Assignment 5: Dependency Theory

Hien Tu - tun1

November 19, 2021

D1 Reasoning with dependencies

1.
- Assume $AB \longrightarrow C, A \longrightarrow D$ and $CD \longrightarrow EF$
- Apply Augmentation on $AB \longrightarrow C$ with A to derive $AB \longrightarrow CA$
- Apply Augmentation on $A \longrightarrow D$ with C to derive $CA \longrightarrow CD$
- Apply Transitivity on $AB \longrightarrow CA$ and $CA \longrightarrow CD$ to derive $AB \longrightarrow CD$
- Apply Transitivity on $AB \longrightarrow CD$ and $CD \longrightarrow EF$ to derive $AB \longrightarrow EF$
- Apply Reflection on $F \subseteq EF$ to derive $EF \longrightarrow F$
- Apply Transitivity on $AB \longrightarrow EF$ and $EF \longrightarrow F$ to derive $AB \longrightarrow F$
- Hence, $AB \longrightarrow F$

2.
- Assume we have rows $r_1, r_2 \in I$ of instance I such that $r_1[XW] = r_2[XW]$
- By $r_1[XW] = r_2[XW]$, we have $r_1[X] = r_2[X]$ and $r_1[W] = r_2[W]$
- Using $X \longrightarrow Y$ and $r_1[X] = r_2[X]$, we conclude that $r_1[Y] = r_2[Y]$
- By $r_1[Y] = r_2[Y]$ and $r_1[W] = r_2[W]$, we have $r_1[YW] = r_2[YW]$
- Using $YW \longrightarrow Z$ and $r_1[YW] = r_2[YW]$, we conclude that $r_1[Z] = r_2[Z]$

- Hence, $r_1[Z] = r_2[Z]$ holds.

3.
- Assume we have $r_1[X]$ for every instance I_1 of R and every row $r_1 \in I_1$
- Using $R[X] \subseteq S[Y]$ and $r_1[X]$, there exists a row in instance I_2 of S with $r_1[X] = r_2[X]$
- Using $S[Y] \subseteq T[Z]$ and $r_2[Y]$, there exists a row in instance I_3 of T with $r_2[Y] = r_3[Z]$
- Thus, for every instance I_1 of R and every row $r_1 \in R$, there exists a row in instance I_3 of T such that $r_1[X] = r_3[Z]$
- Hence, $R[X] \subseteq T[Z]$

4.
- Assume $X \twoheadrightarrow Y$ and $XY \longrightarrow Z$
- Apply Complementation on $X \twoheadrightarrow Y$ to derive $X \twoheadrightarrow Z$ (with Z all attributes of $\mathbf{R}$ not in X and Y)
- Apply Reflexivity on $Z \setminus (X \cup Y) \subseteq Z$ to derive $Z \longrightarrow Z \setminus (X \cup Y)$
- Apply Transitivity on $XY \longrightarrow Z$ and $Z \longrightarrow Z \setminus (X \cup Y)$ to derive $XY \longrightarrow Z \setminus (X \cup Y)$
- Since Z is all attributes of $\mathbf{R}$ not in X and Y, $Z \cap XY = \emptyset$
- Apply Coalescence on $X \twoheadrightarrow Z$, $XY \longrightarrow Z \setminus (X \cup Y)$, $Z \cap XY = \emptyset$ and $Z \setminus (X \cup Y) \subseteq Z$, we conclude that $X \longrightarrow Z \setminus (X \cup Y)$
- Hence, $X \longrightarrow Z \setminus (X \cup Y)$

5. Consider the following table of the schema

person(<u>name</u>, <u>number</u>, birthdate, age)

name	**number**	**birthdate**	**age**
Alice	1	2001-01-01	20
Alice	2	2005-09-05	16

Since name, number are the primary keys, they determine all attributes. Thus, "name, number $\longrightarrow$ birthdate".
We know from the lecture that birthdate determines age since people who have the same birthdate would have the same age, and so "birthdate $\longrightarrow$ age".
Since {birthdate} $\subseteq$ {name, birthdate}, by applying Reflexivity, we get "name, birthdate $\longrightarrow$ birthdate".
Applying Transitivity on "name, birthdate $\longrightarrow$ birthdate" and "birthdate $\longrightarrow$ age", we derive "name, birthdate $\longrightarrow$ age"
Hence, we have "name, number $\longrightarrow$ birthdate" and "name, birthdate $\longrightarrow$ age".
However, from the table, we can see that "name $\longrightarrow$ age" does not hold.
Let X = name, W = number, Y = birthdate, Z = age, we have shown that the inference rule from the question is not sound.

6.
- For the most trivial case where $Y \subseteq X$, we can just use Reflexivity to get $X \longrightarrow Y$. Otherwise,
- Let $Y = \{y_1, y_2, \ldots, y_n\}$ where each y_i is "atomic" attribute (i.e., not a set of attributes)
- Let $X = \{x_1, x_2, \ldots, x_m\}$ where each x_i is "atomic" attribute (i.e., not a set of attributes)
- Since $\mathfrak{S} \models X \longrightarrow Y$, by Decomposition, we know $\mathfrak{S} \models X \longrightarrow y_i$ for all $y_i \in Y$ $(1 \leq i \leq n)$
- Furthermore, since closure is complete, we know $y_i \in X^+$ for all $y_i \in Y$. Thus, $Y \subseteq X^+$
- So, we can write $X^+ = XYZ_1Z_2 \ldots Z_k$
  where $Z_1Z_2 \ldots Z_k$ are sets of attributes different from each other and different from X and Y. They are added during the process of the closure algorithm.
  Note that Z_i may contain attributes of X and/or Y. Z_i could also be strict subset or strict superset of X and/or Y.
- Another way of writing X^+ is $X^+ = X \cup Y \cup Z_1 \cup Z_2 \cup \cdots \cup Z_k$
- If there is no $Z_i \in X^+$, that is $X^+ = X \cup Y$, then it means that $X \longrightarrow Y \in \mathfrak{S}$. Thus, it is obvious that we can derive $X \longrightarrow Y$ from $\mathfrak{S}$

- If there is at least one $Z_i \in X^+$
  - If $X \longrightarrow Y \in \mathfrak{S}$, then it is trivial as shown before
  - If $X \longrightarrow Y \notin \mathfrak{S}$:
    We will prove by induction on the last step taken to get $X \longrightarrow Y$. We will proceed by case analysis over all possible situations of the last step taken to get $X \longrightarrow Y$.
    From the start of the closure algorithm for X^+, we know that $X \longrightarrow Z_i \in \mathfrak{S}$ for some Z_i that is neither X nor Y.
    Case 1:
    If $Z_i \longrightarrow Y \in \mathfrak{S}$, then we can use Transitivity on $X \longrightarrow Z_i$ and $Z_i \longrightarrow Y$ to get $X \longrightarrow Y$.
    Case 2:
    If $Z_i \supset Y$ (Z_i is a strict superset of Y), then, by Decomposition rule, we could get $X \longrightarrow Y$ (and $X \longrightarrow Z_i \setminus Y$).
    Case 3:
    If $Z_i \subset Y$ and we also have $X \longrightarrow Z_j \in \mathfrak{S}$ where $Y = Z_i \cup Z_j$, then we can use Union rule to get $X \longrightarrow Y$.
    Case 4:
    If, from the start of the closure algorithm, instead of having $X \longrightarrow Z_i$, we have $X_k \longrightarrow Z_i$ where $X_k \subset X$.
    * If $X \setminus X_k \cup Z_i = Y$ (Y is made of Z_i and the part of X where X_k does not have), then we can use Augmentation, to get $X_i(X \setminus X_i) \longrightarrow Z_i(X \setminus X_i)$, which is equivalent to $X \longrightarrow Y$.
    * If $X \setminus X_k \cup Z_i \neq Y$, then we will fall into one of the above cases.
- Therefore, $X \longrightarrow Y$ can be dervied using the inference rules R1 - R3.
- Thus, the Armstrong's Axiom is complete.

7. The attribute closure of set of attributes C:
   - Initially, $closure = \{C\}$
   - From $C \longrightarrow A$, since $C \subseteq closure$ and $A \nsubseteq closure$, $closure = \{C, A\}$

- From $AC \longrightarrow E$, since $AC \subseteq closure$ and $E \nsubseteq closure$, $closure = \{C, A, E\}$
- From $E \longrightarrow B$, since $E \subseteq closure$ and $B \nsubseteq closure$, $closure = \{C, A, E, B\}$
- From $AB \longrightarrow D$, since $AB \subseteq closure$ and $D \nsubseteq closure$, $closure = \{C, A, E, B, D\}$
- From $BC \longrightarrow D$, since $BC \subseteq closure$ and $D \subseteq closure$, we don't need to add D into $closure$ more (as D is already in $closure$)
- From $D \longrightarrow A$, since $D \subseteq closure$ and $A \subseteq closure$, we don't need to add A into $closure$ more (as A is already in $closure$)
- Therefore, $C^+ = \{A, B, C, D, E\}$

The attribute closure of set of attributes (EA):

- Initially, $closure = \{E, A\}$
- From $E \longrightarrow B$, since $E \subseteq closure$ and $B \nsubseteq closure$, $closure = \{E, A, B\}$
- From $AB \longrightarrow D$, since $AB \subseteq closure$ and $D \nsubseteq closure$, $closure = \{E, A, B, D\}$
- From $AC \longrightarrow E$, since $AC \nsubseteq closure$, we don't need to add E into $closure$
- From $BC \longrightarrow D$, since $BC \nsubseteq closure$, we don't need to add D into $closure$
- From $C \longrightarrow A$, since $C \nsubseteq closure$, we don't need to add A into $closure$
- From $D \longrightarrow A$, since $D \subseteq closure$ and $A \subseteq closure$, we don't need to add A into $closure$ more (as A is already in $closure$)
- Therefore, $(EA)^+ = \{A, B, D, E\}$

8. Compute X^+ for every $X \subseteq \{A, B, C, D, E\}$

   - $\{\}^+ = \{\}$ since there is no $\{\} \longrightarrow \ldots$ in $\mathfrak{S}$

- $A^+ = \{A\}$ since there is no dependency that would satisfy the closure algorithm
  So we have $A \longrightarrow A$
- $B^+ = \{B\}$ since there is no dependency that would satisfy the closure algorithm
  So we have $B \longrightarrow B$
- $C^+ = \{A, B, C, D, E\}$ from question 7
  So we have $C \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $D^+ = \{A, D\}$ since there is only $D \longrightarrow A$ that would satisfy the closure algorithm
  So we have $D \longrightarrow X$ for all $X \subseteq \{A, D\}$
- $E^+ = \{B, E\}$ since there is only $E \longrightarrow B$ that would satisfy the closure algorithm
  So we have $E \longrightarrow X$ for all $X \subseteq \{B, E\}$
- $(AB)^+ = \{A, B, D\}$ since there is only $AB \longrightarrow D$ that would satisfy the closure algorithm
  So we have $AB \longrightarrow X$ for all $X \subseteq \{A, B, D\}$
- $(AC)^+ = \{A, B, C, D, E\}$ since AC includes C and we can reach others from only C
  So we have $AC \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(AD)^+ = \{A, D\}$ since there is no dependency that would satisfy the closure algorithm
  So we have $AD \longrightarrow X$ for all $X \subseteq \{A, D\}$
- $(AE)^+ = \{A, B, E\}$ since there is only $E \longrightarrow B$ that would satisfy the closure algorithm
  So we have $AE \longrightarrow X$ for all $X \subseteq \{A, B, E\}$
- $(BC)^+ = \{A, B, C, D, E\}$ since BC includes C and we can reach others from only C
  So we have $BC \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(BD)^+ = \{A, B, D\}$ since there is only $D \longrightarrow A$ that would satisfy the closure algorithm
  So we have $BD \longrightarrow X$ for all $X \subseteq \{A, B, D\}$
- $(BE)^+ = \{B, E\}$ since there is no dependency that would satisfy the closure algorithm
  So we have $BE \longrightarrow X$ for all $X \subseteq \{B, E\}$

- $(CD)^+ = \{A, B, C, D, E\}$ since CD includes C and we can reach others from only C
  So we have $CD \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(CE)^+ = \{A, B, C, D, E\}$ since CE includes C and we can reach others from only C
  So we have $CE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(DE)^+ = \{A, B, D, E\}$ there are $D \longrightarrow A$, $E \longrightarrow B$ satisfy the closure algorithm
  So we have $DE \longrightarrow X$ for all $X \subseteq \{A, B, D, E\}$
- $(ABC)^+ = \{A, B, C, D, E\}$ since ABC includes C
  So we have $ABC \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ABD)^+ = \{A, B, D\}$ since there is no dependency that would satisfy the closure algorithm
  So we have $ABD \longrightarrow X$ for all $X \subseteq \{A, B, D\}$
- $(ABE)^+ = \{A, B, D, E\}$ since there is only $AB \longrightarrow D$ that would satisfy the closure algorithm
  So we have $ABE \longrightarrow X$ for all $X \subseteq \{A, B, D, E\}$
- $(ACD)^+ = \{A, B, C, D, E\}$ since ACD includes C
  So we have $ACD \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ACE)^+ = \{A, B, C, D, E\}$ since ACE includes C
  So we have $ACE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ADE)^+ = \{A, B, D, E\}$ since there is only $E \longrightarrow B$ that would satisfy the closure algorithm
  So we have $ADE \longrightarrow X$ for all $X \subseteq \{A, B, D, E\}$
- $(BCD)^+ = \{A, B, C, D, E\}$ since BCD includes C
  So we have $BCD \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(BCE)^+ = \{A, B, C, D, E\}$ since BCE includes C
  So we have $BCE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(BDE)^+ = \{A, B, D, E\}$ since there is only $D \longrightarrow A$ that would satisfy the closure algorithm
  So we have $BDE \longrightarrow X$ for all $X \subseteq \{A, B, D, E\}$
- $(CDE)^+ = \{A, B, C, D, E\}$ since CDE includes C
  So we have $CDE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$

- $(ABCD)^+ = \{A, B, C, D, E\}$ since $ABCD$ includes C
  So we have $ABCD \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ABCE)^+ = \{A, B, C, D, E\}$ since $ABCE$ includes C
  So we have $ABCE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ABDE)^+ = \{A, B, D, E\}$ since there is no dependency that would satisfy the closure algorithm
  So we have $ABDE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ACDE)^+ = \{A, B, C, D, E\}$ since $ACDE$ includes C
  So we have $ACDE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(BCDE)^+ = \{A, B, C, D, E\}$ since $BCDE$ includes C
  So we have $BCDE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- $(ABCDE)^+ = \{A, B, C, D, E\}$ since $ABCDE$ includes C
  So we have $ABCDE \longrightarrow X$ for all $X \subseteq \{A, B, C, D, E\}$
- From the closure, we know that any combination that contains C is the superkey, more specifically, AC, BC, C, CD, CE, ABC, ACD, ACE, BCD, BCE, CDE, $ABCD$, $ABCE$, $ACDE$, $BCDE$, $ABCDE$ are superkeys
- Notice that, from C, we can derive all attributes in **r** C is already minimal, so C is the (candidate) key

9.
- Starting with $\{AB \longrightarrow D, AC \longrightarrow E, BC \longrightarrow D, C \longrightarrow A, D \longrightarrow A, E \longrightarrow B\}$
- From $C \longrightarrow A$, we can use Augmentation with C to derive $C \longrightarrow AC$. We also have $AC \longrightarrow E$. Then, by Transitivity, we conclude $C \longrightarrow E$. Thus, we can add $C \longrightarrow E$ to our set.
  $\{AB \longrightarrow D, AC \longrightarrow E, BC \longrightarrow D, C \longrightarrow A, D \longrightarrow A, E \longrightarrow B, C \longrightarrow E\}$
- From $C \longrightarrow E$, we can use Augmentation with A to derive $AC \longrightarrow AE$. Then, we can use Decomposition to get $AC \longrightarrow A$ and $AC \longrightarrow E$. $AC \longrightarrow A$ is trivial since $A \subseteq AC$. This means that, from $C \longrightarrow E$, we can get $AC \longrightarrow E$. So, we can get rid of $AC \longrightarrow E$.
  $\{AB \longrightarrow D, BC \longrightarrow D, C \longrightarrow A, D \longrightarrow A, E \longrightarrow B, C \longrightarrow E\}$

- From $C \longrightarrow E$ and $E \longrightarrow B$, by Transitivity, we get $C \longrightarrow B$. Then, we get apply Augmentation on $C \longrightarrow B$ with C to derive $C \longrightarrow BC$. We also have $BC \longrightarrow D$. So, by Transitivity, we derive $C \longrightarrow D$. Thus, we can add $C \longrightarrow D$ to our set.
  $\{AB \longrightarrow D, BC \longrightarrow D, C \longrightarrow A, D \longrightarrow A, E \longrightarrow B, C \longrightarrow E, C \longrightarrow D\}$
- From $C \longrightarrow D$, we can apply Augmentation with B to get $BC \longrightarrow BD$. Then, we can use Decomposition to get $BC \longrightarrow B$ and $BC \longrightarrow D$. $BC \longrightarrow B$ is trivial since $B \subseteq BC$. This means that from $C \longrightarrow D$, we can get $BC \longrightarrow D$. So, we can get rid of $BC \longrightarrow D$.
  $\{AB \longrightarrow D, C \longrightarrow A, D \longrightarrow A, E \longrightarrow B, C \longrightarrow E, C \longrightarrow D\}$
- By Transitivity on $C \longrightarrow D$ and $D \longrightarrow A$, we can derive $C \longrightarrow A$. Thus, we can get rid of $C \longrightarrow A$.
  $\{AB \longrightarrow D, D \longrightarrow A, E \longrightarrow B, C \longrightarrow E, C \longrightarrow D\}$
- From here, we cannot "reduce" anymore. Thus, the minimal cover for $\mathfrak{S}$ is $\{AB \longrightarrow D, D \longrightarrow A, E \longrightarrow B, C \longrightarrow E, C \longrightarrow D\}$