TDT4125 Algoritmekonstruksjon

Eksamen, 6. juni 2024, 09:00–13:00

Faglig kontakt Magnus Lie Hetland

Hjelpemiddelkode E

Løsningsforslag

Løsningsforslagene i rødt nedenfor er *eksempler* på svar som vil gi full uttelling. Det vil ofte være helt akseptabelt med mange andre, beslektede svar, spesielt der det bes om en forklaring eller lignende. Om du svarte noe litt annet, betyr ikke det nødvendigvis at du svarte feil!

1 Hva er dualen til følgende lineærprogram?

$$\begin{array}{lrl} \max & 2x_1 + 3x_2 & \\ \text{s.t.} & x_1 + x_2 & \leqslant 4\,, \\ & 2x_1 + x_2 & \geqslant 5\,, \\ & x_2 & \geqslant 0\,. \end{array}$$

$$\begin{array}{lrl} \min & 4y_1 + 5y_2 & \\ \text{s.t.} & y_1 + 2y_2 & = 2\,, \\ & y_1 + y_2 & \geqslant 3\,, \\ & y_1 & \geqslant 0\,, \\ & y_2 & \leqslant 0\,. \end{array}$$

2 Forklar Hopcroft–Karp med egne ord.

Her trenger man ikke en veldig grundig forklaring. Det viktigste er å få frem hvordan den baserer seg på forøkende stier, og at den finner flere (disjunkte) slike samtidig (som er det som gjør den mer effektiv enn Edmonds–Karp).

3 Forklar den ungarske metoden, og hvordan den kan ses på som en versjon av (eller forløper til) primal–dual-metoden.

Her trenger man ikke en veldig grundig forklaring. Det viktigste er å få frem hvordan prisingen på nodene fungerer, dvs., at disse «betaler» for kanten sin, og at man må sørge for å ikke betale for mye, etc., og hvordan disse prisene fungerer som dualvariable.

4 I ryggsekkproblemet (*the knapsack problem*) har man oppgitt vekt og verdi for hver av et sett med gjenstander, samt en kapasitet (maksimal totalvekt) for ryggsekken. Man ønsker å velge en delmengde av gjenstandene som maksimerer den totale verdien, uten å gå over kapasiteten. (Du kan anta at alle gjenstandene får plass i ryggsekken hver for seg.)

Én strategi er å sortere gjenstandene (synkende) etter verdi per vekt, og så velge gjenstander grådig. På et punkt kommer man til en gjenstand man ikke får plass til. Hvis man har lov til å dele opp denne gjenstanden (*fractional knapsack*), vil man da ende med en optimal løsning, men anta at du ikke får lov til det (*0-1 knapsack*). Vurder da følgende to strategier:

Strategi 1: Dropp gjenstanden, og let videre etter gjenstander som får plass (om noen).

Strategi 2: Enten dropper du gjenstanden, eller så tar du den med, og dropper heller alle gjenstander du hadde valgt så langt. Velg det alternativet som gir høyest verdi. Etter dette, fortsetter du grådig, som i strategi 1.

Diskuter de to strategiene. Hva kan du si om ytelsesgarantiene (*performance guarantees*) deres som approksimasjonsalgoritmer?

Merk at det her *ikke* er snakk om *the minimum knapsack problem*, som er diskutert i pensum. Du kan anta at vekter, verdier og kapasitet er positive heltall.

Den første strategien kan gi en vilkårlig dårlig ytelsesgaranti. For eksempel, anta at vi har to gjenstander, der $v_1 = w_1 = 1$ og $v_2 = W - 1, w_2 = W$, der v_i er verdien og w_i er vekten til gjenstand i, og W er kapasiteten. Da vil vi velge gjenstand 1 først, og få en totalverdi på 1, mens det optimale er en totalverdi på $W - 1$. Ytelsesgarantien blir da $1/(W - 1)$.

I den andre strategien har vi sett på k gjenstander som må gi oss mer enn optimal verdi tilsammen, siden de inneholder all den verdien vi får med den grådige løsningen for den fraksjonelle versjonen av problemet. Vi vet at $1, \dots, k - 1$ får plass i ryggsekken, og det samme gjør k. Minst én av disse delmengdene må ha en verdi på minst halvparten av den optimale verdien, så vi får en ytelsesgaranti på $1/2$.

5 Diskuter forskjeller og likheter mellom ytelsesgarantiene vi gir for vanlige approksimasjonsalgoritmer og for onlinealgoritmer (dvs., *approximation ratio* vs. *competitive ratio*).

Her er det sentrale at man for approksimasjon sammenligner resultatet fra den polynomiske approksimasjonsalgoritmen med optimum (for samme problem), mens man for onlinealgoritmer sammenligner resultatet fra onlinealgoritmen med optimum for *offline*-varianten av problemet, der hele instansen er tilgjengelig.

6 I pensum beskrives en 1/2-approksimasjonsalgoritme for MAX CUT. Det nevnes også at denne algoritmen kan derandomiseres. Forklar kort hvordan dette kan gjøres ved å bruke metoden for betingede forventninger (*the method of conditional expectations*).

Algoritmen går ut på å plassere nodene tilfeldig i to delmengder. For hver kant er da sannsynligheten 1/2 for at den krysser snittet, så det forventede antallet blir $|E|/2 \geqslant \mathrm{OPT}/2$.

Derandomisering blir som for MAX SAT, som nevnt (men ikke forklart i detalj) i pensum. Når du velger hvilken mengde du legger en node i, anta at de gjenværende fortsatt skal plasseres tilfeldig, og maksimer den resulterende forventningen.

Vi vil alltid forvente at halvparten av kantene tilknyttet noder som ikke er plassert ennå vil havne over snittet, og dette påvirkes ikke av valget vi gjør. Vi ender dermed med en grådig algoritme, der vi plasserer hver node i den mengden der den har færrest naboer.

7 Du skal løse en versjon av mengdedekkeproblemet (*the set cover problem*), der hvert element har oppgitt hvor mange mengder det minst må dekkes av. (I den vanlige utgaven av problemet, er dette antallet 1 for alle elementer.)

Hvis du følger strategien fra pensum, der du løser dette som et lineærprogram uten heltallsrestriksjoner, og runder opp primalvariable der dualrestriksjonene er stramme, hva blir ytelsesgarantien? Forklar ved hjelp av tilnærmet komplementær slakkhet (*approximate complementary slackness*).

Dette blir som for for den vanlige utgaven av problemet, bortsett fra at b_i ikke lenger blir 1 for alle i. Dermed er ikke feilen lenger gitt av maksfrekvensen $\max_i f_i$, men av $\alpha = \max_i f_i/b_i$, siden dette er det største avviket en primalrestriksjon kan ha fra å være oppfylt ved likhet. Betingelsene for komplementær slakkhet er da oppfylt innen en faktor α, siden de primale restriksjonene *alltid* er oppfylt innen en faktor α, og de duale restriksjonene alltid er oppfylt ved likhet der vi har positive primalvariable.

8 I en versjon av ryggsekkproblemet er gjenstandene plasert i en sekvens, og hver gjenstand har en *farge* i stedet for en verdi. Du ønsker nå å avgjøre om det er mulig å få med seg et sett med gjenstander slik at blant dem som *er igjen*, så ligger alle gjenstander av samme farge ved siden av hverandre. (Etter at man har tatt med seg gjenstander, skal altså de m gjenværende gjenstandene av en gitt farge ligge på indekser $i, i+1, i+2, \ldots, i+m-1$, for en eller annen i.)

Gjenstandene har fortsatt også en vekt, og ryggsekken har en kapasitet (en maksimal totalvekt den kan bære).

Beskriv en polynomisk problemkjerne (*kernel*) for dette problemet, der kapasiteten er parameteren.

En mulig fremgangsmåte er å *slå sammen* gjenstander, der to gjenstander ved siden av hverandre erstattes med en ny gjenstand, som har en vekt lik summen av de to opprinnelige vektene.

Regel 1: Hvis man har to gjenstander av samme farge etter hverandre, slå dem sammen, og behold fargen.

Regel 2: Hvis to gjenstander ved siden av hverandre har farger som bare forekommer én gang, slå dem sammen og gi dem fargen til (f.eks.) den første gjenstanden.

Om man anvender disse reglene så lenge det er mulig, vil ingen nabogjenstander ha samme farge, og vi vil aldri ha to nabogjenstander der begge har en farge som forekommer bare én gang.

Hvis vi fjerner én gjenstand, kan antallet farger som forekommer flere ganger reduseres med maksimalt 2:

Regel 3: Hvis vi har minst $2k+1$ gjenstander med farger som forekommer mer enn én gang, er svaret *nei*.

Ellers kan vi maksimalt ha $2k+2$ gjenstander med farger som forekommer én gang, siden disse må skilles av gjenstander med farger som forekommer flere ganger. Vi har altså maksimalt $4k+3$ gjenstander, som hver krever $\lg(k+1)$, så størrelsen er $O(k \lg k)$.

9 En *partisjonsmatroide* er et mengdesystem $M = (E, S)$ der E er partisjonert i k *kategorier* $E_1, E_2, \ldots, E_k$, og hver kategori E_i har en terskel t_i. En delmengde $X \subseteq E$ er uavhengig (dvs., $X \in S$) når $|S \cap E_i| \leqslant t_i$ for alle i.

Vis at en partisjonsmatroide faktisk er en matroide.

Her kan man bruke en hvilken som helst karakterisering av matroider. En mulighet er å vise t for alle delmengder A av E, så har alle maksimale delmengder (basiser) samme kardinalitet. Dette følger direkte av at for hver delmengde E_i, vil antallet elementer som inngå i en maksimal uavhengig delmengde uansett være $\min\{|A \cap E_i|, t_i\}$.

10 Du organiserer en konferanse, og skal fordele plasser til ulike foredrag, der noen går samtidig. (Dvs., foredrag overlapper enten ikke, ellers så starter og slutter de samtidig.)

Du kan anta at alle deltakerne er enige om hvor bra de de ulike foredragene er, og at dette er representert ved en positiv verdi for hvert foredrag.

Du ønsker å fordele plassene rettferdig, og innfører derfor følgende regel: Hver agent får maksimalt én plass til et foredrag fra de k beste, deretter maksimalt én plass til et foredrag fra de k nest beste, osv.

Gitt denne regelen, vil du maksimere den totale verdien av foredragene som blir valgt. I første omgang begrenser du deg til å se på én av deltakerne. Hvor nært optimum kommer du om du velger grådig? Hva kan du si om en ev. polynomisk algoritme?

Husk at en deltaker ikke kan delta i to foredrag som går samtidig.

Dette er snittet av to partisjonsmatroider. Den ene har en kategori for hvert starttidspunkt (dvs., en mengde med overlappende foredrag), der hver kategori har terskel 1 (siden man ikke kan være i to foredrag samtidig).

Den andre har en kategori for hvert rangeringsnivå, der terskelen er 1 for de k beste, 1 for de k nest beste, osv.

Dermed kan man garantere at man får i hvert fall halvparten av optimum ($\alpha = 1/2$) ved grådig valg. Man kan også løse problemet i polynomisk tid, siden det finnes en polynomisk algoritme for optimering over snittet av to matroider.