Understanding BCNF : Boyce Codd Normal Form

(Ken Goldberg, UC Berkeley IEOR Dept, Jan 2014)

Recall the definition of 3NF:
R is in 3NF if $\forall X \rightarrow Y$ **, either X is a superkey or Y is a prime attribute.**

BCNF is stricter:

R is in BCNF if $\forall X \rightarrow Y$ **, X is a super key.**
(BCNF is stronger, it eliminates second option)

Conditions for violating BCNF:

Consider R(A,B,C)
R is in 3NF but NOT in BCNF if all 5 of these conditions hold:

1) AB $\rightarrow$ C (required by the fact that AB is a Candidate Key)
2) A $\not\rightarrow$ C (A does NOT determine C: otherwise R is not in 3NF)
3) B $\not\rightarrow$ C (similarly, otherwise R is not in 3NF)
4) C $\rightarrow$ B (violates BCNF)
5) C $\not\rightarrow$ A (otherwise given 4, C would be a superkey)

We can normalize R into BCNF:
R1(A,C)
R2(C,B)

Example:

StudentMajor(SID, Major, Advisor)

Note: a student can have more than one Major, and one Advisor for each of their Major, and note that Advisors only advise in one Major

Advisor $\rightarrow$ Major

StudentMajor is in 3NF since Major is a Prime Attribute
but it is NOT in BCNF because Advisor is not a superkey.

To Normalize into BCNF

StudentAdvisors(SID, Advisor)
AdvisorMajor(Advisor, Major)

(Aside: Note:

StudentMajors(SID, Major)

AdvisorMajor(Advisor, Major)

This is in BCNF but does not capture which Advisors a student has.)