

COMP 3311

Database Management Systems

Lab 8

Task 3 Screenshots

Author: Create Submission

Conference Website Author - PC Chair - PC Member -

Create Submission

Title: Test submission

Abstract: Test abstract

Type: Industrial

Add Author

Title: None

Name:

Institution:

Country:

Phone number:

Email:

Add Author

Authors

CONTACT?	TITLE	NAME	INSTITUTION	COUNTRY	PHONE NO	EMAIL
<input checked="" type="checkbox"/>	Dr	Test name	Test Institution	Test country	00000000	test@test
<input type="checkbox"/>	Mr	Test name 2	Test Institution 2	Test country 2	11111111	test@test1

Submit

© 2018 - Conference Website

Author: Find All Author Submissions

Conference Website Author PC Chair PC Member

Find All Submissions For An Author

Enter your email address:

SUBMISSION	TITLE	ABSTRACT	TYPE
1	Example-Driven Schema Mapping	End-users increasingly find the need to perform light-weight, customized data integration. State-of-the-art tools usually require an in-depth understanding of the semantics of multiple schemas. We propose a system, MWeaver, that facilitates data integration for end-users.	research

© 2018 - Conference Website

Author: Find Submission

Conference Website Author PC Chair PC Member

Find Submission

Submission: 1 Search

Title: Example-Driven Schema Mapping

Abstract: End-users increasingly find the need to perform light-weight, customized data integration. State-of-the-art tools usually require an in-depth understanding of the semantics of multiple schemas. We propose a system, MWeaver, that facilitates data integration for end-users.

Submission type: research

Author Information

CONTACT	TITLE	NAME	INSTITUTION	COUNTRY	PHONE NO	EMAIL
✓		Michael J. Cafarella	University of Michigan	USA	16084343702	michjc@umich.edu
	Prof	H. V. Jagadish	University of Michigan	USA	16084343923	jag@umich.edu

© 2018 - Conference Website

PC Chair: Assign Submission To PC Member

[Conference Website](#) [Author](#) [PC Chair](#) [PC Member](#)

Assign Submission To PC Member

Submission:

Title: Example-Driven Schema Mapping

PC Members Assigned To This Submission:

PC CODE	NAME
ec01	Eric Crestan
er01	Elke Rundensteiner
jh01	Jiawei Han

Minimum preference:

PC Members Available To Be Assigned To This Submission:

SELECT	PC CODE	PREFERENCE	SUBMISSIONS ASSIGNED
<input type="checkbox"/>	hj01	2	3
<input type="checkbox"/>	lc01	1	3

© 2018 - Conference Website

PC Chair: Create PC Member

Conference Website Author▼ PC Chair▼ PC Member▼

Create PC Member

PC code: ee01

Title: Prof ▼

Name: Test

Institution: Test

Country: Test

Phone number: 00000000

Email: test@test

Submit

© 2018 - Conference Website

PC Chair: Display PC Member Information

Conference Website Author PC Chair PC Member							
PC Member Information							
	PC CODE	TITLE	NAME	INSTITUTION	COUNTRY	EMAIL	PHONE NO
Edit	bo01	Prof	Beng Chin Ooi	National University of Singapore	Singapore	ooibc@comp.nus.edu.sg	6565166465
Edit	ec01	Dr	Eric Crestan	Yahoo! Labs	USA	ecrestan@yahoo-inc.com	14156786654
Edit	er01	Prof	Elke Rundensteiner	Worcester Polytechnic Institute	USA	rundenst@cs.wpi.edu	15088315815
Edit	fl01	Prof	Frank Lo	Hong Kong University of Science and Technology	China	flo@cse.ust.hk	85223586996
Edit	hj01	Prof	H. V. Jagadish	University of Michigan	USA	jag@umich.edu	16084343923
Edit	jh01	Prof	Jiawei Han	University of Illinois at Urbana-Champaign	USA	hanj@cs.uiuc.edu	12173336903
Edit	jy01	Prof	Jeffrey Yu	Chinese University of Hong Kong	China	yu@se.cuhk.edu.hk	85239438309
Edit	lc01	Prof	Lei Chen	Hong Kong University of Science and Technology	China	leichen@cse.ust.hk	85223586980
Edit	pl01	Dr	Paul Larson	Microsoft Research	USA	palarson@microsoft.com	14257036260
Edit	pp01	Dr	Patrick Pantel	Yahoo! Labs	USA	ppantel@yahoo-inc.com	14156782345
© 2018 - Conference Website							

PC Chair: Edit PC Member Information

Conference Website Author▼ PC Chair▼ PC Member▼

Edit PC Member Information

PC code: bo01

Title: Prof ▼

Name:

Beng Chin Ooi

Institution:

National University of Singapore

Country:

Singapore

Phone number:

6565166465

Email:

ooibc@comp.nus.edu.sg

Update

© 2018 - Conference Website

PC Member: Create Referee Report

Conference Website Author - PC Chair - PC Member -

Create Referee Report

PC code: F01 [Get My Submissions for Review](#) Submission: 0

Title: Towards User-Friendly Entity Resolution

Author(s): Jeffrey Naughton, Rajasekar Krishnamurthy

Use (Y)es or (N)o and if absolutely necessary (M)aybe to answer the following (you must explain (N)o and (M)aybe answers in the comments section).

The paper is relevant to the conference: Yes

The paper is technically correct: Yes

The length and content of the paper are comparable to the expected final version: Yes

For the following categories, please assign integer scores from 1 to 5.

Originality	Impact	Presentation	Technical Depth	OVERALL RATING
1 - Reject	1 - Reject	1 - Reject	1 - Reject	1 - Reject

Reviewer Confidence (0.5-1): 1

Main Contribution(s):

Three strong points of the paper:

Three weak points of the paper:

Overall Summary:

Detailed Comments:

Confidential comments:

Submit Referee Report

© 2018 - Conference Website

PC Member: Display Referee Report

[Conference Website](#) [Author -](#) [PC Chair -](#) [PC Member -](#)

[Display Referee Report/Discussion](#) — [Add Discussion](#)

PC code: Submission:

Title: Optimal Schemes for Robust Web Extraction

Author(s): Aditya Parameswaran, Hector Garcia-Molina, Nilesh Dalvi, Rajeev Rastogi

The paper is relevant to the conference: Y

The paper is technically correct: Y

The length and content of the paper are comparable to the expected final version: Y

Originality	Impact	Presentation	Technical Depth	OVERALL RATING
3	3	3	3	3

Reviewer Confidence: 0.8

Main Contribution(s): A robust wrapper is proposed that is shown to be robust. The paper is clearly presented and well motivated.

Three strong points of the paper: An adversarial and a probabilistic model are proposed.

Three weak points of the paper: The two models are well motivated.

Overall Summary: A good paper on a timely topic.

Detailed Comments

Confidential comments to the PC, if any:

© 2018 - Conference Website

PC Member: Display Reviewing Status

[Conference Website](#) [Author](#) [PC Chair](#) [PC Member](#)

Reviewing Assignments

PC code:

Submissions Reviewed

SUBMISSION	TITLE	ABSTRACT	TYPE
2	Optimal Schemes For Robust Web Extraction	We consider the problem of constructing robust wrappers for web information extraction. We consider two models to study robustness formally; the adversarial model and probabilistic model. We demonstrate that our algorithms can reduce wrapper breakage by up to 500% over existing techniques.	research

Submissions Not Reviewed

SUBMISSION	TITLE	ABSTRACT	TYPE
6	Towards User-Friendly Entity Resolution	We explore the possibility of treating user input as an integral part of the entity resolution process. We design a simple two-stage approach that separates merging and splitting records into two separate stages.	research
9	Druid: A Real-time Analytical Data Store	Druid is an open source data store designed for real-time exploratory analytics on large data sets. It combines column-oriented storage layout, distributed, shared-nothing architecture, and advanced indexing to allow for the arbitrary exploration of billion-row tables with sub-second latencies.	industrial

© 2018 - Conference Website

PC Member: Display Submission Preferences

Conference Website Author PC Chair PC Member

Submission Preferences

PC code:

Submissions For Which You Have Already Specified A Preference:

SUBMISSION	TITLE	ABSTRACT	TYPE	PREFERENCE
7	TsingHUS: A Location-based Service System	TsingHUS aims to provide users with more user-friendly location-aware search experiences. TsingHUS incorporates continuous search to efficiently support continuously moving queries in a client-server system thereby reducing the communication cost between the client and server.	demo	4
8	A Java Stream Computational Model for Big Data	The addition of lambda expressions and a Stream API in Java 8 provide a powerful and expressive query language. We build on Java 8 Stream and add a DistributableStream abstraction that supports federated query execution over an extensible set of distributed compute engines.	industrial	3

Submissions For Which You Have Not Specified A Preference:

PREFERENCE	SUBMISSION	TITLE	ABSTRACT	TYPE
<input type="button" value="Select"/>	1	Example-Driven Schema Mapping	End-users increasingly find the need to perform light-weight, customized data integration. State-of-the-art tools usually require an in-depth understanding of the semantics of multiple schemas. We propose a system, MWeaver, that facilitates data integration for end-users.	research
<input type="button" value="Select"/>	2	Optimal Schemas for Robust Web Extraction	We consider the problem of constructing robust wrappers for web information extraction. We consider two models to study robustness formally: the adversarial model and probabilistic model. We demonstrate that our algorithms can reduce wrapper breakage by up to 500% over existing techniques.	research
<input type="button" value="Select"/>	3	Web-Scale Knowledge Extraction	We propose a classification algorithm and a rich feature set for automatically recognizing layout tables and attribute/value tables. In 79% of our Web tables, our method finds the correct protagonist in its top three returned candidates.	research
<input type="button" value="Select"/>	4	Efficient Fusion of Historical Data	Historical data may include severe data conflicts that prevent researchers from obtaining the correct answers to queries on an integrated historical database. We consider an efficient approach to large-scale historical data fusion.	research
<input type="button" value="Select"/>	5	CETR - Content Extraction via Tag Ratios	Content Extraction via Tag Ratios (CETR) is a method to extract content text from diverse webpages using the HTML document's tag ratios. We evaluate our approach against a large set of alternative methods, which shows that CETR achieves better content extraction performance than existing methods.	research
<input type="button" value="Select"/>	6	Towards User-Friendly Entity Resolution	We explore the possibility of treating user input as an integral part of the entity resolution process. We design a simple two-stage approach that separates merging and splitting records into two separate stages.	research
<input type="button" value="Select"/>	9	Droid: A Real-time Analytical Data Store	Droid is an open source data store designed for real-time exploratory analytics on large data sets. It combines column-oriented storage layout, distributed, shared-nothing architecture, and advanced indexing to allow for the arbitrary exploration of billion-row tables with sub-second latencies.	industrial
<input type="button" value="Select"/>	10	Orca: A Modular Query Optimizer	Orca, a new query optimizer for all Pivotal data management products, is a comprehensive development uniting state-of-the-art query optimization technology with original research resulting in a modular and portable optimizer architecture.	industrial
<input type="button" value="Select"/>	11	Amoeba: A Shape Changing, Big Data Storage System	Amoeba is a distributed storage system which uses adaptive multi-attribute data partitioning to efficiently support ad-hoc as well as recurring queries.	demo

Update Preferences

© 2018 - Conference Website