

DD1352 Algoritmer, datastrukturer och komplexitet, hösten 2017

Lösningsförslag till mästarpöv 2: komplexitet

1. Samodling

Samodlingsproblemet som beslutsproblem:

INDATA: mängd frön F , mängd par av frön $P = \{\{f_i, f_j\}\}$ som inte kan odlas tillsammans, positivt heltal K (mål).

FRÅGA: Finns det K stycken frön $S \subseteq F$ som kan odlas tillsammans?

Vi bevisar att problemet ligger i NP genom att visa hur man lätt kan verifiera att en delmängd S beskriver en lösning. Kolla bara att $S \subseteq F$, att $|S| = K$ och att $s_1, s_2 \in S \Rightarrow (s_1, s_2) \notin P$. Detta går i polynomisk tid.

Vi visar sedan att problemet är NP-svårt genom att reducera problemet oberoende mängd till samodlingsproblemet.

OberoendeMängd($G = (V, E), K'$) =

$F \leftarrow V$

$P \leftarrow E$

$K \leftarrow K'$

return Samodling(F, P, K)

Reduktionen är polynomisk. Låt oss visa att den är korrekt, det vill säga att det existerar en oberoende mängd I med K hörn om och endast om det existerar en kryddodling S med K kryddor som kan odlas tillsammans.

Anta att I är en oberoende mängd i G med K hörn. Låt S vara I s motsvarande kryddor. Ta två godtyckliga kryddor f_i och f_j i S och visa att dom kan samodlas. Att I är en oberoende mängd betyder att

$$v_i, v_j \in I \Rightarrow (v_i, v_j) \notin E \Rightarrow \{f_i, f_j\} \notin P \Rightarrow f_i \text{ och } f_j \text{ kan samodlas.}$$

Åt andra hållet, anta att det existerar en kryddodling $S \subseteq F$ med $|S| = K$ och kryddor som kan odlas tillsammans. Då gäller att

$$f_i, f_j \in S \Rightarrow \{f_i, f_j\} \notin P \Rightarrow (v_i, v_j) \notin E \Rightarrow v_i \text{ och } v_j \text{ är oberoende.}$$

Därmed är NP-fullständigheten bevisad.

2. NP-fullständig avkastning

Avkastningsproblemet som beslutsproblem:

INDATA: positivt heltal d (odlingssäsongens längd), fröpar $F = \{f_i\}$ (där frö f_i är ett par (a_i, s_i) som anger avkastning a_i och växttid s_i), positivt heltal K (mål).

FRÅGA: Finns det ett odlingsschema $S = (S_1, \dots, S_5)$, där $S_i \subseteq \{1, \dots, |F|\}$ talar om vilka fröer som ska odlas i växthus i , så att $S_i \cap S_j = \emptyset$ för alla $i \neq j$, $\sum_{j \in S_i} (s_i + 1) \leq d + 1$, $\sum_{1 \leq i \leq 5} \sum_{j \in S_i} a_j \geq K$?

Vi bevisar att problemet ligger i NP genom att visa hur man lätt kan verifiera att ett odlingsschema S beskriver en lösning. Kolla bara att $S_i \subseteq \{1, \dots, |F|\}$, $S_i \cap S_j = \emptyset$, $\sum_{j \in S_i} (s_i + 1) \leq d + 1$, $\sum_{1 \leq i \leq 5} \sum_{j \in S_i} a_j \geq K$. Detta går i polynomisk tid.

Vi visar sedan att problemet är NP-svårt genom att reducera mängdpartitionering av positiva jämna tal till vårt problem. Vi kommer senare att bevisa att mängdpartitionering av positiva jämna tal är NP-fullständigt.

För varje tal i indata till mängdpartitioneringsproblemet skapar vi en fröpåse som har avkastningen 1 och odlingstiden ett mindre än talet. Eftersom det tar en dag att bereda jorden till nästa grönsak så kommer den totala odlingstiden i ett växthus att vara ett mindre än summan av motsvarande tal i mängdpartitioneringsproblemet. Att det är ett mindre beror på att jorden inte ska beredas efter att sista grönsaken skördats. Vi sätter därför odlingssäsongens längd d till ett mindre än den sökta mängdpartitioneringssumman (som är hälften av totala summan).

Eftersom alla tal i indata är positiva och jämna så är minsta möjliga talet 2, så det går bra att minska talet med ett och fortfarande ha ett positivt tal.

Idén är att låta odlingarna i dom två första växthusen motsvara lösningen till mängdpartitioneringsproblemet. Vi lägger därför till tre extra fröpåsar för att fylla ut odlingen i dom tre resterande växthusen. Dessa fröpåsar är alla märkta med avkastningen 1 och odlingstiden d .

MängdpartitioneringAvJämna($\{x_i\}_1^n$) = // alla x_i är positiva och jämna

```

 $\sigma \leftarrow \sum_1^n x_i$ 
 $F \leftarrow \emptyset$ 
 $d \leftarrow (\sigma/2) - 1$ 
for  $i \leftarrow 1$  to  $n$  do
     $F \leftarrow F \cup \{ \text{new Pair}(1, x_i - 1) \}$ 
 $F \leftarrow F \cup \{ \text{new Pair}(1, d), \text{new Pair}(1, d), \text{new Pair}(1, d) \}$ 
 $K \leftarrow n + 3$ 
return Avkastning( $d, F, K$ )

```

Reduktionen är polynomisk. Låt oss visa att den är korrekt, det vill säga att det existerar en mängdpartitionering av $\{x_i\}$ om och endast om det existerar ett odlingsschema $(S_1, S_2, S_3, S_4, S_5)$ som uppfyller villkoren.

Anta att $\{x_i\}$ partitioneras i $A \cup B$ där $\sum_{x \in A} x = \sum_{x \in B} x = \sigma/2$. Låt S_1 bestå av dom i som motsvarar $x_i \in A$. Låt S_2 bestå av dom i som motsvarar $x_i \in B$. Låt $S_3 = n + 1$, $S_4 = n + 2$, $S_5 = n + 3$ (fröpåsarna med d dagars odlingstid). Då är $S_i \cap S_j = \emptyset$, $\sum_{j \in S_1} (s_j + 1) = \sum_{\{x \in A\}} x = \sigma/2 = d + 1$ och dito för S_2 , $\sum_{1 \leq i \leq 5} \sum_{j \in S_i} a_j = n + 3 \geq n + 3$ eftersom det finns $n + 3$ grönsaker och alla har avkastningen 1. Därför har vi en ja-instans till avkastningsproblemet!

Åt andra hållet, anta att det existerar ett odlingsschema $(S_1, S_2, S_3, S_4, S_5)$ som uppfyller villkoren. Notera först att alla fröpåsar måste vara med i odlingsschemat eftersom den totala avkastningen är (minst) $n + 3$ och summan av alla fröpåsars avkastning är just $n + 3$.

Dom tre fröpåsarna $(1, d)$ fyller var och en upp en hela odlingssäsongen d , så dom måste odlas ensamma i var sitt växthus. Vi kan alltid numrera om växthusen så att dessa fröpåsar odlas i växthus 3, 4 och 5. Alla återstående påsar måste då odlas i växthus 1 och 2.

Låt $A = \{x_i : i \in S_1\}$ och $B = \{x_i : i \in S_2\}$. Då är $\sum_{\{x \in A\}} x = \sum_{j \in S_1} (s_j + 1) \leq d + 1 = \sigma/2$ och dito för B . Därför har vi en ja-instans för mängdpartitioneringsproblemet.

Det som återstår för att visa NP-fullständigheten för avkastningsproblemet är nu bara att visa att mängdpartitioneringsproblemet för positiva jämna tal är NP-svårt.

Vi reducerar det känt NP-fullständiga problemet SUBSET SUM (delmängdssumma)

Indata: En mängd positiva heltal P , positivt heltal K

Fråga: Finns det en delmängd av talen i P vars summa är K ?

Vi skapar en instans av mängdpartitioneringsproblemet för positiva jämna tal på följande sätt (i stort sett samma reduktion som visades på föreläsning 27):

SubsetSum($\{p_i\}_1^n, K$) =
 $\sigma \leftarrow \sum_1^n p_i$

```

 $X \leftarrow \emptyset$  //  $X$  är en multimängd
for  $i \leftarrow 1$  to  $n$  do  $X \leftarrow X \cup \{2p_i\}$ 
if  $K \neq \sigma/2$  then  $X \leftarrow X \cup \{2 \cdot |2K - \sigma|\}$ 
return MängdpartitioneringAvJämna( $X$ ) =

```

Anta att det finns en delmängd S av P med summa exakt K . Om $K = \sigma/2$ blir summan av återstoden av talen också $\sigma/2 = K$, varför vi direkt har en mängdpartitionering av talen i X (som ju bara är dubblingar av ursprungstalen).

Om $K > \sigma/2$ har utöver dom dubblerade ursprungstalen även talet $2 \cdot (2K - \sigma)$ lagts till i X . Summan av alla tal i X blir $2\sigma + 2 \cdot (2K - \sigma) = 4K$. Betrakta dom tal i X som härrör från talen i S . Eftersom varje tal är en dubbling av ursprungstalen så är summan av dessa tal $2K$, och summan av övriga tal är $4K - 2K = 2K$, dvs precis lika stor, varför vi har en mängdpartitionering.

Om $K < \sigma/2$ har utöver dom dubblerade ursprungstalen även talet $2 \cdot (\sigma - 2K)$ lagts till i X . Summan av alla tal i X blir $4\sigma - 4K$. Betrakta dom tal i X som härrör från talen i S och som enligt ovan har summan $2K$. Tillsammans med extratalet $2 \cdot (\sigma - 2K)$ blir summan $2\sigma - 2K$ och summan av övriga tal är $4\sigma - 4K - (2\sigma - 2K) = 2\sigma - 2K$, dvs precis lika stor, varför vi har en mängdpartitionering.

Åt andra hållet, anta att det finns en mängdpartitionering av X i X_1 och X_2 . Om $K = \sigma/2$ kan vi direkt läsa av delmängdssumman K från elementen i X_1 (halverade). Annars, anta vi att extratalet $2 \cdot (\sigma - 2K)$ har hamnat i X_1 . Om $K > \sigma/2$ är summan av alla tal i X $4K$, så summan av talen i X_2 är $2K$ och vi kan läsa av delmängdssumman K från elementen i X_2 (halverade). Om $K < \sigma/2$ är summan av alla tal i X $4\sigma - 4K$, så summan av alla tal i X_1 förutom extratalet är $(4\sigma - 4K)/2 - 2 \cdot (\sigma - 2K) = 2K$, så från dessa tal (halverade) kan vi läsa av delmängdssumman K .

Reduktionen går i polynomisk tid (närmare bestämt linjär tid), så problemet är alltså NP-fullständigt.

3. Konstruktion av plan för optimal avkastning

Problemet är att hitta den plan som maximerar avkastningen. Låt σ vara summan av avkastningarna för alla fröpåsar i indata.

För att hitta det optimala värdet gör vi en binärsökning med K mellan 1 och σ och anropar algoritmen Avkastning($\log \sigma$ gånger).

```

MaxAvkastning( $d, F$ ) =
   $\sigma \leftarrow 0$ 
  for  $(a, s) \in F$  do  $\sigma \leftarrow \sigma + a$ 
   $min \leftarrow 1$ ;  $max \leftarrow \sigma$ 
  while  $min < max$  do
    INV lösningen finns i intervallet  $[min..max]$ 
     $M \leftarrow \lceil (min + max)/2 \rceil$ 
    if Avkastning( $d, F, M$ ) then  $min \leftarrow M$  else  $max \leftarrow M - 1$ 
  return  $min$ 

```

När vi vet den optimala avkastningen ska vi ta reda på vilka fröer som ska odlas och sedan ta reda på vilka som ska odlas i samma växthus. För att få reda på vilka fröer som ska odlas plockar vi bort ett frö i taget och kollar om maximala avkastningen fortfarande är samma. I så fall låter vi fröet förbli bortplockat, annars stoppas det tillbaka igen. Därefter plockar vi ett godtyckligt frö som får bli det första frö som ska odlas i ett växthus. Vi provar att slå ihop detta frö med i tur och ordning alla andra fröer och låter fröinstanser vara så ifall avkastningen fortfarande är maximal. Annars ska båda hopslagna fröerna odlas i detta växthus och vi låter fröerna förbli hopslagna.

```

maxavkastning  $\leftarrow$  MaxAvkastning(d, F)
for f  $\in$  F do
    if Avkastning(d, F - {f}, maxavkastning) then F  $\leftarrow$  F - {f}
for i  $\leftarrow$  1 to 5 do
    INV Det finns en optimal odlingsplan för dom i - 1 första växthusen som ges av S1, ..., Si-1
    Si  $\leftarrow$  j, där fj är ett godtyckligt frö som fortfarande är kvar i F
    F  $\leftarrow$  F - {fj}
    for fk  $\in$  F do
        if Avkastning(d, F - {(a, s), (ak, sk)}  $\cup$  {(a + ak, s + sk + 1}, maxavkastning) then
            F  $\leftarrow$  F - {(a, s), (ak, sk)}  $\cup$  {(a + ak, s + sk + 1)}
            Si  $\leftarrow$  Si  $\cup$  {k}
            a  $\leftarrow$  a + ak; s  $\leftarrow$  s + sk + 1
    return {S1, ..., S5}

```

Algoritmen tar polynomisk tid och gör $\log \sigma + |F| + 5|F|$ anrop av *Avkastning*(), vilket är polynomiskt många i antalet bitar i indata.

Korrektheten för algoritmen inses på följande sätt. Binärsökningen efter optimala värdet är en standardalgoritm som bevisas med en invariant som säger att det sökta värdet finns mellan dom aktuella gränserna i binärsökningen.

Efter binärsökningen vet vi att det finns en lösning med värdet *maxavkastning* men ingen lösning med värdet *maxavkastning* + 1. Under resten av algoritmen utökar vi instansen med fler och fler krav på ett sånt sätt att det alltid finns en lösning med värdet *maxavkastning* som uppfyller kraven.

Resonemanget i samband med algoritmbeskrivningen ovan visar att kraven hårdnar allteftersom tills det bara finns en enda möjlig lösning som kan uppfylla kraven, och det är vår sökta odlingsplan.