

BOGDAN ALEXE

Ph.D. Candidate
Computer Science Department
University of California, Santa Cruz
1156 High Street
Santa Cruz, CA 95060

Phone: +1-831-295-0304
Email: abogdan@cs.ucsc.edu
Web: www.cs.ucsc.edu/~abogdan

Education

University of California, Santa Cruz	Since 2005
Ph.D. Candidate in Computer Science Advisor: Prof. Wang-Chiew Tan Area: Databases Research Interests: Information Integration, Data Exchange Expected graduation date: June 2011	
University of California, Santa Cruz	2005 - 2007
M.Sc. in Computer Science	
Ecole Nationale Supérieure des Télécommunications Paris, France	2004 - 2005
M.Sc. in Computer Science (Joint program with Ecole Polytechnique)	
Ecole Polytechnique Palaiseau, France	2002 - 2004
Majored in Computer Science, Cumulative GPA: 3.28/4 Graduated with the “Engineer of Ecole Polytechnique” Diploma	
“Politehnica” University Bucharest, Romania	1999 - 2004
B.Sc. in Computer Science, Cumulative GPA: 9.41/10	

Selected Graduate Coursework

- Databases (project: relational chase engine for tuple- and equality-generating dependencies)
- Programming Languages (project: type inference for simply-typed lambda calculus)
- Machine Learning (project: text classification techniques)
- Cloud Computing (project: natural language interface to relational databases via crowdsourcing)

Research Experience

Research Intern at IBM Almaden Research Center June - September 2007, 2008
Intelligent Information Integration (Clio) Group
Mentors: Lucian Popa and Mauricio A. Hernández

I contributed to a novel mapping generation mechanism for arbitrary flows of schema mappings in the Clio system. The new mapping generation technique considers independently designed, uncorrelated, smaller schema mappings as building blocks and automatically assembles flows of such mappings into specifications for larger, richer data transformations.

Research Assistant at UC Santa Cruz

Since 2005

Computer Science Department, Databases Group

Advisor: Prof. Wang-Chiew Tan

Contributed to the following projects:

- **Designing and Refining Schema Mappings via Data Examples:** I contributed to the design and prototype implementation of a novel approach to schema mapping design and refinement through data examples. Given a set of data examples, each representing a partial specification of the semantics of the desired schema mapping, the system constructs a schema mapping specified by Global-and-Local-As-View constraints that “fits” the data examples, if such mapping exists. Although the worst-case complexity of the algorithm at the core of the system is high, the system achieves very good performance on real-life scenarios and demonstrates the feasibility of interactively designing schema mappings using data examples.
- **STBenchmark: A Benchmark for Mapping Systems:** I have contributed to the development of STBenchmark, a solution for benchmarking mapping systems. STBenchmark has three components: a suite of basic mapping scenarios that should be readily supported by any mapping system, a mapping scenario generator and instance generator for producing complex mapping scenarios and instances of varying sizes (also usable for stress testing implementations of various information integration algorithms), and a simple usability model used to evaluate the ease of use of a mapping system.
- **MUSE: Mapping Understanding and deSign by Example:** I have contributed to the development of Muse, a schema mapping design wizard that uses data examples to assist designers in understanding and refining a schema mapping towards the desired specification. Muse works on two components of a mapping specification that correspond to the design of grouping semantics and the desired interpretation of ambiguous mappings.
- **SPIDER - A Schema Mappings Debugger:** I have contributed to the development of the SPIDER prototype tool for debugging schema mappings.
- **DBNotes: An Annotation Management System for Relational Databases.** My contributions were the design and implementation of an alternative annotation storage scheme aimed at eliminating data redundancies.

Intern at Bell Labs/Lucent Technologies

April - August 2004

Network Data and Services Research Department

Mentors: Arnaud Sahuguet and Irimi Fundulaki

- **Evaluation Strategies for the XSquirrel Subtree Language and Applications for Federated User Profile Management:** The GUPster system developed at Bell Labs aimed at ensuring privacy-conscious access to global user profile data in a distributed environment. It was based on the XSquirrel language for XML subdocument queries. My main contributions resided in the core of the GUPster system (its access control and data integration modules), in the web service wrappers built around the various data sources, and in the XSquirrel evaluation engine.

Publications

- [1] MapMerge: Correlating Independent Schema Mappings
B. Alexe, M.A. Hernández, L. Popa, and W. Tan
Proceedings of the VLDB Endowment, Volume 3, September 2010, pages 81-92
- [2] Characterizing Schema Mappings via Data Examples
B. Alexe, P.G. Kolaitis, and W. Tan
ACM Symposium on Principles of Database Systems (PODS) 2010, pages 261-272
- [3] STBenchmark: Towards a Benchmark for Mapping Systems
B. Alexe, W. Tan, and Y. Velegrakis
International Conference on Very Large Data Bases (VLDB) 2008, pages 230-244
- [4] Comparing and Evaluating Mapping Systems with STBenchmark
B. Alexe, W. Tan, and Y. Velegrakis
International Conference on Very Large Data Bases (VLDB) 2008, Demo, pages 1468-1471
- [5] MUSE: A System for Understanding and Designing Mappings
B. Alexe, L. Chiticariu, R.J. Miller, D. Pepper, and W. Tan
International Conference on Management of Data (SIGMOD) 2008, Demo, pages 1281-1284
- [6] MUSE: Mapping Understanding and deSign by Example
B. Alexe, L. Chiticariu, R.J. Miller, and W. Tan
International Conference on Data Engineering (ICDE) 2008, pages 10-19
- [7] SPIDER: A Schema Mappings Debugger
B. Alexe, L. Chiticariu, and W. Tan
International Conference on Very Large Data Bases (VLDB) 2006, Demo, pages 1179-1182
- [8] Sub-Document Queries Over XML with XSquirrel
A. Sahuguet and B. Alexe
International Conference on World Wide Web (WWW) 2005, pages 268-277
- [9] User Profile Management in Converged Networks(Episode II):"Share your data, Keep your secrets"
A. Sahuguet, B. Alexe, I. Fundulaki, P.-Y. Laligand, A. Shikfa, and A. Arnail
Conference on Innovative Data Systems Research (CIDR) 2005, pages 200-212
- [10] An Electronic Patient Record "on Steroids": Distributed, Peer-to-Peer, Secure and Privacy-conscious
S. Abiteboul, B. Alexe, O. Benjelloun, B. Cautis, I. Fundulaki, T. Milo, and A. Sahuguet
International Conference on Very Large Data Bases (VLDB) 2004, Demo, pages 1273-1276

Technical Report

- [11] Specification, Composition and Evaluation of Subtree Queries
M. Benedikt, I. Fundulaki, B. Alexe
Bell Labs Technical Report, 2005

Teaching Experience

Teaching Assistant at UC Santa Cruz

April - June 2006

- **Introduction to Programming for Natural Sciences Students** (lower division introductory Java programming class): designed lab assignments, conducted lab sections, held office hours and graded exams

Posters

MUSE: Schema Mapping Understanding and deSign by Example

UC Santa Cruz School of Engineering Research Review Day, October 2007

UC Santa Cruz Graduate Research Symposium, May 2007 (Best Engineering Poster Award)

SPIDER: a Schema Mappings Debugger

UC Santa Cruz Graduate Research Symposium, May 2006

UC Santa Cruz School of Engineering Research Review Day, October 2007

Awards and Honors

University of California, Santa Cruz Chancellor's Fellowship	2005 - 2006
Ecole Polytechnique Foundation Scholarship	2002 - 2004
Ranked 7th (out of 57 teams) in the Southwestern European ACM Programming Contest	2003
Romanian Government Merit-based Scholarship	1999 - 2002
First Prize, Scientific Communications Session of the "Politehnica" University, Bucharest	2000
Ranked First at the National Romanian Physics Competition	1999
Member of the Extended Romanian Team for the International Physics Competition	1999
Bronze Medal at the International Environmental Projects Competition, Istanbul, Turkey	1998
Prizes at regional physics, mathematics and programming competitions in Romania	1995 - 2000

Service to the Community

External Reviewer: SIGMOD 2010, VLDB 2007-2010, ICDE 2008-2010, WebDB 2009, BIRTE 2008, ICMT 2008, DASFAA 2006

Technical Skills

Programming languages: Java, OCaml, C/C++, PHP, Pascal, Basic

Relational Database Systems: IBM DB2, Oracle

Operating Systems: Linux, Solaris, Windows

Others: Verilog, Xilinx Circuit Design Tools, Cisco Networks Design and Administration

Spoken Languages

English: Fluent (TOEFL CBT Score: 300/300, TSE Score: 60/60)

French: Fluent (TCF Level 5), **Spanish and German:** Conversational, **Romanian:** Native