

Jak neuvíznout v bažinách

Daniel Zejda, David Ondráček, Zdeněk Kunc

Úvod

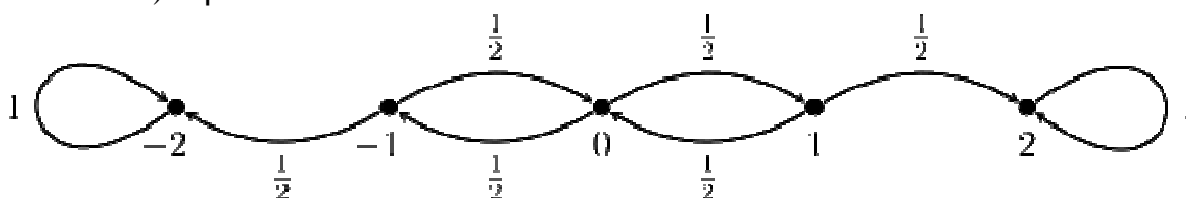
Tenhle miniprojekt nás zlákal zajímavým názvem, pod kterým si lze představit širokou škálu možností, které nemusí vždy nutně souviset s exaktními vědami. Napětí očekáváním jsme se nakonec dozvěděli, že se jedná o programování a následné vyhodnocování údajů, které jsme díky teorii matic, generátorům náhodných čísel a znalosti *Markovových absorbních řetězců* získali a následné porovnání s teoretickými hodnotami vypočítané spoustou matematických vzoreček. Doufali jsme v to, že se dozvíme spoustu zajímavých informací z oborů matematiky, statistiky, programování a dalších a vyzkoušet si tyto nové poznatky při kolektivní práci na větším projektu, který nám bude velkou motivací a v případě zvládnutí i zaslouženou hrdostí. Naše očekávání se splnila.

Markovův absorbní řetězec

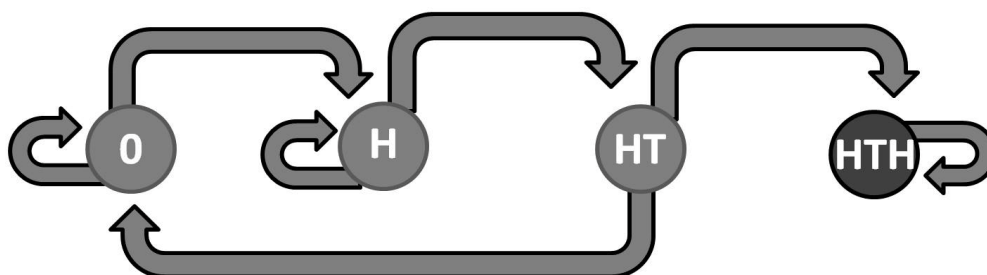
je jím v matematice chápán takový řetězec, ze kterého vždy vede cesta do absorbního stavu přes translační stavy v konečném počtu kroků.

Studované konfigurace

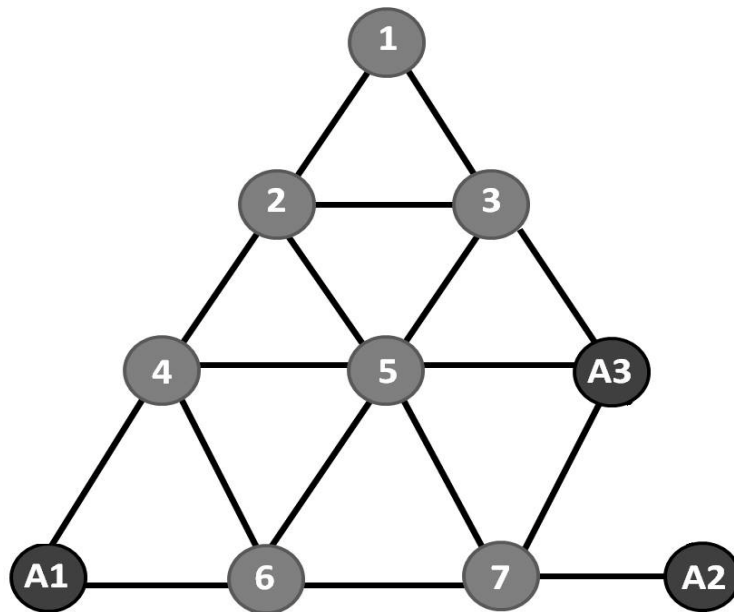
A) Opilec



B) Slova



C) Trojúhelník



Generátory náhodných čísel

Ke generování náhodných čísel jsme použili 3 generátory: Kiss, Xor a Lineárně kongruentní generátor. Testovali jsme jejich „náhodnost“ a používali jak ty volně dostupné na internetu, tak i vlastnoručně naprogramovaný.

Princip simulace

Díky teoretickým znalostem nalezených prostudováním anglické wikipedie jsme byli schopni vytvořit matici pravděpodobnosti přechodů mezi jednotlivými stavy daných problémů, ty jsme poté implementovali do kódu, kde jsme je dále vyhodnocovali. Za pomoci náhodně generovaného čísla, z celkem tří generátorů, které jsme postupně testovali, jsme milionkrát nechali vypočítat ve kterém bodě průběh skončí a kolik kroků bude na toto potřebovat. Z těchto naměřených hodnot jsme dosazením do vzorců pro vypočítání odchylky získali potřebné údaje k jasnému rozhodnutí, že naše generátory byly dostatečně „náhodné“ a výsledky se liší pouze o zanedbatelnou míru od matematicky vypočítaných předpokladů.

Ukázka ze zdrojového kódu, jak se pomocí kumulativních součtů a náhodně generovaného čísla určí, na základě pravděpodobnosti přechodu mezi dvěma stavy, následující stav

```
int random = RANDX();
for (int i = 0 ; i < nt + na; i++)
{
    if( i == 0)
    {
        probCumul[i] = prob[ i ];
    }
    else
    {
```

```

        probCumul[ i ] = prob[ i ] + probCumul[ i - 1 ];
    }

}
Snew = 0;
for(int i =0 ; i < nt + na; i++)
{
    if ( random >= probCumul[ i ] )
    {
        Snew = Snew + 1;
    }
}
if(Snew >= nt)
{
    S1 += (j+ 1);
    S2 += ((j+ 1)*(j+ 1));
    return( Snew - nt + 1 );
}
S = Snew;

```

Experimentální část

A) Pravděpodobnost absorce

Při výpočtech byl využit chi kvadrant test, který zjišťuje relativní odchylku od matematicky vypočítaných údajů, pro dané problémy

| generátor | opilec | slovo | trojuhelník |
|-----------|--------|-------|-------------|
| KISS | 0,76 | x | 0,83 |
| XOR | 0,49 | x | 0,14 |
| LCG | 0,85 | x | 0,59 |

B) Střední doba pohybu

Při výpočtech byl využit z-skore test, který porovnává hodnoty průměrného počtu různých stavů před absorbcí a údajů z matematických výpočtů

| generator | opilec | slovo | trojuhelnik |
|-----------|--------|-------|-------------|
| KISS | -0,32 | -1,98 | -0,29 |
| XOR | 0,97 | 0,29 | 2,97 |
| LCG | 0,8 | 2,03 | 0,82 |

Závěr

K testování pravděpodobnosti absorpce byl použit chi kvadrát test dobré

shody. Při něm jsou porovnány teoretické četnosti s experimentálně určenými.

$$\chi^2 = \sum_{i=1}^k \frac{(X_i - Np_i)^2}{Np_i}$$

Při tom využíváme vztah:

,kde n_k je počet případů. N je celkový počet experimentů. p_k jsou teoretické pravděpodobnosti. Provedli jsme milion pokusů, teoretické pravděpodobnosti jsme určili z matice B . Při testování jsme použili 5% hladinu významnosti. Výsledky jsou shrnuty v tabulce jako příslušné p hodnoty. Testování u úlohy se slovy nebylo využito, neboť má pouze jeden absorpční stav. Vzhledem k tomu, že všechny hodnoty v tabulce jsou vyšší, než 0,05, nemůžeme zamítnout hypotézu o shodě teoretických a experimentálních četností. To znamená, že všechny 3 generátory na uvedených 2 úlohách obstály. Dále byla testována střední doba pohybu po přechodových stavech. Při tom jsme určili z -skóre ze

$$z : x_i \rightarrow \frac{x_i - \bar{x}}{\sigma(x)}$$

vztahu:

,kde \bar{x} průměrné je průměrem z milionu simulací, $\sigma(x)$ teoretické je dáno vektorem t a směrodatná odchylka průměru je rovněž určena z experimentu. Výsledné hodnoty z -skóre jsou vedeny pro všechny 3 úlohy v tabulce. Kritická hodnota z -skóre při hladině významnosti 0,05 je rovna 1,960. K překročení kritické hodnoty došlo u každého generátoru pouze v 1 případě. Vzhledem k tomu, že hodnoty z -skóre jsou vždy menší než 3, nepovažujeme takové překročení za mimořádné.