

SINGLE – VALUED FACTS AND FUNCTIONAL DEPENDENCIES

SINGLE-VALUED FACTS (SVFS) – only one value is associated and implied

Single-valued facts constitute an important part of semantics (meaning of attributes) of a (relational) data model and they offer a detailed analysis of the natural language predicates for the relations in a model.

The Appointments relational data model

GeneralPractitioner (GPId, GPName)

Patient (PatientId, PatientName, GPId)

Appointment (PatientId, ApptDate, ApptTime, ConsId) [Alternate key is (ConsId, ApptDate, ApptTime)]

Consultant (ConsId, ConsName)

Hospital (HospNo, HospName)

ConsultantVisit (ConsId, ApptDate, HospNo)

Single-valued facts represented

1. A general practitioner has a name.
2. A patient has a name.
3. A patient has a general practitioner.
4. A patient on a given date has an appointment at a particular time.
5. A patient on a given date has an appointment with one consultant.
6. A consultant on a given date at a particular time has an appointment with one patient.
7. A consultant has a name.
8. A hospital has a name.
9. A consultant on a given date visits one hospital.

FUNCTIONAL DEPENDENCIES (FDS) – a convenient and precise shorthand for expressing SVFS

Given a relation, R, attribute R.B (which may be composite) is **functionally dependent** on attribute R.A (which may also be composite) iff each value of R.A is associated with precisely with one value of R.B at any one time.

$R.A \rightarrow R.B$

Given a relation, R, attribute R.B (which may be composite) is **fully functionally dependent** on attributes R.A (which may also be composite) if it is functionally dependent on R.A and not functionally dependent on any subset of R.A.

In the notation $R.A \rightarrow R.B$, the left-hand attribute is known as **determinant**.

FDs that express the SVFs of the Appointments relational data model

1. $GPId \rightarrow GPName$
2. $PatientId \rightarrow PatientName$
3. $PatientId \rightarrow GPId$
4. $PatientId, ApptDate \rightarrow ApptTime$
5. $PatientId, ApptDate \rightarrow ConsId$
6. $ConsId, ApptDate, ApptTime \rightarrow PatientId$
7. $ConsId \rightarrow ConsName$
8. $HospNo \rightarrow HospName$
9. $ConsId, ApptDate \rightarrow HospNo$

Question:

What “fully” means in functional dependencies?

BOYCE-CODD NORMAL FORM

BCNF – a form where, within a relation, all determinants are candidate keys

Being in BCNF, the semantics of the SVFs have been attended to by means of a representation involving candidate keys, more than that, it recognizes implicitly the relationship between relations, via primary key/foreign key mechanism.

The ExtendedPatient relation

ExtendedPatient (PatientId, PatientName, GPId, GPName, ApptDate, ApptTime, ConsId, ConsName, HospNo, HospName)

PatientId	PatientName	GPId	GPName	ApptDate	ApptTime	ConsId	ConsName	HospNo	HospName
P01	Balthazar	G1	Leila	12/10	14:00	C1	Louella	H1	Faith
P01	Balthazar	G1	Leila	14/10	13:00	C2	Clementine	H2	Hope
P01	Balthazar	G1	Leila	09/09	09:00	C3	Nectarine	H3	Charity
P02	Cornelius	G2	Fitzdare	09/09	15:00	C1	Louella	H4	Flanders
P02	Cornelius	G2	Fitzdare	14/10	14:00	C2	Clementine	H2	Hope
P03	Samuel	G3	VonB	16/10	09:00	C2	Clementine	H3	Charity
P03	Samuel	G3	VonB	13/10	16:00	C1	Louella	H4	Flanders
P04	Darcy	G2	Fitzdare	12/10	13:00	C1	Louella	H1	Faith
P05	Schultz	G1	Leila	12/09	13:00	C3	Nectarine	H2	Hope
P05	Schultz	G1	Leila	13/09	14:00	C3	Nectarine	H2	Hope
P06	Sammel	G4	Von B	23/11	17:00	C4	Louella	H5	Charity

Non-loss Decomposition

No information has been lost or gained in the series of projections and joins. This is an important technique used in the process of normalization. Each of the individual moves, from one intermediate normal form to another intermediate normal form, can also be characterized as a non-loss decomposition.

join Patient and GeneralPractitioner where GPId = GPId giving A
join A and Appointment where PatientId = PatientId giving B
join B and consultant where ConsId = ConsId giving C
join C and ConsultantVisit where ConsId, ApptDate = ConsId, ApptDate giving D
join D and Hospital where HospNo = HospNo giving ExtendedPatient



project ExtendedPatient over GPId, GPName giving GeneralPractitioner
project ExtendedPatient over PatientId, PatientName, GPId, giving Patient
project ExtendedPatient over PatientId, ApptDate, ApptTime, ConsId giving Appointment
project ExtendedPatient over ConsId, ConsName giving Consultant
project ExtendedPatient over HospNo, HospName giving Hospital
project ExtendedPatient over ConsId, ApptDate, HospNo giving ConsultantVisit

Problems of not being in BCNF

See details in pages 53 - 58

- Instances of single-valued facts recorded in more than one place, e.g. general practitioner identified by g1 named Leila;
- Insertion anomalies *generally, new instances of those determinants are candidate keys may be inserted.*
- Deletion anomalies *generally, instances of those SVFs that are represented by means of candidate keys cannot be deleted without deleting instances of other SVFs; instances of those SVFs that are not represented by means of candidate keys can be deleted without deleting instances of other SVFs.*
- Amendment anomalies *generally, instances of those SVFs represented by means of candidate keys can be amended by the amendment of a single tuple;*

Block IV - Development of Database Systems

Instances of those SVFs that are not represented by means of candidate keys require the amendment of many tuples.

NORMALIZATION

NORMALIZATION: Non-BCNF \rightarrow 1NF \rightarrow 2NF \rightarrow 3NF \rightarrow BCNF

The move to BCNF progressively rearranges the representation of FDs, by means of non-loss decomposition, until all FDs have candidate keys as determinants.

Derivation by Transitivity

If $A \rightarrow B$ and $B \rightarrow C$, then $A \rightarrow C$

E.g. $\text{ConsId}, \text{ApptDate} \rightarrow \text{HospNo}$ and $\text{HospNo} \rightarrow \text{HospName} \Rightarrow \text{ConsId}, \text{ApptDate} \rightarrow \text{HospName}$

Derivation by Augment

If, in a relation R, the FD: $A \rightarrow B$ holds then the FD: $A, Z \rightarrow B, Z$ also holds, where Z is an (composite) attribute of the relation R. E.g. $\text{PatientId}, \text{ApptDate} \rightarrow \text{ConsId}$ (Augmented with ApptDate) $\Rightarrow \text{PatientId}, \text{ApptDate} \rightarrow \text{ConsId}, \text{ApptDate}$

FDs that hold for in relation ExtendedPatient

FD1	$\text{GDIId} \rightarrow \text{GPName}$	
FD2	$\text{PatientId} \rightarrow \text{PatientName}$	
FD3	$\text{PatientId} \rightarrow \text{GPId}$	
FD4	$\text{PatientId}, \text{ApptDate} \rightarrow \text{ApptTime}$	
FD5	$\text{PatientId}, \text{ApptDate} \rightarrow \text{ConsId}$	
FD6	$\text{ConsId}, \text{ApptDate}, \text{ApptTime} \rightarrow \text{PatientId}$	
FD7	$\text{ConsId} \rightarrow \text{ConsName}$	
FD8	$\text{HospNo} \rightarrow \text{HospName}$	
FD9	$\text{ConsId}, \text{ApptDate} \rightarrow \text{HospNo}$	
	<i>[Derivation from the original set]</i>	
FD10	$\text{PatientId} \rightarrow \text{GPName}$	(transitivity: FD3 & FD1)
FD11	$\text{ConsId}, \text{ApptDate} \rightarrow \text{HospName}$	(transitivity: FD9 & FD8)
FD12	$\text{ConsId}, \text{ApptDate}, \text{ApptTime} \rightarrow \text{PatientName}$	(transitivity: FD6 & FD2)
FD13	$\text{ConsId}, \text{ApptDate}, \text{ApptTime} \rightarrow \text{GPName}$	(transitivity: FD6 & FD3)
FD14	$\text{ConsId}, \text{ApptDate}, \text{ApptTime} \rightarrow \text{GPName}$	(transitivity: FD13 & FD1 or FD6 & FD10)
FD15	$\text{PatientId}, \text{ApptDate} \rightarrow \text{HospNo}$	(augmentation: FD5 with ApptDate gives $\text{PatientId}, \text{ApptDate} \rightarrow \text{ConsId}, \text{ApptDate}$, with FD9 give this derivation by transitivity)

[HospName and ConsName must be f.d. on the primary key, so there must be some FDs missing]

FD16	$\text{PatientId}, \text{ApptDate} \rightarrow \text{HospName}$	(transitivity: FD15 & FD8)
FD17	$\text{PatientId}, \text{ApptDate} \rightarrow \text{ConsName}$	(transitivity: FD5 & FD7)

First Normal Form (1NF)

A relation is in first normal form if and only if every non-primary-key attribute is functionally dependent on the primary key.

Not in 1NF: Grades (StudentId, CourseCode, (TMANo, Grade))

Any given occurrence of Grades, identified by a (StudentId, CourseCode) value, may have many (TMANo, Grade) values; that is (TMANo, Grade) is a repeating group.

Block IV - Development of Database Systems

1NF: Grades (StudentId, CourseCode, TMANo, Grade)

ExtendedPatient (PatientId, patientName, GPId, GPName, ApptDate, ApptTime, ConsId, ConsName, HospNo, HospName)

Second Normal Form (2NF)

A relation is in second normal form if and only if it is in 1NF and every non-primary-key attribute is **fully** functionally dependent on the primary key.

Offending FDs to 2NF:

FD2 PatientId \rightarrow PatientName

FD3 PatientId \rightarrow GPId

FD10 PatientId \rightarrow GPName

Putting into 2NF:

To remove the 'offending' FDs into a new, separate relation by a series of projections that give a non-loss decomposition:

- for each different left-hand side occurring in an 'offending' FDs projecting the original relation over the left-hand side and its corresponding right-hand side attribute(s);
- projecting the original relation over all its attributes save those which form the right-hand side(s) of the offending FD(s).

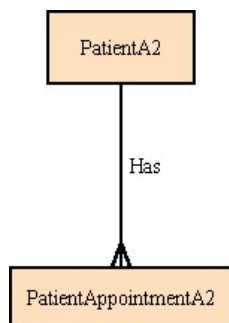
project ExtendedPatientA **over** PatientId, PatientName, GPId, GPName **giving** PatientA2

project ExtendedPatientA **over** PatientId, ApptDate, ApptTime, ConsId, ConsName, HospNo, HospName **giving** PatientAppointmentA2

PatientA2 (PatientId, PatientName, GPId, GPName)

PatientAppointment2 (PatientId, ApptDate, ApptTime, ConsId, ConsName, HospNo, HospName)

An important result of this move to 2NF is that any relation produced in step a) is related to the relation produced in step b) by means of the primary key/foreign key mechanism.



Third Normal form (3NF)

A relation is in third normal form if and only if it is in 2NF and every non-primary-key attribute is non-transitively dependent on the primary.

Transitive Dependency:

Whenever $R.A \rightarrow R.B$ and $R.B \rightarrow R.C$ hold, then the transitive dependency $R.A \rightarrow R.C$ also holds providing neither $R.B \rightarrow R.A$ nor $R.C \rightarrow R.A$ holds. We say C is transitively dependent on attribute A via attribute B.

e.g. GPName is transitively dependent on the primary key PatientId via GPId (Exercise 3.42)

ConsName is transitivity dependent on the primary key (PatientId, ApptDate) via HospNo (Exercise 3.43)

Block IV - Development of Database Systems

Offending FDs to 3NF:

The FDs that have been identified as the R.,B → R.C in the definition of transitive dependency:

FD1 GPID → GPName for patientA2

FD7 ConsId → ConsName for PatientAppointmentA2

FD8 HospNo → HospName for PatientAppointmentA2

Putting into 3NF:

Project PatientA2 **over** GPID, GPName **giving** GeneralPractitionerA3

Project PatientA2 **over** PatientId, PatientName, GPID **giving** PatientA3

GeneralPractitionerA3 (GPID, GPName)

PatientA3 (PatientId, PatientName, GPID)

Project PatientAppointmentA2 **over** ConsId, ConsName **giving** ConsultantA3

Project PatientAppointmentA2 **over** HospNo, HospName **giving** HospitalA3

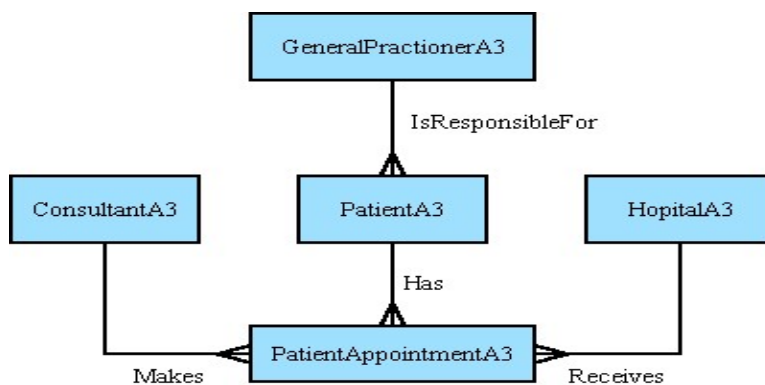
Project PatientAppointmentA2 **over** PatientId, ApptDate, ApptTime, ConsId, HospNo **giving** HospitalA3

ConsultantA3 (ConsId, ConsName)

HospitalA3 (HospNo, HospName)

PatientAppointmentA3 (PatientId, ApptDate, ApptTime, ConsId, HospNo)

As before, the relationships can be inferred from the primary key / foreign key mechanism.



Boyce-Codd Normal Form (BCNF)

A relation is in BCNF if and only if all determinants are candidate keys.

Offending FD to BCNF:

The FDs whose left-hand sides are determinants which are not candidate keys. The only 3NF relation that is not in BCNF is PatientAppointmentA3, by virtue of (ConsId, ApptDate) being a determinant but not a candidate key.

ConsId, ApptDate → HospNo

Project PatientAppointmentA3 **over** ConsId, AppDate, HospNo **giving** ConsultantVisitABCNF

Project PatientAppointmentA3 **over** PatientId, ApptDate, ApptTime, ConsId **giving** PatientAppointmentABCNF

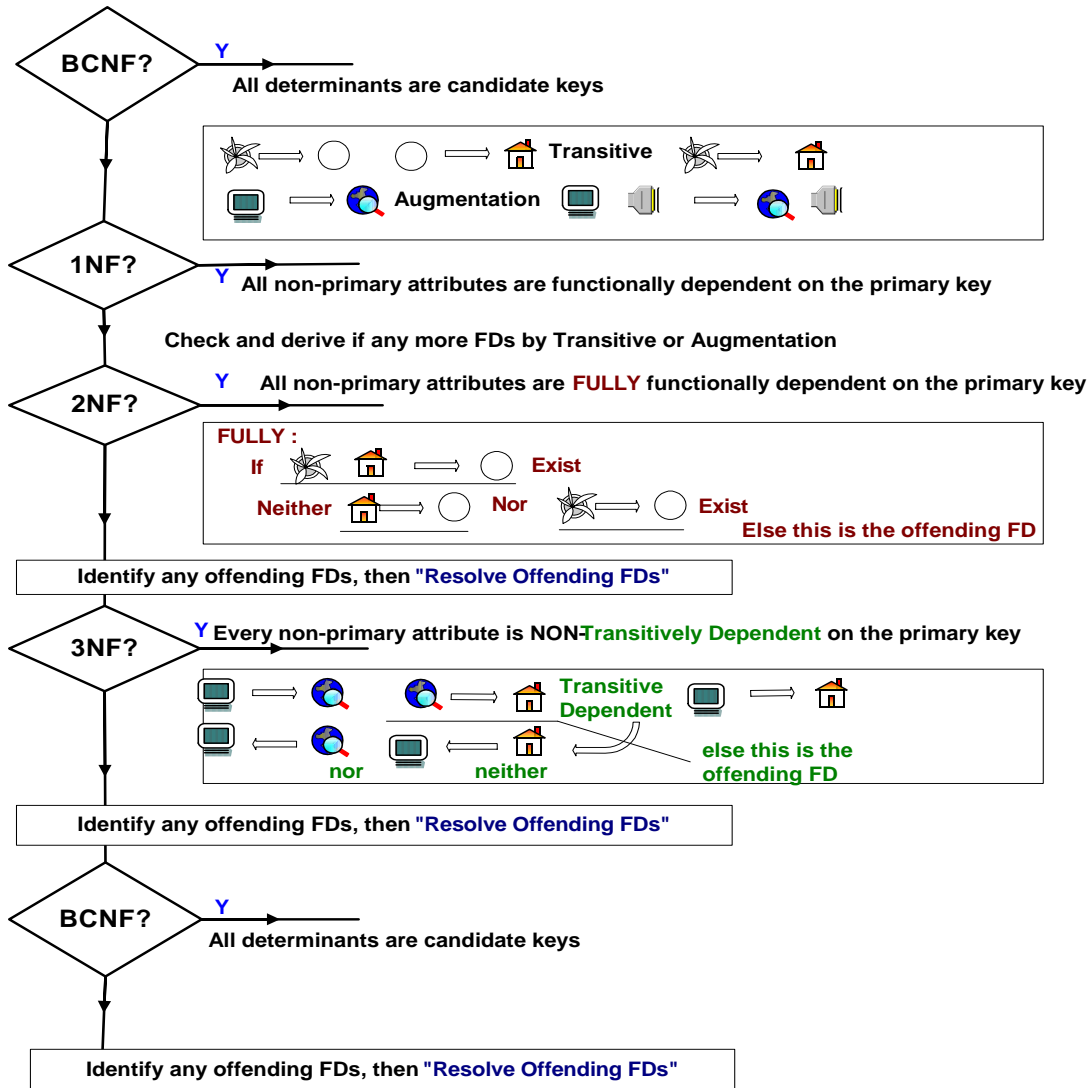
PatientAppointmentABCNF (PatientId, ApptDate, ApptTime, ConsId)

ConsultantVisitABCNF (ConsId, ApptDate, HospNo)

Question:

What is the ER diagram corresponding to this BCNF?

NORMALISATION - THE METHODOLOGY



Resolve Offending FDs

