Lecture 18
Databases

Nadia Polikarpova

SQLizer: strengths

[Yaghmazadeh et al. OOPSLA'17]

Natural interaction model

High accuracy

Novel combination of NLP and type-directed synthesis

- NLP generates the structure and local hints; type-directed synthesis enumerates valid completions
- statistical reasoning + logical reasoning

SQLizer: limitations

[Yaghmazadeh et al. OOPSLA'17]

How should the user evaluate the result?

Works poorly for queries that return empty tables

Does not handle noise in the data part of the query

Semantic parser needs domain-specific training

Synthesizer queries the database (inefficient)

Unfair comparison with Nalir (?)

SQLizer: techniuqes

[Yaghmazadeh et al. OOPSLA'17]

Sketch generation via semantic parsing

- the idea of first generating a sketch, then filling it was also used in Cosi (Fast Synthesis of Fast Collections)
- somewhat similar to Leon where one technique generates the structure, and another one fills in details

Quantitative type inhabitation

- deductive synthesis that also deduces weights
- somewhat similar to Synquid (and other type-directed techniques)

Sketch refinement

- most similar to program repair

Entity matching

[Singh et al. VLDB'17]

Entity matching: which rows correspond to the same person?

Goal: more interpretable results than existing approaches (e.g. decision trees)

Search strategy: Sketch + techniques for handling noise

(a) D_1: **an instance of schema** R

	name	address	email	nation	gender
r_1	Catherine Zeta-Jones	9601 Wilshire Blvd., Beverly Hills, CA 90210-5213	c.jones@gmail.com	Wales	F
r_2	C. Zeta-Jones	3rd Floor, Beverly Hills, CA 90210	c.jones@gmail.com	US	F
r_3	Michael Jordan	676 North Michigan Avenue, Suite 293, Chicago		US	M
r_4	Bob Dylan	1230 Avenue of the Americas, NY 10020		US	M

(b) D_2: **An instance of the schema** S

	name	apt	email	country	sex
s_1	Catherine Zeta-Jones	9601 Wilshire, 3rd Floor, Beverly Hills, CA 90210	c.jones@gmail.com	Wales	F
s_2	B. Dylan	1230 Avenue of the Americas, NY 10020	bob.dylan@gmail.com	US	M
s_3	Micheal Jordan	427 Evans Hall #3860, Berkeley, CA 94720	jordan@cs.berkeley.edu	US	M

Entity matching

[Singh et al. VLDB'17]