קומבינטוריקה למדעי המחשב - הרצאות

גדי אלכסנדרוביץ'

תוכן עניינים

1 מבוא

קומבינטוריקה היא התחום במתמטיקה שעוסק באובייקטים סופיים, וכפי שניתן לנחש מההגדרה העמומה הזו מדובר על תחום רחב ביותר שגבולותיו לא פשוטים להגדרה. בקורס הזה נעסוק בשני תחומים עיקריים: בעיות ספירה (קומבינטוריקה אנומרטיבית), ומבוא לתורת הגרפים.

בעיות ספירה הן בעיות מהצורה הבאה: בהינתן הגדרה כלשהי של קבוצה סופית של אובייקטים, כמה אובייקטים יש בקבוצה? לדוגמה, כמה ידיים אפשריות קיימות במשחק ברידג'? כמה תוצאות אפשריות יש בלוטו? כמה הרכבי בתים אפשריים קיימים בטורניר כדורסל? כמה מצבים אפשריים יש ללוח במשחק שחמט? וכדומה.

לרוב השאלות מנוסחות על קבוצת אובייקטים שמוגדרת באמצעות פרמטר n כלשהו. למשל, בכמה דרכים אפשר להציג את המספר n כסכום של מספרים טבעיים קטנים ממנו? בכמה דרכים ניתן לחלק מצולע משוכלל בעל n צלעות למשולשים? בכמה דרכים יכול דוור מבולבל/מזכירה מבולבלת (תלוי את מי רוצים להעליב) לחלק n מכתבים כך שאף אחד לא יגיע ליעדו?

שאלות כאלו צצות באופן טבעי במדעי המחשב, כשההקשר הנפוץ ביותר הוא ניתוח סיבוכיות של אלגוריתמים, שם הקומבינטוריקה נחוצה הן בכדי להבין את מספר צעדי הריצה שהאלגוריתם מבצע והן את גודל מרחב הקלטים שאיתו הוא מנסה להתמודד. הקשר חשוב אחר הוא בעיות הסתברותיות ואלגוריתמים הסתברותיים; הבנה של ההסתברות של מאורע מסויים דורשת לרוב ניתוח קומבינטורי של כמות התוצאות האפשריות (למשל, ידיעת ההסתברות לזכיה בלוטו דורשת הבנה של כמות התוצאות האפשריות בלוטו).

הפתרון הטוב ביותר לבעיה קומבינטורית הוא כמובן מספר מדויק, ובמקרה של בעיה שתלויה בפרמטר n, נוסחה פשוטה שתלויה ב־n ־ למשל, מספר תתי־הקבוצות של קבוצה מגודל n הוא בדיוק 2^n. בקורס זה תיווצר `אשליה` שרבות הבעיות שניתן למצוא להן נוסחה מדוייקת שכזו, שכן רבות מהבעיות שנציג בקורס אכן יהיו כאלו; בעולם האמיתי נוסחה סגורה שכזו היא נדירה בהרבה, ולעתים קרובות גם איננה מועילה במיוחד. עיקר העניין הוא ב**סדר הגודל** של הפתרון; קירוב אסימפטוטי כלשהו אליו. מכיוון שאנו עוסקים במבוא בלבד, בקורס מבוא זה לא ניגע בכלל בעניינים אלו, שהם לב לבה של הקומבינטוריקה האנומרטיבית.

עבור בעיות שלא קל למצוא להן פתרון סגור פשוט באמצעות שיקולים אלמנטריים נלמד מקצת מהכלים החזקים יותר שמאפשרים התמודדות עם הבעיה ־ עקרון ההכלה וההפרדה, נוסחאות נסיגה ופתרונן ופונקציות יוצרות. כמו כן נדבר על הבינום של ניוטון, משולש פסקל ועקרון שובך היונים, שהם מושגים מתמטיים בסיסיים הקשורים לספירה שהדרך הטובה ביותר להבינם הוא במסגרת הקומבינטוריקה.

תורת הגרפים עוסקת באובייקט שהוא אולי המרכזי במדעי המחשב ־ אוסף של איברים (`צמתים`) שחלקם מחוברים אלו לאלו (על ידי `קשתות`). גרפים ממדלים אינספור בעיות ומושגים במדעי המחשב ־ החל ברשתות חברתיות ותרשימי זרימה של תוכנות, עבור במעגלים בוליאניים וקוונטיים וכלה במפות של מערכת כבישים. לא מעט מהאלגוריתמים הבסיסיים במדעי המחשב מנוסחים על גרפים, ובהתאם לכך אנו רוצים להכיר כאן את ההגדרות הפורמליות והתכונות הבסיסיות שמערבות גרפים, אם כי כמעט ולא נעסוק כאן באלגוריתמים על גרפים. בחלק זה של הקורס הגישה תהיה מעט פורמלית ומדויקת יותר מאשר בחלקו הראשון של הקורס; ננצל את הפשטות היחסית של החומר שבו אנו עוסקים כדי להמחיש את שיטות הלימוד הנפוצה במתמטיקה של "הגדרה־משפט־הוכחה".

חלק I
קומבינטוריקה אנומרטיבית

2 עקרונות ספירה בסיסיים

בפרק זה נציג את `כלי העבודה` הבסיסיים של קומבינטוריקה אנומרטיבית ־ העקרונות המנחים שמשמשים אותנו בפתרון כמעט כל בעיה קומבינטורית, ופתרונות של כמה בעיות יסודיות שמשתמשים בהן לרוב בתור אבן הבניין לפתרון בעיות מורכבות יותר.

בלשון מתמטית פורמלית, בעיית ספירה קומבינטורית היא זו: נתונה קבוצה סופית A ואנו מעוניינים למצוא מהו $|A|$ ־ מספר האיברים ב־A.

2.1 עקרון החיבור ועקרון הכפל

ראשית נציג את **עקרון החיבור**.

דוגמא במשחק לוח הזוי כלשהו השחקן יכול בתורו או להטיל קוביה רגילה או להטיל מטבע. כמה תוצאות אפשריות ישנן?

במקרה זה ישנן 6 תוצאות אפשריות להטלת הקוביה, ו־2 תוצאות אפשריות להטלת המטבע, ולכן בסך הכל יש $6+2=8$ תוצאות אפשריות.

דוגמא כמה מהלכי פתיחה חוקיים יש לשחקן בכלים הלבנים במשחק השחמט?

במקרה זה כל רגלי של הלבן יכול לנוע צעד או שני צעדים קדימה, וכל אחד מהפרשים יכול לנוע אחד משני צעדים אפשריים קדימה. בסך הכל יש $8+8+2+2=20$ מהלכי פתיחה אפשריים.

הכללה מה העקרון המשותף לדוגמאות שלעיל? בכל המקרים ביצענו בחירה מתוך כמה `סוגי` אפשרויות שונים, כשהבחירה היא מסוג **או** ־ או שמזיזים רגלי צעד אחד (8 אפשרויות) או שמזיזים רגלי שני צעדים (8 אפשרויות), או שמזיזים את הפרש הימני (2 אפשרויות) או שמזיזים את הפרש השמאלי (2 אפשרויות). זהו מקרה פרטי של **עקרון החיבור**:

טענה 1.2 (עקרון החיבור) אם קיימות n_1 אפשרויות לבצע בחירה מסוג אחד ו־n_2 אפשרויות לבצע בחירה מסוג שני, אז קיימות n_1+n_2 אפשרויות לבצע בחירה מסוג אחד או מהסוג השני.

בניסוח מתמטי פורמלי, אם A,B הן שתי קבוצות זרות אז $|A\cup B|=|A|+|B|$

כעת נעבור להציג את **עקרון הכפל**.

דוגמא סטודנט צריך לקחת בסמסטר קורס ספורט אחד מבין שחמט וברידג' (מעשה שהיה באמת בטכניון), וקורס מדעי אחד מבין פיזיקה וכימיה. כמה אפשרויות בחירה יש לו?

לסטודנט יש 2 בחירות לקורס ספורט ו־2 בחירות לקורס מדעי וכל זוג אפשרי של בחירות הוא חוקי. ישנן אם כן 4 אפשרויות:

1. שחמט, פיזיקה
2. שחמט, כימיה

3. ברידג', פיזיקה

4. ברידג', כימיה

כלומר, לכל אחת מהאפשרויות לבחירה הראשונה, בן זוגה יכול להיות כל אחת מהאפשרויות בבחירה השניה.

דוגמא במונופול בכל סיבוב מוטלות שתי קוביות. כמה תוצאות אפשריות ישנן?

לכל אחת מהקוביות יש 6 תוצאות אפשריות, ואנחנו מתעניינים בכל **הזוגות** של תוצאה של קוביה אחת ותוצאה של הקוביה השניה, כך שבסך הכל יש לנו 36 זוגות אפשריים מהצורה (i,j) כאשר $1 \le i,j \le 6$. שימו לב שיש הטלות קוביה שונות שמתורגמות לאותו מספר צעדים על הלוח, למשל ההטלה $(2,5)$ מתורגמת ל־7 צעדים, כמו ההטלה $(3,4)$ וההטלה $(5,2)$; כאן אנו סופרים את התוצאות האפשריות של הטלת הקוביה, לא את מספר הצעדים האפשריים בכל סיבוב.

הכללה מה העקרון המשותף לדוגמאות שלעיל? בכל המקרים ביצענו בחירה **דו שלבית**. הבחירה היא מסוג `**וגם**` ־ יש לבחור גם קורס ספורט וגם קורס מדעי. זהו מקרה פרטי של **עקרון הכפל**:

טענה 2.2 (עקרון הכפל) אם קיימות n_1 אפשרויות לבצע בחירה מסוג אחד ו־n_2 אפשרויות לבצע בחירה מסוג שני, אז קיימות $n_1 \cdot n_2$ אפשרויות לבצע בחירה מסוג אחד **וגם** בחירה מהסוג השני.

בניסוח מתמטי פורמלי, אם A, B הן שתי קבוצות (לא בהכרח זרות) אז $|A \times B| = |A| \cdot |B|$ ($A \times B$ הוא אוסף כל הזוגות של איבר מ־A ואיבר מ־B).

גם בעיקרון החיבור וגם בעיקרון הכפל ניסחנו את העיקרון למקרה של זוג בחירות, עם n_1 אפשרויות לראשונה ו־n_2 אפשרויות לשניה. אפשר להכליל את העקרונות גם למספר סופי כלשהו של בחירות עם $n_1, n_2, \ldots, n_k$ אפשרויות; יש $n_1 + n_2 + \ldots + n_k$ אפשרויות לבצע בחירה באחת מבין האפשרויות (עקרון החיבור) ו־$n_1 \cdot n_2 \cdots n_k$ אפשרויות לבצע בחירה k שלבית שבה בשלב i בוחרים מבין n_i אפשרויות.

2.2 תמורות (סידור בשורה)

בכמה דרכים ניתן לסדר n ילדים בשורה?

זהו שימוש ישיר בעקרון הכפל. נפתור את הבעיה בשלוש גישות שונות על מנת להמחיש שניתן להגיע לתוצאה מכמה נקודות מבט שונות (מה שמחזק את האמונה שלנו בנכונות התוצאה).

- בגישה הראשונה נניח שאנו עוברים אחד אחד על הילדים ולכל ילד בוחרים לו מקום בשורה מבין אלו שעדיין פנויים (אנו מניחים מראש שהשורה מחולקת ל־n `תאים`). עבור הילד הראשון יש n בחירות, עבור השני יש רק $n-1$ בחירות (כי מקום אחד כבר תפוס), עבור השלישי $n-2$ בחירות וכן הלאה עד לילד האחרון שיש לו בדיוק בחירה אחת.

- בגישה השניה נניח שאנו עוברים אחד אחד על המקומות בשורה ולכל אחד מהם בוחרים איזה ילד יהיה בו. גם כאן יש n בחירות לילד הראשון, $n-1$ בחירות לילד השני (כי כבר סידרנו את אחד הילדים בשורה ולא ניתן לבחור בו שוב), וכן הלאה עד למקום האחרון שבו יכול להיות רק ילד אחד בלבד ־ זה שנשאר.

- בגישה השלישית נבנה את השורה ילד אחרי ילד מבלי להניח שהשורה מחולקת מראש לתאים, כשבכל פעם אנו שואלים את עצמנו להיכן אפשר להכניס את הילד הבא בתור. יש בחירה 1 לילד הראשון (כי כרגע השורה ריקה), 2 בחירות לילד השני (משמאל או מימין לילד הקיים), 3 בחירות עבור השלישי (משמאל לזוג הקיים, מימין לו או באמצע) וכן הלאה עד n בחירות לילד האחרון.

בכל המקרים קיבלנו את אותה התוצאה: $1\cdot2\cdot3\cdots n$. בגלל השימושיות הרבה של הפעולה הזו במתמטיקה יש לה שם וסימון מיוחד: $n!=1\cdot2\cdots n$ (קרי `$n$ עצרת`).

את $n!$ ניתן להגדיר גם באופן רקורסיבי שהוא מועיל לעתים:

- $0!=1$
- $n!=n\cdot(n-1)!$ לכל $n\geq1$.

הערה 3.2 אין ל־$n!$ נוסחה פשוטה, אך יש לה קירוב מצויין, שהופך לשימושי מאוד כאשר עוסקים בקומבינטוריקה בחיפוש אחרי סדרי גודל בלבד: **נוסחת סטירלינג**, $n!\approx\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}$ (המשמעות הפורמלית של הסימון $\approx$ היא שמנת שני צדדי המשוואה שואפת ל־1 כאשר n שואף לאינסוף, דהיינו $1\underset{n\to\infty}{\to}\left(n!/\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}\right)$). בקורס הזה לא נזדקק לנוסחת סטירלינג אך מומלץ להיות מודעים לקיומה.

דוגמא בכמה דרכים ניתן לסדר n ילדים בשורה אם ידוע שאליס ובוב חברים ורוצים להיות זה ליד זו?

כמקודם, נפתור במספר דרכים:

- נסדר את כל הילדים בשורה למעט אליס: $(n-1)!$ אפשרויות. כעת אליס יכולה להיות משמאל או מימין לבוב, ולכן סך הכל יש $2\left(n-1\right)!$ אפשרויות.
- `נדביק` יחד את אליס ובוב ונחשוב עליהם בתור ילד אחד (בוליס?). נסדר את $n-1$ הילדים (הילדים הרגילים ו`בוליס`) בשורה ונקבל $(n-1)!$ אפשרויות. כעת, יש שתי אפשרויות לסידור הפנימי של `בוליס` (בוב מימין ואליס משמאל או בוב משמאל ואליס מימין) ולכן מעקרון הכפל נקבל $2\left(n-1\right)!$ אפשרויות.

דוגמא אליס ובוב רבו. בכמה דרכים ניתן לסדר את n הילדים בשורה כך שאליס **אינה** ליד בוב?

- נסדר את כל הילדים בשורה למעט אליס - $(n-1)!$ אפשרויות. כעת אליס יכולה להיות בכל מקום למעט ימין או שמאל בוב, ולכן יש לה $n-2$ אפשרויות ומעקרון הכפל נקבל $(n-2)\left(n-1\right)!$.
- `עקרון החיסור`: מספר אפשרויות הסידור בשורה של הילדים הוא $n!$, וראינו כבר כי בדיוק ב־$2\left(n-1\right)!$ מתוך האפשרויות הללו אליס היא ליד בוב. אז מה שנותר הוא

$$n!-2\left(n-1\right)! = n\left(n-1\right)!-2\left(n-1\right)!=\left(n-2\right)\left(n-1\right)!$$

אפשרויות.

2.3 חליפות (בחירה ללא חזרות ועם חשיבות לסדר)

דוגמא יש ספסל עם 5 מקומות ו־20 ילדים. בכמה דרכים אפשר לסדר 5 מבין הילדים על הספסל?

- יש 20 בחירות של ילד למקום הראשון, 19 למקום השני וכן הלאה עד 16 למקום החמישי: $20 \cdot 19 \cdot 18 \cdot 17 \cdot 16$. על פי ההגדרה קל לראות שזה שווה ל־$\frac{20!}{15!}$.
- `עקרון החילוק` ־ נסדר את 20 הילדים בשורה ־ $20!$ אפשרויות. כעת ניקח את חמשת הראשונים ונשים על הספסל בסדר שלהם בשורה. בצורה זו יש לנו **ספירות כפולות** ־ כל סידור של ילדים על הספסל מתאים בדיוק ל־$15!$ מבין אפשרויות הסידור של הילדים בשורה ־ מספר הסידורים הפנימיים של 15 הילדים שאינם במקומות הראשונים. גישה זו קשה יותר להבנה מהגישה הראשונה אך חזקה בהרבה ומסבירה ישירות את התוצאה $\frac{20!}{15!}$.

הכללה הדוגמה שלעיל היא מקרה פרטי של הבעיה הבאה: בכמה דרכים ניתן לבחור k מתוך n אובייקטים ($k \leq n$) כאשר יש חשיבות לסדר שבו נבחרים האיברים?

כפי שראינו בדוגמה, הפתרון הכללי הוא $n \cdot (n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}$.

לצורך פשטות משתמשים לעתים בסימון $P(n,k) = P_n^k = \frac{n!}{(n-k)!}$. **אזהרה!** יש עוד שלל צורות כתיבה מוזרות לסימון (דוגמת P_k^n או ${}_nP_k$) והדרך הבטוחה ביותר להימנע מבלבול היא פשוט לא להשתמש בו כלל.

2.4 צירופים (בחירה ללא חזרות וללא חשיבות לסדר)

בכמה דרכים ניתן לבחור k מתוך n אובייקטים כאשר אין חשיבות לסדר שבו נבחרים האיברים?

נסמן ב־C_n^k את המספר הזה. אז $C_n^k \cdot k! = P_n^k$ על פי עיקרון הכפל ־ קודם בוחרים k איברים בלי חשיבות לסדר (C_n^k אפשרויות) ואז בוחרים אחד מ־$k!$ הסידורים האפשריים שלהם. מספר זה שווה למספר האפשרויות לבחור את האיברים כאשר מלכתחילה מתחשבים בסדר.

מכאן ש־$C_n^k = \frac{P_n^k}{k!} = \frac{n!}{k!(n-k)!}$.

סימון אחר ומקובל בהרבה ל־C_n^k הוא זה: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. בסימון זה נשתמש מכאן ואילך. הסימון מוגדר רק כאשר $k \leq n$; עבור $k > n$ מגדירים $\binom{n}{k} = 0$ (לא ניתן לבחור k מתוך n איברים ללא חזרות אם n קטן מ־k).

2.5 סידור בשורה עם עצמים זהים

נתונים k_1 כדורים מצבע אחד, k_2 כדורים מצבע אחר וכן הלאה עד k_t כדורים מצבע t. נסמן $n = \sum_{i=1}^{t} k_i$. בכמה דרכים ניתן לסדר את הכדורים בשורה?

דרך הפתרון היא לחשוב על כל הכדורים כשונים אלו מאלו, לסדר אותם בשורה ($n!$ אפשרויות) ואז לכל צבע לחלק במספר הסידורים הפנימיים של אותו הצבע ($k_i!$ לכל i).

מקבלים: $\frac{n!}{k_1!k_2!\cdots k_t!}$.

צירופים הם מקרה פרטי כאשר $t=2$ (אפשר לחשוב על כך כאילו בוחרים את המקומות לכדורים מהצבע הראשון, ואז המקומות עבור הצבע השני נקבעים מאליהם).

אפשר לחשוב על $\frac{n!}{k_1!k_2!\cdots k_t!}$ בתור הכללה של צירופים: מספר האפשרויות לבחור k_1 מתוך n אובייקטים עבור קבוצה אחת, k_2 עבור קבוצה שניה וכן הלאה, כאשר אין חשיבות לסדר הבחירה.

לעתים משתמשים בסימון $\binom{n}{k_1,k_2,\ldots,k_t} = \frac{n!}{k_1!k_2!\cdots k_t!}$.

2.6 בחירה עם חזרות ועם חשיבות לסדר

דוגמא בכמה דרכים ניתן לבנות מספר בן 5 ספרות בעזרת הספרות 1,2,3 בלבד?

יש לנו 3 אפשרויות בחירה לספרה הראשונה, 3 לספרה השניה וכן הלאה. על פי עקרון הכפל נקבל $3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^5$.

בדוגמה זו יש חשיבות לסדר הבחירה (13 איננו אותו מספר כמו 31) ויש חזרות ־ ניתן לבחור את אותה ספרה יותר מפעם אחת. המקרה הכללי מטופל באותו האופן:

טענה 4.2 מספר האפשרויות לבחור k מתוך n אובייקטים עם חזרות ועם חשיבות לסדר הוא n^k.

שימו לב כי כאן לא נדרש ש־$k \le n$ (כי ניתן לבחור את אותו איבר מספר פעמים).

2.7 בחירה עם חזרות וללא חשיבות לסדר

דוגמא כמה סדרות **מונוטוניות לא יורדות** באורך k קיימות מעל $1,2,\ldots n$?

דוגמה לסדרה מונוטונית לא יורדת שכזו עבור $k=6, n=7$: $(1,3,3,3,5,7)$.

הבחנה: $a_1,a_2,\ldots,a_k$ היא סדרה מונוטונית לא יורדת מעל $1,2,\ldots n$ אם ורק אם $a_1+0, a_2+1, \ldots, a_k+(k-1)$ היא סדרה מונוטונית עולה מעל $1,2,\ldots,n+(k-1)$.

סדרות מונוטוניות עולות קל בהרבה לספור מבחינה רעיונית: בוחרים את k המספרים שישתתפו בסדרה ללא חשיבות לסדר, והסדרה כבר נקבעת מעצמה על פיהם.

לכן קיבלנו $\binom{n+k-1}{k}$ וזה גם הפתרון לבעיה המקורית של ספירת סדרות מונוטוניות לא יורדות.

שימו לב שזו דוגמא לבחירה עם חזרות (ניתן לבחור את אותו מספר כמה פעמים) וללא חשיבות לסדר (הסידור של המספרים בסדרה נקבע באופן יחיד).

דוגמא מה מספר הדרכים להכנסת k כדורים זהים ל־n תאים שונים?

נוח לחשוב על התהליך באופן הפוך ־ k הכדורים מסודרים בשורה, ויש לבנות סביבם `מחיצות` כדי ליצור n תאים, כך שצריך $n-1$ מחיצות.

ניתן לתאר באופן סכמטי באמצעות סדרה: 01001 כאשר 0 מייצג כדור ו־1 מייצג מחיצה. כאן יש שלושה תאים: בשמאלי יש כדור אחד, באמצעי שניים ובימני אפס.

אם כן, המספר הוא מספר הסדרות הבינאריות עם k אפסים ו־$n-1$ אחדים. כל שנדרש הוא לבחור את מיקום האפסים כך שיש $\binom{n+k-1}{k}$ אפשרויות.

גם כאן הייתה בחירה עם חזרות (ניתן להכניס כדור לאותו תא פעמים רבות) וללא חשיבות לסדר (הכדורים זהים ולכן לא חשוב אם קודם מכניסים אחד לתא 1 ואז לתא 2 או הפוך ־ בסוף בשני התאים יהיה כדור בודד).

דוגמא כמה פתרונות במספרים שלמים אי שליליים יש למשוואה $x_1+x_2+\cdots+x_n = k$?

קל לראות שיש התאמה חח`ע ועל בין הבעיה הזו לבעיה הקודמת: המשתנים הם התאים, וערכו של כל משתנה הוא מספר הכדורים שהוכנסו אליו ולכן גם כאן הפתרון הוא $\binom{n+k-1}{k}$.

2.8 סיכום

- סידור n עצמים בשורה: $n!$

- סידור n עצמים בשורה כאשר הם מחולקים למחלקות זהות בגדלים $k_1,\ldots,k_t$: $\frac{n!}{k_1!\ldots k_t!}$
- בחירות של k מתוך n:

סדר\חזרות	חשוב	לא חשוב
בלי	$P_n^k=\frac{n!}{(n-k)!}$	$C_n^k=\binom{n}{k}$
עם	$PP_n^k=n^k$	$CC_n^k=\binom{n+k-1}{k}$

עוד כמה תרגילים ופתרונים:

דוגמא כמה `ידיים` שונות של 5 קלפים בפוקר ניתן לקבל?
זוהי בחירה של 5 קלפים מתוך 52 ללא חשיבות לסדר וללא חזרה ולכן $\binom{52}{5}=\frac{52!}{5!47!}$

דוגמא כמה אפשרויות למילוי טופס טוטו קיימות?
כאן ישנם 16 טורים שבכל אחד מסמנים 1 או 2 או X, כלומר בחירה עם חשיבות לסדר ועם חזרות של 16 מ־3, ולכן $3^{16}=43,046,721$.

דוגמא מה ההסתברות לזכות בלוטו הישראלי, שבו ממלאים 6 מ־37 מספרים ועוד 1 מ־7 `מספרים חזקים`?
כאן יש לנו שתי בחירות ללא חשיבות לסדר וללא חזרות ואנו מפעילים עליהן את עקרון הכפל ומקבלים $7\cdot\binom{37}{6}=16,273,488$ ולכן סיכויי הזכייה הם $1:16,273,488$.

דוגמא `וקטור בינארי` מאורך n הוא פשוט סדרה של n ערכים שהם 0 או 1.
ברור כי יש 2^n וקטורים בינאריים מאורך n (בחירה עם חזרות מתוך $\{0,1\}$ ועם חשיבות לסדר כי וקטור הוא סדרה).

כמה וקטורים בינאריים קיימים שבהם יש לפחות מופע אחד של 1?

פתרון נפוץ **ושגוי** לשאלות מסוג זה הוא כדלהלן: נבחר אחד מ־n המקומות בתור המקום שבו יופיע ה־1 שאנחנו `מחוייבים` לו, ואז נבחר בחופשיות את הכניסות עבור שאר המקומות, ונקבל $n\cdot 2^{n-1}$ אפשרויות.

דרך לזהות את השגיאה היא לבדוק את הפתרון עבור ערכים קטנים: עבור $n=2$ נקבל מהנוסחה כי ישנם $2\cdot 2^1=4$ וקטורים בינאריים מתאימים, אבל קל לראות כי קיימים רק שלושה: $11,01,10$. ביצענו **ספירה כפולה**.

הספירה הכפולה ספרה את הוקטור 11 פעמיים: פעם אחת נבחר 1 להיות במקום הראשון בשלב הראשון, ובשלב השני נבחר שבמקום השני יופיע גם כן אחד; בפעם האחרת נבחר 1 להיות דווקא במקום השני, ואילו ה־1 שבמקום הראשון נבחר אחר כך. הדבר אינו נוגד את עקרון הכפל שכן עקרון הכפל דורש שכל זוג בחירות יוביל לתוצאה שונה[1], ואילו כאן יש שני זוגות בחירות שונים אפשריים שמובילים לאותו האובייקט בדיוק.

הדרך הנכונה לפתור את התרגיל הזה היא באמצעות **עקרון החיסור**: ישנו רק וקטור בודד מאורך n שלא מכיל 1־ים (הוקטור שכולו אפסים) ולכן יש 2^n-1 וקטורים מאורך n שמכילים 1 לפחות פעם אחת (כי על פי עקרון החיבור, מספר הוקטורים הכולל 2^n שווה לסכום של מספר הוקטורים שלא מכילים 1־ים ומספר הוקטורים שמכילים 1 אחד לפחות).

[1] עקרון הכפל סופר כמה זוגות של בחירות ישנם; השימוש שלנו בעיקרון הכפל מניח במובלע שהאובייקטים שאותם אנחנו סופרים נוצרים על ידי זוגות הבחירות הללו כך שכל אובייקט נוצר בידי זוג אחד בדיוק.

דוגמא כמה פתרונות בשלמים יש למשוואה $x_1 + x_2 + x_3 + x_4 + x_5 = 30$ אם דורשים כי $x_i \geq 1$?

כאן אנחנו רוצים להשתמש בתוצאה המוכרת של בחירות בלי חשיבות לסדר ועם חזרות, אבל שם התוצאה תקפה עבור $x_i \geq 0$.

הרעיון האינטואיטיבי - מחלקים 30 כדורים לחמישה תאים תחת האילוץ שאין תא ריק, אז קודם כל נשים כדור אחד בכל תא ואז נחלק את 25 הכדורים הנותרים באופן חופשי.

בפועל: נשתמש בתעלול ונגדיר משתנים חדשים y_i כך ש-$x_i = y_i + 1$. נציב במשוואה המקורית ונקבל:

$$(y_1 + 1) + (y_2 + 1) + (y_3 + 1) + (y_4 + 1) + (y_5 + 1) = 30$$

ובניסוח שקול:

$$y_1 + y_2 + y_3 + y_4 + y_5 = 25, \quad y_i \geq 0.$$

ולכן הפתרון הוא $\binom{5+25-1}{25} = \binom{29}{25}$.

דוגמא יהא $\mathbb{F}_q$ שדה סופי עם q איברים. כמה מטריצות הפיכות 2×2 מעל $\mathbb{F}_q$ קיימות?

עבור מטריצות 2×2, מטריצה היא הפיכה אם ורק אם אף שורה איננה כפל בסקלר של השורה השניה. עבור השורה הראשונה כל שורה אפשרית פרט לשורה שכולה אפסים היא לגיטמית, ומכיוון שיש q ערכים אפשריים לכל כניסה, מעקרון הכפל יש q^2 שורות אפשריות, ולאחר חיסור שורת האפסים נקבל $q^2 - 1$ אפשרויות.

כעת, בהינתן השורה הראשונה כל אחת מ-q^2 השורות אפשריות עבור השורה השניה פרט לאלו שהן כפל בסקלר של השורה הראשונה. קיימים q סקלרים ואותה שורה לא מתקבלת על ידי כפל בשני סקלרים שונים, כך שיש $q^2 - q$ שורות לגיטימיות בסך הכל.

מעקרון הכפל נקבל שיש $\left(q^2 - 1\right)\left(q^2 - q\right)$ מטריצות הפיכות מהסוג הנדרש.

3 עקרון שובך היונים

עקרון שובך היונים הוא אבחנה מתמטית פשוטה ביותר, ועם זאת הוא כלי שימושי מאוד בפתרון בעיות קיום רבות, לעתים בצורות מפתיעות למדי.

טענה 1.3 (עקרון שובך היונים): אם ב-n שובכים ישנן $n + 1$ יונים, אז קיים שובך שבו יש לפחות שתי יונים.

ניסוח כללי יותר: אם ב-n שובכים ישנן m יונים, אז קיים שובך שבו יש לפחות $\left\lceil \frac{m}{n} \right\rceil$ יונים.

הוכחת הטענה היא בשלילה - אם בכל שובך יש לכל היותר יונה אחת, אז יש בכולם יחד לא יותר מ-n יונים. באופן דומה מוכחת הטענה הכללית.

נפתח בדוגמאות פשוטות:

דוגמא קיימים בעולם שני אנשים לא קרחים בעלי בדיוק אותה כמות שערות על הראש. זה נובע מכך שמספר השערות על הראש נמדד במאות אלפים, בעוד שבעולם ישנם מיליארדי אנשים - האנשים הם היונים, מספרי השערות האפשריים הם השובכים.

דוגמא בחדר עם 366 אנשים קיימים שני אנשים בעלי אותו יום הולדת (אם מתעלמים מחריגים כמו ה-29 בפברואר)[2].

[2] פרדוקס יום ההולדת בתורת ההסתברות מראה שכבר אם ישנם 23 אנשים בחדר וימי ההולדת שלהם מתפלגים באופן אחיד, ההסתברות לשני אנשים בעלי אותו יום הולדת גדולה מחצי.

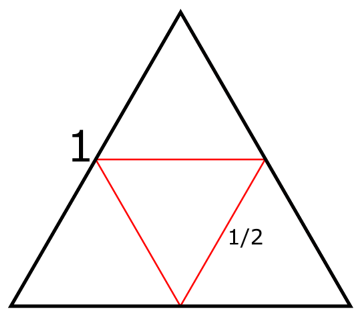
דוגמא בקורס עם למעלה מ־100 סטודנטים מובטח שיהיו שני סטודנטים שיקבלו את אותו הציון (אם כולם ניגשים לבחינה...)

דוגמא לא קיים כיווץ משמר מידע שמקטין כל קובץ: לכל n יש 2^n קבצים מאורך n ביטים ו־$1+2+4+\cdots+2^{n-1}=2^n-1$ קבצים מאורך לכל היותר $n-1$ ביטים ולכן מעקרון שובך היונים כל כיווץ שמקטין כל קובץ בן n ביטים חייב לגרום להתנגשות בין שני קבצים (ולכן המידע לא משומר במלואו ־ לא ניתן לשחזר כל קובץ מהכיווץ שלו). טיעון קצת יותר מחוכם מוכיח שאם הכיווץ מקטין ולו קובץ בודד, הוא חייב **להגדיל** קובץ אחר כלשהו.

נעבור כעת לדוגמאות מורכבות יותר.

דוגמא נתון משולש שווה צלעות עם צלע באורך 1. יש להוכיח כי לכל 5 נקודות במשולש יש שתיים במרחק קטן או שווה ל־$\frac{1}{2}$.

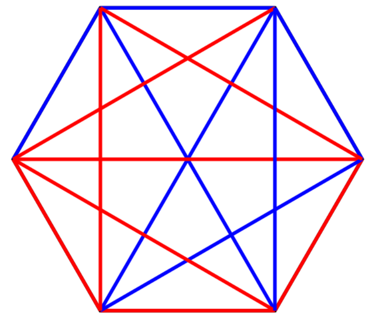
הפתרון: מחלקים את המשולש ל־4 משולשים שווי צלעות שאורך צלעם $\frac{1}{2}$:



המרחק בין שתי נקודות בתוך כל משולש הוא לא יותר מ־$\frac{1}{2}$, ועל פי עקרון שובך היונים יש שתי נקודות באותו משולש (הנקודות הן היונים, המשולשים הם השובכים).

דוגמא שישה אנשים נפגשו במסיבה וחלקם לחצו ידיים אחד לשני. יש להוכיח כי יש שלישייה של אנשים כך שכל חבריה או לחצו את ידיהם של כל חבריהם לשלישייה, או לחצו את ידו של אף חבר בשלישייה.

בניסוח גאומטרי, נתבונן במצולע משוכלל בן 6 קודקודים כך שכל האלכסונים האפשריים נמתחו בו (כלומר, כל קודקוד מחובר בקו לכל קודקוד אחר). נצבע כל אחד מהקווים באדום או כחול; יש להוכיח כי קיים משולש מונוכרומטי (שכל צלעותיו צבועות באותו הצבע).



הפתרון: נתבונן על אדם מס' 1. ישנם 5 קטעים שמחברים אותו עם שאר האנשים, ומעקרון שובך היונים המוכלל יש שלושה הצבועים באותו הצבע (הקווים הם היונים, הצבעים הם השובכים). נניח בלי הגבלת הכלליות שהצבע הזה היה אדום נתבונן על שלושת האנשים המחוברים לקווים הללו. אם קיים ביניהם זוג שהקו שמחבר אותו הוא בצבע אדום, אז ביחד עם אדם מס' 1 קיבלנו את השלישייה שלנו; ואם לא קיים ביניהם זוג כזה אז כל שלושת האנשים הללו מחוברים אלו לאלו עם קווים כחולים, ושוב קיבלנו את השלישייה שלנו.

בהערת אגב נציין שתוצאה זו היא מקרה פרטי של משפט כללי בקומבינטוריקה הנקרא **משפט רמזי**, ובתורו מהווה בסיס לתחום בקומבינטוריקה הנקרא **תורת רמזי**. לא נציג את המשפט בקורס.

דוגמא בכל קבוצה של 12 מספרים טבעיים דו ספרתיים קיימים שניים אשר הפרשם הוא מספר בעל שתי ספרות זהות.

הפתרון: השאריות האפשריות בחלוקה ב־11 של המספרים יהיו השובכים, והמספרים יהיו היונים. בהכרח יש שני מספרים בעלי אותה שארית בחלוקה ב־11 ולכן הפרשם יתחלק ב־11, ומכיוון שהוא מספר דו ספרתי הוא יהיה בעל שתי ספרות זהות.

דוגמא הייצוג העשרוני של כל מספר רציונלי הוא מחזורי.

כדי למצוא את הייצוג העשרוני של מספר רציונלי $\frac{a}{b}$ (עם $a<b$) מבצעים חילוק ארוך; ניתן לתאר זאת כחזרה אינסופית על הצעדים הבאים:

1. $a \leftarrow 10 \cdot a$
2. פלוט את $\left\lfloor \frac{a}{b} \right\rfloor$.
3. $a \leftarrow a\%b$.

האלגוריתם עצמו הוא אינסופי, אבל יש רק מספר סופי של ערכים ש־a יכול לקבל בשלב 3 (הערכים בין 0 ו־$b-1$) ומכאן שהחל משלב מסויים האלגוריתם יחל לחזור על עצמו, שכן ערכו של a בשלב מסויים קובע באופן יחיד את כל המשך האלגוריתם[3].

[3] רעיון זה, לפיו ריצה אינסופית של אלגוריתם שיכול להיות רק במספר סופי של `מצבים`, תחיל חזרות משלב מסויים הוא שימושי ביותר ובא לידי ביטוי, למשל, בלמת הניפוח בקורס באוטומטים ושפות פורמליות, ופתרון של וריאנטים מוגבלים של בעיית העצירה בקורס בתורת החישוביות.

4 הבינום של ניוטון ומשולש פסקל

4.1 הבינום של ניוטון

בבית הספר לומדים את נוסחת הכפל המקוצר $(a+b)^2 = a^2 + 2ab + b^2$.

הנוסחה $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ גם היא מוצגת בבית הספר אבל ככל הנראה זכורה פחות.

נראה כעת כיצד מגיעים לנוסחאות אלו וכיצד שיטה זו מטפלת גם במקרה הכללי של $(a+b)^n$.

ראשית, נשים לב ש־

$$
\begin{aligned}
(a+b)^2 &= (a+b)(a+b) = aa + ab + ba + bb \\
&= a^2 + 2ab + b^2
\end{aligned}
$$

ה־$2ab$ נובע מ־$ab + ba$ ומכך שכפל הוא קומוטטיבי, כלומר $ab = ba$.

באופן דומה:

$$
\begin{aligned}
(a+b)^3 &= (a+b)(a+b)(a+b) \\
&= aaa + aab + aba + abb + baa + bab + bba + bbb
\end{aligned}
$$

כאן ישנם שמונה מחוברים, שכל אחד מתקבל על ידי בחירה של a־ים מחלק מהסוגריים ו־b־ים מהסוגריים הנותרים. למשל, aba מתקבל מבחירה של a בסוגריים הראשונים והאחרונים ו־b באמצעיים.

באופן כללי, $(a+b)^n$ הוא סכום של מחוברים שכל אחד מהם מתקבל מבחירה של i פעמים a מחלק מהסוגריים ו־$n-i$ פעמים b מהנותרים, וזאת לכל $0 \le i \le n$. האיבר $a^i b^{n-i}$ יכול להיבחר בדיוק $\binom{n}{i}$ פעמים ־ מספר האפשרויות לבחור את i זוגות הסוגריים שמתוכם נבחר a (או באופן שקול, $\binom{n}{n-i} = \binom{n}{i}$ אפשרויות לבחור את הסוגריים שמהם יילקחו ה־b־ים).

מכאן אנו מגיעים לנוסחה הכללית:

טענה 1.4 (הבינום של ניוטון) $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$.

בשל נוסחה זו המספרים $\binom{n}{i}$ מכונים לעתים קרובות **מקדמי הבינום**.

ההוכחה שלנו הייתה דוגמא ל**הוכחה קומבינטורית**: במקום להוכיח את הזהות האלגברית על ידי מניפולציות אלגבריות, נתנו הסבר קומבינטורי מניח את הדעת לנכונות הנוסחה. בהמשך, אחרי שנראה את המושג של אינדוקציה מתמטית, נוכל להוכיח את הבינום גם באופן אלגברי טהור.

4.2 משולש פסקל

יש למקדמי הבינום תיאור גרפי נאה הנקרא **משולש פסקל** (אף כי לא פסקל המציא אותו ־ המשולש היה מוכר כבר בימי הביניים, ופסקל בסך הכל תיאר אותו בספר מתמטיקה שכתב):

$$\begin{array}{ccccccccccc} & & & & & 1 & & & & & \\ & & & & 1 & & 1 & & & & \\ & & & 1 & & 2 & & 1 & & & \\ & & 1 & & 3 & & 3 & & 1 & & \\ & 1 & & 4 & & 6 & & 4 & & 1 & \\ 1 & & 5 & & 10 & & 10 & & 5 & & 1 \end{array}$$

בשורה ה־n־ית של המשולש נמצאים המספרים $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}$.
נשים לב למספר תכונות של המשולש ולאחר מכן נוכיח אותן:

1. המשולש סימטרי.
2. שפת המשולש מורכבת כולה מ־1־ים.
3. הכניסות שליד השפה בשורה ה־n הן n.
4. כל איבר במשולש הוא סכום שני האיברים שמעליו (ובמקרה של איברים בשפה, של האיבר היחיד שמעליו).
5. סכום השורה ה־n הוא 2^n (נובע בקלות מנוסחת הבינום, כאשר מציבים בה $a = b = 1$).
6. סכום המקומות הזוגיים בשורה ה־n במשולש הוא 2^{n-1} (ולכן גם סכום המקומות האי זוגיים הוא 2^{n-1}).

נוכיח כל תכונה בשתי דרכים ־ אלגברית (כלומר, על ידי מניפולציה של משוואות) וקומבינטורית (כלומר, על ידי תיאור בעיית ספירה מתאימה).

1. זוהי בעצם הטענה $\binom{n}{i} = \binom{n}{n-i}$.
הוכחה אלגברית: $\binom{n}{i} = \frac{n!}{i!(n-i)!} = \frac{n!}{(n-i)!i!} = \binom{n}{n-i}$.
הוכחה קומבינטורית: מספר הדרכים לבחור i איברים מתוך n הוא כמו מספר הדרכים לבחור אילו $n-i$ איברים מתוך n **לא** לקחת.
2. זוהי בעצם הטענה $\binom{n}{0} = \binom{n}{n} = 1$ (השוויון הראשון נובע מהסימטריה).
הוכחה אלגברית: $\binom{n}{0} = \frac{n!}{0!n!} = \frac{n!}{n!} = 1$
הוכחה קומבינטורית: יש רק דרך אחת לבחור 0 מ־n איברים ־ לא בוחרים אף אחד.
3. זוהי בעצם הטענה $\binom{n}{1} = \binom{n}{n-1} = n$ (שוב, השוויון הראשון נובע מהסימטריה).
הוכחה אלגברית: $\binom{n}{1} = \frac{n!}{1!(n-1)!} = \frac{n\cdot(n-1)!}{(n-1)!} = n$
הוכחה קומבינטורית: יש n דרכים לבחור איבר בודד מתוך n.
4. זוהי בעצם הטענה $\binom{n}{i} = \binom{n-1}{i-1} + \binom{n-1}{i}$ (שנכונה עבור $n, i > 0$).

הוכחה אלגברית:

$$\begin{aligned}\binom{n-1}{i-1}+\binom{n-1}{i} &= \frac{(n-1)!}{(i-1)!\,(n-i)!}+\frac{(n-1)!}{i!\,(n-i-1)!}\\ &= (n-1)!\left[\frac{i}{i!\,(n-i)!}+\frac{(n-i)}{i!\,(n-i)!}\right]\\ &= (n-1)!\cdot\frac{n}{i!\,(n-i)!}=\frac{n!}{i!\,(n-i)!}=\binom{n}{i}\end{aligned}$$

הוכחה קומבינטורית: מספר הדרכים לבחור i איברים מתוך n הוא כמספר הדרכים לבחור $i-1$ איברים מתוך $n-1$ הראשונים ולהוסיף אליהם את האחרון, ועוד מספר הדרכים לבחור i איברים מתוך $n-1$ הראשונים מבלי להוסיף להם את האחרון (נובע מעקרון החיבור). זוהי הוכחה פשוטה וקל בהרבה לזכור אותה מאת ההוכחה האלגברית.

5. זוהי בעצם הטענה $\sum_{i=0}^{n}\binom{n}{i}=2^n$.
הוכחה אלגברית: מהבינום של ניוטון עולה ש־ $\sum_{i=0}^{n}\binom{n}{i}=\sum_{i=0}^{n}\binom{n}{i}1^i1^{n-i}=(1+1)^n=2^n$.
הוכחה קומבינטורית: $\binom{n}{i}$ הוא מספר הוקטורים הבינאריים מאורך n עם בדיוק i אפסים.
2^n הוא מספר הוקטורים הבינאריים הכולל מאורך n, ועל פי עיקרון החיבור הוא שווה לסכום מספרם של כל הוקטורים הבינאריים עם בדיוק i אפסים לכל i.

6. זוהי בעצם הטענה $\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{2i}=2^{n-1}$.
הוכחה אלגברית: לכל $i>0$ ראינו ש־ $\binom{n}{2i}=\binom{n-1}{2i-1}+\binom{n-1}{2i}$, וכמו כן $\binom{n}{0}=\binom{n-1}{0}$ ולכן נקבל

$$\begin{aligned}\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{2i} &=\binom{n-1}{0}+\sum_{i=1}^{\lfloor\frac{n}{2}\rfloor}\left(\binom{n-1}{2i-1}+\binom{n-1}{2i}\right)=\\ &=\binom{n-1}{0}+\sum_{i=1}^{n-1}\binom{n-1}{i}=2^{n-1}\end{aligned}$$

הוכחה קומבינטורית: מספר הוקטורים הבינאריים מאורך n שבהם מספר זוגי של אפסים שווה למספר הוקטורים הבינאריים מאורך n שבהם מספר אי זוגי של אפסים: התאמה חח`ע ועל ביניהם מתקבל מהפיכת הביט הראשון בוקטור. יש סה`כ 2^n וקטורים ולכן בדיוק חצי מהם יהיו עם מספר זוגי של אפסים, כלומר 2^{n-1}.

4.3 המולטינום

ראינו את נוסחת הבינום של ניוטון:

$$(a+b)^n=\sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}$$

אולם מה קורה כאשר יש לנו לא זוג מחוברים, אלא מספר סופי כלשהו של מחוברים? במקרה זה הנוסחה המקבילה נקראת **מולטינום**.

כזכור, בחלק 2.5 ראינו כי מספר האפשרויות לסידור בשורה של n עצמים, כך שיש k_1 עצמים זהים מסוג אחד, k_2 עצמים זהים מסוג שני וכן הלאה עד k_r עצמים זהים מהסוג ה־r ו־$n = k_1 + \ldots + k_r$ הוא

$$\binom{n}{k_1, k_2, \ldots, k_t} = \frac{n!}{k_1!k_2!\cdots k_t!}$$

הדמיון של הסימון $\binom{n}{k_1,k_2,\ldots,k_t}$ לסימון של מקדמי הבינום כמובן אינו מקרי. במקרה שבו יש בדיוק שני סוגי איברים מקבלים בדיוק את נוסחת הבינום. מכאן שההכללה שנציג כעת לבינום אינה מפתיעה במיוחד:

טענה 2.4 יהיו $x_1, \ldots, x_r$ איברים כלשהם ו־n מספר טבעי, אז

$$(x_1 + x_2 + \ldots + x_r)^n = \sum_{k_1+k_2+\ldots+k_r=n} \binom{n}{k_1, k_2, \ldots, k_t} x_1^{k_1} x_2^{k_2} \cdots x_r^{k_r}$$

כאשר הסכום נלקח על פני כל ה־r־יות $(k_1, \ldots, k_r)$ של מספרים טבעיים עבורם $k_1 + k_2 + \ldots + k_r = n$.

הוכחה: ההוכחה היא קומבינטורית, בדומה לבינום של ניוטון. כאשר אנו פותחים את הביטוי $(x_1 + x_2 + \ldots + x_r)^n$ אנו מקבלים סכום של **מונומים** כאשר כל מונום הוא ביטוי מהצורה $x_1^{k_1} x_2^{k_2} \cdots x_r^{k_r}$ שמתקבל מבחירת איבר בודד מכל אחד מ־n זוגות הסוגריים, כך ש־k_i שווה למספר זוגות הסוגריים שמהם נבחר x_i. מכיוון שיש בסך הכל n זוגות סוגריים, בהכרח $k_1 + k_2 + \ldots + k_r = n$ לכל מונום (כי מספר הבחירות הכולל שביצענו שהוביל ליצירת המונום הוא בדיוק n כמספר זוגות הסוגריים).

מספר הפעמים הכולל שבו המונום $x_1^{k_1} x_2^{k_2} \cdots x_r^{k_r}$ מתקבל שווה למספר הדרכים לכתוב סדרה מאורך n בדיוק מהצורה $y_1 y_2 \ldots y_n$ כך ש־x_1 מופיע בסדרה בדיוק k_1 פעמים, x_2 מופיע בדיוק k_2 פעמים וכן הלאה ־ כלומר, סידור בשורה של n איברים שמתוכם k_1 מסוג אחד, k_2 מסוג שני וכן הלאה, ולכן מספר הפעמים שווה אל $\binom{n}{k_1,k_2,\ldots,k_t}$. ■

5 אינדוקציה ורקורסיה

5.1 אינדוקציה מתמטית

אינדוקציה מתמטית היא אחת מטכניקות ההוכחה הבסיסיות והשימושיות ביותר במתמטיקה. אינדוקציה פירושה הסקה מהפרט אל הכלל; במתמטיקה פירושה שדי להוכיח טענה עבור `מקרי בסיס` פשוטים ולהראות כיצד ניתן להסיק את נכונות הטענה עבור מקרה מורכב ממקרים פשוטים יותר, בכדי להוכיח שהטענה נכונה תמיד. ניתן לחשוב על אינדוקציה כעל דומינו: בסיס האינדוקציה מפיל את האבן הראשונה, וצעד האינדוקציה מראה כיצד כל אבן נופלת מפילה את הבאה אחריה. התוצאה היא שכל האבנים נופלות.

נציג מספר סוגים של אינדוקציה:

טענה 1.5 (אינדוקציה על הטבעיים במשתנה יחיד) אם $A_0, A_1, A_2, \ldots$ היא סדרה של טענות, כך שמתקיימים שני התנאים הבאים:

1. (בסיס האינדוקציה) A_0 נכונה.
2. (צעד האינדוקציה) אם A_i נכונה, אז גם A_{i+1} נכונה.

אז כל הטענות $A_0, A_1, A_2, \ldots$ נכונות.

הוכחה: נניח בשלילה כי 1 ו־2 נכונים אך לא כל הטענות $A_0, A_1, A_2, \ldots$ נכונות, ויהא n הטבעי הקטן ביותר כך ש־A_n אינו נכון. בשל תנאי 1 לא ייתכן ש־$n = 0$, ולכן A_{n-1} היא טענה מתוך הסדרה $A_0, A_1, A_2, \ldots$ ומכיוון ש־n היה מינימלי, A_{n-1} כן נכונה ומתנאי 2 עולה שגם A_n נכונה, בסתירה להנחת השלילה. ■

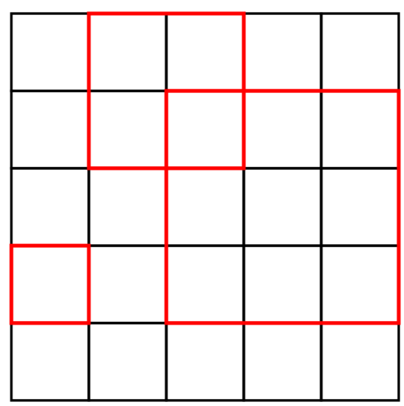
הוכחה זו מסתמכת על כך שלכל תת קבוצה של טבעיים יש איבר מינימלי; תכונה זו מכונה `סדר טוב`, ואינדוקציה ניתנת להגדרה בכל קבוצה שיש בה סדר טוב, אך לא נפרט על כך בקורס.

דוגמא כשהוכחנו שמספר התמורות על n איברים הוא $n!$ השתמשנו באופן מובלע באינדוקציה. נוכיח זאת עכשיו באופן מפורש.

בסיס: מספר האפשרויות לסדר 0 איברים בשורה הוא 1 ("הסידור הריק").

צעד: נניח שמספר האפשרויות לסדר n איברים בשורה הוא $n!$. בהינתן $n+1$ איברים, נסדר את n הראשונים בשורה ואז יש לנו $n+1$ מקומות שונים לשים בהם את האיבר הנוסף (בתחילת השורה, או אחרי כל אחד מ־n האיברים האחרים). לכן מעקרון הכפל, מספר האפשרויות הכולל הוא $n! \cdot (n+1) = (n+1)!$.

דוגמא בלוח $n \times n$ מספר תת־הריבועים הכולל הוא $S_n = \frac{n(n+1)(2n+1)}{6}$ (מספר זה שווה ל־$1^2 + 2^2 + \ldots + n^2$ אבל זו דרך משעממת יותר לנסח את התוצאה).

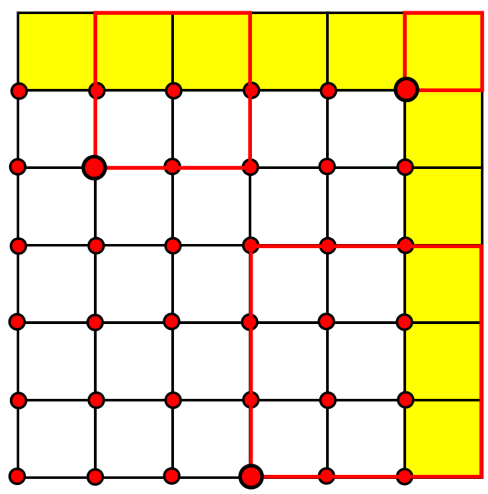


נוכיח זאת באינדוקציה על n:

בסיס: עבור $n = 1$ קיים ריבוע יחיד בלוח: הריבוע 1×1. ואכן, $S_1 = \frac{1(1+1)(2+1)}{6} = \frac{6}{6} = 1$.

צעד: נניח את נכונות הנוסחה עבור n. ניקח לוח $n \times n$ ונרחיב אותו ללוח $(n+1) \times (n+1)$ (נדמיין שהוספנו שורה חדשה למעלה ועמודה חדשה מימין). כל ריבוע בלוח החדש נופל לאחת משתי קטגוריות:

- הריבוע מוכל כולו בלוח ה־$n \times n$ הישן: יש בדיוק S_n ריבועים כאלו.
- הריבוע גולש לשורה/עמודה החדשה: במקרה זה, קיימת התאמה חח"ע ועל בין קבוצת הקודקודים של הריבועים בלוח הקיים והריבועים החדשים (בהינתן קודקוד, יש דרך יחידה להרחיב את הריבוע שאותו קודקוד הוא הפינה השמאלית־תחתונה שלו כך שיגיע אל השורה/עמודה החדשות). בלוח של n משבצות יש $(n+1)\times(n+1)=(n+1)^2$ קודקודים כאלו.



קיבלנו ש־$S_{n+1}=S_n+(n+1)^2$ ומכאן נמשיך עם אלגברה והנחת האינדוקציה:

$$\begin{aligned}
S_n+(n+1)^2 &= \frac{n(n+1)(2n+1)}{6}+(n+1)^2\\
&= \frac{(n+1)\left(2n^2+n+6n+6\right)}{2}\\
&= \frac{(n+1)\left(2n^2+7n+6\right)}{2}\\
&= \frac{(n+1)(n+2)(2n+3)}{2}
\end{aligned}$$

וזו בדיוק התוצאה המבוקשת.

דוגמא נוכיח את נוסחת הבינום של ניוטון, $(a+b)^n=\sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}$, בצורה אלגברית באמצעות אינדוקציה על n.

עבור $n=0$ מתקיים מצד אחד $(a+b)^n=(a+b)^0=1$ (כי $x^0=1$ לכל $x\in\mathbb{R}$, כולל $x=0$, עם ההגדרות בקורס שלנו). מצד שני $\sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}=\binom{0}{0}a^0b^{0-0}=1$, כך שקיבלנו את השוויון המבוקש במקרה הזה.

נניח שהטענה מתקיימת עבור $n-1$ ונוכיח עבור n:

$$
\begin{aligned}
(a+b)^n &= (a+b)\,(a+b)^{n-1} \\
&= (a+b)\sum_{i=0}^{n-1}\binom{n-1}{i}a^ib^{n-1-i} \\
&= \sum_{i=0}^{n-1}\binom{n-1}{i}a^{i+1}b^{n-1-i}+\sum_{i=0}^{n-1}\binom{n-1}{i}a^ib^{n-i} \\
(1) &= \sum_{i=1}^{n}\binom{n-1}{i-1}a^ib^{n-i}+\sum_{i=0}^{n-1}\binom{n-1}{i}a^ib^{n-i} \\
(2) &= \sum_{i=0}^{n}\binom{n-1}{i-1}a^ib^{n-i}+\sum_{i=0}^{n}\binom{n-1}{i}a^ib^{n-i} \\
&= \sum_{i=0}^{n}\left[\binom{n-1}{i-1}+\binom{n-1}{i}\right]a^ib^{n-i} \\
(3) &= \sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}
\end{aligned}
$$

כאשר מעבר (1) נובע מהחלפת משתנה הסכימה (ניתן לחשוב על כך כאילו i מוחלף ב־$i-1$), מעבר (2) נובע מכך ש־$\binom{n-1}{-1}=\binom{n-1}{n}=0$ ומעבר 3 נובע מהנוסחה לסכום מקדמי הבינום שראינו קודם.

דוגמא כשל נפוץ בהוכחות באינדוקציה מתואר על ידי ה`הוכחה` הבאה שכל הסוסים בעלי אותו הצבע. פורמלית, שבכל קבוצה של סוסים, כל הסוסים בעלי אותו צבע. האינדוקציה היא על גודל הקבוצה ומתחילה מ־1.

1. (בסיס) בקבוצה של סוס בודד כל הסוסים באותו הצבע שכן קיים בה רק סוס בודד.
2. (צעד) בהינתן קבוצה בעלת $n+1$ סוסים נוציא את אחד הסוסים החוצה וניוותר עם n סוסים שכולם באותו הצבע. כעת נחזיר את הסוס לקבוצה ונוציא סוס אחר ושוב נקבל קבוצה שבה כל הסוסים בעלי אותו הצבע, ולכן הסוס שהוצאנו בהתחלה הוא בעל אותו צבע כמו היתר.

הרמאות ב`הוכחה` הזו היא בכך שצעד האינדוקציה אינו עובד כאשר $n=1$ (יש לשים לב כי עבור $n>1$ הוא עובד, אך זה חסר משמעות).

טענה 2.5 (אינדוקציה שלמה על הטבעיים במשתנה יחיד) אם $A_0,A_1,A_2,\ldots$ היא סדרה של טענות, כך שמתקיים התנאי הבא:

1. אם A_k נכונה לכל $k<n$ אז A_n נכונה.

אז כל הטענות $A_0,A_1,A_2,\ldots$ נכונות.

אינדוקציה שלמה נבדלת מאינדוקציה `רגילה` בכך שקל יותר להוכיח את צעד האינדוקציה מכיוון שניתן בהוכחת A_n להיעזר בנכונות כל הטענות $A_0,\ldots,A_{n-1}$ ולא רק ב־A_{n-1}

עצמה; עם זאת, לרוב אין בה צורך.

הוכחה: ההוכחה זהה לזו של אינדוקציה רגילה: נניח בשלילה כי תנאי 1 נכון אך לא כל הטענות $A_0, A_1, A_2, \ldots$ נכונות, ויהא n הטבעי הקטן ביותר כך ש־A_n אינו נכון. אז ממינימליות n עולה שלכל $k < n$, הטענה A_k כן מתקיימת ולכן מתנאי 1 גם A_n עצמה מתקיימת, בסתירה להנחת השלילה. ■

הניסוח של אינדוקציה שלמה עשוי להיראות "פשוט מדי" כי את שני התנאים של אינדוקציה רגילה החליף תנאי יחיד, ובסיס האינדוקציה לכאורה נעלם. בפועל, עבור $n = 0$ תנאי (1) אומר "אם A_k נכונה לכל $k < 0$ אז A_0 נכונה" אלא שאין $k < 0$ ולכן כל חלק ה"אם" של הטענה, שבא לתת לנו הנחות שנוכל להיעזר בהן במהלך ההוכחה, אינו בא לידי ביטוי. כלומר, הוכחה של טענה 1 בפרט גוררת הוכחה של A_0 ללא הנחות נוספות ־ כמו מקרה הבסיס של אינדוקציה רגילה.

דוגמא נוכיח שלכל מספר טבעי חיובי קיים פירוק למכפלה של מספרים ראשוניים:

בסיס: עבור $n = 1$ "המכפלה הריקה" היא הפירוק שאנחנו מחפשים (אפשר גם להתחיל את האינדוקציה מ־$n = 2$ למי שקונספט המכפלה הריקה מפריע לו בשלב זה).

צעד: נניח שלכל מספר טבעי קטן מ־n קיים פירוק למכפלה של מספרים ראשוניים. נתבונן ב־n. אם הוא עצמו ראשוני, אז n היא המכפלה המבוקשת; אחרת, $n = a \cdot b$ כך ש־$a, b < n$ שניהם. עבור כל אחד מהם קיים פירוק למכפלה של ראשוניים, כך שמכפלת שתי המכפלות הללו היא הפירוק המבוקש של n.

בדוגמה הזו האינדוקציה השלמה היא הכרחית, שכן אנחנו לא יכולים להפיק פירוק של n מתוך פירוק של $n - 1$ ־ אנחנו נאלצים ללכת יותר אחורה באינדוקציה.

טענה 3.5 (אינדוקציה דו ממדית) אם $A_{n,m}$ היא קבוצה של טענות ($n, m \geq 0$ טבעיים) כך שמתקיים התנאי הבא:

1. לכל $m, n \in \mathbb{N}$ אם $A_{i,j}$ נכונה לכל $0 \leq i \leq n$ ו־$0 \leq j \leq m$ כך ש־$(i, j) \neq (n, m)$, אז גם $A_{n,m}$ נכונה.

אז כל הטענות $A_{n,m}$ נכונות.

הוכחה: ראשית נזכיר את האופן שבו ניתן להגדיר סדר על $\mathbb{N}^2$: **סדר לקסיקוגרפי** שבו $(a_1, b_1) \leq (a_2, b_2)$ אם ורק אם $a_1 < a_2$ או $a_1 = a_2$ וגם $b_1 \leq b_2$. ניתן לראות כי זהו אכן יחס סדר וזהו סדר טוב, במובן זה שלכל תת־קבוצה לא ריקה של $\mathbb{N}^2$ קיים איבר קטן ביותר יחיד. כדי למצוא איבר שכזה בתת־קבוצה, ראשית נסתכל על קואורדינטות ה־a של הזוגות שבקבוצה וניקח את ה־a המינימלי מביניהן; ואחר כך נתבונן על כל הזוגות מהצורה $(a, \ldots)$ וניקח את ה־b המינימלי מבין קואורדינטות ה־b של איברים אלו.

כעת, נניח שתנאי 1 מתקיים אבל לא כל הטענות $A_{n,m}$ נכונות. תהא $A_{n,m}$ הטענה המינימלית בסדר לקסיקוגרפי שאינה נכונה. משמעות הסדר הלקסיקוגרפי היא שכל טענה מהצורה $A_{i,j}$ כך ש־$i < n$ היא בוודאי נכונה כי $(i, j) < (n, m)$, וכמו כן אם $i = j$ אז מהגדרת סדר לקסיקוגרפי $j \leq m$. לכן לכל $0 \leq i \leq n$ ו־$0 \leq j \leq m$ כך ש־$(i, j) \neq (n, m)$ הטענה $A_{i,j}$ נכונה ומתנאי 1 נסיק שגם $A_{n,m}$ נכונה, בסתירה להנחת השלילה. ■

דוגמא נוכיח באינדוקציה שמספר האפשרויות לבחור m מתוך n איברים עם חשיבות לסדר וללא חזרות הוא $P_n^m = \frac{n!}{(n-m)!}$.

צעד: ראשית נטפל במקרי קצה אפשריים. לכל $n \ge 0$ מתקיים $P_n^0 = 1$ כי יש דרך יחידה לבחור 0 מתוך n איברים (זוהי "הבחירה הריקה"). כמו כן לכל $m \ge 1$ מתקיים $P_0^m = 0$ כי אין דרך לבחור $m \ge 1$ מתוך 0 איברים (במקרה הזה לא מדובר על "הבחירה הריקה" כי אנחנו צריכים לקבל קבוצה בת m איברים שכל אברי שייכים לקבוצה הריקה, אבל פשוט לא קיימת קבוצה כזו).

עכשיו נוכל להניח בהמשך כי $n, m \ge 1$ ולכן תהיה לביטויים P_{n-1}^m, P_{n-1}^{m-1} משמעות.

כדי לבחור m מתוך n איברים עם חשיבות לסדר וללא חזרות כאשר $n, m \ge 1$ אפשר לפרק לשני מקרים: ראשית, מספר האפשרויות לבחור m מתוך $n-1$ האיברים הראשונים הוא, על פי הנחת האינדוקציה, $P_{n-1}^m = \frac{(n-1)!}{(n-m-1)!}$. שנית, אם בוחרים את האיבר ה־n־י זה יכול להיות בכל אחד מ־m שלבי הבחירה, כך שיש לנו תהליך של בחירה דו־שלבית: ראשית בוחרים באיזה שלב ייבחר האיבר ה־n־י (m אפשרויות) ושנית בוחרים את $m-1$ האיברים הנותרים מבין $n-1$ האיברים הנותרים לבחירה. מספר האפשרויות הכולל במקרה זה הוא $mP_{n-1}^{m-1} = \frac{m(n-1)!}{(n-1-(m-1))!} = \frac{m(n-1)!}{(n-m)!}$. נותר לנו כעת לחבר את שני המקרים השונים:

$$\begin{aligned}\frac{(n-1)!}{(n-m-1)!} + \frac{m\,(n-1)!}{(n-m)!} &= (n-1)!\left[\frac{n-m}{(n-m)!} + \frac{m}{(n-m)!}\right] \\ &= (n-1)! \cdot \frac{n}{(n-m)!} = \frac{n!}{(n-m)!}\end{aligned}$$

כמבוקש.

נשים לב שבהוכחה האינדוקטיבית הזו נזקקנו לערכים P_{n-1}^m ו־P_{n-1}^{m-1}. כלומר, הניסוח של האינדוקציה הדו־ממדית בתור מעין "אינדוקציה שלמה" ולא רק טענה מהצורה "אם $P_{n,m}$ נכונה אז גם $P_{n+1,m}$ ו־$P_{n,m+1}$ נכונות" היה הכרחי כבר עבור הדוגמא הפשוטה הזו.

5.2 רקורסיה

הגדרה רקורסיבית היא הגדרה של סדרה או פונקציה (או אובייקטים כלליים יותר) שבה כל ערך מוגדר באמצעות הערכים של קודמיו, פרט אולי למספר ערכים התחלתיים שמוגדרים במפורש.

נוסחה רקורסיבית היא לעתים קרובות פשוטה יותר מנוסחה סגורה עבור אותה סדרה (ולעתים ניתן למצוא נוסחה רקורסיבית אף שאין נוסחה סגורה פשוטה) אך היא יותר קשה לחישוב בפועל (שכן כדי לחשב ערך כלשהו באמצעותה יש לחשב קודם כל את הערכים שהוא מסתמך עליהם).

נראה מספר דוגמאות לנוסחאות רקורסיביות המגדירות סדרה, ובנוסף גם הנוסחאות הסגורות המתאימות. בפרט, כל נוסחאות הספירה שתיארנו בפרק 2 ניתנות לתיאור כנוסחאות רקורסיביות (ואת חלקן גם תיארנו כך מבלי לציין במפורש כי זה מה שאנו עושים).

- סדרה חשבונית: $a_n = a_{n-1} + d$ (הנוסחה הסגורה: $a_n = a_1 + (n-1)\,d$).
- סדרה הנדסית: $a_n = a_{n-1} \cdot q$ (הנוסחה הסגורה: $a_n = a_1 \cdot q^{n-1}$).
- סדרת פיבונאצ'י: $a_n = a_{n-1} + a_{n-2}$, עם תנאי התחלה $a_0 = 0, a_1 = 1$ (בהמשך נראה כיצד מוצאים את הנוסחה הסגורה, $a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$).

- סידורים בשורה: $P_n = n \cdot P_{n-1}$ (נוסחה סגורה: $P_n = n!$, כפי שכבר ראינו).
- בחירה בלי חזרות ועם חשיבות לסדר: $P_n^k = P_{n-1}^k + k \cdot P_{n-1}^{k-1}$ עם תנאי ההתחלה $P_n^0 = 1$ לכל n (נוסחה סגורה: $P_n^k = \frac{n!}{(n-k)!}$).
- בחירה עם חזרות ועם חשיבות לסדר: $PP_n^k = n \cdot PP_n^{k-1}$ עם תנאי ההתחלה $PP_n^0 = 1$ (נוסחה סגורה: $PP_n^k = n^k$).
- בחירה בלי חזרות ובלי חשיבות לסדר: $C_n^k = C_{n-1}^{k-1} + C_{n-1}^k$ עם תנאי ההתחלה $C_n^0 = 1$ ו־$C_n^n = 1$ (נוסחה סגורה: $C_n^k = \binom{n}{k}$).
- בחירה עם חזרות ובלי חשיבות לסדר: $CC_n^k = CC_n^{k-1} + CC_{n-1}^k$ עם תנאי ההתחלה $CC_n^0 = 1$ עבור $n \geq 1$ ו־$CC_0^k = 0$ עבור $k > 0$ (נוסחה סגורה: $CC_n^k = \binom{n+k-1}{k}$).

את המקרה האחרון של CC_n^k קל לקבל בעזרת הנוסחא $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ שכבר ראינו:

$$\begin{aligned}
CC_n^k =& \binom{n+k-1}{k} = \binom{n-1+k-1}{k-1} + \binom{n-1+k-1}{k} = \\
=& \binom{n+(k-1)-1}{k-1} + \binom{(n-1)+k-1}{k} = CC_n^{k-1} + CC_{n-1}^k
\end{aligned}$$

נציג כעת דוגמא מעט יותר מורכבת:

דוגמא **הפרת סדר** על n איברים היא תמורה על המספרים $1, 2, \ldots, n$ שבה לכל $1 \leq i \leq n$, המספר ה־i אינו נמצא במקום ה־i. למשל, 312 היא הפרת סדר על 3 איברים ואילו 321 לא (כי 2 נמצא במקום 2).

נסמן ב־D_n את מספר הפרות הסדר על n איברים.

ניתן לחשב את D_n כך: עבור 1, יש לנו $(n-1)$ בחירות של מקום עבורו (כי את מקום 1 לא ניתן לבחור בשבילו). לאחר מכן אנו נותרים עם $n-1$ מספרים שיש לסדר. נאמר ששמנו את 1 במקום i, אז יש שתי אפשרויות: או ש־i יושב במקום 1, או שלא. אם הוא מושם במקום 1, אז אפשר לשכוח הן מ־1 והן מ־i ולטפל ב־$n-2$ המספרים הנותרים באופן בלתי תלוי, כלומר יש D_{n-2} הפרות סדר במקרה זה; ואילו אם i אינו מושם במקום מס' 1, אז אפשר לחשוב על i כאילו הוא עצמו המספר 1 ואסור לו להיות במקום 1, וזהו התנאי הרגיל של הפרות סדר, ולכן יש לנו D_{n-1} הפרות סדר במקרה זה.

קיבלנו את הנוסחה הרקורסיבית $D_n = (n-1)\left[D_{n-1} + D_{n-2}\right]$. כדי להשתמש בנוסחה אנחנו זקוקים לשני ערכים התחלתיים, $D_0 = 1$ ו־$D_1 = 0$ (כי הסדרה הריקה מקיימת באופן ריק את התנאי של הפרת סדר ולכן יש הפרת סדר אחת על 0 איברים, אבל הסדרה שכוללת את האיבר הבודד 1, שהיא הסדרה היחידה מאורך 1, אינה מקיימת את התנאי של הפרת סדר).

בהערת אגב נציין שלא קיימת נוסחה סגורה קומבינטורית, אך בהמשך נראה גישה נוספת לטיפול בבעיה זו שממנה נקבל ש־$D_n = \left[\frac{n!}{e}\right]$ (הסוגריים המרובעים מציינים את פונקצית הערך השלם ־ המספר הטבעי הקרוב ביותר ל־$\frac{n!}{e}$).

6 כלל ההכלה וההפרדה

נתונות שתי קבוצות A, B ואנו מעוניינים לדעת מהו $|A \cup B|$. אם הקבוצות זרות (ללא איברים משותפים) אז $|A \cup B| = |A| + |B|$ - זהו עקרון החיבור. אך מה קורה אם הקבוצות אינן זרות, כלומר הקבוצה $A \cap B$ של האיברים המשותפים לשתיהן אינה ריקה?

במקרה זה הבעיה ב־$|A| + |B|$ הוא שאיברים משותפים ל־A, B נספרים **פעמיים**; פעם כאיברי A ופעם כאיברי B. את הטעות הזו ניתן `לתקן` על ידי כך שמחסרים מהסכום הכולל את מספר האיברים שנספרו פעמיים, כלומר נקבל את הנוסחה:

$$|A \cup B| \;=\; |A| + |B| - |A \cap B|$$

נוסחה זו נכונה לכל זוג קבוצות A, B.

נעבור כעת למקרה של שלוש קבוצות: $|A \cup B \cup C|$. ניתן היה לקוות שגודל הקבוצה יהיה $|A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C|$ אך זה אינו נכון ומעיד על כך שלא ניתן להסתפק בבחינת **זוגות** של קבוצות בלבד. הבעיה היא שאיבר שנמצא בכל שלוש הקבוצות ייספר בחיוב שלוש פעמים (עם $|A|, |B|, |C|$) אבל גם לשלילה שלוש פעמים (עם $|A \cap B|, |A \cap C|, |B \cap C|$) ולכן בסך הכל יוסיף 0 לספירה הכוללת (בזמן שהוא אמור להוסיף 1). לכן כדי לתקן אנו מוסיפים עוד פעם אחת את האיברים שבכל שלוש הקבוצות, כלומר מחברים לסכום את $|A \cap B \cap C|$ ומקבלים את הנוסחה הנכונה:

$$|A \cup B \cup C| \;=\; |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

מכאן אנו מגיעים באופן טבעי למקרה הכללי:

משפט 1.6 (כלל ההכלה וההפרדה) אם $A_1, \ldots, A_n$ הן קבוצות אז

$$\begin{aligned} \left|\bigcup_{i=1}^{n} A_i\right| \;&=\; \sum_{1 \le i \le n} |A_i| - \sum_{1 \le i,j \le n} |A_i \cap A_j| + \\ &+ \sum_{1 \le i,j,k \le n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n-1} |A_1 \cap \cdots \cap A_n| \end{aligned}$$

הוכחה: יש להראות שכל איבר של $\bigcup_{i=1}^{n} A_i$ נספר בדיוק פעם אחת באגף ימין, אחרי שמקזזים ספירות חיוביות ושליליות.

נניח שהאיבר מופיע בדיוק ב־t מתוך n הקבוצות. מספר הפעמים שבהן הוא נספר באותו איבר בסכום שבו רצים על חיתוכים של i קבוצות תלוי ב־i; אם $i > t$ אז האיבר אינו נספר כלל כי בחיתוך של i קבוצות בהכרח משתתפת בחיתוך קבוצה שאינה מכילה אותו. לעומת זאת, אם $i \le t$ אז הוא מופיע בדיוק ב־$\binom{t}{i}$ מהחיתוכים - אלו שבהם משתתפות רק קבוצות שמכילות אותו.

על כן, הספירה עבור אותו איבר היא $\sum_{i=1}^{t} (-1)^{i-1} \binom{t}{i}$, כעת, מהבינום של ניוטון:

$$\sum_{i=1}^{t}(-1)^{i-1}\binom{t}{i} = 1-\sum_{i=0}^{t}(-1)^{i}\binom{t}{i} = 1-(1-1)^{t}=1$$

כנדרש. ∎

לרוב השימוש שלנו לעקרון ההכלה וההפרדה הוא כזה: נתון `עולם` בן n איברים, ומספר קבוצות $A_1,\ldots,A_k$ שאבריהן נלקחים מתוך העולם ואנו חושבים עליהן כעל `תכונות רעות` שהאיברים יכולים לקיים. מטרתנו היא למצוא את כמות האיברים ש**אינם מקיימים** אף תכונה רעה, כלומר את $\left|\bigcap_{i=1}^{k}\overline{A_i}\right|$. מעקרון ההכלה וההפרדה נקבל:

$$\left|\bigcap_{i=1}^{k}\overline{A_i}\right| = n-\left|\bigcup_{i=1}^{k}A_i\right| = n-\sum|A_i|+\cdots+(-1)^{n}|A_1\cap\cdots\cap A_n|$$

ניסוח נוסף שהוא נוח מעט יותר לעבודה הוא זה: אם ישנם n איברים ו־k תכונות $P_1,\ldots,P_k$, נסמן ב־$w(P_i)$ את מספר האיברים שמקיימים את P_i, ב־$w(P_iP_j)$ את מספר האיברים שמקיימים גם את P_i וגם את P_j, וכן הלאה, ולכל מספר טבעי r נשתמש בסימון $w(r)=\sum_{1\le i_1,\ldots,i_r\le k}w(P_{i_1}\cdots P_{i_r})$ (לסימון זה אין משמעות קומבינטורית; אותו איבר יכול להיספר בו כמה וכמה פעמים). נקבל את הניסוח הבא של כלל ההכלה וההפרדה:

משפט 2.6 (כלל ההכלה וההפרדה, ניסוח באמצעות תכונות) יהא $E(0)$ מספר האיברים שאינם מקיימים אף תכונה , אז:

$$E(0) = \sum_{r=0}^{k}(-1)^{r}w(r)$$

דוגמא מבין המספרים $1,2,\ldots,300$, כמה **אינם** מתחלקים ב־3, 7 או 11?

כאן `תכונה רעה` היא התחלקות ב־3, 7 או 11 ־ כלומר, יש לנו שלוש תכונות, שנסמן P_3,P_7,P_{11}. יש 300 מספרים ולכן $w(0)=300$.

קל לראות כי $w(P_3)=\left\lfloor\frac{300}{3}\right\rfloor=100$, $w(P_7)=\left\lfloor\frac{300}{7}\right\rfloor=42$ ו־$w(P_{11})=\left\lfloor\frac{300}{11}\right\rfloor=27$. לכן $w(1)=27+42+100=169$

כמו כן מכיוון ש־$3,7,11$ כולם ראשוניים, מספר מתחלק בכמה מהם רק אם הוא מתחלק במכפלה שלהם. כלומר:

$w(P_3P_7)=\left\lfloor\frac{300}{21}\right\rfloor=14$, $w(P_3P_{11})=\left\lfloor\frac{300}{33}\right\rfloor=9$, $w(P_7P_{11})=\left\lfloor\frac{300}{77}\right\rfloor=3$. לכן $w(2)=3+9+14=26$

ולסיום $w(P_3P_7P_{11})=\left\lfloor\frac{300}{231}\right\rfloor=1$ ולכן $w(3)=1$.

מנוסחת ההכלה וההפרדה נקבל כי כמות המספרים שאינם מתחלקים ב־3,7,11 היא בדיוק

$$\begin{aligned}E(0) &= w(0)-w(1)+w(2)-w(3)\\ &= 300-169+26-1\\ &= 156\end{aligned}$$

הפתרון עשוי להיראות טרחני ומסובך, אך בפועל הכלה והפרדה משפרת משמעותית את הסיבוכיות של אלגוריתם שמבצע אותה אוטומטית; פתרון נאיבי לבעיה שלעיל דורש לעבור על כל 300 המספרים ולבדוק לכל אחד מהם התחלקות; הפתרון עם הכלה והפרדה דורש חישוב 7 פעולות חילוק בלבד ועוד ביצוע של מספר פעולות סיכום שעלותן זניחה. באופן כללי אם הטווח שלנו הוא עד n ואנו בודקים התחלקות ב־k ראשוניים אז פתרון נאיבי דורש $O\left(n\cdot k\right)$ פעולות חילוק, ופתרון עם הכלה והפרדה דורש $O\left(2^{k}\right)$ פעולות כאלו (כל פעולה מתבצעת בזמן שהוא פולינומי ב־$\log n$), כך שעבור k קטן (ובפרט קבוע) ו־n גדול מדובר על פתרון יעיל משמעותית.

דוגמא יהא D_n מספר הפרות הסדר על n איברים: פרמוטציות של $1,\dots,n$ שבהן לכל i המספר i אינו נמצא במקום ה־i. ראינו כבר כיצד למצוא נוסחת נסיגה עבור D_n; ניעזר כעת בעיקרון ההכלה וההפרדה ובקצת חשבון דיפרנציאלי ואינטגרלי כדי למצוא נוסחה סגורה.

התכונה P_i תהיה התכונה `המספר $i$ נמצא במקום ה־$i$`.

הסימטריה של הבעיה מקלה מאוד על חישוב $w\left(r\right)$ במקרה זה. לכל r, ראשית נבחר r מתוך n מקומות שאנחנו רוצים `לקלקל` ($\binom{n}{r}$ אפשרויות), ולאחר מכן נספור את מספר התמורות שבהן המקומות שבחרנו `מקולקלים`. ייתכן שעוד מקומות יהיו מקולקלים אך זה לא משנה עבורנו (זה בדיוק הכוח שבהכלה והפרדה). כדי לקלקל r מקומות אנחנו מציבים בכל אחד מהם את המספר שמקלקל אותו, ואז נותרים לנו $n-r$ מקומות שבהם אפשר לסדר באופן חופשי את המספרים, כלומר יש לנו $\left(n-r\right)!$ אפשרויות.

בסך הכל קיבלנו כי $w\left(r\right)=\binom{n}{r}\left(n-r\right)!=\frac{n!}{r!}$, ומכאן נקבל:

$$D_n=\sum_{r=0}^{n}\left(-1\right)^{r}w\left(r\right)=\sum_{r=0}^{n}\left(-1\right)^{r}\frac{n!}{r!}=n!\sum_{r=0}^{n}\frac{\left(-1\right)^{r}}{r!}$$

כאן מסתיימת הקומבינטוריקה אך בעזרת חשבון דיפרנציאלי ואינטגרלי ניתן להתקדם עוד קצת:

ידוע ש־$e^{x}=\sum_{r=0}^{\infty}\frac{x^{r}}{r!}$, ולכן $e^{-1}=\sum_{r=0}^{\infty}\frac{\left(-1\right)^{r}}{r!}$. מכאן ש־$\sum_{r=0}^{n}\frac{\left(-1\right)^{r}}{r!}$ הוא קירוב של e^{-1}, וגודל הטעות הוא זניח. מכאן ש־$D_n\approx\frac{n!}{e}$, ובפועל ניתן לראות ש־$D_n=\left[\frac{n!}{e}\right]$ (D_n הוא המספר הטבעי הקרוב ביותר ל־$\frac{n!}{e}$) לכל n. מכאן אנו רואים שכלל ההכלה וההפרדה סייע לנו למצוא נוסחה מדוייקת עבור D_n אף אם הוא לא נתן אותה במפורש בעצמו.

7 חלוקות

נחזור כעת לבעיה שכבר עסקנו בה: בכמה דרכים ניתן לחלק n כדורים ל־k תאים, בהינתן אילוצים מסויימים? נראה את הפתרון עבור אילוצים שונים רבים, בתור מעין סיכום ביניים למה שראינו עד כה.

נתחיל בוריאציות האפשריות על שלוש התכונות הבאות:

- האם הכדורים זהים או שונים.
- האם התאים זהים או שונים.
- האם מותר מספר כלשהו של כדורים בכל תא, או לכל היותר כדור אחד בכל תא.

יש 3 תכונות שונות עם 2 אפשרויות לכל תכונה, כך שיהיו לנו $2^{3}=8$ מקרים שונים.

1. n כדורים זהים, k תאים שונים ולכל היותר כדור אחד בכל תא: כאן יש לבחור את n מתוך k התאים שבהם יושמו כדורים.

בגלל שהכדורים זהים אין חשיבות לסדר הבחירה.
בגלל שיש כדור אחד לכל היותר בכל תא אין חזרות בבחירה.
מסקנה: $\binom{k}{n}$ אפשרויות.

2. n כדורים שונים, k תאים שונים ולכל היותר כדור אחד בכל תא: גם כאן בוחרים את n מתוך k התאים שבהם יושמו כדורים.
בגלל שהכדורים שונים יש חשיבות לסדר הבחירה.
בגלל שיש כדור אחד לכל היותר בכל תא אין חזרות בבחירה.
מסקנה: $P_k^n = \binom{k}{n}n!$ אפשרויות.

3. n כדורים שונים, k תאים שונים, מספר כלשהו של כדורים בכל תא: כאן לכל כדור בוחרים אחד מ־n התאים האפשריים.
בגלל שהכדורים שונים יש חשיבות לסדר הבחירה.
בגלל שאין מגבלות על מספר הכדורים בכל תא זו בחירה עם חזרות.
מסקנה: k^n אפשרויות.

4. n כדורים זהים, k תאים שונים, מספר כלשהו של כדורים בכל תא: כאן לכל כדור בוחרים אחד מ־k התאים האפשריים.
בגלל שהכדורים זהים אין חשיבות לסדר הבחירה.
בגלל שאין מגבלות על מספר הכדורים בכל תא זו בחירה עם חזרות.
מסקנה: $CC_k^n = \binom{n+k-1}{n}$ אפשרויות.

5. n כדורים זהים, r תאים זהים, לכל היותר כדור אחד בכל תא: אם $k \ge n$ אז ניתן לחלק את n הכדורים לתאים, ולאחר מכן נקבל n תאים עם כדור אחד ו־$k-n$ תאים בלי אף כדור. מכיוון שכל התאים זהים, זו התוצאה האפשרית היחידה.
אם $k < n$ אז מעיקרון שובך היונים, אין דרך לחלק את הכדורים לתאים בלי שתא כלשהו יכיל שני כדורים, ולכן במקרה זה יש 0 תוצאות. ניתן לסמן זאת $[k \ge n] = \begin{cases} 1 & k \ge n \\ 0 & k < n \end{cases}$.

6. n כדורים שונים, k תאים זהים, לכל היותר כדור אחד בכל תא: המקרה הזה זהה לקודמו. אם $k \ge n$ נוכל לחלק את הכדורים לתאים, אבל לא ניתן לייצר חלוקות שונות זו מזו מכיוון שהתאים זהים; כל מה שנראה הוא שיש תא שיש בו רק את כדור 1, תא אחר עם כדור 2 וכדומה. לכן גם במקרה זה התוצאה היא $[k \ge n]$.

נותרו שני מקרים שבהם התאים **זהים**, ניתן לשים מספר כלשהו של כדורים בכל תא, והכדורים יכולים להיות זהים או שונים. המקרים הללו יהיו מאתגרים יותר ממה שראינו עד כה ויצריכו מושגים חדשים. לפני שנפתור אותם, נעבור לטפל בבעיה נוספת של חלוקת כדורים לתאים: כאשר התאים יכולים להיות שונים או זהים, הכדורים שונים או זהים, ואין תא ריק. כאן יהיו לנו ארבעה מקרים שונים לטפל בהם.

1. n כדורים זהים, k תאים שונים, אין תא ריק.
זה המקרה הקל יחסית. שכבר ראינו קודם בקורס. אם $n < k$ יש 0 אפשרויות (כי אין דרך למלא את k התאים). אם $n \ge k$ נשים כדור אחד בכל תא. מכיוון שכל הכדורים זהים יש רק דרך אחת לעשות זאת, ולאחר מכן נותרנו עם בעיית חלוקה של $n-k$ כדורים זהים ל־k תאים שונים, בלי מגבלות נוספות על מספר הכדורים בכל תא ־ כלומר, בעיה של בחירה של $n-k$ מתוך k בלי חשיבות לסדר ועם חזרות, ולכן $CC_k^{n-k} = \binom{n-1}{k-1}$.

2. n כדורים שונים, k תאים שונים, אין תא ריק.
עבור $k > n$ התשובה היא 0 ולכן נניח כי $k \leq n$.
במקרה שבו הכדורים היו זהים, הפתרון היה לחלק כדור לכל תא ואז לחלק את $n-k$ הנותרים בלי חשיבות לסדר ועם חזרות. נסיון לנקוט בגישה זו כאן יוביל לספירה כפולה (אם $1,2$ באותו תא זה ייספר פעם אחת כאשר 1 ייבחר להיות כדור שמחולק בשלב הראשון ו־2 מחולק בשלב השני, ופעם כש־2 מחולק בשלב הראשון ו־1 בשלב השני).
במקום זאת נפתור באמצעות הכלה והפרדה. התכונה P_i תהיה `התא $i$ ריק`.
את $w(r)$ נחשב באופן הבא: מספר הדרכים לבחור r מתוך k תאים כדי שיהיו ריקים ($\binom{k}{r}$), וחלוקה חופשית של כדורים ל־$k-r$ התאים הנותרים ($(k-r)^n$).
נקבל את הפתרון $T(n,k)=\sum_{r=0}^{k}(-1)^r\binom{k}{r}(k-r)^n$ אפשרויות. למרבה הצער, אין נוסחה סגורה.

3. n כדורים שונים, k תאים זהים, אין תא ריק.
זהו ניסוח שקול ל`חלוקה של קבוצת המספרים $\{1,2,\ldots,n\}$ ל־$k$ קבוצות זרות לא ריקות`. מספר זה, $S(n,k)$, נקרא `מספר סטירלינג מהסוג השני` ומסומן לפעמים $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$.
פתרון: נחלק את הכדורים ל־k תאים **שונים** - $T(n,k)$. כעת נחלק במספר הסדרים הפנימיים של תאים ונקבל

$$S(n,k)=\frac{T(n,k)}{k!}=\frac{1}{k!}\sum_{r=0}^{k}(-1)^r\binom{k}{r}(k-r)^n$$

בשל התפקיד המרכזי יותר במתמטיקה של $S(n,k)$ נוכל כעת לכתוב $T(n,k)=k!S(n,k)$.
ניתן למצוא גם נוסחת נסיגה פשוטה עבור $S(n,k)$. אם אנחנו מחלקים את המספרים $1,\ldots,n$ לקבוצות, נסתכל בקבוצה של n. יש שתי אפשרויות:

- או ש־n לבדו, או שאינו לבדו. אם הוא לבדו, שאר קבוצות המספרים מהוות חלוקה של $n-1$ מספרים ל־$k-1$ קבוצות, כלומר יש $S(n-1,k-1)$ אפשרויות במקרה זה.
- או ש־n אינו לבד. במקרה זה, אפשר לחשוב על הבניה של הרכב הקבוצות באופן הבא: ראשית נבנו k קבוצות שכוללות את $n-1$ המספרים הראשונים, ולבסוף נבחרה קבוצה להוסיף אליה את n (מכיוון ש־n אינו לבד, לא אפשרי המקרה שבו הוא יוצר קבוצה חדשה). זה תהליך דו שלבי, שבשלב הראשון שלו יש $S(n-1,k)$ אפשרויות ובשלב השני k אפשרויות (מספר הקבוצות להוסיף אליהן את n) ולכן מעיקרון הכפל יש במקרה זה בסך הכל $kS(n-1,k)$ אפשרויות.

על פי עקרון החיבור, נקבל $S(n,k)=kS(n-1,k)+S(n-1,k-1)$. תנאי ההתחלה הם $S(k,k)=1$ (תא אחד שכולל כדור אחד), $S(n,0)=0$ אם $n>0$ (לא ניתן לחלק מספר חיובי של כדורים אם אין אף תא שבו אפשר לשים אותם) ו־$S(n,k)=0$ עבור $n<k$ (אין מספיק כדורים כדי להבטיח שלא יהיו תאים ריקים).

4. n כדורים זהים, k תאים זהים ואין תא ריק. מסומן ב־$p_k\left(n\right)$.
זהה למספר טבלאות יאנג: טבלה עם k שורות ו־n משבצות בסך הכל, כך שבכל שורה אין יותר משבצות מאשר בשורה שקדמה לה.
זהה למספר האפשרויות לכתוב את n כסכום של k מספרים טבעיים חיוביים שמסודרים בסדר עולה (למשל: $3=1+2=1+1+1$ הן שלוש אפשרויות החלוקה של 3).
קיימת נוסחת הנסיגה הבאה:
$p_k\left(n\right)=p_{k-1}\left(n-1\right)+p_k\left(n-k\right)$
כאן המחובר הראשון מתאים לסיטואציה שבה באחד התאים יש בדיוק כדור אחד, והמחובר השני מתאים לסיטואציה שבה בכל התאים יש לפחות שני כדורים.
תנאי התחלה:
$p_0\left(0\right)=1$ (חלוקה `ריקה` של אפס כדורים לאפס תאים)
$p_k\left(n\right)=0$ כאשר $k>n$ (אם אין מספיק כדורים כדי לחלק כדור לכל תא, יהיה תא ריק ולכן אין חלוקות מתאימות).
נשים לב לכך שתנאי ההתחלה השני תקף גם כאשר n הוא שלילי.
אם אנחנו משמיטים את הדרישה לכך שיהיו בדיוק k תאים, מספר התוצאות האפשריות הוא $p\left(n\right)=\sum_{k=1}^{n}p_k\left(n\right)$ שהוא מספר האפשרויות הכולל לכתוב את n בתור סכום של טבעיים חיוביים. פונקציה זו, $p\left(n\right)$, נקראת `פונקציית החלוקה` והיא מהפונקציות המפורסמות בקומבינטוריקה ובתורת המספרים ועבודה רבה (החורגת מהיקף הקורס) הושקעה בהבנה של התנהגותה. ניתן למשל למצוא עבורה נוסחת נסיגה בעזרת שימוש בפונקציות יוצרות (אותן נראה בהמשך) ובתוצאה קומבינטורית הנקראת **משפט המספרים המחומשים**.

נחזור כעת אל הבעיות הקודמות שטרם טיפלנו בהן: המקרה של חלוקת n כדורים (שונים או זהים) ל־k תאים **זהים** כך שניתן לשים מספר כלשהו של כדורים בכל תא.

1. n כדורים שונים, k תאים זהים ואין מגבלה על חלוקת הכדורים.
מכיוון שהתאים זהים, לא ניתן להבדיל בין תאים ריקים. נחלק למקרים על פי מספר התאים הלא ריקים. אם יש i תאים לא ריקים, חזרנו אל תנאי הסעיף שנפתר על ידי מספרי סטירלינג, כלומר $S\left(n,i\right)$. לכן מעיקרון החיבור נקבל שהפתרון הכולל הוא $\sum_{i=0}^{k}S\left(n,i\right)$.
אם $k\ge n$ אין חשיבות לתאים עודפים מעבר ל־n תאים שכן הם תמיד יהיו ריקים ולא נוכל להבדיל ביניהם. נהוג לסמן מקרה זה ב־$B\left(n\right)=\sum_{i=0}^{n}S\left(n,i\right)$ ומספר זה, $B\left(n\right)$, נקרא `מספר בל". הוא מייצג את מספר הדרכים שבהן ניתן לחלק את המספרים $\{1,2,\ldots n\}$ למספר **כלשהו** של קבוצות זרות ולא ריקות.

2. n כדורים זהים, k תאים זהים, אין מגבלה על חלוקת הכדורים.
גם כאן ניתן לחשוב על הבעיה בתור כתיבה של המספר n בתור סכום של k טבעיים שמסודרים בסדר עולה, אך כעת גם 0 יכול להשתתף בסכום מספר כלשהו של פעמים.
נשתמש בתעלול הבא: אם $n=a_1+a_2+\ldots+a_k$ כך ש־$0\le a_1\le a_2\le\ldots\le a_k$ אז על ידי הוספת 1 לכל איבר נקבל $n+k=\left(a_1+1\right)+\left(a_2+1\right)+\ldots+\left(a_k+1\right)$ וקיבלנו חלוקה של $n+k$ ל־k טבעיים חיוביים. מספר החלוקות הללו הוא $p_k\left(n+k\right)$, ועל ידי החסרת 1 מכל איבר בחלוקה כזו נחזור לחלוקה מהסוג שממנו פתחנו כך שיש ביניהן התאמה חח"ע ועל, והפתרון לסעיף זה הוא $p_k\left(n+k\right)$.

נסכם בטבלה את כל מה שראינו:

כדורים	תאים	לכל היותר כדור אחד בתא	בלי הגבלה	לפחות כדור אחד בתא
שונים	שונים	P_k^n	k^n	$k!S\left(n,k\right)$
זהים	שונים	$\binom{k}{n}$	CC_k^n	CC_k^{n-k}
שונים	זהים	$[k\geq n]$	$\sum_{i=0}^{k}S\left(n,i\right)$	$S\left(n,k\right)$
זהים	זהים	$[k\geq n]$	$p_k\left(n+k\right)$	$p_k\left(n\right)$

8 פונקציות יוצרות

8.1 מבוא ודוגמאות ראשונות

מרבית הבעיות שבהן עוסקים בקומבינטוריקה הן בעיות ספירה עבור פרמטר n: לכל מספר טבעי $n\geq 0$ מתאים מספר a_n שמתאר את כמות האובייקטים שמתאימים לקריטריונים של בעיית הספירה עבור הפרמטר n הספיציפי. כך למשל D_n תיאר, לכל $n\geq 0$, את מספר הפרמוטציות מגודל n (הגודל הוא הפרמטר) שהן הפרות סדר. המטרה של הקומבינטוריקה היא להבין באופן הטוב ביותר את התנהגות הסדרה a_n.

באופן פורמלי מעט יותר, בבעיית ספירה קומבינטורית נתונה קבוצה A כך שלכל איבר $x\in A$ אנחנו מתאימים **גודל** $|x|$ שהוא מספר טבעי (כולל 0), ומגדירים את הסדרה
$$a_n=|\{x\in A\mid x=n\}|$$
כלומר, a_n סופר את מספר האיברים ב־A מגודל n.

עד כה אופן הפעולה שלנו היה כזה שבו אנחנו `מקפיאים` את n ומנסים למצוא נוסחה ספציפית עבור a_n, לפעמים בהסתמך על איברים אחרים בסדרה (מה שמניב נוסחת נסיגה). **פונקציות יוצרות** הן גישה שונה מהותית לתיאור סדרות, שמצליחה לתפוס את כל הסדרה `בבת אחת`. גישה זו מאפשרת התמודדות כללית וחזקה יותר עם בעיות ספירה רבות ובפרט כאלו שלא ניתן למצוא בהן נוסחה מפורשת עבור a_n, אך במבט ראשון היא גם מבלבלת ונראית `חלשה יותר` מהשיטות שנלמדו עד כה.

הרעיון בפונקציות יוצרות הוא `לשתול` את אברי הסדרה בתור מקדמים ב**טור חזקות אינסופי**; טור שכזה מגדיר פונקציה שלאחר מכן ניתן לבצע עליה מניפולציות אלגבריות סטנדרטיות שמתבצעות על פונקציות ־ חיבור עם פונקציות אחרות, כפל בסקלרים ובפונקציות אחרות, העלאה בחזקה ואפילו גזירה ואינטגרציה. לכל המניפולציות הללו משמעויות קומבינטוריות. המטרה היא למצוא ביטוי מפורש כלשהו לפונקציה היוצרת של הסדרה. מרגע שביטוי שכזה נמצא, ניתן להפיק ממנו לעתים נוסחה מפורשת עבור אברי הסדרה או נוסחת נסיגה עבורם, וגם במקרה שהדבר אינו מתאפשר עדיין יש מידע שניתן להפיק דוגמת קצב הגידול של הסדרה (לא ניכנס לניתוחים הללו בקורס זה).

קיימים מספר סוגים של פונקציות יוצרות ובקורס זה נציג רק את הפשוט שבהם:

הגדרה 1.8 (פונקציה יוצרת) עבור סדרה $\{a_n\}_{n=0}^{\infty}$, **הפונקציה היוצרת** של הסדרה היא הביטוי $\sum_{n=0}^{\infty}a_nx^n$.

בחשבון דיפרנציאלי ואינטגרלי יש חשיבות ל**תחום ההתכנסות** של טורי חזקות כדוגמת $\sum_{n=0}^{\infty}a_nx^n$, אך אנו לא נזדקק בשום שלב להצבת ערכים בתוך x כך שפרטים אלו לא יהיו רלוונטיים עבורינו.

דוגמא הפונקציה היוצרת של הסדרה הסופית $1,2,1$ (שניתן לחשוב עליה כעל הסדרה האינסופית $1,2,1,0,0,\ldots$) היא $f\left(x\right)=1+2x+x^2$.

באופן כללי, פונקציה יוצרת של סדרה היא פולינום אם ורק אם הסדרה היא סופית (מכילה רק אפסים החל ממקום מסויים).

דוגמא לסדרה $\binom{n}{0},\binom{n}{1},\ldots,\binom{n}{n}$, כלומר הסדרה $a_k=\binom{n}{k}$ יש את הפונקציה היוצרת $f(x)=\sum_{i=0}^{n}\binom{n}{i}x^i$.

באמצעות הבינום של ניוטון ניתן לפשט את הביטוי: $f(x)=(1+x)^n$.

דוגמא זו מעידה על אחד מהגורמים לכוחן הרב של פונקציות יוצרות ־ לרוב ניתן לתת להן ביטוי פשוט שקל לבצע בו מניפולציות אלגבריות.

דוגמא לסדרה $0,0,0,\ldots$, כלומר הסדרה $a_n=0$, יש את הפונקציה היוצרת $f(x)=\sum_{n=0}^{\infty}0\cdot x^n=0$.

דוגמא לסדרה $1,1,1,\ldots$, כלומר הסדרה $a_n=1$, יש את הפונקציה היוצרת $f(x)=\sum_{n=0}^{\infty}x^n=\frac{1}{1-x}$.

השוויון האחרון הוא הנוסחה המוכרת לסכום של טור הנדסי אינסופי מתכנס ־ אולם כזכור, איננו מניחים כלום על ההתכנסות של הטור ולכן אנו נדרשים לנימוק שונה שנראה בהמשך.

דוגמא לסדרה $1,\lambda,\lambda^2,\ldots$, כלומר הסדרה $a_n=\lambda^n$, יש את הפונקציה היוצרת $f(x)=\sum_{n=0}^{\infty}(\lambda x)^n=\frac{1}{1-\lambda x}$.

8.2 פעולות על פונקציות יוצרות

ראשית, עלינו להסביר טוב יותר מה האובייקט שאנו משתמשים בו כשאנו עובדים עם פונקציות יוצרות ־ מה שכינינו "טור חזקות" ועכשיו נכנה "טור חזקות **פורמלי**" כדי להדגיש את ההבדל בין זה ובין טורי החזקות שמופיעים בחדו"א ובדיון בהם עוסקים גם ברדיוס התכנסות.

אינטואיטיבית, טור חזקות פורמלי הוא אובייקט הדומה לפולינום, רק שבעוד שבפולינום יש מספר סופי של מקדמים, בטור חזקות פורמלי יש אינסוף.

הגדרה 2.8 טור חזקות פורמלי הוא ביטוי מהצורה $\sum_{n=0}^{\infty}a_nx^n$. האיברים a_n נקראים **המקדמים** של הטור.

כמו עם פולינומים, כך גם על טורי חזקות אפשר להגדיר פעולות אלגבריות:

הגדרה 3.8 (חיבור של טורי חזקות): בהינתן שני טורי חזקות $a(x)=\sum_{n=0}^{\infty}a_nx^n, b(x)=\sum_{n=0}^{\infty}b_nx^n$ **הסכום** שלהם $c(x)=a(x)+b(x)$ הוא טור חזקות $c(x)=\sum_{n=0}^{\infty}c_nx^n$ כך ש־$c_n=a_n+b_n$.

הגדרת החיבור היא פשוטה. הגדרת הכפל מורכבת יותר (כדי לקבל אינטואיציה, כדאי לנסות לכפול פולינומים ולראות מה קורה) אך היא גם הסיבה לכוח של ייצוג סדרות באמצעות פונקציות יוצרות.

הגדרה 4.8 (כפל של טורי חזקות): בהינתן שני טורי חזקות $a(x)=\sum_{n=0}^{\infty}a_nx^n, b(x)=\sum_{n=0}^{\infty}b_nx^n$ **המכפלה** שלהם $c(x)=a(x)b(x)$ היא טור חזקות $c(x)=\sum_{n=0}^{\infty}c_nx^n$ כך ש־$c_n=\sum_{k=0}^{n}a_kb_{n-k}$.

דוגמא נתבונן על שני טורי החזקות

$a(x) = 1 + x + x^2 + x^3 + \dots$

$b(x) = 1 - x$

כלומר, הטורים עבור הסדרות $a_n = 1$ ו־

$$b_n = \begin{cases} 1 & n = 0 \\ -1 & n = 1 \\ 0 & n > 1 \end{cases}$$

אינטואיטיבית, המכפלה של שני אלו תניב **טור טלסקופי**:

$$\begin{aligned} a(x)b(x) &= 1 - x + x(1-x) + x^2(1-x) + \dots \\ &= 1 - x + x - x^2 + x^2 - x^3 + \dots \\ &= 1 \end{aligned}$$

אבל זה אינו נימוק פורמלי; נימוק פורמלי ייעזר בנוסחה $c_n = \sum_{k=0}^{n} a_k b_{n-k}$:

$c_0 = a_0 b_0 = 1 \cdot 1 = 1$

$c_1 = a_0 b_1 + a_1 b_0 = 1 \cdot (-1) + 1 \cdot 1 = -1 + 1 = 0$

$c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0 = 1 \cdot 0 + 1 \cdot (-1) + 1 \cdot 1 = 0$

ובאופן כללי עבור $n \ge 1$, מכיוון שאם $n - k \le 1$ מתקיים $b_{n-k} = 0$:

$c_n = \sum_{k=0}^{n} a_k b_{n-k} = a_{n-1} b_1 + a_n b_0 = 1 - 1 = 0$

זו ההצדקה הפורמלית לכתיב $1 + x + x^2 + \dots = \frac{1}{1-x}$.

באופן דומה ניתן גם להוכיח את $1 + \lambda x + \lambda^2 x^2 + \dots = \frac{1}{1-\lambda x}$.

דוגמא אם $a(x) = \sum_{n=0}^{\infty} a_n x^n$ היא הפונקציה היוצרת של הסדרה $\{a_n\}_{n\ge 0}$ ו־$\lambda \in \mathbb{R}$ הוא סקלר כלשהו, אז $\lambda a(x) = \sum_{n=0}^{\infty} \lambda a_n x^n$ היא הפונקציה היוצרת של הסדרה $\{\lambda a_n\}_{n\ge 0}$ (כאן λ הוא טור החזקות שהאיבר הראשון שלו הוא λ וכל יתר האיברים הם 0).

דוגמא אם $a(x) = \sum_{n=0}^{\infty} a_n x^n$ היא הפונקציה היוצרת של הסדרה $a_0, a_1, a_2, \dots$ אז

$$\begin{aligned} xa(x) &= x \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+1} \\ &= 0 \cdot x^0 + \sum_{n=1}^{\infty} a_{n-1} x^n \end{aligned}$$

היא הפונקציה היוצרת של הסדרה $0, a_0, a_1, \dots$. כלומר, מכפלה ב־x מזיזה את הסדרה צעד אחד ימינה ומכניסה 0 בהתחלה. בדומה, מכפלה ב־x^k תזיז את הסדרה k צעדים ימינה ותכניס k אפסים בהתחלה (כאן x הוא טור החזקות $0 + x + 0x^2 + 0x^3 + \dots$).

נחזור כעת לקומבינטוריקה. כזכור, בעיות הספירה שלנו הן על פי רוב מהצורה הזו: נתונה לנו **מחלקה** של אובייקטים A, ואנו מסמנים ב־a_n את מספר האובייקטים ב־A שה"גודל" שלהם הוא n. למשל, מספר המחרוזות שמורכבות מהתווים $\{1, 2, 3\}$ והן מאורך n הוא 3^n; כאן ה"גודל" של מחרוזת הוא מספר התווים שבה.

כעת נניח שאנו רוצים למצוא את מספר המחרוזות מאורך n שבנויות משני חלקים; בחלק הראשון יש מחרוזת מעל $\{1, 2, 3\}$ ובחלק השני יש מחרוזת מעל $\{a, b\}$. אין מגבלה

על האורך של כל חלק בנפרד: למשל, $12ab$ היא מחרוזת מתאימה מאורך 4 אבל גם 3333 וגם $abab$ הן מחרוזות מתאימות שכאלו.

אם ננסה למצוא את מספר המחרוזות מאורך n באמצעות עיקרון הכפל, נראה כי קודם כל עלינו להחליט כמה אותיות מהמחרוזות יהיו שייכות לחלק הראשון וכמה לחלק השני. אם מספר האותיות ששייכות לחלק הראשון הוא k, אז מספר המחרוזות מעל $\{1,2,3\}$ עבור החלק הראשון הוא 3^k ומספר המחרוזות מאורך $n-k$ מעל $\{a,b\}$ עבור החלק השני הוא 2^{n-k}. נקבל בסך הכל $3^k\cdot 2^{n-k}$, ומכיוון ש־k יכול להיות כל מספר בין 0 ל־n נקבל בסך הכל $\sum_{k=0}^{n}3^k\cdot 2^{n-k}$ - אותו הסכום שהופיע בהגדרת הכפל של פונקציות יוצרות.

התרגיל הזה ממחיש את העיקרון הכללי:

משפט 5.8 יהיו $\{a_n\}_{n\ge0}$ ו־$\{b_n\}_{n\ge0}$ סדרות כך ש־a_n מתאר את מספר האובייקטים מגודל n במחלקה A ו־b_n מתאר את מספר האובייקטים מגודל n במחלקה B, ויהיו $a(x),b(x)$ הפונקציות היוצרות המתאימות.

1. (חיבור) אם $C=A\cup B$ והאיחוד זר, כלומר כל אובייקט ב־C מגודל n הוא או אובייקט מגודל n ב־A או אובייקט מגודל n ב־B ו־$A\cap B=\emptyset$, אז $c(x)=a(x)+b(x)$.

2. (כפל) אם $C=A\times B$ כך שאיבר ב־C מגודל n הוא זוג של איבר מ־A ואיבר מ־B שסכום הגדלים שלהם הוא n, אז $c(x)=a(x)b(x)$.

המשפט נובע מהתכונות הבסיסיות של חיבור וכפל טורי חזקות.

נעבור כעת לדוגמאות.

דוגמא כמה פתרונות במספרים טבעיים יש למשוואה $x_1+\cdots+x_k=n$?

ראינו כבר כי בעיה זו היא דוגמא ל**בחירה ללא חשיבות לסדר ועם חזרות**: אנחנו מתחילים כשבכל המשתנים מוצב הערך 0, ו־n פעמים אנחנו בוחרים אחד מהמשתנים ומגדילים את ערכו ב־1. ראינו גם כי מספר זה הוא $\binom{n+k-1}{n}$.

נחשוב כעת על אותה בעיה מזווית הראייה של פונקציות יוצרות.

אם $A=\mathbb{N}$ כך שהגודל של מספר הוא פשוט המספר עצמו ($|x|=x$), אז האיברים מגודל n ב־$\mathbb{N}^k$ הם בדיוק ה־k־יות של מספרים טבעיים שסכומם n, דהיינו פתרון למשוואה $x_1+\cdots+x_k=n$, כלומר יש $\binom{n+k-1}{k-1}$ איברים מגודל n ב־$\mathbb{N}^k$. מצד שני, הפונקציה היוצרת של $\mathbb{N}$ היא פשוט $1+x+x^2+\cdots=\frac{1}{1-x}$ ולכן הפונקציה היוצרת של $\mathbb{N}^k$ היא העלאה בחזקת k של $\frac{1}{1-x}$. קיבלנו את הזהות השימושית הבאה:

$$\left(1+x+x^2+\ldots\right)^k=\frac{1}{(1-x)^k}=\sum_{n=0}^{\infty}\binom{n+k-1}{k-1}x^n$$

כדי לפשט את הסימונים בהמשך, נשתמש בסימון $CC_k^n=\binom{n+k-1}{k-1}$. בסימון זה, הזהות שמצאנו היא

$$\left(1+x+x^2+\ldots\right)^k=\frac{1}{(1-x)^k}=\sum_{n=0}^{\infty}CC_k^n x^n$$

דוגמא כמה פתרונות **במספרים טבעיים אי זוגיים** יש למשוואה $x_1+\cdots+x_k=n$?
כלומר, כאשר $x_i\in\{1,3,5,\ldots\}$ לכל $1\le i\le k$.
כאן במקום לקחת $A=\mathbb{N}$ ניקח את $A=\{1,3,5,\ldots\}$ ונתבונן ב־k־יות של A^k.
הפונקציה היוצרת של A היא $x+x^3+x^5+\ldots$ (כי חזקות זוגיות של x מייצגות את המספרים הזוגיים, ואנחנו מחפשים פתרונות שכולם במספרים אי זוגיים).
כדי לקבל ביטוי סגור עבור הטור הזה, נשתמש במניפולציות אלגבריות:

$$x+x^3+x^5+\ldots=x\left(1+x^2+x^4+\ldots\right)$$

אותו טיעון שהראה כי $1+x+x^2+\ldots=\frac{1}{1-x}$ מראה כי

$$1+x^2+x^4+\ldots=\frac{1}{1-x^2}$$

(כדי לקבל אינטואיציה, אפשר לסמן $y=x^2$ ואז מקבלים $1+x^2+x^4+\ldots=1+y+y^2+\ldots=\frac{1}{1-y}=\frac{1}{1-x^2}$).
לכן הפונקציה היוצרת עבור מספר הפתרונות היא:

$$\begin{aligned}\left(x+x^3+x^5+\ldots\right)^k&=x^k\left(1+x^2+x^4+\ldots\right)^k\\&=\frac{x^k}{\left(1-x^2\right)^k}\end{aligned}$$

קיבלנו ביטוי סגור פשוט עבור הפונקציה היוצרת: $\frac{x^k}{(1-x^2)^k}$. ביטוי זה מספיק לנו לצרכים רבים ובפרט לתרגילים מסובכים יותר שמתבססים על הנוכחי. עם זאת, אנו רוצים לנסות לחלץ מהפונקציה היוצרת גם נוסחה סגורה עבור מספר הפתרונות, ולכן נפתח את הביטוי לטור. נזכור כי ראינו

$$\frac{1}{(1-y)^k}=\sum_{t=0}^{\infty}CC_k^t y^t$$

(כאן במקום להשתמש ב־x,n כרגיל השתמשנו ב־y,t כדי לא לערבב את הסימונים של הנוסחה הזו שמצאנו קודם עם הסימונים של התרגיל החדש).
לכן אם נציב $y=x^2$ ונכפיל בביטוי x^k, נקבל:

$$\frac{x^k}{\left(1-x^2\right)^k}=x^k\sum_{t=0}^{\infty}CC_k^t x^{2t}=\sum_{t=0}^{\infty}CC_k^t x^{2t+k}$$

נזכור מה אנחנו מעוניינים למצוא: $\frac{x^k}{(1-x^2)^k}=\sum_{n=0}^{\infty}a_nx^n$ ואנו רוצים למצוא נוסחה סגורה ל־a_n. אז נשווה את הביטוי הזה עם הביטוי שמצאנו:

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{t=0}^{\infty} CC_k^t x^{2t+k}$$

אם n אינו מהצורה $2t+k$ אז המקדם של x^n בביטוי מימין הוא 0. לכן נקבל בסופו של דבר:

$$a_n = \begin{cases} 0 & n \neq 2t+k \\ CC_k^t & n = 2t+k \end{cases}$$

כלומר, כאשר $n = 2t+k$ מספר האפשרויות שווה לבחירה עם חזרות וללא חשיבות לסדר של $t = \frac{n-k}{2}$ איברים מתוך k איברים אפשריים: $CC_k^{\frac{n-k}{2}}$.

דוגמא כמה פתרונות במספרים טבעיים יש למשוואה $x_1 + \cdots + x_k = n$ כאשר לכל $1 \le i \le k$ מתקיים $0 \le x_i \le m$ עבור מספר טבעי m?

ההגבלה כאן על גודל הערך ש־x_i יכול לקבל מקשה מאוד על השימוש בפתרון סטנדרטי של בחירה עם חזרות ובלי חשיבות לסדר. דרך אחת להתמודד עם הקושי היא באמצעות **עקרון ההכלה וההפרדה** (תכונה "רעה" היא כשמשתנה מקבל את לפחות את הערך $m+1$).

כאן נציג את ההתמודדות עם הקושי באמצעות שימוש בפונקציות יוצרות. הערכים שמשתנה אחד יכול לקבל הם אברי הקבוצה $A = \{0, 1, \ldots, m\}$ ולכן הפונקציה היוצרת של הערכים שמשתנה אחד יכול לקבל היא $f(x) = 1 + x + x^2 + \ldots + x^m$. אפשר לקבל ביטוי סגור ל־$f(x)$ על ידי הנוסחה הסטנדרטית לטור הנדסי סופי:

$$f(x) = \frac{1 - x^{m+1}}{1 - x}$$

כמקודם, אנחנו מעוניינים ב־ $f^k(x) = \frac{\left(1-x^{m+1}\right)^k}{(1-x)^k}$ - הפונקציה היוצרת שמתארת את מספר הפתרונות למשוואה כאשר יש לנו k משתנים.

ראשית, נטפל במונה. אנחנו יודעים איך לפתוח אותו באמצעות **הבינום של ניוטון**:

$$\begin{aligned}\left(1 - x^{m+1}\right)^k &= \sum_{i=0}^{k} \binom{k}{t} \left(-x^{m+1}\right)^i \cdot 1^{k-i} \\ &= \sum_{i=0}^{k} (-1)^i \binom{k}{i} x^{i(m+1)}\end{aligned}$$

עבור המכנה נתבסס שוב על התוצאה שראינו קודם:

$$\frac{1}{(1-x)^k} = \sum_{t=0}^{\infty} CC_k^t x^t$$

ולכן

$$\frac{\left(1-x^{m+1}\right)^{k}}{\left(1-x\right)^{k}} = \left(1-x^{m+1}\right)^{k}\sum_{t=0}^{\infty}CC_{k}^{t}x^{t}$$
$$= \sum_{i=0}^{k}\left(-1\right)^{i}\binom{k}{i}x^{i(m+1)}\sum_{t=0}^{\infty}CC_{k}^{t}x^{t}$$
$$= \sum_{i=0}^{k}\sum_{t=0}^{\infty}\left(-1\right)^{i}\binom{k}{i}CC_{k}^{t}x^{t+i(m+1)}$$

ושוב, אנו רוצים להשוות את הביטוי הזה לטור $\sum_{n=0}^{\infty}a_nx^n$ ולכל n, לחלץ את הערך של a_n. מכאן נשאלת השאלה: עבור n נתון, מה הערכים של t,i שעבורם מתקיים $n=t+i\left(m+1\right)$? כל זוג ערכים של t,i שכאלו תורמים למקדם של x^n. ועבור n נתון, מתקיים $t=n-i\left(m+1\right)$.

לכן נקבל:

$$a_n = \sum_{\substack{t,i:\\ n=t+i(m+1)}}\left(-1\right)^{i}\binom{k}{i}CC_{k}^{t} = \sum_{i=0}^{k}\left(-1\right)^{i}\binom{k}{i}CC_{k}^{n-i(m+1)}$$

כאשר אם $n-i\left(m+1\right)<0$ המוסכמה היא ש־$CC_{k}^{n-i(m+1)}=0$.

9 פתרון נוסחאות נסיגה

9.1 דוגמא ראשונה: נוסחת נסיגה עם צעד אחד אחורה

9.1.1 הבעיה

נתונים n ישרים במישור במיקום כללי, כלומר אין שלושה ישרים שנחתכים כולם באותה הנקודה. לכמה חלקים הם מחלקים את המישור?

לא קשה לראות שאם $n-1$ ישרים כבר מונחים במישור ומתווסף ישר חדש, הוא מוסיף למישור n חלקים חדשים ־ בכל פעם שבה הוא פוגש את אחד הישרים שכבר קיימים, הוא מחלק לשניים את האיזור שאליו הוא נכנס, ובנוסף לכך הוא מחלק לשניים את האיזור שבו הוא היה לפני שהוא פגש ישר כלשהו. זה נותן לנו את הנוסחה הרקורסיבית הבאה:

$a_0=1$ (המישור ללא ישרים כלל מורכב מחלק בודד)

$a_n=a_{n-1}+n$

אנו מעוניינים להפיק מנוסחת הנסיגה הזו פתרון סגור למשוואה. נציג שלוש דרכים שונות לעשות זאת:

1. הצבה נשנית.
2. שיטת המשוואה האופיינית.
3. פונקציות יוצרות.

9.1.2 שיטת ההצבה הנשנית

בשיטה זו מסתמכים על כך שניתן להציב את המשוואה הרקורסיבית שוב ושוב בעצמה ולאחר שעושים זאת n פעמים תתקבל נוסחה שאינה רקורסיבית. התקווה היא שניתן יהיה לגלות את החוקיות שנוצרת במהלך ההצבות הנשנות הללו (מה שדורש יצירתיות לעתים).

עבור הדוגמה שלנו:

$$\begin{aligned} a_n &= a_{n-1}+n \\ &= a_{n-2}+(n+n)-1 \\ &= a_{n-3}+(n+n+n)-(1+2) \\ &= a_{n-4}+(n+n+n+n)-(1+2+3) \end{aligned}$$

וכן הלאה. בבירור הצורה הכללית כאן היא $a_n = a_{n-k}+kn-(1+2+\cdots+(k-1))$.
נשתמש בנוסחה לסדרה חשבונית: $1+2+\cdots+k-1=\frac{k(k-1)}{2}$, ונקבל:
$a_n = a_{n-k}+kn-\frac{k(k-1)}{2}$.
כדי לסיים נציב $k=n$ ונשתמש בתנאי ההתחלה $a_0=1$ כדי לקבל:
$a_n = 1+n^2-\frac{n(n-1)}{2}=\frac{2+2n^2-n^2+n}{2}=\frac{n^2+n+2}{2}=1+\frac{n(n+1)}{2}=1+\binom{n+1}{2}$
בשלב זה מומלץ לבדוק את נכונות התוצאה על ידי הצבת ערכים קטנים בה.

9.1.3 שיטת המשוואה האופיינית

בשיטה זו אנו מתחילים עם `ניחוש` לצורה הכללית של הפתרון ומהצבה בנוסחה מקבלים את הפתרון במדויק. פורמלית, לאחר שנמצאה צורת הפתרון יש צורך להוכיח באינדוקציה כי זהו אכן הפתרון, שכן ההצבה אינה מוכיחה כי הפתרון נכון אלא רק מראה מה צריכים להיות הפרמטרים המדוייקים של הפתרון אם צורתו היא הצורה שניחשנו.

עבור נוסחת הנסיגה שלנו ננחש שצורת הפתרון הכללי היא $a_n = An^2+Bn+C$, נציב במשוואה הרקורסיבית ונקבל:
$An^2+Bn+C=A(n-1)^2+B(n-1)+C+n$
ואחרי פתיחת סוגריים ופישוט:
$A(2n-1)+B=n$
המשוואה הזו מתקיימת לכל n, ובפרט עבור $n=0,1$, כך שקיבלנו ממנה מייד שתי משוואות:
$-A+B=0$
$A+B=1$
שפתרונן הוא $A=B=\frac{1}{2}$.
כמו כן מתנאי ההתחלה $a_0=1$ נקבל $C=1$.
קיבלנו שצורת הפתרון הכללי היא $a_n=\frac{n^2+n}{2}+1=1+\binom{n+1}{2}$.

9.1.4 שימוש בפונקציות יוצרות

תהא $f(x)$ הפונקציה היוצרת של הסדרה a_n. אז מנוסחת הנסיגה ומתנאי ההתחלה אנו מקבלים את המשוואה הבאה:
$f(x)=xf(x)+\frac{x}{(1-x)^2}+1$
הסבר:

ה־a_{n-1} הוא $xf(x)$ ־ זו ההשפעה של ביצוע `הזזה ימינה` על כל אברי הסדרה על ידי כפל ב־x.
$\sum_{n=0}^{\infty}(n+1)x^n = \frac{1}{(1-x)^2}$ על פי הנוסחאות שראינו קודם, ועל כן $\frac{x}{(1-x)^2}$ = $\sum_{n=1}^{\infty}nx^n$ היא הפונקציה היוצרת של הסדרה בה האיבר ה־n־י הוא n (שיטה אחרת: $\sum_{n=1}^{\infty}nx^n = x\cdot\left(\frac{1}{1-x}\right)' = \frac{x}{(1-x)^2}$).
ה־$+1$ הוא תנאי ההתחלה.
מהמשוואה לעיל נחלץ את $f(x)$ ונקבל:
$f(x) = \frac{x}{(1-x)^3} + \frac{1}{1-x}$
כזכור, הטור של $\frac{1}{(1-x)^3}$ הוא $\sum_{n=0}^{\infty}\binom{n+2}{2}x^n$ ולכן על ידי כפל ב־x מקבלים את

$$\sum_{n=0}^{\infty}\binom{n+2}{2}x^{n+1} = \sum_{n=0}^{\infty}\binom{n+1}{2}x^n$$

הטור של $\frac{1}{1-x}$ הוא $\sum_{n=0}^{\infty}x^n$, ולכן נקבל שפתרון הנוסחה הוא $\binom{n+1}{2}+1$, שוב.

9.2 דוגמא שניה: נוסחת נסיגה עם שני צעדים אחורה

נתבונן על נוסחת פיבונאצ'י,

$$a_n = a_{n-1} + a_{n-2}$$

אנו רוצים למצוא ביטוי סגור ל־a_n כדי להדגים שתי טכניקות כלליות שבהן ניתן לגשת לבעיה הזו.

9.2.1 שיטת המשוואה האופיינית

להבדיל מבדוגמא הקודמת, עבור נוסחת נסיגה כמו $a_n = a_{n-1} + a_{n-2}$ שבה הולכים שני צעדים אחורה, הערכים של a_n גדלים **אקספוננציאלית**:

$$\begin{aligned} a_n &= a_{n-1} + a_{n-2} = (a_{n-2} + a_{n-3}) + a_{n-2} \\ &\geq 2a_{n-2} \geq 4a_{n-4} \geq 8a_{n-6} \geq \ldots \\ &\geq 2^k a_{n-2k} = \ldots = O\left(2^{n/2}\right) \end{aligned}$$

זה מוביל אותנו לנחש פתרון שהוא פונקציה אקספוננציאלית, כלומר פונקציה מהצורה $a_n = \lambda^n$, אולם מהו הערך של λ? אם נציב $a_n = \lambda^n$ בנוסחת הנסיגה $a_n = a_{n-1} + a_{n-2}$, נקבל:

$$\lambda^n = \lambda^{n-1} + \lambda^{n-2}$$

אמנם $\lambda = 0$ מניב את הפתרון הקביל $a_n = 0$ לנוסחת הנסיגה, אולם ברור שזה לא הפתרון שאנחנו מחפשים: בפרט הוא אינו מקיים את תנאי ההתחלה $a_0 = 0, a_1 = 1$. לכן נניח ש־$\lambda \neq 0$ ולכן ניתן לחלק בו, להעביר אגפים ולקבל

$$\lambda^2 - \lambda - 1 = 0$$

זו משוואה ממעלה שניה ואנו יודעים לפתור משוואות כאלו באמצעות נוסחת השורשים:

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

נסמן $\phi_+ = \frac{1+\sqrt{5}}{2}$ ו־$\phi_- = \frac{1-\sqrt{5}}{2}$. המספר ϕ_+ הוא מוכר למדי ומכונה **יחס הזהב**.

העובדה ש־ϕ_+, ϕ_- פותרים את המשוואה $\lambda^2 - \lambda - 1 = 0$ מוכיחה שהם פותרים גם את המשוואה $\lambda^n = \lambda^{n-1} + \lambda^{n-2}$, כך שמצאנו **שני** פתרונות שונים אפשריים לנוסחת הנסיגה $a_n = a_{n-1} + a_{n-2}$:

$a_n = \phi_+^n$

$a_n = \phi_-^n$

לרוע המזל, אף אחד משני פתרונות אלו אינו מקיים את תנאי ההתחלה עבור סדרת פיבונאצ'י, כלומר $a_0 = 0$ ו־$a_1 = 1$. למשל, עבור הפתרון של ϕ_+^n, הערכים הראשונים הם $1, \phi_+$ במקום $0, 1$.

למרבה המזל, בהינתן שני הפתרונות לנוסחת הנסיגה ניתן ליצור מהם אינסוף פתרונות נוספים באמצעות **צירוף לינארי** של הפתרונות הקיימים: לכל $A, B \in \mathbb{R}$, גם $a_n = A\phi_+^n + B\phi_-^n$ יהיה פתרון של נוסחת הנסיגה (ניתן לראות זאת על ידי הצבה ובדיקה ישירה). ננסה, אם כן, לבנות מהפתרונות שמצאנו פתרון חדש לנוסחת הנסיגה שבנוסף יקיים את תנאי ההתחלה. נציב $n = 0$ ו־$n = 1$ ונקבל את זוג המשוואות הבאות:

$0 = A\phi_+^0 + B\phi_-^0 = A + B$

$1 = A\phi_+ + B\phi_-$

מהמשוואה הראשונה נסיק $A = -B$ וכשנציב זאת במשוואה השניה נקבל

$1 = A\left(\phi_+ - \phi_-\right)$

מכיוון ש־$\phi_+ - \phi_- = \frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2} = \sqrt{5}$ נקבל

$A = \frac{1}{\phi_+ - \phi_-} = \frac{1}{\sqrt{5}}$

ולכן הפתרון לנוסחת הנסיגה שמקיים גם את תנאי ההתחלה הוא

$a_n = A\phi_+^n + B\phi_-^n = \frac{\phi_+^n - \phi_-^n}{\sqrt{5}}$

9.2.2 שימוש כללי בשיטת המשוואה האופיינית

הטכניקה שבה השתמשנו עבור פיבונאצ'י ניתנת להכללה עבור כל נוסחת נסיגה **לינארית**, כלומר כזו מהצורה

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}$$

בנוסחת נסיגה לינארית, האיבר a_n הוא **צירוף לינארי** של k איברים קודמים ־ סכום של האיברים הללו כשכל אחד מהם מוכפל בסקלר c. הנוסחה $a_n = a_{n-1} + n$ שראינו קודם אינה לינארית בגלל האיבר החופשי n שאינו כפל במקדם של איבר קודם בנוסחת הנסיגה.

גם הנוסחה $a_n = a_{n-1}^2 + a_n$ איננה לינארית כי האיבר a_{n-1} אינו מופיע כמות שהוא אלא כשהוא מועלה בריבוע.

כאשר נתונה לנו נוסחת נסיגה לינארית, אנו מחפשים לה פתרונות מהצורה $a_n = \lambda^n$, כפי שראינו קודם. הצבה של פתרון כזה בנוסחת הנסיגה מניבה בסופו של דבר את המשוואה

$$\lambda^k - c_1\lambda^{k-1} - \ldots - c_k = 0$$

אם $k = 2$ אנו יכולים לפתור את המשוואה בקלות בעזרת נוסחת השורשים, אבל עבור ערכים גדולים יותר של k המצב קשה יותר (בפרט, עבור $k \geq 5$ לא קיימת נוסחה כללית לפתרון משוואות כאלו) ולעתים קרובות נזקקים לאלגוריתם נומרי (כדוגמת אלגוריתם **ניוטון־רפסון**) שיחשב קירוב טובים לפתרונות.

נוסחת נסיגה שהולכת אחורה k צעדים זקוקה ל־k תנאי התחלה שונים. אם בנוסף לכך קיימים למשוואה k פתרונות שונים $\lambda_1, \ldots, \lambda_k$, אז נכתוב פתרון כללי מהצורה $A_1\lambda_1^n + \ldots + A_k\lambda_k^n$, נשווה לתנאי ההתחלה ונקבל מערכת של k משוואות לינאריות ב־k נעלמים:

$$\begin{aligned} A_1 + A_2 + \ldots + A_k &= a_0 \\ A_1\lambda_1 + A_2\lambda_2 + \ldots + A_k\lambda_k &= a_1 \\ &\vdots \\ A_1\lambda_1^{k-1} + A_2\lambda_2^{k-1} + \ldots + A_k\lambda_k^{k-1} &= a_{k-1} \end{aligned}$$

כאן הנעלמים הם $A_1, \ldots, A_k$. אם מערכת המשוואות פתירה, מצאנו פתרון לנוסחת הנסיגה שעונה על תנאי ההתחלה.

מה קורה אם למשוואה אין מספיק פתרונות? כדי להבין מתי זה קורה ניזכר בטענה כללית על משוואות פולינומיות.

המשפט היסודי של האלגברה קובע כי לפולינום ממעלה n מעל המרוכבים $\mathbb{C}$ קיימים בדיוק n שורשים, **עד כדי ריבוי**. משמעות הדבר היא שניתן לכתוב כל פולינום ממעלה n בתור

$(x - z_1)(x - z_2)\cdots(x - z_n)$

כך ש־$z_1, z_2 \ldots, z_n \in \mathbb{C}$ הם מספרים מרוכבים, לאו דווקא שונים זה מזה. אם "מאגדים" יחד פתרונות זהים, מקבלים את הכתיב

$(x - z_1)^{r_1}(x - z_2)^{r_2}\cdots(x - z_t)^{r_t}$

כך ש־$z_1, \ldots, z_t$ הם השורשים **השונים** של הפולינום, $r_1 + \ldots + r_t = n$ ו־r_i נקרא **הריבוי** של השורש z_i.

עד עכשיו עסקנו רק במקרה שבו היו לנו n שורשים שונים, כלומר הריבוי של כל אחד מהם היה 1, ובמקרה זה אם λ היה שורש אז λ^n היה פתרון של נוסחת הנסיגה.

אם λ הוא שורש מריבוי r, עדיין ניתן לקבל ממנו r פתרונות שונים לנוסחת הנסיגה:

$$\lambda^n, n\lambda^n, n^2\lambda^n, \ldots, n^{r-1}\lambda^n$$

כך שניתן להמשיך לפתור את נוסחת הנסיגה בעזרת פתרונות אלו.

לדוגמא, נתבונן בנוסחת הנסיגה $a_n = 4\left(a_{n-1} - a_{n-2}\right)$. המשוואה האופיינית עבור נוסחה זו היא $0 = \lambda^2 - 4\lambda + 4 = \left(\lambda - 2\right)^2$ ואנו מקבלים את הפתרונות הבאים של נוסחת הנסיגה:

$a_n = 2^n$

$a_n = n \cdot 2^n$

ואכן, אם נציב את $a_n = n \cdot 2^n$ בנוסחת הנסיגה נקבל:

$$\begin{aligned} 4\left(a_{n-1} - a_{n-2}\right) &= 4\left(\left(n-1\right) \cdot 2^{n-1} - \left(n-2\right) \cdot 2^{n-2}\right) \\ &= 4 \cdot 2^{n-2}\left(2\left(n-1\right) - \left(n-2\right)\right) \\ &= 2^n \cdot n = a_n \end{aligned}$$

נוכיח את התכונה המועילה הזו:

טענה 1.9 עבור נוסחת הנסיגה $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}$, אם λ הוא פתרון של המשוואה האופיינית $\lambda^k - c_1\lambda^{k-1} - \ldots - c_k = 0$ מריבוי r, אז $a_n = n^i \lambda^n$ הוא פתרון של נוסחת הנסיגה לכל $0 \le i < r$.

הוכחה: המפתח להוכחה נעוץ בעובדה הבאה: אם λ הוא שורש מריבוי r של פולינום $p\left(x\right)$ אז ניתן לכתוב $p\left(x\right) = \left(x - \lambda\right)^r q\left(x\right)$ (על פי ההגדרה של ריבוי שהצגנו). כעת נגזור את $p\left(x\right)$ על פי כללי הגזירה הרגילים, ונקבל:

$$\begin{aligned} p'\left(x\right) &= r\left(x - \lambda\right)^{r-1} q\left(x\right) + \left(x - \lambda\right)^r q'\left(x\right) \\ &= \left(x - \lambda\right)^{r-1}\left[rq\left(x\right) + \left(x - \lambda\right) q'\left(x\right)\right] \end{aligned}$$

כלומר, $p'\left(x\right) = \left(x - \lambda\right)^{r-1} s\left(x\right)$, ולכן אנו רואים ש־$\lambda$ הוא שורש מריבוי $r - 1$ של הפולינום $p'\left(x\right)$.

נשים לב גם לאבחנה הטריוויאלית לפיה אם λ הוא שורש מריבוי r של $p\left(x\right)$ אז הוא שורש מריבוי r של $x \cdot p\left(x\right)$ (שכן עדיין ניתן לכתוב $xp\left(x\right) = \left(x - \lambda^r\right) xq\left(x\right)$ כאשר $xq\left(x\right)$ הוא הרכיב שאינו מתאפס על ידי λ).

נעבור כעת להוכחת הטענה. ראשית נסמן את הפולינום שמתאים למשוואה האופיינית ב־$p\left(x\right)$:

$$p\left(x\right) = x^k - c_1 x^{k-1} - \ldots - c_k$$

מהנתון שלנו, λ הוא שורש מריבוי r של $p\left(x\right)$.

יהא n טבעי כלשהו. אנו רוצים להראות שלכל $0 \le i < r$ מתקיים

$$n^i \lambda^n = c_1\left(n-1\right)^i \lambda^{n-1} + c_2\left(n-2\right)^i \lambda^{n-2} + \ldots + c_k\left(n-k\right)^i \lambda^{n-k}$$

או במילים אחרות, אנו רוצים להראות ש־λ הוא שורש של הפולינום $q_i(x)$ המוגדר על ידי

$$q_i(x) = n^i x^n - c_1 (n-1)^i x^{n-1} + c_2 (n-2) x^{n-2} + \ldots + c_k (n-k)^i x^{n-k}$$

נראה טענה חזקה יותר: ש־λ הוא שורש מריבוי $r-i$ של $q_i(x)$, ונעשה זאת באינדוקציה על $i = 0, 1, \ldots, r-1$.

ראשית, בדיקה מיידית מראה כי $q_0(x) = x^{n-k} p(x)$ ומכיוון ש־λ מריבוי r ב־$p(x)$, כך גם ב־$q_0(x)$.

כעת, כל q_{i+1} עבור $i \geq 0$ מתקבל מקודמו באופן הבא:

$$\begin{aligned} x \cdot q_i'(x) &= x \left[n^i \cdot n x^{n-1} - c_1 (n-1)^i (n-1) x^{n-2} + \ldots + c_k (n-k)^i (n-k) x^{n-k-1} \right] \\ &= n^{i+1} x^n - c_1 (n-1)^{i+1} x^{n-1} + \ldots + c_k (n-k)^{i+1} x^{n-k} \\ &= q_{i+1}(x) \end{aligned}$$

כפי שראינו, אם λ הוא שורש מריבוי $r-i$ של q_i, אז λ הוא שורש מריבוי $r-i-1$ של q_i', ולכן שורש מריבוי $r-i-1$ של $x \cdot q_i'$. קיבלנו ש־λ הוא שורש מריבוי $r-(i+1)$ של q_{i+1}, כמבוקש. ■

9.3 נוסחאות נסיגה ופונקציות יוצרות רציונליות

פונקציה נקראת **רציונלית** אם היא מהצורה $f(x) = \frac{p(x)}{q(x)}$ כאשר $p(x), q(x)$ הם פולינומים.
למשל, $f(x) = \frac{x^2+3x}{1-x}$ כאשר הפולינומים הם $p(x) = x^2 + 3x$ ו־$q(x) = 1 - x$.

קיים קשר הדוק בין בעיות ספירה שקיימת עבורן נוסחת נסיגה לינארית ובין פונקציות יוצרות:

משפט 2.9 הסדרה $\{a_n\}_{n \geq 0}$ מקיימת את נוסחת הנסיגה הלינארית $a_n = c_1 a_{n-1} + \ldots + c_k a_{n-k}$ אם ורק אם הפונקציה היוצרת שלה $f(x) = \sum_{n=0}^{\infty} a_n x^n$ היא מהצורה $f(x) = \frac{p(x)}{1 - c_1 x - c_2 x^2 - \ldots - c_k x^k}$ כאשר $p(x)$ הוא פולינום ממעלה קטנה מ־k.

במילים אחרות, נוסחת הנסיגה "מקודדת" בתוך המכנה של הפונקציה היוצרת. **הוכחה:** נניח ש־a_n מקיימת את נוסחת הנסיגה. כזכור, לכפל של פונקציה יוצרת ב־x^i יש אפקט של "הזזת" הסדרה שהפונקציה היוצרת מייצגת i מקומות קדימה והכנסת 0 בהתחלה. כלומר

$$x^i f(x) = x^i \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+i} = \sum_{n=i}^{\infty} a_{n-i} x^n$$

ומכאן

$$
\begin{aligned}
c_1 x^1 f(x) + \ldots + c_k x^k f(x) &= \sum_{n=1}^{\infty} c_1 a_{n-1} x^n + \ldots + \sum_{n=k}^{\infty} c_k a_{n-k} x^n \\
&= t(x) + \sum_{n=k}^{\infty} c_1 a_{n-1} x^n + \ldots + \sum_{n=k}^{\infty} c_k a_{n-k} x^n \\
&= t(x) + \sum_{n=k}^{\infty} \left(c_1 a_{n-1} + \ldots + c_k a_{n-k}\right) x^n \\
&= t(x) + \sum_{n=k}^{\infty} a_n x^n \\
&= \sum_{n=0}^{\infty} a_n x^n - p(x) \\
&= f(x) - p(x)
\end{aligned}
$$

כאשר $t(x), p(x)$ הם פולינומים ממעלה לכל היותר $k-1$: $t(x)$ מתקבל מהאיברים של הטורים $\sum_{n=i}^{\infty} c_1 a_{n-i} x^n$ עבורם $x<k$; ואילו $p(x) = \sum_{n=0}^{k-1} a_n x^n - t(x)$.
קיבלנו את השוויון

$$c_1 x^1 f(x) + \ldots + c_k x^k f(x) = f(x) - p(x)$$

נעביר אגפים, נוציא גורם משותף ונקבל

$$f(x)\left(1 - c_1 x - \ldots - c_k x^k\right) = p(x)$$

נחלק ונקבל

$$f(x) = \frac{p(x)}{1 - c_1 x - \ldots - c_k x^k}$$

כמבוקש.
בכיוון השני, אם ל־$f(x)$ יש את הצורה הנ"ל, על ידי היפוך הפעולות החשבוניות שביצענו ניתן לשחזר את נוסחת הנסיגה $a_n = c_1 a_{n-1} + \ldots + c_k a_{n-k}$ עבור $n \geq k$. ■

ההוכחה לא סיפקה לנו נוסחה מפורשת עבור $p(x)$, שהוא החלק של הפונקציה היוצרת שתלוי בתנאי ההתחלה של הסדרה, $a_0, a_1, \ldots, a_{k-1}$, אולם קל יחסית לשחזר את $p(x) = b_0 + b_1 x + \ldots + b_{k-} x^{k-1}$ מתוך הנוסחה $f(x)\left(1 - c_1 x - \ldots - c_k x^k\right) = p(x)$. אם

נכתוב את במפורש את $f(x) = a_0 + a_1x + a_2x^2 + \ldots$, נפתח את הסוגריים ונשווה מקדם מקדם עבור החזקות של x שהן לכל היותר k:

$$\begin{aligned} b_0 &= a_0 \\ b_1 &= a_1 - c_1a_0 \\ b_2 &= a_2 - c_1a_1 - c_2a_0 \\ &\vdots \\ b_{k-1} &= a_{k-1} - c_1a_{k-2} - \ldots - c_{k-1}a_0 \end{aligned}$$

הנוסחאות שלעיל מאפשרות לנו לעשות גם את ההפך: לחשב רקורסיבית את האיברים הראשונים בסדרה מתוך המקדמים של נוסחת הנסיגה, והמקדמים של $p(x)$:

$$\begin{aligned} a_0 &= b_0 \\ a_1 &= b_1 + c_1a_0 \\ a_2 &= b_2 + c_1a_1 + c_2a_2 \\ &\vdots \\ a_{k-1} &= b_{k-1} + c_1a_{k-2} + \ldots + c_{k-1}a_0 \end{aligned}$$

דוגמא עבור נוסחת פיבונאצ'י, $a_n = a_{n-1} + a_{n-2}$, בלי תלות בתנאי ההתחלה הפונקציה היוצרת היא מהצורה $\frac{p(x)}{1-x-x^2}$ כאשר $p(x)$.

אם נבחר את תנאי ההתחלה $a_0 = 0$ ו־$a_1 = 1$ נקבל $p(x) = x$ ולכן את הפונקציה היוצרת $\frac{x}{1-x-x^2}$. אם לעומת זאת נבחר את תנאי ההתחלה $a_0 = a_1 = 1$ כפי שמקובל לעתים לעשות, נקבל $b_0 = 1$ ו־$b_1 = 1 - 1 = 0$ ולכן את הפונקציה היוצרת $\frac{1}{1-x-x^2}$.

חלק II

מבוא לתורת הגרפים

10 גרפים ־ הגדרה ודוגמאות

נפתח במספר בעיות ידועות במתמטיקה.

דוגמא נתונה צורה המורכבת מקודקודים וקווים המחברים אותם. האם ניתן לצייר אותה מבלי להרים את העט מהדף?

זוהי **בעיית הכרעה** אלגוריתמית; מסתבר שקיים אלגוריתם יעיל ופשוט ביותר אשר מכריע אותה, ונראה אותו בהמשך הקורס. הפתרון נעוץ בחשיבה על הצורה כעל **גרף** ותרגום השאלה האם ניתן לצייר אותה מבלי להרים את העט מהדף לשאלה האם קיים בגרף המתאים **מסלול אוילרי**.

דוגמא נתונה מפה כלשהי. האם ניתן לצבוע אותה עם ארבעה צבעים בלבד, כך ששתי מדינות סמוכות אינן צבועות באותו הבא?

זוהי **בעיית ארבעת הצבעים** המפורסמת מאוד, שהייתה פתוחה למעלה ממאה שנים עד שנפתרה (עם תשובה חיובית; די בארבעה צבעים) בשנות ה־70 של המאה ה־20, בסיוע מחשב (שבדק אלפי טענות פרטניות שנדרשו להוכחה הכללית). בניסוחה המתמטי בעיית ארבעת הצבעים היא השאלה **האם כל גרף מישורי הוא 4־צביע**?

דוגמא נתונים שלושה בתים ושלושה מקורות של מים, חשמל וגז. האם ניתן לחבר כל בית לכל שלושת המקורות מבלי ששני חיבורים ייחתכו? (הכל מצוייר על דף נייר והבתים המקורות הם נקודות).

התשובה לבעיה זו היא לא. בניסוחה המתמטי זוהי הטענה ש**הגרף הדו צדדי המלא** $K_{3,3}$ **איננו מישורי**.

דוגמא יש לבנות באופן אקראי ובזמן יעיל מבוך שבו בין כל שני תאים קיים מסלול אחד ויחיד.

בניסוח מתמטי מבקשים כאן לבנות **עץ פורש של גרף מלא**; קיימים אלגוריתמים יעילים לפתרון בעיה זו (ולפתרון בעיה כללית מעט יותר שבה יש `מחיר` לחיבור בין כל שני תאים ורוצים מבוך בעל מחיר מינימלי).

דוגמא נתונים n גברים ו־n נשים כך שכל אישה מעוניינת בחלק מהגברים וכל גבר מעוניין בחלק מהנשים. האם ניתן לחלק את את הגברים והנשים לזוגות באופן מונוגמי כך שיווצרו n זוגות שבהם בני הזוג מעוניינים אלו באלו?

משפט החתונה של הול נותן תיאור מדויק וקל לבדיקה של תנאי שבו הדבר מתאפשר. בניסוח מתמטי זוהי השאלה באילו תנאים יש **שידוך מושלם בגרף דו צדדי**.

נעבור כעת לתיאור פורמלי ואבחנות בסיסיות.

הגדרה 1.10 (גרפים)

- גרף הוא זוג $G=(V,E)$ כאשר V היא קבוצה כלשהי (`קודקודים`) ו־E היא אוסף זוגות של קודקודים (`קשתות`).
- אם יש שתיים או יותר קשתות מצומת v אל צומת u הן נקראות **קשתות מקבילות**.
- אם יש קשת מ־v אל v היא נקראת **חוג עצמי**.
- **גרף פשוט** הוא גרף ללא קשתות מקבילות וחוגים עצמיים.
- **גרף מכוון** הוא גרף שבו קשת מ־v אל u נחשבת שונה מקשת מ־u אל v (במקרה זה יכולה להיות קשת בכל כיוון והן לא ייקראו קשתות מקבילות). כל עוד אנחנו לא אומרים זאת במפורש, כל הגרפים שנעסוק בהם אינם מכוונים.
- בהינתן גרף מכוון G, **גרף התשתית** שלו הוא הגרף המתקבל מ־G על ידי מחיקת כיווני הקשתות (כלומר, על ידי הפיכת G לגרף לא מכוון).
- **דרגה** של צומת $v\in V$, המסומנת $d\left(v\right)$, היא מספר הקשתות בגרף שמחוברות אל v.
- בגרף מכוון, **דרגת הכניסה** של צומת v, המסומנת $d_{in}\left(v\right)$, היא מספר הקשתות שנכנסות אל v; **דרגת היציאה** $d_{out}\left(v\right)$ היא מספר הקשתות שיוצאות מ־v.

- צומת **מבודדת** היא צומת מדרגה 0.
- גרף $G=(V,E)$ הוא **סופי** אם הקבוצות V,E סופיות.

כבר כעת ניתן להוכיח משפט פשוט:

טענה 2.10 בגרף סופי $G=(V,E)$ מתקיים $\sum_{v\in V}d\left(v\right)=2\left|E\right|$ ־ סכום דרגות הקודקודים הוא פעמיים מספר הקשתות.

הוכחה: נספור נקודות חיבור של קודקוד וקשת בשתי דרכים שונות. בדרך הראשונה, נעבור קשת קשת ולכל קשת נוסיף 2 לספירה כי היא מחוברת בדיוק לשני קודקודים ־ קיבלנו $2\left|E\right|$. בדרך השניה נעבור קודקוד קודקוד ולכל קודקוד נוסיף לספירה את כל הקשתות שנוגעות בו ־ קיבלנו $\sum_{v\in V}d\left(v\right)$. ■

נחזור להגדרות:

הגדרה 3.10 (מסלולים, גרפים קשירים)

- **מסלול** בגרף הוא סדרה של צמתים $v_1,v_2,\ldots,v_n$ כך שבין כל שני צמתים סמוכים בסדרה יש קשת (ואם הגרף מכוון, הקשת היא מ־v_i אל v_{i+1}). מסלול יכול להיות גם אינסופי (ואז פשוט אין איבר אחרון לסדרה). מסלול מסומן לרוב בתור $v_1\to v_2\to\cdots\to v_n$.
- **אורך** של מסלול סופי הוא כמספר הקשתות שבהן עוברים במסלול (כל קשת נספרת כמספר הפעמים שעוברים בה), כלומר אורך המסלול $v_1\to\cdots\to v_n$ הוא $n-1$.
- **מעגל** בגרף הוא מסלול שבו צומת ההתחלה זהה לצומת הסיום: $v_1=v_n$ (כל צומת על המעגל יכול לשמש כצומת התחלה וסיום).
- מסלול או מעגל הם **פשוטים** אם הם אינם עוברים באותה צומת יותר מפעם אחת, למעט נקודות ההתחלה והסיום במקרה של מעגל. כמו כן מעגל פשוט נדרש להיות מאורך 3 לפחות.
- גרף הוא **קשיר** אם בין כל שני צמתים בגרף קיים מסלול.
- גרף מכוון הוא **קשיר** אם גרף התשתית שלו קשיר. הוא **קשיר היטב** אם קיים מסלול בגרף מכל צומת אל כל צומת אחר.

משפט 4.10 (אפיון אלטרנטיבי לקשירות של גרף) גרף לא מכוון $G=(V,E)$ הוא קשיר אם ורק אם בכל חתך שלו (חלוקה של V לאיחוד זר של שתי קבוצות לא ריקות $V=X\cup Y$) קיימת קשת מצומת כלשהי ב־X לצומת כלשהי ב־Y (עבור גרף מכוון, הגרף קשיר היטב אם ורק אם בכל חתך יש קשת מ־X אל Y ומ־Y אל X).

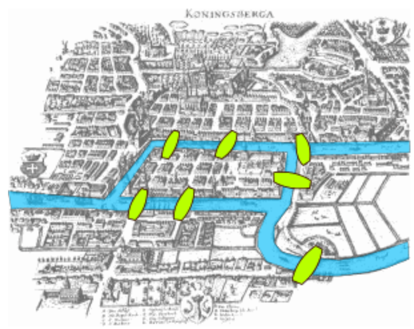
הוכחה: כיוון אחד: נניח כי G קשיר ויהא $V=X\cup Y$ חתך. X,Y לא ריקות אז יש $x\in X,y\in Y$. מכיוון שהגרף קשיר קיים מסלול $v_1\to v_2\to\cdots\to v_n$ כך ש־$v_1=x,v_n=y$.

יהא i האינדקס המינימלי של צומת במסלול v_i כך ש־$v_i\in Y$. מכיוון ש־$v_n=y\in Y$ ו־$v_1=x\notin Y$, הרי ש־$2\le i\le n$. מהמינימליות של i עולה ש־$v_{i-1}\in X$ ולכן (v_{i-1},v_i) היא קשת מ־X אל Y, כנדרש.

כיוון שני: נניח שהקריטריון מתקיים ונוכיח שהגרף קשיר. יהיו $x,y\in V$ כלשהם, ונגדיר קבוצה $U\subseteq V$ בתור קבוצת הצמתים שיש מסלול מ־x אליהם ב־G. בהכרח $x\in U$ כי קיים מסלול מ־x לעצמו באורך 0, ומכאן ש־U לא ריקה. אם $U=V$ אז סיימנו כי $y\in U$; אחרת $V=U\cup(V-U)$ הוא חתך של V ולכן קיימת קשת מ־$u\in U$ אל $v\in V-U$. אבל יש מסלול מ־x אל u ולכן יש מסלול מ־x אל v ($x\to\cdots\to u\to v$), אבל אז $v\in U$ בסתירה לכך ש־$v\in V-U$. מכאן ש־$U=V$, כנדרש. ■

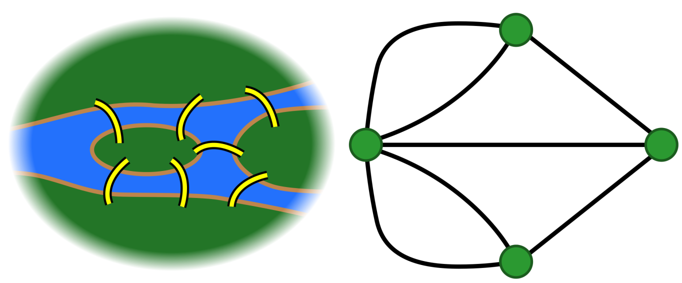
11 מסלולים אוילריים

נתונה מפת העיר קניגסברג של ימיו של אוילר.



את העיר חוצה נהר במספר מקומות ועל הנהר יש גשרים. האתגר: למצוא טיול בעיר שעובר בכל גשר בדיוק פעם אחת.

אפשר למדל את הבעיה באמצעות גרפים ־ קודקוד לכל איזור של העיר וקשת לכל גשר. מתקבל גרף לא מכוון עם קשתות מקבילות.



השאלה: האם קיים בגרף מסלול שעובר בכל קשת בדיוק פעם אחת?

הגדרה 1.11 מסלול בגרף שעובר בכל קשת בדיוק פעם אחת נקרא **מסלול אוילרי**.

מסלול בגרף שעובר בכל צומת בדיוק פעם אחת נקרא **מסלול המילטוני**.

בדומה, מעגל בגרף שעובר בכל קשת בדיוק פעם אחת נקרא **מעגל אוילרי** ומעגל בגרף שעובר בכל צומת (פרט לצומת ההתחלה והסיום) בדיוק פעם אחת נקרא **מעגל המילטוני**.

הבעיה של בדיקה האם קיים מסלול אוילרי בגרף היא פשוטה ונפתרה על ידי אוילר, כפתרון כללי לבעיית הגשרים של קניגסברג (על פתרון זה אומרים שהוא מציין את הולדת תורת

הגרפים). לעומת זאת, הבעיה של בדיקה האם קיים מסלול המילטוני בגרף היא קשה ולא נעסוק בה בקורס זה; הזכרנו מסלולים המילטוניים לצרכי שלמות בלבד.

הגדרה 2.11 גרף G נקרא **אוילרי** אם קיים בו מסלול אוילרי, ונקרא **אוילרי מעגלי** אם קיים בו מעגל אוילרי.

משפט 3.11 (אוילר) יהא G גרף סופי וקשיר, אז:

1. G הוא אוילרי מעגלי אם ורק אם $d\left(v\right)$ זוגית לכל $v\in V$.
2. G הוא אוילרי אם ורק אם $d\left(v\right)$ אי זוגי בדיוק עבור שני צמתים $v_{1},v_{2}\in V$.

הוכחה: ראשית נוכיח את 2 בהינתן ש־1 כבר הוכח. אם ב־G בדיוק שני צמתים מדרגה אי זוגית נוסיף קשת שמחברת אותם (ייתכן שהיא תהיה מקבילה לקשתות שכבר מחברות אותם) וכעת דרגת כל הצמתים בגרף זוגית וקיים בו מעגל אוילרי. ניקח את המעגל ונסיר ממנו את המעבר על פני הקשת שהוספנו (ובכך נכריח את הצמתים האי זוגיים להיות הראשון והאחרון במסלול המתקבל) וקיבלנו מסלול אוילרי ל־G.

בכיוון השני, אם G הוא אוילרי אז ניקח מסלול אוילרי בו, נחבר את צומת ההתחלה והסיום בקשת ונקבל גרף עם מעגל אוילרי, ולכן דרגת כל הצמתים בו זוגית, ולכן לאחר הסרת הקשת שהוספנו נקבל שדרגת בדיוק שני צמתים היא אי זוגית ־ הצמתים שלהם הוספנו קשת.

נעבור כעת להוכחת 1. נניח ש־G הוא אוילרי מעגלי ויהא $v_{1}\to v_{2}\to\cdots\to v_{1}$ מעגל אוילרי בו. נטייל על המעגל ובכל פעם שבה ניכנס לצומת או נצא ממנו נגדיל מונה עבור אותו צומת ב־1. נשים לב שבסיום הטיול על המעגל המונה של של צומת יהיה שווה בדיוק ל־$d\left(v\right)$ שכן אנחנו עוברים על כל הקשתות בגרף, ולכל קשת המחוברת ל־v אנו מגדילים את המונה שלו (אם כאשר אנחנו נכנסים אליו ואם כאשר אנחנו יוצאים).

כעת נשים לב שלכל צומת, מספר הפעמים שאנו נכנסים אליו שווה למספר הפעמים שאנו יוצאים ממנו (כי אחרי כל כניסה לצומת אנחנו גם יוצאים אליו למעט עבור v_{1} שפעם אחת (בהתחלה) אנו יוצאים ממנו מבלי להיכנס, ובפעם השניה אנו נכנסים אליו מבלי לצאת, כך ששוב אנו מתאזנים. מכאן ש־$d\left(v\right)$ זוגי תמיד.

הכיוון השני הוא עיקר ההוכחה. נניח ש־$d\left(v\right)$ זוגי לכל הצמתים בגרף הקשיר G ונוכיח כי קיים בו מעגל אוילרי.

נבחר צומת שרירותי $v\in V$ ונטייל בגרף החל ממנו באופן אקראי לחלוטין, כאשר אנו מוחקים כל קשת שאנו עוברים בה. מכיוון שדרגת כל צומת זוגית, מובטח לנו שבכל פעם שבה אנו נכנסים לצומת שאינה v אנו גם יכולים לצאת ממנה לאחר מכן ולכן איננו יכולים `להיתקע` אלא רק על ידי חזרה אל v. מכאן שהטיול שלנו בגרף ייצור בהכרח מעגל.

לאחר מחיקת כל קשתות המעגל מהגרף עדיין נשמר התנאי שכל הדרגות זוגיות. לכן ניתן לחזור שוב על התהליך (החל מצומת שדרגתה גדולה מ־0) ולקבל מעגל נוסף, וכן הלאה. בכל פעם מספר הקשתות שנותרו בגרף הופך לקטן יותר, ולכן התהליך יסתיים כעבור מספר סופי של צעדים ותתקבל סדרה $C_{1},C_{2},\ldots,C_{k}$ של מעגלים בגרף שכוללים את כל קשתות הגרף.

נשים לב שכל זוג מעגלים בעלי צומת משותף ניתן לאחד באופן הבא: אם u הוא הצומת המשותף, אז נבנה משני המעגלים מעגל שצומת ההתחלה שלו הוא u, לאחריה הולכים כמו על המעגל הראשון ומסיימים ב־u, ואז ממשיכים כמו המעגל השני ומסיימים ב־u (כמובן, המעגל אינו פשוט, אך זה לא נדרש).

כל עוד ניתן לאחד זוג מעגלים מתוך $C_{1},\ldots,C_{k}$, נעשה זאת. אם לבסוף מתקבל רק מעגל אחד, סיימנו; אחרת, תהא C קבוצת הצמתים של אחד המעגלים. מכיוון ש־G קשיר,

קיימת קשת מצומת u ב־C אל צומת $v \in V - C$. מכיוון שכל קשת שייכת למעגל כלשהו, גם הקשת (u,v) שייכת למעגל שאיננו C (כי $v \notin C$) אבל מכאן עולה שהצומת u שייך למעגל הזה ולכן הוא משותף למעגל ול־C, בסתירה לכך שאין שני מעגלים בעלי צומת משותף. ■

קיים ניסוח של המשפט גם עבור גרפים מכוונים:

משפט 4.11 (אוילר, גרסה לגרפים מכוונים) יהא G גרף סופי, מכוון וקשיר.

1. G הוא אוילרי מעגלי אם ורק אם לכל צומת v מתקיים $d_{in}(v) = d_{out}(v)$.
2. G אוילרי אם ורק אם $d_{in}(v) = d_{out}(v)$ לכל זוג צמתים פרט לשני צמתים v, u אשר מקיימים:

   (א) $d_{in}(v) = d_{out}(v) + 1$

   (ב) $d_{out}(u) = d_{in}(v) + 1$

הוכחה: ההוכחה דומה להוכחה של משפט אוילר הרגיל; התנאים של הגרסה המכוונת של המשפט מאפשרים להפעיל את אותה ההוכחה ללא שינוי מהותי. ■

12 גרפי דה־ברויין

ראשית נציג את ההגדרות הפורמליות שלנו עבור מה שבתכנות נקרא **מחרוזת** - סדרה של תווים.

הגדרה 1.12 אלפבית Σ הוא קבוצה סופית. אברי Σ נקראים **אותיות.**
מילה מעל Σ היא סדרה $w = \sigma_1 \dots \sigma_n$ של אותיות מ־Σ.
אורך של מילה הוא מספר האותיות שבה: $|\sigma_1 \dots \sigma_n| = n$.
אוסף כל המילים מאורך n מעל Σ מסומן ב־Σ^n.

עבור אלפבית Σ ו־n נתונים, אנו מתעניינים בסדרה קצרה ככל הניתן של אותיות כך ש**כל** מילה מאורך n מופיעה בתוך הסדרה כאחד מרצפי האותיות שבה, כשרצפים נלקחים בצורה ציקלית (כלומר, אם רצף חורג מעבר לסוף המילה הוא חוזר להתחלה).

הגדרה 2.12 מילה u מאורך n מופיעה במילה $\sigma_1 \dots \sigma_t$ **באופן ציקלי** אם $u = \sigma_i \sigma_{i+1} \dots \sigma_{i+n-1}$ כאשר פעולת החיבור באינדקסים מתבצעת מודולו t, במובן הבא: אם $j > t$ אז $\sigma_j \triangleq \sigma_{j-t}$ (אנו נמנעים מלהגדיר $\sigma_j \triangleq \sigma_{j \bmod t}$ מהטעם הטכני הפשוט שבשיטת הכתיב שלנו אין σ_0 אבל $t \bmod t = 0$).

למשל, עבור $\Sigma = \{0,1\}$ נתבונן על הסדרה 00111010. אם נעבור על אברי הסדרה משמאל לימין ובכל צעד ניקח את רצף 3 האותיות הבאות, נקבל את המילים

$$001, 011, 111, 110, 101, 010, 100, 000$$

אלו כל 8 המילים מאורך 3 מעל Σ, וקיבלנו אותם באמצעות סדרה מאורך 8. קל להשתכנע שסדרה כזו היא אופטימלית:

טענה 3.12 אם w היא מילה מעל Σ כך שכל מילה ב־Σ^n מופיעה באופן ציקלי ב־w, אז $|w| \geq |\Sigma|^n$.

הוכחה: למילים שונות שמופיעות ב־w יש אינדקס שונה עבור האות הראשונה, כך שאם מופיעות ב־w לפחות $|\Sigma|^n$ מילים שונות, יש לפחות $|\Sigma|^n$ אינדקסים שונים לאותיות w. ■

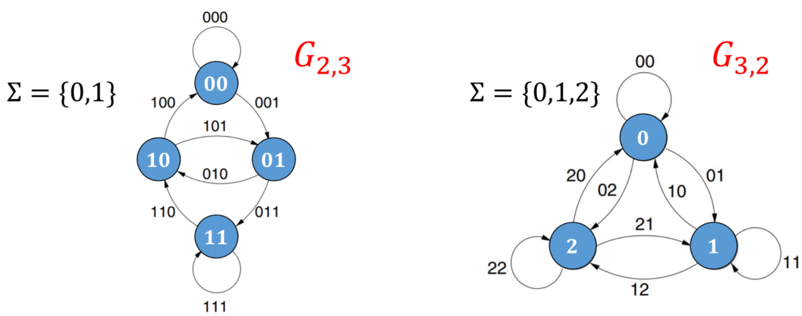
מכאן שאנו מתעניינים במיוחד בסדרות שהן מהאורך האופטימלי $|\Sigma|^n$ ונותנים להן שם:

הגדרה 4.12 סדרת דה־ברויין מעל Σ למילים מאורך n היא סדרה $\sigma_1 \ldots \sigma_t$ כך ש־$t = |\Sigma|^n$ ומתקיים $\{\sigma_i \sigma_{i+1} \ldots \sigma_{i+n-1} \mid i = 1, 2, \ldots, t\} = \Sigma^n$.

כיצד ניתן למצוא סדרות דה־ברויין? כאן באה תורת הגרפים לעזרתנו: עם בניה מתאימה של גרף מתאים, שייקרא **גרף דה־ברויין**, נוכל לקבל את כל סדרות דה־ברויין בתור **מעגלים אוילריים** בגרף.

הגדרה 5.12 גרף דה־ברויין עם פרמטרים k, n, המסומן $G_{k,n}$, הוא גרף מכוון המוגדר באופן הבא:

- ראשית מוגדר אלפבית $\Sigma = \{0, 1, \ldots, k-1\}$.
- $V = \Sigma^{n-1}$
- $E = \Sigma^n$
- הקשת $b_1 b_2 \ldots b_n$ יוצאת מהצומת $b_1 b_2 \ldots b_{n-1}$ ונכנסת לצומת $b_2 b_3 \ldots b_n$.



טענה 6.12 לכל k, n הגרף $G_{k,n}$ הוא אוילרי מעגלי.

הוכחה: על פי משפט אוילר, די להראות שלכל צומת $d_{in}(v) = d_{out}(v)$, ו־$G_{k,n}$ קשיר.

למעשה, $G_{k,n}$ קשיר היטב. נראה מסלול מצומת $u = a_1 a_2 \ldots a_{n-1}$ לצומת $v = b_1 b_2 \ldots b_{n-1}$:

$$a_1 a_2 \ldots a_{n-1} \quad \to a_2 \ldots a_{n-1} b_1 \quad \to \cdots \to b_1 b_2 \ldots b_{n-1}$$

כלומר, בכל צעד מכניסים מצד ימין עוד תו במילה של v ומוציאים תו מהמילה של u. קל לראות שהקשתות המתאימות קיימות.

כדי לראות ש־$d_{in}(v) = d_{out}(v)$ נשים לב להתאמה חח`ע ועל בין קשתות נכנסות ויוצאות מ־$v$: אם $v = a_1 \ldots a_{n-1}$ ו־$\sigma \in \Sigma$ כלשהו, אז נתאים בין הקשת הנכנסת $\sigma a_1 \ldots a_{n-1}$ והקשת היוצאת $a_2 \ldots a_{n-1} \sigma$. קל לראות כי זוהי אכן התאמה חח`ע ועל. ■

דוגמא למשל, כפי שראינו, דוגמא אחת למעגל אוילרי בגרף $G_{2,3}$ היא זו:
$00 \to 01 \to 11 \to 11 \to 10 \to 01 \to 10 \to 00 \to 00$
הקשתות עליהן עוברים במעגל הזה הן:
$001 \to 011 \to 111 \to 110 \to 101 \to 010 \to 100 \to 000$
ניתן להשתמש בכתיבה מקוצרת ולתאר את כל המסלול כסדרה באורך 8 (בכל פעם מתואר התו החדש שמתווסף למחרוזת): 00111010. זו סדרת דה־ברויין שראינו בהתחלה.
כעת נוכל להסיק:

טענה 7.12 לכל k, n יש סדרת דה־ברויין מעל $\Sigma = \{0, 1, \dots, k-1\}$ למילים מאורך n.

הוכחה: נתבונן בגרף דה־ברויין $G_{k,n}$. כפי שראינו, קיים בו מעגל אוילרי. יהיו $e_1, e_2, \dots, e_{k^n}$ הקשתות בגרף לפי הסדר שבו הן מופיעות במעגל. נבנה כעת את המילה $w = (e_1)_1 (e_2)_1 \cdots (e_{k^n})_1$, דהיינו מכל אחת מהקשתות על המעגל ניקח את האות הראשונה.
אנו טוענים כי w היא סדרת דה־ברוין המבוקשת. ראשית, היא מהאורך המתאים: k^n. כעת, אם נראה כי המילה $w_i w_{i+1} \dots w_{i+n-1}$ (כשהאינדקסים נלקחים בצורה ציקלית כפי שתיארנו) היא המילה e_i, סיימנו; מכיוון שהקשתות נלקחו ממעגל אוילרי, כל קשת בגרף (כלומר כל מילה ב־Σ^n) מופיעה כתת־מילה של w.
כדי לראות את הטענה הזו, ראשית נשים לב לכך שאם יש לנו קשת $e = \sigma_1 \dots \sigma_n$ בגרף והיא נכנסת לצומת v כלשהו, אז $v = \sigma_2 \dots \sigma_n$. כעת, כל קשת שיוצאת מ־$v$ היא מהצורה $\sigma_2 \dots \sigma_n \tau$ עבור $\tau \in \Sigma$ כלשהו. מכאן שלכל קשת e' שמופיעה אחרי e במעגל האוילרי, האות הראשונה ב־e' היא σ_2. באותו האופן כשלוקחים סדרה של n קשתות סמוכות שמתחילה ב־e, סדרת האותיות הראשונות בקשתות הללו תהיה $\sigma_1 \dots \sigma_n$, כמבוקש. ■

13 עצים

13.1 הגדרה ואפיונים בסיסיים

הגדרה 1.13 **עץ** הוא גרף פשוט G המקיים את שתי התכונות הבאות:

- G קשיר.
- G חסר מעגלים.

משפט 2.13 התנאים הבאים שקולים[4]:

1. G הוא עץ.
2. G חסר מעגלים ותוספת כל קשת ל־G יוצרת מעגל (G הוא מקסימלי ביחס לתכונה `חסר מעגלים`)
3. G קשיר ומחיקת כל קשת מ־G תהפוך אותו ללא קשיר (G הוא מינימלי ביחס לתכונה `קשיר`).
4. לכל זוג צמתים u, v קיים מסלול פשוט **יחיד** מ־u אל v.

[4] שימו לב לדמיון בין משפט זה למשפט מאלגברה לינארית לפיו ארבעת התנאים הבאים שקולים: קבוצת וקטורים היא בסיס, קבוצת וקטורים היא מקסימלית ביחס לאי־תלות לינארית, קבוצת וקטורים היא מינימלית ביחס לפרישה, וכל וקטור ניתן לכתיבה כצירוף לינארי יחיד של אברי הקבוצה. הדמיון איננו מקרי - מרחבים וקטוריים וגרפים הם שתי הדוגמאות הקלאסיות למטרואידים, שהם אובייקטים בעלי תכונות דומות לאלו.

הוכחה: נוכיח באמצעות שרשרת גרירות.

$1 \Leftarrow 2$: אם G הוא עץ הוא חסר מעגלים על פי הגדרה. נניח כי מוסיפים ל־G את הקשת (u,v); מכיוון ש־G קשיר כבר קיים מסלול בין u אל v, $v \to \cdots \to u$ ומכיוון ש־(u,v) לא הייתה בגרף, המסלול לא עובר דרך (u,v) ולכן ניתן להוסיף אותה בסוף המסלול ולהשלים אותו למעגל $u \to \cdots \to v \to u$.

$2 \Leftarrow 4$: ניקח שני צמתים $u,v \in V$. אם קיים ביניהם מסלול הוא יחיד, שכן שני מסלולים שונים ניתן לשרשר לקבלת מעגל ונתון ש־G חסר מעגלים. לכן נותר להוכיח כי קיים מסלול בין u,v. אם קיימת ב־G הקשת (u,v) אז קיים ביניהם המסלול $u \to v$. אם הקשת (u,v) אינה בגרף, אז הוספתה ל־G תיצור מעגל; מכיוון ש־G הוא חסר מעגלים המעגל חייב לעבור דרך (u,v) ולכן הוא מהצורה $u \to v \to \cdots \to u$ (בלי הגבלת הכלליות ניתן להניח שצומת ההתחלה והסיום הוא u) ומכאן שקיים ב־G כבר מסלול: $v \to \cdots \to u$.

$4 \Leftarrow 3$: G קשיר כי בין כל זוג צמתים קיים מסלול. תהא (u,v) קשת בגרף; מכאן ש־$u \to v$ הוא המסלול היחיד בגרף מ־u אל v, ולכן אם תוסר הקשת (u,v) לא יהיה מסלול מ־u אל v והגרף יפסיק להיות קשיר.

$3 \Leftarrow 1$: G קשיר על פי הנתון. נניח בשלילה שקיים בו מעגל פשוט $u \to v \to \cdots \to w \to u$. מכיוון שהמעגל פשוט $v \neq w$ (כי v אינו יכול להופיע פעמיים במעגל פשוט), אז לאחר הסרת הקשת (u,v) הגרף יוסיף להיות קשיר: כל מסלול בו שהשתמש בקשת (u,v) יכול ללכת במסלול $u \to w \to \cdots \to v$ במקום. ■

הגדרה 3.13 **יער** הוא גרף פשוט וחסר מעגלים (איחוד עצים זרים).
עלה בגרף כלשהו הוא צומת מדרגה 1.

טענה 4.13 ביער סופי בן קשת אחת לפחות קיימים לפחות שני עלים.

הוכחה: ניקח בגרף מסלול פשוט באורך מקסימלי (קיים כזה שכן הגרף סופי ולכן אורך כל מסלול פשוט חסום על ידי מספר הצמתים בגרף). נקודות ההתחלה והסיום חייבות להיות עלים, שאם לא כן ניתן יהיה להאריך את המסלול; אם צומת ההתחלה אינו עלה, הוא מחובר לצומת נוסף שחייב לא להיות על המסלול (שאם לא כן יהיה בגרף מעגל, בסתירה לכך שהוא יער), ולכן ניתן להרחיב את המסלול על ידי הוספת מעבר לאותו הצומת. ■

טענה 5.13 אם $G=(V,E)$ הוא עץ בעל $n=|V|$ צמתים אז $|E|=n-1$.

הוכחה: נוכיח באינדוקציה על n ־ מספר הצמתים.

בסיס האינדוקציה הוא עבור $n=1$: במקרה זה ב־G יש בדיוק $n-1=0$ קשתות (כי קשת דורשת לפחות שני צמתים).

נניח את נכונות הטענה עבור n ונוכיח שבעץ בעל $n+1$ צמתים יש בדיוק n קשתות.

מכיוון שיש בעץ לפחות שני צמתים, כדי שהוא יהיה קשיר בהכרח קיימת בו קשת אחת לפחות ולכן על פי טענה 4.13, ב־G קיים עלה. אם נסיר את העלה מהעץ, נקבל גרף בן n צמתים שגם הוא עץ (שכן הסרת עלה אינה פוגעת בקשירות העץ ואינה יכולה ליצור מעגל). לכן יש בו $n-1$ קשתות על פי הנחת האינדוקציה; ומכאן שב־G יש n קשתות ($n-1$ הקשתות של תת־העץ, ועוד הקשת שמחברת את תת־העץ אל העלה). ■

טענה 6.13 יהא $G=(V,E)$ גרף סופי בעל $n=|V|$ צמתים.

1. G הוא עץ אם ורק אם G חסר מעגלים בעל $n-1$ קשתות.

2. G הוא עץ אם ורק אם G קשיר בעל $n-1$ קשתות.

הוכחה: אם G הוא עץ אז לפי טענה 5.13 הוא בעל $n-1$ קשתות וכמובן שהוא קשיר וחסר מעגלים, כך שנותר רק להוכיח את הכיוון השני בכל אחת משתי הטענות.

נניח ש־G חסר מעגלים בעל $n-1$ קשתות. כל עוד ניתן להוסיף ל־G קשתות מבלי ליצור מעגל, נעשה זאת עד לקבלת גרף G' שבו הוספת כל קשת תיצור מעגל. על פי טענה G',2.13 הוא עץ; ולכן מטענה 5.13 יש בו $n-1$ קשתות, כלומר $G=G'$, ולכן G הוא עץ.

נניח ש־G קשיר בעל $n-1$ קשתות. כל עוד ניתן להסיר מ־G קשת מבלי לפגום בקשירות שלו נעשה זאת עד לקבלת גרף G' שבו הורדת כל קשת תפגום בקשירות של הגרף. פי טענה G',2.13 הוא עץ; ולכן מטענה 5.13 יש בו $n-1$ קשתות, כלומר $G=G'$, ולכן G הוא עץ.

■

13.2 משפט קיילי לספירת עצים

נסמן ב־f_n את מספר העצים על קבוצת הצמתים $V=\{1,2,\ldots,n\}$ (`עצים מסומנים`; כאשר צמתי העץ לא ניתנים להבחנה אלו מאלו הבעיה קשה בהרבה).

משפט 7.13 (קיילי) $f_n=n^{n-2}$.

נציג את ההוכחה של Prüfer. בהוכחה זו מראים התאמה חח`ע ועל בין קבוצת העצים על $V=\{1,2,\ldots,n\}$ וקבוצת המחרוזות מאורך $n-2$ מעל הא`ב $\{1,2,\ldots,n\}$. ההתאמה תוצג באמצעות אלגוריתם המתרגם עץ למחרוזת, והוכחה שהאלגוריתם מבצע התאמה חח"ע ועל.

אלגוריתם TreeToWord: קלט: $G=(V,E)$ כך ש־G הוא עץ.
פלט: מילה $w=\sigma_1\sigma_2\ldots\sigma_{n-2}$.
האלגוריתם:

1. עבור $i=1,2,,\ldots,n-2$:

(א) תהא $(u,v)\in E$ הקשת כך ש־u הוא העלה המינימלי (מבחינת מספרו הסידורי) בגרף G.

(ב) קבעו $\sigma_i=v$ (האות ה־i היא מספרו הסידורי של **השכן** של העלה u ב־G).

(ג) הסירו את u מהגרף G.

בסיום ריצת האלגוריתם הוסרו מהגרף $n-2$ צמתים ו־$n-2$ קשתות, ולכן G נותר עם זוג צמתים בודד שמחובר בקשת. כפי שנראה, אין צורך בטיפול נוסף בצמתים אלו.

נשים לב לכך שהאלגוריתם עובד תמיד; לא מובן מאליו שבשלב א' יהיה תמיד ניתן למצוא עלה, אך זה נובע ישירות מטענה 4.13: בתחילת האלגוריתם ב־G יש $n-1$ קשתות, והאלגוריתם מסיר $n-2$ קשתות במהלך ריצתו, ולכן תמיד קיימת ב־G קשת אחת לפחות ולכן תמיד קיים ב־G עלה אחד לפחות. מכאן שהפונקציה שהאלגוריתם מחשב היא מוגדרת היטב (לכל קלט קיים פלט יחיד).

כעת נותר להוכיח שההתאמה ש־TreeToWord מגדיר היא חח"ע ועל. כלומר, שלכל מילה מאורך $n-2$ מעל $\{1,\ldots,n-2\}$ קיים עץ שמייצר אותה (זה יוכיח שהפונקציה היא על) ושהעץ הזה הוא יחיד (זה יוכיח שהפונקציה היא חח"ע).

טענה 8.13 אם w היא המילה שמתקבלת כפלט של TreeToWord על $G=(V,E)$, אז לכל $v\in V$ מספר המופעים של v ב־w הוא בדיוק $d(v)-1$.

הוכחה: עבור v יש שתי אפשרויות: או שהוא אחד משני הצמתים שנשארים בסיום ריצת האלגוריתם, או שהוא מוסר מהגרף בשלב מסוים. האלגוריתם מסיר מהגרף רק צומת שהוא עלה, כלומר בעל שכן בודד בעץ, ולכן בכל אחד משני המקרים v מגיע באלגוריתם למצב שבו יש לו שכן בודד, ומכאן ש־$d(v)-1$ משכניו מוסרים לפניו, ובכל הפעמים הללו v מתווסף למחרוזת. לאחר הסרת שכנים אלו v הופך בעצמו לעלה, ולכן המקרה היחיד שבו לא יוסר הוא אם גם שכנו הוא עלה, ובמקרה זה שני צמתים אלו הם האחרונים שנותרו בעץ. ■

מסקנה 9.13 אם v אינו מופיע ב־w, אז v הוא עלה בכל עץ שיוצר את w.

טענה 10.13 בהינתן מילה $w\in\{1,2,\ldots,n\}^{n-2}$, קיים ויחיד עץ T היוצר אותה באמצעות האלגוריתם TreeToWord.

הוכחה: נוכיח באינדוקציה על n.

בסיס האינדוקציה עבור $n=2$: במקרה זה w היא המחרוזת הריקה, ואכן קיים עץ יחיד בן שני הצמתים $\{1,2\}$.

צעד: נניח באינדוקציה שלמה כי לכל $n'<n$, ההתאמה בין עצים ומילים של Tree-ToWord היא אכן חח`ע ועל.

יהא $u\in\{1,2,\ldots,n\}$ המינימלי שאינו מופיע ב־w (קיים כזה שכן אורך w הוא $n-2$). מכיוון ש־u הוא הקטן מבין האיברים שאינם מופיעים ב־w, הוא בהכרח העלה המינימלי בכל עץ שיוצר את w. לכן בהכרח w_1 (האות הראשונה ב־w) הוא השכן של u בכל עץ שיוצר את w.

כעת `נקצוץ` את w_1 מ־w לקבלת $w'=w_2\ldots w_{n-2}$ ונסיר מ־V את u לקבלת $V'=V-\{u\}$.

מהנחת האינדוקציה נובע שקיים עץ יחיד $T'=(V',E')$ היוצר את המילה w'. מעץ זה מתקבל T על ידי הוספת הקשת (u,w_1) שראינו כי היא נקבעת באופן יחיד. מכאן ש־T נקבע באופן יחיד על ידי w, כנדרש. ■

הוכחה זו גם מתווה את האופן שבו יעבוד אלגוריתם פענוח שבונה עץ בהינתן מילה:

אלגוריתם WordToTree: קלט: מילה $w=\sigma_1\sigma_2\ldots\sigma_{n-2}$

פלט: עץ $G=(V,E)$

האלגוריתם:

1. אתחלו $E=\emptyset$, $S=V$.
2. עבור $i=1,2,\ldots,n-2$:
   (א) מצאו את u ־ הצומת המינימלי ב־S שאינו מופיע ב־$w_iw_{i+1}\ldots w_{n-2}$.
   (ב) $E\leftarrow E\cup\{(u,w_i)\}$.
   (ג) $S\leftarrow S-\{u\}$.
3. בשלב זה $S=\{u,v\}$. בצעו $E\leftarrow E\cup\{(u,v)\}$.

13.3 עצים מכוונים

נפתח בהגדרות הדרושות כדי להכליל את מושג העץ לגרפים מכוונים:

הגדרה 11.13 **שורש** בגרף מכוון הוא צומת שקיים מסלול ממנו אל כל הצמתים.
עץ מכוון הוא גרף מכוון אשר גרף התשתית שלו הוא עץ וקיים לו שורש.

כמקודם, ניתן לתת מספר אפיונים שקולים לכך שגרף הוא עץ מכוון:

משפט 12.13 יהא G גרף מכוון. הטענות הבאות שקולות:

1. G הוא עץ מכוון.
2. ל־G יש שורש ולכל צומת בגרף המסלול מהשורש אליו הוא יחיד.
3. ל־G יש שורש שדרגת הכניסה שלו 0 ולשאר הצמתים בגרף דרגת כניסה 1.
4. ל־G יש שורש ומחיקת כל קשת מ־G הופכת את G לחסר שורש.
5. גרף התשתית של G קשיר ול־G יש צומת אחד עם דרגת כניסה 0 ולשאר הצמתים דרגת כניסה 1.

הוכחה: נוכיח באמצעות שרשרת גרירות.

$1 \Leftarrow 2$: מכיוון ש־G הוא עץ מכוון קיים לו שורש v. נניח בשלילה כי קיים בגרף צומת u כך שיש שני מסלולים $v \overset{1}{\rightsquigarrow} u, v \overset{2}{\rightsquigarrow} u$, אז בגרף התשתית של G השרשור של שני המסלולים יוצר מעגל ולכן גרף התשתית אינו עץ, בסתירה לכך ש־G עץ מכוון.

$2 \Leftarrow 3$: אם דרגת הכניסה של השורש v אינה 0 זה אומר שיש צומת u כך שהקשת $u \to v$ בגרף; מכיוון שקיים מסלול מ־v אל u קיבלנו שיש שני מסלולים מ־v אל v בגרף ־ המסלול $v \rightsquigarrow u \to v$ והמסלול מאורך 0 שכולל רק את v, בסתירה. באופן דומה, אם ל־u כלשהו יש דרגת כניסה לפחות 2, אז יש u_1, u_2 כך שהקשתות $u_1 \to u$ ו־$u_2 \to u$ בגרף, וקיבלנו שני מסלולים מ־v אל u: המסלול $v \rightsquigarrow u_1 \to u$ והמסלול $v \rightsquigarrow u_2 \to u$. אם ל־$u$ יש לו דרגת כניסה 0 אז לא קיים מסלול אליו מהשורש, ולכן המסקנה היא שדרגת הכניסה של כל u שאינו השורש היא 1.

$3 \Leftarrow 4$: קיום שורש v מובטח מתנאי 3. תהא $u \to w$ קשת כלשהי בגרף. בהכרח המסלול מ־v אל w הוא מהצורה $v \rightsquigarrow u \to w$, כלומר הצעד האחרון שלו משתמש בקשת מ־u אל w, אחרת היינו מקבלים שקיימים שני מסלולים שונים מ־v אל w. לכן אם נמחקת הקשת $u \to w$ מהגרף אין מסלול מ־v אל w ובכך v מפסיק להיות שורש. מכיוון שדרגת הכניסה של v הייתה אפס הוא היה בהכרח השורש היחיד בגרף, ולכן הגרף הפך לחסר שורש.

$4 \Leftarrow 5$: מכיוון שיש ל־G שורש v גרף התשתית בהכרח קשיר (מסלול בין כל שני צמתים בגרף התשתית נבנה משרשור שני המסלולים שמחברים את השורש אליהם בגרף המקורי). ל־v יש דרגת כניסה אפס כי אם יש קשת $u \to v$ אז אפשר להסיר אותה מהגרף ו־v עדיין יישאר שורש. בדומה לכל צומת u אם יש שתי קשתות $u_1 \to u$ ו־$u_2 \to u$ אפשר להסיר אחת מהן ועדיין יוותר מסלול מ־v אל u ומכאן ש־v יוותר שורש (ואם ל־u דרגת כניסה 0 אז לא קיים מסלול אליו מהשורש).

$5 \Leftarrow 1$: כאן עלינו להראות כי גרף התשתית של G הוא חסר מעגלים וכי יש ל־G שורש. יהא v הצומת בעל דרגת הכניסה 0 בגרף. ראשית נוכיח כי v הוא שורש. יהא u צומת כלשהו בגרף, אז קיים מסלול $v \to w_1 \to \cdots \to w_k = u$ בגרף התשתית של G; נותר להראות

שב־G כל הקשתות מכוונות בכיוון המסלול. אם אין קשתות במסלול, סיימנו; בנוסף, הקשת (u, w_1) חייבת להיות מכוונת לכיוון w_1 אחרת דרגת הכניסה של v גדולה מאפס, ומכאן נמשיך באינדוקציה: הקשת (w_{i-1}, w_i) היא מ־w_{i-1} אל w_i על פי הנחת האינדוקציה, ולכן כדי שדרגת הכניסה של w_i תהיה בדיוק 1 הכרחי שהקשת (w_i, w_{i+1}) תהיה מכוונת לכיוון w_{i+1} וכך עד $w_k = u$.

נותר להראות כי גרף התשתית של G אינו כולל מעגלים. נניח בשלילה כי קיים מעגל; אז v אינו יכול להיות עליו (אחרת הייתה קשת נכנסת ל־v ודרגת הכניסה של v לא הייתה 0). נתבונן במסלול מ־v אל צומת u כלשהו על המעגל: $v = u_0 \to u_1 \to u_2 \to \cdots \to u_k = u$, ויהא u_i בעל האינדקס i המינימלי שנמצא על המעגל ($i \geq 1$ שכן ראינו כי v אינו יכול להיות על המעגל). אז מצד אחד נכנסת אליו הקשת מ־u_{i-1} שאינו על המעגל, ומצד שני נכנסת אליו קשת מצומת שכן על המעגל, ולכן דרגת הכניסה שלו היא לפחות 2, בסתירה לנתון. ■

ישנו עוד אפיון אחד שהשמטנו מהרשימה שכן הוא תקף רק עבור גרפים מכוונים **סופיים**:

טענה 13.13 גרף מכוון סופי G הוא עץ בעל שורש r אם ורק אם דרגת הכניסה של r היא 0, דרגת הכניסה של שאר הצמתים בגרף היא 1, וגרף התשתית של G חסר מעגלים.

הוכחה: כיוון אחד קל - אם G הוא עץ מכוון בעל שורש r אז גרף התשתית הוא עץ ולכן חסר מעגלים, וכבר ראינו כי היותו של G עץ גוררת את התכונה הנדרשת על הדרגות (אפיון 3 במשפט 12.13).

בכיוון השני עלינו להראות כי r הוא שורש (מה שגם יגרור את קשירות גרף התשתית). יהא u צומת כלשהו בגרף. נבנה סדרה $\ldots ,u_1, u_2$ באופן הבא: $u_1 = u$ ולכל i, u_{i+1} הוא הצומת שנכנס אל u_i. יש צומת כזה כל עוד $u_i \neq r$, כי דרגת הכניסה של u_i היא 1.

נניח כעת בשלילה כי קיימים $i < j$ כך ש־$u_i = u_j$. אז קיבלנו מעגל בגרף: $u_j \to u_{j-1} \to \cdots \to u_i = u_j$. מכאן שכל אברי הסדרה שונים אלו מאלו. מכיוון שניתן להמשיך את הסדרה כל עוד לא הגענו אל r ויש רק מספר סופי של צמתים בגרף שמספרם חוסם את אורך הסדרה, הסדרה חייבת להגיע אל r, מה שיראה קיום של מסלול מ־r אל u. ■

דוגמה נגדית פשוטה למשפט שלעיל במקרה שבו הגרף אינסופי היא גרף שמורכב מ`שרוך` אינסופי לשני הכיוונים $\ldots \to \bullet \to \bullet \to \bullet \to \cdots$, ועוד צומת מבודד (שישמש בתפקיד r).

13.4 עצים פורשים

13.4.1 הגדרה וקיום

הגדרה 14.13 עבור גרף $G = (V, E)$ (שיכול להיות מכוון או לא מכוון):

- **הגרף המושרה** על G על ידי קבוצת **צמתים** $V' \subseteq V$ הוא הגרף $G' = (V', E')$ בו E' מכילה את קשתות E אשר שני קצותיהן ב־V' ($E' = \{(u, v) \in E \mid u, v \in V'\}$).
- **הגרף המושרה** על G על ידי קבוצת **קשתות** $E' \subseteq E$ הוא $G' = (V', E')$ כאשר V' מכילה את כל הצמתים בהם נוגעות קשתות מ־E' ($V' = (v \in V \mid \exists u \in V : (u, v) \in E \vee (v, u) \in E)$).

הגדרות אלו שימושיות באופן כללי, אך אנו רוצים לדבר על סוג מיוחד של תת־גרף מושרה:

הגדרה 15.13 עץ פורש של גרף $G = (V, E)$ הוא עץ $T = (V, E')$ המושרה על ידי תת קבוצה $E' \subseteq E$.

כלומר, עץ פורש צריך לכלול את כל צמתי הגרף המקורי וחלק מהקשתות, כך שהוא יהיה עץ.

ברור כי לכל גרף לא מכוון קשיר קיים עץ פורש: פשוט מסירים קשתות מהגרף עד שמגיעים למצב שבו הסרת כל קשת הופכת את הגרף לבלתי קשיר, ואז הוא עץ על פי משפט 2.13. פחות ברור המקרה של עץ מכוון:

טענה 16.13 לכל גרף מכוון עם שורש r יש עץ פורש עם שורש r.

הוכחה: לכל צומת $v \in V$ נגדיר את $\text{dist}\,(v)$ להיות אורך המסלול הקצר ביותר מ־r אל v (המספר הזה תמיד מוגדר כי r הוא שורש).

לכל $v \neq r$, נתבונן במסלול מאורך $\text{dist}\,(v)$ מ־r אליו: $r \to \cdots \to u \to v$. נוסיף לעץ שאנו בונים את הקשת $u \to v$.

ברור כי בבניה זו דרגת הכניסה של r תהיה 0 ודרגת הכניסה של כל צומת אחר תהיה 1 (כי מוסיפים לגרף בדיוק קשת אחת שנכנסת אליו). נותר להראות שהגרף המושרה שבנינו קשיר. נניח בשלילה שהוא אינו קשיר, ויהי v צומת מינימלי ביחס ל־$\text{dist}\,(v)$ שאינו ישיג מ־r. אז בגרף המקורי v נמצא על מסלול $r \to \cdots \to u \to v$, ובלי הגבלת הכלליות ניתן להניח שזהו המסלול שבאמצעותו נבחרה הקשת שהוספנו לגרף עבור v. מכאן שהקשת $u \to v$ נמצאת בגרף המושרה שבנינו. כעת, $\text{dist}\,(u) < \text{dist}\,(v)$ (כי u נמצא לפני v במסלול הקצר ביותר שמוביל אל v) ומהמינימליות של v עולה ש־u ישיג מ־r בתת הגרף המושרה שבנינו, אבל מכך נובע שגם v ישיג. ■

13.4.2 ספירת עצים פורשים ־ משפט קירכהוף לגרפים מכוונים

בהינתן גרף מכוון G וצומת r, נשאלת השאלה כמה עצים פורשים יש ל־G עם שורש r. משפט קירכהוף מצביע על שיטה למציאת מספר זה באמצעות חישוב **דטרמיננטה** של מטריצה מיוחדת המותאמת לגרף.

הגדרה 17.13 בהינתן גרף מכוון G ללא חוגים עצמיים על צמתים $V = \{v_1, v_2, \ldots, v_n\}$, נגדיר מספר מטריצות ב־$\mathbb{R}^{n\times n}$ המתאימות לגרף:

- **מטריצת השכנויות** A של הגרף מוגדרת כך ש־$[A]_{ij}$ הוא מספר הקשתות מ־v_i אל v_j ב־G.
- **מטריצת דרגת הכניסה** Δ מוגדרת על ידי $[\Delta]_{ij} = \begin{cases} d_{in}\,(v_i) & i = j \\ 0 & i \neq j \end{cases}$ כלומר זו מטריצה אלכסונית שבה דרגות הכניסה של צמתי G מופיעים על האלכסון.
- **מטריצת הלפסליאן** L מוגדרת על ידי $\Delta - A$, כלומר
  - L_{ii} שווה לדרגת הכניסה של v_i, $d_{in}\,(v_i)$
  - L_{ij} עבור $i \neq j$ שווה למינוס מספר הקשתות ב־G מ־v_i אל v_j (אם מספר הקשתות הוא k_{ij} אז $L_{ij} = -k_{ij}$).

ההנחה שאין ב־G חוגים עצמיים אינה מגבילה אותנו, שכן עץ פורש ממילא אינו יכול להכיל קשת מצומת לעצמו (היא תיצור מעגל), ולכן בהינתן גרף כלשהו, מספר העצים הפורשים שלו זהה למספר העצים הפורשים של הגרף שמתקבל ממנו על ידי הסרת החוגים העצמיים.

עם זאת, ב־G בהחלט יכולות להיות קשתות מקבילות, ואנו סופרים בנפרד עצים פורשים שמשתמשים בקשתות שונות עבור אותם זוגות צמתים.

כדי לנסח את משפט קירכהוף אנו זקוקים להגדרה נוספת מאלגברה לינארית:

הגדרה 18.13 מטריצת המינור L_r של L היא המטריצה המתקבלת מ־L על ידי מחיקת השורה והעמודה ה־r־ים.

וכעת ניתן לעבור לניסוח המשפט:

משפט 19.13 (קירכהוף) יהא G גרף מכוון עם מטריצת לפלסיאן L. מספר העצים הפורשים את G עם שורש v_r הוא בדיוק $\det\left(L_r\right)$.

המקרה הפשוט של המשפט הוא זה שבו מספר הקשתות בגרף קטן מ־$n-1$. במקרה זה אין מספיק קשתות ב־G כדי ליצור עץ פורש (שכן כל עץ פורש הוא בעל $n-1$ קשתות) ולכן מספר העצים הפורשים הוא 0. קל לראות כי $\det\left(L_r\right)=0$ במקרה זה; מכיוון שיש ב־G לכל היותר $n-2$ קשתות, דרגת הכניסה של שני צמתים לפחות היא 0. זה אומר שיש צומת v_i כך ש־$i\neq r$ עבורו $d_{in}\left(v_i\right)=0$ ולכן $L_{ii}=0$ וכמו כן לכל $j\neq i$ מתקיים $L_{ji}=0$ (כי אין קשת מ־v_j שנכנסת אל v_i). כלומר, העמודה ה־i ב־L כולה אפסים, כך שגם ב־L_r יש עמודה שכולה אפסים, ולכן $\det\left(L_r\right)=0$.

נעבור למקרה שבו מספר הקשתות ב־G הוא לפחות $n-1$. לצורך הוכחת המשפט נתבסס על כך שאנו יכולים לכתוב את L_r כמכפלה של שתי מטריצות לא ריבועיות. ראשית, עבור $G=(V,E)$ נסמן ב־$n=|V|$ את מספר הצמתים וב־$m=|E|$ את מספר הקשתות.

כעת נגדיר מטריצות A,B מסדר $(n-1)\times m$ כך שכל שורה מתאימה לצומת של G פרט לצומת r (נניח בלי הגבלת הכלליות ש־$r=n$), על פי הכללים הבאים:

$A_{ik}=B_{ik}=-1$ אם הקשת e_k נכנסת לצומת v_i ב־G.

$A_{ik}=1$ אם הקשת e_k יוצאת מהצומת v_i ב־G.

שאר הכניסות של שתי המטריצות הן 0.

כלומר, כל עמודה של A,B מייצגת קשת בגרף, כך שבשורה שמתאימה לצומת שממנה יוצאת הקשת יש 1 ב־A ו־0 ב־B, וכמו כן בשורה שמתאימה לצומת שאליה נכנסת הקשת יש -1 בשתי המטריצות. נשים לב כי השורה שמתאימה לצומת r של השורש לא מופיעה במטריצות.

טענה 20.13 $L_r=AB^T$.

הוכחה: $L_{ii}=\sum_{k=1}^{m}A_{ik}B_{ki}^{T}=\sum_{k=1}^{m}A_{ki}B_{ki}=\sum_{j:e_k=v_j\to v_i}(-1)\cdot(-1)=$ $d_{in}\left(v_i\right)$.

כמו כן עבור $i\neq j$:

$L_{ij}=\sum_{k=1}^{m}A_{ik}B_{kj}^{T}=\sum_{k=1}^{m}A_{ik}B_{jk}=\sum_{e_k=v_i\to v_j}1\cdot(-1)$, כלומר מינוס מספר הקשתות שנכנסות מ־i אל j. ■

כזכור, באלגברה לינארית מוכיחים את המשפט $\det\left(A\cdot B\right)=\det A\det B$. לרוע המזל, המשפט נכון רק עבור מטריצות ריבועיות, ואילו A,B שלנו אינן ריבועיות; למרבה המזל קיימת הכללה של משפט זה.

משפט 21.13 (קושי־בינה): אם A,B מטריצות מסדר $n\times m$ בהתאמה (כך ש־AB^T היא מטריצה מסדר $n\times n$), וגם $n\leq m$, אז מתקיים $\det\left(AB^T\right)=\sum_{\sigma}\det A_{\sigma}\det B_{\sigma}^{T}$, כאשר σ רץ על כל הקבוצות של n אינדקסים מתוך $\{1,\ldots,m\}$, ו־A_{σ} מייצג את תת־המטריצה

מסדר $n \times n$ שמתקבלת מ־A על ידי מחיקת כל העמודות פרט לאלו שהאינדקסים שלהן ב־σ, ו־B_σ מוגדרת בדומה.

כזכור, כבר טיפלנו במקרה שבו $m < n-1$ (במצב זה $\det(L_r) = 0$) ולכן ניתן להניח כי $m \geq n-1$ ותנאי משפט קושי־בינה מתקיימים.

המשפט מקל עלינו את הוכחת משפט קירכהוף בצורה משמעותית, שכן ל־A_σ ו־B_σ יש משמעות קומבינטורית פשוטה: הן מייצגות את מה שנותר מ־G לאחר שנבחרה בו תת־קבוצה σ של קשתות שהן **מועמדות ליצור עץ**. שימו לב ש־A, B הן מסדר $(n-1) \times m$ ולכן σ הוא בחירה של $n-1$ עמודות (קשתות) מתוך m העמודות (קשתות) האפשריות. מכיוון שעל פי משפט קושי־בינה מתקיים

$$\det L_r = \sum_\sigma \det A_\sigma \det B_\sigma^T = \sum_\sigma \det A_\sigma \det B_\sigma$$

מה שנותר להראות כדי להוכיח ש־$\det L_r$ הוא מספר העצים הפורשים של G עם השורש v_r הוא שאם הקשתות שנבחרו ב־σ יוצרות עץ, המחובר שמתאים להן יהיה 1, ואם אינן יוצרות עץ, המחובר יהיה 0. פורמלית נראה:

1. אם σ מתאים לבחירה של $n-1$ קשתות שיוצרות עץ פורש שהשורש שלו הוא v_r, אז או ש־$\det A_\sigma = \det B_\sigma = 1$ או ש־$\det A_\sigma = \det B_\sigma = -1$ (ולכן $\det A_\sigma \det B_\sigma^T = 1$).

2. אם σ מתאים לבחירה של $n-1$ קשתות שאינן יוצרות עץ פורש, אז $\det A_\sigma = 0$ או $\det B_\sigma = 0$.

לבה הטכני של ההוכחה הוא בשתי התוצאות הבאות מאלגברה לינארית:

- החלפת שתי שורות או שתי עמודות של מטריצה הופכת את סימן הדטרמיננטה שלה אך לא משנה את ערכה המוחלט.

- הדטרמיננטה של מטריצה משולשית תחתונה היא מכפלת האיברים שעל האלכסון.

אם כן, נביא את A_σ, B_σ לצורה של מטריצה משולשית תחתונה על ידי סידור מחדש של השורות והעמודות.

אם σ מגדירה עץ פורש T ב־G, אז נבנה סדרה $u_1, u_2, \ldots$ של צמתים ו־$e_1, e_2, \ldots$ של קשתות באופן הבא: מכיוון ש־T הוא עץ, קיימים בו לפחות שני עלים, כלומר יש בו עלה שאינו v_r. עלה זה יהיה u_1. הקשת שמחברת את u_1 לשאר הגרף תהיה e_1. כעת נמחק את u_1 ו־e_1 מהעץ, ונקבל עץ חדש, שגם בו שני עלים שאחד מהם אינו v_r ומהם נבנה את u_2, e_2 וכן הלאה. הצומת האחרון שיוותר בגרף הוא v_r וממנו פשוט נתעלם.

כעת נסדר מחדש את A_σ, B_σ כך שהשורה הראשונה היא של הצומת u_1, העמודה הראשונה של e_1, השורה השניה של u_2 וכן הלאה.

נתבונן בשורה שמתאימה ל־u_i בכל אחת משתי המטריצות. אנו רוצים להראות שלכל $k > i$ מתקיים שהכניסה ה־ik שווה 0 כדי שהמטריצות יהיו משולשיות תחתונות. אם $u_{ik} \neq 0$ זה אומר שהקשת e_k מחוברת לצומת u_i (נכנסת או יוצאת ממנו) בעץ T, ושקשת זו הוסרה מהעץ רק אחרי הצעד שבו u_i הוסר מהעץ. אבל כאשר u_i מוסר מהעץ הוא עלה, ולכן e_i הייתה הקשת האחרונה שחיברה את u_i לעץ, ומכאן שלא ייתכן ש־e_k הייתה מחוברת אליו. כמו כן, e_i היא קשת שנכנסת ל־u_i, ולכן הכניסה ה־ii במטריצה היא -1 (גם ב־A_σ וגם ב־B_σ). מכאן שאכן נקבל $\det A_\sigma = \det B_\sigma = \pm 1$ במקרה זה.

נניח כעת כי σ אינה מגדירה עץ, ונראה כי $\det A_\sigma = 0$ או $\det B_\sigma = 0$, בהתאם לשני הדברים שיכולים להשתבש.

ראשית, ייתכן ש־σ אינה עץ אפילו בגרף התשתית של G. במקרה זה, מכיוון ש־σ היא קבוצה של $n-1$ קשתות, בהכרח יש בגרף שמושרה מ־σ שני רכיבי קשירות, שכן כל גרף לא מכוון קשיר עם $n-1$ קשתות הוא עץ. נתבונן ברכיב הקשירות שבו v_r אינו נמצא, ובאוסף השורות ב־A_σ שמתאים לצמתים של רכיב קשירות זה. אם נראה כי סכום שורות אלו הוא 0, סיימנו; זה מוכיח כי השורות תלויות לינארית ולכן A_σ היא סינגולרית ו־$\det A_\sigma = 0$. הסכום הוא אפס שכן לכל עמודה במטריצה, אם הקשת שהיא מייצגת לא שייכת לרכיב הקשירות אז אף אחד משני הצמתים שמחוברים אליה לא שייכים לקבוצת הצמתים של השורות שלקחנו, ואז העמודה תהיה שווה ל־0 בכל השורות בקבוצה שלנו; או שהקשת כן שייכת לרכיב הקשירות ואז שני הצמתים המחוברים יהיו בקבוצת השורות שלנו, ומכיוון שבצומת שממנו יוצאת הקשת הערך הוא 1 ובצומת שאליו נכנסת הקשת הערך הוא -1 שוב נקבל שהסכום הוא 0.

נותר לטפל במקרה שבו σ מגדירה עץ בגרף התשתית אך בגרף המכוון זה איננו עץ ששורשו v_r. במקרה זה נראה כי $\det B_\sigma = 0$. נסדר את המטריצה כפי שעשינו במקרה שבו σ כן הגדירה עץ. מכיוון ש־σ לא מגדירה עץ במקרה זה, בהכרח זה מכיוון שיש קשת אחת לפחות שמכוונת בכיוון הלא נכון, כלומר יש i כך ש־e_i **יוצאת** מהצומת v_i, ולכן הכניסה ii במטריצה המסודרת מחדש תהיה 0, ומכאן שנקבל שהדטרמיננטה היא 0. זה מסיים את ההוכחה.

13.4.3 ספירת עצים פורשים ־ משפט קירכהוף לגרפים לא מכוונים

משפט קירכהוף שהוכחנו קודם ניתן לניסוח גם עבור גרפים לא מכוונים; ההוכחה שלו מתבססת על רדוקציה למקרה של גרף מכוון.

נזכיר את המטריצות המעורבות, הפעם בהגדרה עבור גרפים לא מכוונים:

הגדרה 22.13 בהינתן גרף לא מכוון G ללא חוגים עצמיים על צמתים $V = \{v_1, v_2, \ldots, v_n\}$, נגדיר מספר מטריצות ב־$\mathbb{R}^{n\times n}$ המתאימות לגרף:

- **מטריצת השכנויות** A של הגרף מוגדרת כך ש־$[A]_{ij}$ הוא מספר הקשתות בין v_i ו־v_j ב־G.
- **מטריצת הדרגות** Δ מוגדרת על ידי $[\Delta]_{ij} = \begin{cases} d(v_i) & i=j \\ 0 & i\ne j \end{cases}$ כלומר זו מטריצה אלכסונית שבה הדרגות של צמתי G מופיעים על האלכסון.
- **מטריצת הלפסליאן** L מוגדרת על ידי $\Delta - A$, כלומר
  - L_{ii} שווה לדרגה של v_i, $d(v_i)$
  - L_{ij} עבור $j \ne i$ שווה למינוס מספר הקשתות ב־G בין v_i ו־v_j (אם מספר הקשתות הוא k_{ij} אז $L_{ij} = -k_{ij}$).

בניגוד למקרה המכוון, במקרה הלא מכוון מטריצת השכנויות A היא סימטרית, שכן מספר הקשתות בין v_i ו־v_j זהה, כמובן, למספר הקשתות בין v_j ו־v_i.

משפט 23.13 (קירכהוף) יהא G גרף לא מכוון עם מטריצת לפלסיאן L ויהא $v \in V$ צומת כלשהו ב־G. מספר העצים הפורשים את G הוא בדיוק $\det(L_v)$.

בניגוד למקרה של גרפים מכוונים, בגרף לא מכוון אין חשיבות לבחירה של $v \in V$; כפי שנראה, עבור כל $v \in V$ תתקבל אותה התוצאה. **הוכחה:** כמו במקרה המכוון, ניתן להניח שאין ב־G חוגים עצמיים שכן הם ממילא לא יכולים להשתתף באף עץ פורש.

מהגרף הלא מכוון $G = (V, E)$ נבנה גרף מכוון $G' = (V, E')$ באופן הבא: ב־G' אותם צמתים כמו G, ואם ב־G קיימת הקשת (v, u) אז ב־G' יהיו קיימות שתי הקשתות $v \to u$ ו־$u \to v$ (אם יש יותר מקשת אחרת בין v, u בגרף G אז נוסיף שתי קשתות ל־G' עבור כל קשת שכזו).

מטריצת השכנויות של הגרף החדש היא $A' = A$ על פי הגדרה. בנוסף $\Delta' = \Delta$ שכן כל קשת שמחוברת לצומת ב־G הופכת לזוג קשתות, שאחת מהן נכנסת לצומת זו ב־G'. מכאן ש־$L' = L$, ולכן $\det(L'_v) = \det(L_v)$ לכל $v \in V$.

כדי להשלים את ההוכחה צריך לראות שלכל $v \in V$ יש התאמה חח"ע ועל בין עצים פורשים של G ועצים פורשים של G' עם שורש v. בכיוון אחד, בהינתן עץ פורש T ב־G נוסיף כיוונים לקשתות שלו כך שדרגת הכניסה של G תהיה 0 ודרגת הכניסה של כל צומת אחר תהיה 1. ניתן לעשות זאת באופן הבא: לכל קשת (u, w) כלשהי ב־T אנו יודעים שקיים מסלול **יחיד** $u \rightsquigarrow v$. אם w מופיע על המסלול, נכוון את הקשת (u, w) מ־w אל u; אחרת, נכוון אותה מ־u אל w. את הגרף שהתקבל נסמן ב־T'. אנו יודעים שגרף התשתית T שלו הוא עץ, ולכן כדי להשתכנע ש־T' הוא עץ מכוון די להראות ש־v הוא שורש. בהינתן $u \in V$, נתבונן במסלול $u \rightsquigarrow v$ בעץ T ונראה שהוא מסלול גם ב־T'. לשם כך נתבונן במעבר כלשהו $w_1 \to w_2$ על המסלול ב־T ונראה שהקשת (w_1, w_2) מכוונת מ־w_1 אל w_2 גם ב־T'; דבר זה נובע מכך שהמסלול היחיד מ־v אל w_2 ב־T חייב לעבור קודם ב־w_1 (אחרת גם המסלול מ־v אל u לא היה עובר קודם ב־w_1).

הראינו התאמה שלכל עץ פורש T של G מחזירה עץ פורש T' עם השורש v של G'. בכיוון ההפוך, בהינתן עץ פורש T', נחזיר את גרף התשתית שלו T שהוא בוודאי עץ על פי הגדרת עץ פורש בגרף מכוון. התאמה זו היא ההופכית של ההתאמה הקודמת שהראינו, מה שמוכיח שההתאמה היא חח"ע ועל. ■

13.5 למת האינסוף של קניג

13.5.1 תיאור הלמה

נעבור כעת לדון בתכונה של גרפים מכוונים אינסופיים.

משפט 24.13 (למת האינסוף של קניג) יהא G גרף מכוון אינסופי עם שורש r כך שלכל צומת v, $d_{out}(v) < \infty$. אז קיים ב־G מסלול אינסופי שמתחיל ב־r.

לא קשה לראות את הכרחיות התנאי על סופיות הדרגות (למשל, גרף `קיפוד` שבו שורש אחד עם אינסוף שכנים ותו לא הוא דוגמה נגדית, כי כל המסלולים בו הם מאורך 1). **הוכחה:** על פי טענה 16.13 קיים ל־G עץ פורש עם שורש r. העץ אינסופי ודרגת היציאה של כל צומת בו היא סופית.

נבנה מסלול באופן האינדוקטיבי הבא: $v_0 = r$ ולכל i, אם כבר נבנה v_i אז ייבחר v_{i+1} להיות אחד מבניו של v_i שלו אינסוף צאצאים. התכונה שנשמרת באינדוקציה היא של־v_i יש אינסוף צאצאים, מה שמתקיים עבור r, ואם ל־v_i אינסוף צאצאים אז יש לו בן שגם לו אינסוף צאצאים כי מספר הצאצאים של v_i הוא סכום צאצאי בניו.

מכיוון שה־v_i נבחרו בתוך העץ הפורש באופן שיוצר מסלול, לא ייתכן שנחזור פעמיים לאותו v_i שכן זה יסגור מעגל. מכאן שהמסלול אינסופי, כנדרש. ■

13.5.2 דוגמת שימוש - ריצופי Wang

אריח *Wang* הוא ריבוע שצלעותיו צבועות בצבעים כלשהם (לרוב כל צלע מסומנת במספר או אות במקום צבע כדי להקל על הסימון). ריצוף של המישור באמצעות אריחי Wang פירושו כיסוי כל המישור על ידי אריחים הצמודים זה לזה אך אינם עולים זה על זה, כך שכל זוג צלעות סמוכות של אריחים שונים הוא בעל אותו צבע.

בעית ההכרעה הקלאסית של אריחי Wang היא זו: בהינתן קבוצה סופית של אריחים, כך שמכל אריח בקבוצה זמינים לנו אינסוף עותקים שלו, האם ניתן לרצף את המישור או לא באמצעות אריחים אלו?

ניתן להוכיח כי בעיה זו אינה כריעה אלגוריתמית; לא קיים אלגוריתם הקובע לכל קבוצה האם היא מרצפת את המישור או לא. מצד שני, ישנו קריטריון שמאפשר להקל על ההוכחה שקיים ריצוף במקרים מסויימים:

משפט 25.13 (וואנג) בהינתן קבוצת אריחים A קיים ריצוף של המישור באמצעות A אם ורק אם לכל n טבעי קיים ריצוף של ריבוע $n \times n$ באמצעות A.

הוכחה: כיוון אחד קל: אם A מרצפת את המישור, אז לכל n נתבונן על ריבוע בגודל $n \times n$ כלשהו במישור; הוא עצמו יהיה מרוצף באופן חוקי ולכן קיבלנו ריצוף שלו באמצעות A.

בכיוון השני, נגדיר גרף באופן הבא: צמתי הגרף יהיו ריצופים חוקיים באמצעות A של ריבועים בגודל $n \times n$ לכל n טבעי אי זוגי ($1, 3, 5, \dots$). על פי ההנחה, יש אינסוף ריצופים שכאלו (לפחות אחד לכל n) ולכן הגרף אינסופי.

נגדיר קשתות בגרף באופן הבא: יש קשת מהצומת u אל הצומת v אם ורק אם u הוא ריצוף של ריבוע בגודל $2n-1$, v הוא ריצוף של ריבוע בגודל $2n+1$, ו־u מתקבל מ־v על ידי `קילוף` הטבעת החיצונית ביותר.

כמו כן נוסיף לגרף צומת r ונוציא קשת ממנו לכל צומת u שמייצג ריצוף בגודל 1.

בבירור r הוא שורש של הגרף שהתקבל: בהינתן u, `נקלף` אותו שכבה שכבה עד להגעה אל r.

בנוסף, דרגת היציאה של כל צומת בגרף היא סופית - לכל ריצוף יש רק מספר סופי של טבעות שבהן אפשר להקיף אותו.

מלמת האינסוף של קניג נובע כעת כי קיים בגרף מסלול אינסופי. הצמתים במסלול זה מייצגים ריצוף הולך ומתרחב של המישור (כל צומת `מסכים` עם הצמתים שקדמו לו על המשבצות שכבר רוצפו), וריצוף של המישור כולו מתקבל בתור הגבול של סדרה זו (לכל ריבוע במישור, האריח שיונח בריבוע זה מופיע בכל אברי הסדרה החל ממקום מסויים). ■

מסקנה 26.13 קיים ריצוף של המישור בעזרת A אם ורק אם קיים ריצוף של רבע המישור בעזרת A.

חלק III
נושאים מתקדמים

14 מספרי קטלן

נעבור כעת לתאר סדרת מספרים שמופיעה בבעיות ספירה טבעיות רבות, ובפרט שתיים הקשורות לעצים.

1. כמה מסלולים יש ב־$\mathbb{Z}^2$ מ־$(0,0)$ אל (n,n) כאשר הצעדים המותרים הם ימינה ולמעלה והמסלול אף פעם לא מגיע אל מתחת לאלכסון הראשי $x=y$?
2. כמה סדרות סוגריים חוקיות קיימות עם n פותחים ו־n סוגרים? (סדרת סוגריים היא **חוקית** אם כשקוראים אותה משמאל לימין בשום שלב אין יותר סוגריים סוגרים מפותחים, ובסיום מספרם שווה).
3. כמה עצים מכוונים בינאריים יש עם n צמתים?
4. כמה עצים מכוונים בינאריים מלאים יש עם $n+1$ עלים? (בעץ בינארי מלא, לכל צומת שאינו עלה יש שני בנים בדיוק).
5. כמה עצים מכוונים סדורים יש עם $n+1$ צמתים? (בעץ סדור יש חשיבות לסדר הבנים של כל צומת)
6. בכמה דרכים ניתן לחלק מצולע בן $n+2$ צלעות למשולשים?

הפתרון לכל הבעיות הללו הוא C_n - מספר קטלן ה־n־י.

14.1 מסלולי שריג

נפתח בפתרון 1 שיאפשר לנו למצוא נוסחה מפורשת ל־C_n.

מספר מסלולי השריג הכוללים מ־$(0,0)$ אל (n,n) שכוללים צעדים ימינה ולמעלה ואין עליהם מגבלות אחרות הוא $\binom{2n}{n}$ - בוחרים את n הצעדים שבהם נעלה למעלה, ובשאר הצעדים הולכים ימינה.

מסלול `רע` הוא כזה שיורד מתחת לאלכסון $x=y$. נראה כי מספר המסלולים הרעים הוא כמספר המסלולים הכולל מ־$(1,-1)$ אל (n,n), כלומר $\binom{2n}{n-1}$.

כל מסלול רע חייב לפגוש מתישהו את האלכסון המשני $y=x-1$ שכן בהתחלה המסלול מקיים $x=y$ ובכל צעד משנים את אחת הקואורדינטות ב־1.

נסמן ב־p את נקודת המפגש הראשונה של המסלול הרע עם $y=x-1$. כעת נשקף את המסלול בקטע שבין $(0,0)$ אל p ביחס לאלכסון $y=x-1$ (שיקוף שכזה פירושו שהמסלול מתחיל מ־$(1,-1)$, עולה בכל פעם שבה המסלול המקורי הולך ימינה, והולך ימינה בכל פעם שבה המסלול המקורי עולה). קל לראות פורמלית כי גם המסלול המשוקף מגיע אל p, ולכן אפשר לשרשר לו את המשך המסלול הרגיל.

קל לראות כי ההתאמה שתיארנו היא חד חד ערכית, שכן ניתן להפוך אותה על ידי ביצוע שיקוף חוזר. כמו כן נשים לב שההתאמה הפיכה **לכל** מסלול מ־$(1,-1)$ אל (n,n) שכן מסלול שכזה חייב לגעת מתישהו ב־$y=x-1$ (כי הוא מתחיל מתחת לאלכסון זה וצריך לסיים מעליו), ולכן קיבלנו התאמה חח`ע ועל כמבוקש.

קיבלנו כי $C_{\mathrm{n}}=\binom{2n}{n}-\binom{2n}{n-1}$. כדי לפשט את הביטוי נשים לב לכך ש־

$$\begin{aligned}\binom{2n}{n-1} &= \frac{(2n)!}{(n-1)!\,(n+1)!}=\frac{(2n)!}{n!n!}\cdot\frac{n!}{(n-1)!}\cdot\frac{n!}{(n+1)!}\\ &= \binom{2n}{n}\cdot\frac{n}{n+1}\end{aligned}$$

ומכאן:

$$\begin{aligned} C_n &= \binom{2n}{n} - \binom{2n}{n-1} \\ &= \binom{2n}{n}\left(1 - \frac{n}{n+1}\right) \\ &= \frac{1}{n+1}\binom{2n}{n} \end{aligned}$$

זהו ביטוי מפורש למספר קטלן C_n.

14.2 סוגריים מאוזנים ומילות דיק

סוגריים מאוזנים הם מקרה פרטי של מושג כללי יותר:

הגדרה 1.14 מילת דיק מאורך $2n$ מעל הסימנים $\{a,b\}$ היא מילה $w \in \{a,b\}^{2n}$ עם בדיוק n מופעים של a ו־n מופעים של b שהיא בעלת התכונה שבכל רישא של המילה, מספר ה־a־ים גדול או שווה למספר ה־b־ים. פורמלית מסמנים זאת באופן הבא: לכל פירוק $w = uv$ מתקיים ש־$\#_b(u) \geq \#_a(u)$, וכמו כן מתקיים $\#_a(w) = \#_b(w)$.

סדרת סוגריים חוקית היא מילת דיק שבה הסימנים הם $($, $)$ בהתאמה. לצורך נוחות הסימונים נעבור כעת להשתמש בסימנים U, R.

ראשית נשים לב לשקילות הברורה שבין מסלולי שריג וסדרות סוגריים חוקיות - U מציין צעד עלייה למעלה, R מציין צעד ימינה, והתנאי על האלכסון שקול לתנאי על מילות דיק. לכן C_n הוא מספרן של מילות הדיק מאורך $2n$.

כעת נשתמש בדרך ההצגה באמצעות מילות דיק כדי לפתח ביטוי רקורסיבי למספרי קטלן.

בסיס: $C_0 = 1$
צעד: $C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i}$

הנוסחה נובעת מהאבחנה הבאה: לכל מילת דיק לא ריקה w קיימת הצגה יחידה מהצורה $w = UxRy$ כך ש־x, y הן מילות דיק, אולי ריקות. מכאן שמספר ה־w־ים מאורך $2(n+1)$ הוא כמספרם של כל הזוגות x, y של מילות דיק שמקיימות $|x| + |y| = 2n$.

מכאן מתקבלת הנוסחה: $C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i}$ הוא עיקרון החיבור בפעולה, כשמפרידים בין מקרים שונים לפי אורך x (שהוא במקרה הכללי $2i$). בהינתן בחירה של מילת דיק באורך $2i$ עבור x (C_i אפשרויות) אנו בוחרים מילת דיק y מאורך $2(n-i)$ (C_{n-i} אפשרויות) ומשתמשים בעיקרון הכפל לקבלת מספר כל הזוגות של (x,y) שכאלו.

נוכיח את הטענה על קיום הפירוק היחיד $w = UxRy$. ראשית, האות הראשונה ב־w חייבת להיות U כי אם היא תהיה R זה יפר את התנאי שמגדיר מילות דיק.

נתבונן על הפירוקים האפשריים של w לרישא וסיפא, $w = uy$, וניקח פירוק שמקיים $\#_R(u) = \#_U(u)$ כך שאורך u הוא מינימלי מבין כל הפירוקים שמקיימים זאת. נשים לב לכך שבהכרח קיים פירוק אחד לפחות כזה, שכן $u = w$ מקיים זאת על פי הגדרת מילות דיק.

נסמן $u = u'\sigma$. לפי התנאי על מילות דיק $\#_R(u') \leq \#_U(u')$. לכן בהכרח $\sigma = R$ (אחרת היינו מקבלים $\#_R(u) < \#_U(u)$) ולכן ניתן לכתוב $u = UxR$. כעת נראה כי x היא מילת דיק: לכל רישא x' של x מתקיים

$$\#_R(x') = \#_R(Ux') < \#_U(Ux') = \#_U(x') + 1$$

כאשר אי השוויון הוא חזק בזכות המינימליות של u. מכך ש־$\#_R(x') < \#_U(x') + 1$ והעובדה שכל המספרים במשוואה שלמים, נסיק $\#_R(x') \le \#_U(x')$.

כמו כן, $\#_R(x) = \#_R(UxR) - 1 = \#_U(UxR) - 1 = \#_U(x)$

ולכן x מקיימת את התכונה שמספר ה־U וה־R־ים בה מתאזן בסוף. מכאן ש־x היא מילת דיק.

כעת נעבור ל־y. כדי לראות שהיא מילת דיק נסתמך כמקודם על התכונה שמגדירה מילות דיק והעובדה ש־u ו־$w = uy$ הן מילת דיק:

$$\#_R(y) = \#_R(uy) - \#_R(u) = \#_U(uy) - \#_U(u) = \#_U(y)$$

ולכל רישא y' של y, נסתמך על כך ש־uy' היא רישא של w ונקבל:

$$\#_R(y') = \#_R(uy') - \#_R(u) \le \#_U(uy') - \#_U(u) = \#_U(y')$$

כמבוקש.

14.3 עצים בינאריים

הגדרה 2.14 עץ בינארי הוא עץ מכוון שבו דרגת היציאה של כל צומת היא לכל היותר 2.
עץ בינארי מלא הוא עץ בינארי שבו דרגת היציאה של כל צומת היא 0 או 2.

אם יש קשת מצומת u אל צומת v אנו אומרים ש־v הוא **בן** של u. בן יכול להיות אחד מהשניים: **בן ימני** או **בן שמאלי** (אך לא שניהם), ולכן לכל צומת יש ארבע אפשרויות: או שיש לו בן ימני ובן שמאלי, או שיש לו רק בן ימני, או שיש לו רק בן שמאלי, או שאין לו בנים כלל. אנו מבדילים בין עצים שונים על פי הבנים שלהם: למשל, עץ שמורכב משני צמתים שאחד מהם הוא בן ימני של השני ייחשב **שונה** מעץ בעל שני צמתים שאחד מהם הוא בן שמאלי של השני.

כמה עצים בינאריים בעלי n צמתים (לא מסומנים) קיימים?

כאן נוח להשתמש בתיאור רקורסיבי של עצים בינאריים:

טענה 3.14 (עצים בינאריים, הגדרה רקורסיבית)

- הגרף הריק הוא עץ בינארי.
- אם T_1, T_2 הם עצים בינאריים אז גרף מצומת r וקשתות אל T_1, T_2 הוא עץ בינארי (`קשת אל $T$` פירושה קשת אל השורש של T אם T לא ריק, ואחרת אין קשת)

נסמן ב־B_n את מספר העצים הבינאריים על n צמתים. מהתיאור הרקורסיבי שלעיל נובעת הנוסחה הרקורסיבית הבאה:

$B_0 = 1$

$B_{n+1} = \sum_{i=0}^{n} B_i B_{n-i}$ - סכימה על הבחירות האפשריות של עצים בתור בנים ל־r; סכום הצמתים בשני העצים הוא n כי הצומת הנוסף משמש בתפקיד r.

נוסחה זו זהה לחלוטין לנוסחה של מספרי קטלן, דהיינו $B_n = C_n$.

נעבור כעת לטפל באופן דומה בעצים בינאריים מלאים. ההבדל העיקרי הוא שאנחנו סופרים את מספר העצים הבינאריים המלאים בעלי n **עלים**, ולא צמתים. נסמן מספר זה ב־D_n ונוכיח שמתקיים $D_{n+1} = C_n$, כלומר מספר קטלן ה־n־י הוא מספר העצים הבינאריים המלאים בעלי $n+1$ עלים.

טענה 4.14 (עצים בינאריים מלאים, הגדרה רקורסיבית)

- גרף בעל צומת בודד הוא עץ בינארי מלא.
- אם T_1, T_2 הם עצים בינאריים מלאים לא ריקים אז גרף שמורכב מצומת r וקשתות אל T_1, T_2 הוא עץ בינארי מלא.

בעץ בעל צומת אחד יש גם עלה בודד. בעץ שנבנה מתוך T_1 ו־T_2 מספר העלים הוא סכום מספרי העלים של T_1, T_2. זה מוביל לנוסחה הבאה:

$D_1 = 1$

$D_{n+1} = \sum_{i=1}^{n} D_i D_{n+1-i}$ - דומה לנוסחה הקודמת, אבל כעת עץ אינו יכול להכיל 0 עלים.

נוכיח ש־$C_n = D_{n+1}$ באינדוקציה על n.

בסיס: $C_0 = 1 = D_1$

צעד: נניח $C_k = D_{k+1}$ לכל $k \leq n$ ונוכיח עבור $n+1$ תוך שניעזר בהחלפת המשתנה $i = k+1$ (ולכן $k = i-1$):

$$\begin{aligned} C_{n+1} &= \sum_{k=0}^{n} C_k C_{n-k} = \sum_{k=0}^{n} D_{k+1} D_{n+1-k} \\ &= \sum_{i=1}^{n+1} D_i D_{n+2-i} = D_{n+2} \end{aligned}$$

כלומר, מספר העצים הבינאריים המלאים בעלי $n+1$ עלים הוא בדיוק מספר קטלן ה־n.

נעבור כעת לעצים **לא בינאריים** אבל שעדיין יש בהם חשיבות לסדר הבנים. עצים כאלו נקראים **עצים סדורים**. נסמן ב־E_n את מספר העצים הבינאריים הסדורים בעלי n צמתים.

טענה 5.14 $E_{n+1} = C_n$, כלומר מספר קטלן ה־n הוא מספר העצים הסדורים בעלי $n+1$ צמתים.

הוכחה: להבדיל מהמקרים הקודמים בהם נעזרנו בהגדרה האינדוקטיבית של מספרי קטלן, כאן נראה התאמה חח"ע ועל ישירה בין עצים בעלי $n+1$ צמתים ומילות דיק מאורך $2n$ מעל $\{a, b\}$.

בהינתן עץ מכוון כלשהו $T = (V, E)$ עם שורש r נגדיר פונקציה $f : V \to \{a, b\}^*$ באופן הרקורסיבי הבא: לכל צומת v, נסמן את בניו בתור $u_1, \ldots, u_k$ על פי הסדר שלהם (אם v הוא עלה סדרת הבנים תהיה ריקה). ניתן להניח כי f כבר הוגדרה רקורסיבית על בנים אלו.

- אם $r \neq v$ נגדיר $f(v) = af(u_1)f(u_2)\cdots f(u_k)b$
- אם $v = r$ נגדיר $f(v) = f(u_1)f(u_2)\cdots f(u_k)$ (אותה הגדרה, אך ללא ה־a בהתחלה וה־b בסוף)

בבירור f מוגדרת היטב וקל לבדוק שהיא מחזירה מילת דיק על פי בדיקה ישירה של התנאים שמגדירים מילת דיק. כעת, לכל עץ T בעל $n+1$ צמתים נתאים את המילה $f(r)$. יש להוכיח מילה זו היא מאורך $2n$; כדי לעשות זאת מוכיחים באינדוקציה כי לכל $v \neq r$ מתקיים ש־$f(v)$ היא מאורך 2 כפול גודל תת־העץ ששורשו v (כולל v עצמו).

נותר להראות כי ההתאמה הפיכה, כלומר לכל מילת דיק מאורך $2n$ עלינו להראות עץ סדור בעל $n+1$ צמתים שמייצר מילה זו. ההוכחה היא באינדוקציה על n. עבור $n=0$, העץ בעל צומת בודד מייצר את המילה הריקה. נניח כי לכל $k<n$, לכל מילת דיק מאורך $2k$ קיים עץ בעל $k+1$ צמתים שמייצר את המילה הזו. נוכיח את הטענה עבור מילת דיק מאורך $2(n+1)$ עבור $n \geq 0$ כלשהו.

כפי שראינו קודם, מילת דיק מאורך גדול מ־0 היא בעלת הצגה יחידה בצורה $axby$ כאשר x, y מילות דיק. אז $|x|+|y| = 2(n+1)-2 = 2n$. נסמן $|x| = 2i$, אז $|y| = 2(n-i)$. ניתן להשתמש בהנחת האינדוקציה הן עבור x והן עבור y ולקבל שקיימים עבורם עצים בעלי $i+1$ ו־$n+1-i$ צמתים בהתאמה שהאלגוריתם מייצר עליהם את המילים x, y. נסמן עצים אלו ב־T_x, T_y.

כעת נבנה עץ חדש T_w באופן הבא: T_w יכלול צומת r שישמש כשורש העץ, ואליו נחבר את השורש של העץ T_x (כשאל שורש זה מחובר יתר העץ), וכמו כן נחבר אליו על פי הסדר שלהם את כל **הבנים** של השורש של T_y (את השורש של T_y עצמו לא נחבר).

בעץ החדש שקיבלנו מספר הצמתים הוא כמספר צמתי T_x ועוד מספר צמתי T_y למעט אחד (השורש) ועוד אחד (השורש החדש). בסך הכל יש בעץ $n+2 = (i+1)+(n+1-i)$ צמתים, בהתאם לכך ש־w היא מאורך $2(n+1)$.

כדי לראות כי האלגוריתם מחזיר את w על T_w, נסמן את שורש T_w ב־r, את שורשח T_x, T_y ב־r_x, r_y ואת בניו של r_y ב־$v_1, \ldots, v_k$.

אנו יודעים ש־$x = f(T_x) = f(r_x)$ כאשר r_x הוא שורש, ולכן נסיק שאם r_x **איננו** שורש בעץ כלשהו, אז $f(r_x) = axy$.

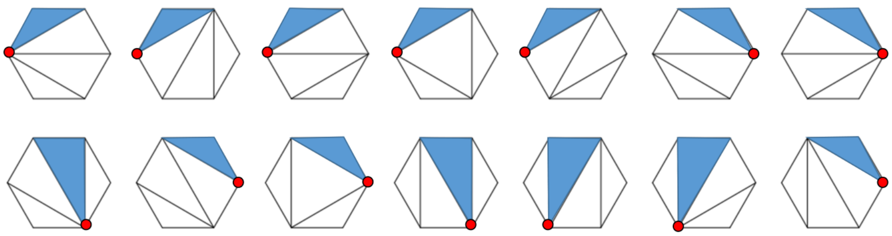
כמו כן אנו יודעים ש־$y = f(T_y) = f(r_y) = f(v_1)\ldots f(v_k)$. כעת נסיק:

$$
\begin{aligned}
f(T_w) &= f(r) = f(r_x)f(v_1)\ldots f(v_k) \\
&= axbf(v_1)\ldots f(v_k) \\
&= axby
\end{aligned}
$$

כמבוקש. ■

14.4 שילושים של מצולע קמור

בהינתן מצולע קמור, **שילוש** שלו הוא אוסף של אלכסונים שמועברים בתוכו מבלי לחתוך אלו את אלו ומחלקים אותו למשולשים.



נתון מצולע קמור בן n צלעות. כמה חלוקות אפשריות שלו למשולשים יש?

נמספר את קודקודי המצולע ב־$1,2,3,\ldots,n$ לפי סדר הופעתם על המצולע. נתבונן על הצלע שמחברת את הקודקודים $1,n$: אנחנו יודעים שבכל חלוקה של המצולע למשולשים, הצלע הזו היא חלק ממשולש. כל אחד מ־$n-2$ הקודקודים האחרים של המצולע יכול לשמש בתור הקודקוד השלישי של המשולש. אם בחרנו את הקודקוד i ואנו מוחקים את המשולש מהמצולע, מתקבלים שני מצולעים קמורים חדשים: המצולע שמורכב מהקודקודים $1,2,\ldots,i$ והצלע $(1,i)$, והמצולע שמורכב מהקודקודים $i,i+1,\ldots,n$ והצלע (i,n). אפשר כעת לטפל רקורסיבית בכל אחד מהמקרים, אולם נשים לב לכך שאם $i=2$ או $i=n-1$, אחד משני המצולעים החדשים שנקבל הוא "מנוון"; הוא כולל רק שני קודקודים. בסיטואציה כזו, של מצולע בעל 2 קודוקדים, אנו מגדירים את מספר השילושים שלו להיות 1 ("השילוש הריק").

אם נסמן ב־T_n את מספר השילושים של מצולע קמור בעל n צלעות, אז הדיון לעיל מראה כי מתקיימת נוסחת הנסיגה

$T_2=1$

$T_n=\sum_{i=2}^{n-1}T_iT_{n+1-i}$

כאשר i רץ על הבחירות האפשריות לקודקוד השלישי שאליו מחברים את $1,n$, מה שמוביל לשילוש של מצולע בעל i קודקודים ($1,2,\ldots i$) ומצולע בעל $n+1-i$ קודקודים ($i,i+1,\ldots,n$).

קל לראות כעת באינדוקציה כי $C_n=T_{n+2}$, שכן

בסיס: $C_0=1=T_2$

צעד: נניח $C_k=T_{k+2}$ לכל $k\le n$ ונוכיח עבור $n+1$, תוך שניעזר בהצבה $i=k+2$ (ולכן $k=i-2$):

$$\begin{aligned}C_{n+1}&=\sum_{k=0}^{n}C_kC_{n-k}=\sum_{k=0}^{n}T_{k+2}T_{(n-k)+2}=\\&=\sum_{i=2}^{n+2}T_iT_{(n-i+2)+2}=\sum_{i=2}^{n+2}T_iT_{(n+3)+1-i}\\&=T_{n+3}\end{aligned}$$

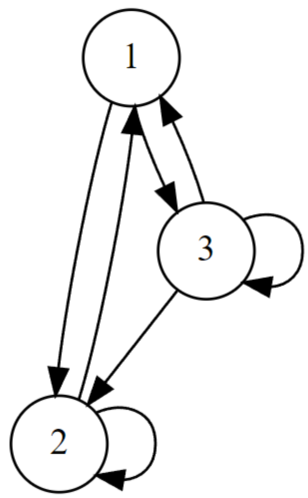
כמבוקש.

15 ספירת מסלולים בגרף

15.1 מבוא

בעיות ספירה רבות ניתנות לרדוקציה אל הבעיה של **ספירת מסלולים בגרף**. נציג כאן דוגמא ספציפית אחת. נניח שאנו מעוניינים לדעת מה מספרן של המילים $w \in \{1,2,3\}^n$ שהרצף 11 והרצף 23 לא מופיע בהן. פתרון הבעיה בעזרת כלים אלמנטריים או בעזרת עקרון ההכלה וההפרדה לא נראה מבטיח; נוסחת נסיגה נראית מתבקשת יותר. מה שנציג כאן הוא דרך שיטתית למצוא נוסחת נסיגה שכזו.

נבנה גרף שצמתיו מסומנים ב־$1,2,3$ כך שיש קשת מ־i אל j אם לרצף ij מותר להופיע במחרוזת:



מסלול מאורך $n-1$ בגרף מתאר מילה חוקית מאורך n. נשים לב לכך שמסלול כזה יכול להתחיל בכל אחד מהצמתים, ולכן עלינו לפתור את השאלה: בהינתן צומת v_i בגרף, כמה מסלולים מאורך n שמתחילים מ־v_i קיימים? לצורך כך, נזכיר מושג שראינו קודם:

הגדרה 1.15 בהינתן גרף לא מכוון G ללא חוגים עצמיים על צמתים $V=\{v_1,v_2,\dots,v_m\}$ **מטריצת השכנויות** $A\in\mathbb{R}^{m\times m}$ של הגרף מוגדרת כך ש־$[A]_{ij}$ הוא מספר הקשתות בין v_i ו־v_j ב־G.

עבור הגרף שבאיור, מטריצת השכנויות היא

$$A=\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

אפשר לחשוב על המטריצה הזו כמייצגת את מספר המסלולים מאורך 1 בין צמתי הגרף: מסלול מאורך 1 כולל קשת בודדת, ולכן מספר הקשתות בין v_i ו־v_j שווה למספר המסלולים ביניהם.

אם נכפול את A בעצמה (נזכיר בהמשך את ההגדרה הפורמלית של כפל מטריצות) נקבל

$$A^2 = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix}$$

נתבונן על ה־3 בכניסה ה־$(3,2)$ של המטריצה; הוא סופר את שלושת המסלולים האפשריים מאורך 2 מ־v_3 אל v_2:

$v_3 \to v_3 \to v_2$
$v_3 \to v_2 \to v_2$
$v_3 \to v_1 \to v_2$

באופן דומה, כל כניסה במטריצה סופרת את המסלולים מאורך 2. הרעיון כעת ברור: $[A^n]_{ij}$ סופר את המסלולים מאורך n מ־v_i אל v_j. נותר רק להוכיח זאת.

15.2 הוכחת הטענה הכללית

ניגש להוכחת המשפט המרכזי שלנו:

משפט 2.15 אם $G = (V,E)$ הוא גרף (מכוון או לא מכוון) עם קבוצת צמתים $V = \{v_1,\dots,v_m\}$ ו־A מטריצת השכנויות של הגרף, אז $[A^n]_{ij}$ הוא מספר המסלולים מאורך n מ־v_i אל v_j בגרף.

מכיוון שהמשפט והוכחתו מתבססים על חזקות של מטריצות, נזכיר את ההגדרה:

הגדרה 3.15 אם A מטריצה מסדר $n \times t$ ו־B מטריצה מסדר $t \times m$ אז AB היא מטריצה מסדר $n \times m$ המוגדרת על ידי

$$[AB]_{ij} = \sum_{k=1}^{t} [A]_{ik} [B]_{kj}$$

אנו מגדירים חזקות של מטריצות באופן רקורסיבי באמצעות כפל: $A^0 = I$ ו־$A^{n+1} = A^n \cdot A$, כאשר I היא מטריצת היחידה: $[I]_{ij} = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$.

ההגדרה של כפל מטריצות אינה קלה לעיכול, אולם דווקא התוצאה שנוכיח כעת היא אחת מהדרכים לקבל אינטואיציה עבורה. לכן נעבור כעת להוכחת המשפט. **הוכחה:** ההוכחה תהיה באינדוקציה על n.

בסיס: עבור $n = 0$, מכיוון שקיים מסלול יחיד מאורך 1 מצומת v_i לעצמו (המסלול v_i שכולל צומת יחיד ואפס קשתות) וקיימים 0 מסלולים מ־v_i אל v_j עבור $i \neq j$, נקבל על פי הגדרה ש־$A^0 = I$ אכן מתארת את מספר המסלולים המבוקש.

נניח את נכונות הטענה עבור n ונוכיח עבור $n+1$. יהיו $1 \le i,j \le m$ כלשהם: נוכיח ש־$\left[A^{n+1}\right]_{ij}$ הוא מספר המסלולים מאורך $n+1$ מ־v_i אל v_j.

בהינתן מסלול $v_i \rightsquigarrow v_j$ שהוא מאורך לפחות 1, ניתן לסמן ב־v_k את הצומת הלפני אחרון במסלול ולקבל את המסלול $v_i \rightsquigarrow v_k \to v_j$ כך שהמסלול $v_i \rightsquigarrow v_k$ הוא מאורך n. על פי הנחת האינדוקציה, מספר המסלולים מ־v_i אל v_k הוא $[A^n]_{ik}$, ועל פי הגדרת A

מספר הקשתות מ־v_k אל v_j הוא $[A]_{kj}$. כל מסלול מהצורה $v_i \rightsquigarrow v_k \to v_j$ מתקבל על ידי בחירה בלתי תלויה של מסלול $v_i \rightsquigarrow v_k$ וקשת $v_k \to v_j$ ולכן מעיקרון הכפל מספר המסלולים הכולל מהצורה $v_i \rightsquigarrow v_k \to v_j$ הוא $[A^n]_{ik}[A]_{kj}$.

כעת, על פי עיקרון החיבור, מספר המסלולים הכולל $v_i \rightsquigarrow v_j$ מאורך $n+1$ הוא סכום על מספר המסלולים מהצורה $v_i \rightsquigarrow v_k \to v_j$ לכל הערכים האפשריים של $1 \le k \le m$. כלומר מספר זה הוא

$$\sum_{k=1}^{m}[A^n]_{ik}[A]_{kj} = \left[A^{n+1}\right]_{ij}$$

על פי הגדרת כפל מטריצות, מה שמסיים את ההוכחה. ■

אם כן, כאשר אנו מתבוננים על הביטוי $[AB]_{ij} = \sum_{k=1}^{t}[A]_{ik}[B]_{kj}$ שמגדיר כפל מטריצות אנו רואים כי ניתן לתת לו משמעות קומבינטורית: הסכום $\sum_{k=1}^{t}$ הוא **עיקרון החיבור** בפעולה, והמכפלה $[A]_{it}[B]_{tj}$ היא **עיקרון הכפל** בפעולה.

15.3 שימוש: פונקציות יוצרות ונוסחאות נסיגה

בפני עצמה, התוצאה שראינו עד כה לא מקדמת אותנו יותר מדי ־ כפל מטריצות הוא פעולה יקרה יחסית ואנו מעוניינים בייצוג טוב יותר לסדרת המספרים שאנחנו מחפשים מאשר הייצוג המובלע ב־A. למרבה המזל, קיימת טכניקה כללית שבהינתן A מוצאת את הפונקציה היוצרת $f_{ij}(x)$ של מספר המסלולים מ־v_i אל v_j מכל אורך n.

נסתמך ללא הוכחה על משפט מאלגברה לינארית:

משפט 4.15 אם B היא מטריצה הפיכה, אז $\left[B^{-1}\right]_{ij} = (-1)^{i+j}\frac{\det B_{ji}}{\det B}$ כאשר B_{ji} היא המטריצה המתקבלת מ־B על ידי מחיקת השורה ה־j והעמודה ה־i.

נעבור כעת למשפט שברצוננו להוכיח. ראשית, נציג את האינטואיציה המובילה אליו. במקום להסתכל על המטריצה A נתבונן על המטריצה xA, שבה כל הכניסות של A מוכפלות ב־x. נקבל ש־$\left[(xA)^n\right]_{ij} = [A^n]_{ij}x^n$ הוא מספר המסלולים מאורך n מ־v_i אל v_j כשמספר זה מוכפל ב־x^n. כלומר, מתקיים

$f_{ij}(x) = \sum_{n=0}^{\infty}\left[(xA)^n\right]_{ij}$

נתבונן במטריצה F המוגדרת על ידי $[F]_{ij} = f_{ij}(x)$, אז מתקיים

$F = \sum_{n=0}^{\infty}(xA)^n$

כזכור, ראינו כבר כי ניתן להוכיח את השוויון הפורמלי $\sum_{n=0}^{\infty}x^n = \frac{1}{1-x}$. "קפיצת האמונה" שעלינו לבצע כאן נובעת מכך שנעשה כעת את אותו הדבר עבור **מטריצה**. דהיינו

$$\sum_{n=0}^{\infty}(xA)^n = (I - xA)^{-1}$$

ההוכחה זהה לחלוטין לזו שכבר ראינו: נכפול את $\sum_{n=0}^{\infty}(xA)^n$ ב־$(I - xA)$, ונקבל טור טלסקופי אינסופי שבו כל האיברים מתבטלים מלבד ה־I בהתחלה.

מכאן נסיק שמתקיים $F = (I - xA)^{-1}$, ומהמשפט מאלגברה לינארית שציטטנו קודם, נקבל את התוצאה:

$$
\begin{aligned}
f_{ij}(x) = [F]_{ij} &= \left[(I-xA)^{-1}\right]_{ij} \\
&= (-1)^{i+j}\,\frac{\det(I-xA)_{ji}}{\det(I-xA)}
\end{aligned}
$$

כאשר, כזכור, $(I-xA)_{ji}$ הוא המטריצה $I-xA$ לאחר שנמחקו ממנה השורה ה־j והעמודה ה־i.

קיבלנו שהפונקציה היוצרת של מספר המסלולים מ־v_i אל v_j היא **פונקציה רציונלית** במשתנה x. יותר מכך: המכנה $\det(I-xA)$ אינו תלוי ב־v_i, v_j והוא זהה לכל זוג צמתים בגרף. כזכור, המכנה של פונקציה יוצרת רציונלית מגדיר נוסחת נסיגה עבור סדרת המספרים שהיא מייצגת; המונה קובע את תנאי ההתחלה.

חישוב בפועל של הביטוי $(-1)^{i+j}\frac{\det(I-xA)_{ji}}{\det(I-xA)}$ ומציאת נוסחאות הנסיגה ממנו אינו קל לביצוע ידנית, אך למרבה המזל ניתן לביצוע בקלות באמצעות מחשב וספריית מתמטיקה התומכת בחישוב סימבולי; זה הופך את פתרון הבעיה הקומבינטורית כולה לבעיה של מציאת הפונקציה A שמתארת את הבעיה.

15.4 חזרה אל הדוגמא

בדוגמא הקונקרטית שלנו המטריצה הרלוונטית הייתה

$$
A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}
$$

ואנו מתעניינים במספר המסלולים **הכולל**, מכל צומת לכל צומת.

ראשית נחשב את $I-xA$:

$$
\begin{aligned}
I - xA &= \begin{pmatrix} 1 & -x & -x \\ -x & 1-x & 0 \\ -x & -x & 1-x \end{pmatrix} \\
\det(I-xA) &= -x\begin{vmatrix} -x & 1-x \\ -x & -x \end{vmatrix} + (1-x)\begin{vmatrix} 1 & -x \\ -x & 1-x \end{vmatrix} \\
&= -x\left(x^2 + x - x^2\right) + (1-x)\left(1-x-x^2\right) \\
&= -x^2 + \left(1-x-x^2\right) - \left(x - x^2 - x^3\right) \\
&= 1 - 2x - x^2 + x^3
\end{aligned}
$$

חישוב הדטרמיננטה של 9 המינורים של המטריצה הוא מהיר יחסית כי אלו דטרמיננטות של מטריצות 2×2. מקבלים:

$$
(I-xA)^{-1} = \frac{1}{1-2x-x^2+x^3}\begin{pmatrix} (1-x)^2 & x & x(1-x) \\ x(1-x) & 1-x-x^2 & x^2 \\ x & x(1+x) & 1-x-x^2 \end{pmatrix}
$$

הפונקציה היוצרת המבוקשת שלנו היא $\sum_{ij} f(x)$, כלומר המכנה שלה הוא $1-2x-x^2+x^3$ והמונה שלה הוא סכום כל אברי המטריצה; חישוב מראה שהסכום הזה הוא $3+x-x^2$, כך שקיבלנו את הפונקציה היוצרת

$$f(x)=\frac{3+x-x^2}{1-2x-x^2+x^3}$$

(כל החישוב הנ"ל ניתן לביצוע באמצעות מחשב).

עבור נוסחת הנסיגה, חישוב המונה אינו באמת הכרחי; די במכנה כדי לקבל את נוסחת הנסיגה $a_n=2a_{n-1}+a_{n-2}-a_{n-3}$ ולחשב באופן מפורש שלושה תנאי התחלה על ידי ספירת כל האפשרויות:

- $a_0=1$ (המחרוזת הריקה)
- $a_1=3$ (המחרוזות $1,2,3$)
- $a_2=7$ (כל המחרוזות מאורך 9 למעט $11,23$).

הטכניקה שראינו ל"חילוץ" האיברים הראשונים בסדרה מתוך המונה והמכנה ניתנת למימוש באמצעות מחשב, ותניב בצורה אוטומטית את אותם תנאי התחלה, כאשר לוקחים בחשבון שמלכתחילה עסקנו בבעיה עם היסט של 1: בנינו גרף כך ש־a_n שווה למספר המסלולים מאורך $n-1$ בו, ולכן הפונקציה היוצרת מתאימה לסדרה $3,7,16\dots$.

16 נוסחת נסיגה עבור פונקציית החלוקה

16.1 הגדרות

בפרק 7 ראינו דוגמאות לבעיות שונות ומשונות של חלוקת כדורים לתאים. הבעיה המאתגרת ביותר שעמדה בפנינו הייתה בעיית החלוקה של n כדורים זהים למספר כלשהו של תאים ריקים כך שאין תא ריק. סימנו מספר זה ב־$p(n)$ ואנו מכנים את $p(n)$ בשם **פונקציית החלוקה**.

דרך אחרת לחשוב על $p(n)$, שהיא הדרך המקובלת יותר, היא בתור הפונקציה שסופרת **חלוקות של מספרים טבעיים**. כלומר, n סופרת את מספר הדרכים שבהם ניתן להציג את n בתור סכום של מספרים טבעיים כך שאין חשיבות לסדר המחוברים. למשל, עבור $n=5$ קיימות החלוקות הבאות:

1. $5=1+1+1+1+1$
2. $5=1+1+1+2$
3. $5=1+2+2$
4. $5=1+1+3$
5. $5=1+4$
6. $5=2+3$

כלומר, במקרה זה $p(5) = 7$, אבל אין דרך שיטתית ברורה למצוא מספר זה. הקושי אינו מקרי - לא קיימת נוסחה סגורה עבור $p(n)$, והבנת ההתנהגות האסימפטוטית של $p(n)$ היא בעיה חשובה בקומבינטוריקה. עם זאת, קיימת **נוסחת נסיגה** עבור $p(n)$ שמאפשרת לפשט מאוד את חישוב $p(n)$ בפועל ביחס לגישה שמייצרת במפורש את כל החלוקות של n וסופרת אותן. נוסחת הנסיגה הזו שונה באופיה מאלו שראינו עד כה בקורס, שכן אנו ראינו רק נוסחאות נסיגה ש"הולכות מספר צעדים קבוע אחורה" בעוד נוסחת הנסיגה עבור $p(n)$ חוזרת אחורה עד ההתחלה. ליתר דיוק, היא מחברת ומחסרת ערכים מסויימים של $p(k)$ עבור $k < n$.

ראשית, נציג את התוצאה באופן מפורש ומדויק, ואז נעבור לשאלה כיצד ניתן להוכיח אותה - הוכחת שתיעזר גם בפונקציות יוצרות וגם בפתרון "ישיר" של בעיה קומבינטורית.

נוסחת הנסיגה עבור $p(n)$ היא מהצורה

$p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + \dots$

החוקיות של ההליכה לאחור של הסדרה נעוצה בסדרת מספרים שנקראת **מספרים מחומשים** ומוגדרת כך:

$t(k) = \frac{k(3k-1)}{2}$

כאשר כאן k הוא **מספר שלם** שונה מאפס - הוא יכול להיות גם מספר חיובי וגם מספר שלילי.

חישוב ישיר של אברי הסדרה עבור $k = 1, -1, 2, -2, 3, -3, \dots$ יניב את $1, 2, 5, 7, 12, 15, \dots$ וכן הלאה: אלו ה"צעדים אחורה" שנוסחת הנסיגה מבצעת. השאלה האם האיברים יחוברו או יחוסרו תלויה בזוגיות של k: אם k הוא אי זוגי האיברים יופיעו עם סימן שלילי ואחרת עם סימן חיובי, כלומר אנו כופלים את האיברים ב-$(-1)^{k+1}$.

16.2 פונקציה יוצרת עבור בעיית החלוקה

נתבונן בפונקציה היוצרת של סדרת החלוקות,

$$f(x) = \sum_{n=0}^{\infty} p(n) x^n = 1 + x + 2x^2 + 3x^3 + 5x^4 + \dots$$

עבור פונקציה זו נוכיח את הנוסחה הבאה:

$\sum_{n=0}^{\infty} p(n) x^n = \prod_{n=1}^{\infty} \frac{1}{1-x^n}$

הביטוי באגף ימין כולל מכפלה אינסופית, שהיא מושג שטרם הגדרנו באופן פורמלי ולכן נעשה זאת כעת. כזכור, כאשר הגדרנו מכפלה של **שני** טורי חזקות פורמליים, הסתמכנו על כך שלמרות האינסופיות של שניהם, המקדם של כל איבר במכפלה נקבע על ידי התבוננות על מספר **סופי** של איברים:

$c_n = \sum_{k=0}^{n} a_k b_{n-k}$

לרוע המזל, באופן כללי דבר כזה לא חייב לקרות כשמסתכלים על מכפלה אינסופית. למשל, נתבונן על המכפלה הבאה:

$(1+x)(1+x)(1+x)\cdots$

כדי לקבל את x^2 בוחרים שני זוגות סוגריים ומהם לוקחים x, ולוקחים 1 מכל היתר. יש אינסוף זוגות סוגריים, ולכן אנו בוחרים 2 מתוך אינסוף - מכאן שלא נוכל לקבל מקדם סופי עבור x^2. כדי שהגדרה של מכפלה תעבוד בצורה שאנו מעוניינים בה, צריך שלכל n, יהיה רק מספר **סופי** של איברים במכפלה שכוללים חזקות של x מהצורה x^k כאשר $1 \le k \le n$

(עבור $k=0$ אנו מקבלים את האיבר 1 שאין בו בעיה כי הכפלה בו לא משנה דבר ולכן הוא בעצם הדרך שלנו לגרום לזוג סוגריים מסוים "לא להשתתף").

במקרה שלנו זה בדיוק מה שקורה: כזכור, ראינו בקורס ש $\frac{1}{1-x^n}=1+x^n+x^{2n}+x^{3n}+\ldots$

ולכן החזקה x^m תוכל להיווצר רק על ידי כפל של איברים שונים מ־1 מתוך זוגות סוגריים המתאימים ל־$\frac{1}{1-x^n}$ כך ש־$n\le m$, ויש רק מספר סופי של כאלו. לכן המכפלה $\prod_{n=1}^{\infty}\frac{1}{1-x^n}$ מוגדרת היטב, אך עדיין צריך להראות שהיא שווה אל $\sum_{n=0}^{\infty}p(n)x^n$.

יהא $n\ge 0$ כלשהו. אנו רוצים להראות כי המקדם של x^n בפיתוח של $\prod_{k=1}^{\infty}\frac{1}{1-x^k}$ הוא בדיוק $p(n)$. לשם כך נביט על אותו חלק של $\prod_{k=1}^{\infty}\frac{1}{1-x^k}$ שכולל איברים שיכולים לתרום למכפלה גורם שונה מ־1:

$\prod_{k=1}^{n}\frac{1}{1-x^k}=\left(1+x+x^2+\ldots\right)\cdots\left(1+x^k+x^{2k}+\ldots\right)$

כאן יש לנו n זוגות סוגריים, כאשר בזוג הסוגריים ה־i־י נמצא הטור $1+x^i+x^{2i}+\ldots$

האיבר x^n מתקבל פעם אחת בדיוק מכל דרך שבה ניתן לבחור איבר מכל זוג סוגריים, כך שסכום החזקות הכולל של האיברים שנבחרו הוא n. כל דרך כזו מתאימה לחלוקה אחת של n: אם האיבר i מופיע בסכום בדיוק t פעמים, אנו בוחרים את האיבר x^{ti} מתוך זוג הסוגריים ה־i.

למשל, עבור x^5, נסתכל על המכפלה

$$\left(1+x+x^2+\ldots\right)\left(1+x^2+x^4+\ldots\right)\left(1+x^3+\ldots\right)\left(1+x^4+\ldots\right)\left(1+x^5+\ldots\right)$$

החלוקה $5=3+1+1$ מתקבלת על ידי בחירת x^2 מזוג הסוגריים **הראשון**, x^3 מהשלישי, ו־1 מהיתר.

החלוקה $5=3+2$ מתקבלת על ידי בחירת x^2 מזוג הסוגריים **השני**, x^3 מהשלישי ו־1 מהיתר.

החלוקה $5=5$ מתקבלת על ידי בחירת x^5 מזוג הסוגריים **החמישי** ו־1 מהיתר.

באופן זה כל חלוקה מתקבלת בדיוק פעם אחת, וכל בחירת איברים מהסוגריים שיוצרת את x^5 מתאימה לאחת מהחלוקות, ולכן המקדם של x^n בפתיחת הסוגריים אכן שווה אל $p(n)$.

16.3 קבלת נוסחת הנסיגה מהפונקציה היוצרת

ראינו את המבנה של הפונקציה היוצרת:

$$\sum_{n=0}^{\infty}p(n)x^n=\prod_{n=1}^{\infty}\frac{1}{1-x^n}$$

בהמשך נראה משפט, הנקרא **משפט המספרים המחומשים**, שיראה לנו כי מתקיים

$$\prod_{n=1}^{\infty}\left(1-x^n\right)=\sum_{k=-\infty}^{\infty}\left(-1\right)^k x^{t(k)}$$

כאשר $t(k)=\frac{k(3k-1)}{2}$
אם ניקח את ההופכי של שני אגפי המשוואה, נקבל

$$\prod_{n=1}^{\infty}\frac{1}{1-x^n}=\frac{1}{\sum_{k=-\infty}^{\infty}(-1)^k x^{t(k)}}=\frac{1}{1-x-x^2+x^5+x^7-\ldots}$$

מבחינה פורמלית, מה שזה מוכיח הוא

$$f(x)\left(1-x-x^2+x^5+x^7-\ldots\right)=1$$

כזכור, במשפט 2.9 ראינו כי אם $f(x)=\frac{p(x)}{1-c_1x-\ldots-c_kx^k}$ כאשר $f(x)$ היא הפונקציה היוצרת של הסדרה $\{a_n\}_{n=0}^{\infty}$ אז הסדרה מקיימת את נוסחת הנסיגה $a_n=c_1a_{n-1}+\ldots+c_ka_{n-k}$. אותה הוכחה עובדת גם במקרה שבו המכנה אינו פולינום אלא טור חזקות, ובמקרה זה נוסחת הנסיגה תלך אחורה מספר בלתי חסום של צעדים (לכל n, הנוסחה תוכל ללכת אחורה עד a_0). כלומר, נקבל במקרה הזה
$p(n)=p(n-1)+p(n-2)-p(n-5)-p(n-7)+\ldots$ כפי שאמרנו שנקבל.

16.4 משפט המספרים המחומשים

סיימנו עם החלק בהוכחה שעוסק בפונקציות יוצרות, ומה שנותר לנו כעת הוא להוכיח כי מתקיים

$$\prod_{n=1}^{\infty}(1-x^n)=\sum_{k=-\infty}^{\infty}(-1)^k x^{t(k)}$$

כאשר $t(k)=\frac{k(3k-1)}{2}$.
על פניו נראה שאנחנו עדיין עוסקים בפונקציות יוצרות, אך בפועל יש לביטויים המופיעים בשוויון משמעות קומבינטורית פשוטה, שתאפשר לנו להוכיח את הזהות הזו באופן **קומבינטורי**.
כל ביטוי מהצורה a_nx^n שמתקבל מפתיחת הביטוי $\prod_{k=1}^{\infty}\left(1-x^k\right)$ מתקבל על ידי סכום של איברים שנבחרים באופן הבא:

- בוחרים מספרים טבעיים **שונים זה מזה**, $n_1<n_2<\ldots<n_r$ כך ש־$n=n_1+n_2+\ldots+n_r$.
- מכל סוגריים מהצורה $\left(1-x^{n_i}\right)$ בוחרים את $-x^{n_i}$
- מכל יתר זוגות הסוגריים בוחרים את 1

בצורה הזו, האיבר המתקבל מפתיחת הסוגריים מורכב ממכפלה של אינסוף 1־ים ובנוסף לכך המכפלה $\left(-x^{n_1}\right)\left(-x^{n_2}\right)\ldots\left(-x^{n_r}\right)=(-1)^r x^{n_1+\ldots+n_r}=(-1)^r x^n$
נשים לב שיש רק מספר **סופי** של דרכים לכתוב את t בתור סכום של טבעיים שונים זה מזה (מספר זה חסום מלמעלה על ידי $p(n)$, שהוא מספר הדרכים הכולל לכתוב את n כסכום של טבעיים, לאו דווקא שונים אלו מאלו). לכן אנחנו מקבלים סכום **סופי** של איברים

מהצורה $(-1)^r x^n$. אם r אי זוגי, אז האיבר שמתווסף לסכום הוא x^n ואם r הוא זוגי אז האיבר שמתווסף לסכום הוא $-x^n$, מה שמוביל אותנו אל הסימון הבא:

בהינתן n טבעי, נסמן ב־$q(n)$ את מספר הדרכים הכולל לכתוב את n כסכום של טבעיים שונים זה מזה. נסמן ב־$q_{even}(n)$ את מספר החלוקות הללו שבהן מספר המחוברים זוגי, וב־$q_{odd}(n)$ את מספר החלוקות שבהן מספר המחוברים אי זוגי. בפרט, $q(n) = q_{even}(n) + q_{odd}(n)$.

הדיון שקיימנו קודם על האיברים שמתקבלים מפתיחת $\prod_{k=1}^{\infty}\left(1-x^k\right)$ הראה לנו שהמקדם של x^n יהיה שווה אל $q_{even}(n) - q_{odd}(n)$ שכן כל חלוקה של n למספר זוגי של מחוברים תורמת את x^n לסכום וכל חלוקה למספר אי זוגי של מחוברים תורמת את $q_{odd}(n)$ לסכום.

במילים אחרות, ראינו שמתקיימת המשוואה

$$\prod_{k=1}^{\infty}\left(1-x^k\right) = \sum_{n=0}^{\infty}\left[q_{even}(n) - q_{odd}(n)\right]x^n$$

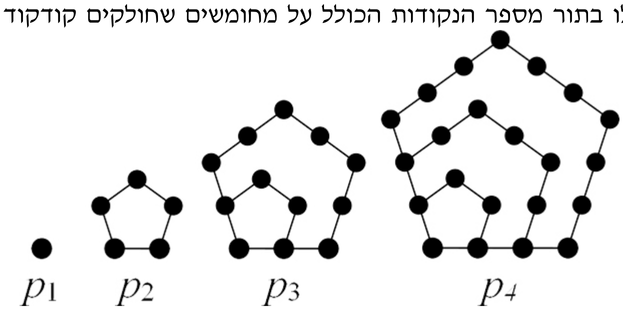
משפט המספרים המחומשים עוסק במספר $q_{even}(n) - q_{odd}(n)$. הוא אומר כי כמעט לכל המספרים מתקיים $q_{even}(n) = q_{odd}(n)$, והמספרים היחידים שעבורם הסיטואציה שונה הם מספרים מהצורה $t(k) = \frac{k(3k-1)}{2}$, ועבורם ההפרש הוא ± 1. פורמלית:

משפט 1.16 (משפט המספרים המחומשים) לכל n טבעי מתקיים

$$q_{even}(n) - q_{odd}(n) = \begin{cases} 1 & n = \frac{k(3k-1)}{2},\ k \text{ even} \\ -1 & n = \frac{k(3k-1)}{2},\ k \text{ odd} \\ 0 & \text{else} \end{cases}$$

ובניסוח קומפקטי, $\prod_{n=1}^{\infty}\left(1-x^n\right) = \sum_{k=-\infty}^{\infty}(-1)^k x^{\frac{k(3k-1)}{2}}$

המשפט נקרא על שם המספרים מהצורה $\frac{k(3k-1)}{2}$ שמופיעים בו. כאשר $k \ge 1$ מספרים אלו נקראים **מספרים מחומשים**, כאשר השם מגיע מהאופן שבו ניתן להגדיר את המספרים הללו בתור מספר הנקודות הכולל על מחומשים שחולקים קודקוד משותף:



p_1 p_2 p_3 p_4

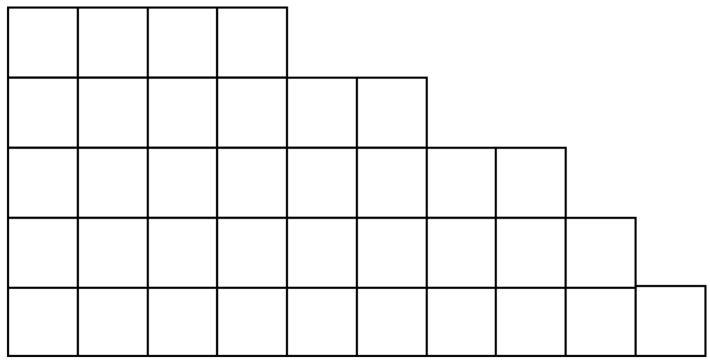
לא נזדקק לדרך הצגה זו בהמשך אז לא נעסוק בה. כאשר מציבים גם ערכים שליליים בנוסחה $\frac{k(3k-1)}{2}$, סדרת המספרים המתקבלת נקראת **מספרים מחומשים מוכללים**.

16.5 טבלאות יאנג והוכחת משפט המספרים המחומשים

אנו רוצים להוכיח את משפט המספרים המחומשים, כלומר לנתח את ערכו של $q_{even}(n) - q_{odd}(n)$. ההוכחה שנציג היא של המתמטיקאי האמריקאי פרנקלין מסוף המאה ה־19,

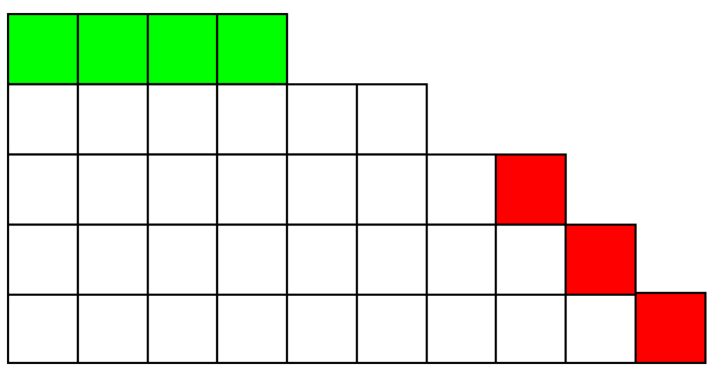
והיתרון שלה הוא בויזואליות הרבה שלה שמקלה על הבנה של מה שעומד מאחורי המשפט.

ההוכחה של פרנקלין נעזרת ב**טבלאות יאנג**, שהן דרך ציורית להציג חלוקות. חלוקה של n מוצגת בתור מעין ערימה של n ריבועים המסודרים בשורות עם נקודת התחלה משותפת, כך שכל שורה ארוכה ממש מזו שמעליה. כל שורה מתאימה לאחד מהמחוברים בחלוקה של n. כך למשל החלוקה $37 = 4 + 6 + 8 + 9 + 10$ תוצג באמצעות טבלת יאנג הבאה:



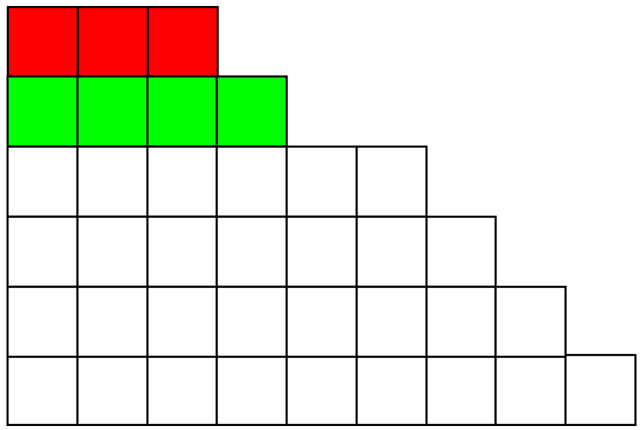
לצורך הוכחת המשפט, ננסה לבנות **פונקציה חח"ע ועל** בין קבוצת החלוקות עם מספר זוגי של מחוברים וקבוצת החלוקות עם מספר אי זוגי של מחוברים. ויזואלית באמצעות טבלאות יאנג, נדגים תהליך שלוקח טבלה אחת וממיר אותה בטבלה אחרת עם אותו מספר ריבועים ועם שורה אחת נוספת או שורה אחת פחות (כלומר, עם זוגיות שונה של מספר השורות). התהליך הזה יהיה ההופכי של עצמו, מה שיבטיח שהוא חח"ע ועל, למעט עבור לכל היותר מקרה קצה אחד שיצוץ רק כאשר n הוא מספר מחומש.

בהינתן טבלת יאנג, נצבע בירוק את כל אברי השורה העליונה (הקצרה ביותר) ובאדום את אברי **האלכסון** הימני ביותר, ונסמן ב־r את מספר הריבועים הירוקים וב־s את מספר הריבועים האדומים:

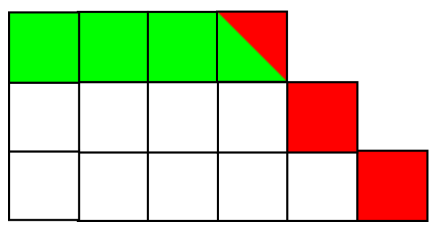


באופן פורמלי, אם נסמן את אורך השורות בטבלה ב־$(\lambda_1, \lambda_2, \ldots, \lambda_k)$ כך ש־$\lambda_1 > \lambda_2 > \ldots > \lambda_k$, אז נסמן $r = \lambda_k$ ו־s, אורך האלכסון, יהיה מספר השורות הרצופות בתחתית הטבלה כך שאורך כל אחת מהן הוא בדיוק 1 פחות אורך הקודמת. כלומר, הוא המספר הגדול ביותר עבורו מתקיים $(\lambda_1, \lambda_2, \ldots, \lambda_s) = (\lambda_1, \lambda_1 - 1, \ldots, \lambda_s - (s-1))$.

אם $s < r$, אפשר לקחת את הריבועים האדומים ולבנות מהם שורה עליונה חדשה. מכיוון ש־$s < r$ בדרך כלל יתקיים התנאי של טבלאות יאנג לפיו כל שורה קצרה יותר מקודמתה. הטבלה שבדוגמא תשתנה באופן הבא:



קיים מקרה קצה קריטי אחד שבו לא ניתן לבצע את הפעולה הזו: אם השורה העליונה משתתפת באלכסון, פירוש הדבר הוא שנקטין את הגודל שלה ב־1, וזה עלול להוביל לכך שלא נוכל להניח עליה את השורה החדשה. הנה דוגמא לאופן שבו זה מתרחש עבור החלוקה $15 = 4 + 5 + 6$:

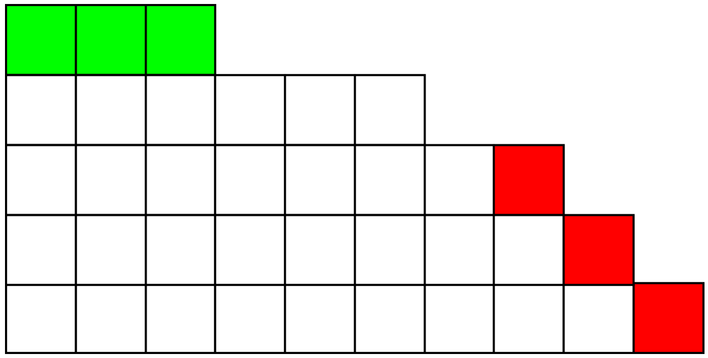


כאן $r = 4$ ו־$s = 3$, אבל אחרי סילוק הריבועים האדומים נקבל $r' = 3$ ולא ניתן להניח על שורה מאורך 3 עוד שורה מאותו האורך.

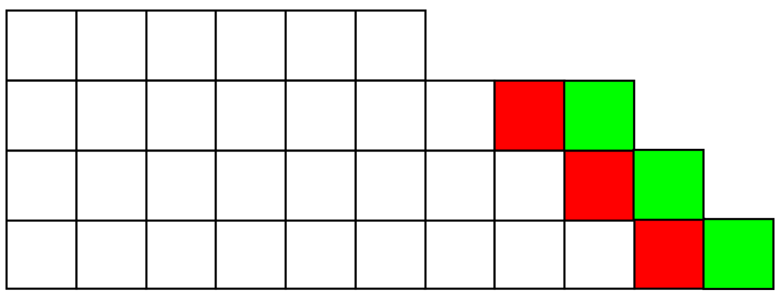
אם נסמן ב־k את מספר השורות הכולל, אפשר לראות שהתנאים לכך שהסיטואציה הבעייתית תתרחש הם ש־$s = k$ (ולכן יש ריבוע אדום גם בשורה העליונה ביותר) ו־$r = s+1$. נטפל במקרה זה בהמשך.

אם $s > r$ אפשר **להסיר** את השורה העליונה ולהוסיף איבר אחד ממנה לכל אחת מהשורות התחתונות: מכיוון ש־$s > r$, אנחנו יודעים שיש לנו מספיק שורות לחלק להן ריבועים; יש s שורות בסך הכל, ואפילו אם השורה העליונה שאנו מפרקים היא אחת מהן, עדיין נישאר עם $s - 1 \geq r$ שורות שאפשר להוסיף להם את r הריבועים הירוקים. אם נפעיל את הכלל הזה על הדוגמא שלעיל, נחזור לסיטואציה המקורית.

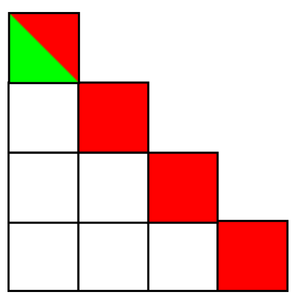
הסיטואציה שבה $r = s$ מורכבת יותר. נתבונן בדוגמא עבור הטבלה של $36 = 3 + 6 + 8 + 9 + 10$:



כאן $s = r$. בסיטואציה כזו לא ניתן להפוך את אברי האלכסון לשורה חדשה, כי שורה זו תהיה מאותו אורך כמו זו שמתחתיה. לכן הפעולה היחידה שניתן לעשות היא לקחת את הריבועים הירוקים ולהשתמש בהם כדי להאריך את השורות התחתונות – כלומר, לכל ריבוע אדום מוצמד ריבוע ירוק:



בדוגמא שלעיל לא הייתה בעיה עם הפעולה הזו, אבל במקרה אחד יכולה להיות בעיה – אם יש ריבוע ירוק שהוא גם אדום. הנה דוגמא למקרה כזה:



באופן כללי, יהיה ריבוע שהוא גם ירוק וגם אדום אם מספר השורות הכולל k שווה לאורך האלכסון האדום s, כך שיש ריבוע על האלכסון גם בשורה העליונה. עם זאת, בדוגמה שראינו אין בעיה כי אפשר להסיר את הריבוע הירוק ולהוסיף אותו לשורה התחתונה ולקבל טבלת יאנג תקינה. הבעיה היחידה יכולה לצוץ אם אורך השורה העליונה שווה לאורך האלכסון האדום, כי אז אחרי שנוסיף ריבועים ירוקים לאלכסון, עדיין נישאר עם ריבוע בודד שאין לנו מה לעשות איתו (לא ניתן להשאיר אותו למעלה כי כך יוצא שלא שינינו את מספר השורות, מה שהיה כל מטרת התהליך).

כלומר, תהיה לנו בעיה רק אם מספר השורות הכולל k **שווה** למספר הריבועים הירוקים והאדומים, $s = r = k$.

ראינו כי המקרים הבעייתיים היחידים שבהם הפעולה אינה מוגדרת הם אלו שבהם $s = k$ ו־$r = s$ או $r = s+1$. אלו המקרים שמהם יגיעו המספרים המחומשים; ראשית נסיים לטפל במקרים האחרים.

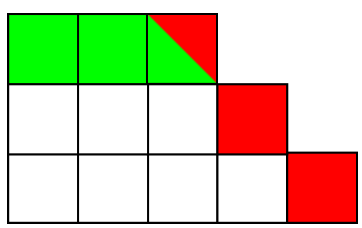
טענה 2.16 אם $s \neq k$ או ש־$s = k$ אבל $r \neq s, s+1$ אז הפעלה כפולה של הפעולה מבטלת את עצמה.

הוכחה: ראשית נניח ש־$r < s$, כלומר אנחנו בסיטואציה שבה נבנתה שורה חדשה על ידי לקיחת s הריבועים של האלכסון. נסמן ב־s', r' את אורך האלכסון ואורך השורה העליונה החדשים, אז על פי הבניה $r' = s$ וכמו כן בהכרח $s' \geq s$ שכן גם אחרי שהסרנו ריבוע אחד מ־s השורות התחתונות, האורכים שלהם הם עדיין רצופים, ואולי התווספו שורות נוספות אל האלכסון. כמו כן, מספר השורות הקודם קיים $k \geq s$ ו־$k' = k + 1$ כי הוספנו שורה אחת, כך ש־$k' > s = r'$ ולכן אנחנו לא בסיטואציה שבה יש ריבוע ירוק שהוא גם אדום. לכן הפעלת הפעולה שוב תסיר את השורה החדשה ותיצור ממנה את האלכסון שהסרנו קודם - כלומר, הפעולה ביטלה את עצמה.

כעת נניח ש־$r \geq s$ אבל $s \neq k$. בסיטואציה הזו אנו מרחיבים את r השורות התחתונות על ידי הוספת משבצות השורה העליונה. מכיוון שכל השורות שהוספנו אליהן משבצות היו קודם באלכסון, הן גם כעת באלכסון (כי הגדלנו את כולן באותה המידה) ונפתח פער של לפחות שתי משבצות מהשורות הנותרות (יש שורות נותרות כי $s < k$) כך ש־$s' = r$, ואנו יודעים ש־$r < r'$ על פי הכלל שבו השורה העליונה קטנה מהשורה הבאה אחריה. לכן אנחנו בסיטואציה $s' < r'$ והפעלה נוספת של הפעולה תסיר את משבצות האלכסון שהוספנו ותבנה מהם מחדש את השורה העליונה. ■

ראינו שפרט לסיטואציה שבה $s = k$ ו־$r \in \{s, s+1\}$ הפעולה שתיארנו מוגדרת היטב ומהווה את ההופכית של עצמה. נראה כי סיטואציות כאלו יכולות להתרחש רק כאשר מספר המשבצות הכולל n הוא מספר מחומש, ויתר על כן - עבור כל מספר מחומש, רק סיטואציה אחת כזו יכולה להתרחש ולכן עבור כל שאר החלוקות של n, מספר החלוקות הזוגיות שווה למספר החלוקות האי־זוגיות.

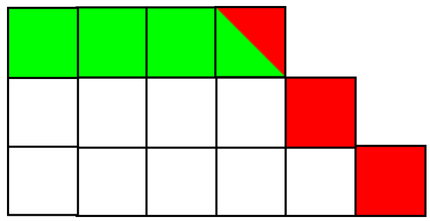
ראשית נניח כי $s = r = k$. דוגמא לסיטואציה כזו עבור החלוקה $12 = 3 + 4 + 5$:



מתוך $s = r = k$ קל לנו לחשב את מספר המשבצות הכולל. השורה העליונה היא בעלת k משבצות, ויש בדיוק k שורות, כך שסכום אורכיהן הוא

$$\begin{aligned} k + (k+1) + (k+2) + \ldots + (k + (k-1)) &= k^2 + (1 + 2 + \ldots + (k-1)) \\ &= k^2 + \frac{(k-1)\,k}{2} = \frac{2k^2 + k^2 - k}{2} \\ &= \frac{k\,(3k-1)}{2} \end{aligned}$$

מקרה זה מתאים בדיוק לנוסחת המספרים המחומשים, כאשר מציבים בה k חיובי (מספר מחומש "רגיל", להבדיל ממספר מחומש "מוכלל").
דוגמא למקרה בעייתי בו $s=k$ ו־$r=s+1$ ראינו קודם, עם החלוקה $15=4+5+6$:



גם כאן, קל לחשב את מספר המשבצות הכולל, שכן השורה הראשונה היא $k+1$ ויש k שורות, ולכן נקבל

$$\begin{aligned}(k+1)+(k+2)+\ldots+(k+k) &= k^2+(1+2+\ldots+k)\\ &= k^2+\frac{k(k+1)}{2}=\frac{2k^2+k^2+k}{2}\\ &= \frac{3k^2+k}{2}=\frac{k(3k+1)}{2}\end{aligned}$$

כאשר k חיובי. אפשר היה לסיים כאן, אבל אנחנו מנסים ליצור אחידות בנוסחאות ולכן רוצים להראות שאפשר לקבל את אותו מספר בדיוק על ידי הצבה של מספר **שלילי** בנוסחה של המספרים המחומשים. פורמלית, אם קיבלנו את המספר $\frac{t(3t+1)}{2}$, עבור t חיובי כלשהו, אנחנו רוצים להראות שקיים k שלילי כך ש־$\frac{t(3t+1)}{2}=\frac{k(3k-1)}{2}$. לצורך כך נבחר פשוט $k=-t$, ונקבל

$$\frac{k(3k-1)}{2}=\frac{-t(-3t-1)}{2}=\frac{-t(-3t-1)}{2}=\frac{3t^2+t}{2}=\frac{t(3t+1)}{2}$$

כפי שרצינו.
לסיום, נזכיר את הנוסחה שאנו מוכיחים:

$$q_{even}(n)-q_{odd}(n)=\begin{cases}1 & n=\frac{k(3k-1)}{2},\ k\text{ even}\\ -1 & n=\frac{k(3k-1)}{2},\ k\text{ odd}\\ 0 & \text{else}\end{cases}$$

בנוסחה הזו יש פיסת אינפורמציה אחת נוספת: הזוגיות של k קובעת מי תהיה החלוקה ה"עודפת": אם k זוגי, אז יש חלוקה זוגית אחת נוספת, ואם k אי זוגי יש חלוקה אי זוגית אחת נוספת. כפי שראינו זה עתה, k הוא **מספר השורות** במקרה הבעייתי היחיד, ולכן הזוגיות של k היא אכן זוגיות החלוקה העודפת, כמבוקש.