DASC 5333 Database Systems for Data Science
CSCI 4333 Design of Database Systems
Fall 2023
Suggested Solution to Homework #7

[1] See the separate file.

[2]

[a] R(A,B,C,D) {D->C, C->B}
[b] R(A,B,C,D) {AB->C, C->D}
[c] R(A,B,C,D) {A->B, B->ACD}
[d] R(A,B,C,D) {AB->C, AD->C}
[e] R(A,B,C,D) {A->B, B->A, AC->D}

[a] R(A,B,C,D) {D->C, C->B}

    CK:[1] AD

    Highest NF: 1NF

    Reason: D -> C violates 2NF since D is a proper subset of a CK (AD), and C is non-prime.

[b] R(A,B,C,D) {AB->C, C->D}

    CK:[1] AB

    Highest NF: 2NF

    Reason: C->D violates 3NF as C is not a superkey, and D is non-prime.

[c] R(A,B,C,D) {A->B, B->ACD}

    CK: [1] A, [2] B

    Highest NF: BCNF

    Reason: A and B in the LHS of FDs are superkeys.

[d] R(A,B,C,D) {AB->C, AD->C}

    CK:[1] ABD

    Highest NF: 1NF

    Reason: AB->C violates 2NF since AB is a proper subset of ABC, a CK, and C is non-prime.

[e] R(A,B,C,D) {A->B, B->A, AC->D}

    CK:[1] AC, [2] BC

    Highest NF: 3NF

    Reason: A-> B and B->A both violate BCNF as they are not superkeys.

[3] For GO(GroupId, GroupName, GroupEMail, GroupChairId, GroupChairLName, GroupChairFName, GroupMemberId, GroupMemberMajor)

(a) FD:

FD1: GroupId -> GroupName, GroupEMail, GroupChairId

FD2: GroupName -> GroupId

FD3: GroupChairId -> GroupChairLName, GroupChairFName

(b) CK: [1] {GroupId, GroupMemberId, GroupMemberMajor}, [1] {GroupName, GroupMemberId, GroupMemberMajor}

(c) 1NF. FD1 violates 2NF; FD2 violates BCNF; FD3 violates 3NF.

(d) Decomposition into BCNF component relations:

Group(GroupId, GroupName, GroupEMail, GroupChairId) { GroupId -> GroupName, GroupEMail, GroupChairId, GroupName -> GroupId}

GroupChair(GroupChairId, GroupChairLName, GroupChairFName) { GroupChairId -> GroupChairLName, GroupChairFName}

GroupMember(GroupId, GroupMemberId) {}

GroupeMemberMajor(GroupMemberId, GroupMemberMajor) {}

[4] For R(A,B,C,D,E), F = {A->B, AB->D, AD->E, C->D}

(a) Canonical cover: F' = {A->BDE, C->D}

Candidate key: [1] AC
Prime attributes: A, C;
Non-prime attributes: B, D, E

(b) 1NF. A->BDE and C->D both violate 2NF.

(c) Decomposition:

R1(A,B,D,E) {A->BDE} in BCNF
R2(C->D) {C->D} in BCNF
R3(A,C) {} in BCNF

[5] It is known that for R(A,B,C,D,E):

1. R has exactly two candidate keys
2. A is a candidate key.
3. D and E are non-prime attributes.

There are three scenarios for the second candidate keys:

1. B => 24 SK {A, AB, AC, AD, AE, ABC, ABD, ABE, ACD, ACE, ADE, ABCD, ABCE, ABDE, ACDE, ABCDE, B, BC, BD, BE, BCD, BCE, BDE, BCDE}
2. C => 24 SK {You figure out}
3. BC => 20 SK {A, AB, AC, AD, AE, ABC, ABD, ABE, ACD, ACE, ADE, ABCD, ABCE, ABDE, ACDE, ABCDE, BC, BCD, BCE, BCDE}

Thus, the number SK: {20,24}