Curriculum vitae et studiorum

Carlo Sartiani

Personal Information

Name: Carlo Sartiani
Citizenship: Italian
Current Position: Assistant Professor at Dipartimento di Matematica e Informatica, Università della Basilicata

Contacts

Office: Dipartimento di Matematica e Informatica - Università della Basilicata - Via dell'Ateneo Lucano - Potenza - Italy
Email: sartiani@di.unipi.it - carlo.sartiani@unibas.it
Home Page: http://www.di.unipi.it/∼sartiani

Education

01/2004 — PhD in Computer Science from Università of Pisa

PhD Thesis:*Efficient Management of Semistructured XML Data*; Supervisor: Prof. Giorgio Ghelli; International Referees: Alberto Mendelzon and Dan Suciu.

11/1998 - 10/2002 — PhD candidate in Computer Science at Dipartimento di Informatica - Università di Pisa

PhD Short Courses:

1) Semantics of Programming Languages II (Prof. Ugo Montanari);
2) Database Programming Languages (Prof. Antonio Albano, Prof. Giuseppe Castagna, Prof. V. Benzaken) (Bertinoro);
3) Mobility Languages (Davide Sangiorgi) (Bertinoro);
4) Petri Nets (Prof. G. Balbo) (Bertinoro);
5) Syntactic Rule Formats and Defining Higher-Order Languages (Prof. Karen Bernstein);
6) Applied Cryptography (Prof. Paolo Ferragina, Prof. Fabrizio Luccio, Prof. Giuseppe Persiano);
7) Adaptive Data Structures (Prof. Ah Chung Tsoi);
8) Data mining & Knowledge Discovery in Databases (Prof. Jiawei Han, Prof. Dino Pedreschi, Fosca Giannotti);
9) Constraint Reasoning and Programming (Prof. Thom Fruehwirth).

Summer Schools:

1) International School for Computer Science Researchers, Lipari, 2000;
2) International Summer School on Applied Semantics (APPSEM'2000);
3) International School for Computer Science Researchers, Lipari, 2001;
4) International School for Computer Science Researchers, Lipari, 2002;
5) EDBT'02 SUMMER SCHOOL - Distributed Databases on the Net: Models, Languages and Infrastructures.

07/1998 — Master Degree in Computer Science from Università di Pisa

Phd Thesis - Abstract

The last few years have seen the rapid emerging of the eXtensible Markup Language (XML). XML was designed as a simplification of SGML, and it has become the most widely used universal data representation format. In particular, the ability of XML to easily represent data with irregular structure has imposed XML as the standard incarnation for *semistructured* data, i.e., data with irregular, unstable, or even unknown structure.
In the context of XML data management systems, the estimation of query cardinality is becoming more and more important: the information provided by a query result estimator can be used as input to the query optimizer, as an early feedback to user queries, as well as input for determining an optimal storage schema.
This Thesis describes the result size estimation model of Xtasy, a prototype XML database management system. Unlike other existing models, which focus on very limited subsets of XQuery, the proposed model covers the FLWR core of XQuery, and estimate not only the *raw* cardinality of query results, but also their distribution.

Professional Activity

2004 - 2005	Postdoctoral researcher at Dipartimento di Informatica - Università di Pisa;
2005 - 2006	Temporary contract holder at Dipartimento di Informatica - Università di Pisa;
2006 - 2008	Postdoctoral researcher at Dipartimento di Informatica - Università di Pisa;
2008 -	Assistant professor at Dipartimento di Matematica e Informatica - Università della Basilicata;
May 2007 - June 2007	Visiting researcher at Laboratoire de Recherce en Informatique (LRI) - Université Paris Sud - France.
April 2008	Visiting researcher at Laboratoire de Recherce en Informatique (LRI) - Université Paris Sud - France.

Referee Activity

Organization	XOODB 2009 - First International Workshop on XML and Object-oriented Database Technology - In conjunction with ECOOP 2009;
PC Member	IADC 2007 Special Special Track on Advances in Querying Non-Conventional Data Sources; DataX 2008, Third International Workshop on Database Technologies for Handling XML Information on the Web; DataX 2009, Fourth International Workshop on Database Technologies for Handling XML Information on the Web;
Journals	ACM Transactions on Database Systems (ACM TODS), Data Knowledge and Engineering (DKE), Software: Practice and Experience, Journal of the American Society for Information Science and Technology (JASIST), Journal of Digital Information Management (JDIM), IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), ACM Transactions on Programming Languages (TOPLAS);
International Conferences	ICFP 2001, DBPL 2001, DBPL 2003, DBPL 2005, ICDE 2004, ICDE 2005, POPL 2005, BDA 2006, EDBT 2006, DataX 2004, ADBIS 2007, EDBT 2008;
French Conferences	BDA 2006;
Italian Conferences	SEBD 2000, SEBD 2004, SEBD 2005, SEBD 2006, SEBD 2007.

Teaching Activity

2006	**Semistructured Data and XML - Track inside the Database Management Systems course of the International Master on Information Technology of Scuola Superiore Sant'Anna.** This track was primarily focused on XML as both a data exchange and data representation format. The audience of the course comprised graduate students, and the lectures covered both theoretical and practical aspects of XML data management.
2006	**Database Lab.** During this course we taught undergraduate students how to design and implement the information system of a large organization. Much emphasis was given to the principles and techniques of conceptual and relational design, as well as to their practical implementation.
2007	**Database Lab.** During this course we taught undergraduate students how to design and implement the information system of a large organization. Much emphasis was given to the principles and techniques of conceptual and relational design, as well as to their practical implementation.
2008	**Database Lab.**
2008	**Procedural Languages Lab.**
2009	**OO Languages Lab.**
2009	**Formal Languages and Compilers**

Advising Activity

2004	Co-advisor of Giacoma Monreale's Bachelor thesis;
2004-2005	Co-advisor of Nicola Gioia's Master thesis;
2005-2006	Co-advisor of Paolo Tomei's Master thesis;
2005-2006	Co-advisor of Giovanni Pardini's Bachelor thesis;
2005-2006	Co-advisor of Federico De Faveri's Bachelor thesis;
2005-2006	Co-advisor of Luca Pardini's Bachelor thesis;
2006-2007	Co-advisor of Michele Freschi's Bachelor thesis;
2006-2007	Co-advisor of Donato Ferrante's Bachelor thesis;
2007-2008	Co-advisor of Marco Volpetti's Bachelor thesis;
2007-2008	Co-advisor of Giovanni Viscuso's Bachelor thesis;
2007-2008	Co-advisor of Giacomo Bachi's Bachelor thesis;
2007-2008	Co-advisor of Giacomo Falchetti's Bachelor thesis;
2007-2008	Co-advisor of Luca Pardini's Master thesis.

Foreign Languages

Fluency in both spoken and written English.

Research Interests

My research interests lie in the fields of database system and programming languages. The guiding idea of my past and current research activity is to extend the technology of current database systems to new contexts and settings, and to explore the intersections and the interactions between databases and programming languages. These principles also guide my future research plans.

Future Research Directions

Databases and new storage media The last few years have seen the emerging of a new kind of mass storage devices called *Solid-State Drives* (SSDs). These new peripherals are based on flash memory and/or static RAM chips, and contain no moving mechanisms.

SSDs are usually used to store photos, videos, and music files, as their limited power consumption fits very well the needs of portable devices.

The absence of moving parts and, consequently, the reduced power consumption of SSDs have made solid-state drives attractive for mobile computing as well as for data centers and computer farms. In particular, the lower energy requirements wrt conventional hard drives (and the resulting lower heat generation) allow data centers to significantly decrease the cooling and air conditioning costs, hence shrinking the overall operational costs (as well as the greenhouse gas emissions).

Being based on a very different technology, SSDs have very different performances wrt conventional hard drives. In particular, SSDs have very low access (*seek*) and latency times ($\sim$ 1 ms), while their transfer rate is a few magnitude orders lower than that of hard drives. This implies that any I/O-bound application should be modified and adapted to this new context to maximize its performance and to avoid issues related to the very different cost model.

Database systems are a typical example of applications whose performance is strictly related to the cost model of mass storage peripherals. Current DBMSs feature very sophisticated query optimizers (e.g., the Starburst optimizer family of IBM DB2) that allow them to efficiently process and execute very heavy workloads. These optimizers strongly rely on the cost model of current hard drives, that favours sequential scans wrt random accesses. It is quite evident, hence, that all optimization heuristics and strategies (used to decrease the portion of search space to be explored) should be revised when moving to mass storage peripherals, like SSDs, with efficient random accesses and slow sequential scans.

This research project aims at exploring how current DBMSs should be modified to better exploit the features and the potential of solid-state drives. In particular, we will focus on three main research tasks:

1. we will analyze the experimental performance of current DBMSs on a standard database benchmark when the database and the indexes reside on a SSD; these experimental results will be compared with the results we will gain on a conventional hard-drive setting;

2. we will explore the applicability of techniques coming from main memory databases as an alternative to standard persistent techniques;

3. we will try to design a new cost model for query optimizers, taking SSDs into account, as well as new optimization heuristics to be used inside query optimizers.

All these research activities will integrated by extensive experimental analyses.

Bidirectional Schema Mappings The astonishing growth of available information, on both the Web and local computers, has significantly redefined the requirements and the needs of modern data management: indeed, the problem of finding the right piece of information at the right time has evolved and shifted from a relatively centralized context to an highly decentralized and distributed one. While traditional data management has focused on data hosted on a single data source (e.g., a relational or an object-oriented database), the fragmentation of information on multiple heterogeneous sources, often autonomous and accessible only through forms or low-level web services, implies the definition and development of new techniques and tools for their access.

This paradigm shift significantly impacts on current data integration solutions. Indeed, data integration systems are shifting from strongly controlled and centralized approaches to decentralized

and loose solutions, where the query reformulation process is distributed across the whole network of integrated data sources. In such contexts, it is very important the ability to traverse the data source graph in any direction, so to find the best (and cheapest) reformulation path. As schema mappings are usually expressed through unidirectional formalisms (see, for instance, the formalism used in Clio), decentralized data integration systems adopt very computationally expensive mapping inversion algorithms, which can be applied only in very limited situations.

A similar problem arises in data synchronization, where synchronization rules must be easily invertible to properly propagate updates backward and forward. To this aim, bidirectional programming languages have been defined in the recent past and applied to the data synchronization problem and to the view update problem.

This research aims at using the mechanisms of bidirectional programming languages, which are strongly procedural, to define a new class of inherently bidirectional schema mappings. These mappings, expressing a bidirectional transformation between two data sources, do not need to be inverted and can be used to go back and forth between the data sources.

As the mechanisms of bidirectional programming languages are inherently procedural, while schema mappings are declarative specifications of actual transformations, a challenging point of this research is to find a simple, yet robust way to hide the procedural aspects inside a (new) declarative specification language. We expect that a key issue in bridging these two realms is the development of a graphical tool that helps the mapping designer in defining bidirectional mappings.

Current Research Activities

Gamma (Co-investigator) This project is a joint activity with Dario Colazzo, currently at LRI, France, and started in the last quarter of 2004.

The basic idea of the project is to study the problem of maintaining schema mappings in XML p2p systems, so to develop techniques for automatically discovering corrupted mappings. This problem is very important, as query answering depends on the quality of schema mappings.

Our first approach, presented at PlanX 2005, was to derive the correctness of a mapping from the correctness of rewritten queries over the target schema: if a query Q_1 on $\mathcal{S}$ is rewritten by a mapping m into a query Q_2 over $\mathcal{T}$, and Q_2 does not match $\mathcal{T}$, then m is for sure incorrect.

This approach, even if very simple, has two main drawback. First, it is not complete, as some errors cannot be captured. Second, it depends on the query answering algorithms, hence it is computationally very expensive. To overcome these issues, we developed a second approach, presented at DBPL 2005, which is complete (**ALL** errors are captured) and is (almost) tractable. The basic idea of this technique is to infer (once) the output type of a mapping and to compare it with the target schema, according to a type projection relation.

Type projection is a semantic notion based on the usual intuition of a schema mapping as a transformation + a projection. While in the DBPL 2005 paper we analyzed the relationship between type projection and semantic subtyping as well as its decidability properties, in a recent paper (PPDP 2006) we provided a polynomial algorithm (up to type normalization) for type projection and discussed its properties.

We are currently extending our approach so to cover a wider class of schemas and mappings. In particular, we are enhancing the type projection algorithm to manage recursive schemas too; recursion, even though not frequently used in data-centric applications, is widely employed in XML-based document processing applications.

A prototype implementation of this approach, with a nice graphical tool, is available at the following URL:

`http://www.di.unipi.it/~sartiani/projects/gamma.html`.

Decision Problems for XML Types with Interleaving and Counting (Co-investigator) This project is a joint activity with Giorgio Ghelli (Dipartimento di Informatica, Pisa) and Dario Colazzo (LRI, Paris).

Neven et al. showed that decision problems for single-type eDTDs with interleaving and counting are computationally very hard; in particular, membership is NP-complete, while inclusion is EXPSPACE-complete. This project aims at identifying restrictions on the grammar of types that ensure the tractability of these problems.

We designed a restriction, called *conflict freedom*, that guarantees the polynomial (quadratic) complexity of inclusion when satisfied by both the subtype and the supertype. This restriction is satisfied by most hand-crafted XSDs used in practical scenarios.

We also extend this approach to membership and shown that membership is polynomial when the type is conflict-free. Our membership algorithm has been compared against other membership (included the standard SAX validator) and showed excellent scalability properties.

Of course, it is not realistic to assume that both the subtype and the supertype satisfy the conflict freedom restriction, in particular when the subtype, as during type-checking, is automatically inferred by a type-checker (for instance, when the type of a function argument is compared against the signature of the function). Recognizing this issue, we further investigated the problem and noticed that, by lifting the restriction on the subtype, we can still obtain a polynomial (quadratic) algorithm, provided that the supertype is still conflict-free. Hence, we extended our approach to a more realistic setting where the subtype is unrestricted and can be inferred by a type-checker or a compiler.

As a further step of our research, we are currently investigating an interesting optimization to our approach. We observed, indeed, that we can combine our approach with a form of *syntactic* subtyping (with linear complexity) that only looks at the syntactical form of types. We identified several applicability conditions for syntactic subtyping and designed an optimized algorithm that uses syntactic subtyping whenever is possible and falls back to our standard approach when syntactic subtyping fails, without the need of backtracking.

We are now in the process of evaluating the experimental performance of our optimized algorithm wrt our standard approach.

As a final step of our research, we are exploring the extension of the type language with intersection.

Past Research Activities

Semistructured Data and XML (Principal investigator for query processing; co-investigator for type system design) In the context of the Italian PRIN DataX Project, I studied the most prominent issues involving the storage, management, and querying of semistructured and XML data. In particular, I focused my research on two key problems: the efficient management of great amounts of XML data, with the twofold aim of extending the technology of existing DBMSs to XML data and of designing new, special-purpose techniques; and the design of a type system for analyzing the correctness of queries wrt a schema.

In the context of the first activity, I designed a logical query algebra, used for representing a significant fragment of the XQuery query language, as well as a set of physical operators for executing XML queries on top of persistently stored XML data. The physical algebra was endowed with an estimation framework for statically predicting the distribution of XML data in query results: this framework was validated on several testbeds, and proved to be quite accurate. The final result of this activity was the Xtasy data management system for XML data.

In the context of the second activity, I cooperated with Dario Colazzo in the design of a type system for a fragment of XQuery. This type system serves both the purpose of inferring an output type for a query and of checking for the correctness of a query wrt a schema, i.e., whether the structural requirements of a query match the source schema. Most of the techniques we developed in this activity are now part of the "official" type system for XQuery designed by the World Wide Web Consortium.

As a collateral activity, I cooperated with Scuola Normale Superiore in designing a query language for XML-encoded ancient and medieval documents. This query language (Tequyla) incorporates some novel techniques such the ability to filter out elements describing the graphical organization of the text and, more generally, to support various levels of markup relevance.

Web Services (Participant) Most web services are currently represented by means of WSDL documents. The purpose of this research, funded by Microsoft Corporation under the BigTop grant, was to design an alternative representation model for web services, based on XML, supporting:

- semantic web service search;
- integration with high-level distributed programming languages.

In particular, we designed a specialized query language for XML data, inspired by XQuery, which can be used to query existing web service representations from inside the Highwire programming language. The novelty of this language was its ability to seamlessly and transparently manage both ordered and unordered data, as well as to integrate with behavioral type systems.

Peer-to-Peer (P2P) Systems (XPeer) (Principal investigator) This project, funded by a FIRB GRID.IT grant, aimed at extending XML query processing facilities from a centralized context to data dispersed over the Internet. The key idea motivating the project was that the Internet itself can be seen as a formidable, massively distributed data repository, containing user-supplied information about near all knowledge fields. This repository is characterized by some ground properties, mostly induced by the behavior of data providers (typically, net-users) and by the characteristics of data being provided. These properties can be described as *heterogeneity*, *autonomy*, and *no administration*.

Till now, database technology has not been able to replicate the success of the Internet in building large or global-scale databases. As pointed out by the CAP theorem, the reasons of this failure are mostly related to common features of current database systems, such as ACID transactions, that are not adequate to a global-scale environment (and they are sometimes even an obstacle).

Therefore, the main objective of the project was the design and implementation of XPeer, a data sharing system for massively distributed XML data. XPeer allows users to publish and query *heterogeneous* information without any significant administration efforts. To this end, XPeer is based on a p2p architecture that is able to *self-organize* and *self-manage* its own administrative layers without the intervention of a database or system administrator.

Unlike similar projects, XPeer recognizes that heterogeneity is unavoidable and assists the user in *surviving* heterogeneity, i.e., XPeer assumes that data are potentially heterogeneous and tries to dispatch any given query to any potentially relevant peer, while retaining a good degree of selectivity in query dissemination. As today data integration technology does not offer scalable and fully autonomic solutions to the problem of integrating heterogeneous data sources, XPeer supports an automatic *query-to-schema* matching technique that allows queries to be run on data with a limited degree of heterogeneity.

The contribution of XPeer is twofold. First, XPeer offers very selective query dissemination solutions, that allow the system to deliver a query only to a small superset of the peers containing relevant data, hence reducing both communication and execution costs. Second, its architectural design allows for a scalable management of heterogeneity, and can be easily adapted to more sophisticated data integration techniques based on schema mappings and query reformulation.

XPeer has been fully implemented and deployed on a local cluster. We are in the process of releasing the source code of XPeer under an open source license.

Teaching Statement

Teaching Methodology

I believe that a successful computer science course (at both undergraduate and graduate levels) should be intellectually hard. Such a course should combine a strong understanding of the theoretical aspects of the subject with a massive experimental activity. By doing so, the course should allow the students to appreciate the joy of theoretical studies, while achieving a good experience in developing projects and facing real life challenges.

As a consequence, my teaching methodology is inspired by three principles: choosing challenging course topics; combining theoretical topics with several projects and experiments; and strongly interacting with students. These principles should turn in successful, yet hard, courses, able to appeal the interest of top students.

Teaching Interests

I would feel comfortable teaching most introductory courses in computer science and a variety of advanced courses, including but not limited to:

- Undergraduate level: Introduction to Database Systems, Design and Implementation of Database Systems, Programming Languages.
- Graduate level: Advanced Implementation Techniques for Database Systems, Theory of Database Systems, Distributed Databases and Concurrency Control.

Referees

First Referee

Name: Prof. Giorgio Ghelli.
Address: Dipartimento di Informatica - Università di Pisa
Largo P. Pontecorvo 3, 56127, Pisa, Italy
email: ghelli@di.unipi.it

Second Referee

Name: Prof. Antonio Albano.
Address: Dipartimento di Informatica - Università di Pisa
Largo P. Pontecorvo 3, 56127, Pisa, Italy
email: albano@di.unipi.it

Additional Referee

Name: Prof. Dan Suciu.
Address: Computer Science & Engineering
University of Washington
Box 352350
Seattle, WA 98195-2350
email: suciu@cs.washington.edu

Bibliometrical Indexes

- H-index (as computed on 23rd January, 2009): 8
- G-index (as computed on 23rd January, 2009): 14

Publications

Journal Papers

2008

1. Dario Colazzo, Giorgio Ghelli, and Carlo Sartiani. Efficient Inclusion for a Class of XML Types with Interleaving and Counting. Accepted for publication In Information Systems Journal (Elsevier), 06/26/2008.

2006

1. Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Types for Path Correctness of XML Queries. Journal of Functional Programming, Volume 16, Issue 4, 2006.

2002

1. Dario Colazzo, Carlo Sartiani, Antonio Albano, Paolo Manghi, Giorgio Ghelli, Luca Lini, and Michele Paoli. A Typed Text Retrieval Query Language for XML Documents. *Journal of American Society for Computer Science and Technology (JASIST), Special Issue on XML and Information Retrieval*, 53(6):467–488, 2002.

Technical Reports

2007

1. Giorgio Ghelli, Dario Colazzo and Carlo Sartiani. Efficient Inclusion for a Class of XML Types with Interleaving and Counting. Technical report TR-07-14, Dipartimento di Informatica, Università di Pisa, June 2007.
2. Dario Colazzo and Carlo Sartiani. Efficient Subtyping for Unordered XML Types. Technical report TR-07-03, Dipartimento di Informatica, Università di Pisa, February 2007.

Conference/Workshop Papers

2009

1. Dario Colazzo, Giorgio Ghelli and Carlo Sartiani. Efficient Asymmetric Inclusion Between Regular Expression Types. To Appear In *Proceedings of the 12th International Conference on Database Theory (ICDT 2009), St. Petersburg, Russia, 23-25 March 2009*

2008

1. Giorgio Ghelli, Dario Colazzo and Carlo Sartiani. Linear Time Membership for a Class of XML Types with Interleaving and Counting. In *Proceedings of the ACM 17th Conference on Information and Knowledge Management (CIKM 2008), Napa Valley, California, October 26-30, 2008*
2. Giorgio Ghelli, Dario Colazzo and Carlo Sartiani. Linear Time Membership for a Class of XML Types with Interleaving and Counting (Extended Abstract). In *Proceedings of the International Workshop on Programming Languages and XML (PLAN-X 2008), colocated with POPL 2008, San Francisco, California - 9 January 2008.*

2007

1. Giorgio Ghelli, Dario Colazzo and Carlo Sartiani. Efficient Inclusion for a Class of XML Types with Interleaving and Counting. In *Proceedings of the Eleventh International Symposium on Database Programming Languages (DBPL 2007), Vienna, Austria. September 23-24, 2007.*
2. Giovanni Conforti, Giorgio Ghelli, Paolo Manghi and Carlo Sartiani. Scalable Query Dissemination in XPeer (Extended Version). In *Proceedings of the Eleventh International Database Engineering & Applications Symposium (IDEAS 2007), Banff, Canada. September 6-8, 2007.*
3. Dario Colazzo and Carlo Sartiani. XML Type Projection: A Maximum Flow Approach (Extended Abstract). In *Proceedings of the Fifteenth Italian Symposium on ADVANCED DATABASE SYSTEMS - Sistemi Evoluti per Basi di Dati (SEBD-2007), Torre Canne, Italy, 17-20 June, 2007.*
4. Giovanni Conforti, Giorgio Ghelli, Paolo Manghi and Carlo Sartiani. Scalable Query Dissemination in XPeer. In *Proceedings of the ICDT Workshop on Emerging Research Opportunities in Web Data Management (EROW 2007), Barcelona, Spain. January 13, 2007.*

2006

1. Dario Colazzo and Carlo Sartiani. An Efficient Algorithm for XML Type Projection. In *Proceedings of the Eighth ACM-SIGPLAN International Symposium on Principles and Practice of Declarative Programming (PPDP'06), Venice, Italy, 10-12 July 2006.*
2. Dario Colazzo and Carlo Sartiani. Mapping Maintenance in XML P2P Databases (Extended Abstract). In *Proceedings of the Fourteenth Italian Symposium on ADVANCED DATABASE SYSTEMS - Sistemi Evoluti per Basi di Dati (SEBD-2006), Portonovo, Italy, 18-21 June, 2006.*
3. Carlo Sartiani. A Query Algebra for XML P2P Databases. In *Proceedings of the Eleventh International Workshop on Foundations of Models and Languages for Data and Objects (FMLDO), (co-located with EDBT) 30-31 March, 2006. Munich, Germany.*

2005

1. Dario Colazzo and Carlo Sartiani. Mapping Maintenance in XML P2P Databases, 2005. In *Proceedings of the Tenth International Symposium on Database Programming Languages (DBPL-2005), (co-located with VLDB) 28-29 August, 2005. Trondheim, Norway.*
2. Carlo Sartiani. Correctness of Query Results in XML P2P Databases. In *Proceedings of the Thirteenth Italian Symposium on ADVANCED DATABASE SYSTEMS - Sistemi Evoluti per Basi di Dati (SEBD-2005), Bressanone, Italy, 19-22 June, 2005*
3. Dario Colazzo and Carlo Sartiani. Detecting Corrupted Schema Mappings in XML P2P Databases. In *Proceedings of the Thirteenth Italian Symposium on ADVANCED DATABASE SYSTEMS - Sistemi Evoluti per Basi di Dati (SEBD-2005), Bressanone, Italy, 19-22 June, 2005*

4. Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Types for Path Correctness of XML Queries. In *Proceedings of the Thirteenth Italian Symposium on ADVANCED DATABASE SYSTEMS - Sistemi Evoluti per Basi di Dati (SEBD-2005), Bressanone, Italy, 19-22 June, 2005*
5. Dario Colazzo and Carlo Sartiani. Typechecking Queries for Maintaining Schema Mappings in XML P2P Databases, 2005. In *Proceedings of PLAN-X 2005: Programming Language Technologies for XML, colocated with POPL 2005.*

2004

1. Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Types for Path Correctness of XML Queries. In *Proceedings of the Nineth International Conference on Functional Programming (ICFP), Snowbird, Utah, September 19-22, 2004*, 2004.
2. Carlo Sartiani. On the Correctness of Query Results in XML P2P Databases. In *Proceedings of the Fourth IEEE International Conference on Peer-to-Peer Computing (P2P2004), 25-27 August, Zurich, Switzerland, 2004*, 2004.
3. Giovanni Conforti, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. A self-organizing XML P2P database system. In *Proceedings of the Twelfth Italian Symposium on Advanced Database Systems, SEBD 2004, S. Margherita di Pula, Cagliari, Italy, June 21-23, 2004*, 2004.
4. Carlo Sartiani. A Query Algebra for XML P2P Databases. In *Proceedings of the Thirteenth International World Wide Web Conference, WWW 2004 (Alternate Track Papers & Posters), New York, NY, May 17-22, 2004*, 2004.
5. Carlo Sartiani, Giorgio Ghelli, Paolo Manghi, and Giovanni Conforti. XPeer: A self-organizing XML P2P database system. In *Proceedings of the First EDBT Workshop on P2P and Databases (P2P&DB 2004), 2004*, 2004.

2003

1. Carlo Sartiani. A General Framework for Estimating XML Query Cardinality. In *Proceedings of the 9th International Workshop on Data Bases and Programming Languages - DBPL03, Potsdam, Germany, September 6-8, 2003*, 2003.
2. Carlo Sartiani. A Framework for Estimating XML Query Cardinality. In *Proceedings of the Sixth International Workshop on the Web and Databases (WebDB 2003), San Diego, California, June 12-13, 2003*, 2003.
3. Carlo Sartiani. Evaluating Nested Queries on XML Data. In *Proceedings of the 7th International Database Engineering and Applications Symposium (IDEAS 2003), Hong Kong, China, July 16-18, 2003*, 2003.

2002

1. Carlo Sartiani and Antonio Albano. Yet Another Query Algebra For XML Data. In Mario A. Nascimento, M. Tamer Özsu, and Osmar Zaïane, editors, *Proceedings of the 6th International Database Engineering and Applications Symposium (IDEAS 2002), Edmonton, Canada, July 17-19, 2002*, 2002.
2. Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Types For Correctness of Queries Over Semistructured Data. In *Proceedings of the Fifth International Workshop on the Web and Databases (WebDB 2002), Madison, Wisconsin, June 6-7, 2002*, 2002.
3. Giovanni Conforti, Giorgio Ghelli, Antonio Albano, Dario Colazzo, Paolo Manghi, and Carlo Sartiani. The Query Language TQL. In *Proceedings of the Fifth International Workshop on the Web and Databases (WebDB 2002), Madison, Wisconsin, June 6-7, 2002*, 2002.

2000

1. Antonio Albano, Dario Colazzo, Giorgio Ghelli, Paolo Manghi and Carlo Sartiani. A Type System For Querying XML Documents. In *Proceedings of ACM SIGIR 2000 Workshop On XML and Information Retrieval, Athens, Greece, 2000.*
2. Carlo Sartiani. Basi di Dati Semistrutturate e XML. In *XML Italia 2000, Atti del Convegno, Pisa, 9-10 Maggio, 2000, Consiglio Nazionale delle Ricerche, Istituto per le Applicazioni Telematiche ISBN 88-7781-272-9.*