הרצאה 5

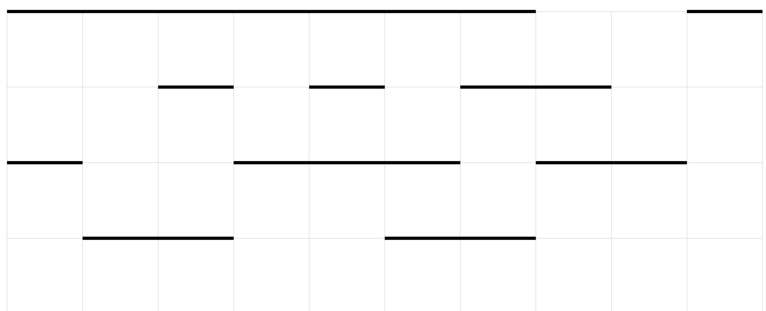
אלגוריתמים חמדניים

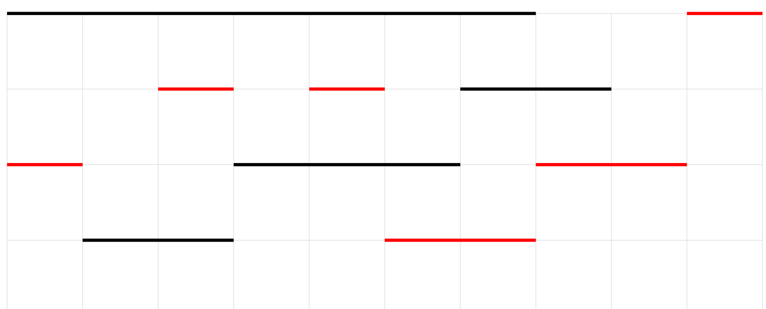
הקדמה

לעיתים קרובות אפשר לייצג בעיות אופטימיזציה כקבוצה של אלמנטים כאשר פתרון חוקי הוא תת קבוצה של אלמנטים שמקיימת תכונות מסויימות. למשל, עץ פורש מינימלי. בדרך כלל יש פונקציית מחיר / רווח לכל תת קבוצה והמטרה שלנו היא למזער / למקסם את הערך הזה.
אלגוריתם חמדן, באופן לא פורמלי, הוא כזה שבונה פתרון (תת קבוצה של אלמנטים) באופן איטרטיבי ובכל שלב מוסיף / מסיר מהקבוצה

קבוצת אינטרוולים בלתי תלויה בגודל מקסימלי

נתונים n אינטרוולים $A = \{a_1, \ldots, a_n\}$, נסמן ב-$s(a_i)$ את זמן ההתחלה של האינטרוול a_i וב-$e(a_i)$ את זמן הסיום שלו. לכל אינטרוול מתקיים ש-$s(a_i), e(a_i) \in \mathbb{R}_+$ וכן $s(a_i) < e(a_i)$ רוצים למצוא תת קבוצה בגודל מקסימלי $I \subseteq A$ כך שהאינטרוולים ב-I זרים בזוגות, כלומר לכל $a, b \in I$, אחד התנאים מתקיים: $e(a) < s(b)$ או ש-$e(b) < s(a)$.
דוגמה:





אלגוריתם חמדן:

1. אתחול: $I \leftarrow \emptyset$, $\bar{e} \leftarrow 0$
2. עבור כל אינטרוול a בסדר לא יורד של ערכי $e(a)$:
   (א) אם $s(a) \geq \bar{e}$
      i. $I \leftarrow I \cup \{a\}$
      ii. $\bar{e} \leftarrow e(a)$

לפני שנוכיח נכונות נראה דוגמאות לגישות חמדניות שלא עובדות:
לבחור את האינטרוול עם זמן התחלה הכי מוקדם

לבחור את האינטרוול הכי קצר

לבחור את האינטרוול שנחתך עם הכי מעט אינטרוולים

הוכחת נכונות: נוכיח את הטענה הבאה, בכל צעד של האלגוריתם קיימת קבוצה בגודל מקסימלי, I' כך ש-I רישא שלה ביחס למיון ע"פ ערכי e.
בסיס: באתחול טריוויאלי
צעד: נבחן את הקבוצות I ו-I' בצעד ה-$i+1$. לפי הנחת האינדוקציה הקבוצות, ממוינות על פי ערכי e נראות כך:

$$\begin{aligned} I &= \{\alpha_1,\ldots,\alpha_i,\alpha_{i+1}\} \\ I' &= \{\alpha_1,\ldots,\alpha_i,\beta_1,\ldots,\beta_k\} \end{aligned}$$

נסתכל על הפתרון

$$I'' = \{\alpha_1,\ldots,\alpha_i,\boldsymbol{\alpha_{i+1}},\ldots,\beta_k\}$$

מכיוון ש-I' פתרון חוקי האינטרוולים שם זרים בזוגות ולכן גם האינטרוולים ב-I'' למעט אולי α_{i+1}. מכיוון שהאלגוריתם בונה פתרון חוקי אז אנחנו יודעים ש-α_{i+1} זר ל-$\alpha_1,\ldots,\alpha_i$ ובגלל האופי החמדני של האלגוריתם מתקיים ש-$e(\alpha_{i+1}) \leq e(\beta_1) \leq s(beta_2) \leq,\ldots,\leq s(\beta_k)$ ולכן I'' פתרון בגודל מקסימלי כך ש-I רישא שלו.

שיבוץ משימות

נתונות n משימות $A=\{a_1,\ldots,a_n\}$ נסמן ב-$t(a_i)$ את הזמן הנדרש לביצוע משימה a_i וב-$d(a_i)$ את זמן הסיום הרצוי של המשימה. בהינתן סדר ביצוע המשימות (פרמוטציה) $\pi : A \to [n]$. נסמן ב-$\delta(a_i)$ את זמן הסיום של המשימה a_i, כלומר

$$\delta(a_i) = \sum_{i \leq \pi(a_i)} t(\pi^{-1}(i))$$

נסמן ב-$l(a_i) := \delta(a_i) - d(a_i)$ את האיחור בביצוע משימה a_i. רוצים למצוא סדר שממזער את האיחור המקסימלי, כלומר

$$\arg\min_{\pi}\{\max_{i} l(a_i)\}$$

האלגוריתם החמדן יבצע את המשימות בסדר לא יורד של זמני הסיום הרצויים.
הוכחת נכונות
נוכיח באינדוקציה את הטענה הבאה:
לכל i קיים פתרון אופטימלי שמבצע את i המשימות הראשונות לפי זמני הסיום שלהן.
בסיס: טריוויאלי
צעד: נסתכל על המשימה, a, שזמן הסיום שלה הוא ה-$i+1$ לפי סדר לא יורד. אם הפתרון האופטימלי מבצע את המשימה הזאת בזמן $i+1$ סיימנו, אחרת הוא מבצע אותה בזמן $j > i+1$ נסתכל על סדר ביצוע המשימות מזמן $i+1$ עד זמן j:

$$b_{i+1},\ldots,b_{j-1},a$$

נבחן פתרון שמבצע את המשימות הללו בסדר הבא:

$$a,b_{i+1},\ldots,b_{j-1}$$

נבדוק את האיחור המקסימלי של משימות אלו (האיחור המקסימלי של יתר המשימות לא השתנה) ונניח בשלילה שהוא גדל (אחרת סיימנו). אם זמן הסיום גדל זה חייב להיות בגלל אחת מהמשימות $b_{i+1},\ldots,b_{j-1}$, נסמן אותה ב-b. נסמן את זמן הסיום שלה לפי הסדר החדש ב-$\delta'(b)$ אנחנו יודעים אבל ש-$\delta'(b) \leq \delta(a)$ וגם ש-$d(a) \leq d(b)$ ולכן $\delta'(b) - d(b) \leq \delta(a) - d(a)$.