Dietmar Seipel
José María Turull-Torres (Eds.)

LNCS 2942

Foundations of Information and Knowledge Systems

Third International Symposium, FoIKS 2004
Wilhelminenburg Castle, Austria, February 2004
Proceedings

Springer

Dietmar Seipel José María Turull-Torres (Eds.)

Foundations of Information and Knowledge Systems

Third International Symposium, FoIKS 2004
Wilheminenburg Castle, Austria, February 17-20, 2004
Proceedings

Springer

Volume Editors

Dietmar Seipel
University of Würzburg
Department of Computer Science
Am Hubland, 97074 Würzburg, Germany
E-mail: seipel@informatik.uni-wuerzburg.de

José María Turull-Torres
Massey University
Department of Information Systems
Private Bag 11222, Palmerston North, New Zealand
E-mail: J.M.Turull@massey.ac.nz

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

CR Subject Classification (1998): H.2, H.3, H.5, I.2.4, F.3.2, G.2

ISSN 0302-9743
ISBN 3-540-20965-4 Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Protago-TeX-Production GmbH
Printed on acid-free paper SPIN: 10983706 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 2942
Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Lecture Notes in Computer Science

For information about Vols. 1–2858
please contact your bookseller or Springer-Verlag

Vol. 2857: M.A. Nascimento, E.S. de Moura, A.L. Oliveira (Eds.), String Processing and Information Retrieval. Proceedings, 2003. XI, 379 pages. 2003.

Vol. 2858: A. Veidenbaum, K. Joe, H. Amano, H. Aiso (Eds.), High Performance Computing. Proceedings, 2003. XV, 566 pages. 2003.

Vol. 2859: B. Apolloni, M. Marinaro, R. Tagliaferri (Eds.), Neural Nets. Proceedings, 2003. X, 376 pages. 2003.

Vol. 2860: D. Geist, E. Tronci (Eds.), Correct Hardware Design and Verification Methods. Proceedings, 2003. XII, 426 pages. 2003.

Vol. 2861: C. Bliek, C. Jermann, A. Neumaier (Eds.), Global Optimization and Constraint Satisfaction. Proceedings, 2002. XII, 239 pages. 2003.

Vol. 2862: D. Feitelson, L. Rudolph, U. Schwiegelshohn (Eds.), Job Scheduling Strategies for Parallel Processing. Proceedings, 2003. VII, 269 pages. 2003.

Vol. 2863: P. Stevens, J. Whittle, G. Booch (Eds.), «UML» 2003 – The Unified Modeling Language. Proceedings, 2003. XIV, 415 pages. 2003.

Vol. 2864: A.K. Dey, A. Schmidt, J.F. McCarthy (Eds.), UbiComp 2003: Ubiquitous Computing. Proceedings, 2003. XVII, 368 pages. 2003.

Vol. 2865: S. Pierre, M. Barbeau, E. Kranakis (Eds.), Ad-Hoc, Mobile, and Wireless Networks. Proceedings, 2003. X, 293 pages. 2003.

Vol. 2866: J. Akiyama, M. Kano (Eds.), Discrete and Computational Geometry. Proceedings, 2002. VIII, 285 pages. 2003.

Vol. 2867: M. Brunner, A. Keller (Eds.), Self-Managing Distributed Systems. Proceedings, 2003. XIII, 274 pages. 2003.

Vol. 2868: P. Perner, R. Brause, H.-G. Holzhütter (Eds.), Medical Data Analysis. Proceedings, 2003. VIII, 127 pages. 2003.

Vol. 2869: A. Yazici, C. Şener (Eds.), Computer and Information Sciences – ISCIS 2003. Proceedings, 2003. XIX, 1110 pages. 2003.

Vol. 2870: D. Fensel, K. Sycara, J. Mylopoulos (Eds.), The Semantic Web - ISWC 2003. Proceedings, 2003. XV, 931 pages. 2003.

Vol. 2871: N. Zhong, Z.W. Raś, S. Tsumoto, E. Suzuki (Eds.), Foundations of Intelligent Systems. Proceedings, 2003. XV, 697 pages. 2003. (Subseries LNAI)

Vol. 2873: J. Lawry, J. Shanahan, A. Ralescu (Eds.), Modelling with Words. XIII, 229 pages. 2003. (Subseries LNAI)

Vol. 2874: C. Priami (Ed.), Global Computing. Proceedings, 2003. XIX, 255 pages. 2003.

Vol. 2875: E. Aarts, R. Collier, E. van Loenen, B. de Ruyter (Eds.), Ambient Intelligence. Proceedings, 2003. XI, 432 pages. 2003.

Vol. 2876: M. Schroeder, G. Wagner (Eds.), Rules and Rule Markup Languages for the Semantic Web. Proceedings, 2003. VII, 173 pages. 2003.

Vol. 2877: T. Böhme, G. Heyer, H. Unger (Eds.), Innovative Internet Community Systems. Proceedings, 2003. VIII, 263 pages. 2003.

Vol. 2878: R.E. Ellis, T.M. Peters (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003. Part I. Proceedings, 2003. XXXIII, 819 pages. 2003.

Vol. 2879: R.E. Ellis, T.M. Peters (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003. Part II. Proceedings, 2003. XXXIV, 1003 pages. 2003.

Vol. 2880: H.L. Bodlaender (Ed.), Graph-Theoretic Concepts in Computer Science. Proceedings, 2003. XI, 386 pages. 2003.

Vol. 2881: E. Horlait, T. Magedanz, R.H. Glitho (Eds.), Mobile Agents for Telecommunication Applications. Proceedings, 2003. IX, 297 pages. 2003.

Vol. 2882: D. Veit, Matchmaking in Electronic Markets. XV, 180 pages. 2003. (Subseries LNAI)

Vol. 2883: J. Schaeffer, M. Müller, Y. Björnsson (Eds.), Computers and Games. Proceedings, 2002. XI, 431 pages. 2003.

Vol. 2884: E. Najm, U. Nestmann, P. Stevens (Eds.), Formal Methods for Open Object-Based Distributed Systems. Proceedings, 2003. X, 293 pages. 2003.

Vol. 2885: J.S. Dong, J. Woodcock (Eds.), Formal Methods and Software Engineering. Proceedings, 2003. XI, 683 pages. 2003.

Vol. 2886: I. Nyström, G. Sanniti di Baja, S. Svensson (Eds.), Discrete Geometry for Computer Imagery. Proceedings, 2003. XII, 556 pages. 2003.

Vol. 2887: T. Johansson (Ed.), Fast Software Encryption. Proceedings, 2003. IX, 397 pages. 2003.

Vol. 2888: R. Meersman, Zahir Tari, D.C. Schmidt et al. (Eds.), On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. Proceedings, 2003. XXI, 1546 pages. 2003.

Vol. 2889: Robert Meersman, Zahir Tari et al. (Eds.), On The Move to Meaningful Internet Systems 2003: OTM 2003 Workshops. Proceedings, 2003. XXI, 1096 pages. 2003.

Vol. 2890: M. Broy, A.V. Zamulin (Eds.), Perspectives of System Informatics. Proceedings, 2003. XV, 572 pages. 2003.

Vol. 2891: J. Lee, M. Barley (Eds.), Intelligent Agents and Multi-Agent Systems. Proceedings, 2003. X, 215 pages. 2003. (Subseries LNAI)

Vol. 2892: F. Dau, The Logic System of Concept Graphs with Negation. XI, 213 pages. 2003. (Subseries LNAI)

Vol. 2893: J.-B. Stefani, I. Demeure, D. Hagimont (Eds.), Distributed Applications and Interoperable Systems. Proceedings, 2003. XIII, 311 pages. 2003.

Vol. 2894: C.S. Laih (Ed.), Advances in Cryptology - ASIACRYPT 2003. Proceedings, 2003. XIII, 543 pages. 2003.

Vol. 2895: A. Ohori (Ed.), Programming Languages and Systems. Proceedings, 2003. XIII, 427 pages. 2003.

Vol. 2896: V.A. Saraswat (Ed.), Advances in Computing Science – ASIAN 2003. Proceedings, 2003. VIII, 305 pages. 2003.

Vol. 2897: O. Balet, G. Subsol, P. Torguet (Eds.), Virtual Storytelling. Proceedings, 2003. XI, 240 pages. 2003.

Vol. 2898: K.G. Paterson (Ed.), Cryptography and Coding. Proceedings, 2003. IX, 385 pages. 2003.

Vol. 2899: G. Ventre, R. Canonico (Eds.), Interactive Multimedia on Next Generation Networks. Proceedings, 2003. XIV, 420 pages. 2003.

Vol. 2900: M. Bidoit, P.D. Mosses, CASL User Manual. XIII, 240 pages. 2004.

Vol. 2901: F. Bry, N. Henze, J. Maluszyński (Eds.), Principles and Practice of Semantic Web Reasoning. Proceedings, 2003. X, 209 pages. 2003.

Vol. 2902: F. Moura Pires, S. Abreu (Eds.), Progress in Artificial Intelligence. Proceedings, 2003. XV, 504 pages. 2003. (Subseries LNAI).

Vol. 2903: T.D. Gedeon, L.C.C. Fung (Eds.), AI 2003: Advances in Artificial Intelligence. Proceedings, 2003. XVI, 1075 pages. 2003. (Subseries LNAI).

Vol. 2904: T. Johansson, S. Maitra (Eds.), Progress in Cryptology – INDOCRYPT 2003. Proceedings, 2003. XI, 431 pages. 2003.

Vol. 2905: A. Sanfeliu, J. Ruiz-Shulcloper (Eds.), Progress in Pattern Recognition, Speech and Image Analysis. Proceedings, 2003. XVII, 693 pages. 2003.

Vol. 2906: T. Ibaraki, N. Katoh, H. Ono (Eds.), Algorithms and Computation. Proceedings, 2003. XVII, 748 pages. 2003.

Vol. 2908: K. Chae, M. Yung (Eds.), Information Security Applications. Proceedings, 2003. XII, 506 pages. 2004.

Vol. 2910: M.E. Orlowska, S. Weerawarana, M.P. Papazoglou, J. Yang (Eds.), Service-Oriented Computing – ICSOC 2003. Proceedings, 2003. XIV, 576 pages. 2003.

Vol. 2911: T.M.T. Sembok, H.B. Zaman, H. Chen, S.R. Urs, S.H.Myaeng (Eds.), Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access. Proceedings, 2003. XX, 703 pages. 2003.

Vol. 2912: G. Liotta (Ed.), Graph Drawing. Proceedings, 2003. XV, 542 pages. 2004.

Vol. 2913: T.M. Pinkston, V.K. Prasanna (Eds.), High Performance Computing – HiPC 2003. Proceedings, 2003. XX, 512 pages. 2003.

Vol. 2914: P.K. Pandya, J. Radhakrishnan (Eds.), FST TCS 2003: Foundations of Software Technology and Theoretical Computer Science. Proceedings, 2003. XIII, 446 pages. 2003.

Vol. 2916: C. Palamidessi (Ed.), Logic Programming. Proceedings, 2003. XII, 520 pages. 2003.

Vol. 2918: S.R. Das, S.K. Das (Eds.), Distributed Computing – IWDC 2003. Proceedings, 2003. XIV, 394 pages. 2003.

Vol. 2919: E. Giunchiglia, A. Tacchella (Eds.), Theory and Applications of Satisfiability Testing. Proceedings, 2003. XI, 530 pages. 2004.

Vol. 2920: H. Karl, A. Willig, A. Wolisz (Eds.), Wireless Sensor Networks. Proceedings, 2004. XIV, 365 pages. 2004.

Vol. 2921: G. Lausen, D. Suciu (Eds.), Database Programming Languages. Proceedings, 2003. X, 279 pages. 2004.

Vol. 2922: F. Dignum (Ed.), Advances in Agent Communication. Proceedings, 2003. X, 403 pages. 2004. (Subseries LNAI).

Vol. 2923: V. Lifschitz, I. Niemelä (Eds.), Logic Programming and Nonmonotonic Reasoning. Proceedings, 2004. IX, 365 pages. 2004. (Subseries LNAI).

Vol. 2924: J. Callan, F. Crestani, M. Sanderson (Eds.), Distributed Multimedia Information Retrieval. Proceedings, 2003. XII, 173 pages. 2004.

Vol. 2926: L. van Elst, V. Dignum, A. Abecker (Eds.), Agent-Mediated Knowledge Management. Proceedings, 2003. XI, 428 pages. 2004. (Subseries LNAI).

Vol. 2927: D. Hales, B. Edmonds, E. Norling, J. Rouchier (Eds.), Multi-Agent-Based Simulation III. Proceedings, 2003. X, 209 pages. 2003. (Subseries LNAI).

Vol. 2928: R. Battiti, M. Conti, R. Lo Cigno (Eds.), Wireless On-Demand Network Systems. Proceedings, 2004. XIV, 402 pages. 2004.

Vol. 2929: H. de Swart, E. Orlowska, G. Schmidt, M. Roubens (Eds.), Theory and Applications of Relational Structures as Knowledge Instruments. Proceedings. VII, 273 pages. 2003.

Vol. 2931: A. Petrenko, A. Ulrich (Eds.), Formal Approaches to Software Testing. Proceedings, 2003. VIII, 267 pages. 2004.

Vol. 2932: P. Van Emde Boas, J. Pokorný, M. Bieliková, J. Štuller (Eds.), SOFSEM 2004: Theory and Practice of Computer Science. Proceedings, 2004. XIII, 385 pages. 2004.

Vol. 2933: C. Martín-Vide, G. Mauri, G. Păun, G. Rozenberg, A. Salomaa (Eds.), Membrane Computing. Proceedings, 2003. VIII, 383 pages. 2004.

Vol. 2935: P. Giorgini, J.P. Müller, J. Odell (Eds.), Agent-Oriented Software Engineering IV. Proceedings, 2003. X, 247 pages. 2004.

Vol. 2937: B. Steffen, G. Levi (Eds.), Verification, Model Checking, and Abstract Interpretation. Proceedings, 2004. XI, 325 pages. 2004.

Vol. 2938: Z. Zhang, C. Zhang, Agent-Based Hybrid Intelligent Systems. XV, 196 pages. 2004. (Subseries LNAI).

Vol. 2942: D. Seipel, J.M. Turull-Torres (Eds.), Foundations of Information and Knowledge Systems. Proceedings, 2004. X, 321 pages. 2004.

Vol. 2944: K. Aberer, M. Koubarakis, V. Kalogeraki (Eds.), Databases, Information Systems, and Peer-to-Peer Computing. Proceedings, 2003. X, 249 pages. 2004.

Vol. 2945: A. Gelbukh (Ed.), Computational Linguistics and Intelligent Text Processing. Proceedings, 2004. XVIII, 651 pages. 2004.

Vol. 2946: R. Focardi, R. Gorrieri (Eds.), Foundations of Security Analysis and Design II. VII, 267 pages. 2004.

Vol. 2950: N. Jonoska, G. Păun, G. Rozenberg (Eds.), Aspects of Molecular Computing. XI, 391 pages. 2004.

Preface

This volume contains the papers presented at the 3rd International Symposium on Foundations of Information and Knowledge Systems (FoIKS 2004), which was held in Castle Wilhelminenberg, Vienna, Austria, from February 17th to 20th, 2004.

FoIKS is a biennial event focussing on theoretical foundations of information and knowledge systems. It aims at bringing together researchers working on the theoretical foundations of information and knowledge systems and attracting researchers working in mathematical fields such as discrete mathematics, combinatorics, logics, and finite model theory who are interested in applying their theories to research on database and knowledge base theory.

FoIKS took up the tradition of the conference series Mathematical Fundamentals of Database Systems (MFDBS) which enabled East–West collaboration in the field of database theory. The first FoIKS symposium was held in Burg, Spreewald (Germany) in 2000, and the second FoIKS symposium was held in Salzau Castle (Germany) in 2002. Former MFDBS conferences were held in Dresden (Germany) in 1987, Visegrád (Hungary) in 1989, and in Rostock (Germany) in 1991. Proceedings of these previous events were published by Springer-Verlag as volumes 305, 364, 495, 1762, and 2284 of the LNCS series, respectively.

In addition the FoIKS symposium was intended to be a forum for intensive discussions. For this reason the time slots for long and short contributions were 50 and 30 minutes, respectively, followed by 20 and 10 minutes for discussions, respectively. Furthermore, participants were asked in advance to prepare to act as correspondents for the contributions of other authors. There were also special sessions for the presentation and discussion of open research problems.

The FoIKS 2004 call for papers solicited contributions dealing with any foundational aspect of information and knowledge systems, e.g.,

- mathematical foundations: discrete methods, boolean functions, finite model theory
- database design: formal models, dependency theory, schema translations, desirable properties
- query languages: expressiveness, computational and descriptive complexity, query languages for advanced data models, classifications of computable queries
- semi-structured databases and WWW: models of Web databases, querying semi-structured databases, Web transactions and negotiations
- security in data and knowledge bases: cryptography, steganography, information hiding

- integrity and constraint management: verification, validation, and enforcement of consistency, triggers
- information integration: heterogenous data, views, schema dominance and equivalence
- database and knowledge base dynamics: models of transactions, models of interaction, updates, consistency preservation, concurrency control
- intelligent agents: multi-agent systems, autonomous agents, foundations of software agents, cooperative agents
- logics in databases and AI: non-classical logics, spatial and temporal logics, probabilistic logics, deontic logic, logic programming
- knowledge representation: planning, description logics, knowledge and belief, belief revision and update, non-monotonic formalisms, uncertainty
- reasoning techniques: theorem proving, abduction, induction, constraint satisfaction, common-sense reasoning, probabilistic reasoning, reasoning about actions.

The programme committee received 64 submissions. Each paper was carefully reviewed by at least two experienced referees, and most of the papers were reviewed by three referees. Fourteen papers were chosen for long presentations and four papers for short presentations. This volume contains versions of these papers polished based on the comments made in the reviews. A few papers will be selected for further extension and publishing in a special issue of the journal Annals of Mathematics and Artificial Intelligence.

We would like to thank all authors who submitted papers and all workshop participants for the fruitful discussions. We are grateful to the members of the programme committee and the external referees for their timely expertise in carefully reviewing the papers, and we would like to express our thanks to our hosts for the beautiful week in the pleasant surroundings of Castle Wilhelminenberg near Vienna.

February 2004

Dietmar Seipel
José María Turull-Torres

Programme Committee Co-chairs

Dietmar Seipel	University of Würzburg (Germany)
José María Turull-Torres	Massey University (New Zealand)

Programme Committee

Leopoldo Bertossi	Carleton University, Ottawa (Canada)
Joachim Biskup	University of Dortmund (Germany)
François Bry	University of Munich (Germany)
Samir Chopra	University of New South Wales (Australia)
Jürgen Dix	University of Manchester (UK)
Thomas Eiter	Vienna University of Technology (Austria)
Jörg Flum	University of Freiburg (Germany)
Burkhard Freitag	University of Passau (Germany)
Stephen J. Hegner	Umeå University (Sweden)
Lauri Hella	University of Tampere (Finland)
Sven Hartmann	Massey University (New Zealand)
Gyula Katona	Hungarian Academy of Sciences (Hungary)
Gabriele Kern-Isberner	FernUniversität Hagen (Germany)
Hans-Joachim Klein	University of Kiel (Germany)
Alberto Mendelzon	University of Toronto (Canada)
Jack Minker	University of Maryland (USA)
Kotagiri Ramamohanarao	University of Melbourne (Australia)
Vladimir Sazonov	University of Liverpool (UK)
Klaus-Dieter Schewe	Massey University (New Zealand)
Thomas Schwentick	Philipps University at Marburg (Germany)
Dietmar Seipel	University of Würzburg (Germany)
Bernhard Thalheim	University of Kiel (Germany)
José María Turull-Torres	Massey University (New Zealand)

External Referees

Matthias Beck	Christoph Beierle
Gerhard Bloch	Gerd Brewka
Yijia Chen	Marina de Vos
Janos Demetrovics	Scott Dexter
Claus Dziarstek	Heinz-Dieter Ebbinghaus
Klaus Fellbaum	Flavio A. Ferrarotti
Hans-Peter Gumm	Michael Guppenberger
Wolfgang Hesse	Marbod Hopfner
Eyke Hüllermeier	Roland Kaschek
Kathrin Konczak	Boris Konev
Georg Lausen	Jens Lechtenbörger
Sebastian Link	Alexei Lisitsa
Thomas Meyer	Michael J. Minock
Bernhard Nebel	Thomas Nitsche
Maurice Pagnucco	Jari Palomaki
Gerald Pfeifer	Flavio Rizzolo
Attila Sali	Torsten Schlieder
Petra Schwaiger	Alexei Tretiakov
Bela Uhrin	Alejandro A. Vaisman
Wiebe van der Hoek	Mark Weyer

Organization

Thomas Eiter	Vienna University of Technology (Austria)

Table of Contents

Invited Talks

Regular Papers

Hypergraph Transversals

Georg Gottlob

Institut für Informationssysteme, Technische Universität Wien
Favoritenstraße 9–11, A-1040 Wien, Austria
gottlob@dbai.tuwien.ac.at

Hypergraph Transversals have been studied in Mathematics for a long time, cf. [2]. Generating minimal transversals of a hypergraph is an important problem which has many applications in Computer Science, especially in database Theory, Logic, and AI. We briefly survey some results on problems which are known to be related to computing the transversal hypergraph, where we focus on problems in database theory, propositional Logic and AI (for a more detailed survey and further references cf. [10]).

A *hypergraph* $\mathcal{H} = (V, E)$ consists of a finite collection E of sets over a finite set V. The elements of E are called *hyperedges*, or simply *edges*. A *transversal* (or *hitting set*) of $\mathcal{H}$ is a set $T \subseteq V$ that meets every edge of E. A transversal is *minimal*, if it does not contain any other transversal as a subset. The set T of all minimal transversals of $\mathcal{H} = (V, E)$, constitutes together with V also a hypergraph $Tr(\mathcal{H}) = (V, T)$, which is called the *transversal hypergraph* of $\mathcal{H}$. The famous Transversal Hypergraph Problem (TRANS-HYP) is then as follows: Given two hypergraphs $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V, F)$ on a finite set V, decide whether $\mathcal{G} = Tr(\mathcal{H})$ holds. The corresponding computation problem is called TRANSVERSAL ENUMERATION (TRANS-ENUM) and is phrased as follows: Given a hypergraph $\mathcal{H} = (V, E)$ on a finite set V, compute the transversal hypergraph $Tr(\mathcal{H})$.

From the point of computability in polynomial time, the decisional and the computational variant of the transversal hypergraph problem are in fact equivalent: It is known that, for any class $\mathcal{C}$ of hypergraphs, TRANS-ENUM is solvable in *polynomial total time* (or *output-polynomial time*), i.e., in time polynomial in the combined size of $\mathcal{H}$ and $Tr(\mathcal{H})$, if and only if TRANS-HYP is in the class P for all pairs $(\mathcal{H}, \mathcal{G})$ such that $\mathcal{H} \in \mathcal{C}$ [3].

The precise complexity of TRANS-HYP is not known to date, and is in fact open for more than 20 years now. Accordingly, it is unknown whether TRANS-ENUM can be solved in output-polynomial time.

The problems TRANS-HYP and TRANS-ENUM have a large number of applications in many areas of Computer Science, including Distributed Systems, Databases, Boolean Circuits and Artificial Intelligence. There, they have important applications in Diagnosis, Machine Learning, Data Mining, and Explanation Finding, see e.g. [9,12,17,20,21,22,24] and the references therein.

We call a decision problem Π TRANS-HYP-*hard*, if problem TRANS-HYP can be reduced to it by a standard polynomial time transformation. Furthermore, Π is TRANS-HYP-*complete*, if Π is TRANS-HYP-hard and, moreover, Π can be polynomially transformed into TRANS-HYP; that is, Π and TRANS-HYP are

D. Seipel and J.M. Turull-Torres (Eds.): FoIKS 2004, LNCS 2942, pp. 1–5, 2004.

equivalent modulo polynomial time transformations. We use analogous terminology of TRANS-ENUM-*hardness* and TRANS-ENUM-*completeness* for computations problems, i.e., problems with output (for more details, cf. [10]).

Let us first discuss issues of *structural* complexity. In a landmark paper, Fredman and Khachiyan [15] proved that TRANS-HYP can be solved in time $n^{o(\log n)}$, and thus in quasi-polynomial time. This shows that the problem is most likely not co-NP-complete, since no co-NP-complete problem is known which is solvable in quasi-polynomial time; if any such problem exists, then all problems in NP and co-NP can be solved in quasi-polynomial time.

A natural question is whether TRANS-HYP lies in some lower complexity class based on other resources than just runtime. In a recent paper [11], it was shown that the complement of this problem is solvable in polynomial time with *limited nondeterminism*, i.e, by a nondeterministic polynomial-time algorithm that makes only a poly-logarithmic number of guesses in the size of the input. For a survey on complexity classes with limited nondeterminism, and for several references see [18]. More precisely, [11] shows that non-duality of a pair $\mathcal{G}, \mathcal{H})$ can be proved in polynomial time with $O(\chi(n) \cdot \log n)$ suitably guessed bits, where $\chi(n)$ is given by $\chi(n)^{\chi(n)} = n$; note that $\chi(n) = o(\log n)$.

This result is surprising, because most researchers dealing with the complexity of the transversal hypergraph thought so far that these problems are completely unrelated to limited nondeterminism.

A large number of tractable restrictions of TRANS-HYP and TRANS-ENUM are known in the literature, e.g. [7,5,4,8,9,13,11,16,23,27], and references therein.

Examples of tractable classes are instances $(\mathcal{H}, \mathcal{G})$ where $\mathcal{H}$ has the edge sizes bounded by a constant, or where $\mathcal{H}$ is *acyclic*. Various "degrees" of hypergraph acyclicity have been defined in the literature [14]. The most general notion of hypergraph acyclicity (applying to the largest class of hypergraphs) is α-acyclicity; less general notions are (in descending order of generality) β-, γ-, and Berge-acyclicity (see [14]). In [9], it was shown that Hypergraph transversal instances with β-acyclic $\mathcal{H}$ are tractable. In [11], this tractability result has been recently improved to instances where $\mathcal{H}$ is α-acyclic and simple. This result is a corollary to a more general tractability result for hypergraphs whose *degeneracy* is bounded by a constant; simple, α-acyclic hypergraphs have degeneracy 1.

Furthermore, [11] shows that instances $(\mathcal{H}, \mathcal{G})$ of TRANS-HYP where the vertex-hyperedge incidence graphs of $\mathcal{H}$ (or of $\mathcal{G}$) have *bounded treewidth* are solvable in polynomial time.

In the sequel, we mention a few problems closely related to TRANS-HYP or TRANS-ENUM. These and many other such problems are discussed in detail in [10,9].

Functional Dependency Inference

Given a database relation instance r and a set of functional dependencies (FDs) F, deciding whether F characterizes precisely the FDs holding on r, i.e., deciding whether F^+ is identical with the set $FD(r)$ of all FDs valid on r, is TRANS-HYP-hard and in co-NP. The precise complexity is currently open. The same problem

becomes TRANS-HYP-complete if F is in Boyce-Codd Normal Form (BCNF), i.e., iff all left hand sides of the FDs in F are keys. Given a relation instance r, generating a cover F of FDs suchthat $F^+ = FD(r)$ is TRANS-ENUM-hard. Vice-versa, given a set F, generating a so-called Armstrong relation, i.e., a relation instance r incorporating precisely all FDs of F^+ is also TRANS-ENUM hard. For more details on the Dependency Inference problem and on related problems cf. [24,25,19,26,10], as well as the extended version of [9].

Data Mining: Maximal Frequent Sets

Given a $0/1$ $m \times n$ matrix A and an integral threshold t, associate with each subset $C \subseteq \{1, \ldots, n\}$ of column indices the subset $R(C)$ of all rows $r \in \{1, \ldots, m\}$ in A such that $A(r, j) = 1$ for every $j \in C$. Then C is called *frequent*, if $|R(C)| \geq t$, and C is called *infrequent*, if $|R(C)| < t$. Let us denote by $F_t(A)$ and $\hat{F}_t(A)$ the sets of all frequent and infrequent column sets C in A, respectively.

The generation of frequent and infrequent sets in A is a key problem in knowledge discovery and data mining, which occurs in mining association rules, correlations, and other tasks. Of particular interest are the maximal frequent sets $M_t \subseteq F_t$ and the minimal infrequent sets $I_t \subseteq \hat{F}_t$, since they mark the boundary of frequent sets (both maximal and minimal under set inclusion). The following result has been recently proved in [6]: *The problem of computing, given a 0/1 matrix A and a threshold t, the sets M_t and I_t is* TRANS-ENUM-*complete.*

Theory Approximation: Horn Envelope

A logical theory Σ is *Horn*, if it is a set of Horn clauses, i.e., disjunctions $l_1 \vee \cdots \vee l_m$ of literals l_i such that at most one of them is positive. Semantically, Horn theories are characterized by the property that their set of models, $mod(\Sigma)$, is closed under intersection, i.e., $M, M' \in mod(\Sigma)$ implies $M \cap M' \in mod(\Sigma)$; here, $M \cap M'$ is the model M'' which results by atomwise logical conjunction of M and M', i.e., $M'' \models a$ iff $M \models a$ and $M' \models a$, for every atom a.

Any theory Σ has a unique *Horn envelope*, which is the strongest (w.r.t. implication) Horn theory Σ' such that $\Sigma \models \Sigma'$. The Horn envelope might be represented by different Horn theories, but there is a unique representation, which we denote by $HEnv(\Sigma)$, which consists of all prime clauses of Σ'. The following result was established in [22], where the TRANS-ENUM hardness part was proved in [21]: *The problem of computing, given the models $mod(\Sigma)$ of a propositional theory Σ, the Horn envelope $HEnv(\Sigma)$ is* TRANS-ENUM-*complete.*

Dualization of Boolean Functions

There is a well-known and close connection of TRANS-ENUM to the well-known dualization problem of Boolean Functions: given a CNF ϕ of a Boolean function f, compute a prime CNF ψ of its dual f^d, i.e., the function which has value 1 in input vector $b = (b_1, \ldots, b_n)$ iff f has value 0 on the input vector $\overline{b} =$

$(b_1 \oplus 1, \ldots, b_n \oplus 1)$. Special algorithms for the dualization problem can be tracked down at least to the 60's of the past century, cf. [1]. It is not hard to see that this problem is intractable; in fact, its decisional varian DUAL which given two CNFs φ and ψ of Boolean functions f and g, respectively. consists in deciding whether φ and ψ represent a pair (f, g) of dual Boolean functions is co-NP-complete, where hardness holds even if ψ is asserted to be a prime CNF of g.

In case of monotone Boolean functions, the dualization problem is equivalent to determining the transversal hypergraph. Let MONOTONE DUALIZATION designate the subcase of DUALIZATION where f is a monotone Boolean function, and similarly MONOTONE DUAL the subcase of DUAL where f is a monotone Boolean function.

The following statement summarizes well-known results that are part of the folklore: MONOTONE DUALIZATION *is* TRANS-ENUM-*complete, and* MONOTONE DUAL *is* TRANS-HYP-*complete.* For a proof see [10].

Restricted Versions of the Satisfiability Problem

We denote by IMSAT (Intersecting Monotone SAT) the restriucted version of the classical SAT problem where each clause contains either only positive literals or only negative literals, and where every pair of a positive and a negative clause resolves, i.e., there is at least one atom which occurs unnegated in the positive clause and negated in the negative claues. The following holds [9,10]: IMSAT *is co*-TRANS-HYP-*complete.* This complexity result holds even if we restrict IMSAT instances to clause sets $\mathcal{C}$ where the negative clauses are precisely all clauses C^- such that $C^- = \{\neg u : u \in C^+\}$ for some positive clause $C^+ \in \mathcal{C}$.

References

1. C. Benzaken. Algorithme de dualisation d'une fonction booléenne. *Revue Francaise de Traitment de l'Information – Chiffres*, 9(2):119–128, 1966.
2. C. Berge. Hypergraphs. North Holland, 1989.
3. C. Bioch and T. Ibaraki. Complexity of identification and dualization of positive Boolean functions. *Information and Computation*, 123:50–63, 1995.
4. E. Boros, K. Elbassioni, V. Gurvich, and L. Khachiyan. An efficient incremental algorithm for generating all maximal independent sets in hypergraphs of bounded dimension. *Parallel Processing Letters*, 10(4):253–266, 2000.
5. E. Boros, V. Gurvich, and P. L. Hammer. Dual subimplicants of positive Boolean functions. *Optimization Methods and Software*, 10:147–156, 1998.
6. E. Boros, V. Gurvich, L. Khachiyan, and K. Makino. On the complexity of generating maximal frequent and minimal infrequent sets. In *Proc. STACS-02*, LNCS 2285, pp. 133–141, 2002.
7. E. Boros, P. Hammer, T. Ibaraki, and K. Kawakami. Polynomial time recognition of 2-monotonic positive Boolean functions given by an oracle. *SIAM J. Comput.*, 26(1):93–109, 1997.
8. C. Domingo, N. Mishra, and L. Pitt. Efficient read-restricted monotone CNF/DNF dualization by learning with membership queries. *Machine Learning* 37, 1999.