# Příklady aplikací bayesovských sítí

Jiří Vomlel

Ústav teorie informace a automatizace (ÚTIA)
Akademie věd České republiky
http://www.utia.cz/vomlel

Praha, 24. listopadu 2014

# Obsah přednášky

- Příklad bayesovské sítě a možných výpočtů

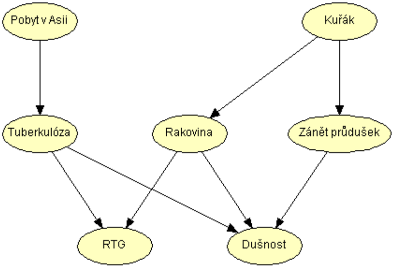
# Obsah přednášky

- Příklad bayesovské sítě a možných výpočtů
- Aplikace 1: Technická diagnostika

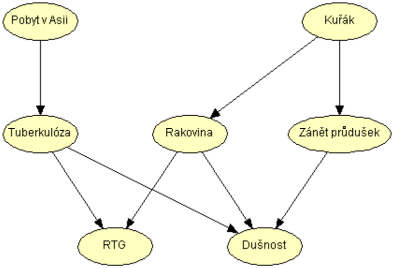
## Obsah přednášky

- Příklad bayesovské sítě a možných výpočtů
- Aplikace 1: Technická diagnostika
- Aplikace 2: Adaptivní testování znalostí

# Zjednodušený diagnostický model - Plicní klinika



# Zjednodušený diagnostický model - Plicní klinika




Zadáváme pouze:

P(Pobyt v Asii)
P(Kuřák)
P(Tuberkulóza | Pobyt v Asii)
P(Rakovina | Kuřák)
P(Zánět průdušek | Kuřák)
P(RTG | Tuberkulóza, Rakovina)
P(Dušnost | Tuberkulóza, Rakovina, Zánět průdušek)

## P(RTG | Tuberkulóza, Rakovina)

Nejprve předpokládejme, že RTG je pozitivní právě tehdy, když pacient má tuberkulózu nebo rakovinu.

| RTG | Tuberkulóza | Rakovina | p |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 |
| 1 | 1 | 1 | 1 |

# $P(\mathrm{RTG} \mid \mathrm{Tuberkulóza}, \mathrm{Rakovina})$

Připustíme, že RTG může být pozitivní i z jiného důvodu a nemusí být nutně pozitivní ikdyž pacient má tuberkulózu nebo rakovinu.

| RTG | Tuberkulóza | Rakovina | $p$ | $p'$ | |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | $p_0$ | 0.95 |
| 0 | 0 | 1 | 0 | $p_0 * p_1$ | 0.019 |
| 0 | 1 | 0 | 0 | $p_0 * p_2$ | 0.019 |
| 0 | 1 | 1 | 0 | $p_0 * p_1 * p_2$ | 0.00038 |
| 1 | 0 | 0 | 0 | $1 - p_0$ | 0.05 |
| 1 | 0 | 1 | 1 | $1 - p_0 * p_1$ | 0.981 |
| 1 | 1 | 0 | 1 | $1 - p_0 * p_2$ | 0.981 |
| 1 | 1 | 1 | 1 | $1 - p_0 * p_1 * p_2$ | 0.99962 |
| | | | | $p_0, p_1, p_2 \in \langle 0, 1 \rangle$ | |

# $P(RTG \mid Tuberkulóza, Rakovina)$

Připustíme, že RTG může být pozitivní i z jiného důvodu a nemusí být nutně pozitivní ikdyž pacient má tuberkulózu nebo rakovinu.

| RTG | Tuberkulóza | Rakovina | $p$ | $p'$ | |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | $p_0$ | 0.95 |
| 0 | 0 | 1 | 0 | $p_0 * p_1$ | 0.019 |
| 0 | 1 | 0 | 0 | $p_0 * p_2$ | 0.019 |
| 0 | 1 | 1 | 0 | $p_0 * p_1 * p_2$ | 0.00038 |
| 1 | 0 | 0 | 0 | $1 - p_0$ | 0.05 |
| 1 | 0 | 1 | 1 | $1 - p_0 * p_1$ | 0.981 |
| 1 | 1 | 0 | 1 | $1 - p_0 * p_2$ | 0.981 |
| 1 | 1 | 1 | 1 | $1 - p_0 * p_1 * p_2$ | 0.99962 |
| | | | | $p_0, p_1, p_2 \in \langle 0, 1 \rangle$ | |

Tento lokální model v bayesovské síti se nazývá „logické NEBO se šumem“ (noisy-or).

# $P(\mathrm{RTG} \mid \mathrm{Tuberkulóza}, \mathrm{Rakovina})$

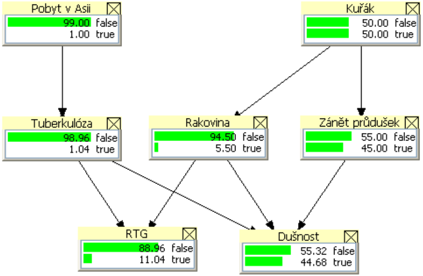
Připustíme, že RTG může být pozitivní i z jiného důvodu a nemusí být nutně pozitivní ikdyž pacient má tuberkulózu nebo rakovinu.

| RTG | Tuberkulóza | Rakovina | $p$ | $p'$ | |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | $p_0$ | 0.95 |
| 0 | 0 | 1 | 0 | $p_0 * p_1$ | 0.019 |
| 0 | 1 | 0 | 0 | $p_0 * p_2$ | 0.019 |
| 0 | 1 | 1 | 0 | $p_0 * p_1 * p_2$ | 0.00038 |
| 1 | 0 | 0 | 0 | $1 - p_0$ | 0.05 |
| 1 | 0 | 1 | 1 | $1 - p_0 * p_1$ | 0.981 |
| 1 | 1 | 0 | 1 | $1 - p_0 * p_2$ | 0.981 |
| 1 | 1 | 1 | 1 | $1 - p_0 * p_1 * p_2$ | 0.99962 |
| | | | | $p_0, p_1, p_2 \in \langle 0, 1 \rangle$ | |

Tento lokální model v bayesovské síti se nazývá „logické NEBO se šumem“ (noisy-or).

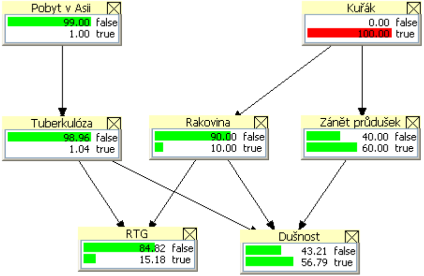
Pro jeho definici potřebujeme $k+1$ hodnot $p_0, p_1, \ldots, p_k$, kde $k$ je počet rodičů dané proměnné.

# Příklad výpočtu v bayesovské síti



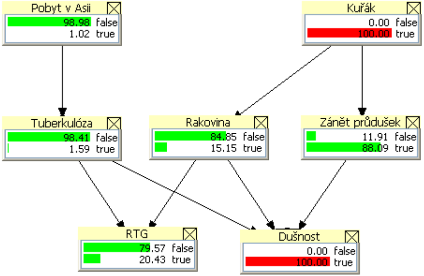

apriorní pravděpodobnost $P(X)$ všech veličin $X$

# Příklad výpočtu v bayesovské síti



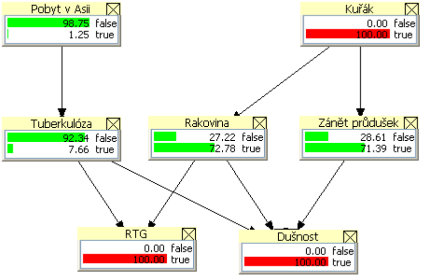

podmíněná pravděpodobnost $\mathrm{P}(X|\mathrm{Kuřák})$

# Příklad výpočtu v bayesovské síti



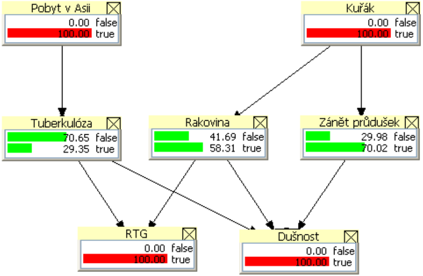

podmíněná pravděpodobnost $\mathrm{P}(\mathrm{X}|\mathrm{Kuřák},\ \mathrm{Dušnost})$

# Příklad výpočtu v bayesovské síti




podmíněná pravděpodobnost $P(X|\text{Kuřák}, \text{Dušnost}, \text{RTG})$

# Příklad výpočtu v bayesovské síti




podmíněná pravděpodobnost

P(X|Kuřák, Dušnost, RTG, Pobyt v Asii)

# Technická diagnostika – optimální strategie opravy

- Příčiny problému (závady) $C \in \mathcal{C}$.

# Technická diagnostika – optimální strategie opravy

- Příčiny problému (závady) $C \in \mathcal{C}$.
- Akce $A \in \mathcal{A}$ - opravné kroky, které mohou odstranit závadu.

# Technická diagnostika – optimální strategie opravy

- Příčiny problému (závady) $C \in \mathcal{C}$.
- Akce $A \in \mathcal{A}$ - opravné kroky, které mohou odstranit závadu.
- Otázky $Q \in \mathcal{Q}$ - kroky, které mohou pomoci identifikovat, kde je závada.

# Technická diagnostika – optimální strategie opravy

- Příčiny problému (závady) $C \in \mathcal{C}$.
- Akce $A \in \mathcal{A}$ - opravné kroky, které mohou odstranit závadu.
- Otázky $Q \in \mathcal{Q}$ - kroky, které mohou pomoci identifikovat, kde je závada.
- Ke každé akci i otázce je přiřazena cena ($c_A$ značí cenu akce $A$, $c_Q$ cenu otázky $Q$).
  Cena může znamenat:
  – dobu potřebnou k provedení akce či otázky,
  – cenu za náhradní díl, který použijeme
  – rizikovost akce
  – kombinaci výše uvedeného.

# Technická diagnostika laserové tiskárny

**Trouble**: světlý tisk.
**Troubleshooter**: doporučí kroky, které vedou k odstranění "trouble"

# Technická diagnostika laserové tiskárny

**Trouble**: světlý tisk.
**Troubleshooter**: doporučí kroky, které vedou k odstranění "trouble"

| Akce a otázky | cena |
|---|---|
| $A_1$: `Remove, shake and reseat toner` | 5 |
| $A_2$: `Try another toner` | 15 |
| $A_3$: `Cycle power` | 1 |
| $Q_1$: `Is the printer configuration page printed light?` | 2 |

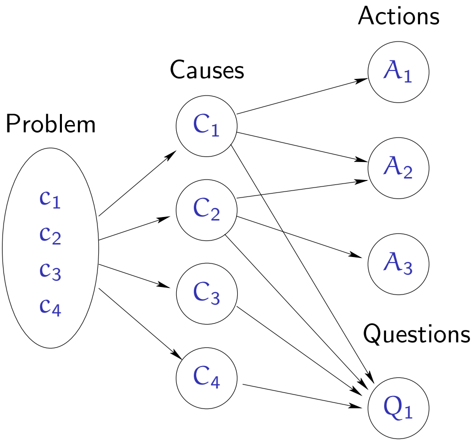
# Technická diagnostika laserové tiskárny

**Trouble**: světlý tisk.
**Troubleshooter**: doporučí kroky, které vedou k odstranění "trouble"

| Akce a otázky | cena |
|---|---|
| $A_1$: `Remove, shake and reseat toner` | 5 |
| $A_2$: `Try another toner` | 15 |
| $A_3$: `Cycle power` | 1 |
| $Q_1$: `Is the printer configuration page printed light?` | 2 |

| Možné závady při světlém tisku | $P(C_i)$ |
|---|---|
| $C_1$: `Toner low` | 0.4 |
| $C_2$: `Defective toner` | 0.3 |
| $C_3$: `Corrupted dataflow` | 0.2 |
| $C_4$: `Wrong driver setting` | 0.1 |

# Bayesovská síť pro problém světlého tisku



# Očekávaná cena opravy - ECR

- Strategie může skončit neúspěšně - např. jsme vyčerpali všechny možné akce:
  - uplatní se penalizace $c(\boldsymbol{e}_\ell)$
  - penalizací může být např. cena, kterou zaplatíme za zavolání servisních techniků

# Očekávaná cena opravy - ECR

- Strategie může skončit neúspěšně - např. jsme vyčerpali všechny možné akce:
  - uplatní se penalizace $c(\boldsymbol{e}_\ell)$
  - penalizací může být např. cena, kterou zaplatíme za zavolání servisních techniků
- Strategie může končit vyřešením problému $c(\boldsymbol{e}_\ell) = 0$.

# Očekávaná cena opravy - ECR

- Strategie může skončit neúspěšně - např. jsme vyčerpali všechny možné akce:
  - uplatní se penalizace $c(\boldsymbol{e}_\ell)$
  - penalizací může být např. cena, kterou zaplatíme za zavolání servisních techniků
- Strategie může končit vyřešením problému $c(\boldsymbol{e}_\ell) = 0$.
- Získaná evidence

$$\boldsymbol{e} \;=\; \left\{ \begin{array}{l} A = yes/no, A \in \text{Provedené akce}, \\ Q = yes/no, Q \in \text{Zodpovězené otázky} \end{array} \right\}$$

# Očekávaná cena opravy - ECR

- Strategie může skončit neúspěšně - např. jsme vyčerpali všechny možné akce:
  - uplatní se penalizace $c(\boldsymbol{e}_\ell)$
  - penalizací může být např. cena, kterou zaplatíme za zavolání servisních techniků
- Strategie může končit vyřešením problému $c(\boldsymbol{e}_\ell) = 0$.
- Získaná evidence

$$\boldsymbol{e} \;=\; \left\{ \begin{array}{l} A = yes/no, A \in \text{Provedené akce}, \\ Q = yes/no, Q \in \text{Zodpovězené otázky} \end{array} \right\}$$

- $P(\boldsymbol{e})$ ... pravděpodobnost evidence $\boldsymbol{e}$

# Očekávaná cena opravy - ECR

- Strategie může skončit neúspěšně - např. jsme vyčerpali všechny možné akce:
  - uplatní se penalizace $c(\boldsymbol{e}_\ell)$
  - penalizací může být např. cena, kterou zaplatíme za zavolání servisních techniků
- Strategie může končit vyřešením problému $c(\boldsymbol{e}_\ell) = 0$.
- Získaná evidence

$$\boldsymbol{e} = \left\{ \begin{array}{l} A = yes/no, A \in \text{Provedené akce}, \\ Q = yes/no, Q \in \text{Zodpovězené otázky} \end{array} \right\}$$

- $P(\boldsymbol{e})$ ... pravděpodobnost evidence $\boldsymbol{e}$
- $t(\boldsymbol{e})$ ... celková cena provedených akcí a otázek

# Očekávaná cena opravy - ECR

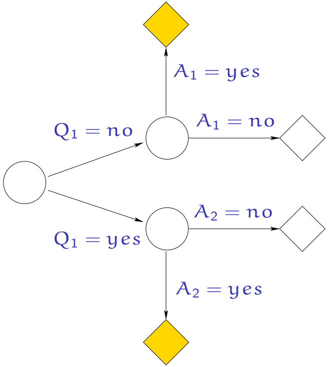
- Strategie může skončit neúspěšně - např. jsme vyčerpali všechny možné akce:
  - uplatní se penalizace $c(\boldsymbol{e}_\ell)$
  - penalizací může být např. cena, kterou zaplatíme za zavolání servisních techniků
- Strategie může končit vyřešením problému $c(\boldsymbol{e}_\ell) = 0$.
- Získaná evidence

$$\boldsymbol{e} = \left\{ \begin{array}{l} A = yes/no, A \in \text{Provedené akce}, \\ Q = yes/no, Q \in \text{Zodpovězené otázky} \end{array} \right\}$$

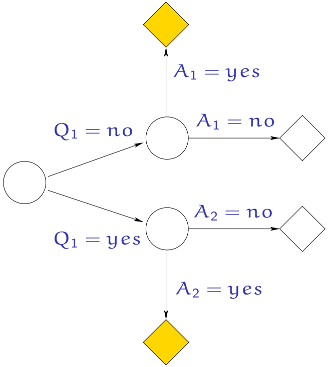
- $P(\boldsymbol{e})$ ... pravděpodobnost evidence $\boldsymbol{e}$
- $t(\boldsymbol{e})$ ... celková cena provedených akcí a otázek

$$\mathrm{ECR}(\boldsymbol{s}) = \sum_{\ell \in \text{Listy strategie } \boldsymbol{s}} P(\boldsymbol{e}_\ell) \cdot (t(\boldsymbol{e}_\ell) + c(\boldsymbol{e}_\ell))$$

# Ohodnocení strategie opravy - ECR



# Ohodnocení strategie opravy - ECR




| Strategie | očekávaná cena opravy (ECR) |
|---|---|
| $Q_1 \begin{cases} A_1 \\ A_2 \end{cases}$ | $\begin{array}{rlcl} & p(Q_1 = no, A_1 = yes) & \cdot & (c_{Q_1} + c_{A_1} + 0) \\ + & p(Q_1 = no, A_1 = no) & \cdot & (c_{Q_1} + c_{A_1} + c_{CS}) \\ + & p(Q_1 = yes, A_2 = yes) & \cdot & (c_{Q_1} + c_{A_2} + 0) \\ + & p(Q_1 = yes, A_2 = no) & \cdot & (c_{Q_1} + c_{A_2} + c_{CS}) \end{array}$ |

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $s^\star$ takovou, že pro všechny strategie $s$ platí

$$\mathrm{ECR}(s^\star) \leqslant \mathrm{ECR}(s) \ .$$

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $s^\star$ takovou, že pro všechny strategie $s$ platí

$$\mathrm{ECR}(s^\star) \leqslant \mathrm{ECR}(s) \; .$$

**Polynomiálně řešitelná úloha**

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $\boldsymbol{s}^\star$ takovou, že pro všechny strategie $\boldsymbol{s}$ platí

$$\mathrm{ECR}(\boldsymbol{s}^\star) \quad \leqslant \quad \mathrm{ECR}(\boldsymbol{s}) \ .$$

**Polynomiálně řešitelná úloha**

(1) každá akce řeší právě jednu závadu,

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $s^\star$ takovou, že pro všechny strategie $s$ platí

$$\mathrm{ECR}(s^\star) \;\leqslant\; \mathrm{ECR}(s)\;.$$

**Polynomiálně řešitelná úloha**

(1) každá akce řeší právě jednu závadu,

(2) všechny akce jsou navzájem podmíněně nezávislé při známých příčinách

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $s^\star$ takovou, že pro všechny strategie $s$ platí

$$\mathrm{ECR}(s^\star) \leqslant \mathrm{ECR}(s) \ .$$

**Polynomiálně řešitelná úloha**

(1) každá akce řeší právě jednu závadu,

(2) všechny akce jsou navzájem podmíněně nezávislé při známých příčinách

(3) platí, že zařízení má v jednom okamžiku pouze jednu závadu (single fault assumption)

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $s^\star$ takovou, že pro všechny strategie $s$ platí

$$\mathrm{ECR}(s^\star) \leqslant \mathrm{ECR}(s) \ .$$

**Polynomiálně řešitelná úloha**

(1) každá akce řeší právě jednu závadu,

(2) všechny akce jsou navzájem podmíněně nezávislé při známých příčinách

(3) platí, že zařízení má v jednom okamžiku pouze jednu závadu (single fault assumption)

**NP-těžká úloha**

# Cíl technické diagnostiky

**Cíl:** nalézt strategii $s^\star$ takovou, že pro všechny strategie $s$ platí

$$\mathrm{ECR}(s^\star) \leqslant \mathrm{ECR}(s) \ .$$

**Polynomiálně řešitelná úloha**

(1) každá akce řeší právě jednu závadu,

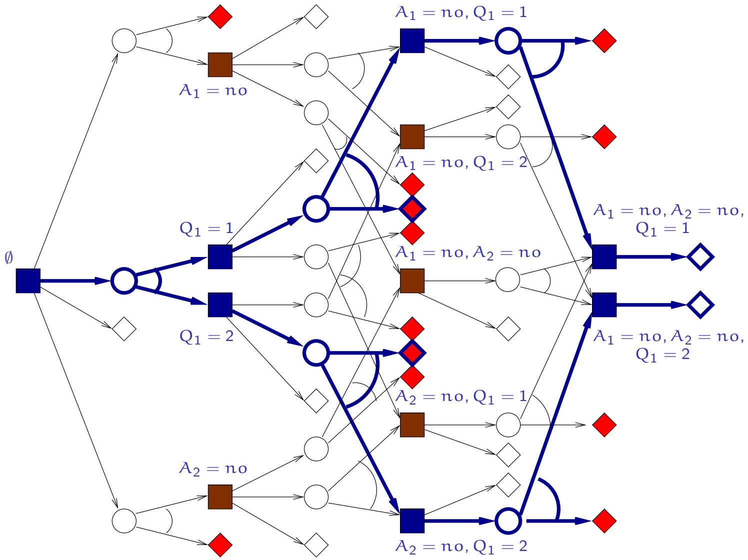
(2) všechny akce jsou navzájem podmíněně nezávislé při známých příčinách

(3) platí, že zařízení má v jednom okamžiku pouze jednu závadu (single fault assumption)

**NP-těžká úloha**

(1) některé akce řeší více než dvě závady

# Reprezentace prostoru všech řešení AND/OR grafem



## Algoritmy pro nalezení optimálního řešení

- využití metod prohledávání stavového prostoru:

## Algoritmy pro nalezení optimálního řešení

- využití metod prohledávání stavového prostoru:
- metoda větví a mezí (branch and bound)

# Algoritmy pro nalezení optimálního řešení

- využití metod prohledávání stavového prostoru:
- metoda větví a mezí (branch and bound)
- metody dynamického programování

# Algoritmy pro nalezení optimálního řešení

- využití metod prohledávání stavového prostoru:
- metoda větví a mezí (branch and bound)
- metody dynamického programování
- $A^\star$ pro AND/OR grafy

# Algoritmy pro nalezení optimálního řešení

- využití metod prohledávání stavového prostoru:
- metoda větví a mezí (branch and bound)
- metody dynamického programování
- $A^\star$ pro AND/OR grafy

Prohledávání prostoru všech řešení je výpočetně náročné.

# Suboptimální řešení v reálném čase

**BATS troubleshooter**:

- Vyvinutý v rámci společného projektu Hewlett-Packard a Aalborg University.

# Suboptimální řešení v reálném čase

**BATS troubleshooter**:

- Vyvinutý v rámci společného projektu Hewlett-Packard a Aalborg University.
- Využívá několik heuristik založených na poměru $p/c$.

# Suboptimální řešení v reálném čase

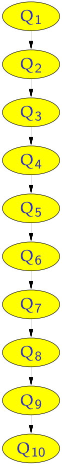
**BATS troubleshooter**:

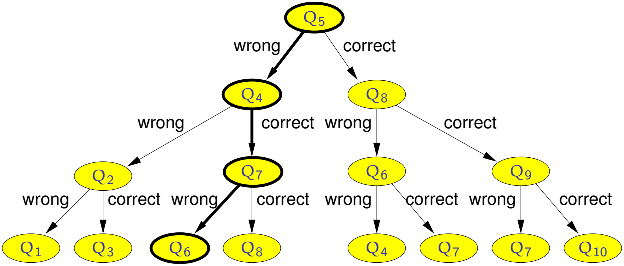
- Vyvinutý v rámci společného projektu Hewlett-Packard a Aalborg University.
- Využívá několik heuristik založených na poměru $p/c$.

**Porovnání optimálního řešení s BATS troubleshooterem**

| Problem | $\mid\mathcal{A}\mid$ | $\mid\mathcal{Q}\mid$ | OPTIM | BATS |
|---|---|---|---|---|
| 53 | 6 | 2 | **433.238** | 443.305 |
| Tray | 9 | 3 | **129.214** | 129.214 |
| Overrun | 11 | 3 | **106.204** | 112.456 |
| Load | 12 | 3 | **38.3777** | 38.4229 |
| Pjam | 13 | 4 | **124.323** | 124.365 |
| Scatter | 14 | 4 | **115.410** | 115.862 |
| NotDupl | 9 | 9 | **70.6740** | 73.5984 |
| Spots | 16 | 5 | **161.385** | 162.246 |
| MIO1 | 10 | 10 | **250.452** | 253.310 |

# Adaptivní test znalostí



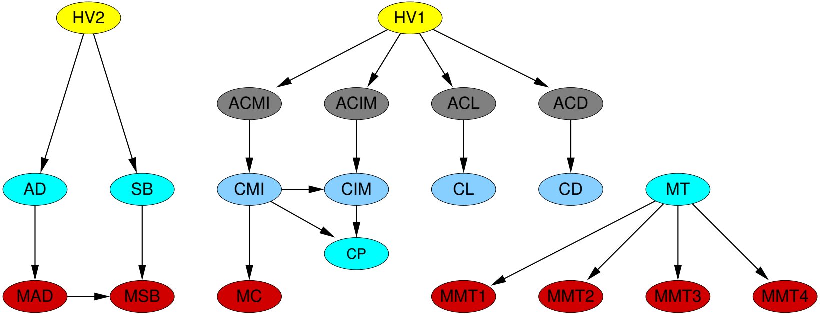
# Příklad testu: Matematické operace se zlomky

| | Dovednost | Příklad |
|---|---|---|
| **CP** | Porovnávání (společný čitatel nebo jmenovatel) | $\frac{1}{2} > \frac{1}{3}, \quad \frac{2}{3} > \frac{1}{3}$ |
| **AD** | Sčítání (společný jmenovatel) | $\frac{1}{7} + \frac{2}{7} = \frac{1+2}{7} = \frac{3}{7}$ |
| **SB** | Odčítání (společný jmenovatel) | $\frac{2}{5} - \frac{1}{5} = \frac{2-1}{5} = \frac{1}{5}$ |
| **MT** | Násobení | $\frac{1}{2} \cdot \frac{3}{5} = \frac{3}{10}$ |
| **CD** | Převod na společný jmenovatel | $\left(\frac{1}{2}, \frac{2}{3}\right) = \left(\frac{3}{6}, \frac{4}{6}\right)$ |
| **CL** | Krácení | $\frac{4}{6} = \frac{2 \cdot 2}{2 \cdot 3} = \frac{2}{3}$ |
| **CIM** | Převod na smíšené zlomky | $\frac{7}{2} = \frac{3 \cdot 2 + 1}{2} = 3\frac{1}{2}$ |
| **CMI** | Převod na nepravé zlomky | $3\frac{1}{2} = \frac{3 \cdot 2 + 1}{2} = \frac{7}{2}$ |

# Mylné postupy (Misconceptions)

| | Příklad | Četnost výskytu v datech |
|---|---|---|
| **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% |

# Model studenta



# Model pro úlohu $T_1$

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

# Model pro úlohu $T_1$

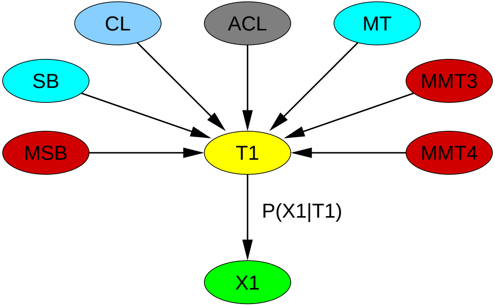
$$\left(\frac{3}{4}\cdot\frac{5}{6}\right)-\frac{1}{8}=\frac{15}{24}-\frac{1}{8}=\frac{5}{8}-\frac{1}{8}=\frac{4}{8}=\frac{1}{2}$$

$$\mathrm{T1} \quad \Leftrightarrow \quad \mathit{MT}\ \&\ \mathit{CL}\ \&\ \mathit{ACL}\ \&\ \mathit{SB}\ \&\ \neg \mathit{MMT3}\ \&\ \neg \mathit{MMT4}\ \&\ \neg \mathit{MSB}$$

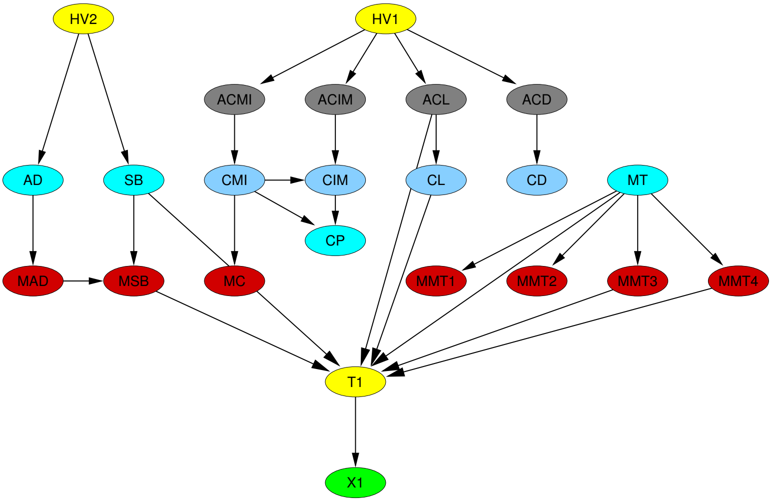
# Model pro úlohu $T_1$

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

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



# Model studenta spojený s modelem pro úlohu $T_1$



## Adaptivní test

Opakujeme dva základní kroky:

1. odhadujeme úroveň jednotlivých znalostí studenta

# Adaptivní test

Opakujeme dva základní kroky:

1. odhadujeme úroveň jednotlivých znalostí studenta
2. vybíráme vhodnou otázku na základě odhadu úrovně znalostí

# Adaptivní test

Opakujeme dva základní kroky:

1. odhadujeme úroveň jednotlivých znalostí studenta
2. vybíráme vhodnou otázku na základě odhadu úrovně znalostí

Entropie jako míra informace:

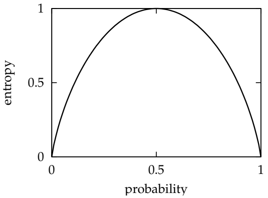
# Adaptivní test

Opakujeme dva základní kroky:

1. odhadujeme úroveň jednotlivých znalostí studenta
2. vybíráme vhodnou otázku na základě odhadu úrovně znalostí

Entropie jako míra informace:

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



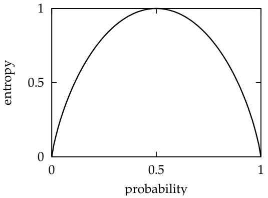
# Adaptivní test

Opakujeme dva základní kroky:

1. odhadujeme úroveň jednotlivých znalostí studenta
2. vybíráme vhodnou otázku na základě odhadu úrovně znalostí

Entropie jako míra informace:

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




- Čím je entropie nižší tím více toho o znalostech studenta víme.

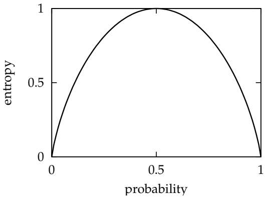
# Adaptivní test

Opakujeme dva základní kroky:

1. odhadujeme úroveň jednotlivých znalostí studenta
2. vybíráme vhodnou otázku na základě odhadu úrovně znalostí

Entropie jako míra informace:

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




- Čím je entropie nižší tím více toho o znalostech studenta víme.
- Hladový algoritmus: V každém kroku vybereme otázku, která nejvíce snižuje entropii.

# Kvalita predikce jednotlivých dovedností