תיאוריית תכנון סכמות
למסדי נתונים יחסיים
חלק 1
Design Theory for
Relational Databases
Part 1

הקדמה

Introduction

המטרה

- לדעת מתי סכמה של יחס אינה טובה
  - תלוי באילוצים
  - הסוג הנפוץ ביותר של אילוצים הוא תלויות פונקציונליות (הכללה של מושג המפתח)
- אם סכמה אינה טובה, צריך למצוא דרך להחליף אותה בסכמה או סכמות אחרות שהן טובות יותר

מתי סכמה אינה טובה?

- דוגמה:

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- המידע לגבי ראש מחלקה מופיע הרבה פעמים – פעם אחת עבור כל סטודנט השייך למחלקה

מדוע שכפול מידע אינו טוב?

- ראשית, בזבוז של מקום בזיכרון
- שנית, כאשר אותו מידע מופיע במספר מקומות במסד הנתונים, צריכה להיות מודעות לכך בזמן ביצוע עדכונים
  - חייבים לעדכן את המידע בכל המקומות בהם הוא מופיע

דוגמה לעדכון שגוי

מעדכנים את ראש המחלקה של CS רק ברשומה אחת, וכתוצאה מכך יש נתונים סותרים לגבי ראש המחלקה של CS

S	D	H
Levy	CS	Katz
Cohen	Math	Bush
Barak	CS	Rubin

- כאשר מידע משוכפל במספר מקומות במסד הנתונים, הניסיון מראה שלאורך זמן יש סיכוי שהמידע יעודכן רק בחלק מהמקומות, וכתוצאה מכך יהיו במסד הנתונים פריטי מידע הסותרים זה את זה

מדוע נוצר מצב של שכפול מידע?

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- למחלקה יש רק מנהל אחד אבל יש הרבה סטודנטים
- לכן המידע על מנהל המחלקה נשמר עבור כל סטודנט השייך למחלקה

7

תלות פונקציונלית

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- האילוץ "למחלקה יש רק מנהל אחד" נקרא *תלות פונקציונלית* (ת"פ) ומסומן ע"י $D \rightarrow H$

8

מתי ת"פ מתקיימת ביחס?

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- יחס r
- הת"פ $D \rightarrow H$ מתקיימת ביחס r אם כל שתי רשומות ששוות בעמודה D שוות גם בעמודה H

9

דוגמה של ת"פ שאינה מתקיימת ביחס נתון

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Rubin
Barak	CS	Rubin

- יחס r'
- הת"פ $D \rightarrow H$ *אינה* מתקיימת ביחס r', כי שתי הרשומות הראשונות שוות בעמודה H אבל שונות בעמודה D
- הת"פ $D \rightarrow H$ מתקיימת בשני היחסים: r ו- r'

10

כשמתקיימות ת"פ מסוימות אז מתקיימות גם ת"פ אחרות

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- היחס r
- מקיים את הת"פ $D \rightarrow H$
- ומקיים את הת"פ $S \rightarrow D$
- לפיכך, היחס r מקיים גם את הת"פ $S \rightarrow H$

11

אם מוסיפים רשומה ליחס r...

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Levy	Math	Bush
Barak	CS	Rubin

- היחס r_1
- מקיים את הת"פ $D \rightarrow H$
- אבל אינו מקיים את הת"פ $S \rightarrow D$ (בגלל הרשומה הראשונה והשלישית)
- היחס r_1 גם אינו מקיים את הת"פ $S \rightarrow H$

12

אבל...

Student
Department
Head

- היחס r_2

S	D	H
Levy	CS	Rubin
Cohen	Math	Rubin
Levy	Math	Rubin
Barak	CS	Rubin

- מקיים את הת"פ $D \rightarrow H$
- ואינו מקיים את הת"פ $S \rightarrow D$
- אבל היחס r_2 *כן* מקיים את הת"פ $S \rightarrow H$

תהליך תכנון הסכמה כולל

- קביעת העמודות של היחס, לדוגמה {S,D,H}
- קביעת הת"פ שהיחס אמור לקיים
  - אם לכל מחלקה מנהל יחיד, אז הת"פ $D \rightarrow H$ אמורה להתקיים ביחס
  - אם לכל סטודנט מחלקה יחידה, אז הת"פ $S \rightarrow D$ אמורה להתקיים ביחס
  - אבל אם סטודנט יכול להיות שייך למספר מחלקות, אז הת"פ $S \rightarrow D$ אינה אמורה להתקיים ביחס

לאחר שקובעים את הת"פ

- צריך להחליט (בדרך אלגוריתמית) מהן הת"פ הנוספות שמתקיימות בכל היחסים האפשריים עבור הסכמה
- צריך להחליט האם קיום (או אי קיום) של ת"פ מסוימות גורם לשכפול מידע (ולכן הסכמה אינה טובה)
- אם יש ת"פ שגורמות לשכפול מידע, אז צריך למצוא דרך להחליף את הסכמה המקורית בסכמה או סכמות אחרות שבהן אין אפשרות לשכפול מידע

אפיון האפשרות לשכפול מידע בעזרת ת"פ (דוגמה)

Student
Department
Head

S	D	H
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- למחלקה יש רק מנהל אחד אבל יש הרבה סטודנטים
- כלומר, הת"פ $D \rightarrow H$ מתקיימת, אבל הת"פ $D \rightarrow S$ אינה מתקיימת

ת"פ בין קבוצות אטריביוטים

- לעיתים מספר עמודות קובעות ערכים של מספר עמודות אחרות

דוגמה

Student
Course
Grade

- היחס

S	C	G
Levy	OS	90
Cohen	DB	82
Levy	DB	92
Levy	Logic	87

- מקיים את הת"פ $SC \rightarrow G$
  - אינו מקיים את הת"פ $S \rightarrow G$
  - ואינו מקיים את הת"פ $C \rightarrow G$

פורמלית, את הת"פ $SC \rightarrow G$ היה צריך לכתוב $\{S,C\} \rightarrow G$

דוגמה (המשך)

- היחס

S	C	G
Levy	OS	90
Cohen	DB	82
Levy	DB	92
Levy	Logic	87

Student
Course
Grade

- גם מקיים את הת"פ
  SC → GC

פורמלית, את הת"פ
SC → GC היה צריך
לכתוב {S,C} → {G,C}

עתה נעבור להגדרות פורמליות של המושגים שהוצגו בדוגמאות

סימונים

- אותיות מסוף האלף-בית, כלומר $X,Y,Z,\ldots$, מסמנות קבוצות של אטריביוטים
- אותיות מתחילת האלף-בית, כלומר $A,B,C,\ldots$, מסמנות אטריביוטים בודדים
- איחוד של שתי קבוצות X ו-Y מסומן ע"י XY
- קבוצת אטריביוטים $\{A,B,C,D\}$ מסומנת ע"י $ABCD$

תלות פונקציונאלית (ת"פ)

- תלות פונקציונאלית (ת"פ) מסומנת ע"י $X \rightarrow Y$ ופירושה ש-X *קובע* את Y, כלומר בהינתן ערכים עבור האטריביוטים של X, הערכים עבור האטריביוטים של Y נקבעים באופן אחד ויחיד
- ת"פ הנה *אילוץ (constraint)* על היחסים האפשריים, כלומר אם קובעים שת"פ צריכה להתקיים, אז פירושו של דבר שמגבילים את היחסים האפשריים רק לאלה שמקיימים את הת"פ

דוגמאות לת"פ

- הסימון $Y \rightarrow X$ מציין ת"פ בצורתה הכללית ביותר, כלומר קבוצת אטריביוטים X קובעת קבוצת אטריביוטים Y
- הסימון $A \rightarrow X$ מציין ת"פ בצורה קצת פחות כללית, כלומר קבוצת אטריביוטים X קובעת אטריביוט בודד A
- דוגמאות לת"פ
  - SD → H (הקבוצה SD קובעת את האטריביוט H)
  - AB → CDG (הקב' AB קובעת את הקב' CDG)
  - SD → HDS (הקב' SD קובעת את הקב' HDS)

סימונים נוספים

- רשומות מסומנות ע"י r,s,t,u,v
- בהינתן קבוצת אטריביוטים X ורשומה t, ההטלה של t על X מסומנת ע"י t[X]
  - כלומר $t[X]=\pi_X(\{t\})$
- בהינתן אטריביוט בודד A, הערך של רשומה t עבור A מסומן ע"י t[A], כלומר $t[A]= \pi_A(\{t\})$
- t[X] גם נקרא *ערך ה-* X של t

שוויונות ואי-שוויונות בין רשומות

- השוויון t[X] = r[X] פירושו שהרשומות r ו- t שוות על כל האטריביוטים של X
- אי-השוויון $r[X] \neq t[X]$ פירושו שקיים ב- X לפחות אטריביוט אחד A, כך שלרשומות r ו- t ערכים שונים עבור A, כלומר $r[A] \neq t[A]$

סכמות ויחסים

- סכמה של יחס הנה קבוצה R של אטריביוטים
  - במסד נתונים אין שני יחסים שונים עם אותה קבוצה של אטריביוטים (קרי, עם אותה סכמה)
  - קבוצת האטריביוטים R היא גם השם של הסכמה
- יחס r עבור סכמה R הנו קבוצה סופית של רשומות, שעמודותיהן מתאימות לאטריביוטים של R
- סכמה הנה קבועה, בעוד שהיחס עבורה משתנה עם הזמן, כתוצאה ממחיקה, הוספה ושינוי של רשומות

הגדרה: יחס *מקיים* (או *מספק*) ת"פ

- תהי R סכמה ויהיו X ו-Y קבוצות של אטריביוטים המוכלות ב-R, כלומר $X, Y \subseteq R$
- יחס r מעל העמודות של R *מקיים* ת"פ $X \rightarrow Y$ אם לכל שתי רשומות s ו-t, ששייכות ל-r, מתקיים:
  - אם $s[X] = t[X]$, אז $s[Y] = t[Y]$
    - כלומר אם שתי הרשומות שוות על כל עמודה של X, אז הן שוות על כל עמודה של Y

מתי יחס *אינו מקיים* ת"פ?

- כדי להראות שיחס r עבור R *אינו מקיים* ת"פ $X \rightarrow Y$, צריך למצוא ביחס r שתי רשומות s ו- t, ששוות על X ושונות על Y, כלומר שתי רשומות s ו- t שמקיימות:
  - $s[X] = t[X]$ וגם
  - $s[Y] \neq t[Y]$
  - אי השוויון הנ"ל פירושו שיש אי שוויון לפחות על אטריביוט אחד של Y

דוגמה

- היחס

A	B	C	D	E	F
1	1	1	1	1	1
1	1	1	0	0	0

- מקיים את הת"פ AC → BA
- מקיים את הת"פ ABD → CFE
- ואינו מקיים את הת"פ AC → BF

ת"פ הן חלק מהסכמה, כלומר הן קובעות מהם היחסים האפשריים

- נתונה סכמה R עם קבוצה F של תלויות פונקציונליות (הכוללות רק אטריביוטים של R)
- יחס r הנו *אפשרי* עבור R ו- F אם העמודות של r מתאימות ל- R והיחס r מקיים את כל הת"פ של F
- $SAT(R,F)$ מסמן את קבוצת היחסים האפשריים עבור הסכמה R וקבוצת הת"פ F

סיכום עד כה

- קבענו איך רושמים ת"פ
  - $X \rightarrow Y$
- הגדרנו מתי יחס נתון מקיים ת"פ $X \rightarrow Y$

תלויות פונקציונליות טריוויאליות

Trivial Functional Dependencies

ת"פ טריוויאלית

- ת"פ הנה *טריוויאלית* אם היא מתקיימת בכל יחס (שמכיל את האטריביוטים שלה)
- *טענה:* ת"פ $X \rightarrow Y$ הנה טריוויאלית אם ורק אם $Y \subseteq X$
- דוגמאות לת"פ טריוויאליות
  - $D \rightarrow D$
  - $SD \rightarrow S$

הוכחה הטענה – אם

- נניח ש- $Y \subseteq X$ ויהי r יחס שעמודותיו כוללות את כל האטריביוטים של $X \rightarrow Y$
- צריך להראות ש-r מקיים את $X \rightarrow Y$, כלומר
  - תהיינה s ו-t שתי רשומות של r, כך ש- $s[X] = t[X]$
  - צריך להראות ש- $s[Y] = t[Y]$ וברור ששווין זה מתקיים מכיוון ש- $Y \subseteq X$

הוכחה – רק אם

- צ"ל: ת"פ $X \rightarrow Y$ הנה טריוויאלית רק אם $Y \subseteq X$
- נניח ש- $Y \not\subset X$
- נראה שיש יחס r שאינו מקיים את $X \rightarrow Y$
- יהי A אטריביוט כך ש- $A \in Y - X$
- יהי r יחס שבו שתי רשומות s ו-t, ששוות על כל האטריביוטים של X ושונות על A
  - ברור ש-r אינו מקיים את $X \rightarrow Y$

היחס r מההוכחה של "רק אם"

שאר העמודות	A		X	
לא חשוב מה הערכים בעמודות אלה	1	1	1	1
	0	1	1	1

$A \in Y - X$

- היחס r אינו מקיים את הת"פ $X \rightarrow Y$ מכיוון ש-A הוא אטריביוט ששיך ל-Y אבל לא ל-X

גרירה של תלויות פונקציונליות

Implication of
Functional Dependencies

37

מוטיבציה למושג הגרירה

- אם יחס r מקיים קבוצה F של ת"פ, אזי הוא חייב לקיים גם ת"פ נוספות
- לדוגמה, אם יחס מקיים את הת"פ $A \to B$ ו- $B \to C$ אז הוא מקיים גם את $A \to C$

38

הגדרה: גרירה של ת"פ

- קבוצה F של ת"פ *גוררת* ת"פ $X \to Y$, סימון $F \models X \to Y$, אם כל יחס r שמקיים את F מקיים גם את $X \to Y$

39

כשמתקיימות ת"פ מסוימות אז מתקיימות גם ת"פ אחרות

Student	Department	Head
Levy	CS	Rubin
Cohen	Math	Bush
Barak	CS	Rubin

- כל יחס שמקיים את הת"פ $S \to D$ ו- $D \to H$
- חייב לקיים גם את הת"פ $S \to H$
- ולכן $S \to D, D \to H \models S \to H$

40

דוגמה נגדית לגרירה

- כדי להראות שקבוצה F איננה גוררת ת"פ $X \to Y$ צריך למצוא *דוגמה נגדית*, כלומר יחס r שמקיים את F ואינו מקיים את $X \to Y$
- נראה ש- $A \to B, B \to C \not\models C \to A$ ע"י הדוגמה הנגדית

A	B	C
1	1	1
0	0	1

41

עוד דוגמה של דוגמה נגדית

- נראה ש- $A \to B, C \to B \not\models C \to A$ היחס הימני אינו דוגמה נגדית (מדוע?), אבל היחס השמאלי הנו דוגמה נגדית

A	B	C
1	1	1
0	1	1

A	B	C
1	1	1
0	0	1

42

גזירה של תלויות פונקציונליות

Derivations of Functional Dependencies

43

מוטיבציה למושג הגזירה

- המטרה היא לדעת האם קבוצת ת"פ F גוררת את הת"פ $X \to Y$
- צריך לבדוק אם יש דוגמה נגדית
  - אבל יש אינסוף יחסיים אפשריים, שכל אחד מהם יכול להיות דוגמה נגדית
- גזירה היא תהליך סינטקטי שמאפשר למצוא בזמן סופי את כל הת"פ ש-F גוררת
- תהליך הגזירה מבוסס על האקסיומות של ארמסטרונג

44

האקסיומות של ארמסטרונג

- *רפלקסיביות:* אם $Y \subseteq X$ אז $X \to Y$
- *אוגמנטציה (הוספה):*
  אם $X \to Y$ אז $XZ \to YZ$
- *טרנזיטיביות:*
  אם $X \to Y$ ו- $Y \to Z$ אז $X \to Z$

45

הערה

- רפלקסיביות היא *אקסיומה*, כלומר מציינת מהן כל הת"פ הטריוויאליות (קרי, שמתקיימות בכל יחס)
- אוגמנטציה וטרנזיטיביות הן *כללי היסק*, כלומר מציינות איך ניתן להסיק ת"פ נוספות מת"פ נתונות
- לשלושתן נהוג לקרוא אקסיומות

46

גזירה של ת"פ

- בהינתן קבוצה F של ת"פ, ניתן *לגזור* מהן ת"פ נוספות בעזרת האקסיומות, לדוגמה תהי:
  $F = \{ A \to C, B \to D \}$
- להלן גזירה של $AB \to ABCD$ מ-F

1.	$A \to C$	(נתון ב-F)
2.	$AB \to ABC$	(אוגמנטציה של 1 עם AB)
3.	$B \to D$	(נתון ב-F)
4.	$ABC \to ABCD$	(אוגמנטציה של 3 עם ABC)
5.	$AB \to ABCD$	(טרנזיטיביות של 2 ו- 4)

47

הגדרה של גזירה

- גזירה של ת"פ $X \to Y$ מקבוצה F של ת"פ היא סדרה סופית של ת"פ, שמקיימת את שני הדברים הבאים:
  - כל ת"פ בסדרה שייכת ל- F, נובעת מרפלקסיביות או נובעת מת"פ קודמות בסדרה, לפי אוגמנטציה או טרנזיטיביות
  - בנוסף, הסדרה מסתיימת בת"פ $X \to Y$

48

גזירה לעומת גרירה

- הסימון $F \vdash X \to Y$ מציין שקיימת גזירה של הת"פ $X \to Y$ מקבוצת הת"פ F
- הסימון $F \models X \to Y$ מציין ש- F גוררת את $X \to Y$, כלומר אין דוגמה נגדית
- גזירה נועדה כדי לגלות האם יש גרירה, לפיכך נרצה להראות שהמושגים של גרירה וגזירה שקולים, כלומר

$$F \models X \to Y \Leftrightarrow F \vdash X \to Y$$

נאותות של גזירות

Soundness of Derivations

נאותות של גזירות

- נאותות פירושה שאם יש גזירה אז יש גם גרירה, כלומר

$$F \models X \to Y \Leftarrow F \vdash X \to Y$$

נאותות של האקסיומות

- כדי להראות נאותות של גזירות, צריך להראות תחילה שכל אקסיומה היא נאותה
- הצורה הכללית של אקסיומה:
  - אם נתונות ת"פ מסוימות ומתקיימים תנאים מסוימים, אז אפשר לגזור ת"פ חדשה
  - הערה: לפעמים נתונות רק ת"פ ואין תנאים או להיפך
- אקסיומה הנה נאותה אם הת"פ הנתונות גוררות את הת"פ החדשה

נאותות של רפלקסיביות

- *רפלקסיביות:* אם $Y \subseteq X$ אז $X \to Y$
- נאותות של רפלקסיביות פירושה שאם $Y \subseteq X$ אז $X \to Y$ נגררת מהקבוצה הריקה של ת"פ
  - כלומר, $X \to Y$ מתקיימת בכל יחס
- הראנו קודם שת"פ הנובעת מרפלקסיביות הנה טריוויאלית, כלומר מתקיימת בכל יחס, ולפיכך אקסיומת הרפלקסיביות הנה נאותה

נאותות של אוגמנטציה

- צריך להראות שהת"פ $X \to Y$ גוררת את הת"פ $XZ \to YZ$
- לפיכך, נניח שהיחס r מקיים את הת"פ $X \to Y$
- צריך להראות שהיחס r מקיים את הת"פ

$$XZ \to YZ$$

הוכחת נאותות של אוגמנטציה

- נניח שהת"פ $X \to Y$ מתקיימת ביחס r ונראה שגם $XZ \to YZ$ מתקיימת ב- r
- תהיינה $s,t \in r$ רשומות כך ש- $s[XZ] = t[XZ]$
- צריך להראות ש- $s[YZ] = t[YZ]$
- נתון שהת"פ $X \to Y$ מתקיימת ב-r, כי $s[Y] = t[Y]$ ו- $s[X] = t[X]$ (כי נתון $s[XZ] = t[XZ]$)
- ברור ש-$s[Z] = t[Z]$, ולפיכך $s[YZ] = t[YZ]$

נאותות של טרנזיטיביות

- צריך להראות שהת"פ $X \to Y$ ו- $Y \to Z$ גוררות את הת"פ $X \to Z$
- נניח שהת"פ $X \to Y$ ו- $Y \to Z$ מתקיימות ביחס r
- צריך להראות שגם $X \to Z$ מתקיימת ב-r

נאותות של טרנזיטיביות

- נניח שהת"פ $X \to Y$ ו- $Y \to Z$ מתקיימות ביחס r ונראה שגם $X \to Z$ מתקיימת ב-r
- תהיינה $s,t \in r$ רשומות כך ש- $s[X] = t[X]$
- צריך להראות ש-$s[Z] = t[Z]$
- נתון שהת"פ $X \to Y$ מתקיימת ב-r, כי $s[Y] = t[Y]$ ונתון ש- $s[X] = t[X]$
- לפיכך, $s[Z] = t[Z]$, כי הת"פ $Y \to Z$ מתקיימת ב-r

משפט הנאותות

- *משפט:* $F \models X \to Y \Leftarrow F \vdash X \to Y$
- *הוכחה:* ההוכחה הפורמלית היא באינדוקציה על אורך הגזירה. הטענות על הנאותות של כל אחת מהאקסיומות, שהוכחו בשקפים הקודמים, משמשות כטענות עזר בהוכחה האינדוקטיבית.

אקסיומות נוספות

Additional Axioms

אקסיומות נוספות

- *איחוד:*
  אם $X \to Y$ ו- $X \to Z$, אז $X \to YZ$
- *פירוק:*
  אם $X \to YZ$, אז $X \to Y$ וגם $X \to Z$
- *פסודו-טרנזיטיביות:*
  אם $X \to Y$ ו- $YW \to Z$, אז $XW \to Z$

נאותות של האקסיומות הנוספות

- נאותות של האקסיומות החדשות אפשר להראות באחת משתי הדרכים הבאות:
  - באופן ישיר, כפי שהראנו עבור האקסיומות הבסיסיות, או
  - על ידי גזירת האקסיומות הנוספות מהאקסיומות הבסיסיות

61

גזירה של איחוד

1.	$X \to Y$	נתון
2.	$XZ \to YZ$	אוגמנטציה של 1 עם Z
3.	$X \to Z$	נתון
4.	$X \to XZ$	אוגמנטציה של 3 עם X
5.	$X \to YZ$	טרנזיטיביות של 4 ו- 2

62

גזירה של פירוק

1.	$X \to YZ$	נתון
2.	$YZ \to Y$	רפלקסיביות
3.	$X \to Y$	טרנזיטיביות של 1 ו- 2

באופן דומה מראים שאפשר לגזור את
$X \to Z$

63

גזירה של פסודו-טרנזיטיביות

1.	$X \to Y$	נתון
2.	$XW \to YW$	אוגמנטציה של 1 עם W
3.	$YW \to Z$	נתון
4.	$XW \to Z$	טרנזיטיביות של 2 ו- 3

64

סגור של קבוצת אטריביוטים

The Closure of a Set of Attributes

65

הגדרת הסגור של קבוצת אטריביוטים

- *הסגור* של קבוצת אטריביוטים X ביחס לקבוצת ת"פ F, סימון X_F^+, הוא

$$X_F^+ = \{A \mid F \vdash X \to A\}$$

כלומר, קבוצת כל האטריביוטים A כך שניתן לגזור את $X \to A$ מ- F

- אם F ברור מההקשר, אז מסמנים X^+

66

הלמה של הסגור

- *למה:* $X \to Y$ גזירה מ-F אם ורק אם $Y \subseteq X_F^+$
- *הוכחה:* הלמה נובעת מההגדרה של הסגור ומהאקסיומות של האיחוד והפירוק.

- הערה: מהגדרת הסגור ומאקסיומת הרפלקסיביות נובע ש- $X \subseteq X^+$

שלמות של האקסיומות

Completeness of the Axioms

שלמות של האקסיומות

- שלמות פירושה שאם יש גרירה אז יש גם גזירה, כלומר

$$F \models X \to Y \Rightarrow F \vdash X \to Y$$

- בצורה שקולה, שלמות פירושה שאם אין גזירה אז אין גרירה, כלומר יש דוגמה נגדית

משפט השלמות

- *משפט:* $F \models X \to Y \Rightarrow F \vdash X \to Y$
- *הוכחה:*
- נניח שיש ת"פ $X \to Y$ שאיננה נגזרת מ-F, כלומר $F \not\vdash X \to Y$
- נראה שאין גרירה ע"י כך שנבנה דוגמה נגדית, כלומר יחס r שמקיים את F ואינו מקיים את $X \to Y$

אינטואיציה לגבי מבנה היחס r: היחס r אינו מקיים את $X \to Y$

X		שאר האטריביוטים				
1	1					
1	1					

- לכן חייבות להיות שתי רשומות ששוות על X (ושונות לפחות על אחד האטריביוטים של Y)

היחס r מקיים את כל הת"פ של F

X^+				שאר האטריביוטים		
1	1	1	1			
1	1	1	1			

- לכן שתי הרשומות חייבות להיות שוות על X^+ כי יחס שמקיים את F מקיים גם את $X \to X^+$
  - מכיוון שלפי הלמה של הסגור, $F \vdash X \to X^+$
  - ולכן לפי משפט הנאותות, $F \models X \to X^+$

האינטואיציה מובילה להגדרה הבאה של היחס r

שאר האטריביוטים				X^+		
1	1	1	1	1	1	1
0	0	0	1	1	1	1

- שאר האטריביוטים הנם כל האטריביוטים שמופיעים ב-F או ב-Y ואינם מופיעים ב-X^+
- ב-Y יש לפחות אטריביוט אחד שמופיע בין שאר האטריביוטים (מדוע?)
  - לכן הת"פ $X \to Y$ אינה מתקיימת ב-r

המשך ההוכחה הפורמלית: הדוגמה הנגדית

- הדוגמה הנגדית היא יחס r מעל קבוצת האטריביוטים שמופיעים ב-F
- ביחס שתי רשומות ששוות על כל האטריביוטים של X^+ ושונות על כל אטריביוט שאינו שייך ל-X^+

עובדה 1: $X \to Y$ איננה מתקיימת

- נתון ש- $F \not\vdash X \to Y$
- לכן, לפי הלמה של הסגור, $Y \not\subset X^+$
- לפיכך, הת"פ $X \to Y$ איננה מתקיימת ביחס r

עובדה 2: F מתקיימת ב-r

- שים לב: משתי העובדות נובע שהיחס r הוא אכן דוגמה נגדית לגרירה $F \models X \to Y$
- כדי להראות שכל הת"פ של F מתקיימות ביחס r, צריך לדון בשני מקרים

מקרה ראשון

- נניח ש- $V \to W \in F$ ת"פ, כך ש- $V \not\subset X^+$
- ברור ש- $V \to W$ מתקיימת ב-r (כי אין ב-r שתי רשומות שונות ששוות על כל האטריביוטים של V)

מקרה שני

- נניח ש- $V \to W \in F$ ת"פ, כך ש- $V \subseteq X^+$
- לפי הלמה של הסגור, ניתן לגזור מ- F את הת"פ $X \to V$
- אבל הת"פ $V \to W$ שייכת ל- F
- לכן, ע"י טרנזיטיביות, אפשר לגזור את $X \to W$
- לפיכך, לפי הלמה של הסגור, $W \subseteq X^+$, ולכן:
- שתי הרשומות של r שוות על כל האטריביוטים של W; מסקנה: $V \to W$ מתקיימת ב-r