BRICS Newsletter

Basic Research in Computer Science — No 2, December 94

In this Issue

Welcome 1

BRICS Themes 2
- Complexity Theory '94 2
- Logic in Semantics '95 2
  - Inductive and Co-inductive Techniques in the Sem. of Func. Programs 2
  - The modal mu-Calculus 2
  - Synthetic Domain Theory 2
  - Workshop on Full Abstraction 3

Reports on Conferences and Workshops 3

Reports on Courses 6

Newly Appointed Researchers 6

Coming Events 7
- Theme Events 7
- Dynamic Algorithms 7
- TACAS 7
- TAPSOFT 7

Dissertation Abstracts 7
- Block Ciphers - Analysis, Design and Applications 7
- Timed Modal Specifications 8
- A HOL Basis for Reasoning about Functional Programs 9

New Reports 10

Notes Series 11

Technical Contributions 12
- Of Simple Problems and Hard Solutions . 12

Calendar of Events 16

BRICS Address and World Wide Web 16

Welcome

Welcome to the second issue of the BRICS newsletter. Its purpose is to inform you of appointments, publications, courses and other activities within BRICS. Further details can be obtained by contacting the addresses below.

BRICS has now been running for almost a year. When our first researcher arrived in January we didn't yet have our own offices—those involved with BRICS were rather scattered. We now have over forty researchers associated to BRICS and our own two corridors within the Computer Science department in Aarhus, as well a smaller BRICS contingent in Aalborg. We are well equipped with machines, and well staffed with energetic young researchers, who are responsible in no small measure for BRICS' success, socially and intellectually.

This year has seen the successful completion of one research theme in Complexity Theory and an excellent start to another, Logic in Semantics, which continues into '95. There have been numerous BRICS seminars and courses. To date we have 48 BRICS research reports. The companion series of BRICS notes is for less polished but nevertheless important material, often associated with the courses we hold. A start is being made on a BRICS lectures notes series; the intention here is to produce high quality lecture notes on topics in the foundations of Computer Science which are to be freely available for Computer Science and Mathematics departments worldwide. In some respects we are still finding our way, of course. However all in all we are pleased with how things are developing, and look forward to another exciting year.

BRICS Themes

Complexity Theory '94

In the months of August and September, 1994 BRICS hosted the meeting *Complexity Theory - A Task Force* where a small number of researchers who have made fundamental contributions to the area of complexity theory participated.

The visitors were: Miklos Ajtai, Michael Ben-Or, Allan Borodin, Oded Goldreich, Russell Impagliazzo, Mark Jerrum, Dexter Kozen, Noam Nisan, Mike Paterson, Alexander Razborov, Shmuel Safra, Adi Shamir, Avi Wigderson, and Uri Zwick. In addition a few PhD students were visiting BRICS during the meeting. They were: Amos Beimel, Anna Gal, Robert Szelepcsenyi, and Amnon Ta-Shma.

During the meeting several activities took place. A number of talks for different audiences including people from industry were given. As part of the meeting, a conference *Complexity Theory - Present and Future* took place during the 3rd week of August. As a lead-in to the conference, there were in the previous week some introductory seminars aimed at PhD students and others who are not "experts" in complexity theory. Slides, overviews and papers on which the conference talks were based can be found in NS-94-4.

Logic in Semantics '95

The theme for '95 is *Logic in Semantics* concentrating on such topics as: proof principles for reasoning about functional programs and recursive types (e.g., Andy Pitts' course, Andy Gordon's seminar), synthetic domain theory (e.g., Pino Rosolini's course, Edmund Robinson's visit), full abstraction and stable and sequential domain theory (e.g., Full Abstraction Workshop). In order to avoid clashes with the programme at the Isaac Newton Institute at Cambridge, activities will take place before late summer '95, with two activities (Andy Pitts' mini-course and the Peripatetic Seminar on Sheaves and Logic) having already taken place in late '94.

In connection with the theme, Edmund Robinson (University of Sussex) will be visiting BRICS for two months from mid January '95, it is hoped that Thomas Ehrhard (Laboratorie de Mathematiques Discretes, Marseilles) will visit for two months around the Full Abstraction Workshop (though this is not yet confirmed) and Bart Jacobs will visit for a month from mid January. In addition, there will be a number of short-term visits, including from: John Power, Pino Rosolini, Thomas Streicher, Andy Gordon, Dusko Pavlović, Stefano Kasangian, Marek Zawadowski.

Activities in *Logic in Semantics* include:

Inductive and Co-inductive Techniques in the Semantics of Functional Programs

Andrew Pitts, Computer Laboratory, Cambridge University, delivered 9 lectures 21 November – 2 December, 1994, on recent advances in formal techniques for establishing observational equivalence of functional programs. Both operational and denotational methods and the relationship between them were considered. Copies of lecture notes are available in the BRICS Notes series, NS-94-5.

The modal mu-Calculus

Igor Walukiewicz, BRICS, will hold a mini-course of around 5 lectures a week in February on the modal mu-calculus, culminating with his solution to the long-standing open problem of the completeness of Kozen's proof system—it is!

Synthetic Domain Theory

Pino Rosolini, University of Genova, is visiting 5–15 March and will give a mini-course (5 lectures) on synthetic domain theory.

Workshop on Full Abstraction

Glynn Winskel is organising a workshop on full abstraction for PCF and related languages for the two weeks starting April 18, 1995. Guests at this period include: Samson Abramsky, Pierre-Louis Curien, Thomas Ehrhard, Matthias Felleisen, Luke Ong, Achim Jung, Pasquale Malacaria, Peter O'Hearn, Jon Riecke, Allen Stoughton, Kurt Sieber, Antonio Bucciarelli, Radha Jagadeesan, Thierry Coquand, Vladimir Sazonov. ≣

Reports on Conferences and Workshops

Fourth Scandinavian Workshop on Algorithm Theory

On July 6–8 1994, Aarhus University hosted SWAT '94, Fourth Scandinavian Workshop on Algorithm Theory. The workshop was sponsored by the Danish Natural Science Research Council, Aarhus University and BRICS.

A total of 34 papers were presented, including three invited lectures. The approximate 80 participants enjoyed one of the best summers Aarhus has had in recent years.

The proceedings have appeared as vol. 824 in the Springer Lecture Notes Series. The January 1995 issue of Nordic Journal of Computing will contain selected papers from the Workshop. ≣

Complexity Theory - Present and Future

See the description of the '94 theme on Complexity Theory above. ≣

The 6th Nordic Workshop on Programming Theory

Aarhus, Denmark, 17–19 October 1994.

We here bring excerpts from a report written by *Kārlis Cerāns*, Inst. of Mathematics and Computer Science, University of Latvia, for the EATCS Bulletin (to appear in the February issue).

During three sunny days of October, 1994, Aarhus, the second largest city of Denmark ("The world's smallest big city", according to a tourist brochure) was hosting the Nordic Workshop on Programming Theory. The main objective of the Workshop was to bring together researchers from the Nordic and Baltic countries interested in programming theory, in order to improve mutual contacts and cooperation.

This year's NWPT was already the 6th in the annual series of similar Workshops, held in previous years in Uppsala (Sweden, 1989), Aalborg (Denmark, 1990), Göteborg (Sweden, 1991), Bergen (Norway, 1992) and Turku (Finland, 1993) under the name of "Nordic Workshop on Program Correctness".

NWPT'94 has attracted 63 participants, coming from Norway (3), Sweden (11), Finland (5), Denmark (33), Latvia (1), Lithuania (2), Estonia (1), England (3) and Germany (4). And it was a very excellent event indeed!

An important contribution to the success of the Workshop was the financial sponsorship provided by grants from BRICS (a centre for Basic Research In Computer Science established in cooperation between the Danish National Research Foundation and Universities of Aarhus and Aalborg) and the Danish Science Research Council. It should be noted in particular that these grants allowed to reduce or even fully cover for some participants not only the expenses for their workshop attendance, but also those for travelling to the conference site. Among the "fully sponsored" were all participants from Baltic countries, for what, as being one of them, I want to express my personal most sincere thanks both to the sponsors and the organizers of the Workshop.

In the scientific program of the Workshop per-

haps one would notice, first, in all respects excellent invited 60 minute lectures given at the Workshop by

- Bernhard Steffen from University of Passau (Germany) on *Finite model checking and beyond* (with a remarkable explanation on how to model - check context free processes, and why this problem is very natural),
- Ib Holm Sørensen (B-Core limited, UK) on *The B-technologies: A system for computer aided programming* (when asked during the demo session, why just "B", among other explanations he offered was that "B" is the most *B*eautiful letter of the alphabet), and
- Matthew Hennessy from University of Sussex, UK on *Higher order processes and their models.*

One could mention that it is a tradition at this series of Nordic Workshops to have invited speakers from outside the Nordic region - for the sake of making more direct contacts with "non-local" programming theory authorities. This could be especially important for students who are in the beginning of their research career, and who were quite many at the Workshop.

Besides the invited talks the main part of the Workshop's scientific program was filled in by contributed presentations - running mostly in 2 parallel sessions and each one lasting for 30 minutes. Among the areas attracting the attention of the contributed presentations, one could mention semantics and analysis of concurrency (including value-passing process calculi, broadcasting systems, categorical models of concurrency), refinement calculus, specification and verification of parametric and real time systems, several aspects of functional and logic programming (synthesis, implementation, typing), algebraic specification. And, of course, there was ALF[1] (its younger offsprings HALF and GANDALF were just mentioned, in passing). A more detailed inspiration about the Workshop's program can be obtained from the collection of the abstracts of the presentations, which has been published in BRICS Notes series, and have been distributed at the Workshop.

At the Workshop also B. Steffen and T. Margaria demonstrated their tool for *High-level synthesis of heterogeneous analysis systems.* The demo on *B-toolkit* was given by I. H. Sørensen. The "Göteborg tool" *ALF* was demonstrated by C. Coquand and L. Magnusson, and *Epsilon* (an "Aalborg tool") was shown by K. Larsen and J. Niedermann.

What has been said so far about the scientific program of the Workshop should have convinced the reader that this event scientifically indeed was great. However, as usual, the highly interesting presentations have made only a part of the success of the Workshop. Indeed, the possibilities of meeting the old colleagues and making new contacts in a relaxing and friendly atmosphere, discussions both on specific research and also more general problems have contributed a lot. One can hope that these discussions will strengthen the existing and lead to new scientific and human cooperation instances.

So, the Workshop is over, though not quite. Inspired from the discussions, the participants who gave their talks at the Workshop now are encouraged to submit a paper (extended abstract) for the Proceedings, which are going to be published by BRICS. It is also planned to invite a few of the Workshop participants to prepare later full versions of their papers for a special issue of *Nordic Journal of Computing.*

Finally, let us all say once again thanks to the organizers of NWPT'94 for their work in making such a really great event, and see you all next autumn in Göteborg, where NWPT '95 is to take place!

All papers of the proceedings, NS-94-6, are electronically available (see Abstract of 94-6 in Notes Series of Publications of BRICS WWW).

[1] Another Logical Framework, from Göteborg.

56th Peripatetic Seminar on Sheaves and Logic

The *56th Peripatetic Seminar on Sheaves and Logic* (PSSL), organized and sponsored by BRICS jointly with the Mathematics Department, was held on 3 and 4 December 1994 at the University of Aarhus.

The Peripatetic Seminar is perhaps not so widely known in the Computer Science community. *Vladimiro Sassone*, BRICS, has kindly provided the following information.

Background

The seminar is, as the name suggests, a "wandering little conference". Founded by Dana Scott in the seventies, it is held two or three times a year, at varying places, over a weekend. Talks start on Saturday morning and end by lunchtime on Sunday. Participants provide their own funding, there is no program committee, and there are no proceedings.

Topics

It is not easy to give a comprehensive account of the topics the seminar wants to cover, and it would actually be inappropriate to try to list them precisely: it seems a natural consequence of its wandering spirit that the character of the talks varies with the place where the seminar is held. Initially, the focus was on the applications of category theory to *mathematical logic* (in particular, intuitionistic logic), but also on *categorical algebra* and *topology*. *Topos theory*, a general framework for both logic and topology, is still an important subject in the seminar. Later, applications of category theory in *computer science* came to be another relevant source. Currently, the keywords "*logic*" and "*category theory*"—with gradually more emphasis on the latter—both in mathematics and in computer science cover a large part of the intersection of the interests of the participants.

56th PSSL

The 56th PSSL focussed in particular on topos theory, locales and quantales, homology theory, sketches and geometric theories, fibrations, monoidal and autonomous categories, cofree coalgebras, intuitionistic linear logic and linear λ-calculus, and semantics of concurrency. The meeting was very successful, as all the talks met high standards of quality. Some of them presented important new results, and working relationships were established which will result in some participants visiting BRICS next year. There were 20 participants coming from six different countries: Denmark (9), England (4), France (2), Germany (1), Holland (3), and Italy (1).

Local Organization

C. Hermida, A. Kock, J. van Oosten, V. Sassone, and S. Soloviev.

List of the Talks at the 56th PSSL

C. Mulvey, Pure States

T. Plewe, Absolutely Borel Locales

I. Moerdijk, "Makkai"

D. Pavlovic, How couniversal is the Chu construction?

P. Johnstone, What can you do with a finitary sketch?

V. Sassone, An approach to the category of Petri net computations

T. Streicher, Partial left exact categories and partial toposes

J. van Oosten, Fibrations and Calculi of Fractions

R. Guitart, On the Definition of Homology

A. Kock, Some extensive quantities in locale theory

S. Soloviev, Proof of a conjecture by S. Mac Lane

P. Damphousse, On endofunctors of Ens fixing objects

B. Jacobs, Mongruences and cofree coalgebras

T. Braüner, The Girard translation extended with recursion

Reports on Courses

Introduction to Complexity Theory

Introductory seminars prior to the conference *Complexity Theory - Present and Future* are described in connection with the '94 theme on page 2.

Induction Based on Rippling and Proof Planning

David Basin, Max-Planck-Institute, Saarbrücken, gave 11 August a full one-day course with 37 participants from both academia and industry. Slides and papers from the course material, NS-94-2, are also obtainable via WWW and FTP.

Linear Time Temporal Logic and Buchi Automata

P. S. Thiagarajan, SPIC Science Foundation, Madras, INDIA, gave introductory lectures in the period 11–13 October 1993 on Linear Time Temporal Logic, Buchi Automata, Decision Procedures, and Model Checking Algorithms. He also covered generalisations to Mazurkiewicz Trace Theory and Buchi Asynchronous Automata.

Inductive and Co-inductive Techniques in the Semantics of Functional Programs

Some 25 people attended this course by *Andy Pitts* mentioned previously in the section about the theme for 95 on Logic in Semantics.

Introduction to Categories

As part of the PhD course “Logic in Semantics”, *Jaap van Oosten* gave an introductory course in Category Theory and wrote some 50 pages of course notes which will appear in the BRICS Notes Series. The notes contain material on elementary category theory from the most basic definitions through adjoints and monads.

Newly Appointed Researchers

Devdatt Dubhashi
Analysis of random structures and algorithms. Computational complexity of logical and algebraic theories.

Kåre Jelling Kristoffersen
Real-time process calculi and model checking. Research assistant currently extending and implementing EPSILON.

Coming Events

Theme Events

For the courses *The modal mu-Calculus* and *Synthetic Domain Theory* as well as the Workshop on Full Abstraction, we refer to the '95 theme activities on page 2.

Dynamic Algorithms

Giuseppe F. Italiano, University of Salerno, will visit BRICS in February '95. Professor Italiano is one of the leading experts on Dynamic Algorithms and he'll give a short course (8 lectures) on this topic.

TACAS

19–20 May, TACAS, *Workshop on Tools and Algorithms for the Construction and Analysis of Systems*, University of Aarhus, organised by Arne Skou and Kim G. Larsen (<{ask,kgl}@iesd.auc.dk>).

TAPSOFT

22–26 May, TAPSOFT '95, *Sixth International Joint Conference on the Theory and Practice of Software Development*, University of Aarhus; organised by P. D. Mosses (chair <tapsoft@daimi.aau.dk>), M. Nielsen, M. I. Schwartzbach.

Dissertation Abstracts

Block Ciphers - Analysis, Design and Applications

by Lars Ramkilde Knudsen

In this thesis we study cryptanalysis, applications and design of secret key block ciphers. In particular, the important class of *Feistel ciphers* is studied, which has a number of rounds, where in each round one applies a cryptographically weak function.

Applications

The main application of block ciphers is that of encryption. We study the available modes of operation for encryption, introduce a new taxonomy for attacks on block ciphers and derive a new theoretical upper bound for attacks on block ciphers. Also another important application of block ciphers is studied; as building blocks for cryptographic hash functions. Finally we examine how to use block ciphers as building blocks in the design of digital signature schemes. In particular we analyze Merkle's proposed scheme and show that under suitable and reasonable conditions, Merkle's scheme is secure and practical.

Cryptanalysis

We study the most important known attacks on block ciphers, linear cryptanalysis and differential cryptanalysis and introduce a new attack based on simple relations. Differential cryptanalysis makes use of so-called *differentials* (A, B), i.e. a pair of plaintexts with difference A, which after a certain number of rounds result in a difference B with a non-negligible probability. This fact can be used to derive (parts of) the secret key. Ideas of how to find the best such differentials are given. Also it is shown that higher order differentials, where more than two plaintexts are considered at a time, and partial differentials, where only a part of (A, B) can be predicted, both have useful applications. The above attacks and our new methods of attacks on block ciphers, is applied to the specific block ciphers, DES, LOKI'91, s^2-DES, xDES[1] and xDES[2].

Attacks on hash functions based on block ciphers are studied and new attacks on a large class of hash functions based on a block cipher, including three specific proposed schemes, are given. Also a fourth scheme, the AR Hash function, belonging to another class of hash functions based on block ciphers is studied. The scheme is faster than the known standard ones and was used in practice by German banks. It is shown that the scheme is completely insecure.

Design

We discuss principles for the design of secure block ciphers. For both linear and differential cryptanalysis we establish lower bounds on the complexities of success of attacks. It is furthermore shown that there exist functions, which can be used to construct block ciphers provable secure against both linear and differential attacks, the two most important attacks known to date. Furthermore we define so-called *strong key schedules*. A block cipher with a strong key schedule is shown to be secure against attacks based on simple relations and the improved immunity to other attacks is discussed. Also we give a simple design of a strong key schedule. A well-known and wide-spread way of improving the security of a block cipher is by means of multiple encryption, i.e. where a plaintext block is processed several times using the same (component) block cipher, but with different keys. We study the methods of multiple encryption and give a new proposal of a scheme, which is provable as secure as the component block cipher using a minimum number of component keys.

See DAIMI PB - 485, Department of Computer Science, University of Aarhus.

Timed Modal Specifications — A theory for verification of real–time concurrent systems

by Jens Christian Godskesen

In this thesis we have worked mainly in two areas:

- analysis of four behavioural equivalences of the real–time processes algebra TCCS with respect to the existence and non–existence of expansion theorems and with respect to reduction of parallel composition.
- definition of a new theory Timed Modal Specifications (TMS), for real–time and concurrent systems. The theory is a conservative extension of TCCS and has been implemented in the automatic verification tool EPSILON.

As part of the analysis of TCCS we consider the alternative model for real–time systems, Timed Graphs. We define an algebra of timed graphs and prove that any process belonging to a specific class of TCCS processes may be represented as a graph in the algebra. The main result of the work with timed graphs is The Gap Theorem. It is used to prove that for the strongest behavioural equivalence of TCCS one can not hope to reduce the number of parallel components in a TCCS process expression.

We prove that only for the weakest of the four behavioural equivalences of TCCS there exists an expansion theorem. Hence, we have demonstrated that the traditional techniques know from untimed process algebra can not be directly used for deciding and for axiomatizing the three strongest TCCS equivalences since these techniques eliminate parallel composition by interleaving and non–determinism.

The theory of TMS is developed in a process algebraic setting with the overall motivation of aiming at generality. Generality is obtained by the introduction of partial or loose specifications. That

is, systems that allow for various behavioural inequivalent implementations may be specified. Looseness of specifications is achieved through the introduction of two different modes of events: events which are required and events which are allowed. The introduction of looseness permits for the definition of refinement orderings extending in a natural way the behavioural equivalences of TCCS.

Since we know that not all the TCCS behavioural equivalences can be decided using directly existing techniques from untimed process algebra we can not hope to prove decidability of the TMS refinement orderings using these techniques. Hence decidability of the refinement orderings must be proven using a new technique. We prove decidability of two of the TMS refinement orderings in full detail using a symbolic technique, for the two remaining refinement orderings we outline ideas of how to prove decidability. The proofs are in some sense constructive in that they may be viewed as directions of how to develop decision procedures. Taking advantage of this constructiveness TMS has been implemented in the automatic verification tool EPSILON.

Based on the symbolic techniques used in the proofs for deciding whether a specification T is refined by a specification S one may give constructive proofs of how to synthesize diagnostic information in case T is not refined by S. The diagnostic information is based on a logical characterization of the refinement orderings and the diagnostics provided is a formula satisfied by one specification but not the other. We give such constructive proofs in full detail for two of the TMS refinement orderings, for the two remaining refinement orderings we sketch ideas of how to perform the constructive proofs. The generation of diagnostic information has been implemented in EPSILON.

Finally, we apply TMS and EPSILON on two examples: a train crossing and a timed stop–and–wait protocol. In the appendix a short user's guide to EPSILON can be found.

The thesis is published as IESD report R 94-2039, Department of Mathematics and Computer Science, Aalborg University.

A HOL Basis for Reasoning about Functional Programs

by Sten Agerholm

Domain theory is the mathematical theory underlying denotational semantics. This thesis presents a formalization of domain theory in the Higher Order Logic (HOL) theorem proving system along with a mechanization of proof functions and other tools to support reasoning about the denotations of functional programs. By providing a fixed point operator for functions on certain domains which have a special undefined (bottom) element, this extension of HOL supports the definition of recursive functions which are not also primitive recursive. Thus, it provides an approach to the long-standing and important problem of defining non-primitive recursive functions in the HOL system.

Our philosophy is that there must be a direct correspondence between elements of complete partial orders (domains) and elements of HOL types, in order to allow the reuse of higher order logic and proof infrastructure already available in the HOL system. Hence, we are able to mix domain theoretic reasoning with reasoning in the set theoretic HOL world to advantage, exploiting HOL types and tools directly. Moreover, by mixing domain and set theoretic reasoning, we are able to eliminate almost all reasoning about the bottom element of complete partial orders that makes the LCF theorem prover, which supports a first order logic of domain theory, difficult and tedious to use. A thorough comparison with LCF is provided.

The advantages of combining the best of the domain and set theoretic worlds in the same system are demonstrated in a larger example, showing the correctness of a unification algorithm. A major part of the proof is conducted in the set theoretic setting of higher order logic, and only at a late stage of the proof domain theory is intro-

duced to give a recursive definition of the algorithm, which is not primitive recursive. Furthermore, a total well-founded recursive unification function can be defined easily in pure HOL by proving that the unification algorithm (defined in domain theory) always terminates; this proof is conducted by a non-trivial well-founded induction. In such applications, where non-primitive recursive HOL functions are defined via domain theory and a proof of termination, domain theory constructs only appear temporarily.

Published as BRICS report RS-94-44 below.

New in the BRICS Report Series

ISSN 0909-0878

20 Peter D. Mosses and Mart´n Musicante. *An Action Semantics for ML Concurrency Primitives.* July 1994. 21 pp. Appears in Proc. FME '94 (Formal Methods Europe, Symposium on Industrial Benefit of Formal Methods), LNCS 873, 1994.

21 Søren Riis. *Count(q) does not imply Count(p).* July 1994. 55 pp.

22 Torben Braüner. *A General Adequacy Result for a Linear Functional Language.* August 1994. 39 pp. Presented at MFPS '94.

23 Søren Riis. *Finitisation in Bounded Arithmetic.* August 1994. 31 pp.

24 Søren Riis. *A Fractal which violates the Axiom of Determinacy.* August 1994. 3 pp.

25 Søren Riis. *Bootstrapping the Primitive Recursive Functions by 47 Colors.* August 1994. 5 pp.

26 Søren Riis. *Count(q) versus the Pigeon-Hole Principle.* August 1994. 3 pp.

27 Torben Braüner. *A Model of Intuitionistic Affine Logic from Stable Domain Theory (Revised and Expanded Version).* September 1994. 19 pp. Full version of paper appearing in: ICALP '94, LNCS 820, 1994.

28 Oded Goldreich. *Probabilistic Proof Systems.* September 1994. 19 pp.

29 Ronald Cramer and Ivan Damgård. *Secure Signature Schemes Based on Interactive Protocols.* September 1994. 24 pp.

30 Thore Husfeldt. *Fully Dynamic Transitive Closure in Plane Dags with one Source and one Sink.* September 1994. 26 pp.

31 Noam Nisan and Amnon Ta-Shma. *Symmetric Logspace is Closed Under Complement.* September 1994. 8 pp.

32 Alexander Aiken, Dexter Kozen, and Ed Wimmers. *Decidability of Systems of Set Constraints with Negative Constraints.* October 1994. 33 pp.

33 Vladimiro Sassone. *Strong Concatenable Processes: An Approach to the Category of Petri Net Computations.* October 1994. 40 pp. To appear in TAPSOFT '95.

34 Henrik Reif Andersen, Colin Stirling, and Glynn Winskel. *A Compositional Proof System for the Modal μ-Calculus.* October 1994. 18 pp. Appears in: Proceedings of LICS '94, IEEE Computer Society Press.

35 Gerth Stølting Brodal. *Partially Persistent Data Structures of Bounded Degree with Constant Update Time.* November 1994. 24 pp.

36 Alexander A. Razborov. *On provably disjoint* **NP***-pairs.* November 1994. 27 pp.

37 Jaap van Oosten. *Fibrations and Calculi of Fractions.* November 1994. 21 pp.

38 Ivan B. Damgård and Lars Ramkilde Knudsen. *Enhancing the Strength of Conventional Cryptosystems.* November 1994. 12 pp.

39 Ivan Damgård, Oded Goldreich, and Avi Wigderson. *Hashing Functions can Simplify*

Zero-Knowledge Protocol Design (too). November 1994. 18 pp.

40 Luca Aceto and Anna Ingólfsdóttir. *CPO Models for GSOS Languages — Part I: Compact GSOS Languages.* December 1994. 70 pp. An extended abstract of the paper will appear in: *Proceedings of CAAP '95*, LNCS, 1995.

41 Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. *On Data Structures and Asymmetric Communication Complexity*. December 1994. 17 pp.

42 Dany Breslauer and Leszek Gąsieniec. *Efficient String Matching on Coded Texts.* December 1994. 20 pp.

43 Luca Aceto and Alan Jeffrey. *A Complete Axiomatization of Timed Bisimulation for a Class of Timed Regular Behaviours (Revised Version)*. December 1994. 18 pp. To appear in *Theoretical Computer Science.*

44 Sten Agerholm. *A HOL Basis for Reasoning about Functional Programs.* December 1994. PhD thesis. viii+224 pp.

45 Jørgen H. Andersen, Kåre J. Kristoffersen, Kim G. Larsen, and Jesper Niedermann. *Automatic Synthesis of Real Time Systems.* December 1994. 17 pp.

46 Amos Beimel, Anna Gál, and Mike Paterson. *Lower Bounds for Monotone Span Programs.* December 1994. 14 pp.

47 Kim G. Larsen, Bernhard Steffen, and Carsten Weise. *A Constraint Oriented Proof Methodology based on Modal Transition Systems.* December 1994. 13 pp.

48 Jens Chr. Godskesen and Kim G. Larsen. *Synthesizing Distinguishing Formulae for Real Time Systems.* December 1994. 21 pp.

BRICS Notes Series, 1994

ISSN 0909-3206

1 Peter D. Mosses, editor. *Proc. 1st International Workshop on Action Semantics* (Edinburgh, 14 April, 1994), May 1994. 145 pp.

2 David Basin. *Induction Based on Rippling and Proof Planning. Mini-Course.* August 1994. 62 pp.

3 Sven Skyum, editor. *Complexity Theory: Present and Future* (Aarhus, Denmark, 15–18 August, 1994), September 1994. v+213 pp.

4 Peter D. Mosses, editor. *Abstracts of the 6th Nordic Workshop on PROGRAMMING THEORY* (Aarhus, Denmark, 17–19 October, 1994), October 1994. v+52 pp.

5 Andrew M. Pitts. *Some Notes on Inductive and Co-Inductive Techniques in the Semantics of Functional Programs,* DRAFT VERSION. December 1994. vi+135 pp.

6 Uffe H. Engberg, Kim G. Larsen, and Peter D. Mosses, editors. *Proceedings of the 6th Nordic Workshop on Programming Theory* (Aarhus, Denmark, 17–19 October, 1994), December 1994.

Technical Contributions

Of Simple Problems and Hard Solutions

by Devdatt Dubhashi

1 Of Balls and Bins

Suppose we throw a certain number of balls into a certain number, n, of bins, uniformly and independently at random. For $i \in [n]$, let X_i be the random variable denoting the number of balls in the ith bin. The question is: how are the X_i related? Although the balls are thrown uniformly and independently at random, these variables are *not* independent; in particular, their sum is fixed, being equal to the number of balls thrown. Intuitively, the variables are *negatively correlated* in the manner indicated by the following innocous looking statements:

$$\Pr[X_1 \geq 3 \mid X_2 \geq 5] \leq \Pr[X_1 \geq 3] \quad (1)$$

$$\Pr[X_1 \geq 3 \mid X_2 \geq 5, X_3 \geq 4] \leq \Pr[X_1 \geq 3] \quad (2)$$

$$\Pr[X_1 + X_2 \geq 5 \mid X_3 + X_6 + X_{17} \geq 6] \leq \Pr[X_1 + X_2 \geq 5]. \quad (3)$$

$$\Pr[X_1 \geq 3, X_2 \geq 4 \mid X_3 \geq 5, X_4 \geq 6] \leq \Pr[X_1 \geq 3, X_2 \geq 4 \mid X_3 \geq 5] \leq \Pr[X_1 \geq 3, X_2 \geq 4]. \quad (4)$$

$$\Pr[X_1 \geq 3 \mid X_2 \geq 4, X_3 \geq 5] \leq \Pr[X_1 \geq 3 \mid X_2 \geq 4]. \quad (5)$$

Although these statements appear almost self–evident, they seem to be surprisingly hard to prove. In particular, a direct counting argument with binomial coefficients seems to lead nowhere.

2 Of Personal Struggles

The problem was brought to my attention at the Max–Planck–Institut für Informatik, Saarbrücken by Torben Hagerup; indeed it would require a person of his meticulous zeal to even consider this a serious problem deserving of careful proofs! It was required in the analysis of algorithms for load balancing developed by his student Thomas Lauer. Thomas later told me that he had mentioned the problem also to Kurt Mehlhorn, who at first look declared that he would have a proof in five minutes! The next morning he had admitted the problem was not trivial. Interestingly, he said that had Thomas simply made those statements in his thesis, he would have gladly passed over them without a second thought, but now that he had brought up the issue, a proof had to be found!

In the summer of '94, Desh Ranjan from the New Mexico State University was visiting MPI for a period of three months. We knew each other very well from our student days at Cornell, and this was a welcome opportunity to interact again. I got him interested in the problem and we were confident of disposing it off in a week or two, so that we could then move on to "more substantial" problems. The weeks turned into months of fruitless scribblings of calculations on reams and reams of paper, and our misplaced confidence was soon humbled! One day I mentioned the possibility of applying the celebrated "FKG Inequality" to the problem. It seemed an intriguing possibility, but I had not been able to successfully apply it. I watched with grim amusement as he went through the same sequence of unsuccessful steps as I had done. One morning though, he rushed into the office completely out of breath – "It works!" he gasped. He had the glimmerings of an idea that turned into the proof in the next section.

3 Of Elegant Proofs

Given a finite, distributive lattice L, a function $f : L \to \mathsf{R}$ is said to be non–decreasing (non–increasing) if $x \leq_L y$ implies $f(x) \leq f(y)$ (respectively, $x \leq_L y$ implies $f(x) \geq f(y)$). A function $\mu : L \to \mathsf{R}^+$ on a distributive lattice L, is called *log–supermodular* if

$$\mu(x)\mu(y) \leq \mu(x \vee y)\mu(x \wedge y)$$

for all $x, y \in L$. Motivated by a problem in statistical mechanics, Fortuin, Kasteleyn and Ginibre, and independently Sarkar, proved the following celebrated inequality [4, 6, 5, 1]:

Theorem 1 (FKG Inequality) *Let L be a finite distributive lattice and let $\mu : L \to \mathsf{R}^+$ be a log–supermodular function. Then if $f, g : L \to \mathsf{R}^+$ are both non–decreasing or both non–increasing, we have*

$$(\sum_{x\in L} \mu(x)f(x)) \cdot (\sum_{x\in L} \mu(x)g(x)) \quad \leq$$
$$(\sum_{x\in L} \mu(x)f(x)g(x)) \cdot (\sum_{x\in L} \mu(x)) \qquad .$$

If one of the functions is non–decreasing and the other is non–increasing then the reverse inequality holds.

We shall use the FKG inequality to give a short proof of assertions (1) and (3) in § 1. A possible *configuration* of the experiment can be represented by a vector $\mathbf{a} := (a_1, \cdots, a_m)$, with $a_i \in [n]$ for each $i \in [m]$. This is the configuration where ball i goes into bin a_i for each $i \in [m]$. Define the lattice L to be the set of all such configurations ordered component–wise:

$$\mathbf{a} \leq_L \mathbf{b} \quad \leftrightarrow \quad a_i \leq b_i, \text{ for each } i \in [m].$$

It turns out that this in fact defines a distributive lattice, with join and meet given by the following equation on the components:

$$(a \vee b)_i := \max(a_i, b_i) \quad \text{and} \quad (a \wedge b)_i := \min(a_i, b_i).$$

Distributivity follows because of the following property of the integers:

$$\min(a, \max(b, c)) = \max(\min(a, b), \min(a, c)),$$

$$\max(a, \min(b, c)) = \min(\max(a, b), \max(a, c)).$$

Now define $\mu : L \to \mathsf{R}^+$ by $\mu(\mathbf{a}) := 1/n^m$ for each $\mathbf{a} \in L$. So defined, μ is trivially log–supermodular, and crucially, it makes each configuration equally likely, representing the fact that the balls are thrown uniformly and independently at random into the bins.

For a configuration $\mathbf{a}$, we introduce naturally, for $i \in [n]$,

$$X_i(\mathbf{a}) := |\{j \mid a_j = i\}|.$$

This gives the number of balls in the ith bin in configuration $\mathbf{a}$.

Let $I, J \subseteq [n]$ be two index sets such that either $I \cap J = \emptyset$ or $I \cup J = [n]$; with no loss of generality, we can arrange it by renumbering, so that $J := \{1, \cdots, |J|\}$ and $I := \{n - |I| + 1, \cdots, n\}$. Let t_I, t_J be arbitrary non–negative integers. Define $f, g : L \to \mathsf{R}^+$ as indicator functions by

$$f(\mathbf{a}) := \begin{cases} 1, & \text{if } \sum_{i\in I} X_i(\mathbf{a}) \geq t_I; \\ 0, & \text{otherwise.} \end{cases}$$

$$g(\mathbf{a}) := \begin{cases} 1, & \text{if } \sum_{j\in J} X_j(\mathbf{a}) \geq t_J; \\ 0, & \text{otherwise.} \end{cases}$$

The definition of the lattice order ensures that f is non–decreasing while g is non–increasing. Also, note that

$$\begin{aligned} \sum_{\mathbf{a}\in L} \mu(\mathbf{a})f(\mathbf{a}) &= \frac{1}{n^m} \cdot |\{\mathbf{a} \in L : \sum_{i\in I} X_i(\mathbf{a}) \geq t_I\}| \\ &= \Pr[\sum_{i\in I} X_i \geq t_I]. \end{aligned}$$

Similarly

$$\sum_{\mathbf{a}\in L} \mu(\mathbf{a})g(\mathbf{a}) = \Pr[\sum_{j\in J} X_j \geq t_J]$$

$$\sum_{\mathbf{a}\in L} \mu(\mathbf{a})f(\mathbf{a})g(\mathbf{a}) = \Pr[\sum_{i\in I} X_i \geq t_I, \sum_{j\in J} X_j \geq t_J].$$

Applying the FKG Inequality, we get the following correlation inequality on the random variables $X_i, i \in [n]$:

Theorem 2 *Let $I, J \subseteq [n]$ be index sets such that $I \cap J = \emptyset$ or $I \cup J = [n]$, and let t_I, t_J be arbitrary non–negative integers. Then*

$$\Pr[\sum_{i\in I} X_i \geq t_I \mid \sum_{j\in J} X_j \geq t_J] \leq \Pr[\sum_{i\in I} X_i \geq t_I].$$

REMARK 1: By taking I, J to be singletons, this also implies that $\Pr[X_i \geq t_i \mid X_j \geq t_j] \leq \Pr[X_i \geq t_i]$ for any distinct $i, j \in [n]$ and any non–negative integers t_i, t_j.

4 Of Frustrations

Once we had proven the theorem in the previous section – which yields assertions (1) and (3) in § 1 – we were confident that we would be able to swiftly extend it to prove also the remaining assertions there. In this, once again, our confidence was belied. Try as we might, we could not come up with a suitable definition of a lattice to yield the required assertions. In retrospect, it was a happy miracle that the lattice we had for the proof actually satisfied the properties needed! In the end, we did manage to give proofs of those statements, but by an entirely different route, see [2].

Another frustrating experience is trying to convince people about its worth. I remember the reaction of the audience when Desh gave a *Mittagseminar* at MPI. It was a – in retrospect predictable – reaction of disbelief: Why should one need to employ a sophisticated theorem on lattices to prove simple things about balls and bins. Juris Hartmanis who was also in the audience, asked in incredulous surprise: "Didn't Bernoulli do this already?"! Some people went further and asked: "What is there to prove at all?"! Perhaps only an audience that is itself subjected to months of frustrating calculations on reams of paper would appreciate the proof.

5 Of Other Victims

After our months of struggle in the summer, we sporadically came across literature that showed that we were not the only victims of innocent looking problems. In this section, I relate another similar story.

Suppose that we are conducting a tennis tournament in two leagues, an A league, with players $a_1, \ldots, a_m$ and a B league with players $b_1, \ldots, b_n$. At first there are some intra–league matches, that determine some ranking within each of the two leagues. We can think of a partial order P as being established with relations of the form $a_i < a_j$ or $b_i < b_j$. But there are *no* relations of the form $a_i < b_j$. Let us denote by $\Pr[a_1 < b_1 \mid P]$, the conditional probability that a_1 loses to b_1 given the partial order P. (More precisely, this is the ratio of the number of linear extensions of P in which $a_1 < b_1$ to the number of all linear extensions of P.) Suppose now, that some inter–league matches are played, and all information garnered through those are of the form $a_i < b_j$. Call the resulting partial order P'. Intuitively, the partial order P' seems to propagate the prejudice that the a's are inferior to the b's. Hence, if we denote by $\Pr[a_1 < b_1 \mid P']$ the corresponding conditional probability with the new partial order, Graham, A.C. Yao and F. Yao [3] conjectured that

$$\Pr[a_1 < b_1 \mid P] \leq \Pr[a_1 < b_1 \mid P'].$$

Shepp [7] proved this conjecture by an application of the FKG Inequality. He also discovered the following counter–intuitive curiousity. Suppose the original order P also had information of the form $a_i < b_j$. Intuitively, this would go even more to reinforce the belief that the a's are inferior to the b's. However, the conjecture is *false* in this case!

Now, suppose, once again that we are in the setting of a tennis tournament. Suppose we have information that $x < z$. How does that alter the state of affairs without any knowledge at all? Rival and Sands made the reasonable conjecture that the event $x < y$ for some other y is reinforced by this belief, so

$$\Pr[x < y \mid x < z] \geq \Pr[x < z].$$

Shepp [8] used a devilishly clever application of the FKG Inequality to transform this so–called xyz–Conjecture into the xyz–Theorem. On the other hand, there are the following surprises: it is not true that

$$\Pr[x < u < v \mid x < y] \geq \Pr[x < u < v].$$

Winkler [9] defines two partial orders P_1 and P_2 to be *universally correlated*, $P_1 \uparrow P_2$, if

$$\Pr[P_1 \mid P_2] \geq \Pr[P_1].$$

Thus, Shepp's result is $x < y \uparrow x < z$. Winkler [9] gives a characterisation for determining when two partial orders are universally correlated.

6 Of Hopes

We would like to believe Shepp's hope that there must be a way to establish correlation inequalities via systematic applications of the FKG inequality. In particular, we would like to be able to prove the other statements in § 1 by an application of the FKG inequality. Moreover, we would like to answer more complex questions involving *mixed* conditions. Thus, what can one say about $\Pr[X_1 \geq 5 \mid X_2 \geq 4, X_3 \leq 2]$ versus the unconditional probability?

References

[1] N. Alon, P. Erdös and J. Spencer, *The Probabilistic Method*, John Wiley, 1992.

[2] D. Dubhashi and D. Ranjan: "Correlation Inequalities for the Probabilistic Analysis of Algorithms", Max–Planck–Institut für Informatik, Technical Report 143, August 1994.

[3] R. Graham, A.C. Yao and F.F. Yao: "Some Monotonicity Properties of Partial Orders", SIAM J. Of Discrete and Algebraic Methods, 1:3, pp. 251–258, 1980.

[4] C.M. Fortuin, J. Ginibre and P.N. Kasteleyn, "Corelational Inequalities for Partially Ordered Sets", *Communications of Mathematical Physics* 22, pp. 89–103, 1971.

[5] R.L. Graham, "Application of the FKG Inequality and its Relatives" in A. Bachem, M. Grötschel and B. Korte Ed., *Mathematical Programming: The State of the Art*, Springer–Verlag, 1983.

[6] T.K. Sarkar, "Some Lower Bounds of Reliability", Tech. Report. 124, Dept. of Operations Research and Statistics, Stanford University, 1969.

[7] L.A. Shepp, "The FKG Inequality and Some Monotonicity Properties of Partial Orders", *SIAM Journal on Algebraic and Discrete Methods* 1, pp. 295–299, 1980.

[8] L.A. Shepp, "The XYZ Conjecture and the FKG Inequality" *Annals of Probability* 10, pp. 824-827, 1982.

[9] P.M. Winkler, "Correlations among Partial Orders", *SIAM Journal on Algebraic and Discrete Methods*, 4(1), pp. 1–7, 1983.

Calendar of Events

Date	Event
Jan –	'95 theme on Logic in Semantics.
Feb	*Giuseppe F. Italiano*, University of Salerno; 8 lectures on Dynamic Algorithms
Feb	*Igor Walukiewicz*, BRICS; around 5 lectures on: The modal mu-Calculus.
5–15 Mar	*Pino Rosolini*, University of Genova; 5 lectures on: Synthetic Domain Theory.
18 Apr –	Two week Workshop on Full Abstraction.
19–20 May	TACAS.
22–26 May	TAPSOFT '95.

BRICS Address and World Wide Web

For further information, hard copies of information material and reports of the BRICS Series as well as enrolment into the BRICS newsgroup, please contact **BRICS** at

Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: <BRICS@daimi.aau.dk>

or, in writing, at

Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK - 8000 Aarhus C
Denmark.

You can get access to information from BRICS through World Wide Web (WWW) and anonymous FTP. To connect to the BRICS WWW entry, open the URL:

```
http://www.daimi.aau.dk/BRICS/
```

The BRICS WWW entry contains updated information about most of the topics covered in the newsletter as well as access to electronic copies of information material and reports of the BRICS Series (look under Publications).

To access the information material and reports of the BRICS Series via anonymous FTP do the following:

```
ftp ftp.daimi.aau.dk
cd pub/BRICS
get README.
```

BRICS *Newsletter*

ISSN 0909-6043

Editors: Glynn Winskel & Uffe H. Engberg
Lay-out: Uffe H. Engberg
Publisher: BRICS
Print: Institute of Mathematics
University of Aarhus