



# Umělá inteligence a rozpoznávání

Expertní genealogický systém

Vypracoval: Zdeněk JANEČEK

Datum: 12. května 2013

# 1 Zadání

Vytvořte genealogický expertní systém. Vstupem bude informace databáze základních informací o osobě [id, jméno, matka, otec, pohlaví, partner]. Program bude umět odpovídat na otázky: Kdo je můj pokrevní příbuzný? Kolik má moje prarabata sester? Jaký je vztah mezi Petrem a Pavlem?... a další podobné.

# 2 Analýza

Rodinný strom je z hlediska grafové teorie  $n$ -ární strom, protože mám dva rodiče a neomezeně mnoho dětí. Pokud zanedbám manželství a postavím strom jen ze vztahů rodič–potomek, bude acykličnost zachována. Manželství není dokonce stálé a stačí si ho tedy pouze zaznamenat. Společné děti manželů, také nejsou jisté, tudíž je praktičtější ukládat si potomky u každého rodiče zvlášť.

Abych mohl ukládat rodinný strom je třeba stanovit kořen. Máme tedy dvě varianty. Kořenem bude jeden společný potomek, a nebo společná matka. Obě varianty jsou rovnocenné, protože ve výsledku v programu procházím strom nezávisle na kořenovém uzlu. Zvolil jsem jako superuzel matku, která dostala id 0.

Pro ukládání dat o rodině jsem vytvořil jednoduchý formát souboru. Na obrázku 1 vidíte následující položky:

0. ID
1. Jméno
2. (červeně) id matky
3. (žluté) id otce
4. (zelené) true – muž, false – žena
5. (modře) partner

Když víme počet položek z prvního řádku, algoritmus je pak jednoduchý:

1. načti atributy
2. když neexistuje odkazovaná osoba, vytvoř základ
3. přidej se jako dítě rodiče a nastav partnera

9

[1, Kocourek Petr, 4, 3, true, nil]  
[4, Kocourková Ivana, 8, 9, false, 3]

Obrázek 1: Formát souboru s rodinnými daty.

4. pokud ještě neexistuji, vytvoř se
5. opakuj krok 1 pro další osobu

### 3 Design

Aplikace bude konsolová. Při psaní do terminálu se hodí escape sekvence, které umožňují mazat konkrétní kusy obrazovky a tím dosáhnout určité interaktivity. Budou stačit pouhé dva způsoby vstupu a to id osoby a požadovaný vztah. V hlavičce programu se bude vypisovat formulovaná otázka.

Uvnitř programu zastupuje třída `NTree` rodinný strom. Tato třída je konstruována s parametrem souboru s rodinnými daty. Při vytvoření se soubor načte a uloží pole všech instancí třídy `Node`. Každý `Node` představuje jednu osobu. Procházení stromu je pak rychlé, protože, můžu využít jak tabulku tak odkaz na sousední uzel.

Další část je práce s expertními daty. Bylo třeba navrhnout jak reprezentovat rodinný vztah. Využil jsem predikátové logiky, která je dostatečně univerzální pro práci s množinou prvků. Definoval jsem například:

$$P(x, a) \wedge P(a, b) \wedge C(b, c) \wedge N(c, a) \rightarrow c \text{ je sourozenec rodičů}$$

Vystačil jsem si se třemi funkcemi říkájící:

**P(a,b)** rodiče termu  $a$  jsou  $b$ .

**C(a,b)** děti termu  $a$  jsou  $b$ .

**N(a,b)** term  $a$  neobsahuje prvky z  $b$ .

Soubor jsem navrhl podobně. Stačilo se zbavit závorek a konjunkcí, které budou platit vždy a není tedy nutné je ukládat. Soubor má formát podle obrázku 2.

```
2
[3,prarodič,dědeček,babička]
PxaPab
[4,sourozenec rodičů,strýc,teta]
PxaPabCbcNca
```

Obrázek 2: Formát souboru s expertními daty.

## 4 Shrnutí

Při psaní jsem několikrát lehce změnil návrh. V počátku jsem o expertním systému nevěděl vůbec nic. Ve výsledku jsem použil již ověřené postupy.