Closure of F

SAPIENZA
UNIVERSITÀ DI ROMA

F^A

- problem: computing the set of dependencies F^+ that **is satisfied** by **each legal instance of** a schema *R* over which a set of functional dependencies F is defined

- we have concluded that trivially $F \subseteq F^+$, as an instance r is **legal only if it** satisfies **all** dependencies in F

- what about the other dependencies in F^+?

- we introduce another set, easy (but time consuming) to compute: $\mathbf{F^A}$

Armstrong's axioms

F^A is a set of functional dependencies on R, so that:

- if $X \to Y \in F$ then $X \to Y \in F^A$
- if $Y \subseteq X \in R$ then $X \to Y \in F^A$ (**reflexivity**)
- if $X \to Y \in F^A$ then $XZ \to YZ \in F^A$, for each $Z \in R$ (**augmentation**)
- if $X \to Y \in F^A$ and $Y \to Z \in F^A$ then $X \to Z \in F^A$ (**transitivity**)

we will show that $F^+ = F^A$, i.e., that the closure of a set of functional dependencies F can be obtained from F by recursively applying the **Armstrong's axioms**

A few simple observations

- if $Y \subseteq X \in R$ then $X \rightarrow Y \in F^A$ (reflexivity)
- Name∈(Name, Surname) so, obviously, if two tuples have the same values for (Name, Surname) then they will surely have the same value for Name, therefore (Name, Surname)→Name is always satisfied

- if $X \rightarrow Y \in F^A$ then $XZ \rightarrow YZ \in F^A$, for each $Z \in R$ (augmentation)
- if TaxCode→Surname is satisfied, it means that if two tuples have the same TaxCode then they must also have the same Surname
- if we also consider the Address attribute, we will surely have that if two tuples are identical on (TaxCode, Address), they will be identical on (Surname, Address) too
- so, if TaxCode→Surname is satisfied
- then TaxCode, Address→Surname, Address is also satisfied

A few simple observations

- if $X \rightarrow Y \in F^A$ and $Y \rightarrow Z \in F^A$ then $X \rightarrow Z \in F^A$(transitivity)
- if Matriculation→TaxCode is satisfied then **if** two tuples have the same Matriculation, **then** they must also have the same TaxCode
- if TaxCode→Surname is satisfied then **if** two tuples have the same TaxCode, **then** they must also have the same Surname
- if both the above dependencies are satisfied, and two tuples have the same Matriculation, then they must also have the same TaxCode; but then they must also have the same Surname!
- so, if both dependencies are satisfied, then whenever two tuples have the same Matriculation, then they must also have the same Surname: Matriculation→Surname

Before we proceed...

- we introduce three more rules derived from the axioms:
  - if $X \rightarrow Y \in F^A$ and $X \rightarrow Z \in F^A$ then $X \rightarrow YZ \in F^A$ (**union rule)**
  - if $X \rightarrow Y \in F^A$ and $Z \subseteq Y$ then $X \rightarrow Z \in F^A$ (**decomposition rule)**
  - if $X \rightarrow Y \in F^A$ and $WY \rightarrow Z \in F^A$ then $WX \rightarrow Z \in F^A$ (**pseudotransitivity rule**)

Theorem

- **Theorem** Let F be a set of functional dependencies; the following implications apply:

a) if $X \rightarrow Y \in F^A$ and $X \rightarrow Z \in F^A$ then $X \rightarrow YZ \in F^A$

b) if $X \rightarrow Y \in F^A$ and $Z \subseteq Y$ then $X \rightarrow Z \in F^A$

c) if $X \rightarrow Y \in F^A$ and $WY \rightarrow Z \in F^A$ then $WX \rightarrow Z \in F^A$

- **Demonstration**

we're dealing with sets, so XX=X

- **(a)**
- if $X \rightarrow Y \in F^A$**,** by the axiom of augmentation we have $X \rightarrow XY \in F^A$; similarly**,** if $X \rightarrow Z \in F^A$**,** by the axiom of augmentation we have $XY \rightarrow YZ \in F^A$; since $X \rightarrow XY \in F^A$ and $XY \rightarrow YZ \in F^A$, by the axiom of transitivity we have $X \rightarrow YZ \in F^A$
- **(b)**
- if $Z \subseteq Y$ then, by the axiom of reflexivity, we have $Y \rightarrow Z \in F^A$; since $X \rightarrow Y \in F^A$ and $Y \rightarrow Z \in F^{A,}$ by the axiom of transitivity we have $X \rightarrow Z \in F^A$
- **(c)**
- if $X \rightarrow Y \in F^A$, by the axiom of augmentation we have $WX \rightarrow WY \in F^A$; since $WX \rightarrow WY \in F^A$ and $WY \rightarrow Z \in F^A$, by the axiom of transitivity we have $WX \rightarrow Z \in F^A$

Observation

we observe that:

- for the **union** rule, if $X \to A_i \in F^A$, i=1, ..., n then $X \to A_1, \ldots, A_i \ldots A_n \in F^A$
- for the **decomposition** rule, if $X \to A_1, \ldots, A_i \ldots A_n \in F^A$ then $X \to A_i \in F^A$, i=1, ..., n

so

- $X \to A_1, \ldots, A_i \ldots A_n \in F^A \Leftrightarrow X \to A_i \in F^A$, i=1, ..., n

if and only if

Closure a set of attributes

- **Definition**

let R be a schema, F a set of dependencies on R, and X a subset of R

the closure of X with respect to F, denoted X^+_F (or simply X^+, if no ambiguity arises) is defined as:

$$X^+_F = \{A \mid X \rightarrow A \in F^A\}$$

- in practice, all those attributes that are functionally determined by X, by possibly applying Armstrong's axioms

trivially: $X \subseteq X^+_F$ (by reflexivity!)

Lemma

Lemma let R be a schema and F a set of functional dependencies on R: $X \rightarrow Y \in F^A \Leftrightarrow Y \subseteq X^+$

- **Demonstration**

Let $Y = A_1, A_2, \ldots, A_n$

- **if** ($\Leftarrow$)

since $Y \subseteq X^+$, for each i=1,...,n we have that $X \rightarrow A_i \in F^A$ by the **union** rule, $X \rightarrow Y \in F^A$

- **only if** ($\Rightarrow$)

since $X \rightarrow Y \in F^A$, by the **decomposition** rule we have that, for each i=1,...,n, $X \rightarrow A_i \in F^A$, i.e., $A_i \in X^+$, for each i=1,...,n, and, therefore, $Y \subseteq X^+$

Theorem: $F^+ = F^A$

- **Theorem** let R be a schema and F a set of functional dependencies on R:
  - $F^+ = F^A$
- **Proof** (we prove it by double inclusion)

contains

$F^+ \supseteq F^A$: let $X \rightarrow Y$ be a functional dependency in F^A; we prove that $X \rightarrow Y \in F^+$ **by induction** on i, the number of applications of one of the Armstrong's axioms

basis: i=0; in such a case $X \rightarrow Y \in F^A$ is also in F and, therefore, $X \rightarrow Y$ is in F^+

induction step: i>0; by hypothesis, every functional dependency obtained from F by applying Armstrong's axioms a number of times less than or equal to i-1 is in F^+; we have to prove this is also true a number of applications equal to i

Theorem: $F^+ = F^A$ (continued)

(case 1)

- $X \rightarrow Y$ was obtained in F^A by applying reflexivity, as $Y \subseteq X$
- let r be an instance of R and let t_1 and t_2 be two tuples of r, such that
- $t_1[X]=t_2[X]$
- trivially, we have that $t_1[Y]=t_2[Y]$, so $X \rightarrow Y \in F^+$

Theorem: $F^+ = F^A$ (continued)

(case 2)

- $X \rightarrow Y \in F^A$ was obtained by applying augmentation to a functional dependency $V \rightarrow W \in F^A$, that was obtained, in turn, by recursively applying Armstrong's axioms a number of times $\leq i-1$
- thus, for the inductive hypothesis $V \rightarrow W \in F^+$
- it will then be $X=VZ$ and $Y=WZ$, for some $Z \subseteq R$
- let r be a legal instance of R and let t_1 and t_2 be two tuples of r, such that $t_1[X]=t_2[X]$
- trivially, we have that $t_1[V]=t_2[V]$ and $t_1[Z]=t_2[Z]$
- for the inductive hypothesis from $t_1[V]=t_2[V]$ follows that $t_1[W]=t_2[W]$ and from $t_1[W]=t_2[W]$ and $t_1[Z]=t_2[Z]$ follows that $t_1[Y]=t_2[Y]$

Theorem: $F^+ = F^A$ (continued)

(case 3)

- $X \rightarrow Y \in F^A$ was obtained by applying the axiom of transitivity to two functional dependencies $X \rightarrow Z$, $Z \rightarrow Y \in F^A$, obtained, in turn, by recursively applying Armstrong's axioms a number of times less than or equal to i-1
- thus, for the inductive hypothesis $X \rightarrow Z$, $Z \rightarrow Y \in F^+$
- let r be a legal instance of R and let t_1 and t_2 be two tuples of r, such that $t_1[X]=t_2[X]$
- for the inductive hypothesis from $t_1[X]=t_2[X]$ follows $t_1[Z]=t_2[Z]$; from $t_1[Z]=t_2[Z]$, again for the inductive hypothesis follows $t_1[Y]=t_2[Y]$

Theorem: $F^{+}=F^{A}$ (continued)

$F^{+}\subseteq F^{A}$ (by contradiction)

let's suppose that there exists a functional dependency $\mathbf{X}\rightarrow\mathbf{Y}\in\mathbf{F}^{+}$, such that $\mathbf{X}\rightarrow\mathbf{Y}\notin\mathbf{F}^{A}$; we will use a particular legal instance of R, to show that this assumption leads to a contradiction

let's consider the following instance r of R:

$$
r \quad \begin{array}{|c|c|c|c|c|c|c|c|}
\multicolumn{4}{c}{\overbrace{\hspace{12em}}^{X^{+}}} & \multicolumn{4}{c}{\overbrace{\hspace{12em}}^{R-X^{+}}} \\
\hline
1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \\
\hline
1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\
\hline
\end{array}
$$

r has two tuples, identical on attributes in $\mathbf{X}^{+}$**,** and different on the others ($\mathbf{R}$-$\mathbf{X}^{+}$):

- we show that r is legal
- using r, we show that if $X\rightarrow Y\in F^{+}$, then it cannot happen that $\mathbf{X}\rightarrow\mathbf{Y}\notin\mathbf{F}^{A}$ (if we assume that, we have a contradiction)

Theorem: $F^+ = F^A$ (continued)

- r is a legal instance
  - let's assume that r is not legal, so there exists at least a dependency $V \rightarrow W \in F$ (so, also in F^A) that it is not satisfied by r
  - it means that there are 2 tuples in r that are equal on V and different on W this implies that $V \subseteq X^+$ and $W \cap (R - X^+) \neq \emptyset$
  - since $V \subseteq X^+$, by the Lemma, we know that $X \rightarrow V \in F^A$; therefore, by the axiom of transitivity, since $X \rightarrow V$ and $V \rightarrow W$, follows that $X \rightarrow W \in F^A$
  - again by the Lemma , $W \subseteq X^+$, which contradicts $W \cap (R - X^+) \neq \emptyset$
  - so, r is a legal instance

Theorem: $F^{+}=F^{A}$ (continued)

- if $X \rightarrow Y \in F^{+}$, then it cannot happen that $X \rightarrow Y \notin F^{A}$
  - suppose that $X \rightarrow Y \in F^{+}$ and $X \rightarrow Y \notin F^{A}$
  - we know that r is a legal instance
  - since the dependencies in F^{+} are satisfied by every legal instance, then r satisfies $X \rightarrow Y$ (that is, if 2 tuples in r are identical on X, then they must be identical on Y)
  - we know that $X \subseteq X^{+}$
  - are there 2 tuples in r, which are identical on X? yes, there are!
  - so, they must be identical on Y too, so $Y \subseteq X^{+}$
  - then, by the lemma, $X \rightarrow Y \in F^{A}$, and we have a contradiction

Theorem: $F^{+}=F^{A}$ (final remarks)

- it is useful to note that the proof of this theorem relies on two very important facts:
  - the link that exists between the set of dependencies **F^+** and the legal instances: on one hand, **if** an instance is legal then it also satisfies all dependencies in **F^+;** on the other hand, **F^+** is the set of dependencies satisfied by each legal instance (to check if an instance is legal we need to check that it satisfies all the dependencies in F)
  - the link that exists between the closure X^+ of a set of attributes X and the subset of dependencies in F^A (which, we have just seen, is equal to F^+) that have X as determinant, i.e., $X \rightarrow Y \in F^A \Leftrightarrow Y \subseteq X^+$ which is equivalent to saying that:

    $X \rightarrow Y \in F^+ \Leftrightarrow Y \subseteq X^+$ and in particular that $X \rightarrow Y$ must be satisfied by every legal instance

Why do we need to know F^+?

- we now have a way to identify all dependencies in F^+: those that can be inserted into F^A starting from F and applying Armstrong's axioms and derived rules
- computing $F^A=F^+$ takes exponential time in R:
  - if we just consider the axiom of reflexivity, each possible subset of R leads to one dependency, and since the possible subsets of R are $2^{|R|}$, then $|F^+| >> 2^{|R|}$

Why do we need to know F^+?

- the definition of Third Normal Form (3NF) relies on F^+
- to obtain a schema that is in 3NF from another one which is not, we decompose the initial one into "smaller" ones
- in doing that, we would like to "maintain" the dependencies of the original schema, so we would like to preserve F^+