

HYPERGRAPHS AS A CONCEPTUAL MODEL FOR MEDIATING TRANSLATIONS BETWEEN DATA MODELS

ARUNA BANSAL



AMAR NATH & SHASHI KHOSLA SCHOOL OF INFORMATION
TECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY DELHI
MARCH 2023

© Indian Institute of Technology Delhi - 2023
All rights reserved.

HYPERGRAPHS AS A CONCEPTUAL MODEL FOR MEDIATING TRANSLATIONS BETWEEN DATA MODELS

by

ARUNA BANSAL

Amar Nath & Shashi Khosla School of Information Technology

Submitted

in fulfillment of the requirements of the degree of
Doctor of Philosophy

to the



Indian Institute of Technology Delhi

March 2023

DEDICATED TO

My family.

Thank you for your eternal support.

Certificate

This is to certify that the thesis titled **HYPERGRAPHS AS A CONCEPTUAL MODEL FOR MEDIATING TRANSLATIONS BETWEEN DATA MODELS** being submitted by **Ms. ARUNA BANSAL** for the award of **Doctor of Philosophy in Information Technology** is a record of bona fide work carried out by her under my guidance and supervision at the Amar Nath and Shashi Khosla School of Information Technology, Indian Institute of Technology Delhi. The thesis has reached the standard fulfilling the requirements of the regulations related to the degree. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma.

Kolin Paul

Professor

Dept. of Computer Science and Engg. and

Head of AN & SK School of Information Technology

Indian Institute of Technology Delhi

New Delhi- 110016

Acknowledgements

The research experience at IIT Delhi for this degree has been an enlightening, enriching, insightful, frustrating, and painful experience. I am indebted to several people who contributed directly or indirectly on my way to realizing this thesis. I take this opportunity to thank as many of them as I can. Although I might have missed many names, I remain grateful for every piece of advice, every word of encouragement, and every blessing that has kept me going.

First of all, I am very thankful to God for letting me through all the difficulties and giving me the ability and strength to do my research and finish it. I would not have dreamt of embarking on this journey without his blessings.

I wish to express my deepest gratitude to my advisor, **Prof. Sanjiva Prasad**, for his invaluable guidance, fruitful discussion, and support during this research. I also thank him for his constant feedback on my presentations, which has helped me refine the skills necessary to present my research work. His criticism and thought-provoking review motivated me towards continuous improvement. I also acknowledge him for all the travel-related support, thanks to which I could attend various conferences & workshops and present my research papers.

I would like to express my great appreciation to my research committee members, **Prof. Kolin Paul**, **Srikanta Bedathur**, and **Niladri Chatterjee**, for their insightful comments, excellent suggestions, and feedback on my work. A special thanks go to Prof. Kolin Paul, who stepped

in as my caretaker supervisor, and Prof. Srikanta Bedathur for giving his expert and constructive feedback on my work and thesis to give it more value. I would like to express my deep gratitude to Prof. Niladri Chatterjee for constantly encouraging and morally supporting me when I was in the worst time of my Ph.D. journey and for guiding me with the right administrative decisions. His words of encouragement helped me maintain the required levels of motivation and enabled me to overcome disappointing phases. I am also thankful to **Prof. Sorav Bansal** and **Dr. B. Chandra** for their helpful suggestions during the initial phase of the research. I am immensely thankful to the Dean of Academics, **Prof. Narayanan D. Karur**, for supporting me with the various administrative tasks required to submit my thesis.

Thanks to the financial assistance I was given, which made my Ph.D. journey relaxed. I acknowledge the financial support given by IIT Delhi, especially the former Dean, **Prof. Shantanu Chaudhary**, and the former Associate Dean of Students' Welfare, **Prof. Anjalee Multani**, for approving me for the institute fellowship, and **Prof. Prem Kalra** for providing me with the project assistantship for a short duration.

I am grateful to my thesis examiners and anonymous conference reviewers (who reviewed my research papers) for providing detailed comments and suggestions that helped me add more value to my thesis and finalize it more interestingly.

I acknowledge the help provided by the IIT administrative and technical staff for their often underappreciated efforts. Last but not least, I would like to thank the lab staff, Manju Chopra and Vandana Ahluwalia, for all the lab-related technicalities, and the administrative staff at the Department of Computer Science and School of Information Technology, especially Ramesh Kumar and Suresh, for their continuous support and for facilitating various administrative matters.

I am immensely blessed to have had a wonderful and great set of friends at IIT Delhi with whom I shared my joys and sorrows during the learning years of my life. My earnest gratitude to my friends Himanshu Gandhi, Dr. Geeta Yadav, and Dr. Madhulika Mohanty, who acted as my informal co-guides during my research publication & thesis writing phases, for their motivation

and professional guidance and for proofreading my manuscripts and providing feedback. I am also grateful to my friends Pratibha Jagnere, Dr. Dipanjan Chakarvarty, Dr. Piyas Kedia, Vijay Kumar, Priyanka Chouhan, Dr. Chandrika Bhardwaj, Dr. Shubhani Gupta, Aarushi Ranjan, and Dr. Brity Baby for their companionship. The never-forgetful time spent with my friends for the never-ending discussions, presentation rehearsals, manuscript reviews, experience-sharing sessions, and wonderful lab environment kept me motivated and reinvigorated throughout the journey. Our professional and personal discussions and numerous cups of tea throughout these years provided great companionship and helped maintain the required energy and enthusiasm.

My daughter Tiya was in the early phase of her schooling, and my son Pulkit was born while I was pursuing my Ph.D. It was very challenging for me to deal with. I pursued both the motherhood and Ph.D. journeys together. I could handle both things because of my strong pillar, my husband, Atul Bansal. I am blessed with his love, efforts, faith in me, and support throughout the journey, which enabled me to overcome all the ups and downs of the research. He has always encouraged me to keep dreaming and setting high goals and has helped me in every way possible throughout this time. Additionally, I am ever grateful to the other pillars of my success, my in-laws Satvir Singh Bansal and Nirmala Devi, for sustaining my dream by supporting my decision to do a Ph.D. and for their constant, valuable encouragement and understanding through the journey. A special thanks to my mother-in-law for helping and encouraging me during the difficult time of my Ph.D. and helping me when I have to attend conferences. I am thankful to them for never imposing family responsibilities and understanding when I could not attend a family function or do household chores on time.

I express my gratitude to my parents, Kishan Chand Gupta (who left the world with a dream of seeing my name with “Dr.”) and Laxmi Gupta, for their love, confidence, and support. They have been extremely supportive of all my endeavors from a very young age, and I am grateful to them for all the warmth. I would also like to thank my siblings for their unconditional love and trust in me, which helped me finish this thesis.

This thesis is a result of not only my hard work but also God’s blessings, my father’s ambition, my mother’s encouragement, my husband’s dedication, faith, love, and care, my parent-in-laws’

dreams, support, and immense patience, and my kids' joyful presence. Therefore, I would like to dedicate this thesis to my family, including all of them.

Aruna Bansal

Abstract

Hypergraphs are rich candidate structures that provide a natural mathematical way to represent the richness of diversified data and complex relationships in hierarchical, relational, navigational, and semi-structured settings, e.g., bibliographic paper submissions, mHealth, and social media applications. The representation and modeling of information in terms of entities and relationships are challenging in hypergraphs. Further, representing and formulating queries on such hypergraphs, and obtaining results for those queries are other challenges. Some hypergraph-oriented representations and data models are mentioned in the literature, but they do not pay enough attention to entities and relationships. Additionally, the existing approaches work with low-level models and do not unify the essence of other data models.

In this thesis, a *Hypergraph Data Model* (or *HgDM*) proposes a notion of *schematically well-formed typed nested* hypergraphs that integrates the benefits of various data models for representing data, metadata, and complex relationships. The HgDM is formally presented as a high-level hypergraph data model that can be directly represented in object-oriented frameworks. The queries are considered diagrams of ideas, and objects in a database are matched to

the diagrams.

Hypergraphs have a trivial embedding from other data models, such as hyperedges can directly represent relations of a relational model, edges of graphs and tree structures, and objects of object-oriented databases. However, generalizing hypergraphs in other data models and vice versa in a lossless way is another challenge for which the thesis proposes a solution by presenting translation patterns. The thesis proposes *Hypergraph Mediator* (or *HgMed*) for hypergraph-based mediation to other data models. The HgMed presents a notion of simulation relations between schematic diagrams of different data models modulo path morphisms. On the basis of this, the correctness of representations and translations between schema and schematic queries can be determined. In order to preserve structural properties that may be otherwise lost in a translation, we use annotations that help compensate for the lossiness of translations.

Furthermore, hypergraph-oriented applications and databases may operate in distributed environments with a requirement for availability while coping with high network latencies. Replication is the commonly used approach to achieving a high degree of availability, facilitating local query processing. However, replication requires expensive (often infeasible) concurrency control to ensure consistency. Moreover, administrative and security policies may prohibit certain parts of the database from being fully replicated at certain sites. The thesis proposes a solution to fully or partially replicate hypergraphs across multiple replicas as a data type in conformance with certain distribution policies while ensuring data availability, network latency, and a weak notion of consistency. In this direction, the thesis presents *Hypergraphs as Conflict-free Replicated Data Type* (or *HgCRDT*), and *Hypergraphs as Conflict-free Partially Replicated Data Type* (or *HgCPRDT*) that ensure strong eventual consistency.

Apart from these works, the thesis proposes and presents *HgQL* that semantically and syntactically integrates the richness of hypergraphs in GraphQL (a query language used for APIs). Traditional GraphQL supports graphs and hierarchical structures. The thesis enhances the functionality of GraphQL by embracing the richness of hypergraphs so that GraphQL can be employed with hypergraph-oriented databases.

Additionally, the thesis presents details of a prototype implementation system (named *Hypergraph Queriable* system or *HQ*) that demonstrates the applicability of the above approaches.

In brief, the thesis aims at presenting *hypergraphs* as a unifying conceptual data model within which different data models (i.e., hierarchical, graph, and relational data models) can be expressed, keeping most of their strengths while avoiding most of their shortcomings. The thesis also aims at leveraging this conceptual hypergraph data model to generalize other data models by mediating to and from hypergraphs in a lossless way. The thesis represents hypergraphs as a data type to update information among multiple replicas of a distributed domain. Therefore, the thesis proposes and presents HgDM for hypergraph-based representation and modeling; HgMed for hypergraph-based mediation to other data models; HgCRDT and HgCPRDT for updating hypergraphical information in distributed settings; HgQL for integrating schematic hypergraphs into GraphQL; and HQ, a prototype system that implements our proposed approaches.

सार

हाइपरग्राफ समृद्ध उम्मीदवार संरचनाएं हैं जो पदानुक्रमित, संबंधपरक, नेविगेशनल और अर्ध-संरचित सेटिंग्स में विविध डेटा और जटिल संबंधों की समृद्धि का प्रतिनिधित्व करने के लिए एक प्राकृतिक गणितीय तरीका प्रदान करती हैं, उदाहरण के लिए, ग्रंथ सूची पेपर सबमिशन, एमहेल्थ और सोशल मीडिया एप्लिकेशन। हाइपरग्राफ में संस्थाओं और संबंधों के संदर्भ में सूचना का प्रतिनिधित्व और मॉडलिंग चुनौतीपूर्ण है। इसके अलावा, ऐसे हाइपरग्राफ पर प्रश्नों का प्रतिनिधित्व करना और तैयार करना, और उन प्रश्नों के परिणाम प्राप्त करना अन्य चुनौतियाँ हैं। साहित्य में कुछ हाइपरग्राफ-उन्मुख अभ्यावेदन और डेटा मॉडल का उल्लेख किया गया है, लेकिन वे संस्थाओं और संबंधों पर पर्याप्त ध्यान नहीं देते हैं। इसके अतिरिक्त, मौजूदा दृष्टिकोण निम्न-स्तरीय मॉडल के साथ काम करते हैं और अन्य डेटा मॉडल के सार को एकीकृत नहीं करते हैं।

इस थीसिस में, एक *Hypergraph Data Model* (या *HgDM*) योजनाबद्ध रूप से अच्छी तरह से निर्मित टाइप किए गए नेस्टेड हाइपरग्राफ की एक धारणा का प्रस्ताव करता है जो डेटा, मेटाडेटा, और का प्रतिनिधित्व करने के लिए विभिन्न डेटा मॉडल के लाभों को एकीकृत करता है। जटिल रिश्ते। *HgDM* को औपचारिक रूप से एक उच्च-स्तरीय हाइपरग्राफ डेटा मॉडल के रूप में प्रस्तुत किया जाता है जिसे सीधे ऑब्जेक्ट-ओरिएंटेड फ्रेमवर्क में प्रदर्शित किया जा सकता है। प्रश्नों को विचारों का आरेख माना जाता है, और डेटाबेस में वस्तुओं का आरेखों से मिलान किया जाता है।

हाइपरग्राफ में अन्य डेटा मॉडल से एक तुच्छ एम्बेडिंग होती है, जैसे कि हाइपरडेज सीधे रिलेशनल मॉडल के संबंधों, ग्राफ के किनारों और ट्री संरचनाओं और ऑब्जेक्ट-ओरिएंटेड डेटाबेस की वस्तुओं का प्रतिनिधित्व कर सकते हैं। हालांकि, अन्य डेटा मॉडल में हाइपरग्राफ को सामान्य बनाना और इसके विपरीत दोषरहित तरीके से एक और चुनौती है जिसके लिए थीसिस अनुवाद पैटर्न पेश करके समाधान का प्रस्ताव करती है। थीसिस अन्य डेटा मॉडल के लिए हाइपरग्राफ-आधारित मध्यस्थता के लिए *Hypergraph Mediator* (या *HgMed*) प्रस्तावित करती है। *HgMed* विभिन्न डेटा मॉडल मॉडुलो पथ आकारिकी के योजनाबद्ध आरेखों के बीच सिमुलेशन संबंधों की धारणा प्रस्तुत करता है। इसके आधार पर, स्कीमा और योजनाबद्ध प्रश्नों के बीच अभ्यावेदन और अनुवाद की शुद्धता निर्धारित की जा सकती है। संरचनात्मक गुणों को संरक्षित करने के लिए जो अनुवाद में अन्यथा खो सकते हैं, हम एनोटेशन का उपयोग करते हैं जो अनुवाद के नुकसान की भरपाई करने में मदद करते हैं।

इसके अलावा, हाइपरग्राफ-उन्मुख एप्लिकेशन और डेटाबेस उच्च नेटवर्क लेटेंसी से निपटने के दौरान उपलब्धता की आवश्यकता के साथ वितरित वातावरण में काम कर सकते हैं। स्थानीय क्लेरी प्रसंस्करण को सुविधाजनक बनाने, उच्च स्तर की उपलब्धता प्राप्त करने के लिए प्रतिकृति आमतौर पर उपयोग की जाने वाली विधि है। हालाँकि, प्रतिकृति को निरंतरता सुनिश्चित करने के लिए महंगे (और अक्सर अव्यवहार्य) संगामिति नियंत्रण की आवश्यकता होती है। इसके अलावा, प्रशासनिक और सुरक्षा नीतियां डेटाबेस के कुछ हिस्सों को कुछ साइटों पर पूरी तरह से दोहराए जाने से रोक सकती हैं। थीसिस डेटा उपलब्धता, नेटवर्क विलंबता और स्थिरता की कमजोर धारणा को सुनिश्चित करते हुए कुछ वितरण नीतियों के अनुरूप डेटा प्रकार के रूप में कई प्रतियों में हाइपरग्राफ को पूरी तरह या आंशिक रूप से दोहराने के लिए एक समाधान प्रस्तावित करता है। इस दिशा में, थीसिस *Hypergraphs as Conflict-free Replicated Data Type* (या *HgCRDT*), और *Hypergraphs as Conflict-free Partially Replicated Data Type* (या *HgCPRDT*) प्रस्तुत करती है। जो मजबूत अंतिम स्थिरता सुनिश्चित करते हैं।

इन कार्यों के अलावा, थीसिस *HgQL* को प्रस्तावित और प्रस्तुत करता है जो सिमेंटिकली और सिंटैक्टिक रूप से ग्राफक्यूएल (एपीआई के लिए उपयोग की जाने वाली एक क्लेरी भाषा) में हाइपरग्राफ की समृद्धि को एकीकृत करता है। पारंपरिक ग्राफक्यूएल ग्राफ और पदानुक्रमित संरचनाओं का समर्थन करता है। थीसिस हाइपरग्राफ की समृद्धि को अपनाकर ग्राफक्यूएल की कार्यक्षमता को बढ़ाती है ताकि ग्राफक्यूएल को हाइपरग्राफ-उन्मुख डेटाबेस के साथ नियोजित किया जा सके।

इसके अतिरिक्त, थीसिस एक प्रोटोटाइप कार्यान्वयन प्रणाली (नामित *Hypergraph Queriable system* या *HQ* का विवरण प्रस्तुत करता है जो उपरोक्त दृष्टिकोणों की प्रयोज्यता को प्रदर्शित करता है।

संक्षेप में, थीसिस का उद्देश्य *hypergraphs* को एक एकीकृत वैचारिक डेटा मॉडल के रूप में प्रस्तुत करना है, जिसके भीतर विभिन्न डेटा मॉडल (यानी, पदानुक्रमित, ग्राफ, और रिलेशनल डेटा मॉडल) को व्यक्त किया जा सकता है, जबकि उनकी अधिकांश शक्तियों से परहेज करते हुए उनकी अधिकांश शक्तियों को बनाए रखा जा सकता है। कमियों। थीसिस का उद्देश्य इस वैचारिक हाइपरग्राफ डेटा मॉडल का लाभ उठाना है ताकि दोषरहित तरीके से हाइपरग्राफ से मध्यस्थता करके अन्य डेटा मॉडल को सामान्य बनाया जा सके। थीसिस एक वितरित डोमेन के कई प्रतिकृतियों के बीच जानकारी को अद्यतन करने के लिए डेटा प्रकार के रूप में हाइपरग्राफ का प्रतिनिधित्व करती है। इसलिए, थीसिस हाइपरग्राफ-आधारित प्रतिनिधित्व और मॉडलिंग के लिए एचजीडीएम का प्रस्ताव और प्रस्तुत करता है; अन्य डेटा मॉडलों के लिए हाइपरग्राफ-आधारित मध्यस्थता के लिए *HgMed*; वितरित सेटिंग्स में हाइपरग्राफिकल जानकारी को अपडेट करने के लिए एचजीसीआरडीटी और एचजीसीपीआरडीटी; ग्राफक्यूएल में योजनाबद्ध हाइपरग्राफ को एकीकृत करने के लिए एचजीक्यूएल; और मुख्यालय, एक प्रोटोटाइप प्रणाली जो हमारे प्रस्तावित दृष्टिकोणों को लागू करती है।

Contents

Certificate	i
Acknowledgements	iii
Abstract	vii
List of Figures	xxi
List of Tables	xxvi
Glossary	xxvii
Acronyms	xxxii
Symbols	xxxiii
1 Introduction	1
1.1 Data models	2

1.2	Hypergraphs as a unifying data model	5
1.3	Challenges associated with hypergraphs	7
1.4	Objective of the thesis	7
1.5	A summary of the contributions	8
1.6	Thesis organization	12
2	Background and Related Work	15
2.1	Review of database models	15
2.1.1	Navigational database models	16
2.1.2	Relational database models	17
2.1.3	Object-oriented database models	17
2.1.4	NoSQL database models	18
2.2	Hypergraphs in databases	20
2.2.1	Hypergraphs in relational schema analyses	21
2.2.2	Hypergraphs mediated translations	22
2.2.3	Hypergraphs as representational frameworks	26
2.3	Other relevant work in databases	30
2.3.1	Property graphs	30
2.3.2	GraphQL	31
2.4	Review of Conflict-free Replicated Data Types or CRDTs	33

2.4.1	Conflict-free Replicated Data Types or CRDTs	33
2.4.2	Replication and implementation approaches for CRDTs	35
2.4.3	Data types in the portfolio of CRDTs	37
2.5	CRDTs: Related Work	39
2.5.1	Advanced CRDTs	40
2.5.2	CRDTs in databases	41
2.5.3	CRDTs with partial replication	42
3	Representation of Hypergraphs	45
3.1	Introduction	45
3.2	Organization of the chapter	46
3.3	Hypergraph-based data representation	47
3.4	Hypergraphs as a data type	54
3.5	Representation of schematic typed hypergraph objects	54
3.6	Queries on hypergraphs	57
3.7	Implementation Details	63
3.7.1	Building a hypergraph database	63
3.7.2	Query processing on hypergraphical objects	67
3.8	Summary of the Chapter	68

4	Formalising a Hypergraph Data Model (HgDM) and Hypergraph Queries	71
4.1	Types and schema in HgDM	71
4.2	Hypergraph Data Model or HgDM	74
4.2.1	Formalizing the hypergraph schema	74
4.2.2	Hypergraph assignment functions	78
4.2.3	Formalizing hypergraph instances	79
4.2.4	Defining the HgDM	80
4.3	Integrity constraints associated with HgDM	81
4.3.1	Other constraints	83
4.4	Formalising queries in hypergraphs	84
4.5	Summary of the Chapter	86
5	Translations between Hypergraphs and Other Data Models in HgMed	87
5.1	Introduction	87
5.2	Organization of the chapter	88
5.3	Translation patterns in HgMed	88
5.3.1	Schematic mappings from hypergraph into other databases	88
5.3.2	Reverse mappings from the transformed databases into a hypergraph schema	94
5.4	Correctness of translations	95

5.5	Paths, traces and simulations in hypergraphs	99
5.6	Translation-related implementation details	104
5.6.1	Translations between hypergraphs and a relational database	105
5.6.2	Translations between hypergraphs and semi-structured documents	106
5.6.3	Translation between hypergraphs and a typed property graph model	106
5.6.4	Reverse translations from the translated data models into a hypergraph	106
5.6.5	Experimental Evaluation	107
5.7	Summary of the Chapter	107
6	Hypergraph Data Model in Distributed Settings	109
6.1	Introduction	109
6.2	Contributions and organization of the chapter	110
6.3	Design of <i>Hypergraphs as a CRDT (HgCRDTs)</i>	111
6.4	Specification of <i>HgCRDTs</i>	113
6.4.1	Query operations in HgCRDTs	115
6.4.2	Update operations in HgCRDTs	116
6.5	Introducing Conflict-free Partially Replicated Hypergraphs – <i>HgCPRDTs</i>	120
6.5.1	Auxiliary operations in HgCPRDT	121
6.5.2	Query operations in HgCPRDT	122
6.5.3	Update Operations in HgCPRDT	122

6.6	Proof of correctness	127
6.7	An example using the HgCPRDT approach	131
6.8	Implementation details	135
6.9	Summary of the Chapter	135
7	Conclusion and Future Directions	137
7.1	Future Directions	139
	Bibliography	141
A	Implementation Details	157
A.1	System details of HQ	157
A.2	Building a hypergraph database	158
A.2.1	Formulation of a hypergraph schema	158
A.2.2	Mapping instructions for generating hypergraph objects	161
A.2.3	Storing typed hypergraphical objects	162
A.3	Executing queries on a hypergraph database	163
A.3.1	Query formulation	163
A.3.2	Queries processing on hypergraph objects	165
A.4	Ensuring consistency of hypergraphical objects	167

B HgQL: Supporting the Schematic Hypergraphs	169
B.1 Formalizing hypergraph schema in HgQL	170
B.1.1 Example of an HgQL schema	171
B.1.2 Well-formed HgQL schemas	172
B.2 Formalizing hypergraph queries in HgQL	173
B.2.1 Query representation in HgQL	173
B.2.2 Query response in HgQL	174
B.2.3 Well-formed HgQL queries	175
B.2.4 Properties of HgQL Queries	175
B.2.5 Query evaluation in HgQL	176
List of Publications	181
Biography	183

List of Figures

1.1	Bibliographic dataset in three different database models, i.e., a tree showing XML data (top diagram), a property graph (bottom left diagram), and relational data (bottom right diagram).	4
1.2	Hypergraph is an embedding and a generalization of other data models.	6
1.3	Summary of thesis contributions.	9
2.1	Representation of a traditional CRDT specification consisting of payloads and different operations.	35
2.2	Specification of State-based 2P-Set	38
3.1	Representation of bibliographic article structure into a hypergraph schema capturing <i>article-coauthors</i> and <i>article-InJournalIssue</i> relationships.	47
3.2	Representing two bibliographic article records as hypergraph instances following the schema of Figure 3.1.	48
3.3	The bibliographic dataset in XML.	49
3.4	The bibliographic dataset in a property graph.	50
3.5	An unnormalized relational table capturing the bibliographic data.	51

3.6	Capturing the bibliographic data in three relational tables in First-NF with some decomposition, violating functional dependencies, and ignoring entity-relationship existence.	51
3.7	The highly decomposed relational data requiring lots of joins.	52
3.8	Different design choices to represent authors in the hypergraph.	53
3.9	A class diagram showing hypergraph classes for the hypergraph schema of Figure 3.1.	56
3.10	Query example: a query in hierarchical, semi-structured, graph, and relational models asking for “ <i>getting the title of all the articles authored by Radhia Cousot and Patrick Cousot where either only both of them have coauthored the article, or Radhia Cousot was the first author, or Radhia Cousot appeared in the author list just before Patrick Cousot</i> ”.	59
3.11	Sample of an input dataset.	60
3.12	Query example: query results in hierarchical, semi-structured, graph, and relational models, for the query of Figure 3.10.	60
3.13	Query example: graphical representation of a query and its results on hypergraphs equivalent to Figure 3.10 and Figure 3.12.	61
3.14	Query example: a possible navigation path on the hypergraph query (refer to Figure 3.13) selecting the <code>Author</code> , <code>Authors</code> and the <code>Article</code> objects based on the given constraints, and projecting the <code>title</code> of the matched <code>Article</code> objects.	62
3.15	Translating a source XML dataset to a Hypergraph Database.	64
3.16	Translation of a hypergraph schema (Figure 3.1 into the dynamically generated hypergraph classes (Figure 3.9).	65
3.17	Generating hypergraph objects using a source dataset, hypergraph schema, and mappings.	66

3.18	Query Processing on hypergraph objects.	68
4.1	Collections supported in hypergraphs.	72
4.2	TYPES for the hypergraph of Figure 3.1.	73
4.3	Mapping of the hypergraph schema (Figure 3.1) into formalized hypergraph schema defined over the types given in Figure 4.2.	77
4.4	Example of other hypergraph schema defined over the types given in Figure 4.2.	77
4.5	Instantiating a few hypergraph instances using assignment functions.	79
4.6	Formalized hypergraph instances obtained as a result of assignments (refer to Figure 4.5).	80
4.7	Hypergraph query in First-Order Logic representing the query diagram as shown in Figure 3.13.	84
4.8	Sub-hypergraph showing the query matchings.	85
5.1	Semantic mappings from hypergraphs into other database models.	90
5.2	Relational database diagram showing schema of the transformed tables.	91
5.3	DTD obtained from the hypergraph schema translation.	92
5.4	Property graph classes generated from the hypergraph schema.	93
5.5	Our transformation approach.	95
5.6	Encoding a hypergraph into XML and hierarchical structures.	97
5.7	Embeddings: Reverse encoding from the transformed XML variants into hypergraph structures.	98
5.8	Translation from hypergraph into other data models.	105

6.1	Hypergraph data model in distributed settings: a summary of the contribution.	110
6.2	A broad overview of the HgCRDT specification consisting of a payload and different types of operations.	114
6.3	A broad overview of the HgCPRDT specification consisting of a payload and different types of operations.	120
6.4	Example: Addition of reviewers in the hypergraph schema capturing the bibliographic article.	132
6.5	The partial distribution of an in-process article among three remote sites using HgCPRDT operations (mentioned as a number with details given in Table 6.10).	134
6.6	Payload at different replicas before and after executing all the operations of Table 6.10.	135
A.1	Example showing hypergraph schema definitions syntactically in HgQL.	159
A.2	Mapping instructions to convert an XML document into hypergraph schemas for further generating schema instances.	162
A.3	An example of a hypergraph query in HgQL.	164
A.4	An example of a hypergraph query result in HgQL.	165
B.1	Semantic mapping of hypergraph schema definitions (refer to Figure A.1) as per HgQL schema definition.	172
B.2	Examples showing non-equivalent hypergraph queries (left side) with their results (right side). The top left and the bottom left queries are equivalent to $\beta\{\beta'(\sigma)\{\psi'\}\}$, and $\beta(\sigma)\{\beta'\{\psi'\}\}$ query expressions, respectively.	178
B.3	Semantics of an HgQL query.	179

List of Tables

2.1	Applications of Hypergraphs	20
2.2	A summary of the significant work related to the hypergraph-based representations.	26
2.3	A summary of the significant work in CRDTs.	39
2.4	A summary of the CRDT's related work with a comparison.	40
3.1	Operations on Hypergraphs	55
3.2	Experimental evaluation: generating hypergraph objects from the XML article records. Note that the processing time is an average of three runs.	67
6.1	Query Operations in HgCRDT	115
6.2	Add Operations in HgCRDT	117
6.3	Remove Operations in HgCRDT	118
6.4	Modify Operation in HgCRDT	119
6.5	Payload and Auxiliary Operations in HgCPRDT	121
6.6	Add Operations in HgCPRDT	124

6.7	Remove Operations in HgCPRDT	125
6.8	Modify Operation in HgCPRDT	126
6.9	Distribution policy to share atoms with replica S1, S2, and S3.	131
6.10	HgCPRDT operations initiated and shared by a source replica to downstream replicas for Figure 6.5.	133

Glossary

DTD entity	A mechanism to define replacement values..
XML element	A container holding other elements, text, attributes, and/or media objects..
XML attribute	A property of an XML element..
Hypergraph entity	A real-time entity having its own existence and is independent on a relationship..
Hypergraph relationship	A complex entity, a role, or a structured relationship linking other simple or complex entities that are required to build the relationships.
Hypergraph	A collection of a set of hyperedges, and a set of vertices.
Hypergraph Vertex	A representation of a real-time entity or a simple entity.
Hyperedge	A representation of a complex entity, a role, or a relationship. A hyperedge consists of a set of intrinsic and extrinsic attributes.
Hypergraph element	A hypergraph vertex or a hyperedge.

Hypergraph atom	An abstract name used for a vertex or a hyperedge.
Schema types in a hypergraph	A schema defined for a type having a unique schema type name, and a set of attributes.
Vertex type	A unique type that classifies vertex instances.
Vertex schema type	A schema definition used to describe a unique vertex type.
Hyperedge type	A unique type that classifies hyperedge instances.
Hyperedge schema type	A schema definition used to describe a unique hyperedge type. It constitutes a set of <i>intrinsic</i> and <i>extrinsic</i> attributes.
Scalar Type	Data types supported by a programming/database languages. We use String, Int, Float, Boolean, Char, Long, and Short as the scalar types.
Attributes in hypergraph	Fields or attributes (consisting of a name and a type) that helps in defining a vertex/hyperedge..
Internal attributes	The attributes that help to define a relationship. These attributes are of either scalar types or nested structure types.
External attributes	The attributes that define the properties that are required for a relationship or based on which or using which the relationship is built. Each extrinsic attribute points to one or more hypergraph elements of the same type.
Vertex instance	A vertex object that carries data.

Hyperedge instance	A hyperedge object that has values for its intrinsic attributes and referential-links for its extrinsic attributes.
RDF Triple	A logical statement of a relationship, in which, each predicate defines a relationship of a subject to an object.
Tree	A restricted form of a graph, are well-suited for describing parent-child relationships.
Schema	A template used to capture information from the user.
Entity-Relationship model	A model that naturally describes entities and complex relationships.
Relational algebra	A family of algebras with a well-founded semantics used for modeling and querying the records stored in relational tables.

Acronyms

API	Application Programming Interface.
CRDTs	Conflict-free Replicated Data Types.
CmRDTs	Commutative Replicated Data Types.
CvRDTs	Convergent Replicated Data Types.
DM	Data model.
HQ	Hypergraph Queriable system.
HgDM	Hypergraph Data Model.
HgCRDT	Hypergraphs as Conflict-free Replicated Data Types.
HgCPRDT	Hypergraph as Conflict-free Partially Replicated Data Types.
HQ	Hypergraph system.
HgMed	Hypergraph-based Mediations.
GraphQL	Graph Query Language.
GDB	Graph Database.
REST	Representational State Transfer.

JSON	JavaScript Object Notation.
URI	Uniform Resource Identifiers.
URL	Uniform Resource Locator.
XML	eXtensible Markup Language.
XPath	XML Path.
DTD	Document Type Definition.
FOL	First-Order Logic.
DAG	Directed Acyclic Graph.
UML	Unified Modeling Language.
P2P	Peer-to-Peer model.
Groovy	Graphically Represented Object-Oriented data model with Values.
SQL	Structured Query Language.
XQuery	XML Query Language.

Symbols

\mathbb{G}_M	Hypergraph Database Model.
\mathbb{G}_S	Hypergraph Schema.
\mathbb{G}_I	Hypergraph Instances.
\mathbb{G}_A	Hypergraph Assignment functions.
Σ_V	A set of vertex schemas.
Σ_H	A set of hyperedge schemas.
V	A set of vertex instances.
H	A set of hyperedge instances.
λ	A function that builds a vertex instance.
δ	A function that builds a hyperedge instance.
VA	A payload set for added vertices.
HA	A payload set for added hyperedges.
VR	A payload set for removed vertices.
HR	A payload set for removed hyperedges.
$\text{EDGES}(\mathbb{G}_S)$	An edge set in \mathbb{G}_S .

$x_i \xrightarrow{f} x_j$	An edge stating that (f, x_j) is an attribute field in object type x_i .
$\lambda: \text{TYPES} \rightarrow \text{OTN}$	A labelling function that maps names of object types.
$\text{trace}(p)$	A trace of a path p .
$\text{TRACES}(\mathbb{G}_S)$	A set of all traces of hypergraph schema \mathbb{G}_S .
$\mathbb{G}_{S1} \succeq \mathbb{G}_{S2}$	\mathbb{G}_{S1} simulates \mathbb{G}_{S2} .
$T_1 \sqsubseteq T_2$	Object subtyping.
$he(\text{mutable atom set } U)$	A hyperedge with a set of mutable atoms.
OTN	A set of names intended to denote objects.
FN	A set of names intended to denote the names of fields within object types.
TYPES	A set over OTN and FN .
CONS	A set of constraints.
$\phi(x_1, \dots, x_n)$	An expression used for a hypergraph query.
ψ	A query expression in HgQL.
\mathbb{S}	A hypergraph schema in HgQL.
Hyperedges	A set of hypergraph types in HgQL.
Attributes	A set of scalar type attribute names in HgQL.
References	A set or referential type attribute names in HgQL.
Constraints	A set of constraints in HgQL.
Vertex	A set of vertex types in HgQL.
Scalar	A set of scalar types in HgQL.
SubTypes	A set of sub-structure types in HgQL.

<code>Vals</code>	A set of scalar values.
<code>List</code>	A list type in HgQL.
<code>UList</code>	A list type without duplicates in HgQL.
<code>Set</code>	A set type in HgQL.
<code>MultiSet</code>	A bag type in HgQL.
<code>Non-Null</code>	Hypergraph Schema.
<i>attributes_S</i>	A function that assigns a set of scalar type fields.
<i>references_S</i>	A function that assigns a set of reference type fields.
<i>constraints_S</i>	A function that assigns a set of constraints.
<i>types_S</i>	A function that assigns scalar types.
<i>referenceTypes_S</i>	A function that assigns reference types.
<i>values</i>	A function that assigns values.
<i>rootFields_S</i>	A function that assigns fields in <i>root</i> .