

Relation Schema:

Students(SID, name, address)

$f_{\text{sid} \rightarrow \text{name}}(\text{SID}) \rightarrow \text{name}$

functional dependency

$f(\text{Set A of attributes}) \rightarrow \text{Set B of attributes}$


If you know the values of all the attributes in A, for a row in the relation \rightarrow then there is at most one value for each attr in B for the same row



Apply:

[applicant]

SSN	college	name	HSCode	HSCity	sports
1	'NCsu'	'mary'	23	'Garner'	
1	'Duke'	'mary'	23	'Garner'	
1	'NCsu'	'mary'	74	'Raleigh'	
1	'Duke'	'mary'	74	'Raleigh'	


 12 tuples if 3 sports in each high school



1. Redundancy
2. Update anomaly
3. Deletion anomaly



$R(A, B, C, D)$
 $A \rightarrow B$

Meaning: functional dependency "A determines B" holds on R

 $BC \rightarrow D$

Not OK:

R:

A	B	C	D
a	b	c	q
a	c	g	k
a	b	c	b
f	d	h	m

R: A B C D

a	b	c	d
f	b	c	d
a	b	g	h

OK



Apply(SSN, college, name,
HSCode, HSCity, sports)

SSN \rightarrow name

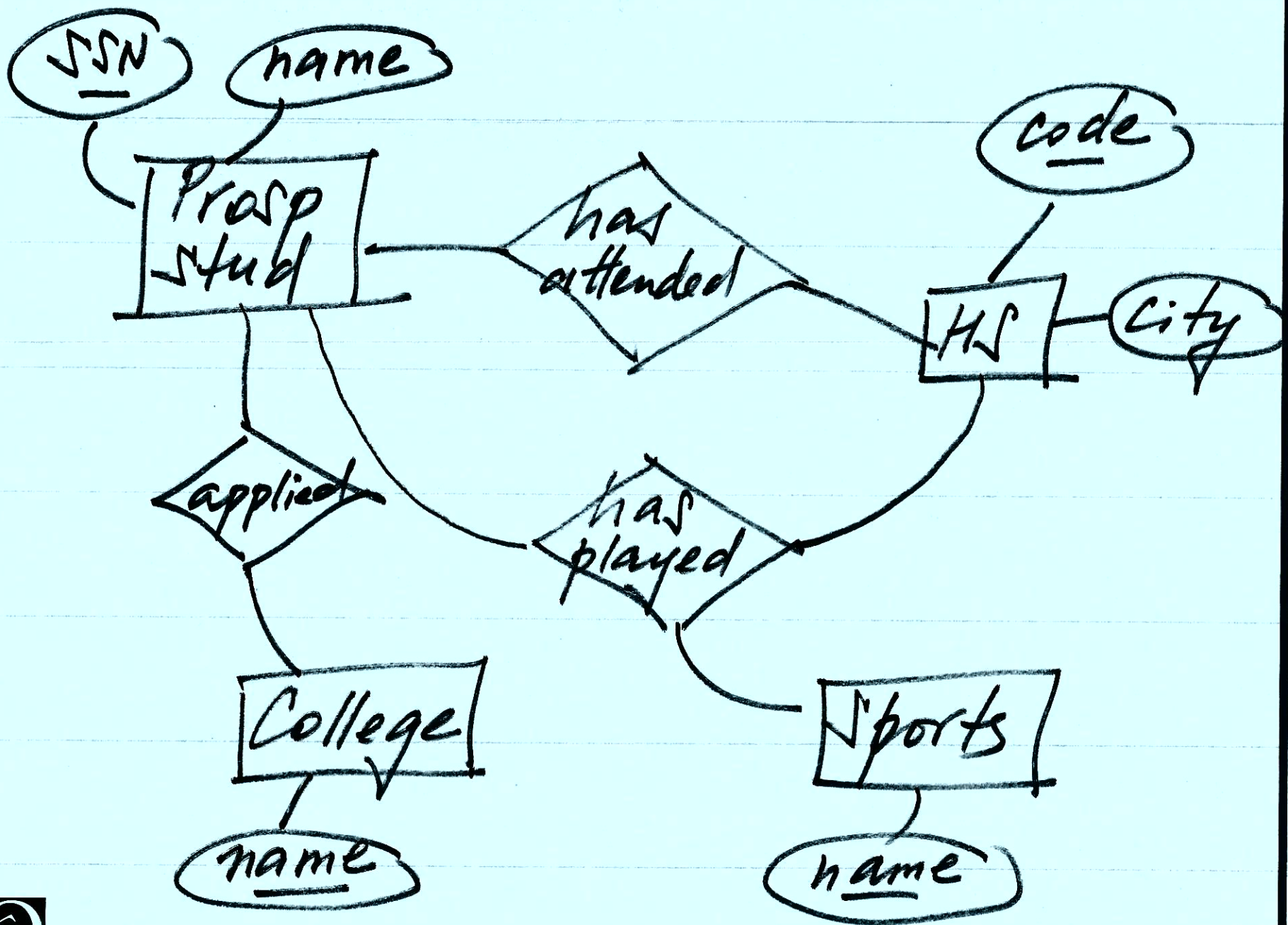
HSCode \rightarrow HSCity

ApplicatIT(SSN, name) \Leftarrow 1 row from
the Apply
relation

HS(HSCode, HSCity) \Leftarrow 2 rows from
the Apply relation

ApplyNew(SSN, college, HSCode, sports) \Leftarrow 12 rows from
the Apply relation





$PS(\underline{SSN}, name) \Leftarrow 1 \text{ tuple for Mary}$
 $HS(\underline{code}, city) \Leftarrow 2 \text{ tuples } \checkmark$
 $Sports(\underline{name}) \Leftarrow 3 \text{ tuples}$
 $Campus(\underline{name}) \Leftarrow 2 \text{ tuples}$
 $HasAttended(\underline{SSN}, \underline{code}) \Leftarrow 2 \text{ tuples}$
 $HasPlayed(\underline{SSN}, \underline{code}, \underline{name}) \Leftarrow 6 \text{ tuples}$
 $Applies(\underline{SSN}, \underline{name}) \Leftarrow 2 \text{ tuples}$



$R(A, B, C)$
 $A \rightarrow B$
 $B \rightarrow C$

 $R: A \quad B \quad C$

 $\underline{a} \quad \underline{d} \quad \underline{g}$
 $b \quad k \quad h$
 $c \quad \text{~~a~~} \quad \text{~~g~~}$
 $\underline{a} \quad \underline{d} \quad \underline{g}$

Better schema:

 $R1(A, B)$
 $R2(B, C)$
 \Downarrow
 $R1: A \quad B$

 $a \quad d$
 $b \quad k$
 $c \quad d$
 $R2: B \quad C$

 $\text{d} \quad \text{g}$
 $k \quad h$


$R(A, B, C)$ $A \rightarrow B$ $B \rightarrow C$

\checkmark
 $A \rightarrow C$ yes
"transitive
FD"

$C \overset{?}{\rightarrow} B$ no



Using FDs to prove

1. ~~#~~ Showing that a candidate FD holds (or not)
2. Deriving FDs on same relation
3. Find keys of relation
4. Find FDs in projections of relation
5. Decompose relation schemas - normalize them (3NF, BCNF)



Key of relation schema
 is \subseteq ~~it~~ a subset of
 its attributes such that:

(1) it functionally determines
 all the other attributes
 — "superkey"

(2) none of ~~its~~ ^{K 's} attributes
 can be removed with (1)
 preserved

