Assignment 8: Bernstein Algorithm and BCNF

Jose Manipon, 500906166

Amiel Castillo, 500883624

Anthony Trinh, 500831193

Patient Table: (Healthcard #, First name, Last name, DOB, Gender, Address, Phone, Email, Hospital ID)

HEALTHC...	F_NAME	L_NAME	DATE_OF...	GENDER	AGE	ADDRESS...	ADDRESS...	ADDRESS...	ADDRESS...	ADDRESS...	ADDRESS...	PHONE	EMAIL
1234567890	Ann	Smith	97-01-28	F	23	1290 Ba...	(null)	Toronto	ON	M1B2X4	CAD	6471234567	ann.smi...
1234567891	Bob	Brown	00-02-21	M	20	11 Clif...	(null)	Toronto	ON	M2C1L5	CAD	6471234568	bob.bro...
1234567892	Carl	Jones	99-07-11	M	21	103 Roa...	(null)	Oakville	ON	L6L2X6	CAD	4162578564	carl.jo...
1234567893	Dan	Miller	98-04-07	M	22	56 Stee...	(null)	Pickering	ON	L1V0A1	CAD	4162874587	dan.mil...
1234567894	Eve	Williams	98-01-22	F	22	23 Gree...	(null)	Markham	ON	L1C3P2	CAD	6478985674	eve.wil...
1234567895	Stan	Murphy	99-01-29	M	21	1111 Ba...	(null)	Toronto	ON	M1B2X4	CAD	6471212311	stan.mu...
1234567896	Alice	Yum	20-03-21	F	1	11 Stee...	(null)	Toronto	ON	M2C3CD	CAD	6471234231	yum.ali...
1234567897	Steve	Jan	20-07-12	M	100	2 Ins S...	(null)	Oakville	ON	L6L2X	CAD	4162532313	steve.j...
1234567898	Mike	Lanny	90-08-07	M	30	55 Stee...	(null)	Pickering	ON	L1DED	CAD	4162875561	mike.la...
1234567899	Camy	Hun	32-01-30	F	88	2 Hopki...	(null)	Markham	ON	L1QD2P	CAD	6478983164	camy.hu...
1234567900	Anthony	Trinh	99-03-12	M	21	135 Hil...	(null)	Mississ...	ON	L5B3Z2	CAD	6478983164	anthony...

Step 1: Determine functional dependencies

Health card # → First name
Health card # → Last name
Health card # → Address
Health card # → Phone
Health card # → Email
Health card # → DOB
Health card # → Age
First name, Last name → Health card #
Last name, Health card # → First name
Last name, Health card # → Age
Last name, Health card # → DOB
Last name, Health card # → Address
Last name, Health card # → Phone
Last name, Health card # → Email

Step 2: Find redundancies

We have to get rid of the redundant dependencies

Health card # → First name
Health card # → Last name
Health card # → Address
Health card # → Phone
Health card # → Email
Health card # → DOB
Health card # → Age
First name, Last name → Health card #
Last name, Health card # → First name
Last name, Health card # → Age
Last name, Health card # → DOB
Last name, Health card # → Address
Last name, Health card # → Phone
Last name, Health card # → Email

The last six FD's are redundant, so we can get rid of them.
There are no partial dependencies since in the dependency First name, Last name → Health card #, first names and last names are not unique, so they cannot alone determine the health card #.

Step 3: Find Keys

We have two candidate keys.
{Health Card #}
{First name, Last name}
{DOB}

Step 4: Find relations
R1(Health card #, First name, Last name, DOB, Age1, Address, Phone, Email)
R2(First name, Last name, Health card #)

Since R2 is a subset of R1, we can eliminate R2 and thus our final schema is R1.

_____

Medicine Table: (Medicine ID, Dosage, IUPAC name, Generic name, Inventory, Price, Expiration date, Manufacturer, Hospital ID)

	MEDICINE_ID	DOSAGE	IUPAC_NAME	GENERIC_NAME	INVENTORY	PRICE	EXPIRATION_DATE	MANUFACTURER	HOSPITAL_ID
1	191919	20	N-(4-hydroxyphenyl)acetamide	acetaminophen	500	5	30-01-01	Tylenol	12
2	191920	100	8-Chloro-1-methyl-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepine	alprazolam	200	50	20-12-25	Tylenol	12
3	191921	10	(S,S)-2-methylamino-1-phenylpropan-1-ol	pseudoephedrine	100	50	30-10-22	Johnson and Johnson	12
4	191922	50	(S)-2-Amino-3-[4-(4-hydroxy-3,5-diiodophenoxy)-3,5-diiodophenyl]propanoic acid	synthroid	50	25.5	29-06-01	Johnson and Johnson	12
5	191923	50	(3R,5R)-7-[2-(4-Fluorophenyl)-3-phenyl-4-(phenylcarbamoyl)-5-propan-2-ylpyrrol-1-yl]-3,5-dihydroxyheptanoic acid	lipitor	200	60	20-12-01	Johnson and Johnson	12
6	191924	75	(S,S)-2-methylamino-1-phenylpropan-1-ol	pseudoephedrine	100	48.75	29-05-02	Sinopharm	12
7	191925	125	(3R,5R)-7-[2-(4-Fluorophenyl)-3-phenyl-4-(phenylcarbamoyl)-5-propan-2-ylpyrrol-1-yl]-3,5-dihydroxyheptanoic acid	lipitor	1200	62.5	27-12-01	Sinopharm	12

Step 1: Determine functional dependencies

Medicine ID → Dosage, IUPAC name, Generic name, Inventory, Price, Expiration date, Manufacturer,
                    Hospital ID
Hospital ID → Inventory, Price
IUPAC name → Generic name
Manufacturer, Hospital ID, Dosage → Inventory, Price

Step 2: Find redundancies

Manufacturer, Hospital ID, Dosage → Inventory, Price and *Hospital ID → Inventory, Price* are redundant, so we will remove the latter.

Step 3: Find keys

Candidate keys are
{Medicine ID}
{Hospital ID}
{IUPAC Name}

Step 4: Find relations

R1(Medicine ID,Dosage, IUPAC name, Generic name, Inventory, Price, Expiration date, Manufacturer,
      Hospital ID)
R2(Hospital ID, Inventory, Price)
R3(IUPAC Name, Generic Name)

Since R2, R3 are subsets of R1, our final schema is R1.

Appointment Table: (Appointment ID, Appointment date, Appointment time, Room #, Nurse ID, Healthcard #, Doctor ID, Hospital ID)

APPOINTMENT_ID	APPOINTMENT_DATE	APPOINTMENT_TIME	ROOM_NO	NURSE_ID	HEALTHCARD_NO	DOCTOR_ID	HOSPITAL_ID
10031	20-12-12	1300	1	555879	1234567890	141524	123
10012	20-10-20	1400	2	555880	1234567897	141526	124
10013	20-10-22	1500	3	555881	1234567892	141525	125

Step 1: Determine functional dependencies

Appointment ID → Appointment date, Appointment time, Room #, Nurse ID, Healthcard #, Doctor ID, Hospital ID
Appointment date, Appointment time, Healthcard # → Appointment ID
Appointment date, Appointment time, Doctor ID → Appointment ID
Appointment date, Appointment time, Nurse ID → Appointment ID

Step 2: Find redundancies

Appointment date, Appointment time, Healthcard # → Appointment ID, Appointment date, Appointment time, Doctor ID → Appointment ID, and Appointment date, Appointment time, Nurse ID → Appointment ID are redundant, so we will remove the latter.

Step 3: Find keys

Candidate keys are
{ Appointment ID }
{ Appointment date, Appointment time, Healthcard # }
{ Appointment date, Appointment time, Doctor ID }
{ Appointment date, Appointment time, Nurse ID }

Step 4: Find relations

R1(Appointment ID, Appointment date, Appointment time, Room #, Nurse ID, Healthcard #, Doctor ID, Hospital ID)
R2(Appointment date, Appointment time, Healthcard # → Appointment ID)
R3(Appointment date, Appointment time, Doctor ID → Appointment ID)
R4(Appointment date, Appointment time, Nurse ID → Appointment ID)

Since R2, R3, R4 are subsets of R1, we can eliminate R2, R3, R4.
Our final schema is R1.

_____

Hospital Table: (Hospital ID Hospital Name, Address, Phone)

	HOSPITAL_ID	HOSPITAL_NAME	ADDRESS_STREET	ADDRESS_STREET2	ADDRESS_CITY	ADDRESS_PR...	ADDRESS_POSTAL...	ADDRESS_COUNTRY	PHONE
1	124	Michael Garron Hospital	825 Coxwell Ave	(null)	East York	ON	M4C3E7	CAD	4167893127
2	125	Toronto Western Hospital	399 Bathurst St	(null)	Toronto	ON	M5T2S8	CAD	4167898713
3	123	Toronto General Hospital	200 Elizabeth St	(null)	Toronto	ON	M5G2C4	CAD	4167895297

Step 1: Determine the FD's that violate BCNF

Hospital ID → Hospital Name, Address, Phone

This is already in BCNF

_____

Employee Table: (Employee ID, First name, Last name, DOB, Gender, Age, Address, Phone, Email, Hospital ID)

EMPLOYEE_ID	F_NAME	L_NAME	DATE_OF_BIRTH	GENDER	AGE	ADDRESS_STREET	ADDRESS_STREET2	ADDRESS_CITY	ADDRESS_PROVINCE	ADDRESS_POSTALCODE	ADDRESS_COUNTRY	PHONE	EMAIL	HOSPITAL_ID
323952	Andrew	Musak	87-07-12	M	33	121 Collins Rd.	(null)	Toronto	ON	M1V1N3	CAD	4168794561	dr.andrew.musak33@gmail.com	123
323678	Mindy	Ramirez	90-10-10	F	30	126 Sisao St.	(null)	Toronto	ON	M1P4R2	CAD	6475213125	mindy.ramirez@gmail.com	124
323679	John	Cooper	89-08-10	M	27	B9 Prospect Street	(null)	Toronto	ON	M1PX13	CAD	6475213256	cooper.john@gmail.com	124
323680	Michelle	Stanley	91-10-06	F	30	202 Enble Street	(null)	Toronto	ON	M1D13C	CAD	6473125676	stanley.michelle@gmail.com	123
323681	Samuel	Honey	90-10-12	M	31	A4 Laven Street	(null)	Toronto	AL	M2X3B1	CAD	64753561234	samuel.honey@gmail.com	125
323682	Ruby	Yul	92-10-11	F	26	C6 Thorne Street	(null)	Toronto	ON	M2R32X	CAD	6475352132	ruby.yul@gmail.com	125
323683	Robert	Smith	82-12-15	M	38	123 Bay Rd.	(null)	Toronto	ON	M1V1N2	CAD	416279454	rsmith@gmail.com	123

Step 1: Determine the FD's that violate BCNF
Employee ID → First name, Last name, DOB, Gender, Age, Address, Phone, Email, Hospital ID
DOB → Age

The last FD violates BCNF

Step 2: Decompose the tables
Now we have…
R1(DOB,Age)
R2(Employee ID, First name, Last name, DOB, Gender, Address, Phone, Email, Hospital)

Now, both of these tables are in BCNF.

_____

Doctor Table: (Doctor ID, Doctor License Expiry, Employee ID)

DOCTOR_ID	DOCTORLICENSE_EXPIRY	EMPLOYEE_ID
141524	30-03-06	323952
141525	20-03-12	323678
141526	12-04-06	323679

Step 1: Determine the FD's that violate BCNF

Doctor ID → Doctor License Expiry, Employee ID

No FD violates BCNF.

Nurse Table: (Nurse ID, Nurse License Expiry, Employee ID)

NURSE_ID	NURSELICENSE_EXPIRY	EMPLOYEE_ID
555879	25-09-10	323680
555880	22-09-11	323681
555881	23-10-10	323682

Step 1: Determine the FD's that violate BCNF

Nurse ID → NurseLicense Expiry, Employee ID

No FD violates BCNF.

_____

Patient Table: (Healthcard #, First name, Last name, DOB, Gender, Address, Phone, Email, Hospital ID)

HEALTHCA...	F_...	L_NAME	DATE_OF_BIRTH	GENDER	AGE	ADDRESS_STREET	ADDRESS_STREET2	ADDRESS_CITY	ADDRESS_PROVINCE	ADDRESS_POSTALCODE	ADDRESS_COUNTRY	PHONE	EMAIL	HOSPITAL_ID
1234567890	Ann	Smith	97-01-28	F	22	1290 Bayview Rd.	(null)	Toronto	ON	M1B2X4	CAD	6471234567	ann.smith97@gmail.com	123
1234567891	Bob	Brown	00-02-21	M	20	11 Clifton St.	(null)	Toronto	ON	M2C1L5	CAD	6471234568	bob.brown00@hotmail.com	124
1234567892	Carl	Jones	99-07-11	M	21	103 Roadhouse Rd.	(null)	Oakville	ON	L6L2X6	CAD	4162578564	carl.jones99@gmail.com	125
1234567893	Dan	Miller	98-04-07	M	22	56 Steeling Ave.	(null)	Pickering	ON	L1V0A1	CAD	4162874587	dan.miller98@gmail.com	125
1234567894	Eve	Williams	98-01-22	F	22	23 Greentint Cres.	(null)	Markham	ON	L1C3P2	CAD	6478985674	eve.williams22@hotmail.com	124
1234567895	Stan	Murphy	99-01-29	M	30	111 Bayview Rd.	(null)	Toronto	ON	M1B2X4	CAD	6471212311	stan.murphy@gmail.com	123
1234567896	Alice	Yum	20-03-21	F	21	11 Steeling St.	(null)	Toronto	ON	M2C3CD	CAD	6471234231	yum.alice@hotmail.com	123
1234567897	Steve	Jan	20-07-12	M	23	2 Ins Street	(null)	Oakville	ON	L6L2X	CAD	4162532313	steve.jan@gmail.com	124
1234567898	Mike	Lanny	90-08-07	M	40	55 Steeling Ave.	(null)	Pickering	ON	L1DED	CAD	4162875561	mike.lanny@gmail.com	125
1234567899	Camy	Hun	32-01-30	F	30	2 Hopkins Street	(null)	Markham	ON	L1QD2P	CAD	6478983164	camy.hun@hotmail.com	123

Step 1: Determine the FD's that violate BCNF

Healthcard # → First name, Last name, DOB, Gender, Age, Address, Phone, Email, Hospital ID)

Step 2: Decompose the tables

Now we have…
R1(DOB,Age)
R2(Heatlhcard #, First name, Last name, DOB, Gender, Address, Phone, Email, Hospital ID)

Now, both of these tables are in BCNF.

_____

Appointment Table: (Appointment ID, Appointment date, Appointment time, Room #, Nurse ID, Healthcard #, Doctor ID, Hospital ID)

APPOINTMENT_ID	APPOINTMENT_DATE	APPOINTMENT_TIME	ROOM_NO	NURSE_ID	HEALTHCARD_NO	DOCTOR_ID	HOSPITAL_ID
10031	20-12-12	1300	1	555879	1234567890	141524	123
10012	20-10-20	1400	2	555880	1234567897	141526	124
10013	20-10-22	1500	3	555881	1234567892	141525	125

Step 1: Determine the FD's that violate BCNF

Appointment ID → Appointment date, Appointment time, Room #, Nurse ID, Healthcard #, Doctor ID, Hospital ID

No FD violates BCNF.

Diagnosis Table: (Diagnosis ID, Results, Appointment ID)

	DIAGNOSIS_ID	RESULTS	APPOINTMENT_ID
1	123456	Stomach cancer	10012
2	123457	Covid-19	10013

Step 1: Determine the FD's that violate BCNF

Diagnosis ID → Results, Appointment ID

No FD violates BCNF.

Invoice Table: (Invoice #, Medicine ID, Date Issued, Amount owed, appointment ID)

INVOICE_NO	MEDICINE_ID	DATE_ISSUED	AMOUNT_OWED	APPOINTMENT_ID
111112	3	20-10-20	100	10012
111113	3	20-10-22	10	10013

Step 1: Determine the FD's that violate BCNF

Invoice # → Medicine ID, Date issued, amount owed, appointment ID

No FD violates BCNF.

Prescription Table: (Prescription #, Appointment ID, Medicine ID, Diagnosis ID)

PRESCRIPTION_NO	DATE_ISSUED	APPOINTMENT_ID	MEDICINE_ID	DIAGNOSIS_ID
1231231	20-10-20	10012	191919	123456
1231232	20-10-22	10013	191919	123457
1231233	20-10-22	10013	191920	123457

Step 1: Determine the FD's that violate BCNF
Prescription # → Date Issued, Medicine ID, Diagnosis ID, Appointment ID
No FD violates BCNF.

Medical History Table: (Healthcard #, Appointment ID, Diagnosis ID, Medical Desc)

HEALTHCARD_NO	APPOINTMENT_ID	DIAGNOSIS_ID	MEDICAL_DESC
1234567897	10012	123456	This guy has stomach cancer, he is allergic to tylonel so prescribe him something else
1234567892	10013	123457	Make sure they stay at home for two weeks

Step 1: Determine the FD's that violate BCNF
Healthcard #, Appointment ID, Diagnosis ID → Medical Desc
No FD violates BCNF.

Medicine Table: (Medicine ID, Dosage, IUPAC Name, Generic Name, Inventory, Price, Expiration date, Manufacturer, Hospital ID)

	MEDICINE_ID	DOSAGE	IUPAC_NAME	GENERIC_NAME	INVENTORY	PRICE	EXPIRATION_DATE	MANUFACTURER	HOSPITAL
1	191919	20	N-(4-hydroxyphenyl)acetamide	acetaminophen	500	5	30-01-01	Tylenol	
2	191920	100	8-Chloro-1-methyl-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepine	alprazolam	200	50	20-12-25	Tylenol	
3	191921	10	(S,S)-2-methylamino-1-phenylpropan-1-ol	pseudoephedrine	100	50	30-10-22	Johnson and Johnson	
4	191922	50	(S)-2-Amino-3-[4-(4-hydroxy-3,5-diiodophenoxy)-3,5-diiodophenyl]propanoic acid	synthroid	50	25.5	29-06-01	Johnson and Johnson	
5	191923	50	(3R,5R)-7-[2-(4-Fluorophenyl)-3-phenyl-4-(phenylcarbamoyl)-5-propan-2-ylpyrrol-1-yl]-3,5-dihydroxyheptanoic acid	lipitor	200	60	20-12-01	Johnson and Johnson	
6	191924	75	(S,S)-2-methylamino-1-phenylpropan-1-ol	pseudoephedrine	100	48.75	29-05-02	Sinopharm	
7	191925	125	(3R,5R)-7-[2-(4-Fluorophenyl)-3-phenyl-4-(phenylcarbamoyl)-5-propan-2-ylpyrrol-1-yl]-3,5-dihydroxyheptanoic acid	lipitor	1200	62.5	27-12-01	Sinopharm	

Step 1: Determine the FD's that violate BCNF

Medicine ID → Dosage, IUPAC Name, Generic name, Inventory, Price, Expiration date, Manufacturer, Hospital ID

Hospital ID → Inventory, Price

IUPAC Name → Generic Name

The latter two FD's violate BCNF

Step 2: Decompose the Tables

R1(Hospital ID, Inventory, Price)

R2(Medicine ID, Dosage, IUPAC Name, Generic Name, Expiration date, manufacturer, Hospital ID)

R3(IUPAC Name, generic name)

R4(Medicine ID, Dosage, IUPAC Name, Expiration date, Manufacturer, Hospital ID)

So finally, the tables decomposed into BCNF are R1, R3, R4. We omit R2 since it is a subset of R4.

In conclusion, all tables are now BCNF/3NF since all of their attributes are independent of one another, and depend on ONLY on their respective tables Primary Key.