

基于等价数据库状态的逻辑缺陷检测方法

宋建森, 窦文生, 郑莹莹, 高钰, 崔紫玉, 王伟, 魏峻

Detecting Schema-Related Logic Bugs in Relational DBMSs via Equivalent Database Construction

51st International Conference on Very Large Databases (VLDB 2025)

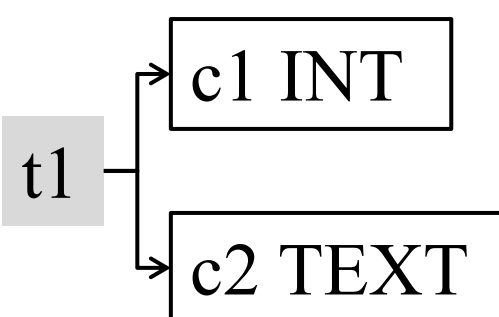
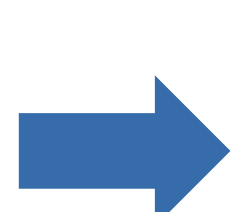
联系方式: 宋建森 songjiansen20@otcaix.iscas.ac.cn

研究背景

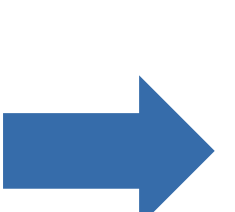
关系型数据库系统使用数据库模式描述数据库中数据的组织方式

```
CREATE TABLE t1 (  
  c1 INT,  
  c2 TEXT  
);
```

DDL语句



数据库模式 S



table_name	column_name	data_type	... is_nullable
t1	c1	integer	YES
t1	c2	text	YES

系统表 INFORMATION_SCHEMA

问题描述

数据库模式相关信息的不正确生成或者使用会造成不正确的执行结果

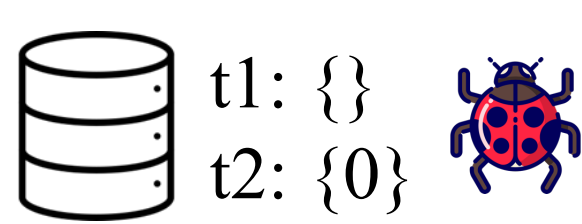
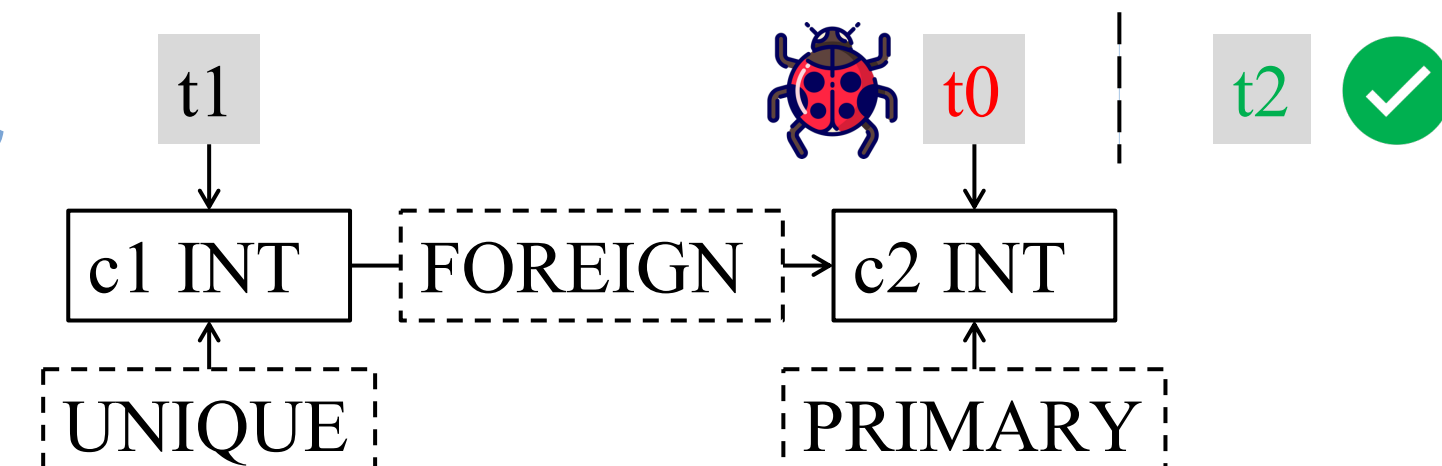
```
CREATE TABLE t0 (c2 INT, PRIMARY KEY (c2));  
CREATE TABLE t1 (c1 INT UNIQUE,  
  FOREIGN KEY (c1) REFERENCES t0(c2));  
ALTER TABLE t0 RENAME AS t2, ALGORITHM COPY;  
INSERT INTO t2 VALUE (0);  
INSERT INTO t1 VALUE (0);  
SELECT CAST(IFNULL(c1, '0') AS DATE) FROM t1;
```

测试用例

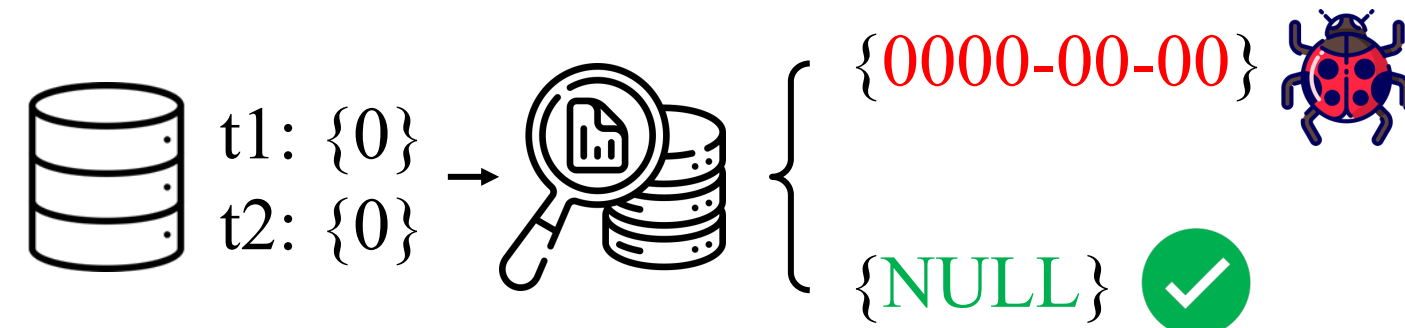
创建数据库模式

存储数据

查询数据



存储不正确的数据



返回不正确的查询结果

研究方案和思路

通过构造等价数据库状态来检测与数据库模式相关的逻辑错误

CREATE

CREATE

ALTER

db₁

创建数据库 db₁

CREATE

CREATE

db₂

合成数据库 db₂

INSERT

INSERT

生成测试语句 S

在 db₁ 上执行 S

在 db₂ 上执行 S

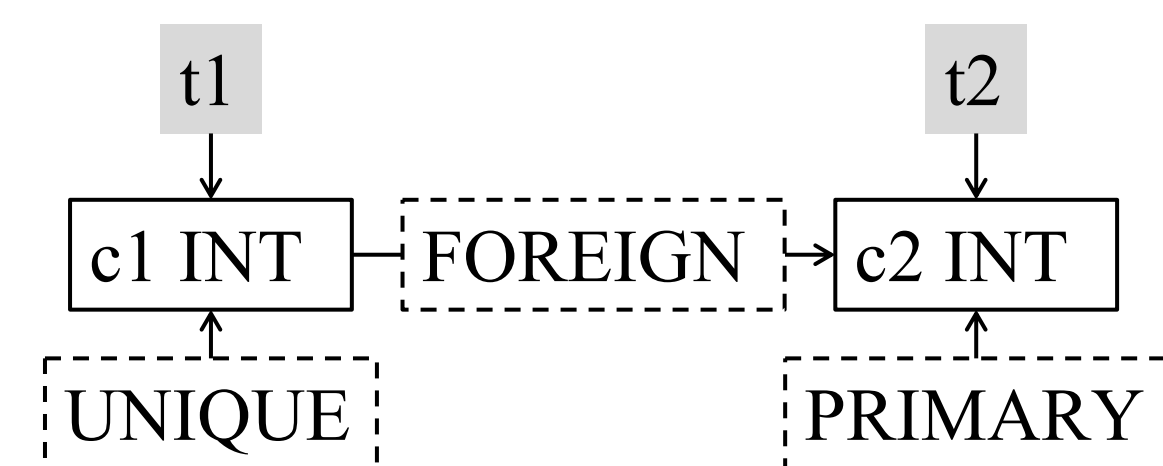
db'₁

rs₁

db'₂

rs₂

比较语句执行结果



合成CREATE TABLE语句

```
CREATE TABLE t2 (c2 INT PRIMARY KEY);  
CREATE TABLE t1 (  
  c1 INT UNIQUE,  
  FOREIGN KEY (c1) REFERENCES t2(c2)  
);
```

等价的DDL语句序列

实验结果

- 共计检测34个缺陷, 其中29个被确认为新的缺陷, 9个已被修复
- 相关工作至多能够检测13个本方法发现的逻辑缺陷

数据库系统	提交	确认	修复	重复	误报
MySQL	14	14	2	0	0
SQLite	0	0	0	0	0
PostgreSQL	0	0	0	0	0
MariaDB	11	7	3	3	0
CockroachDB	2	1	0	1	0
TiDB	7	7	4	0	0
总共	34	29	9	4	0

数据库系统	本方法	NoREC	Radar	DQP
MySQL	13	0	10	8
SQLite	0	0	0	0
PostgreSQL	0	0	0	0
MariaDB	5	0	1	0
CockroachDB	1	0	1	0
TiDB	2	0	1	1
总共	21	0	13	9