

Algoritmer, datastrukturer och komplexitet

Övning 10

Anton Grensjö
grensjo@csc.kth.se

27 november 2015

Kursplanering

Ö9: NP-fullständighetsbevis

F27: NP-fullständighetsreduktioner

F28: NP-fullständighetsreduktioner

Ö10: NP-fullständighetsbevis

Bonusdatum labb 4 måndag 23/11

F29: Approximationsalgoritmer

F30: Approximationsalgoritmer

Ö11: Approximationsalgoritmer

Idag

- Approximationsalgoritmer

Repetition av viktiga begrepp

Vad gör man om man **måste** lösa ett NP-fullständigt problem?

1. Begränsa problemet - vissa indatatyper kan vara enklare.
 2. Lös för små indata med exponentiell algoritm.
 3. Använd en **approximationsalgoritm**, som garanterat ger en lösning som är nära den optimala.
 4. Använd en heuristik, som förhoppningsvis ger en bra lösning.
- **Approximationskvoten** är ett mått på hur bra lösningar en approximationsalgoritm ger.

$$\text{Minimeringsproblem: } \frac{APPROX}{OPT} \geq 1$$

$$\text{Maximeringsproblem: } \frac{OPT}{APPROX} \geq 1$$

Repetition av viktiga begrepp

- $APX = \{\text{Problem som kan approximeras inom någon konstant}\}$
- Det finns problem som $\notin APX$, som alltså inte kan approximeras inom någon konstant. Exempel:
 - TSP
 - Max klick
- Väntevärde
 - Definition: Om X är en diskret stokastisk variabel så gäller

$$E(X) = \sum_k k \cdot P(X = k)$$

- Egenskap: Väntevärdet är linjärt.

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

Approximation av oberoende mängd

INDEPENDENT SET-B: Problemet att hitta en maximal mängd oberoende hörn i en graf vars gradtal (i varje hörn) är högst B .

Visa att problemet ligger i *APX*.

Enkel approximationsalgoritm:

```

1: function APPROX-INDEPENDENT-SET-B( $V, E$ )
2:    $V' \leftarrow \emptyset$ 
3:    $W \leftarrow V$ 
4:   for  $v \in W$  do
5:      $V' \leftarrow V' \cup \{v\}$ 
6:      $W \leftarrow W - \{w \in W : (w, v) \in E\} - \{v\}$ 
7:   return  $V'$ 

```

Vilken approximationskvot får vi?

Approximation av oberoende mängd

- 1: $V' \leftarrow \emptyset$
- 2: $W \leftarrow V$
- 3: **for** $v \in W$ **do**
- 4: $V' \leftarrow V' \cup \{v\}$
- 5: $W \leftarrow W - \{w \in W : (w, v) \in E\} - \{v\}$

- Efter algoritmen är V' en oberoende mängd hörn.
- Vi ser att V' också är en dominerande mängd hörn.
- Vi vill visa att $|V_{opt}|$ är högst B gånger större än $|V'|$.
 - Betrakta ett hörn $v \in V'$. Om $v \notin V_{opt}$ så tillhör i värsta fall alla v 's grannar V_{opt} . Men dessa är som värst B stycken.
 - Detta gäller för varje hörn i V' .
 - Totalt måste alltså $|V_{opt}| \leq B \cdot |V'|$.
 - Vårt resonemang stämmer eftersom V' är en dominerande mängd. Det vill säga, varje hörn i V ligger antingen i V' eller är granne till en nod i V' . Därmed har vi täckt in alla fall.
- Slutsats: Problemet kan approximeras inom en faktor B .

Probabilistisk alla-inte-lik-satisfiering

MAX NOT-ALL-EQUAL 3-CNF SAT:

INMATNING: En CNF-formel som består av klausulerna c_1, c_2, \dots, c_m där varje klausul är en disjunktion av exakt tre literaler (variabler eller negerade variabler). Variablerna heter x_1, x_2, \dots, x_n .

LÖSNING: En variabeltilldelning.

MÅLFUNKTION: Antalet klausuler som innehåller minst en sann literal och minst en falsk literal.

PROBLEM: Maximera målfunktionen.

Detta problem är NP-svårt, så vi vill approximera det inom en konstant i polynomisk tid.

- Konstruera en probabilistisk approximationsalgoritm för problemet med förväntad approximationskvot $4/3$.
- Analysera tidskomplexitet och förväntad approximationskvot.

Probabilistisk alla-inte-lik-satisfiering

ALGORITHM: Sätt varje variabel till ett slumpmässigt valt värde, med uniform sannolikhetsfördelning.

APPROXIMATIONSKVOT:

A_j = händelsen att klausulen c_j har både sanna och falska literaler

$$X_j = \begin{cases} 1 & \text{om } A_j \text{ händer} \\ 0 & \text{annars} \end{cases}, \quad X = \sum_{j=1}^m X_j$$

$$E(X) = E\left(\sum_{j=1}^m X_j\right) = \sum_{j=1}^m E(X_j) = \sum_{j=1}^m \left(\sum_{k=0}^1 k \cdot P(X_j = k)\right) =$$

$$= \sum_{j=1}^m 1 \cdot P(A_j) = \sum_{j=1}^m \left(1 - \frac{2}{8}\right) = \sum_{j=1}^m \frac{3}{4} = \frac{3m}{4}$$

$$\frac{OPT}{APPROX} \leq \frac{m}{3m/4} = \frac{4}{3}$$

Linjär tid, linjär slump.

Approximation av linjära olikheter

MAX SAT LR \geq : Givet en uppsättning linjära olikheter av typen \geq , som t.ex:

$$\begin{cases} 2x_1 - 8x_3 + 3x_8 & \geq 3 \\ -4x_2 + x_3 - 2x_5 + 7x_7 & \geq -4 \\ x_4 + 2x_6 - 4x_7 & \geq 0 \end{cases}$$

hitta en variabeltilldelning som satisfierar så många olikheter som möjligt.

Konstruera en approximationsalgoritm som approximerar MAX SAT LR \geq inom faktorn 2.

Approximation av linjära olikheter

```

1: function APPROX-MAX-SAT-LR≥( $X, E$ )
2:   while  $E \neq \emptyset$  do
3:     if det finns olikheter i  $E$  med en enda variabel then
4:        $U \leftarrow \{x \in X : x \text{ är ensam variabel i minst en olikhet i } E\}$ 
5:       Välj godtyckligt  $y \in U$ .
6:        $F(y) \leftarrow \{e \in E : e \text{ bara innehåller variabeln } y\}$ 
7:       Ge  $y$  ett värde som satisfierar så många olikheter i  $F(y)$ 
           som möjligt.
8:        $E \leftarrow E - F(y)$ 
9:     else
10:      Välj godtyckligt  $y \in X$ 
11:       $y \leftarrow 0$ 
12:      Evaluera om olikheterna i  $E$  som innehåller  $y$ .
13:       $X \leftarrow X - \{y\}$ 
14:   return de gjorda tilldelningarna

```

Approximation av linjära olikheter

- Tilldelningen av y på rad 7 gör alltid att minst hälften av olikheterna i $F(y)$ är satisfierade. Varför?
- Alltså måste minst hälften av olikheterna i hela systemet vara satisfierade i slutet.
- Eftersom högst alla **kan** satisfieras blir approximationskvoten

$$\frac{OPT}{APPROX} = \frac{|E|}{\frac{1}{2}|E|} = 2$$

- Polynomisk tidskomplexitet.

Övre gräns för approximation av homogena bipolära olikheter

MAX HOM BIPOLAR SAT LR[≥]: Samma problem som MAX SAT LR[≥], men variablerna får bara anta värdena -1 och 1 och alla olikheter är homogena. Exempel:

$$\begin{cases} 2x_1 - 8x_3 + 3x_8 & \geq 0 \\ -4x_2 + x_3 - 2x_5 + 7x_7 & \geq 0, \quad x_i \in \{-1, 1\}, \\ x_4 + 2x_6 - 4x_7 & \geq 0 \end{cases}$$

Visa att:

- MAX HOM BIPOLAR SAT LR[≥] kan approximeras inom faktorn 2.
- MAX HOM BIPOLAR SAT LR[>] kan approximeras inom faktorn 4.

Övre gräns för homogena bipolära olikheter

- a) Visa att MAX HOM BIPOLAR SAT LR \geq kan approximeras inom faktorn 2.
- Tag en godtycklig bipolär vektor \mathbf{x} .
 - Varje olikhet satisfieras antingen av \mathbf{x} eller $-\mathbf{x}$.
 - Välj den av \mathbf{x} och $-\mathbf{x}$ som satisfierar flest. Detta är minst hälften.
- b) Visa att MAX HOM BIPOLAR SAT LR $>$ kan approximeras inom faktorn 4.
- Samma algoritm som för a) fungerar inte när vi har $>$ istället för \geq .
 - Dela in i två delproblem:
 - 1 Hitta en lösning för vilken minst hälften av vänsterleden är nollskilda.
 - 2 Olikheterna med vänsterskilt vänsterled kan nu behandlas som i a).
 - Modifiera algoritmen för MAX SAT LR \geq för att hitta en lösning \mathbf{x} där minst hälften av vänsterleden är nollskilda.

Övre gräns för homogena bipolära olikheter

b) Visa att MAX HOM BIPOLAR SAT LR[>] kan approximeras inom faktorn 4.

- Modifiera algoritmen för MAX SAT LR[≥] för att hitta en lösning x där minst hälften av vänsterleden är nollskilda.

```

1: while  $E \neq \emptyset$  do
2:   if det finns olikheter i  $E$  med en enda variabel then
3:      $U \leftarrow \{x \in X : x \text{ är ensam variabel i minst en olikhet i } E\}$ 
4:     Välj godtyckligt  $y \in U$ .
5:      $F(y) \leftarrow \{e \in E : e \text{ bara innehåller variabeln } y\}$ 
6:     Ge  $y$  ett värde som gör att vänsterledet är nollskilt för så
       många  $F(y)$  som möjligt.
7:      $E \leftarrow E - F(y)$ 
8:   else
9:     Välj godtyckligt  $y \in X$ 
10:     $y \leftarrow 0$ 
11:    Evaluera om olikheterna i  $E$  som innehåller  $y$ .
12:     $X \leftarrow X - \{y\}$ 

```

Övre gräns för homogena bipolära olikheter

- Med hjälp av den algoritmen får vi alltså ut en tilldelning \mathbf{x} sådan att minst hälften av olikheternas vänsterled är nollskilda.
- Givet ett sådant \mathbf{x} , betrakta också $-\mathbf{x}$.
- Enligt samma resonemang som för a) måste varje nollskild olikhet vara uppfylld antingen i \mathbf{x} eller i $-\mathbf{x}$.
- Välj den av \mathbf{x} och $-\mathbf{x}$ som uppfyller flest.

Approximationskvot:

$$\frac{OPT}{APPROX} = \frac{|E|}{\left(\frac{1}{2} \cdot \frac{1}{2}\right) |E|} = 4$$

Undre gräns för approximation av binära olikheter

MAX BINARY SAT LR[≥]: Samma problem som MAX SAT LR[≥], men där variablerna bara får anta värdena 0 och 1.

Visa att MAX BINARY SAT LR[≥] \notin APX.

- Reducera ett problem som inte ligger i APX till MAX BINARY SAT LR[≥] med en approximationsbevarande reduktion. Vilket?
- Vi reducerar MAX CLIQUE.
- Låt $G = (V, E)$ vara indata till MAX CLIQUE.
- För varje hörn $v_i \in V$, konstruera en variabel x_i och olikheten

$$x_i - \sum_{j \in N(v_i)} x_j \geq 1,$$

där $j \in N(v_i) \iff v_j \neq v_i \wedge (v_i, v_j) \notin E$.

- Notera: den i :te olikheten är sann om $x_i = 1$ och $x_j = 0 \ \forall j \in N(v_i)$.
- Om vi har en maximal lösning till systemet av olikheter, så är $\{v_i \mid \text{olikheten konstruerad från } v_i \text{ är satisfierad}\}$ en maximal klick.

Undre gräns för approximation av binära olikheter

- Om vi har en s -klick $V' \subseteq V$ så kan vi få en binär olikhet som satisfierar de s motsvarande variablerna genom att låta $x_i = 1$ om $v_i \in V'$, och $x_i = 0$ annars.
- Om vi har en binär lösning \mathbf{x} som satisfierar s olikheter så kan vi få en s -klick genom att låta V' bestå av alla hörn v_i som motsvarar satisfierade olikheter.
- Problemen är alltså ekvivalenta, och en maximal lösning av det ena ger en maximal lösning av det andra.
- Reduktionen är approximationsbevarande.
- Reduktionen är dessutom kostnadsbevarande, eftersom reduktionen bevarar målfunktionens värde.

Slutsats: $\text{MAX BINARY SAT LR}^{\geq} \notin \text{APX}$.

Vad kan vi säga om $\text{MAX BINARY SAT LR}^>$ och $\text{MAX BINARY SAT LR}^=?$

Nästa gång

- Blandade uppgifter från gamla tentor/muntor