

# Sudoku na 100%

Vypracoval: František Kodídek

# Téma

Za téma jsem si vybral řešení sudoku pomocí Constrain Satisfaction, přesněji Constrain Satisfaction Problem (CSP).

CSP je definováno pomocí konečné sady proměnných, funkce která mapuje každou proměnou na konečnou doménu a konečné sady omezení.

Nejdříve je potřeba nadefinovat Sudoku jako CSP. Na to jsou potřeba 3 vlastnosti: proměnná, doména a omezení.

Každá proměnná je část problému, které je potřeba přiřadit odpovídající hodnotu.

Doména určuje, které hodnoty lze přiřadit konkrétní proměnné, a nakonec omezení které určuje jaké existující hodnoty v doméně mohou být použity.

- **Proměnná:** Prázdný „čtvereček“ v sudoku
- **Doména:** Bude jakákoliv hodnota od 1 do 9.
- **Omezení:** Stejně číslo se nesmí vyskytovat v řádku, sloupci a 3x3 čtverci

Dále jsem také použil AC-3 (Arc Consistency Algorithm #3). Na CSP lze také pohlížet jako orientovaný graf, kdy stavy jsou proměnné s hranami mezi nimi jako symetrické omezení. AC-3 právě tato spojení zkoumá a odstraní ty hodnoty z domény, které neodpovídají omezením mezi dvěma body. Jestli jsou nějaké hodnoty odstraněny z domény všechny hrany k odstraněné hodnotě jsou přidány do kolekce. Jelikož je množina domény a množina hran konečná nebo aspoň jedna hodnota je při jednom kroku odstraněna, je zaručeno konečnost algoritmu.

## Implementace

Generování sudoku je sám o sobě také problém, jelikož jednoduše může dojít ke špatnému vyplnění hodnot, což má za následek, že vygenerované sudoku nepůjde vyřešit. Jestli bude zase vyplněno málo políček, může řešení sudoku trvat velmi dlouho. Po vygenerování sudoku je z „desky“ vytvořen CSP. Dále je CSP předáno AC3 pro optimalizaci. Pokud nějaké nejasnosti po AC3 stejně zůstaly. Jsou vyřešeny pomocí backtracking search.

## Experiment

Řešení sudoku je velmi náchylné na nesplnitelnost a čím více vodítek vytvoříte tím se snižuje šance na splnitelnost sudoku. I řešení splnitelných sudoku je velmi variabilní kdy například AC3 má komplexitu až  $ed^3$  (e je počet hran a d je největší doména). Zkoušel jsem vytáčet průměrné hodnoty řešení, ale variabilita byla příliš obrovská a tím pádem průměr z 20 vzorků by mohl trvat velmi dlouho a nejednalo by se o „kvalitní“ průměr

