Data Normalization

Intercollege CP Tournament Analysis Database

Data and Applications Project: Phase 3

Amul Agrawal (2019101113)
Mihir Bani (2019113003)
Snehal Kumar (2019101003)

---

1. After mapping ER to relational model

Note: Relations/Entities are in **bold,** relationships in underline and attributes in *italics*.

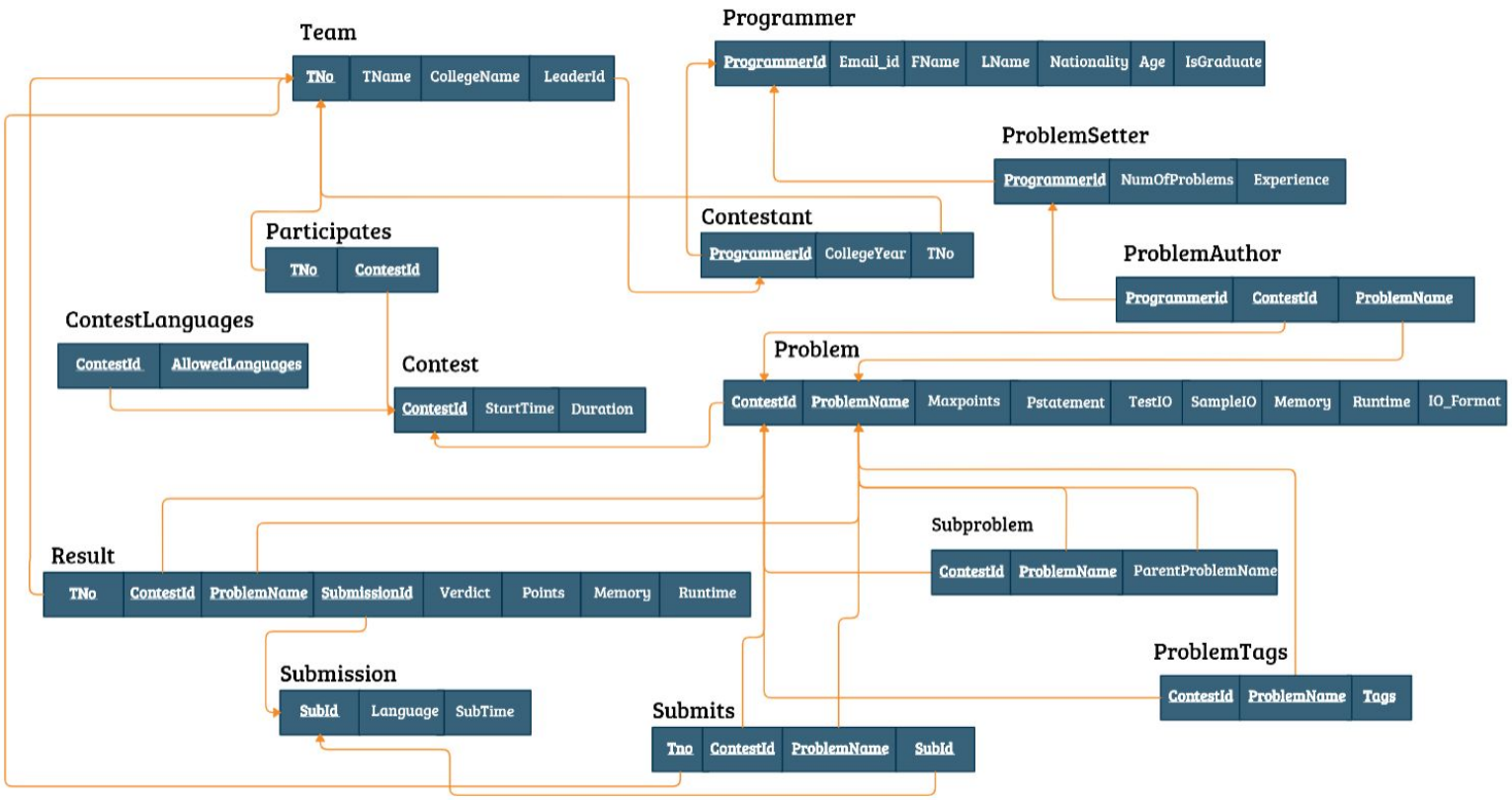
Entity Transitions: All entities got converted to a unique Relation.

Relationship Transitions:

- LED_BY is added to **Team** through *LeaderId.*
- HAS is added to **Contestant** through *Tno.*
- PARTICIPATES is converted to **Participates.**
- CONTAINS is added to **Problem** through *ContestID*.
- WRITES is converted to **ProblemAuthor**.
- IS_SUBPART is converted to **Subproblem**.
- SUBMITS is converted to **Submits.**
- The multivalued attributes *Tags (*from **Problem***)* and *AllowedLanguages* (from **Contest**) are represented as a separate relation named **ProblemTags** and **ContestLanguages.**
- We removed the attribute *TotalScore*, *Rank* from **Participates**, as these were derived attributes and will be calculated during the query.

  Clarification:
  Although *TotalScore* (from **Participates)** was saved as a normal attribute due to increased time complexity since *Rank* is a derived attribute on *TotalScore,* after clarification we decided to keep it as derived and remove it from the Relational Model.

Note (in picture): Primary Keys are represented in bolds. The lines in orange represent relation mappings, blue represent functional dependencies.

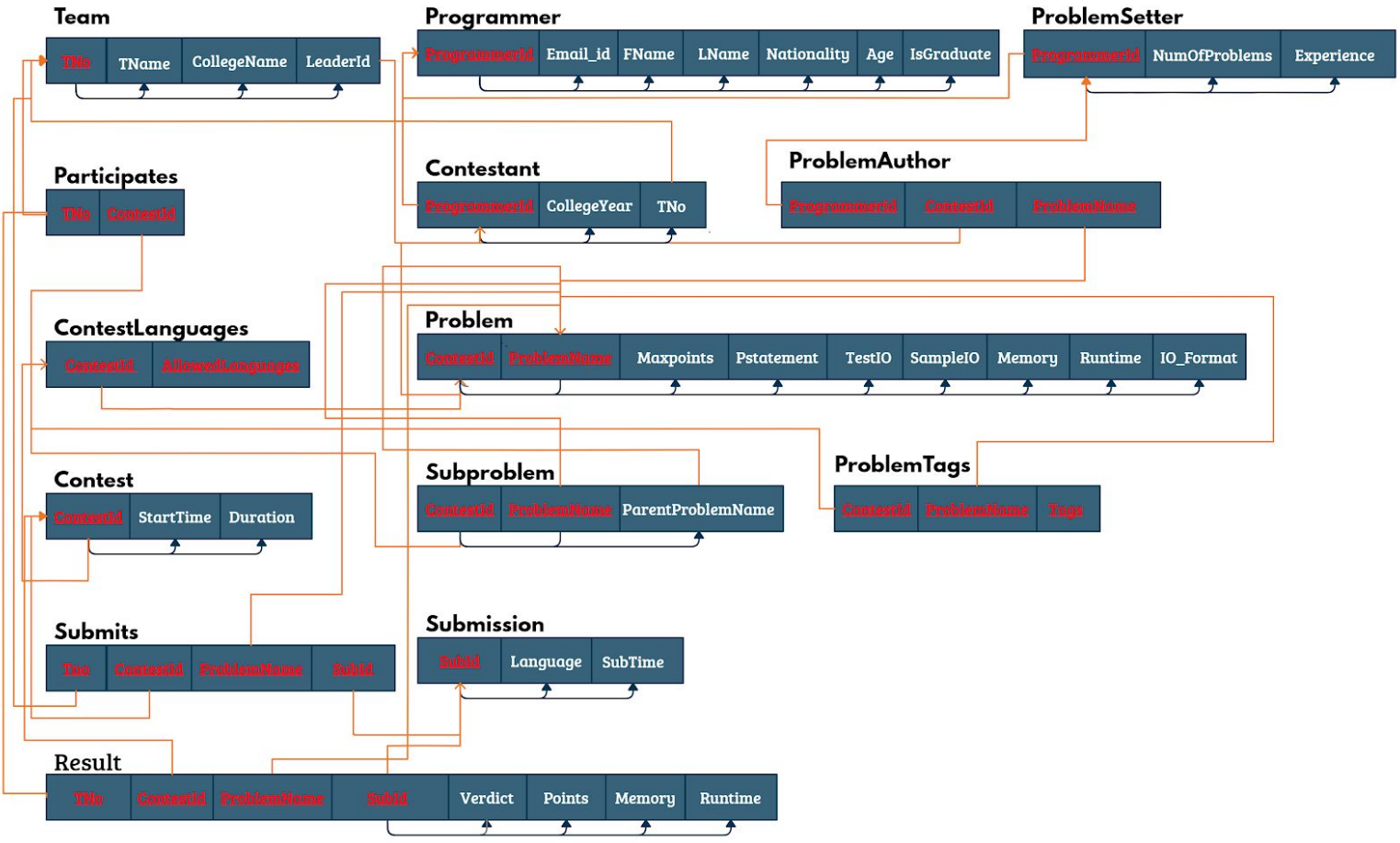


2. Relational model after conversion to 1NF

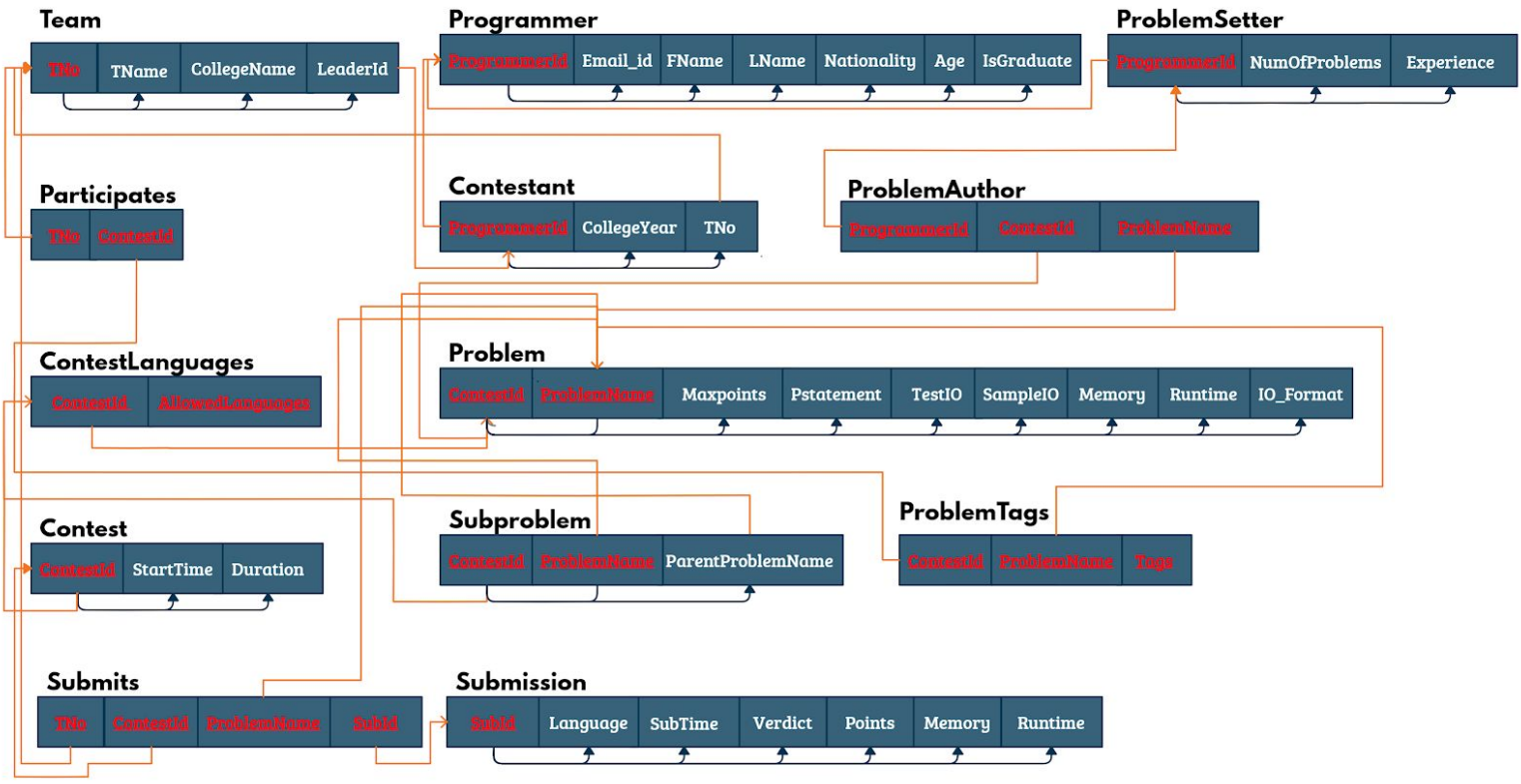
Transition:

- Mapped the relations between entities using the keys in validation with referential integrity constraint and included the functional dependencies between the attributes.

Note(in pictures): Primary Keys are represented as **Key**.The lines in orange represent relation mappings, blue represent functional dependencies.



3. Relational model after conversion to 2NF



Transition:

- Due to redundancy of data and violation of 2NF in **Result** (*SubId* is sufficient to functionally determine the attributes), entity **Result** was removed, attributes of **Result** were added to **Submission**.

4. Relational model after conversion to 3NF

Since the model was already in the 3NF after conversion to 2NF, no further changes were required.

```
INSERT INTO CONTEST VALUES
(1, '2020-01-01 18:00:00', '02:00:00'),
(2,'2019-08-01 15:00:00','02:00:00'),
(3,'2020-01-06 16:00:00', '02:00:00') ;

INSERT INTO CONTESTLANGUAGES VALUES
(1, 'C++'), (1, 'C'), (1, 'Python'),
(2,'JAVA'), (2,'Python'), (2,'C++'), (2,'RUBY'),
(3,'C'), (3,'C#'), (3,'C++'), (3,'RUBY');
```

```
INSERT INTO PROGRAMMER VALUES
(1,'numberdedopls@gmail.com', 'Bhavyajeet' ,'Singh','India','20','0'),
(2,'sirplsgivemarks@gmail.com', 'Jaidev', 'Sriram','India',20,'0'),
(3,'maampls@gmail.com', 'Jyoti' , 'Sunkara','India','20','0'),
(4,'Charaswati@gmail.com', 'Ashish' , 'Gupta','India','19','0'),
(5,'ArnabGoswami@gmail.com','Rutvij', 'Menavlikar','India','19','0'),
(6,'BONDOP@gmail.com', 'Anvay' , 'Karmore','India','19','0'),
(7,'multiplexer@gmail.com', 'Tejas' , 'Chaudhari','India','19','0'),
(8,'sidthesloth@gmail.com', 'Jaywant' , 'Patel','India','12','0'),
(9,'randibaaz@gmail.com', 'Aditya' , 'Verma','India','19','0'),
(10,'tyrantmf@gmail.com', 'Xi' ,'Jinping','China','29','1'),
(11,'momowala@gmail.com', 'Kim','Jong-Un','North Korea','25','1'),
(12,'omshaantiom@jaishreeram.com', 'Dalai','Lama','Nepal','35','1'),
(13,'orangeblob@whodoesntliketheirdaughtersass.com', 'Donald','Trump','USA','69','1');

INSERT INTO TEAM VALUES
(1,'3NOOBS','IIIT Hyderabad'),
(2,'HighHopes','IIIT Hyderabad'),
(3,'Manforce', 'IIT Hyderabad');

INSERT INTO CONTESTANT VALUES
(1,3,1),(2,3,1),(3,3,1),
(4,2,2),(5,2,2),(6,2,2),
(7,2,3),(8,2,3),(9,2,3);

INSERT INTO PARTICIPATES VALUES
(1,1),(2,1),(3,1),(1,2),(2,2),(3,2),(2,3),(3,3);


INSERT INTO PROBLEM VALUES
(1,'A',500,'Problem statement for problem A contest 1','sample test','IOFORMAT','sample
io',500,'00:30:00'),
(1,'B',750,'Problem statement for problem B contest 1','sample test','IOFORMAT','sample
io',500,'00:40:00'),
(1,'C',1000,'Problem statement for problem C contest 1','sample test','IOFORMAT','sample
io',500,'00:50:00'),
```

(1,'D',2000,'Problem statement for problem D contest 1','sample test','IOFORMAT','sample io',500,'00:10:00'),
(2,'A',500,'Problem statement for problem A contest 2','sample test','IOFORMAT','sample io',500,'00:30:00'),
(2,'B',750,'Problem statement for problem B contest 2','sample test','IOFORMAT','sample io',500,'00:40:00'),
(3,'A',500,'Problem statement for problem A contest 3','sample test','IOFORMAT','sample io',500,'00:10:00'),
(3,'B',750,'Problem statement for problem B contest 3','sample test','IOFORMAT','sample io',500,'00:20:00');

INSERT INTO SUBMITS VALUES
(1,1,'A',1),(1,1,'B',3),(1,1,'B',4),
(2,1,'A',2),(2,1,'B',6),(2,1,'C',7),
(3,1,'A',5),
(1,2,'A',8),
(2,2,'A',9),(2,2,'B',10),
(3,2,'B',11),
(1,3,'A',12),
(2,3,'A',13),(2,3,'B',14),(2,3,'B',15)
(3,3,'A',16);

INSERT INTO SUBMISSION VALUES
(1,'C++','00:12:00','AC',500,125,1),(2,'C++','00:16:00','WA',0,125,1),(3,'C++','00:17:00','AC',500,125,1),
(4,'C++','00:15:00','AC',500,125,1),(5,'C++','00:20:00','AC',750,125,1),(6,'C++','01:40:00','AC',1000,125,2),
(7,'C','00:25:00','AC',500,125,1),
(8,'C++','00:21:00','AC',500,125,1),
(9,'C++','00:12:00','AC',500,125,1),(10,'C++','00:40:00','AC',700,125,1),
(11,'JAVA','01:10:00','AC',750,125,1),
(12,'C++','00:08:00','AC',500,125,1),
(13,'C++','00:12:00','AC',500,125,1),(14,'C++','00:30:00','RE',0,125,1),(15,'C++','00:45:00','AC',750,125,1),
(16,'RUBY','01:12:00','WA',0,125,1);