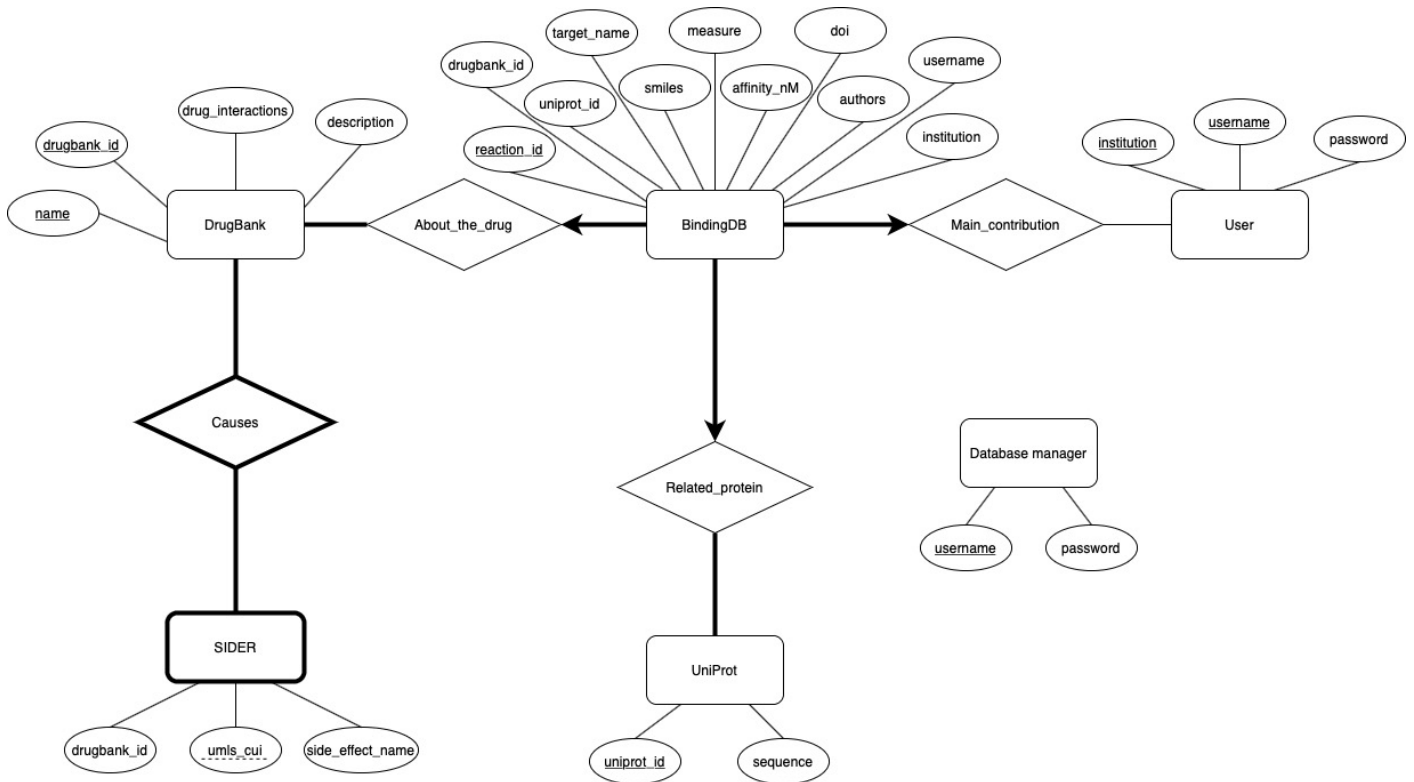# 2021 Spring - CMPE321 Project 2 Report
Mehmet Selahaddin Şentop - 2018400105
Ömer Faruk Süve - 2017400147

Part 1

Part 2

Relational schemas and tables are as follows:

- **User**(*username:* **string**, *institution:* **string**, *password:* **string**)

- **Database manager**(*username:* **string**, *password:* **string**)

- **DrugBank**(*drugbank_id:* **string**, *name:* **string**, *drug_interactions:* **list<string>**, *description:* **string**)

- **SIDER**(*ulms_cui:* **string**, *drugbank_id:* **string**, *side_effect_name:* **string**)

- **BindingDB**(*reaction_id:* **integer**, *drugbank_id:* **string**, *uniprot_id:* **string**, *target_name:* **string**, *smiles:* **string**, *measure:* **string**, *affinity_nM:* **real**, *doi:* **string**, *authors:* **string**, *username:* **string**, *institution:* **string**)

- **UniProt**(*uniprot_id:* **string**, *sequence:* **string**)

User(username: string, institution: string, password: string)		
username	institution	password
Abramovitz	Merck Frosst Centre for Therapeutic Research	abramovitz3232
Afzelius	AstraZeneca R&D	afzelius4343
Ahlin	Uppsala University	ahlin5151

Database manager(username: string, password: string)	
username	password
selen.parlar	selen.parlar
riza.ozcelik	riza.ozcelik0
arzucan_ozgur	arzucan_135

DrugBank(drugbank_id: string, name: string, drug_interactions: list<string>, description: string)			
drugbank_id	name	drug_interactions	description
DB00459	Acitretin	['DB00304', 'DB00367', 'DB00162']	An oral retinoid effective ...
DB00523	Alitretinoin	['DB00304', 'DB00367', 'DB00162']	An important regulator of ...
DB00964	Apraclonidine	['DB06237']	Apraclonidine, also known as ...

SIDER(ulms_cui: string, drugbank_id: string, side_effect_name: string)		
ulms_cui	drugbank_id	side_effect_name
C0000737	DB00459	Abdominal pain
C0702166	DB00459	Acne
C0155626	DB00459	Acute myocardial infarction

BindingDB(reaction_id: integer, drugbank_id: string, uniprot_id: string, target_name: string, smiles: string, measure: string, affinity_nM: real, doi: string, authors: string, username: string, institution: string)										
reaction_id	drugbank_id	uniprot_id	target_name	smiles	measure	affinity_nM	doi	authors	username	institution
50876947	DB00925	O15245	Adrenergic Alpha	CC(CO...	IC50	15100	10.1021/...	Ahlin, G;...	Ahlin	Uppsala University
50876963	DB01162	O15245	Adrenergic Alpha	COc1c...	IC50	23700	10.1021/...	Ahlin, G; ...	Ahlin	Uppsala University
50739658	DB01132	O95342	Bile salt export pump	CCc1c...	IC50	400	10.1002/...	Aleo, MD;	Aleo	Pfizer Inc

UniProt(uniprot_id: string, sequence: string)	
uniprot_id	sequence
P04278	MESRG...
Q99808	MTTSH...
P51574	MGMSK...

Part 3

Now we will go through all the relations and check their suitability of BCNF.

1- User(username(U): **string**, institution(I): **string**, password(P): **string**):
K = {U, I} is the primary key.
Non-trivial functional dependencies:
K → P
Since k is a key for the entity, the requirements for BCNF are met.

2- Database manager(username(U): **string**, password(P): **string**):
U is the primary key.
Non-trivial functional dependencies:
U → P
Since k is a key for the entity, the requirements for BCNF are met.

3- DrugBank(drugbank_id(Did): **string**, drug_name(Dn): **string**, drug_interactions(Di): **list<string>**, drug_description(Des): **string**):
Did and Dn are both keys since any one of them uniquely determines the entity.

Non-trivial functional dependencies:

Did → DnDiDes

Dn → DidDiDes

Since Did and Dn are both keys for the entity, the requirements for BCNF are met.

4- SIDER(*ulms_cui(U):* **string**, *drugbank_id(D):* **string**, *side_effect_name(S):* **string**) :

K1 = {U, D} and K2 = {S, D} are both keys since they uniquely determine SIDER entities.

Non-trivial functional dependencies:

K1 → S

K2 → U

Since K1 and K2 are both keys for the entity, the requirements for BCNF are met.

5- BindingDB(*reaction_id(R):* **integer**, *drugbank_id(Did):* **string**, *uniprot_id(Uid):* **string**, *target_name(T):* **string**, *smiles(S):* **string**, *measure(M):* **string**, *affinity_nM(An):* **real**, *doi(D):* **string**, *authors(A):* **string**, *username(U):* **string**, *institution(I):* **string**):

R is the primary key.

Non-trivial functional dependencies:

R → DidUidTSMAnDAUI

Did → S

S → Did

D → A

Since Did, S and D are not superkeys, the requirements of BCNF are not met. Likewise, since D is not a superkey and A is not part of some key for the relation, the requirements of 3NF are also not met. We can decompose the relation into BCNF when the following functional dependencies hold:

R → DidUidTMAnDUI

Did → S

S → Did

D → A

Where the new and updated attributes and entities are (the parts that are newly added are written in <mark>green</mark>):

BindingDB(*reaction_id(R):* **integer**, *drugbank_id(Did):* **string**, *uniprot_id(Uid):* **string**, *target_name(T):* **string**, *measure(M):* **string**,*affinity_nM(An):* **real**, *doi(D):* **string**, *username(U):* **string**, *institution(I):* **string**)
Non-trivial functional dependencies:
R → DidUidTMAnDUI
Since R is a key, the requirements of BCNF are met.

Links(*doi(D):* **string**, *authors(A):* **string**)
D is a key for Links since it completely determines the entity.
Non-trivial functional dependencies:
D → A
Since D is a key, the requirements of BCNF are met.

DrugBank(*drugbank_id(Did):* **string**, *drug_name(Dn):* **string**, *drug_interactions(Di):* **list<string>**, *drug_description(Des):* **string,** *smiles(S):* **string**)
Did, Dn and S are all keys since any one of them uniquely determines the entity.
Non-trivial functional dependencies:
Did → DnDiDesS
Dn → DidDiDesS
S → DidDnDiDes
Since Did, Dn and S are all keys for the entity, the requirements for BCNF are met.

- Dependency preserving:

As seen above, the functional dependencies are preserved and distributed among entities. Hence the decomposition is dependency preserving.

- Lossless-join:

-The attributes are preserved, i.e. neither new attributes are introduced nor existing attributes are discarded. Instead, some attributes are shifted to other entities which preserves the union of attributes after decomposition.

-After the decomposition, the intersection of the remaining attributes in **BindingDB** with the removed attributes contains keys for entities **Links** (which is D) and **DrugBank** (which is S).

Because of the conditions above that are shown to hold, the decomposition is lossless-join.

6- UniProt(*uniprot_id(Uid): **string**, sequence(S): **string***):

Uid is the primary key.

Non-trivial functional dependencies:

Uid → S

Since Uid is a key for the entity, the requirements for BCNF are met.

Updated tables for updated and new entities are below:

BindingDB(*reaction_id:* **integer**, *drugbank_id:* **string**, *uniprot_id:* **string**, *target_name:* **string**, *measure:* **string**, *affinity_nM:* **real**, *doi:* **string**, *username:* **string**, *institution:* **string**)								
reaction_id	drugbank_id	uniprot_id	target_name	measure	affinity_nM	doi	username	institution
50876947	DB00925	O15245	Adrenergic Alpha	IC50	10.1021/...	Ahlin	Ahlin	Uppsala University
50876963	DB01162	O15245	Adrenergic Alpha	IC50	10.1021/...	Ahlin	Ahlin	Uppsala University
50739658	DB01132	O95342	Bile salt export pump	IC50	10.1002/...	Aleo	Aleo	Pfizer Inc

Links(*doi:* **string**, *authors:* **string**)	
doi	authors
10.1021/...	Ahlin, G;...
10.1021/...	Ahlin, G; ...
10.1002/...	Aleo, MD;

DrugBank(*drugbank_id:* **string**, *drug_name:* **string**, *drug_interactions:* **list<string>**, *drug_description:* **string**, *smiles:* **string**)				
drugbank_id	drug_name	drug_interactions	drug_description	smiles
DB00459	Acitretin	['DB00304', 'DB00367', 'DB00162']	An oral retinoid effective ...	CC(CO...
DB00523	Alitretinoin	['DB00304', 'DB00367', 'DB00162']	An important regulator of ...	COc1c...
DB00964	Apraclonidine	['DB06237']	Apraclonidine, also known as ...	CCc1c...