DD2350 Algoritmer, datastrukturer och komplexitet

Uppgifter till övning 11

Approximationsalgoritmer

På denna övning är det också **inlämning av skriftliga lösningar av teoriuppgifterna till labb 5** och muntlig redovisning av teoriuppgifterna.

Approximation av oberoende mängd Låt INDEPENDENT SET$-B$ vara problemet att hitta en maximal mängd oberoende hörn i en graf vars gradtal (i varje hörn) är begränsat av konstanten B. Visa att detta problem ligger i APX, det vill säga kan approximeras inom någon konstant i polynomisk tid.

Probabilistisk alla-inte-lika-satisfiering Det NP-svåra problemet MAX NOT-ALL-EQUAL 3-CNF SAT definieras på följande sätt.

INMATNING: En CNF-formel som består av klausulerna $c_1, c_2, \ldots, c_m$ där varje klausul är en disjunktion av exakt tre literaler (variabler eller negerade variabler). Variablerna heter $x_1, x_2, \ldots, x_n$.
LÖSNING: En variabeltilldelning.
MÅLFUNKTION: Antalet klausuler som innehåller minst en sann literal och minst en falsk literal.
PROBLEM: Maximera målfunktionen.

Eftersom detta problem är NP-svårt så vill vi ha en algoritm som approximerar det inom en konstant i polynomisk tid. Konstruera en probabilistisk approximationsalgoritm för problemet som ger en förväntad approximationskvot på 4/3. Analysera din algoritms tidskomplexitet och förväntad approximationskvot. Du behöver inte slumpeliminera algoritmen.

Approximation av linjära olikheter MAX SAT LR$^{\geq}$ (maximum satisfiable linear subsystem) är problemet att, givet en uppsättning linjära olikheter av typen $\geq$ (till exempel $2x_1 - 8x_3 + 3x_8 \geq 3$), hitta en variabeltilldelning som satisfierar så många olikheter som möjligt. Konstruera en approximationsalgoritm som approximerar MAX SAT LR$^{\geq}$ inom faktorn 2.

Övre gräns för approximation av homogena bipolära olikheter
MAX HOM BIPOLAR SAT LR$^{\geq}$ (maximum homogeneous bipolar satisfiable linear subsystem) är samma problem som MAX SAT LR$^{\geq}$ men variablerna får bara anta värdena 1 och -1, och alla olikheter är homogena, det vill säga saknar konstanttermer (har noll i högerledet).

Visa att MAX HOM BIPOLAR SAT LR$^{\geq}$ kan approximeras inom faktorn 2 och att MAX HOM BIPOLAR SAT LR$^{>}$ kan approximeras inom 4.

Undre gräns för approximation av binära olikheter MAX BINARY SAT LR$^{\geq}$ (maximum binary satisfiable linear subsystem) är samma problem som MAX SAT LR$^{\geq}$ där variablerna bara får anta värdena 0 och 1.

Visa att MAX BINARY SAT LR$^{\geq}\notin$APX.

Lösningar

Lösning till Approximation av oberoende mängd
Givet en graf $G = (V, E)$ vars gradtal är begränsat av B, konstruera en oberoende mängd hörn på följande sätt.

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

Efter algoritmen är V' en oberoende mängd hörn. Det är också lätt att se att V' är en dominerande mängd. Den optimala oberoende mängden hörn V'_{opt} kan högst vara B gånger större än $|V'|$. För att se detta behöver man bara titta på ett av hörnen i V' och dess grannar i V. Om alla dess grannar är med i V'_{opt} så blir det B stycken, och värre kan det inte bli. Eftersom V' är en dominerande mängd måste varje hörn i V'_{opt} antingen vara med i V' eller vara granne till ett hörn i V'. Det betyder att V'_{opt} är högst B gånger större än V', det vill säga att approximationskvoten är B, som är en konstant. □

Lösning till Probabilistisk alla-inte-lika-satisfiering
Algoritmen är helt enkelt: sätt varje variabel till ett slumpmässigt valt värde. Det ska vara lika sannolikt att variabeln sätts till sant som till falskt. Om vi nu tittar på en godtycklig klausul så är det bara i två fall av åtta (nämligen falskt-falskt-falskt och sant-sant-sant) som den inte ska räknas. Väntevärdet för antalet klausuler som räknas är

$$\begin{aligned}
&E[\text{antal klausuler som har både sanna och falska literaler}] = \\
&= m \cdot Pr[\text{en godtycklig klausul är varken helt sann eller falsk}] = \\
&= m \cdot (1 - Pr[\text{en godtycklig klausul är antingen helt sann eller helt falsk}]) = \\
&= m \cdot (1 - 2/8) = 3m/4.
\end{aligned}$$

Eftersom högst m klausuler kan räknas blir den förväntade approximationskvoten

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

Algoritmen tar linjär tid och kräver linjär slump i antalet variabler. □

Lösning till Approximation av linjära olikheter
Anta att inmatningen ges som en mängd X med rationella variabler och en mängd E med linjära olikheter över X.

while $E \neq \emptyset$ **do**
    if (det finns olikheter i E med en enda variabel) **then**
        $U \leftarrow \{x \in X : x \text{ är ensam variabel i minst en olikhet i } E\}$
        Välj godtyckligt $y \in U$.
        $F(y) \leftarrow \{e \in E : e \text{ bara innehåller variabeln } y\}$

Ge y ett värde som satisfierar så många olikheter i $F(y)$ som möjligt.
$E \leftarrow E - F(y)$
else
Välj godtyckligt $y \in X$.
$y \leftarrow 0$
Evaluera om olikheterna i E som innehåller y.
$X \leftarrow X - \{y\}$

Algoritmen tilldelar alltid y ett värde som satisfierar minst hälften av olikheterna i $F(y)$. Alltså konstruerar den en lösning som satisfierar minst hälften av olikheterna i indata. Eftersom högst alla olikheterna kan satisfieras blir approximationskvoten för algoritmen 2. Tidskomplexiteten är polynomisk eftersom varje variabel och term bara behandlas en gång. □

Lösning till Övre gräns för approximation av homogena bipolära olikheter
Vi börjar med approximation av Max Hom Bipolar Sat LR$^{\geq}$. Ta en godtycklig bipolär vektor $\mathbf{x}$ och titta på antalet satisfierade olikheter om variablerna sätts till $\mathbf{x}$ respektive $-\mathbf{x}$. Om vänsterledet av en olikhet är positivt för $\mathbf{x}$ så är det negativt för $-\mathbf{x}$ och vice versa. Därför kommer en av dessa två vektorer att satisfiera minst häften av olikheterna, det vill säga uppnå approximationskvoten 2.

Denna triviala algoritm fungerar inte för Max Hom Bipolar Sat LR$^{>}$, för många relationer kan vara noll för båda vektorerna. Därför börjar vi med att söka upp en lösning med många nollskilda relationer. Approximationsalgoritmen för Max Sat LR$^{\geq}$ ovan kan modifieras så att den hittar en lösning $\mathbf{x}$ för vilken minst hälften av relationerna är nollskilda. Samma sak gäller då också $-\mathbf{x}$, och någon av dessa två vektorer måste då satisfiera minst en fjärdedel av alla relationer. □

Lösning till Undre gräns för approximation av binära olikheter
Visa detta genom att reducera Max Clique till Max Binary Sat LR$^{\geq}$ med en approximationsbevarande reduktion. Eftersom vi vet att Max Clique inte tillhör APX så kommer det att innebära att Max Binary Sat LR$^{\geq}\notin$APX.

Låt $G = (V, E)$ vara indata till Max Clique. För varje hörn $v_i \in V$ så konstruerar vi en variabel x_i och olikheten

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

där j är med i $N(v_i)$ om $v_j \neq v_i$ och $(v_i, v_j) \notin E$ (det vill säga v_j inte är granne med v_i). Nu har vi ett system med $|V|$ olikheter. Observera att den i-te olikheten är satisfierad om och endast om $x_i = 1$ och $x_j = 0$ för alla $j \in N(v_i)$.

Det är lätt att verifiera att om man får en s-klick $V' \subseteq V$ så kan man få en binär lösning som satisfierar dom s motsvarande olikheterna genom att sätta $x_i = 1$ om $v_i \in V'$ och $x_i = 0$ annars. Å andra sidan, om man får en binär lösning $\mathbf{x}$ som satisfierar s olikheter så kan man få en s-klick genom att låta V' bestå av alla hörn v_i som motsvarar satisfierade olikheter.

Denna reduktion är inte bara *approximationsbevarande* utan också *kostnadsbevarande* eftersom den bevarar målfunktionens värde helt och hållet.

Notera att denna reduktion också fungerar för $A\mathbf{x} > \mathbf{0}$ och $A\mathbf{x} = \mathbf{1}$, så vi kan alltså visa att Max Binary Sat LR$^{>}\notin$APX och att Max Binary Sat LR$^{=}\notin$APX. □
