8.3.09

לוגיקה למדמ"ח - שיעור 1

פרטים כלליים

המרצה: ארנון אברון
קבלה: יום ד' סוף-פ (ואפשר לתפוס אותו בזמנים אחרים).
יש תרגילים. אין חובת הגשה (ולא נכנס לציון).
באתר הקורס יש סיכומים שעברו אישור של ארנון.
חדר המרצה: שרייבר 111
הדבר שהכי מעצבן אותו: טלפונים מצלצלים בשיעור.

ארנון אומר שהקורס יתבסס על למדות השם שיצא לו.

נקודות שבכל זאת יכולות להיות בעייתיות:
- קורס מאוד מדוייק (פורמלי)
- בקורס עיקר עיסוקינו הוא בשפות, אך דברים שאנו אומרים עליהם נאמרים בך שפה מסויימת ולעתים יש תלות (לדוגמא בשימוש באותם סימנים) לתיאור השפה ולתיאור השפה שמתארת שפות.

דוגמא: = אז לעתים נסמן אותו בצורה שונה (כגון ≈) עבור הקשר מסוי"ם.

נשתמש בטרמינולוגיה: שפה ומטא-שפה.
הסיבה שזה קשה כי אנו לא אוהבים להשתמש בהרבה סימנים.

קצת היסטוריה

מתחילה באריסטו (כמו רוב הדעות המדוייקים) שכתב על 5 נושאים. ספרי הלוגיקה שלו התקיימו הכי הרבה זמן לעומת ספריו האחרים (כמו לדוגמא בגיאולוגיה). הוא התעסק בסילוגיזם, הסקות - מהצורה:
כל אדם הוא בן תמותה, יהודי הוא אדם ⇐ יהודי הוא בן תמותה.

על הנושא הזה נקדיש מעט זמן. פריצת הדרך הבאה היא
במאה ה-19, בעבודה של בול (מבוליאניות) שרצה לעשות "ימ"
אלגברי בהיסקים לוגיים.
פותחה הלוגיקה מסדר ראשון, עם לוגיקות... נמצאו פרדוקסים
שגרמו לשבר במתמטיקה.

לכן הילברט רצה להוכיח "אפשר לסמוך" על הלוגיקה הקלאסית, וזה
הצריך להגדיר בדיוק מה זה הוכחה, וזה הוביל לפיתוח תורת
החישוביות (עד לפני המצאת המחשב).

<u>מה הפאם שהקורס הזה אמור לתת לנו?</u>

- ביטוי יבש דקלרטיבי בצורה פורמלית שהמחשב מסוגל להבין
(מה שאנחנו בעצם עושים כאשר אנחנו ~~מ~~ מתכנתים).
~~אם~~ זה נוטה לקשר בין קלט התכנית לפלט, להוכיח
נכונות של תוכניות וכו'.
יש מערכות יבש פשוטות, מסדי נתונים פשוטים ולדוגמא שפת
כמו SQL מבוססות על לוגיקה מסדר ראשון. ~~יש~~ מערכות
יותר מתוחכמות מסוגלות להמליץ על תרופה לחולה. ברמה הכי
יותר גבוהה זה מחשב מסוגל לעשות מתמטיקה במקומנו.

- מודל חישובי:
אנו מכירים מודלים חישוביים כמו מכונות טיורינג וכו'.
נכיר מודלים נוספים.

- עוד מגוון דברים אחרים שלא נוטה להכנס אליהם עכשיו...

~~ה~~
הרכיבים העקריים של מערכת לוגית.

I) שפה פורמלית: ~~[illegible]~~ שפה - מחרוזת של סימנים. פורמלית -
בדור של מחרוזת הסימנים בשפה או לא.
(תחביר מאוד מדויק)

השפה מורכבת מ:

1) א"ב. (בדר"כ סופי, אבל אנו הכי הרבה בן מנייה), נסמנו ב-L.

2) קטגוריות סינטקטיות (תחביריות) [כמו שבעברית יש שם תואר, פועל, שם מספר וכו' - בעברית בסה"כ 9 קטגוריות. וזה רק קטגוריות של מילים, יש גם קטגוריות של משפטים].

אחת ~~~~ מהקטגוריות תהיה קטגוריה של נוסחאות, נסמן אותה ב-ο (או-מיקרון). בשפת מסדר ראשון יהיו לנו לדוגמא שמות עצם. כמתים עצמם גם דוגמא לקטגוריה שונה היות.

[הבעיה המרכזית של הלוגיקה: איזה עובדות נובעות מקבוצה של הנחות]

II) יחס נביעה (consequence relation) - נסמנו ב- $\vdash$

יחס נביעה זה הוא יחס בין $\underbrace{\text{קבוצות של נוסחאות}}_{\text{משמאל ליחס}}$ ל$\underbrace{\text{נוסחאות}}_{\text{מימין ליחס}}$ המקיים את הדברים הבאים:

1) [קוראים לו לפעמים "רפלקסיביות", אבל זה לא אומר שהוא רפלקסיבי]:

$A\in T \Leftarrow T\vdash A$. (אם A הוא חלק מקבוצת הנחות אז הוא נובע ממנו).

2) "מונוטוניות":

אם $T\vdash A$ ו- $T\subseteq S$ אז $S\vdash A$. (אם אני מוסיף עובדות אני עדיין יכול להסיק דברים שיכולתי לפני ההוספה).

3) "טרנזיטיביות":

אם $T\vdash A$ ו- $T\cup\{A\}\vdash B$ אז $T\vdash B$. (אפשר להסתמך על משפטים שהוכחנו כבר וכאילו ניתן להתייחס אליהם כהנחות).

איך מגדירים יחס נביעה. יש בגדול 2 דרכים:

- הגדרה הסמנטית
- הגדרה הסינטקטית

אנו נשתמש בשניהם (~~?~~ לא דיון עליהם עדיין).

הגישור בין הלוגיקה הפורמלית ובין שימושים בה נעשה ע"י תהליך שקוראים לו <u>הצרנה</u> (פורמליזציה). (כאשר מחליטים מה משמעות הסימנים)

דוגמא: נניח שרצף הסימנים $q \vee \neg q$ הוא נוסחה בשפה שלנו. ~~?~~ כנ"ל עבור p.

ונניח ~~?~~ שנתון עבור יחס הנביעה: $\{p, q \vee \neg q\} \vdash q$

(הוא לא נותן את {} אבל נזכור שהם ~~?~~ שם). ☆ קיפצו לרגע לעמוד הבא

האם הצרנה היא נכונה או לא היא שאלה מתמטית (כמו האם תרגום של "המלך" הוא נכון). זה לא אומר שלא נשאל על זה במבחן, אבל לא ניתן למחשב לעשות את זה, אלא אנחנו נבדוק בעצמנו האם ההצרנה היא מה שהתכוונו.

<u>תחשיב הפסוקים הקלאסי</u>

יש 2 סיבות למה אנו מקדישים את החלק הראשון של הקורס לנושא זה, למרות שזו שפה חלשה:

1. ~~?~~ פדגוגית - את הרעיונות נרצה להדגים בשפה פשוטה
2. נלמד גם מעט על לוגיקות לא-קלאסיות, ועיקר ההבדלים הם ברמה של תחשיב הפסוקים.

נסמן את השפה הזו ב- L_{PC}. נגדיר אותה:

(נשנה קצת את הדוגמא לרגע, נגיד: $\{\neg p, p \vee q\} \vdash q$)

אם כך נקבל ש: p - צ'צ' תהיה ראש הממלכה.

q - גיגי יהיה ראש הממלכה.

הסימנים שלנו מקבלים משמעות, ונקבל ~~...~~ (בהגות ש צ'צ' או גיגי יהיו ראש ממלכה וצ'צ' לא תהיה) שגיגי יהיה ראש הממלכה.

I. השפה:

1) א"ב: - פסוקים אטומיים / משתנים אטומיים: $p_0, p_1, p_2, \ldots$

- קשרים: [illegible] $\wedge, \vee, \rightarrow, \neg$

- סוגריים: ימיני- "(", שמאלי- ")".

עד כאן יש לנו ב הגדרות דקדוקיות (תחביריות). (*)

והגדרה המרכזית, הנוסחאות!

- נוסחאות: מוגדרות באופן רקורסיבי כך:

(1) p פסוק אטומי הוא נוסחה (או פסוק, בתחשיב הפסוקים הן מילים נרדפות).

(2) אם φ ו- ψ ~~הן~~ נוסחאות אז גם:

(*) $\neg\varphi$, $(\varphi\vee\psi)$, $(\varphi\wedge\psi)$, $(\varphi\rightarrow\psi)$

הן גם נוסחאות.

בצורה מסוימת (מאד) שתשמש כסימון:

$$\varphi \equiv p \mid \neg\varphi \mid (\varphi\vee\varphi) \mid (\varphi\wedge\varphi) \mid (\varphi\rightarrow\varphi)$$

סימון:- נשתמש ב- p,q,r בתור משתנים עבור נוסחאות אטומיות.

- נשתמש ב- φ,ψ,A,B בתור משתנים עבור נוסחאות.

נשים לב ש: φ,ψ למשל הם משתנים במטא-שפה, ואילו $p_0, p_1, p_2, \ldots$ הם משתנים בשפה!

כנ"ל עבור p,q,r הם לא משתנים בשפה, הם משתנים במטא-שפה עבור $\underbrace{p_0, p_1, \ldots}_{\underbrace{\text{המשתנים של השפה}}_{\text{הפנימיים}}}$ (משתנים עבור משתנים)

למה שמנו סוגריים ב-(*)? כדי שנוכל ~~להבין~~ לקרוא φ נוסחה באופן יחיד

(שים לב שעבור ¬ לא שמנו. יכולנו, אבל גם בלי סוגריים נבין להבחין באופן יחיד). דוגמא:

$\neg p \to q \lor r$

$(\neg p) \to (q \lor r)$ $\qquad$ $(\neg p \to q) \lor r$

לא ניתן הגדרה בדיוק מה זה קריאה באופן יחידה, אבל פחות או יותר שנוכל לבנות תחבירי עץ איך הפסוק נבנה (ושהוא יהיה יחיד)

ההוכחה לקריאה יחידה היא ארוכה ומסובכת

בסיכום: "מרוב סוגריים לא רואים את היער".

לכן נעשה מספר קיצורים: (לא תמיד נשתמש בהם)

- סוגריים החיצוניים תמיד נוריד
- נגדיר קדימויות ואז במקרים מסויימים לא נצטרך סוגריים

והיא:

$\neg$

$\lor, \land$

$\to$

וכעת את הפסוק: $\neg p \to q \lor r$ ניתן להבין רק כך:

$((\neg p) \to (q \lor r))$

- $A \to B \to C$ שמשמעותו $A \to (B \to C)$

יש גרסה בה אין סוגריים בכלל! הדבר נקרא השיטה הפולנית, בה כותבים את הקשר בהתחלה, כלומר במקום $A \lor B$ נרשום $\lor A B$ (זה כבר מתקרב לשפת כמו LISP, סקים). בשיטה הפולנית אין סוגריים בכלל וכל דבר ניתן לקריאה בצורה יחידה!

החסרון בשיטה הוא שלאדם ~~קצת~~ קשה לקרוא את הנוסחאות.

מוסר השכל: יש הרבה דרכים להציג את הנוסחאות (בכתיב המקוצר).

<u>טענה</u> (שנביא כדי להראות דוגמא איך להוכיח דבר על שפה)
בכל נוסחה, מספר הסוגריים הימניים שווה למספר הסוגריים השמאליים.

<u>הוכחה</u>
מוכח הכל נעשה בדר"כ באינדוקציה (כמו ברוב למה כי
הגדרת הנוסחאות נעשתה בצורה רקורסיבית).
מספר דרכים, נראה שתי הוכחות.

<u>הוכחה א'</u>
אינדוקציה על n - מספר הסימנים בנוסחה. (אינדוקציה רגילה לא
עובדת כאן ולכן נשתמש באינדוקציה שלמה).
בסיס: $n=1$ (אין לנו נוסחאות עם 0 סימנים!)
במקרה זה φ חייבת להיות פסוק אטומי ולכן אין בה סוגריים
בכלל, ובפרט מספר הסוגריים השמאליים והימניים שווים ל-0
(ולכן שווים).
נניח שהטענה נכונה לכל φ שמספר הסימנים בה $\leq n$, ונניח שמספר
הסימנים ב-φ הוא $n+1$. כעת יש לבצע חלוקה למקרים:
(φ לא יכולה להיות כאן פסוק אטומי)
א) $\varphi = \neg\psi$ אז מספר הסוגריים הימניים של φ הוא כמו של ψ וגם
מספר הסוגריים השמאליים, אבל מס' הסימנים ב-ψ הוא n
ולכן ניתן להפעיל עליו את הנחת האינדוקציה, ומכאן נקבל
את הטענה על φ.

שיפור כתיבת ההוכחה!
נסמן ב-$l(\varphi)$ את מספר הסוגריים השמאליים ב-φ
" " $r(\varphi)$ " " הימניים "

" $s(\varphi)$ " , " הסימנים " ב-φ.

כעת ננסה את א' שוב:

א) $\varphi = \neg\psi$, במקרה זה $s(\varphi)=s(\psi)+1$ ולכן $s(\psi)=n$,

$$(i) \quad l(\varphi) = l(\psi)$$
$$(ii) \quad r(\varphi) = r(\psi)$$

כיוון ש: $s(\psi)=n$ הרי לה"נ לפי הנחת האינדוקציה:

$$(iii) \quad l(\psi) = r(\psi)$$

ולכן מ-(i)-(iii) נובע כי $l(\varphi)=r(\varphi)$.

ב) $\varphi=(\psi_1 \vee \psi_2)$, במקרה זה $s(\psi_1)<n$, $s(\psi_2)<n$.

$$l(\varphi) = 1 + l(\psi_1) + l(\psi_2)$$
$$r(\varphi) = r(\psi_1) + r(\psi_2) + 1$$

להנ"ח עבור ψ_1, ψ_2:

$$l(\psi_1) = r(\psi_1), \quad l(\psi_2) = r(\psi_2)$$

ומשם נקבל: $l(\varphi)=r(\varphi)$

נשארו עוד 2 מקרים. אפשר פשוט להעתיק ממקרה ב', אבל בגלל ש"אין לי זמן ואני עצלן", יש לנו 2 אפשרויות:

[1 ג) $\varphi=(\psi_1 \wedge \psi_2)$, דומה למקרה ב'

ד) $\varphi=(\psi_1 \rightarrow \psi_2)$, דומה למקרה ב'

כשעושים את זה נראה לעין שהמקרים באמת דומים...

[2 נשנה את ב' ל:

ב) $\varphi=(\psi_1 * \psi_2)$, כש: $* \in \{\vee, \wedge, \rightarrow\}$

זה היה סימון אחד להוכחה ~~[מחוק]~~ שהשתמש באינדוקציה כמו שעשינו בתיכון. בדר"כ לא נעשה את זה כך. בדר"כ נשתמש באינדוקציה מבנית. (שתסתמך על המבנה של ההגדרה). ההבדל הוא פחות או יותר בצורת הכתיבה שבשונה השנייה תהיה יותר נוחה.

את הנוסחה הגדרנו בצורה רקורסיבית: (תזכורת)

בסיס - כל נוסחה אטומית היא נוסחה

הצעדים האינדוקטיביים: (1) אם φ נוסחה אז גם $\neg\varphi$ נוסחה

(2) אם φ ו- ψ נוסחאות אז גם $(\varphi\vee\psi)$ נוסחה

(3) ______ , ______ $(\wedge)$ "

(4) ______ , ______ $(\to)$ "

<u>הוכחה ב'</u>

באינדוקציה מבנית:

1. אם φ פסוק אטומי אז $l(\varphi)=0=r(\varphi)$
2. נניח נכונות הטענה ל- ψ_1 ו- ψ_2.

אם $\varphi=(\psi_1 * \psi_2)$ (כאשר $* \in \{\vee,\wedge,\to\}$) אז:

$$l(\varphi)=1+l(\psi_1)+l(\psi_2)=r(\psi_1)+r(\psi_2)+1=r(\varphi)$$

כנ"ל

כאן חסכנו בלתת את המדד של מספר הסימנים.

כעת, איך נראה ש $(\neg p_1\to(p_3\vee p_{17}))$ נוסחה בנויה כהלכה?

נוכל להראות איך אנו בונים אותה:

$$p_1,\ \neg p_1,\ p_3,\ p_{17},\ (p_3\vee p_{17}),\ (\neg p_1\to(p_3\vee p_{17}))$$

כל שלב ~~[מחוק]~~ נוכל שהדבר הוא נוסחה, לפי זה שאלו לפניו הן נוסחאות ושימוש בכללים מההגדרה

(טומר אחד ל פסוק)

מדד נוסף שתמיד נוח להוכיח בשימוש באינדוקציה מתיכון במקום באינדוקציה מבנית הוא אורך סדרת הבנייה (בדוגמא הקודמת 6).

יש השתמשים בעץ בנייה במקום סדרה:

$$\begin{tikzcd}[column sep=small]
p_1 \arrow[d, dash] & & p_3 \arrow[dr, dash] & & p_{17} \arrow[dl, dash] \\
\neg p_1 \arrow[dr, dash] & & & (p_3 \vee p_{17}) \arrow[dll, dash] & \\
& (\neg p_1 \to (p_3 \vee p_{17})) & & &
\end{tikzcd}$$

(וכאן נצטרך לכן עוד מדדים כמו גובה העץ).

הרעיון של המדדים זה אם אנו מתקשים להוכיח את הטענה כטענה של המספרים הטבעיים ואז צריך מדד, שהוא מספר טבעי שניתן לכל נוסחא. (כי לרוב אנו מרגישים יותר בטוחים עם זה)
תמיד נוח למצוא מדד ולהוכיח לגביו ~~[illegible]~~ אך זה יוצא יותר מסורבל מהוכחה באינדוקציה מבנית.

* נשים לב שההוכחות שלנו לא היו ממש פורמליות (השתמשו בעברית ובמילים כמו "אם" "אז", והעברית היא לא שפה פורמלית). אבל קשה להתחמק מזה - גם בהגדרת מתי $A \wedge B$ נכון אנו משתמשים במילה העברית "וגם".

<u>יחס הנביעה של תחשיב הפסוקים הקלאסי</u>

(נזכר שאמרנו שיש 2 דרכים לשפות כאלה, סמנטית וסינטקטית)
באופן טב', הגדרה סמנטית של יחס-נביעה מתבססת על המושג "מודל": מה שקובע הוא שיש אוסף של דברים שנקראים מבנים

נוסף יחס $\models$ שנקרא "יחס הסיפוק" בין נוסחאות למבנים.

כאשר הנוסחא φ מסתפקת, $M \models \varphi$ אומרים ש-"M מספקת את φ" או ש-"φ מסתפקת ע"י M" או ש-"M הינו מודל של φ".

אח"כ יש לנו מודל של קבוצת פסוקים / תורה T שמוגדר להיות המבנה מודל של כל אחד מאיברי T. (פורמלית: $(M \models T \iff \forall \varphi \in T.\ M \models \varphi)$)

ולבסוף מגדירים $T \models \varphi$ אם כל מודל של T הוא גם מודל של φ.

↑ זה הוא יחס הנביעה!

תרגיל: להוכיח שיחס הנביעה שהגדרנו הוא באמת כזה! (מקיים את התנאים מההגדרה של יחס נביעה).

ט הפסקה האחרונה היא מאוד טל"ית (לדוגמא, בכלל לא הגדרנו מה זה מודל), אבל היא נועדה ~~לה~~ להדגים את הרעיון ונבין תכלס מה זה מהדוגמאות

שאלה בכיתה: "אתה יכול לחזור על מה זה מודל?"

תשובה: "לא, כי לא הגדרתי את זה אפילו פעם אחת..."

במקרה של תחשיב הפסוקים "המבנים" הם פונקציות מקבוצת הפסוקים האטומיים אל $\{t,f\}$. פונ' כאלה נקראות "השמות".

יחס הסיפוק כאן (בתחשיב הפסוקים), שנסמנו $\models_{CPL}$, מוגדר בצורה הבאה:

$v \models p$ אם"ם $v(p)=t$

$v \models A \vee B$ אם"ם v מספק את A או v מספק את B

$v \models A \wedge B$ אם"ם v מספק את A וגם v מספק את B

$v \models A \to B$ אם"ם v אינו מספק את A או v מספק את B.

$v \models \neg A$ אם"ם v אינו מספק את A.

כשאנו מגדירים כך אנו מניחים שאנו יודעים מה המשמעות של או, וגם, אינו.

ניתן לראות כאן איך הכנסנו משמעות לסימנים בשפה (בהגדרת $v \models A$ ניתן לראות שיש קשר בין $\neg$ לבין המילה "לא" בעברית).

דרך אחרת להגדיר אותו דבר עצמו:

לכל אחד מהקשרים אנו מתאימים פונקציה:

$$g_\neg : \{t,f\} \to \{t,f\}$$

$$g_\vee, g_\wedge, g_\to : \{t,f\}^2 \to \{t,f\}$$

קראנו להם גם לוחות של אמת (כי הן ~~...~~ מקבוצה סופית לסופית ונוח לתאר בטבלה)

עוד דרך יותר קומפקטית ~~...~~ לתאור מטריצות:

$A \to B$	B: t	B: f
A: t	t	f
A: f	t	t

נרחיב את v השמה v לפונקציה שגם אותה נסמן ב-V (כמו שאמרנו בהתחלה, יכול לגרום לבלבול, אבל זה מאוד נוח) באופן הבא: (בהגדרה אינדוקטיבית)

$$V(\neg A) = g_\neg(V(A))$$

$$V(A \vee B) = g_\vee(V(A), V(B))$$

$$V(A \wedge B) = \ldots$$

$$V(A \to B) = \ldots$$

(ובמקרה עבור p [illegible]: $V(p) = v(p)$, החדשה / הישנה)

$(T \models_{sem} \varphi)$

כעת, אומרים ש-v היא מודל של φ אם $V(\varphi) = t$

אם v מודל של T הוא גם מודל של φ כלומר אם כל השמה שנותנת ערך t לכל הפסוקים ב-T נותנת ערך t גם ל-φ.