שימושי זרימה

פרק 13-7.5 ב-Kleinberg/Tardos

שידוך בגרף דו-צדדי
עיבוד תמונות

בעיית השידוך

באתר שידוכים רשומים m נשים ו-n גברים. תוכנת האתר מאתרת זוגות מתאימים. בהינתן האוסף של ההתאמות האפשריות, יש לשדך מספר מירבי של זוגות.

הצגת הנתונים: גרף דו-צדדי $G = (X,Y,E)$. הקבוצה X מייצגת את קבוצת הנשים, Y את קבוצת הגברים, והקשתות E את אוסף הזוגות המתאימים.

שידוך בגרף דו-צדדי

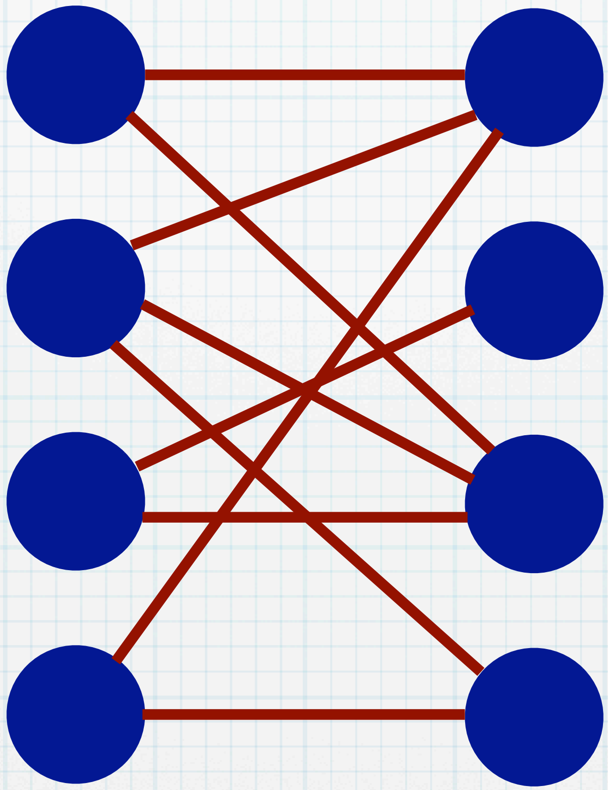
הגדרה: שידוך הוא קבוצה של קשתות M עבורה לכל צומת x יש לכל היותר קשת אחת ב-M שנוגעת בו, כלומר $|\{e \in M: x \in e\}| \le 1$.

הגדרה: שידוך M נקרא שידוך מושלם אם לכל צומת $x \in X \cup Y$ מתקיים $|\{e \in M: x \in e\}| = 1$.

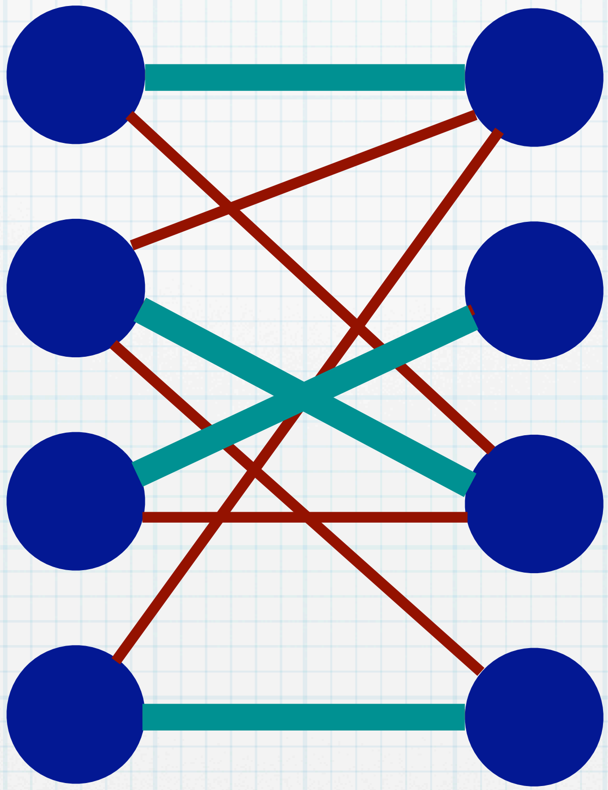
הקלט: גרף דו-צדדי $G = (X,Y,E)$.

הפלט: שידוך M עבורו $|M|$ מירבי (שידוך מקסימום).

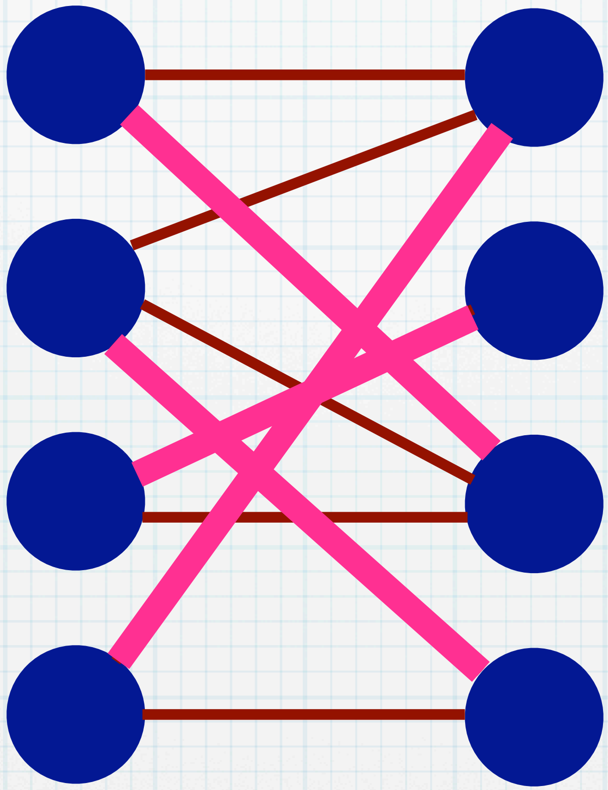
הדגמה



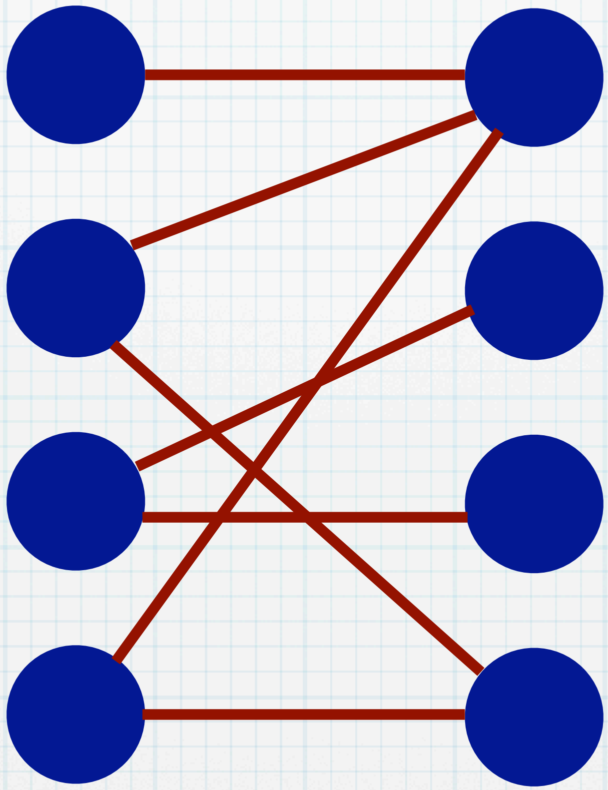
הדגמה



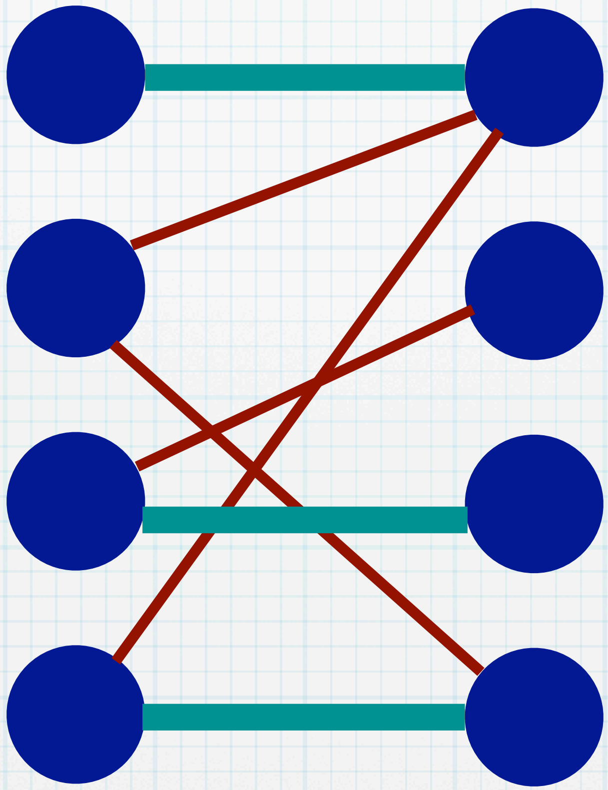
הדגמה



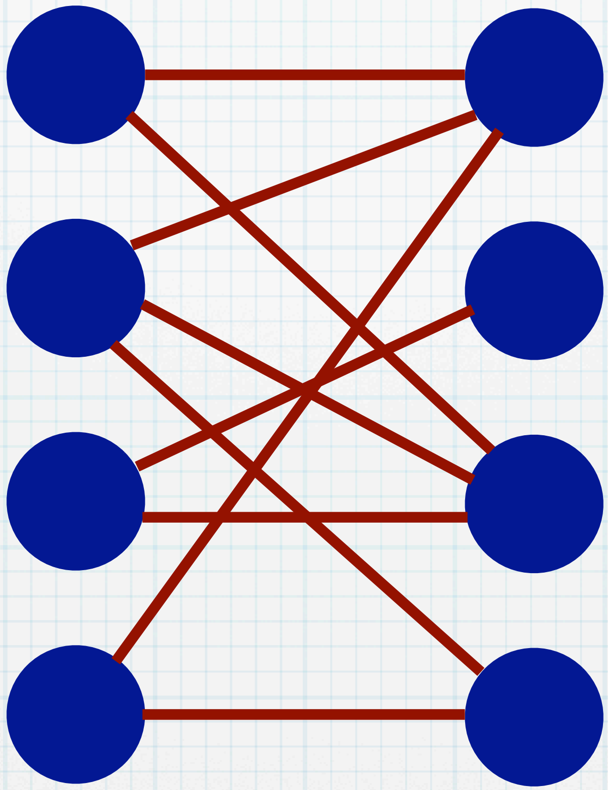
הדגמה



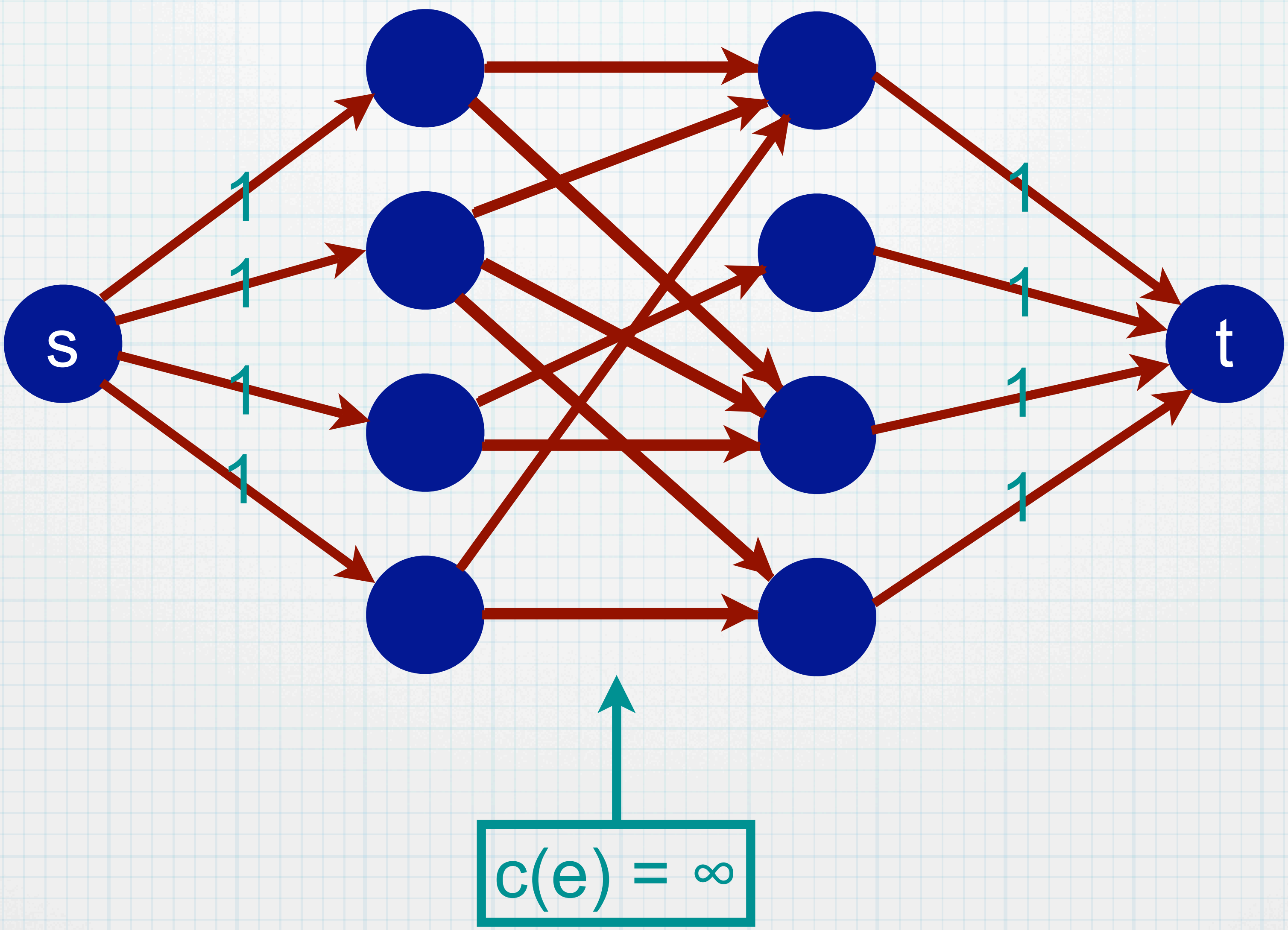
הדגמה



רדוקציה לבעיית זרימה



רדוקציה לבעיית זרימה



הרדוקציה

נבנה רשת זרימה H,c,s,t:

$H = (X \cup Y \cup \{s,t\}, F)$
$F = \{(s,x): x \in X\} \cup \{(x,y) \in X \times Y: \{x,y\} \in E\} \cup \{(y,t): y \in Y\}$
$c(e) \in \{1, \infty\}$
$c(e) = 1 \text{ iff } e = (s,x) \vee e = (y,t)$

נמצא זרימת מקסימום בשלמים f^*.

פתרון בעיית השידוך:
$M^* = \{\{x,y\}: (x,y) \in F \wedge f^*(x,y) > 0\}$

משידוך לזרימה

יהי M שידוך ב-G. נגדיר $f_M: F \to \{0,1\}$ על ידי

$$f_M(e) = 1 \Leftrightarrow (e = (x,y) \wedge \{x,y\} \in M) \vee$$
$$(e = (s,x) \wedge \exists y, \{x,y\} \in M) \vee$$
$$(e = (y,t) \wedge \exists x, \{x,y\} \in M)$$

אזי f_M היא זרימה ב-H,c,s,t.

הוכחה: אילוצי הקיבול נשמרים כי $c(e) \geq 1$ לכל קשת $e \in F$. אם $x \in X$ אזי

$$f_M(s,x) = 1 \Leftrightarrow \exists y, \{x,y\} \in M \Leftrightarrow f_M(x,y) = 1$$

ולכן מתקיים ב-x כלל שימור הזרימה. באופן דומה הכלל מתקיים בכל $y \in Y$.

הוכחת נכונות

משפט: אם f זרימה בשלמים ברשת H,c,s,t אזי הקבוצה $M = \{\{x,y\}: (x,y) \in F \wedge f(x,y) > 0\}$ היא שידוך ב-G עבורו $|f| = |M|$. אם M שידוך ב-G אזי f_M היא (כפי שהראנו) זרימה ב-H,c,s,t המקיימת $|f_M| = |M|$.

הוכחה: החלק השני של המשפט קל יותר:

$$|f_M| = \sum_{x\in X} f_M(s,x) = |\{x \in X: f_M(s,x) = 1\}| =$$
$$|\{x \in X: \exists y, \{x,y\} \in M\}| = |M|$$

לגבי החלק הראשון של המשפט, נשים לב שעבור כל $X \in x$ ועבור כל $Y \in y$ מתקיים

$$\sum_{e\in in(x)} f(e) \in \{0,1\} \wedge \sum_{e\in out(y)} f(e) \in \{0,1\}$$

לכן, בהסתמך על כך ש-f זרימה בשלמים מתקיים כי לכל $(x,y) \in X \times Y \cap F$ ערך הזרימה $f(x,y) \in \{0,1\}$, וכמו כן, לכל $x \in X$ (ובאופן דומה ל-$y \in Y$) מתקיים כי $|\{y \in Y: f(x,y) = 1\}| \leq 1$. לכן M שידוך, ו-$|f| = |M|$ $\therefore$

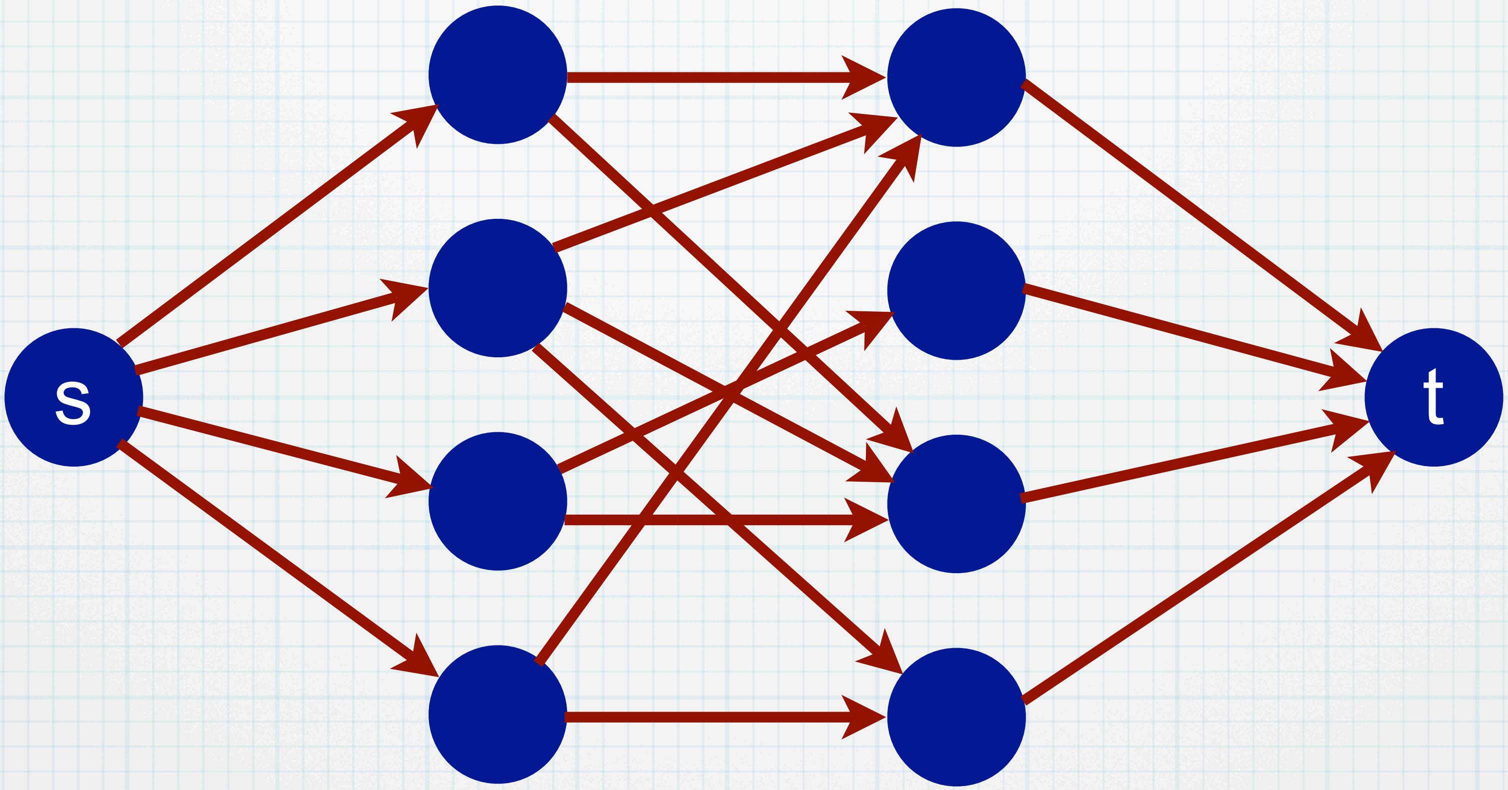
מסלולים מתחלפים

נתבונן בזרימה בשלמים f ב-Hc,s,t המתאימה לשידוך M. מסלול משפר ברשת השיורית
מתחיל בקשת (s,x) לצומת לא משודך $x \in X$
ומסתיים בקשת (y,t) מצומת לא משודך $y \in Y$.
הקטע בין הקשתות הללו הוא סדרה האורך אי-זוגי של קשתות, שבה הקשתות במקומות הזוגיים נמצאות ב-M והקשתות במקומות האי-זוגיים אינן ב-M (הכוונה לגרסה הלא מכוונת של כל קשת כזו).

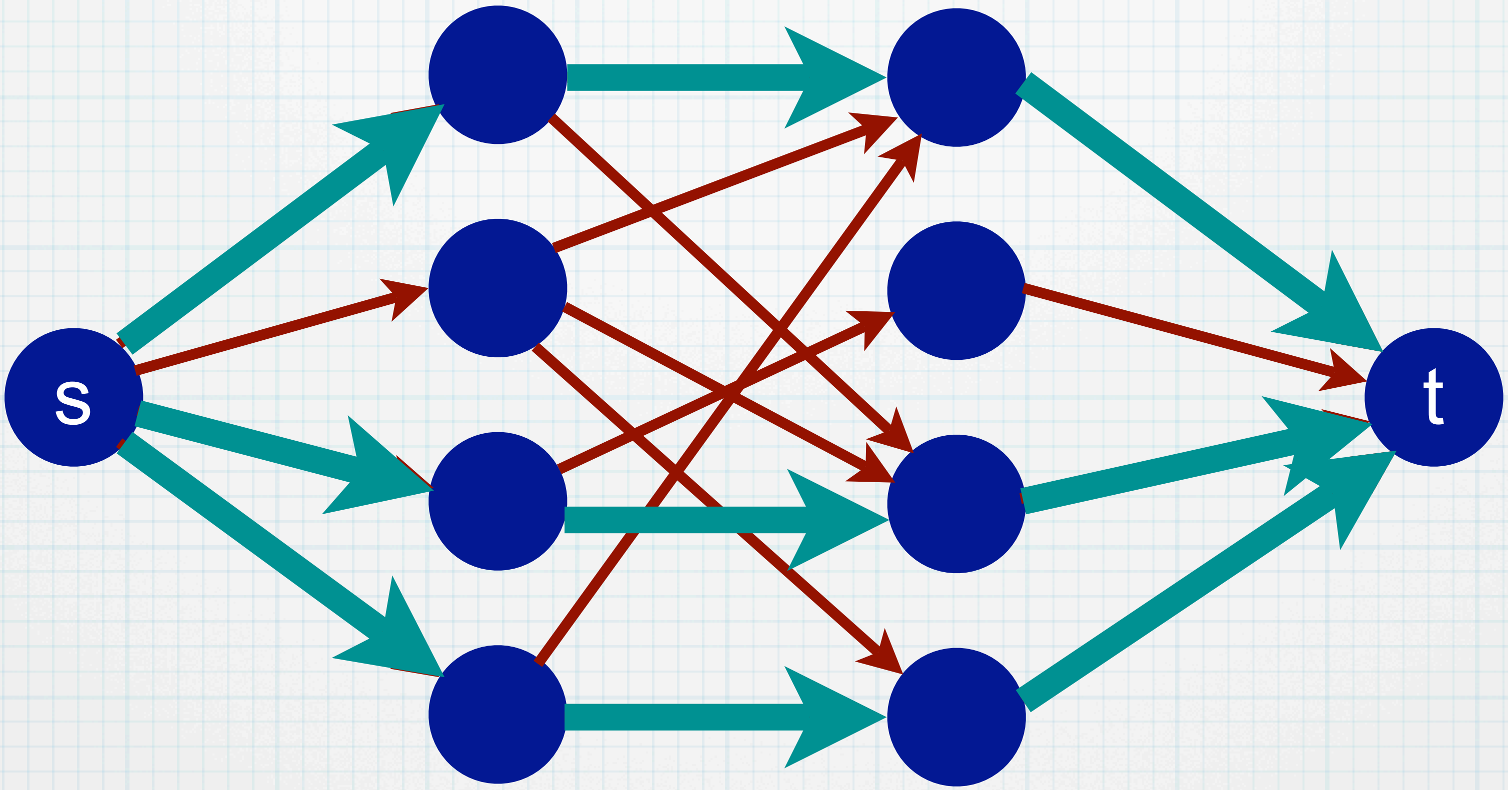
כל מסלול מתחלף כזה ב-G שמתחיל ומסתיים בצומת לא משודך ניתן להתאים למסלול משפר.

מסקנה: M שידוך מקסימום אםם אין ב-G מסלול מתחלף שמתחיל ומסתיים בצומת לא משודך.

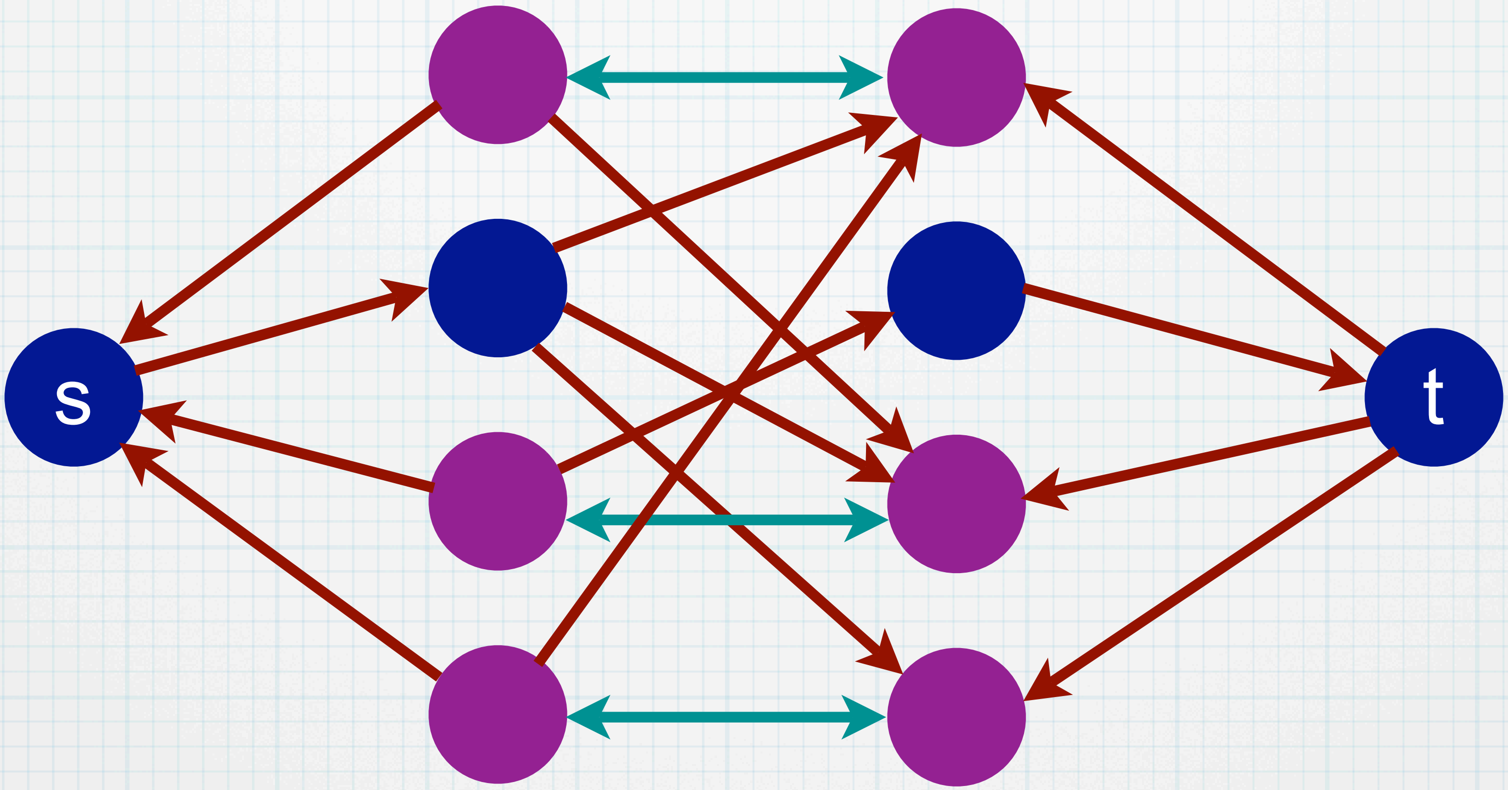
הדגמה



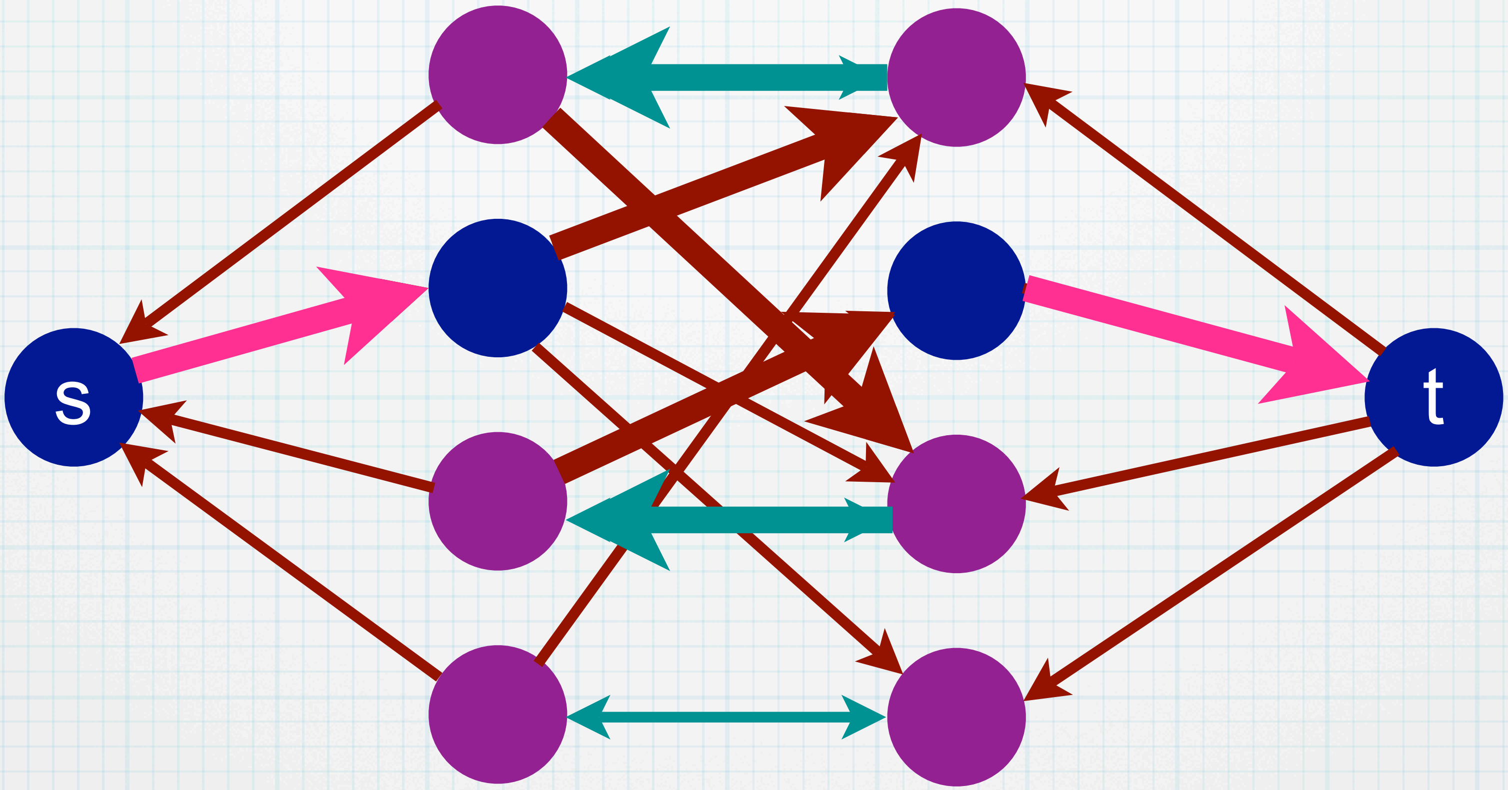
הדגמה



(המשך)



(המשך)



שיפור השידוך

נתבונן בגרף הנפרש על ידי הקשתות
$\{(x,y) \in X \times Y: \{x,y\} \in M^* \setminus M\} \cup$
$\{(y,x) \in Y \times X: \{x,y\} \in M \setminus M^*\}$
זהו איחוד של מסלולים ומעגלים מתחלפים. כל המעגלים הם באורך זוגי, ולכן אם $|M| < |M^*|$ יש בגרף זה מסלול מתחלף באורך אי-זוגי שמתחיל ומסתיים בצומת לא משודך. מספר המסלולים הללו הוא בדיוק $|M^*| - |M|$.

כל המסלולים והמעגלים הללו זרים בצמתים. לכן האורך הממוצע של מסלול כזה הוא לכל היותר:
$$(|X| + |Y|) / (|M^*| - |M|) - 1$$
בוודאי יש מסלול שאורכו לכל היותר הממוצע.
לכן, ב-$H_f = H_{f_M}$ יש מסלול משפר באורך
$$(|X| + |Y|) / (|M^*| - |M|) + 1$$

סיבוכיות

נסמן V = X ∪ Y.

נמצא זרימת מקסימום f* ב-H,c,s,t בעזרת אלגוריתם דיניץ. אם הזרימה הנוכחית f מקיימת $|f| \geq |f^*| - \sqrt{|V|}$ אזי נותרו עוד לכל היותר $\sqrt{|V|}$ איטרציות.

עד לקבלת f כזו תמיד יש מסלול משפר באורך של לכל היותר $\sqrt{|V|}+1$. לפי הניתוח מהשעור שעבר, מספר האיטרציות עד שהמרחק של t מ-s גדול מ- $\sqrt{|V|}$ הוא לכל היותר $\sqrt{|V|}$.

במקרה זה ניתן לממש כל איטרציה בזמן O(|E|).
(בכל מסלול משפר יותר ממחצית הקשתות רוויות.)

סה"כ $O\left(\sqrt{|V|} \cdot |E|\right)$

משפט Hall

נניח ש-$|Y| \leq |X|$. עבור קבוצה $X' \subseteq X$ נסמן ב-$N(X')$ את קבוצת שכני X ב-G. כלומר:

$$N(X') = \{y \in Y : \exists x \in X', \{x,y\} \in E\}$$

משפט: יש ב-G שידוך בגודל $|X|$ אם לכל $X' \subseteq X$ מתקיים $|N(X')| \leq |X'|$.

הוכחה: יהי $(S, V \setminus S)$ חתך מינימום ברשת H,c,s,t. אזי:

$$|M^*| = |f^*| = c(S, X \cup Y \cup \{s,t\} \setminus S) \leq c(\{s\}, X \cup Y \cup \{t\}) = |X|$$

שימו לב כי קיבול הקשתות מ-X ל-Y הוא ∞ ולכן חייב להיות כי $N(S \cap X) \subset S$. כלומר,

$$c(S, X \cup Y \cup \{s,t\} \setminus S) = |X \setminus S| + |Y \cap S| \geq |X \setminus S| + |N(S \cap X)|$$

אם $|S \cap X| \geq |N(S \cap X)|$ נקבל $|M^*| \geq |X|$, ולכן $|M^*| = |X|$.

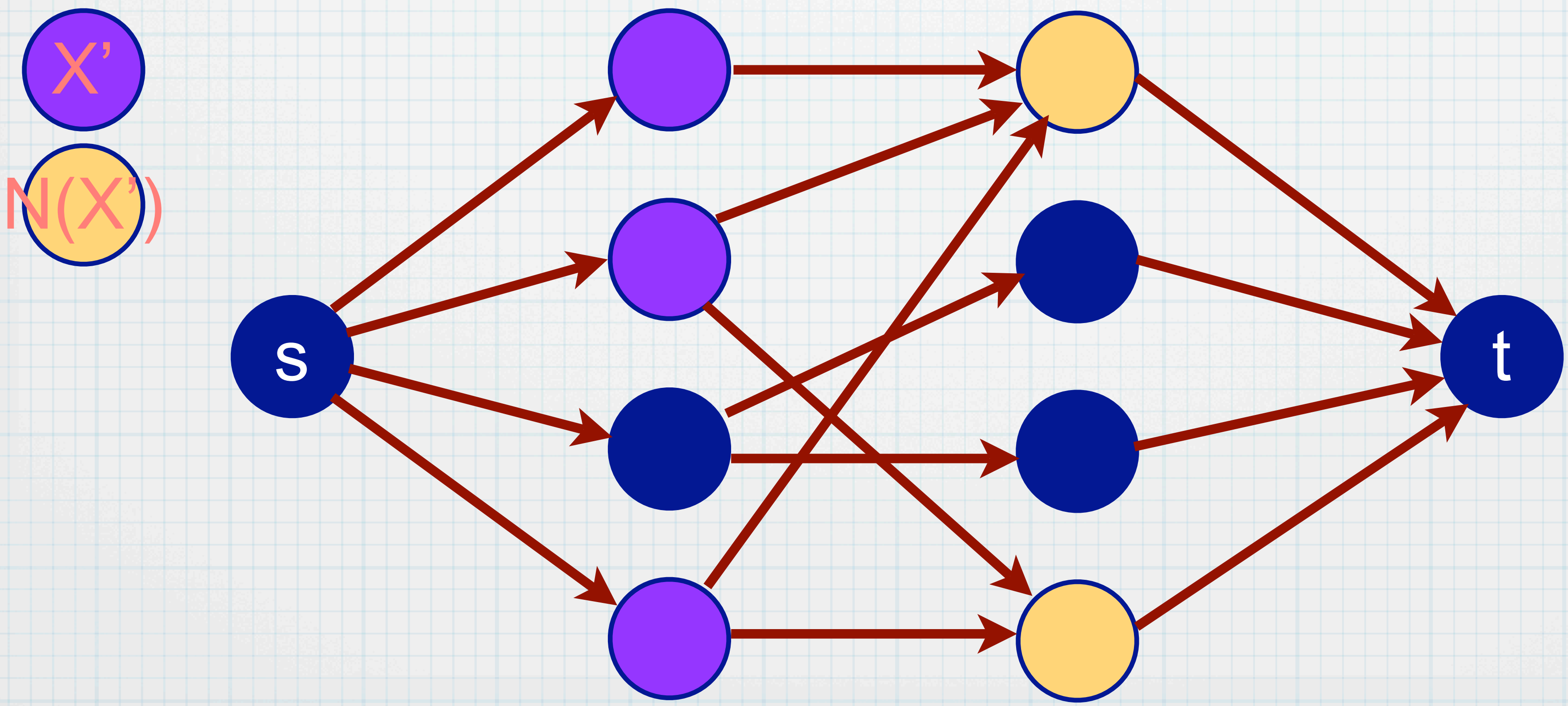
המשך הוכחת משפט Hall

מצד שני, אם יש $X' \subseteq X$ עבורה $|N(X')| < |X'|$ אזי:

$$c(\{s\} \cup X' \cup N(X') \,,\, (X \setminus X') \cup (Y \setminus N(X')) \cup \{t\}) =$$

$$|X \setminus X'| + |N(X')| < |X|$$

ולכן $|M^*| < |X|$. $\therefore$



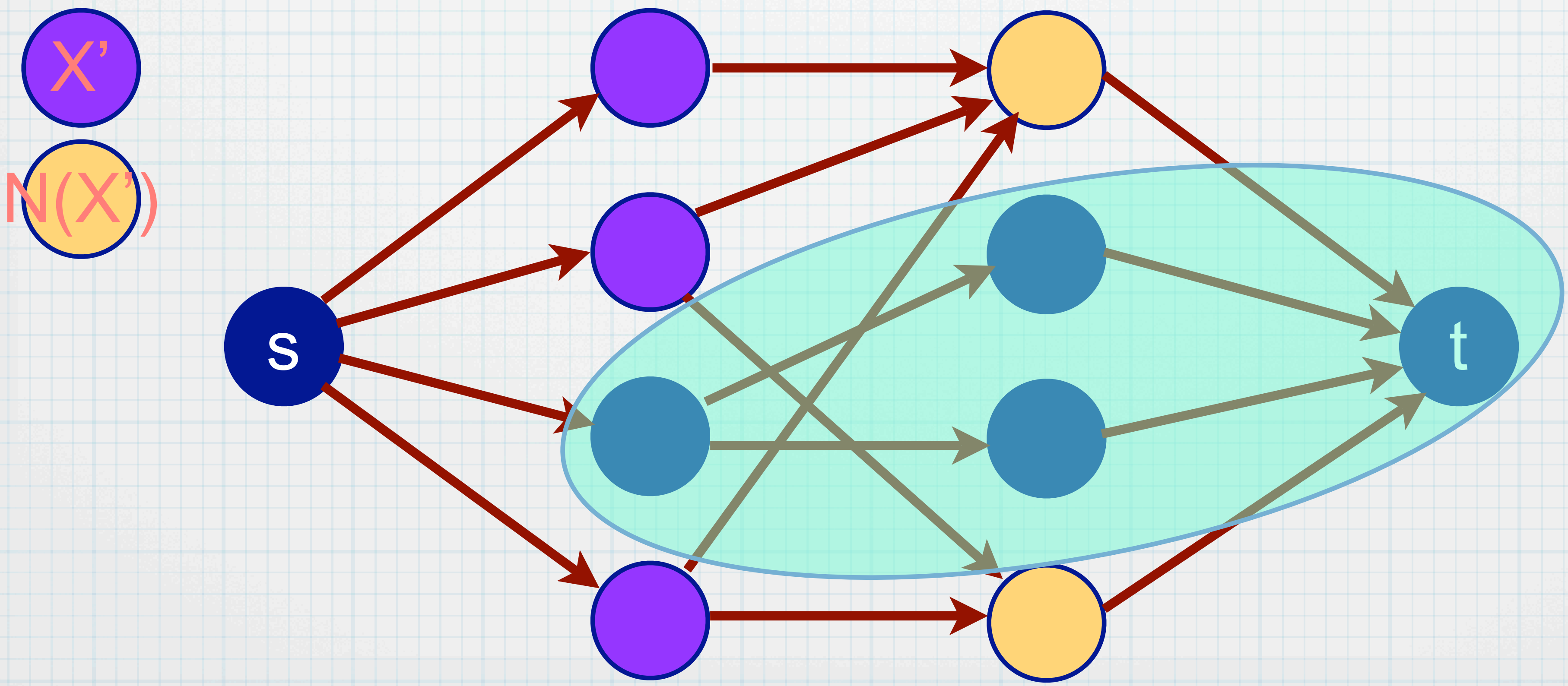
המשך הוכחת משפט Hall

מצד שני, אם יש $X' \subseteq X$ עבורה $|N(X')| < |X'|$ אזי:

$$c(\{s\} \cup X' \cup N(X'),\ (X \setminus X') \cup (Y \setminus N(X')) \cup \{t\}) =$$

$$|X \setminus X'| + |N(X')| < |X|$$

ולכן $|M^*| < |X|$. $\therefore$



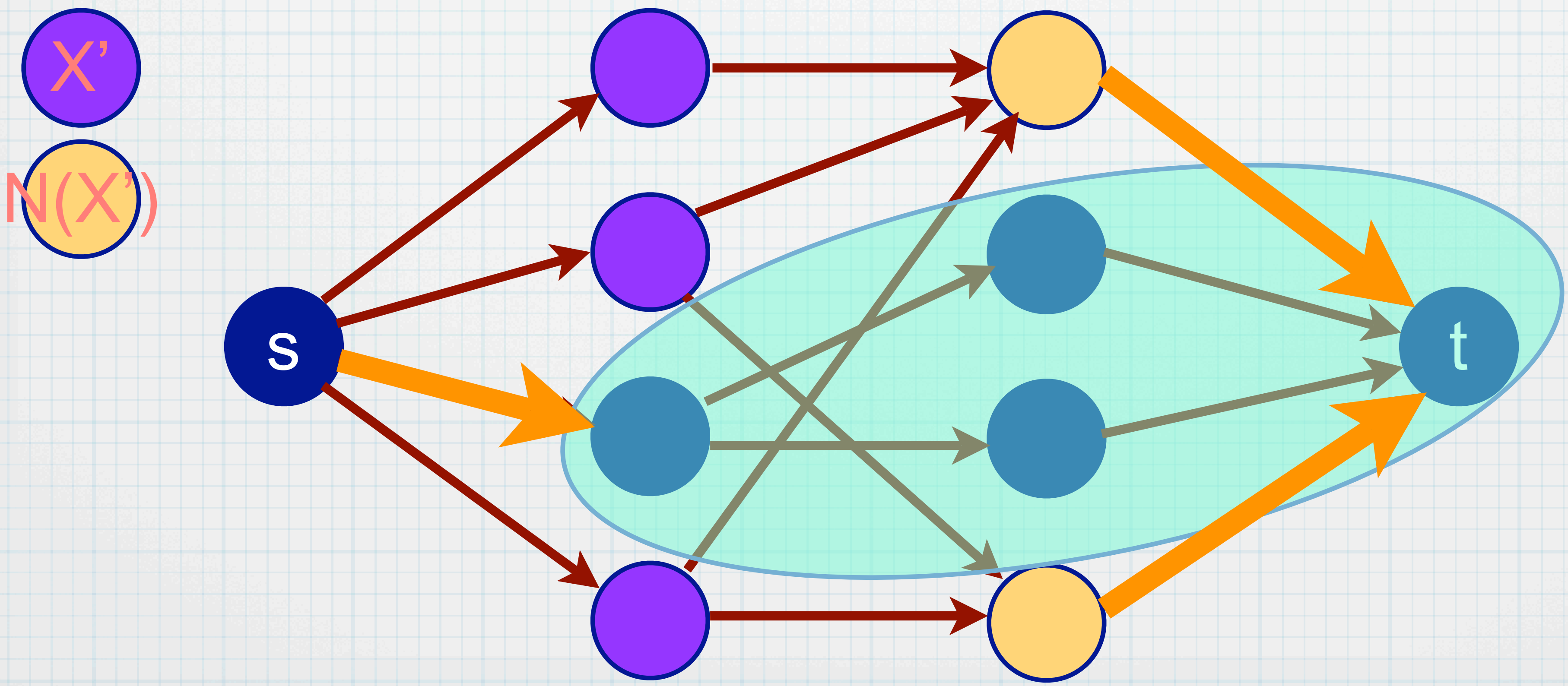
המשך הוכחת משפט Hall

מצד שני, אם יש $X' \subseteq X$ עבורה $|N(X')| < |X'|$ אזי:

$$c(\{s\} \cup X' \cup N(X') \,,\, (X \setminus X') \cup (Y \setminus N(X')) \cup \{t\}) =$$

$$|X \setminus X'| + |N(X')| < |X|$$

ולכן $|M^*| < |X|$. $\therefore$



נתונה תמונה דיגיטלית של $n \times m$ פיקסלים. רוצים להפריד בין האוביקטים המצולמים לבין הרקע. לצורך זה נתונים: (א) לכל פיקסל p, הסבירות a_p שהוא שייך לאחד האוביקטים והסבירות b_p שהוא שייך לרקע; (ב) לכל שני פיקסלים שכנים p ו-q, מידת הדמיון ביניהם $c_{p,q}$.

המטרה: למצוא חלוקה טובה ביותר של הפיקסלים לקבוצת האוביקטים A וקבוצת הרקע B. ערך החלוקה הוא:

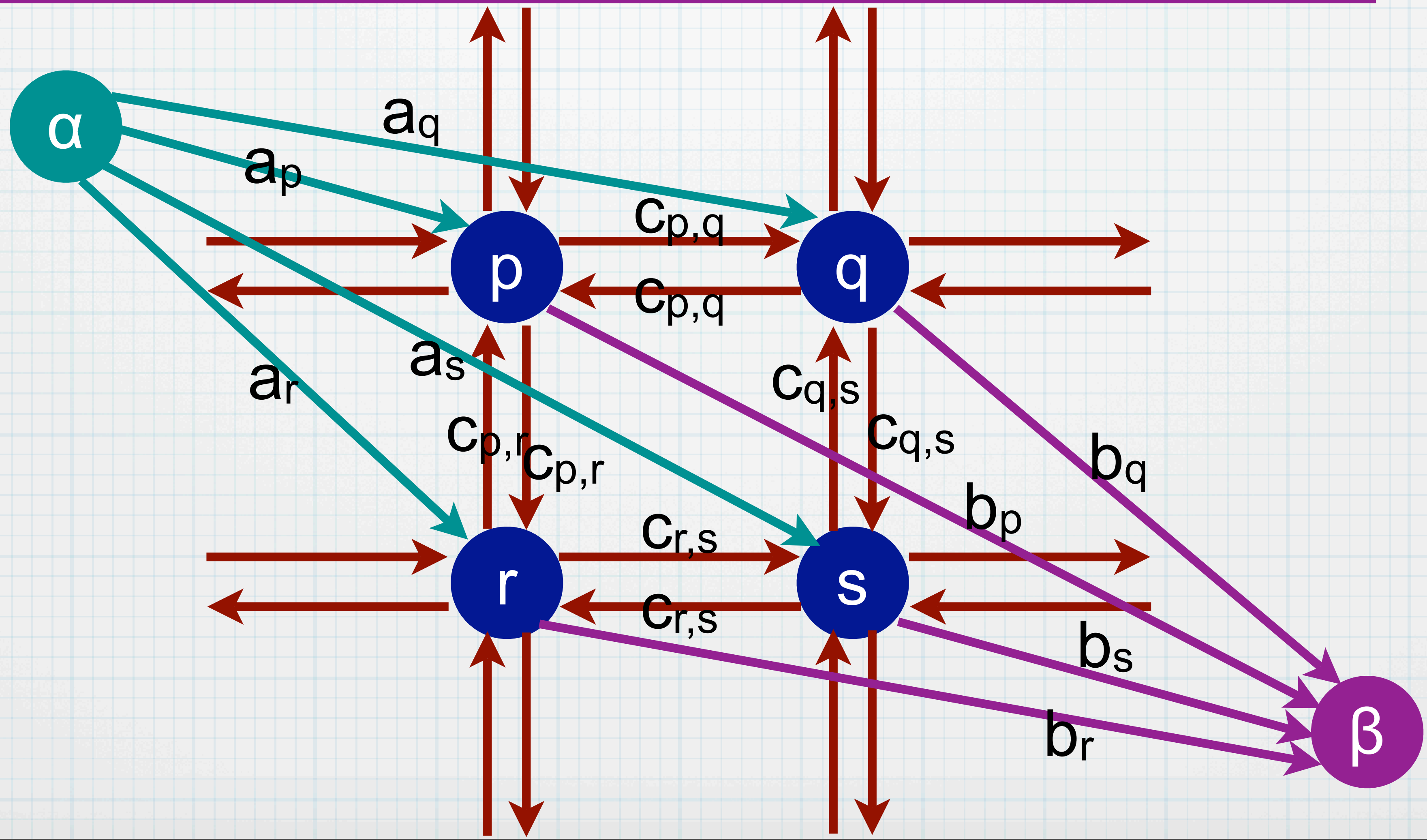
$$v(A,B) = \sum_{p\in A} a_p + \sum_{q\in B} b_q - \sum_{p,q \text{ שכנים מופרדים}} c_{p,q}$$

רוצים למצוא חלוקה שערכה מירבי.

מידול כגרף

במקום למצוא מקסימום של פונקצית המטרה המקורית, אפשר למצוא מינימום של:

$\sum_p(a_p+b_p)-v(A,B) = \sum_{p\in B} a_p + \sum_{q\in A} b_q + \sum_{p,q}$ שכנים מופרדים $c_{p,q}$



האלגוריתם

נמצא חתך מינימום עם α כמקור ו-β כבור. קיבול חתך הוא בדיוק $v(A,B)$ - V, באשר A קבוצת הפיקסלים בצד של α ו-B קבוצת הפיקסלים בצד של β.