Zkoumání parametrů řešiče Rubikovy kostky

Viktor Číhal

Úvod

V této práci jsem se rozhodl zkoumat určité parametry svého řešiče Rubikovy kostky. Zdrojový kód a notebook s experimenty se dá nalézt na `https://github.com/Reblexis/rubik-solver`.

Popis řešiče

Algoritmus

Řešič funguje na základě opakovaného prohledávání stavového prostoru metodou DFS s dodatečným přidáváním náhodných tahů, pokud není nalezen zlepšující stav kostky. K rozhodnutí o tom, zda je nalezen zlepšující stav, se používá metrika, počítající správné kostičky v kostce. Hledání je rozděleno do 2 fází. V první fázi se hledá řešení, které dostane kostku do podgrupy G1 a ve druhé fázi se pomocí menšího počtu různých tahů (a tedy i větší hloubky) hledá již úplné řešení.
Řešič tedy většinou nenalezne nejkratší řešení.

Parametry

Při spouštění lze nastavovat tyto parametry:

- Počet tahů zamíchání kostky (n)
- Časový limit pro hledání řešení (v milisekundách) (t)
- Maximální hloubka hledání v první fázi ($d1$)
- Maximální hloubka hledání ve druhé fázi ($d2$)

Více informací se nachází v README.md souboru.

Vliv počtu tahů při míchání na dosažené skóre

Rozhodl jsem se prozkoumat tento vztah, neboť se hodí vědět kolik stačí tahů při míchání, aby byla kostka ‘dostatečně náhodná’.

Graf závislosti

Skóre je počítáno jako počet správně umístěných kostiček v kostce (tedy na správné pozici a zároveň správně otočených). Pro každou testovanou hodnotu n spustíme 100-krát řešič a ostatní parametry zafixujeme na $t = 100$, $d1 = 4$ a $d2 = 5$.
Algoritmus sice řeší kostku jiným způsobem než člověk, ale je založený na podobném principu postupného přesouvání kostiček na správné pozice.

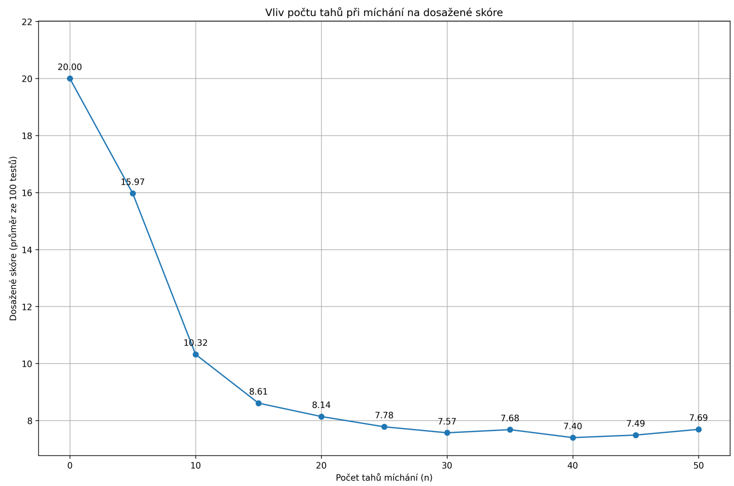

Figure 1: Vliv počtu tahů při míchání na dosažené skóre

Z grafu můžem vypozorovat, že 15 míchacích tahů nejspíš nestačí a od 20 míchacích tahů by již kostka mohla být dostatečně náhodná, ale stále je nějaký rozdíl mezi 20 a více tahy zamíchání.
Jelikož je maximální možný počet tahů k nejkratšímu vyřešení kostky 20 (God's Number), mohl by tento počet stačit.

Test normality distribuce

Raději ale tuto hypotézu otestujeme. Nejprve provedeme test, zda mají distribuce skóre normální rozdělení. Provedeme tedy Shapiro-Wilkův test s nulovou hypotézou, že data pocházejí z normálního rozdělení.
Spustíme tedy 300-krát řešič jednotlivě pro hodnoty $n = 15$, $n = 20$ a $n = 50$ a zafixujeme $t = 100$, $d1 = 4$ a $d2 = 5$. Z testu nám vyšla p-hodnota výrazně menší než 0.05, což znamená, že zamítáme nulovou hypotézu a můžeme tedy říct, že distribuce skóre není normální.

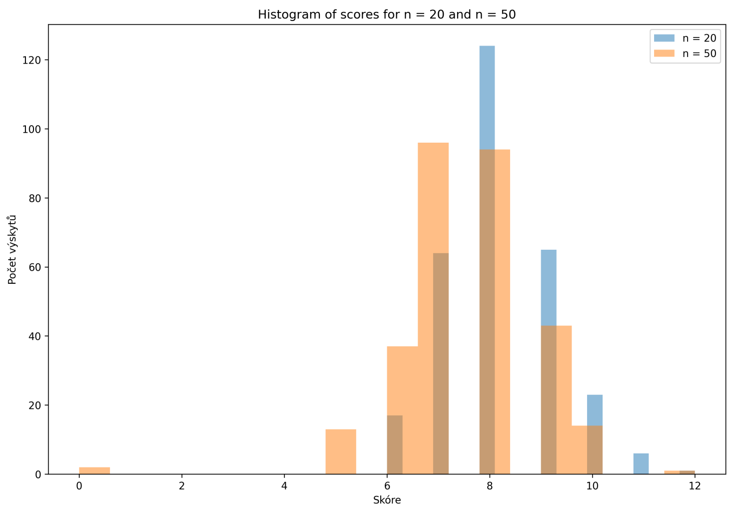

Figure 2: Histogram skóre pro $n = 15$, $n = 20$ a $n = 50$

KS test

Jelikož distribuce není normální, použiji dvouvýběrový KS test, kde porovnám postupně $n = 15$ a $n = 20$ s $n = 50$.