

Modelování obtížnosti logických a výukových úloh

Radek Pelánek

- otázky: Co ovlivňuje obtížnost problémů? Jak obtížnost predikovat?
- výzkum na FI, spolupráce Petr Jarušek a další
- využití modelů:
 - výpočetní
 - statistické
- ilustrace kontextu modelování: sběr dat, evaluace, aplikace

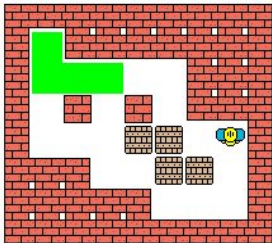
Obtížnost a výpočetní modelování

případové studie Sokoban, Sudoku:

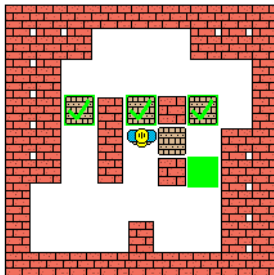
- sběr dat
- analýzy dat, hledání metrik obtížnosti
- výpočetní modely

Sokoban

4 min



49 min

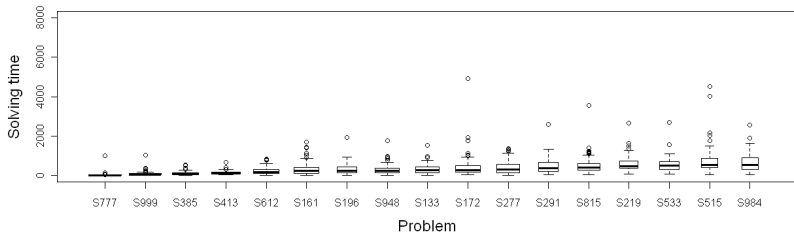


Sběr dat

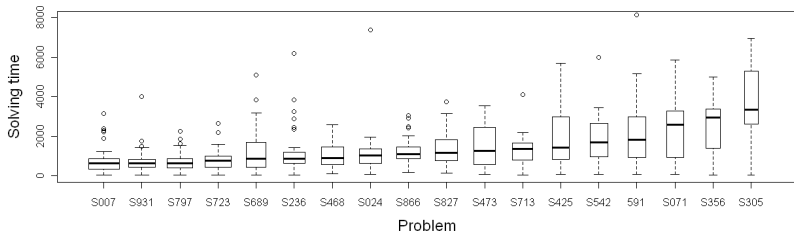
- vlastní web
- logování všech tahů
- 35 levelů, vždy 4 bedny
- cca 2000 odehraných her, data robustní

Výsledky – čas

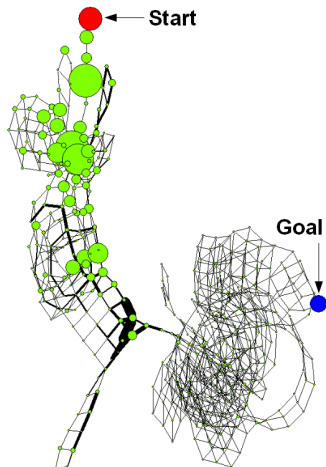
Sokoban solving time (problems 1-17)



Sokoban solving time (problems 18-35)



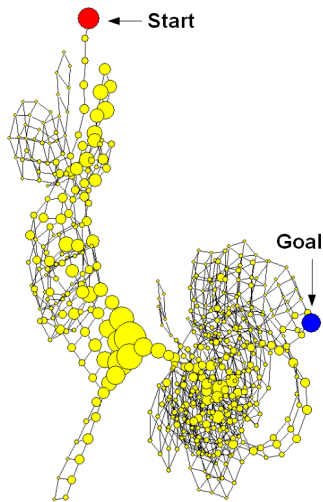
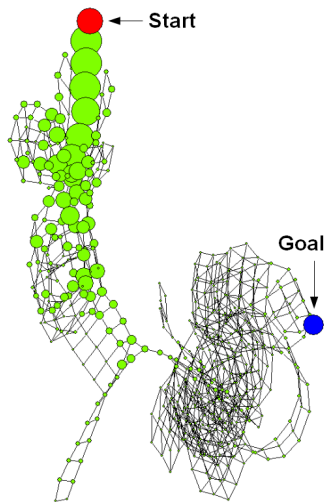
Stavový prostor – čas strávený lidmi



Výpočetní model

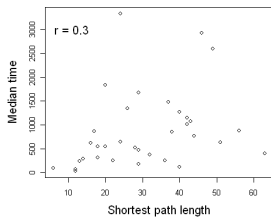
- dynamické procházení stavového prostoru
- snaha simulovat chování člověka (nikoliv myšlení)
- jednoduchý model, kombinace dvou tendencí
 - „dobrý nápad“ – ideální posun směrem k cíli
 - bloudění – náhoda
- čím blíže k cíli, tím méně náhody

Lidé a model

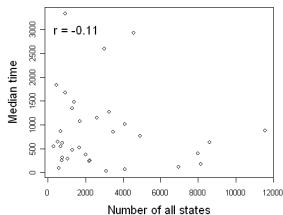


Metriky obtížnosti

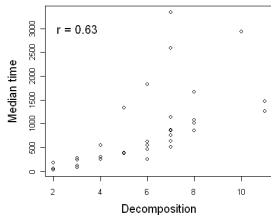
Shortest path length and median time



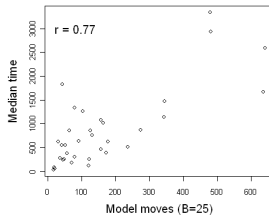
Number of all states and median time



Decomposition



Model moves and median time (B=25)

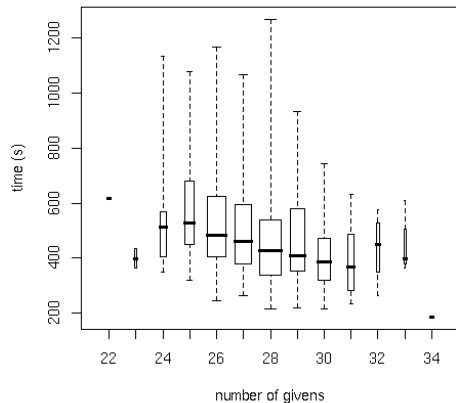


Sudoku

| | | | | | | | | |
|---|---|---|---|--|---|---|---|---|
| | 7 | 5 | | | 1 | 8 | | 4 |
| | | 9 | | | | | | 3 |
| | | | | | 5 | 1 | 2 | |
| | | 4 | | | 8 | 6 | 3 | |
| | | | | | | | | |
| | 2 | 3 | 1 | | | 4 | | |
| | 4 | 7 | 9 | | | | | |
| 1 | | | | | | 9 | | |
| 3 | | 6 | 4 | | | 2 | 8 | |

- `fed-sudoku.eu`:
 - cca 1000 zadání, časy pro 100 řešitelů
 - data od webmastera
- `sudoku.org.uk`:
 - cca 1000 zadání, cca 1000 řešitelů, jen průměrný čas
 - stažené skriptem z webu
- `czech-sudoku.com`:
 - použito cca 20 zadání
 - logy her (každý tah)
 - stažené z webu / od webmastera

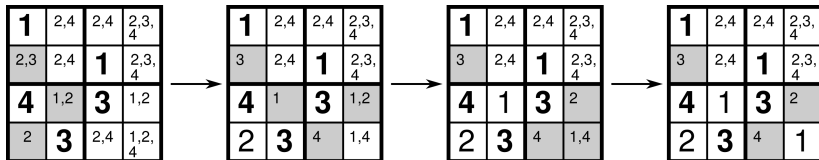
Naivní metrika: počet čísel v zadání



Výpočetní model

- „logické techniky“ pro odvození další pozice (hidden single, naked single, hidden pair, X-wing, ...)
- náhodně vybere jedno z možných doplnění a pak pokračuje

Výpočetní model: příklad



Výpočetní modely: principy

- běžný přístup: hodně logických technik, mnoho parametrů
- náš model: málo technik, málo parametrů
 - snazší vyhodnocení
 - lepší přenositelnost
 - lepší „vhled“

Jednoduchý výpočetní model

- jen dvě základní logické techniky (hidden single, naked single)
- selže jednoduchá logika \Rightarrow prohledávání
 - hledá pole, kde je potřeba nejmenší počet kroků k vyvrácení špatných kandidátů
 - aproximace složitějších logických technik

Model vs lidé: srovnání pro konkrétní hru

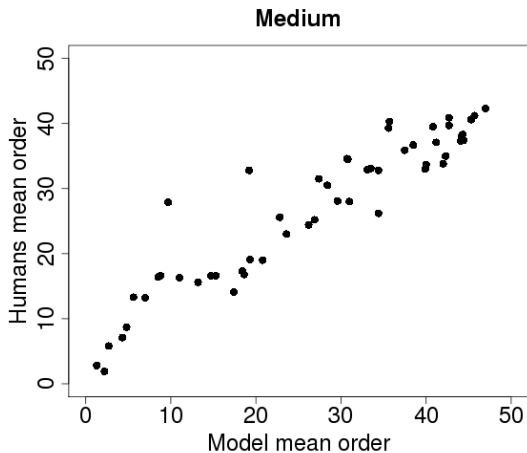
lidé

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 6 | | 1 | | | 2 | | 8 | |
| 9 | | | | 4 | | 1 | | 7 |
| | 8 | | | | | | | |
| | 2 | | 9 | | | | | |
| | | 5 | 6 | | 8 | 3 | | |
| | | | | | 4 | | 5 | |
| | | | | | | | 3 | |
| 5 | | 8 | | 2 | | | | 1 |
| | 4 | | 7 | | | 8 | | 6 |

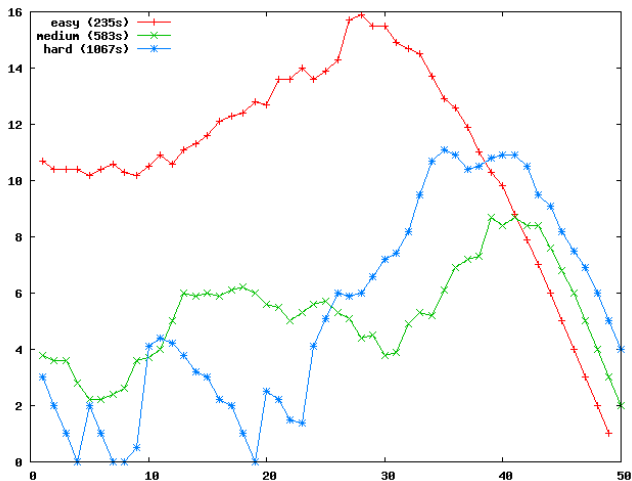
model

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 6 | | 1 | | | 2 | | 8 | |
| 9 | | | | 4 | | 1 | | 7 |
| | 8 | | | | | | | |
| | 2 | | 9 | | | | | |
| | | 5 | 6 | | 8 | 3 | | |
| | | | | | 4 | | 5 | |
| | | | | | | | 3 | |
| 5 | | 8 | | 2 | | | | 1 |
| | 4 | | 7 | | | 8 | | 6 |

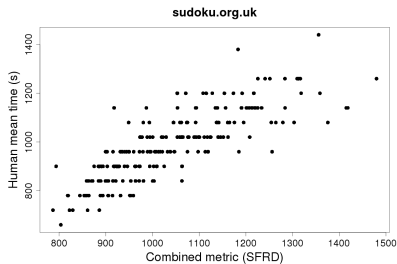
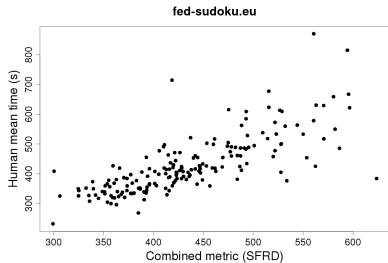
Model vs lidé: srovnání pro konkrétní hru



Počet možných „tahů“ a obtížnost



Výsledky: kombinovaná metrika



koeficient determinace r^2

| metric | fed-sudoku.eu | | sudoku.org.uk | |
|------------------|---------------|--------|---------------|--------|
| | all | simple | all | simple |
| number of givens | 6% | 5% | 2% | 12% |
| Serate | 49% | 30% | 74% | 28% |
| Serate LM | 61% | 36% | 75% | 43% |
| Fowler's | 47% | 28% | 76% | 41% |
| Refutation sum | 47% | – | 70% | – |
| Dependency | 45% | 54% | 49% | 62% |
| Combined (RD) | 54% | – | 78% | – |
| Combined (SFRD) | 66% | 57% | 91% | 66% |

Shrnutí zkušeností – Sokoban, Sudoku

- „statické“ metriky nefungují
- „dynamické“ výpočetní modely
 - jednoduché, abstraktní modely, málo parametrů, fungují docela dobře
 - nejsou úplně přímočaré – spousta „dobrých nápadů“ nefungovala
- metrika „vzorek lidí“ – i pro úlohy s jednoduchými pravidly těžké překonat

- `tutor.fi.muni.cz`
- současně:
 - sběr dat o řešení
 - využití dat pro predikce, doporučování úloh
- predikce pomocí statistického modelu
- inspirace: doporučující systémy (recommender systems), např. Amazon, Netflix
- asi 30 úloh (logické, matematické, informatické)

Tutor: úlohy

tutor.fi.muni.cz

PROBLEM SOLVING

TUTOR



Body: 4464



Jste přihlášen jako **radek** v individuálním módu

Můj účet

přepnout na Výukový mód

Odhlásit

PROBLÉMY

STATISTIKY

VÝSLEDKOVKA

Informatické



Interaktivní
Python



Konečné
automaty

```
#include <stdio.h>
#define N 5

main() {
    int x, y;

    printf(" ");
    for (x = 1; x <= N;
        printf("???", x);
        printf("\n\n");
}
```

Programování
v C

pes zajíc
kočka rys
husa kozel

$^*[a-z](3,4)^*$

Regulární
výrazy



Robot Karel



Robotanik



Želví grafika

Matematické

| | X | Y | Z |
|---|---|---|---|
| A | 1 | 1 | 1 |
| B | 0 | 0 | |
| C | | 1 | 1 |

Binární
křížovka



Grafář (nová
verze)



Matematické
pexeso 2



Obrazce



Rozbitá
kalkulačka



Transformace



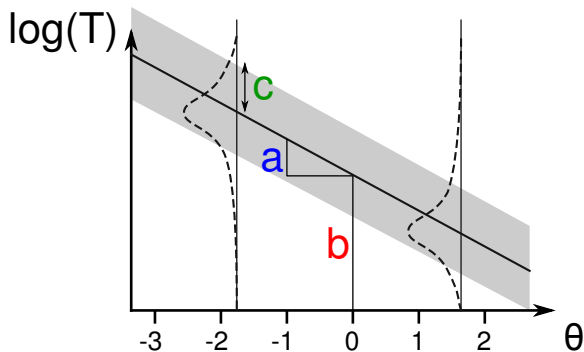
Výpočetní
stromy

Tutor: předpovědi

tutor.fi.muni.cz

| | | | | | | | | |
|---|---|--|---|---|--|--|---|---|
| Kuželosečky - hyperboly Neřešeno Předpověď 1:12 | Komplexní čísla - násobení Vyřešeno Čas 0:58 | Logaritmy a mocniny - vzorečky Neřešeno Předpověď 1:14 | Komplexní čísla - mocniny i Vyřešeno Čas 1:24 | Vlastnosti funkcí Neřešeno Předpověď 1:17 | Kuželosečky 2 Vyřešeno Čas 1:05 | Zlomky Neřešeno Předpověď 1:19 | Komplexní čísla - absolutní hodnoty Vyřešeno Čas 0:45 | Logaritmy - hodnoty 2 Neřešeno Předpověď 1:23 |
| Kvadratické rovnice - řešení Vyřešeno Čas 1:45 | Vzdálenosti Vyřešeno Čas 1:16 | Kuželosečky Neřešeno Předpověď 1:35 | Množiny - základní operace Neřešeno Předpověď 1:42 | Kombinační čísla Neřešeno Předpověď 1:42 | Kvadratická funkce 2 Neřešeno Předpověď 1:44 | Definiční obory a obory hodnot Neřešeno Předpověď 1:46 | Logaritmy - vzorečky Vyřešeno Čas 2:42 | Množiny Neřešeno Předpověď 1:56 |
| Směs Neřešeno Předpověď 2:02 | Derivace - goniometrické funkce Vyřešeno Čas 2:15 | Součty Vyřešeno Čas 1:10 | Kombinační čísla - vzorečky Neřešeno Předpověď 2:17 | Komplexní čísla Neřešeno Předpověď 2:17 | Úhly 2 Neřešeno Předpověď 2:29 | Nerovnosti Neřešeno Předpověď 2:39 | Kuželosečky - kružnice Neřešeno Předpověď 2:56 | Limity funkcí Vyřešeno Čas 2:36 |

Model obtížnosti úloh



Odhad parametrů

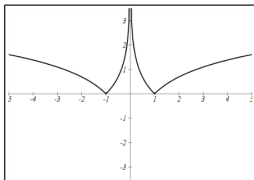
- dostupná data: uživatel i , vyřešil úlohu j v čase t_{ij}
- potřebujeme současně odhadnout:
 - schopnosti uživatelů θ
 - parametry problémů a, b, c
- metody strojového učení (stochastic gradient descent)
- analogické doporučujícím systémům (např. Netflix – hodnocení filmů)
- vyhodnocení: úspěšnost predikcí (RMSE)

stejná základní obtížnost

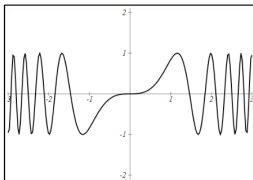
vysoká diskriminace

vysoká náhodnost

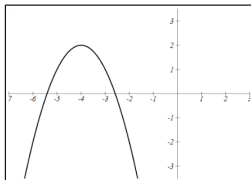
"na jistotu"



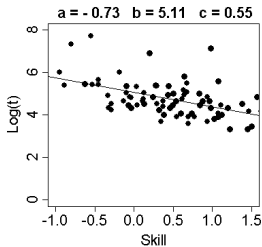
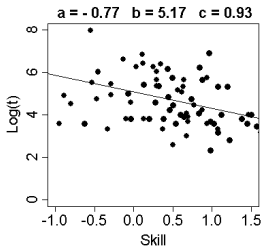
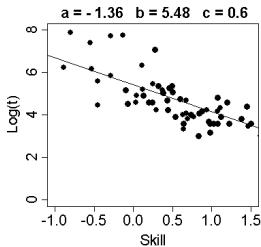
$\text{abs}(\log(\text{abs}(x)))$



$\sin(x^3)$



$-(x+4)^2+2$



- modely:
 - výpočetní – simulace chování člověka, specifické pro problém, využitelné pro nápovědy
 - statistické – popisné, metody strojového učení, široce aplikovatelné
- kvantitativní vyhodnocení nad daty
- možnost zapojení – máme spousty dat ...