

07. Grådighet og stabil matching

Forelesning 7

Grådige algoritmer består av en serie med valg, og hvert valg tas *lokalt*: Algoritmen gjør alltid det som ser best ut her og nå, uten noe større perspektiv. Slike algoritmer er ofte enkle; utfordringen ligger i å finne ut om de gir rett svar.

Pensum

- Kap. 15. Greedy algorithms: Innl. og 15.1–15.3
- Kap. 25. Matching in bipartite graphs: 25.2

Læringsmål

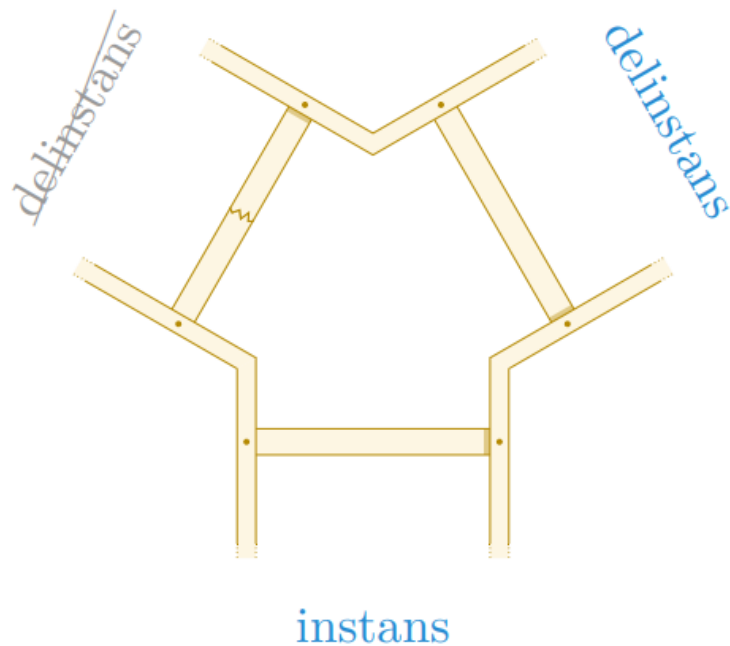
- [G₁] Forstå designmetoden *grådighet*
- [G₂] Forstå *grådighetsegenskapen* (*the greedy-choice property*)
- [G₃] Forstå eksemplene *aktivitet-utvelgelse* og *det kontinuerlige ryggsekkproblemet*
- [G₄] Forstå HUFFMAN og *Huffman-koder*
- [G₅] Vite hva en stabil matching er
- [G₆] Forstå GALE-SHAPLEY
- [G₇] Forstå hvorfor GALE-SHAPLEY finner beste matching for én side og verste for den andre

Grådighet

Grådighet er en strategi for å løse problemer ved å ta grådige valg tidlig, og senere bevise at valget man tok var det optimale. Hittil har vi brukt forskjellige egenskaper ved problemet til å bestemme hvilken strategi vi vil bruke for å løse det.

- Har problemet klare steg som hverken overlapper eller splitter seg? Inkrementelt design.
- Har problemet uavhengige delproblemer? Splitt og hersk.
- Har problemet overlappende problemer? Dynamisk programmering.

I tilfeller hvor dynamisk programmering må foreta valg, kan en grådig løsning ta et valg så lenge man kan bevise at det var det rette valget.



Vi bygger kun videre i den mest lovende retningen

Det er ikke mulig å bruke grådighet i alle løsninger. Problemet (og delproblemene) må oppfylle grådighetsegenskapen:

- Den globale optimale løsningen kan nås ved å ta et optimalt lokalt valg.
- Ved å ta et optimalt lokalt valg, må vi ikke ha eliminert alle mulige optimale løsninger

Dessuten vil det ikke gi mening å bruke grådighet på et problem uten optimal delstruktur.

Om vi ikke har ...

1. **Grådighetsegenskapen**
vil et grådig valg kunne gjøre at vi ikke lenger har håp om optimalitet
2. **Optimal delstruktur**
vil vi kunne måtte løse ting på en helt annen måte etter første valg

Kontinuerlige ryggsekkproblemet

Det kontinuerlige ryggsegg bygger videre på det binære ryggsekkproblemet. Forskjeller er at du nå har mulighet til å velge antallet av verdien istedenfor å enten ta eller forkaste verdien.

Løsningen til det kontinuerlige ryggsekkproblemet er å ta så mye av det med høyest verdi som mulig, fordi et annet valg vil aldri klare å ta igjen det grådigste.

Input: Verdier v_1, \dots, v_n , vekter w_1, \dots, w_n og en kapasitet W .

Output: Indekser i_1, \dots, i_k og en fraksjon $0 \leq \epsilon \leq 1$ slik at $w_{i_1} + \dots + w_{i_{k-1}} + \epsilon \cdot w_{i_k} \leq W$ og totalverdien $v_{i_1} + \dots + v_{i_{k-1}} + \epsilon \cdot v_{i_k}$ er maksimal.



66



Aktivitetsutvalg

Aktivitetsutvalgproblemet handler om å velge de beste tidene på en timeplan av et utvalg med aktiviteter slik at timeplanen er så full som mulig. Dette kan for eksempel være planlegging av en forelesningsal med forskjellige forelesninger og andre aktiviteter som har lyst til å bruke salen. Da gjelder det å velge de beste aktivitetene slik at salen er brukt så mye som mulig.

Input: Intervaller $[s_1, f_1), \dots, [s_n, f_n)$.

Output: Flest mulig ikke-overlappende intervaller.

Løsningen vil være å velge aktiviteten som slutter tidligst.

Huffman

Huffman er en komprimeringsalgoritme som lager en enkoding basert på frekvensen til bokstavene eller tegnene i teksten. Målet med Huffman enkodingen er å lage en enkoding med kortest kodelengde. Dette gjøres ved å bygge et Huffman-tre og deretter traversere treet for å generere enkodingen basert på inputen.

Input: Alfabet $C = \{c, \dots\}$ med frekvenser $c.freq$.

Output: Binær koding som minimerer forventet kodelengde $\sum_{c \in C} (c.freq \cdot length(code(c)))$.

Det optimale treet gir bokstaver som blir brukt minst de lengste kodene.

Gale-Shapley

Gale-Shapley er en stabil matchingalgoritme som matcher kvinner og menn (eller andre entiteter som studenter og skoler). En matchingalgoritme er stabil om det ikke finnes noen blokkerende par. Et blokkerende par er et par som foretrekker hverandre ovenfor parteren de har blitt tildelt.

Gale-Shapley er designet slik at matchingen er optimal for kvinnene, og pessimal for mennene.