CSIE 7242: Advanced Topics in Database Theory

(Semester 2, 2015–2016)

Course website

`http://www.csie.ntu.edu.tw/~tonytan/teaching/2015b-adb/2015b-adb.html`

Description

In this course we will cover some topics in database theory which in my opinion are interesting and relevant to real world applications. Some materials require a bit of background knowledge such as treewidth and hypergraphs, and we will cover them in the class too. I hope that this course can provide necessary exposure for students to the latest and interesting topics in database theory.

Prerequisite

Database systems, basic discrete mathematics. Some knowledge of automata theory and mathematical logic is useful, but not necessary. We will try to make this course as self contained as possible.

Syllabus (tentative)

	Week	Topics	Important dates
I	1	First-order logic in database	HW 1 out
	2	Relational algebra	–
	3	Ehrenfeucht-Fraïssé games	–
	4	Datalog queries	–
	5	–	HW 1 due
II	6	Treewidth of graphs/databases*	HW 2 out
	7	Tree automata*	–
	8	Query evaluations on bounded treewidth databases	–
	9	–	HW 2 due
III	10	Hypergraphs*	HW 3 out
	11	Acyclic (join) queries	–
	12	Width of queries	–
	13	–	HW 3 due
IV	14	Entropy*	HW 4 out
	15	Counting with entropy*	
	16	Bounds on join	–
	17	Some recent join algorithms	–
	18	–	HW 4 due

The asterisked topics denote background knowledge required for our next topic.

Grading

- Four homeworks weigh 20% each.
- Participation and discussion in all four parts, each part weighs 5%.

References

References to textbooks and papers will be introduced along the way. Here are a sample of them.

(1) S. Abiteboul, R. Hull, V. Vianu. *Foundations of Databases*. Addison-Wesley, 1995.

Available online: `http://webdam.inria.fr/Alice/`

(2) A. Atserias, M. Grohe, D. Marx. Size Bounds and Query Plans for Relational Joins. *SIAM Journal on Computing*, 42(4):1737–1767 (2013).

(3) E. Codd. A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6):377–387 (1970).

(4) H. Comon, M. Dauchet, R. Gilleron and C. Löding, F. Jacquemard, D. Lugiez, S. Tison, M. Tommasi *Tree Automata Techniques and Applications*.

In `http://tata.gforge.inria.fr/`

(5) H. Ngo, C. Ré, A. Rudra. Skew strikes back: new developments in the theory of join algorithms. *SIGMOD Record* 42(4): 5-16 (2013).

(6) T. Veldhuizen. Triejoin: A Simple, Worst-Case Optimal Join Algorithm. Proceedings of the 17^{th} *International Conference on Database Theory (ICDT)* 2014, 96–106.

(7) M. Yannakakis. Algorithms for Acyclic Database Schemes. Proceedings of the 7^{th} international conference on *Very Large Data Bases (VLDB)* 1981, 82–94.