<u>בעיית הסוכן הנוסע המטרי (Metric TSP) - הדיקות של ניתוח האלגוריתמים</u>

נתון:

- גרף $G = (V, E, w)$, כאשר $|V| = n = 2k+1$, $E = \{(v_i, v_{i+1}) \mid 1 \le i < n\} \cup \{(v_i, v_{i+2}) \mid 1 \le i < n - 1\}$,
ו- $w(v_i, v_{i+1}) = 1$, $w(v_i, v_{i+2}) = 1 + \varepsilon$, $0 < \varepsilon < 1$.

- גרף מלא $G' = (V', E')$, כאשר $V' = V$, $E' = \{(v_i, v_j) \mid v_i, v_j \in V', i \neq j\}$, והמטריקה d על G' מוגדרת על ידי מרחקי המסלולים הקצרים ביותר ב-G.

מטרה:

1. הצג והשווה את התוצאות של האלגוריתמים "פעמיים MST" ו-Christofides לעלות הפתרון האופטימלי.
2. תאר פונקציה $f(n)$ כך שאם $\varepsilon = f(n)$, הניתוח של אלגוריתם Christofides הופך להדוק ככל ש-n גדל.

פתרון:

1. פתרון אופטימלי ל-TSP:

למה 1: הפתרון האופטימלי ל-TSP ב-G' הוא מעגל המילטוני בצורה ($v_1, v_3, \ldots, v_n, v_{n-1}, v_{n-3}, \ldots, v_2, v_1$), עם עלות של $(n - 1) \cdot (1 + \varepsilon) + 1 - \varepsilon$.

הוכחה (באינדוקציה על n):

מקרה בסיס: עבור $n = 3$, הפתרון האופטימלי ל-TSP הוא (v_1, v_3, v_2, v_1) עם עלות של $3 + \varepsilon$, שווה ל-$(3 - 1) \cdot (1 + \varepsilon) + 1 - \varepsilon$.

צעד אינדוקטיבי: נניח שהלמה מתקיימת עבור $n' < n$. נוכיח אותה עבור n.

נניח שקיים פתרון אופטימלי ל-TSP עם עלות הקטנה מ-$(n - 1) \cdot (1 + \varepsilon) + 1 - \varepsilon$. נניח שפתרון זה מתחיל ב-v_1 ומכיל את הקשת (v_1, v_j), כאשר $j > 2$.

מקרה 1 ($j > 2$): נתבונן בתת-גרף G_j שנוצר על ידי $\{v_j, v_{j+1}, \ldots, v_n\}$. לפי ההנחה האינדוקטיבית, הפתרון האופטימלי ל-TSP ב-G_j הוא בעל עלות $(n - j) \cdot (1 + \varepsilon) + 1 - \varepsilon$. כדי לחבר את v_{j+1} בחזרה ל-v_1, אנו זקוקים למסלול בעל עלות הקטנה מ-$(j - 1) \cdot (1 + \varepsilon) - (j - 1)/2 \cdot (1 + \varepsilon)$. עם זאת, המסלול הקצר ביותר מ-v_{j+1} ל-v_1 בתת-הגרף שנוצר על ידי $\{v_1, v_2, \ldots, v_{j-1}\}$ הוא בעל עלות של לפחות $(j - 2) \cdot (1 + \varepsilon) + 1 - \varepsilon$, שהיא גדולה יותר מהעלות הנדרשת עבור $j > 3$, מה שמוביל לסתירה.

מקרה 2 ($j = 2$): נתבונן בתת-גרף G_2 שנוצר על ידי $\{v_2, v_3, \ldots, v_n\}$. לפי ההנחה האינדוקטיבית, הפתרון האופטימלי ל-TSP ב-G_2 הוא בעל עלות $(n - 2) \cdot (1 + \varepsilon) + 1 - \varepsilon$. חיבור v_3 ל-v_1 במקום ל-v_2 מוביל לעלות של $(n - 2) \cdot (1 + \varepsilon) + 2$, שווה לעלות האופטימלית, מה שסותר את ההנחה של פתרון בעל עלות נמוכה יותר.

לכן, הפתרון האופטימלי ל-TSP ב-G' הוא $(v_1, v_3, \ldots, v_n, v_{n-1}, v_{n-3}, \ldots, v_2, v_1)$ עם עלות של $(n - 1) \cdot (1 + \varepsilon) - \varepsilon + 1$.

2. אלגוריתם Twice MST:

אלגוריתם Twice MST מוצא עץ פורש מינימלי (MST) ב-G', מכפיל את קשתותיו, ומוצא מסלול אוילרי בגרף המתקבל. המסלול מקוצר כדי לקבל מעגל המילטוני.

עלות MST ב-G': $n - 1$

עלות הפתרון של Twice MST: $2(n - 1) + (n - 1)/2 \cdot (1 + \varepsilon)$

יחס של Twice MST לפתרון האופטימלי:

$(2(n - 1) + (n - 1)/2 \cdot (1 + \varepsilon)) / ((n - 1) \cdot (1 + \varepsilon) - \varepsilon + 1)$

$\geq (3n - 3 + (n - 1)\varepsilon/2) / (n - 1 + (n - 1)\varepsilon)$

$= (3 + \varepsilon/2) / (1 + \varepsilon)$

$\geq 3/2$ (עבור $0 < \varepsilon < 1$)

לכן, אלגוריתם Twice MST מספק קירוב של 1.5 לפתרון האופטימלי של TSP ב-G'.

3. אלגוריתם Christofides:

אלגוריתם Christofides מוצא MST ב-G', מחשב התאמה מושלמת במשקל מינימלי (minimum-weight perfect matching) על הקודקודים בעלי דרגה אי-זוגית ב-MST, ובונה מסלול אוילרי בגרף הרב-קודקודי המתקבל. המסלול מקוצר כדי לקבל מעגל המילטוני.

במקרה זה, הקודקודים היחידים בעלי דרגה אי-זוגית ב-MST הם v_1 ו-v_n. ההתאמה המושלמת במשקל מינימלי מוסיפה את הקשת (v_1, v_n) עם משקל של $(n - 1)/2 \cdot (1 + \varepsilon)$.

הניתוח של אלגוריתם Christofides זהה לאלגוריתם Twice MST, ומוביל לקירוב של 1.5 לפתרון האופטימלי של TSP ב-G'.

4. הדיקות של אלגוריתם Christofides:

כדי למצוא פונקציה $f(n)$ כך שאם $\varepsilon = f(n)$, הניתוח של אלגוריתם Christofides הופך להדוק ככל ש-n גדל, נשווה את היחס בין עלות Christofides לעלות האופטימלית ל-1.5 ונפתור עבור ε:

$$(3n - 3 + (n - 1)\varepsilon/2) / (2n - 2 + (n - 1)\varepsilon) = 1.5$$

$$(n - 1)\varepsilon/2 = 0.5(n - 1)\varepsilon$$

$$\varepsilon = 1/(n - 1)$$

לכן, אם נבחר $f(n) = 1/(n - 1)$, הניתוח של אלגוריתם Christofides הופך להדוק ככל ש-n גדל.

לחלופין, ניתן לבחור $f(n) = 1/(n + 1)$, שגם מקיים את התנאי, מכיוון ששתי הפונקציות שואפות ל-0 ככל ש-n שואף לאינסוף, מה שגורם ליחס לשאוף ל-1.5.

גבול היחס כאשר n שואף לאינסוף (באמצעות $f(n) = 1/(n - 1)$ או $f(n) = 1/(n + 1)$):

$\lim (n\to\infty) ((3n - 3 + (n - 1)\varepsilon/2) / (2n - 2 + (n - 1)\varepsilon))$, כאשר $\varepsilon = f(n)$

$\lim (n\to\infty) ((6n - 5) / (4n - 2)) =$

$1.5 =$

לסיכום, אלגוריתמי Twice MST ו-Christofides מספקים קירוב של 1.5 לפתרון האופטימלי של TSP בגרף הנתון 'G. על ידי בחירת $\varepsilon = f(n) = 1/(n - 1)$ או $\varepsilon = f(n) = 1/(n + 1)$, הניתוח של אלגוריתם Christofides הופך להדוק ככל ש-n גדל, כאשר היחס שואף ל-1.5 ככל ש-n שואף לאינסוף.

2.

בעיית כיסוי הקבוצה (Set Cover Problem):

נתון:

יהי $U = \{1, 2, ..., n\}$ היקום של האיברים, כאשר n הוא חזקה של 2 (כלומר, $n = 2^k$ עבור איזשהו מספר שלם k).

נבנה את משפחת הקבוצות S באופן הבא:

1. נכלול את הקבוצות $\{1, 2, ..., n/2\}$ ו- $\{n/2 + 1, ..., n\}$ ב-S. שתי קבוצות אלה מהוות פתרון אופטימלי (OPT) בגודל 2.

2. עבור כל איבר i ב-U, ניצור קבוצה $S_i = \{i\}$.

3. עבור $j = 1$ עד $\log(n) - 1$:

 - נחלק את U ל- $n / 2^j$ תתי-קבוצות, כל אחת בגודל 2^j.

 - עבור כל תת-קבוצה בחלוקה זו, נוסיף את הקבוצה המתאימה ל-S.

דוגמה עבור $n = 8$:

$U = \{1, 2, 3, 4, 5, 6, 7, 8\}$

$OPT = \{\{1, 2, 3, 4\}, \{5, 6, 7, 8\}\}$ (גודל = 2)

S = {

 {1}, {2}, {3}, {4}, {5}, {6}, {7}, {8},

 {1, 2}, {3, 4}, {5, 6}, {7, 8},

 {1, 2, 3, 4}, {5, 6, 7, 8}

}

האלגוריתם החמדן תמיד יבחר תחילה בקבוצות היחידה {i}, מכיוון שיש להן את מספר האיברים הלא מכוסים המקסימלי (1 כל אחת). לאחר בחירת כל קבוצות היחידה, האלגוריתם יבחר בקבוצות בגודל 2, לאחר מכן 4, וכן הלאה, עד שכל האיברים יכוסו.

סך הקבוצות שנבחרו על ידי האלגוריתם החמדן הוא:

n (קבוצות יחידה) + n/2 (קבוצות בגודל 2) + n/4 (קבוצות בגודל 4) + ... + 1 (קבוצה בגודל n/2)

$= n + n/2 + n/4 + ... + 1$

$2n - 1 =$

$\Theta(n) =$

מכיוון ש- $|OPT| = 2$ ו- $n = 2^k$, יש לנו:

גודל הפתרון החמדן = $\Theta(n) = \Theta(2^k) = \Theta(2^{(\log n)}) = \Theta(|OPT| \cdot \log n)$

לכן, משפחת דוגמאות זו מדגימה כי הניתוח של האלגוריתם החמדן הוא הדוק, מכיוון שגודל הפתרון החמדן הוא $\Omega(|OPT| \cdot \log n)$.

<u>Approximations Algorithms & FPTAS</u>

<u>Bin Packing</u>

1. תיאור האלגוריתם:

- עבור כל פריט, נסה לארוז אותו בפח הפתוח הנוכחי.
- אם הוא לא מתאים, סגור את הפח הזה ופתח חדש.
- הנח את הפריט בפח החדש והמשך לפריט הבא.

2. הוכחת יחס הקירוב:

אנחנו רוצים להראות ש- $|ALG| \le 2 \cdot |OPT| - 1$, כאשר $|ALG|$ הוא מספר הפחים שמשתמש בהם האלגוריתם ו-$|OPT|$ הוא מספר הפחים בפתרון האופטימלי.

הוכחה:

א) נתבונן בכל שני פחים סמוכים בפתרון ALG.

ב) סכום הפריטים בשני פחים אלה חייב להיות גדול מ-1 (קיבולת של פח יחיד). אחרת, כל הפריטים מהפח השני היו מתאימים בפח הראשון.

ג) משמעות הדבר היא שעבור כל זוג פחים ב-ALG (למעט אולי הפח האחרון), הגודל הכולל של הפריטים גדול מ-1.

ד) לכן, הגודל הכולל של כל הפריטים גדול מ-$(|ALG| - 1) / 2$.

ה) מכיוון ש-OPT לא יכול לארוז פריטים בגודל כולל הגדול מ-$|OPT|$, יש לנו:

$$|OPT| > (|ALG| - 1) / 2$$

ו) פתרון עבור $|ALG|$:

$$|ALG| < 2 \cdot |OPT| + 1$$

ז) מכיוון ש-$|ALG|$ הוא מספר שלם, אנחנו יכולים לכתוב:

$|ALG| \le 2 \cdot |OPT|$

זה מוכיח את יחס הקירוב של 2.

3. הדוקות הניתוח:

כדי להראות שהניתוח הדוק, עלינו לספק דוגמה שבה היחס $|ALG| / |OPT|$ מתקרב ל-2.

דוגמה:

נתבונן במקרה הבא:

- $2n$ פריטים בגודל $(1/2 + \varepsilon)$
- $2n$ פריטים בגודל $(1/2 - \varepsilon)$

כאשר n הוא מספר שלם גדול ו-ε הוא מספר חיובי קטן הקרוב ל-0.

א) פתרון אופטימלי (OPT):

OPT יכול לארוז את הפריטים הללו ב-$n + 1$ פחים:

- n פחים עם פריט אחד מכל גודל $(1/2 + \varepsilon) + (1/2 - \varepsilon) = 1$
- פח 1 עם שני הפריטים הנותרים בגודל $(1/2 + \varepsilon)$

אז, $|OPT| = n + 1$

ב) אלגוריתם ההתאמה הבאה (ALG):

ALG יארוז את הפריטים כך:

- n פחים עם שני פריטים בגודל $(1/2 + \varepsilon)$
- n פחים עם שני פריטים בגודל $(1/2 - \varepsilon)$

אז, $|ALG| = 2n$

ג) יחס:

$|ALG| / |OPT| = 2n / (n + 1)$

כאשר n שואף לאינסוף, היחס הזה מתקרב ל-2.

דוגמה זו מראה שהניתוח הדוק, כיוון שקיימים מקרים שבהם הביצועים של אלגוריתם ההתאמה הבאה יכולים להיות קרובים באופן שרירותי לפי שניים מהפתרון האופטימלי.

לסיכום, הוכחנו ש-$|ALG| \leq 2 \cdot |OPT| - 1$ והדגמנו שהניתוח הזה הדוק על ידי מתן דוגמה שבה היחס מתקרב ל-2.

One more Bin Packing for Dessert

משפט: בהינתן קבוצה U של פריטים עם משקלים $w_u \geq \delta$ לכל $u \in U$ ופרמטר $k > 0$, קיימת סכימת עיגול היוצרת קבוצה חדשה U' של פריטים כך ש:

1. $|OPT(U)| \leq |OPT(U')|$
2. $|OPT(U')| \leq |OPT(U)| + \lfloor |OPT(U)|/k \rfloor + 1$

כאשר $OPT(U)$ ו-$OPT(U')$ מסמנים את הפתרונות האופטימליים לאריזת המכולות עבור המופעים המקורי והמעוגל, בהתאמה.

הוכחה:

חלק 1: $|OPT(U)| \le |OPT(U')|$

נתבונן בסכימת העיגול שבה לכל פריט $u \in U$ עם משקל $w_u \in ((i-1)\delta/k, i\delta/k]$, אנו יוצרים פריט חדש $u' \in U'$ עם משקל $w_{u'} = i\delta/k$.

תהי S פתרון אופטימלי עבור המופע המקורי U. נוכל לבנות פתרון תקף S' עבור המופע המעוגל U' על ידי החלפת כל פריט $u \in S$ בפריט המעוגל המתאים לו $u' \in U'$. מכיוון ש-$w_u \le w_{u'}$ לכל $u \in U$, סך המשקל של הפריטים בכל מכולה של S' הוא לפחות סך המשקל של הפריטים במכולה המתאימה של S. לכן, S' הוא פתרון תקף עבור U', ולכן $|OPT(U)| \le |OPT(U')|$.

חלק 2: $|OPT(U')| \le |OPT(U)| + \lfloor |OPT(U)|/k \rfloor + 1$

תהי S פתרון אופטימלי עבור המופע המקורי U, ותהי S' פתרון אופטימלי עבור המופע המעוגל U'. נראה שנוכל להפוך את S' לפתרון תקף עבור U על ידי פתיחה של לכל היותר $\lfloor |OPT(U)|/k \rfloor + 1$ מכולות נוספות.

נתבונן בכל מכולה B ב-S'. סך המשקל של הפריטים ב-B הוא לכל היותר $1 + \delta/k$ מכיוון שתהליך העיגול מגדיל את המשקל של כל פריט לכל היותר ב-δ/k. אם נחליף כל פריט $u' \in B$ בפריט המקורי שלו u, סך המשקל של הפריטים ב-B יקטן לכל היותר ב-δ/k. לכן, סך המשקל של הפריטים המקוריים ב-B הוא לכל היותר 1.

כעת, נחלק את הפריטים המקוריים מ-S' לשני סוגים של מכולות:

1. סוג 1: מכולות המכילות פריטים מלכל היותר k מכולות של S.
2. סוג 2: מכולות המכילות פריטים מיותר מ-k מכולות של S.

עבור כל מכולה מסוג 1, נוכל לסדר מחדש את הפריטים המקוריים לכל היותר k מכולות של S, מכיוון שסך המשקל של הפריטים בכל מכולה הוא לכל היותר 1.

עבור כל מכולה מסוג 2, נפתח מכולה חדשה ונציב את כל הפריטים המקוריים ממכולה זו במכולה החדשה. מכיוון שיש לכל היותר $\lfloor |OPT(U)|/k \rfloor$ מכולות מסוג 2, נפתח לכל היותר $\lfloor |OPT(U)|/k \rfloor$ מכולות חדשות.

לבסוף, ייתכן שנצטרך לפתוח מכולה נוספת אחת כדי להכיל פריטים נותרים כלשהם.

לכן, נוכל להפוך את S' לפתרון תקף עבור U על ידי פתיחה של לכל היותר $\lfloor |OPT(U)|/k \rfloor + 1$ מכולות נוספות.
מכאן נובע ש-$|OPT(U')| \leq |OPT(U)| + \lfloor |OPT(U)|/k \rfloor + 1$, והוכחנו את המשפט.

בפתרון זה, הוכחנו ישירות את החסם העליון על $|OPT(U')|$ על ידי הפיכת הפתרון האופטימלי עבור המופע המעוגל לפתרון תקף עבור המופע המקורי, תוך מעקב אחר המכולות הנוספות הנדרשות בתהליך.