts: expert knowledge is used to get the structure and parameters of the model
- A dataset of records is collected and a machine learning method is used to to construct a model and estimate its parameters.
- A combination of previous two: e.g. experts helps with the stucture, data are used to estimate parameters.

An example of a strategy:



X_3 is more difficult question than X_2 which is more difficult than X_1.

Building strategies using the models

For all terminal nodes $\ell \in \mathcal{L}(\mathbf{s})$ of a strategy $\mathbf{s}$ we define:

- steps that were performed to get to that node (e.g. questions answered in a certain way). It is called collected evidence $\mathbf{e}_\ell$.
- Using the probabilistic model of the domain we can compute probability of getting to a terminal node $P(\mathbf{e}_\ell)$.
- Also during the process, when we have collected certain evidence $\mathbf{e}$ we can update the probability of getting to a terminal node, which now corresponds to conditional probability $P(\mathbf{e}_\ell)$

Building strategies using the models

For all terminal nodes $\ell \in \mathcal{L}(\mathbf{s})$ of a strategy $\mathbf{s}$ we have also defined:

- an evaluation function $f : \cup_{\mathbf{s} \in \mathcal{S}} \mathcal{L}(\mathbf{s}) \mapsto \mathbb{R}$.

For each strategy we can compute:

- expected value of the strategy:

$$E_f(\mathbf{s}) = \sum_{\ell \in \mathcal{L}(\mathbf{s})} P(\mathbf{e}_\ell) \cdot f(\mathbf{e}_\ell)$$

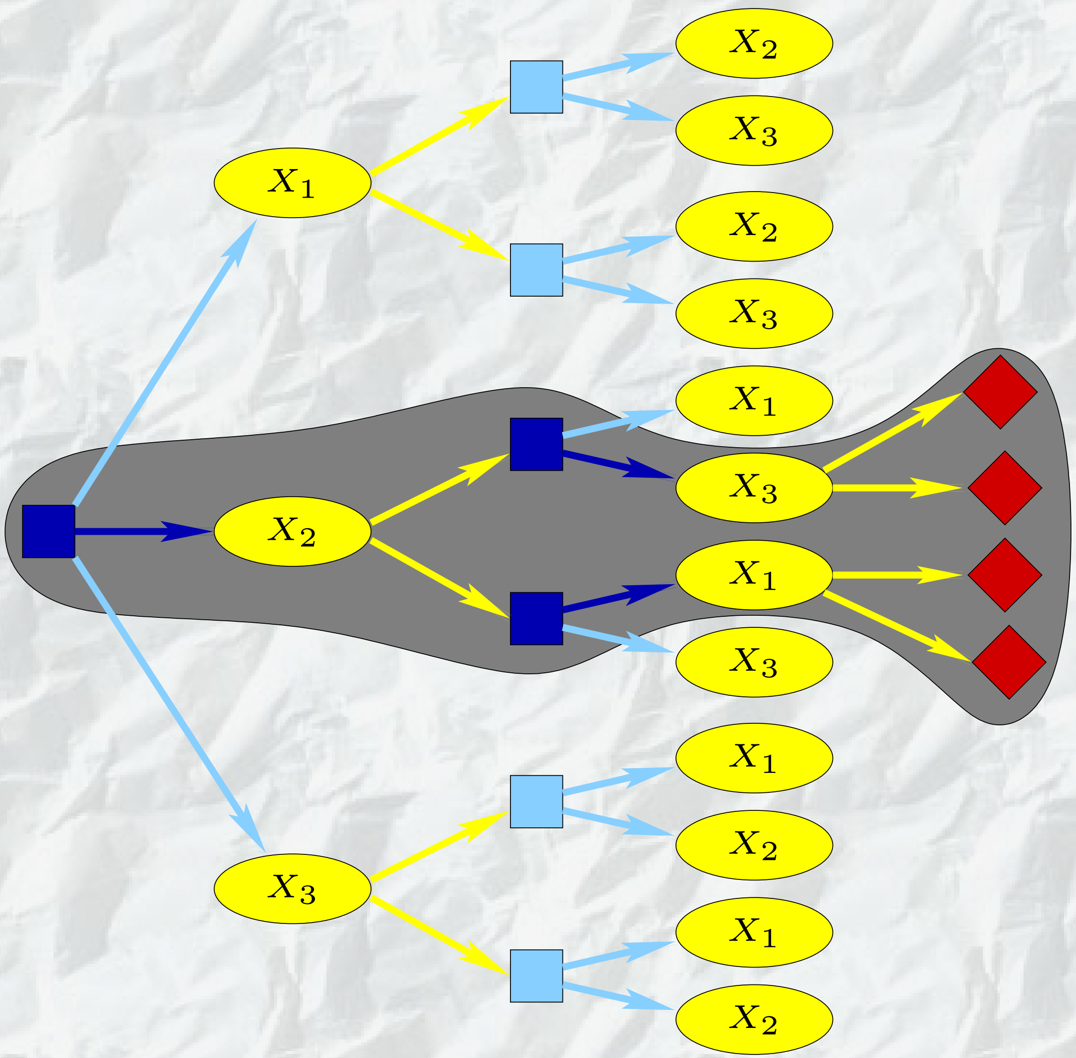
The goal:

- find a strategy that maximizes (minimizes) its expected value

Using entropy as an information measure

“The lower the entropy of a probability distribution the more we know.”

$$H\left(P(\mathbf{S})\right) \;=\; -\sum_{\mathbf{s}} P(\mathbf{S}=\mathbf{s}) \cdot \log P(\mathbf{S}=\mathbf{s})$$



Entropy in node n

$$H(\mathbf{e}_n) = H(P(\mathbf{S} \mid \mathbf{e}_n))$$

Expected entropy at the end of test $\mathbf{t}$

$$E_H(\mathbf{t}) = \sum_{\ell \in \mathcal{L}(\mathbf{t})} P(\mathbf{e}_\ell) \cdot H(\mathbf{e}_\ell)$$

$\mathcal{T}$ ... the set of all possible tests (e.g. of a given length)

A test $\mathbf{t}^\star$ is **optimal** iff

$$\mathbf{t}^\star = \arg\min_{\mathbf{t} \in \mathcal{T}} E_H(\mathbf{t}) \ .$$

Application 1: Adaptive test of basic operations with fractions

Examples of tasks:

$$T_1: \quad \left(\frac{3}{4}\cdot\frac{5}{6}\right)-\frac{1}{8} \quad = \quad \frac{15}{24}-\frac{1}{8}=\frac{5}{8}-\frac{1}{8}=\frac{4}{8}=\frac{1}{2}$$

$$T_2: \quad \frac{1}{6}+\frac{1}{12} \quad = \quad \frac{2}{12}+\frac{1}{12}=\frac{3}{12}=\frac{1}{4}$$

$$T_3: \quad \frac{1}{4}\cdot 1\frac{1}{2} \quad = \quad \frac{1}{4}\cdot\frac{3}{2}=\frac{3}{8}$$

$$T_4: \quad \left(\frac{1}{2}\cdot\frac{1}{2}\right)\cdot\left(\frac{1}{3}+\frac{1}{3}\right) \quad = \quad \frac{1}{4}\cdot\frac{2}{3}=\frac{2}{12}=\frac{1}{6}\ .$$

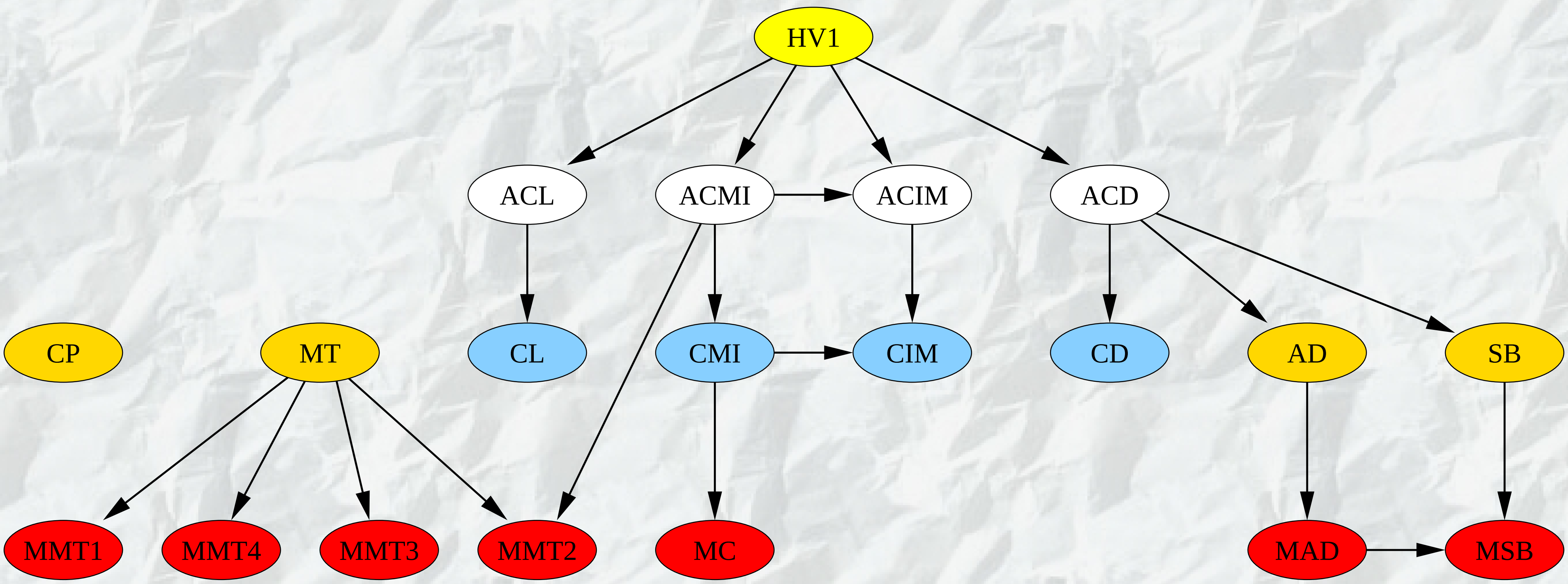
Elementary and operational skills

CP	Comparison (common numerator or denominator)	$\frac{1}{2} > \frac{1}{3}, \quad \frac{2}{3} > \frac{1}{3}$
AD	Addition (comm. denom.)	$\frac{1}{7} + \frac{2}{7} = \frac{1+2}{7} = \frac{3}{7}$
SB	Subtract. (comm. denom.)	$\frac{2}{5} - \frac{1}{5} = \frac{2-1}{5} = \frac{1}{5}$
MT	Multiplication	$\frac{1}{2} \cdot \frac{3}{5} = \frac{3}{10}$
CD	Common denominator	$\left(\frac{1}{2}, \frac{2}{3}\right) = \left(\frac{3}{6}, \frac{4}{6}\right)$
CL	Cancelling out	$\frac{4}{6} = \frac{2 \cdot 2}{2 \cdot 3} = \frac{2}{3}$
CIM	Conv. to mixed numbers	$\frac{7}{2} = \frac{3 \cdot 2+1}{2} = 3\frac{1}{2}$
CMI	Conv. to improp. fractions	$3\frac{1}{2} = \frac{3 \cdot 2+1}{2} = \frac{7}{2}$

Misconceptions

Label	Description	Occurrence
MAD	$\frac{a}{b} + \frac{c}{d} = \frac{a+c}{b+d}$	14.8%
MSB	$\frac{a}{b} - \frac{c}{d} = \frac{a-c}{b-d}$	9.4%
MMT1	$\frac{a}{b} \cdot \frac{c}{b} = \frac{a \cdot c}{b}$	14.1%
MMT2	$\frac{a}{b} \cdot \frac{c}{b} = \frac{a+c}{b \cdot b}$	8.1%
MMT3	$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot d}{b \cdot c}$	15.4%
MMT4	$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b+d}$	8.1%
MC	$a\frac{b}{c} = \frac{a \cdot b}{c}$	4.0%

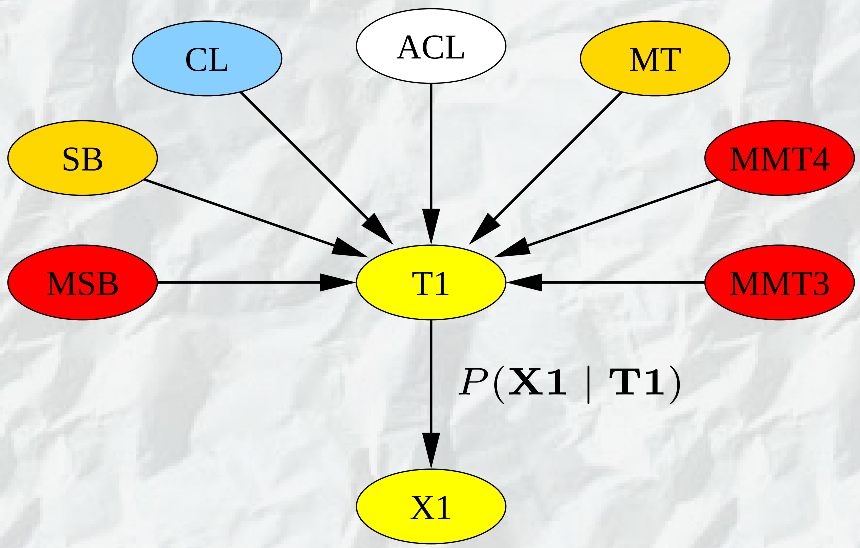
Student model

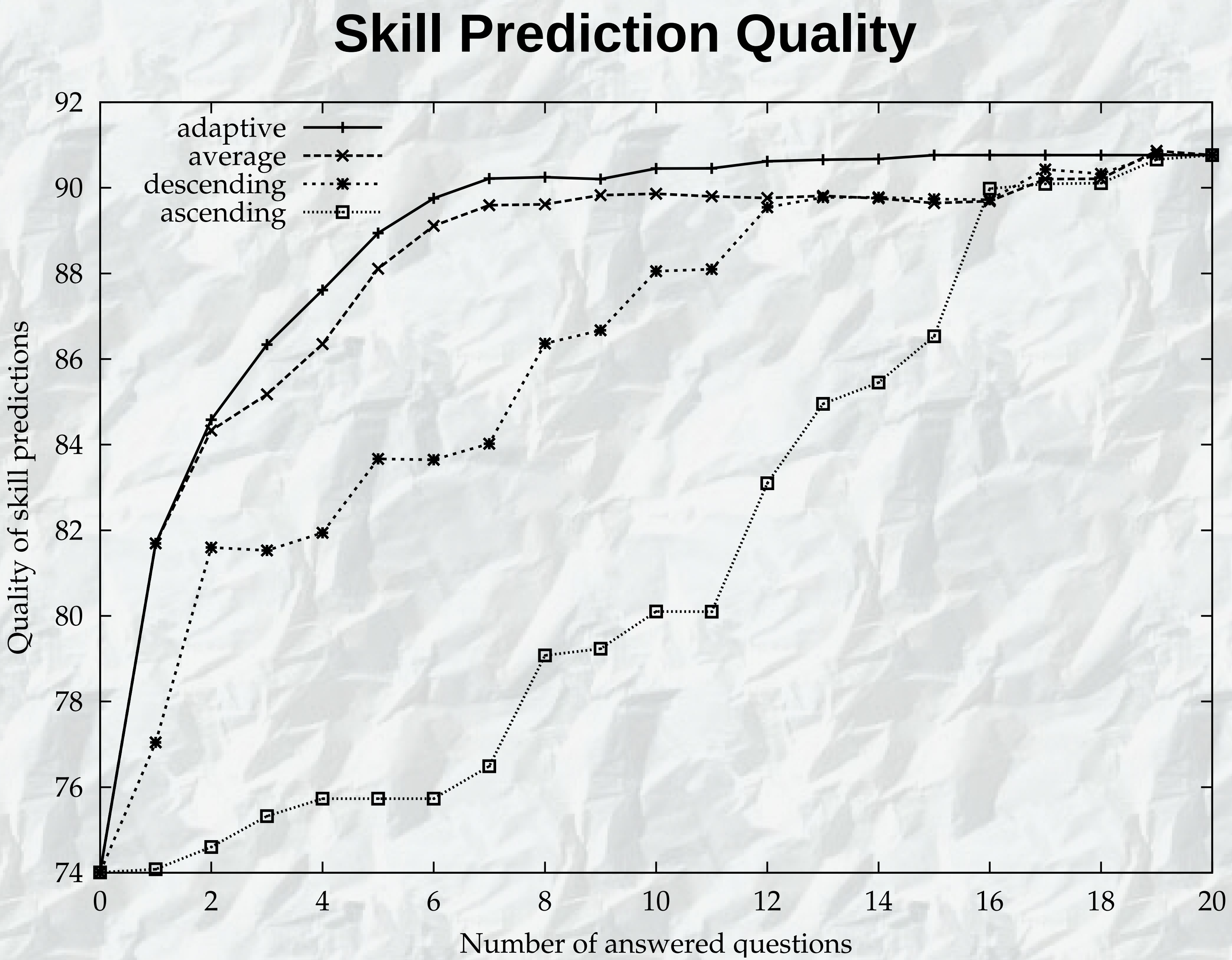


Evidence model for task $T1$

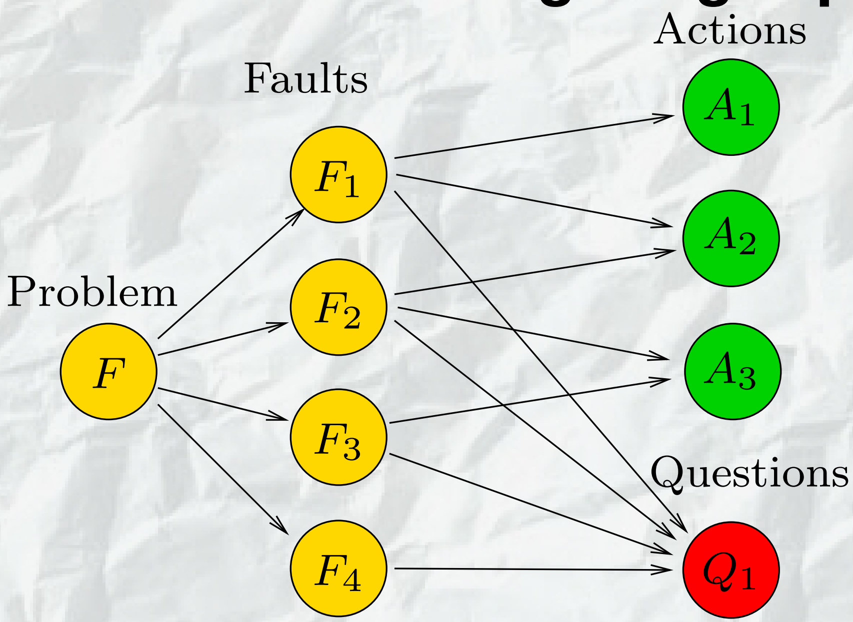
$$\left(\frac{3}{4}\cdot\frac{5}{6}\right)-\frac{1}{8}=\frac{15}{24}-\frac{1}{8}=\frac{5}{8}-\frac{1}{8}=\frac{4}{8}=\frac{1}{2}$$

$$T1 \quad \Leftrightarrow \quad MT \;\&\; CL \;\&\; ACL \;\&\; SB \;\&\; \neg MMT3 \;\&\; \neg MMT4 \;\&\; \neg MSB$$



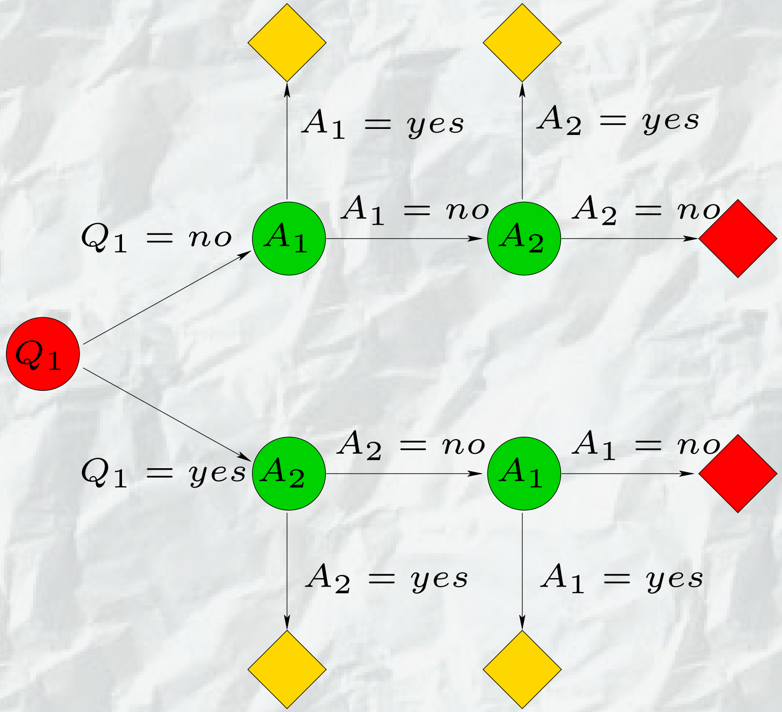

Skill Prediction Quality
adaptive
average
descending
ascending
Quality of skill predictions
Number of answered questions
92
90
88
86
84
82
80
78
76
74
0
2
4
6
8
10
12
14
16
18
20

Application 2: Troubleshooting - Light print problem



- Problems: F_1 *Distribution problem*, F_2 *Defective toner*, F_3 *Corrupted dataflow*, and F_4 *Wrong driver setting*.
- Actions: A_1 *Remove, shake and reseat toner*, A_2 *Try another toner*, and A_3 *Cycle power*.
- Questions: Q_1 *Is the configuration page printed light?*

Troubleshooting strategy



The task is to find a strategy $\mathbf{s} \in \mathcal{S}$ minimising expected cost of repair

$$E_{CR}(\mathbf{s}) = \sum_{\ell \in \mathcal{L}(\mathbf{s})} P(\mathbf{e}_\ell) \cdot (\, t(\mathbf{e}_\ell) + c(\mathbf{e}_\ell) \,) \quad .$$

Going commercial...

- Hugin Expert A/S.
  software product: Hugin - a Bayesian network tool.
  http://www.hugin.com/
- Educational Testing Service (ETS)
  the world's largest private educational testing organization
  In 2000/2001 more than 3 millions students took the ETS's largest exam SAT. Research unit doing research on adaptive test using Bayesian networks: http://www.ets.org/research/
- SACSO Project
  Systems for Automatic Customer Support Operations
  - research project of Hewlett Packard and Aalborg University.
  The troubleshooter offered as DezisionWorks by Dezide Ltd.
  http://www.dezide.com/