C2142 Návrh algoritmů pro přírodovědce

12. Těžké problémy.

Tomáš Raček

Typy problémů

V rámci teoretické analýzy nejčastěji rozlišujeme dva typy problémů:

Rozhodovací problém

- ověření, zdali něco platí, nebo ne
- př. Existuje v grafu G cesta mezi vrcholy s a t délky nejvýše 10?
- odpověď: ANO $\times$ NE

Optimalizační problém

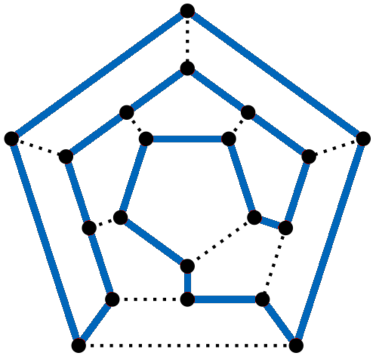
- cílem je nalezení nejlepšího řešení z množiny přípustných řešení
- př. Jaká je nejkratší cesta v grafu G mezi vrcholy s a t?
- odpověď: konkrétní nejkratší cesta $\times$ cesta neexistuje

Poznámka. Pokud existuje polynomiální algoritmus pro rozhodovací problém, existuje i pro jeho optimalizační variantu (a naopak).

Problém obchodního cestujícího (TSP)

Problém. Nalezněte nejkratší cestu, která prochází všemi zadanými městy a začíná a končí ve stejném městě.

Alternativní definice. Nalezněte v hranově ohodnoceném grafu Hamiltonovskou kružnici (= obsahující všechny vrcholy) minimální délky.



TSP – možnosti řešení

Hrubá síla. Vygeneruji a ověřím délky všech možných cest.

- složitost přístupu $\mathcal{O}(n!)$
- v praxi nepoužitelné

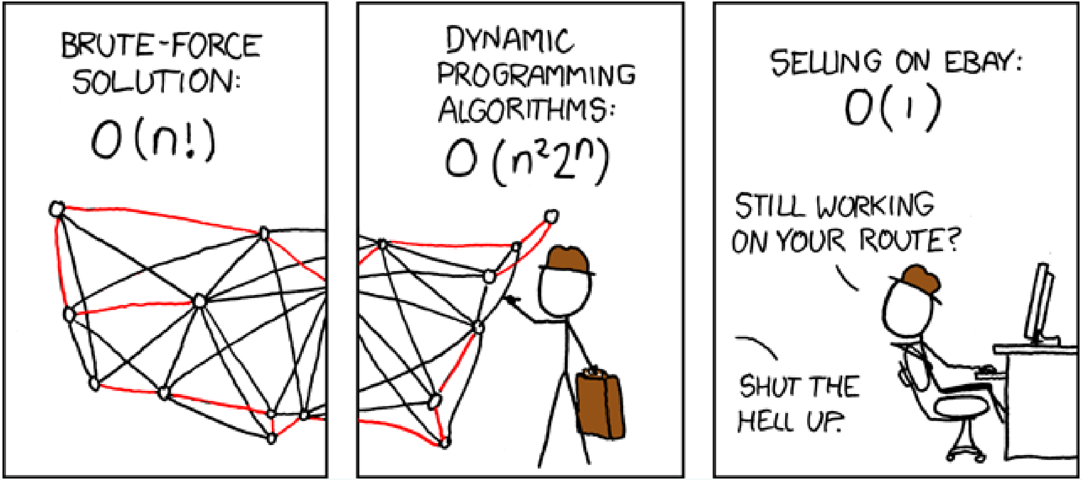
Dynamické programování

- výrazně netriviální
- složitost $\mathcal{O}(n^2 2^n)$

Aktuální stav řešení TSP.

- nevíme, zdali existuje algoritmus se složitostí nižší než $\mathcal{O}(2^n)$
- v roce 2006 se podařilo najít řešení pro instanci problému o velikosti 85 900 měst $\rightarrow$ 136 CPU let výpočtů

xkcd.com/399



BRUTE-FORCE SOLUTION:
O(n!)
DYNAMIC PROGRAMMING ALGORITHMS:
O(n²2ⁿ)
SELLING ON EBAY:
O(1)
STILL WORKING ON YOUR ROUTE?
SHUT THE HELL UP.

Třídy problémů

Pozorování. Velká část dosud prezentovaných problémů byla bez větších problémů prakticky řešitelná. Opakem je například TSP.

V rámci teorie pak můžeme přemýšlet, zdali lze problémy dělit do kategorií podle složitosti jejich řešení.

Nejčastěji rozlišujeme dvě třídy problémů:

- **P** třída problémů řešitelných v polynomiálním čase
- **NP** třída problémů, pro které lze ověřit řešení v polynomiálním čase

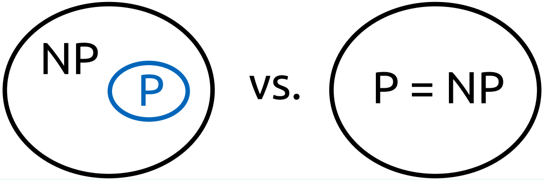
Poznámka. V rámci zařazování problémů do těchto tříd vždy uvažujeme jejich rozhodovací varianty.

Příklad. TSP je ve třídě NP, nejkratší vzdálenost v grafu je v P i v NP.

P vs. NP

Zamyšlení. Zjevně platí, že každý problém ve třídě P patří i do třídy NP, tedy $\mathrm{P} \subseteq \mathrm{NP}$.

Otázka. Platí to ale i naopak ($\mathrm{NP} \subseteq \mathrm{P}$)? Pokud ano, pak $\mathrm{P} = \mathrm{NP}$.



P vs. NP

- otevřený problém, jeden z největších v matematice a informatice
- jeden ze sedmi problémů milénia (Millennium Prize Problem $\rightarrow$ odměna 1 milion dolarů)

P = NP

If P = NP, then the world would be a profoundly different place than we usually assume it to be. There would be no special value in ’creative leaps,’ no fundamental gap between solving a problem and recognizing the solution once it’s found. Everyone who could appreciate a symphony would be Mozart; everyone who could follow a step-by-step argument would be Gauss...

Scott Aaronson, MIT

NP-úplné problémy

Pozorování. I v rámci třídy NP jsou problémy, které jsou různě těžké.

NP-úplné problémy jsou nejtěžší problémy ve třídě NP.

- každý problém v NP lze převést na NP-úplný problém v polynomiálním čase (existuje polynomiální redukce)
- rozhodovací varianta TSP je NP-úplný problém
- pro žádný NP-úplný problém není znám polynomiální algoritmus

Možnosti řešení P vs. NP

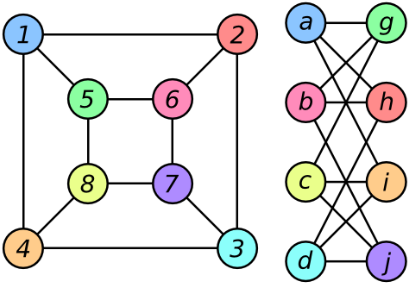
1. Ukázat, že pro některý NP-úplný problém nelze zkonstruovat polynomiální algoritmus. Pak $\mathrm{P} \neq \mathrm{NP}$.
2. Nalézt polynomiální algoritmus pro libovolný NP-úplný problém. Pak $\mathrm{P} = \mathrm{NP}$.

Problémy v NP – příklady

Zamyšlení. Předpokládejme $\mathrm{P} \neq \mathrm{NP}$. Existují problémy, které jsou v NP, ale nejsou NP-úplné?

Pravděpodobně následující:

- prvočíselný rozklad
- izomorfismus grafů



Prvočíselný rozklad

Úkol. Rozložte následující číslo na prvočísla:

1350664108659952233496032162788059699388814756056670275244851438515265106048595338339402871505719094417982072821644715513736804197039641917430464965892742562393410208643832021103729587257623585096431105640735015081875106765946292055636855294752135008528794163773285339061097505443349998111500569772368909275633

- ekvivalentní rozluštění RSA-1024
- odměna 100 000 dolarů
- soutěž skončila v roce 2007

Travelling salesman (2012)

Travelling Salesman

Drama / Mysteriózní / Thriller / Sci-Fi
USA, 2012, 80 min

Hrají: Steve West

Obsah

Čtveřice geniálních matematiků objeví v průběhu úspěšného výzkumu problému P versus NP algoritmus rapidně zrychlující výpočetní operace. Jejich objev může mít obrovské důsledky. Jak pozitivní, v podobě mohutné akcelerace biologického a medicínského vývoje, tak negativní, neboť nový algoritmus mj. umožňuje překonat moderní šifrování během několika vteřin. Poté, co vláda Spojených států nabídne každému z nich 10 milion dolarů za exkluzivní přístup k jejich části algoritmu, musí se čtveřice vypořádat s morálními i praktickými problémy, které jejich rozhodnutí přináší. *(Slaboproud)*

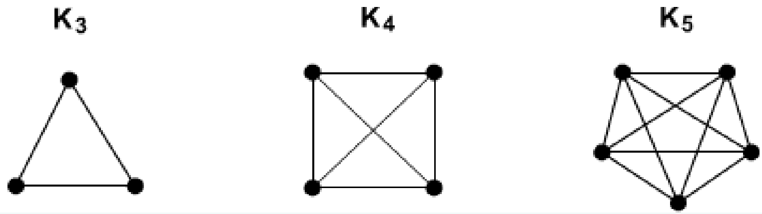
Vybrané příklady NP-úplných problémů I

Problém splnitelnosti výrokových formulí

- formule výrokové logiky s proměnnými $A_1, \ldots, A_n$
- Existuje přiřazení proměnných takové, že se zadaná formule vyhodnotí na TRUE?
- Příklad: $(\neg A_1 \vee A_2) \wedge A_3 \wedge \neg A_1$ je splnitelná např. pro $A_1 = 0$, $A_2 = 0$, $A_3 = 1$,

Klika

- Existuje v grafu klika (= podgraf, který je úplným grafem) o k vrcholech?



Vybrané příklady NP-úplných problémů II

Problém dvou loupežníků

- Lze rozdělit multimnožinu nezáporných čísel na dvě tak, že v obou bude součet obsažených čísel stejný?

Izomorfismus podgrafu

- Je graf H izomorfní nějakému podgrafu grafu G?

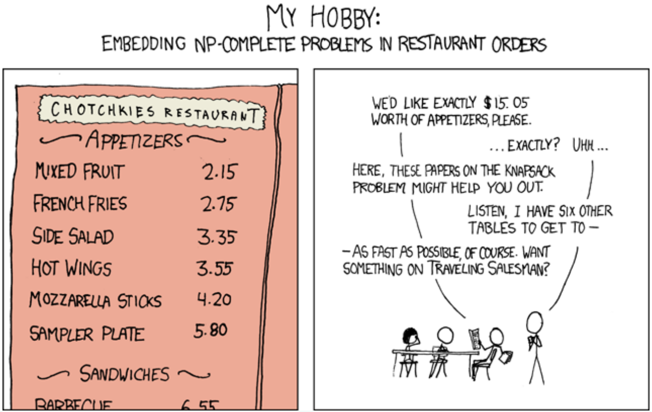
Problém batohu

- mějme batoh o nosnosti W a n předmětů, každý o hmotnosti w_i a hodnotě v_i
- Lze do batohu umístit předměty o celkové hodnotě alespoň V?

Součet podmnožiny

- Lze najít podmnožinu zadané množiny celých čísel takovou, že součet jejích prvků je nula?

xkcd.com/287



General solutions get you a 50% tip.

P vs. NP – poznámky

Možné výsledky

- $\mathrm{P} = \mathrm{NP}$, ale nejlepší algoritmus pro TSP se složitostí $\Omega(n^{100})$
- $\mathrm{P} \neq \mathrm{NP}$, ale algoritmus pro TSP se složitostí $\mathcal{O}(2^{0{,}00\ldots01\cdot n})$

Možnosti řešení

1. přijmout exponenciální algoritmus
2. omezit se na speciální případy (př. izomorfismus stromů je v P)
3. přijmout suboptimální řešení (užitím hladových algoritmů, heuristik)