Použití bayesovských sítí při testování znalostí

Jiří Vomlel

Laboratoř inteligentních systémů
Vysoká škola ekonomická Praha

Tato prezentace je k dispozici na:
http://www.utia.cas.cz/vomlel/

Obsah

- Testování znalostí může být “velký obchod”.
- Model studenta a evidenční modely.
- Jednoduchý příklad ukazující, že je důležité modelovat závislosti mezi dovednostmi.
- Co je to pevný test a adaptivní test?
- Optimální and krátkozrace optimální testy.
- Konstrukce krátkozrace optimálního testu.
- Test základních operací se zlomky.
- Výsledky experimentů.
- Závěr.

Educational Testing Service (ETS)

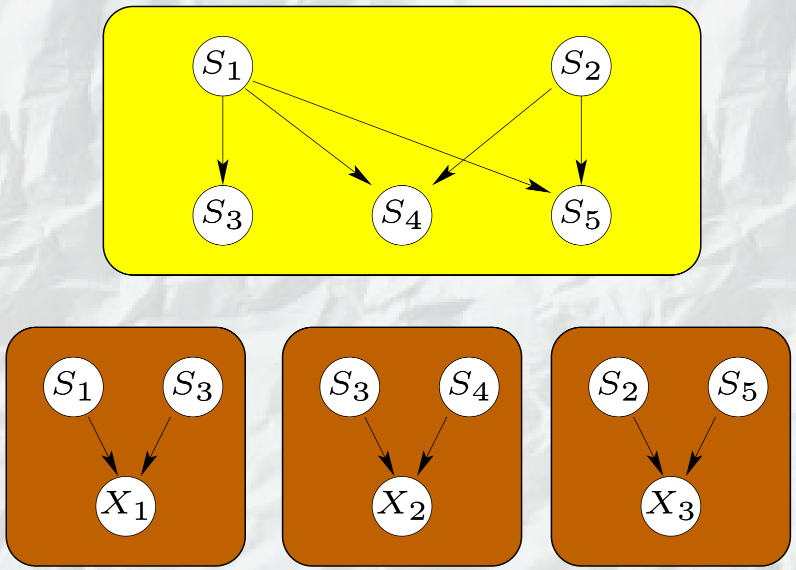
- Educational Testing Service (ETS) je největší soukromá organizace testující znalosti. Má 2300 stálých zaměstnanců.
- Počet účastníků v jednotlivých testech ve škol. roce 2000-2001:

3 185 000	SAT I Reasoning Test and SAT II: Subject Area Tests *SAT test je v USA standarně používán při přijímacích zkouškách na univerzitu.*
2 293 000	PSAT: Preliminary SAT/National Merit Scholarship Qualifying Test
1 421 000	AP: Advanced Placement Program
801 000	The Praxis Series: Professional Assessments for Beginning Teachers and Pre-Professional Skills Tests
787 000	TOEFL: Test of English as a Foreign Language
449 000	GRE: Graduate Record Examinations General Test

Model studenta a modely otázek

(R. Almond and R. Mislevy, 1999)

- jeden **model studenta** popisující vztahy mezi dovednostmi, schopnostmi, špatnými přístupy (misconceptions), apod.
- **modely otázek**, jeden model pro každou otázku.



Příklad jednoduché diagnostické úlohy

Diagnóza přítomnosti či nepřítomnosti tří dovedností

$$S_1, S_2, S_3$$

pomocí tří otázek

$$X_{1,2}, X_{1,3}, X_{2,3} \quad .$$

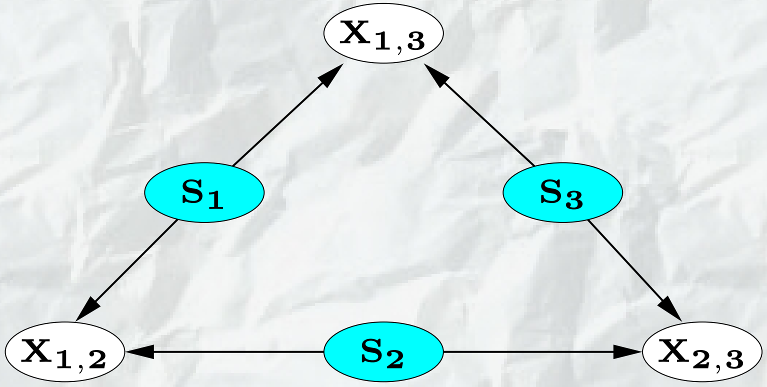
jejichž závislost na dovednostech je definována pomocí podmíněných pravděpodobností

$$P(X_{i,j} = 1 | S_i = s_i, S_j = s_j) \quad = \quad \begin{cases} 1 & \text{jestliže } (s_i, s_j) = (1,1) \\ 0 & \text{jinak.} \end{cases}$$

Předpokládejme, že všechny odpovědi byly špatně, t.j.

$$X_{1,2} = 0, \;\; X_{1,3} = 0, \;\; X_{2,3} = 0 \quad .$$

Usuzování za předpokladu nezávislosti dovedností



Dovednosti jsou navzájem nezávislé

$$P(S_1, S_2, S_3) \quad = \quad P(S_1) \cdot P(S_2) \cdot P(S_3)$$

a pro $i = 1, 2, 3, s_i = 0, 1$

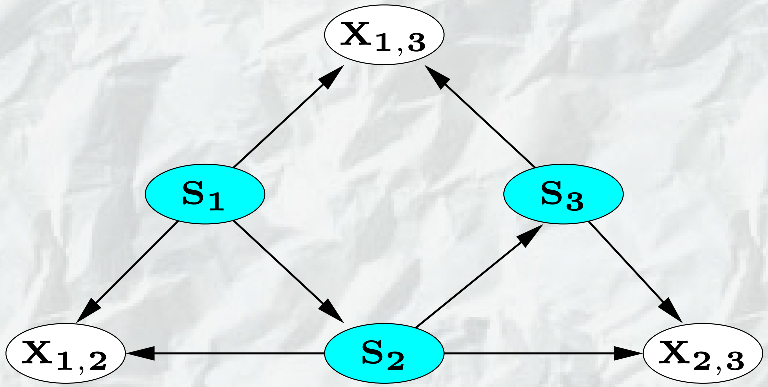
$$P(S_i = s_i) \quad = \quad \frac{1}{2}$$

Potom pravděpodobnosti pro $j = 1, 2, 3$ jsou:

$$P(S_j = 0 \mid X_{1,2} = 0, X_{1,3} = 0, X_{2,3} = 0) = 0.75 \ \ ,$$

t.j. nemůžeme s jistotou rozhodnout, které dovednosti student má a které nemá.

Modelování závislosti mezi dovednostmi



s deterministickou hierarchií

$$S_1 \Rightarrow S_2,\ S_2 \Rightarrow S_3$$

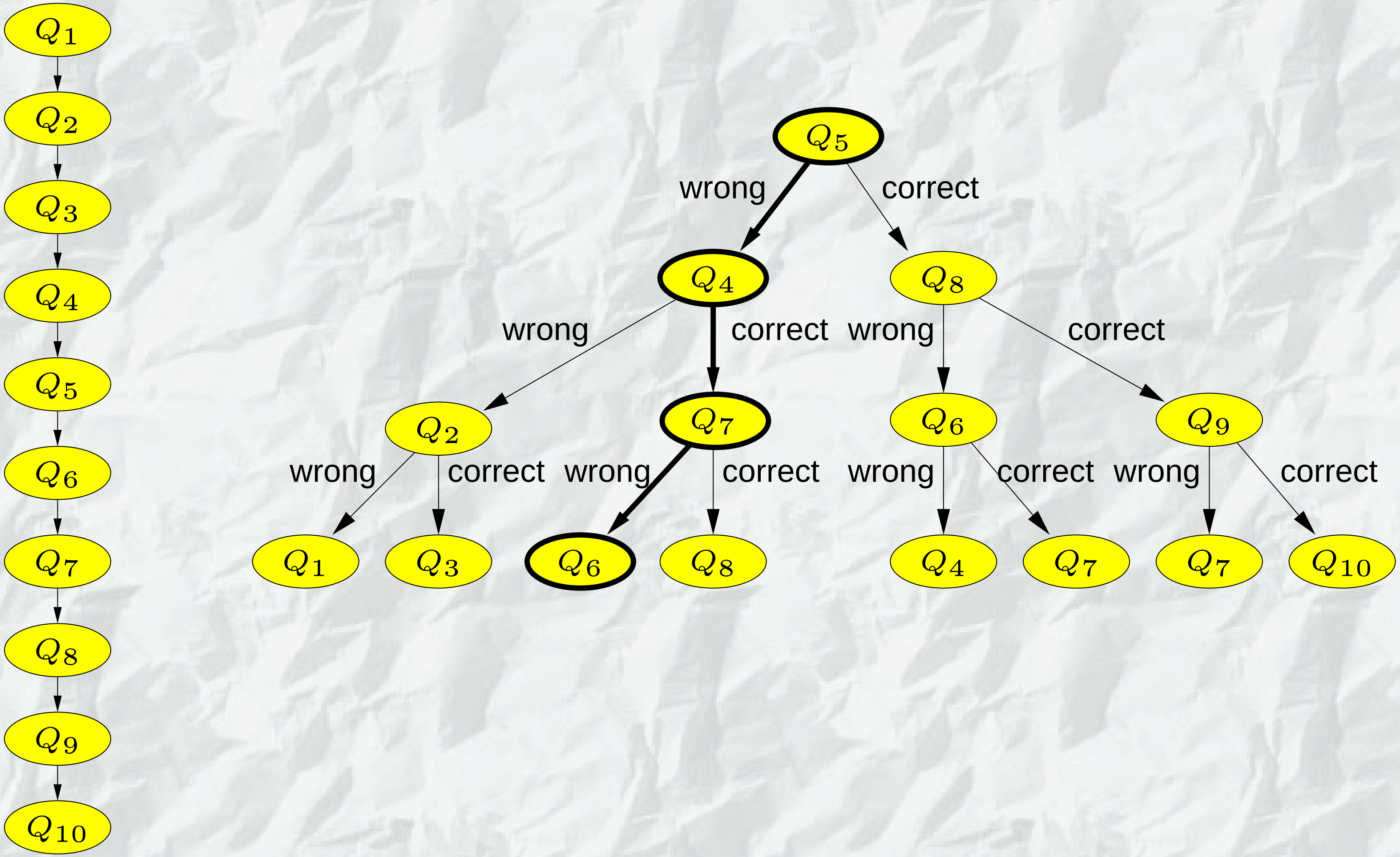
$$\begin{aligned}
P(S_1 = 0 \mid X_{1,2} = 0, X_{1,3} = 0, X_{2,3} = 0) &= 1 \\
P(S_2 = 0 \mid X_{1,2} = 0, X_{1,3} = 0, X_{2,3} = 0) &= 1 \\
P(S_3 = 0 \mid X_{1,2} = 0, X_{1,3} = 0, X_{2,3} = 0) &= 0.5
\end{aligned}$$

Povšimněme si, že pro $i = 1, 2, 3$

$$P(S_i \mid X_{1,2} = 0, X_{1,3} = 0, X_{2,3} = 0) = P(S_i \mid X_{2,3} = 0)\ ,$$ t.j.

$X_{2,3} = 0$ dává **stejnou informaci** jako $X_{1,2} = 0, X_{1,3} = 0, X_{2,3} = 0$.

Pevný test a adaptivní test



Počítačové adaptivní testování

Computerized Adaptive Testing (CAT)

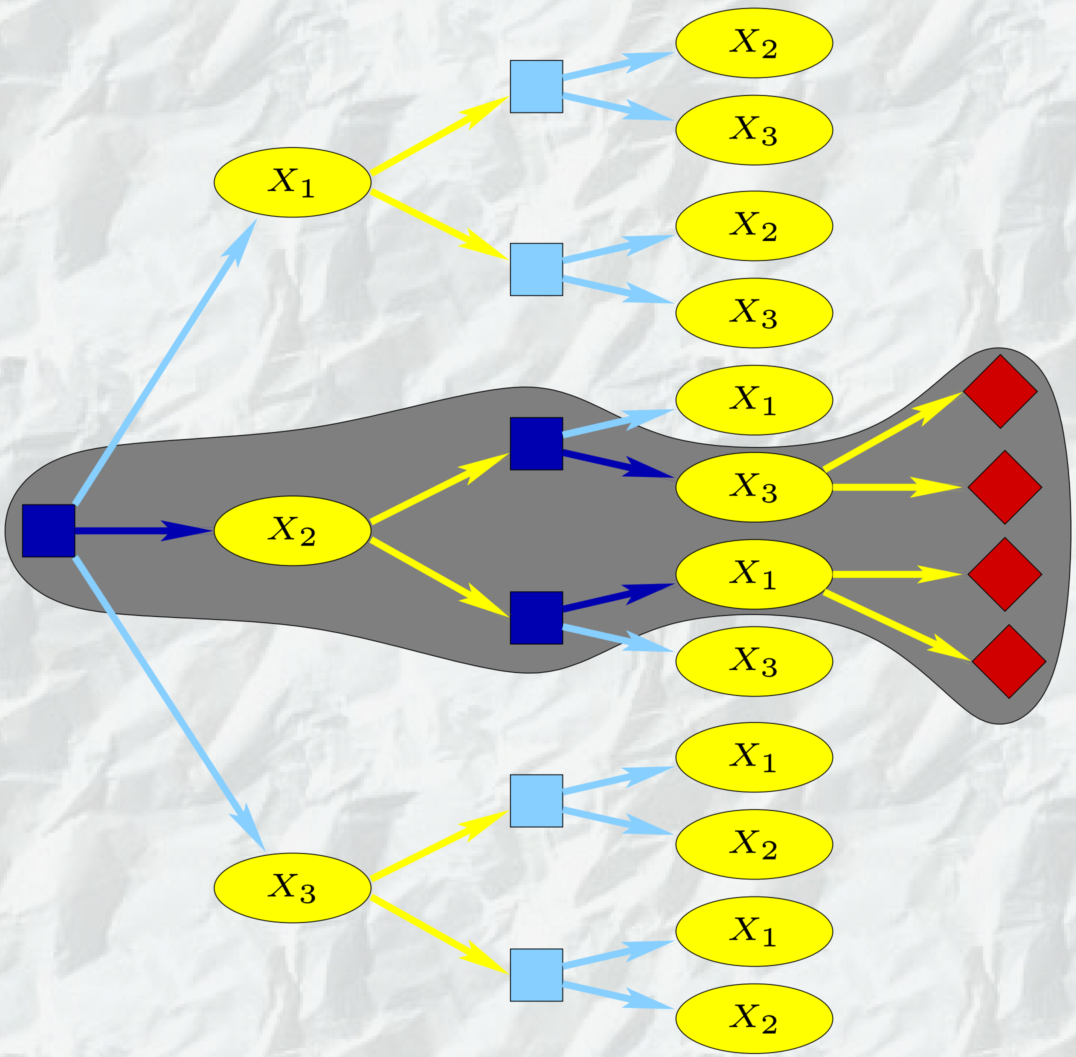
Cílem je test, který přináší **co nejvíce informací o každém studentovi**.

Dva základní kroky
1. odhad úrovně znalostí studenta
2. výběr otázky odpovídající úrovni studenta

Entropie (míra neuspořádanosti) pravděpodobnostního rozdělení $P(\mathbf{S})$

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

“Čím je entropie nižší, tím více víme o testovaném studentovi.”



Entropie v uzlu n

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

Očekávaná entropie na konci testu $\mathbf{t}$

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

$\mathcal{T}$ množina přípustných testů (např. testů dané délky)

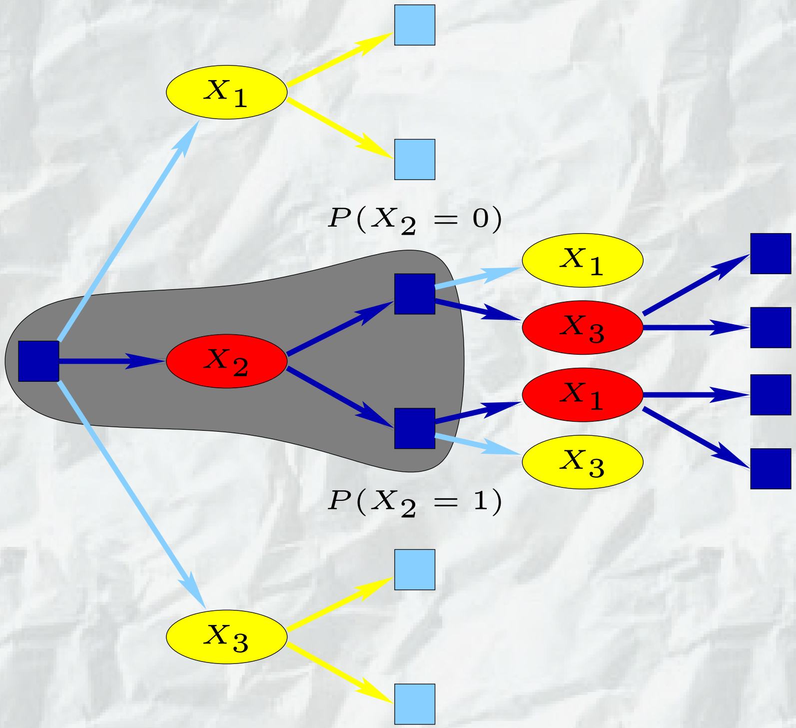
Test $\mathbf{t}^\star$ je **optimální** jestliže

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

Krátkozrace optimální test $\mathbf{t}$ je test kde každá vybraná otázka $X^\star$ v testu $\mathbf{t}$ minimalizuje očekávanou entropii po zodpovězení otázky:

$$X^\star = \arg\min_{X \in \mathcal{X}} E_H(\mathbf{t}_{\downarrow X}) \ ,$$

t.j. při výběru otázky se chováme tak, jakoby test měl po vybrané otázce skončit.



$$
\begin{aligned}
e\_list &= \{\{X_2 = 0\}, \{X_2 = 1\}\} \\
new\_e\_list &= \{\{X_2 = 0, X_3 = 0\}, \\
&\quad \{X_2 = 0, X_3 = 1\}, \\
&\quad \{X_2 = 1, X_1 = 0\}, \\
&\quad \{X_2 = 1, X_1 = 1\}\}
\end{aligned}
$$

Krátkozraká konstrukce testu

$e\_list := \{\emptyset\};$

$test := \{\ \};$

for $position := 1$ to $test\_lenght$ do

  $new\_e\_list := \{\ \};$

  for all $\mathbf{e} \in e\_list$ do

    $i := most\_informative\_X(\mathbf{e});$

    for all values x_j of X_i do

      $\mathbf{e}_{i,j} := \{\mathbf{e} \cup \{X_i = x_j\}\};$

      $append(new\_e\_list, \mathbf{e}_{i,j});$

      $append(test, \mathbf{e} \rightarrow \mathbf{e}_{i,j});$

  $e\_list := new\_e\_list;$

$return(test);$

Adaptivní test základních operací se zlomky

Příklady úloh:

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

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

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

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

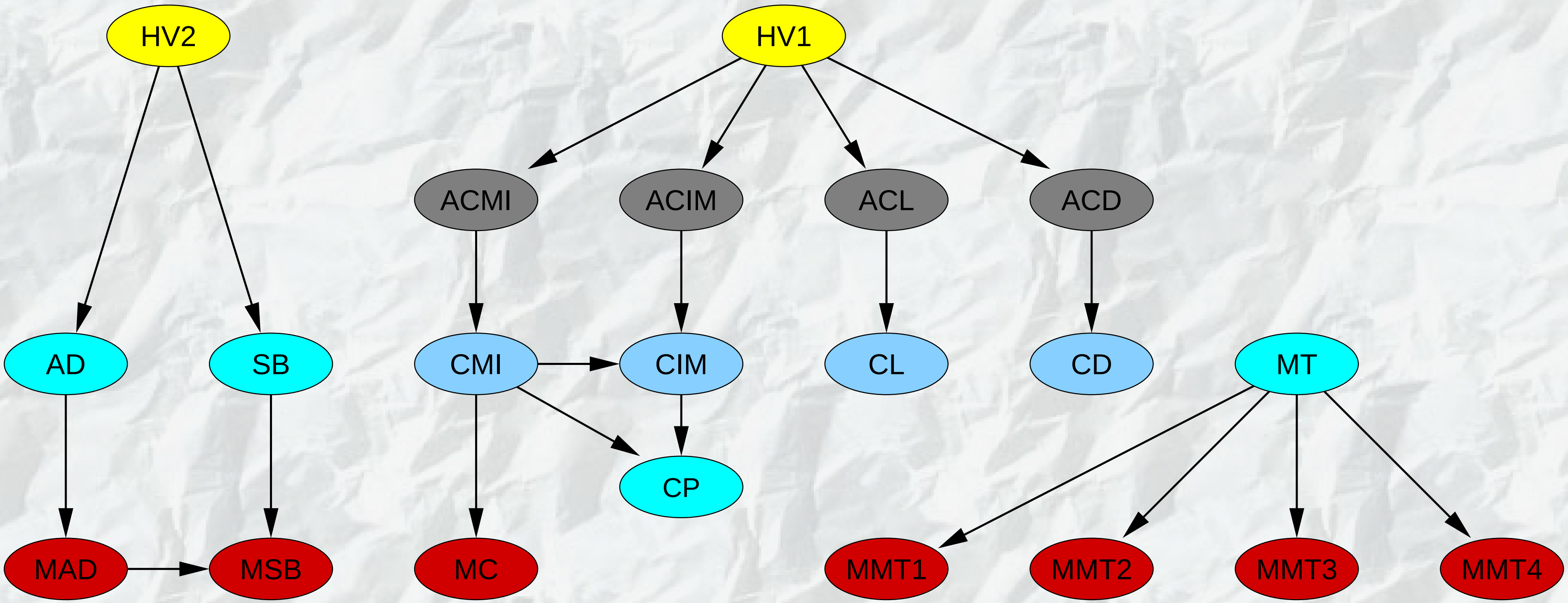
Elementární a operační dovednosti

CP	Porovnávání (spol. čitatel nebo jmenov.)	$\frac{1}{2} > \frac{1}{3}, \quad \frac{2}{3} > \frac{1}{3}$
AD	Sčítání (spol. jmenov.)	$\frac{1}{7} + \frac{2}{7} = \frac{1+2}{7} = \frac{3}{7}$
SB	Odčítání (spol. jmenov.)	$\frac{2}{5} - \frac{1}{5} = \frac{2-1}{5} = \frac{1}{5}$
MT	Násobení	$\frac{1}{2} \cdot \frac{3}{5} = \frac{3}{10}$
CD	Společ. 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 složené zlomky	$\frac{7}{2} = \frac{3 \cdot 2 + 1}{2} = 3\frac{1}{2}$
CMI	Převod na nepravé zlomky	$3\frac{1}{2} = \frac{3 \cdot 2 + 1}{2} = \frac{7}{2}$

Špatné postupy

Označení	Popis	Výskyt
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%

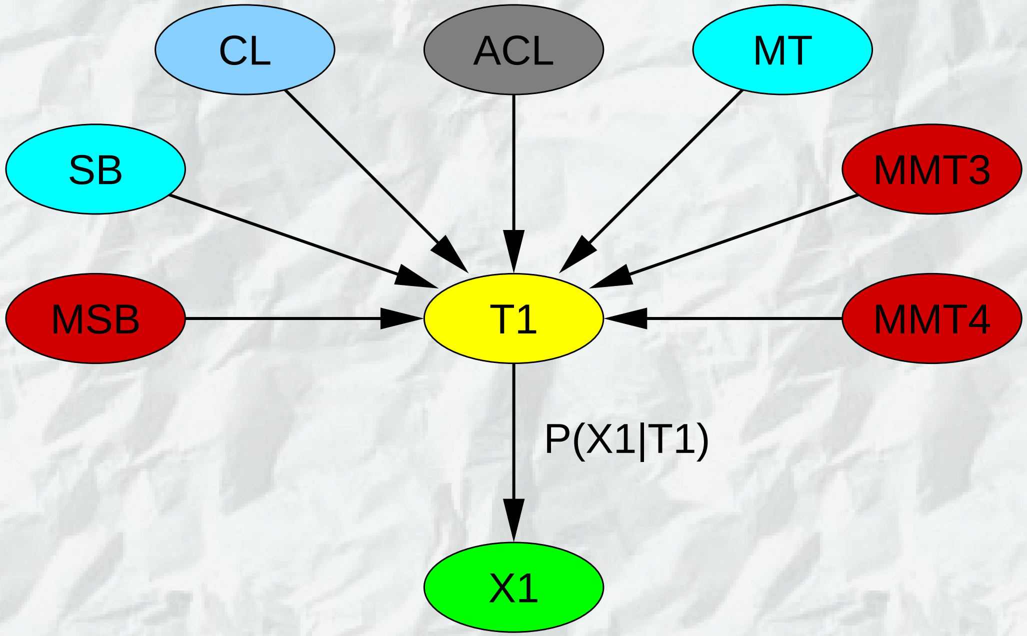
Model studenta



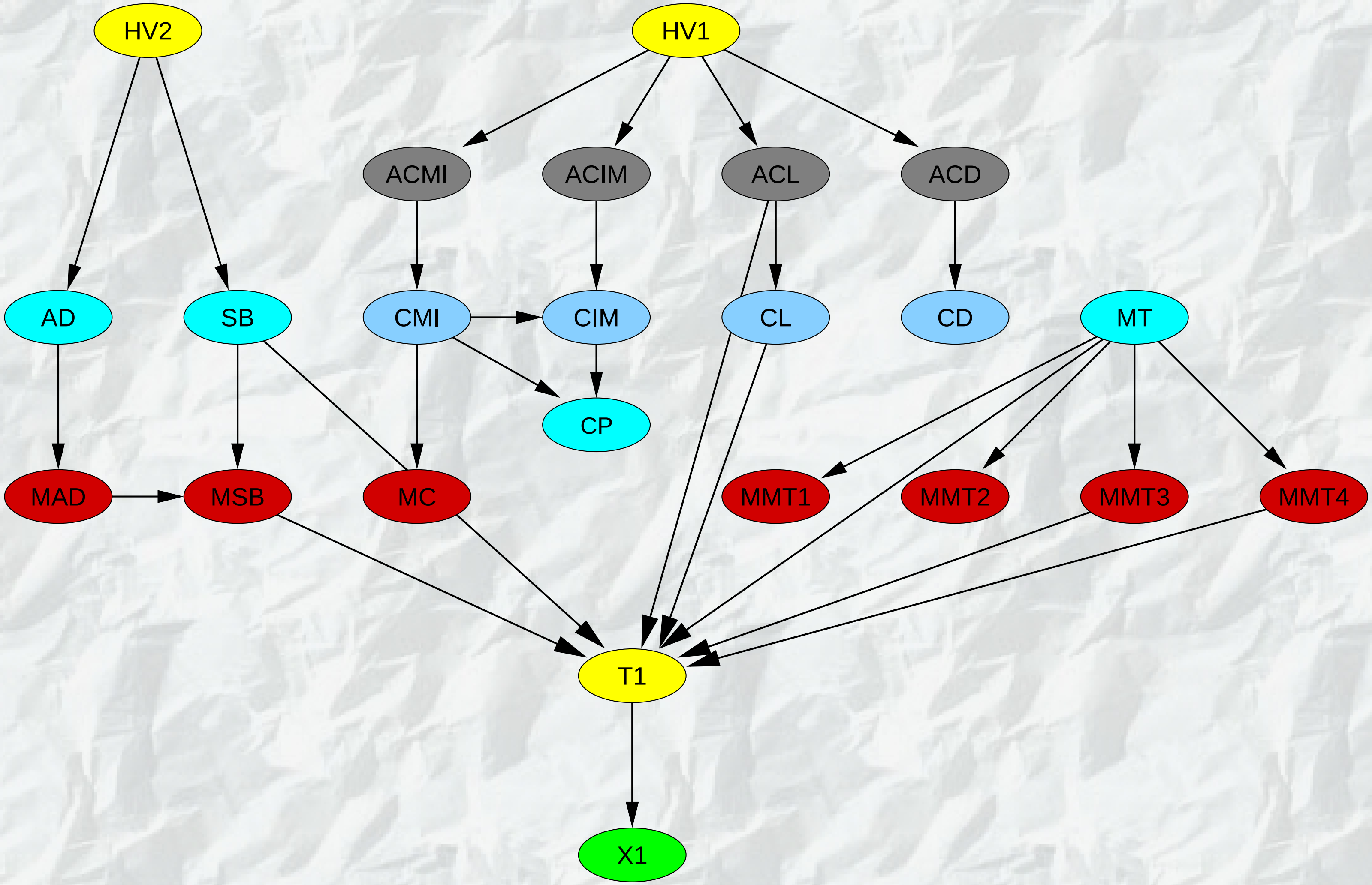
Model úlohy T1

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

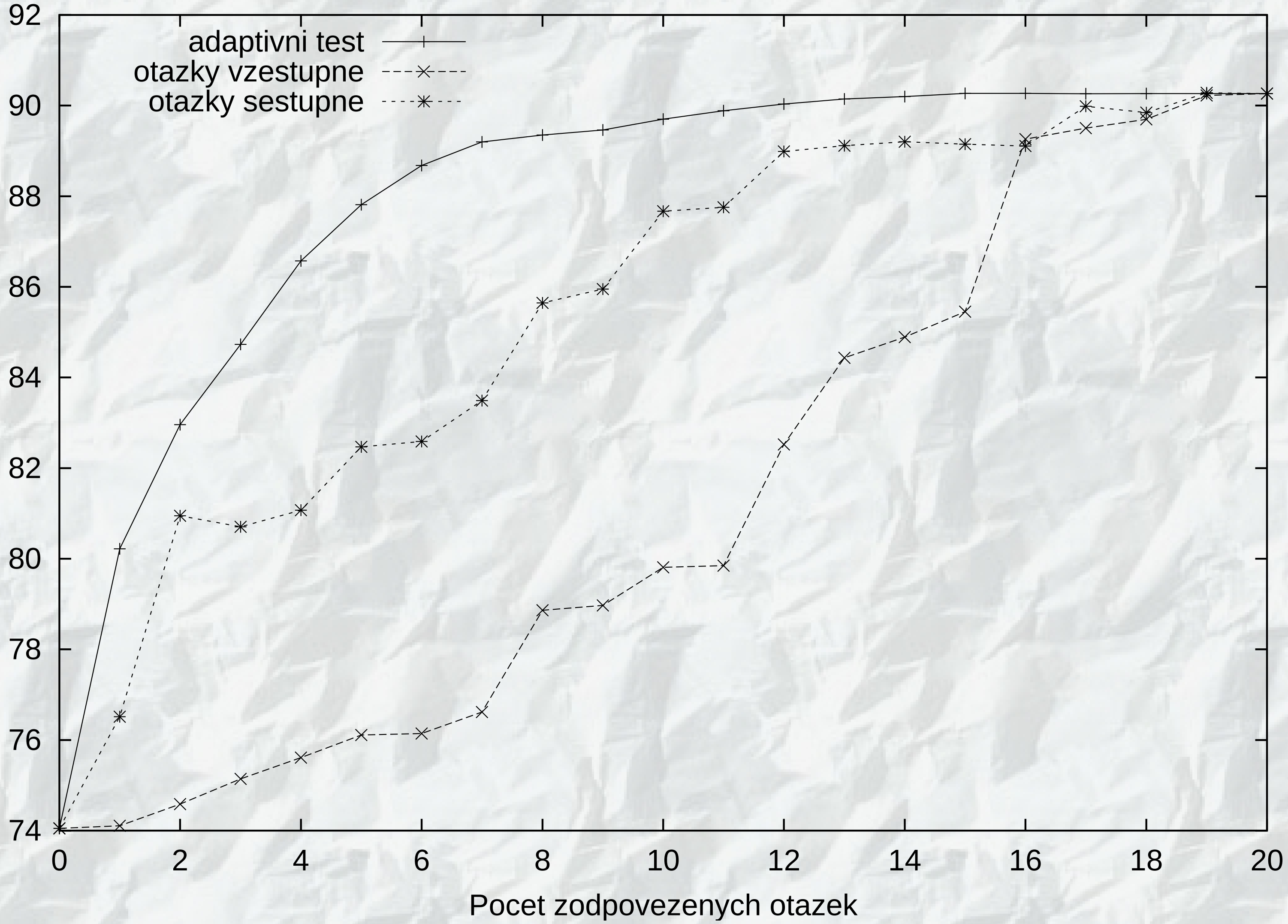
$$T1 \quad \Leftrightarrow \quad MT \,\&\, CL \,\&\, ACL \,\&\, SB \,\&\, \neg MMT3 \,\&\, \neg MMT4 \,\&\, \neg MSB$$



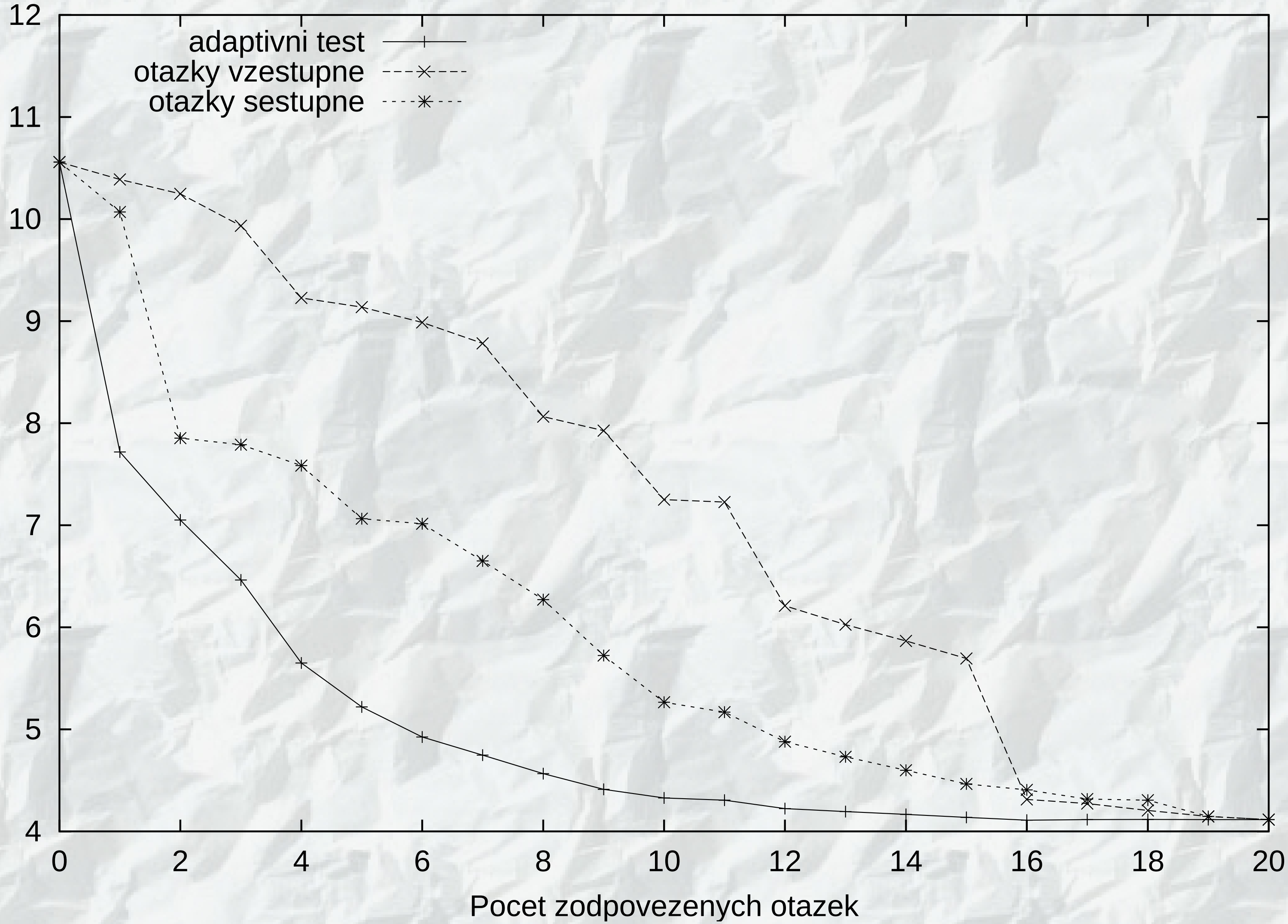
Model studenta spojený s modelem úlohy



Kvalita předpovědi dovedností



Vývoj entropie



Závěr

- Ukazuje se, že při testování znalostí lze s výhodou použít **bayesovské sítě**.
- Adaptivní testy mohou podstaně **snížit počet otázek**, na které je třeba se studenta zeptat.