

Cvičenia pre informatikov, 10.12.2020
Celočíslené lineárne programovanie

Praktické programy na NP ťažké problémy

Nájdu optimálne riešenie, mnohé inštancie zrátaju v rozumnom case, ale môžu bežať aj veľmi dlho

- ILP: CPLEX, Gurobi (komerčné), SCIP (nekomerčný)
- SAT: Minisat, Lingeling, glucose, CryptoMiniSat, painless
- TSP: Concorde

Iné NP ťažké problémy môžeme previesť na jeden z týchto problémov

ILP: Integer linear programming

Lineárne programovanie:

reálne premenné x_1, \dots, x_n

minimalizujeme $\sum_i a_i x_i$ pre dané váhy a_1, \dots, a_n

podmienky v tvare $\sum_i b_i x_i \leq c$

Dá sa riešiť v polynomiálnom čase

Celočíselné lineárne programovanie:

Pridáme podmienku, že niektoré premenné sú celočíselné alebo binárne

NP-ťažký problém

Zapíšme známe NP ťažké problémy ako ILP

Knapsack

Máme dané predmety s hmotnosťami $w_1 \dots w_n$ a cenami $c_1 \dots c_n$.
Ktoré z nich vybrať, aby celková hmotnosť bola najviac T a cena bola čo najvyššia?

Zapíšme známe NP ťažké problémy ako ILP

Set cover

Máme n podmnožín $S_1 \dots, S_n$ množiny $U = \{1 \dots m\}$.

Chceme vybrať čo najmenší počet zo vstupných množín tak, aby ich zjednotenie bola cela množina U .

Protein threading

Proteín A má známu sekvenciu aj štruktúru, proteín B iba sekvenciu.

Chceme zarovnať A a B tak, že ak sú dve aminokyseliny blízko v A , tak ich ekvivalenty v B by mali byť "kompatibilné".

V A určíme "jadrá", ktoré majú zostať zachované bez inzercií a delécií a v rovnakom poradí.

Jadrá sú oddelené slučkami, ktorých dĺžka sa môže ľubovoľne meniť a ktorých zarovnanie nebudeme skórovať.

Protein threading - formulácia

Vstup: sekvencia $B = b_1 \dots b_n$,

dĺžky m jadier $c_1 \dots c_m$,

skórovacie tabuľky

– E_{ij} : ako dobre $b_j \dots b_{j+c_i-1}$ sedí do sekvencie jadra i ,

– E_{ijkl} : ako dobre by jadrá i a k interagovali, keby začínali na poz. j, ℓ .

Úloha: zvoliť začiatky jadier $x_1 < x_2 < \dots < x_m$ tak, aby

– išli v správnom poradí a bez prekryvov,

– dosiahli maximálne možné skóre

Poznámka: nevravíme, ako zvoliť jadrá a skórovacie tabuľky, čo je modelovací, nie algoritmický problém

ILP

Označenie: sekvencia $B = b_1 \dots b_n$,

dĺžky m jadier $c_1 \dots c_m$,

skórovacie tabuľky

– E_{ij} : ako dobre $b_j \dots b_{j+c_i-1}$ sedí do sekvencie jadra i ,

– E_{ijkl} : ako dobre by jadrá i a k interagovali, keby začínali na poz. j, ℓ .

Riešenie:

