ב. יסודות תורת הקבוצות

ב.1 מושגי יסוד

מושג הקבוצה הינו אחד המושגים היסודיים של המתמטיקה. למעשה זהו *המושג היסודי* של המתמטיקה המודרנית. בהתאם, הלשון של תורת הקבוצות היא הלשון בה מוצגים כיום ענפי המתמטיקה השונים, ומשתמשים בה רבות גם במדעי המחשב. לימוד לשון זו, יחד עם לימוד התוצאות היסודיות של תורת הקבוצות, הינם המטרות של חלק זה.

מהי, אם כן, קבוצה? ובכן, קבוצה (set) הינה אוסף של עצמים, המהווה עצם בעצמו. לעצמים, מהם מורכבת קבוצה, קוראים *איברי* הקבוצה, ועל כל אחד מהם אומרים, שהוא *שייך* לקבוצה, או שהקבוצה *כוללת* אותו.

את מה שנכתב בפסקה הקודמת אין לראות בגדר *הגדרה* של מושג הקבוצה. מסופקתני, אם המושג "אוסף", המופיע ב"הגדרה" זו, הינו ברור יותר מאשר מושג הקבוצה עצמו. בסופו של דבר, המושגים של "קבוצה", "איבר", ו"שייכות" הינם מושגים יסודיים של תורת הקבוצות, ממש כמו ש"נקודה", "ישר" ו"המצאות נקודה על ישר" הם מושגים יסודיים בגיאומטריה[1] . יש לראות אפוא בפסקה האחרונה הסבר מסייע בלבד. אשר לתוספת "המהווה עצם בעצמו" המופיעה שם - כרגע היא נראית סתומה, מן הסתם. משמעותה והצורך בה יובהרו בהמשך.

הנקודות הבאות חשובות להבנת מושג הקבוצה:

(א) אין הגבלה מראש על מה שיכול לשמש כאיבר בקבוצה. כל דבר, אותו רואים אנו כעצם, כשיר לצורך כך. כך מורכבת קבוצת חברי כנסת מבני-אדם, בעוד הקטע [0,1] (קבוצת המספרים בין 0 ל- 1, כולל 0 ו- 1) הינו קבוצה של מספרים. השאלה, מה יכול לשמש איבר בקבוצה, זהה למעשה לשאלה, מה מוכנים אנו לראות כעצם לגיטימי. התשובה עשויה להיות תלויית קונטקסט. כשעוסקים, למשל, במיתולוגיות, הרי אפילו אלי האולימפוס יהיו (ביחד) קבוצה. בקורס זה, באופן טבעי, נתעניין בעיקר בעצמים מתמטיים. העקרונות שנלמד יהיו נכונים אבל באופן כללי.

[1] אוקלידס, בספרו המקורי, "הגדיר" כל מושג, כולל המושגים היסודיים, אולם בנוגע למושגים היסודיים, כמו "נקודה", אין לראות ב"הגדרות" אלו יותר מהסבר, המסייע לקורא להבין במה מדובר.

(ב) קבוצות אינן חייבות להיות הומוגניות כמו בדוגמאות, שהבאנו בהערה הקודמת. כך הקבוצה, המורכבת מחברי כנסת יחד עם המספרים הטבעיים בין 91 למאה, היא קבוצה לכל דבר, ויש בה 130 איברים.

(ג) קבוצות יכולות להיות סופיות (כמו בדוגמא של הכנסת) או אין-סופיות (כמו בדוגמא של הקטע $[0,1]$).

(ד) כמו שהכרזנו ב"הגדרת" מושג הקבוצה, קבוצות נחשבות בעצמן כעצמים לגיטימיים. אי לכך כל קבוצה יכולה לשמש כאיבר של קבוצה אחרת. דוגמא מהחיים: כיתה בבית הספר הינה קבוצה של תלמידים. ממילא קבוצת הכיתות של אותו בית ספר הינה קבוצה, שכל איבריה הם קבוצות בעצמם.

(ה) אנו רואים קבוצה כנקבעת לחלוטין על-ידי איבריה ולא על-ידי שום דבר אחר (כמו למשל, הסדר, שבו הם מוצגים). כך, אם נסדר יום אחד את התלמידים בכיתה מסוימת לפי סדר אלף-ביתי של שמם, וביום אחר לפי מספרי הזהות שלהם, הרי שבשני המקרים מדובר עדיין באותה קבוצה (בהנחה ששום תלמיד לא נגרע בינתיים, ושום תלמיד חדש לא התווסף). הניסוח המדויק של רעיון זה מתבטא בעיקרון הבא:

עקרון האקסטנציונליות:
שתי קבוצות הן שוות, אם ורק אם יש להן בדיוק אותם איברים.

כיוון שמושג ההשתייכות של עצם לקבוצה הינו אחד המושגים הבסיסיים של תורת הקבוצות, נכניס לו סימון מיוחד $\in$ (או ε)[2]. מכאן ואילך, כשנרצה להגיד שהעצם t הינו איבר של הקבוצה A נכתוב בקיצור $t \in A$ (לדוגמא: $\frac{1}{2} \in [0, 1]$). במקום $\neg(t \in A)$ נכתוב בדרך כלל $t \notin A$ (לדוגמה $2 \notin [0,1]$). קיצורים מקובלים אחרים, דומים לאלו הקשורים ב- $\leq$, הינם:

לכתוב $\forall x \in A\ (...)$ – במקום $\forall x(x \in A \Rightarrow . . .)$

לכתוב $\exists x \in A\ (...)$ – במקום $\exists x(x \in A \wedge . . .)$

בעזרת הסימן החדש נוכל עכשיו לבטא את עיקרון האקסטנציונליות כך:

$$\textbf{(Ext1):}\quad \forall A\ \forall B\ (A = B \Leftrightarrow (\forall x.\ \ x \in A \Leftrightarrow x \in B))$$

[2] $\in$ ו- ε הם שתי צורות מקובלות של האות היוונית "אפסילון", ובשתיהן ניתן להיתקל בספרות.

אם נשתמש בשקילויות לוגיות (הגדרת $\Leftrightarrow$ ושקילות (15a) מטבלה א.3), נקבל שזה שקול ל:

$$\textbf{(Ext2)}: \quad \forall A \; \forall B \; (A = B \Leftrightarrow [(\forall x.\, x \in A \Rightarrow x \in B) \wedge (\forall x.\, x \in B \Rightarrow x \in A)]\,)$$

הניסוח האחרון (במיוחד שני הקוניוקטים בתוך הסוגריים המרובעים) מביא אותנו באופן טבעי להכנסת מושג בסיסי חדש של תורת הקבוצות:

הגדרה:

נאמר שקבוצה A היא ***תת-קבוצה*** של קבוצה B (או ש- A ***חלקית*** ל- B , או ש- B ***מכילה*** את A) אם כל איבר של A הינו איבר של B. A נקראת ***חלקית ממש*** ל- B אם A חלקית ל- B אך לא שווה לה.

סימונים:

$A \subseteq B$ פירושו ש- A חלקית ל- B

$A \subset B$ פירושו ש- A חלקית ממש ל- B

בשפה הפורמלית, לכן, הגדרת $\subseteq$ ו- $\subset$ הינה:

$$A \subseteq B =_{Df} \forall x \, (x \in A \Rightarrow x \in B)$$
$$A \subset B =_{Df} A \subseteq B \wedge A \neq B$$

(הסימון $=_{Df}$ פירושו, כזכור, שאגף ימין מהווה הגדרה של אגף שמאל).

אם נתבונן עתה ב- **(Ext2)**, הניסוח השני של עקרון האקסטנציונליות, נראה שבעזרת הסימונים החדשים ניתן לנסחו כך:

$$\forall A \; \forall B \, (A = B \;\; \Leftrightarrow \;\; (A \subseteq B \; \wedge \; B \subseteq A \,))$$

במלים: שתי קבוצות הינן שוות, אם כל אחת מהן הינה חלקית לשנייה. מזה נגזרת הצורה הנפוצה ביותר להוכחת שוויון של שתי קבוצות A ו- B. ההוכחה בצורה זו מתחלקת לשני חלקים: מראים לחוד ש- A חלקית ל- B ולחוד ש- B חלקית ל- A (דוגמאות נראה בפרק הבא).

הערות על סימונים:

1. כמקובל במקרים אחרים, נכתוב לעתים קרובות $A \not\subset B$, $A \not\subseteq B$ במקום $\neg(A \subseteq B)$, $\neg(A \subset B)$, בהתאמה.

2. לרוע המזל, השימוש בסימנים $\subseteq$ ו- $\subset$ ובטרמינולוגיה למעלה אינו אחיד בספרות. יש האומרים "חלקי או שווה" היכן שאנו אומרים "חלקי" , ו"חלקי" במקום בו אנו אומרים "חלקי ממש". בדומה, יש כאלה המשתמשים בסימן $\subset$ היכן שאנו משתמשים ב $\subseteq$, ו- $\underset{\neq}{\subset}$ במקום שאנו כותבים $\subset$.

3. ממש כמו לגבי $\leq$ ו- $\in$ גם לגבי $\subseteq$ ו- $\subset$ מקובלים הקיצורים
$$\forall x \subseteq A(...) =_{Df} \forall x(x \subseteq A \Rightarrow ...)$$
$$\exists x \subseteq A(...) =_{Df} \exists x(x \subseteq A \wedge ...)$$
וכנ"ל לגבי $\subset$.

דוגמאות:

(1) קבוצת המספרים הזוגיים חלקית לקבוצת המספרים הטבעיים. יתר על כן: היא חלקית ממש לקבוצה זו.

(2) קבוצת הנקודות בפנים מעגל מסוים חלקית (ממש) לקבוצת הנקודות במישור.

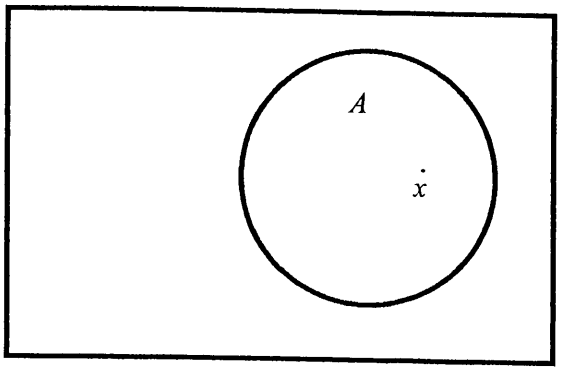
(3) כל קבוצה הינה חלקית לעצמה. ואכן, לפי הגדרה $A \subseteq A$ אם ורק אם $\forall x\ (x \in A \Rightarrow x \in A)$, וזוהי אמת לוגית. לעומת זאת $A \not\subset A$.

אזהרה:

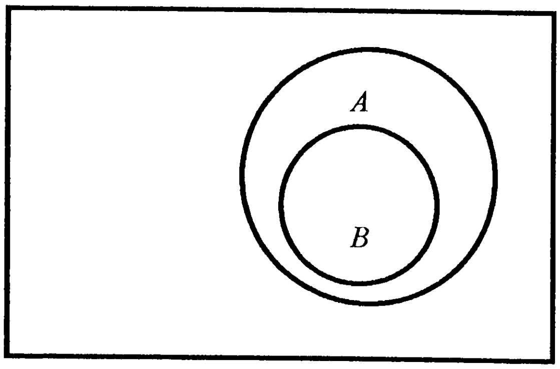
יש להיזהר מאד מלבלבל בין מושג השייכות ($\in$) ובין מושג החלקיות ($\subseteq$) . קבוצת המספרים הזוגיים היא , כאמור , חלקית לקבוצת המספרים הטבעיים, אך היא איננה שייכת לה, כיוון שהיא עצמה איננה מספר טבעי. אזהרה זו הינה חשובה, בגלל שהשפות הטבעיות מטשטשות לעתים קרובות את ההבדל בין שני היחסים. לדוגמה: כשאנו אומרים ש"היהודים הם בני אדם", כוונתנו לכך שקבוצת היהודים הינה ***חלקית*** לקבוצת בני האדם. לעומת זאת באומרנו כי "היהודים הם עם ככל העמים" כוונתנו היא, שקבוצת היהודים ***שייכת*** למשפחת העמים (שהיא קבוצה, שאבריה הם בעצמם קבוצות). מכאן, שמשמעות מלה כמו "הם" יכולה לפעמים להיות "$\in$" ולפעמים "$\subseteq$" , בהתאם לקונטקסט!

כדי *להמחיש* את המושגים היסודיים של תורת הקבוצות משתמשים לעתים באמצעי הידוע *כדיאגרמת וון*. בדיאגרמות אלו מיוצג עולם העצמים, שבו יש לנו עניין, על-ידי מלבן במישור. הקבוצות השונות של עולם זה מיוצגות על ידי עיגולים או אליפסות, הפנימיים למלבן זה (ולעתים על ידי חלקים שלהם, או קבוצה סופית של חלקים

כאלה). העצמים בעולם זה מיוצגים על-ידי הנקודות בפנים המלבן. כך למשל מומחשות העובדות ש- $x \in A$ וש- $B \subseteq A$ באופן הבא:



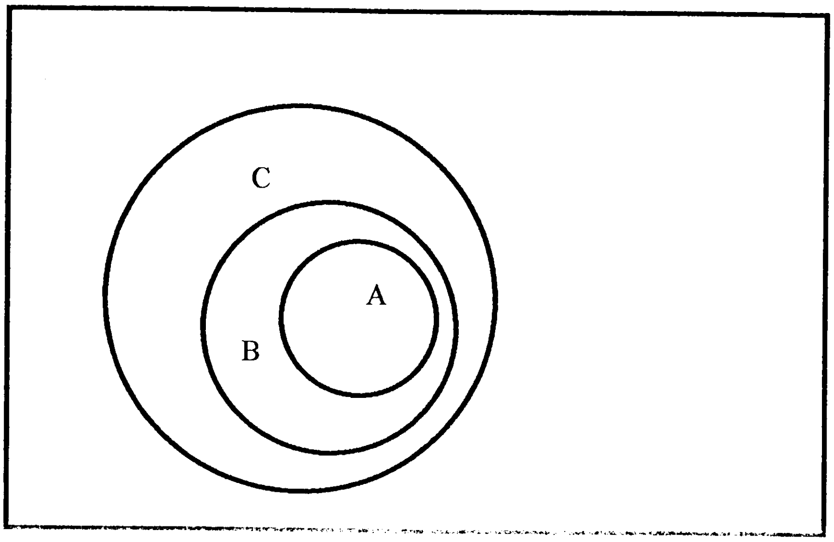
$x \in A$



$B \subseteq A$

דוגמה *לאינטואיציה*, שהשימוש בדיאגרמות וון יכול לספק נותנת לנו הדיאגרמה הבאה:

ציור ב.1: $A \subseteq B \wedge B \subseteq C$



הדיאגרמה ממחישה מצב שבו A, B ו- C הינן קבוצות כך ש- $A \subseteq B$ וגם $B \subseteq C$. התבוננות בה מגלה (אם נמחק בדמיוננו את קו השפה של B), שבמקרה זה $A \subseteq C$. זהו אכן עקרון פשוט וחשוב, הידוע בשם:

טרנזיטיביות יחס ההכלה:

אם $A \subseteq B$ ו- $B \subseteq C$ אז $A \subseteq C$

(או באופן פורמלי: $\forall A \forall B \forall C (A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C)$).

חשוב להבין, שלמרות שהשימוש בדיאגרמה הוביל אותנו לאינטואיציה בדבר עיקרון זה ואף נותן אינטואיציה למה הוא נכון, *אין* הוא מהווה *הוכחה*, כי הוא מטפל בסוג מיוחד מאוד של קבוצות. ניתן, אפוא, הוכחה אמיתית:

הוכחת טרנזיטיביות ההכלה:

עלינו להראות שלכל x, אם $x \in A$ אז $x \in C$ יהי לכן $x \in A$ כיוון ש- $A \subseteq B$ נובע מזה ש- $x \in B$ מזה ומהעובדה ש- $B \subseteq C$ נובע ש- $x \in C$ מש"ל.

%%

למען אלו שלא קראו את פרק א.4, ולמען אלה שרוצים לקראו שנית, אך עם דוגמה אחרת, נביא עתה ניתוח לוגי מפורט של ההוכחה שהבאנו עתה, ומה בדיוק נעשה בה[3] (בעתיד לא נעשה עוד ניתוחים כאלה!).

ובכן, הטענה שאנו רוצים להוכיח היא (בסגנון הגיאומטריה בתיכון):

(I) נתון: (1) $A \subseteq B$

(2) $B \subseteq C$

צ"ל: $A \subseteq C$

הצעד הראשון בהוכחה שהבאנו הוא צעד, שנעשה באופן *לא-מפורש*. כותב ההוכחה מניח במקרים כאלו, שגם הקורא עושה צעד זה לעצמו באופן אוטומטי. הצעד הינו החלפת הסימנים והמושגים, המופיעים בנתונים ובמה שצריך להוכיח, במשמעותם המקורית. במקרה שלפנינו משמעות $A \subseteq B$, כפי שניתנה בהגדרה למעלה, היא, כזכור: $\forall x\,(x \in A \Rightarrow x \in B)$. מכאן, שבעצם הבעיה היא הבעיה הבאה:

(II) נתון: (1) $\forall x\,(x \in A \Rightarrow x \in B)$

(2) $\forall x\,(x \in B \Rightarrow x \in C)$

צ"ל: $\forall x\,(x \in A \Rightarrow x \in C)$

ההוכחה למעלה מתחילה למעשה בנקודה זו. הבה ננתחה משפט אחר משפט.

---

[3] הסברים *למה* הכללים הלוגיים אכן תקפים ניתן למצוא בפרק א.4. בדוגמא כאן נתרכז בהבנת צורת השימוש בהם.

"עלינו להראות שלכל x, אם $x \in A$ אז $x \in C$"
לכאורה אין כאן אלא ניסוח ב"עברית מתמטית" של מה שצריך להוכיח. למעשה, יש כאן , כמו שנראה מייד, גם רמז, שאנו עומדים להשתמש בכלל 13 (וגם 1) מטבלה א.4, עמוד 26.

"יהי לכן $x \in A$"
משפט קצר ותמים-למראה זה טומן בחובו הפעלה של שני כללים! ראשית, מה שאנו רוצים להוכיח הוא משפט אוניברסלי מהצורה $\forall x\varphi$ (כש- φ היא כאן הנוסחה $x \in A \Rightarrow x \in C$). משפט כזה מוכיחים בדרך כלל בעזרת כלל 13 (כל המספרים – לפי טבלה א.4). אם נקרא כלל זה ***מלמטה למעלה***, הוא אומר, שכדי להוכיח משהו מהצורה $\forall x\varphi$, עלינו להוכיח פשוט את φ – ובלבד שהמשתנה x בו מדובר אינו מופיע חופשי בהנחות. במקרה שלנו המשתנה x אומנם ***מופיע*** בהנחות (1) ו- (2), אך ***אינו*** מופיע בהם ***חופשי***. לכן הפעלת כלל 13 הינה אפשרית[4]. מספיק לכן להוכיח את הנוסחה $x \in A \Rightarrow x \in C$ מהנתונים (1) ו- (2). נוסחה זו היא בעלת צורה של גרירה, וגרירה מוכיחים בדרך כלל בעזרת כלל 1. כלל זה קובע (אם נקרא גם אותו מלמטה למעלה), שכדי להוכיח נוסחה מהצורה $\varphi \Rightarrow \psi$ עלינו להוסיף את φ לרשימת הנתונים, ולנסות להוכיח את ψ מרשימת הנתונים המורחבת. במקרה שלנו צריכים אנו לכן להוסיף את $x \in A$ לרשימת הנתונים ולנסות להוכיח את $x \in C$ מרשימת הנתונים החדשה.

נסכם אפוא: כדי להוכיח את (II) למעלה, די אם נוכיח את הטענה הבאה:

(III) נתון: (1) $\forall x\,(x \in A \Rightarrow x \in B)$
(2) $\forall x\,(x \in B \Rightarrow x \in C)$
(3) $x \in A$

צ"ל: $x \in C$

אם נצליח, אז בעזרת כלל (1) נוכל להסיק, ש- $x \in A \Rightarrow x \in C$ נובע מהנתונים (1) ו-(2) (בלבד!), ואז לפי כלל (13), ש- $\forall x(x \in A \Rightarrow x \in C)$ נובע גם הוא מנתונים אלו. כך נקבל הוכחה של (II).

[4] לו היה x מופיע חופשי באחת ההנחות, היה עלינו להשתמש תחילה בכלל (α) ולהחליף את $\forall xA$ ב- $\forall y(A(y/x))$, כש-y משתנה חדש בתכלית – ומכאן להתקדם באותה צורה.

כדאי לשים לב לשני הדברים הבאים:

(א) הכללים (1) ו- (13) הם כללים, שבאופן מעשי משתמשים בהם כדי לעשות *רדוקציה* של הבעיה לבעיה פשוטה יותר. במקום להוכיח טענה מסוימת אנו מוכיחים משהו אחר, שהוא *מספיק*, כי אם נצליח, אז כללים (1) ו- (13) יסיימו את המלאכה. לכן, למרות שבהוכחה מלאה, פורמלית, כללים אלו מופעלים בדרך כלל *בסוף*, ***ההתיחסות*** אליהם בהוכחות לא פורמליות נעשית (ברמז בלבד!) *בהתחלה*! רמז לכך במקרה שלנו ניתן במלים "עלינו להראות", בהן נפתח המשפט הראשון (בעברית) בהוכחה. יש בהן רמז, שאנו מתחילים בהוכחת משהו, ***שמספיק*** לצורך הוכחת מה שאנו רוצים להוכיח באמת (ולכן, בדרך כלל, זהו רמז לשימוש בכללים (13) ו- (1)).

(ב) ביטויים כמו "*יהי* $x \in A$", "יהי $\varepsilon > 0$" וכדומה מרמזים ***תמיד*** על שימוש בכלל 13, וכמעט תמיד – בצירוף עם כלל (1).

נמשיך עתה בניתוח הלוגי של ההוכחה.

"כיוון ש- $A \subseteq B$, נובע מזה ש- $x \in B$"
גם משפט זה טומן בחובו שימוש בשני כללים. תחילה אנו מסיקים, בעזרת כלל 14, מהנתון (1) ב- III (שאינו אלא $A \subseteq B$, כזכור) ש- $x \in A \Rightarrow x \in B$. מנוסחה אחרונה זו והנתון החדש (3) ($x \in A$) נובע $x \in B$ בעזרת כלל (2).

"מזה ומהעובדה ש- $B \subseteq C$ נובע $x \in C$"
הסיפור כאן דומה. מ"זה" (כלומר: מ- $x \in B$) ומנתון (2) ($B \subseteq C$) נובע, בעזרת כלל 14 ובעקבות הפעלה של כלל (2), $x \in C$.

"מש"ל"
ה- "מש"ל" אומר, בעצם, שסיימנו הוכחת III. מכל מה שהסברנו פירוש הדבר, שאנו יכולים להוכיח גם את II וממילא את I. אין התייחסות מפורשת לכך, ולעולם גם לא תהיה!

הערה
המעמיקים ישימו לב כאן, שלמעשה, אפילו כדי לקבל את הצורה (I) מהניסוח הפורמלי של המשפט מספר שורות קודם לכן, הופעלו עוד כללים לוגיים רבים: תחילה אנו אומרים, שבמקום להוכיח

$$\forall A \forall B \forall C (A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C)$$

די, בגלל כלל 13, להוכיח את $A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C$. בשביל זה די, לאור כלל (1), להניח את $A \subseteq B \wedge B \subseteq C$ ולהוכיח את $A \subseteq C$. את הנתון $A \subseteq B \wedge B \subseteq C$ אנו מפרקים מיד, לפי כלל (4), לשני נתונים: $A \subseteq B$ ו- $B \subseteq C$. רק אז מקבלים אנו, שמספיק להוכיח את (I) !

%%

לסיום הפרק, שתי הערות עדינות-משהו:

(1) קוראה חדת-עין הבחינה אולי בבעייתיות מסוימת בצורה, בה ניסחנו את עקרון האקסטנציונליות למעלה (ב- **Ext1**, למשל). כאשר כתוב שם $\forall x...$ הכוונה היא אכן *לכל* עצם אפשרי. לעומת זאת כאשר כתוב , באותה שורה, $\forall A \forall B....$ הכוונה היא רק לכל *קבוצה* A ולכל *קבוצה* B. האמת היא, שבתורת הקבוצות המתמטית הטהורה אין הבדל, כי שם כל עצם מתמטי הינו קבוצה (כן, אפילו המספרים השונים מוגדרים שם כקבוצות מסוימות!) . עם זאת, בקורס הנוכחי אנו מניחים את אפשרות קיומם של עצמים, שאינם קבוצות. למעלה מזאת: אנו אכן נתייחס למספרים, נקודות במישור ודברים אחרים כאל עצמים "אטומיים", שאינם קבוצות. כדי לנסח לכן את **Ext1** ומשפטים אחרים בצורה מדויקת, היה עלינו להכניס יחס (או "פרדיקט") מיוחד, כמו $\text{set}(A)$, שמובנו: "A הינה קבוצה". היה עלינו אז לנסח את **Ext1** כך:

$$\forall A \forall B\ (\text{set}(A) \wedge \text{set}(B) \Rightarrow [\mathrm{A} = B \Leftrightarrow (\forall x.\ x \in A \Leftrightarrow x \in B)])$$

כדי למנוע סרבול זה, לא נשתמש בפרדיקט $\text{set}(X)$. במקום זאת נשתמש בשיטה של משתנים מיוחדים עבור קבוצות. מכאן ואילך (ולמעשה כבר לכל אורך פרק זה) אותיות לטיניות גדולות A, B, C ...X, Y, Z ישמשו כמשתנים עבור קבוצות בלבד. רשמית לכן:

כאשר נכתוב $\forall A(...)$ הכוונה היא ל: $\forall A(\text{set}(A) \Rightarrow ...)$

כאשר נכתוב $\exists A(...)$ הכוונה היא ל: $\exists A(\text{set}(A) \wedge ...)$

(2) כמו כן נלך גם אנו לעתים קרובות בעקבות הנוהג המקובל ונשמיט כמתים אוניברסליים בהתחלת טענה. משהו כמו **Ext1** ייכתב לכן, בדרך כלל, פשוט כך:

$$A = B \Leftrightarrow \forall x.\ x \in A \Leftrightarrow x \in B$$

ב.2 : הגדרת קבוצות וסימונן

בפרק הקודם, עת רצינו להתייחס לקבוצה מסוימת, השתמשנו בלשון הבאה: "קבוצת כל המספרים הטבעיים הזוגיים", "קבוצת המספרים הממשיים בין 0 ל- 1" וכו'. קבוצות הוגדרו על-ידי *תכונה* מסוימת, המשותפת לכל איבריהן, ולאיבריהן בלבד יש תכונה זו. כיוון שזו הצורה הנפוצה ביותר להגדרת קבוצות, הוכנס עבורה סימן מיוחד. במקום "קבוצת כל העצמים x, שיש להם תכונה ψ", כותבים : "$\{x\,|\psi\}$" . בביטוי זה x חייב להיות משתנה, ו- ψ היא נוסחה, ש- x הוא אחד ממשתניה החופשיים (לא בהכרח היחיד). למשל:

(א) $\{n \mid (\text{הינו טבעי } n) \wedge \exists k. (\text{הינו טבעי } k) \wedge (n = 2\cdot k)\}$ הינה קבוצת כל המספרים הטבעיים הזוגיים.

(ב) $\{x \mid (\text{הינו ממשי } x) \wedge 0 \le x \le 1\}$ הינה קבוצת המספרים הממשיים בין 0 ל- 1.

את הנוסחאות בעברית "n הינו טבעי", "x הינו ממשי" נחליף בהמשך בסימונים קצרים יותר. איתם או בלעדיהם מקובל לקצר ביטויים כמו השניים האחרונים על ידי שימוש במשתנים מסוג מיוחד. כך מקובל מאוד, שמשתנים המתחילים באותיות i,j,k,l,m,n הינם משתנים עבור המספרים הטבעיים. לכן נכתוב לעתים קרובות $\{n \mid \exists k.\ n = 2k\}$ במקום הביטוי הארוך יותר למעלה (עוד נשוב לנקודה זו).

האופרטור $\{\ldots \mid \_\_\}$ הוא אופרטור קשירה. את המשתנה הנקשר כותבים משמאל לקו ההפרדה " | " שבסוגריים המסולסלים. טווח הקשירה הוא הצד הימני של קו הפרדה זה. כיוון שכך, חלים כאן כל העקרונות שעליהם דיברנו בפרק הקודם. כך למשל כלל (α) ישים, ונובע ממנו ש:

$$\{x \mid \psi\} = \{y \mid \psi(y/x)\}$$

ובלבד ש- y כשיר להפעלת כלל (α) כאן. כך לדוגמה:

$$\{n \,|\, \exists k.\ n = 2k\} = \{j \,|\, \exists k.\ \mathrm{j} = 2k\} = \{j \,|\, \exists i.\ j = 2i\}$$
$$\{x \,|\, 0 \le x \le 1\} = \{z \,|\, 0 \le z \le 1\}$$

התכונה היסודית של $\{x\,|\,\psi\}$ ניתנת על ידי:

$$\forall x (x \in \{x \,|\, \psi)\} \Leftrightarrow \psi)$$

מה ששקול, לפי כלל (α) עבור $\forall$, ל :

$$\forall y (y \in \{x \,|\, \psi\} \Leftrightarrow \psi(y/x))$$

מזה נובע, לפי הכלל הבסיסי של $\forall$ (כלל 14), שלכל ביטוי t, החופשי להצבה במקום x ב- ψ מתקיים:

כלל (β) : $\boxed{t \in \{x \mid \psi\} \Leftrightarrow \psi(t/x)}$

דוגמאות

1. $\frac{1}{4} \in \{x \mid 0 \le x \le 1\} \Leftrightarrow 0 \le \frac{1}{4} \le 1$

כיוון שצד ימין של האקויולנציה נכון, דהיינו : אכן $0 \le \frac{1}{4} \le 1$, אז גם צד שמאל שלה הוא נכון, כלומר $\frac{1}{4} \in \{x \mid 0 \le x \le 1\}$.

2. $2 \in \{x \mid 0 \le x \le 1\} \Leftrightarrow 0 \le 2 \le 1$

כיוון שכאן צד ימין של האקויולנציה *אינו* נכון, גם צד שמאל אינו נכון, כלומר: $2 \notin \{x \mid 0 \le x \le 1\}$.

3. **נסמן את קבוצת המספרים הטבעיים ב- $\mathbf{N}$. אזי[5]:**

$$8 \in \{n \mid n \in \mathbf{N} \wedge \exists k \in \mathbf{N}.\ n = 2k\} \Leftrightarrow 8 \in \mathbf{N} \wedge \exists k \in \mathbf{N}.\ 8 = 2k$$

עתה כאן אגף ימין של השקילות הוא נכון. זאת, משום ששני הקוניוקטים בו הם נכונים. $8 \in \mathbf{N}$ (כלומר 8 אכן מספר טבעי) וכמו כן $4 \in \mathbf{N}$ ו- $8 = 2 \cdot 4$, ולכן $\exists k \in \mathbf{N}.8 = 2k$. מכאן שאכן $8 \in \{n \mid n \in \mathbf{N} \wedge \exists k \in \mathbf{N}.\ n = 2k\}$. במלים אחרות: 8 שייך לקבוצת המספרים הזוגיים (דהיינו: 8 הינו מספר זוגי).

הערה: במקום לכתוב "n הינו טבעי" כמו למעלה, כתבנו פה "$n \in \mathbf{N}$" .

4. **הבה נבדוק האם**

$$1.5 \in \{x \mid \forall y.\ y \in \{x \mid x^2 < 2\} \Rightarrow x \ge y\} \qquad (*)$$

לפי הכלל הבסיסי (β) זה נכון אם ורק אם

$$\forall y.\ y \in \{x \mid x^2 < 2\} \Rightarrow 1.5 \ge y$$

---

[5] נזכיר ש : $\exists k \in \mathbf{N}....$ פירושו: $\exists k(k \in \mathbf{N} \wedge ...)$

(כאן ψ היא הנוסחה "$\forall y.\ y \in \{x | x^2 < 2\} \Rightarrow x \geq y$", והצבנו את 1.5 במקום המופע החופשי היחיד שיש ל-x בה, שהוא המופע האחרון בה של x). שוב לפי הכלל הבסיסי (β), מופעל כאן עתה על $y \in \{x | x^2 < 2\}$, מה שקיבלנו הוא נכון אם ורק אם:

$$(**) \quad \forall y.\ y^2 < 2 \Rightarrow 1.5 \geq y$$

טענה אחרונה זו הינה נכונה: 1.5 אכן גדול מכל מספר, שריבועו קטן מ- 2. מכאן שגם הטענה (*), בה התחלנו, הינה נכונה.

הערה:

בתרגום לעברית, (*) אומר, ש- 1.5 שייך לקבוצת החסמים מלעיל של קבוצת המספרים, שריבועם קטן מ- 2 (כלומר 1.5 הוא חסם מלעיל של קבוצת המספרים, שריבועם קטן מ- 2).

שאלה:

בקבוצה הנ"ל יש איבר קטן ביותר. מיהו?

הערה נוספת:

ב- (*) המשתנה x נקשר במספר מקומות על ידי אופרטורי קשירה שונים. כרגיל, מי שחושש להתבלבל יכול להפעיל את כלל (α) ולהימנע ממצב עניינים זה. אפשר להעביר תחילה בעזרתו את (*) ל:

$$1.5 \in \{z | \forall y.\ y \in \{x | x^2 < 2\} \Rightarrow z \geq y\}$$

או ל:

$$1.5 \in \{x | \forall y.\ y \in \{z | z^2 < 2\} \Rightarrow x \geq y\}$$

את הפישוטים בעזרת כלל (β) ניתן להתחיל מאחת הצורות הללו. התוצאה הסופית, (**) לא תהיה שונה (בדקו!).

הערה שלישית ואחרונה:

אין ספק שכשלעצמו, (*) קשה מאד לקריאה והבנה. מבחינה זו הניסוח בעברית עדיף. לעומת זאת, כשאנו באים לברר את שאלת *הנכונות*, הרי הרבה יותר קל להשיב נכונה על השאלה אם (**) נכון, מאשר אם הטענה, כמו שהיא מנוסחת בעברית, היא נכונה (ורמת הסיבוכיות כאן אינה גבוהה במיוחד!). עתה, הנקודה המרכזית כאן היא, שהמעבר מ- (*) ל- (**) נעשה לפי כללים מדויקים, שהפעלתם ***אינה דורשת כל מחשבה ואינטליגנציה, אלא רק רכישת מיומנות.*** פישוט (*) ל- (**) דומה לפישוט ביטויים אלגבריים על פי נוסחאות, הנעשה בחטיבת הביניים (למעשה הוא מסובך

פחות!) , או לגזירת פונקציות בסוף התיכון. אפשר ללמוד לגזור פונקציות נכונה בלי להבין כלל את מושג הנגזרת. בתהליך הלימוד נעשות שגיאות לא מעטות (למשל: שגיאות הנובעות מאי הפעלה נכונה של הכלל לגזירת פונקציה מורכבת), אבל בסופו של דבר כולם לומדים לגזור נכון. זו רק שאלה של זמן ותרגול. לימוד הפעלה נכונה של כללים כמו (α) ו (β) הוא דומה, אך הוא פשוט לאין ערוך!

אנו משאירים לבדיקת הקוראים את נביעת העובדות הפשוטות (אך חשובות) הבאות מן ההגדרות:

(i) $\{x|\psi\}\subseteq\{x|\varphi\}\Leftrightarrow\forall x(\psi\Rightarrow\varphi)$

(ii) $\{x|\psi\}=\{x|\varphi\}\Leftrightarrow\forall x(\psi\Leftrightarrow\varphi)\Leftrightarrow\forall x(\psi\Rightarrow\varphi)\wedge\forall x(\varphi\Rightarrow\psi)$

(iii) $[\forall x\in\{y|\varphi\}.\ \psi]\Leftrightarrow[\forall x.\ \varphi(x/y)\Rightarrow\psi]$

(iv) $[\exists x\in\{y|\varphi\}.\ \psi]\Leftrightarrow[\exists x.\ \varphi(x/y)\wedge\psi]$

ב- (iii) וב- (iv) x חייב, כמובן, להיות חופשי להצבה במקום y בנוסחה φ.

דוגמאות:

$$[\forall x\in\{z|0\leq z\leq 1\}.\ x^2\leq 1]\Leftrightarrow[\forall x.\ 0\leq x\leq 1\Rightarrow x^2\leq 1]$$

$$[\exists a\in\{z|0\leq z\leq 1\}.\ a=a^2]\Leftrightarrow[\exists a.\ 0\leq a\leq 1\wedge a=a^2]$$

הקדשנו לא מעט זמן לסימון $\{x|\psi\}$ בגלל חשיבותו המרובה. נקודה שלא התייחסנו אליה עד כה היא: אלו נוסחאות ψ קבילות לצורך הגדרת קבוצות בצורה $\{x|\psi\}$? כדי לראות, שהשאלה חשובה היא, ושהגבלות כלשהן הינן נחוצות, נתבונן בדוגמאות הבאות.

פרדוקס ברי (Berry):

תהי B ה"קבוצה" הבאה:

B = {n | הוא מספר טבעי, שאי-אפשר להגדיר בעברית בפחות מ- 20 מלים n }

בתוך הסוגריים המסולסלות, מימין ל- " | ", רשומה נוסחה בשפה העברית. בכך אין פסול: בשום מקום עד כה לא התייחסנו לשפה, בה רשומה הנוסחה ψ בביטוי כמו$\{x|\psi\}$, ולגבינו לכן כל נוסחה בעלת מובן, בכל שפה, הינה באה בחשבון. כמו כן נוסחה זו מגדירה אכן, כמדומה, תכונה של עצמים. לכל העצמים, שאינם מספרים

טבעיים, אין ,כמובן, תכונה זו. ברור לעומת זאת, שלרוב המספרים הטבעיים *יש* תכונה זו. בשפה העברית יש מספר סופי של מלים, ולכן רק מספר סופי של ביטויים כשירים דקדוקית, המכילים פחות מעשרים מלים. מתוך אלה רק חלק מגדיר מספרים טבעיים (למשל, הביטוי: " המספר הזוגי החיובי הקטן ביותר", המגדיר את המספר הטבעי 2, ויש בו חמש מלים בלבד). מכאן, שיש מספר סופי של מספרים, להם אין התכונה, ולכל השאר (אין-סוף מספרים) יש. עתה, לפי אחד העקרונות הבסיסיים של המתמטיקה, בכל קבוצה לא-ריקה של מספרים טבעיים יש מספר מינימלי. בפרט נכון הדבר ל-B, אם אכן קבוצה היא. נניח אפוא, ש- b הוא המספר המינימלי ב- B. b הוא לכן ***המספר הטבעי הקטן ביותר שאי אפשר להגדיר בפחות מעשרים מלים*** - והרגע הגדרנו אותו בעשר מלים (המלים המודגשות במשפט האחרון). עשר, כידוע, קטן מעשרים. קיבלנו לכן, שאת b אי אפשר להגדיר בפחות מעשרים מלים, אך אפשר להגדיר בעשר מלים. הכיצד??

למצב עניינים, שאינו מתקבל על הדעת, קוראים פרדוקס[6] , ופרדוקסים הינם דבר מטריד, הדורש התייחסות. ניתוח לעומק של פרדוקס ברי מראה, שמקור הבעיה הוא בכך, שניסוחים בשפה העברית (או כל שפה טבעית אחרת), שהינם תקינים דקדוקית, הם לא תמיד גם בעלי מובן, אפילו אם נדמה כך. במקרה שלפנינו הביטוי "מספר טבעי שאי אפשר להגדיר בפחות מעשרים מלה" אינו צירוף מלים, המגדיר תכונה ברורה, שלכל מספר יש או אין. השפה העברית פשוט עשירה ומורכבת מידי. בתורת הקבוצות האקסיומטית משתמשים לכן בשפה, שהיא גם מצומצמת הרבה יותר וגם פורמלית. בעיות מסוג פרדוקס ברי אינן מתעוררות לגבי שפה זו. אנו לא נתאר כאן בצורה מדויקת את השפה של תורת הקבוצות האקסיומטית. נבטיח עם זאת לקורא, שלהגדרת קבוצות נשתמש אנו רק בשפה מתמטית ברורה, שלכל הנוסחאות בה יהיה מובן חד-משמעי.

בעיה קשה הרבה יותר מציב לפני תורת הקבוצות הפרדוקס הבא:

פרדוקס ראסל:

ותהיו S ה"קבוצה" $\{x \mid x \notin x\}$. אזי

$$S = \{x \mid x \notin x\}$$

[6] ברצוני לנצל הזדמנות זו על מנת להמליץ על ספרה של ענת בילצקי אודות פרדוקסים (בהוצאת האוניברסיטה המשודרת של גל"צ). ספרון מאלף ומהנה זה מתאר הן את הפרדוקסים של תורת הקבוצות והן פרדוקסים מפורסמים בתחומים אחרים. עלי להוסיף, אמנם, שדעות המחבר כאן ודעות המחברת שם חלוקות בתכלית בכל הנוגע לתוכן הפרק, המסיים את ספרה!

לכן: $\forall y.\ y\in S \Leftrightarrow y\in\{x \mid x\notin x\}$

מכאן (בעזרת כלל (β)) $\forall y.\ y\in S \Leftrightarrow y\notin y$

לכן (כלל 14, עם $t = S$) $S\in S \Leftrightarrow S\notin S$

הטענה האחרונה אינה יכולה אבל להיות נכונה. ל- $S\in S$ ול- $S\notin S$ יש תמיד ערך אמת שונה, ולכן הן אינן יכולות להיות שקולות. דבר זה ניתן לברר בקלות בעזרת טבלת האמת של $\Leftrightarrow$ (גם בלעדיה ברור, שמהטענה האחרונה נובע, שאם S שייך ל- S אז S לא שייך ל- S ולהפך – מצב עניינים אבסורדי). הכיצד??

הערה:

השימוש בשם S בתיאור הפרדוקס נועד לצרכים פסיכולוגיים בלבד, ואין הוא למעשה חיוני. בלעדיו היה הפרדוקס נראה כך:

(β) : $\forall y.\ y\in\{x \mid x\notin x\} \Leftrightarrow y\notin y$

לכן: $\{x \mid x\notin x\}\in\{x \mid x\notin x\} \Leftrightarrow \{x \mid x\notin x\}\notin\{x \mid x\notin x\}$

(זאת לפי כלל 14, כשלוקחים $t = \{x \mid x\notin x\}$). המשפט האחרון הוא שוב סתירה לוגית, כי שום טענה אינה יכולה להיות שקולה לשלילתה.

הבעיה שיוצר פרדוקס ראסל קשה שבעתיים מזו שמציג פרדוקס ברי. מצד אחד ברור ממנו, שהביטוי $\{x \mid x\notin x\}$ אינו יכול להגדיר קבוצה. מצד שני הנוסחה $x\notin x$ (כלומר $\neg(x\in x)$) הינה תקינה בהחלט בשפה שנשתמש בה, וגם בשפת תורת הקבוצות האקסיומטית. כיוון שכך, עלינו להשלים עם העובדה, שלא עבור כל נוסחה תקינה ψ מגדיר הביטוי $\{x \mid \psi\}$ *קבוצה* (כלומר אוסף, שהוא בעצמו *עצם*, וממילא הכללים (β) והכלל הבסיסי של $\forall$ חלים עליו). נשאלת אפוא השאלה, כיצד נחליט, איזו ψ "כשרה" היא ואיזו לא? תשובה מקובלת ניתנת במסגרת תורת הקבוצות האקסיומטית הידועה בשם ZF (Zermelo-Fraenkel). ZF כוללת רשימה של אקסיומות, שמרביתן פשוט מבטיחות, לגבי נוסחאות רבות, שהן אכן כשרות לשימוש. אקסיומטיקה זו, כשלעצמה, אינה מטרת קורס זה. ברם, כמעט לכל אקסיומה של ZF מתאים *סימון* נפוץ עבור קבוצות. סימונים אלו וכללים גם וכללים במסגרת מה שכל מי שמשתמש בשפה של תורת הקבוצות חייב לדעת. נביא אפוא עתה את העקרונות של ZF עם הסימונים התואמים. עיקר ענייננו, נדגיש שוב, הוא *בסימונים*: אנו נכניס לשימוש סוגים חדשים של ביטויים ונתאר, מתי ביטויים אלה מייצגים קבוצות, ואילו קבוצות הם מייצגים.

התואמים. עיקר ענייננו, נדגיש שוב, הוא *בסימונים*: אנו נכניס לשימוש סוגים חדשים של ביטויים ונתאר, מתי ביטויים אלה מייצגים קבוצות, ואילו קבוצות הם מייצגים.

עיקרון I:

יש מספר אוספים חשובים שעליהם נניח *במפורש*, שאכן קבוצות הם, ואף נעניק להם שם וסימון מיוחד. בין קבוצות אלה נמנה את:

(א) קבוצת המספרים הטבעיים (כולל 0) אותה נסמן ב- **N**.

(ב) קבוצת המספרים הטבעיים החיוביים ממש (ללא אפס), אותה נסמן ב- $\mathbf{N}^+$.

(ג) קבוצת המספרים השלמים $\{...-2,-1,0,1,2...\}$. סימונה: **Z**.

(ד) קבוצת המספרים הרציונליים (שברים אמיתיים או מדומים, חיוביים או שליליים, כולל 0). סימונה: **Q**.

(ה) קבוצת המספרים הממשיים. סימונה: **R**.

(ו) קבוצת המספרים המרוכבים. סימונה: **C**.

הערות:

(1) כאקסיומה יש להניח, למעשה, רק את קיומה (כקבוצה) של **N**. את כל השאר אפשר לבנות (ולהוכיח את עובדת היותן קבוצות) על סמך אקסיומה זו (בעזרת שאר האקסיומות). לא נעשה זאת בקורס זה[7].

(2) מה שאנו קוראים פה "שמות מיוחדים" נקרא בשפות תכנות "reserved words".

קבוצה חשובה במיוחד נוספת, שאת קיומה נניח (ולמעשה אף *נוכיח* בשלב מאוחר יותר) היא הקבוצה *הריקה*, דהינו הקבוצה, שאין בה איברים כלל (לפי עיקרון האקסטנציונליות תתכן לכל היותר קבוצה אחת כזו, ולכן השימוש בהא הידוע ביחס אליה מוצדק). חשיבותה של קבוצה זו מקנה לה את הזכות לסימון מיוחד משלה. הסימון המקובל הוא $\varnothing$ ("פי", בפא לא דגושה).

את $\varnothing$ ניתן להגדיר כך:

$$\varnothing = \{x \mid x \neq x\}$$

תכונה מיוחדת מאד של $\varnothing$ הינה, שהיא קבוצה חלקית לכל קבוצה אחרת: $\varnothing \subseteq A$ לכל קבוצה A. דבר זה נובע מיד מההגדרות. $\varnothing \subseteq A$ פירושו

$$\forall x.\ x \in \varnothing \Rightarrow x \in A$$

---

[7] המתעניינים בבניית **R**, למשל, יוכלו לעיין בפרקים הראשונים בספרו של מייזלר: "חשבון אינפיניטסימלי".

אבל אם x עצם כלשהו, אז, מהגדרה, $x\notin\varnothing$ מעובדה זו נובע לפי טבלת האמת של $\Rightarrow$, ש- $x\in\varnothing \Rightarrow x\in A$ (כיוון שערך האמת של $x\in\varnothing$ הוא f, ערך האמת של הגרירה כולה הוא t). זה נכון *לכל* עצם x, ולכן $\forall x.\ x\in\varnothing \Rightarrow x\in A$.

הסבר *אינטואיטיבי* למה $\varnothing \subseteq A$, ניתן לתת כך: אם לוקחים קבוצה A ו"זורקים" ממנה איברים בזה אחר זה, נקבל בכל שלב קבוצות מצומצמות יותר ויותר, אך כולן חלקיות כמובן לקבוצה המקורית A. הדעת נותנת אפוא, שמה שמתקבל בסוף התהליך, עת "זרקנו" את כל האיברים ונותרנו עם הקבוצה הריקה, גם הוא קבוצה חלקית של הקבוצה המקורית[8]. (כדאי לשים לב, שגם פה האינטואיציה ש- $\varnothing \subseteq A$ תואמת, ואפילו מכתיבה, את טבלת האמת של $\Rightarrow$).

לסיום עניין השמות המיוחדים נציין, שפרט לשמות הקבועים שמנינו כאן, מקובל לתת מדי פעם שמות *זמניים* לקבוצות בעלות חשיבות בקונטקסט מסוים בו עוסקים (כמו בניסוח משפט או בהוכחה כלשהי). כותבים אז משהו כמו: "נסמן $A = \{x|\psi\}$", או, בניסוח שקול: "נגדיר קבוצה A על-ידי $\forall x(x\in A \Leftrightarrow \psi)$" (את ה- "$\forall x$" משמיטים בדרך כלל בניסוחים אלה. במקום "A" יכול להופיע כאן כל שם אחר. ψ חייבת, כמובן, להיות נוסחה כך ש- $\{x|\psi\}$ היא קבוצה). אנו עצמנו השתמשנו באמצעי זה, עת הענקנו ל"קבוצת" ראסל $\{x|x\notin x\}$ (שאומנם אינה קבוצה כלל) את השם הזמני S. הסיבה שם היתה הרצון להגדיל בהירות ולקצר ניסוחים, וזוהי תמיד הסיבה להכנסת שמות זמניים. לשמות זמניים קוראים בשפות תכנות "קונסטנטות". הכוונה שם היא לשמות, שתוקפם חל רק במסגרת תכנית מסוימת, והוכרזו ככאלה על-ידי כותב אותה תכנית.

עיקרון II:

אם $t_1,\dots,t_n$ הם ביטויים המייצגים עצמים, אז $\{t_1,\dots,t_n\}$ הוא ביטוי המייצג קבוצה. ביטוי זה הינו קיצור של $\{x|x = t_1 \vee x = t_2 \vee \dots \vee x = t_n\}$, כאשר x הוא משתנה כלשהו, שאינו מופיע חופשי ב- $t_1,\dots,t_n$. לכן עבור כל ביטוי s מתקיים:

$$\boxed{s \in \{t_1,\dots,t_n\} \Leftrightarrow s = t_1 \vee \dots \vee s = t_n}$$

---

[8] המדקדקים יוכלו לטעון, שטיעון זה תקף לקבוצות סופיות בלבד. כיוון שמדובר בהסבר אינטואיטיבי, שאיננו, וגם לא מתימר להיות, *הוכחה*, אין עניין זה מדאיג פה. יתר על כן: ברגע שהשתכנענו, כי $\varnothing$ הינה חלקית לכל קבוצה סופית, הרי טרנסיטיביות יחס ההכלה, אותה הראינו למעלה, כופה זאת גם על קבוצות אין-סופיות.

דוגמאות:

(i) הביטוי $\{1,2,3\}$ מייצג קבוצה בת שלושה איברים: המספרים 1,2,3.

(ii) הביטוי $\{\varnothing\}$ מייצגת קבוצה בת איבר אחד. איברה היחיד הוא הקבוצה $\varnothing$. (יש לשים לב, ש- $\{\varnothing\}$ עצמה *אינה* ריקה: יש בה איבר!).

(iii) הביטוי $\{\varnothing,\{\varnothing\}\}$ מיצג קבוצה בת שני איברים: איבריה הם $\varnothing$ ו- $\{\varnothing\}$ (כדאי לשים לב, ש- $\{\varnothing,\{\varnothing\}\}$ היא קבוצה, שאיבריה הם בעצמם קבוצות!).

לקבוצות כמו $\{\varnothing\}$ ו- $\{2\}$, דהינו: קבוצות בעלות איבר יחיד, נוהגים לקרוא ***סינגלטונים***. לעומת זאת, לקבוצה מהצורה $\{t,s\}$ נוהגים לקרוא ***הזוג הלא סדור*** של t ו- s. $\{1,3\}$ ו- $\{\varnothing, \{\varnothing\}\}$ הן דוגמאות לזוגות לא סדורים. כדאי לשים לב, שלפי עיקרון האקסטנציונליות:

(א) $\{t,s\} = \{s,t\}$ (ומכאן השם זוג לא סדור).

(ב) $\{t\} = \{t,t\}$ (ולכן , עקרונית, כל סינגלטון הוא "זוג" לא סדור).

עיקרון III: עיקרון הקומפרהנסיה המוגבל:

אם –

(1) S ביטוי המייצג קבוצה

(2) ψ נוסחה

(3) x משתנה, שאינו חופשי ב- S

אז: הביטוי $\{x{\in}S \mid \psi\}$ הוא ביטוי המייצג קבוצה. ביטוי זה הוא קיצור של $\{x \mid x{\in}S \wedge \psi\}$. לכן:

$$\boxed{t \in \{x{\in}S \mid \psi\} \Leftrightarrow t \in S \wedge \psi(t/x)}$$

עבור כל ביטוי t, החופשי להצבה במקום x ב- ψ.

דוגמאות:

(i) אוסף הפתרונות הממשיים של המשוואה $\sin ax = 0$ הוא הקבוצה

$$\{x \in \mathbf{R} \mid \sin ax = 0\}$$

(כאן ψ מכילה את הפרמטר a, והביטוי מגדיר קבוצה תחת ההנחה, שמציבים במקום a ערכים מ- $\mathbf{R}$).

(ii) קבוצת המספרים הטבעיים הזוגיים היא הקבוצה:

$$\{n \in \mathbf{N} \mid \exists k \in \mathbf{N}.\ n = 2k\}$$

(iii) $\{x \in \mathbf{N} \mid x \neq x\} = \varnothing$. בדומה, לכל A $\{x \in A \mid x \neq x\} = \varnothing$

(iv) $A = \{x \in A \mid x = x\} = \{x \mid x \in A \wedge x = x\} = \{x \mid x \in A\}$.

(השוויון האחרון נובע מעיקרון האקסטנציונליות).

הערות:

(1) ברור ש- $\{x \in A \mid \psi\} \subseteq A$, כלומר: $\{x \in A \mid \psi\}$ הינה תמיד קבוצה חלקית ל - A.

(2) עיקר דוגמה (iv) הוקדש כדי להראות שהביטוי $\{x \mid x \in A\}$ מייצג קבוצה (בתנאי ש- A מייצג קבוצה, כמובן). מ-(iv) נובע אבל גם עיקרון חשוב, אם כי פשוט ביותר, הידוע בשם עיקרון (η)[9]. עיקרון זה קובע, שאם S ביטוי המייצג קבוצה, אז:

$$\boxed{\{x \mid x \in S\} = S}$$

(בתנאי ש- x אינו חופשי בביטוי S).

(3) דוגמה מספר (iii) מראה, שהעובדה, שאכן $\varnothing$ הינה קבוצה, נובעת למעשה מעיקרון הקומפרהנסיה המוגבל, ואינה מצריכה אקסיומה מיוחדת. כמו כן היא מציגה מזווית ראיה שונה את העובדה ש- $\varnothing \subseteq A$ לכל A (על סמך הערה (1)).

(4) **%%** בניסוח עיקרון הקומפרהנסיה המוגבל השמטנו אי אלו תנאים, שעלולים היו לבלבל. גם תנאי (3) שם הוא אולי סתום מעט. כדי להבין אותם נדגיש שוב, שאנו עוסקים בתיאור ***שפה*** ובמשמעות ביטוייה. כשמיישמים את עיקרון הקומפרהנסיה המוגבל, הרי S יהיה ***ביטוי***, אולי קצר יותר, אולי ארוך – ואולי אחד שמכיל משתנים חופשיים. ψ תהיה תמיד ***נוסחה***, ואף היא עשויה להכיל פרמטרים (כלומר: משתנים חופשיים). הביטוי $\{x \in S \mid \psi\}$ עשוי אפוא להכיל פרמטרים. הוא יוכל, כמובן, לייצג קבוצה קונקרטית רק כאשר מציבים ערכים קונקרטיים במקום הפרמטרים – וגם אז ייתכן, שנקבל ביטוי, המייצג קבוצה רק עבור ערכים מסוימים של הפרמטרים. כדוגמה ניקח את דוגמה מס' (i) למעלה: בביטוי $\{x \in \mathbf{R} \mid \sin ax = 0\}$ מופיע הפרמטר a. הביטוי מייצג קבוצות חלקיות ל- $\mathbf{R}$ רק כאשר מציבים במקום a מספרים ממשיים. למשל כש- $a = \pi$ נקבל את $\{x \in \mathbf{R} \mid \sin \pi x = 0\}$, שהיא אכן קבוצה חלקית ל- $\mathbf{R}$ ($\mathbf{Z}$, למעשה).

---

[9] (η) היא האות היוונית אטה. השמות α, β ו η לעקרונות, שכנינו כאן בשם זה, שאולים מתחשיב λ, כיוון שמדובר בעקרונות מאד דומים לאלה שם. על סימון λ ועקרונות α, β ו- η שלו נלמד בהמשך.

כשאנו מבינים כל זאת, נעבור לתנאי (3) בניסוח עיקרון הקומפרהנסיה, ונדגים את נחיצותו. נעיין אפוא בדוגמה הבאה: $S = \{y \mid y = x\}$ ו- $\psi = x \notin x$. עבור כל ערך קונקרטי של המשתנה החופשי x בביטוי S, ייצג ביטוי זה קבוצה (הסינגלטון $\{x\}$). תנאים (1) ו (2) מתקיימים פה לכן. תנאי (3) לא. ואכן :

$$\{x \in S \mid \psi\} = \{x \in \{y \mid y = x\} \mid x \notin x\}$$

$= \{x \mid x \in \{y \mid y = x\} \wedge x \notin x\}$ (לפי הגדרת הסימון $\{x \in \ldots \mid \_\}$)

$= \{x \mid x = x \wedge x \notin x\}$ (לפי עיקרון (β), מופעל על $x \in \{y \mid y = x\}$)

$= \{x \mid x \notin x\}$ (כי $x = x$ נכון לכל x)

ברם, $\{x \mid x \notin x\}$ הינו ביטוי, ***שאינו*** מייצג קבוצה, כמו שראינו עת תיארנו את פרדוקס ראסל.

%%

תרגילים

(א) הוכח: $\{x \in \{\varnothing, \{\varnothing\}\} \mid x \neq \varnothing\} = \{\{\varnothing\}\}$

(ב) הוכח: $\forall p.\ p \in \mathbf{N} \wedge p \neq 0 \Rightarrow 9p \in \{x \in \{n \in \mathbf{N} \mid \exists k \in \mathbf{N}.\ n = pk\} \mid x \geq 9\}$

(ג) $\forall x \in \{x \in A \mid \varphi\}.\psi \Leftrightarrow \forall x \in A(\varphi \Rightarrow \psi)$

(ד) $\exists x \in \{x \in \mathrm{A} \mid \varphi\}.\psi \Leftrightarrow \exists x \in A(\varphi \wedge \psi)$

(ה) הביטוי $\{x \mid x = x\}$ אינו מיצג קבוצה (במלים אחרות : הוא אינו מוגדר).

עיקרון IV: עיקרון ההחלפה

אם –

(1) S ביטוי המייצג קבוצה

(2) x משתנה שאינו חופשי ב- S

(3) t ביטוי

אז: $\{t \mid x \in S\}$ הוא ביטוי המייצג קבוצה. ביטוי זה הוא קיצור של $\{y \mid \exists x \in S.\ y = t\}$, כאשר y הוא משתנה כלשהו, שאינו חופשי ב- S או ב- t. מכאן שאם t' ביטוי ש- x אינו מופיע בו חופשי, אז:

$$\boxed{t' \in \{t \mid x \in S\} \Leftrightarrow \exists x \in S.\ t' = t}$$

דוגמאות

(א) $\{2n \mid n\in\mathbf{N}\}$ היא קבוצת המספרים הזוגיים.

(ב) $\{\{n\} \mid n\in\mathbf{N}\}$ היא קבוצת כל הסינגלטונים של מספרים טבעיים (כלומר: קבוצת כל הקבוצות המהוות סינגלטון, שאיברו היחיד הוא מספר טבעי).

אם ננתח את (ב) לפי עיקרון ההחלפה, הרי הביטוי S הוא כאן השם המיוחד (והקבוע) $\mathbf{N}$ (שמייצג קבוצה!), בתפקיד x משחק כאן המשתנה n, ו- t הינו $\{n\}$. $\{\{n\} \mid n\in\mathbf{N}\}$ הוא, לכן, קיצור של $\{z \mid \exists n\in\mathbf{N}.z = \{n\}\}$ (למה דווקא z ? ובכן, למה לא?).

אינטואיטיבית:

$$\{\{n\} \mid n\in\mathbf{N}\}=\{\{0\},\{1\},\{2\},...\}$$

נדגיש, עם זאת, ששפתנו הרשמית אינה כוללת ביטויים כמו זה שבאגף ימין של "המשוואה" האחרונה, כיוון שאין כללים ברורים, מה 3 הנקודות מייצגות בכל מקרה ומקרה. עם זאת, כיוון שהקוראים בני אדם הם ולא מחשבים, סביר להניח, כי יוכלו לנחש, למה הכוונה כאן!

עיקרון V: עיקרון קבוצה החזקה

אם A קבוצה אז אוסף כל הקבוצות החלקיות שלה הוא גם קבוצה. אוסף זה נקרא *קבוצת החזקה* של A.

סימון : אם S ביטוי המייצג קבוצה (כולל המקרה ש- S הוא משתנה!) אז $P(S)$ מייצג את קבוצת החזקה של קבוצה זו. (סימון אחר: 2^S).

הקבוצה $P(A)$ מוגדרת אפוא בצורה הבאה:

$$P(A) = \{x \mid x\subseteq A\}$$

ולכן אם S ביטוי המייצג קבוצה, ו- t הינו ביטוי, אז:

$$\boxed{t \in P(S) \Leftrightarrow t\subseteq S}$$

הערות

(1) כרגיל, בכותבנו הגדרה כמו זו של $P(A)$ למעלה, יש להתייחס אליה (כשמפעילים אותה) כאילו היה כתוב: $\forall A.\ P(A) = \{x \mid x \subseteq A\}$. הערה דומה תחול על כל הגדרה באותו סגנון בעתיד.

(2) ברוב המכריע של ספרי הלימוד כותבים פשוט: "אם A קבוצה אז $P(A)$ מסמן את קבוצת החזקה של A" (ולא את הדבר המסורבל שאנו כתבנו). ניסוח זה איננו מדויק, למעשה. אם ניקח, למשל, את קבוצת חברי הכנסת, הרי לא נוכל לשים אותם בסוגריים ולכתוב את האות "P" לפני התוצאה. מה שנוכל לשים בסוגריים הוא ***ביטוי*** לשוני, המתאר או ***מייצג*** את קבוצת חברי הכנסת, לא את הקבוצה עצמה. עם זאת, יש להודות שהניסוח שלנו, המדבר על ביטויים, אכן ***מסורבל*** הוא. כיוון שכך, נתחיל גם אנו להשתמש יותר ויותר בנוסח המקובל, הפשוט יותר לקריאה. תקוותנו ואמונתנו, שבשלב זה יבינו תמיד הקוראים, מה היה צריך לכתוב ***באמת***.

עובדות ברורות:

(1) $A \in P(A)$

(2) $\varnothing \in P(A)$

(3) $\{x \in A \mid \varphi\} \in P(A)$

דוגמאות :

(1) $P(\varnothing) = \{\varnothing\}$

(2) $P(\{1\}) = \{\varnothing, \{1\}\}$

(3) $P(\{1,2\}) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$

הערה:
בהקשר של הדוגמה הראשונה כדאי לשים לב, ש- $\varnothing \subseteq \varnothing$, ולכן $\varnothing \in \{\varnothing\}$.
כמו כן, $\varnothing \subseteq \{\varnothing\}$. עם זאת, $\varnothing \notin \varnothing$.

הדוגמאות מעלות בצורה ברורה את ההשערה הבאה, שאת הוכחתה באינדוקציה נשאיר לקוראים (כמו כן עוד נחזור למשפט למטה, בחלק שיוקדש לקומבינטוריקה).

משפט: אם A קבוצה סופית בת n איברים אז ב- $P(A)$ יש 2^n איברים.

בכך סיימנו כמעט לסקור את העקרונות הבסיסיים, המאפשרים יצירת קבוצות באופן קונסטרוקטיבי (כולל סימון מתאים). נותר עוד עיקרון אחד, שייסקר (במסגרת נושאים נוספים) בפרק הבא.

לסיום נזכיר עוד את הסימונים המקובלים הבאים מחדו"א עבור קבוצות חלקיות חשובות של $\mathbf{R}$:

$$(-\infty,a) = \{x\in \mathbf{R} \mid x < a\} \qquad (a,b) = \{x\in \mathbf{R} \mid a < x < b\}$$
$$(-\infty,a] = \{x\in \mathbf{R} \mid x \le a\} \qquad (a,b] = \{x\in \mathbf{R} \mid a < x \le b\}$$
$$(a,\infty) = \{x\in \mathbf{R} \mid x > a\} \qquad [a,b) = \{x\in \mathbf{R} \mid a \le x < b\}$$
$$[a,\infty) = \{x\in \mathbf{R} \mid x \ge a\} \qquad [a,b] = \{x\in \mathbf{R} \mid a \le x \le b\}$$

נספח: על משוואות ושאילתות

במערכות בסיסי נתונים (DBMS) מסתכלים על $\{x \mid \varphi\}$ כעל ***שאילתה***: מצא את כל העצמים המקיימים את התנאי φ. בדרך-כלל רוצים, שלשאילתה יהיה מספר סופי של תשובות, כך שאפשר יהיה לסדרן בטבלה (וכן שאפשר יהיה למצוא תשובות אלו באופן אפקטיבי). אחת הדרכים המרכזיות להשיג זאת היא להגביל את השאילתות המותנות למה שנקרא "שאילתות בטוחות". צורת שאילתות אלו היא $\{x\in A \mid \varphi\}$, כאשר A היא קבוצה סופית.

בהסתכלות על $\{x\in A \mid \varphi\}$ כעל שאילתה אין, למעשה, כל חידוש עבורנו. כשאנו נדרשים לפתור בביה"ס התיכון משוואה כמו $\sin x = 0$, אנו נדרשים למצוא את ***קבוצת*** הפתרונות הממשיים של משוואה זו. תיאור מדויק של הקבוצה המבוקשת ניתן על ידי: $\{x\in \mathbf{R} \mid \sin x = 0\}$. ברם תאור זה, מדויק ככל שיהיה, לא כל כך מספק את המורה (ודי בצדק). ברור מעמיק יותר מגלה, שמה שאנו באמת נדרשים לעשות, הוא לתת תיאור פשוט יותר, בעל צורה מסוימת, ***קנונית***, לקבוצת הפתרונות. מהי האורה הנחשבת לקנונית בתיכון הוא דבר המשתנה לפי סוג הבעיה:

(א) כשמספר הפתרונות הוא סופי אנו רוצים ייצוג מהצורה $\{a_1, \ldots, a_n\}$ למשל:

$$\{x\in \mathbf{R} \mid x^2 - 3x + 2 = 0\} = \{1,2\}$$

(ב) במשוואות טריגונומטריות רוצים "הצגה פרמטרית", כמו :

$$\{x\in \mathbf{R} \mid \sin x = 0\} = \{k\pi \mid k\in\mathbf{Z}\}$$

(או איחוד של קבוצות כאלו. על "איחוד" נדבר בפרק הבא).

(ג) ב"אי-שוויונים" רוצים הצגה כאיחוד של קטעים :

$$\{x\in \mathbf{R} \mid x^2 > 1\} = \{x\in \mathbf{R} \mid x < -1 \vee x > 1\} \quad [= (-\infty,-1) \cup (1,\infty)]$$

אנו רואים אפוא, ש"משוואות" ו"אי-שוויונים" אינם אלא ייצוגים מסוימים של קבוצות של מספרים, ו"פתרונם" הוא ייצוג אותן קבוצות בצורה "קנונית", נוחה יותר לשימוש. ה"נעלמים" במשוואות הינם לכן משתנים קשורים (על-ידי $\{\ldots \mid \_\ \}$). ב"משוואות עם פרמטרים" הפרמטרים אכן כשמם כן הם: משתנים חופשיים.

את המשפט המרכזי על פתרון משוואות ריבועיות נוכל לנסח עתה כך:

$$\forall a\in \mathbf{R}\ \forall b\in \mathbf{R}\ \forall c\in \mathbf{R}\ \Big[a \neq 0 \Rightarrow$$

$$\Big((b^2 - 4ac < 0 \Rightarrow$$

$$\{x\in \mathbf{R} \mid ax^2 + bx + c = 0\} = \varnothing) \wedge$$

$$(b^2 - 4ac \geq 0 \Rightarrow$$

$$\{x\in \mathbf{R} \mid ax^2 + bx + c = 0\} = \left\{\frac{-b+\sqrt{b^2-4ac}}{2a}, \frac{-b-\sqrt{b^2-4ac}}{2a}\right\})\Big)\Big]$$

ב.3. פעולות יסודיות על קבוצות

בפרק זה נכיר את הפעולות הבסיסיות הקשורות לקבוצות ונלמד את תכונותיהן.

(א) חיתוך, איחוד, משלים, הפרש

הגדרה

א. ***האיחוד*** של שתי קבוצות היא הקבוצה המכילה את כל העצמים השייכים לפחות לאחת משתי הקבוצות.
האיחוד של שתי קבוצות A ו- B מסומן ב- $A\cup B$.

ב. ***החיתוך*** של שתי קבוצות היא הקבוצה המכילה את כל העצמים השייכים לשתי הקבוצות גם יחד.
החיתוך של שתי קבוצות A ו- B מסומן ב- $A\cap B$.

ג. ***ההפרש*** של שתי קבוצות, A ו- B, הוא קבוצת העצמים השייכים ל- A אך לא שייכים ל- B.
ההפרש של A ו- B מסומן ב- $A-B$.

ד. ***ההפרש הסימטרי*** של שתי קבוצות הוא אוסף העצמים הנמצאים בדיוק באחת משתי הקבוצות (כלומר: נמצאים באחת אך לא נמצאים באחרת).
ההפרש הסימטרי של A ו- B מסומן ב- $A \Delta B$.

הגדרה קצרה יותר של הפעולות $\cup$, $\cap$, $-$, Δ ניתנת על-ידי:

1. $A\cup B = \{x \mid x\in A \vee x\in B\}$
2. $A\cap B = \{x \mid x\in A \wedge x\in B\} = \{x\in A \mid x\in B\}$
3. $A-B = \{x \mid x\in A \wedge x\notin B\} = \{x\in A \mid x\notin B\}$
4. $A \Delta B = \{x \mid (x\in A \wedge x\notin B) \vee (x\in B \wedge x\notin A)\} = (A-B)\cup(B-A)$

הערות:

1. נזכיר שוב, כי להגדרות מהסוג של 1-4 יש להתייחס כאילו הן טענות עם כמתים כוללים $\forall A\forall B$ בתחילתם (וממילא כלל α וכללי $\forall$ ישימים).

2. מהשוויון השני, הן ב- 2 והן ב- 3, נובע, שאם A ו- B הן קבוצות, אז גם $A \cap B$ ו- $A - B$ הן קבוצות (לפי עקרון הקומפרהנסיה המוגבל). בהנחה שאיחוד של שתי קבוצות נותן קבוצה, נובע מזה ומהשוויון השני ב- 4, ש- $A \Delta B$ קבוצה (כאשר A ו- B קבוצות). לעומת זאת אין העובדה, ש- $A \cup B$ היא קבוצה, נובעת משום עיקרון, שלמדנו עד כה. נחוץ לכך עקרון מיוחד:

עקרון האיחוד, נוסח חלש:

האיחוד של שתי קבוצות הוא קבוצה. (ממילא, כש- A ו- B ביטויים המגדירים קבוצות אז הביטוי $\{x | x \in A \vee x \in B\}$ מגדיר קבוצה).
ניסוח אחר לעקרון זה הוא: אם הביטויים $\{x \mid \varphi\}$ ו- $\{x \mid \psi\}$ מגדירים קבוצות, אז גם הביטוי $\{x \mid \varphi \vee \psi\}$ מגדיר קבוצה.

דוגמאות:

(א) נניח:

$$A = [0,2] = \{x \in \mathrm{R} |\ 0 \le x \le 2\}$$
$$B = [1,4] = \{x \in \mathrm{R} |\ 1 \le x \le 4\}$$

אזי:

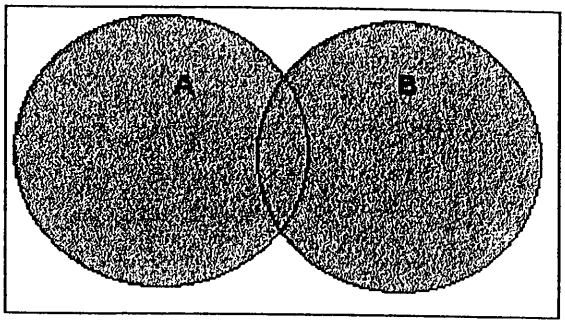
$$A \cup B = [0,4] = \{x \in \mathrm{R} |\ 0 \le x \le 4\}$$
$$A \cap B = [1,2] = \{x \in \mathrm{R} |\ 1 \le x \le 2\}$$
$$A - B = [0,1) = \{x \in \mathrm{R} |\ 0 \le x < 1\}$$
$$B - A = (2,4] = \{x \in \mathrm{R} |\ 2 < x \le 4\}$$
$$A \Delta B = \{x \in \mathrm{R} |\ 0 \le x < 1 \vee 2 < x \le 4\}$$

(ב)

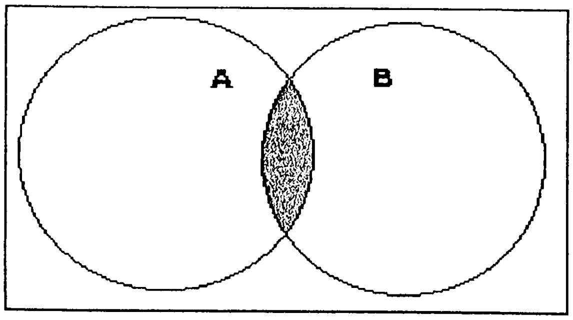
$$\mathrm{P(N)} \cap \mathrm{P}(\{\varnothing, 1\}) = \{ \varnothing, \{1\} \}$$
$$\mathrm{P}(\{\varnothing, 1\}) - \mathrm{P(N)} = \{ \{\varnothing\}, \{\varnothing, 1\} \}$$
$$\mathrm{P(N)} \cup \mathrm{P}(\{\varnothing, 1\}) = \{x \mid x \subseteq \mathrm{N} \vee x = \{\varnothing\} \vee x = \{\varnothing, 1\}\}$$

(ג) המחשה של האופרציות בעזרת דיאגרמות ואן אפשר למצוא בציור מס' ב.2 באיור הבא.

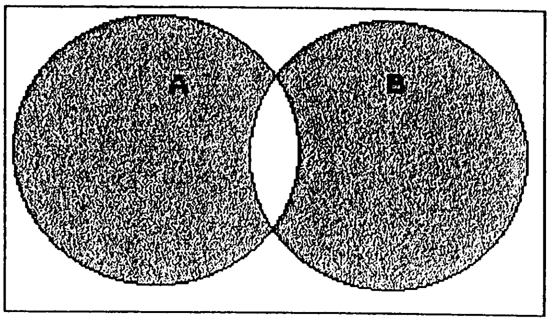
ציור ב.2: דיאגרמת וון עבור הפעולות



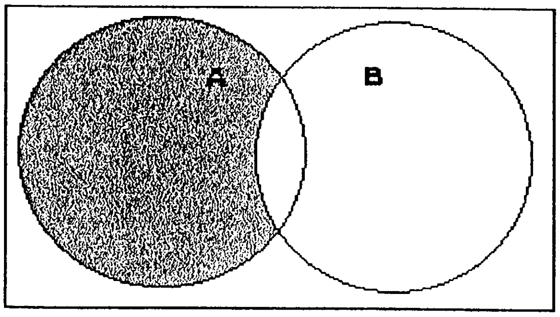
$A \cup B$



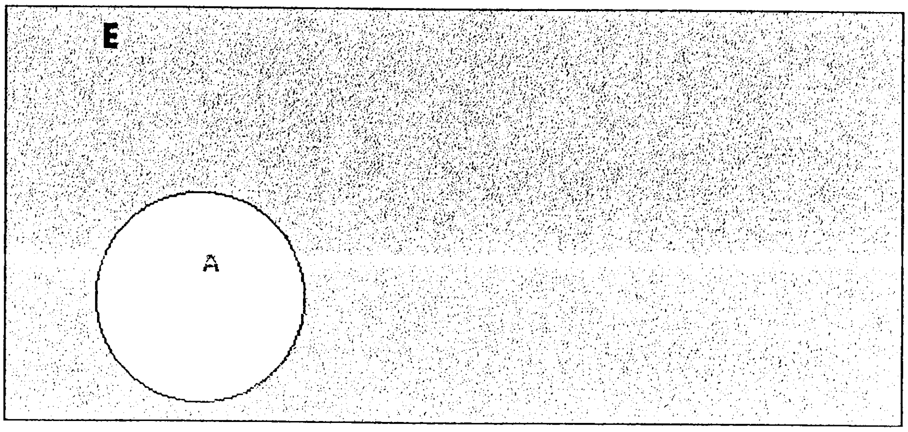
$A \cap B$



$A \Delta B$



$A - B$



$\overline{A}_E$

ההגדרות של חיתוך ואיחוד מראות על קשר הדוק בינן ובין הקשרים הלוגיים $\wedge$ ו- $\vee$. לעומת זאת, יש לשים לב, שאם A קבוצה, אז $\{x|\ x \notin A\}$ ***אינו*** ביטוי המגדיר קבוצה. ואכן, לו היה מגדיר קבוצה, אז לפי עקרון האיחוד גם $A \cup \{x|\ x \notin A\}$ היה מגדיר קבוצה. אבל:

$$A \cup \{x|\ x \notin A\} = \{x|\ x\in A\} \cup \{x|\ \neg\,(x \in A)\} =$$

$$= \{x|\ x\in A \vee \neg\,(x\in A)\} = \{x|\ x = x\}$$

וזאת כיון ש- $x\in A \vee \neg(x\in A)$ הינה אמת לוגית, הנכונה לכל x ו- A. ברם, ראינו למעלה, שהביטוי $\{x|\ x = x\}$ אינו ביטוי המגדיר קבוצה. מכאן, שגם הביטוי $\{x|\ x \notin A\}$ אינו מגדיר קבוצה. עם זאת, ברוב הסיטואציות בהם עוסקים בקבוצות, יש איזו "קבוצת גג" E, שכל הקבוצות המעניינות אותנו חלקיות לה. במקרה זה, אם $A \subseteq E$, נוהגים לקרוא ל- $E - A$ ***"המשלים של A ביחס ל- E"*** ומסמנים קבוצה זו ב- $\overline{A_E}$ או A_E^c. למעשה, בדרך כלל משמיטים את ה- "E" בה מדובר, מדברים פשוט על "המשלים של A" ומסמנים זאת ב- $\overline{A}$ או ב- A^c. יש לזכור אבל, שזהותה המדויקת של $\overline{A}$ תלויה בשאלה: מי היא הקבוצה E אותה לוקחים אנו כקבוצת הגג.

המחשה של $\overline{A_E}$ על ידי דיאגרמתוון ניתן למצוא בציור ב.1.

(ב) התכונות הבסיסיות של הפעולות הבסיסיות

בטבלה מספר ב.1 (עמ' 87) מרוכזות התכונות היסודיות של הפעולות, שהוגדרו בסעיף הקודם (למעט ההפרש הסימטרי, שחשיבותו פחותה). אין צורך להיבהל ממספרן הרב: מרביתן מובנות מאליהן. אנו אכן לא נטרח להוכיח את כל החוקים בטבלה. את מרביתן נשאיר כתרגילים לקוראים. כאן נסתפק רק במספר דוגמאות מייצגות.

♦ דוגמה 1:

נוכיח את 5b :

הוכחה א':

הוכחה זו מסתמכת על השקילויות בטבלה א.3 (עמודים 20-21) הקוראים מתבקשים למצוא באילו שקילויות (פרט להגדרת $\subseteq$ ו- $\cup$) אנו משתמשים בכל שלב.

$$
\begin{aligned}
A \cup B \subseteq C \quad & \Leftrightarrow \quad \forall x.\ x \in A \cup B \Rightarrow x \in C \\
& \Leftrightarrow \quad \forall x\,((x \in A \vee x \in B) \Rightarrow x \in C) \\
& \Leftrightarrow \quad \forall x\,((x \in A \ \Rightarrow x \in C) \wedge (x \in B \Rightarrow x \in C)) \\
& \Leftrightarrow \quad \forall x\,(x \in A \ \Rightarrow x \in C) \wedge \forall x\ (x \in B \Rightarrow x \in C) \\
& \Leftrightarrow \quad A \subseteq C \wedge B \subseteq C
\end{aligned}
$$

הוכחה ב':

זוהי צורת ההוכחה הסטנדרטית. היא נעשית בעזרת כללי ההיסק הלוגים של טבלה א.4 (עמ' 26). עם זאת, כדי להבין אותה אין צורך להכיר טבלה זו!

ראשית, כיוון שמדובר בהוכחת שקילות, אנו מפרקים זאת לשני חלקים:

הוכחה ש: $A \cup B \subseteq C \ \Rightarrow \ A \subseteq C \wedge B \subseteq C$ (הכיוון $\Rightarrow$)

והוכחה ש: $A \cup B \subseteq C \ \Leftarrow \ A \subseteq C \wedge B \subseteq C$ (הכיוון $\Leftarrow$)

הוכחת ($\Rightarrow$) :

נניח ש- $A \cup B \subseteq C$. נוכיח ש- $A \subseteq C$ וש- $B \subseteq C$. נראה למשל, ש- $A \subseteq C$ (הוכחת $B \subseteq C$ היא דומה). יהי אפוא $x \in A$. כיוון ש- $A \subseteq A \cup B$, נובע מזה ש- $x \in A \cup B$. מזה ומהנתון $A \cup B \subseteq C$ נובע ש- $x \in C$.

הוכחת ($\Leftarrow$) :

נניח ש- $A \subseteq C$ וש- $B \subseteq C$. נראה ש- $A \cup B \subseteq C$. נניח לכן ש- $x \in A \cup B$ ונראה ש- $x \in C$. כיוון ש- $x \in A \cup B$, יש שתי אפשרויות :

אפשרות (i): $x \in A$. כיוון שנתון ש- $A \subseteq C$, נובע מזה ש- $x \in C$.

אפשרות (ii): $x \in B$. כיוון שנתון ש- $B \subseteq C$, נובע מזה ש- $x \in C$.

בשני המקרים $x \in C$, כמבוקש.

הערות:

1. במהלך ההוכחה השתמשנו בתכונה 2a מטבלת התכונות היסודיות. הנחנו, אם כך, שתכונה זו כבר הוכחה (ההוכחה היא אכן טריוויאלית במיוחד). למעשה, היה קצר יותר שלא להשתמש בה כלל ולכתוב כך:

   "יהי אפוא $x \in A$. מזה נובע $x \in A \vee x \in B$ ולכן $x \in A \cup B$".

   מקובל יותר אבל בטקסטים לכתוב בצורה שכתבנו למעלה.

2. "החלוקה למקרים" אינה אלא שימוש בכלל 6 מטבלה א.4.

טבלה ב.1: תכונות יסודיות של האופרציות הבסיסיות

1a $A \cap B \subseteq A$	1b $A \cap B \subseteq B$
2a $A \subseteq A \cup B$	2b $B \subseteq A \cup B$
3a $C - A \subseteq C$	3b $A \subseteq B \Rightarrow C - B \subseteq C - A$
4a $A \subseteq B \Leftrightarrow A \cap B = A$	4b $A \subseteq B \Leftrightarrow A \cup B = B$
5a $C \subseteq A \cap B \Leftrightarrow C \subseteq A \wedge C \subseteq B$	5b $A \cup B \subseteq C \Leftrightarrow A \subseteq C \wedge B \subseteq C$
6a $A \cap A = A$	6b $A \cup A = A$
7a $A \cap B = B \cap A$	7b $A \cup B = B \cup A$
8a $(A \cap B) \cap C = A \cap (B \cap C)$	8b $(A \cup B) \cup C = A \cup (B \cup C)$
9a $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	9b $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
10a $A \cap \varnothing = \varnothing$	10b $A \cup \varnothing = A$
11a $A - A = \varnothing$	11b $A - \varnothing = A$
12. $C - (C - A) = C \cap A$	
13a $C - (A \cap B) = (C - A) \cup (C - B)$	13b $C - (A \cup B) = (C - A) \cap (C - B)$
14a $A \cap (C - A) = \varnothing$	14b $A \cup (C - A) = C \cup A$
12 *. $\overline{\overline{A}} = A$	
13a * $\overline{A \cap B} = \overline{A} \cup \overline{B}$	13b * $\overline{A \cup B} = \overline{A} \cap \overline{B}$
14a * $A \cap \overline{A} = \varnothing$	14b * $A \cup \overline{A_E} = E$
15. $A - B = A \cap \overline{B}$	

הוכחה ג':

גם בהוכחה זו אנו מפרקים את מה שצריך להוכיח לשני חלקים, אך משתמשים בשביל כל אחד בתכונות פשוטות יותר מהטבלה (אותן צריך, כמובן, להוכיח קודם).

הוכחת ($\Rightarrow$) :

נניח ש- $A \cup B \subseteq C$, אז לפי 4b: $A \cup B \cup C = C$, כיוון ש- $A \subseteq A \cup B \cup C$ ו- $B \subseteq A \cup B \cup C$, נובע מזה ש- $A \subseteq C$ וש- $B \subseteq C$.

הוכחת ($\Leftarrow$) :

נניח ש- $A \subseteq C$ וש- $B \subseteq C$, אז לפי 4b: $A \cup C = C$ ו- $B \cup C = C$.

לפי כללים 6-8:

$$(A \cup B) \cup C = (A \cup B) \cup (C \cup C) = (A \cup C) \cup (B \cup C) = C \cup C = C$$

לכן, לפי 4b, $A \cup B \subseteq C$.

מוסר השכל: לטענות טריוויאליות אפשר למצוא הרבה הוכחות טריוויאליות ...

♦ דוגמה 2:

נוכיח: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ 9a.

הוכחה א':

$$\begin{aligned} x \in A \cap (B \cup C) \quad &\Leftrightarrow \quad x \in A \ \wedge \ x \in B \cup C \\ &\Leftrightarrow \quad x \in A \ \wedge \ (x \in B \vee x \in C) \\ &\Leftrightarrow \quad (x \in A \ \wedge \ x \in B) \vee (x \in A \wedge x \in C) \\ &\Leftrightarrow \quad x \in A \cap B \vee \ x \in A \cap C \\ &\Leftrightarrow \quad x \in (A \cap B) \cup (A \cap C) \end{aligned}$$

הוכחה ב':

($\subseteq$) :

נניח $x \in A \cap (B \cup C)$ אז $x \in A$ וגם $x \in B \cup C$, לכן $x \in B$ או $x \in C$.

אפשרות (i): $x \in B$.

במקרה זה, כיוון ש- $x \in A$, נקבל ש- $x \in A \cap B$.

כיוון ש- $A \cap B \subseteq (A \cap B) \cup (A \cap C)$, נקבל ש- $x \in (A \cap B) \cup (A \cap C)$.

אפשרות (ii): $x \in C$

במקרה זה , כיוון ש- $x \in A$, נקבל ש- $x \in A \cap C$.

לכן גם כאן נקבל ש- $x \in (A \cap B) \cup (A \cap C)$.

$(\supseteq)$:

לשם הגיוון, נוכיח כיוון זה בעזרת תכונות קודמות בדף.

ובכן, כיוון ש- $A \cap B \subseteq A$ ו- $A \cap C \subseteq A$, נובע, לפי 5b, ש-

$$(A \cap B) \cup (A \cap C) \subseteq A$$

כמו כן, $A \cap B \subseteq B \subseteq B \cup C$ ו- $A \cap C \subseteq C \subseteq B \cup C$.

לכן, שוב לפי 5b, מתקיים גם:

$$(A \cap B) \cup (A \cap C) \subseteq B \cup C$$

משתי העובדות שהוכחנו נובע, לפי 5a, ש:

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)$$

הערה:

בעתיד, כאשר נרצה להשתמש בתכונות יסודיות מהדף (כמו 5a ו- 5b), לא נציין בפירוש, באיזו תכונה אנו משתמשים, אלא פשוט נשתמש בה!

הערות:

1. לחלק מהחוקים יש שמות, שמן הראוי לזוכרם. הכללים ב- 6 נקראים החוקים האידמפוטנטיים, הכללים ב- 7: החוקים הקומוטטיביים (חוקי החילוף), הכללים ב- 8: החוקים האסוציאטיבים (חוקי הקיבוץ), הכללים ב- 9: החוקים הדיסטריבוטיבים (חוקי הפילוג), והכללים ב- 13: חוקי דה-מורגן. כדאי לציין, שלחלק מהשקילויות הלוגיות יש שמות דומים. דמיון זה אינו מקרי, כמובן, אלא מצביע על קשר אמיץ בין החוקים שם וכאן. כך, למשל, אם נבדוק את ההוכחה הראשונה שהבאנו לחוק הפילוג 9a, נראה, שפרט להגדרת חיתוך ואיחוד, הדבר היחיד, שהשתמשנו בו שם, הוא חוק פילוג לוגי מתאים (שקילות 4a בטבלה א.3, עמוד 20).

2. בגלל חוק הקיבוץ 8 הרשינו לעצמנו בהוכחות לכתוב $A \cup B \cup C$, בלי לפרט אם הכוונה ל- $(A \cup B) \cup C$ או ל- $A \cup (B \cup C)$ גם בהמשך ננהג כך (גם לגבי $\cap$ כמובן). כללית, נובע מחוקים 6-8, שבביטוי מהצורה $A_1 \cup A_2 \cup \ldots \cup A_n$ (או $A_1 \cap A_2 \cap \ldots \cap A_n$) אין חשיבות למיקום הסוגריים, לסדר האיברים או למספר הפעמים, שכל אחד מהם מופיע.

3. את חוקי פעולת ההפרש אין, למעשה, צורך לזכור. במקום זאת די לזכור את חוקי פעולת המשלים (להם יש כללים לוגים מקבילים). כל חוקי פעולת ההפרש ינבעו אז מהזהות (15):

$$A - B = A \cap \overline{B}$$

כאשר המשלים כאן, $\overline{B}$, יכול להילקח יחסית לכל קבוצה, המכילה גם את A וגם את B ($A \cup B$, למשל). לדוגמה:

$$C - (A \cup B) = C \cap (\overline{A \cup B}) = C \cap (\overline{A} \cap \overline{B}) =$$
$$= (C \cap \overline{A}) \cap (C \cap \overline{B}) = (C - A) \cap (C - B)$$

בחישוב זה פעולת המשלים נלקחה יחסית ל- $A \cup B \cup C$ (למשל), והשתמשנו בחוקי האיחוד, החיתוך והמשלים בלבד.
מאידך גיסא, כל החוקים של פעולת המשלים הם מקרים פרטיים של חוקי ההפרש, כאשר מניחים אנו ש- $A \subseteq E$ ו- $B \subseteq E$ (E – הקבוצה ביחס אליה נלקח המשלים). לדוגמה, *12 קובע, בעצם, שכאשר $A \subseteq E$ אז $E - (E - A) = A$. זה נובע מיד מ- 12 ומ- 4a.

4. הפעולות הבסיסיות על קבוצות ידועות גם בשם הפעולות הבוליאניות. הכללים שבטבלה ב.1 נכונים לא רק לקבוצות, אלא לכל מבנה מתמטי המהווה מה שנקרא "אלגברה בוליאנית". לא נוכל כאן, לצערנו, להרחיב את הדיבור על אלגבראות בוליאניות מעבר לרמז זה.

(ג) הכללת הפעולות של איחוד וחיתוך

עד כה הגדרנו איחוד של *שתי* קבוצות וחיתוך של *שתי* קבוצות. נכליל עתה פעולות אלו למקרה כללי הרבה יותר.

הגדרה:

תהי F קבוצה, שכל איבריה אף הן קבוצות (לקבוצה הזו קוראים לעתים קרובות, מסיבות בלתי ברורות, ***משפחה*** של קבוצות). ***האיחוד של F*** מוגדר כקבוצת כל העצמים, שנמצאים בלפחות אחת מהקבוצות ב- F. אם F אינה ריקה, אז ***החיתוך של F*** מוגדר כקבוצת כל העצמים, הנמצאים בכל איברי F.

סימונים:

אם A ביטוי המיצג קבוצה, אז "$\cup(S)$" המייצג את האיחוד של קבוצה זו, בעוד "$\cap(S)$" ביטוי המייצג את החיתוך שלה.

את ההגדרות אפשר לסכם כך:

$$\cup (A) = \{ x \mid \exists\ z \in A \,.\, x \in z \}$$

$$\cap (A) = \{ x \mid \forall z \in A \,.\, x \in z \}$$

יש לשים לב, שהביטוי $\cap (F)$ אינו מוגדר כש- F קבוצה ריקה. זאת, כיוון שלפי ההגדרה הפורמלית לעיל, $\cap (\varnothing)$ היה צריך להיות אוסף כל העצמים, ואוסף זה אינו קבוצה. לעומת זאת אם $B \in F$, אז :

$$\cap (F) = \{ x \in B \mid \forall z \in F \,.\, x \in z \}$$

ולכן $\cap (F)$ הינו קבוצה, כש- F אינה ריקה (לפי עיקרון הקומפרהנסיה המוגבל). לעומת זאת, כדי להבטיח ש- $\cup (F)$ הינו קבוצה אנו זקוקים לעקרון מיוחד:

VI. עיקרון האיחוד (נוסח מלא):

אם F היא קבוצה אז כך גם האיחוד שלה[10].

בניסוח המתייחס לשפה ננסח את העיקרון כך:
אם A ביטוי המייצג קבוצה, אז כך גם הביטוי $\{x \mid \exists\, z \in A.\, x \in z\}$, דהיינו: $\cup (A)$.

דוגמאות:

(1) נניח A ו- B קבוצות, ותהי $F = \{A, B\}$. אז:

$$\cup (F) = A \cup B \quad \text{ו-} \quad \cap (F) = A \cap B \qquad \text{(לבדוק!)} .$$

פעולות האיחוד והחיתוך הכלליות, שהגדרנו כאן, הן לכן הכללה של הפעולות, שהגדרנו קודם.

(2) $$\cup (\{ \{n\} \mid n \in \mathbf{N}\}) = \mathbf{N}$$

הוכחה:

לכיוון האחד, נניח $x \in \mathbf{N}$. אז $x \in \{x\}$, ו- $\{x\} \in \{\{n\} \mid n \in \mathbf{N}\}$. לכן :

$$\exists z \,.\, z \in \{\{n\} \mid n \in \mathbf{N}\} \wedge x \in z \qquad (z = \{x\})$$

---

[10] קורא קשוב שם לב, אני מקווה, שהושמט כאן התנאי ש- F היא קבוצה של קבוצות. למעשה, הן ההגדרות והן העיקרון ישימים גם כש- F היא קבוצה כלשהי, לאו דווקא "משפחה". באופן מעשי משתמשים בהם, אבל, רק במקרה המשפחתי.

כלומר:

$$x \in \cup(\{\{n\} \mid n \in \mathbf{N}\})$$

לכיוון ההפוך, נניח $x \in \cup(\{\{n\} \mid n \in \mathbf{N}\})$ אז יש z ב- $\{\{n\} \mid n \in \mathbf{N}\}$ כך ש- $x \in z$. z זה הוא מהצורה $\{n\}$ עבור איזה $n \in \mathbf{N}$. מזה נובע, שקיים n ב- $\mathbf{N}$ כך ש- $x \in \{n\}$. אבל $x \in \{n\}$ פירושו ש- $x = n$. ממילא, אם $n \in \mathbf{N}$, אז גם $x \in \mathbf{N}$.

(3)

$$\bigcup\left(\left\{\{x \in \mathbf{R} \mid \frac{1}{n} \le x \le 1 - \frac{1}{n}\} \mid n \in \{k \in \mathbf{N} \mid k \ge 2\}\right\}\right) = \{x \in \mathbf{R} \mid 0 < x < 1\}$$

$$\bigcap\left(\left\{\{x \in \mathbf{R} \mid \frac{1}{n} \le x \le 1 - \frac{1}{n}\} \mid n \in \{k \in \mathbf{N} \mid k \ge 2\}\right\}\right) = \{\tfrac{1}{2}\}$$

הבנת דוגמה זו הינה תרגיל טוב בהבנת הנקרא בשפה, אותה אנו לומדים. הבה נסביר אותה בפרוטרוט. ובכן מדובר פה ב- $\cup(F)$ וב- $\cap(F)$, כאשר:

$$F = \left\{\{x \in \mathbf{R} \mid \frac{1}{n} \le x \le 1 - \frac{1}{n}\} \mid \mathrm{n} \in \{\mathrm{k} \in \mathbf{N} \mid \mathrm{k} \ge 2\}\right\}$$

עתה לביטוי F הצורה, שמרשה עיקרון ההחלפה, כלומר צורתו היא $\{t \mid n \in A\}$, כאשר:

$$t = \{\mathrm{x} \in \mathbf{R} \mid \frac{1}{n} \le x \le 1 - \frac{1}{n}\}$$

$$A = \{k \in \mathbf{N} \mid k \ge 2\}$$

A עצמה היא קבוצה לפי עיקרון הקומפרהנסיה המוגבל. t הינו ביטוי, שבו n מופיע כמשתנה חופשי (הוא לא חופשי ב- F כולו, כמובן). לכל $n \in \mathbf{N}$ ספציפי t מגדיר קבוצה (שוב, לפי עיקרון הקומפרהנסיה המוגבל). קבוצה זו מסמנים, כזכור, בצורה: $[\frac{1}{n}, 1 - \frac{1}{n}]$. מכאן שהביטוי F אכן מיצג קבוצה לפי עקרון ההחלפה, ו- $\cup(F)$ מוגדר אפוא. הקבוצה שהוא מיצג היא קבוצת כל המספרים הממשיים השיכים לקטע כלשהו מהצורה: $[\frac{1}{n}, 1 - \frac{1}{n}]$, כש- $n \in \{k \in \mathbf{N} \mid k \ge 2\}$ (כלומר: כש- $n \ge 2$). כיוון שלכל מספר x בין 0 ל- 1 (לא כולל 0 ו- 1), אפשר למצוא $\mathrm{n} \ge 2$ כך ש- $\frac{1}{n} \le x \le 1 - \frac{1}{n}$, ורק למספרים אלו תכונה זו, $\cup(F)$ הינו הקטע הפתוח $(0,1)$ (כלומר: $\{x \in \mathbf{R} \mid 0 < x < 1\}$). את העובדה ש- $\cap(F) = \{\frac{1}{2}\}$ נשאיר כתרגיל לקוראים.

הערה:

קבוצה כמו F בדוגמה האחרונה מתוארת בדרך-כלל בספרות כך :

$$F = \{[\frac{1}{n}\ , 1 - \frac{1}{n}]\ \mid n \in \mathbf{N}\ \wedge n \geq 2\}$$

יש כאן שימוש בסימון המקוצר המקובל עבור קטעים על הישר הממשי, יחד עם שילוב של הסימונים עבור עקרונות ההחלפה והקומפרהנסיה. השילוב מרשה לרשום: $\{t \mid x \in A \wedge \varphi\}$ במקום: $\{t \mid x \in \{x \in A \mid \varphi\}\}$.

סימונים מקובלים נוספים :

1. כאשר F סופית, $F = \{A_1, ..., A_n\}$, אז במקום $\cup(F)$ מקובל לכתוב $A_1 \cup A_2 \cup \ldots \cup A_n$. למשל: $A \cup B \cup C \cup D = \cup(\{A, B, C, D\})$. בדומה, במקום $\cap(F)$ כותבים $A_1 \cap A_2 \cap \ldots \cap A_n$. קל לראות שסימון זה מתאים לפירוש הקודם, שנתנו לביטוי $A_1 \cup A_2 \cup \ldots \cup A_n$.
2. כמו שכבר ראינו בדוגמאות (2) ו- (3), ברוב המקרים בהם עוסקים ב"משפחה" של קבוצות, ובמיוחד כשעושים חיתוך או איחוד של משפחה כזו, המשפחה מתוארת לפי עיקרון ההחלפה. במלים אחרות: היא מיוצגת בצורה $\{A(\mathrm{i}) \mid i \in I\}$. במקרה זה מקובל לכתוב $\underset{i \in I}{\cup} A(i)$ במקום $\cup(\{A(i) \mid i \in I\})$. לחיתוך יש סימון דומה.

   למשל:

   בדוגמא מספר 2 כותבים: $\bigcup\limits_{n \in N} \{n\}$ (זה שווה, כזכור, ל- **N**).

   בדוגמא מספר 3 כותבים: $\bigcup\limits_{n \in \{k \in \mathbf{N} \mid k \geq 2\}} \left\{x \in \mathbf{R} \mid \frac{1}{n} \leq x \leq 1 - \frac{1}{n}\right\}$

   או אפילו (בעזרת (β)): $\bigcup\limits_{n \geq 2} \left\{x \in \mathbf{R} \mid \frac{1}{n} \leq x \leq 1 - \frac{1}{n}\right\}$

   ובעוד יותר קיצור: $\bigcup\limits_{n \geq 2} \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$

המקרים הנפוצים ביותר הם כש- $I = \{1, 2, \ldots, n\}$ או כש- $I = \{n \in \mathbf{N} \mid n \geq k\}$ לאיזה k טבעי. במקרים אלו כותבים לעתים $\bigcup\limits_{i=1}^{n} A(i)$ ו- $\bigcup\limits_{i=k}^{\infty} A(i)$ (בהתאמה). בסימון דומה משתמשים עבור $\cap$.

את הדוגמא השלישית, למשל, נכתוב בדרך כלל כך:

$$\bigcup_{n=2}^{\infty}\left[\frac{1}{n},1-\frac{1}{n}\right]=(0,1)\ ,\quad \bigcap_{n=2}^{\infty}\left[\frac{1}{n},1-\frac{1}{n}\right]=\left\{\frac{1}{2}\right\}$$

כאשר משתמשים בסימונים הנ"ל יש לזכור את שתי העובדות הבסיסיות הבאות:

$$\boxed{\begin{array}{lcl} x\in\bigcup\limits_{i\in I}A_i & \Leftrightarrow & \exists\, i\in \mathrm{I}\,.\, x\in A_i \\[2ex] x\in\bigcap\limits_{i\in I}A_i & \Leftrightarrow & \forall\, i\in \mathrm{I}\,.\, x\in A_i \end{array}}$$

(הלכנו כאן בעקבות הנוהג המקובל בכל הטקסטים וכתבנו A_i במקום $A(i)$ בכתיבה זו קוראים ל- "i" שב- A_i "אינדקס" ול- I "קבוצת אינדקסים". אין למונחים אלו שום חשיבות, למעשה).

%% הערה:

אם נתרגם את $x\in\bigcup\limits_{i\in I}A_i$ (למשל) לפי ההגדרות, באופן שיטתי, נקבל משהו מסובך יותר:

$$x\in\bigcup_{i\in I}A_i \ \Leftrightarrow\ \exists Y(x\in Y\wedge \exists i\in I.Y=A_i)$$

זה שקול למה שכתוב למעלה לפי כללים לוגיים, הקשורים בשוויון, אותם לא ניסחנו במפורש בקורס זה. עם זאת, די בחשיבה אינטואיטיבית כדי להבין שקילות זו. למעשה, אפשר להבין את נכונות העקרונות שבמסגרת למעלה ישירות מהגדרת "איחוד של משפחה של קבוצות" ו "חיתוך של משפחת קבוצות", כמו שניסחנו אותם בעברית.
%%

תכונות של הפעולות המוכללות:

הפעולות המוכללות של איחוד וחיתוך מקיימות הכללות טבעיות של כמה מהכללים בטבלה מספר ב.1. הכללות אלו מסוכמות בטבלה מספר ב.2. המספרים בה תואמים את אלה שבטבלה מספר ב.1. הכללים מתוארים בצורה המקובלת בספרות (שהיא השימושית יותר). בתוך הסוגרים המרובעים נמצאת אבל הצורה הבסיסית יותר של הכלל. אנו נוכיח, לדוגמה, את כלל 9a:

$$A\cap(\bigcup_{i\in I}B_i)=\bigcup_{i\in I}(A\cap B_i)$$

$(\subseteq)$:

נניח $x \in A \cap (\bigcup_{i\in I} B_i)$. אז $x \in A$ ו- $x \in \bigcup_{i\in I} B_i$. מכאן שקיים $i \in I$ כך ש- $x \in B_i$. עבור אותו i , $x \in A \cap B_i$, ולכן $x \in \bigcup_{i\in I} (A \cap B_i)$.

$(\supseteq)$:

נניח $x \in \bigcup_{i\in I} (A \cap B_i)$. אז קיים $i \in I$ כך ש- $x \in A \cap B_i$. מכאן ש- $x \in A$ ו- $x \in B_i$. מהעובדה האחרונה נובע, ש- $x \in \bigcup_{i\in I} B_i$. מזה ומ- $x \in A$ נובע ש- $x \in A \cap (\bigcup_{i\in I} B_i)$.

טבלה ב.2: תכונות יסודיות של האופרציות המוכללות

1	$\forall j \in I.\ \bigcap_{i\in I} A_i \subseteq A_j$	$[\ \forall X \in F.\ \cap(F) \subseteq X\ ]$
2	$\forall j \in I.\ A_j \subseteq \bigcup_{i\in I} A_i$	$[\ \forall X \in F.\ X \subseteq \cup(F)\ ]$
5a	$C \subseteq \bigcap_{i\in I} B_i \Leftrightarrow \forall i \in I.\ C \subseteq B_i$	$[\ C \subseteq \cap(F) \Leftrightarrow \forall X \in F.\ C \subseteq X\ ]$
5b	$\bigcup_{i\in I} B_i \subseteq C \Leftrightarrow \forall i \in I.\ B_i \subseteq C$	$[\cup(F) \subseteq C \Leftrightarrow \forall X \in F.\ X \subseteq C\ ]$
8a	$(\bigcap_{i\in I_1} A_i) \cap (\bigcap_{i\in I_2} A_i) = \bigcap_{i\in I_1\cup I_2} A_i$	$[(\cap(F_1)) \cap (\cap(F_2)) = \cap(F_1 \cup F_2)\ ]$
8b	$(\bigcup_{i\in I_1} A_i) \cup (\bigcup_{i\in I_2} A_i) = \bigcup_{i\in I_1\cup I_2} A_i$	$[\ (\cup(F_1)) \cup (\cup(F_2)) = \cup(F_1 \cup F_2)\ ]$
9a	$A \cap (\bigcup_{i\in I} B_i) = \bigcup_{i\in I} (A \cap B_i)$	$[\ A \cap (\cup F) = \cup(\{A \cap X \mid X \in F\})\ ]$
9b	$A \cup (\bigcap_{i\in I} B_i) = \bigcap_{i\in I} (A \cup B_i)$	$[\ A \cup (\cap F) = \cap(\{A \cup X \mid X \in F\})\ ]$
13a*	$\overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i}$	$[\overline{\cap(F)} = \cup(\{\overline{X} \mid X \in F\})\ ]$
13b*	$\overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i}$	$[\overline{\cup(F)} = \cap(\{\overline{X} \mid X \in F\})\ ]$

(ד) קבוצות זרות, איחוד זר

הגדרה:

קבוצות A ו- B נקראות זרות זו לזו (Disjoint) אם אין להן שום איבר משותף, דהיינו אם $A \cap B = \varnothing$.

סימון:

כאשר A ו- B זרות מקובל (אך לא הכרחי) לסמן את האיחוד שלהן בצורה הבאה: $A \uplus B$. קוראים לזה אז ***"איחוד זר"***. יש להדגיש כי אין הבדל בין ערך הביטוי $A \uplus B$ (כשהוא מוגדר) ובין ערך $A \cup B$. השימוש ב- $\uplus$ בא רק לציין שאנו יודעים (או מניחים לצורך החישוב) ש- A ו- B זרות זו לזו. למעשה, $\cup$ ו- $\uplus$ שונים זה מזה רק בתחום ההגדרה שלהם: הביטוי $A \cup B$ מוגדר עבור כל שתי קבוצות A ו- B, בעוד ש- $A \uplus B$ מוגדר רק כאשר A ו- B קבוצות זרות.

את המושג והסימון ניתן (ומקובל) להכליל גם לאיחוד של יותר משתי קבוצות. כללית, הביטוי $\biguplus_{i \in I} A_i$ מוגדר אם (ורק אם):

$$\forall\, i \in I \quad \forall\, j \in I \,.\, A_i \cap A_j = \varnothing$$

(דהיינו – אם איברי הקבוצה $\{A_i | \, i \in I\}$ זרים זה לזה בזוגות). כאשר $\biguplus_{i \in I} A_i$ מוגדר, אין ערכו שונה מזה של $\bigcup_{i \in I} A_i$.

דוגמאות:

$\biguplus_{x \in Z} [x, x+1) = \mathbf{R}$, בעוד $\biguplus_{x \in Z} [x, x+1]$ הוא ביטוי בלתי מוגדר, כיוון ש(לדוגמה): $[2,3] \cap [3,4] \neq \varnothing$.

(ה) זוגות סדורים, מכפלה קרטזית

הגדרה:

הזוג הסדור של a ו- b הוא הקבוצה $\{\{a\},\{a,b\}\}$, והוא מסומן ב $<a,b>$ (בטקסטים אחדים כותבים (a,b), אך איננו רוצים ליצור כאן בלבול עם הסימן המקובל לקטעים פתוחים על הישר הממשי).

התכונה העיקרית של זוגות סדורים ניתנת על ידי הטענה הבאה:

טענה:

$$<a,b> = <c,d> \Leftrightarrow a = c \wedge b = d$$

הוכחה: תרגיל.

מהותו של הזוג סדור $<a,b>$ הינה, שהוא יוצר עצם חדש משני עצמים, שבו הסדר של שני העצמים הללו הינו ***חשוב***. זהו ההבדל בינו ובין הזוג הלא סדור $\{a,b\}$ (שגם הוא עצם הנוצר על-ידי קומבינציה של שני עצמים). לגבי זוגות לא סדורים $\{a,b\} = \{b,a\}$, לעומת זאת, $<a,b> \neq <b,a>$ (אלא אם $a = b$, כמובן).

הגדרה:
המכפלה הקרטזית של שתי קבוצות A ו- B (סימון $A \times B$) היא קבוצת כל הזוגות הסדורים $<a,b>$ כך ש- $a \in A$ ו- $b \in B$.

$$A \times B = \{ z \mid \exists a \in A \, \exists b \in B. z = <a,b> \}$$

בעיה:
האם נוכל להראות ש- $A \times B$ אכן קבוצה, כאשר ידוע ש- A ו- B קבוצות?

התשובה חיובית:
ברור שאם $a \in A$ ו- $b \in B$ אז $a \in A \cup B$ ו- $b \in A \cup B$.
לכן $\{a\} \subseteq A \cup B$ ו- $\{a,b\} \subseteq A \cup B$. מכאן ש- $\{a\} \in P(A \cup B)$ ו- $\{a,b\} \in P(A \cup B)$.
לכן $\{\{a\},\{a,b\}\} \subseteq P(A \cup B)$ וממילא $<a,b> \in P(P(A \cup B))$. מכאן ש:

$$A \times B = \{ z \in P(P(A \cup B)) \mid \exists a \in A \, \exists b \in B. z = <a,b> \}$$

ולכן $A \times B$ קבוצה לפי עיקרון הקומפרהנסיה המוגבל (יחד עם עיקרון האיחוד ועיקרון קבוצת החזקה).

אזהרה: בדרך-כלל $A \times B \neq B \times A$ ו- $A \times (B \times C) \neq (A \times B) \times C$.

סימונים נוספים:

1. כאשר $A=B$ מקובל לכתוב A^2 במקום $A \times A$. כך לדוגמה: $<-3, \pi> \in \mathbf{R}^2$.

2. במקום ביטוי מהצורה: $\{z \mid \exists x \exists y.\ z = <x,y> \wedge \varphi\}$
מקובל לכתוב את הדבר הבא: $\{<x, y> \mid \varphi\}$.
בצורת כתיבה זו קושר האופרטור $\{\ \ldots \mid \ldots\ \}$ בבת אחת גם את x וגם את y !
כלל (β) מקבל כאן את הצורה:

$$<t, s> \in \{<x, y> \mid \varphi\} \Leftrightarrow \varphi(t/x, s/y)$$

(ובלבד ש- t חופשי להצבה במקום x ב- φ ו- s חופשי שם להצבה במקום y).
קבוצות חלקיות ל- $A \times B$ המוגדרות לפי עיקרון הקומפרהנסיה המוגבל מיוצגות בדרך כלל על ידי ביטויים מהצורה:

$$\{<x, y> \in A \times B \mid \varphi\}$$

זהו כמובן קיצור של:

$$\{z \in A \times B \mid \exists x \exists y.\ z = <x, y> \wedge x \in A \wedge y \in B \wedge \varphi\}$$

לכן:

$$<t, s> \in \{<x, y> \in A \times B \mid \varphi\}$$

אם ורק אם: $t \in A \wedge s \in B \wedge \varphi(t/x, s/y)$.

דוגמה:

קבוצת הנקודות במישור, הנמצאות בפנים עיגול היחידה (העיגול שמרכזו בראשית הצירים ורדיוסו 1) מיוצגת בגיאומטריה אנליטית כך:

$$\{<x, y> \in \mathbf{R}^2 \mid x^2 + y^2 < 1\}$$

מושג חשוב נוסף הקשור בזוגות סדורים הוא מושג *ההיטל* :

הגדרה:

הפעולות π_1 ו- π_2 על זוגות סדורים מוגדרות על ידי הנוסחאות:

$$\pi_2(<a, b>) = b \qquad\qquad \pi_1(<a, b>) = a$$

פעולות אלו מספקות את *ההיטלים* של הזוג $<a, b>$. כדאי לציין, שאם z הינו זוג סדור, אז $z = <\pi_1(z), \pi_2(z)>$.

הכללות:

1. ***n-יות*** – אלו מוגדרות, באופן כללי, בצורה הבאה:

$$< a_1, \ldots, a_n> = < a_1, < a_2 < \ldots < a_{n-1}, a_n > \ldots >>>$$

למשל הרביעייה $<a,b,c,d>$ מוגדרת כך:

$$< a,b,c,d > = < a, < b, < c,d >>>$$

התכונה החשובה של n-יות היא, שוב, ש:

$$<a_1, \ldots, a_n> = <b_1, \ldots, b_n> \Leftrightarrow a_1 = b_1 \wedge a_2 = b_2 \wedge \ldots \wedge a_n = b_n$$

(תרגיל: להוכיח זאת באינדוקציה על n, תוך שימוש במקרה $n = 2$).

2. אם $A_1, A_2, \ldots, A_n$ הינן קבוצות, אז ***מכפלתן הקרטזית***: $A_1 \times A_2 \times \ldots \times A_n$ היא קבוצת כל ה- ***n-יות*** $<a_1,\ldots, a_n>$, כך ש-

$$a_1 \in A_1 \wedge a_2 \in A_2 \wedge \ldots \wedge a_n \in A_n$$

תרגיל: $A_1 \times A_2 \times \ldots \times A_n = A_1 \times (A_2 \times (\ldots \times A_n)) \ldots)$

גם כאן כותבים A^n כאשר $A_1 = A_2 = \ldots = A_n = A$.

3. את הסימון $\{<x, y> \mid \varphi\}$ אפשר להכליל ל- ***n-יות*** כך:

$$\{<x_1, \ldots, x_n> \mid \varphi\}$$

או בשימוש בעיקרון הקומפרהנסיה המוגבל:

$$\{<x_1, \ldots, x_n> \in A_1 \times \ldots \times A_n \mid \varphi\}$$

נשאיר לקוראים את ניסוח הגרסה המתאימה של כלל (β).

4. פונקציות ההיטל π_i^n מוגדרות בצורה הבאה:

$$\pi_i^n(<a_1, \ldots, a_n>) = a_i$$

למשל: $\pi_2^5(<-3, 2, \varnothing, \{\varnothing\}, e>) = 2$

טבלה מספר ב.3 מסכמת את הסימונים השונים עבור קבוצות, שלמדנו עד כה (כן נכללים בה אי-אלו סימונים נוספים, אותם נלמד בהמשך).

טבלה ב.3: סימונים עבור קבוצות

I. <u>שמות מיוחדים (וקבועים)</u>: $\mathbf{N}$, $\mathbf{N}^+$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, $\varnothing$

II. <u>קבוצות סופיות מפורטות</u>: $\{t_1, t_2, \ldots, t_n\}$

III. <u>עיקרון הקומפרהנסיה המוגבל</u>: $\{x \in S \mid \varphi\}$

IV. <u>עיקרון ההחלפה</u>: $\{t \mid x \in S\}$

V. <u>קבוצת החזקה</u>: $P(S)$ או 2^S

VI. <u>פעולות בינאריות</u>: $A \cup B$ איחוד, $A \cap B$ חיתוך, $A - B$ הפרש, $A \Delta B$ הפרש סימטרי, $A \times B$ מכפלה קרטזית

VII. <u>משלים יחסי</u>: $\overline{A}_E$ או A^c_E ולפעמים פשוט $\overline{A}$ או A^c

VIII. <u>הכללת הפעולות הבינאריות</u>:

$\bigcup(A)$ $\quad$ $\bigcap(A)$

$\bigcap_{\alpha \in I} A_\alpha$ $\quad$ $\bigcup_{\alpha \in I} A_\alpha$

סימונים נוספים, שיילמדו בהמשך:

$A \to B$ או B^A – קבוצת הפונקציות מ- A ל- B

$A/\equiv$ – קבוצת המנה של A ביחס ל- $\equiv$

ב.4. פונקציות

I. מושג הפונקציה

בפרק הזה נדון באחד המושגים הבסיסיים ביותר של המתמטיקה ושל מדעי המחשב: מושג הפונקציה. זהו, ללא ספק, אחד הפרקים החשובים ביותר בספר זה.

הגדרה

(1) ***פונקציה*** מעל קבוצה A היא התאמה, המתאימה לכל איבר x ב- A עצם אחד ויחיד. עצם זה נקרא ***ערך הפונקציה*** ב- x.

(2) ***פונקציה חלקית מעל קבוצה*** A היא פונקציה מעל תת-קבוצה כלשהי של A.

הערות

1. בעצם, אין אלו "הגדרות" ממש, כי מושג ה"התאמה", המופיע בהן, אינו בהיר הרבה יותר (אם בכלל) ממושג הפונקציה עצמו. יש לראותן יותר בגדר הסבר. מאוחר יותר ניתן הגדרה פורמלית במונחי תורת הקבוצות בלבד.
2. ברור מה"הגדרה", שכל פונקציה מעל A היא גם פונקציה חלקית מעל A. לפעמים מדגישים את ייחודן של הפונקציות מעל A על-ידי זה, שמכנים אותן "פונקציות טוטאליות" (מלאות) מעל A. אצלנו לא יהיה אבל הבדל בין המושגים של "פונקציה" ושל "פונקציה מלאה".
3. עוד ברור מה"הגדרה", כי מושג הפונקציה (מלאה או חלקית) כולל שני מרכיבים עיקריים:

(א) הקבוצה עליה הפונקציה מוגדרת. קבוצה זו נקראת ***התחום*** של הפונקציה.
(ב) ההתאמה עצמה, שבדרך-כלל ניתנת על-ידי ***כלל התאמה*** כלשהו.

סימונים

1. את תחום ההגדרה של פונקציה (מלאה או חלקית) f מסמנים ב- $\mathrm{Dom}(f)$.

2. את ערך הפונקציה f עבור איבר t של תחום ההגדרה מסמנים בדרך-כלל בצורה $f(t)$ (באי-אלו שפות מחשב, כמו LISP ו- SCHEME, כותבים דווקא (f, t)). בביטויים מהצורה $f(t)$ מקובל לקרוא ל- t ה"ארגומנט" של הביטוי, ול- $f(t)$ כולו "ערך הפונקציה f עבור הארגומנט t".

אינפורמציה חשובה נוספת על פונקציות (מלאות/חלקיות) היא מאין הן לוקחות את הערכים שלהן.

הגדרה

(1) תהי f פונקציה. קבוצה B נקראת ***טווח*** של f אם מתקיים:

$$\forall x \in \mathrm{Dom}(f).\, f(x) \in B$$

(כלומר, B כוללת את כל הערכים ש-f "מקבלת").

(2) פונקציה חלקית מעל A, ש- B היא טווח שלה, נקראת ***פונקציה (חלקית) מ- A ל- B***.

סימונים

1. הסימון המקובל לכך, ש- "f היא פונקציה מ- A ל- B" הוא $f : A \to B$.
2. בעקבות מספר שפות מחשב מתקדמות (כמו SML), נסמן אנו את קבוצת הפונקציות מ- A ל- B ב- $A \to B$. עם סימון זה אין הבדל, למעשה, בין $f : A \to B$ ובין $f \in A \to B$.[11]
   סימון מקובל אחר בספרות עבור $A \to B$ (שגם אנו נשתמש בו לעתים) הוא B^A.
3. אנו נסמן את קבוצת הפונקציות החלקיות מ- A ל- B ב- $A \overset{p}{\to} B$. סימון זה הוא עבור ספר זה בלבד, ואין לו בסיס בטקסטים אחרים.

הערות

(1) $A \to B$ ו- $A \overset{p}{\to} B$ הן אכן קבוצות כאשר A ו- B קבוצות. דבר זה יתברר בהמשך. (כמו כן נובע גם מהאקסיומות של תורת הקבוצות, שלכל פונקציה יש טווח. זהו, למעשה, ניסוח אלטרנטיבי לעיקרון ההחלפה, אך לא ניכנס לפרטים).

---

[11] באותן שפות מחשב משתמשים גם ב- ":" במקום ב- "$\in$". כך כותבים, למשל, 3:Int במקום $3 \in \mathbf{Z}$. לפיכך, משמעות "$f : A \to B$" בשפות אלו היא *בדיוק* זו של תורת הקבוצות (אם כי היסטורית, סימון זה קדם בהרבה לסימונים ולמושגים של תורת הקבוצות).

(2) כדאי לשים לב, שלא הגדרנו את מושג הטווח של פונקציה f, אלא "טווח" סתם. בעניין זה, *אין* ספר זה תואם את ההגדרות הרשמיות במרבית הטקסטים, אבל הוא תואם מאוד את האופן בו משתמשים *באמת* במושגים. כך, למשל, מעטים המתמטיקאים, אם בכלל, שירצו באמת ובתמים להבדיל בין "שתי" הפונקציות הבאות:

$f : \mathbf{Z} \to \mathbf{Z}$ המוגדרת על-פי הכלל $f(z) = z^2$.

$g : \mathbf{Z} \to \mathbf{N}$ המוגדרת על-פי הכלל $g(z) = z^2$.

למרות זאת, נדבר גם אנו לפעמים על ה"טווח" של פונקציה מסוימת. זה יקרה כשחשוב יהיה לציין תחום מסוים, ממנו היא לוקחת את ערכיה.

בין הטווחים הרבים של פונקציה יש אחד, שיש לו חשיבות מיוחדת. זהו ה*מינימלי* ביניהם:

הגדרה

ה*תמונה* של פונקציה f היא הקבוצה הבאה:

$$\mathrm{Im}(f) = \{f(x) \mid x \in \mathrm{Dom}(f)\}$$

דהיינו:

$$\mathrm{Im}(f) = \{y \mid \exists x \in \mathrm{Dom}(f).\ y = f(x)\}$$

הערות

(1) ברור שאם $f : A \to B$, אז $\mathrm{Im}(f) = \{y \in B \mid \exists x \in A.\ y = f(x)\}$.

(2) קבוצה B היא טווח של פונקציה f אם"ם $\mathrm{Im}(f) \subseteq B$. בפרט $\mathrm{Im}(f)$ היא טווח של f.

דוגמאות

(1) נגדיר $f : \mathbf{R} \to \mathbf{R}$ על-ידי $f(x) = x^2$. ההצהרה "$f: \mathbf{R} \to \mathbf{R}$" אומרת כאן, שהתחום של הפונקציה הוא $\mathbf{R}$, ושהפונקציה מקבלת ערכים ממשיים בלבד. "$f(x) = x^2$" נותן את כלל ההתאמה, והוא אכן קומפטיבילי עם ההצהרה הקודמת: ניתן להפעיל כלל זה על כל מספר ממשי והתוצאה תהיה אכן מספר ממשי (יש לבדוק דברים כאלה!). לאותו f נכון גם ש- $f : \mathbf{R} \to \{x \in \mathbf{R} \mid x \geq -1\}$, כיוון שכמו $\mathbf{R}$, גם $\{x \in \mathbf{R} \mid x \geq -1\}$ הוא טווח אפשרי של הפונקציה f. התמונה של f, לעומת זאת, היא $\{x \in \mathbf{R} \mid x \geq 0\}$. ברור לכן, שמתקיים גם $f : \mathbf{R} \to \{x \in \mathbf{R} \mid x \geq 0\}$.

(2) נגדיר $g: \mathbf{R} \overset{p}{\rightarrow} \mathbf{R}$ על-ידי הכלל $g(x) = \frac{1}{x}$. זוהי אכן רק פונקציה חלקית מ- $\mathbf{R}$ ל- $\mathbf{R}$, כי $g(0)$ אינו מוגדר. התחום של g הינו $\{x \in \mathbf{R} \mid x \neq 0\}$. לכן $g: \{x \in \mathbf{R} \mid x \neq 0\} \rightarrow \mathbf{R}$ (כאן כבר מדובר בפונקציה מלאה, אם כי עדיין נכון גם ש- $g: \{x \in \mathbf{R} \mid x \neq 0\} \overset{p}{\rightarrow} \mathbf{R}$). גם התמונה של g היא $\{x \in \mathbf{R} \mid x \neq 0\}$.

הערות

(1) שתי הדוגמאות היו של פונקציות ממשיות, דהיינו: פונקציות מקבוצות חלקיות של $\mathbf{R}$ אל $\mathbf{R}$. היסטורית, מושג הפונקציה התפתח מפונקציות כאלו, והן אלו העומדות במרכז החשבון הדיפרנציאלי והאינטגרלי. נדגיש אבל כבר כאן, שכל ***קבוצה יכולה להיות תחום של פונקציה, וכל קבוצה יכולה לשמש כטווח.*** עוד מעט נראה אכן דוגמאות רבות של פונקציות, שאינן מ- $\mathbf{R}$ ל- $\mathbf{R}$.

(2) בבעיות של "חקירת פונקציות" בחדו"א נותנים למעשה כלל התאמה, המגדיר פונקציה חלקית מ- $\mathbf{R}$ ל- $\mathbf{R}$. חקירת הפונקציה החלקית נפתחת במציאת תחום ההגדרה של פונקציה חלקית זו.

(3) תמיד $A \rightarrow B \subseteq A \overset{p}{\rightarrow} B$, ולמעשה $A \rightarrow B \subset A \overset{p}{\rightarrow} B$ (אלא אם כן $A = \varnothing$, אך במקרה זה כדאי לדון רק אחרי שניתן את ההגדרה הרשמית של פונקציות).

(4) אם $f: A \overset{p}{\rightarrow} B$ אז $f: \mathrm{Dom}(f) \rightarrow B$ (ומתקיים, כמובן, $\mathrm{Dom}(f) \subseteq A$).

שוויון פונקציות

העיקרון לגבי שוויון פונקציות דומה לזה של שוויון קבוצות: ממש כמו שזהות קבוצה אינה תלויה באופן בו מגדירים אותה אלא רק בזהות איבריה, כך גם זהות פונקציה אינה תלויה באופן בו אנו מתארים את כלל ההתאמה, אלא רק בזהות תחומה והערכים שהיא מקבלת. מכאן מקבלים אנו את ***עיקרון האקסטנציונליות עבור פונקציות***:

$$\boxed{f = g \iff \mathrm{Dom}(f) = \mathrm{Dom}(g) \wedge \forall x \in \mathrm{Dom}(f).\, f(x) = g(x)}$$

דוגמה טריוויאלית: אם התחום של f ו- g הוא $\mathbf{R}$ ו-

$$g(x) = x^2 + 2x + 1 \qquad f(x) = (x+1)^2$$

אז $f = g$.

II. הגדרת פונקציות וסימונן

הדרכים לסימון פונקציות ולהגדרתן דומות מאוד לאלו הנהוגות לגבי קבוצות, אותן ראינו בפרק ב.2. הבה נסקור אותן.

שיטת סימון מס' 1: שמות קבועים

לפונקציות חשובות רבות יש שמות *קבועים* מיוחדים. המספר כאן גדול לאין ערוך מזה של הקבוצות בעלי שמות מיוחדים, ולא ננסה כלל לתת רשימה מלאה. הנה דוגמאות: sin, tg, $\sqrt{\ }$, $|\,|$, $!$, t. כיוון שסימונים אלו הוכנסו לשימוש לפני פיתוח התיאוריה הכללית, לעתים קרובות אין משתמשים בהם באופן הסטנדרטי. כך כותבים אנו $\sqrt{9}$ ולא $\sqrt{\ }(9)$ $(\sqrt{\ }:\{x\in\mathbf{R}\mid x\geq 0\}\to\mathbf{R})$, $7!$ ולא $!(7)$ $(!:\mathbf{N}\to\mathbf{N})$, $|-3|$ ולא $|\,|(-3)$ $(|\,|:\mathbf{C}\to\mathbf{R})$, $\begin{bmatrix}1&0\\3&5\end{bmatrix}^t$ ולא $t\left(\begin{bmatrix}1&0\\3&5\end{bmatrix}\right)$ (הפונקציה t המתאימה לכל מטריצה את המטריצה המוחלפת שלה, והיא פונקציה מקבוצת כל המטריצות הממשיות (נאמר) אל קבוצת כל המטריצות הממשיות). כדאי לציין, שדבר זה מתוקן באי-אלו שפות מחשב, בהן כותבים, למשל, SQRT(9) במקום $\sqrt{9}$ או fact(9) במקום $7!$.

שיטת סימון מס' 2: שמות זמניים

הדרך הסטנדרטית ביותר בטקסטים להתייחס אל פונקציות, שאין להן שם קבוע, היא על-ידי מתן שמות *זמניים*. בדרך זו נקטנו גם אנו בשתי הדוגמאות הראשונות, שהבאנו למעלה ($f(x)=x^2$, $g(x)=1/x$). השמות הפופולריים ביותר הם f, g, h, F, G, H (אולי בתוספת אינדקסים: f_1, g_2). הניסוח הסטנדרטי הוא: "תהי $f:A\to B$ מוגדרת על-ידי $f(x)=\ldots$". נזכיר, שבצורת הגדרה זו המשתנה x הינו קשור על-ידי כמת אוניברסלי סמוי $\forall$, ויש להשתמש בהגדרה כאילו כתוב: "$\forall x\in A.\ f(x)=\ldots$" (ההקבלה בקבוצות: הגדרת קבוצה A על-ידי: "$\forall x.\ x\in A\Leftrightarrow\varphi(x)$").

שיטת סימון מס' 3: טבלאות

לגבי קבוצות *סופיות*, ראינו את הדרך הפשוטה של *פירוט* איבריהן, כמו $\{1,2,3,4\}$. דרכי פירוט דומות קיימות לגבי פונקציות, אם תחום הפונקציה סופי. שימושית וקומפקטית במיוחד היא דרך התיאור בעזרת *טבלה*. לדוגמה, פונקציה, שהתחום שלה

הוא קבוצת הסטודנטים ב"מתמטיקה בדידה" והטווח – המספרים הטבעיים בין 0 ל- 100 [12], מתוארת בעזרת טבלה כך:

סטודנט	ציון
סטיבן הנדרי	85
ג'ימי ווייט	90
סטיב דייויס	100
דייגו מרדונה	20

(זה היה בסמסטר קיץ, ומספר הסטודנטים היה קטן במיוחד).

שיטת סימון מס' 4: סימון למדא

הדרך של מתן שמות זמניים אינה נחשבת (ובצדק!) מספקת ויעילה עבור קבוצות, ולכן הומצאה שיטת האבסטרקציה: $\{x|\varphi\}$. לגבי פונקציות המצב דומה. שיטת השמות הזמניים אינה מספקת ולעתים קרובות אינה יעילה. למעשה, היא עוד הרבה פחות יעילה לגבי פונקציות מאשר לגבי קבוצות. אי לכך המציאו הלוגיקאים שיטת אבסטרקציה גם לצורך הגדרת פונקציות. השיטה ידועה בשם "סימון למדא", כיוון שמשתמשים בה באות היוונית λ (למדא). סימון זה טרם נקלט, למרבה הצער, אצל המתמטיקאים, אך משתמשים בו הרבה במדעי המחשב התיאורטיים, ובעיקר בשפות תכנות פונקציונליות. הרעיון: אם t ביטוי, אז $\lambda x.t$ מתאר את הפונקציה (אולי חלקית), המתאימה לכל עצם a את הערך של t עבור $x = a$. (בצורה אחרת: $\lambda x.t$ היא הפונקציה, שעבור x מסוים כקלט מחזירה כפלט את תוצאת חישוב t עבור x זה). לדוגמה: $\lambda x.x^2$ היא הפונקציה המתאימה לכל מספר x את ריבועו.

$\lambda x.t$ נותן בעצם רק את *הכלל*. כדי שסימון למדא יתאר פונקציה באופן מלא, מכניסים גם את התחום לסימון וכותבים $\lambda x \in A.t$ (או $\lambda x{:}A.t$) כדי לציין ש- A היא התחום (זה מקביל מאוד לסימון $\{x \in A\ |\varphi\}$ עבור קבוצות). כשרוצים לציין גם טווח, כותבים $\lambda x.t{:}A \rightarrow B$ או $\lambda x \in A.t \in B$, אך אנו נסתפק, בדרך-כלל, בסימון $\lambda x \in A.t$. מציאת טווח לפונקציה תיעשה על-ידי ניתוח הביטוי t. כשיהיה ברור באיזה תחום מדובר, נשמיט אפילו אותו ונכתוב רק $\lambda x.t$.

[12] כמובן, היינו צריכים לומר "אחד הטווחים", אבל במקרה הנוכחי אין זה נוח, וזה גם מחטיא את המטרה. השימוש בהא הידוע ("הטווח") נוח יותר. במקרה זה הוא בא לציין, שמראש ידוע שהציונים לא יחרגו מהטווח הנ"ל, אך כל איבר בו אפשרי (אם כי לא בהכרח מתקבל).

♦ דוגמאות ראשונות

הפונקציה בדוגמה הראשונה למעלה היא $\lambda x \in \mathbf{R}.x^2$. הפונקציה החלקית בדוגמה השנייה היא $\lambda x \in \mathbf{R}.\frac{1}{x}$. אם ברצוננו לתארה כפונקציה, עלינו לדייק ולכתוב: $\lambda x \in (\mathbf{R}\text{-}\{0\}).\frac{1}{x}$. (אם לא יצוין בפירוש אחרת, ביטויי-למדא ישמשו אצלנו לתיאור פונקציות בלבד).

האופרטור λ הוא אופרטור קשירה. המשתנה שהוא קושר נכתב מיד אחריו (כמו במקרה של $\forall$ ו- $\exists$). טווח הקשירה נקבע כמו במקרה של ι ו- μ (כלומר: הקטע הארוך ביותר אחרי הנקודה, שהוא תקין תחבירית, כשהוא עומד בפני עצמו). כיוון ש- λ הוא אופרטור קשירה, כלל (α) תופש, כמובן. כך אין כל הבדל בין $\lambda x \in \mathbf{R}.x^2$ ובין $\lambda y \in \mathbf{R}.y^2$. שני הביטויים זהים במשמעותם (ומתארים אותה פונקציה בדיוק!)

כשם שהעובדה היסודית ביותר הקשורה בסימון $\{x|\varphi\}$ היא כלל (β):

$$t \in \{x|\varphi\} \Leftrightarrow \varphi(t/x)$$

כך העובדה היסודית בדבר סימון למדא, זו *שמגדירה* אותו בעצם, היא כלל (β), עבורו:

$$\boxed{(\beta) \qquad (\lambda x.t)(s) = t(s/x)}$$

בתנאי, כמובן, ש- s הוא ביטוי החופשי להצבה ב- t במקום x. עבור $\lambda x \in A.\, t$ יש כלל דומה (ותנאי נוסף: ש- s הוא ביטוי המייצג איבר של A).

דוגמאות:

$$(\lambda x \in \mathbf{R}.x^2)(3) = 3^2 = 9$$

$$(\lambda x \in \mathbf{R}.x^2)(a+b) = (a+b)^2$$

$$(\lambda x.\int_0^x xydy)(z+1) = \int_0^{z+1}(z+1)ydy$$

אבל

$$(\lambda x.\int_0^x xydy)(y+1) \neq \int_0^{y+1}(y+1)ydy$$

למעשה:

$$(\lambda x.\int_0^x xydy)(y+1) = (\lambda x.\int_0^x xzdz)(y+1) = \int_0^{y+1}(y+1)z\,dz$$

הסברים:

כלל (β) משמש לצורך פישוט של ביטויים. צורת הפעלתו היא כדלקמן: בהינתן ביטוי מהצורה $(\lambda x.t)(s)$ (דהיינו: אפליקציה של ביטוי-λ לארגומנט s), מוחקים את "$\lambda x.$" מהביטוי $\lambda x.t$, ובמה שנשאר מציבים את s במקום כל מופע חופשי של x.

בחישוב האחרון השתמשנו באמצע בכלל (α) כדי להפוך את הביטוי $y+1$ שלא היה חופשי להצבה, לביטוי, שהוא כן חופשי להצבה.

כמו כל ביטוי או נוסחה אחרים, גם ביטוי-למדא יכול להכיל פרמטרים. אם נסמן את $\{x \in \mathbf{R} \mid x > 0\}$ ב- $\mathbf{R}^+$, אז הביטוי $\lambda x \in \mathbf{R}^+.x^a$ הוא ביטוי עבור פונקציה, שערכו המדויק תלוי בערך הפרמטר a. כאשר $a = 2$, זו תהיה הפונקציה $\lambda x \in \mathbf{R}^+.x^2$, וכש- $a = 1/2$ – הפונקציה $\lambda x \in \mathbf{R}^+.\sqrt{x}$. את כלל (β) ניתן, כמובן, להפעיל גם כשיש פרמטרים. לדוגמה:

$$(\lambda x \in \mathbf{R}^+.x^a)(3) = 3^a$$

(כיוון שזו ***זהות***, הכוונה בעצם ל- $\forall a \in \mathbf{R}[(\lambda x \in \mathbf{R}^+.x^a)(3) = 3^a]$).

נצעד עתה צעד נוסף קדימה, ונשים לב שהביטוי $\lambda a \in \mathbf{R}.\ \lambda x \in \mathbf{R}^+.x^a$ מתאר פונקציה, שהתחום שלה הוא $\mathbf{R}$, והיא מתאימה לכל $a \in \mathbf{R}$ ***פונקציה*** מ- $\mathbf{R}^+$ ל- $\mathbf{R}$ (או ל- $\mathbf{R}^+$):

$$\lambda a \in \mathbf{R}.\ \lambda x \in \mathbf{R}^+.x^a : \mathbf{R} \to (\mathbf{R}^+ \to \mathbf{R})$$

כך, למשל: $((\lambda a \in \mathbf{R}.\ \lambda x \in \mathbf{R}^+.x^a)(2))(3) = (\lambda x \in \mathbf{R}^+.x^2)(3) = 3^2 = 9$

דוגמה נוספת:

$$\lambda n \in \mathbf{N}.\ \lambda x \in [0,1].\ x^n(1-x^n) : \mathbf{N} \to ([0,1] \to [0,1])$$

הביטוי כאן משמאל מגדיר פונקציה, שמתאימה לכל מספר טבעי פונקציה מסוימת מהקטע $[0,1]$ אל $[0,1]$. כך, למשל:

$$\underbrace{((\lambda n \in \mathbf{N}.\ \lambda x \in [0,1].\ x^n(1-x^n))(3))}(0.2) = (\lambda x \in [0,1].\ x^3(1-x^3))(0.2) =$$

$$= 0.2^3(1-0.2^3)$$

(הערה: ב- " $\underbrace{\qquad}$ " סימנו את אותו חלק בביטוי המורכב, עליו הופעל כלל (β)).

לסיום הצגה ראשונה זו של סימון λ, נביא אנלוגיה נוספת בין הסימון $\{x|\varphi\}$ ובין הסימון $\lambda x.t$. לגבי $\{x|\varphi\}$ היה לנו, כזכור, כלל η, שקבע כי $\{x|x \in S\} = S$, בתנאי ש- x אינו חופשי ב- S. לגבי λ יש לנו כלל דומה. הוא נובע מההבחנה הבאה: אם f היא פונקציה, אז הפונקציה, המחזירה לכל x בתחום של f את $f(x)$, זהה לפונקציה f (כי זו עושה, כמובן, אותו דבר בדיוק). פורמלית:

$$\forall f(\lambda x.\, f(x) = f)$$

בנוסחה זו אין התייחסות לתחום של $\lambda x.\, f(x)$, אבל ברור, שתחום ההגדרה של הפונקציה, שביטוי זה מתאר, זהה לתחום ההגדרה של f.

מהנוסחה האחרונה נובע, בעזרת הכלל היסודי של $\forall$, שאם t ביטוי עבור פונקציות, שאינו מכיל את x כפרמטר, אזי:

$$\boxed{(\eta) \qquad \lambda x.\, t(x) = t}$$

(וכמובן $\lambda x \in A.\, t(x) = t$ בתנאי הנוסף, ש- t מגדיר פונקציה, שהתחום שלה הוא A).

דוגמאות:

(א) $\lambda x.\, \sin x = \sin$

(ב) אם g משתנה עבור פונקציות, אז $\lambda x.\, g(x) = g$.

(ג) נקח $t = \lambda y \in \mathbf{R}.y^2$, אז נקבל (אם נרשום גם את התחום של $\lambda x.\, t(x)$):

$$\lambda x \in \mathbf{R}.\, (\lambda y \in \mathbf{R}.\, y^2)(x) \overset{\eta}{=} \lambda y \in \mathbf{R}.\, y^2$$

את האינסטנציה הזו של כלל (η) יכולנו לקבל, במקרה זה, ישירות מכללי (α) ו- (β) כך:

$$\lambda x \in \mathbf{R}.\, \underbrace{(\lambda y \in \mathbf{R}.\, y^2)(x)} \overset{\beta}{=} \lambda x \in \mathbf{R}.\, x^2 \overset{\alpha}{=} \lambda y \in \mathbf{R}.\, y^2$$

הכללים α, β ו- η (ביחד עם עיקרון האקסטנציונליות, עליו ניתן להראות שהוא שקול לכלל η) הם הכללים היסודיים של ביטויי-λ. היסטורית, למעשה, השמות (והניסוח המפורש של הכללים) נעשו קודם כל עבור ביטויים אלה דווקא. טבלה מס' ב.4 מסכמת את הכללים הללו הן עבור קבוצות והן עבור פונקציות. כמו כן הוכנסו לטבלה שלושה כללים פשוטים נוספים, המשמשים לפישוטים, אותם הכרנו בפרק הקודם.

טבלה ב.4: כללי היסוד לפישוטים

(α) $\{y\|\varphi\} = \{x\|\varphi(x/y)\}$ בתנאי ש- x משתנה חדש, שאינו מופיע ב- φ.	(α) $\lambda y.\ t = \lambda x.\ t(x/y)$ בתנאי ש- x משתנה חדש, שאינו מופיע ב- t.
(β) $s \in \{x\|\varphi\} \Leftrightarrow \varphi(s/x)$ בתנאי ש- s חופשי להצבה במקום x ב- φ.	(β) $(\lambda x.\ t)(s) = t(s/x)$ בתנאי ש- s חופשי להצבה במקום x ב- t.
(η) $\{x \mid x \in S\} = S$ בתנאי ש- x אינו מופיע חופשי ב- S.	(η) $\lambda x.\ t(x) = t$ בתנאי ש- x אינו מופיע חופשי ב- t.
(Ext) $\forall A \forall B[A = B \Leftrightarrow \forall x.\ x \in A \Leftrightarrow x \in B]$	(Ext) $\forall f \forall g\ [f = g \Leftrightarrow \forall x.\ f(x) = g(x)]$ וביתר דיוק: $\forall f \forall g\ [f = g \Leftrightarrow \mathrm{Dom}(f) = \mathrm{Dom}(g) \wedge$ $\wedge\ \forall x \in \mathrm{Dom}(f).\ f(x) = g(x)]$
$\pi_2(<x, y>) = y$ $\quad$ $\pi_1(<x, y>) = x$ $z = <\pi_1(z)\ ,\ \pi_2(z)>$ (הנוסחה האחרונה נכונה בתנאי ש- z הוא זוג סדור).	

שיטת סימון מס' 5: סימון למדא עבור פונקציות של כמה משתנים

במתמטיקה נוהגים לעתים קרובות להתעסק בפונקציות של שני משתנים או יותר. כך, למשל, אפשר למצוא בטקסטים בחדו"א הגדרה של פונקציה g, הנראית כך:

$$g(x, y) = xy^2$$

רשמית, רואים בדרך-כלל בפונקציות כאלה לא משהו מסוג חדש, אלא פונקציות ***במובן הרגיל***, שהארגומנטים שלהן הם ***זוגות*** של מספרים. פונקציה g כמו זו שבדוגמה האחרונה היא אפוא פונקציה מ- $\mathbf{R} \times \mathbf{R}$ אל $\mathbf{R}$. למעשה, היה צריך, לאור זאת, לכתוב בעצם:

$$g(<x, y>) = xy^2$$

הסימון המקובל בטקסטים רחוק אבל, לרוע המזל, מלהיות עקבי, וכמעט תמיד כותבים $g(x, y)$, היכן שהיה צריך לכתוב $g(<x, y>)$.

את הסימון $\{\_\ |...\}$ עבור קבוצות מסגלים, כזכור, לצורה מיוחדת, נוחה, כשרוצים להגדיר קבוצה של זוגות (כותבים: $\{<x,y>|\ \varphi(x,y)\}$). בדומה, גם את סימון-למדא מרחיבים לצורך טיפול נוח בפונקציות, שהתחום שלהן הוא מכפלה קרטזית (או קבוצה כלשהי אחרת של זוגות). כאשר רוצים לתאר את ***הכלל*** בלבד, כותבים: $\lambda x,y.\ t$

(לדוגמה: $\lambda x,y.xy^2$). כאשר רוצים לציין גם את התחום, ותחום זה הוא המכפלה הקרטזית $A \times B$, כותבים:

$$\lambda x \in A\ ,\ y \in B.\ t$$

לדוגמה: הביטוי $\lambda x \in \mathbf{R}\ , y \in \mathbf{R}.\ xy^2$ מייצג את הפונקציה g למעלה מ- $\mathbf{R}^2$ אל $\mathbf{R}$.

את הסימון המיוחד עבור פונקציות של שני משתנים ניתן להכליל ללא קושי לפונקציות של מספר כלשהו של משתנים. כללית, $\lambda x_1 \in A_1, x_2 \in A_2, \ldots, x_n \in A_n.\ t$ הוא ביטוי המתאר פונקציה, שהתחום שלה הוא $A_1 \times A_2 \times \ldots \times A_n$ ומתאימה לכל $<x_1, x_2, \ldots, x_n>$ ב- $A_1 \times \ldots \times A_n$ את הערך המתאים של הביטוי t.

נקודות שחשוב לזכור כאן:

(i) בביטוי כמו $\lambda x \in A\ , y \in B.\ t$, האופרטור λ קושר בבת-אחת בתוך t גם את x וגם את y.

(ii) חשוב לשים לב להבדל בין השימוש בפסיק לבין השימוש בנקודה: פסיקים מפרידים בין המשתנים השונים ש- λ קושר, בעוד הנקודה מציינת את תחילתו של טווח הקשירה.

(iii) חשוב מאוד להבדיל בין $\lambda x \in A\ , y \in B.\ t$ ובין $\lambda x \in A.\ \lambda y \in B.\ t$. כך, למשל, $\lambda x \in \mathbf{R}, y \in \mathbf{R}.\ xy^2$ מציין פונקציה מ- $\mathbf{R} \times \mathbf{R}$ אל $\mathbf{R}$, בעוד $\lambda x \in \mathbf{R}.\ \lambda y \in \mathbf{R}.\ xy^2$ מציין פונקציה מ- $\mathbf{R}$ אל $\mathbf{R} \to \mathbf{R}$! (כשמכניסים, לדוגמה, 2 כקלט לביטוי האחרון, התוצאה היא $\lambda y \in \mathbf{R}.\ 2y^2$, וזו היא פונקציה מ- $\mathbf{R}$ ל- $\mathbf{R}$).

(iv) כלל (β) לפונקציות של שני משתנים נראה כך:

$$(\lambda x,y.\ t)\ (<u,v>) = t(u/x,\ v/y)$$

בתנאי ש- u חופשי להצבה במקום x ב- t, ו- v חופשי להצבה במקום y ב- t (במקום $\lambda x,y.\ t$ יכול כאן, כמובן, להיות רשום משהו מהצורה המלאה יותר: $\lambda x \in A, y \in B.\ t$).

דוגמאות:

$$(\lambda x \in \mathbf{R}\ , y \in \mathbf{R}.\ xy^2)\ (<3,2>) = 3 \cdot 2^2$$
$$(\lambda x \in \mathbf{R}\ , y \in \mathbf{N}.\ x^y)\ (<x+1,\ 2y>) = (x+1)^{2y}$$

מעשית, נכתוב גם אנו לעתים קרובות $(\lambda x,y.\ t)(u,v)$ במקום $(\lambda x,y.\ t)(<u,v>)$. למשל: $(\lambda x \in \mathbf{R}\ , y \in \mathbf{R}.\ xy^2)(3,2)$ או $(\lambda x \in \mathbf{R}\ , y \in \mathbf{N}.\ x^y)(x+1,\ 2y)$.

את כלל (β) ניתן להכליל בצורה דומה לפונקציות של n משתנים ($n \in \mathbf{N}^+$).

(v) שתי פונקציות חשובות במיוחד בהקשר הנוכחי הן פונקציות ההטל, π_1 ו- π_2[13], אותן הכרנו כבר בסוף הפרק הקודם:

$$\pi_1 = \lambda x,y.\ x \qquad \pi_2 = \lambda x,y.\ y$$

ולכן:

$$\boxed{\pi_1(<x,y>) = x \qquad \pi_2(<x,y>) = y}$$

אם נדייק, מה שתיארנו כאן הוא שוב רק ה*כלל*. כדי שנוכל לדבר על *פונקציות*, עלינו לציין מה התחום. לאמיתו של דבר אין, בעצם, פונקציות "π_1" ו- "π_2", אלא לכל שתי קבוצות A ו- B יש לנו:

$$\pi_1^{A,B} = \lambda x \in A\ ,\ y \in B.\ x : A \times B \to A$$
$$\pi_2^{A,B} = \lambda x \in A\ ,\ y \in B.\ y : A \times B \to B$$

שיטת סימון מס' 6: סדרות

בחשבון דיפרנציאלי ואינטגרלי עיקר העניין, בדרך-כלל, הוא בפונקציות מ- $\mathbf{R}$ ל- $\mathbf{R}$. התכונות החשובות ביותר בו של פונקציות כאלו הן רציפות וגזירות. במתמטיקה בדידה, לעומת זאת, יש עניין גדול הרבה יותר בפונקציות, שהתחום שלהן הוא $\mathbf{N}$ או $\mathbf{N}^+$. לפונקציות כאלו קוראים *סדרות*[14]. לפעמים קוראים בשם זה גם לפונקציות, שהתחום שלהן הוא מהצורה $\{1,\ldots,n\}$ (או אולי $\{0,\ldots,(n-1)\}$). שם אחר לפונקציות כאלו הוא "רשימות"). המספר n נקרא אז אורך הסדרה או הרשימה.

עבור סדרות מקובלים מאוד, לדאבוננו, סימונים, שהינם שונים מאלה המקובלים לגבי שאר הפונקציות. מקובל, למשל, לכתוב "f_n" במקום $f(n)$ (ולקרוא לזה "האיבר ה- n-י של הסדרה" במקום: "ערכה של f עבור הארגומנט n"). יתר על כן, במקום להשתמש בסימון-למדא ולכתוב $\lambda n \in \mathbf{N}^+.\ t(n)$ (או רק $\lambda n.\ t(n)$, אם ברור שהתחום הוא $\mathbf{N}^+$), כותבים $\{t_n\}_{n=1}^{\infty}$ (או $(t_n)_{n=1}^{\infty}$). לדוגמה: במקום לכתוב $\lambda n \in \mathbf{N}^+.\frac{1}{n}$, כותבים בטקסטים בחדו"א $\{\frac{1}{n}\}_{n=1}^{\infty}$. אין למעשה שום הבדל במובן של שני הסימונים. אי לכך ברור,

---

[13] כדאי להעיר שב- LISP (או SCHEME) כותבים, משום מה, CAR במקום π_1 ו- CDR במקום π_2.

[14] שני סוגים של סדרות, שיש להם חשיבות רבה גם בחשבון אינפיניטסימלי, הן סדרות עם טווח $\mathbf{R}$ (ונקראות "סדרות של מספרים ממשיים"), וסדרות עם טווח $\mathbf{R} \xrightarrow{p} \mathbf{R}$ (ונקראות "סדרות של פונקציות ממשיות").

שבביטוי מהצורה $\{t(n)\}_{n=1}^{\infty}$ המשתנה n הינו קשור. כלל (α) תופס אפוא כאן: אין הבדל בין הסדרה $\{\frac{1}{n}\}_{n=1}^{\infty}$ והסדרה $\{\frac{1}{k}\}_{k=1}^{\infty}$.

לעובדה, שעבור סדרות משתמשים בצורת סימון שונה, יש מספר השלכות לא נעימות. ראשית, למרות שאין הבדל בין $\lambda n \in \mathbf{N}^{+}.\ t(n)$ ובין $\{t(n)\}_{n=1}^{\infty}$, השימוש בכלל (β) אינו מוכר עבור הסימון האחרון, ולכן צורות הניסוח של טענות וחישובים בטקסטים בחדו"א מסובכות לעתים קרובות ללא כל הצדקה. יתר על כן: קיימת נטייה גדולה שם להוריד את הסוגריים המסולסלים ה"מעצבנים", ולכתוב: $a_n \to a$ במקום $\lambda n \in N^{+}.\ a_n \to a$ (או, לפחות, $\{a_n\}_{n=1}^{\infty} \to a$ או $\lim\limits_{n\to\infty} a_n = a$). צורת כתיבה זו פשוט מזמינה בלבול: כך בנוסחה $\frac{1}{n} < \varepsilon$ הביטוי "$\frac{1}{n}$" מייצג *מספר* (שערכו תלוי בערכו של n), בעוד ב"נוסחה" $\frac{1}{n} \to 0$ הביטוי "$\frac{1}{n}$" לא מייצג בעצם מספר, כי אם את ה*סדרה* $\lambda n \in \mathbf{N}^{+}.\ \frac{1}{n}$. אין אבל שום דבר בצורת הביטויים, שמרמז על ההבדל המהותי, או על כך, ש- n הינו קשור ב- "$\frac{1}{n} \to 0$" וחופשי ב- "$\frac{1}{n} < \varepsilon$". יש לנהוג אפוא זהירות יתרה בסימונים המקובלים (בעיקר כשיש פרמטרים נוספים), ובכל מקרה של חשש לבלבול – להשתמש בסימונים שלמדנו (ובמיוחד בסימון-למדא)! (למען ההגינות נציין, שכאשר מדובר בסדרות של *פונקציות*, הרי יש לסימון המקובל יתרון פסיכולוגי מסוים. אם $f: \mathbf{N} \to (\mathbf{R} \to \mathbf{R})$, הרי נוח יותר לעתים לדבר על $f_7(\sqrt{2})$, למשל, מאשר על $((f(7))(\sqrt{2})$.)

III. הגדרת פונקציות כסוג מיוחד של קבוצות

בסעיף זה נראה, כיצד ניתן להגדיר פונקציות במונחים היסודיים של תורת הקבוצות. מכאן ואילך יהיו ההגדרות, שנכניס בסעיף זה, ההגדרות הרשמיות של המושגים הקשורים בפונקציות (הגם שאינן חיוניות לצורך הבנת מרבית מה שבא בהמשך). מושג המפתח הוא המושג הבא:

הגדרה:
הגרף של פונקציה f, שתחומה A, הוא הקבוצה $G(f) = \{<x, f(x)> \mid x \in A\}$.

לגרף $G(f)$ של פונקציה f יש התכונות הבאות:

(א) אם $z \in G(f)$ אז קיימים a ו- b כך ש- $z = <a,b>$.

(ב) אם $<a, b_1> \in G(f)$ ו- $<a, b_2> \in G(f)$ אז $b_1 = b_2$.

בתורת הקבוצות (ובמתמטיקה המודרנית בכלל) נהוג ***לזהות*** פונקציה עם הגרף שלה (שהוא קבוצה, כמובן). בהתאם, מגדירים פונקציה כקבוצה, שיש לה התכונות שפירטנו לעיל.

הגדרה:

1. פונקציה f היא קבוצה של זוגות סדורים, כך שאם $<a,b_1> \in f$ ו- $<a, b_2> \in f$, אז $b_1 = b_2$.
2. אם f פונקציה, אז הקבוצה $\{x | \exists y. <x,y> \in f\}$ נקראת ה***תחום*** של f ומסומנת ב- $\mathrm{Dom}(f)$. כאשר $A = \mathrm{Dom}(f)$, אומרים ש- f היא פונקציה ***מעל*** A. אם $\mathrm{Dom}(f) \subseteq A$, אומרים ש- f היא פונקציה חלקית מעל A.

תרגיל

הראו ש- $\mathrm{Dom}(f) \in P(\cup(\cup(f)))$, ולכן $\mathrm{Dom}(f)$ אכן קבוצה.

הגדרה:

נניח ש- f פונקציה. אם $f \subseteq A \times B$, אומרים ש- f היא פונקציה חלקית מ- A ל- B, וש- B היא טווח של f. אם, בנוסף, $A = \mathrm{Dom}(f)$, אומרים ש- f פונקציה מ- A ל- B, ומסמנים $f : A \to B$ או $f \in A \to B$. (הגדרה מדויקת של $A \to B$ וכן של $A \xrightarrow{p} B$ כקבוצות ניתנת בטבלה ב.5 למטה).

תרגיל

הראו ש- $(A \to B) \in P(P(A \times B))$.

נשים לב, שאם $f : A \to B$, אז:

$$\forall x \in A\ \exists! y \in B.\ <x,y> \in f$$

כאשר $\exists! y. \psi$ פירושו, שיש y יחיד שיש לו התכונה ψ, דהיינו:

$$\exists! y.\ \psi =_{Df} \exists y.\ \psi \wedge \forall y_1\ \forall y_2.\ \psi(y_1/y) \wedge \psi(y_2/y) \Rightarrow y_1 = y_2$$

או, בניסוח שקול לוגית אחר:

$$\exists! y.\ \psi =_{Df} \exists y(\psi \wedge \forall z.\ \psi(z/y) \to z = y))$$

סימון:

אם f פונקציה ו- $x \in \mathrm{Dom}(f)$, אז נסמן ב- $f(x)$ את ה- y היחיד, כך ש- $<x,y> \in f$:

$$f(x) =_{Df} \iota y.\ <x,y> \in f$$

נפנה עתה לסוגייה של שוויון פונקציות. בתחילת פרק זה ניסחנו את עיקרון האקסטנציונליות עבור פונקציות. באותו שלב מילא עיקרון זה, בעצם, תפקיד של הגדרה: הוא הגדיר מתי שתי פונקציות נחשבות שוות. עתה, כאשר פונקציות הינן קבוצות, הרי שוויון ביניהן כמוהו כשוויון בין כל שתי קבוצות אחרות, דהיינו $f=g$ אם"ם $\forall x.\, x \in f \Leftrightarrow x \in g$. בשלב זה הופך, אפוא, העיקרון הנ"ל ***לטענה***, שיש להוכיחה:

טענה (עיקרון האקסטנציונליות עבור פונקציות):

אם f ו- g פונקציות, אז:

$$f=g \Leftrightarrow [\mathrm{Dom}(f)=\mathrm{Dom}(g) \wedge \forall x \in \mathrm{Dom}(f).\, f(x)=g(x)]$$

הוכחה:

($\Rightarrow$) נניח $f=g$. אז לפי עקרונות הלוגיקה, כל מה שנכון לגבי f נכון לגבי g, ולהיפך, בפרט $\mathrm{Dom}(f)=\mathrm{Dom}(g)$, ולכל x $f(x)$ מוגדר אם"ם $g(x)$ מוגדר והם שווים (בתנאי שאחד מהם אכן מוגדר).

($\Leftarrow$) נניח ש- $\mathrm{Dom}(f)=\mathrm{Dom}(g)$, ושלכל $x \in \mathrm{Dom}(f)$ מתקיים ש- $f(x)=g(x)$. נוכיח ש- $f=g$, כלומר: ש- $f \subseteq g$ ו- $g \subseteq f$. נראה, לדוגמה, ש- $f \subseteq g$ (ההוכחה ש- $g \subseteq f$ היא דומה). נניח אפוא, ש- $z \in f$. כיוון ש- f פונקציה, קיימים x ו- y כך ש- $z=\langle x,y\rangle$. מהגדרת $\mathrm{Dom}(f)$ נובע לכן, ש- $x \in \mathrm{Dom}(f)$ ו- $y=f(x)$. כיוון ש- $\mathrm{Dom}(f)=\mathrm{Dom}(g)$, נובע מזה ש- $x \in \mathrm{Dom}(g)$, ולכן (לפי הנתון השני) $y=f(x)=g(x)$. אבל העובדה ש- $x \in \mathrm{Dom}(g)$ ו- $y=g(x)$ פירושה, מהגדרותינו, ש- $\langle x,y\rangle \in g$. במלים אחרות: $z \in g$.

לבסוף, סימון λ גם הוא אינו, עקרונית, אלא סימון מקוצר. למעשה:

$$\lambda x.\, t = \{z \mid \exists x.\, z=\langle x, t\rangle\}$$
$$\lambda x \in A.\, t = \{\langle x, t\rangle \mid x \in A\}$$

הערות:

א. t יכיל כאן, בדרך-כלל, מופעים חופשיים של x.

ב. $\lambda x.\, t$ הוא פונקציה אח"ח היטוי באגף ימין של הגדרתו הרשמית הינו ביטוי כשר, המתאר קבוצה. עם זאת, בתפקידו כמתאר ***כלל***, יש ל- $\lambda x.\, t$ מובן גם אם אינו קבוצה.

ג. $\lambda x \in A.\, t$ מייצג תמיד פונקציה ***חלקית*** מעל A. הוא מייצג ***פונקציה*** מעל A אם t מוגדר עבור כל $x \in A$.

ד. אפשר להראות ללא קושי, שעם ההגדרות הרשמיות למעלה, כללי α, β ו- η עבור λ נובעים מהכללים המקבילים עבור קבוצות.

טבלה ב.5 מסכמת את ההגדרות הרשמיות הבסיסיות הקשורות בפונקציות. היא כוללת גם את ההגדרות הרשמיות של הרכבת פונקציות ושל פונקציה הפוכה (מושגים אותם נלמד בהמשך).

טבלה ב.5: פונקציות – הגדרות בסיסיות

(1)	קבוצה f נקראת ***פונקציה*** אם מתקיים: (א) $\forall z \in f \exists a\, \exists b.\ z = \langle a,b\rangle$ (ב) $\forall a\, \forall b_1\, \forall b_2.\ \langle a, b_1\rangle \in f \wedge \langle a, b_2\rangle \in f \Rightarrow b_1 = b_2$
(2)	תהי f פונקציה, אז: (i) $\underline{\mathrm{Dom}(f)} = \{x \mid \exists y.\ \langle x,y\rangle \in f\}$ (ii) f היא <u>פונקציה חלקית מעל A</u> אם $\mathrm{Dom}(f) \subseteq A$ (iii) f היא <u>פונקציה מעל A</u> אם $\mathrm{Dom}(f) = A$ (iv) f היא <u>פונקציה חלקית מ- A ל- B</u> אם $f \subseteq A \times B$ (v) f היא <u>פונקציה מ- A ל- B</u> אם $f \subseteq A \times B \wedge \mathrm{Dom}(f) = A$
(3)	אם f פונקציה ו- $x \in \mathrm{Dom}(f)$, אז: $\underline{f(x)} =_{Df} \iota y.\ \langle x,y\rangle \in f$
(4)	אם $f: A \to B$ ו- $g: B \to C$, אז: $\underline{g \circ f} = \{\langle a,c\rangle \in A \times C \mid \exists b \in B.\ \langle a,b\rangle \in f \wedge \langle b,c\rangle \in g\}$
(5)	אם $f: A \to B$ פונקציית שקילות, אז: $\underline{f^{-1}} = \{\langle b,a\rangle \in B \times A \mid \langle a,b\rangle \in f\}$
(6)	**טענה:** אם f ו- g פונקציות, אז: $f = g \Leftrightarrow [\mathrm{Dom}(f) = \mathrm{Dom}(g) \wedge \forall x \in \mathrm{Dom}(f).\ f(x) = g(x)]$
(7)	(i) $\lambda x.\ t = \{z \mid \exists x.\ z = \langle x, t\rangle\}$ (ii) $\lambda x \in A.\ t = \{\langle x, t\rangle \mid x \in A\}$
(8)	(i) $A \to B = \{f \in P(A \times B) \mid$ (f פונקציה) $\wedge\ \mathrm{Dom}(f) = A\}$ (ii) $A \xrightarrow{p} B = \{f \in P(A \times B) \mid$ (f פונקציה) $\wedge\ \mathrm{Dom}(f) \subseteq A\}$

IV. דוגמאות של פונקציות חשובות

א. תהי A קבוצה. פונקצית ***הזהות*** על A היא הפונקציה:

$$i_A = \lambda x \in A.\, x$$

במלים אחרות: $i_A(x) = x$ לכל $x \in A$.

ב. תהיינה A ו- B קבוצות. פונקציה ***קבועה*** מ- A ל- B היא פונקציה מהצורה $\lambda x \in A.\, c$, כאשר c איבר של B. לדוגמה: $\lambda x \in R.2$ היא פונקציה המחזירה לכל מספר ממשי את הערך 2 ("$y = 2$" כותבים לפעמים בחדו"א).

חשוב להבדיל בין c, שהוא איבר של B, ובין פונקציה הקבועה עליו, שאינה איבר של B, אלא איבר בקבוצה מהצורה $A \to B$. כך, למשל, אין זה נכון שהנגזרת של $\lambda x \in R.2$ היא אפס! מה שנכון הוא:

$$(\lambda x \in \mathbf{R}.\, 2)' = \lambda x \in \mathbf{R}.\, 0$$

ג. תהי A קבוצה. ***הפונקציה האופיינית*** של A מסומנת בדרך-כלל על-ידי χ_A ומוגדרת בצורה הבאה:

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

הגדרה מסוג זה נקראת הגדרה לפי מקרים. אלטרנטיבית נוכל להשתמש בסגנון ההגדרה הבא:

$$\chi_A = \lambda x.\ \text{if } x \in A \text{ then } 1 \text{ else } 0$$

כך או כך, ברור ש:

$$\forall x(x \in A \iff \chi_A(x) = 1)$$

מכאן, שהפונקציה האופיינית של A כשמה כן היא: היא מאפיינת את A לחלוטין. הכרה מושלמת של χ_A פירושה הכרה מושלמת של A, ולהפך. ניתן היה לכן לקחת דווקא את מושג הפונקציה כמושג יסודי, ולראות בקבוצות סוג מיוחד של פונקציות: אלו עם טווח $\{0,1\}$ (בתורת הפונקציות הַחֲשִׁיבוֹת (computable) זו אכן הפרוצדורה המקובלת).

שאלה אחת בקשר ל- χ_A נשארה פתוחה למעלה: מהו, בעצם, תחום ההגדרה שלה? ממבט ראשון זהו "כל העולם". ברם, ראינו, שאוסף כל העצמים אינו מהווה

קבוצה ולכן אינו יכול להיות תחום הגדרה של פונקציה. χ_A אינה, לכן, פונקציה במובן שהגדרנו למעלה.[15] ניתן אבל להופכה לכזו אם מגבילים אותה לאיזו קבוצה מקיפה E, כך ש- $A \subseteq E$ אם נגדיר, במקרה כזה,

$$\chi_A^{(E)} = \lambda x \in E.\ \text{if } x \in A \text{ then } 1 \text{ else } 0$$

אז:

$$\chi_A^{(E)} \in E \to \{0,1\}$$

הערה:

בדרך-כלל משמיטים את ההתייחסות ל- E, אם היא ברורה מהקונטקסט, או אם זהותה המדויקת אינה חשובה (השווי לדיון ב- $\overline{A}$ בפרק ב.3).

ד. ניתן ללכת צעד נוסף, מעבר למה שנעשה בדוגמה הקודמת, ולהגדיר את הפונקציה $\lambda A \in P(E).\chi_A^{(E)}$. זוהי פונקציה מ- $P(E)$ אל $E \to \{0,1\}$. במלים אחרות: זוהי פונקציה, שהארגומנטים שלה (ה"קלטים", בלשון מדעי המחשב) הם קבוצות, וערכיה (ה"פלטים") הם עצמם פונקציות. נדגיש שוב, שפונקציות כמו $\lambda A \in P(E).\chi_A^{(E)}$ הן עצמים לגיטימיים לחלוטין ואף מועילים. לרוע מזלנו, לימודינו בבית-הספר התיכון הרגילו אותנו להבין במושג "פונקציה" רק התאמה מקבוצה אחת של מספרים אל קבוצה אחרת של מספרים. הכרחי לכן להתגבר על הרגל זה ולעכל את הרעיון, שאפשריות וקיימות פונקציות *מכל* קבוצה (אפילו קבוצה של קבוצות או קבוצה של פונקציות!) אל *כל* קבוצה אחרת! אגב, אין כאן רק "מוזרות מתמטית". בשפות מחשב פונקציונליות, כמו LISP, SCHEME או SML, פונקציות הינן אכן "אזרחיות מדרגה ראשונה", דהיינו: הן יכולות לשמש הן בתפקיד של קלט והן בתפקיד של פלט, ממש כמו מספרים!

הבה נשתמש בדוגמה הנוכחית למטרה נוספת. נשים לב, שהפונקציה, שהוגדרה בה (אם נכתוב אותה באופן מלא, ללא השימוש בשם "$\chi_A^{(E)}$"), היא:

$$(*) \quad \lambda A \in P(E).\ \lambda x \in E.\ \text{if } x \in A \text{ then } 1 \text{ else } 0$$

הבה ננתח ביטוי זה כדוגמה, כיצד נעשה ניתוח כזה באופן כללי. ובכן, הביטוי (*) מתחיל ב- "$\lambda A \in P(E)$". זה מלמד, ש- (*) מייצג פונקציה, שהתחום שלה הוא $P(E)$. ההמשך (מה שבא אחרי הנקודה הראשונה) מלמד, מה פונקציה זו מתאימה לכל איבר A בתחום זה, ומניתוחו נוכל למצוא טווח שלה. ובכן, מה ש- (*) מתאים

[15] χ_A היא פונקציה במובן רחב יותר מזה שהגדרנו, מובן שבו הדגש הוא על הכלל, *החוקיות*. סימון למדא נועד, בעצם, קודם כל, לתאר פונקציות במובן רחב זה!

ל- $A \in P(E)$ הוא משהו, שמתחיל ב- "$\lambda x \in E$". מכאן למדים אנו, ש- (*) מתאים לכל A ב- $P(E)$ פונקציה מסוימת, שהתחום שלה E. זהותה של פונקציה זו תתברר מתוך ניתוח ההמשך: $\text{if } x \in A \text{ then } 1 \text{ else } 0$, אנו רואים מיד, שערך ביטוי זה יכול להיות 0 או 1 (תלוי בערכו של x). מכאן אנו למדים, שטווח ודאי של הפונקציה, ש- (*) מתאים לאיזה $A \in P(E)$, הוא $\{0,1\}$. סיכום כל זאת מראה, ש-(*) הוא פונקציה, שתחומה $P(E)$, ומתאימה לכל $A \in P(E)$ פונקציה, שתחומה E, ו- $\{0,1\}$ הוא טווח שלה. (*) מגדיר לכן איבר של $P(E) \to (E \to \{0,1\})$. (*) כולל, כמובן, גם אינפורמציה מדויקת על ***פעולתו*** של איבר זה (כפונקציה מעל $P(E)$).

דוגמה:

ניקח את E בתור $\mathbf{N}$, ונסמן את קבוצת המספרים הטבעיים הזוגיים ב- $\mathbf{N}_{\text{even}}$. אז:

$$\big(\underbrace{(\lambda A \in P(\mathbf{N}).\ \lambda x \in \mathbf{N}.\ \text{if } x \in A \text{ then } 1 \text{ else } 0)\,(\mathbf{N}_{\text{even}})}\big)\,(7)$$

$$\overset{\beta}{=} (\lambda x \in \mathbf{N}.\ \text{If } x \in \mathbf{N}_{\text{even}} \text{ then } 1 \text{ else } 0)\,(7)$$

$$\overset{\beta}{=} \text{If } 7 \in \mathbf{N}_{\text{even}} \text{ then } 1 \text{ else } 0$$

$$= 0$$

כמו כן:

$$(\lambda A \in P(\mathbf{N}).\ \lambda x \in \mathbf{N}.\ \text{If } x \in A \text{ then } 1 \text{ else } 0)\,(\{n \in \mathbf{N} \mid n \le 20\})$$

$$= \lambda x \in \mathbf{N}.\ \text{If } x \in \{n \in \mathbf{N} \mid n \le 20\} \text{ then } 1 \text{ else } 0$$

$$= \lambda x \in \mathbf{N}.\ \text{If } x \le 20 \text{ then } 1 \text{ else } 0$$

זוהי, כמובן, פונקציה מ- $\mathbf{N}$ אל $\{0,1\}$ (הלא היא $\chi_{\{n \in \mathbf{N} \mid n \le 20\}}$).

ה. אופרטור הגזירה D (המסומן בדרך-כלל על-ידי תג אחרי שם הפונקציה, אותה גוזרים) הינו דוגמה מצוינת לפונקציה, המקבלת פונקציות (חלקיות) מ- $\mathbf{R}$ ל- $\mathbf{R}$ כקלטים, ומחזירה פונקציות כאלה כפלטים:

$$D : (\mathbf{R} \overset{p}{\to} \mathbf{R}) \to (\mathbf{R} \overset{p}{\to} \mathbf{R})$$

כך, לדוגמה:

$$D(\lambda x.\ x^2) = \lambda x.\ 2x$$

(השמטנו כאן, כמקובל, את ההתייחסות ל- $\mathbf{R}$, ולא כתבנו $\lambda x \in \mathbf{R}.\ 2x$, כי התחום כאן ברור, ובמקרה כזה מוטב לקצר).

השימוש בסימון $D(f)$ לציון הנגזרת של הפונקציה f מקובל באמת רק בשיטות מתמטיות מתקדמות (בהן חיוני להתייחס אל D כאל עצם, עליו מבצעים פעולות שונות). בדרך-כלל, כידוע, כותבים f' במקום $D(f)$. נוהג זה אינו שונה, במהותו, מהנוהג לכתוב $n!$ במקום $!(n)$, וברוב המקרים אינו יוצר בעיות מיוחדות. מה שיוצרת גם יוצרת בעיות היא העובדה, שזהויות כמו $(\lambda x.\ x^2)' = \lambda x.\ 2x$ כותבים בטקסטים בחדו"א בצורה המאוד לא-מוצלחת הבאה: $(x^2)' = 2x$. ביטויים כמו "x^2" ו- "$2x$" הינם באופן טבעי ביטויים עבור *מספרים*, וערכם תלוי בערך המשתנה החופשי x, המופיע בהם. בנוסחאות כמו $(x^2)' = 2x$, לעומת זאת, משתמשים ב- "x^2" וב- "$2x$" כביטויים עבור *פונקציות*. למעלה מזאת: המשתנה x קשור בהם (אי אפשר, למשל, להציב במקומו ערכים קונקרטיים: "$(7^2)' = 2 \cdot 7$" הוא, כפי שכבר ראינו, צירוף *חסר משמעות*. כמו כן, כלל α ישים: ל- $(z^2)' = 2z$ אותה משמעות כמו ל- $(x^2)' = 2x$). למרות זאת, אין רואים שום אופרטור קשירה, שקושר את x! השימוש הכפול בביטוי כמו "x^2", פעם כמספר, פעם כפונקציה, ואי-השימוש באופרטור קשירה לציון משתנים קשורים, הם (שניהם ביחד וכל אחד לחוד) ערובה ודאית לשגיאות ובלבולים – ובמיוחד כשמשתנים נוספים מעורבים בעניין (כמו ב"נוסחה" $(x^a)' = ax^{a-1}$). יתר על כן: זה גורם לכך, שאפילו ניסוחן של עובדות פשוטות מאוד הופך עניין מסורבל וקשה. נניח, לדוגמה, שאנו רוצים לבטא בנוסחה את העובדה, ש- z הוא מספר, שהנגזרת של "x^2" בו שווה ל- $z+1$. איך נעשה זאת? בטקסטים הרגילים, הדרך היחידה הפתוחה בפנינו היא להשתמש בשם זמני, תוך היעזרות במלים בעברית, למשל: "z הוא מספר, כך שאם $f(x) = x^2$, אז $f'(z) = z + 1$". עם סימון נכון, המאפשר להבדיל בין הביטוי "x^2" עבור מספרים, ו- "$\lambda x.\ x^2$" עבור פונקציות, נכתוב פשוט, בנוסחה אחת:

$$(\lambda x.\ x^2)'(z) = z + 1$$

(או, אפילו, אם יתחשק לנו, $(\lambda z.\ z^2)'(z) = z + 1$: אין הרי קשר בין שני המופעים הראשונים של z כאן, שהם קשורים, ובין השניים האחרונים, שהם חופשיים!). ועוד: כאשר משתמשים אנו בסימון נכון, ברור לנו, שהנוסחה $(\lambda x.\ x^2)' = \lambda x.\ 2x$ שקולה ל- $(\lambda x.\ x^2)' = \lambda z.\ 2z$. לעומת זאת, הנוסחה $(x^2)' = 2z$ תיחשב כמעט בוודאות כשגויה. למה? דומה, שמפסיקים להתייחס בה ל- "$2z$" כאל ביטוי המייצג פונקציה, או (מה שגרוע יותר) – לא מתייחסים ל- "$2x$" ול- "$2z$" כביטויים המייצגים אותה פונקציה (למרות שברור, שכאשר מגדירים $f(x) = 2x$ ו- $g(z) = 2z$, אזי f ו- g הן אותה פונקציה – ואין מורה, שלא ידגיש זאת!).

לסיום עניין הנגזרת, נחזור לבעיה של הנוסחה $(x^a)' = ax^{a-1}$, בה דשנו כבר בפרק א.5. שם, כזכור, פירשנו אותה כמבטאת זהות בין שני מספרים (שערכם תלוי בערך של a ושל x). כעת נוכל לתת לה פירוש אחר, הקרוב יותר לאופן, שבו משתמשים בה בדרך-כלל. הפעם נפרשה כמבטאת שוויון בין שתי ***פונקציות*** (שזהותן תלויה בערך של a). באופן מלא ונכון צריכה הנוסחה להיכתב כך:

$$\forall a[(\lambda x.\ x^a)' = \lambda x.\ ax^{a-1}]$$

גם בפירוש זה ברור מיד, למה לא נוכל להציב במקום a ביטוי, המכיל את x כמשתנה חופשי: גם כאן ביטוי כזה אינו חופשי להצבה במקום a בתוך הסוגריים המרובעים, ולכן כלל (14) אינו ישים עבורו. הפעם, אבל, נקשר x על-ידי אופרטורי למדא, לא על-ידי "$\forall x$".

ו. האינטגרל המסוים הינו פונקציה. כדי להבין זאת, הבה נהרהר רגע מה אנו צריכים לדעת כדי שנוכל להתחיל במלאכת חישוב אינטגרל מסוים. ובכן, עלינו לדעת, מהי הפונקציה לה אנו עושים אינטגרל, ואת הגבול התחתון והעליון שלו ("מאין ועד היכן עושים את האינטגרל"). עלינו לקבל אפוא פונקציה ושני מספרים כקלט, ולקבל בחזרה מספר כפלט. ברור, לכן, שאינטגרל מסוים הוא פונקציה "של שלושה משתנים" (שהיא חלקית, כיוון שלא לכל פונקציה אפשר לעשות אינטגרל מכל מספר לכל מספר). ליתר דיוק:

$$\int : ((\mathbf{R} \overset{p}{\to} \mathbf{R}) \times \mathbf{R} \times \mathbf{R}) \overset{p}{\to} \mathbf{R}$$

לדוגמה, בסימון האחיד לפונקציות, הביטוי $\int_1^2 x^2dx$ ייכתב כך: $\int(\lambda x.\ x^2, 1, 2)$.

נשים לב, שהמשפט היסודי של החשבון הדיפרנציאלי והאינטגרלי אומר ש:

$$\forall f \in \mathbf{R} \overset{p}{\to} \mathbf{R}\ \forall a \in \mathbf{R}\ \forall y \in \mathbf{R}\ [(a,y) \text{ -רציפה ב } f \Rightarrow (D(\lambda x. \int(f, a, x)))(y) = f(y)]$$

ז. גם אינטגרל לא מסוים הינו פונקציה. הפעם עלינו להכניס כקלט פונקציה (חלקית) מ- $\mathbf{R}$ ל- $\mathbf{R}$, ומה שנקבל כפלט יהיה *קבוצה* שלמה של פונקציות. לכן:

$$\int : (\mathbf{R} \overset{p}{\to} \mathbf{R}) \to P(\mathbf{R} \overset{p}{\to} \mathbf{R})$$

(אנו משתמשים, כנהוג, באותו הסימן עבור האינטגרל המסוים והלא מסוים. זהו נוהג, שיש לו גם חסרונות וגם מעלות). למשל, כשאנו כותבים:

$$\int x^2 dx = \frac{x^3}{3} + c$$

כוונתנו, בעצם, היא ש:

$$\int (\lambda x.\, x^2) = \{\lambda x.\, \frac{x^3}{3} + c \mid c \in \mathbf{R}\}$$

V. מושגים בסיסיים הקשורים בפונקציות

א. צמצום של פונקציה

נניח $f: A \to B$ ו- $X \subseteq A$ אז f/X, ***הצמצום של f ל- X***, היא הפונקציה $\lambda x \in X.\, f(x)$. f/X היא אפוא פונקציה הזהה ל- f (כמעט), אך עם תחום הגדרה מצומצם יותר.

נשים לב:

1. $f/X : X \to B$
2. במונחי תורת הקבוצות, $f/X \subseteq f$. למעשה: $f/X = \{<x, f(x)> \mid x \in X\}$
3. $f/\mathrm{dom}(f) = f$
4. אם $X \subseteq A$ אז $(\lambda x \in A.\, t) / X = \lambda x \in X.\, t$

נעיר עוד, שניתן להכליל כל זאת לפונקציות חלקיות, בלי לשנות דבר.

ב. מקור ותמונה

נניח $f: A \to B$.

1. כאשר $f(x) = y$ (עבור איזה $x \in A$ ו- $y \in B$), אז קוראים ל- y ***התמונה*** של x לפי y, בעוד x מכונה ***מקור*** של y לפי f (נשים לב, שלכל x ב- A יש תמונה יחידה ב- B, בעוד מספר המקורות של $y \in B$ אינו מוגבל, ויכול גם להיות 0).

2. נגדיר $\hat{f} : P(A) \to P(B)$ על-ידי:

$$\hat{f} = \lambda X \in P(A).\, \{f(x) \mid x \in X\} \; (= \lambda X \in P(A).\, \{y \in B \mid \exists x \in X.\, f(x) = y\})$$

$\hat{f}$ היא פונקציה, המתאימה לכל קבוצה חלקית של A את קבוצת התמונות של איבריה לפי f. אם $X \subseteq A$, אז ל- $\hat{f}(X)$ קוראים ***התמונה*** של X לפי f. מקובל מאוד לסמן את התמונה של X לפי f ב- $f(X)$ במקום $\hat{f}(X)$. זהו נוהג גרוע, כיוון שהוא עלול לגרום לדו-משמעויות ולבלבולים. ברם, כיוון שהוא ***מאוד*** מקובל,

נאמץ אותו בדרך-כלל גם אנו (סימון עדיף, מקובל למדי, הוא $f[X]$, ומדי פעם נשתמש גם בו).

נדגיש שוב: אם $x \in A$, אז $f(x)$ התמונה של x לפי f, היא איבר ב- B ($f(x) \in B$). לעומת זאת, אם $X \subseteq A$, אז $f(X)$ היא איבר ב- $P(B)$ (כלומר $f(X) \subseteq B$). כאשר $x \in A \cap P(A)$ (כלומר, כאשר x גם שייך ל- A וגם חלקי ל- A בעת ובעונה אחת), אז הסימון $f(x)$ הוא דו-משמעי, וזוהי צרה. (הסימון $f[x]$ אינו יוצר בעיה זו!).

דוגמאות:

(א) $(\lambda x \in R.x^2)((-\infty, 3]) = [9,\infty)$

(ב) אם $f = \lambda x \in P(\mathbf{N}).\, 1$ אז:

$f(\varnothing) = 1$ כאשר מסתכלים על $\varnothing$ כעל איבר של $P(\mathbf{N})$.

$f(\varnothing) = \varnothing$ כאשר המובן של "$f(\varnothing)$" הוא "$\hat{f}(\varnothing)$", ומסתכלים על $\varnothing$ כעל קבוצה חלקית של $P(\mathbf{N})$.

במצב כזה נעדיף לכתוב: $f(\varnothing) = 1$ ו- $f[\varnothing] = \varnothing$.

(ג) אם $f: A \to B$ אז $f(A) = \mathrm{Im}(f)$.

3. נניח שוב, ש- $f: A \to B$. נגדיר $\breve{f}: P(B) \to P(A)$ על-ידי:

$$\breve{f} = \lambda X \in P(B).\ \{x \in A \mid f(x) \in X\}$$

$\breve{f}$ היא פונקציה, המתאימה לכל קבוצה חלקית של B את קבוצת מקורותיה לפי f. ל- $\breve{f}(X)$ קוראים *המקור* של X לפי f, ומסמנים אותה בדרך-כלל ב- $f^{-1}(X)$. סימון עדיף, אם כי מקובל פחות, הוא $f^{-1}[X]$ (אנו נשתמש בשניהם).

נדגיש: כאשר $x \in B$, אזי ייתכנו לו מספר מקורות לפי f, וייתכן גם, שאין לו שם שום מקור. לעומת זאת, כש- $X \subseteq B$, אז יש ל- X מקור יחיד לפי f. המונח "מקור של x לפי f" הופך להיות דו-משמעי כש- $x \in B \cap P(B)$. בהמשך נראה, שגם *הסימון* $f^{-1}(x)$ (אך לא $f^{-1}[x]$!) עלול להיות דו-משמעי במקרים מסוימים.

דוגמאות:

(1) אם $f = \lambda x \in \mathbf{R}.\, x^2$, אז

$$f^{-1}([0,1)) = (-1,1)$$

$$f^{-1}([9, \infty)) = (-\infty, -3] \cup [3, \infty)$$

$$f^{-1}([-\frac{1}{2}, 1)) = (-1,1)$$

הבה נפרט את הדוגמה האחרונה. כיוון שכאן $f: \mathbf{R} \to \mathbf{R}$ ו- $[-\frac{1}{2}, 1) \subseteq \mathbf{R}$,

אז לפי הגדרה:

$$\begin{aligned} f^{-1}([-\tfrac{1}{2}, 1)) &= \{x \in \mathbf{R} \mid f(x) \in [-\tfrac{1}{2}, 1)\} \\ &= \{x \in \mathbf{R} \mid x^2 \in [-\tfrac{1}{2}, 1)\} \\ &= \{x \in \mathbf{R} \mid -\tfrac{1}{2} \le x^2 < 1\} \\ &= \{x \in \mathbf{R} \mid x^2 < 1\} \\ &= \{x \in \mathbf{R} \mid -1 < x < 1\} \\ &= (-1, 1) \end{aligned}$$

(2) נניח $g = \lambda x \in \mathbf{N}.\, 1$. אז $g: \mathbf{N} \to \mathbf{N}$ ומתקיים:

$$g^{-1}(\{1,2\}) = \mathbf{N}$$
$$g^{-1}(\{2,3\}) = \varnothing$$

תרגיל

נניח $f: A \to B$ הוכיחי:

(i) $\forall X \in P(A).\, X \subseteq f^{-1}(f(X))$

(ii) $\forall X \in P(B).\, f(f^{-1}(X)) \subseteq X$

מצאי דוגמאות בהן ההכלה ב- (i) וב- (ii) היא הכלה ממש.

ג. הרכבה של פונקציות

אם $f: B \to C$ ו- $g: A \to B$, אז ***ההרכבה*** שלהן, $f \circ g$, הינה הפונקציה $\lambda x \in A.\, f(g(x))$.

נשים לב:

1. $f \circ g: A \to C$
2. $\forall x \in A.\, (f \circ g)(x) = f(g(x))$
3. שתי ההבחנות הקודמות מהוות ביחד הגדרה שקולה ל- $f \circ g$.
4. הרכבה $f \circ g$ של שתי פונקציות f ו- g מוגדרת אם"ם התמונה של g חלקית לתחום של f.

5. ייתכן ש- $f \circ g$ תהיה מוגדרת, בעוד $g \circ f$ לא.
6. לפי כלל (α), יכולנו בהגדרת $f \circ g$ להשתמש במשתנה אחר במקום x לדוגמה: $f \circ g = \lambda y. f(g(y))$. יש אבל להיזהר מלהשתמש במשתנה, המופיע חופשי בביטויים המייצגים את הפונקציות f ו- g.

דוגמאות:

(1)
$$(\lambda x. x^2) \circ (\lambda x. x+1) = \lambda x. (\lambda x. x^2)\underbrace{((\lambda x. x+1)(x))}$$
$$= \lambda x. \underbrace{(\lambda x. x^2)(x+1)}$$
$$= \lambda x. (x+1)^2$$

הסברים:

א. שוב כתבנו לשם הנוחות רק $\lambda x. x^2$ במקום $\lambda x \in \mathbf{R}. x^2$ וכו'.

ב. כדי להבין את החישוב, סימנו בכל שלב ב- $\underbrace{\qquad}$ את תת הביטוי עליו מופעל כלל (β).

ג. לצורך הנוחות יכולנו להפעיל גם את כלל (α) בכל מקום בו x נקשר. היה אכן יכול להיות ברור יותר, אולי, אם היינו מתחילים כך:
$$(\lambda x. x^2) \circ (\lambda x. x+1) = (\lambda y. y^2) \circ (\lambda z. z+1) = \ldots$$

בדוגמה הבאה נדגים זאת.

(2)
$$(\lambda x. x+1) \circ (\lambda x. x^2) \overset{\alpha}{=} \lambda x. (\lambda y. y+1)\underbrace{((\lambda z. z^2)(x))}$$
$$\overset{\beta}{=} \lambda x. \underbrace{(\lambda y. y+1)(x^2)}$$
$$\overset{\beta}{=} \lambda x. x^2 + 1$$

תכונות של פעולת ההרכבה:

חוק הקיבוץ:

אם $h: A \to B$, $g: B \to C$, $f: C \to D$, אז
$$f \circ (g \circ h) = (f \circ g) \circ h$$

הוכחה

תחום ההגדרה של שני האגפים הוא A (ו- D הוא טווח).

כמו כן, לכל $x \in A$ מתקיים:

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x)))$$

$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$$

לכן $(f \circ (g \circ h))(x) = ((f \circ g) \circ h)(x)$ לכל $x \in A$. מכאן שהפונקציות שוות.

כרגיל, חוק הקיבוץ מאפשר לנו לכתוב פשוט $f \circ g \circ h$ בלי לשים סוגריים, כי מקום הסוגריים אינו משנה. כאשר $f : A \to A$, הרי ניתן להרכיב את f עם עצמו מספר כלשהו של פעמים: $f \circ f \circ \ldots \circ f$. ביטוי זה מקצרים ל- f^n (כאשר n הוא מספר הפעמים ש-f מופיע כאן). למשל: $f^4 = f \circ f \circ f \circ f$.

מה בדבר החוק הקומוטטיבי (חוק החילוף)? ובכן, כאמור למעלה, בדרך כלל $g \circ f$ אינו מוגדר כלל כאשר $f \circ g$ מוגדר. שניהם מוגדרים רק כש- $f : A \to B$ ו- $g : B \to A$. אפילו אז ברוב המקרים $f \circ g \neq g \circ f$, כמו שמראות כבר שתי הדוגמאות הפשוטות, שהבאנו למעלה.

שני מקרים, שחשוב (וקל) לזכור, בהם כן מתקיים החוק הקומוטטיבי, הם:

(i) אם $f : A \to A$, אז $f^m \circ f^n = f^n \circ f^m$. הסיבה: שני האגפים שווים פשוט ל- f^{m+n}. למשל:

$$f^3 \circ f^4 = (f \circ f \circ f) \circ (f \circ f \circ f \circ f) = f \circ f \circ f \circ f \circ f \circ f \circ f$$

(ההוכחה של הנוסחה הכללית היא דומה. נעיר, שגם הנוסחה $(f^n)^m = f^{n \cdot m}$ מתקיימת כאן, כש- $n,m \in \mathbf{N}^+$. ההוכחה שוב קלה).

(ii) אם $f : A \to A$, אז $f \circ i_A = i_A \circ f = f$

ואכן לכל $x \in A$:

$$(f \circ i_A)(x) = f(i_A(x)) = f(x)$$

$$(i_A \circ f)(x) = i_A(f(x)) = f(x)$$

נהוג להגדיר $f^0 = i_A$ כאשר $f : A \to A$. עם הגדרה זו הנוסחאות $f^m \circ f^n = f^{m+n}$ ו- $(f^n)^m = f^{n \cdot m}$ נכונות לכל n ו- m ב- $\mathbf{N}$.

את העובדות על i_A, שראינו כאן, אפשר להכליל באופן הבא:

טענה:

אם $f : A \to B$, אז $f \circ i_A = f$ ו- $i_B \circ f = f$.

את ההוכחה נשאיר כתרגיל לקורא.

%%

תת-סדרות

דוגמה חשובה לשימוש בהרכבת פונקציות היא מה שנקרא ***"תת-סדרות"***. אם a סדרה של מספרים ממשיים (כלומר $a : \mathbf{N}^+ \to \mathbf{R}$), ו- n הינה ***סדרה*** (***לא*** מספר!) עולה של מספרים טבעיים (כלומר $n : \mathbf{N}^+ \to \mathbf{N}$, ומתקיים ש- $\forall i \forall j.\ i < j \Rightarrow n_i < n_j$ (דהיינו: $\forall i \forall j.\ i < j \Rightarrow n(i) < n(j)$), אז הרכבתן $a \circ n$ נקראת תת-סדרה של a.

לדוגמה: אם $a = \lambda n \in \mathrm{N}^+.\frac{1}{n}$ $\left(= \left\{\frac{1}{n}\right\}_{n=1}^{\infty}\right)$ ו- $n = \lambda k \in \mathbf{N}^+.\, k^2$ $\left(= \left\{k^2\right\}_{k=1}^{\infty}\right)$,

אז

$$\begin{aligned} a \circ n &= \lambda k \in \mathbf{N}^+.\, a(n(k)) && (= \lambda k \in \mathbf{N}^+.\, a_{n_k}) \\ &= \lambda k \in \mathbf{N}^+.\, a(k^2) && (= \lambda k \in \mathbf{N}^+.\, a_{k^2}) \\ &= \lambda k \in \mathbf{N}^+.\frac{1}{k^2} \\ &\overset{\alpha}{=} \lambda n \in \mathbf{N}^+.\frac{1}{n^2} \end{aligned}$$

המסקנה היא, ש- $\lambda n.\frac{1}{n^2}$ היא תת-סדרה של $\lambda n.\frac{1}{n}$ (או, במלים יותר סטנדרטיות: $\left\{\frac{1}{n^2}\right\}_{n=1}^{\infty}$ היא תת-סדרה של $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$).

כמו שנרמז בסוגריים בחישובים למעלה, במקום לכתוב $a(n(k))$ או $(a \circ n)(k)$, כותבים a_{n_k}. חשוב מאוד להבין, שהשוני בסימון אינו שוני בתוכן. גם כשכותבים "a_{n_k}", a ו- n הם שמות של סדרות (או משתנים עבור סדרות), בעוד k – משתנה עבור מספרים טבעיים. $\left\{a_{n_k}\right\}_{k=1}^{\infty}$ אינו אלא $\lambda k.\, a(n(k))$!

הערה:

כדאי לשים לב, שבחישוב למעלה אין שום קשר בין המשתנה n בשורה האחרונה $(\lambda n \in \mathrm{N}.\frac{1}{n^2})$ לבין ה- n שבביטוי $a \circ n$!

%%

ד. חד-חד-ערכיות, על, שקילות, פונקציה הפוכה

הגדרה:

1. פונקציה f נקראת ***חד-חד-ערכית*** (***ח.ח.ע.***) אם

$$\forall x \in \mathrm{Dom}(f)\ \forall y \in \mathrm{Dom}(f).\ (f(x) = f(y) \Rightarrow x = y)$$

או, בניסוח שקול לוגית:

$$\forall x \in \mathrm{Dom}(f)\ \forall y \in \mathrm{Dom}(f)\ (x \neq y \Rightarrow f(x) \neq f(y))$$

2. פונקציה $f: A \to B$ נקראת **על**[16] B אם B הינה התמונה של A לפי f, דהיינו: כל ערך של B מתקבל. פורמלית, f היא על B אם:

$$\forall y \in B\ \exists x \in A.\ y = f(x)$$

3. פונקציה $f: A \to B$, שהיא גם ח.ח.ע. וגם על B, נקראת ***פונקציית שקילות*** בין A ל- B (או מ- A על B). f היא לכן פונקציית שקילות בין A ל- B אם:

$$\forall y \in B\ \exists! x \in A.\ y = f(x)$$

דוגמאות:

1. נתחיל בפונקציות מ- **R** ל- **R**.

(i) $\lambda x.\ e^x$ היא ח.ח.ע., אך לא על **R** (כי $e^x > 0$ לכל x).

(ii) $\lambda x.\ x^2$ היא לא ח.ח.ע. ולא על **R** ($(\lambda x.\ x^2)(-1) = (\lambda x.\ x^2)(1) = 1$ ו- $\forall x.\ x^2 > 0$).

(iii) $f = \lambda x.\ x^2(x-1)$ היא על **R**, אך היא אינה ח.ח.ע. ($f(0) = f(1) = 0$, ולכן f אינה ח.ח.ע.. שהיא על **R** נובע מהעובדה, שהיא רציפה וש-

$$\lim_{x \to -\infty} f(x) = -\infty \quad , \lim_{x \to \infty} f(x) = \infty$$

זה מאפשר להפעיל את משפט ערך הביניים של החשבון הדיפרנציאלי).

(iv) $\lambda x.\ x+1$ היא גם ח.ח.ע. וגם על **R**.

---

[16] לא לבלבל עם "מעל" A ! (באנגלית "on" ו- "onto").

2. לכל A הפונקציה i_A היא פונקציית שקילות מ- A על A.

3. הפונקציה: $\lambda x \in \{\varnothing,\{\varnothing\}\}.\ \text{if } x = \varnothing \text{ then } 1 \text{ else } 2$ היא פונקציית שקילות מ- $\{\varnothing,\{\varnothing\}\}$ על $\{1,2\}$. היא איננה על $\{1,2,3\}$ (אם זה נלקח בתור הטווח), אך עדיין ח.ח.ע. בטווח זה (ובכל טווח אחר).

4. $\lambda A \in P(E).\chi_A^{(E)}$ (ראה סעיף (IV) למעלה) היא פונקציית שקילות בין $P(E)$ ובין $E \to \{0,1\}$. (*תרגיל:* להוכיח זאת!).

5. אם f היא פונקציה ח.ח.ע., אז f היא פונקציית שקילות בין התחום של f והתמונה של f.

המשפט הבא הוא שימושי, עת רוצים להוכיח, שפונקציה מסוימת היא פונקציית שקילות:

משפט:

$f: A \to B$ היא פונקציית שקילות מ- A על B אם"ם קיימת פונקציה $g: B \to A$, כך ש- $g \circ f = i_A$ ו- $f \circ g = i_B$. יתר על כן: אם פונקציה כזו קיימת, אז היא יחידה.

הוכחה:

$(\Rightarrow)$ נניח שפונקציה g כנ"ל קיימת. נראה ש-f פונקציית שקילות מ- A על B.

ח.ח.ע.:

נניח ש- $f(x) = f(y)$ ל- y,x כלשהם ב- A.

$$\begin{aligned}
&\Rightarrow g(f(x)) = g(f(y))\\
&\Rightarrow (g \circ f)(x) = (g \circ f)(y)\\
&\Rightarrow i_A(x) = i_A(y)\\
&\Rightarrow x = y
\end{aligned}$$

על B:

יהי $y \in B$. נקח $x = g(y)$. אז $x \in A$ ומתקיים:

$$f(x) = f(g(y)) = (f \circ g)(y) = i_B(y) = y$$

לכן $\exists x \in A.\ y = f(x)$.

בטרם נוכיח את הכיוון ההפוך נראה, שאם g כנ"ל קיימת, אז היא יחידה. נניח אפוא, ש- $h : B \to A$ מקיימת אותם תנאים (כלומר: $h \circ f = i_A$, $f \circ h = i_B$). נראה ש- $h = g$. יהי אפוא $y \in B$ כלשהו. אזי:

$$y = i_B(y) = (f \circ h)(y) = (f \circ g)(y)$$

ולכן
$$f(h(y)) = f(g(y))$$

הראינו אבל מקודם, שאם g כנ"ל קיימת, אז f הינה ח.ח.ע.. לכן, מהשוויון האחרון נובע, ש- $h(y) = g(y)$. כיוון שזה נכון לכל $y \in B$, הרי $h = g$.

($\Leftarrow$) נניח $f : A \to B$ ח.ח.ע. ועל B. נגדיר:

$$g = \lambda y \in B.\ \iota x \in A.\ f(x) = y$$

(כלומר: g מתאימה לכל y ב- B את האיבר היחיד ב- A כך ש- $f(x) = y$. איבר כזה קיים, כי f הינה על B, והוא יחיד, כי f ח.ח.ע.). אז:

$$f \circ g = \lambda y \in B.\ f(g(y)) = \lambda y \in B.\ f(\iota x \in A.\ f(x) = y) = \lambda y \in B.\ y = i_B$$

$$g \circ f = \lambda x \in A.g(f(x)) \overset{(\alpha)}{=} \lambda z \in A.\underbrace{g(f(z))}$$

$$\overset{(\beta)}{=} \lambda z \in A.\iota x \in A.f(x) = f(z)$$

$$= \lambda z \in A.z$$

$$= i_A$$

(הסתמכנו כאן על העובדה הטריביאלית, שאם f היא ח.ח.ע., אז $z = (\iota x \in A.\ f(x) = f(z))$, כלומר: ה- x היחידי, שערך f בו שווה אז לערך של f ב- z, הוא z עצמו).

מ.ש.ל.

הגדרה:

אם $f : A \to B$ היא פונקציית שקילות, אז לפונקציה g, שקיומה הובטח במשפט הקודם, קוראים הפונקציה *ההפוכה* של f, ומסמנים אותה ב- f^{-1}. (במקום להגיד, ש- f פונקציית שקילות, אומרים לכן לעתים קרובות, ש- f *הפיכה*).

נשים לב:

(i) $f^{-1} : B \to A$ (או, בצורה כללית יותר: אם f ח.ח.ע., אז $f^{-1} : \mathrm{Im}(f) \to B$).

(ii) $f^{-1} \circ f = i_A \ , \ f \circ f^{-1} = i_B$.

(iii) f^{-1} מתאפיינת על-ידי התכונה:

$$\boxed{\forall x \in A.\ \forall y \in B.\ y = f(x) \Leftrightarrow x = f^{-1}(y)}$$

(iv) לפי ההגדרות הרשמיות ו- (iii):

$$f^{-1} = \{<y,x> \mid <x,y> \in f\}$$

(v) אם $f : A \to A$ היא פונקציית שקילות, אז את (ii) אפשר לרשום בצורה:

$$f \circ f^{-1} = f^{-1} \circ f = f^0$$

אזהרות:

1. הביטוי f^{-1} מייצג פונקציה אך ורק כש- f פונקציה ח.ח.ע.. f^{-1} הינה אז פונקציה מהתמונה של f על התחום של f.

2. במקום אחר בפרק זה השתמשנו בביטוי f^{-1} במובן אחר. הבה נבהיר אפוא: אם $f : A \to B$, אז עבור $x \in B$, הביטוי $f^{-1}(x)$ הוא בעל מובן, רק אם f הינה ח.ח.ע. ועל B. בהתקיים תנאי זה, $f^{-1}(x) \in A$ ומסמן את הערך של הפונקציה ההפוכה ל-f עבור הארגומנט x. כאשר $X \subseteq B$, לעומת זאת, $f^{-1}(X) \in P(A)$ ומוגדר אפילו אם f איננה ח.ח.ע. ועל (ולמעשה אפילו כש-f רק חלקית!). $f^{-1}(X)$ מייצג אז את המקור של X לפי f. לבסוף, אם f ח.ח.ע. ועל B, ו- $x \in B \cap P(B)$, אז הביטוי $f^{-1}(x)$ הוא דו-משמעי!

דוגמאות:

1. $(\lambda x \in \mathbf{R}.\ x + 1)^{-1} = \lambda x \in \mathbf{R}.\ x - 1$

שאלה 1

כיצד מוכיחים זאת?

תשובה

מרכיבים את $\lambda x.\ x + 1$ ואת $\lambda x.\ x - 1$ בשתי הצורות האפשריות, ורואים, שבשתיהן יוצא $i_{\mathbf{R}}$ (כלומר: $\lambda x \in \mathbf{R}.\ x$).

שאלה 2

כיצד מוצאים זאת?

תשובה

מציבים $y = x + 1$ ומנסים "לבודד" את x, דהיינו: לבטא את x בעזרת y. כאן זה קל: $x = y - 1$. המסקנה היא, ש- $\lambda y.\, y - 1$ היא הפונקציה ההפוכה, או, בעזרת כלל (α): $\lambda x.\, x - 1$.

---

2. $i_A^{-1} = i_A$

3. $(\lambda x \in \{\varnothing, \{\varnothing\}\}.\ \text{if } x = \varnothing \text{ then } 1 \text{ else } 2)^{-1} = \lambda x \in \{1,2\}.\ \text{if } x = 1 \text{ then } \varnothing \text{ else } \{\varnothing\}$

4. $$(\lambda A \in P(E).\chi_A^{(E)})^{-1} = \lambda f \in E \to \{0,1\}.f^{-1}(\{1\})$$
$$= \lambda f \in E \to \{0,1\}.\{x \in E \mid f(x) = 1\}$$

עובדות פשוטות על הפונקציה ההפוכה

טענה:

(1) אם $f : A \to B$ פונקציית שקילות, אז $f^{-1}: B \to A$ גם היא פונקציית שקילות, ו- $(f^{-1})^{-1} = f$.

(2) אם $f : A \to B$ ו- $g : B \to C$ הן פונקציות שקילות, אז כך גם $g \circ f$, ומתקיים:

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

הוכחה

(1) מהשוויונות $f \circ f^{-1} = i_B$ ו- $f^{-1} \circ f = i_A$ נובע, שיש פונקציה $g : A \to B$, כך ש- $f^{-1} \circ g = i_A$, $g \circ f^{-1} = i_B$ ($g = f$). מהמשפט הקודם נובע לכן, ש- f^{-1} היא פונקציית שקילות, וש- f היא הפונקציה ההפוכה לה, דהיינו $(f^{-1})^{-1} = f$.

(2) ברור ש- $f^{-1} \circ g^{-1} : C \to A$. כמו כן:

$$(g \circ f)(f^{-1} \circ g^{-1}) = (g \circ (f \circ f^{-1})) \circ g^{-1} = (g \circ i_B) \circ g^{-1} = g \circ g^{-1} = i_C$$

$$(f^{-1} \circ g^{-1})(g \circ f) = (f^{-1} \circ (g^{-1} \circ g)) \circ f = (f^{-1} \circ i_B) \circ f = f^{-1} \circ f = i_A$$

לכן, לפי המשפט הקודם, $g \circ f$ הפיכה, ו- $f^{-1} \circ g^{-1}$ היא הפונקציה ההפוכה.

מסקנה:

אם $f:A\to A$ היא פונקציית שקילות, אז $(f^{-1})^n=(f^n)^{-1}$.

הוכחה:

$$(f^n)^{-1}=(f\circ f\circ\ldots\circ f)^{-1}\overset{(2)}{=}f^{-1}\circ f^{-1}\circ\ldots\circ f^{-1}=(f^{-1})^n$$

סימון: f^{-n} יסמן את $(f^{-1})^n$ ואת $(f^n)^{-1}$ (השווים לפי המסקנה האחרונה).

בשלב זה מתחיל אולי להתחוור לקוראים פשר הסימון f^{-1} לפחות כאשר $f:A\to A$ ו-f הינה הפיכה, מוביל סימון זה לנכונות כללי החזקות הבאים עבור כל n ו- m *שלמים* (כלומר $n,m\in\mathbf{Z}$):

$$\boxed{\begin{aligned} f^n\circ f^m&=f^{n+m}\\ (f^n)^m&=f^{n\cdot m}\end{aligned}}$$

תרגיל

הוכח זהויות אלו.

אזהרה:

הזהות $(f\circ g)^n=f^n\circ g^n$ ***אינה*** נכונה בדרך-כלל, אפילו אם גם f וגם g הן פונקציות מ- A ל- A (היא נכונה אבל כאשר $f\circ g=g\circ f$).

לסיום סעיף זה נביא מספר הכללות חשובות של עובדות, שראינו קודם. את ההוכחות נשאיר לקוראים.

טענה:

1. אם $f:A\to B$ ו- $g:B\to C$ הן ח.ח.ע., אז $g\circ f$ היא ח.ח.ע..
2. אם $f:A\to B$ היא על B ו- $g:B\to C$ היא על C, אז $g\circ f$ היא על C.
3. אם $f:A\to B$, $g:B\to A$ ו- $g\circ f=i_A$, אז f ח.ח.ע.,ו- g היא על A.

טבלה ב.6 מסכמת את ההגדרות של המושגים החשובים ביותר הקשורים בפונקציות.

טבלה ב.6: פונקציות – מושגים חשובים

(1)		פונקציית ***הזהות*** מעל קבוצה A היא $i_A = \lambda x \in A.\, x$.
(2)		אם $A \subseteq E$, אז הפונקציה ***האופיינית*** של A יחסית ל- E היא: $\chi_A^{(E)} = \lambda x \in E.\ \text{if } x \in A \text{ then } 1 \text{ else } 0$
(3)		אם $f: A \to B$ ו- $X \subseteq A$, אז ***הצמצום*** של f ל- X הוא: $f / X = \lambda x \in X.\, f(x)$ $\quad (f / X : X \to B)$
(4)		אם $f: A \to B$ ו- $g: B \to C$, אז ***ההרכבה*** $g \circ f$ היא: $g \circ f = \lambda x \in A.\, g(f(x))$ $\quad (g \circ f: A \to C)$
(5)	(i)	אם $f: A \to B$ ו- $X \subseteq A$, אז ***התמונה*** של X לפי f היא: $f[X] = \{f(x) \mid x \in A\}$ $\quad (f[X] \subseteq B)$
	(ii)	אם $f: A \to B$ ו- $X \subseteq B$, אז ***המקור*** של X לפי f הוא: $f^{-1}[X] = \{x \in A \mid f(x) \in X\}$ $\quad (f^{-1}[X] \subseteq A)$
(6)	(i)	פונקציה f מעל A נקראת ***חד-חד-ערכית*** (חח"ע) אם: $\forall x \in A\ \forall y \in A.\, f(x) = f(y) \Rightarrow x = y$
	(ii)	פונקציה $f: A \to B$ נקראת ***על*** B אם $B = f[A]$, כלומר: $\forall y \in B\ \exists x \in A.\, y = f(x)$
	(iii)	$f: A \to B$ נקראת ***פונקציית שקילות*** (בין A ל- B) אם היא חח"ע ועל B.
(7)		אם $f: A \to B$ היא פונקציית שקילות, אז ***הפונקציה ההפוכה ל-*** f היא: $f^{-1} = \lambda x \in B.\, \iota y \in A.\, f(y) = x$ $\quad (f^{-1} : B \to A)$

ה. פונקציית Curry ושימושיה

אחת הדרכים המקובלות בשפות מחשב מתקדמות לטיפול בפונקציות של שני משתנים (או יותר), היא לעשות להן רדוקציה לפונקציה של משתנה אחד. דבר זה נעשה על-ידי כך שמספקים את הקלטים באופן סדרתי, זה אחר זה, ולא את כולם בבת-אחת. כאשר מגיע הקלט הראשון, המחשב יוצר בעזרתו פונקציה ***חד***-מקומית. פונקציה חד-מקומית זו מופעלת בהגיע הקלט השני, וכך מתקבל הפלט הסופי. מה שהמחשב עושה, אפוא, הוא לקחת פונקציה דו-מקומית $f: A \times B \to C$ ולעשות לה טרנספורמציה לפונקציה ב- $A \to (B \to C)$, הנקראת f^{curry}. הרעיון עליו מבוססת טרנספורמציה זו הוא פשוט ביותר. לפי כלל (η), אם $f \in A \times B \to C$, אזי

$$f = \lambda x \in A, y \in B.\, f(x,y)$$

מה שהטרנספורמציה עושה, בעיקרו של דבר, הוא להחליף את הפסיק בנקודה:

$$f^{\text{Curry}} = \lambda x \in A.\ \lambda y \in B.\, f(x,y)$$

לדוגמה, אם f היא פונקציית החיבור +, אז:

$$+^{\text{curry}} = \lambda x \in \mathbf{R}.\ \lambda y \in \mathbf{R}.\ x + y$$

לכן

$$+^{\text{curry}}(3) \overset{\beta}{=} \lambda y \in \mathbf{R}.\ 3 + y$$

אנו רואים, אם כך, שבעוד הפונקציה + אינה יכולה לעשות דבר, לפני שיעמדו לרשותה שני הארגומנטים, הפונקציה $+^{\text{curry}}$ עושה משהו מועיל עם קלט אחד בלבד: עם הקלט 3, למשל, היא יוצרת את הפונקציה $\lambda y \in \mathbf{R}.\ 3 + y$. עתה:

$$(+^{\text{curry}}(3))(4) = (\lambda y \in \mathbf{R}.\ 3 + y)\,(4) \overset{\beta}{=} 3 + 4 = 7$$

אם נסכם: בצורה של $+^{\text{curry}}$ פונקציית החיבור מקבלת תחילה את הקלט הראשון (3 בדוגמה) ומחזירה פונקציה ב- $\mathbf{R} \to \mathbf{R}$ (פונקציית החיבור עם 3, בדוגמה שלנו). עם קבלת הקלט השני (4 בדוגמה שלנו) מופעלת פונקציה זו מ- $\mathbf{R} \to \mathbf{R}$ על קלט זה, וכך מתקבלת התוצאה הסופית.

הטרנספורמציה מ- f ל- f^{Curry} היא בעצמה פונקציה, ששמה המלא הוא פונקציית Curry. אנו נסמן אותה, מטעמי חיסכון, ב- Cu. נפרט:

$$Cu : (A \times B \to C) \to (A \to (B \to C))$$

ו- $Cu(f) = f^{\text{Curry}}$ עבור $f \in A \times B \to C$.

או, ביתר פירוט:[17]

$$Cu = \lambda f \in A \times B \to C.\ \lambda x \in A.\ \lambda y \in B.\ f(x,y)$$

הפונקציה Cu היא פונקציית שקילות בין $A \times B \to C$ ו- $A \to (B \to C)$. נוכיח זאת על-ידי שנראה, שיש לה פונקציה הפוכה. פונקציה הפוכה זו ידועה בשם UnCurry, ואנו נסמנה בקיצור ב- U. הגדרת U הינה:

$$U = \lambda g \in A \to (B \to C).\ \lambda x \in A, y \in B.\ (g(x))(y)$$

קל לוודא, שאכן $U : (A \to (B \to C)) \to (A \times B \to C)$. נראה, שהיא הפוכה ל- Cu. דבר זה לא כרוך ביותר מאשר חישוב פשוט בעזרת הכללים α, β ו- η:

(i) אם $h \in A \to (B \to C)$, אז

$$\begin{aligned}
(Cu \circ U)(h) &= Cu(\underbrace{U(h)}) \\
&\overset{\beta}{=} \underbrace{Cu(\lambda x \in A, y \in B.(h(x))(y))} \\
&\overset{\beta}{=} \lambda x \in A.\lambda y \in B.\underbrace{(\lambda x \in A, y \in B.(h(x))(y))(x,y)} \\
&\overset{\beta}{=} \lambda x \in A.\underbrace{\lambda y \in B.(h(x))(y)} \\
&\overset{\eta}{=} \underbrace{\lambda x \in A.h(x)} \\
&\overset{\eta}{=} h
\end{aligned}$$

מכאן ש- $Cu \circ U = i_{A \to (B \to C)}$.

הערה:

השורה השלישית בחישוב זה עלולה להיות מבלבלת. יכולנו, לכן, לשנות לפניה את הגדרת U בעזרת כלל (α) ל:

$$U = \lambda g \in A \to (B \to C).\ \lambda a \in A, b \in B.\ (g(a))(b)$$

17 למעשה, היינו צריכים לכתוב $Cu^{A,B,C}$, כיוון שלכל שלוש קבוצות A,B,C יש את פונקציית Curry שלהן (ממש כשם שלכל קבוצה A יש את פונקציית הזהות שלה, ולכן הסימון i_A). לצורך נוחות הקריאה, אנו משמיטים את האזכור המפורש של A, B ו- C.

ואז היינו מקבלים בשורה השלישית:

$$\lambda x \in A.\lambda y \in B.\underbrace{(\lambda a \in A, b \in B.(h(a))(b))(x,y)}$$

ההמשך היה ללא שינוי.

(ii) אם $h \in A \times B \to C$, אז

$$\begin{aligned}
(U \circ Cu)(h) &= U(\underbrace{Cu(h)}) \\
&\overset{\alpha,\beta}{=} U(\lambda a \in A.\, \lambda b \in B.\, h(a,b)) \\
&\overset{\beta}{=} \lambda x \in A, y \in B.\,(\underbrace{(\lambda a \in A.\, \lambda b \in B.\, h(a,b))(x)})\,(y) \\
&\overset{\beta}{=} \lambda x \in A, y \in B.\,\underbrace{(\lambda b \in B.\, h(x,b))\,(y)} \\
&\overset{\beta}{=} \lambda x \in A, y \in B.\, h(x,y) \\
&\overset{\eta}{=} h
\end{aligned}$$

מכאן ש- $U \circ Cu = i_{A \times B \to C}$.

מ.ש.ל.

לפונקציית Curry יש חשיבות רבה במתמטיקה (הגם שהיא לא ידועה שם בשם זה). כדוגמה נביא את נושא ***הנגזרת***. המושג הבסיסי בנושא נגזרות, ממנו נגזרים כל השאר, הוא המושג של ***נגזרת של פונקציה בנקודה***. זהו ***מספר***, שאינטואיטיבית שווה לשיפוע המשיק של גרף הפונקציה באותה נקודה. עתה, כדי שנוכל לחפש נגזרת של פונקציה בנקודה, עלינו לדעת שני דברים:

(א) באיזו פונקציה מדובר.

(ב) באיזו נקודה מדובר.

נסמן את הפעולה של מציאת נגזרת של פונקציה בנקודה על-ידי D^*. D^* היא פונקציה המצפה לשני קלטים: הראשון – פונקציה חלקית מ- **R** ל- **R**, השני – מספר ממשי. הפלט הוא מספר ממשי. מכאן:

$$D^* : ((\mathbf{R} \overset{p}{\to} \mathbf{R}) \times \mathbf{R}) \overset{p}{\to} \mathbf{R}$$

כאשר D^* מוגדר, כידוע, על-ידי:

$$D^*(f, x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

דהיינו:

$$D^* = \lambda f \in \mathbf{R} \overset{p}{\to} \mathbf{R},\ x \in R.\ \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

בשלב הבא בלימודי חדו"א עוברים לדבר על ***הפונקציה הנגזרת***, f' (או $D(f)$) של פונקציה f. אם f פונקציה חלקית מ- **R** ל- **R**, אז נגזרתה, $D(f)$, גם היא פונקציה חלקית מ- **R** ל- **R**, ומתקיים:

$$D(f) = \lambda x \in \mathbf{R}.\ D^*(f, x)$$

מכאן ש- D היא פונקציה ב- $(\mathbf{R} \overset{p}{\to} \mathbf{R}) \to (\mathbf{R} \overset{p}{\to} \mathbf{R})$, ומתקיים:

$$D = \lambda f \in \mathbf{R} \overset{p}{\to} \mathbf{R}.\ \lambda x \in \mathbf{R}.\ D^*(f, x)$$

אנו רואים אפוא, ש- $D = Cu(D^*)$: הפעולה של גזירת פונקציות אינה אלא התוצאה של הפעלת פונקציית Curry על הפעולה של מציאת נגזרת של פונקציה בנקודה.

דוגמה זו מבליטה באופן מיוחד את יעילותו האפשרית של הרעיון לטפל בקלטים אחד אחד, במקום בבת-אחת. מציאה ישירה של נגזרת של פונקציה בנקודה הינה בעיה לא פשוטה של חישוב גבולות. למשל:

$$D^*(\lambda x.\ x^2, 3) = \lim_{h \to 0} \frac{(3+h)^2 - 9}{h} = \lim_{h \to 0} \frac{6h + h^2}{h} = 6$$

לעומת זאת, בחישובים מעשיים אנו תחילה מטפלים ב- $(\lambda x.\ x^2)$ ומחשבים את נגזרתה, שהיא פונקציה, ***ואז*** מפעילים פונקציה נגזרת זו על 3:

$$D(\lambda x.\ x^2) = \lambda x.\ 2x$$

ולכן

$$D^*(\lambda x.\ x^2, 3) = (D(\lambda x.\ x^2))(3) = (\lambda x.\ 2x)(3) = 2 \cdot 3 = 6$$

%%

שימוש נוסף: נגזרות חלקיות

אחד הכלים החשובים בחקר "פונקציות של שני משתנים" הוא מושג הנגזרת החלקית $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$. הרעיון ב- $\frac{\partial F}{\partial x}$, למשל, הוא שגוזרים את F כאילו היתה פונקציה של x בלבד, כאשר y משמש כפרמטר. ב- $\frac{\partial F}{\partial y}$ מתחלפים, כמובן, התפקידים. לדוגמה:

$$\text{(i)}\quad \frac{\partial x^y}{\partial x} = yx^{y-1} \qquad\qquad \text{(ii)}\quad \frac{\partial x^y}{\partial y} = x^y \ln x$$

מה שכתבנו למעלה הוא צורה סטנדרטית של הצגת נגזרות חלקיות וזהויות הקשורות בהן. נבדוק עתה מה באמת מתרחש כאן.

ממבט ראשון התשובה פשוטה. בחישוב $\frac{\partial x^y}{\partial y}$, למשל, אנו מסתכלים בפונקציה $\lambda x.\, x^y$ (שבה x הוא משתנה, ו- y – פרמטר) וגוזרים אותה. לכן:

$$\frac{\partial x^y}{\partial y} = D(\lambda x.\, x^y)$$

תשובה פשוטה זו היא מטעה! ראשית, אם נתבונן ב"זהות" המתקבלת:

$$D(\lambda y.\, x^y) = x^y \cdot \ln x$$

הרי ברור, שמשהו כאן לא בסדר. $D(\lambda y.\, x^y)$ הוא ביטוי המתאר ***פונקציה*** (שבו x מופיע כפרמטר). $x^y \cdot \ln x$ ***אינו*** ביטוי עבור פונקציה. זהו ביטוי עבור ***מספר***.

מה שנכון הוא ש-

$$D(\lambda y.\, x^y) = \lambda y.\, x^y \ln x$$

נוסחה זו היא אכן מדויקת – אך לא לכך הכוונה בנגזרות חלקיות! נגזרת חלקית של "פונקציה של שני משתנים" אמורה להיות ***בעצמה*** "פונקציה של שני משתנים", לא רק אחד. השימוש בנגזרות חלקיות שניות, כמו $\frac{\partial^2 F}{\partial x \partial y}$, מלמד על כך. מה שהנוסחה (ii) בעמוד הקודם רוצה באמת לבטא, הוא ש:

$$\frac{\partial}{\partial y}(\lambda x, y.\, x^y) = \lambda x, y.\, x^y \ln x$$

זה מחזיר אותנו לשאלה המקורית: מה זה בעצם $\frac{\partial}{\partial y}$?

התשובה היא, שראשית, את מה שקיבלנו על-ידי הפעלת D, $\lambda y.\, x^y \ln x$, אנו רוצים להפוך לפונקציה של שני משתנים. לכך נחוצה אבסטרקציה על x. אנו נרצה לעבור לכן ל- $\lambda x.\, \lambda y.\, x^y \ln x$ על מנת לקשור את x. זה אבל לא מספיק. מה שמתקבל הוא פונקציה חלקית ב- $\mathbf{R} \xrightarrow{p} (\mathbf{R} \xrightarrow{p} \mathbf{R})$, ולא ב- $\mathbf{R} \times \mathbf{R} \xrightarrow{p} \mathbf{R}$. לכן יש עוד להפעיל את U, הפונקציה ההפוכה ל- Curry. ואכן:

$$U(\lambda x.\, D(\lambda y.\, x^y)) = \lambda x,y.\, x^y \ln x$$

כללית,

$$\frac{\partial F}{\partial y} = U(\lambda x.\, D(\lambda y.\, F(x,y))) = U(\lambda x.\, D((Cu(F))(x)))$$

ומכאן

$$\frac{\partial}{\partial y} = \lambda F \in \mathbf{R} \times \mathbf{R} \xrightarrow{p} \mathbf{R}.\, U(\lambda x.\, D(\lambda y.\, F(x, y)))$$

לגבי $\frac{\partial}{\partial x}$, הנוסחה מסובכת טיפה יותר:

$$\frac{\partial}{\partial x} = \lambda F \in \mathbf{R} \times \mathbf{R} \xrightarrow{p} \mathbf{R}.\, U^*(\lambda y.\, D(\lambda x.\, F(x, y)))$$

כאשר

$$U^* = \lambda G \in \mathbf{R} \xrightarrow{p} (\mathbf{R} \xrightarrow{p} \mathbf{R}).\, \lambda x \in R, y \in R.\, G(y, x)$$

כדאי לשים לב, שכיוון שהמשתנים x ו- y בנוסחאות אלו הם קשורים, ניתן להחליפם במשתנים אחרים. אין לכן, למעשה, כל קשר בין הפונקציה $\frac{\partial}{\partial y}$ (למשל) ובין המשתנה y! ב- $\frac{\partial}{\partial y}$ הכוונה בעצם הינה לגזירה "לפי המקום השני".

%%

ב.5. יחסים

I. הגדרת המושג "יחס"

מושג ה"יחס" (relation) הוא אחד המושגים הבסיסיים של המתמטיקה והידע האנושי בכלל.

דוגמאות:

(א) יחס הסדר בין מספרים ממשיים.

(ב) יחס ההתחלקות בין מספרים טבעיים ("a מתחלק ב- b").

(ג) יחס החפיפה בין משולשים.

(ד) היחס "בין" בין נקודות ("הנקודה A נמצאת בין הנקודות B ו- C").

(ה) יחס קרבת המשפחה בין בני האדם.

(ו) יחס ההרשמה בין סטודנטים וקורסים.

כרגיל, בתורת הקבוצות, אנו משתדלים להגדיר כל מושג בעזרת המושגים הבסיסיים של קבוצה ושייכות. בנוגע ליחס, מה שחשוב הוא מי מתייחס למי. כשמדובר ביחסים *בינאריים* (כלומר: דו-מקומיים), ניתן לתאר כל עובדה מסוג זה על-ידי ***זוג סדור***, בו הרכיב הראשון מתייחס לרכיב השני ביחס בו מדובר. זה מוביל להגדרה הבאה:

הגדרה

(1) יחס (בינארי) R הוא קבוצה של זוגות סדורים, דהיינו: $\forall z \in R\ \exists a \exists b.\ z = \langle a,b \rangle$.

(2) קבוצה R היא ***יחס מקבוצה A לקבוצה B***, אם $R \subseteq A \times B$, כלומר:

$$\forall z \in R\ \exists a\ \exists b.\ z = \langle a,b \rangle \wedge a \in A \wedge b \in B$$

(3) יחס מ- A ל- A נקרא גם יחס ***על*** A.

תרגיל

כל יחס R הוא יחס על - $\cup(\cup(R))$.

דוגמאות:

1. $< \; = \{<a,b> \in \mathbf{R}\times\mathbf{R} \mid \exists z \in \mathbf{R}^+.\ b = a + z^2\}$

2. $\backslash \; = \{<a,b> \in \mathbf{N}\times\mathbf{N} \mid \exists k \in \mathbf{N}.\ b = k\cdot a\}$

3. אם T היא קבוצת המשולשים במרחב, אזי:
$\cong \; = \{<a,b> \in T\times T \mid$ b חופף ל- $a\}$

4. אם S קבוצת הסטודנטים ו- Co קבוצת הקורסים, אז יחס ההרשמה בין סטודנטים ובין קורסים הוא:
$\mathrm{Reg} = \{<x,y> \in S\times Co \mid$ y רשום ל- $x\}$

הערות:

1. מקובל מאוד לכתוב aRb במקום $<a,b> \in R$. כך כותבים $a<b$ במקום $<a,b> \in <$, $\Delta ABC \cong \Delta DEF$ במקום $<\Delta ABC, \Delta DEF> \in \cong$ וכו'.

2. רוב היחסים המעניינים במתמטיקה הם בינאריים, אך יש יוצאים מן הכלל. כך, למשל, היחס "בין" בין נקודות (דוגמה (ד) למעלה) הוא יחס ***טרנארי*** (תלת-מקומי). את ההגדרות שנתנו ניתן להכליל בקלות ליחסים כלליים יותר:

יחס n-מקומי הוא קבוצה של n-יות (עצמים מהצורה $<a_1, \ldots, a_n>$).
יחס n-מקומי על $A_1, \ldots, A_n$ הוא קבוצה חלקית של $A_1\times A_2\times\ldots\times A_n$
(בפרט יחס חד-מקומי על קבוצה A הוא פשוט קבוצה חלקית של A).
יחס n-מקומי על A הוא קבוצה חלקית של A^n.

בקורס זה נעסוק בעיקר ביחסים בינאריים.[18]

3. %% הערנו בפרק הקודם, שמושג הפונקציה הכללי ביותר חורג לעתים מהמושג הרשמי. דבר זה נכון ביתר שאת לגבי יחסים באופן כללי. אנו מדברים על יחס השוויון $=$, על יחס ההכלה $\subseteq$ ועל יחס השייכות $\in$, למרות שאי-אפשר לראות בהם קבוצות של זוגות, והבנתם קודמת להבנת מושג "הזוג הסדור" (הרי יהיה זה מטופש לטעון, שהנוסחה "$1 \in \mathrm{N}$" אינה אלא קיצור של "$<1,\mathrm{N}> \in \in$", שהיא בעצמה רק קיצור של "$<<1,\mathrm{N}>,\in> \in \in$", שהיא בעצמה...). לא ניכנס כאן לדיון

[18] תחום יישומי חשוב במדעי המחשב, שבו יש חשיבות רבה ליחסים n-מקומיים כלשהם, הוא "בסיסי נתונים טבלאיים". בבסיסי נתונים כאלה מייצגים יחסים בשיטת הטבלה, שתוארה עבור פונקציות בפרק הקודם.

מעמיק בסוגייה זו. נציין רק, שברוב המכריע של המקרים, המושגים השונים שנגדיר וההגדרות עצמן טובים גם ליחסים כמו $\in$, $=$ ו- $\subseteq$, כלומר ליחסים במובן רחב יותר מזה של ההגדרה הרשמית.

%%

4. אם נבדוק את ההגדרות הרשמיות של פונקציות בפרק הקודם (בטבלה מס' ב.5), ניווכח, שפונקציות הן סוג מיוחד של יחסים. ליתר דיוק:

(i) ***פונקציה*** היא יחס f המקיים את תנאי החד-ערכיות:

$$\forall a \forall b_1 \forall b_2 . <a,b_1> \in f \wedge <a,b_2> \in f \Rightarrow b_1 = b_2$$

(זהו תנאי (1)(ב) בטבלה, בעוד תנאי (1)(א) שם אומר, ש- f יחס).

(ii) ***פונקציה חלקית מ- A ל- B*** היא יחס מ- A ל- B, המקיים את תנאי החד-ערכיות הנ"ל.

(iii) ***פונקציה מ- A ל- B*** היא פונקציה חלקית f מ- A ל- B, המקיימת:

$$\forall a \in A \; \exists b \in B . <a,b> \in f$$

כזכור, כאשר יחס f הוא פונקציה, אנו נוהגים לכתוב $y = f(x)$ במקום $<x,y> \in f$ או xfy

II. פעולות על יחסים

א. היפוך יחס:

אם R יחס, אז היחס ההפוך, R^{-1}, מוגדר על-ידי:

$$R^{-1} = \{<b,a> \mid <a,b> \in R\}$$

ברור ש- R^{-1} אכן יחס כש- R יחס. יתר על כן, אם R הינו יחס מ- A ל- B, אז R^{-1} הינו יחס מ- B ל- A. R^{-1} מתאפיין על-ידי העובדה הבאה:

$$bR^{-1}a \Leftrightarrow aRb$$

דוגמאות:

(1) $<^{-1} = >$

(2) היחס ההפוך ליחס "הורה של" הוא היחס "הילד של", כלומר: x הוא הורה של y אם"ם y הוא ילד של x.

ב. הרכבת יחסים

אם S יחס מ- A ל- B, ו- R יחס מ- B ל- C, או הרכבתם, $R \circ S$, היא היחס הבא מ- A ל- C:

$$R \circ S = \{<a,c> \in A \times C \mid \exists b \in B.\ <a,b> \in S \wedge <b,c> \in R\}$$

במלים אחרות:

$$\forall a \in A\ \forall c \in C.\ a\ R \circ S\ c \Leftrightarrow \exists b \in B.\ aSb \wedge bRc$$

דוגמאות:

(1) ההרכבה של היחס "הורות" עם עצמו נותנת את יחס ה"סבאות".

(2) ההרכבה של היחס "הורה של" עם היחס "אח/אחות של" נותנת את היחס "דוד/דודה של".

שימו לב:

השוואה של ההגדרות, שניתנו פה למושגים של הרכבת יחסים ושל היחס ההפוך, להגדרות של אותם מושגים עבור פונקציות (בטבלה ב.5), מראה שההגדרות ***זהות***. מה שהוגדר כאן אינו אלא הכללה של ההגדרות עבור פונקציות. עם זאת, יש לציין את העובדות הבאות:

א. אם f ו- g הן פונקציות, אז הרכבתן $f \circ g$ אינה סתם יחס, ***אלא היא פונקציה בעצמה***. כאן יש, לכן, התאמה מושלמת בין ראיית $f \circ g$ כהרכבה של פונקציות (כמו בפרק הקודם), ובין ראיית $f \circ g$ כהרכבה של יחסים.

ב. המצב בעניין f^{-1}, לעומת זאת, עדין יותר. היחס ההפוך R^{-1} מוגדר עבור ***כל*** יחס R. מכאן, שאם f פונקציה (ולכן יחס), אז היחס f^{-1} מוגדר אף הוא. ברם, ***יחס זה אינו תמיד פונקציה בעצמו***: תנאי החד-ערכיות יכול שלא להתקיים עבור f^{-1}. רק אם הפונקציה f הינה חד-חד-ערכית, יהיה היחס f^{-1} פונקציה גם הוא. הוא יהווה פונקציה, שהתחום שלה שווה לתמונה של f, והתמונה שלה שווה לתחום של f. היוצא מזה הוא, שאם f פונקציה, אז בעוד ***הנוסחאות*** $<x,y> \in f^{-1}$ (או $xf^{-1}y$) הן בעלות משמעות עבור ***כל*** ערך של x ו- y, הרי ***הביטוי*** $f^{-1}(x)$ הוא בעל משמעות רק בתנאי ש- f הינה פונקציה ח.ח.ע., ו- x שייך לתמונה של f.

התרגיל הבא מראה, שלפעולות המוכללות של הרכבה והפיכה יש תכונות דומות לאלה, שהיו להן במקרה המיוחד של פונקציות:

תרגיל

א. $(R \circ S) \circ T = R \circ (S \circ T)$

ב. $(R^{-1})^{-1} = R$

ג. $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$

ד. אם R יחס מ- A ל- B, אז $i_B \circ R = R \circ i_A = R$.

הנושא של פעולות על יחסים (לא רק בינאריים) הינו בעל אפליקציות רבות (למשל – במערכות בסיסי נתונים), אבל חורג ממה שנזדקק לו בהמשך קורס זה.[19]

III. תכונות של יחסים

בסעיף זה נציג תכונות חשובות, שעשויות להיות ליחסים.

1. (א) יחס R על קבוצה A נקרא ***רפלקסיבי*** (חוזר) על A אם $\forall x \in A.\, xRx$.

(ב) יחס R נקרא ***אי-רפלקסיבי*** אם $\forall x.\, \neg(xRx)$. היחס נקרא אי-רפלקסיבי על A אם $\forall x \in A.\, \neg(xRx)$.

דוגמאות:

(א) יחסים רפלקסיביים: $=$, $\leq$ על $\mathbf{R}$, $\backslash$ על $\mathbf{N}$, יחס הדמיון על קבוצת המשולשים במישור.

(ב) יחסים אי-רפלקסיביים: $\neq$, $<$ על $\mathbf{R}$, יחס ההורות.

(ג) דוגמה ליחס שאינו רפלקסיבי, אך גם אינו אי-רפלקסיבי:

$$S = \{<x,y> \in \mathbf{R}^2 \mid y = x^2\}$$

יחס זה אינו רפלקסיבי (כי $<2,2> \notin S$), אך גם אינו אי-רפלקסיבי (כי $<1,1> \in S$).

[19] חומר נוסף (כולל פעולות, שלא הגדרנו כאן) ניתן למצוא בחלק המוקדש לתורת הקבוצות בקורס "מתמטיקה בדידה" של האוניברסיטה הפתוחה.

הערה:

נשים לב, שרפלקסיביות יחס אינה משהו אבסולוטי, אלא תלויה בקבוצה, עליה רואים את היחס כמוגדר בקונטקסט מסוים. היחס $\{<n,n> \mid n \in \mathbf{N}\}$ הוא רפלקסיבי בתור יחס על $\mathbf{N}$, אך אינו רפלקסיבי בתור יחס על $\mathbf{R}$. ברם, כמעט תמיד, כשנתייחס ליחס R, נציין גם קבוצות A ו- B, שיש באותו קונטקסט לראות את R כיחס מהאחת לשנייה. בדרך-כלל נציין לכן פשוט, ש- R הוא יחס רפלקסיבי או לא-רפלקסיבי (להבדיל מאי-רפלקסיבי!), ויהיה ברור אז מהקונטקסט, יחסית לאיזו קבוצה מדובר.

%%

כדי להימנע מסיבוכים מסוג זה, מכניסים טקסטים רבים את הקבוצות A ו- B לתוך ***ההגדרה*** של יחס. לפי טקסטים אלו, אם R_1 הוא יחס מ- A_1 ל- B_1, ו- R_2 הוא יחס מ- A_2 ל- B_2, הרי די ש- $A_1 \neq A_2$, או ש- $B_1 \neq B_2$, כדי שהיחסים R_1 ו- R_2 ייחשבו כשונים – אפילו אם R_1 ו- R_2 שווים כקבוצות (של זוגות). לגישה זו מחיר משלה. אם ניישם אותה לפונקציות (דבר מקובל מאוד גם הוא), הרי מתברר, שאין פונקציה אקספוננציאלית $\lambda x \in \mathbf{R}.\ e^x$ אחת, אלא הרבה. אנו נאלצים להבדיל, למשל, בין הפונקציה $\mathrm{Exp}_1 = \lambda x.\ e^x\colon \mathbf{R} \to \mathbf{R}$ ובין $\mathrm{Exp}_2 = \lambda x.\ e^x\colon \mathbf{R} \to \mathbf{R}^+$ (כש- $\mathbf{R}^+ =_{df} \{x \in \mathbf{R} \mid x > 0\}$). איש אינו עושה אבל הפרדה כזו, כיוון שהיא היתה הופכת את כתיבתם וקריאתם של טקסטים לבלתי נסבלת! המחיר, שאנו משלמים כאן, הוא זניח לעומת המחיר של הגישה האלטרנטיבית. %%

2. יחס R נקרא ***טרנסיטיבי*** אם:

$$\forall x\ \forall y\ \forall z.\ xRy \wedge yRz \Rightarrow xRz$$

דוגמאות:

$=$, $\leq$, $>$, חפיפה, דמיון, התחלקות וצאצאות הם יחסים טרנסיטיביים (לדוגמה: אם $a \setminus b$, ו- $b \setminus c$, אז $a \setminus c$. אם a צאצא של b, ו- b צאצא של c, אז a צאצא של c).
לעומת זאת, יחס ההורות, יחס האהבה, $\neq$ והיחס S למעלה, אינם טרנסיטיביים.

אזהרה:

כשאנו אומרים, שיחס האהבה, למשל, אינו טרנסיטיבי, ***אין*** כוונתנו ש*בכל* מקרה בו a אוהב את b ו- b אוהב את c, a בהכרח אינו אוהב את c, אלא רק ש- a, b ו- c כאלו ***קיימים***.

3. (א) יחס R נקרא ***סימטרי*** אם $\forall x\ \forall y.\ xRy \Rightarrow yRx$.
   (ב) יחס R נקרא ***אנטי סימטרי במובן החזק*** אם $\forall x\ \forall y.\ xRy \Rightarrow \neg(yRx)$.

(ג) יחס R נקרא ***אנטי סימטרי*** (או אנטי סימטרי במובן החלש) אם

$$\forall x\ \forall y.\ xRy \wedge yRx \Rightarrow x = y$$

דוגמאות:

$=$, $\neq$, חפיפה, יחס הדמיון, יחס הנישואין, היחס $\{<x,y> \in \mathbf{R}^2 \mid x^2 + y^2 = 1\}$ הם יחסים סימטריים. $<$ ויחס ההורות הם אנטי סימטריים במובן החזק. $\leq$, $=$, $\subseteq$, יחס ההתחלקות והיחס S למעלה, הם אנטי סימטריים (כך, למשל, אם $a \leq b$ ו- $b \leq a$, אז $a = b$).

4. יחס R, שהינו רפלקסיבי על A, אנטי סימטרי וטרנסיטיבי, נקרא ***יחס סדר חלקי*** על A אם, בנוסף, מתקיים ש- $\forall x \in A\ \forall y \in A.\ xRy \vee yRx$, אז אומרים ש- R סדר ***מלא*** (או ***טוטאלי***, או ***לינארי***) על A.

דוגמאות:

$=$, $\leq$ על $\mathbf{R}$, $\subseteq$ על $P(E)$ (E קבוצה כלשהי) ויחס ההתחלקות על $\mathbf{N}$ הם יחסי סדר חלקיים. מתוכם $\leq$ הוא יחס סדר מלא על $\mathbf{R}$. כמו כן, היחס:

$$\{<x,y> \in \mathbf{R}^2 \mid x \leq y \wedge \sin \pi x = \sin \pi y\}$$

הינו יחס סדר חלקי על $\mathbf{R}$ (הוכיחי!), שהינו סדר מלא על $\mathbf{N}$, אך לא על $\mathbf{R}$ כולו.

להגדרות הללו יש גרסאות חזקות יותר:

יחס ***סדר חלקי חזק*** (strict) על A הוא יחס טרנסיטיבי ואי-רפלקסיבי על A. יחס כזה נקרא ***מלא*** (או ליניארי, או טוטאלי) על A אם

$$\forall x \in A\ \forall y \in A.\ xRy \vee x = y \vee yRx$$

דוגמאות:

$<$ הוא יחס סדר חזק ומלא על $\mathbf{R}$, בעוד $\subset$ הינו יחס סדר חלקי חזק על $P(E)$, שאינו מלא אם ב- E שני איברים לפחות (אם $a \in E$, $b \in E$, ו- $a \neq b$, אז $\{a\} \neq \{b\}$, $\{a\} \not\subset \{b\}$, ו- $\{b\} \not\subset \{a\}$).

תרגיל

(1) אם R יחס סדר חלקי על A, אז היחס $R^* = \{<x,y> \in A^2 \mid xRy \wedge x \neq y\}$ הינו יחס סדר חלקי חזק על A. אם, בנוסף, R הוא מלא על A, אז גם R^* מלא על A.

(2) אם R יחס סדר חלקי חזק על A, אז היחס $R' = \{<x,y> \in A^2 \mid xRy \vee x = y\}$ הוא יחס סדר חלקי על A. אם, בנוסף, R הינו מלא על A, אז כך גם R'.

5. יחס על קבוצה A, שהוא רפלקסיבי על A, סימטרי וטרנסיטיבי, נקרא יחס *שקילות* (אקויולנציה) על A.

דוגמאות: $=$, דמיון משולשים, חפיפת משולשים.

דוגמאות נוספות:

(א) $\equiv(2) = \{<a,b> \in \mathbf{N} \times \mathbf{N} \mid 2 \setminus |a-b|\}$.

למשל: $<11,5> \in \equiv(2)$ ו- $<5,11> \in \equiv(2)$, כיוון ש- $|11-5|$ ו- $|5-11|$ מתחלקים ב- 2. לעומת זאת, $<4,7> \notin \equiv(2)$, כיוון ש- $|4-7|$ אינו מתחלק ב- 2.

מקובל לסמן $a \equiv b (\mathrm{mod}\ 2)$ במקום $a \equiv (2)\ b$ או $<a,b> \in \equiv(2)$.

הערה: באופן דומה מגדירים את היחס $\equiv(n)$ לכל מספר טבעי n.

(ב) $Q^* = \{<<a,b>, <c,d>> \in (\mathbf{N}^+ \times \mathbf{N}^+)^2 \mid ad = bc\}$.

תרגיל

להוכיח ש- $\equiv(2)$ (ו- $\equiv(n)$ באופן כללי) הוא יחס שקילות על $\mathbf{N}$, בעוד Q^* הוא יחס שקילות על $\mathbf{N}^+ \times \mathbf{N}^+$.

התכונות החשובות השונות, שהגדרנו, מסוכמות בטבלה ב.7 בעמוד הבא. אנו מניחים שם, ש- R יחס על A, ולכן התוספת "על A" הושמטה מהמונחים השונים.

IV. יחסי שקילות ופירוקים, קבוצות מנה

הגדרה

תהי A קבוצה. *פירוק*[20] (partition) של A היא קבוצה F של קבוצות חלקיות לא ריקות של A (כלומר $F \subseteq P(A) - \{\varnothing\}$), המקיימת:

(i) $\cup(F) = A$

(ii) $\forall X \in F\ \forall Y \in F.\ X \neq Y \Rightarrow X \cap Y = \varnothing$

---

[20] יש המשתמשים במונח "חלוקה" במקום "פירוק".

טבלה ב.7: תכונות חשובות של יחסים

		יהי R יחס על A $(R \subseteq A^2)$. ההגדרות בטבלה הן יחסית ל- A.
.1	(א)	R נקרא ***רפלקסיבי***, אם: $\forall a \in A.\ aRa$
	(ב)	R נקרא ***אי-רפלקסיבי***, אם: $\forall a \in A.\ \neg(aRa)$
.2	(א)	R נקרא ***סימטרי***, אם: $\forall a \in A\ \forall\ b \in A.\ aRb \Rightarrow bRa$
	(ב)	R נקרא ***אנטי סימטרי***, אם: $\forall a \in A\ \forall\ b \in A.\ (aRb \wedge bRa) \Rightarrow a = b$
	(ג)	R נקרא ***אנטי סימטרי חזק***, אם: $\forall a \in A\ \forall\ b \in A.\ aRb \Rightarrow\ \neg(bRa)$
.3		R נקרא ***טרנסיטיבי***, אם: $\forall a \in A\ \forall b \in A\ \forall c \in A.\ (aRb \wedge bRc) \Rightarrow aRc$
.4	(א)	R נקרא ***יחס סדר חלקי***, אם R רפלקסיבי, אנטי-סימטרי וטרנסיטיבי.
	(ב)	R נקרא ***יחס סדר מלא***, אם R יחס סדר חלקי, ובנוסף מתקיים: $\forall a \in A\ \forall b \in A.\ aRb \vee bRa$
	(ג)	R נקרא ***יחס סדר חלקי חזק***, אם R אי-רפלקסיבי וטרנסיטיבי.
	(ד)	R נקרא ***יחס סדר מלא חזק***, אם R יחס סדר חלקי חזק, ובנוסף: $\forall a \in A\ \forall b \in A.\ aRb \vee a = b \vee bRa$
.5		R נקרא ***יחס שקילות*** (או יחס ***אקויולנציה***), אם R הינו רפלקסיבי, סימטרי וטרנסיטיבי.

הערות:

1. תנאי (ii) שקול ל:

$$\forall X \in F \, \forall Y \in F. \, X \cap Y \neq \varnothing \Rightarrow X = Y$$

זוהי, בדרך-כלל, צורה נוחה יותר להוכחה.

2. משמעות התנאים היא, ש- F הינו פירוק של A אם"ם כל איבר ב- A שייך לאיבר אחד בדיוק של F (שייך – בגלל (i), לאחד בלבד – בגלל (ii)). בטרמינולוגיה אחרת: F הוא פירוק של A אם $A = \cup(F)$ ו- $X \neq \varnothing$ לכל $X \in F$.

דוגמה

$\{\{1,3,5\},\{6\},\{2,4,7,8\}\}$ היא פירוק של $\{1,2,3,4,5,6,7,8\}$.

סימונים:

1. את קבוצת הפירוקים של קבוצה A נסמן ב- $\mathrm{par}(A)$.
2. את קבוצת יחסי השקילות על קבוצה A נסמן ב- $\mathrm{Eq}(A)$.

תרגילים

1. $\mathrm{par}(A) \in P(P(P(A)))$
2. $\mathrm{Eq}(A) \in P(P(A \times A))$

נראה עתה, שלמושג הפירוק קשר הדוק למושג של יחס שקילות:

הגדרה:

1. יהי R יחס שקילות על קבוצה A, ונניח $x \in A$. ***מחלקת השקילות של x*** לפי R היא הקבוצה:

$$[x]_R = \{y \in A \mid xRy\}$$

2. יהי R יחס שקילות על A. ***קבוצת המנה של A לפי R*** היא הקבוצה:

$$A/R = \{[x]_R \mid x \in A\}$$

במלים אחרות: A/R היא קבוצת מחלקות השקילות לפי R של איברי A.

טענה

יהי R יחס שקילות על A. אז לכל x ו- y ב- A מתקיים:

(0) $x \in [x]_R$

(1) $y \in [x]_R \Leftrightarrow xRy$

(2) $y \in [x]_R \Leftrightarrow yRx$

(3) $y \in [x]_R \Leftrightarrow x \in [y]_R$

(4) $[y]_R \subseteq [x]_R \Leftrightarrow xRy$

(5) $[x]_R = [y]_R \Leftrightarrow xRy$

(6) אם $Z \in A/R$ ו- $x \in Z$, אז $Z = [x]_R$.

הוכחה

(0) מיידי מהגדרת $[x]_R$ והעובדה, ש- R רפלקסיבי.

(1) זוהי בעצם הגדרת $[x]_R$.

(2) מיידי מ- (1) ומסימטריות R.

(3) מיידי מ- (2) ומהגדרת $[y]_R$: $x \in [y]_R \overset{(1)}{\Leftrightarrow} yRx \overset{(2)}{\Leftrightarrow} y \in [x]_R$.

(4) ($\Rightarrow$) . נניח $[y]_R \subseteq [x]_R$. כיוון ש- $y \in [y]_R$ (לפי (0)), נובע מזה, ש- $y \in [x]_R$, ולכן, לפי (1), xRy.

($\Leftarrow$). נניח xRy. יהי $z \in [y]_R$, אז yRz (מ- (1)). מזה ומ- xRy נובע xRz, בגלל ש- R טרנסיטיבי. לכן $z \in [x]_R$ לפי (1). סה"כ $[y]_R \subseteq [x]_R$.

(5) ($\Rightarrow$). אם $[x]_R = [y]_R$, אז $[y]_R \subseteq [x]_R$, ולכן xRy לפי (4).

($\Leftarrow$). נניח xRy. אז גם yRx (כי R סימטרי). לכן מ- (4) נובע, ש- $[y]_R \subseteq [x]_R$ וגם ש- $[x]_R \subseteq [y]_R$. סה"כ $[x]_R = [y]_R$.

(6) נניח $Z \in A/R$. אז $Z = [y]_R$ לאיזה $y \in A$. לכן, אם $x \in Z$, אז $x \in [y]_R$. מכאן, לפי (2), xRy. לכן, לפי (5), $[x]_R = [y]_R$. כלומר: $[x]_R = Z$.

מסקנות:

1. אם R יחס שקילות ו- $x \in A$, אז $x \in Z$ $\iota Z \in A/R.$ $= [x]_R$ (מ- (0) ו- (6)).
2. $\lambda x \in A.\, [x]_R$ היא פונקציה מ- A **על** A/R.

משפט

אם R יחס שקילות על A, אז A/R היא פירוק של A.

הוכחה:

כיוון ש- $[x]_R \subseteq A$ לכל $x \in A$, A/R היא קבוצה של קבוצות חלקיות של A. אף אחת מהן אינה ריקה, כי $x \in [x]_R$ לכל $x \in A$ (סעיף (0) בטענה האחרונה). מהעובדה, ש- $\forall x \in A.\, x \in [x]_R$ נובע גם, שכל איבר x ב- A שייך לאחת הקבוצות ב- A/R (הלא היא $[x]_R$). לכן $A \subseteq \cup(A/R)$. מצד שני, נובע מהעובדה ש- $X \subseteq A$ לכל $X \in A/R$, ש- $\cup(A/R) \subseteq A$ (טבלה ב.2, כלל $(5)_b$). סה"כ $\cup(A/R) = A$. נראה, לבסוף, שמדובר באיחוד זר. נניח, אפוא, ש- $X_1 \in A/R$, $X_2 \in A/R$ ו- $X_1 \cap X_2 \neq \varnothing$. יהי $x \in X_1 \cap X_2$. מהסעיף השישי בטענה האחרונה: $X_1 = [x]_R$ ו- $X_2 = [x]_R$. לכן $X_1 = X_2$.

דוגמאות

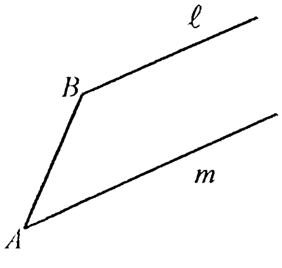
1. ב- $\mathbf{N}/\equiv(2)$ יש שתי מחלקות שקילות:

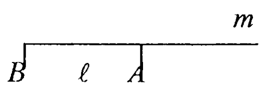
קבוצת המספרים הזוגיים: $N_{\text{even}} = \{2n \mid n \in \mathbf{N}\}$

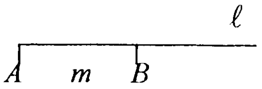
קבוצת המספרים האי-זוגיים: $N_{\text{odd}} = \{2n+1 \mid n \in \mathbf{N}\}$

מתקיים ש- $\mathbf{N}_{\text{even}} = [0]_{\equiv(2)} = [2]_{\equiv(2)} = [4]_{\equiv(2)} = \ldots$, $\mathbf{N}_{\text{odd}} = [1]_{\equiv(2)} = [3]_{\equiv(2)} = \ldots$

2. נגדיר על קבוצת הקרניים במישור את היחס D הבא: $\ell D m$ אם: הקרן ℓ מוכלת בקרן m, או הקרן m מוכלת בקרן ℓ, או ℓ ו- m מקבילות ונמצאות באותו צד של הישר, המחבר את נקודת ההתחלה שלהן:







קל לברר, שזהו יחס שקילות על קבוצת הקרניים במישור. למחלקת השקילות $[\ell]_D$ של הקרן ℓ נוהגים לקרוא במקרה זה ***הכיוון*** של ℓ. היחס היסודי בין עצמים ומחלקות השקילות שלהן (סעיף (5) בטענה למעלה):

$$[\ell]_D = [m]_D \Leftrightarrow \ell D m$$

מְתרגם כאן לעובדה, שקרן ℓ וקרן m נמצאות ביחס D אם ורק אם יש להן ***אותו כיוון***. מחלקת השקילות של קרן ℓ היא קבוצת כל הקרניים, שיש להן אותו כיוון כמו ℓ (גם אינטואיטיבית וגם לפי ההגדרה הרשמית כאן).

דוגמה 2 מדגימה תהליך נפוץ למדי של יצירת ישויות (או עצמים) חדשים על-ידי מעבר מעצמים בקבוצה מסוימת למחלקות השקילות שלהן, לפי יחס שקילות מסוים. העניין לא מסתיים אבל בדרך-כלל בכך. על העצמים החדשים מגדירים בדרך-כלל פעולות ויחסים (כדי שיהיה להם שימוש) משלהם. הדבר נעשה כמעט תמיד על-ידי שימוש ב***מייצגים***. לצורך עניין זה, כל איבר של מחלקת שקילות נחשב ***מייצג*** של אותה מחלקת שקילות (לפי סעיף (6) של הטענה למעלה, x הוא מייצג של Z אם"ם $Z = [x]_R$). ניקח, לדוגמה, את המושג של ***זווית בין כיוונים***. היא נמדדת, כידוע, כך: לוקחים קרן כלשהי בכיוון האחד וקרן כלשהי בכיוון השני. הזווית בין הקרניים היא הזווית בין הכיוונים. כדי שהגדרה זו תהיה בעלת משמעות, הכרחי אבל, שהתוצאה הסופית ***לא תהיה תלויה בקרניים, שבחרנו כמייצגות של הכיוונים*** (תנאי זה נקרא "אי-תלות במייצגים", ועוד נחזור אליו רבות בהקשרים אחרים). בדוגמה של כיוונים קל להוכיח, שאכן כך הדבר.

3. **%%** נחזור ליחס Q^*, שהבאנו כדוגמה ליחס שקילות על $(\mathbf{N}^+)^2$:

$$<a,b> Q^* <c,d> \;\Leftrightarrow\; ad = bc$$

מהי, למשל, מחלקת השקילות של $<3,5>$? ובכן, קל לברר, ש-

$$[<3,5>]_{Q^*} = \{<3,5>, <6,10>, <9,15>, \ldots\} = \{<3k, 5k> \mid k \in \mathbf{N}^+\}$$

עתה, במקום לכתוב $[<3,5>]_{Q^*}$, אנו כותבים, עוד מאז ימי בית-הספר היסודי, $\frac{3}{5}$.

כללית, אנו מסמנים:

$$\frac{a}{b} =_{Df} [< a, b >]_{Q^*}$$

למחלקות שקילות אלו אנו קוראים "מספרים רציונליים חיוביים". נשים לב, שאכן השוויון

$$\frac{a}{b} = \frac{c}{d}$$

פירושו:

$$[<a,b>]_{Q^*} = [<c,d>]_{Q^*}$$

וזה קורה אם"ם $<a,b> Q^* <c,d>$ (לפי האיפיון המרכזי בטענה למעלה).

כלומר אם"ם: $ad = bc$.

את קבוצת המספרים הרציונליים החיוביים אנו מסמנים ב- Q^+. Q^+ היא, אם כך, קבוצת מחלקות השקילות של היחס Q^*, דהיינו $(\mathbf{N}^+ \times \mathbf{N}^+) / Q^*$.

הנקודה המרכזית מכל זה היא, שקבוצת המספרים הרציונליים החיוביים נבנית כקבוצת מנה של יחס שקילות מסוים על $\mathbf{N}^+ \times \mathbf{N}^+$, ומספר רציונלי הוא, בעצם, קבוצה אינסופית של זוגות. כל זוג כזה יכול לשמש כמייצג של המספר הרציונלי. כך, למשל, הזוג $\langle 6,10 \rangle$ יכול לשמש כמייצג של המספר הרציונלי $\frac{3}{5}$ (כי $\frac{6}{10} = \frac{3}{5}$. המדובר פה, נשים לב, בשוויון של ***קבוצות***!).

כאמור למעלה, לעתים קרובות, כשמדובר בקבוצת מנה, מעוניינים אנו להגדיר עליה פעולות ויחסים, הדומים לאלה של קבוצת המקור. כך, לדוגמה, מעוניינים אנו להגדיר חיבור של מספרים רציונליים, הדומה בתכונותיו לחיבור של מספרים טבעיים ומכליל אותו. הדרך הרגילה לעשות דברים כאלו היא על-ידי שימוש במייצגים. עם זאת, יש תמיד להיזהר ולהראות, שמה שהגדרנו אינו תלוי במייצגים, בהם אנו משתמשים. כדי להבין למה הכוונה, נדמה בנפשנו מישהו המגדיר "פעולה" חדשה $\oplus$ על Q^+ על-ידי:

$$\frac{a}{b} \oplus \frac{c}{d} =_{Df} \frac{a+c}{b+d}$$

הגדרה כזו יש להבין כך: כדי לבצע $\oplus$ על-ידי שני איברים X ו- Y ב- Q^+, עלינו למצוא מייצג $\langle a,b \rangle$ עבור X (דהיינו $X = \frac{a}{b}$) ומייצג $\langle c,d \rangle$ עבור Y (ולכן $Y = \frac{c}{d}$), ולחשב ב- $\mathbf{N}^+$ את $a + c$ ואת $b + d$. תוצאת החיבור $\oplus$ של X ו- Y היא אז $\frac{a+c}{b+d}$ – מחלקת השקילות של $\langle a + c,\ b + d \rangle$. הצרה בהגדרה זו הינה, שבחירות שונות של המייצגים של X ו- Y נותנות תשובות שונות לשאלה, כמה שווה $X \oplus Y$. לדוגמה, למרות ש- $\frac{1}{2} = \frac{2}{4}$, הרי $\frac{1}{2} \oplus \frac{1}{3} \neq \frac{2}{4} \oplus \frac{1}{3}$! במקרים כגון אלו נהוג לומר, ש- $\oplus$ ***אינה מוגדרת היטב***.

לעומת זאת, אם נגדיר חיבור על Q^+ באופן הבא:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

אזי נקבל פעולה מוגדרת היטב, כיוון שאם $\langle a_1, b_1 \rangle$ מייצג אחר של $\frac{a}{b}$ (כלומר:

אם $<a_1, b_1> Q^* <a,b>$) ו-$<c_1, d_1>$ מייצג אחד של $\frac{c}{d}$ (כלומר: $<c_1, d_1> Q^* <c,d>$),

אזי

$$<a_1 d_1 + b_1 c_1, b_1 d_1> Q^* <ad + bc, bd>$$

וממילא:

$$\frac{ad + bc}{bd} = \frac{a_1 d_1 + b_1 c_1}{b_1 d_1}$$

במלים אחרות: החלפת המייצג $<a,b>$ של $\frac{a}{b}$ ב- $<a_1, b_1>$ והמייצג $<c,d>$ של $\frac{c}{d}$ ב- $<c_1, d_1>$, לא תשנה את תוצאת הפעולה.

%%

*נספח:
עוד על הקשר בין יחסי שקילות ופירוקים

ראינו בפרק האחרון, שאם R יחס שקילות, אז A/R הינה פירוק של A. מכאן, שכל יחס שקילות על A מגדיר באופן טבעי פירוק של A. נראה עתה, ראשית, שגם ההיפך נכון:

הגדרה

יהי F פירוק של A. נגדיר:

$$A/F = \{<x,y> \in A^2 \mid \exists Z \in F.\ x \in Z \wedge y \in Z\}$$

(כלומר $x(A/F)y$ אם"ם x ו- y נמצאים באותו איבר של F).

טענה

אם F פירוק של A, אז A/F יחס שקילות על A.

הוכחה

(א) רפלקסיביות: מהגדרת A/F, לכל $x \in A$ מתקיים:

$$<x,x> \in A/F \Leftrightarrow \exists Z \in F.\ x \in Z \wedge x \in Z$$
$$\Leftrightarrow \exists Z \in F.\ x \in Z$$

וזה נכון, כיוון ש- F פירוק.

(ב) **סימטריות:** לכל $x,y \in A$:

$$\begin{aligned}<x,y> \in A/F &\Leftrightarrow \exists Z \in F.\ x \in Z \wedge y \in Z\\ &\Leftrightarrow \exists Z \in F.\ y \in Z \wedge x \in Z\\ &\Leftrightarrow <y,x> \in A/F\end{aligned}$$

(ג) **טרנסיטיביות:** נניח $<x,y> \in A/F$ ו- $<y,z> \in A/F$. אז קיים $Z_1 \in F$ כך ש- $x \in Z_1 \wedge y \in Z_1$ וקיים $Z_2 \in F$ כך ש- $y \in Z_2 \wedge z \in Z_2$. מכאן ש- $y \in Z_1 \cap Z_2$. זה גורר (כיוון ש- F פירוק), ש- $Z_1 = Z_2$. מכאן ש- $x \in Z_1 \wedge z \in Z_1$, ולכן קיים $Z \in F$ $(Z = Z_1)$, כך ש- $x \in Z \wedge z \in Z$. במלים אחרות: $<x,z> \in A/F$.

סה"כ קיבלנו, שלכל יחס שקילות על A מתאים פירוק של A, ולכל פירוק של A מתאים יחס שקילות על A. מתעוררת עתה שאלה טבעית: נניח, שאנו מתחילים ביחס שקילות על A. ליחס זה מתאים פירוק של A. לפירוק זה, מצדו, מתאים יחס שקילות על A. האם זה יחס השקילות בו התחלנו? שאלה דומה מתעוררת, כמובן, אם מתחילים בפירוק F. התשובה לשתי השאלות הינה חיובית:

משפט

1. $A/(A/R) = R$ (כאשר R יחס שקילות על A).
2. $A/(A/F) = F$ (כאשר F פירוק של A).

הוכחה

1. נניח ש- R הוא יחס שקילות על A. אז A/R פירוק של A. אם נציב $F = A/R$ בהגדרת A/F, נקבל:

$$A/(A/R) = \{<x,y> \in A^2 \mid \exists Z \in A/R.\ x \in Z \wedge y \in Z\}$$

$$= \{<x,y> \in A^2 \mid \exists Z \in A/R.\ Z = [x]_R \wedge Z = [y]_R\}$$

(לפי סעיף (6) בטענה המרכזית בעמוד 149 על A/R ו- $[x]_R$)

$$= \{<x,y> \in A^2 \mid [x]_R = [y]_R\}$$

(לפי עקרונות לוגיים ברורים הקשורים בשוויון)

$$= \{<x,y> \in A^2 \mid <x,y> \in R\}$$

(לפי סעיף (5) של אותה טענה)

$$= \{<x,y> \mid <x,y> \in R\}$$

(כי $R \subseteq A^2$)

$$\overset{(\text{ח})}{=} R$$

2. נוכיח תחילה את *הלמה* הבאה: אם F_1 ו- F_2 פירוקים של A ו- $A/F_1 = A/F_2$, אז $F_1 = F_2$.

הוכחת הלמה:

נוכיח, למשל, ש- $F_1 \subseteq F_2$ (ההוכחה ש- $F_2 \subseteq F_1$ היא דומה). יהי אפוא $U \in F_1$. כיוון ש- U לא ריק, יש איבר x ב- U. כיוון ש- F_2 פירוק, יש $W \in F_2$ כך ש- $x \in W$. נראה ש- $U = W$, ומזה ינבע ש- $U \in F_2$. בשביל זה מספיק שנוכיח, ש- $U \subseteq W$ ו- $W \subseteq U$. נוכיח, למשל, ש- $U \subseteq W$ (ההוכחה ש- $W \subseteq U$ דומה). יהי אפוא $y \in U$, נראה ש- $y \in W$. כיוון ש- $x \in U$, $y \in U$ ו- $U \in F_1$, נובע, מהגדרת A/F_1, ש- $<x,y> \in A/F_1$. כיוון ש- $A/F_1 = A/F_2$, נקבל ש- $<x,y> \in A/F_2$. מהגדרת A/F_2, פירוש הדבר, שקיים $Z \in F_2$ כך ש- $x \in Z$ ו- $y \in Z$. אבל x גם שייך ל- W, ולכן $x \in W \cap Z$. כיוון ש- F_2 פירוק ו- $W,Z \in F_2$, נובע מזה, ש- $Z = W$. כיוון ש- $y \in Z$, פירוש הדבר, ש- $y \in W$. הראינו, שאם $y \in U$, אז $y \in W$ דהיינו: $U \subseteq W$. באותו אופן מוכיחים ש- $W \subseteq U$. סה"כ $W = U$.

סיום הוכחת 2:

אם F פירוק של A, אז מהמשפט הקודם, A/F יחס שקילות על A. לכן, מחלק 1 של משפט זה נובע ש:

$$A/(A/(A/F)) = A/F$$

משוויון זה ומהלמה (כש- $F_1 = A/(A/F)$ ו- $F_2 = F$) נובע

$$A/(A/F) = F$$

מ.ש.ל.

את המשפט האחרון נוכל להציג מנקודת ראות של פונקציות. נזכיר, ש- $\mathrm{Eq}(A)$ היא קבוצת יחסי השקילות על A ו- $\mathrm{par}(A)$ היא קבוצת הפירוקים של A.

נוכל עתה להתבונן בפונקציות הבאות:

$$H = \lambda R \in \mathrm{Eq}(A).\ A/R \qquad (H : \mathrm{Eq}(A) \to \mathrm{par}(A))$$

$$G = \lambda F \in \mathrm{par}(A).\ A/F \qquad (H : \mathrm{par}(A) \to \mathrm{Eq}(A))$$

המשפט האחרון אומר, בעצם, ש- $H \circ G = i_{\mathrm{par}(A)}$ ו- $G \circ H = i_{\mathrm{Eq}(A)}$.

מסקנה:

$\lambda R \in \mathrm{Eq}(A).\ A/R$ היא פונקציית שקילות בין $\mathrm{Eq}(A)$ ו- $\mathrm{par}(A)$, בעוד $\lambda F \in \mathrm{par}(A).\ A/F$ היא הפונקציה ההפוכה שלה.