COUNCIL OF THE
EUROPEAN ASSOCIATION FOR
THEORETICAL COMPUTER SCIENCE

PRESIDENT: Burkhard Monien
VICE PRESIDENTS: Giorgio Ausiello
                   Don Sannella
                   Paul Spirakis
Treasurer: Dirk Janssens
Bulletin Editor: Maria Serna

Luca Aceto  Iceland  Antonin Kucera  Czech Republic
Susan Albers Germany  Jan van Leeuwen  The Netherlands
Krzysztof Apt Netherlands  Madhavan Mukund  India
Pierre-Louis Curien France  Mogens Nielsen  Denmark
Josep Diaz  Spain  Catuscia Palamidessi  France
Zoltan Esik  Hungary  David Peleg  Israel
Fedor Fomin  Norway  Giuseppe Persiano  Italy
Lance Fortnow USA  Jean-Eric Pin  France
Leslie Ann Goldberg United Kingdom  Vladimir Sassone  United Kingdom
Giuseppe F. Italiano Italy  Andrzej Tarlecki  Poland
Christos Kaklamanis Greece  Thomas Wilke  Germany
Juhan Karhumaki Finland  Gerhard Woeginger  The Netherlands

Past Presidents:
EATCS COUNCIL MEMBERS

EMAIL ADDRESSES

Luca Aceto ............................................. luca@ru.is
Susanne Albers ...................................... albers@informatik.hu-berlin.de
Giorgio Ausiello ..................................... ausiello@dis.uniroma1.it
Krzysztof Apt ......................................... apt@cwi.nl
Pierre-Louis Curien ................................. curien@pps.jussieu.fr
Josep Diaz ............................................. diaz@lsi.upc.es
Zoltán Ésik ............................................ ze@inf.u-szeged.hu
Fedor Fomin ............................................ fomin@ii.uib.no
Lance Fortnow ....................................... fortnow@eecs.northwestern.edu
Leslie Ann Goldberg .............................. L.A.Goldberg@liverpool.ac.uk
Giuseppe F. Italiano ............................... italiano@disp.uniroma2.it
Dirk Janssens ......................................... Dirk.Janssens@ua.ac.be
Christos Kaklamanis ............................... kakl@ceid.upatras.gr
Juhani Karhumäki ..................................... karhumak@cs.utu.fi
Antonin Kucera ....................................... tony@fi.muni.cz
Jan van Leeuwen .................................... jan@cs.uu.nl
Burkhard Monien ..................................... bm@upb.de
Madhavan Mukund .................................. madhavan@cmi.ac.in
Mogens Nielsen ....................................... mn@brics.dk
Catuscia Palamidessi .............................. catuscia@lix.polytechnique.fr
David Peleg .......................................... peleg@wisdom.weizmann.ac.il
Giuseppe Persiano .................................. giuper@dia.unisa.it
Jean-Eric Pin ......................................... Jean-Eric.Pin@liafa.jussieu.fr
Don Sannella ......................................... dts@dcs.ed.ac.uk
Vladimiro Sassone .................................. vs@ecs.soton.ac.uk
Maria Serna .......................................... mjserna@lsi.upc.edu
Paul Spirakis ......................................... spirakis@cti.gr
Andrzej Tarlecki ..................................... tarlecki@mimuw.edu.pl
Thomas Wilke ........................................ wilke@ti.informatik.uni-kiel.de
Gerhard Wöeginger ................................. g.j.woeginger@math.utwente.nl
Bulletin Editor: Maria Serna, Barcelona, Spain
Cartoons: DADARA, Amsterdam, The Netherlands

The bulletin is entirely typeset by PDF\TeX\ and Con\TeX\ in TX\font s.

All contributions are to be sent electronically to bulletin@eatcs.org

and must be prepared in \LaTeX\ using the class beatcs.cls (a version of the standard \LaTeX\ article class). All sources, including figures, and a reference PDF version must be bundled in a ZIP file.

Pictures are accepted in EPS, JPG, PNG, TIFF, MOV or, preferably, in PDF. Please, consult http://www.eatcs.org/bulletin/howToSubmit.html.

We regret we are unfortunately not able to accept submissions in other formats, or indeed submission not strictly adhering to the page and font layout set out in beatcs.cls. We shall also not be able to include contributions not typeset at camera-ready quality.

The details can be found at http://www.eatcs.org/bulletin, including class files, their documentation, and guidelines to deal with things such as pictures and overfull boxes. When in doubt, email bulletin@eatcs.org.

Deadlines for submissions of reports are January, May and September 15th, respectively for the February, June and October issues. Editorial decisions about submitted technical contributions will normally be made in 6/8 weeks. Accepted papers will appear in print as soon as possible thereafter.

The Editor welcomes proposals for surveys, tutorials, and thematic issues of the Bulletin dedicated to currently hot topics, as well as suggestions for new regular sections.

The EATCS home page is http://www.eatcs.org
# Table of Contents

## EATCS Matters
- Letter from the President ........................................... 3
- Letter from the Editor .................................................. 6
- The EATCS Award 2010 .................................................. 7
- Obituary: Amir Pnueli (1941-2009) .................................. 9

## Institutional Sponsors
- EATCS Sponsors .......................................................... 13

## EATCS News
- News from New Zealand, by C.S. Calude .......................... 17

## The EATCS Columns
- The Algorithmic Game Theory Column, by M. Mavronikolas
  - Some Algorithmic Questions About Nash Equilibria, by M. Mavronikolas ........................................... 35
- The Complexity Column, by J. Torán
  - Multilinear Polynomials Modulo Composites, by A. Chattopadhyay ....................................................... 52
- The Concurrency Column, by L. Aceto
  - Temporal Logics for Games, by F. Laroussinie .................. 79
- The Distributed Computing Column, by P. Fotouhi
  - Robust Simulation of Shared Memory: 20 Years After by H. Attiya ......................................................... 99
- The Logic in Computer Science Column, by Y. Gurevich
  - How to Base Probability Theory on Perfect-Information Games, by G. Shafer, V. Vovk, R. Chyczhyla .............. 115

## Reports from Conferences
- CS&P 2009, by M. Kudlek .............................................. 151
- DCM 2009, by M. Kudlek .............................................. 153
- DLT 2009, by M. Kudlek .............................................. 155
- FCT 2009, by M. Kudlek .............................................. 160
- IPMP in Honor of Academician Arto Salomaa, by M. Kudlek ................................................................. 164
- NCMA 2009 by M. Kudlek .............................................. 169
- PETRI NETS 2009, by J. Kiehn and M. Koutny .................. 171
The 8th International Conference on Unconventional Computation 2009, by S. Stepney ...................................... 173
The 7th International Conference on Words (WORDS 2009), by A. Carpi and C. De Felice ............................................. 176
ABSTRACTS OF PHD THESES .................................................. 181
EATCS LEAFLET ................................................................. 187
Letter from the President

Dear EATCS members,

It is again my great pleasure, as usual in this time of the year, to announce the assignment of the EATCS Award: the EATCS Award 2010 has been assigned to Kurt Mehlhorn for his decisive influence on the developments of algorithms, and, more generally of computer science as a whole in many ways. The laudatio, published in this issue of the bulletin, illustrates his distinguished scientific career ranging from his fundamental contributions to a wealth of algorithmic topics like data structures, computational geometry, parallel computing, and complexity theory, to his outstanding success in shaping the field of algorithmic engineering, most prominently represented by LEDA, the Library of Efficient Data Types and Algorithms. The proposal has been made by the selection committee consisting of Eugenio Moggi, Pavlos Spirakis and Emo Welzl (chair), and it has been unanimously approved by the EATCS Council members. The Award will be presented in a ceremony that will take place during ICALP 2010 in Bordeaux. On behalf of the whole EATCS community I would like to offer our congratulations to Kurt for this well-deserved award!

The organization of ICALP 2010 is proceeding well. The submission deadline is approaching. The Program Committees are chaired by Pavlos Spirakis (track A), Samson Abramsky (track B) and Friedhelm Meyer auf der Heide (track C). As in the last year, track A is dedicated to
Algorithms, Complexity and Games, track B to Logic, Semantics, Automata and Theory of Programming, and track C to Foundations of Networked Computations. The conference will also host at least four satellite workshops: AlgoGT (Workshop on Algorithmic Game Theory: Dynamics and Convergence in Distributed Systems), DYNAS 2010 (International Workshop on Dynamic Networks, Algorithms and Security), ALGOSENSORS 2010 (International Workshop on Algorithmic Aspects of Wireless Sensor Networks), and SDKB 2010 (Semantics in Data and Knowledge Bases). A call asking for additional workshop proposals will be launched this month. The Conference Chairs Claude Kirchner and Cyril Gavoille together with their team are doing an excellent job. For further details you may take a look at the ICALP 2010 website: http://icalp10.inria.fr/

The organization of ICALP 2011 has already started. ICALP will be organized next year in Zurich by the ETH team Michael Hoffmann, Juraj Hromkovic, Ueli Maurer, Angelika Steger, Emo Welzl and Peter Widmayer. The conference will have the same tracks as this year. The program committees will be chaired by Jiri Sgall (track A), Luca Aceto (track B) and Monika Henzinger (track C).

You may have already seen that the EATCS website has been modified a lot. It is easier to survey now and the members-only section contains additional information. On the one hand you can find information about EU research funding (the secretarial office will regularly place information about new calls here) and on the other hand there is now a section for young researchers where universities can
advertise their Postdoc- and maybe also other positions. In this context, I would like to encourage you to send corresponding news to our secretarial office.

Together with Maria Serna (as EATCS bulletin editor) and Ioannis Chatzigiannakis (as EATCS secretary) we have agreed that from the June issue on the bulletin will be published fully electronically using the "open journal system". Note that the June issue is the 101st issue of our bulletin, so there will be a real change for this special issue.

On the council meeting in Rhodes, we agreed to run a poll about the distribution of the bulletin. I propose to delay this poll until we have collected first experiences with the electronic version of our bulletin.

I have to end this letter with a sad message. On November 2, 2009, our friend and colleague Amir Pnueli passed away as a result of a brain hemorrhage. Amir Pnueli was one of the most influential scientists of our community. For his achievements he received the ACM Turing Award in 1996 "for seminal work introducing temporal logic into computing science and for outstanding contributions to program and system verification". You can find the obituary in this bulletin issue. On behalf of EATCS, I wish to express his family and all his colleagues the deepest mourning of our community.

Burkhard Monien, Paderborn
February 2010
Letter from the Bulletin Editor

Dear Reader,

First of all, I wish all of you a happy and successful 2010. This is the 100th issue of the Bulletin. As in each centenary this is the time to celebrate that so many years have past with a fruitful contribution to the EATCS community. I would like to thank all the contributors of the Bulletin for their efforts. However, we have decided to look forward and not celebrate with this 100th issue but with the first issue of the hundreds to come Bulletin issues.

Our columns deliver the usual richness and variety of interesting contents. We start with a survey on algorithmic questions about Nash equilibria by Marios Mavronicolas ("Algorithmic Game Theory Column"), and proceed through an analysis of the problem of computing boolean functions by multilinear polynomials by Arkadev Chattopadhyay ("Computational Complexity Column"). An introduction to the modeling and specification of open systems using games by Françoise Larroussinie ("Concurrency Column"). Hagit Attiya contributes with an overview of robust simulations of shared memory on message passing systems ("Distributed Computing Column"), an excursion on the way to compute probabilities based on perfect information games by Glenn Shafer, Vladimir Vovk and Roman Chychyla ("Logic in Computer Science Column").

Let me draw your attention to the many reports on conferences included in this issue, and express my thanks for the effort of the authors in keeping track of the main ideas discussed at them.

I hope you'll enjoy the contents of this 100th Bulletin issue,

Maria Serna, Barcelona
February 2009
THE EATCS AWARD 2010

LAUDATIO FOR KURT MEHLHORN

The EATCS Award is awarded annually to honor a scientist with widely recognized contributions to the field of theoretical computer science throughout a distinguished scientific career. The Committee, consisting of Emo Welzl (Chair), Pavlos Spirakis and Eugenio Moggi in charge of evaluating the nominations to the 2010 EATCS Award has come to the decision to honor

PROFESSOR KURT MEHLHORN

with the EATCS Award 2010 for his decisive influence on the developments of algorithms, and, more generally, of computer science as a whole in many ways. The decision has been unanimously approved by the EATCS Council. The Award will be assigned during a ceremony that will take place in Bordeaux (France) during ICALP 2010 (July 5-12, 2010).

Most remarkable are Kurt Mehlhorn’s fundamental contributions to a wealth of algorithmic topics: Data structures, computational geometry, geometric computing and computer algebra, parallel computing, VLSI-design, complexity theory, combinatorial optimization, and graph algorithms.

One outstanding contribution is his shaping of the field of algorithm engineering, most prominently represented by LEDA, the Library of Efficient Data Types and Algorithms, which was initiated and originally written by Kurt Mehlhorn and Stefan Näher. The LEDA book is a shining example of theory meets practice with its interweaving of theoretical analysis and careful software engineering. Besides efficiency, the library impresses by the treatment of correctness, let it be in the direction of robust, consistent and at the same time efficient geometric computation or in the direction of the employment of certificates which allow checking of results independent of the correctness of algorithms and their implementation. Nowadays LEDA is used extensively both in academia and industry and it has stimulated the development of numerous other more specialized libraries for combinatorial algorithms.
Equally impressive to its scientific contributions is his service to the computer science community. He was a driving force behind the establishment of the Max Planck Institute for Informatics and the Research Center for Computer Science Schloss Dagstuhl, both indispensable hubs for the exchange of ideas in computer science. He is also an initiator for ESA, the European Symposium on Algorithms and for a series of European projects (ALCOM, CGAL) integrating European algorithm research.

Kurt Mehlhorn is a dedicated teacher. His talks are excellent examples of clarity. He has written several books, including a three volume EATCS Monograph on Data Structures and Algorithms, which has been a prime source for students and researchers in the field for many years. An impressive list of his PhD students have by now become prominent scientists.

His dedication to science and education has made him a role model for colleagues and young scientists. Kurt Mehlhorn clearly deserves this distinction which we are both happy and proud to award to him.

The EATCS Award 2010 Committee
Amir Pnueli was born in Nahalal, Israel, on April 22, 1941 and passed away in New York City on November 2, 2009 as a result of a brain hemorrhage.

Amir founded the computer science department in Tel-Aviv University in 1973. In 1981 he joined the department of mathematical sciences at the Weizmann Institute of Science, where he spent most of his career and where he held the Estrin family chair. In recent years, Amir was a faculty member at New York University where he was a NYU Silver Professor.

Throughout his career, in spite of an impressive list of achievements, Amir remained an extremely modest and selfless researcher. His unique combination of excellence and integrity has set the tone in the large community of researchers working on specification and verification of programs and systems, program semantics and automata theory. Amir was devoid of any arrogance and those who knew him as a young promising researcher were struck how he maintained his modesty throughout his career.

Amir has been an active and extremely hard working researcher till the last moment. His sudden death left those who knew him in a state of shock and disbelief.

For his achievements Amir received the 1996 ACM Turing Award. The citation reads: For seminal work introducing temporal logic into computing science and for outstanding contributions to program and system verification. Amir also received, for his contributions, the 2000 Israel Award in the area of exact sciences. More recently, in 2007, he shared the ACM Software System Award. Furthermore, Amir had received the honoris causa doctorate degree from several of universities.

Temporal logic is by now a routine approach to specification and verification of concurrent, reactive, hybrid and real-time systems. The origin of this approach can be traced back to a single paper written by Amir, The Temporal Logic of Programs, 18th Annual Symposium on Foundations of Computer Science (FOCS) 1977, pp. 46-57.

At that time, the main approach to program verification was the assertional method of Tony Hoare, extended in 1976 to parallel programs by Susan Owicki and David Gries, and, independently, by Leslie Lamport. This method was applicable to parallel algorithms but using it one could prove only certain properties.
Amir recognized that an all-out verification of parallel programs called for new ideas. Temporal Logic, proposed by Amir Pnueli, is a formalism that allows for expressing relevant program properties adequately, simply, and elegantly. By suggesting proof methods that allow one to systematically establish these properties he demonstrated that temporal logic can be used in a natural way for verification of parallel programs. In his proposal he achieved a synthesis between two fields from different sciences: mathematical logic and computer science.

Amir’s early work led to Lamport’s classification of the properties of concurrent programs into the categories of safety and liveness properties (for Amir, liveness is further divided into guarantee, response, persistence, and reactivity properties). Temporal logic became a simple and intuitive specification formalism of programs and systems. This in turn led the ground for the subsequent work on model checking, an automatic approach to verification of hardware components and finite-state systems. Nowadays, hardware manufacturers commonly use tools based on temporal logic for hardware verification.

Once the temporal logic approach to program specification and verification has matured, Amir published, in the nineties, jointly with Zohar Manna, two books on this subject devoted, respectively, to program specification and verification of safety properties. Unfortunately, the book about the main novelty of temporal logic—verification of liveness properties—exists only as a draft in Amir’s computer.

Since then this approach turned out to be equally well applicable to reactive, hybrid and real-time program and systems. Amir was the instigator to these and other new areas of applications of temporal logic, and always at the forefront of the development of any such application.

In recent years Amir worked on numerous new topics, to wit translation validation, where instead of verifying a translator (e.g., a compiler) one verifies each translation, verification of infinite-state systems (e.g., systems of numerous processes with similar behavior), automation of deductive reasoning, software verification (e.g., programs with heaps), verification of distributed systems, applications of mathematical and logical methods to formal specifications, and synthesis of reactive, real-time, discrete, continuous, and hybrid systems. His recent (2006) work on synthesis of synchronous “General Reactivity 1” properties resulted in a chip, which was hanging on the wall in his NYU office.

Throughout his career Amir has been a highly prolific researcher, often engaging in a collaboration with others. The DBLP website lists 250 papers of his, written with 136 coauthors. To many of us Amir has been an inspiring collaborator and mentor, as well as a warm and always supportive friend. We shall miss him dearly.

Krzysztof R. Apt and Lenore D. Zuck
INSTITUTIONAL SPONSORS
BiCi, Bertinoro international Center for informatics
Bertinoro, Italy

BRICS, Basic Research in Computer Science
Aarhus, Denmark

CTI, Computer Technology Institute
Greece

Elsevier
Amsterdam, The Netherlands

Microsoft Research
Cambridge, United Kingdom

Springer-Verlag
Heidelberg, Germany
1 Scientific and Community News


1. The latest CDMTCS research reports are ([http://www.cs.auckland.ac.nz/staff-cgi-bin/mjd/secondcgi.pl](http://www.cs.auckland.ac.nz/staff-cgi-bin/mjd/secondcgi.pl)):

368. C. S. Calude and E. Calude. The Complexity of the Four Colour Theorem. 08/2009


370. E. Calude. The Complexity of Goldbach’s Conjecture and Riemann’s Hypothesis. 08/2009
A Dialogue about Computations and Natural Sciences with Professor Giuseppe Longo

Giuseppe Longo, a well-known mathematician and computer scientist, is a professor at the École Normale Supérieure in Paris and Directeur de Recherche at CNRS. Professor Longo has extensively published in many areas including logic and theory of computation, type theory, category theory and their applications to computer science, interfaces between mathematics, physics, biology, philosophy of mathematics and cognitive sciences.

He is the editor of the book series "Visions des sciences", Hermann, Paris, and is serving as an editor to the following academic journals: “Mathematical Structures in Computer Science” (editor-in-chief), “Information and Computation”, “Theoretical Informatics and Applications”, “JUCS”, “La Nuova Critica”, “The European Review”, “Journal of Mind Theory”. Professor Longo has supervised 33 (research-oriented) Master Theses and 15 PhD Theses. He has been an invited lecturer at 30 international conferences and he gave more than 150 seminar talks in universities or research institutions in Europe, USA and Asia. Professor Longo is a member of the Academia Europaea since 1992.

Cristian Calude: Your academic career took you from Pisa (where you graduated “Laurea” (cum laude) in Mathematics and then spent 15 years as an academic) to US (UC Berkeley, MIT and Carnegie Mellon for three years), and then to ENS in Paris (since 1989). How enriching have these moves been?

Giuseppe Longo: Learning from others, the exchange with others is crucial, in scientific work. Very few researchers can do relevant work without interaction. I learned a lot from these very enriching contexts, beginning with the extraordinary milieu of mathematics and informatics in Pisa, in the ’70s and ’80s, and the subsequent American experience. Some collaborations, in particular at MIT and Carnegie Mellon with many, but also in Britain (R. Hindley) and Holland (H.
Barendregt) were fundamental for me. And then the complex network of interactions with colleagues of three disciplines I am enjoying in Paris, in particular with the physicists F. Bailly and T. Paul. The main lesson I try to give to my students is that "two interacting brains think and produce much more than the double". But collaborating is very hard: good researchers are very careful in choosing collaborators and the exchange itself is difficult.

There is a growing fashion instead in the use of words referring to “competition” in research. But this is not how science goes: the difficult and productive side is collaborating, exchanging, learning from others and... to go further on, together. If, time to time, one has to compete for finite resources or for a position, this is part of the game, not the purpose nor the joy of science.

Stressing competition of teams and individuals is a real disaster for scientific research. And it is largely borrowed from the current cultural hegemony of the financial markets, where traders are in continual competition and they compete on a mostly empty economic/productive content. As a further imitation, many institutions entrust “Independent Evaluation Agencies” (like those that evaluated with the highest score Enron in 2001 and Lehman Brothers in 2008, until the “day before”...) for judging scientific work. And self-appointed, “science independent” agents provide automatic indexes for classifying researchers. I proposed an “Editors’ Note: Bibliometrics and the Curators of Orthodoxy” (downloadable from my web page) for Mathematical Structures in Computer Science, the CUP journal I direct, on this theme. It was approved by all the 34 members of the board, from 11 different countries. Yes, in contrast to competition, collaboration and exchange are fundamental and moving enhances them greatly.

CC: Moving is a joy and pain. How hard was for your family, especially, your daughter?

GL: Hard, but stimulating. My wife started a new carrier, first by a master in Pittsburgh, when I was teaching at CMU, then a PhD in Paris, finally a university position in France, but this was tough on her. My daughter moved between the age of 5 and 8 between three very different school systems (USA, Italy, France), not easy for a child. Now she is trilingual, though, and she can ... “adjust” to almost any life context. And she is doing a beautiful thesis work on Italian Quattrocento, in Paris, with very frequent trips to Italy.

CC: We share two main interests: incompleteness and randomness. Let’s talk about incompleteness first.

GL: OK, but I prefer to tackle the issue by relating it to some Mathematics of Physics, in a preliminary way. In a short note of 2001, I suggested that Poincaré’s three-body theorem is an epistemological predecessor of Gödel’s undecidability result, in particular because Hilbert’s completeness conjecture is a
meta-mathematical revival of Laplace idea of the predictability of formally (equationally) determined systems. For Laplace, once the equations are given, you can completely derive the future states of affairs (with some, preserved, approximation). Or, more precisely, in “Le système du monde”, he claims that the mathematical mechanics of moving particles, one by one, two by two, three by three . . . compositionally and completely “covers” or makes understandable the entire Universe. And, as for celestial bodies, by this progressive mathematical integration, “We should be able to deduce all facts of astronomy”, says he.

The challenge, for a closer comparison, is that Hilbert was speaking about a purely mathematical “yes or no” questions, while unpredictability shows up in the relation between a physical system and a mathematical set of equations (or evolution function). That is, in order to give unpredictability, Poincaré’s Negative Result, as he called his proof of the non-analyticity of the equations for three-body system, needs a reference to physical measure. Measure is always, in classical (and relativistic) physics, an interval, that is an approximation. And non-observable initial fluctuations may give observable, thus unpredictable, evolutions, in presence, typically, of non-linearity of the mathematical modelling (main reasons: the initial interval expands exponentially - this is measured by Lyapounov exponents—and it is “mixed”—its order is not preserved).

In order to relate consistently unpredictability to undecidability, one needs to effectivize the dynamical spaces and measure theory (along the lines of Lebesgues’s measure), the loci for dynamic randomness. This allows to have a sound and purely mathematical treatment of the epistemological issue (and obtain a convincing correspondence between unpredictability and undecidability). I will go back to the work on this while answering your next question, on randomness.

As for Gödel’s incompleteness, when studying Poincaré’s theorem, I understood that the two results share also a methodological point: they both destroy the conjecture of predictability/completeness from inside. Poincaré does not need to refer concretely to a physical process that would not be predictable, by measuring it “before and after”. He shows, from the pure analysis of the equations, that the resulting bifurcations and homoclinic intersections (between stable and unstable manifolds) lead to deterministic unpredictability (of course, the equations are derived in reference to three bodies in their gravitational fields, similarly as Peano Axioms are invented in reference to the ordered structure of numbers).

Gödel as well, by playing the purely formal game, formally constructs an undecidable sentence, with no reference what so ever, in the statements and proofs in his 1931 paper, to “semantics”, “truth” or alike, that is to the underlying mathematical structure.

Modern “concrete incompleteness” theorems (that is, Gödel-Girard’s normalisation, Paris-Harrington or Friedman-Kruskal theorems) resemble instead Laskar’s results of the ’90s, where “concrete unpredictability” is shown for the solar sys-
tem. In reference to the best possible astronomical measures, Laskar shows that the evolution of our beloved system is provably unpredictable, globally, beyond one million years (one hundred years, when considering only Earth). Similarly, concrete incompleteness was given by proving (unprovability and) truth over the (standard) model.

More generally, I view the incompleteness of our formal (and equational) approaches to knowledge a fundamental epistemological issue. And this why we permanently need new science: by inventing new principles of conceptual constructions we change directions, propose new intelligibilities, grasp or organise new fragments of the World. There no such a thing as “the final solution to the foundational problem” in mathematics (as Hilbert dreamed—a true nightmare), nor in other sciences.

And finally, then, my other current interest, biology. The “incompleteness” of the molecular theories for understanding life phenomena is a similar issue. No way to understand/derive completely embryogenesis nor phylogenesis (evolution) by looking only at the four letters of the bases of DNA (the formal language of molecular biology). More precisely, in this very different context, “completeness” philosophically corresponds to the largely financed myth that the stability and the organisation of the DNA and the subsequent molecular cascades completely determine the stability and the organisation of the cell and the organism. This is false, since the stability and the organisation of the cell and the organism causally contribute to the stability and the organisation of the DNA and the subsequent molecular cascades. Thus the analysis of the global structure of the cell (and the organism) must parallel the absolutely crucial molecular analyses. The hard philosophical point to explain now, to my friends in molecular biology, is that “incomplete” does not mean useless (well, I worked most of my life in Type Theory, lambda-calculus and related formal systems . . . ), but that we badly need also an autonomous theory of organism and further develop the (fantastic) darwinian theory of evolution.

By the way, randomness plays a crucial role in evolution, but also, and it is increasingly believed so, in embryogenesis. But . . . what kind of randomness? physics, classical/quantum, proposes two distinct notions of randomness . . .

CC: What aspects of randomness interest you?

GL: Classical (physical) randomness is unpredictability of deterministic systems in finite time (dice trajectories are perfectly determined: they follow the Hamiltonian, a unique geodetics; yet, they are very sensitive to initial and contour conditions . . .: it is, in general, not worth writing the equations). Now, Martin-Löf’s (and Chaitin’s) number-theoretic randomness is for infinite sequences. How may this yield a connection then between Poincaré’s unpredictability and Gödel’s undecidability?
As I said, physical randomness, as deterministic unpredictability, is a matter at the interface “equations/process” and shows up at finite time. Yet, also physical randomness may be expressed as a limit or asymptotic notion and, by this, it may be soundly turned into a purely mathematical issue: this is Birkhoff’s ergodicity (for any observable, limit time averages coincide with space averages). That is, physical randomness, as a mathematical limit property, lives in formal systems of equations or evolution functions: in their measurable spaces, they may engender infinite random trajectories or generic points, in the ergodic sense. And this sense applies in (weakly chaotic) dynamical systems, within the frame Poincaré’s geometry of dynamical systems.

As for algorithmic randomness, Martin-Löf randomness is a “Gödelian” notion of randomness, as it is based on recursion theory and yields a strong form of undecidability for infinite 0-1 sequences (in short, a sequence is random if it passes all effective statistical tests). Recently, M. Hoyrup and C. Rojas, under Galatolo’s and my supervision, proved that dynamic randomness (a la Poincaré, thus, but at the purely mathematical limit, in the ergodic sense), in suitable effectively given measurable dynamical systems, is equivalent to (a generalisation of) Martin-Löf randomness (Schnorr’s randomness). This is a non-obvious result, based also on a collaboration with P. Gacs, spreading along two “entangled” doctoral dissertations (defended in June 2008, a nice example on how two collaborating individuals may produce more than the “double”).

In the last three years, I have been teaching a course at ENS, in Paris (and once in Rome III), along these parallel lines, from Poincaré and Gödel to algorithmic randomness. The course (the program is on my web page) was organised with a colleague in quantum mechanics at ENS, Thierry Paul: we alternated one two hours lecture each and he introduced the EPR paradox (Einstein’s and others’ paper on entanglement) and its modern consequences, quantum computing. As Thierry moved to Polytechnique, I took up part of his lectures since we are doing some joint work on a logical and (modern) physical understanding of EPR. By the way, EPR is dedicated to prove the incompleteness (!) of QM. Their argument is (beautiful, but) wrong as it is based on the impossibility of entanglement.

As for quantum randomness, note now that, because of entanglement, it differs from classical: if two classical dice interact and then separate, the probabilistic analysis of their values are independent. When two quanta interact and form a “system”, they can no longer be separated: measures on them give correlated probabilities of the results (mathematically, they violate Bell’s inequalities).

CC: What is the link between computability, continuity and Church-Turing thesis?

GL: The idea hinted in the book and in several papers with Bailly and Paul, two physicists, is that the mathematical structures, constructed for the intelligibility of physical phenomena, according to their continuous (mostly in physics) or discrete
nature (generally in computing), may propose different understandings of Nature. In particular, the “causal relations”, as structures of intelligibility (we "understand Nature" by them), are mathematically related to the use of the continuum or the discrete and may deeply differ (in modern terms: they induce different symmetries and symmetry-breakings).

But what discrete (mathematical) structures are we talking about? I believe that there is one clear mathematical definition of “discrete”: a structure is discrete when the discrete topology on it is “natural”. Of course, this is not a formal definition, but in mathematics we all know what “natural” means. For example, one can endow Cantor’s real line with the discrete topology, but this is not “natural” (you do not do much with it); on the other hand, the integer numbers or a digital data base are naturally endowed with the discrete topology (even though one may have good reasons to work with them also under a different structuring).

Church’s thesis, introduced in the ’30s after the functional equivalence proofs of various formal systems for computability, concerns only computability over integers or discrete data types. As such, it is an extremely robust thesis: it ensures that any sufficiently expressive finitistic formal system over integers (a Hilbertian-type logic-formal system) computes exactly the recursive functions, as defined by Gödel, Kleene, Church, Turing . . . . This thesis therefore emerged within the context of mathematical logic, as grounded on formal systems for arithmetic computations, on discrete data types.

The very first question to ask is the following: If we broaden the formal framework, what happens? If we want to refer to continuous (differentiable) mathematical structures, the extension to consider is to the computable real numbers. Are, then, the various formalisms for computability over real numbers equivalent, when they are maximal? An affirmative response could suggest an extension of Church thesis to computability on “continua”. Of course, the computable reals are countably many, but they are dense in the “natural” topology over Cantor’s reals, a crucial difference, as we shall see.

With this question, we begin to get near to physics, since it is within spatial and often also temporal continuity that we represent dynamical systems, that is, most mathematical models for classical physics. This does not imply that the World is continuous, but only that we have said many things thanks to continuous tools as very well specified by Cantor (but his continuum is not the only possible one: Lawvere and Bell, say, proposed another without points).

Now, from this equivalence of formalisms, at the heart of Church’s thesis, there remains very little when passing to computability over real numbers: the theories proposed are demonstrably different, in terms of computational expressiveness (the classes of defined functions). The various systems (recursive analysis, whose ideas were first developed by Lacombe and Grezgorczyk, in 1955-57; the Blum, Shub and Smale, BSS, system; the Moore-type recursive real functions; differ-
ent forms of “analog” systems, among which threshold neurones, the GPAC . . . )
yield different classes of “continuous” computable functions. Some recent work
established links, reductions between the various systems (more precisely: pair-
wise relations between subsystems and/or extensions), yet, the full equivalence as
in the discrete case is lost. Moreover, and this is crucial, these systems have no
“universal function” in Turing’s sense. And this, for a fundamental reason, which
has to be analysed closely.
If one endows non-trivial space continua with the interval topology (the “real”
topology), there is no way to have an isomorphism between spaces of different
dimension (see below). This isomorphism, instead, is needed for having the uni-
versal map and, in general, for computability on the discrete. Its work spaces may
be of any finite dimension: they are all effectively isomorphic, “Cartesian dimen-
sion” does not matter!
This is highly unsuitable for physics. First, the dimensional analysis is a funda-
mental tool (one cannot confuse energy with force, nor with the square of energy
. . . ). Second, dimension is a topological invariant, in all space manifolds for clas-
sical and relativistic physics. This is shown by the fact that if two such spaces
have isomorphic open subsets, then they have the same dimension. This is one of
the most beautiful correspondence between mathematics and physics. Take phys-
ical measure, which is always an interval (it is approximated, by principle, clas-
sically), as a “natural” starting point for the metric (thus the interval topology),
then you prove that this crucial notion for physics, dimension, is a topological
invariant. Discrete computability destroys this: a cloud of isolated points has no
dimension, per se, and you may, for all theoretical purposes, encode them on a
line. When you have dimension back, in computability over continua, where the
trace of the interval topology maintains good physical properties, you loose the
universal function and the equivalence of systems. Between the theoretical world
of discrete computability and physico-mathematical continua there is a huge gap.
One cannot even extend to the second a sound form of Church Thesis. While I
believe that one should do better than Cantor as for continua, I would not give a
penny for a physical theory where dynamics takes place only on discrete spaces,
departing from physical measure, dimensional analysis and the general relevance
of dimensions in physics (again, from heat propagation to mean field theory, to
relativity theory . . . space dimension is crucial).
As for the relevance of the discrete, quantum mechanics started exactly by the
discovery of a key (and unexpected) discretisation of light absorption or emission
spectra of atoms. Then, a few dared to propose a discrete lower bound to mea-
sure of action, that is of the product energy × time. It is this physical dimension
that bares a discrete structure. Clearly, one can then compute, by assuming the
relativistic maximum for the speed of light, a Planck’s length and time. But in
no way space and time are thus organised in small “quantum boxes”. And this
The Bulletin of the EATCS

is the most striking and crucial feature of quantum mechanics: the “systemic” or entanglement effects, which yield inseparability of observables. No discrete space topology is natural. That is, these quantum effects are the opposite of a discrete, separated organisation of space, while being at the core of its scientific originality. In particular, they motivate quantum computing (as well as our analysis of quantum randomness above). As a matter of fact, Thierry Paul and I claim that the belief in an absolutely separable topology of space continua is Einstein’s mistake in EPR.

A final remark. In general, the discrete is not an approximation of classical continua. In even weakly chaotic systems, a difference by approximation (below measure, typically) quickly leads to major (observable) differences in evolutions. And the approximation relation is at most reversed. In some, not all, dynamical systems, the “shadowing lemma” holds: for any discrete trajectory, there is continuous one approximating it. The quantification is the other way round. This is an important result in Numerical Analysis, as it guarantees that a discrete trajectory on the screen is not meaningless: one can find a continuous one approximating it. Of course, it does not start with the same initial point, in general.

In summary, continua, Cantorian or not, take care rather well (they are not an absolute) of the approximated nature of physical measure, which is represented as an interval: the unknowable fluctuation is within the interval. Classically, I insist, the relevance of measure is derived from Poincaré’s results (changes below measure induce major differences over time). And physical measure is our only form to access “reality”. The arithmetizing foundation of mathematics went along another (and very fruitful) direction, based on perfectly accessible data types. Poincaré firmly opposed to the underlying philosophy of knowledge, by deep, but informal, reflections on “action” in the physical world.

CC: We have entropy and negative entropy. You invented anti-entropy.

GL: Traditionally, information is considered as negentropy (Brillouin). Then, by definition:

1. the sum of a quantity of information (negentropy) and an equal quantity of entropy gives 0;

2. information (Shannon, but also Kolmogorov) is “insensitive to coding” (one can “encrypt” and “decrypt” as much as one wishes but the information content will not be lost/gained, in principle).

I believe however that this notion, of which the applications are numerous, is not sufficient for an analysis of the living state of matter. DNA (usually considered as digital information) is the most important component of the cell, as I said, but it is necessary to analyse the organisation of the organism, as an observable specific
to biological theorisation.
Here, the collaboration with Francis Bailly, a physicist also interested in biology, has been very important. He actually was my teacher in many aspects of natural sciences (Francis recently passed away: a recorded conference in his memory may be accessed from my web page). Concerning biological (morphological) complexity, we have proposed the notion of anti-entropy to define it (or quantify it in terms of complexity of cellular, functional and phenotypical differentiation).
In short, biological complexity may be understood as “information specific to the form”, including the intertwining and enwrapping of levels of organisation. Its use in metabolic balance equations has produced a certain number of results mentioned in a recent long article. We have, in particular, examined systems far from equilibrium and analysed diffusion equations of biomass over biological complexity as anti-entropy, following Schrödinger’s “operational method” in quantum mechanics. This has enabled to operate a mathematical reconstruction of this diffusion, which corresponds to the paleontological data presented by Gould for the evolution of species.

Anti-entropy is compatible with information as negentropy, but it must be considered as a strict extension, in a logical sense, of the thermodynamics of entropy. Typically, the production of entropy and that of anti-entropy are summed in an “extended critical singularity”, an organism, never zero, in contrast to Brillouin’s and others’ negentropy. As it is linked to spatial forms, anti-entropy is “sensitive to coding”, contrarily to digital information (it depends on the dimensions of embedding manifolds, on folds, on singularities . . . ).

In short, over the last six years, in several collaborations and by supervising four theses, we have compared physical (dynamic) randomness with algorithmic randomness (at the center of algorithmic theories of information); we worked at a theory of “extended criticality” (living objects persist in an “extended critical state”); we have added anti-entropy (a “geometrical extension” of the notion of information) to thermodynamic (in)equalities and balance equations; we have begun modelling biological rhythms and time in two-dimensional manifolds, a sort of non-trivial geometrization of time (and, perhaps, a quite useful one, for the digital simulation of cardiac rhythms I am developing with a PhD student, M. Montevil).

The scientific finality of this work may also entail some epistemological consequences, I hope: it should participate to the epistemological debate regarding the notion of information, the updating of its theoretical principles, as part of the many existing interactions with physics and biology. A a possible outcome of these interactions could be to start thinking to . . . the next machine. Aren’t we a little tired of this, nice, but rather old “Discrete Data Types Machine”?

CC: Is it possible to summarise the ideas of your book Mathématiques et sciences de la nature. La singularité physique du vivant (Hermann, Paris, 2006) with F.
The Bulletin of the EATCS

Bailly? Are you planning an English version?

GL: Yes, there is an ongoing translation in English. In the book, Francis and I attempt to identify the organising concepts of some physical and biological phenomena, by means of an analysis of the foundations of mathematics and of physics, in the aim of unifying phenomena, of bringing different conceptual universes into dialog. The analysis of the role of “order” and of symmetries in the foundations of mathematics is linked to the main invariants and principles, among which the geodesic principle (a consequence of symmetries), which govern and confer unity to the various physical theories. Moreover, we attempt to understand causal structures, a central element of physical intelligibility, in terms of symmetries and their breakings. The importance of the mathematical tool is also highlighted, enabling us to grasp the differences in the models for physics and biology which are proposed by continuous and discrete mathematics, such as computational simulations.

As for biology, being particularly difficult and not as thoroughly examined at a theoretical level, we propose a “unification by concepts”, an attempt which should always precede mathematisation, that we later tried in some papers. This constitutes an outline for unification also basing itself upon the highlighting of conceptual differences, of complex points of passage, of technical irreducibilities of one field to another. Indeed, a monist point of view such as ours should not make us blind: we, the living objects, are surely just big bags of molecules or, at least, this is our main metaphysical assumption. The point though is: which theory can help us to better understand these bags of molecules, as they are, indeed, rather funny (singular?), from the physical point of view. Technically, this singularity is expressed by the notion of “extended criticality”, a notion that logically extends the pointwise critical transitions in physics.

CC: In what sense do you think physical or biologically processes “compute”?

GL: The Discrete State Machines that compute are a remarkable invention, based on a long history. As I hint in the paper “Critique of Computational Reason in the Natural Sciences”, this story begins with the invention of the alphabet, probably the oldest experience of discretisation. The continuous song of speech, instead of being captured by the design of concepts and ideas (by recalling “meaning”, like in ideograms), is discretised by annotating phonetic pitches, an amazing idea (the people of Altham, in Mesopotamia, 3300 B.C.). Meaning is reconstructed by the sound, which acts as a compiler, either loud or in silence (but only after the IV century A.D. we learned to read “within the head” !).

I insist that the crucial feature of alphanumeric discretisation is the invention of a discrete coding structure, which is far from obvious. Think also of the originality of Gödel-numbering, an obvious practice now, but another remarkable invention.
Turing’s work followed: the Logical Computing Machine (LCM), as he first called it, at the core of our science (right/left, 0, 1 . . .). Of course, between the alphabet and Turing, you also have Descartes “discretisation” of thought (stepwise reasoning, along a discrete chain of intuitive certitudes . . .) and much more. When, after 1948 or so, Turing gets again interested in physics, he changed the name to his LCM: in the 1950 and 1952 papers, he calls it Discrete State Machine (this is what matters for its physical behaviour). And twice in his 1950 paper (the “imitation game”), he calls it “Laplacian”. Its evolution is theoretically predictable, even if there may be practical unpredictability (too long programs to be grasped, says he).

So, we invented an incredible stable processor, which, by working on discrete data types, does what it is expected to do. And it iterates, very faithfully. Primitive recursion and portability of software are forms of iterability: iteration and update of a register, do what you are supposed to do, respectively, even in slightly different contexts, over and over again. For example, program the evolution function of the most chaotic strange attractor you know. Push “restart”: the digital evolution, by starting on the same initial digits, will follow exactly the same trajectory (on a paper on Turing’s imitation game I discuss the simulation of the double pendulum, a chaotic device). This makes no physical sense, but it is very useful (also in meteorology: you may restart your turbulence, exactly, and try to better understand how it evolves . . .). Of course, you may imitate unpredictability by some pseudo-random generator or by . . . true physical randomness, added ad hoc. But this is cheating the observer, in the same way Turing’s imitation of a woman’s brain is meant to cheat the observer, not to “model” the brain. He says this explicitly, all the while working in his 1952 paper, at a model of morphogenesis, as (non-)linear dynamics. Observe, finally, that our colleagues in networks and concurrency are so good that programming in network is reliable: programs do what they are supposed to do, they iterate and . . . give you the web page you want, identically, one thousands time, one million times. And this is hard, as physical space-time, which we better understand by continua and continuous approximations, steps in, yet still on discrete data types, which allow perfect iteration.

Those who claim that the Universe is a big digital computer, miss the originality of this machine of ours. It is like believing that, when we speak, we produce sequences of letters: this is a cartoon’s vision of language and misses the originality of our invention, the alphabet, an early musical notation (chinese children have a different view: their cartoons’ bubbles evoke concepts). When we construct computers, we make the far from obvious miracle of producing a reliable, thus programmable, physical device, iterating as we wish and any time we wish, even in networks. One should not miss the principles that guided this invention, as well as the principles by which we understand physical dynamics.

By the way, are the main physical constants, $G, c, h$, computable (real numbers)?
It depends on the choice of the reference system and the metrics, of course. So, fix $h = 1$. Then, you have to renormalise all metrics and re-calculate, by equations, dimensional analyses and physical measure, $G$ and $c$. But physical measure will always give an interval, as we said, or, in quantum frame, the probability of a value. If one interprets the classical measure interval as a Cantorian continuum, the best way, so far, to grasp fluctuations, then . . . where are $G$ and $c$?

Non-computable reals form a set of Lebesgues measure 1 . . . . Yet, the most striking mistake of many "computationalists" is to say: but, then, some physical processes would super-compute (compute non-computable functions)! No, this is not the point. Most physical processes, simply do not define a mathematical function. In order to have a classical process to define a function, you have to fix a time for input, associate a (rational) number to the interval of measure and . . . let the process go. Then you wait for the output time and measure again. In order, for the process, to define $f(x) = y$, at a rational input $x$ it must always associate a rational output $y$. But if you restart, say, your physical double pendulum on $x$, that is within the interval of the measure which gave you $x$, a minor (thermal, say) fluctuation, below that interval $x$, will yield a different observable result $y'$ after very short time. So, a good question would be, instead: consider a physical process that defines a (non-trivial) function, is this function computable?

The idea then would be that the process is sufficiently insensitive to initial conditions (some say: robust) as to actually define a function. But, then one should be able to partition the World in little cubes of the smallest size, according to the best measure as for insensitivity (fluctuations below that measure do not affect the dynamics). If the Accessible World is considered finite (but . . . is it?), then one can make a list out of the finite input-output relation established by the given process. This is a "program": is it compressible?

As for biology, what can I say? 60% of fecundations in mammals fail (to not reach a birth): a very bad performance for the DNA as a program. While iterability is at the core of software (and hardware) design, our fantastic invention, the key principle for understanding life, at the phenotypic level, is variability, a form of non-iterability. It is crucial for evolution, but also ontogenesis, that a cell is never identical to the mother cell. So, the principles of intelligibility are the exact opposite: the failure of most fecundations corresponds to the possibility that a mutant better fits a changing environment (affecting the mother’s womb, say). Of course, some molecular processes iterate, but there is an increasing tendency to analyse molecular cascades in terms of statistical phenomena (and this is where good computational imitations may help to understand, by some use of pseudo-randomness or by networks interactions). This opens the way to an increasing role of epigenetics and, thus, to the relevance of downwards regulating effects, from the cell and the organism to DNA expression.
CC: You argue that incomputability phenomena are more important to physics than computable ones? After all, the laws of physics seem more computable than incomputable?

GL: Your second question refers to the effectiveness of our mathematical writing of physical invariants: of course, equations, evolution functions ... are given by sums, products, exponents, derivations, integrations ... all effective operations. Moreover, no one is so crazy to put an incomputable real as a coefficient or exponent in an equation (even if $\hbar$ could be so ...). This gives us remarkable approximations and, most often, qualitative information: Poincaré’s geometry of dynamical systems or Hadamard’s analysis of the geodetic flow on hyperbolic surfaces, do not give predictions, but very relevant global information (by attractors, for example, or regularities in flows ... that we beautifully see today, as never before, by fantastic approximations, “shadowed” on our computers screens).

I do not know (absolute) laws of Nature, but our constructive theorising on the phenomenal veil, at the interface between us and the World. This active constructions are of course effective (we use the alphabet, effective operations and codings, I insist). While predictable processes are not many in Nature: you can predict a few forthcoming Eclipses, at human time scale, but the Solar System is chaotic in astronomical times, as Poincaré proved and Laskar quantified (and computed!). Unpredictable ones are the mathematical and computational challenge. And a computable physical process is, by definition, deterministic and predictable. In order to predict (pre-dicere, “to say in advance” in Latin), just “say” or write the corresponding program and compute in advance; more precisely, the results discussed above, by showing the equivalence of unpredictability and (strong) undecidability, ML-randomness, prove this fact, by logical duality. Unpredictability may pop-out in networks and this because of physical space-time (we then make them computable and predictable-reliable by forcing semaphores, handling interleaving ...). In Nature, many (most, fortunately) processes escape predictions, thus our computations. Fortunately, otherwise there would be no change, nor life in particular: randomness is crucial. And when we compute unpredictable evolutions, we just approximate their initial part, as I said, or give qualitative information, both very relevant tasks. But engineers put some more cement than computed, to take care of vibrations below measure ....

Now, the only mathematical way I know to define randomness, in classical physics, is Birkhoff’s ergodicity. But it is very specific (certain dynamics). Otherwise, randomness is given in terms of probability measure. But this is unsatisfactory, as probability gives a measure of randomness, not a definition. It is the theory of algorithms, thanks to Martin-Löf, Chaitin and you, that gave a fully general, mathematical, notion of randomness, as a strong form of incomputability, independently of probability theory. Again, physical (classical) randomness is de-
terministic unpredictability and, by the results above and more in the literature, the role of computational randomness further comes to the limelight. In particular, it provides a very flexible theory of randomness: you can adjust the class of effective randomness tests (Martin-Löf, Schnorr... and many more). Our joint hope is that this may help to better grasp, for example, the mathematical difference between classical and quantum randomness.

CC: If all papers and books would be destroyed by a disaster, but you could keep just one, which one would you choose? Why?

GL: I do not think I would survive to this, but, just to give a partly random answer: Weyl’s “Philosophy of Mathematics and of Natural Sciences”. Along the lines of “Das Kontinuum”, it radically departs from the Hilbertian alphabetic myths. Mathematics, actually human thought, for formalists and computationalists, is reducible to the matching and replacement of sequences of letters: no geometric judgements, no association of gestalts..., this is why this Laplatican mechanics of thought is incomplete. The proof has also a “geometric structure”, a remark by Poincaré, and, by this, it is “sensitive to codings”. This is also why its formal coding is incomplete. Reasoning is not a chain whose strength is that of the weakest ring, as Descartes claimed, but a network, a rope made of many interlaced wires, as suggested by Peirce, reinforcing each other and coupling to meaning and to forms of life. And Weyl globally develops a deep and broad philosophy of knowledge, well beyond the parody of his views in predicativist or intuitionistic terms.

CC: If all your papers and books would be destroyed by a disaster, but you could keep just one, which one would you choose? Why?

GL: The recent paper on “anti-entropy”, where some (minor) aspects of Evolution are mathematically described, because... it is the last one and because, while working at it, I increasingly learned to love Darwin.

CC: Many thanks.
The EATCS Columns
THE ALGORITHMIC GAME THEORY COLUMN

by

MARIOS MAVRONICOLAS

Department of Computer Science, University of Cyprus
75 Kallipoleos St., CY-1678 Nicosia, Cyprus
mavronic@cs.ucy.ac.cy

SOME ALGORITHMIC QUESTIONS ABOUT NASH EQUILIBRIA

Marios Mavronicolas†

†Department of Computer Science, University of Cyprus
P. O. Box 20537, Nicosia CY-1678, Cyprus.

Abstract

One of the most central results in Algorithmic Game Theory is that computing a (mixed) Nash equilibrium for a strategic game is $\mathbb{PPAD}$-complete, even if there are only two players. This result emerged out of a sequence of breakthrough papers in the last few years.

What happens if one is looking for an output Nash equilibrium that satisfies some additional properties? For several natural properties, the problem becomes $\mathbb{NP}$-complete. In this note, we host a detailed exposition for a proof of one such result; the proof is originally due to Conitzer and Sandholm [8].
1 Introduction

One of the most important problems in Algorithmic Game Theory concerns the complexity of Nash equilibria [22, 23] in (finite) strategic games. Computing a Nash equilibrium is an immensely significant problem; this is so due to the central importance of Nash equilibrium as a solution concept in Game Theory and thanks to the increased importance of Game Theory as the normative framework for analyzing strategic interactions in contexts such as the Internet. Indeed, the problem of computing a Nash equilibrium has attracted an enormous amount of interest and attention in the last few years (see, e.g., [1, 2, 3, 5, 6, 7, 8, 9, 11, 12, 13, 14, 16, 17, 20] and [21] for a more general survey on Algorithmic Game Theory).

Through a sequence of breakthrough papers in the last few years, it is now known that the problem of computing a Nash equilibrium is complete for the complexity class $PPAD$ introduced by Papadimitriou almost twenty years ago [24]. The complexity class $PPAD$ is the set of all total functions whose totality is established by invoking the following lemma (known as the Parity Argument) on a graph whose set of vertices is the solution space of the function’s instance:

In a directed graph with one unbalanced vertex (a vertex with out-degree different from its indegree), there is another unbalanced vertex.

The Parity Argument can be specialized to the case where every vertex has both out-degree and in-degree at most one to read:

In a directed graph where every vertex has both out-degree and in-degree at most one, if there is a source (a vertex with in-degree zero), then there is a sink (a vertex with out-degree zero).

Formally, the complexity class $PPAD$ is the set of all total search problems polynomial-time reducible to the computational problem $ExOhOfTheLine$; two boolean circuits are given as input to this problem, representing the predecessor and successor relation in a concisely represented (but exponentially large) graph, along with a source, and we seek either another source or a sink. A total search problem in $PPAD$ is $PPAD$-complete if all problems in $PPAD$ reduce to it.

Originally, Daskalakis, Goldberg and Papadimitriou [12] establish that computing a Nash equilibrium for a four-player strategic game is $PPAD$-complete. Soon thereafter, the result was independently strengthened by Chen and Deng [2] and Daskalakis and Papadimitriou [13] to three-player strategic games. The remaining case of two-player games was finally settled for its $PPAD$-completeness by Chen and Deng [3]. Furthermore, Chen, Deng and Teng [5] showed that relaxing the output’s requirement to approximate Nash equilibria still retains the
problem \(PPAD\)-complete. These \(PPAD\)-completeness results essentially determine the complexity of Nash equilibrium computation in strategic games; they hold assuming we are only interested in computing any Nash equilibrium, even if there are multiple ones. (The multiplicity of Nash equilibria is often considered an "algorithmic fault".) For an exposition of these (and other) complexity results, we refer the reader to the two excellent surveys \([4, 19]\).

It is very natural to ask how sensitive these general complexity results are to particular requirements on the output Nash equilibrium. From the perspective of a Game Theorist, such requirements may sometimes define a more relevant computational problem as it forms an abstraction of the well-studied equilibrium selection problem— for an account on this problem, recall the monumental book of Harsanyi and Selten \([18]\). For example, one may be interested in computing the best Nash equilibrium for a strategic game (for some appropriate definition of "best").

Not surprisingly, the complexity of (mixed) Nash equilibria satisfying additional properties had been considered since the much earlier work of Gilboa and Zemel \([15]\); there, it had been shown that for many interesting properties, the problem becomes \(NP\)-complete. We note that there are already many complexity results on syntactic restrictions of Nash equilibria such as pure and uniform Nash equilibria— see, for example, \([1, 17]\). Despite the significance of the equilibrium selection problem for general (mixed) Nash equilibria, not much has been discovered yet about its complexity.

In a more recent work, Conitzer and Sandholm \([8]\) provide a single reduction from the archetypical SAT problem to simultaneously prove most of the results by Gilboa and Zemel \([15]\). This note is offering a detailed exposition of the reduction due to Conitzer and Sandholm \([8]\). It is hoped that reductions similar in spirit may well-suffice for establishing the \(NP\)-completeness of computing Nash equilibria with natural properties other than those considered in \([8, 15]\); we strongly encourage the interested reader to formulate other natural properties of (mixed) Nash equilibria and attempt to settle their complexity using variants of the reduction due to Conitzer and Sandholm \([8]\). Such reductions are envisioned as useful for the development of the complexity theory of equilibrium selection.

2 Definitions

Throughout, we shall refer to a strategic game \(SG = \langle [n], \{S_i\}_{i \in [n]}, \{U_i\}_{i \in [n]} \rangle\), where (1) \([n]\) is a finite set of players, indexed by natural numbers, and (2) for each player \(i \in [n]\), (2/a) \(S_i\) is the strategy set for \(i\), which is a finite set of strategies, and (2/b) \(U_i\) is a function \(U_i : S \to \mathbb{R}\), called the utility function for \(i\).
2.1 Profiles

A profile \( s \) is a tuple of strategies, one for each player; denote \( S = \times_{k \in [n]} S_k \). A partial profile \( s_{-i} \) is a tuple of \( n - 1 \) strategies, one for each player other than \( i \); it results by eliminating the strategy \( s_i \) of player \( i \in [n] \) from the profile \( s \). For a profile \( s \) and a strategy \( t_i \in S_i \) of player \( i \), denote as \( s_i \circ t_i \) the profile obtained by substituting strategy \( t_i \) for strategy \( s_i \) in the profile \( s \). Denote \( S_{-i} = \times_{k \in [n]\setminus i} S_k \).

Fix any player \( i \in [n] \). Consider a partial profile \( s_{-i} \). We denote the set of player’s \( i \) best strategies when other players choose strategies according to \( s_{-i} \) by \( B_i(s_{-i}) \). Formally, we define the correspondence \( B_i : S_{-i} \to 2^{S_i} \), called the Best Improvement Correspondence (or Best-Response) of player \( i \), by

\[
B_i(s_{-i}) = \{s_i \in S_i \mid U_i(s_{-i} \circ s_i) \geq U_i(s_{-i} \circ t_i) \text{ for all strategies } t_i \in S_i \}. \]

Intuitively, any strategy in \( B_i(s_{-i}) \) is at least as good for player \( i \) as any other strategy assuming that the other players’ strategies are given by \( s_{-i} \). Now, define the correspondence \( B : S \to 2^S \) with \( B(s) = \times_{i \in [n]} B_i(s_{-i}) \) for any profile \( s \in S \).

2.2 Mixed Profiles

A mixed strategy for player \( i \in [n] \) is a probability distribution \( \sigma_i \) on its strategy set \( S_i \); so, a mixed strategy for player \( i \) is a function \( \sigma_i : S_i \to [0, 1] \) such that \( \sum_{s \in S_i} \sigma_i(s) = 1 \). In other words, \( \sigma_i \in \Sigma(S_i) \), the space of all probability distributions on \( S_i \), which is denoted as \( \Sigma_i \). The support (or carrier) of player \( i \) in the mixed strategy \( \sigma_i \), denoted as \( \text{Support}(\sigma_i) \), is the set of strategies \( s \in S_i \) such that \( \sigma_i(s) > 0 \). Note that a pure strategy is the degenerate case of a mixed strategy where the player deterministically chooses a single strategy (with probability 1).

A mixed profile \( \sigma = (\sigma_i)_{i \in [n]} \) is a tuple of mixed strategies, one for each player. Clearly, \( \sigma \in \times_{i \in [n]} \Sigma_i \), the product space of the individual probability distributions on the strategy sets of the players, which will be denoted as \( \Sigma \). Under a mixed profile \( \sigma \), the utility of each player becomes a random variable. So, associated with the mixed profile \( \sigma \) is the expected utility for each player \( i \in [n] \), denoted as \( U_i(\sigma) \) and defined as the expectation of her utility; so,

\[
U_i(\sigma) = \sum_{s \in \Sigma(S_i)} \left( \prod_{k \in [n]} \sigma_k(s_k) \right) \cdot U_i(s). \]

Say that a profile \( s \) is enabled in the mixed profile \( \sigma \) if it is assigned non-zero probability by it.

A partial mixed profile \( \sigma_{-i} \) is a tuple of \( n - 1 \) mixed strategies, one for each player other than \( i \). For a mixed profile \( \sigma \) and a mixed strategy \( \tau_i \) of player \( i \in [n] \),
denote as $\sigma_{-i} \circ \tau_i$ the mixed profile obtained by substituting the mixed strategy $\tau_i$ for the mixed strategy $\sigma_i$ in the mixed profile $\sigma$. Denote $\Sigma_{-i} = \times_{k \in [n]} \Sigma(S_k)$.

A mixed profile is uniform if for each player $i \in [n]$, $\sigma_i$ is a uniform distribution on $\text{Support}(\sigma_i)$; so, a uniform mixed profile is uniquely determined by the supports of the players. A uniform mixed profile is a natural compromise between (pure) profiles, which are degenerately uniform, and mixed profiles: supports need not be singleton sets, but they still uniquely determine the distributions.

We now extend the notion of Best Improvement Correspondence to mixed profiles. For a partial mixed profile $\sigma_{-i} \in \Sigma_{-i}$, we denote as $B_i(\sigma_{-i})$ the player $i$'s best mixed strategies when other players choose mixed strategies according to $\sigma_{-i}$. Formally, we define the correspondence $B_i : \Sigma_{-i} \rightarrow 2^{\Sigma_i}$ with

$$B_i(\sigma_{-i}) = \{ \sigma_i \in \Sigma_i \mid U_i(\sigma_{-i} \circ \sigma_i) \geq U_i(\sigma_{-i} \circ \tau_i) \text{ for all mixed strategies } \tau_i \in \Sigma_i \}.$$ 

Finally, for a mixed profile $\sigma \in \Sigma$, define the correspondence $B : \Sigma \rightarrow 2^\Sigma$ with $B(\sigma) = \times_{i \in [n]} B_i(\sigma_{-i})$.

### 2.3 Pure and Mixed Nash Equilibria

A pure Nash equilibrium, or Nash equilibrium for short, is a profile $s \in S$ such that for each player $i \in [n]$, for each strategy $t_i \in S_i$, $U_i(s) \geq U_i(s_{-i} \circ t_i)$. The following fact is obvious:

**Proposition 1** (Pure Nash Equilibrium is Best Response). A profile $s$ is a pure Nash equilibrium if and only if $s \in B(s)$.

A mixed Nash equilibrium is a mixed profile $\sigma$ such that for each player $i \in [n]$, for each mixed strategy $\tau_i \in \Delta(S_i)$, $U_i(\sigma) \geq U_i(\sigma_{-i} \circ \tau_i)$. For mixed Nash equilibria, an analog of Proposition 1 holds:

**Proposition 2.** A mixed profile $\sigma$ is a mixed Nash equilibrium if and only if for each player $i \in [n]$, for each pure strategy $t_i \in S_i$, $U_i(\sigma) \geq U_i(\sigma_{-i} \circ t_i)$.

### 3 The Complexity Result and its Proof

We start with the definitions for some properties $C$ to be satisfied by some given (mixed) Nash equilibrium $\sigma$ for the strategic game $SG$.

- $C_1$: $\sigma$ is non-unique; that is, there is an additional, distinct Nash equilibrium $\tau$.
- $C_2$: $\sigma$ is non-pure.
- $C_3$: $\sigma$ is non-degenerately uniform.
$C_4$: For a constant $u$, for each player $i \in [n]$, $U_i(\sigma) \geq u$.

$C_5$: For a constant $u$, $\sum_{i\in[n]} U_i(\sigma) \geq u$.

$C_6$: For an integer $r$, $|\text{Support}(\sigma)| \geq r$.

$C_7$: For some player $i \in [n]$, for a strategy $s \in S_i$, $s \in \text{Support}(\sigma_i)$.

We prove:

**Theorem 1.** For each property $C_j$ with $j \in [8]$, the decision problem for a Nash equilibrium with property $C_j$ is $NP$-complete.

**Proof.** Membership in $NP$ is obvious, since on a given profile $\sigma$, it can be verified in polynomial time that (1) $\sigma$ is a Nash equilibrium, and (2) $\sigma$ has the property $C_j$ with $j \in [8]$.

We prove $NP$-hardness by reduction from SATISFIABILITY \cite{10}. Consider a propositional formula $\Phi$, which is a conjunction of clauses $C = \{c_1, \ldots, c_k\}$ over a set of variables $V = \{x_1, \ldots, x_m\}$, as an instance of SATISFIABILITY. To describe the reduction, we need some notation.

Denote $L = \{x_1, \bar{x}_1, \ldots, x_m, \bar{x}_m\}$, the set of literals corresponding to the variables in $V$. We will use lower-case letters $\ell, \ell_1, \ell_2, \ldots, v, v_1, v_2, \ldots$ and $c, c_1, c_2, \ldots$ to denote literals from $L$, variables from $V$ and clauses from $C$, respectively. For a literal $\ell \in L$, denote as $v(\ell)$ the corresponding variable from $V$. For a player $i \in [2]$ and a mixed profile $\sigma$, denote $\sigma_i(L) = \sum_{\ell \in L} \sigma_i(\ell)$. We construct a bimatrix game $\text{SBMG}(\Phi) = \langle [2], \{S_i\}_{i\in[2]}, \{U_i\}_{i\in[2]} \rangle$ from $\Phi$, as follows:

- For each player $i \in [2]$, $S_i = V \cup L \cup C \cup \{f\}$.

- The utility functions of players 1 and 2 are shown in the following table:
For each profile \( s = (s_1, s_2) \) not included in the table, set \( U_1(s_1, s_2) := U_2(s_2, s_1) \). (Note that this ensures that SBMG is a symmetric game.)

Clearly, the construction can be carried out in polynomial time. Here are some initial observations on the constructed game BMG:

1. (a) For each profile \( s \), (a) \( \sum_{i \in [2]} U_i(s) \leq 2 \), (b) with \( \sum_{i \in [2]} U_i(s) = 2 \) if and only if \( s_1 \in L, s_2 \in L \) and \( s_1 \neq s_2 \). This implies that: (c) For each profile \( s \) such that \( s_i \in V \in C \) for some player \( i \in [2] \), \( \sum_{i \in [2]} U_i(s) < 2 \).

2. By Observation (1/a), it follows that (a) for each mixed profile \( \sigma \),

\[
E_{s \sim \sigma} \left( \sum_{i \in [2]} U_i(s) \right) \leq 2.
\]

So, (b) consider a property \( C \) satisfied by at least one profile enabled in \( \sigma \), so that expectations conditioned on \( C \) are well-defined. Then,

\[
E_{s \sim \sigma} \left( \sum_{i \in [2]} U_i(s) \mid s \text{ satisfies } C \right) \leq 2.
\]

We first identify a (pure) Nash equilibrium for the bimatrix game \( \text{SBMG}(\Phi) \).

**Claim 1.** \( f = (f, f) \) is a Nash equilibrium; moreover, it is the only Nash equilibrium in which one player chooses \( f \) with probability 1.
Proof. Note that \( f \in B(f) \); that is, \( f \) is best-response to itself. Proposition 1 implies that \( f \) is a pure Nash equilibrium. Finally, note that when one player chooses \( f \) with probability 1, then \( f \) is the unique best response of the other player. \( \square \)

We continue to prove that the construction provides the claimed reduction. We start by proving:

**Lemma 1.** The game \( SBMG(\Phi) \) has a Nash equilibrium \( \sigma \) other than \( f \) if and only if \( \Phi \) is satisfiable.

**Proof.** We prove the claim through a sequence of simple milestones. We first prove some necessary conditions about mixed Nash equilibria (other than \( f \)) for the bimatrix game \( SBMG(\Phi) \).

**Claim 2.** Consider a mixed Nash equilibrium \( \sigma \) other than \( f \). Then, for each player \( i \in [2] \), \( \text{Support}(\sigma_i) \subseteq L \cup \{f\} \).

**Proof.** Assume, by way of contradiction, that there is a player \( i \in [2] \) such that \( \text{Support}(\sigma_i) \cap (V \cup C) \neq \emptyset \); so, player \( i \) chooses some strategy \( s \in V \cup C \) with probability \( \sigma_i(s) > 0 \).

Since \( \sigma \) is different from \( f \), Claim 1 implies that player \( i \) chooses some strategy \( v \) from \( V \cup C \cup L \) with probability \( \sigma_i(v) > 0 \). So, the profile \( (s,v) \) is enabled in \( \sigma \). Since the profile \( (s,v) \) satisfies the condition \( C^* : s_1 \neq f \) and \( s_2 \neq f \), it follows that \( C^* \neq \emptyset \). Hence, conditional expectations over \( C^* \) are well-defined.

By the construction of the utility functions, we have that \( U_i((s,v)) < 2 \). Since the profile \( (s,v) \) is enabled in \( \sigma \), Observation (2b) implies that

\[
\mathbb{E}_{\sigma_i} \left( \sum_{s \in [2]} U_i(s) \mid s_1 \neq f \text{ and } s_2 \neq f \right) < 2.
\]

By linearity of expectation, it follows that there is a player \( i \in [2] \) such that

\[
\mathbb{E}_{\sigma_i} (U_i(s) \mid s_1 \neq f \text{ and } s_2 \neq f) < 1.
\]

Since the two players are symmetric, take (without loss of generality) that \( i = 1 \).

By the Law of Conditional Expectations and the independence of the choices of the two players, we get that

\[
U_1(\sigma) = \mathbb{E}_{\sigma_2} (U_1(s)) = \mathbb{E}_{\sigma_2} (U_1(s) \mid s_1 \neq f \text{ and } s_2 \neq f) + \mathbb{E}_{\sigma_2} (U_1(s) \mid s_1 \neq f \text{ and } s_2 = f) + \mathbb{E}_{\sigma_2} (U_1(s) \mid s_1 = f \text{ and } s_2 \neq f) + \mathbb{E}_{\sigma_2} (U_1(s) \mid s_1 = f \text{ and } s_2 = f).
\]

42
By the construction of the utility functions, we get that
\[ E_{s \sim \sigma}(U_1(s) | s_1 \neq f \text{ and } s_2 = f) \cdot \sigma_1(f) \cdot \sigma_2(f) \]
\[ = \begin{cases} 0, & \text{if } \sigma_1(f) \cdot \sigma_2(f) = 0 \\ -2 \cdot \sigma_1(f) \cdot \sigma_2(f), & \text{otherwise} \end{cases} \]
\[ \leq 0, \]
\[ E_{s \sim \sigma}(U_1(s) | s_1 = f \text{ and } s_2 \neq f) \cdot \sigma_1(f) \cdot \sigma_2(f) \]
\[ = \begin{cases} 0, & \text{if } \sigma_1(f) \cdot \sigma_2(f) = 0 \\ 1 \cdot \sigma_1(f) \cdot \sigma_2(f), & \text{otherwise} \end{cases} \]
\[ \leq \sigma_1(f) \cdot \sigma_2(f), \]
and
\[ E_{s \sim \sigma}(U_1(s) | s_1 = f \text{ and } s_2 = f) \cdot \sigma_1(f) \cdot \sigma_2(f) \]
\[ = \begin{cases} 0, & \text{if } \sigma_1(f) \cdot \sigma_2(f) = 0 \\ 0 \cdot \sigma_1(f) \cdot \sigma_2(f), & \text{otherwise} \end{cases} \]
\[ = 0. \]
It follows that
\[ E_{s \sim \sigma}(U_1(s)) < \frac{\sigma_1(f) \cdot \sigma_2(f) + \sigma_1(f) \cdot \sigma_2(f)}{\sigma_2(f)}. \]
However, when player 1 switches in \( \sigma \) to strategy \( f \),
\[ U_1(\sigma \cdot \sigma \circ f) = 0 \cdot \sigma_2(f) + 1 \cdot \sigma_2(f) \]
\[ = \sigma_2(f); \]
by Proposition 2, this contradicts the fact that \( \sigma \) is a Nash equilibrium. \( \square \)

We continue to strengthen the necessary condition for mixed Nash equilibria (other than \( f \)) shown in Claim 2:

**Claim 3.** Consider a mixed Nash equilibrium \( \sigma \) other than \( f \). Then, for each player \( i \in [2] \), \( \text{Support}(\sigma_i) \subseteq L \).

**Proof.** Assume, by way of contradiction, that there is a player \( i \in [2] \) such that both \( \sigma_i(f) > 0 \) and \( \sigma_i(L) > 0 \). This implies that both \( \sigma_i(f) < 1 \) and \( \sigma_i(L) < 1 \). Since the two players are symmetric, take, without loss of generality, that \( i = 2 \). Claims 1 and 2 imply that \( \sigma_i(L) > 0 \).
By the Law of Conditional Expectations and the independence of the choices of the two players, Claim 2 yields that

$$U_1(\sigma) = \mathbb{E}_{\theta \sim \sigma}(U_1(s))$$

$$= \mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 \in L \text{ and } s_2 \in L) \cdot \sigma_1(L) \cdot \sigma_2(L) +$$

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 \in L \text{ and } s_2 = f) \cdot \sigma_1(L) \cdot \sigma_2(f) +$$

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 = f \text{ and } s_2 \in L) \cdot \sigma_1(f) \cdot \sigma_2(L) +$$

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 = f \text{ and } s_2 = f) \cdot \sigma_1(f) \cdot \sigma_2(f).$$

By the construction of the utility functions, we get that

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 \in L \text{ and } s_2 \in L) \cdot \sigma_1(L) \cdot \sigma_2(L)$$

$$= \begin{cases} 0, & \text{if } \sigma_1(L) \cdot \sigma_2(L) = 0 \\ \text{(some convex combination of } 1 \text{ and } -2) \cdot \sigma_1(L) \cdot \sigma_2(L), & \text{otherwise} \\ \leq \sigma_1(L) \cdot \sigma_2(L). \end{cases}$$

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 \in L \text{ and } s_2 = f) \cdot \sigma_1(L) \cdot \sigma_2(f)$$

$$= \begin{cases} 0, & \text{if } \sigma_1(L) \cdot \sigma_2(f) = 0 \\ -2 \cdot \sigma_1(L) \cdot \sigma_2(f), & \text{otherwise} \\ = -2 \sigma_1(L) \cdot \sigma_2(f). \end{cases}$$

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 = f \text{ and } s_2 \in L) \cdot \sigma_1(f) \cdot \sigma_2(L)$$

$$= \begin{cases} 0, & \text{if } \sigma_1(f) \cdot \sigma_2(L) = 0 \\ 1 \cdot \sigma_1(f) \cdot \sigma_2(L), & \text{otherwise} \\ \leq \sigma_1(f) \cdot \sigma_2(L), \end{cases}$$

and

$$\mathbb{E}_{\theta \sim \sigma}(U_1(s) | s_1 = f \text{ and } s_2 = f) \cdot \sigma_1(f) \cdot \sigma_2(f)$$

$$= \begin{cases} 0, & \text{if } \sigma_1(f) \cdot \sigma_2(f) = 0 \\ 0 \cdot \sigma_1(f) \cdot \sigma_2(f), & \text{otherwise} \\ = 0. \end{cases}$$

Since both $\sigma_1(L) > 0$ and $\sigma_2(f) > 0$, it follows that

$$U_1(\sigma) \leq \sigma_1(L) \sigma_2(L) - 2 \sigma_1(L) \sigma_2(f) + \sigma_1(f) \sigma_2(L)$$

$$= \sigma_2(L) - 2 \sigma_1(L) \sigma_2(f)$$

$$< \sigma_2(L).$$
The Bulletin of the EATCS

However, when player 1 switches to strategy $f$,

$$U_1(\sigma_\downarrow f) = 0 \cdot \sigma_2(f) + 1 \cdot \sigma_2(L) = \sigma_2(L);$$

by Proposition 2, this contradicts the fact that $\sigma$ is a Nash equilibrium. □

We now continue to prove a further refinement of Claim 3.

**Claim 4.** Consider a Nash equilibrium $\sigma$ other than $f$. Then, for each player $i \in [2]$ and for each literal $\ell \in L$, $\sigma_i(\ell) + \sigma_i(\overline{\ell}) \geq \frac{1}{m}$.

**Proof.** Assume, by way of contradiction, that there is a player $i \in [2]$ and a literal $\ell \in L$ such that $\sigma_i(\ell) + \sigma_i(\overline{\ell}) < \frac{1}{m}$. Since the two players are symmetric, take (without loss of generality) that $i = 2$.

Consider now player 1. By Claim 3 and the construction of the utility functions, it follows that $U_1(\sigma) \leq 1$. However, since $m > 2$,

$$U_1(\sigma_\downarrow 1 \downarrow f) = (2 - n) \cdot (\sigma_2(\ell) + \sigma_2(\overline{\ell})) + 2 \cdot (1 - (\sigma_2(\ell) + \sigma_2(\overline{\ell}))) \quad \text{(by Claim 3)}$$

$$> (2 - m) \cdot \frac{1}{m} + 2 \cdot \left(1 - \frac{1}{m}\right) \quad \text{(by assumption)}$$

$$= 1;$$

by Proposition 2, this contradicts the fact that $\sigma$ is a Nash equilibrium. □

Claim 3 implies that for each player $i \in [2]$, $\sum_{\ell \in L} \left(\sigma_i(\ell) + \sigma_i(\overline{\ell})\right) = 1$. Hence, Claim 4 immediately implies:

**Corollary 1.** Consider a Nash equilibrium $\sigma$ other than $f$. Then, for each player $i \in [2]$ and for each literal $\ell \in L$, $\sigma_i(\ell) + \sigma_i(\overline{\ell}) = \frac{1}{m}$.

We continue with another necessary condition for mixed Nash equilibria (other than $f$).

**Claim 5.** Consider a Nash equilibrium $\sigma$ other than $f$. Then, for each profile $s$ chosen according to $\sigma$, $\mathbb{P}_{\sigma}(s_1 = s_2) = 0$.

**Proof.** Assume, by way of contradiction, that there is a profile $s$ chosen according to $\sigma$ such that $\mathbb{P}_{\sigma}(s_1 = s_2) > 0$. By the Law of Conditional Expectations, we get that

$$U_1(\sigma) = \mathbb{E}_{s \sim \sigma}(U_i(s))$$

$$= \mathbb{E}_{s \sim \sigma}(U_1(s) | s_1 \neq s_2) \cdot \mathbb{P}(s_1 \neq s_2) + \mathbb{E}_{s \sim \sigma}(U_1(s) | s_1 = s_2) \cdot \mathbb{P}(s_1 = s_2).$$
Corollary 2. Consider a Nash equilibrium $\sigma$ other than $\mathbf{f}$. Then, for each pair of a player $i \in [n]$ and a literal $\ell \in \mathbb{L}$, $\sigma_i(\ell) \cdot \sigma_i(\overline{\ell}) = 0$.

Proof. Assume, by way of contradiction, that there is a player $i \in [2]$ and a literal $\ell \in \mathbb{L}$ such that both $\sigma_i(\ell) > 0$ and $\sigma_i(\overline{\ell}) > 0$; so, $\min(\sigma_i(\ell), \sigma_i(\overline{\ell})) > 0$. Note that for a profile $s = (s_1, s_2)$ chosen according to $\sigma$,

$$P_{\sigma}(s_1 = \overline{s_2}) \geq P_{\sigma}(s_1 = \ell) + P_{\sigma}(s_1 = \overline{\ell} \text{ and } s_2 = \ell) \quad \text{(by Claim 3)}$$

$$= \sigma_i(\ell) \cdot \sigma_i(\overline{\ell}) + \sigma_i(\overline{\ell}) \cdot \sigma_i(\ell) \quad \text{(by independence)}$$

$$\geq \min(\sigma_i(\ell), \sigma_i(\overline{\ell})) \cdot \max(\sigma_i(\ell), \sigma_i(\overline{\ell}))$$

$$> 0 \quad \text{(by assumption and Corollary 1)},$$

a contradiction to Claim 5. \qed

An immediate consequence of Corollary 1, Claim 5 and Claim 6 follows:

Corollary 2. Consider a Nash equilibrium $\sigma$ other than $\mathbf{f}$. Then, for each literal $\ell \in \mathbb{L}$, either (1) $\sigma_i(\ell) = \sigma_i(\overline{\ell}) = \frac{1}{2 \mathit{m}}$ (and $\sigma_i(\overline{\ell}) = \sigma_i(\ell) = 0$) or (2) $\sigma_i(\ell) = \sigma_i(\overline{\ell}) = 0$ (and $\sigma_i(\overline{\ell}) = \sigma_i(\ell) = \frac{1}{2 \mathit{m}}$). Furthermore, for each player $i \in [2]$, $U_i(\sigma) = 1$.  

46
Corollary 3. Consider a Nash equilibrium $\sigma$ other than $f$. Then, $\sigma$ induces a truth assignment $(\ell_1, \ldots, \ell_m)$ for $\Phi$.

We are now ready to show that satisfiability of $\Phi$ and existence of a (mixed) Nash equilibrium other than $f$ for the symmetric bimatrix game $\text{SBMG}(\Phi)$ are equivalent. In the one direction, we establish:

Claim 7. If $\Phi$ is not satisfiable, then there is no Nash equilibrium other than $f$.

Proof. Assume, by way of contradiction, that there is a Nash equilibrium $\sigma$ other than $f$. By Corollary 2, for each player $i \in [2]$, $U_i(\sigma) = 1$. By Corollary 3, $\sigma$ induces a truth assignment $(\ell_1, \ldots, \ell_m)$ for $\Phi$. Since $\Phi$ is not satisfiable, there is a clause $c$ such that for each index $j \in [m]$, $\ell_j \notin c$. Hence, by the construction of the utility functions, for each player $i \in [2]$, $U_i(\sigma \odot c) = 2$; by Proposition 2, this contradicts the fact that $\sigma$ is a Nash equilibrium. □

We continue with the opposite direction.

Claim 8. If $\Phi$ is satisfiable, then there is a Nash equilibrium other than $f$.

Proof. Consider a truth assignment $(\ell_1, \ldots, \ell_m)$ that satisfies $\Phi$. Clearly, it induces a mixed profile $\sigma$ such that for each player $i \in [2]$ and for each index $j \in [m]$, $\sigma_i(\ell_j) = \frac{1}{m}$. Note that for each player $i \in [2]$, $U_i(\sigma) = 1$. We will prove that $\sigma$ is a Nash equilibrium.

Consider an arbitrary player $i \in [2]$. There are four ways for player $i$ to switch in $\sigma$ to a strategy $s \in S_i$; we shall examine them all.

1. $s = v \in V$: Clearly, there is a literal $\ell_j$ in the satisfying assignment which corresponds to $v$; that is, $v = v(\ell_j)$. So, by the construction of the utility functions and the mixed profile $\sigma$,
   \[
   U_i(\sigma \odot v) = (2 - m) \cdot \sum_{\ell \in L_i(v) \neq v} \sigma(\ell) + 2 \cdot \sum_{\ell \in L_i(v)} \sigma(\ell) \\
   = (2 - m) \cdot \frac{1}{m} + 2 \cdot \frac{m - 1}{m} \\
   = 1.
   \]

2. $s = c \in C$: Since $(\ell_1, \ldots, \ell_m)$ is a satisfying assignment for $\Phi$, there is some index $j \in [m]$ such that $\ell_j \in c$. Note that by the construction of the mixed profile $\sigma$, $P_{\sigma}(s_\ell \in c) \geq \frac{1}{m}$ (since $\ell_j \in c$ and $\sigma(\ell_j) = \frac{1}{m}$). This implies that
\[ P_{\sigma}(s_i \not\in c) \leq 1 - \frac{1}{m}. \] So, by the construction of the utility functions, we get that

\[ U_i(\sigma_{-i} \circ c) = (2 - m) \cdot P_{\sigma}(s_i \in c) + 2 \cdot P_{\sigma}(s_i \not\in c) \leq (2 - m) \cdot \frac{1}{m} + 2 \cdot \left(1 - \frac{1}{m}\right) \quad \text{(since } m \geq 2) \]

(3). \( s = \ell_k \in L \): In this case, \( P_{\sigma}(s_i \not\in c) = P_{\sigma}(s_i = \ell_k) \) and \( P_{\sigma}(s_i = c) = P_{\sigma}(s_i = \bar{\ell}_k) \). So, by the construction of the utility functions, we get that

\[ U_i(\sigma_{-i} \circ \ell_k) = 1 \cdot P_{\sigma}(s_i \not\in \ell_k) + (-2) \cdot P_{\sigma}(s_i = \ell_k). \]

There are now two subcases.

(3/a). \( \ell_k = \ell_j \), for a literal \( \ell_j \) in the satisfying assignment: Note that \( P_{\sigma}(s_i \not\in \ell_k) = P_{\sigma}(s_i \not\in \ell_j) = 1 \) and \( P_{\sigma}(s_i = \ell_k) = P_{\sigma}(s_i = \ell_j) = 0 \). It follows that \( U_i(\sigma_{-i} \circ \ell_k) = 1 \).

(3/b). \( \ell_k = \ell_j \), where \( \ell_j \) is a literal in the satisfying assignment: So, \( P_{\sigma}(s_i \not\in \ell_k) = P_{\sigma}(s_i \not\in \ell_j) = 1 - \frac{1}{m} \) and \( P_{\sigma}(s_i = \ell_k) = P_{\sigma}(s_i = \ell_j) = \frac{1}{m} \). It follows that

\[ U_i(\sigma_{-i} \circ \ell_k) = 1 \cdot \left(1 - \frac{1}{m}\right) + (-2) \cdot \frac{1}{m} = 1 - \frac{3}{m}. \]

(4). \( s = f \): By the construction of the utility functions, \( U_i(\sigma_{-i} \circ f) = 1 \).

Since for all strategies \( s \in S_i \), \( U_i(\sigma) \geq U_i(\sigma_{-i} \circ s) \), Proposition 2 implies that \( \sigma \) is a mixed Nash equilibrium. □

The claim follows now from Claims 7 and 8. □

Note that in the case where there is a Nash equilibrium \( \sigma \) other than \( f \), Claim 1 implies Constraint \( C_1 \); Corollary 2 implies Constraints \( C_2 \) and \( C_3 \); Corollary 2 implies Constraints \( C_4, C_5 \) and \( C_6 \) (with \( u = 1, u = 2 \) and \( r = 2 \), respectively); Corollary 3 implies Constraints \( C_7 \) and \( C_8 \) (with \( s \in L \) and \( s = f \), respectively). Hence, Lemma 1 implies the claim. □
4 Epilogue and Open Problems

We have presented a detailed exposition of a proof due to Conitzer and Sandholm [8] that computing a Nash equilibrium with certain natural properties is \(NP\)-complete. In an extended version of their work [9], Conitzer and Sandholm present an improved version of their reduction from [8] to establish that several natural optimization problems about Nash equilibria (such as maximizing the social welfare of a Nash equilibrium) are inapproximable: unless \(P = NP\), there is no polynomial time algorithm returning a Nash equilibrium that is close to obtaining the optimal value for the social welfare. This implies an improvement over the earlier results of Gilboa and Zemel [15]. More important, we believe that the single reduction employed by Conitzer and Sandholm [8, 9] will be useful for settling the complexity in other instances of computation of mixed Nash equilibria with certain properties.

There is still a lot to be understood regarding the complexity of Nash equilibria with additional properties. Most important, we are missing a dichotomy between easy and hard properties. In particular, there are no results yet on the complexity of the decision problem for an irrational Nash equilibrium: given a strategic game (with three or more players), is there an irrational Nash equilibrium? This important problem has been communicated to us by Elias Koutsoupias during the recent SAGT 2009 conference.

References


The Bulletin of the EATCS


More than 20 years ago, Smolensky proved exponential lower bounds for the size of constant size circuits over unbounded AND, OR and MOD$_m$ gates for a prime power $m$. Since then many researchers have tried to extend this result to modular gates MOD$_m$ with $m$ having at least two distinct prime factors. Despite considerable effort, the problem remains elusive. Arkadev Chattopadhyay surveys in this column the progress made on a recent approach, based on a simpler model, for attacking the problem.

Multilinear Polynomials Modulo Composites

Arkadev Chattopadhyay*

Abstract

Understanding the power of constant-depth circuits that are allowed to use MOD$_m$ gates, where $m$ is an arbitrary but fixed positive integer, is a fundamental and inviting problem in theoretical computer science. Despite intensive efforts for more than twenty five years, this problem remains wide open.

*University of Toronto, Department of Computer Science, arkadev@cs.toronto.edu. The author is partially supported by a postdoctoral fellowship of the Natural Sciences and Engineering Research Council (NSERC) of Canada.
In this column, we focus our attention on the related, but much simpler, model of computing a boolean function by multilinear polynomials over the ring \( \mathbb{Z}_m \), when \( m \) is a composite number. As widely known, it is essential to understand this model in order to make progress with constant-depth circuits with MOD gates. We survey some recent results in this natural model that yield superpolynomial lower bounds on the size of some restricted circuits with MOD\(_m\) gates. The ingredients that get used in these results are perhaps more interesting. Some natural next steps emerge from these results that are also of independent mathematical interest. It is hoped that progress along these lines is feasible and would provide further insight into the general problem.

1 Introduction

Eric Allender [2] starts his recent survey of the state-of-affairs in proving lower bounds on circuit size by noting that his earlier survey [1] remains depressingly current. While it is true that we cannot pitifully find a function in EXP that cannot be computed by linear size depth-three circuits comprising only MOD\(_6\) gates, the time honored George Polya principle of considering simpler problems seems to again provide ways to making meaningful progress. In this article, we further argue that such considerations have raised (and sometimes solved) natural and appealing problems that can be stated in pure mathematical terms. This holds the promise that tools from mainstream mathematics can be further exploited in the context of understanding the computational power of mod counting.

A series of interesting works on constant-depth circuits have recently appeared. Here, we just focus on the ones that are motivated by circuits having MOD\(_m\) gates, where \( m \) is an arbitrary number. Note that MOD\(_m\) is a boolean function that is defined below:

**Definition 1.1.** Let \( A \subseteq \mathbb{Z}_m \) be some accepting set. Then, for each \( x \in \{0, 1\}^n \), \( \text{MOD}_A^m(x) = 1 \) if \( \sum_{i=1}^n x_i \equiv a \text{ (mod } m) \) for some \( a \in A \), otherwise the function outputs 0.

By default, the accepting set \( A \) is \( \mathbb{Z}_m - \{0\} \) and in this case it is dropped from the superscript. The class of functions computed by polynomial size and constant-depth circuits\(^1\) of unbounded fan-in having AND, OR and MOD\(_m\) gates is called ACC\(^0[\text{mod } m] \). The union of these classes over all fixed positive integer \( m \) is defined to be the complexity class ACC\(^0\). As is the convention, we overload these terms to also mean the underlying circuits with no restrictions on size. Understanding the

\(^1\)The input layer of all boolean circuits considered in this article have access to each variable and its negation, in addition to boolean constants 0 and 1.
computational limitations of $\text{ACC}^0$ is a major goal of computational complexity that remains unfulfilled.

Smolensky [40] in the late eighties, building upon the elegant work of Razborov [39], showed that $\text{ACC}^0[p]$ circuits require exponential size to compute $\text{MOD}_q$, if $p$ is a prime and $k$ any fixed positive number and $q$ has a prime factor different from $p$. A simple exercise then shows that $\text{MAJORITY}$ cannot be computed in sub-exponential size by such circuits. Indeed, one can very well imagine the excitement this generated back at the time. Smolensky made the following very tempting conjecture:

**Conjecture 1.2** (Smolensky). For any fixed positive integer $m$, $\text{ACC}^0[m]$ circuits needs exponential size to compute $\text{MOD}_q$, if $m,q$ are co-prime numbers.

At the moment, we seem to be far from proving (or disproving) Smolensky’s conjecture. One may be inclined to think that circuits that are restricted to have only $\text{MOD}_m$ gates (and constant depth, denoted by $\text{CC}^0[m]$) are easier to deal with? Such a thought is especially appealing, given the following fact about prime moduli: for any prime $p$, circuits of constant-depth having only $\text{MOD}_p$ gates cannot compute all functions. In particular, they cannot compute a high degree function (over $\mathbb{Z}_p$) like $\text{OR}$, $\text{AND}$ and $\text{MOD}_q$, no matter how much size is allowed. Indeed, this is a very strong computational limitation and follows surprisingly easily from the fact that $\mathbb{Z}_p^*$ is a group. In contrast, depth-2 such circuits having only $\text{MOD}_m$ gates can compute everything:

**Fact 1.3** (Folklore, (see [7])). Let $m$ be any number that has at least two distinct prime factors. Then, every $n$-variate boolean function $f$ can be computed by a depth-two circuit of size $2^n$ having only $\text{MOD}_m$ gates.

In a recent work, Hansen and Koucký [32] observe that one can combine Fact 1.3 with the Razborov-Smolensky idea of approximating AND/OR gates by low degree polynomials over any finite field to yield the following interesting result:

**Theorem 1.4** (implied in [32]). Every quasipolynomial size circuit $C$ comprising $\text{AND}$, $\text{OR}$ and $\text{MOD}_m$ gates of depth $d$ can be approximated very well by a quasipolynomial size circuit $C'$ of depth $O(d)$ comprising only $\text{MOD}_m$ gates, i.e. $\Pr_x [C'(x) \neq C(x)] \leq 1/q\text{poly}(n)$.

This hints that proving Smolensky’s conjecture for circuits with only $\text{MOD}_m$ gates may be as hard as proving the general case. Indeed, Smolensky obtains his result for $\text{ACC}^0[p]$ by showing the stronger result that they cannot even approximate well $\text{MOD}_q$. This strengthening is crucial to his argument. Theorem 1.4,
The Bulletin of the EATCS

on the other hand, shows that such a strengthened result (against \( \text{MOD}_q \)) for the special case of \( \text{CC}^0[m] \) circuits is sufficient to deal with general \( \text{ACC}^0[m] \) circuits.

Nevertheless, the intuition that \( \text{CC}^0[m] \) circuits are weaker and hence easier to deal with, may not be entirely lost. For a boolean function \( f \), let the support set of \( f \), denoted by \( \text{supp}(f) \), be the set of points in the cube where \( f \) evaluates to 1. The support set of a \( \text{MOD}_m \) gate is large in size and is in some sense uniformly spread out in the cube. Can the following be true?

**Conjecture 1.5** (Large Support Set\(^2\), appears in [17]). There exists a function \( h : \mathbb{N} \rightarrow \mathbb{N} \), such that any non-constant function computed by a \( \text{CC}^0[m] \) circuit of size \( s \) and depth \( d \) has a support set of size at least \( 2^{\Omega(\log s h(d))} \).

Indeed, the Large Support Set Conjecture is true in a very strong sense when \( m \) is a prime \( p \) (or a prime power). The argument goes through polynomials over \( \mathbb{Z}_p \) and we point this out in Section 2 after the statement of Conjecture 2.7.

Note that in particular, the Large Support Set Conjecture implies that AND (or OR) cannot be computed in small size by \( \text{CC}^0[m] \) circuits. This is dual to the celebrated result that \( \text{MOD}_m \) cannot be computed easily by \( \text{AC}^0 \) circuits. Such a possibility has long been conjectured by McKenzie, Péladau and Thérien [33]. The relative hardness of Smolensky’s Conjecture and the Large Support Set Conjecture is not clear. Unfortunately, both seem out of hand for the moment.

In this article, we focus our attention on a very basic and natural model of computation: that of multilinear polynomials over the ring \( \mathbb{Z}_m \). It is well known that understanding this model is absolutely necessary before significant progress on above conjectures can be made. Indeed, Razborov [39] and Smolensky [40] introduced computation by polynomials over the prime field \( \mathbb{Z}_p \) as a key ingredient in their arguments for lower bounds on constant-depth circuits\(^3\). Unfortunately, as reviewed in the next section, understanding polynomials over \( \mathbb{Z}_p \) already presents significant difficulties and several questions remain open. Our study of polynomials is motivated by Smolensky’s Conjecture and the Large Support Set Conjecture. In particular, we aim to prove sort of their analogs in the polynomial world.

Before we proceed further, it is important to point out that polynomials over reals are also a very natural and interesting model of computing boolean functions. It is indeed extremely relevant for understanding constant-depth circuits. For lack of space and the sake of focus, we leave out this topic here. The interested reader can consult the excellent survey by Beigel [8] to get pointers to the older literature and more recent works like [37]. Beigel [8] also discusses polynomials over finite rings, but the survey is somewhat dated and broader in scope than ours. Here we

---

\(^2\)In the thesis [17], where this conjecture originates, it is called the Small Support Set Conjecture referring to the fact that functions with a small support set are difficult for \( \text{CC}^0[m] \) circuits.

\(^3\)In fact their methods also work over the ring \( \mathbb{Z}_{p^k} \), where \( p \) is a prime and \( k \) is fixed positive integer.
survey some recent (and some not so recent) works on polynomials over \( \mathbb{Z}_m \) and point out some of the challenges that lie ahead.

## 2 Computation by Polynomials

An interesting thing to observe is that every function \( f : \{0, 1\}^n \to \mathbb{Z}_m \) is expressible as a multilinear polynomial over \( \mathbb{Z}_m \). To see this one merely has to verify that each so called delta function is expressible by such a polynomial. More precisely, for each \( w \in \{0, 1\}^n \), define the delta function \( \delta_w : \{0, 1\}^n \to \mathbb{Z}_m \) as \( \delta_w(x) = 1 \) if \( w = x \) and otherwise \( \delta_w(x) = 0 \). Consider the set of functions \( \Delta = \{ \delta_w \mid w \in \{0, 1\}^n \} \).

It is easy to see that every function \( f \) can be uniquely expressed as a \( \mathbb{Z}_m \)-linear combination of such functions. On the other hand,

\[
\delta_w(x) = \left( \prod_{i: w_i = 1} x_i \right) \left( \prod_{i: w_i = 0} (1 - x_i) \right).
\]

The simple identity above implies that every \( \mathbb{Z}_m \)-valued function over the boolean cube is expressible as a multilinear polynomial over the ring \( \mathbb{Z}_m \). Indeed, a simple counting argument shows that the polynomial corresponding to each such function is unique. This enables us to view each boolean function as an algebraic object. Natural measures of the complexity of this object are its degree and the number of monomials appearing in it. Formalizing things, let \( \deg_m(f) \) denote the degree of the polynomial representing the boolean function \( f \) over \( \mathbb{Z}_m \). In our discussion, polylogarithmic degree will be considered small and \( n^{\Omega(1)} \) degree will be high. Exhibiting a function of high degree is not hard. For example,

\[
\text{AND}(x) = x_1 x_2 \cdots x_n
\]

\[
\text{OR}(x) = 1 - \prod_{i=1}^n (1 - x_i)
\]

showing that \( \deg_m(\text{OR}) = \deg_m(\text{AND}) = n \). On the other hand, demanding a polynomial \( P \) to satisfy \( P(x) = f(x) \) for each point \( x \) in the cube seems too restrictive. A more natural definition, at least from a computational perspective, was introduced in the very interesting work of Barrington, Beigel and Rudich [6]. Let \( A \subseteq \mathbb{Z}_m \) be an accepting set. Then \( P \) represents \( f \) w.r.t \( A \) if it satisfies the following property for each \( x \) in the boolean cube: \( P(x) \in A \mod m \) iff \( f(x) = 1 \). The first thing to note about this model, is that there is not a unique polynomial computing \( f \) w.r.t some fixed accepting set \( A \). A straightforward counting argument shows that there are exactly \( |A|^{\text{supp}(f)}(m - |A|)^{2n - |\text{supp}(f)|} \) polynomials representing \( f \) w.r.t. the accepting set \( A \).
Definition 2.1. Let \( \deg_{A_m}(f) \) denote the minimal degree among degrees of polynomials representing \( f \) w.r.t accepting set \( A \). The generalized degree of \( f \), denoted by \( \text{gen-deg}_{m}(f) \), is then defined to be the degree of \( f \) w.r.t. to the best accepting set, i.e.

\[
\text{gen-deg}_{m}(f) = \min\{ \deg_{A_m}(f) : A \subseteq \mathbb{Z}_m \}.
\]

While it is immediate that \( \text{gen-deg}_{m}(f) \leq \deg_{m}(f) \) for every \( f \), it is a central question in the theory of polynomial representations to determine how much degree savings can generalized representation achieve over exact representation in the ring \( \mathbb{Z}_m \). For general \( m \), it seems fairly non-trivial to get good estimates of \( \deg_{A_m}(f) \) for even a simple \( f \) like OR and AND. However, when \( m \) is a prime \( p \) (or a prime power), tight bounds can be obtained in a simple and elegant fashion. The fact that \( \mathbb{Z}_p^* \) is a group turns out to be very useful:

Fact 2.2 (Fermat’s Gift). Let \( p \) be any prime. For every \( x \not\equiv 0 \pmod{p} \), \( x^{p-1} \equiv 1 \pmod{p} \).

This gift is great for booleanization. Let \( P \) be any polynomial and \( A \) any accepting set. Let \( Q(x) = \sum_{a \in A} 1 - (P(x) - a)^{p-1} \). Using Fermat’s Gift, it is easy to verify that \( Q(x) \) is \( 0/1 \) valued modulo \( p \) and \( P(x) \in A \pmod{p} \) iff \( Q(x) \equiv 1 \pmod{p} \). Thus, if \( P \) represented \( f \) w.r.t \( A \), then \( Q \) is the unique polynomial corresponding to \( f \). Noting that degree of \( Q \) is larger than \( P \) by a factor of at most \( p-1 \), one gets linear lower bounds on the degree of \( P \) if the function represented is a hard function like OR and AND (recall equation (2.1)):

Fact 2.3. For any prime \( p \), \( \text{gen-deg}_p(f) \geq \deg_p(f)/(p - 1) \). In particular,

\[
\text{gen-deg}_p(\text{OR}), \text{gen-deg}_p(\text{AND}) \geq \frac{n}{p - 1}.
\]

Unfortunately, when \( m \) contains two distinct prime factors, Fermat’s gift stops working. One could hope that given any accepting set \( A \subseteq \mathbb{Z}_m \), there is some univariate \( 0/1 \) valued polynomial \( R \) over \( \mathbb{Z}_m \) corresponding to the characteristic function of the set \( A \). Indeed, Fermat’s gift yields such a polynomial when \( m \) is prime. Having some such \( R \) would be enough for proving lower bounds on the generalized degree of \( f \) over \( \mathbb{Z}_m \). This hope gets killed for the following reason: let \( m = p_1p_2 \) be a product of two distinct primes. Recall, via chinese remaindering, the map \( a \mapsto ((a \mod p_1), (a \mod p_2)) \) forms a bijection between \( \mathbb{Z}_m \) and \( \mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2} \). Thus, 0 and 1 in \( \mathbb{Z}_m \) correspond to tuples \((0, 0) \) and \((1, 1) \) in \( \mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2} \).

Fact 2.4. Let \( m = p_1p_2 \) be a product of two distinct primes. Then the characteristic function of the set \( A = \{1\} \) (and the set \( A = \{0\} \)) has no (univariate) polynomial representation over \( \mathbb{Z}_m \).
Proof. Assume for the sake of contradiction that $R$ is such a polynomial. Applying the Chinese Remaindering Theorem, $R$ gives rise to two polynomials, $R_{p_1}$ over $\mathbb{Z}_{p_1}$ and $R_{p_2}$ over $\mathbb{Z}_{p_2}$, with the property that $R(x) \equiv (R_{p_1}(x \mod p_1), R_{p_2}(x \mod p_2))$. Now $R(0) \equiv 0 \pmod{m}$. Hence, $R_{p_1}(0) \equiv 0 \pmod{p_1}$. Similarly, $R_{p_2}(0) \equiv 0 \pmod{p_2}$. Observing that $R(1) \equiv 1 \pmod{m}$ and applying a similar argument yields the following: $R_{p_1}(1) \equiv 1 \pmod{p_1}$ and $R_{p_2}(1) \equiv 1 \pmod{p_2}$. Thus, combining things back via chinese remaindering, $R((0, 1)) \equiv (0, 1) \pmod{m}$ and $R((1, 0)) \equiv (1, 0) \pmod{m}$. However, as $R$ is the exact representation of the characteristic function of $A = \{1\}$, $R((0, 1)) \equiv R((1, 0)) \equiv (0, 0) \pmod{m}$, leading us to the required contradiction. $\square$

Fact 2.4 has turned out to be somewhat of a serious blow to proving lower bounds on the composite degree of boolean functions. To some extent, this is explained by a surprising upper bound discovered by Barrington, Beigel and Rudich [6].

**Theorem 2.5** (Barrington, Beigel and Rudich). Let $m$ have $t$ distinct prime factors. Let $A = \{1\}$ and $A' = \mathbb{Z}_m - \{0\}$. Then, $\text{deg}_m^A(\text{AND}) = O(n^{1/t})$ and $\text{deg}_m^A(\text{OR}) = O(n^{1/t})$.

The above theorem shows that composite moduli can obtain non-trivial computational advantage over their primal counterparts when the accepting set is carefully chosen. Even more surprisingly, the above theorem has been exploited in explicit constructions in combinatorics [29, 22] and very recently in obtaining efficient locally decodable codes by Efremenko [20].

Tardos and Barrington [42] obtained the following lower bound on the generalized degree of the OR function.

**Theorem 2.6** ([42]). Let $m$ have $t \geq 2$ distinct prime factors, and let $q$ be the smallest maximal prime power divisor of $m$. Then, $\text{gen-deg}_m(\text{OR})$ is at least $\left(\frac{1}{q-1} - o(1)\right) \log n^\frac{1}{t-1}$.

The above lower and upper bounds on the degree for OR and AND has not been improved in more than ten years and it is an important challenge to narrow down the gap between them. On the other hand, we speculate the following:

**Conjecture 2.7.** Let $P$ be a multilinear polynomial of degree $d$ over $\mathbb{Z}_m$. Let $a \in \mathbb{Z}_m$ be such that there exists an $x_0 \in \{0, 1\}^n$ with $P(x_0) \equiv a \pmod{m}$. Then the number of points in the cube at which $P$ evaluates to $a$ is at least $2^{n - O(\sqrt{d})}$, where $m = p_1 \cdots p_t$ and each $p_i$ is a distinct prime.

It is simple to verify that this conjecture implies that for such square-free $m$, $\text{gen-deg}_m(\text{OR}), \text{gen-deg}_m(\text{AND}) = n^{\Omega(1/t)}$. The conjecture above admits a natural
The Bulletin of the EATCS

modification to composites with repeated prime factors. We do not state that formally to keep the discussion simple and focused on the essential problem that lies ahead. Before we end this section, it is worth mentioning that the above conjecture is known to be true for prime moduli (see for example [5]). Using Ramsey Theory, Péladeau and Thérien [38] prove a result that easily implies this conjecture for arbitrary \( m \) as long as the degree \( d \) is a constant.

2.1 Computing \( \text{MOD}_q \)

The advantage of composites over primes is not limited to computing AND and OR. Among other things, Bhatnagar et.al.[19] showed that one can compute the \( \text{THRESHOLD}_k \) function by polynomials of degree \( O(n^{1/r + \epsilon}) \) over \( \mathbb{Z}_m \), if \( m \) has \( t \) distinct prime factors and \( k \) is a constant. This is a generalization of the upper bound due to Barrington et.al. as OR is just \( \text{THRESHOLD}_1 \). Bhatnagar et.al. wondered if interesting degree upper bounds could be proved for the simple function \( \text{MOD}_q \). Hansen [31], disproving a conjecture of Bhatnagar et.al. [19], showed the following:

**Theorem 2.8** (Hansen). Assume \( m = p_1 \cdots p_t \) and \( q \) are co-prime satisfying the following condition: there exists positive integers \( b_1, \ldots, b_t \) such that \( \sum_{i=1}^t \frac{1}{b_i} < 1 \) and \( p_i \geq qb_i \) for all \( i \). Then \( \text{gen-deg}_m(\text{MOD}_q) = O(n^{1/t}) \).

Tardos and Barrington’s [42] technique can be adapted (see for example [13]) to prove an \( \Omega((\log n)^{1/(r-1)}) \) lower bound on \( \text{gen-deg}_m(\text{MOD}_q) \). Such bounds degrade with the number of distinct prime factors of \( m \). In a breakthrough work, Bourgain [11] proved an \( \Omega(\log n) \) lower bound on the generalized \( \text{MOD}_m \)-degree of \( \text{MOD}_q \). Bourgain’s method is interesting due to several reasons. First, it proves something stronger, showing that the correlation between the boolean function computed by a sub-logarithmic degree polynomial over \( \mathbb{Z}_m \), w.r.t. an accepting set, and \( \text{MOD}_q \) is exponentially small. Such a correlation bound was not known even for polynomials modulo primes, a model which one typically assumes we understand well. The result, very significantly, improves upon a long line of work (see, for example, [21, 12, 24, 4, 26]). Second, Bourgain’s method boils down to estimating certain exponential sums. This is an elementary but powerful technique that has spawned more recent progress [15, 30, 18]. Due to its importance, we include a proof of Bourgain’s result. Our treatment follows that of Chattopadhyay [14, 16], that is very close to the method of [11, 27] but is slightly simpler and sharper.

**Definition 2.9.** For any \( b \in \{0, \ldots, q-1\} \), define the \( b \)th \( \text{MOD}_q \)-residue class of...
\[ M_q(b) = \{ x = (x_1, \ldots, x_n) \in \{0,1\}^n \mid \sum_{i=1}^{n} x_i = b \pmod{q} \} \]  

**Definition 2.10.** For any polynomial \( P \) over \( \mathbb{Z}_m \) and \( a \in \mathbb{Z}_m \), let \( P^{-1}(a) \) define the set of points in \( \{0,1\}^n \) where \( P \) evaluates to \( a \).

An intuition about a random and uniform set is that each of the \( M_q(b) \) residue classes are equally represented in such a set. Bourgain’s result essentially shows that if \( P \) has low degree, then \( P^{-1}(a) \) appears pseudorandom \(^4\) to the \( \text{MOD}_q \) function. In other words, either each of the \( \text{MOD}_q \) residue classes are almost equally represented in \( P^{-1}(a) \) or the set is a very small fraction of the cube.

**Lemma 2.11** (Bourgain’s Uniformity Lemma). For all positive co-prime integers \( m, q \), there exists a positive constant \( \gamma = \gamma(q) < 1 \) such that for every polynomial \( P \) of degree \( d \) over \( \mathbb{Z}_m \) and every \( a \in \mathbb{Z}_m \), the following holds:

\[ \left| \Pr \left[ x \in (P^{-1}(a) \cap M_q(b)) \right] - \frac{1}{q} \Pr \left[ x \in P^{-1}(a) \right] \right| \leq \exp \left( - \frac{\gamma m}{(m^2 m^{-1})^d} \right) \]  

Before we start the proof, let us recall an elementary fact about the primitive roots of unity that we make repeated use of henceforth. Let \( e_m(y) \) denote the primitive \( m \)-th root of unity raised to the \( y \)th power, i.e. \( \exp \left( \frac{2\pi i y}{m} \right) \), where \( j \) is the complex square root of \(-1\). Then,

**Fact 2.12.** If \( y = 0 \) then, \( \frac{1}{m} \sum_{\alpha=0}^{m-1} e_m(\alpha y) \) is 1 and the expression is 0 otherwise.

Armed with this basic fact, we prove the Uniformity Lemma below:

**Proof of Uniformity Lemma.** We write \( \Pr \left[ x \in (P^{-1}(a) \cap M_q(b)) \right] \) as an exponential sum. Thus,

\[ \Pr \left[ x \in (P^{-1}(a) \cap M_q(b)) \right] = \mathbb{E}_{x \in \{0,1\}^n} \left[ \left( \frac{1}{m} \sum_{\alpha=0}^{m-1} e_m(\alpha (P(x) - a)) \right) \left( \frac{1}{q} \sum_{\beta=0}^{q-1} e_q(\beta(x_1 + \cdots + x_n - b)) \right) \right] \]  

The method employed by Bourgain to prove this result is closely related to method employed commonly in communication complexity to estimate the discrepancy of a function. Indeed, the quantity in the LHS of (2.3) is closely related to the discrepancy of \( \text{MOD}_q \) function w.r.t. polynomial mappings modulo \( m \). The interested reader can find more details on this point of view in [16, 17].
The Bulletin of the EATCS

Expanding the sum inside the second multiplicand and treating the case of \( \beta = 0 \) separately, one gets

\[
(2.4) = \frac{1}{q} \mathbb{E} \left[ \frac{1}{m} \sum_{x=0}^{m-1} e_m(\alpha(P(x) - a)) \right] + \frac{1}{mq} \sum_{\alpha \in [m], \beta \in [q] - \{0\}} S^{m\beta}(\alpha, \beta, P) e_m(-a\alpha)e_q(-b\beta) \tag{2.5}
\]

where,

\[
S^{m\beta}(\alpha, \beta, P) = \mathbb{E}_{x \in \{0,1\}^n} \left[ e_m(aP(x)) \cdot e_q(\beta(x_1 + \cdots + x_n)) \right] \tag{2.6}
\]

Observing that the first term in (2.5) is simply \( \frac{1}{q} \Pr \{ x \in P^{-1}(a) \} \) and \( |e_m(-a\alpha)| = |e_q(-b\beta)| = 1 \), we get :

\[
\left| \Pr \{ x \in (P^{-1}(a) \cap M_q(b)) \} - \frac{1}{q} \Pr \{ x \in P^{-1}(a) \} \right| \leq \frac{1}{mq} \sum_{\alpha \in [m], \beta \in [q] - \{0\}} |S^{m\beta}(\alpha, \beta, P)| \tag{2.7}
\]

The Uniformity Lemma 2.11 gets proved by the bound on \( |S^{m\beta}(\alpha, \beta, P)| \) provided below. The bound below is the main technical contribution of Bourgain. □

**Lemma 2.13.** For each pair of co-prime integers \( m, q > 1 \) there exists a constant \( \gamma = \gamma(q) \) such that for every polynomial \( P \) of degree \( d > 0 \) in \( \mathbb{Z}_m \) and numbers \( \alpha \in [m], \beta \in [q] - \{0\} \), the following holds :

\[
|S^{m\beta}(\alpha, \beta, P)| \leq \exp \left( -\frac{\gamma n}{(m^{2m-1})^d} \right) \tag{2.8}
\]

Before we begin our formal calculations, we note that a slightly weaker estimate of \( |S^{m\beta}(\alpha, \beta, P)| \) was first obtained by Bourgain [11] and later generalized by Green et al [27]. The case when \( P \) is a linear polynomial was essentially dealt with in [12] and forms our base case\(^5\) just as in [11, 27].

In order to explain the intuition behind our calculations, we develop some definitions and notations. Let \( f : \{0,1\}^n \to \mathbb{Z}_m \) be any function. Consider any set \( I \subseteq [n] \). Note that each binary vector \( v \) of length \( |I| \) can be thought of as a partial assignment to the input variables of \( f \) by assigning \( v \) to the variables in \( I \) in a natural way. Let \( f^{(v)} \) be the subfunction of \( f \) on variables not indexed in \( I \)

\(^5\)We revisit this base case later in Section 3.4.1.
induced by the partial assignment \( v \) to variables indexed in \( I \). For any sequence \( Y = \{y_1, \ldots, y_t\} \) having \( t \) boolean vectors from \([0, 1]^n\), let \( f_Y \) be the function defined by \( f_Y(x) = f(x) + \sum_{i=1}^t f(x \oplus y_i) \), where the sum is taken in \( \mathbb{Z}_n \). Let \( I[Y] \subseteq [n] \) be the set of those indices on which every vector in \( Y \) is zero and \( J[Y] \) be just the complement of \( I[Y] \). Then, the following observation will be very useful in the ensuing calculation:

**Observation 2.14.** Let \( P \) be a polynomial of degree \( d \) in \( n \) variables over \( \mathbb{Z}_n \). Then, for each sequence \( Y \) of \( m \) boolean vectors in \([0, 1]^n\), the polynomial \( P^{Y[|Y|]} \) is a polynomial of degree \( d - 1 \) in variables from \( I[Y] \) for each vector \( v \in \{0, 1\}^{|I[Y]|} \).

**Proof of Lemma 2.13.** We drop the superscript from \( S^{m+1} \) to avoid clutter in the following discussion. We shall induce on the degree \( d \) of the polynomial. Our IH is that there exists a positive real constant \( \mu_{d-1} < 1 \) such that for all polynomials \( R \) of degree at most \( d - 1 \) and for all \( n \geq 0 \) we have \( |S(\alpha, \beta, R)| \leq 2^d \mu_{d-1} \). The base case of \( d = 0 \) is easily verified and is dealt with in earlier works on correlation. Note that \( \mu_0 \) depends only on \( q \). Our inductive step will yield a relationship between \( \mu_{d-1} \) and \( \mu_d \) that will also give us our desired explicit bound of (2.8).

As in [11, 27], we raise \( S \) to its \( m \)th power. Our point of departure from these work, is to write \( (S)^m \) in a slightly different way.

\[
(S)^m = \mathbb{E}_{y_1, \ldots, y_{m-1} \in \{0, 1\}^n} \mathbb{E}_x \left[ c_m \left( P(x) + \sum_{j=1}^{m-1} P(x \oplus y^j) \right) \right] \times \\
\times \mathbb{E}_y \left[ \sum_{i=1}^n x_i + \sum_{i=1}^n (x_i \oplus y^1_i) + \cdots + \sum_{i=1}^n (x_i \oplus y^{m-1}_i) \right] \quad (2.9)
\]

Let \( Y \) be the sequence of length \( m - 1 \) formed by a given set of vectors \( y^1, \ldots, y^{m-1} \). We denote by \( u \) and \( v \) respectively the projection of \( x \) to \( I[Y] \) and \( J[Y] \). Let \( n_I \) and \( n_J \) be the cardinality of \( I[Y] \) and \( J[Y] \) (note that \( n_I + n_J = n \)). Then, one can verify

\[
(2.9) = \mathbb{E}_{y_1, \ldots, y_{m-1} \in \{0, 1\}^n} \mathbb{E}_{v \in \{0, 1\}^n} \left[ c_m \left( Q^{y_1, \ldots, y^{m-1}}(v) \right) c_q(n_J) \times \\
\times \mathbb{E}_{u \in \{0, 1\}^n} \left[ c_m \left( P^{Y[|Y|]}(u) \right) c_q(m \sum_{i=1}^{n_I} u_i) \right] \right] \quad (2.10)
\]

where \( Q^{y_1, \ldots, y^{m-1}} \) is some polynomial that is determined by \( y^1, \ldots, y^{m-1} \) and polynomial \( P \).
The key thing to note is that Observation 2.14 implies $p_Y(v)$ to be a polynomial of degree at most $d-1$ over $u$ for every sequence $Y = y^1, \ldots, y^{m-1}$ and every vector $v$. Thus, the inside sum of (2.10) over the variable $u$ can be estimated using our inductive hypothesis. Noting that the number of sequences $Y$ for which $|I_Y| = k$ is exactly $\binom{n}{k}(2^{m-1} - 1)^{k-1}$ and using the triangle inequality with the binomial theorem, we get.

$$|S|^m \leq \sum_{k=0}^{n} \binom{n}{k}(2^{m-1} - 1)^{k-1} \sum_{i=0}^{2^{m-1}} \mu_d = 2^n \left(1 - \frac{1 - \mu_{d-1}}{2^{m-1}}\right)^n$$  \hspace{1cm} (2.11)

The rest of the calculation proceeds exactly as in Green et. al. [27]. We repeat it here for the sake of self-containment. Taking the $m$th root of both sides of (2.11), using the inequality $(1 - x)^{1/m} \leq 1 - x/m$ if $0 \leq x < 1$ and $m > 1$ after rearranging, we obtain

$$1 - \mu_d \geq \frac{1 - \mu_{d-1}}{m2^{m-1}} \geq \frac{1 - \mu_0}{(m2^{m-1})^\gamma} \hspace{1cm} (2.12)$$

Substituting $\gamma = 1 - \mu_0$, one gets $\mu_d \leq \exp(-\frac{\gamma}{(m2^{m-1})^\gamma})$. This immediately yields (2.8) in Lemma 2.13.

3 \hspace{1cm} Computation by a System of Polynomials

It is natural to extend the notion of computation of a boolean function by a single polynomial to the notion of computation by a system of polynomials. Apart from the fact that systems of polynomials are central objects of interest in branches of pure mathematics like algebraic geometry, the study of their computational power is motivated from proving lower bounds in both boolean and arithmetic circuits. As before, the fact that our polynomials are over a ring $\mathbb{Z}_m$ (rather than a field) and that we are interested in their behavior over the boolean cube, presents difficulties.

Let $\mathcal{P}$ be a system of polynomials $P_1, \ldots, P_s$, each over $\mathbb{Z}_m$ and let $A_1, \ldots, A_s$ be their respective accepting sets. The boolean function computed by $\mathcal{P}$, denoted by $f^\mathcal{P}$, is simply given by the following: for any $x \in \{0, 1\}^n$, $f^\mathcal{P}(x)$ = 1 if $P_i(x) \in A_i \mod m$ for each $1 \leq i \leq s$, otherwise $f^\mathcal{P}(x) = 0$. The degree of the system $\mathcal{P}$, denoted by $\deg(\mathcal{P})$, is the degree of a maximal degree polynomial in $\mathcal{P}$, i.e. $\max(\deg(P_i) : i \leq s)$.

**Definition 3.1.** The $s$-simultaneous $\text{MOD}_m$-degree of a boolean function $f$, denoted by $\deg_s^m(f)$, is the degree of a minimal degree system of $s$ polynomials computing $f$. 

63
Of course, making progress on proving degree lower bounds for a system of polynomials in general is a harder problem than proving lower bounds on the degree of a single polynomial. It may thus seem pointless to work with systems of polynomials before resolving questions from the previous section. However, consider the following: we know that a linear polynomial over $\mathbb{Z}_m$ cannot represent any of AND, OR and MOD$_q$ function. In fact, from results in the previous section, we know that one provably needs almost logarithmic degree to represent them. Thus, one may hope to answer questions of the following type: How large a lower bound on $s$ can we prove so that $\deg s^m(f) > 1$? As we will see that even for this case, proving strong lower bounds on $s$ can be non-trivial. Additionally, such lower bounds yield new lower bounds on the size of some restricted circuits for which no other methods are currently known.

3.1 Linear Systems

Let $L = \{\ell_1, \ldots, \ell_t\}$ be a set of $n$-variate linear forms over $\mathbb{Z}_m$. Such a set forms a linear map $L : \mathbb{Z}_m^n \to \mathbb{Z}_m^t$. Conversely, given such a linear map, there exists a corresponding set of linear forms. For $v \in \mathbb{Z}_m^t$, let $K_L^t(v)$ represent the set of points in $\{0, 1\}^n$, that satisfy $\ell_i = v_i$ for all $1 \leq i \leq t$. Then, we show the following:

**Theorem 3.2** (Chattopadhyay, Goyal, Pudlák and Thérien [15]). For every positive integer $m$, there exists a positive constant $c$ such that the following holds. Let $L : \mathbb{Z}_m^n \to \mathbb{Z}_m^t$ be a linear map. For any $v \in \mathbb{Z}_m^t$, if $K_L(v)$ is non-empty, then

$$|K_L(v)| \geq \frac{2^n}{c^t}. \quad (3.1)$$

A simple averaging argument shows that for every $L : \mathbb{Z}_m^n \to \mathbb{Z}_m^t$, there exists a $v \in \mathbb{Z}_m^t$ such that $K_L(v)$ has size at least $2^n/m^t$. Theorem 3.2 is a kind of concentration result in the sense that it shows that every $K_L^t(v)$ is of size close to the average size if it is non-empty. We note that the results in [43], based on methods introduced in [7], imply a lower bound of $(\frac{\alpha}{\alpha - 1})^n \cdot \frac{1}{\alpha}$ on the size of $K_L^t(v)$ when it is non-empty, and $\alpha$ is an increasing function of $m$. This is still exponentially weaker than what is given by (3.1).

3.2 An Excursion

Before we prove Theorem 3.2, we draw on a notion from combinatorial group theory. Consider a fixed finite abelian group $G$. The *Davenport constant* of $G$, denoted by $s(G)$, is the smallest integer $k$ such that every sequence of elements of $G$ of length at least $k$, has a non-empty subsequence that sums to zero. The pigeon-hole-principle shows that $s(G)$ is finite if $G$ is finite. This is because if we
have a sequence of length larger than $|G|^2$, then some element $a$ of $G$ is repeated at least $|G|$ times. The sub-sequence formed by the first $|G|$ instances of $a$ indeed sums to zero as the order of every element in $G$ divides $|G|$. Thus, $s(G) \leq |G|^2$, which gives a quadratic upper bound on the Davenport constant w.r.t. the size of the group.

For specific groups, one can show much better bounds. For instance, if the group is $\mathbb{Z}_p$, then one can show, using the polynomial method, that $s(\mathbb{Z}_p)$ is $p$. Clearly, the lower bound follows by considering the sequence of $(p - 1)$ occurrences of the identity element. Such a sequence has no non-empty subsequence summing to zero. The upper bound can be established as follows: Let $a_1, \ldots, a_p$ be a sequence of elements from $\mathbb{Z}_p$. Assume that no zero-sum subsequence of it exists. In other words, the polynomial $a_1x_1 + \cdots + a_px_p$ over $\mathbb{Z}_p$ evaluates to zero only at one point in the boolean cube $\{0,1\}^p$, which is the all zero point. Thus, applying Fermat’s Gift, the polynomial $P = 1 - (a_1x_1 + \cdots + a_px_p)^{p-1}$ is exactly the OR function of $p$ boolean variables over $\mathbb{Z}_p$. However, recall that equation (2.1) shows that the degree of the OR polynomial is $p$. This contradiction finishes the argument.

Olson [35] showed a more general statement: Let $G$ be an abelian $p$-group of the form $\mathbb{Z}_{p^1} \oplus \mathbb{Z}_{p^2} \oplus \cdots \oplus \mathbb{Z}_{p^s}$, where $\oplus$ denotes direct sum. He shows that $s(G) = 1 + \sum_{i=1}^{s} (p^i - 1)$ in this case. We show a little later that $s(\mathbb{Z}_p^n)$ is at most $c(m)n$, where $c(m)$ is a constant that just depends on $m$. Before doing that, we recall another result by Olson [36] that connects $s(G)$ with the set of boolean solutions to the equation $g_1x_1 + \cdots + g_nx_n = 0$, denoted by $K(G, n)$, where each $g_i \in G$.

**Theorem 3.3** (Olson’s Theorem). $|K(G, n)| \geq \max\{1, 2^{n+1-s(G)}\}$.

*Proof adapted from [36].* We prove this by induction on $n$. For $n \leq s(G) - 1$, the theorem is vacuously true. Assuming it is true for $n$, we prove it for $n + 1$. Let the equation be $g_1x_1 + \cdots + g_{n+1}x_{n+1} = 0$. By the definition of $s(G)$, there is a subsequence of $g_1, \ldots, g_{s(G)}$ that has a subsequence that sums to zero. W.l.o.g., assume this subsequence to be $g_1, \ldots, g_t$. Then consider the equation $(-g_t)x_2 + \cdots + (-g_t)x_t + g_{t+1}x_{t+1} + \cdots + g_{n+1}x_{n+1} = 0$. By our hypothesis, this equation on $n$ variables has at least $2^{n+1-s(G)}$ solutions. For each such solution point $u$, we obtain a solution to the original equation over $n + 1$ variables in which the value of $x_1$ is set to $1$ in the following way: $x_1 = 1$, for $2 \leq i \leq t$, $x_i$ is set to the value that is the complement of its value in $u$, and for $t < i \leq n + 1$, $x_i$ is set to its corresponding value in $u$. Finally, extend the solutions of $g_2x_2 + \cdots + g_{n+1}x_{n+1} = 0$ to our original equation by $s!$ imply fixing $x_1 = 0$ to obtain at least another $2^{n+1-s(G)}$ solutions. Thus, we have at least $2^{n+2-s(G)}$ solutions in total, proving the theorem. □
3.3 A Simple Fourier Analytic Argument

The usefulness of Olson’s Theorem for our purpose is evident from its following immediate corollary:

**Corollary 3.4.** Let $L : \mathbb{Z}_m \to \mathbb{Z}_m^t$ be a linear map. Then, for all $v \in \mathbb{Z}_m^t$ such that $K^L(v)$ is non-empty, we have $|K^L(v)| \geq 2^{m+1-n|L|^t}$. 

**Proof.** Let $L \equiv \{\ell_1, \ldots, \ell_t\}$ be the underlying linear forms, where $\ell_i = a_{i,1}x_1 + \cdots + a_{i,n}x_n$. As $K^L(v)$ is non-empty, there exists $b \in \{0, 1\}^n$ such that $\ell_i(b) = v_i$. Consider $\ell'_i = a'_{i,1}x_1 + \cdots + a'_{i,n}x_n$, where $a'_{i,j} = a_{i,j}$ if $b_j = 1$ and otherwise $a'_{i,j} = a_{i,j}$, for each $1 \leq j \leq n$ and $1 \leq i \leq t$. Define $L' \equiv \{\ell'_1, \ldots, \ell'_t\}$. Then, it is straightforward to verify that sets $K^L(v)$ and $K^{L'}(0')$ are in one-to-one correspondence with each other. The result follows by observing that Olson’s Theorem implies $K^{L'}(0')$ has size at least $2^{m+1-n|L|^t}$.  

In view of Corollary 3.4, it is sufficient to establish an $O(t)$ upper bound on $s(\mathbb{Z}_m^t)$ for proving Theorem 3.2. Although, to the best of our knowledge, determining the exact bound on $s(\mathbb{Z}_m^t)$ is still open, the linear upper bound that we seek follows from the independent work of Meshulam [34] and Therien [43]. We include a proof of this, using simple Fourier analysis over groups of the form $\mathbb{Z}_m^t$. Recall, from the proof of Bourgain’s Theorem in Section 2.1, $e_m(y)$ denotes the primitive $m$-th root of unity raised to the $y$th power.

**Theorem 3.5.** If $m$ is even, $s(\mathbb{Z}_m^t) \leq ct$, where $c = \frac{\log m}{\log m - \log(m-1)}$ is a constant.

**Proof.** Let $L \equiv \{\ell_1, \ldots, \ell_t\}$ be a linear map from $\mathbb{Z}_m^t$ to $\mathbb{Z}_m^t$, such that $K^L(0')$ is a singleton set, i.e. contains only the point $0'$. Let $\lambda_S : \mathbb{Z}_m^t \to \{0, 1\}$ denote the characteristic function for any set $S \subseteq \mathbb{Z}_m^t$. Then, using Fact 2.12, one writes

$$\lambda_{[0,1]}(x) \equiv \frac{1}{m^t} \prod_{j=1}^{t} \left[ \sum_{a=0}^{m-1} e_m(ax_j) + \sum_{a=0}^{m-1} e_m(a(x_j - 1)) \right]$$

$$= \frac{1}{m^t} \prod_{j=1}^{t} \left[ \sum_{a=0}^{m-1} (1 + e_m(-a)) e_m(ax_j) \right].$$

Let $m = 2\ell$. Then clearly for $a = \ell$, we have $(1 + e_m(a)) = 1 + e_m(\pi) = 0$ using a basic trigonometric identity. Thus, noting that $|\text{supp}(f(g))| \leq |\text{supp}(f)| \cdot |\text{supp}(g)|$, we see that $|\text{supp}(\lambda_{[0,1]}(\ell))| \leq (m - 1)\ell$. Further,

$$\lambda_{K^{L}(0')}(x) \equiv \left[ \prod_{j=1}^{t} \left( \frac{1}{m} \sum_{a=0}^{m-1} e_m(a\ell_j(x)) \right) \right] \lambda_{[0,1]}(x).$$
Thus, one concludes
\[ |\text{supp}(\hat{\lambda}_K^{L}(0^t))| \leq m't|\text{supp}(\lambda_{\{0,1\}^t})| \leq m'(m - 1)^t.\]

Applying the Uncertainty Principle from Fourier Analysis, we get
\[ m'(m - 1)^t \geq |Z_m^s| = m^t,\]
whence the result follows.

The case of an odd \( m \) can be dealt with by the following simple trick. Multiply each linear form \( \ell_i \) by 2. Viewing each modified linear form to be over \( \mathbb{Z}_{2m} \) (instead of over \( \mathbb{Z}_m \)), we obtain a new map \( L' : \mathbb{Z}_{2m}^t \rightarrow \mathbb{Z}_m^t \). It is easily verified that sets \( K^L(0^t) \) and \( K^{L'}(0^t) \) are in one-to-one correspondence with each other. Hence, applying Theorem 3.5 to \( K^{L'}(0^t) \) yields bounds on \( K^L(0^t) \) as well, though with a very slight worsening of the constant \( c \).

**Corollary 3.6.** For every \( m, s(\mathbb{Z}_m^t) \leq c t, \) where \( c = \frac{\log(2m)}{\log(2m) - \log(2m - 1)} \) is a constant that just depends on \( m \).

Combining Corollary 3.4 with bounds on \( s(\mathbb{Z}_m^t) \) as given above, we immediately derive Theorem 3.2 which states that the size of each non-empty \( K^L(v) \) is at least \( 2^n c^t \).

**Remark 3.7.** Here, we point out a consequence of Theorem 3.2 for \( \text{CC}^0[m] \) circuits. It easily yields a linear lower bound on the size of such circuits\(^6\) for computing \( \text{AND} \). Such a bound was first obtained by Thérien [43]. It also makes some progress toward the Large Support Set Conjecture (see Conj. 1.5). While there it is conjectured that the size of the support set of a function computed by a \( \text{CC}^0[m] \) circuit decays polynomially w.r.t. the size of the circuit, Theorem 3.2 yields an exponential decay. Recently, Allender and Koucký [3] have shown that a lower bound of the form \( n^{1 + \gamma} \) on the size of a \( \text{CC}^0[m] \) circuit computing \( \text{AND} \) (MOD\(_q\)), for any constant \( \gamma > 0 \) that does not depend on the depth of the circuit, is enough to imply a superpolynomial lower bound on \( \text{CC}^0[m] \) circuits computing \( \text{AND} \) (MOD\(_q\)).

### 3.4 Computing MOD\(_q\)

Until recently, it was not known if a linear system \( \mathcal{L} = \{\ell_1, \ldots, \ell_t\} \) over \( \mathbb{Z}_m \) with arbitrary accepting sets \( \{A_1, \ldots, A_t\} \) could compute \( \text{MOD}_q \), even for \( t = o(n) \).

---

\(^6\)In fact, as the bound is information theoretic, one need not impose any restriction on the depth of a circuit.
A stronger result of [15] (and implicit in the independent work of Hansen [30]),
showed that even polynomial systems of low degree and small size fail to correlate
well with \( \text{MOD}_q \).

**Definition 3.8.** The \( \mathbb{Z}_q \)-discrepancy of a boolean function \( f \),
denoted by \( \text{disc}_q(f) \), is given by the following:

\[
\text{disc}_q(f) \equiv \left| \Pr[f(x) = 1 \land x \in M_q(b)] - \frac{1}{q} \Pr[f(x) = 1] \right|
\]

The theorem below, first obtained in [15] and independently in [30], shows
that low-degree polynomial systems of small size have exponentially small \( \mathbb{Z}_q \)-discrepancy.

**Theorem 3.9 (Polynomial Uniformity).** For all positive co-prime integers \( m, q \),
there exists a positive constant \( \gamma = \gamma(m, q) < 1 \) such that the following holds:
let \( \mathcal{P} = \{P_1, \ldots, P_t\} \) be a \( n \)-variate polynomial system of degree \( d \) over \( \mathbb{Z}_m \), with
accepting sets \( \{A_1, \ldots, A_t\} \). Then,

\[
\text{disc}_q(f^\mathcal{P}) \leq (m - 1)\exp(-n/\gamma^d).
\]

The above result follows from a simple use of exponential sums, hinting at
their untapped potential in this context.

**Remark 3.10.** The special case of the Polynomial Uniformity Theorem, obtained
by restricting the system to be linear, already leads to an interesting consequence
for circuits with \( \text{MOD}_m \) gates. Using this, [15] shows that circuits (of arbitrary
depth) comprising only \( \text{MOD}_m \) gates cannot compute \( \text{MOD}_q \) in sub-linear size, if
\( (m, q) \) are co-prime. This significantly improves upon the earlier result of Smolensky
[41] that showed such circuits need \( \Omega(\log n) \) size. Further, [15] combine this
special case of the Polynomial Uniformity Theorem with graph-theoretic argu-
ments to prove that such circuits of bounded depth need superlinear number of
wires to compute \( \text{MOD}_q \). This, in some sense, is the strongest known lower bound
for general \( \text{CC}^0[m] \) circuits.

Very recently, Chattopadhyay and Wigderson [18] have been able to signif-
ically improve Theorem 3.9 for the case of linear systems under the condition
that \( m \) is precisely a product of two primes.

**Theorem 3.11 (Two-Prime Uniformity).** Let \( m, q \) be coprime positive integers,
with \( m = p_1p_2 \) and each \( p_i \) is a prime. There exists a positive constant \( \gamma = \gamma(m, q) < 1 \) such that the following holds: let \( \mathcal{L} = \{L_1, \ldots, L_t\} \) be a \( n \)-variate
linear system over \( \mathbb{Z}_m \), with accepting sets \( \{A_1, \ldots, A_t\} \). Then,

\[
\text{disc}_q(f^\mathcal{L}) \leq \exp(-\gamma n).
\]
An interesting thing to note is that the constant $\gamma$ in equation (3.3) above is independent of the size $t$ of the system. The argument of [18] is complicated and combines ideas of using exponential sums from [15], estimates of Bourgain (Lemma 2.13 in this article) with the notion of matrix rigidity from the ingenious work of Grigoriev and Razborov [28] in \textit{arithmetic circuits}. While space constraints will not allow us to cover the entire argument, we describe some details of the main ideas involved in proving the Two-Prime Uniformity Theorem.

### 3.4.1 Singleton Accepting Sets

To begin with, let us assume that each accepting set is a singleton set. In this case, w.l.o.g each $A_i \equiv \{0\}$. Then, as before, one can write

$$
\Pr_x[f_L(x) = 1 \land x \in M_q(b)] = \mathbb{E}_{x \in \{0,1\}^n} \left[ \left( \frac{1}{m} \sum_{a=0}^{m-1} e_m(a(\ell_i(x) - a_i))) \right) \left( \frac{1}{q} \sum_{b=0}^{q-1} e_q(b(x_1 + \cdots + x_n - b)) \right) \right]
$$

(3.4)

Mimicking arguments used in the proof of Bourgain’s Uniformity Lemma to go from (2.4) to (2.7), we obtain,

$$
disc_q(f_L) \leq \frac{1}{m'} \sum_{j=1}^{m'} \mathbb{E}_{x \in \{0,1\}^n} \left[ e_m(r_j(x))e_q(b(x_1 + \cdots + x_n)) \right]
$$

where, each $r_j$ is a linear polynomial obtained by a $\mathbb{Z}_m$-linear combination of $\ell_i$’s. Writing $r_j(x) = a_{j,1}x_1 + \cdots + a_{j,n}x_n$, we can separate variables and obtain

$$
\left| \mathbb{E}_{x \in \{0,1\}^n} \left[ e_m(r_j(x))e_q(b(x_1 + \cdots + x_n)) \right] \right| = \prod_{j=1}^t \left| \mathbb{E}_{x \in \{0,1\}^{n_j}} \left[ e_m(a_{j,1}x_1)e_q(bx_1) \right] \right| \leq \exp(-an)
$$

for some $0 < a < 1$, where the last inequality is a simple exercise to derive using the fact that $m, q$ are co-prime. Thus, in the singleton case there is no dependence on $t$ the number of polynomials in $L$.

For general accepting sets, the first thing to do is to break down our original system into all possible singleton accepting set systems: we write $f_L \equiv \sum_{j=1}^t f_{L_j}$, where $L_j$ is a singleton system verifying if $x$ satisfies $\ell_i(x) = a_{i,j}$ for $1 \leq i \leq t$ and $a_{i,j} \in A_i$. Here $s \leq (m - 1)^t$ as we may assume that each $A_i$ is a proper subset.
This decomposition of $f^L$, along with an application of triangle inequality allows us to deal with individual $f^{L_i}$ in the manner prescribed above for singleton accepting sets. It is straightforward to verify that it proves Theorem 3.9 for the restricted case of linear systems.

Remark 3.12. The careful reader may have noted that fortified with Bourgain’s estimates from (2.8) in Lemma 2.13 for degree $d$ polynomials, each step of the above argument readily adapts to polynomial systems of degree $d$ yielding the Polynomial Uniformity Theorem. Further, it is worth pointing out that this technique yields much stronger results for singleton polynomial systems just as in the case of singleton linear systems described above. These stronger bounds yield exponential lower bounds for depth-four circuits of type \( \text{MAJ} \circ \text{AND} \circ \text{MOD}_m^{(0)} \circ \text{AND}_{\log n} \) (see Theorem 6 in [18]).

3.4.2 Low Rank Systems

The first thing to note is that arguments in the previous section for linear systems of small size almost instantaneously generalize to systems of low rank. Of course, we have to define rank properly as we are over the ring $Z_m$ with zero divisors. The definition we need is simply the following: the $Z_m$-rank of $L$ is the smallest positive integer $r$ such that there exists $r$ linear forms in $L$ that generate every other linear form in the system as some $Z_m$-linear combination of them. W.l.o.g, let these basis forms be $\ell_1, \ldots, \ell_r$.

Observation 3.13. Let $L$ be a linear system of rank $r$. Then, $\text{disc}_q(f^L) \leq (m-1)^r \exp(-\gamma n)$, where $\gamma = \gamma(m, q)$ is a constant.

Proof. Assume w.l.o.g., that $\ell_1, \ldots, \ell_r$ span the remaining $t - r$ forms in $L$. Thus, the $r$-tuple $(\ell_1(x), \ldots, \ell_r(x))$ at any point $x$, determines $\ell_j(x)$ for any $\ell_j \in L$. Hence, we can write $f^L \equiv \sum_{t \times J} f^{L_i}$, as before, going over all possible $r$-tuples of values of the singletons composing $A_i$ for $i \leq r$, and keeping only those tuples for which satisfying the first $r$ equations implies satisfying the remaining $t - r$ equations determined by them. Thus, $|J| \leq (m-1)^r$ and we conclude as in the proof of (linear subcase of) Theorem 3.9.

Hence, if our system has sublinear rank we can prove very good bounds on the discrepancy. A tempting intuition from linear algebra suggests that systems with high (i.e. linear) rank should be almost unsatisfiable and hence their solution set cannot correlate well with a nearly balanced function like $\text{MOD}_q$. This may not be true because our domain of interest is the boolean cube and not $Z_m^n$. Indeed, the following example confirms this fear: let $L$ have $n$ linear forms, with the $i$th linear form being just $x_i$. Each accepting set $A_i \equiv \{0, 1\}$. Thus, the rank of this system is $n$, but every point in our boolean domain satisfies it!
On the other hand, this counter example represents a natural class of systems, those that are sparse. We say $L$ is $k$-sparse if each $\ell_i \in L$ has at most $k$ non-zero coefficients (out of the possible $n$) appearing in it. The following shows that sparse systems have low $\mathbb{Z}_q$ discrepancy.

**Lemma 3.14.** Let $L$ be a $k$-sparse linear system in $\mathbb{Z}_m$. Then, $\text{disc}_q(f^L) \leq \exp(-n/\gamma^k)$ for some constant $\gamma(m,q)$, if $m,q$ are co-prime.

**Proof.** Consider any linear form $\ell_i$ in the system, with its accepting set $A_i$. As $L$ is $k$-sparse, the boolean function $f^{\ell_i}$ depends on at most $k$ variables. Hence, there is a polynomial $P_i$ of degree at most $k$ over $\mathbb{Z}_m$ that exactly represents it, i.e. $P_i(x) = f^{\ell_i}(x)$ for all $x \in \{0,1\}^n$. Replacing each $\ell_i$ by its corresponding $P_i$ thus yields a singleton polynomial system $P$ of degree at most $k$. The argument gets finished by mimicking the arguments in Section 3.4.1 (see also Remark 3.12 in that section).

### 3.4.3 Low Rigid Rank

It turns out that we can combine low rank and sparsity such that we can handle linear systems which can be made to have low rank after a sparse change to each linear form. This is inspired by Valiant’s famous notion of rigidity [45, 46], used to attack (so far unsuccessfully) size-depth trade-offs for computing linear systems over fields. We use the following definition:

We say $L$ is $(k,r)$-sparse if its associated linear forms $\ell_1, \ldots, \ell_t$ satisfy the following property: each $\ell_i$ can be written as $\ell'_i + L_i$ such that the set $\{L_i|1 \leq i \leq t\}$ has rank $r$ and every $\ell'_i$ is $k$-sparse.

**Lemma 3.15.** Let $L$ be a linear system that is $(k,r)$-sparse. Then, there exists a constant $\gamma$ such that $\text{disc}_q(f^L) \leq m^r \exp(-n/\gamma^k)$, when $m,q$ are co-prime numbers.

**Proof.** As before, we look at the possible evaluations of the various linear forms. Let $t$ be the size of $L$, and let $\ell_i = \ell'_i + L_i$. Wlog, assume that $L_1, \ldots, L_t$ are the linearly independent forms that span every other $L_i$. Then our idea is to split the sum into at most $m^r$ different ones, corresponding to the possible evaluations of $L_1, \ldots, L_t$. Let $u$ be any such evaluation in $\mathbb{Z}_m$. Given $u$, we know what each $L_i$ evaluates to in $\mathbb{Z}_m$, for all $i \leq t$. Hence, we know the set of values in $\mathbb{Z}_m$, denoted by $A^u_i$, that $\ell'_i$ could evaluate to so that $\ell_i$ evaluates to some element in $A_i$. Since, $\ell'_i$ depends on at most $k$ variables, there exists a multilinear polynomial $P^u_i$ over $\mathbb{Z}_m$ of degree at most $k$ such that $P^u_i(x) = 0$ (mod $m$) iff $\ell'_i(x) \in A^u_i$. These observations
allow us to write the following:

\[
\text{disc}_q^t(f^L) = \left| \sum_{u \in [m]^t} B_u \left( \prod_{j=1}^t \frac{1}{m} \sum_{a=0}^{m-1} e_m(\alpha_L(x) - u_j) \right) \times \left( \prod_{i=1}^t \frac{1}{m} \sum_{a=0}^{m-1} e_m(\alpha_{P_i}(x)) \right) e_q(b \sum_{i=1}^n x_i) \right|
\]

Expanding out the product of sums into sum of products,

\[
\text{disc}_q^t(f^L) \leq \sum_{u \in [m]^t} \frac{1}{m^{rt}} \sum_{i=1}^m \sum_{j=1}^m \left| B_u \left( e_m(R_i^t(x) + Q_i^r(x)) e_q(b \sum_{i=1}^n x_i) \right) \right|
\]

where each \( Q_i^r(x) \) is a polynomial of degree at most \( k \) obtained by a \( \mathbb{Z}_m \)-linear combination of the \( t \) polynomials \( P_1^r, \ldots, P_t^r \), and each \( R_i^t \) is a linear polynomial obtained by the \( i \)th \( \mathbb{Z}_m \)-linear combination of the \( L_i \)'s. Thus, applying the bounds from Bourgain’s estimate (2.8), we are done. \( \square \)

At this point, could we hope that systems that are not \((k,r)\)-sparse, i.e. do not have low rigid rank are hardly satisfiable over the cube? Indeed, such a hope is generated from a beautiful result of Grigoriev and Razborov [28]: they manage to show that if a linear system \( L \) over a field \( \mathbb{F}_q \) has high rigid rank, then an exponentially small fraction of the set of points in the boolean cube satisfy the system. To show this, they introduce an ingenious notion of rank called communication rank. Porting their argument to our setting raises an obvious difficulty: they work over a field and we work over the ring \( \mathbb{Z}_m \).

However, in [18], we show that their argument can be generalized to our setting in the following sense: let \( m = p_1 \cdots p_s \) be a product of \( s \) distinct primes. Let \( L \equiv \{ \ell_1, \ldots, \ell_t \} \) be a linear system having \( t \) linear forms in \( \mathbb{Z}_m \). Via Chinese remaindering, any linear form \( \ell_i \) projects to \( s \) linear forms \( \ell_i^1, \ldots, \ell_i^s \), where \( \ell_i^j \) is in the field \( \mathbb{Z}_{p_j} \). Hence, \( L \) naturally projects to \( s \) linear systems \( L^1, \ldots, L^s \), with \( L^j \) in \( \mathbb{Z}_{p_j} \). Indeed, one could consider the rank and sparsity of each \( L^j \) in the field \( \mathbb{Z}_{p_j} \). Motivated by this, we say \( L \) in \( \mathbb{Z}_m \) is \( r \)-simple if the set of linear forms can be partitioned into \( s \) sets \( J_1, \ldots, J_s \), with the following property: the projection of the set of linear forms in \( J_j \) to \( \mathbb{Z}_{p_j} \) forms a \((sm, (sm+1)r)\)-sparse system.

**Theorem 3.16** (Chattopadhyay and Wigderson [18], extending Grigoriev and Razborov [28]). Let \( L = \{ \ell_1, \ldots, \ell_t \} \) be a system of \( t \) linear forms, in \( n \) variables, over \( \mathbb{Z}_m \), where \( m \) is a fixed positive integer with no repeated prime factors. If \( L \) is not \( r \)-simple, then

\[
\Pr \_{x \in [0,1]^n} \left[ \bigwedge_{i=1}^t \ell_i(x) \in A \right] \leq \exp(-\Omega(r)),
\]

72
The Bulletin of the EATCS

where each $A_i \subseteq \mathbb{Z}_m$ is an arbitrary set.

We remark that the proof of the above theorem uses very different techniques than any that we have covered here. In particular, it involves an interesting combination of elementary additive combinatorics and linear algebra. Theorem 3.16 provides a rank-sparsity condition under which the system becomes highly unsatisfiable. It is worth noting that apart from assuming that $m$ is square-free, it does not limit the number of prime factors of $m$. Extending ideas from the proof of Lemma 3.15, [18] complements the above Theorem by the following:

**Lemma 3.17.** Let $\mathcal{L}$ be a linear system over $\mathbb{Z}_m$ with $m = p_1p_2$. Let linear forms in $\mathcal{L}$ admit a partition into sets $J_1$ and $J_2$ such that the set of linear forms in $J_i$ are $(k, r)$-sparse over $\mathbb{Z}_{p_i}$ for each $i \leq 2$. Then, if $m, q$ are co-prime,

$$\text{disc}_{q}^{d}(f^{J}) \leq m^{2r} \exp\left(- \frac{n}{\gamma^{k+1}} \right).$$

where $\gamma = \gamma(m, q)$ is a constant.

Unfortunately, the argument in [18] for proving the above works only for the case when $m$ is precisely a product of two primes. It is not hard to combine Lemma 3.17 and Theorem 3.16 to prove the Two-Prime Uniformity Theorem. We do not waste space filling in more details as the interested reader can find the full argument in [18].

4 Conclusion

We argued that the world of low-degree multilinear polynomials modulo a composite is a very natural and fascinating setting to explore the power of modular counting. Fundamental questions on the degree needed to represent simple functions remain wide open. No serious bottleneck is known that prevents us from making progress on them. We believe that with more efforts these problems can be solved in the not too distant future.

On the other hand, mysteriously $\log n$ comes up as a common barrier in different settings. For instance, it shows up in the argument of Tardos and Barrington [42] seemingly for one reason and in Bourgain’s [11] argument for seemingly another. Is it just coincidental? More intriguingly, by the result of Beigel and Tarui [10] (improving upon an earlier work of Yao [48]), we know that every function in $\text{ACC}^0$ can be written as $f(P(x_1, \ldots, x_n))$, where $f$ is a symmetric function and $P$ is an integer polynomial of polylogarithmic degree with coefficients of magnitude at most quasipolynomial. Again $\Omega(\log n)$-degree bounds can be proven for $P$ (via multiparty communication complexity) to decompose a simple function like GIP.
that can be trivially computed in ACC$^0$[2]. Improving over $\log n$ is wide open!
While it is conceivable that going past $\log n$ degree is difficult for a general symmetric function $f$, it is remarkable that we are stuck, more or less, at the same place even when $f$ is a very special symmetric function like $\text{MOD}_m^a$. Viola and Wigderson [47], using the language of Gower’s norm [23], try to suggest an answer. In a related work, Chattopadhyay [16] argues that troubles on both fronts emanate from the technique of repeatedly raising the sum in question to a fixed power, until the degree of the polynomial crashes to linear. While these provide some clue, we feel that the mystery is not entirely solved.

Assuming that going past $\log n$ degree is a difficult task, we have explored questions about systems of polynomials of degree well below $\log n$. Even understanding linear systems over a modulus that is just a product of two distinct primes has proved non-trivial and has generated interesting mathematics. It is hard to believe that this cannot be pushed to three and more prime factors. Finally, can one generalize Bourgain’s result to the setting of a system of polynomials over $\mathbb{Z}_m$, where each polynomial has appropriately low degree. We know this is true at least when $m$ is a prime power by techniques discussed in this article.

References

The Bulletin of the EATCS


This installment of the concurrency column is devoted to a very informative survey, contributed by François Laroussinie, of recent work on the modelling and specification of open systems using games and alternating-time temporal logics. In particular, the paper focuses on fundamental semantic questions for those specification formalisms, such as the kind of properties that can be stated in various types of logics for games, and on the computational complexity of their model-checking problems. Enjoy it!

I wish the readers of the concurrency column a happy, healthy and productive 2010. I trust that this will be yet another year of interesting developments within our research area, which will culminate with CONCUR 2010 in Paris—a conference that François Laroussinie will co-chair with Paul Gastin. I wish them the best of luck with their work.

TEMPORAL LOGICS FOR GAMES

François Laroussinie
LIAFA, Univ. Paris Diderot – Paris 7 & CNRS UMR 7089, IUF
Francois.Laroussinie@liafa.jussieu.fr

1 Introduction

Model-checking is now a well-established method in the area of formal verification. It has been initiated roughly 25 years ago with the seminal papers by Pnueli,
Clarke and Sifakis [25, 13, 26], and since then many theoretical and practical results have been obtained. In this framework, the behavior of the system to be analyzed is described with some formal model, the property to be verified is stated with a specification language and the analysis is done automatically with a tool (a model checker).

One can use many different types of formalism to represent the behavior of a system. This choice depends on the nature and the features of the system we deal with: Which kind of data is manipulated by the system? Is it necessary to handle real-time constraints? Are there probabilities to consider? The same holds for the specification language: for example, there is a wide family of temporal logics allowing to express different kinds of properties. Of course there is a trade-off between expressiveness and efficiency: at the very end, one aims at using a model-checker to decide whether the specification is satisfied by the model and then the efficiency of the decision procedures is crucial.

An important aspect of complex systems is that they are usually composed by several components that interact together. One can consider them as a whole and verify that the resulting system satisfies some property: given a complete description of some protocol (with a sender, several receivers, a network,...) one can ask whether there is no deadlock (with CTL, this could be expressed by $\text{AG} \neg \text{deadlock}$ where $\text{deadlock}$ would be an atomic proposition characterizing the deadlocked states). We could also verify that every new message is followed by a reception with $\text{AG} (\text{new-message} \Rightarrow \text{AF} \text{reception})$.

Sometimes it is interesting to consider one component $C$ (or a subset of components) and to express properties over it within its environment (i.e. the other components): one can then analyze its ability to act upon the whole system in order to ensure some property. For example, it can be useful to know whether $C$ can always ensure that some request will be granted. By always, we mean whatever the other components do. Thus we need a formalism to model the interaction of components and a specification language that allows us to express this kind of property. And note that it cannot be easily stated with classical temporal logics because it requires an existential quantification over the actions of $C$ ("$C$ can ensure...") and an universal quantification over the actions of the remaining part of the system ("whatever the other components do"). Such problems occur when considering the verification of open systems that have to behave correctly whatever their environments do.

Games are a natural and interesting model to represent this kind of problem. Consider for example the classical train crossing problem. Trains and cars can arrive at the crossing, and the gate has to be controlled in order to (1) avoid crash and (2) ensure liveness in the whole system (neither trains nor cars can be blocked forever). Such a problem can be seen as a game: one player deals with trains, another one drives cars and the last one, let’s say Alice, has to control (open or
close) the gate. The question is then: Has Alice a winning strategy in this game? If so, the corresponding control problem has a solution.

Then we need special specification languages to deal with games: it is important to be able to state the existence of strategies for a given player. This has motivated the introduction of Alternating-time Temporal Logics (ATL, ATL*, . . . ) where explicit quantification over the strategies is possible [5]. The aim of this document is to give an overview of results about these temporal logics. We will mainly focus on semantic questions (Which kind of properties can be stated? How can we increase the expressive power?), and we will also consider complexity results for model checking problems.

**Plan of the paper.** In Section 2 we give the formal definition of Concurrent Game Structures. In Section 3 the temporal logic ATL and its variant ATL* are presented. Several questions about the expressive power and the complexity of these logics will be discussed. In Sections 5 and 6, we will present two extensions of ATL: the first one uses strategy contexts to express complex properties over strategies and the second one allows us to add real-time constraints in the specifications.

## 2 Concurrent Game Structures

Let \( \mathcal{AP} \) be a set of atomic propositions. Concurrent Game Structures are a multiplayer extension of Kripke structures:

**Definition:** [Concurrent Game Structure [5]]

A Concurrent Game Structure (CGS) is a 7-tuple \( \mathcal{S} = (Q, q_0, \ell, \text{Agt}, \mathcal{M}, \mathcal{Mv}, \text{Edg}) \) where:

- \( Q \) is a finite set of control states (or locations) and \( q_0 \in Q \) is the initial location;
- \( \ell : Q \to \mathcal{P}(\mathcal{AP}) \) is a labeling of locations with atomic propositions;
- \( \text{Agt} = \{A_1, \ldots, A_k\} \) is a finite set of agents (or players);
- \( \mathcal{M} \) is a finite alphabet of moves;
- \( \mathcal{Mv} : Q \times \text{Agt} \to \mathcal{P}(\mathcal{M}) \) gives the set of possible moves for every player in every location;
- \( \text{Edg} : Q \times \mathcal{M}^k \to Q \) is the transition table of \( \mathcal{S} \): Edg\((q, m_1, \ldots, m_k) \) is the new location of \( \mathcal{S} \) when \( A_i \) plays the move \( m_i \) for \( i = 1, \ldots, k \) from location \( q \).
The size $|C|$ of a CGS $C$ is defined as $|Q| + |Edg|$, where $|Edg|$ is the size of the transition table.

From the current location $q$, any player $A_i$ independently chooses a move $m_i \in \mathcal{M}_v(q, A_i)$ and the transition table $Edg$ provides the new location $q' = Edg(q, m_1, \ldots, m_k)$. We use $\text{Next}(q, A_i, m_i)$ to denote the set of locations that are reachable from $q$ when Player $A_i$ chooses $m_i$ (i.e. $q' \in \text{Next}(q, A_i, m_i)$ iff $\exists \bar{m} \in \mathcal{M}_k$ with $\bar{m}(j) \in \mathcal{M}_v(q, A_j)$ for any $j$, $\bar{m}(i) = m_i$, and $q' = Edg(q, \bar{m})$). And $\text{Next}(q)$ is the set of locations reachable from $q$ for some set of moves of the agents (i.e. $q' \in \text{Next}(q)$ iff $\exists \bar{m} \in \mathcal{M}_k$ with $\bar{m}(i) \in \mathcal{M}_v(q, A_i)$ for any $i$, and $q' = Edg(q, \bar{m})$).

**Example:** Figure 1 presents a CGS corresponding to the game "Rock-paper-scissors". There are two players and each one has to choose R (rock), P (paper) or S (scissors). The alphabet of moves is $\{R, P, S\}$. The transitions of the CGS are labeled with the corresponding set of moves ($\langle R, P \rangle$ means that the move of Player 1 (resp. Player 2) is R (resp. P).

![Figure 1: Rock-paper-scissors](image)

Note that the transition table may be exponential in the number of agents. This has motivated the introduction of a succinct encoding of $Edg$ by using Boolean functions [18]. In this model, namely the implicit CGSs, one can specify the possible transitions from $q$ as a sequence of "If-Then-Else" tests whose conditions are Boolean combinations of propositions of the form "Agent $i$ chooses $m_i$". Formally the transition function from $q$ is defined by a finite sequence $((\varphi_0, q_0), \ldots, (\varphi_n, q_n))$ s.t.: $q_i \in Q$ and $\varphi_i$ is a Boolean combination of propositions “$A_j \rightarrow m$” (i.e. “Agent $A_j$ plays $m$”) and $\varphi_n = \top$. Then we define $\text{Edg}(q, m_1, \ldots, m_k)$ as $q_j$ with $1 \leq j = \min_i(m_i \ldots m_k \models \varphi_j)$. An example of implicit CGS with three players is shown in Figure 2 with $\mathcal{M}_v(q_0) \overset{\text{def}}{=} \{(A_1 \rightarrow 1, q_0), (A_2 \rightarrow 2, q_1), (\top, q_2)\}$: the right part of the figure corresponds to the equivalent "explicit" CGS. We will see in Section 4 that this encoding changes the complexity of model checking.

---

1 with the convention: $m_1 \ldots m_k \models "A_j \rightarrow m"$ iff $m_j = m$. 

82
Alternating transition systems. In [4], another game model is studied: the Alternating Transition Systems (ATS). In this setting, a move of a player consists in a subset of successor locations. $Mv(q,A)$ is then a set of subsets of $Q$. When every player has chosen a move, the new location is the intersection of every move. By construction, this intersection has to be a singleton (this point makes it quite difficult to design an ATS in practice). Translations from ATSs to CGSs and back are possible (w.r.t. alternating bisimulation [6]) but may be expensive [21].

Turn-based games. Finally note also that the turn-based CGSs are an important subclass with interesting semantic and algorithmic properties. In a turn-based CGS, for every location $q$, there is at most one player $A_i$ who has several choices: $|Mv(q,A_i)| = 1$ for any $j \neq i$. Thus the set of locations is partitioned into $k$ sets $Q_1, \ldots, Q_k$: $Q_i$ contains locations that "belong" to Player $A_i$ (who decides the successor location).

Coalitions, executions and strategies. A coalition $A$ is a subset of agents. A move for $A$ is a move for every player in $A$. The previous definitions for $\text{Next}$ can easily be extended to deal with coalitions instead of single players. Given $m$ a move for $A$ and $m'$ a move for $A \setminus A_i$, we use $m \cdot m'$ to denote the corresponding move for $A \setminus A_i$. Of course we have $\text{Next}(q,A\setminus A_i) = \{\text{Edg}(q,m \cdot m')\}$.

An execution is an infinite sequence $\rho = q_0 \rightarrow q_1 \rightarrow q_2 \ldots$ such that $q_{i+1} \in \text{Next}(q_i)$ for any $i$. We use $\rho[i]$ to denote the state $q_i$ and $\rho[0\ldots i]$ (or $\rho_i$) for the prefix $q_0 \rightarrow q_1 \rightarrow \ldots \rightarrow q_i$.

A strategy for Player $A_i$ is a function $f_{A_i}$ that maps any finite prefix $q_0 \ldots q_j$ of some execution to a possible move for $A_i$, i.e. satisfying $f_{A_i}(q_0 \ldots q_j) \in Mv(q_j, A_i)$. We use $\text{Strat}(A_i)$ to denote the set of strategies for $A_i$. This notion can also be extended to coalitions.

A memoryless (or state-based) strategy $f_{A_i}$ only depends on the current state of the system (i.e. the last state of the prefix): $f_{A_i}$ is then a mapping from $Q$ to...
More generally we could also consider \( k \)-bounded-memory strategies for any integer \( k \) \([23, 29]\), this is a way to characterize the resources needed by a strategy.

Given a strategy \( F_A \) for a coalition \( A \), we use \( \text{Out}(q, F_A) \) to denote the set of executions enforced by the strategy \( F_A \): 
\[
q_0 \rightarrow q_1 \rightarrow q_2 \ldots \in \text{Out}(q, F_A) \iff q_0 = q \text{ and for any } i \text{ we have } q_{i+1} \in \text{Next}(q_i, A, F_A(q_0, \ldots, q_i)).
\]
If \( F_{\emptyset} \) is a strategy for the empty coalition, \( \text{Out}(q, F_{\emptyset}) \) is the set of all executions starting from \( q \) (denoted by \( \text{Exec}(q) \)).

### 3 Temporal logics for games

Temporal logic (TL) is an interesting framework to express properties over reactive systems \([25]\): the temporal modalities offer a natural way to deal with the ordering of events along executions. There are many different TLs differing from the nature of the underlying models, the allowed temporal modalities, etc. First we can mention two main families: the linear-time temporal logics and the branching-time temporal logics. In the first case, a system is viewed as a set of executions and formulas are interpreted over these runs. This is the case of the well-known LTL. One can specify that "any problem is followed by an alarm" by using the following formula: 
\[
\varphi \overset{\text{def}}{=} G (\text{problem} \Rightarrow F \text{ alarm}).
\]
Modality \( G \) (resp. \( F \)) means "Always along the run" (resp. "Eventually along the run"). When we say that an LTL formula \( \varphi \) holds for a Kripke structure \( S \), we usually mean that every run of \( S \) satisfies \( \varphi \). There is an implicit universal quantification over the executions of \( S \).

Branching-time temporal logics (such as CTL or CTL\(^*\)) are interpreted over states of Kripke structures. Every state may have several successors. In addition to classical temporal modalities, we can use the existential (\( E \)) or universal (\( A \)) quantification over the runs starting from the current state. The previous property can be stated in CTL with the following formula: 
\[
AG (\text{problem} \Rightarrow AF \text{ alarm}).
\]
The formula \( AG (\text{problem} \Rightarrow EF \text{ alarm}) \) is very different: it specifies that from any problem state, it is possible to find a run leading to alarm. This property cannot be expressed in linear-time temporal logic (see \([15]\) for a detailed overview of temporal logics). Adding path quantifications allows us to express complex properties on the behavior of reactive systems. In this framework CTL\(^*\) is very powerful and contains both CTL and LTL.

When considering games, it is natural to deal with the strategies for agents to enforce temporal properties. For example, in the game Rock-paper-scissors, one could ask whether there exists a strategy for Player 1 to reach the state 1-Win. Such a query is in fact an existential quantification over a subset of runs (generated by a strategy for Player 1) followed by an universal quantification over the runs of
The Bulletin of the EATCS

this subset. It cannot be expressed by using the existential or universal quantification over paths and this has motivated the introduction of modalities $\langle A \rangle$ in [5]: $\langle A \rangle \Phi$ means that "there exists a strategy for $A$ such that $\Phi$ holds for any run". This provides a new family of temporal logics: $\text{ATL}$, $\text{ATL}^\star$... that correspond to branching-time TLs where $\text{E}$ and $\text{A}$ have been replaced by modalities $\langle A \rangle$ for $A \subseteq \text{Agt}$.

Formally $\text{ATL}$ is defined as follows:

**Definition:** [Syntax of $\text{ATL}$]

\[
\text{ATL} \ni \phi, \psi ::= P | \neg \phi, \phi \lor \psi, \text{ } \langle A \rangle \phi_p
\]

\[
\phi_p ::= X \phi, \phi U \psi, \phi R \psi
\]

with $P \in \text{AP}$ and $A \subseteq \text{Agt}$.

As usual we can easily define $\Rightarrow, \land, \top, \bot$... We will use $F \phi$ (resp. $G \phi$) as a shorthand for $\top U \phi$ (resp. $\neg F \neg \phi$). $\text{ATL}$ formulas are interpreted over states of CGS (the case of Boolean operators is omitted):

\[
q \models_S \langle A \rangle \phi_p \text{ iff } \exists F_A \in \text{Strat}(A). \forall \rho \in \text{Out}(q, F_A) \text{ we have } \rho \models_S \phi_p.
\]

\[
\rho \models_S X \phi \text{ iff } \rho[1] \models_S \phi,
\]

\[
\rho \models_S \phi U \psi \text{ iff } \exists i. \rho[i] \models_S \psi \text{ and } \forall 0 \leq j < i \text{ we have } \rho[j] \models_S \phi,
\]

\[
\rho \models_S \phi R \psi \text{ iff } \forall i : (\exists j < i. \rho(j) \models_S \phi) \lor \rho(i) \models_S \psi
\]

$\text{ATL}^\star$ is defined in the same manner as $\text{CTL}^\star$ with $\langle A \rangle$ modalities: there is no restriction on the embedding of temporal modalities and any strategy quantifier $\langle A \rangle$ can be followed by a general path formula defined as:

\[
\phi_p, \psi ::= \phi, \phi_p \lor \psi_p, \neg \phi_p, X \phi_p, \phi_p U \psi_p
\]

For example, $\langle A \rangle G F P$ is an $\text{ATL}^\star$ formula.

Note that classical quantifications $\text{E}$ and $\text{A}$ can be easily expressed with modalities $\langle A \rangle$:

\[
\text{E} \Phi \equiv \langle \text{Agt} \rangle \Phi \quad \text{A} \Phi \equiv \langle \emptyset \rangle \Phi
\]

Indeed the existential quantification corresponds to a case where all agents cooperate together in order to satisfy a property. Conversely if $\Phi$ is true when nobody tries to make it to be true, it means that $\Phi$ is always true. This implies that $\text{ATL}$ contains $\text{CTL}$, and $\text{ATL}^\star$ contains $\text{CTL}^\star$ (note also that a Kripke structure can be seen as a one-player CGS).
Example of properties. Now we give examples of properties that can be expressed with ATL.

- \(\langle\langle \text{Controller} \rangle\rangle G (\neg P_b)\): there exists a strategy for the controller to ensure that \(P_b\) is never true.

- \(\langle\langle A \rangle\rangle F (\neg \langle\langle B \rangle\rangle F P \land \neg \langle\langle C \rangle\rangle F P)\): there exists a strategy for \(A\) to reach a state where neither \(B\) nor \(C\) can manage alone reach \(P\).

- \(\langle\langle A \rangle\rangle F (\neg \langle\langle B \rangle\rangle F P \land \neg \langle\langle C \rangle\rangle F P \land \langle\langle B, C \rangle\rangle F P)\): there exists a strategy for \(A\) to reach a state where neither \(B\) nor \(C\) can manage alone reach \(P\), but where \(B\) and \(C\) can cooperate together to reach \(P\).

- \(\langle\langle \text{Sender, Receiver} \rangle\rangle F (\text{msg-ok})\): there exists a strategy for the sender and the receiver to reach a state where \(\text{msg-ok}\) is true.

Dealing with strategy quantifiers is not always easy. For example, consider two path formulas \(\Phi_1\) and \(\Phi_2\), it is clear that \(E(\Phi_1 \lor \Phi_2)\) is equivalent to \(E\Phi_1 \lor E\Phi_2\) (and \(A(\Phi_1 \land \Phi_2) \equiv A\Phi_1 \land A\Phi_2\)). But such rules are not true when considering strategy quantifiers:

\[
\langle\langle A \rangle\rangle (\Phi_1 \lor \Phi_1) \not\Rightarrow\Leftarrow \langle\langle A \rangle\rangle \Phi_1 \lor \langle\langle A \rangle\rangle \Phi_1
\]

CGSs are not determined for temporal properties \([5, 16]\): if a coalition \(A\) does not have a strategy to enforce \(\Phi\), it does not imply that \(\text{Agt} \setminus A\) has a strategy to ensure \(\neg \Phi\). Thus \(\neg \langle\langle A \rangle\rangle \varphi\) is not equivalent to \(\langle\langle \text{Agt} \setminus A \rangle\rangle \neg \varphi\). For example, in the "Rock-paper-scissors" game, nobody has a strategy to win or to prevent a defeat and thus we have: \(q_0 \not\models \langle\langle A_1 \rangle\rangle X \text{Win}\) and \(q_0 \not\models \langle\langle A_2 \rangle\rangle X \neg \text{Win}\).

Finally note that memoryless strategies are sufficient to deal with ATL specifications \([5, 28]\) but this is not the case for ATL*: for example, the formula \(\langle\langle A \rangle\rangle (F P_1 \land F P_2)\) – "there exists a strategy for \(A\) to reach \(P_1\) and to reach \(P_2\)" – may require to choose two different moves in the same location in order to reach first \(P_1\) and then \(P_2\).

In \([9]\) there is an extension of ATL* with strategy quantifiers of the form \(\langle A,k \rangle\) with \(k \in \mathbb{N}\) in order to quantify over strategies using a memory of size \(k\). These modalities increase the expressive power (and even the distinguishing power) of ATL*.
Other specification languages. In addition to ATL and ATL*, there are different specification languages in the literature. First we can mention the Alternating-time \( \mu \)-Calculus (AMC) \([5]\) based on the modality \( \langle\langle A \rangle\rangle \) X and fixed-point operators, its semantics is similar to that of propositional \( \mu \)-calculus.

Game Logic \([5]\) allows us to deal explicitly with the trees obtained by the choice of a given strategy for a coalition. In GL, there is an existential quantifier over the strategies of coalition \( A (\exists A) \), and path quantifiers \( (\exists \) and \( \forall = \neg \exists \neg) \) to deal with the paths in the underlying tree corresponding to the chosen strategy. The syntax is as follows (GL contains the tree formulas, and GL\(_p\) contains the path formulas):

**Definition:** [Syntax of GL \([5]\)]

\[
\begin{align*}
\text{GL} \ni \phi, \psi &::= P | \neg \phi | \phi \lor \psi | \exists A \phi \\
\text{GL}_t \ni \phi, \psi &::= \phi | \neg \phi | \phi \lor \psi | \exists \phi \\
\text{GL}_p \ni \phi, \psi &::= \phi | \neg \phi | \phi \lor \psi | X \phi | \phi U \psi
\end{align*}
\]

with \( A \subseteq \text{Agt} \) and \( P \in \text{AP} \).

Semantics is natural: \( \exists A \phi \) holds for a location \( q \) iff there is a strategy \( F_A \in \text{Strat}(A) \) such that the tree \( T_{F_A} \) produced by \( F_A \) from \( q \) (i.e. whose branches belong to \( \text{Out}(q, F_A) \)) satisfies \( \phi \).

For example, \( \exists A (\exists G P_1 \land \exists G P_2) \) holds for \( q \) iff there is a strategy \( F_A \) for \( A \) such that there exist a run where \( P_1 \) is always true and another one verifying \( P_2 \): both runs belong to \( \text{Out}(q, F_A) \). This formula cannot be stated with ATL* \([5]\). We will see in Section 5 an extension of ATL allowing to express such properties.

### 4 Model checking

Given a CGS, \( S \) and an ATL formula \( \Phi \), the model checking problem consists in deciding whether the initial location of \( S \) satisfies \( \Phi \) or not.

The decision procedure computes the subset \( [\psi] \) of locations of \( S \) satisfying \( \psi \) for any subformula \( \psi \) in \( \Phi \). It proceeds in a bottom-up manner starting by atomic propositions and then dealing with outermost formulas. Boolean operators can be easily handled. The set \( [\langle\langle A \rangle\rangle \Psi] \) corresponds to the controllable predecessors of \( [\Psi] \) by \( A \), denoted \( CPre(A, [\Psi]) \). Given a set of locations \( S \) and a coalition \( A \), we have:

\[
CPre(A, S) \overset{\text{def}}{=} \{ q \in Q | \exists m_A \in Mv(q, A) \text{ such that } \text{Next}(q, A, m_A) \subseteq S \}
\]

From any location in \( CPre(A, S) \), \( A \) can enforce to reach a location in \( S \).

---

\(^2\)We will consider other logics in Section 5.
And the treatment of $\langle\langle A \rangle\rangle_U$ and $\langle\langle A \rangle\rangle_R$ is based on a standard fixed-point computation in which we use the controllable predecessors. Given $\Psi \overset{\text{def}}{=} \langle\langle A \rangle\rangle P_1 U P_2$, the set $\llbracket \Psi \rrbracket$ is defined as the least fixed-point of the function: $f : 2^0 \rightarrow 2^0$ defined as follows:

$$f(Z) \overset{\text{def}}{=} \llbracket P_2 \rrbracket \cup (\llbracket P_1 \rrbracket \cap CPre(A,Z))$$

(1)

The formula $\Psi$ can also be expressed by using the following Alternating-time $\mu$-calculus formula: $\mu Z. (P_2 \lor (P_1 \land \langle\langle A \rangle\rangle X Z))$.

The difference from CTL model checking is then the use of $\langle\langle A \rangle\rangle X$ instead of $EX$: we need to consider controllable predecessors instead of classical predecessors. Note also that if we consider turn-based games, it is possible to express $\langle\langle A \rangle\rangle P_1 U P_2$ with the standard propositional $\mu$-calculus because for every state we can use either the existential modality $EX$ or the universal modality $AX$ depending on the "owner" ($A$ or $\bar{A}$) of the state.

The complexity of $ATL$ model checking depends on the complexity of the computation $CPre$ (see [21] for these results for ATS, CGS and implicit CGS). Finally we have:

**Theorem:** Model checking $ATL$ formulas...

- is P-complete on CGSs [5];
- is $\Delta^p_1$-complete on implicit CGSs [21];
- is $\Delta^p_2$-complete on ATSs [21].

For $ATL^*$, we have:

**Theorem:** [5, 21] Model checking $ATL^*$ is 2EXPTIME-complete on ATSs as well as on CGSs and implicit CGSs.

**Satisfiability.** For $ATL$ (and $ATL^*$) the satisfiability problem ("given $\Phi$, does there exists a CGS satisfying $\Phi$?") can be defined in several ways [30]:

(Pb 1.) Given $\Phi$ and a finite set of agents $Agt$, does there exist a multi-agent model over $Agt$ satisfying $\Phi$?

(Pb 2.) Given $\Phi$, does there exist a finite set of agents $Agt$ such that there is a multi-agent model over $Agt$ satisfying $\Phi$?

(Pb 3.) Given $\Phi$, does there exist a multi-agent model over $Agt_0$ (i.e. the set of agents occurring in $\Phi$) satisfying $\Phi$?
These three variations are discussed in [30]. The importance of the agents is highlighted by the following example: consider
\[ \neg \langle A \rangle X P \land \neg \langle A \rangle X P' \land \langle A \rangle X (P \lor P') \]. This formula is not satisfiable in a model where \( A \) is the unique player. But it is clearly satisfiable when considering two players. In [30] it is shown that for ATL specifications, we have: (Pb 1.) and (Pb 3.) are polynomially reducible to each other and (Pb 2.) can be reduced to (Pb 1.) by considering one extra player in addition to \( \text{Ag}_b \). These remarks could also be extended to ATL\(^*\). Finally we have:

**Theorem:**

- Satisfiability problems for ATL are EXPTIME-complete [30].
- Satisfiability problems for ATL\(^*\) are 2EXPTIME-complete [27].

Finally note that in [30] the problem is considered for Alternating Transition Systems (and not for CGSs), but as there exist translations between ATS and CGS (where the set of agents is preserved), the decision problems for the existence of ATS or CGS have the same complexity.

## 5 Extension with strategy contexts

In this section, we present an extension of ATL and ATL\(^*\) with *strategy contexts* that allow us to express complex properties over strategies [9]. The basic idea is to deal with properties of the following form: given a strategy \( F_A \) for \( A \), does there exist (1) a strategy \( F_B \) for \( B \) such that the combination of \( F_B \) with \( F_A \) (denoted "\( F_B \circ F_A \)") ensures some property \( \Phi \), and (2) a strategy \( F_C \) for \( C \) such that "\( F_C \circ F_A \)" ensures some property \( \Psi \)? The choices of \( F_B \) and \( F_C \) are independent (we do not assume any kind of cooperation between \( B \) and \( C \)) but they rely on the choice of \( F_A \).

Consider the following example. A server \( S \) has to treat the requests of \( n \) agents \( A_1, \ldots, A_n \). We want to check whether there is a strategy for \( S \) in order to grant any request of the agents. Expressing such a property with ATL or ATL\(^*\) is not possible. Something like \( \langle S \rangle G \left( \bigwedge_i \text{req}_i \Rightarrow \langle A_i \rangle F \text{grant}_i \right) \) is not correct because the strategy for \( S \) is not taken into account in the choices of \( A_1, A_2, \ldots \). This would mean that every agent can enforce the server to grant its request. The property \( \langle S \rangle G \left( \bigwedge_i \text{req}_i \Rightarrow \langle S, A_i \rangle F \text{grant}_i \right) \) is not right: because we want to specify that one given strategy for \( S \) ensures the property for every agent. Finally \( \langle S, A_1, \ldots, A_n \rangle \left( G \bigwedge_i \text{req}_i \Rightarrow F \text{grant}_i \right) \) does not work because we do not want to assume that agents cooperate together. This has motivated the use of
new modalities $\langle A \rangle$ that are interpreted in strategy contexts containing the strategies fixed by outermost quantifiers: the previous property could then be expressed by the following formula:

$$\langle S \rangle \left( \bigwedge_{i=1}^{n} (\text{req}_i \Rightarrow \langle A_i \rangle F \text{grant}_i) \right)$$

The innermost modalities $\langle A_i \rangle$ are strategy quantifiers within the context of the strategy for $S$ selected by the outermost modality $\langle S \rangle$.

Another example to illustrate the semantics of the new modality $\langle A \rangle$ is given in Figure 3. The left part of the figure gives a graphical interpretation of the formula $\langle A \rangle G \langle B \rangle F P$ in a location $q$: the first strategy quantifier $\langle A \rangle$ selects a set of runs satisfying $G \langle B \rangle F P$, thus from every state $q'$ along these executions there is a strategy for $B$ to ensure $F P$ (and this strategy does not rely on the previous strategy for $A$). In the right part of figure, we consider the formula $\langle A \rangle G \langle B \rangle F P : B$ has to select a set of executions among those provided by the strategy fixed by $A$.

![Figure 3: Interpretation of formulas in strategy contexts.](image)

To formally define the new modalities $\langle - \rangle$, we first introduce some notions over the strategies. A strategy context is a strategy for a subset of agents in $\text{Agt}$. Let $F$ be a strategy for $A \subseteq \text{Agt}$. We use $\text{dom}(F)$ to denote $A$ (the domain of $F$). Given $B \subseteq \text{Agt}$, $(F_A)_B$ (resp. $(F_A)_B$) denotes the restriction of $F_A$ to the coalition $A \cap B$ (resp. $A \setminus B$). It is also possible to combine two strategies $F \in \text{Strat}(A)$ and $F' \in \text{Strat}(B)$, resulting in a strategy $F \circ F' \in \text{Strat}(A \cup B)$ defined as follows: $(F \circ F')_B(q_0, \ldots, q_m)$ equals $F_A(q_0, \ldots, q_n)$ if $A_j \in A$, and it equals $F'_B(q_0, \ldots, q_n)$ if $A_j \in B \setminus A$. Finally, given a strategy $F$ and a finite execution $\rho$,
The Bulletin of the EATCS

we define the strategy $F^\rho$ corresponding to the behaviour of $F$ after $\rho$ as follows:

$F^\rho(\pi) = F(\rho \cdot \pi)$.

Now we define ATLsc:

**Definition**: [Syntax of ATLsc [9]]

$$\text{ATLsc} \ni \phi, \psi : \equiv \prod \phi, \sum \phi, \vee \psi, \langle A \rangle \phi, \langle A \rangle \phi$$

$$\phi_p : \equiv \mathbf{X} \phi, \mathbf{U} \psi, \mathbf{R} \psi$$

with $P \in \text{AP}$ and $A \subseteq \text{Agt}$.

ATLsc formulas are interpreted over states of a CGS $S$ within a context $F$:

$q \models_{S,F} \langle A \rangle \phi_p$ iff $\exists F_A \in \text{Strat}(A). \forall \rho \in \text{Out}(q, F_A \circ F)$ we have $\rho \models_{S,F,F} \phi_p$,

$\rho \models_{S,F} \mathbf{X} \phi$ iff $\rho[1] \models_{S,F,F} \phi$,

$\rho \models_{S,F} \mathbf{U} \psi$ iff $\exists \rho[i] \models_{S,F,F} \psi$ and $\forall 0 \leq j < i$ we have $\rho(j) \models_{S,F,F} \psi$,

$\rho \models_{S,F} \mathbf{R} \psi$ iff $\forall i : (\exists j < i, \rho(j) \models_{S,F,F} \psi) \lor \rho(i) \models_{S,F,F} \psi$.

The modality $\langle A \rangle$ is used to remove the strategy for $A$ from the current strategy context. This modality allows us to easily express ATL strategy quantifiers $\langle A \rangle$:

$$\langle A \rangle \Phi \equiv \langle A \rangle \Phi$$

We define ATLsc as the variant of ATL with $\langle \cdot \rangle$ modalities.

Note that ATLsc formula like $\langle A_1 \rangle (\langle A_2 \rangle \Phi \land \neg \langle A_3 \rangle \Phi')$ can be written in ATLsc as follows: $\langle A_1 \rangle \bot \mathbf{U} (\langle A_2 \rangle \Phi \land \neg \langle A_3 \rangle \Phi')$. Thus we can nest strategy quantifiers in ATLsc.

In [9] several results about the expressiveness of ATLsc and ATLsc are given. In particular, we can see that these logics have a stronger distinguishing power than ATL (and ATLsc): alternating-bisimilar CGSs can be distinguished with $\langle A \rangle$ modalities.

Moreover, examples of complex properties – like Nash equilibrium and winning secure equilibrium – are given in [9]. Note that these properties have been used to motivate the introduction of Strategy Logic (SL) in [12]. SL extends linear-time temporal logics with first-order quantification over strategies, this logic has been studied for the two-player turn-based games.

The model checking problem for ATLsc is decidable. Indeed ATLsc can be translated into $\text{oD}_{\mu}$ [24] that is a powerful extension of $\mu$-calculus with special modalities to deal with strategies.

Finally note also that ATLsc and ATLsc are incomparable with the Alternating-time $\mu$-calculus and strictly more expressive than GL [9].

There exist other logics that have been introduced to deal with such complex properties over strategies. In [1], a variant of ATL called IATL is proposed: the
The main idea is to consider irrevocable strategies (the difference with strategy contexts is that with IA TL as soon as a player has chosen a strategy, he cannot modify it in the sequel). Decision procedures are given when one considers memoryless strategies. Stochastic Game Logic [8] is a probabilistic extension of ATL that uses a variant of strategy contexts, for stochastic games (model checking is undecidable in the general case, but proved decidable when restricting to memoryless strategies).

6 Timed extensions

Classical model checking framework has been adapted to handle timed verification: both models and specification languages have been extended to deal with real-time constraints. Several models have been proposed; the most interesting is probably the well-known timed automata introduced by Alur and Dill [3]. Temporal logics have also been extended in several manners to measure time elapsing between system events.

In this section we will consider a simple real-time extension of CGS with integer durations and two variants of games over dense time. In these models, we have a notion of duration (denoted \(\text{Duration}(\pi)\)) associated with any finite prefix \(\pi\) of an execution. We will only present them informally and mention the main related results.

From the specification language point of view, we will consider an extension of ATL where real-constraints are associated with \(U\) modalities (exactly as it is done for many timed temporal logics as TCTL [2]).

6.1 Timed ATL

TATL has been introduced in [17].

**Definition:** [Syntax of TATL]

\[
\begin{align*}
\text{TATL} \ni & \phi, \psi ::= \quad P \mid \neg \phi \mid \phi \lor \psi \mid \langle \langle A \rangle \rangle \phi_p \\
\phi_p & ::= \quad \phi, U_{\leq c} \psi, \mid \phi, R_{\leq c} \psi
\end{align*}
\]

with \(P \in \text{AP}\) and \(A \subseteq \text{Agt}\).

The integer constants are assumed to be encoded in binary (this is important for the complexity results). TATL formulas are interpreted over states in a timed game model. The semantics is defined as follows:
The Bulletin of the EATCS

\[ q \models S \langle \langle A \rangle \rangle \phi_p \iff \exists F_A \in \text{Strat}(A). \forall p \in \text{Out}(q, F_A) \text{ we have } \rho \models S \phi_p, \]

\[ \rho \models S \phi_s U \sim_d \psi_s \iff \exists p. \rho[p] \models S \psi_s \text{ and } \text{Duration}(\rho[p]) \sim d \]

\[ \rho \models S \phi_s R \sim_d \psi_s \iff \rho \models S (\neg \phi_s U \sim_d \neg \psi_s). \]

Note that we do not consider X modality because as soon as we consider dense time models, the notion of successor state is not relevant.

In the definition of the semantics of U, we use p and p' to denote positions along the run \( \rho \) and \( \prec_p \) is the precedence relation over these positions. In the discrete-time case, a position is an integer but for the dense-time case a position is defined as a pair \((k, \delta) \in \mathbb{N} \times \mathbb{R}_{\geq 0}\) to represent the configuration reached after the \( k \)-th action-transition followed by a \( \delta \) time units delay.

### 6.2 Discrete time

First we consider a simple timed variant of CGSs: A Durational CGS (DCGS) is a CGS where integer durations are associated with every transition [20]. More precisely the transition table is of the form \( \text{Edg} : Q \times \mathbb{N} \to I \times Q \) where \( I \) denotes the set of intervals over \( \mathbb{N} \cup \{\infty\} \). For every transition \( \text{Edg}(q, m) = ([d; D], q') \), the final choice of the duration among the interval \([d; D]\) is done by an additional agent. This allows us to use these agents in the strategy quantifiers within the formulas. Note also that we call TDCGS (Tight DCGS) the class of DCGS where every interval is restricted to a single value.

Given a formula \( \phi \overset{\text{def}}{=} \langle \langle A \rangle \rangle P_1 U \sim_d P_2 \) with \( P_1, P_2 \in \text{AP} \), we can label the locations satisfying \( \phi \) by computing the minimal duration \( d' \) required by \( A \) to enforce reaching \( P_2 \) along a path satisfying \( P_1 \). This can be done recursively over the number of turns in the game. See [20] for a complete description of the algorithms.

**Theorem:** [Model checking DCGS and TDCGS [20]]

- Model checking TATL formula over DCGS or Tight DCGS is EXPTIME-complete.
- Model checking TATL_{\leq \geq} over DCGS or Tight DCGS is P-complete.
EXPTIME-hardness. The complexity lower bound for TATL is based on the complexity of model checking $\langle\langle A \rangle\rangle F \omega P$ in a tight DCGS. It corresponds to a two-players game – the countdown game – in a weighted graph $(V, E, w)$ with $w : E \rightarrow \mathbb{N}_{>0}$. A configuration of this game is a pair $(q, \alpha)$ with $q \in V$ and $\alpha \in \mathbb{N}$. At every turn, Player $A_1$ chooses an integer $d$ such that (1) $0 < d \leq \alpha$ and (2) $\exists (q, r) \in E$ s.t. $w(q, r) = d$. Then Player $A_2$ chooses a transition from $q$ whose duration is $d$ and leading to some $q'$. The new configuration is then $(q', \alpha - d)$. A configuration $(q, 0)$ is winning for $A_1$ (for any $q$). A configuration $(q, \alpha)$ with $\alpha > 0$ and no transition from $q$ whose duration is less than $\alpha$ is winning for $A_2$. Deciding whether there is a winning strategy for $A_1$ in such a game is EXPTIME-complete [19]. It can easily be reduced to a model checking problem of $\langle\langle A \rangle\rangle F \omega P$ over a Tight DCGS.

6.3 Dense time

Here we present two variants of game models based on timed automata (other models exist, see for example [22, 7]).

Timed Games. The most well-known timed models for games are the Timed Games Automata (TGA) introduced in [14]. As in timed automata, there are clocks that progress synchronously with time and every transition is guarded by a constraint over the clocks. The main difference with TA is that the set of transitions is partitioned in two subsets: one for each player (NB: this is a two-player game).

At every turn, from a configuration $(q, v)$ where $q$ is a location and $v$ is a valuation for the clocks, Player $A_1$ has to choose a delay $d_i$ and a transition $t_i$ from his set of transitions (starting from $q$). Then the new configuration of the game is computed as follows: assume $d_1 < d_2$ (resp. $d_2 < d_1$) then the transition $t_1$ (resp. $t_2$) is fired after $d_1$ (resp. $d_2$) time units. If $d_1 = d_2$, the system chooses non-deterministically to perform $t_1$ or $t_2$ after $d_1$ time units.

Consider the example below:

![Diagram of a timed game automaton]

If Player $A_1$ plays $(1.5, c_1)$ and $A_2$ chooses $(0.8, c_2)$ from $(q_1, 0)$, then the new configuration is $(q_1, 0)$ after 0.8 time units. If $A_1$ chooses $(0.5, c_1)$ and $A_2$ plays $(0.8, c_2)$ then the new configuration is $(q_2, 0.5)$. 
The Bulletin of the EATCS

On this example, one can see that $A_2$ can avoid to reach $q_2$ by always choosing $(0, c_2)$ since Player $A_1$ has to play $(d, c_1)$ with $d > 0$... But such a strategy for $A_2$ is not fair: it is a Zeno strategy, it makes time converge. In TGA, we usually require strategies to be non-Zeno: a player cannot block time elapsing. For example, in [14] it is assumed that a player win a game only if its winning objective is satisfied and either time diverges or this player is not responsible for the convergence of time. With such a requirement the previous strategy for $A_2$ is not winning any more even if his control objective is to avoid $q_2$. And in this game, Player $A_1$ has a strategy to reach $q_2$.

Timed CGS. Timed CGSs are CGS with real-time constraints [10]. The main difference compared with the previous TGA is that the concurrency of CGS is preserved: the next location depends on the choices of all agents. An example is given in the Figure below.

In these games, every player chooses a delay $d$ (after which he wants to play) and a move function from $\mathbb{R}^+$ in $M$ that gives a move for any possible delay. This function allows the player to participate to the move even if another player asks to play before $d$. From the choices $(d_i, f_i)$ for $i = 1, \ldots, k$, one can deduce the new location as follows: the system will delay for $d \overset{\text{def}}{=} \min(d_1, \ldots, d_k)$ time units and then perform the transition $\text{Edg}(q_1, f_1(d), \ldots, f_k(d))$.

Consider the previous example. First note that some transitions are missing: for example, when $x < 2$ there is no move $(1.2)$ from $q_1$. In this case we assume that these moves correspond to a self-loop in the TCGS. From $(q_1, x = 1.2)$, assume that Player $A_1$ chooses $(d_1, f_1)$ with $d_1 = 0.9$ and $f_1(d) = 2$ if $d \leq 0.5$ and $f_1(d) = 1$ when $d > 0.5$. Moreover assume that $A_2$ chooses $(d_2, f_2)$ with $d_2 = 1.4$ and $f_2(d) = 2$ when $d \leq 1$ and $f_2(d) = 1$ otherwise. Then with these moves, the system will perform the transition $q_2 \rightarrow q_1$ labeled with $(1.2)$ after 0.9 time units.

Model checking TATL specifications over TCGS is decidable (one can build a finite "region" CGS that is time-abstract game-bisimilar to the infinite CGS corresponding to the semantics of the TCGS).

And we have:
Theorem:

- Model checking TATL over TGA is EXPTIME-complete [14].
- Model checking TATL over TCGS is EXPTIME-hard and can be done in EXPSPACE [10].

Finally note that there is a tool – UppAal Tiga \(^3\) – to automatically analyze a variant of TGA. One can check reachability properties with an on-the-fly algorithm [11].

Acknowledgments

Special thanks to Nicolas Markey, Arnaud Da Costa, Thomas Brihaye and Ghassan Oreiby for all the work we have done together on ATL and its variants. Many thanks to Luca Aceto and Béatrice Bérard for the interesting discussions about temporal logics.

References


\(^3\)http://www.cs.aau.dk/~adavid/tiga/
The Bulletin of the EATCS


97


THE DISTRIBUTED COMPUTING COLUMN

by

PANAGIOTA FATOUROU

Department of Computer Science, University of Crete
P.O. Box 2208 GR-714 09 Heraklion, Crete, Greece
and
Institute of Computer Science (ICS)
Foundation for Research and Technology (FORTH)
N. Plastira 100. Vassilika Vouton
GR-700 13 Heraklion, Crete, Greece
fat@fatourou@csd.uoc.gr

ROBUST SIMULATION OF SHARED MEMORY:
20 YEARS AFTER*

Hagit Attiya
Department of Computer Science, Technion
and School of Computer and Communication Sciences, EPFL

Abstract
The article explores the concept of simulating the abstraction of a shared memory in message passing systems, despite the existence of failures. This abstraction provides an atomic register accessed with read and write operations. This article describes the Attiya, Bar-Noy and Dolev simulation, its origins and generalizations, as well as its applications in theory and practice.

*This article is based on my invited talk at SPAA 2008.
1 Introduction

In the summer of 1989, I spent a week in the bay area, visiting Danny Dolev, who was at IBM Almaden at that time, and Amotz Bar-Noy, who was a postdoc at Stanford, before going to PODC in Edmonton. Danny and Amotz have already done some work on equivalence between shared memory and message-passing communication primitives [14]. Their goal was to port specific renaming algorithms from message-passing systems [10] to shared-memory systems, and therefore, their simulation was tailored for the specific construct—stable vectors—used in these algorithms.

Register constructions were a big fad at the time, and motivated by them, we were looking for a more generic simulation of shared-memory in message passing systems.

In PODC 1990, we have published the fruits of this study in an extended abstract of a paper [8] describing a simulation of a single-writer multi-reader register in a message-passing system. Inspired by concepts from several areas, the paper presented a simple algorithm, later nicknamed the ABD simulation, that supports porting of shared-memory algorithms to message-passing systems. The ABD simulation allowed to concentrate on the former model, at least for studying computability issues.

The simulation, vastly extended to handle dynamic changes in the system and adverse failures, served also as a conceptual basis for several storage systems, and for universal service implementations (state-machine replication).

In this article, I describe the ABD simulation and its origins, and survey the follow-up work that has spanned from it.

2 Inspirations

This section describes the context of our simulation, discussing specifications, algorithms and simulations, which provided inspiration for our approach. The section also lays down some basic terminology used later in this article.

2.1 Sharing Memory

When we started to work on this research project, we were well-aware of the paper of Upfal and Wigderson [50] simulating a parallel random access machine

\[1\] During the same visit, Danny and I also worked on the first snapshot algorithm with bounded memory; these ideas, tremendously simplified and improved by our coauthors later, lead to the atomic snapshots paper [2].

\[2\] Indeed, our title is a takeoff on the title of their paper.
(PRAM) [24] on a synchronous interconnect, like the one used, for example, in the NYU Ultracomputer [30].

Upfal and Wigderson assume a synchronous system, which does not allow any failures. The paper further assumes a complete communication graph (clique) or a concentrator graph and shows how to emulate a PRAM step, namely, a permutation where each node reads from some memory location, or writes to some memory location.

Because many nodes may access the same memory location, their simulation replicates the values of memory locations and stores each of them in several places, in order to reduce latency. To pick the correct value, a node accesses a majority of the copies, and the intersection between these majority sets ensures that the correct value is obtained.

Upfal and Wigderson concentrate mostly on avoiding communication bottlenecks, and hence the emphasis in their paper is on spreading copies in a way that minimizes the load across nodes. As we shall discuss (Section 3.3), this consideration will make a comeback also in the context of asynchronous, failure-prone systems.

2.2 Handling Asynchrony

In contrast to Upfal and Wigderson, who assumed that nodes do not fail and that communication is synchronous, we were interested in asynchronous systems where nodes may fail. Our original simulation assumed that failed nodes simply crash and stop taking steps; later work addressed the possibility of malicious, Byzantine failures (see Section 4.2).

The combination of failures and asynchrony poses an additional challenge of a possible partitioning of the system. It can easily be seen that if more than a majority of the nodes may fail, then two operations may proceed without being aware of each other. This might cause a read operation to miss the value of a preceding write operation. Therefore, the construction assumes that a majority of the nodes do not fail; thus, the number of nodes $n$ is more than twice the number of possible failures $f$; that is, $n > 2f$.

Armed with this assumption, it is possible to rely on Thomas’s majority consensus approach [49] for preserving the consistency of a replicated database.\(^3\)

The majority consensus algorithm coordinates updates to the database replica sites, by having sites (nodes) vote on the acceptability of update requests. For a request to be accepted and applied to all replicas, only a majority need approve it; an update is approved only if the information upon which the request is based

---

\(^3\) This paper makes an interesting read because it deals with concurrency control prior to the introduction of the notion of serializability [46].
is valid; validity, in turn, is determined according to version numbers associated with data values.

This indicates that version numbers may provide the key for coping with failures in an asynchronous system.

2.3 Atomic Registers

A critical step in the simulation was deciding which abstraction to use; indeed, Bar-Noy and Dolev simulated a very specific communication construct, which is not sufficiently generic to use in other contexts, while Upfal and Wigderson simulated a full-fledged PRAM with the associated cost and complication. Luckily for us, the hippest research trend at the time were register constructions, namely, algorithms to simulate a register with certain characteristic out of registers with weaker features.

The “holy grail” of this research area was an atomic multi-writer multi-reader register, and many papers presented algorithms for simulating it from weaker types of registers. This is a fairly general data structure accessed by read and write operations, allowing all processes to write to and read from the register, and ensuring that operations appear to occur instantaneously (more on this below).

A multi-writer multi-reader atomic register is a very general and convenient to use abstraction. However, since many of these algorithms were wrong, or at best, complicated, we decided to simulate a slightly weaker kind of register, atomic single-writer multi-reader register, which can be written only by a single node. This decision turned out to simplify our algorithm considerably; a multi-writer register could still be simulated by porting the shared-memory algorithms. Later, it turned out that simulating an even weaker type of register, with a single writer and a single reader, could lead to a more efficient simulation [7].

To simulate a register in a message-passing system, one must provide two procedures: one for read and the other for write. These procedures translate the operation into a sequence of message sending and receiving, combined with some local computation. When these procedures are executed together by several nodes their low-level events (message sending and receiving) are interleaved, and we need to state their expected results.

The expected behavior in interleaved scenarios is specified through linearizability [34]. Like sequential consistency [37], linearizability requires the results of operations to be as if they were executed sequentially. Furthermore, linearizability also demands that this order respects the order of non-overlapping operations, in which one operation completes before the other starts (we say that the early operation precedes the later one, and that the later one follows the early one).
The Bulletin of the EATCS

3 The ABD Simulation in a Nutshell

One reason the ABD simulation is well-known is due to its simplicity, at least in the unbounded version. This section explains the basic behavior of the algorithm.

We consider a simple system with one node being the writer of a register; all other nodes are readers. All $n$ nodes also store a copy of the current value of the register; this is done in a separate thread from their roles as a writer or reader.

Each value written is associated with an unbounded version number, an integer denoted $\text{version}#$.

To write a value, the writer sends a message $\text{write}(\text{val}, \text{version}#)$, with the new value and an incremented version number, to all nodes and waits for $n-f$ acknowledgments. Figure 1(a) shows an example of the communication pattern: the writer sends the message to all seven nodes, one node does not receive the message (indicated by a dashed line); of the nodes that receive the message, one does not respond at all, another responds but the message is delayed (indicated by a dashed line), so the writer receives acknowledgments from four nodes (a majority out of seven).

To read a value, a reader queries all nodes, and, after receiving at least $n-f$ responses, picks the value with the highest version number. Figure 1(b) shows an example of the communication pattern: the reader sends the message to all seven nodes; all nodes receive the message, two do not respond at all, while another responds but the message is delayed (indicated by a dashed line), so the reader receives values from four nodes (a majority out of seven).

The key to the correctness of the algorithm is to notice that each operation communicates with a majority of the nodes: since $n > 2f$ it follows that $n-f > n/2$. Thus, there is a common node communicating with each pair of write and read operations. (As illustrated by Figure 1.) This node ensures that the value of the latest preceding (non-overlapping) write operation is forwarded to the later operations.
Figure 2: Old-new inversion between read operations; the dark nodes holds the new value of a write operation. In (a), process $p$ already stores the new value of the register; in (b), the first read operation receives the new value, while in (c), a later read operation does receive the new value. Note that the write operation does not complete.

read operation; the read will pick this value, unless it receives an even later value (with a higher version number).

A formal proof can be found in the original paper or in [13, Chapter 10].

A slight complication happens because two non-overlapping reads, both overlapping a write, might obtain out-of-order values. This can happen if the writer sends a new value (with a higher version number) to all nodes, and node $p$ gets it first. The early read operation might already obtain the new value from $p$ (among the $n - f$ it waits for), while the later read operation does not wait to hear from $p$ and returns the old value. (See Figure 2.)

This behavior is avoided by having a reader write back the value it is going to return, in order to ensure atomicity of reads.

An argument based on communication with a majority of nodes shows that the value returned by a read operation is available to each following read operation, which returns it unless it has obtained an even later value. (See more details in [8] or [13].)

### 3.1 Bounding the version numbers

A large part of the original paper is spent on bounding the version numbers that are appended to each register value that is sent. The key to bounding the version numbers is knowing which of them are currently in use in the system; once this set is known, the writer can make sure that a newly generated sequence number is larger than all existing ones [35]. Tracking the version numbers is complicated by the fact that, although all version numbers are generated by a single node (the writer), they are forwarded with the values exchanged between readers. Track-
ing the version numbers is therefore done by having a reader “record” a version number before forwarding it with a value.

Recording is akin to writing, but it need not preserve atomicity, and hence, can be implemented in a much simpler manner, without forwarding or having to recursively record forwarded values. (See [8] for more details.)

A construction of a single-writer single-reader atomic register [7] avoids this complication, since it needs only to linearize reads from the same node. This reduces the cost of the bounded simulation, even when a multi-reader register is then simulated on top of the single-reader register.

3.2 Immediate Implications

The simulation allowed to port many algorithms designed for shared memory to message-passing systems. This includes atomic snapshots [2], approximate agreement [12], failure detectors [51], and condition-based consensus [44].

This has reduced the interest in the asynchronous message-passing model with crash failures. This model has become virtually “obsolete” when studying computability, as argued for example, by Borowsky and Gafni [16], Herlihy and Shavit [33], as well as Mostefaoui, Rajsbaum and Raynal [44].

3.3 The Quorum Point-of-View

The consistency mechanisms of the ABD simulation can be easily decoupled from its communication pattern. Specifically, the communication pattern of communicating with a majority of the nodes can be replaced with the more general concept of communication with a quorum. This conceptual modification appeared in a paper of Lynch and Shvartsman [39], which extended the unbounded ABD simulation to a multi-writer register, in a more dynamic situation. A similar idea appeared, at about the same time, in work by Malkhi and Reiter [41], dealing with Byzantine failures. (Both works are discussed in more detail in the next section.)

A quorum system is a collection of subsets of nodes, with the property that each pair of sets have at least one common node, that is, each pair of sets have a nonempty intersection. It is quite obvious that the ABD simulation can be paraphrased so that each operation must communicate with a quorum. Indeed, the simple majority quorum system, containing all sets with a majority of nodes, is a straightforward example of a quorum system.

Quorums have been used in the context of data and file replication [27, 29]. These works further separate between write quorums and read quorums, so that every read quorum intersects with every write quorum. A simple example of a read-write quorum system is when nodes are organized in a grid (two-dimensional
array), the read quorums contain all the sets of nodes in the same column, and the write quorums contain all the sets of nodes in the same row.

Clearly, the ABD simulation can be modified so that each write operation (all by the same node) communicates with some write quorum, while each read operation communicates with some read quorum.

This refactoring admits a separation of concerns and paves the way to optimizing the communication pattern without changing the overall workings of the algorithm. For example, it is possible to choose a different quorum system when fewer processes may fail, or so as to optimize the performance features of the quorums, e.g., their load and availability [45].

4 Making the Simulation More Robust

Concentrating on the communication pattern, through a quorum system, allowed to make the register simulation more robust, and most notably, to handle dynamic system changes and tolerate more malicious, Byzantine failures.

4.1 Dynamic Changes

Lynch and Shvartsman [39, 40] address dynamic systems, where nodes can join and leave the system, in their reconfigurable atomic memory service for dynamic networks.

A key concept in these simulations is the notion of a configuration, which includes the set of nodes participating in the simulation, together with the set of read and write quorums. Clearly, when a configuration is fixed, a register can be simulated by running the ABD simulation with the read and write quorums, essentially as described before. Therefore, the core challenge of the algorithm is in reconfiguring the system when changes happen.

Reconfiguration modifies the set of participating nodes, and installs a new quorum system appropriate for the new set. Originally [39], Lynch and Shvartsman relied on a special node to manage reconfiguration. In a later paper [40], they presented a decentralized algorithm, nicknamed RAMBO, in which nodes propose new configurations and use a safe consensus protocol to decide on an agreed new configuration. Safe consensus ensures that nodes agree on the same configuration, which was proposed by one of them, but it does not guarantee termination. (This is necessary since full-fledged consensus cannot be solved in an asynchronous system [23].) A neat feature of the algorithm is that the safe consensus algorithm is implemented from shared registers, which, in turn, are simulated over the message-passing system.
At certain points during the execution of the algorithm, several configurations exist: Some configuration might already be agreed on by some nodes, while other nodes might still hold to previous configurations. A key idea in RAMBO is to communicate with a representative quorum from every known configuration. This leads to a component that needs to “garbage collect” old configurations, so as to reduce the amount of communication after transitional periods.

Improvements to RAMBO have appeared in several works, for example, [17, 21, 28].

Very recently, Aguilera et al. [5] presented a variant of RAMBO that sidesteps the need to reach consensus during reconfiguration, assuming the number of reconfigurations is finite.

4.2 Byzantine Failures

Another interesting research thread deals with Byzantine failures. These failures model the erroneous behavior that is caused by non-deterministic software errors or even malicious attacks.

Malkhi and Rieter [41] considered the behavior of the simulation when some nodes may experience Byzantine failures. When $f$ Byzantine failures may occur, we need to assume that $n > 3f$; otherwise, it is possible to violate the known lower bound on the ratio of Byzantine failures that can be tolerated [22].

In the simplest case, we can take quorums in which any set of $2f + 1$ nodes is a quorum; this ensures that each pair has a large ($> f$) intersection, containing at least one nonfaulty node. When information from correct nodes is self-verifying (e.g., values are digitally signed), this dissemination quorum system suffices for simulating a shared register since this nonfaulty node will forward the correct value.

Malkhi and Rieter also define opaque masking quorum systems, which allow to simulate a shared register without assuming that data values are self-verifying. Such quorum systems assure that a node gets many copies of the current value of the register, all with the same version number. This requires a higher ratio of nonfaulty nodes.

Abraham, Chockler, Keidar and Malkhi [1] present two simulations that only assume $n > 3f$, but provide weaker properties: one simulates only a safe register, while the other simulates a regular register, and is guaranteed to terminate only if the number of writes is finite. Aiyer, Alvisi and Bazzi [6] simulate an atomic register, with Byzantine readers and (up to one-third of) Byzantine servers; their protocol relies on a secret sharing scheme.

Like vanilla quorum systems, it is possible to design more sophisticated dissemination and opaque masking quorum systems, so as to optimize various parameters, e.g., [42].
5 Application: Replicated Services

Many systems cannot afford to guarantee a majority of nonfaulty processing nodes, seemingly implying that fault-tolerance cannot be obtained. However, systems contain other types of components, for example, independent disk drives. Because these components are cheaper than computers, it is feasible to replicate them in order to achieve fault tolerance. Disk drivers are not programmable, but they can respond to client requests; clients may stop taking steps, and disks may stop responding to requests.

Disk Paxos [26] implements an arbitrary fault-tolerant service on such a storage area network containing processing nodes and disks; it provides consistency with any number of asynchronous non-Byzantine node failures.

Disk Paxos is based on a shared-memory version of the classic Paxos algorithm [38]; this algorithm is then ported to a storage area network using an optimized version of the ABD simulation: To write a value, a processor writes the value to a majority of the disks. To read, a processor reads a majority of the disks. (This provides a somewhat weaker register, called regular, which however, suffices for their shared-memory Paxos algorithm.) The algorithm avoids the version numbers used in the ABD simulation by piggybacking on the version numbers of the Paxos algorithm.

A somewhat similar approach was taken with erasure-coded disks [48], where redundancy is used beyond simple replication to tolerate (non-malicious) disk errors, in addition to partitioning and non-responding disks. It incorporates a quorum reconfiguration algorithm, which allows client requests to proceed unimpeded. This algorithm is, in some sense, obstruction-free, since a request may abort when it encounters a concurrent request, yielding an efficient read operation (within a single communication round). This concept was later abstracted by Aguilera et al. to define an abortable object [4], which may abort an operation when there are concurrent operations.

A service can be replicated even when servers are Byzantine, by relying on the register simulation tolerating Byzantine failures, see for example, [47].

6 Closing Remarks

One of my goals in this article was to show how picking the right abstraction can bring forth many applications. Finding the right abstraction is in many ways a key for designing good systems: it should hide enough capabilities "under its hood" to provide good leverage in application design, yet, not too much, so the implementation is efficient (or easily admits optimizations).

The are several other research directions that were not discussed in detail here;
many of them still pose significant challenges.

Many papers studied the response time or the message complexity of the simulation and improved it, especially in the best case, i.e., where the system is well-behaved, by having few failures or maintaining synchronization [19, 32], or by reducing the amount of storage it requires [31]. While there are many pieces of information about the complexity of the simulation [20], it is still open to characterize the exact conditions and scenarios that admit a “fast read” (within a single-communication round) or a “fast write”.

The exact bounds for a simulation tolerating Byzantine failures are also not clear yet: What is the highest ratio of failures that can be tolerated with a constant number of communication rounds, while still providing linearizability and always terminating.

Several papers, e.g., [15, 18, 25], attempt to port the simulation to modern network technologies, which are more ad-hoc and contain mobile nodes, and to sensor networks. The dynamic nature of these systems means these algorithms have to rely on the more robust simulations described in Section 4.1, increasing the importance of developing better understanding of the storage and communication requirements in these systems. These might be traded off with the ratio of faulty nodes that can be tolerated. A related question is to handle a more uniform model, where the number of nodes is unknown in advance.

Needless to say, it is generally unknown how to tolerate Byzantine failures in dynamic systems.

We have described, in brief, how the ABD simulation contributes to service and storage replication, assuming servers might fail. We did not address the problem of dealing with client failures, which may lead to incorrect operations applied to the register; one approach was to consider this as a faulty shared memory [3, 36], and to apply concepts from there [11]. Other approaches deal with faulty clients in a more direct manner, e.g., [43], but this aspect deserves further study.

Acknowledgements: I would to thank Keren Censor and Seth Gilbert for helpful comments.

References


The Bulletin of the EATCS


111


112
The Bulletin of the EATCS


HOW TO BASE PROBABILITY THEORY ON PERFECT-INFORMATION GAMES

Glenn Shafer*  Vladimir Vovk†  Roman Chychyla‡

Abstract

The standard way of making probability mathematical begins with measure theory. This article reviews an alternative that begins with game theory. We discuss how probabilities can be calculated game-theoretically, how probability theorems can be proven and interpreted game-theoretically, and how this approach differs from the measure-theoretic approach.

*Address: Rutgers Business School, 1 Washington Park, Newark, New Jersey 07102, USA, and Computer Learning Research Centre, Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. Email: gshafer@rutgers.edu

†Partially supported by EPSRC grant EP/F002998/1. Address: Computer Learning Research Centre, Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. Email: vovk@cs.rhul.ac.uk

‡Address: Rutgers Business School, 1 Washington Park, Newark, New Jersey 07102, USA. Email: stormydrive@gmail.com
1 Introduction

We can make probability into a mathematical theory in two ways. One begins with measure theory, the other with the theory of perfect-information games. The measure-theoretic approach has long been standard. This article reviews the game-theoretic approach, which is less developed.

In §2, we recall that both measure theory and game theory were used to calculate probabilities long before probability was made into mathematics in the modern sense. In letters they exchanged in 1654, Pierre Fermat calculated probabilities by counting equally possible cases, while Blaise Pascal calculated the same probabilities by backward recursion in a game tree.

In §3, we review the elements of the game-theoretic framework as we formulated it in our 2001 book [22] and subsequent articles. This is the material we are most keen to communicate to computer scientists.

In §4, we compare the modern game-theoretic and measure-theoretic frameworks. As the reader will see, they can be thought of as dual descriptions of the same mathematical objects so long as one considers only the simplest and most classical examples. Some readers may prefer to skip over this section, because the comparison of two frameworks for the same body of mathematics is necessarily an intricate and second-order matter. It is also true that the intricacies of the measure-theoretic framework are largely designed to handle continuous time models, which are of little direct interest to computer scientists. The discussion of open systems in §4.5 should be of interest, however, to all users of probability models.

In §5, we summarize what this article has accomplished and mention some new ideas that have been developed from game-theoretic probability.

We do not give proofs. Most of the mathematical claims we make are proven in [22] or in papers at http://probabilityandfinance.com.

2 Two ways of calculating probabilities

Mathematical probability is often traced back to two French scholars, Pierre Fermat (1601–1665) and Blaise Pascal (1623–1662). In letters exchanged in 1654, they argued about how to do some simple probability calculations. They agreed on the answers, but not on how to derive them. Fermat’s methodology can be regarded as an early form of measure-theoretic probability, Pascal’s as an early form of game-theoretic probability.

Here we look at some examples of the type Pascal and Fermat discussed. In §2.1 we consider a simple case of the problem of points. In §2.2 we calculate the probability of getting two heads in succession before getting two tails in
succession when flipping a biased coin.

2.1 The problem of points

Consider a game in which two players play many rounds, with a prize going to the first to win a certain number of rounds, or points. If they decide to break off the game while lacking different numbers of points to win the prize, how should they divide it?

Suppose, for example, that Peter and Paul are playing for 64 pistoles, Peter needs to win one more round, and Paul needs to win two. If Peter wins the next round, the game is over; Peter gets the 64 pistoles. If Paul wins the next round, then they play another round, and the winner of this second round gets the 64 pistoles. Figure 1 shows Paul’s payoffs for the three possible outcomes: (1) Peter wins the first round, ending the game, (2) Paul wins the first round and Peter wins the second, and (3) Paul wins two rounds.

Figure 1: Paul wins either 0 or 64 pistoles.

If they stop now, Pascal asked Fermat, how should they divide the 64 pistoles? Fermat answered by imagining that Peter and Paul play two rounds regardless of how the first comes out. There are four possible cases:

1. Peter wins the first round, Peter the second. Peter gets the 64 pistoles.
2. Peter wins the first round, Paul wins second. Peter gets the 64 pistoles.
3. Paul wins the first round, Peter the second. Peter gets the 64 pistoles.
4. Paul wins the first round, Paul the second. Paul gets the 64 pistoles.

Paul gets the 64 pistoles in only one of the four cases, Fermat said, so he should get only 1/4 of the 64 pistoles, or 16 pistoles.

Pascal agreed with the answer, 16 pistoles, but not with the reasoning. There are not four cases, he insisted. There are only three, because if Peter wins the first
round, Peter and Paul will not play a second round. A better way of getting the answer, Pascal argued, was to reason backwards in the tree, as shown in Figure 2. After Paul has just won the first round, he has the same right as Peter to the 64 pistoles, and so his position is worth 32 pistoles. Before the first round, then, he is looking at 0 or 32, and this is worth 16.

![Pascal's Backward Recursion](image)

Figure 2: Pascal’s backward recursion, with the game’s value to Paul in each situation.

Pascal and Fermat did not use the word “probability”. But they gave us methods for calculating probabilities. In this example, both methods give $1/4$ as the probability for the event that Paul will win the 64 pistoles.

Fermat’s method is to count the cases where an event $A$ happens and the cases where it fails. If we consider all the cases equally likely, we call the ratio of the number where it happens to the total the event’s probability. This was the classical definition of probability. In the 20th century, it was generalized to a measure-theoretic definition, in which an event is identified with a set and its probability with the measure of the set.

Pascal’s method, in contrast, treats a probability as a price. We assume that the two players have agreed to bet on each round at even odds, regardless of the outcomes of previous rounds or other information. Let $A$ be the event that Paul wins both rounds. We see from Figure 2 that if Paul has 16 pistoles at the beginning, he can bet it in a way that he will have 64 pistoles if $A$ happens, 0 if $A$ fails. (He bets the 16 pistoles on winning the first round, losing it if he loses the round, but doubling it to 32 if he does win, in which case he bets the 32 on winning the second round.) Rescaling so that the prize is 1 rather than 64, we see that $1/4$ is what he needs at the beginning in order to get a payoff equal to 1 if $A$ happens and 0 if $A$ fails. This suggests a general game-theoretic definition of probability for a game in which we are offered opportunities to gamble: the probability of an event is the cost of a payoff equal to 1 if the event happens and 0 if it fails.
2.2 Two heads before two tails

Let us apply Fermat’s and Pascal’s competing methods to a slightly more difficult problem. Suppose we repeatedly flip a coin, with the probability of heads being 1/3 each time (regardless of how previous flips and other events come out). What is the probability we will get two successive heads before we get two successive tails?

Fermat’s combinatorial method is to list the ways the event (two heads before two tails) can happen, calculate the probabilities for each, and add them up. The number of ways we can get two heads before two tails is countably infinite; here are the first few of them, with their probabilities:

- HH: \( \left( \frac{1}{3} \right)^2 \)
- THH: \( \left( \frac{1}{3} \right)^2 \frac{2}{3} \)
- HTHH: \( \left( \frac{1}{3} \right)^3 \frac{2}{3} \)
- THTHH: \( \left( \frac{1}{3} \right)^3 \left( \frac{2}{3} \right)^2 \)
- HTHTHH: \( \left( \frac{1}{3} \right)^4 \left( \frac{2}{3} \right)^2 \)
  etc.

Summing the infinite series, we find that the total probability for two heads before two tails is 5/21.

To get the same answer game-theoretically, we start with the game-theoretic interpretation of the probability 1/3 for a head on a single flip: it is the price for a ticket that pays 1 if the outcome is a head and 0 if it is a tail. More generally, as shown in Figure 3, \( \frac{1}{3}x + \frac{2}{3}y \) is the price for \( x \) if a head, \( y \) if a tail.

![Figure 3: The game-theoretic meaning of probability 1/3 for a head.](image)

119
Let $A$ be the event that there will be two heads in succession before two tails in succession, and consider a ticket that pays 1 if $A$ happens and 0 otherwise. The probability $p$ for $A$ is the price of this ticket at the outset. Suppose now that we have already started flipping the coin but have not yet obtained two heads or two tails in succession. The value of the ticket will have changed, but it should depend only on whether the most recent flip was a head (Situation H) or a tail (Situation T). Let us write $a$ for the value of the ticket in Situation H and $b$ for its value in Situation T.

![Diagram](image)

**Figure 4:** The value of a ticket that pays 1 if $A$ happens and 0 if $A$ fails varies according to the situation.

As we show in Figure 4, a head on the next flip following Situation H would be the second head in succession, and the ticket would pay 1, whereas a tail would put us in Situation T, where the ticket is worth $b$. Applying the rule of Figure 3 to this situation, we get

$$a = \frac{1}{3} + \frac{2}{3}b.$$  

In Situation T, on the other hand, a head puts us in Situation H, and with a tail the ticket pays 0. This gives

$$b = \frac{1}{3}a.$$  

Solving these two equations in $a$ and $b$, we obtain $a = \frac{3}{7}$ and $b = \frac{1}{7}$.

Figure 5 describes the initial situation, before we start flipping the coin. With probability $\frac{1}{3}$, the first flip will put us in a situation where the ticket is worth $\frac{3}{7}$; with probability $\frac{2}{3}$, it will put us in a situation where it is worth $\frac{1}{7}$. So the initial value is

$$p = \frac{1}{3} \cdot \frac{3}{7} + \frac{2}{3} \cdot \frac{1}{7} = \frac{5}{21},$$

in agreement with the combinatorial calculation.
2.3 Why did Pascal and Fermat get the same answers?

We will address a more general version of this question in §4.3, but on this first pass let us stay as close to our two examples as possible. Let us treat both examples as games where we flip a coin, either fair or biased, with a rule for stopping that determines a countable set Ω of sequences of heads and tails as possible outcomes. In our first example, Ω = {H, TH, TT}, where H represents Peter’s winning, and T represents Paul’s winning. In our second example, Ω = {HH, TT, HTT, THH, HTHH, THTT, . . .}. 1

Suppose p is the probability for heads on a single flip. The measure-theoretic approach assigns a probability to each element ω of Ω by multiplying together as many ps as there are Hs in ω and as many (1 − p)s as there are Ts. For example, the probability of HTHH is $p^3(1 - p)$. The probability for a subset A of Ω is then obtained by adding the probabilities for the ω in A.

The game-theoretic approach defines probability differently. Here the probability of A is the initial capital needed in order to obtain a certain payoff at the end of the game: 1 if the outcome ω is in A, 0 if not. To elaborate a bit, consider the capital process determined by a certain initial capital together with a strategy for gambling. Formally, such a capital process is a real-valued function $L$ defined on the set $S$ consisting of the sequences in Ω and all their initial segments, including the empty sequence $\square$. For each $x \in S$, $L(x)$ is the capital the gambler would have right after $x$ happens if he starts with $L(\square)$ and follows the strategy. In the game where we wait for two heads or two tails in succession, for example, $L(HTHT)$ is the capital the gambler would have after HTHT, where the game is not yet over, and $L(HH)$ is the capital he would have after HH, where the game is over. We can rewrite our definition of the probability of A as$^2$

\[ L(\omega) = I_A(\omega) \quad \text{for all } \omega \in \Omega. \]

1To keep things simple, we assume that neither of the infinite sequences HTHHT... or THTHT... will occur.

2In our second example (two heads before two tails), there is more than one capital process $L$ satisfying the condition that $L(\omega) = I_A(\omega)$ for all $\omega \in \Omega$. Strictly speaking, we should say that...
P(A) := L(□), where L is a capital process with

\[ L(\omega) = I_A(\omega) \text{ for all } \omega \in \Omega. \]  \hspace{1cm} (1)

Here \( I_A \) is the indicator function for \( A \), the function on \( \Omega \) equal to 1 on \( A \) and 0 on \( \Omega \setminus A \).

We can use Equation (1) to explain why Pascal's method gives the same answers as Fermat's.

1. If you bet all your capital on getting a head on the next flip, then you multiply it by \( 1/p \) if you get a head and lose it if you get a tail. Similarly, if you bet all your capital on getting a tail on the next flip, then you multiply it by \( 1/(1-p) \) if you get a tail and lose it if you get a head. So Equation (1) gives the same probability to a single path in \( \Omega \) as the measure-theoretic approach. For example, if \( A = \{ \text{HTHH} \} \), we can get the capital \( I_A \) at the end of the game by starting with capital \( p^3(1-p) \), betting it all on H on the first flip, so that we have \( p^2(1-p) \) if we do get H; then betting all this on T on the second flip, so that we have \( p^2 \) if we do get T, and so on, as in Figure 6.

2. We can also see from Equation (1) that the probability for a subset \( A \) of \( \Omega \) is the sum of the probabilities for the individual sequences in \( A \). This is because we can add capital processes. Consider, for example, a doubleton set \( A = \{ \omega_1, \omega_2 \} \), and consider the capital processes \( L_1 \) and \( L_2 \) that appear in Equation (1) for \( A = \{ \omega_1 \} \) and \( A = \{ \omega_2 \} \), respectively. Starting with capital \( P(\{ \omega_1 \}) \) and playing one strategy produces \( L_1 \) with final capital \( I_{\{ \omega_1 \}}(\omega) \) for all \( \omega \in \Omega \), and starting with capital \( P(\{ \omega_2 \}) \) and playing another strategy produces \( L_2 \) with final capital \( I_{\{ \omega_2 \}}(\omega) \) for all \( \omega \in \Omega \). So starting with capital \( P(\{ \omega_1 \}) + P(\{ \omega_2 \}) \) and playing the sum of the two strategies produces the capital process \( L_1 + L_2 \), which has final capital \( I_{\{ \omega_1 \}}(\omega) + I_{\{ \omega_2 \}}(\omega) = I_{\{ \omega_1, \omega_2 \}}(\omega) \) for all \( \omega \in \Omega \).

We can generalize Equation (1) by replacing \( I_A \) with a real-valued function \( \xi \) on \( \Omega \). This gives a formula for the initial price \( \mathbb{E}(\xi) \) of the uncertain payoff \( \xi(\omega) \):

\[ \mathbb{E}(\xi) := L(□), \text{ where } L \text{ is a capital process with} \]

\[ L(\omega) = \xi(\omega) \text{ for all } \omega \in \Omega. \]  \hspace{1cm} (2)

Christian Huygens explained the idea of Equation (2) very clearly in 1657 [14], shortly after he heard about the correspondence between Pascal and Fermat.

\footnote{P(A) is \( L(□) \) for the unique bounded capital process satisfying this condition. We ask the reader’s indulgence to neglect this point for a moment; alternatively, the reader can concentrate on our first example (the problem of points).}

\footnote{This means that we always make both the bets specified by the first strategy and the bets specified by the second strategy.}
The fundamental idea of game-theoretic probability is to generalize Equation (2) as needed to more complicated situations, where there may be more or fewer gambles from which to construct capital processes. Because we cannot count on finding a capital process whose final value will always exactly equal the uncertain payoff $\xi(\omega)$, let alone a unique one, we write instead

$$E(\xi) = \inf \{ L(\square) | L \text{ is a capital process } \& \ L(\omega) \geq \xi(\omega) \text{ for all } \omega \in \Omega \}, \tag{3}$$

and we call $E(\xi)$ the \textit{upper price} of $\xi$.\footnote{As we explain in §3.1, there is a dual and possibly smaller price for $\xi$, called the \textit{lower price}. The difference between the two is somewhat analogous to the bid-ask spread in a financial market.} In games with infinite horizons, where play does not necessarily stop, we consider instead capital processes that equal or exceed $\xi$ asymptotically, and in games where issues of computability or other considerations limit our ability to use all our current capital on each round, we allow some capital to be discarded on each round. But Pascal’s and Huygens’s basic idea remains.

### 3 Elements of game-theoretic probability

Although game-theoretic reasoning of the kind used by Pascal and Huygens never disappeared from probability theory, Fermat’s idea of counting equally likely cases became the standard starting point for the theory in the 19th century and then evolved, in the 20th century, into the measure-theoretic foundation for probability now associated with the names of Andrei Kolmogorov and Joseph Doob \cite{15, 10, 23}. The game-theoretic approach re-emerged only in the 1930s, when Jean Ville used it to improve Richard von Mises’s definition of probability as limiting frequency \cite{18, 19, 28, 1}. Our formulation in 2001 \cite{22} was inspired by Ville’s work and by A. P. Dawid’s work on prequential probability \cite{8, 9} in the 1980s.
Whereas the measure-theoretic framework for probability is a single axiomatic system that has every instance as a special case, the game-theoretic approach begins by specifying a game in which one player has repeated opportunities to bet, and there is no single way of doing this that is convenient for all possible applications. So we begin our exposition with a game that is simple and concrete yet general enough to illustrate the power of the approach. In §3.1, we describe this game, a game of bounded prediction, and define its game-theoretic sample space, its variables and their upper and lower prices, and its events and their upper and lower probabilities. In §3.2, we explain the meaning of upper and lower probabilities. In §3.3, we extend the notions of upper and lower price and probability to situations after the beginning of the game and illustrate these ideas by stating the game-theoretic form of Lévy’s zero-one law. Finally, in §3.4, we discuss how our definitions and results extend to other probability games.

### 3.1 A simple game of prediction

Here is a simple example, borrowed from Chapter 3 of [22], of a precisely specified game in which probability theorems can be proven.

The game has three players: Forecaster, Skeptic, and Reality. They play infinitely many rounds. Skeptic is the one who gambles; Forecaster sets the prices, and Reality determines the outcomes. At the beginning of the game, Skeptic chooses the amount of money he will risk: his initial capital $K_0$. Forecaster begins each round by announcing a number $\mu$, and Reality ends the round by announcing a number $\gamma$. After Forecaster announces $\mu$ and before Reality announces $\gamma$, Skeptic is allowed to buy any number of tickets (even a fractional or negative number), each of which costs $\mu$ and pays back $\gamma$. For simplicity, we require both $\gamma$ and $\mu$ to be in the interval $[0,1]$. Each player hears the others’ announcements as they are made (this is the assumption of perfect information).

**Protocol 1. Bounded prediction**

Skeptic announces $K_0 \in \mathbb{R}$.

FOR $n = 1, 2, \ldots$:

- Forecaster announces $\mu_n \in [0,1]$.
- Skeptic announces $M_n \in \mathbb{R}$.
- Reality announces $\gamma_n \in [0,1]$.

$K_n := K_{n-1} + M_n (\gamma_n - \mu_n)$.

Forecaster’s move $\mu_n$ can be thought of as his estimate or prediction of $\gamma_n$. If Forecaster makes good predictions, Skeptic should not make very much money (see §3.2).

---

5We can change the sets from which $\mu_n$ and $\gamma_n$ are chosen while preserving the idea that $\mu_n$
The Bulletin of the EATCS

There are no probabilities in Protocol 1, only limited opportunities to bet. But we can define prices and probabilities in Pascal’s sense. The following definitions and notation will help.

- A path is a sequence $\mu_1y_2\mu_2y_2\ldots$, where the $\mu$s and $y$s are all in $[0,1]$.
- We write $\Omega$ for the set of all paths, and we call $\Omega$ the sample space.
- An event is a subset of $\Omega$, and a variable is a real-valued function on $\Omega$.
- We call the empty sequence $\Box$ the initial situation.
- We call a sequence of the form $\mu_1y_1\ldots\mu_my_m$ a betting situation.
- We call a sequence of the form $\mu_1y_1\ldots\mu_my_m$ a clearing situation. We write $S$ for the set of all clearing situations. We allow $n = 0$, so that $\Box \in S$.
- A strategy $S_{strat}$ for Skeptic specifies his capital in the initial situation ($K_0$ in $\Box$) and his move $S_{strat}(\mu_1y_1\ldots\mu_{n-1}y_{n-1}\mu_n)$ for every betting situation $\mu_1y_1\ldots\mu_{n-1}y_{n-1}\mu_n$.
- Given a strategy $S_{strat}$ for Skeptic, we define a function $L_{S_{strat}}$ on $S$ by

$$L_{S_{strat}}(\Box) := S_{strat}(\Box)$$

and

$$L_{S_{strat}}(\mu_1y_1\ldots\mu_my_m) := L_{S_{strat}}(\mu_1y_1\ldots\mu_{n-1}y_{n-1}) + S_{strat}(\mu_1y_1\ldots\mu_{n-1}y_{n-1}\mu_n)(y_n - \mu_n).$$

We call $L_{S_{strat}}$ the capital process determined by $S_{strat}$. If Skeptic follows $S_{strat}$, then $L_{S_{strat}}(\mu_1y_1\ldots\mu_my_m)$ is his capital $K_n$ after clearing in the situation $\mu_1y_1\ldots\mu_my_m$.

- We write $\mathcal{L}$ for the set of all capital processes that are bounded below.
- Given $\omega \in \Omega$, say $\omega = \mu_1y_2\mu_2y_2\ldots$, we write $\omega^n$ for the clearing situation $\mu_1y_1\ldots\mu_my_m$.

is a prediction of $y_n$ and staying within the general framework that we will discuss in §3.4. As explained in [22], it is straightforward to replace $[0,1]$ with some other bounded interval, and more complicated games can be formulated to handle unbounded intervals. In §3.2, we discuss the case where the two-point set $\{0,1\}$ replaces the larger set $[0,1]$ as the move space for Reality and perhaps also for Forecaster.

6In [28] and [22], such a capital process is called a martingale.
In the spirit of Equation (3) in §2.3, we say that the upper price of a bounded variable $\xi$ is

$$\mathbb{E}(\xi) := \inf\{L(\Box) | L \in \mathcal{L} \text{ and } \liminf_{n \to \infty} L(\omega^n) \geq \xi(\omega) \text{ for all } \omega \in \Omega\}. \quad (4)$$

We get the same number $\overline{\mathbb{E}}(\xi)$ if we replace the $\liminf$ in (4) by $\limsup$ or $\lim$. In other words,

$$\mathbb{E}(\xi) = \inf\{L(\Box) | L \in \mathcal{L} \text{ and } \limsup_{n \to \infty} L(\omega^n) \geq \xi(\omega) \text{ for all } \omega \in \Omega\} = \inf\{L(\Box) | L \in \mathcal{L} \text{ and } \lim L(\omega^n) \geq \xi(\omega) \text{ for all } \omega \in \Omega\}. \quad (5)$$

(The inequality $\lim_{n \to \infty} L(\omega^n) \geq \xi(\omega)$ means that the limit exists and satisfies the inequality.) For a proof, which imitates the standard proof of Doob’s convergence theorem, see [24]. The essential point is that if a particular strategy for Skeptic produces capital that is sufficient in the sense of $\limsup$ but oscillates on some paths rather than approaching a limit, Skeptic can exploit the successive upward oscillations, thus obtaining a new strategy whose capital tends to infinity (or to a number exceeding $\sup\xi$) on these paths.

If someone from outside the game pays Skeptic $\mathbb{E}(\xi)$ at the beginning of the game, Skeptic can turn it into $\xi(\omega)$ or more at the end of the game. (Here we neglect, for simplicity, the fact that the infimum in (5) may not be attained.) So he can commit to giving back $\xi(\omega)$ at the end of the game without risking net loss. He cannot do this if he charges any less. So $\mathbb{E}(\xi)$ is, in this sense, Skeptic’s lowest safe selling price for $\xi$.

We set $\mathbb{E}(\xi) := -\mathbb{E}(-\xi)$ and call $\mathbb{E}(\xi)$ the lower price of $\xi$. Because selling $-\xi$ is the same as buying $\xi$, $\mathbb{E}(\xi)$ is the highest price at which Skeptic can buy $\xi$ without risking loss.

The names “upper” and “lower” are justified by the fact that

$$\mathbb{E}(\xi) \leq \overline{\mathbb{E}}(\xi). \quad (6)$$

To prove (6), consider a strategy $\text{Sstrat}_1$ that begins with $\mathbb{E}(\xi)$ and returns at least $\xi$ and a strategy $\text{Sstrat}_2$ that begins with $\overline{\mathbb{E}}(-\xi)$ and returns at least $-\xi$. (We again neglect the fact that the infimum in (5) may not be attained.) Then $\text{Sstrat}_1 + \text{Sstrat}_2$ begins with $\mathbb{E}(\xi) + \overline{\mathbb{E}}(-\xi)$ and returns at least 0. This implies that $\mathbb{E}(\xi) + \overline{\mathbb{E}}(-\xi) \geq 0$, because there is evidently no strategy for Skeptic in Protocol 1 that turns a negative initial capital into a nonnegative final capital for sure. But $\mathbb{E}(\xi) + \overline{\mathbb{E}}(-\xi) \geq 0$ is equivalent to $\mathbb{E}(\xi) \leq \overline{\mathbb{E}}(\xi)$.

As we noted in §2.3, probability is a special case of price. We write $\mathbb{P}(A)$ for $\mathbb{E}(I_A)$, where $I_A$ is the indicator function for $A$, and we call it $A$’s upper probability.
Similarly, we write \( \underline{P}(A) \) for \( \underline{E}(I_A) \), and we call it \( A \)'s lower probability. We can easily show that

\[
0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1 \tag{7}
\]

for any event \( A \). The inequality \( \underline{P}(A) \leq \overline{P}(A) \) is a special case of (6). The inequalities \( 0 \leq \overline{P}(A) \) and \( \underline{P}(A) \leq 1 \) are special cases of the general rule that \( \underline{E}(\xi_1) \leq \overline{E}(\xi_2) \) whenever \( \xi_1 \leq \xi_2 \), a rule that follows directly from (4). Notice also that

\[
\overline{P}(A) = 1 - \underline{P}(A^c) \tag{8}
\]

for any event \( A \), where \( A^c := \Omega \setminus A \). This equality is equivalent to \( \overline{E}(I_{A^c}) = 1 + \overline{E}(-I_A) \), which follows from the fact that \( I_{A^c} = 1 - I_A \) and from another rule that follows directly from (4): when we add a constant to a variable \( \xi \), we add the same constant to its upper price.

If \( \underline{E}(\xi) = \underline{E}(\xi) \), then we say that \( \xi \) is priced; we write \( \overline{E}(\xi) \) for the common value of \( \underline{E}(\xi) \) and \( \overline{E}(\xi) \) and call it \( \xi \)'s price. Similarly, if \( \overline{P}(A) = \overline{P}(A) \), we write \( \overline{P}(A) \) for their common value and call it \( A \)'s probability. This is the game-theoretic definition of probability.

### 3.2 The interpretation of upper and lower probabilities

According to the 19th century philosopher Augustin Cournot, as well as many later scholars [20], a probabilistic theory makes contact with the world only by predicting that events assigned very high probability will happen. Equivalently, those assigned very low probability will not happen.

In the case where we have only upper and lower probabilities rather than probabilities, we can make these predictions:

1. If \( \underline{P}(A) \) is equal or close to one, \( A \) will happen.
2. If \( \overline{P}(A) \) is equal or close to zero, \( A \) will not happen.

It follows from (8) that Conditions 1 and 2 are equivalent. We see from (7) that these conditions are consistent with Cournot’s principle. When \( \underline{P}(A) \) is exactly one, \( \overline{P}(A) \) is also one, and \( A \) has probability one according to our game-theoretic definition. When \( \underline{P}(A) \) is close to one, \( \overline{P}(A) \) is also close to one, and so it is reasonable to say that \( A \) nearly has probability one. Similarly, when \( \overline{P}(A) \) is close to zero, we may say that \( A \) has probability equal or close to zero.

In order to see more clearly the meaning of game-theoretic probability equal or close to zero, let us write \( \mathcal{L}^+ \) for the subset of \( \mathcal{L} \) consisting of capital processes that are nonnegative—i.e., satisfy \( L(\omega^n) \geq 0 \) for all \( \omega \in \Omega \) and \( n \geq 0 \). We can then write

\[
\overline{P}(A) := \inf \{ L(\square) \mid L \in \mathcal{L}^+ \text{ and } \lim_{n \to \infty} L(\omega^n) \geq 1 \text{ for all } \omega \in A \}. \tag{9}
\]
When \( P(A) \) is very close to zero, (9) says that Skeptic has a strategy that will multiply the capital it risks by a very large factor \((1/L(\omega))\) if \( A \) happens. (The condition that \( L(\omega') \) is never negative means that only the small initial capital \( L(\omega) \) is being put at risk.) If Forecaster does a good job of pricing the outcomes chosen by Reality, Skeptic should not be able to multiply the capital he risks by a large factor. So \( A \) should not happen.

If an event has lower probability exactly equal to one (and hence probability one), we say that it happens almost surely. Here are two events that happen almost surely in Protocol 1:

- The subset \( A_1 \) of \( \Omega \) consisting of all sequences \( \mu_1 y_1 \mu_2 y_2 \ldots \) such that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (y_i - \mu_i) = 0. 
\]

The assertion that \( A_1 \) happens almost surely is proven in Chapter 3 of [22]. This is a version of the strong law of large numbers: in the limit, the average of the outcomes will equal the average of the predictions.

- The subset \( A_2 \) of \( \Omega \) consisting of all sequences \( \mu_1 y_1 \mu_2 y_2 \ldots \) such that if \( \lim_{n \to \infty} |J_n, a, b| = \infty \), where \( a \) and \( b \) are rational numbers and \( J_n, a, b \) is the set of indices \( i \) such that \( 0 \leq i \leq n \) and \( a \leq \mu_i \leq b \), then

\[
a \leq \lim \inf_{n \to \infty} \frac{\sum_{i \in J_n, a, b} y_i}{|J_n, a, b|} \quad \text{and} \quad \lim \sup_{n \to \infty} \frac{\sum_{i \in J_n, a, b} y_i}{|J_n, a, b|} \leq b.
\]

The assertion that \( A_2 \) happens almost surely is an assertion of calibration: in the limit, the average of the outcomes for which the predictions are in a given interval will also be in that interval. See [36].

In [22], we also give examples of events in Protocol 1 that have lower probability close to one but not exactly equal to one. One such event, for example, is the event, for a large fixed value of \( N \), that \( \frac{1}{N} \sum_{i=1}^{N} (y_i - \mu_i) \) is close to zero. The assertion that this event will happen is a version of Bernoulli’s theorem, sometimes called the weak law of large numbers.

The almost sure predictions we make (\( A \) will happen when \( P(A) = 1 \), and \( A \) will not happen when \( P(A) = 0 \)) will be unaffected if we modify the game by restricting the information or choices available to Skeptic’s opponents. If Skeptic has a winning strategy in a given game, then he will still have a winning strategy when his opponents are weaker. Here are some interesting ways to weaken Skeptic’s opponents in Protocol 1:
• **Probability forecasting.** Require Reality to make each $y_n$ equal to 0 or 1. Then $\mu_n$ can be interpreted as Forecaster’s probability for $y_n = 1$, and the strong law of large numbers, (10), says that the frequency of 1s gets ever closer to the average probability.

• **Exact forecasting.** Require both Forecaster and Reality to choose from the two-point set $\{0, 1\}$ instead of the larger set $[0, 1]$. In this case, $\mu_n$ can be interpreted as Forecaster’s attempt to guess $y_n$ exactly. Skeptic has an obvious strategy (choose $M_n$ very large when $\mu_n = 0$; choose $M_n$ negative but very large in absolute value when $\mu_n = 1$) that makes an arbitrarily large amount of money without risk whenever $y_n \neq \mu_n$ for some $n$. So $y_n = \mu_n$ for all $n$ almost surely.

• **Fixing the probabilities.** Require Forecaster to follow some strategy known in advance to the other players. He might be required, for example, to make all the $\mu_n$ equal to $1/2$. In this case, assuming that Reality is also required to set each $y_n$ equal to 0 or 1, we have the familiar picture of the fair coin, and (10) says that the frequency of 1s will converge to $1/2$.

• **Requiring Reality’s neutrality.** Prevent Reality from playing strategically. This can be done by hiding the other players’ moves from Reality, or perhaps by requiring that Reality play randomly (whatever we take this to mean).

Preventing strategic play by Reality makes Protocol 1 better resemble conventional conceptions of probability, but neither this nor other ways of weakening Skeptic’s opponents invalidate any theorems we can prove in the protocol about upper probabilities being small or about lower probabilities being large. These theorems assert that Skeptic has a strategy that achieves certain goals regardless of his opponents’ moves. Additional assumptions about how his opponents move (stochastic models for their behavior, for example) might enable us to prove that Skeptic can accomplish even more, perhaps raising some lower prices or lowering some upper prices, but they will not invalidate any conclusions about what happens almost surely or with high probability.

It is also noteworthy that the almost sure predictions will not be affected if some or all of the players receive additional information in the course of the game. If Skeptic can achieve a certain goal regardless of how the other players move, then it makes no difference if they have additional information on which to base their moves. We will comment on this point further in §4.5.

The framework also applies to cases where Forecaster’s moves $\mu_n$ and Reality’s moves $y_n$ are the result of the interaction of many agents and influences. One such case is that of a market for a company’s stock, $\mu_n$ being the opening price of
the stock on day $n$, and $y_n$ its closing price. In this case, Skeptic plays the role of a day trader who decides how many shares to hold after seeing the opening price. Our theorems about what Skeptic can accomplish will hold regardless of the complexity of the process that determines $\mu_n$ and $y_n$. In this case, the prediction that $A$ will not happen if $P(A)$ is very small can be called an efficient market hypothesis.

Before leaving the topic of interpretation, we should acknowledge that games played following protocols similar to Protocol 1 do have uses that do not involve Cournot’s principle. One example, very important in finance, is option pricing. In many cases, there are strategies for Skeptic (“hedging strategies”) that replicate given variables (“options”) regardless of whether the market is efficient (i.e., whether events with small upper probability happen). One simple example is that of a forward contract [25].

### 3.3 Price and probability in a situation

We have defined upper and lower prices and probabilities for the initial situation, but the definitions can easily be adapted to later situations. Given a situation $s$ let us write $\Omega(s)$ for the set of paths for which $s$ is a prefix. Then a bounded variable $\xi$’s upper price in the situation $s$ is

$$E(\xi | s) := \inf \{ L(s) | L \in \mathcal{L} \text{ and } \lim_{n \to \infty} L(\omega^n) \geq \xi(\omega) \text{ for all } \omega \in \Omega(s) \}. $$

This definition can be applied both when $s$ is a betting situation and when $s$ is a clearing situation.

We may define $E(\xi | s)$, $E(A | s)$, and $E(A | s)$ in terms of $E(\xi | s)$, just as we have defined $E(\xi)$, $E(A)$, and $E(A)$ in terms of $E(\xi)$. We will not spell out the details. Notice that $E(\xi)$, $E(\xi)$, $E(A)$, and $E(A)$ are equal to $E(\xi | \Box)$, $E(\xi | \Box)$, $E(A | \Box)$, and $E(A | \Box)$, respectively.

In [24], we show that if the upper and lower prices for a bounded variable $\xi$ are equal, then this remains true almost surely in later situations: if $E(\xi) = E(\xi)$, then $E(\xi | \omega^n) = E(\xi | \omega^n)$ for all $n$ almost surely.

The game-theoretic concepts of probability and price in a situation are parallel to the concepts of conditional probability and expected value in classical probability theory.\(^7\) As an illustration of the parallelism, consider Paul Lévy’s zero-one law [17]. In its classical form, it says that if an event $A$ is determined by a sequence $X_1, X_2, \ldots$ of variables, its conditional probability given the first $n$ of these variables tends almost surely, as $n$ tends to infinity, to one if $A$ happens and to

---

\(^7\)In this paragraph, we assume that the reader has some familiarity with the concepts of conditional probability and expected value, even if they are not familiar with the measure-theoretic formalization of the concept that we will review briefly in §4.1.
zero if $A$ fails [3]. More generally, if a bounded variable $\xi$ is determined by $X_1, X_2, \ldots$, the conditional expected value of $\xi$ given the first $n$ of the $X_i$ tends to $\xi$ almost surely. In [24], we illustrate the game-theoretic concepts of price and probability in a situation by proving the game-theoretic version of this law. In its most general form, the game-theoretic version says that

$$\lim_{n \to \infty} \mathbb{E}(\xi | \omega^n) \geq \xi(\omega)$$

(11)

almost surely. If $\xi$’s initial upper and lower prices are equal, so that its upper and lower prices are also equal in later situations almost surely, we can talk simply of its price in situation $s$, $\mathbb{E}(\xi | s)$, and by applying (11) to $\xi$ and $-\xi$, we obtain that

$$\lim_{n \to \infty} \mathbb{E}(\xi | \omega^n) = \xi(\omega)$$

(12)

almost surely. This is Lévy’s zero-one law in its game-theoretic form.

3.4 Other probability games

The game-theoretic results we have discussed apply well beyond the simple game of prediction described by Protocol 1. They hold for a wide class of perfect-information games in which Forecaster offers Skeptic gambles, Skeptic decides which gambles to make, and Reality decides the outcomes.

Let us assume, for simplicity, that Reality chooses her move from the same space, say $Y$, on each round of the game. Then a gamble for Skeptic can be specified by giving a real-valued function $f$ on $Y$: if Skeptic chooses the gamble $f$ and Reality chooses the outcome $y$, then Skeptic’s gain on the round of play is $f(y)$. Forecaster’s offer on each round will be a set of real-valued functions on $Y$ from which Skeptic can choose.

Let us call a set $\mathcal{C}$ of real-valued functions on a set $Y$ a pricing cone on $Y$ if it satisfies the following conditions:

1. If $f_1 \in \mathcal{C}$, $f_2$ is a real-valued function on $Y$, and $f_2 \leq f_1$, then $f_2 \in \mathcal{C}$.

8For those not familiar with Lévy’s zero-one law, here is a simple example of its application to the problem of the gambler’s ruin. Suppose a gambler plays many rounds of a game, losing or winning 1 pistole on each round. Suppose he wins each time with probability $2/3$, regardless of the outcomes of preceding rounds, and suppose he stops playing only if and when he goes bankrupt (loses all his money). A well known calculation shows that when he has $k$ pistoles, he will eventually lose it all with probability $(1/2)^k$. Suppose he starts with 1 pistole, and let $Y(n)$ be the number of pistoles he has after round $n$. Then his probability of going bankrupt is equal to $1/2$ initially and to $(1/2)^{Y(n)}$ after the $n$th round. Lévy’s zero-one law, applied to the event $A$ that the gambler goes bankrupt, says that with probability one, either he goes bankrupt, or else $(1/2)^{Y(n)}$ tends to zero and hence $Y(n)$ tends to infinity. The probability that $Y(n)$ oscillates forever, neither hitting 0 nor tending to infinity, is zero.
2. If \( f \in \mathcal{C} \) and \( c \in [0, \infty) \), then \( cf \in \mathcal{C} \).
3. If \( f_1, f_2 \in \mathcal{C} \), then \( f_1 + f_2 \in \mathcal{C} \).
4. If \( f_1, f_2, \ldots \in \mathcal{C} \), \( f_1(y) \leq f_2(y) \leq \cdots \) for all \( y \in \mathcal{Y} \), and \( \lim_{n \to \infty} f_n(y) = f(y) \) for all \( y \in \mathcal{Y} \), where \( f \) is a real-valued function on \( \mathcal{Y} \), then \( f \in \mathcal{C} \).
5. If \( f \in \mathcal{C} \), then there exists \( y \in \mathcal{Y} \) such that \( f(y) \leq 0 \).

Let us write \( \mathcal{C}_\mathcal{Y} \) for the set of all pricing cones on \( \mathcal{Y} \).

If we require Skeptic to offer a pricing cone on each round of the game, then our protocol has the following form:

**Protocol 2. General Prediction**

**Parameter:** Reality’s move space \( \mathcal{Y} \)

Skeptic announces \( \mathcal{X}_0 \in \mathbb{R} \).

FOR \( n = 1, 2, \ldots \):

- Forecaster announces \( \mathcal{C}_n \in \mathcal{C}_\mathcal{Y} \).
- Skeptic announces \( f_n \in \mathcal{C}_n \).
- Reality announces \( y_n \in \mathcal{Y} \).

\( \mathcal{X}_n := \mathcal{X}_{n-1} + f_n(y_n) \).

The probability games studied in [22] and in the subsequent working papers at [http://probabilityandfinance.com](http://probabilityandfinance.com) are all essentially of this form, although sometimes Forecaster or Reality are further restricted in some way. As we explained in §3.2, our theorems state that Skeptic has a strategy that accomplishes some goal, and such theorems are not invalidated if we give his opponents less flexibility. We may also alter the rules for Skeptic, giving him more flexibility or restricting him in a way that does not prevent him from following the strategies that accomplish his goals.

In the case of Protocol 1, the outcome space \( \mathcal{Y} \) is the interval \([0, 1]\). Forecaster’s move is a number \( \mu \in [0, 1] \), and Skeptic is allowed to choose any payoff function \( f \) that is a multiple of \( y - \mu \). It will not invalidate our theorems to allow him also to choose any payoff function that always pays this much or less, so that his choice is from the set

\[
\mathcal{C} = \mathcal{C}(\mu) = \{ f : [0, 1] \to \mathbb{R} | \text{there exists } M \in \mathbb{R} \text{ such that } f(y) \leq M(y - \mu) \text{ for all } y \in [0, 1] \}.
\]

This is a pricing cone; Conditions 1–5 are easy to check. So we have an instance of Protocol 2.

As we have just seen, Condition 1 in our definition of a pricing cone (the requirement that \( f_2 \in \mathcal{C} \) when \( f_1 \in \mathcal{C} \) and \( f_2 \leq f_1 \)) is of secondary importance;
we can often make it true by adding gambles that Skeptic will not want to take. Conditions 2 and 3 are more essential; they express the linearity of probabilistic pricing. Condition 4 plays the same role as countable additivity (sometimes called continuity) in measure-theoretic probability; it is needed for limiting arguments such as the ones used to prove the strong law of large numbers. Condition 5 is the condition of coherence; it rules out sure bets for Skeptic.

At first glance, it might appear that Protocol 2 might be further generalized by allowing Reality’s move space to vary from round to round. This would not be a substantive generalization, however. If Reality is required to choose from a set $Y_n$ on the $n$th round, then we can recover the form of Protocol 2 by setting $Y$ equal to the union of the $Y_n$; the fact that Reality is restricted on each round to some particular subset of the larger set $Y$ does not, as we noted, invalidate theorems about what Skeptic can accomplish.

4 Contrasts with measure-theoretic probability

For the last two hundred years at least, the mainstream of probability theory has been measure-theoretic rather than game-theoretic. We need to distinguish, however, between classical probability theory, developed during the nineteenth and early twentieth centuries, and the more abstract measure-theoretic framework, using $\sigma$-algebras and filtrations, that was developed in the twentieth century, in large part by Kolmogorov [15] and Doob [10]. Classical probability theory, which starts with equally likely cases and combinatorial reasoning as Fermat did and extends this to continuous probability distributions using the differential and integral calculus, is measure-theoretic in a broad sense. The more abstract Kolmogorov-Doob framework is measure-theoretic in narrower sense: it uses the modern mathematical theory of measure.

Although there is a strong consensus in favor of the Kolmogorov-Doob framework among mathematicians who work in probability theory per se, many users of probability in computer science, engineering, statistics, and the sciences still work with classical probability tools and have little familiarity with the Kolmogorov-Doob framework. So we provide, in §4.1, a concise review of the Kolmogorov-Doob framework. Readers who want to learn more have many excellent treatises, such as [2, 26], from which to choose. For additional historical perspective on the contributions of Kolmogorov and Doob, see [23, 13].

In §4.2 and §4.3, we discuss some relationships between the game-theoretic and measure-theoretic pictures. As we will see, these relationships are best described not in terms of the abstract Kolmogorov-Doob framework but in terms of the concept of a forecasting system. This concept, introduced by A. P. Dawid in 1984, occupies a position intermediate between measure theory and game the-
ory. A forecasting system can be thought of as a special kind of strategy for Forecaster, which always gives definite probabilities for Reality’s next move. The Kolmogorov-Doob framework, in contrast, allows some indefiniteness, inasmuch as its probabilities in new situations can be changed arbitrarily on any set of paths of probability zero. The game-theoretic framework permits a different kind of indefiniteness; it allows Forecaster to make betting offers that determine only upper and lower probabilities for Reality’s next move. In §4.2, we discuss how the game-theoretic picture reduces to a measure-theoretic picture when we impose a forecasting system on Forecaster. In §4.3, we discuss the duality between infima from game-theoretic capital processes and suprema from forecasting systems.

In §4.4, we discuss how continuous time can be handled in the game-theoretic framework. In §4.5, we point out how the open character of the game-theoretic framework allows a straightforward use of scientific theories that make predictions only about some aspects of an observable process.

4.1 The Kolmogorov-Doob framework

The basic object in Kolmogorov’s picture [15, 23] is a probability space, which consists of three elements:

1. A set $\Omega$, which we call the sample space.

2. A $\sigma$-algebra $\mathcal{F}$ on $\Omega$ – i.e., a set of subsets of $\Omega$ that contains $\Omega$ itself, contains the complement $\Omega \setminus A$ whenever it contains $A$, and contains the intersection and union of any countable set of its elements.

3. A probability measure $P$ on $\mathcal{F}$ – i.e., a mapping from $\mathcal{F}$ to $[0, \infty)$ that satisfies

   (a) $P(\Omega) = 1$,
   (b) $P(A \cup B) = P(A) + P(B)$ whenever $A, B \in \mathcal{F}$ and $A \cap B = \emptyset$, and
   (c) $P(\bigcap_{i=1}^{\infty} A_i) = \lim_{i \to \infty} P(A_i)$ whenever $A_1, A_2, \ldots \in \mathcal{F}$ and $A_1 \supseteq A_2 \supseteq \cdots$.

Condition (c) is equivalent, in the presence of the other conditions, to countable additivity: if $A_1, A_2, \ldots$ are pairwise disjoint elements of $\mathcal{F}$, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Only subsets of $\Omega$ that are in $\mathcal{F}$ are called events. An event $A$ for which $P(A) = 1$ is said to happen almost surely or for almost all $\omega$.

A real-valued function $\xi$ on the sample space $\Omega$ that is measurable (i.e., $\{\omega \in \Omega \mid \xi(\omega) \leq a\} \in \mathcal{F}$ for every real number $a$) is called a random variable. If the Lebesgue integral of $\xi$ with respect to $P$ exists, it is called $\xi$’s expected value and is denoted by $E_P(\xi)$. 

134
We saw examples of probability spaces in §2. In the problem of two heads before two tails, \( \Omega = \{\text{HH, TT, HTT, THH, HTTH, THTT, \ldots}\} \), and we can take \( \mathcal{F} \) to be the set of all subsets of \( \Omega \). We defined the probability for an element \( \omega \) of \( \Omega \) by multiplying together as many \( p \)s as there are Hs in \( \omega \) and as many \((1 - p)\)s as there are Ts, where \( p \) is the probability of getting a head on a single flip. We then defined the probability for a subset of \( \Omega \) by adding the probabilities for the elements of the subset.

In general, as in this example, the axiomatic properties of the probability space \((\Omega, \mathcal{F}, P)\) make no reference to the game or time structure in the problem. Information about how the game unfolds in time is hidden in the identity of the elements of \( \Omega \) and in the numbers assigned them as probabilities.

Doob [10] suggested bringing the time structure back to the axiomatic level by adding what is now called a filtration to the basic structure \((\Omega, \mathcal{F}, P)\). A filtration is a nested family of \( \sigma \)-algebras, one for each point in time. The \( \sigma \)-algebra \( \mathcal{F}_t \) for time \( t \) consists of the events whose happening or failure is known at time \( t \). We assume that \( \mathcal{F}_t \subseteq \mathcal{F} \) for all \( t \), and that \( \mathcal{F}_t \subseteq \mathcal{F}_u \) when \( t \leq u \); what is known at time \( t \) is still known at a later time \( u \). The time index \( t \) can be discrete (say \( t = 0, 1, 2, \ldots \)) or continuous (say \( t \in [0, \infty) \) or \( t \in \mathbb{R} \)).

Kolmogorov and Doob used the Radon-Nikodym theorem to represent the idea that probabilities and expected values change with time. This theorem implies that if \( \xi \) is a random variable in \((\Omega, \mathcal{F}, P)\), \( \mathbb{E}_P(\xi) \) exists and is finite, and \( \mathcal{G} \) is a \( \sigma \)-algebra contained in \( \mathcal{F} \), then there exists a random variable \( \zeta \) that is measurable with respect to \( \mathcal{G} \) and satisfies

\[
\mathbb{E}_P(\xi I_A) = \mathbb{E}_P(\zeta I_A)
\]

for all \( A \in \mathcal{G} \). This random variable is unique up to a set of probability zero: if \( \zeta_1 \) and \( \zeta_2 \) are both measurable with respect to \( \mathcal{G} \) and \( \mathbb{E}_P(\xi I_A) = \mathbb{E}_P(\zeta_1 I_A) = \mathbb{E}_P(\zeta_2 I_A) \) for all \( A \in \mathcal{G} \), then the event \( \zeta_1 \neq \zeta_2 \) has probability zero. We write \( \mathbb{E}_P(\xi | \mathcal{G}) \) for any version of \( \zeta \), and we call it the conditional expectation of \( \xi \) given \( \mathcal{G} \).

In the case where each element \( \omega \) of \( \Omega \) is a sequence, and we learn successively longer initial segments \( \omega_1, \omega_2, \ldots \) of \( \omega \), we may use the discrete filtration \( \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \), where \( \mathcal{F}_n \) consists of all the events in \( \mathcal{F} \) that we know to have happened or to have failed as soon as we know \( \omega^n \). In other words,

\[
\mathcal{F}_n := \{ A \in \mathcal{F} \mid \text{if } \omega_1 \in A \text{ and } \omega_2 \notin A, \text{ then } \omega^n_1 \neq \omega^n_2 \}.
\]

It is also convenient to assume that \( \mathcal{F} \) is the smallest \( \sigma \)-algebra containing all the \( \mathcal{F}_n \). In this case, the measure-theoretic version of Lévy’s zero-one law says that for any random variable \( \xi \) that has a finite expected value \( \mathbb{E}_P(\xi) \),

\[
\lim_{n \to \infty} \mathbb{E}_P(\xi | \mathcal{F}_n)(\omega) = \xi(\omega)
\]
This is similar to the game-theoretic version of the law, Equation (12) in §3.3:

$$\lim_{n \to \infty} \mathbb{E}(\xi | \omega^n) = \xi(\omega)$$

almost surely. In the game-theoretic version, it is explicit in the notation that \( \mathbb{E}(\xi | \omega^n) \) depends on \( \omega \) only through what is known at the end of round \( n \), namely \( \omega^n \). In the measure-theoretic version, we know that the value \( \mathbb{E}_P(\xi | \mathcal{F}_n)(\omega) \) of the random variable \( \mathbb{E}_P(\xi | \mathcal{F}_n) \) depends on \( \omega \) only through \( \omega^n \) because this random variable is measurable with respect to \( \mathcal{F}_n \).

There are additional differences between the measure-theoretic and game-theoretic concepts. In the game-theoretic picture, a variable \( \xi \) may have only upper and lower prices, \( \mathbb{E}(\xi | s) \) and \( \mathbb{E}(\xi | s) \), but these are well defined even if the probability of arriving in the situation \( s \) was initially zero. Moreover, in the special case where upper and lower prices are equal, they behave as expected values are supposed to behave: \( \mathbb{E}(\xi_1 + \xi_2 | s) = \mathbb{E}(\xi_1 | s) + \mathbb{E}(\xi_2 | s) \), etc. In contrast, the measure-theoretic quantity \( \mathbb{E}_P(\xi | \mathcal{F}_n)(\omega) \) is undefined (i.e., can be chosen arbitrarily) if \( \omega^n \) has initial probability zero, and the abstract definition (13) does not guarantee that the quantities \( \mathbb{E}_P(\xi | \mathcal{F}_n)(\omega) \) will behave like expected values when \( \omega \) is fixed and \( \xi \) is varied, or even that they can be chosen so that they do so for all \( \omega \).

The extent to which conditional expectations can fail to behave like expected values was a matter of some consternation when it was discovered in the 1940s and 1950s [23]. But in the end, the awkward aspects of the concept of conditional expectation have been tolerated, because the measure-theoretic framework is very general, applying to continuous as well as discrete time, and the usefulness of its theorems for sensible probability models is not harmed by the existence of less attractive models that also satisfy its axioms.

### 4.2 Forecasting systems

In many applications of probability to logic and computer science, we consider an infinite sequence of 0s and 1s. If we write \( \mu(y_1 \ldots y_n) \) for the probability that the sequence will start with \( y_1 \ldots y_n \), then we should have:

- \( 0 \leq \mu(y_1 \ldots y_n) \leq 1 \), and

\[ \text{for almost all } \omega. \]

If we do not assume that \( \mathcal{F} \) is the smallest \( \sigma \)-algebra containing the \( \mathcal{F}_n \), then we can say only that \( \lim_{n \to \infty} \mathbb{E}(\xi | \mathcal{F}_n)(\omega) = \mathbb{E}(\xi | \mathcal{F}_\infty)(\omega) \) for almost all \( \omega \), where \( \mathcal{F}_\infty \) is the smallest \( \sigma \)-algebra containing the \( \mathcal{F}_n \). Lévy’s own statement of his law, first published in 1937 [17], was simpler: He wrote that \( \lim_{n \to \infty} E_n(\xi) = \xi \) almost surely, where \( E_n(\xi) \) is \( \xi \)’s expected value after \( \omega_1 \ldots \omega_n \) are known. Lévy had his own theory of conditional probability and expected value, slightly different from Kolmogorov’s [3].
The Bulletin of the EATCS

- \( \mu(y_1 \ldots y_n) = \mu(y_1 \ldots y_n0) + \mu(y_1 \ldots y_n1) \)

for all finite sequences \( y_1 \ldots y_n \) of zeroes and ones. Let us call a function \( \mu \) satisfying these two rules a binary probability distribution.

Standard expositions of the Kolmogorov-Doob framework show how to construct a probability space \((\Omega, \mathcal{F}, P)\) from a binary probability distribution \( \mu \):

- \( \Omega \) is the set of all infinite sequences of zeroes and ones: \( \Omega = \{0, 1\}^\infty \).
- \( \mathcal{F} \) is the smallest \( \sigma \)-algebra of subsets of \( \Omega \) that includes, for every finite sequence \( y_1 \ldots y_n \) of zeroes and ones, the set consisting of all \( \omega \in \Omega \) that begin with \( y_1 \ldots y_n \).
- \( P \) is the unique probability measure on \( \mathcal{F} \) that assigns, for every finite sequence \( y_1 \ldots y_n \) of zeroes and ones, the probability \( \mu(y_1 \ldots y_n) \) to the set consisting of all \( \omega \in \Omega \) that begin with \( y_1 \ldots y_n \).

Given a random variable \( \xi \) in the probability space \( (\Omega, \mathcal{F}, P) \) constructed in this way, let us write \( \mathbb{E}_\mu(\xi) \) instead of \( \mathbb{E}_P(\xi) \) for its expected value.

Let us call a binary probability distribution \( \mu \) positive if \( \mu(y_1 \ldots y_n) \) is always strictly positive. In this case, conditional probabilities for \( y_n \) given the preceding values \( y_1 \ldots y_{n-1} \) are well defined. Let us write \( \mu_{y_1 \ldots y_{n-1}}(y_n) \) for these conditional probabilities:

\[
\mu_{y_1 \ldots y_{n-1}}(y_n) := \frac{\mu(y_1 \ldots y_{n-1}y_n)}{\mu(y_1 \ldots y_{n-1})} \quad (14)
\]

for any sequence \( y_1 \ldots y_n \) of zeroes and ones.

Now consider the variation on Protocol 1 where Reality must choose each of her moves \( y_n \) from \( \{0, 1\} \) (rather than from the larger set \( \{0, 1\} \)). In this case, Forecaster’s move \( \mu_n \) can be thought of as Forecaster’s probability, after he has seen \( y_1 \ldots y_n \), that Reality will set \( y_n \) to equal 1. This thought points to how Forecaster can use a positive binary probability distribution \( \mu \) as a strategy in the game: he sets his move \( \mu_n \) equal to \( \mu_{y_1 \ldots y_{n-1}}(1) \). If we assume that Forecaster plays this strategy, then we can replace him by the strategy in the protocol, reducing it to the following:

**Protocol 3. Using a positive binary probability distribution as a strategy for probability prediction**

**Parameter:** Positive binary probability distribution \( \mu \)

- Skeptic announces \( K_0 \in \mathbb{R} \).
- FOR \( n = 1, 2, \ldots : \)
  - Skeptic announces \( M_n \in \mathbb{R} \).
  - Reality announces \( y_n \in \{0, 1\} \).
  - \( K_n := K_{n-1} + M_n(y_n - \mu_{y_1 \ldots y_{n-1}}(1)) \).
The sample space for this protocol is the space we just discussed: $\Omega = \{0, 1\}^\infty$. The upper price in this protocol of a bounded variable, if it is measurable, is the same as its expected value in $(\Omega, \mathcal{F}, P)$ ([22], Proposition 8.5).

In the case of a binary probability distribution $\mu$ that is not positive, the denominator in Equation (14) will sometimes be zero, and so $\mu$ will not determine a strategy for Forecaster in our game. To avoid this difficulty, it is natural to replace the concept of a binary probability distribution with the concept of a forecasting system, which gives directly the required probabilities $\mu_{y_1, \ldots, y_n}(y_n)$. A binary probability distribution $\mu$ can be constructed from such a system:

$$
\mu(y_1 \ldots y_n) := \mu_{y_1}(y_2) \cdot \cdots \cdot \mu_{y_1, \ldots, y_{n-1}}(y_n).
$$

If $\mu_{y_1, \ldots, y_{n-1}}(y_n) = 0$ for some $y_1 \ldots y_{n-1}y_n$, then the forecasting system carries more information than the binary probability distribution.

The concept of a forecasting system generalizes beyond probability prediction (the variation on Protocol 1 where the $y_n$ are all either zero or one) to Protocol 2. Fix a $\sigma$-algebra $\mathcal{G}$ on Reality’s move space $\mathcal{Y}$, and write $P_{\mathcal{G}}$ for the set of all probability measures on $(\mathcal{Y}, \mathcal{G})$. Write $\mathcal{Y}^*$ for the set of all finite sequences of elements of $\mathcal{Y}$. In symbols: $\mathcal{Y}^* := \cup_{n=0}^\infty \mathcal{Y}^n$. Then a forecasting system is a mapping $\mu$ from $\mathcal{Y}^*$ to $P_{\mathcal{G}}$ that is measurable in an appropriate sense. Such a system $\mu$ determines a measure-theoretic object on the one hand and game-theoretic object on the other:

- It determines a probability measure $P$ on the sample space $\mathcal{Y}^\infty$, and in each later situation a probability measure whose expected values form conditional expectations with respect to $P$ and that situation.

- It determines a strategy for Forecaster in the protocol: in the situation $y_1 \ldots y_n$, Forecaster announces the pricing cone consisting of every real-valued function $g$ on $\mathcal{Y}$ such that $f \leq g$ for some random variable $g$ on $(\mathcal{Y}, \mathcal{G})$ such that $E_{\mu_{y_1, \ldots, y_n}}(g) \leq 0$.

The two objects agree on global pricing: the game-theoretic upper price of a bounded random variable on $\mathcal{Y}^\infty$ will be equal to its expected value with respect to $P$ ([22], Proposition 8.6).

With respect to our game-theoretic protocols, however, the pricing cones determined by a forecasting system are rather special. In Protocol 1, for example, Forecaster is asked to give only a single number $\mu_n$ as a prediction of $y_n \in [0, 1]$, not a probability distribution for $y_n$. The pricing cone thus offered to Skeptic (tickets that cost $\mu_n$ and pay $y_n$) is much smaller than the pricing cone defined by a probability distribution for $y_n$ that has $\mu_n$ as its expected value. In Protocol 2,
The Bulletin of the EATCS

Forecaster has the option on each move of offering a pricing cone defined by a probability distribution for Reality’s move, but he also has the option of offering a smaller pricing cone.

4.3 Duality

Using the concept of a forecasting system, we can see how game-theoretic and measure-theoretic probability are dual to each other. The quantity $\mathbb{E}(\xi)$ represented in Equation (4) as an infimum over a class of capital processes is also a supremum over a class of forecasting systems.

As a first step to understanding this duality, consider how pricing cones on $Y$ are related to probability measures on $Y$. For simplicity, assume $Y$ is finite, let $\mathcal{G}$ be the $\sigma$-algebra consisting of all subsets of $Y$, and again write $P_Y$ for the set of all probability measures on $(Y, \mathcal{G})$. Given a pricing cone $C$ on $Y$, set

$$P_C := \{P \in P_Y \mid \mathbb{E}_P(f) \leq 0 \text{ for all } f \in C\}. \tag{15}$$

Given a real valued function $\xi$ on $Y$, we can show that

$$C = \{f : Y \to \mathbb{R} \mid \mathbb{E}_P(f) \leq 0 \text{ for all } P \in P_C\} \tag{16}$$

and that

$$\sup\{\mathbb{E}_P(\xi) \mid P \in P_C\} = \inf\{\alpha \in \mathbb{R} \mid \exists f \in C \text{ such that } \alpha + f(y) \geq \xi(y) \text{ for all } y \in Y\}$$

$$= \inf\{\alpha \in \mathbb{R} \mid \xi - \alpha \in C\}. \tag{17}$$

Equations (15) and (16) express one aspect of a duality between pricing cones and sets of probability measures. Equation (17) says that an upper price defined by taking an infimum over a pricing cone can also be obtained by taking a supremum over the dual set of probability measures.\(^{10}\)

The concept of a filtration, because of the way it handles probabilities conditional on events of probability zero, does not lend itself to simple extension of (17) to a probability game with more than one round. Simple formulations in discrete time are possible, however, using the concept of a forecasting system.

For simplicity, assume again that $Y$ is finite, and let us also assume that the game ends after $N$ rounds. Write $Y^*$ for the set of all finite sequences of elements of $Y$ of length less than $N$. In symbols: $Y^* := \cup_{n=0}^{N-1} Y^n$. A forecasting system with horizon $N$ is a mapping from $Y^*$ to $P_Y$. Here, as in the binary case we just studied

\(^{10}\)Because of the finiteness of $Y$ and Condition 4 in our definition of a pricing cone, the infimum and the supremum in (17) are attained.
more closely, a forecasting system $\mu$ determines a probability measure $P_\mu$ on $\mathcal{Y}^N$ that has the probabilities given by $\mu$ as its conditional probabilities when these are well defined. Let us write $\mathcal{F}_{\mathcal{Y},N}$ for the set of all forecasting systems with horizon $N$.

We modify Protocol 2 by stopping play after round $N$ and fixing a strategy for Forecaster, say $\mathcal{F}_{\text{strat}}$, that ignores the moves by Skeptic and chooses $C_n$ based only on Reality’s previous moves $y_1 \ldots y_{n-1}$; this means that $\mathcal{F}_{\text{strat}}$ is a mapping from $\mathcal{Y}^*$ to $\mathcal{C}_\mathcal{Y}$. Since Forecaster’s strategy is fixed, we may remove him from the protocol, writing it in this form:

**Protocol 4. Finite horizon & fixed forecasts**

**Parameters:** $N$, Reality’s move space $\mathcal{Y}$, Forecaster’s strategy $\mathcal{F}_{\text{strat}}$

Skeptic announces $\mathcal{X}_0 \in \mathbb{R}$.

FOR $n = 1, 2, \ldots, N$:

- Skeptic announces $f_n \in \mathcal{F}_{\text{strat}}(y_1 \ldots y_{n-1})$.
- Reality announces $y_n \in \mathcal{Y}$.
- $\mathcal{X}_n := \mathcal{X}_{n-1} + f_n(y_n)$.

In this finite-horizon protocol, $\Omega = \mathcal{Y}^N$, and our definition of the upper price of a bounded variable $\xi$, (4), simplifies to

$$\overline{\mathbb{E}}(\xi) := \inf\{L(\omega) \mid L \in \mathcal{L} \text{ and } L(\omega) \geq \xi(\omega) \text{ for all } \omega \in \Omega\}.$$  

We can show that

$$\overline{\mathbb{E}}(\xi) = \sup\{\mathbb{E}_\mu(\xi) \mid \mu \in \mathcal{F}_{\mathcal{Y},N} \text{ and } \mu_{y_1 \ldots y_n} \in \mathcal{P}_{\mathcal{F}_{\text{strat}}(y_1 \ldots y_n)} \text{ for all } (y_1 \ldots y_n) \in \mathcal{Y}^*\}.$$  

This is the duality we announced at the outset: the infimum over initial stakes for different capital processes available to Skeptic that attain $\xi$ equals the supremum over expected values of $\xi$ for different forecasting systems that respect the offers made to Skeptic. See [7] for proofs and further comments on this duality.

### 4.4 Continuous time

It would be out of place to emphasize continuous-time processes in an introduction to game-theoretic probability for computer scientists. But these processes are very important in the measure-theoretic framework, and we would be selling the game-theoretic framework short if we did not take the time to point out that it can make a contribution in this domain.

How can we adapt the idea of a probability game to the case where Reality chooses a continuous-time path $y_t$ instead of merely a sequence of moves $y_1, y_2, \ldots$? One answer, which uses non-standard analysis, was developed in [22]. In more
recent work, which seems more promising, one supposes that Skeptic divides his capital among many strategies, all of which make bets at discrete points in time, but some of which operate at a much higher frequency than others. This approach has been dubbed high-frequency limit-order trading by Takeuchi [27].

Some of the continuous-time results require surprisingly little structure: we merely assume that Reality outputs a continuous path \( y_t \) that Skeptic observes as time passes, and that at each time \( t \) Skeptic is allowed to buy an arbitrary number of tickets (negative, zero, or positive) that will pay him \( S_{t'} - S_t \) at a future time \( t' \) of his choice. (Imagine that \( S_t \) is the price at time \( t \) of a security traded in an idealized financial market.) This assumption, combined with our definition of almost sure (an event happens almost surely if there is a strategy for Skeptic that multiplies the capital it risks by an arbitrarily large factor when the event fails) allows us to derive numerous qualitative properties that have been proven for Brownian motion and other martingales in the measure-theoretic framework. For example, we can show that \( S_t \) almost surely has no point of increase [33].

We can also show that \( S_t \) will almost surely have the jaggedness of Brownian motion in any interval of time in which it is not constant [34, 27, 31]. It appears that volatility is created by trading itself: if the price is not constant, there must be volatility. In general, a result analogous to that obtained by Dubins and Schwarz in 1965 for continuous martingales in measure-theoretic probability holds in this game-theoretic picture for \( S_t \): any event that is invariant under transformations of the time scale has a game-theoretic probability, which is equal to its probability under Brownian motion [11, 32].

We can add additional structure to this game-theoretic picture by adding another player, Forecaster, who offers Skeptic additional betting opportunities. In this way, we can construct game-theoretic analogs to well known stochastic processes, including counting processes and Brownian motion [29]. The game-theoretic treatment of stochastic differential equations, sketched using non-standard analysis in [22], has yet to be undertaken in the high-frequency limit-order trading model.

The contribution here goes beyond showing that game-theoretic probability can obtain results already obtained by measure-theoretic probability. The game-theoretic approach clarifies the assumptions needed: the notion that Reality be-
haves stochastically is reduced to the assumption that Skeptic cannot multiply
the capital he risks by a large or infinite factor. And because Skeptic tests Reality by betting at discrete points of time, the game-theoretic approach makes the
continuous-time picture directly testable.

4.5 Open systems

An important aspect of the game-theoretic framework for probability is the open
character of the protocols with which it works. Our protocols require only that
the three players move in the order given and that Skeptic see the other players’
moves. The players may receive other information, some of it private. Our theo-
rems, such as the law of large numbers and Lévy’s zero-one law, are not affected
by such additional information.

In some applications, it is useful to make additional information explicit. We
sometimes elaborate Protocol 2, for example, by having Reality give the other
players information $x_n$ before they move on the $n$th round. If we write $X$ for the
space from which this information is drawn, the protocol looks like this:

Protocol 5. Prediction with auxiliary information

Parameters: Reality’s information space $X$, Reality’s move space $Y$

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.

FOR $n = 1, 2, \ldots$:

1. Reality announces $x_n \in X$.
2. Forecaster announces $c_n \in C_y$.
3. Skeptic announces $f_n \in C_x$.
4. Reality announces $y_n \in Y$.

\[ \mathcal{K}_n := \mathcal{K}_{n-1} + f_n(y_n). \]

Putting the protocol in this form allows us to discuss strategies for Forecaster and
Skeptic that use the $x_n$, but it does not invalidate the theorems for Protocol 2 that
we have discussed. These theorems say that Skeptic can achieve certain goals
using only the information about past $y_n$, regardless of how his opponents move
and regardless of their additional information.

In many scientific and engineering applications of probability and statistical
theory, only certain aspects $y_1, y_2, \ldots$ of a process are given probabilities, while
other aspects $x_1, x_2, \ldots$, although they may affect the probabilities for the $y$, are not
themselves given probabilities. Examples include:

- Quantum mechanics, where measurements $y_n$ have probabilities only after
we decide on the circumstances $x_n$ under which we make measurements.

See section 8.4 of [22].
• Genetics, where probabilities for the allele $y_n$ of the next child are specified only after the next parents to have a child, $x_n$, are specified.

• Decision analysis, where in general outcomes $y_n$ have probabilities only after decisions $x_n$ have been made.

• Regression analysis, where each new outcome $y_n$ is modeled only conditionally on a vector $x_n$ of predictor variables.

In these examples, we can say we are using measure theory. Our model, we can say, is a class of probability measures – all the probability measures for $x_1 y_1 x_2 y_2 \ldots$ in which the conditional probabilities for $y_n$ given $x_1 y_1 \ldots x_{n-1} y_{n-1} x_n$ satisfy certain conditions, the conditional probabilities for $x_n$ given $x_1 y_1 \ldots x_{n-1} y_{n-1}$ not being restricted at all. This formulation is, however, pragmatically and philosophically awkward. Pragmatically awkward because in practice one may slip from this assumption to the potentially misleading assumption that the $x_n$ are all fixed in advance. Philosophically awkward because we may not really want to say that the $x_n$ follow some completely unknown or unspecified probability model. What is the content of such a statement?

The game-theoretic approach deals with these examples more straightforwardly. We specify bets on each $y_n$ based on what is known just before it is announced. Using Cournot’s principle we can give these bets an objective interpretation: no opponent will multiply the capital they risk by a large factor. Or we can settle for a subjective interpretation, either by weakening Cournot’s principle (we believe that no opponent will multiply the capital they risk by a large factor) or by asserting, in the spirit of de Finetti, that we are willing to make the bets. There is no need to imagine unspecified bets on the $x_n$.

5 Conclusion

In this article, we have traced game-theoretic probability back to Blaise Pascal, and we have explained, with simple examples, how it generalizes classical probability. In particular, we have stated game-theoretic versions of the strong law of large numbers, Lévy’s zero-one law, and the law of calibration. We have also spelled out various relationships with the measure-theoretic framework for probability.

When a field of mathematics is formalized in different ways, the different frameworks usually treat topics at the core of the field similarly but extend in different directions on the edges. This is the case with the game-theoretic and measure-theoretic frameworks for probability. They both account for the central
results of classical probability theory, and the game-theoretic framework inherits very naturally the modern branches of measure-theoretic probability that rely on the concept of a martingale. But outside these central topics, the two frameworks offer more unique perspectives. Some topics, such as ergodic theory, are inherently measure-theoretic and seem to offer little room for fresh insights from the game-theoretic viewpoint. In other areas, the game-theoretic framework offers important new perspectives. We have already pointed to new perspectives on Brownian motion and other continuous-time processes. Other topics where the game-theoretic viewpoint is promising include statistical testing, prediction, finance, and the theory of evidence.

In the thesis he defended in 1939 [28], Jean Ville explained how we can test a probabilistic hypothesis game-theoretically. The classical procedure is to reject the hypothesis if a specified event to which it assigns very small probability happens. Ville pointed out that we can equivalently specify a strategy for gambling at prices given by the hypothesis and reject the hypothesis if this strategy multiplies the capital it risks by a large factor. In other words, we reject the hypothesis if a nonnegative capital process — a nonnegative martingale, in the now familiar terminology that Ville introduced — becomes many times as large as its initial value. Ville also pointed out that we can average martingales (this corresponds to averaging the gambling strategies) to obtain a more or less universal martingale, one that becomes very large if observations diverge from the probabilities in any important way. In the 1960s, Per Martin-Löf and Claus-Peter Schnorr rediscovered, formalized, and developed the idea of a universal test or universal supermartingale. The game-theoretic framework allows us to make these ideas practical. As we showed in [22], we can construct martingales that test violations of classical laws. The notion of a universal test is only an ideal notion; Martin-Löf’s universal test and Schnorr’s universal supermartingale are not computable. But by combining gambling strategies that test classical laws implied by a statistical hypothesis, we can construct martingales that are more or less universal in a practical sense.

In 1976 [16], Leonid Levin realized that for any test, including any universal test, there is a forecasting system guaranteed to pass the test. Levin’s terminology was different, of course. His picture was not game-theoretic; instead of a forecasting system, he considered something like a probability measure, which he called a semimeasure. He showed that there is a semimeasure with respect to which every sequence of outcomes looks random.

13A capital process is a supermartingale rather than a martingale if the strategy for Skeptic that produces it is allowed to discard some capital at each step (which is not allowed in the case of martingales).

14Levin’s terminology was different, of course. His picture was not game-theoretic; instead of a forecasting system, he considered something like a probability measure, which he called a semimeasure. He showed that there is a semimeasure with respect to which every sequence of outcomes looks random.
versions of Levin’s idea. For a wide class of prediction protocols and computable
game-theoretic laws of probability, one can construct a computable forecasting
system that produces forecasts that conform to the law. By choosing suitable
laws of probability, we can ensure that our forecasts agree with reality in all the
ways we specify. We call this method of defining forecasting strategies defensive
forecasting. It works well in many settings. It extends to decision problems,
because the decisions that are optimal under forecasts that satisfy appropriate laws
of probability will have satisfactory empirical performance, and it compares well
with established methods for prediction with expert advice [36, 30, 6].

We noted some of game-theoretic probability’s implications for the theory of
finance in §4.4. Other work has shown that versions of some of the standard re-
sults in finance can be obtained from the game-theoretic framework alone, with-
out the introduction of stochastic assumptions. In [35], an empirical version of
CAPM, which relates the average returns from securities to their correlations with
a market portfolio, is derived game-theoretically. In [37], observed correlations in
stock returns are subjected to purely game-theoretic tests, and it is concluded that
apparent inefficiencies are due to transaction costs.

A central question in the theory of evidence is the meaning and appropriate-
ness of the judgements involved in updating and the combination of evidence.
What judgements are involved, for example, when we use Bayes’s theorem, Wal-
ley’s rule for updating upper and lower probabilities, or Dempster’s rule for com-
bining belief functions? A game-theoretic answer to these questions is formulated
in [21].

Acknowledgments
We are grateful to Yuri Gurevich and Akimichi Takemura for encouragement.

References
[1] Laurent Bienvenu, Glenn Shafer, and Alexander Shen. On the history of marting-
gales in the study of randomness. Electronic Journal for History of Probability and
Previous editions appeared in 1979 and 1986.


The Bulletin of the EATCS


BEATCS no 100       THE EATCS COLUMNS


REPORTS FROM CONFERENCES
REPORT ON CS&P 2009

The 18th Workshop on Concurrency, Specification, and Programming

Manfred Kudlek

CS&P 2009, the XVIIIth in this series, was held from September 28-30, at Kraków-Przegorzały. Conference site was Dom Gościnny Przegorzały, Uniwersytet Jagielloński (Guest House of the Jagiellonian University), situated on top of a hill, about 6 km from Kraków centre, where also all participants stayed, and breakfast, lunch and dinner were served.

It was organized by Uniwersytet Warszawski and Humboldt Universität zu Berlin. The organizing and program committee consisted of Hans-Dieter Burkhard, Piotr Chrząstowski-Wachtel, Ludwik Czaja, Manfred Kudlek, Gabriela Lindemann, Wojciech Penczek, Louchka Popova-Zeugmann, Andrzej Salwicki, Holger Schlingloff, Andrzej Skowron, Zbigniew Suraj, and Marcin Szczuka.

The international workshop was attended by 58 participants from 8 countries, as given below:

<table>
<thead>
<tr>
<th>Country</th>
<th>BG 1</th>
<th>FR 1</th>
<th>RU 6</th>
<th>SK 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DE 11</td>
<td>PL 33</td>
<td>SE 1</td>
<td>UK 4</td>
</tr>
</tbody>
</table>

The scientific program consisted of 61 contributions, presented in two parallel tracks, reflecting a more theoretical part (mathematical models of concurrency, specification languages, theory of programming, parallel algorithms, model checking and testing), and a more practical part (multi-agent systems, rough sets, object-oriented approaches, knowledge management, knowledge discovery and data mining, soft computing, applications). Their distribution by countries and number of authors is shown below.

<table>
<thead>
<tr>
<th>Country</th>
<th>BG</th>
<th>FR</th>
<th>NL</th>
<th>SE</th>
<th>UK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>CA</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>DE</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

The program can be found at http://csp2009.mimuw.edu.pl/program.php.

The workshop was opened on Monday morning by Ludwik Czaja and Marcin Szczuka, followed by a presentation by Andrzej Skowron on the history of
CS&P which started in the mid 70’s of last century as a bilateral meeting between Berlin and Warsaw.

Only some of good and interesting contributions can be mentioned, as those by Łukasz Mikulski on nonviolence Petri nets, by Michał Knapik on time Petri nets with discrete-time semantics, by Vladimir A. Bashkin on a single-periodic base for reachability in 1-counter nets, by Józef Winkowski on multiplicative transition systems, by Maciej Koutny on minimal regions of elementary net systems with localities, by Nataliya Griboskaya on timed transition systems, and by Roman R. Rędziejowski on associative omega-products of traces.

The proceedings, edited by Ludwik Czaja and Marcin Szczyka, containing all contributions, have been published in 2 volumes as reports of Warsaw University.

The social program consisted of a conference dinner at the conference site on Tuesday evening, and an excursion on Wednesday afternoon to the famous salt mine at Wieliczka near Kraków.

Weather was mostly fine, with some showers and highest temperatures just above 20°.
**REPORT ON DCM 2009**

The 5th Workshop on Developments in Computational Models

Manfred Kudlek

DCM 2009, the 5th workshop in this series, was held on Rhodos on July 11, 2009, as a satellite workshop of ICALP 2009. Conference site was resort hotel Rodos Palace at Ixiá, about 5 km from the centre of Rodos City.

It was organized by Barry Cooper and Vincent Danos.

DCM 2009 was attended by 24 participants from 11 countries. Details are given in the following table (the figures might be slightly different due to insufficient information):

<table>
<thead>
<tr>
<th>Country</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE</td>
<td>1</td>
</tr>
<tr>
<td>CA</td>
<td>1</td>
</tr>
<tr>
<td>CH</td>
<td>1</td>
</tr>
<tr>
<td>DE</td>
<td>2</td>
</tr>
<tr>
<td>FR</td>
<td>6</td>
</tr>
<tr>
<td>GR</td>
<td>1</td>
</tr>
<tr>
<td>IT</td>
<td>6</td>
</tr>
<tr>
<td>NL</td>
<td>1</td>
</tr>
<tr>
<td>RU</td>
<td>2</td>
</tr>
<tr>
<td>SE</td>
<td>1</td>
</tr>
<tr>
<td>UK</td>
<td>2</td>
</tr>
</tbody>
</table>

The scientific program consisted of 2 invited lectures and 12 contributions. It can be found at the web site [http://www.pps.jussieu.fr/~danos/dcm09/](http://www.pps.jussieu.fr/~danos/dcm09/). The talk by Mark Hogarth had to be cancelled.

The workshop was opened on Saturday morning by Barry Cooper.

The first invited lecture, ‘Anyons, Topology and Quantum Computation’ (not anions!, from ‘any’ phase) by Prakash Panagaden was an excellent survey of the meeting point of physics, biology and computer science, on spin statistics and underlying groups as SO(3), SU(2), the braid group, their topology and representations, spin statistics theorems, and influence of entanglement.

Also the second one, ‘Random Constraint Satisfaction Problems’ by Amin Coja-Oghlan, was a very good overview on (‘not a new model of computation, but computational challenge’) relation between statistical mechanics and computer science, spin glasses, random k-SAT, efficient algorithms, replica symmetry and its breaking, and in particular relations to constraint satisfaction problems.

Also to mention are the good presentations by Simon Perdrix on information flow in secret sharing protocols for a system of quantum states, Luca Bernadinello on orthomodular lattices induced by concurrency relation, occurrence nets, and causally closed sets, Walid Gomaa on characterization of polynomial time complexity of rational and real functions, Stefan Leijnen on biological and cultural...
evolution, and size of the creative part of a society, Farid Ablayev and Alexander Vasiliev (a joint talk) on algorithms for quantum branching programs, and by Paola Bonizzoni on circular word languages generated by complete splicing systems, and pure unitary languages.

Özcan Kahramanoğulları dedicated his talk to Emmanuelle Caron, one of his co-authors, who had passed away short time before, and we had a minute of silence commemorating her.

At the end Prakash Panangaden gave some some remarks, in particular on the great variety of fields on the workshop.

All papers can be found at the web site of DCM 2009, under Programme.

DCM 2009 was a successful workshop, of high scientific level. It is intended to put pictures on the web site of ICALP 2009.
Report on DLT 2009

The 13th International Conference on Developments in Language Theory

Manfred Kudlek

DLT 2009, the 13th in this series of conferences on Theoretical Computer Science, took place in Stuttgart from June 29-July 3, 2009. Conference site was the Telekom Hotel, attached to the Campus of Stuttgart University, and in the Computer Science Building nearby.

The organizing committee consisted of Volker Claus, Volker Diekert (chair), Ulrich Hertrampf, Michael Matthiesen, and Dirk Nowotka, assisted by the organizing staff Benjamin Hoffmann, Steffen Kopecki, Manfred Kufleitner, Jürgen Laun, Alexander Lauser, Heike Photien, Horst Prote, and Philipp Rieger.

Supporters were Universität Stuttgart, DFG (Deutsche Forschungsgemeinschaft) and, infos (Informatik-Forum Stuttgart). DLT 2009 was attended by 96 participants from 22 countries, details given in the following table (C country, P number of participants, some had 2 affiliations):

<table>
<thead>
<tr>
<th>C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR</td>
<td>1</td>
</tr>
<tr>
<td>CA</td>
<td>7</td>
</tr>
<tr>
<td>CH</td>
<td>1</td>
</tr>
<tr>
<td>CY</td>
<td>1</td>
</tr>
<tr>
<td>CZ</td>
<td>1</td>
</tr>
<tr>
<td>DE</td>
<td>35</td>
</tr>
<tr>
<td>ES</td>
<td>2</td>
</tr>
<tr>
<td>FI</td>
<td>6</td>
</tr>
<tr>
<td>FR</td>
<td>10</td>
</tr>
<tr>
<td>GR</td>
<td>1</td>
</tr>
<tr>
<td>IT</td>
<td>13</td>
</tr>
<tr>
<td>IS</td>
<td>1</td>
</tr>
<tr>
<td>JP</td>
<td>1</td>
</tr>
<tr>
<td>HU</td>
<td>1</td>
</tr>
<tr>
<td>IL</td>
<td>1</td>
</tr>
<tr>
<td>NL</td>
<td>1</td>
</tr>
<tr>
<td>PL</td>
<td>1</td>
</tr>
<tr>
<td>RO</td>
<td>1</td>
</tr>
<tr>
<td>SK</td>
<td>1</td>
</tr>
<tr>
<td>UK</td>
<td>2</td>
</tr>
<tr>
<td>US</td>
<td>3</td>
</tr>
</tbody>
</table>

The scientific program consisted of 5 invited lectures and 35 presentations, selected from 69 submissions (5 others were withdrawn). A detailed statistics by countries and number of authors is given in the two tables below.

<table>
<thead>
<tr>
<th>C</th>
<th>I</th>
<th>S</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>½</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CY</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EG</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FI</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HU</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IN</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IS</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MY</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NL</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RO</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RU</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZA</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Buletin of the EATCS no 100, pp. 155–159, February 2010
© European Association for Theoretical Computer Science
The presentations were given in plenary sessions, with the exception of late Tuesday afternoon with 2 parallel sessions of 3 talks each. The program can be found at http://www-dlt2009.informatik.uni-stuttgart.de/.

The conference was opened on Tuesday morning by Volker Diekert.

With the first invited talk 'Weighted versus Probabilistic Logics' Paul Gastin (co-author Benedikt Bollig) gave an excellent and very interesting overview on weighted automata, weighted MSO logic, and weighted CTL* and probabilistic CTL*, their relations to each other, and model checking, satisfiability and expressiveness related to them.

Christos A. Kapoutsis gave a clear and very interesting second one with 'Size Complexity of Two-way Finite Automata, a survey on the relations between P/NP, L/NL and 2DFA/2NFA ('all of you must know since I am here'), in particular on lifeness problems and reductions. His last slide showed Bertrand Russell with pipe and sheet of blank paper, still empty in the evening (thinking on his paradoxon).

The third invited lecture 'Factorization Forests' by Mikołaj Bojańczyk was an excellent, well illustrated survey on applications of factorization forests, based on the corresponding theorem of Imre Simon, to get fast algorithms for query questions as infix pattern matching, stronger versions of the Kleene theorem, and production of factorization forests by transducers. He started with 'Probably you will not learn anything new'. Mainly about the work by Imre Simon.

Benjamin Steinberg (co-author Jorge Almeida) presented a very good, fast and interesting fourth one 'Matrix Mortality and the Pin-ˇCerný Conjecture, starting with 'thanks that so many people are present so early in the morning. In the second part, if you want, can leave or whatever you want to do'. The first part was on representation theory, the second one (and now something completely different'), on synchronizing groups, theorems of Pin, Dubuc, Rystsov, Černý’s Cayley graph, and irreducible representations of groups. He finished with ‘Vielen Dank für Ihre Aufmerksamkeit’.

Another excellent invited lecture was the fifth one 'Post Correspondence Problem and Small Dimensional Matrices', given by Tero Harju. He started with 'If you somehow hear music, it’s not me', referring to the following colloquium for Volker Claus. He presented an interesting survey on decidability problems for integer matrices, using Claus instances of the Post correspondence problem, referring to results by Ehrenfeucht, Pavlenko, Matiyaseviich, Sénizergues, Cețin, Halava and Hirvensalo.
On Friday Galina Jirásková received the Infos Best Paper Award from Michael Matthiesen for her contribution ‘Magic Numbers and Ternary Alphabet’ which she then presented in a good talk.

To mention are also the and the good and interesting presentations by Marie-Pierre Béal on synchronizing words and the Černý conjecture, by Elyot Grant on closure operators for languages, by Josh Gunter on complexity of avoidability of partial word sets, by Olivier Carton on left and right synchronous relations, by Shinnosuke Seki on an extension of the Lyndon-Schützenberger result for DNA strands, and by Aleksí Saarela on the complexity of Hmelevskii’s theorem ans satisfiability of equations with 3 unknowns.

Other good and interesting talks were given by Fabio Mogavero on branching time temporal logics with minimal model quantifiers, by Victor Mitrana on iterated superposition of regular languages, by Giovanna Rosone on balanced words with simple Burrows-Wheeler transform, and by the three speakers of the single author session on Wednesday afternoon, Arseny M. Shur on 2-sided bounds for growth of power-free languages, Rodrigo de Souza on equivalence decidability of transducers, and by Georg Zetzsche onerasing in Petri net languages and matrix grammars.

Also to mention are the excellent presentation by Markus Holzer, with his traditional conference tie (but without corresponding shirt), on tigh bounds for de- scriptional complexity of regular expressions, and the good and interesting ones by Szilárd Zsolt Fazeksa on powers of regular languages, by Pierluigi Frisco on multihead 2 way finite and pushdown automata with 1 state, by Holger Petersen on efficiency of simulations of time-bounded counter macine storages, and by Victor Selivanov on first order definability in infix order on words.

The last session on Tuesday was entirely Finnish, all speakers and the chairman Juhani Karhumäki from Finland who remarked that questions can also be in Finnish.

The proceedings, edited by Volker Diekert and Dirk Nowotka, containing all invited lectures as well as all contributions, have been published as Springer LNCS 5583.

In the business meeting on Thursday Volker Diekert talked on DLT 2009 and thanked all speakers, as well as the organizing staff which he also presented. Then Sheng Yu presented the place and time for DLT 2010: UWO, University of Western Ontario, and August 17-20, 2010. Finally, Giancarlo Mauri presented Università degli Studi, Milano-Bicocca as place for DLT 2011.

In the breaks coffee, tea, juice and cakes were offered. Lunch was served in the restaurant of Telekom Hotel where also most of the participant stayed. Access to internet was available on more than 20 PC’s, 5 reserved exclusively for DLT participants.

The social program started with a welcome reception on Tuesday evening.
Warm and cold buffet, beer, mineral water, and wine was offered. On Thursday afternoon we had an excursion to the old university town Tübingen. There we had a guided tour in three groups through the old town and the castle on the hill. Following that we enjoyed a two hour Stocherkahn (boat moved by poking) tour on the river Neckar, in boats covering about 20 people and moved by long poles. On the boats we got beer and pretzels. After that we had the conference dinner in Casino am Neckar. There we got a mainly Italian buffet, fruits, dessert, cheese, bread, wine, mineral water, and coffee. Later on Janusz Brzozowski thanked the organizers for the successful conference. It lasted until 22 h, and we arrived back at Stuttgart around 23 h.

Pictures of DLT 2009 can be found at the web site of the conference.

Weather was fine, except for some thunderstorm on Friday afternoon, with highest temperatures near 30° C.

Thus DLT 2009 was a successful conference, of high level, and well organized in a nice atmosphere.

In connection to DLT 2009 there was also the satellite workshop Automata and Algorithmic Logic (AAL) which took place on June 28-29, 2009 in the Computer Science Building. It was organized by Thomas Colcombet, Dietrich Kuske, and Markus Lohrey, and sponsored by European Science Foundation and Vereinigung von Förderern und Freunden der Universität Leipzig. It was attended by about 20 participants. The scientific program consisted of 3 invited lectures and 11 contributions. It can be found at http://www.informatik.uni-leipzig.de/lohrey/AAL.html.

The invited lectures were ‘Transitive Closure Logic, Nested Tree Walking Automata, and XPath’ by Balder ten Cate, ‘Proving Non-automaticity’ by Bakhadyr Khoussainov, and ‘Definability Questions for MSO’ by Christoph Löding.

After the end of DLT 2009, on Friday afternoon, the Festkolloquium on Occasion of the 65th Birthday of Volker Claus was held, organized by Volker Diekert, with many participants from Stuttgart and all over Germany. The program can also be found at the DLT web site. Volker Diekert opened the colloquium, welcoming all guests, in particular Günter Hotz. In the talks the scientific and academic life of Volker Claus was honoured, among others his Nikolaus lectures, in talks by Johannes Blömer, Dieter Fritsch, Hans-Ulrich Heiß, Peter Widmayer, and in a performance by his students. The laudatio was given by Egon Börger. During the colloquium Volker Claus also received the University Medal of Carl von Ossietzky University Oldenburg from Hans-Jürgen Appelrath, and the first copy of a Testschrift in his honour (Informatik als Dialog zwischen Theorie und Anwendung, edited by Volker Diekert and Karsten Weicker, published by Vieweg+Teubner), containing many articles by his friends and
colleagues. Finally, the honoured also talked about events in his life, and thanked all organizers, speakers and guests. Pictures can be found at http://theopics.fmi.uni-stuttgart.de/v/Claus65/.
REPORT ON FCT 2009

The 17th International Symposium on Foundations of Computation Theory

Manfred Kudlek

FCT 2009, the 17th in this series of conferences on Theoretical Computer Science, was held at Wrocław from September 2-4, 2009. Conference site was a new building (D20), room 10B, of Politechnika Wrocławska (Wrocław University of Technology).

The conference was organized by Instytut Matematyki i Informatyki, Politechnika Wrocławska (Institute of Mathematics and Computer Science of Wrocław University of Technology) and Instytut Informatyki, Uniwersytet Wrocławski (Institute of Computer Science, University of Wrocław).

The organizing committee consisted of Witold Charatonik, Jacek Cichoń, Maciej Gębala (chair), Małgorzata Januszkiewicz, Małgorzata Korzeniowska, Mirosław Kutyłowski, Paweł Zieliński, as well as Przemysław Błaśkiewicz, Tomek Cichocki, Wiktor Dolecki, Mateusz Golicz, Maciej Kamiński, Michał Koza, Anna-Lauks-Dutka, Jakub Lemiesz, and Michał Wrona.

FCT 2009 was sponsored by IBM Poland, mil (Microtech International), Wrocław City, Wrocław University of Technology, University of Wrocław, and TI (Trusted Information Consulting).

In connection with FCT 2009 also two workshops took place:

Non-Classical Models of Automata and Applications (NMCA) from August 31 to September 1, 2009, with 3 invited talks and 14 contributions.

Dynamic Networks: Algorithms and Security (DYNAS) on September 5, with 3 invited talks and 3 contributions.

FCT 2009 was attended by 56 participants from 17 countries, as shown below.

<table>
<thead>
<tr>
<th>Country</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>2</td>
</tr>
<tr>
<td>DE</td>
<td>9</td>
</tr>
<tr>
<td>IN</td>
<td>2</td>
</tr>
<tr>
<td>NL</td>
<td>1</td>
</tr>
<tr>
<td>RU</td>
<td>1</td>
</tr>
<tr>
<td>UK</td>
<td>2</td>
</tr>
<tr>
<td>CA</td>
<td>1</td>
</tr>
<tr>
<td>FR</td>
<td>4</td>
</tr>
<tr>
<td>IT</td>
<td>2</td>
</tr>
<tr>
<td>PL</td>
<td>11+9</td>
</tr>
<tr>
<td>SE</td>
<td>1</td>
</tr>
<tr>
<td>US</td>
<td>3</td>
</tr>
<tr>
<td>CH</td>
<td>1</td>
</tr>
<tr>
<td>IE</td>
<td>1</td>
</tr>
<tr>
<td>JP</td>
<td>3</td>
</tr>
<tr>
<td>RO</td>
<td>1</td>
</tr>
<tr>
<td>SK</td>
<td>2</td>
</tr>
</tbody>
</table>

The scientific program consisted of 3 invited talks and 29 contributions, selected from 67 submissions, as given in the following statistics:
The Bulletin of the EATCS

The distribution by number of authors was as follows:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>17</td>
<td>21</td>
<td>19</td>
<td>9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>8</td>
<td>8</td>
<td>11</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The program can be found at http://fct2009.im.pwr.wroc.pl/.

The presentation of Subhas Kumar Ghosh, Koushik Sinha was cancelled because of visa problems.

The conference was opened on Wednesday morning by Mirosław Kutyłowski, welcoming the participants, thanking the program committee, and talking shortly on the Technical University.

The first invited lecture ‘How to Guard the Guards Themselves’ by Moti Yung was an excellent personal view on modern cryptology, in particular computer security. His key message was ‘Apply cryptographic methods for protecting cryptography itself!’ In details he gave a survey on threshold cryptography (distributed crypto systems, availability), proactive security (time, rerandomizing), local key protection, self-protecting systems (forward, key-insulated, intrusion resilient security), on mathematical and physical problems of security, on side channel attacks (formal system and model theory for it), and on various abstract models (e.g. for leakage). He finished with the need of a holistic view for the interaction between engineering (design) and theory.

The second, very special, invited lecture, ‘In memoriam Prof. Dr. math Ingo Wegener’, was given by Martin Dietzfelbinger on the research and teaching activities of Ingo Wegener (December 4, 1950-November 26, 2008). There are 138 entries of publications. Among other honours Ingo Wegener became member of Leopoldina, GI-Fellow (2004), received the Konrad Zuse medal, and became member of Wissenschaftsrat. His scientific life can be characterized by the phases:

0: (1979-1985) Search algorithms, monograph 1, Suchprobleme (handwritten and typed by secretary).


2\(\frac{1}{2}\): Heap sort.

3: Analysis of metaheuristics, simulated annealing.

He was a gifted and devoted teacher at all levels, attracting students, distributing hand written lecture notes, giving black board lectures. Twice he won the university medal for excellent teaching. He supervised more than 130 diploma theses, 19 PhD theses, and 8 habilitations, and wrote four textbooks:


His last works were on tight bounds for blind search on the integers, and on precision, local search and unimodal functions.


Thomas A. Henzinger (co-authors Krishnendu Chatterjee, Laurent Doyen) presented an excellent third invited lecture with ‘Alternating Weighted Automata’ or ‘From Boolean to Quantitative System Specifications’. In it he talked on Boolean verification agenda (structure, formulae), assigning values to behaviours, to systems and properties, and to pairs of such, on language inclusion for Markov decision processes and weighted automata, quantitative simulation, emptiness and universality, expressiveness, and closure properties, games, and on quantitative synthesis and robust systems.

To mention are also the good and interesting presentations by Fabian Wagner on the complexity of reachability of basic non-planar directed graphs, by Johan M. M. van Rooij, wearing a T-shirt with Tamil characters (India), on computing role assignment of chordal graphs, by Florin Manea on combinatorics of partial words, by Andreas Goerdt on random ordering constraints of SAT, and by Daniel Wagner on three-valued Markov chains and completeness for a PCTL fragment.

A very good talk on bisimulation for higher dimensional automata models was given by Roman Dubtsov presenting the contribution of Elena Oshevskaya, who was present but had problems with her voice. Another excellent presentation was given by Turlough Neary on three small weakly universal Turing machines, simulating rule 110.

Other interesting and good talks were presented by Pietro Cenciarelli on dynamics and behaviour of depletable channels, by Michaela Slaats on parameterized regular infinite games and pushdown strategies for them, in the *colourful*
session, by Michele Zito on tree martingales and the empire chromatic number of random trees, by Rafał Witkowski on an algorithm for multicolouring hexagonal graphs, by Adam Roman on complexity of the road colouring problem, and by Huaming Zhang on NP-completeness of source-target orientation of planar graphs.

The proceedings of FCT 2009, edited by Witold Charatonik, Maciej Gębala and Mirosław Kutylowski have been published as Springer LNCS 5699. They contain all contributions, as well as the invited talks by Moti Yung (extended abstract), by Martin Dietzfelbinger (an obituary), and by Thomas A. Henzinger.

FCT 2009 was a conference of high scientific level and well organized. It could have deserved more participants, in a town with more than 20% of its population of 630000 being students.

In the breaks coffee, tea, soft drinks and snacks were offered. Lunch was also served in the hall in front of the lecture room.

The social program consisted of a guided tour through Wrocław on Thursday afternoon, visiting some of the many churches and crossing many bridges ending at the Market Square (Rynek) in a heavy shower, where the conference dinner took place in the evening in the restaurant Karczma Lwowska.

Weather was good, with highest temperatures between 20 and 25°C.

Next FCT will take place in 2011 in Oslo.
REPORT ON SYMPOSIUM

Information Processing - Modern Perspectives
in honour of the 75th birthday of

ACADEMICIAN ARTO SALOMAA

and Salomaa Symposium - Technical Sessions

Manfred Kudlek

This extraordinary symposium in honour of the 75th birthday of Arto Salomaa, was held on May 25, 2009, at Turku. Conference site was the Auditorium of PharmaCity. It was attended by about 80 participants from at least 14 countries.

The symposium was organized by Grzegorz Rozenberg, who also was chairing the program, and Juhani Karhumäki, as well as Arto Lepistö and Elisa Mikkola. It took place under the auspices of EATCS and Academia Europea, and was sponsored by Turun Yliopisto (UTU, University of Turku), Suomalainen Tiedeakatemia (Finnish Academy of Sciences and Letters), TUCS (Turku Centre for Computer Science), and Suomen Akatemia (Academy of Finland). The program consisted of 6 talks, 2 greeting addresses, and a closing address.

The symposium was started by Juhani Karhumäki, giving general information on the meeting. Then Keijo Virtanen, Rector of UTU (University of Turku), opened the symposium, talking briefly on the scientific life of the honorary who was born on June 6, 1934 in Turku, that weather is always as fine, in particular in May (spring had just started), congratulating Arto, and welcoming all participants.

Grzegorz Rozenberg then gave an exhaustive overview on Arto’s life who’s father had been also professor at Turku, in Philosophy. He presented Arto’s scientific career, in Logics, Automata Theory, Formal Languages, Combinatorics of Words, Cryptography and others, figures as more than 400 articles, more than 20 books, lectures at more than 150 institutes, awards, and that Arto has not really retired 10 years ago. He also mentioned other interests as classical music, soccer, and of course Sauna on which Arto wrote the famous article in the EATCS Bulletin. Not to forget his engagement in his family, Kaarina, his wife he married more than 50 years ago, and their children Kai, also a Computer Scientist, and Kirsti, working in Medicine, as well as three grand children. Finally he mentioned his own relations with Arto, and finished with ‘Thank you, Tarzan, for all’ and greeting Arto’s entire family, all being present.
Hermann Maurer with ‘How Information Technology is Influencing how we Live and Learn’ started with another H. M. talking with his hologram on Arto, Juhani, and MSW (Maurer, Salomaa, Wood), and showed some old pictures. He gave an excellent talk on new applied fields in computer science, their development and dangers, such as search engines, having monopoly and conglomerating too many data, their temptation of misuse, in particular on GOOGLE (‘GOOGLE knows too much about us, even if you don’t use it’), being the most powerful detective agency. Another topic was Wikipedia, possibly being erroneous, incomplete, beyond of control of person involved, biased, contradictory, but much better than others, as well as encyclopedias, bookstores, music industry, newspapers, mobile phones with many functions (he showed a wool egg pig, etc.

In the second part he raised the question ‘What to do?’. Basic language learning automatically, what and when? His conclusion somehow was ‘How much technology do humans need to be happy? (Almost) none’, and the difference between what we have and what we know should not be too big. Since he is also a Science Fiction writer he was referring to some of his books (published in the series XPERTEN, some translated into English, as The Paranet) with the citation ‘and then the internet collapsed one day’. Finally, he thanked Arto for their friendship, showing more pictures of MSW (they wore a T-shirt with these initials).

Cristian Calude with ‘Randomness, Logic, and Computation’ started with a movie from 1991 showing Arto with music of Sibelius in the background. Then he gave a very interesting survey on magical numbers as the Borel number $B$ from 1927, the Turing halting number $H$ from 1936, and Chaitin’s number $\Omega$ from 1975, related to the sizes of halting machines, and a variant related to self-delimiting universal Turing machines, on their (in)computability and randomness definable by them. He also presented complexities for axiomatic systems as ZFC and PA to compute the first $N$ bits of $\Omega$ (the first 40 are known, for answering Riemann hypothesis $>5000$ are needed), and that $\Omega$ is provably random in PA. Finally he showed an $\Omega$ video clip.

Emo Welzl gave a nice and very interesting talk on ‘Satisfiability - Combinatorics and Algorithms’. He started with ‘Ceci n’est pas un tableau noir’ and that Salosauna is different from Sauna. He presented a distant view of the topic (constraints and Lovász’ lemma on independent events), models (hypergraph colorability, CNF and k-CNF satisfiability, linear equation systems, bounds of cooperation), neighbourhood terms (shared variables, algorithmic complexity of the problem), conflicts of neighbours, degree bounds of variables, bounding quality of formulas (linear CNF formulas), and some relaxations of conditions.

After the lunch break, Markku Mattila, President of Academy of Finland, talked about the academic life of Arto, that he was appointed a member of the Finnish Academy by the President of Finland in 2001 (only 12 academicians at
that time). He also informed us on the activities of the Academy, the scientific research in Finland, thanked all organizers and guests, wished much success for Arto in future, and that this event is appropriate to celebrate Arto’s work.

**Tao Jiang** with the third talk ‘Discovering Regulatory Motifs from both DNA Sequence and Gene Expression Data’ gave a very interesting presentation from molecular biology. He started with ‘In China only Arto’s book (Formal Languages) is available’ and thanked Arto for supporting his career. He presented motifs, patterns, decidability problems for pattern languages (equivalence, inclusion), the central dogma of molecular biology genes encoded in DNA, transcribed to RNA, and then translated to proteins, the structure of genes, motivation and general framework (genotypes and phenotypes), finding motifs from data (deterministically, probabilistic) with an example yielding happy birthday, Arto, and other methods for that (Gibbs sampling, ChIP chip (chromation immunoprecipitation)).

In the next excellent presentation ‘The Impressive Past, the Golden Days, and the Bright Future of Automata Theory’, **Juhani Karhumäki** gave a survey on the history of automata theory (AT). He started with the remark that this is obvious (a tautology) for researchers in the field, and for others challenging to justify. ‘That’s my goal!’. He mentioned AT as corner stone of (T)CS and discrete mathematics, overlapping with computability, formal language theory (FL), and complexity. The impressive past (before 1960) begun with Hilbert’s dream in 1900 (‘every meaningful problem can be solved’), reaching Gödel’s incompleteness result, and ended in the new science of computability (Turing, Church, Kleene, Post), as well as in AT and FL (McCullough, Pitts, Kleene, Rabin, Scott, Chomsky, Hopcroft, Ullman, Salomaa). In the golden days (1959-1980) the strong connection between AT and complexity theory was born (programming language design, parsers, compilers, automata models, generative power, decidability problems), ICALP was born (Nivat, Salomaa), a number of fundamental books were published (Salomaa, Hopcroft-Ullman, Rozenberg-Salomaa), AT became diversicated (algebraic AT, formal power series, L systems), and fundamental problems were recognized, some of solved only later (closure of CS under complement, relation between deterministic and non-deterministic versions of LBA, equivalence problem for DPDA). In the bright future (1980-1995) combinatorics of words (Lothaire, Salomaa) were investigated, as well as nonperiodic tiling problems, AT to compute real functions, DLT was started in Turku (1993), and the handbook of FL was published. In the last part A Revival and the Future he mentioned DNA and quantum computing, dynamical systems, WWW research, applications in speech recognition, image manipulation, model checking, all requiring basic research. He finished with ‘Nothing is as applicable as a good theory’. Thanks Arto!

In the last very interesting talk ‘Recent Research by Arto Salomaa’ **Sheng Yu**, starting with ‘Happy 75th birthday’ (from his Department) and showing several
pictures, e.g. Arto working in a hotel), gave a survey on Arto’s research in the period 2001-2008, on decomposition of languages (prime languages), subword histories and Parikh mapping matrices (subwords and scattered subwords), and state complexity of combined operations. He concluded with ‘his spirit and ability is at least as before retirement. We want to celebrate 10, 20, 30 years in future.

Finally, not as last speaker, Grzegorz Rozenberg read a number of birthday messages (Giorgio Ausiello (EATCS), Lila Kari, Werner Kuich, Azaria Paz, Gheorghe Păun, Cunsheng Ding, Oscar Ibarra, Juraj Hromkovič, et. al. Adresses also were given by Hermann Maurer (Academia Europea), Hannu Tenhunen (TUCS), and Solomon Marcus. At the end he thanked the organizers.

The last speaker was Arto, thanking all the organizers and guests, also mentioning Alexandru Mateescu who once saved his life, but passed away several years ago. At the very end Arto got standing ovations for several minutes.

In the evening about 30 participants met for a dinner at Kaskenahde, a fish restaurant (the best one in Turku as Juhani told since he has an appartment in the same house). The cook, Kouko Takalo welcomed us with ‘Especially made for you’. We were served a number of fish dishes. Juhani congratulated Arto, and Ralph Back from Åbo University offered Arto a present as others also did. Arto thanked them and the all the guests, mentioning that he was got his name after Arthur Schopenhauer, has met and shaking hands with Kurt Gödel, and also Risto Ryti and Carl Gustav Emil Mannerheim, Finnish presidents in Second World war (not shaking hands with them), and that he is very optimistic about future.

Happy Birthday, Arto, and more such events in future to celebrate your work!

The technical sessions were held in the afternoon of May 26, 2009, at UTU, Computer Science Department. They were attended by 30 participants from 14 countries (AT, CA, CH, DE, EE, ES, FI, FR, HU, IL, NL, NZ, RO, US). The scientific program consisted of 6 presentations.

In the first very interesting talk ‘The Problems around Černý Conjecture and Road Coloring’ Avraham Trakhtman presented recent results on problems related to the Černý conjecture, in particular on synchronizing edge colourings of strongly connected graphs, generalizations to minimize the synchronizing word, k-synchronizing colouring, and a polynomial time algorithm for the road colouring problem.

Hellis Tamm, with ‘On Minimality of Biseparable Automata’, gave a very nice presentation on residual finite state automata (every state defines a residual language as accepted set), especially on the subclass of biseparable automata...
(closed under reversal), and new results on minimality such as a tight lower bound, and a transition minimality of unambiguous reversible biseparable automata.

Manfred Kudlek with ‘Concurrent Finite Automata, Multiset Pushdown Automata, and Related Language Classes’ gave an overview on finite automata using Petri nets as control, accepting word languages, the so defined language classes and their relations to other classes, and on pushdown automata working on multisets instead on words, as well as the multiset language classes defined in this way and their relation to Parikh images of word language classes.

Solomon Marcus with ‘Correctness and Meaning: Cooperation or Conflict’ presented a very interesting talk on relations between syntax, semantics, pragmatics one one side and correctness, meaning, relation on the other side, the difference between Tarski (truth values) and Bourbaki (structures), between Chomsky (semantic fallacy) and e.g. Salomaa (leading to clear meaning). He also showed that most of Hilbert’s 23 problems include semantic aspects, and that only 8 are completely solved (1,2,4,6,8,9,18,23).

With ‘Lower Bounds for the State Complexity of Nested Word Automata’ Kai Salomaa gave a very nice presentation on lower bounds of state complexity for a deterministic finite automaton working on nested words, equivalent to a given non-deterministic one with \( O(n) \) states, namely \( 2^n \), discussing also methods to find them, and considering the state complexities of basic operations on nested word languages for deterministic and non-deterministic automata versions.

‘Competence or Efficiency’, given by Erzsébet Csuhaı-Várjú, presented interesting new results on generative power and size complexity of context-free cooperating distributed grammar systems where the cooperation protocols are based on competence (number of different non-terminals in current sentential form which can be rewritten by rules of the component) or efficiency (fewer different non-terminals when rules are applied by the component).

The symposium was closed by Juhani Karhumäki, thanking all participants. Weather was fine, with highest temperatures over 20°C.
REPORT ON NCMA 2009

Workshop on Non-Classical Models of Automata and Applications

Manfred Kudlek

The Workshop Non-Classical Models of Automata and Applications (NCMA) was held in Wroclaw, from August 31 to September 1, 2009, in connection with FCT 2009. Conference site was Polytechnika Wrocławska (Wrocław University of Technology), Building D-1, Room 215.

NCMA was organized by Henning Bordihn, Rudolf Freund, Markus Holzer, Martin Kutrib, and Friedrich Otto, and locally by Mirosław Kutyłowski, Małgorzata Janusziewicz, and Małgorzata Korzeniowska.

The workshop was attended by 27 participants from 13 countries, details given below.

AT 1 | DE 6 | IT 6 | PL 1 | US 1
CA 2 | ES 2 | JP 2 | RO 1 |
CZ 2 | FR 1 | NL 1 | SK 1 |

NCMA was supported by AutoMathA and European Science Foundation (ESF).

The scientific program consisted of 3 invited talks and 14 contributions, as given below.

C I S C I S C I S
AT 2 1 3 | DE 6 1 3 | IT 6 1 3 | PL 1 3 |
CA 2 1 1 | HU 2 1 3 | SK 1 3 |
CZ 2 2 2 | IT 2 2 2 | US 1 2 3 |
DE 1 1 3 | JP 1 3 |
ES 1 3 | NL 1 3 |

NCMA was opened on Monday morning by Friedrich Otto and Mirosław Kutyłowski.

The first excellent, clear and well illustrated invited lecture ‘Automata Walking over Trees and Graphs’ was given by Hendrik J. Hoogeboom. He started with ‘Walking on bridges and not getting lost, too may universities in this city’, and showing a Bonsai tree. It was a survey on acceptance by tiling (global labelling) and walking (local control), structures (strings, trees, pictures), pebbles, power of walking and tiling, and relations to logic. He finished with a nice picture on tossing pebbles, actually the title of a book by him, encouraging to study nested pebbles.
The second invited lecture ‘Streaming Tree Automata and XPath’ was presented by Joachim Nieren (co-authors Olivier Gauwin, Sophie Tison). He presented an overview and results (‘Some not even in the PhD thesis of Olivier Gauwin’) on XML streams (linearization of XML documents, reading of elements only, incremental processing, bounded memory), tasks (validation, query answering, transformation), and applications (peer to peer data exchange, data bases, firewalls), as well on space and time complexity, and objectivities. Further topics were query answering only, streamability of query classes, nested word automata, and application to forward XPath.

Klaus Sutner gave an excellent and very interesting third invited talk ‘Cellular Automata, Decidability and Phase Space’. In it he gave a survey on relations between computation and physics, ζ-automaticity, computation and cellular automata. In particular he presented the principle of computational equivalence of Wolfram, that of Landauer, as well programs of Hilbert (to axiomatize all of physics), Born and Deutsch. Then he talked on 1-dimensional cellular automata and their relation to FOL, and on ECA 110 (‘universal in some sense’). Finally he stated a number of open problems on relations between logics and cellular automata.

To mention are also the good and interesting talks by Peter Leupold on hierarchies and decidability of finite automata working on DNA strings, by Tomáš Masopust on some regulated non-determinism for pushdown automata, by Rudolf Freund on new transition modes for P systems, starting with ‘Now you have half an hour time for a nap. Please don’t snore too loud’, by Maria Paola Bianchi on deterministic and quantum unary finite automata, and by Carlo Mereghetti on the relation between pebble machines and log-space bounded Turing machines.

The workshop was closed on Tuesday afternoon by Rudolf Freund, thanking all participants, speakers, the local organization, and ESF.

The proceedings, edited by Henning Bordihn, Rudolf Freund, Markus Holzer, Martin Kutrib, and Friedrich Otto, have been published as Band 256 of Österreichische Computer Gesellschaft (Austrian Computer Society). They contain all contributions and invited talks, although that of Joachim Niehren only as short abstract.

Next NMCA will be held on August 23-24, 2010, at Jena.

In the breaks coffee, tea, juice, mineral water, and snacks were offered.

The social program consisted of two conference dinners at Pod Gryfami at the central place Rynek, on Monday and Tuesday evening.

Pictures of NMCA can be found at [http://www.informatik.uni-giessen.de/ncma2009](http://www.informatik.uni-giessen.de/ncma2009).

Most participants stayed at the hotel Jana Pawła II. Weather was fine, with highest temperatures between 25 and 30 °C.
REPORT ON PETRI NETS 2009

22-26 June, 2009, Paris, France

Jetty Kleijn  Maciej Koutny
Leiden University  Newcastle University

The 30th International Conference on Application and Theory of Petri Nets and Other Models of Concurrency was organized by Fabrice Kordon and his team from the MeFoSyLoMa group at the campus de Jussieu of the Université Pierre et Marie Curie. MeFoSyLoMa stands for Formal Methods for Software and Hardware Systems and involves laboratories from different Parisian institutions. The Petri net conference was co-located with RSP 2009, the 20th IEEE/IFIP International Symposium on Rapid System Prototyping. There were eleven satellite events. The collaborative effort of MeFoSyLoMa took care of a smoothly run operation and provided friendly support for participants and lecturers.

The two conferences shared five keynotes: Joseph Sifakis (Turing Award 2007) who lectured on component-based construction of heterogeneous real-time systems; Grzegorz Rozenberg who presented reaction systems as a formal framework for processes based on biochemical reactions; Bernard Courtois considered past and future of prototyping of custom circuits and systems; Bill Tonti discussed rapid system deployment using the IEEE technology navigator. Unfortunately, the fifth speaker, Gabriel Juhas, fell ill just a few days before the conference and was unable to attend.

The program committee of Petri Nets, chaired by Giuliana Franceschinis and Karsten Wolf, accepted 19 papers (including 5 tool papers) out of 46 submissions from 20 different countries. The proceedings have appeared as volume 5606 in the LNCS series and some authors were invited to publish an extended, full version of their paper in a special issue of the journal Fundamenta Informaticae. The best paper award (presented at the closing session) went to Fernando Rosa-Velardo and David de Frutos-Escrig for their paper on decidability results for Petri nets with name creation and replication.

Both Petri Nets and RSP were officially opened on Wednesday. All satellite events took place on Monday and Tuesday June 22 and 23. There was an introductory tutorial on Petri Nets (Jörg Desel, Susanna Donatelli, Kurt Jensen and Jetty Kleijn) as a precursor to the modular Petri Net Course envisaged for 2010. In addition, the Workshops and Tutorials committee chaired by Susanna Donatelli and Maciej Koutny, selected four more advanced tutorials: Continuous Petri Nets, a special tutorial in honour of Laura Recalde, who passed away in December 2008, organized by Manuel Silva, Alessandro Giua, and Serge Haddad; BioModel Engineering - from Structure to Behaviour, organized by Monika Heiner, David
Gilbert, and Rainer Breitling; Evaluating Concurrent Software Architectures Using Petri Nets, organized by Rob Pettit; PNML, the Petri net Markup Language, Theory and Practice, organized by Ekkart Kindler and Lom Hillah.

RSP had two tutorials: Computer Aided Formal Verification and Validation (by Bret Michael) and System-Level Modeling and Validation of Continuous/Discrete Systems (by Gabriela Nicolescu).

Four workshops were organized in association with the Petri net conference: PNSE’09, International Workshop on Petri Nets and Software Engineering (chair Daniel Moldt); ORGMOD’09, International Workshop on Organizational Modeling (chairs Daniel Moldt, Olivier Boissier and Michael Köhler-Bußmeier); TiSto’09, International Workshop on Timing and Stochasticity in Petri Nets and Other Models of Concurrency (chairs Andras Horvath and Olivier Roux); APNOC’09 International Workshop on Abstractions for Petri Nets and Other Models of Concurrency (chairs Natalia Sidorova and Alexander Serebrenik).

Authors of some of the best workshop papers will be invited to submit a revised version to be considered for inclusion in a volume of ToPNoC, Transactions on Petri Nets and other models of Concurrency, an LNCS subseries.

Petri Nets, RSP and the associated events had a total of 222 attendees, including 75 students. Almost half of the participants came from France. We all enjoyed very much the reception offered by the city of Paris in the magnificent Hôtel de Ville. On Thursday, the visit of the Architecture Museum in the Palais de Caillot, close to the Tour Eiffel, had excellent timing as it took place during heavy showers. Afterwards, the sky was clear again when we had a lovely dinner cruising the Seine.


Petri Nets 2010, organized by João Fernandes, will take place in Braga, Portugal. It will be co-located with ACSD, the International Conference on Application of Concurrency to System Design.

REPORT ON UC 2009

The 8th International Conference on Unconventional Computation

Susan Stepney

The 8th International Conference on Unconventional Computation (UC 2009) took place at the University of the Azores, 7–11 September 2009.

UC 2009 was organised by the University of the Azores and the Centre for Discrete Mathematics and Theoretical Computer Science of the University of Auckland, under the auspices of the European Association for Theoretical Computer Science (EATCS). The conference received support from the University of Azores, from Centro de Matemática e Aplicações Fundamentais (CMAF) of the University of Lisbon, from the Regional Government of the Azores, from FLAD—Associação para a Mobilidade Antero de Quental, from Fundação para a Ciência e a Tecnologia (FCT), and from the Banco Internacional do Funchal (BANIF).

The fully international list of participants includes 87 names, coming from all parts of the globe: Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Hungary, Israel, Italy, Japan, New Zealand, Poland, Portugal, Romania, Russia, South Korea, Spain, UK, and USA.


There were three tutorials, by: Manuel Lameiras Campagnolo (Technical University of Lisbon): “Analogue Computation”, James Crutchfield (University of California at Davis): “Computational Mechanics: Natural Computation and Self-Organization”, Martin Davis (New York University and Berkeley): “Diophantine Equations”.

These, together with the scientific programme of technical presentations of the published papers, comprised of three full days and two half days, which
included a poster session where the authors had the opportunity to make a short informal presentation of their work.

Proceedings of UC 2009 are published in the Springer series as LNCS volume 5715 (ISBN 978-3-642-03744-3). The volume contains abstracts and extended abstracts of the invited papers and tutorials, 18 refereed 14 page full papers, and 5 one page poster papers. Several authors have been invited to submit extended versions for a special issue of Springer’s Natural Computing journal.

In addition to the main technical conference, there were three parallel associated workshops on related unconventional topics (thoughtfully scheduled so as not to coincide with main conference’s invited speakers): Hypercomputation Workshop 2009, Novel Computing Substrates Workshop, Physics and Computation 2009.

Additionally, Stephan Wolfram gave a presentation on “Mining the Computational Universe”, via videolink, arranged as part of the Physics and Computation workshop, but hosted as a plenary session in the main auditorium.

Topics in the main conference session ranged over many aspects of unconventional computation, including computing with real numbers and other mathematical constructs, quantum computing, optical computing, billiard ball computing, membrane computing, neurocomputing, and slime mold computing. The workshops focused on specialised themes, and included tutorials and refereed technical contributions.

Because of the parallel conference and workshop sessions, I attended only a selection of the presentations. Particular highlights for me include the following.

E. Beggs’ plenary on physical oracles helped expose some of the practical difficulties of computing with real numbers. His group’s research provides persuasive evidence to support the conjecture that each successive digit of precision takes exponentially more time to extract from the system. P. Prusinkiewicz’s plenary looked at the topology of computation, and showed how this could illuminate certain classical algorithms: from the sieve of Eratosthenes to the rather more recent shunting algorithm of Dijkstra, and then calculating the convex hull of a Sierpinski fractal curve!

J. Crutchfield’s tutorial on Computational Mechanics helped elucidate some of the properties of epsilon-machines and their use in complexity measurements. He had some exciting results about the relationship of these to “reverse epsilon machines” that retrodict the past from the future. S. Wolfram’s videotalk left me wondering if I have been too dismissive of the message from A New Kind Of Science. It is not just cellular automata, but many diverse kinds of seemingly simple computational devices, that he is exhaustively mining, exposing interesting similarities and correspondences.

In the main conference, J. Jones, in his presentation “Approximating the Behaviours of Physarum polycephalum for the Construction and Minimisation of
The Bulletin of the EATCS

Synthetic Transport Networks”, described an agent-based model of this single-celled creature that can be used to solve certain geometric problems. He presented a compelling link between the model’s behaviour (and hence possibly the biological behaviour) and classical graph hierarchy results.

C. Horsham spoke at the Hypercomputation workshop on “Hybrid hypercomputing: towards a unification of quantum and classical computation”. This approach includes a new graphical formalism for quantum circuits, and unifies quantum and non-unitary (measurement) operations in a way that helps expose the underlying structure and commonalities of these two paradigms.

V. Kendon spoke at the Novel Computing Substrates Workshop on “Analogue computation with microwaves”, a first step towards a novel analogue computational device. The long term goal is to progress the understanding of continuous quantum variable computation. In the same workshop K.-P. Zauner explored the interface between organic and inorganic components in a hybrid computer, and defined matter as animate “if it uses information processing to persist.”

Hopefully this brief survey gives some indication of the vast range of topics making up Unconventional Computation that were covered at this conference. I am sure there were also many other excellent talks in other parallel sessions; maybe I attended these in other parallel universes.

The social programme included a (slightly alarming) lecture on the active geology of the islands, by Dr. Gabriela Queiroz, Director of the Centre of Volcanology and Geological Risks Assessment at the University of Azores (accompanied by an illustrated article in the proceedings), and trips to a magnificent arboretum, active hot springs, past beautiful hedges of hydrangeas to stunning volcanic calderas, and the beach. The conference dinner fully embraced the local geology, by using geothermal energy to cook the food underground, which provided an intriguing sulphurous suggestion to the enjoyable evening. Delegates at the dinner were entertained by a troupe of traditional Azorean folkdancers, and some delegates were persuaded to join in one of the “easier” dances.

Many thanks for an excellent event go to the local organisers: José Félix Costa (Technical University of Lisbon, and Swansea University), and Elisabete Freire, Matthias Funk, Luís Mendes Gomes, and Hélia Guerra (University of Azores).

REPORT ON WORDS 2009
The 7th International Conference on Words

Arturo Carpi
Dipartimento di Matematica e Informatica, Università di Perugia, Italy
Clelia De Felice
Dipartimento di Informatica ed Applicazioni, Università di Salerno, Italy

WORDS is a biannual conference devoted entirely to combinatorics on words (i.e. finite or infinite sequences of symbols taken from a finite alphabet), aimed at gathering together active researchers in this field in order to present recent results, to exchange ideas and to cooperate on the topics of interest. Combinatorics on words is an interesting and relatively new topic in discrete mathematics. For decades, words have been a useful tool for formulating and proving results in different areas and especially in computer science. Recently, combinatorics on words has developed into an important research topic of its own. Lothaire’s books “Combinatorics on Words” (1983), “Algebraic Combinatorics on Words” (2002), “Applied Combinatorics on Words” (2005) have played an important role in the development of this field of interest.

The 7th edition of the international conference “WORDS” took place at the University of Salerno (Fisciano, Italy), from September 14 to September 18, 2009. There were 89 participants from 12 different countries: Australia, Belgium, Canada, the Czech Republic, Finland, France, Israel, Italy, Russia, Spain, Sweden and the USA. Among them there were 6 invited speakers and 34 other speakers. The Programme Committee consisted of 14 members from 6 countries. It was co-chaired by the authors of this report. Each talk session started with an invited lecturer followed by a number of selected communications. There was also a short open problem session on Monday afternoon.

Each day of WORDS 2009 was devoted to an aspect of combinatorics on words. Monday, the first day of the conference, was dedicated to applications. The opening invited lecture by A. Fraenkel was concerned with “Flora”, a game similar to the more well-known “Nim”. Fraenkel showed how winning strategies may be found by iterated applications of the floor function to linear combinations of the golden ratio and its square. Applications of combinatorics of words to game theory was also the topic of the talks given by M. Rigo (joint work with E. Duchene) and by C. Selmi (joint work with M.Arfi, B. Ould M Lemine) respectively. A surprising application of Sturmian words to musical theory was presented in the communication by M. Dominguez (joint work with T. Noll) whereas A. Juhasz’s presentation dealt with applications of combinatorics of words to a longstanding
open problem in the theory of groups. R. Grossi gave an interesting invited talk on
the applications of combinatorics of words to text indexing, combinatorial pattern
matching, and data compression.

Tuesday was the day of the conference that was dedicated more specifically to
Words and Automata Theory. The Tuesday morning session was also dedicated
to Imre Simon, the renowned scientist who recently passed away, with a short
talk in commemoration of him given by J. Sakarovitch. In his invited lecture (a
joint work with N. Rampersad and Z. Xu), J. Shallit presented several results con-
cerning the computational complexity of some decision problems for finite state
automata. There were 9 selected communications on various aspects of Automata
and Formal Languages Theory and their relation with Combinatorics on Words.
Inspired by problems in the area of compilers and processor design, M. Goldwurm
(joint work with L. Breveglieri and S. Crespi Reghizzi) presented an efficient algo-
rithm for the membership problem of certain regular trace languages. In his talk (a
joint work with S. Holub), J. Kortelainen showed that given two non-commuting
words on an alphabet $A$, one can partition the free semigroup generated by $A$ in
two regular semigroups, each of them containing one of the two given words.
Self-generating and semi-linear sets of integers were considered, respectively, in
the talks given by T. Karki (joint work with A. Lacroix and M. Rigo) and by F.
D’Alessandro (joint work with B. Intrigila and S. Varricchio). S. Crespi Reghizzi
(joint work with D. Mandrioli) presented new connections between some known
classes of context-free languages, O. Klima presented a combinatorial proof of
Simon’s algebraic characterization of piecewise testable languages and S. Lomb-
dardy (joint work with J. Sakarovitch) presented a nice result on the cross section
property. The notion of unambiguity was considered under various aspects: M.
Anselmo (joint work with M. Madonia) studied the minimization of unary unam-
biguous automata, S. Julia (joint work with T. Vinh Duc) studied the problem of
deciding whether an $\omega$-power of a regular set is generated by an $\omega$-code.

Applications of combinatorics of words to tiling was the topic of the session on
Wednesday, opened by the interesting invited lecture given by C. Reutenauer (joint
work with I. Assem and D. Smith) followed by two talks on relations between
words, tiling and polyominoes given by S. Brocchi (joint work with A. Frosini,
R. Pinzani, S. Rinaldi) and A. Garon (joint work with A. Blondin Massé, S. Brlek
and S. Labbé) respectively.

Sturmian words and some related notions, such as episturmian, Christoffel,
balanced and 3iet words were the subject of many contributions.

A. Restivo, in his invited lecture (a joint work with G. Rosone), pointed out
that balanced words can be characterized as the words with optimal performance
in text compression based on the Burrows-Wheeler transform. As is well known,
there are several methods for generating Sturmian words and each of them sug-
gests a generalization of this notion to alphabets with more than two letters. For instance, since Sturmian words code the exchange of two intervals, one may consider words coding the exchange of three intervals: this was the case in the communications of P. Ambroz (joint work with Z. Masakova and E. Pelantova) and E. Pelantova (joint work with P. Ambroz, A. Frid, Z. Masakova). Sturmian words are also characterized as infinite words such that the function counting the total number of distinct factors of length $n$ and $n + 1$ is constant, and indeed the communication of S. Starosta studied infinite ternary words with the above property (joint work with L. Balkova and E. Pelantova). The fixed points of certain transformations useful in the construction of Sturmian and episturmian words were studied in the contributions by D. Jamet (joint work with G. Paquin, G. Richomme and L. Vuillon) and by H. Uscka-Wehlou. Some applications of Sturmian words to digital lines and curves were considered by X. Provencal and J.-P. Borel. Pairwise well-formed words, related to Christoffel words, were considered in the communication by D. Clampitt (joint work with T. Noll).

Several contributions concerned the topic of repetitions and pattern avoidance. In his invited lecture (a joint work with L. Ilie and L. Tinta), M. Crochemore presented recent advances on the so-called “runs” conjecture, stating that a word of length $n$ can contain at most $n$ repetitions of maximal length.

There were three communications concerning the repetition threshold and a generalization of it (M. Rao; F. Fiorenzi - joint work with P. Ochem; E. Vaslet): in particular, M. Rao presented a proof of the 37-year-old Dejean conjecture in the few still open cases. There were also three communications concerning word or pattern avoidance by S. Puzynina, D. Merlini (joint work with R. Sprugnoli), A. Mikhailova and two communications on partial words by F. Blanchet-Sadri (a joint work with R. Mercas and K. Wetzler and a joint work with B. Shirey).

Other aspects of combinatorics on words that have been considered in the selected communications were abelian properties of words (K. Saari, joint work with G. Richomme and L. Q. Zamboni), connections between morphic and automatic words and number theory (D. Krieger, joint work with Y. Bugeaud and J. Shallit) and relations between (infinite) permutations and (infinite) words (A. Frid, joint work with S. V. Avgustinovich and T. Kamae). Classical combinatorics of permutations of a finite set was also the topic of the talk by V. Vajnovszki.

All participants manifested their appreciation of the scientific program and the overall organization of the event. The 7th edition of WORDS was under the auspices of the European Association for Theoretical Computer Science (EATCS). The conference was also sponsored by the ESF Project “Automata: from Mathematics to Applications (AutoMathA)”, by the MIUR project (PRIN) “Mathematical aspects and emerging applications of automata and formal languages”, by the Dipartimento di Informatica e Applicazioni “R.M. Capocelli” (University of
Salerno, Italy) and by the Faculty of Science (University of Salerno, Italy). A special issue of Theoretical Computer Science A will be devoted to extended versions of selected papers from WORDS 2009. The web site of the conference has been available since October 2008 at the URL: http://words2009.dia.unisa.it
ABSTRACTS OF

PHD THESIS
Abstract

Graph transformation has many application areas in computer science, such as software engineering or the design of concurrent and distributed systems. Being a visual modeling technique, graph transformation has the potential to play a decisive role in the development of increasingly larger and complex systems. However, the use of visual modeling techniques alone does not guarantee the correctness of a design. In context of rising standards for trustworthy systems, there is a growing need for the verification of graph transformation systems and programs. The research of appropriate methods for this purpose is the topic of this thesis.

The primary goal is to obtain the capability to decide graphical program specifications. These specifications consist of a graphical precondition, a graph program, and a graphical postcondition. As usual, such a specification is said to be correct, if all those system states satisfy the postcondition that are reachable by applying the program on a start state satisfying the precondition. In the considered programs, the selection, deletion, addition and deselection of a graph's nodes and edges are the elementary constructs that can be composed to more complex programs by non-deterministic choice, sequential composition and iteration. The resulting programming language is computationally complete and is able to model transactions that deal with an unbounded number of nodes and edges. As language for the specification of state properties, graph conditions are investigated and used. We show that graph conditions provide an intuitive formalism for first-order structural properties and are suited to infer knowledge about the behavior of graph transformation systems and programs.

According to Dijkstra, the correctness of program specifications can be shown by constructing a weakest precondition of the program relative to the postcondition and checking whether the specified precondition implies the weakest pre-
condition. Hence the correctness problem of program specifications is reduced to an implication problem of conditions. In this thesis, it is shown how to construct weakest preconditions for graph programs and graph conditions. Following a dual approach, a sound and complete satisfiability algorithm for graph conditions is investigated and a fragment of conditions is identified, for which the algorithm decides. On the other hand, a resolution-based calculus for graph conditions is presented and its soundness is proven. Implementations of the aforementioned deciders for conditions are compared with existing theorem provers and satisfiability solvers for first-order logic by verifying three case studies: a railroad control, an access control for computer systems, and, as an external example, a car platoon maneuver protocol.

The research is done within the framework of the so-called weak adhesive high-level replacement categories. Therefore, the results will be applicable to different kinds of graph replacement systems and Petri nets, providing theoretical fundamentals and general concepts for the development of correct transformation-based systems and programs.

Table of Contents

1 Introduction ......................................................... 1
2 Preliminaries .......................................................... 7
   2.1 The category of graphs ........................................... 7
   2.2 Weak adhesive HLR categories ................................. 15
3 Conditions ............................................................. 19
   3.1 Nested conditions ............................................... 19
   3.2 Comparison of $M$- and $A$-satisfiability .................... 23
   3.3 Comparison of graph conditions and graph formulas .......... 31
   3.4 Summary and discussion ....................................... 42
4 Programs and transformation systems ............................... 43
   4.1 Programs with interface ........................................ 43
   4.2 Elimination of $\text{Fix}$ statements .............................. 51
   4.3 Related concepts ............................................... 53
   4.4 Summary and discussion ....................................... 55
5 Correctness of program specifications .............................. 57
   5.1 Program specifications .......................................... 57

184
The Bulletin of the EATCS

5.2 Basic transformations on conditions ......................... 58
5.3 Weakest liberal preconditions ................................. 67
5.4 Weakest invariants ............................................. 74
5.5 Related concepts ............................................... 79
5.6 Summary and discussion ....................................... 82

6 The implication problem of conditions ........................... 83
6.1 Implication, tautology and satisfiability ....................... 83
6.2 Satisfiability solver SeekSat .................................... 87
6.3 Theorem prover ProCon .......................................... 98
6.4 Related concepts ............................................... 110
6.5 Summary and discussion ....................................... 112

7 Case studies .......................................................... 115
7.1 General set-up .................................................. 115
7.2 Railroad control system ........................................ 117
7.3 Car platoon maneuver protocol ................................ 124
7.4 Access control for computer systems ........................... 133
7.5 Summary and discussion ....................................... 141

8 Conclusion ................................................................. 145
8.1 Results ............................................................. 145
8.2 Open problems and future work ................................ 147

A Partial monomorphisms ............................................. 149
B Proof ................................................................. 153
Bibliography ............................................................. 157
Index ................................................................. 173
Curriculum Vitae ...................................................... 177

Author’s correspondence address
Karl-Heinz Pennemann
University of Oldenburg
Department of Computing Science
D-26111 Oldenburg, Germany
Pennemann@Informatik.Uni-Oldenburg.de
http://formale-sprachen.informatik.uni-oldenburg.de/persons/pennemann/
EATCS

HISTORY AND ORGANIZATION

EATCS is an international organization founded in 1972. Its aim is to facilitate the exchange of ideas and results among theoretical computer scientists as well as to stimulate cooperation between the theoretical and the practical community in computer science.

Its activities are coordinated by the Council of EATCS, which elects a President, Vice Presidents, and a Treasurer. Policy guidelines are determined by the Council and the General Assembly of EATCS. This assembly is scheduled to take place during the annual International Colloquium on Automata, Languages and Programming (ICALP), the conference of EATCS.

MAJOR ACTIVITIES OF EATCS

- Organization of ICALP;
- Publication of the “Bulletin of the EATCS;”
- Award of research and academic careers prizes, including the “EATCS Award,” the “Gödel Prize” (with SIGACT) and best papers awards at several top conferences;
- Active involvement in publications generally within theoretical computer science.

Other activities of EATCS include the sponsorship or the cooperation in the organization of various more specialized meetings in theoretical computer science. Among such meetings are: ETAPS (The European Joint Conferences on Theory and Practice of Software), STACS (Symposium on Theoretical Aspects of Computer Science), MFCS (Mathematical Foundations of Computer Science), LICS (Logic in Computer Science), ESA (European Symposium on Algorithms), Conference on Structure in Complexity Theory, SPAA (Symposium on Parallel Algorithms and Architectures), Workshop on Graph Theoretic Concepts in Computer Science, International Conference on Application and Theory of Petri Nets, International Conference on Database Theory, Workshop on Graph Grammars and their Applications in Computer Science.

Benefits offered by EATCS include:
- Subscription to the “Bulletin of the EATCS;”
- Reduced registration fees at various conferences;
- Reciprocity agreements with other organizations;
- 25% discount when purchasing ICALP proceedings;
- 25% discount in purchasing books from “EATCS Monographs” and “EATCS Texts;”
- Discount (about 70%) per individual annual subscription to “Theoretical Computer Science;”
- Discount (about 70%) per individual annual subscription to “Fundamenta Informaticae;”

(1) THE ICALP CONFERENCE

ICALP is an international conference covering all aspects of theoretical computer science and now customarily taking place during the second or third week of July. Typical topics discussed during recent ICALP conferences are: algorithms, computational complexity, game theory, automata theory, formal language theory, logic, semantics, and theory of programming languages, foundations of networked computation, parallel, distributed, and external memory computing, foundations of logic programming, models of concurrent, distributed and mobile systems, software specification, computational geometry, data types and data structures, models for complex networks, theory of security.
THE BULLETIN OF THE EATCS

SITES OF ICALP MEETINGS:

- Paris, France 1972
- Saarbrücken, Germany 1974
- Edinburgh, Great Britain 1976
- Turku, Finland 1977
- Udine, Italy 1978
- Graz, Austria 1979
- Noordwijkerhout, The Netherlands 1980
- Haifa, Israel 1981
- Aarhus, Denmark 1982
- Barcelona, Spain 1983
- Antwerp, Belgium 1984
- Nafplion, Greece 1985
- Rennes, France 1986
- Karlsruhe, Germany 1987
- Tampere, Finland 1988
- Stresa, Italy 1989
- Warwick, Great Britain 1990
- Madrid, Spain 1991
- Wien, Austria 1992
- Lund, Sweden 1993
- Jerusalem, Israel 1994
- Szeged, Hungary 1995
- Paderborn, Germany 1996
- Bologna, Italy 1997
- Aalborg, Denmark 1998
- Prague, Czech Republic 1999
- Genève, Switzerland 2000
- Heraklion, Greece 2001
- Malaga, Spain 2002
- Eindhoven, The Netherlands 2003
- Turku, Finland 2004
- Lisabon, Portugal 2005
- Venezia, Italy 2006
- Wroclaw, Poland 2007
- Reykjavik, Iceland 2008
- Rhodes, Greece 2009

(2) THE BULLETIN OF THE EATCS

Three issues of the Bulletin are published annually, in February, June and October respectively. The Bulletin is a medium for rapid publication and wide distribution of material such as:

- EATCS matters;
- Technical contributions;
- Columns;
- Surveys and tutorials;
- Reports on conferences;
- Reports on computer science departments and institutes;
- Open problems and solutions;
- Abstracts of Ph.D.-Theses;
- Entertainments and pictures related to computer science.

Contributions to any of the above areas are solicited, in electronic form only according to formats, deadlines and submissions procedures illustrated at http://www.eatcs.org/bulletin. Questions and proposals can be addressed to the Editor by email at bulletin@eatcs.org.

(3) OTHER PUBLICATIONS

EATCS has played a major role in establishing what today are some of the most prestigious publication within theoretical computer science.

These include the EATCS Texts and the EATCS Monographs published by Springer-Verlag and launched during ICALP in 1984. The Springer series include monographs covering all areas of theoretical computer science, and aimed at the research community and graduate students, as well as texts intended mostly for the graduate level, where an undergraduate background in computer science is typically assumed.

Updated information about the series can be obtained from the publisher.

The editors of the series are W. Brauer (Munich), J. Hromkovic (Aachen), G. Rozenberg (Leiden), and A. Salomaa (Turku). Potential authors should contact one of the editors.

EATCS members can purchase books from the series with 25% discount. Order should be sent to:
Prof. Dr. G. Rozenberg, LIACS, University of Leiden, P.O. Box 9512, 2300 RA Leiden, The Netherlands who acknowledges EATCS membership and forwards the order to Springer-Verlag.

The journal *Theoretical Computer Science*, founded in 1975 on the initiative of EATCS, is published by Elsevier Science Publishers. Its contents are mathematical and abstract in spirit, but it derives its motivation from practical and everyday computation. Its aim is to understand the nature of computation and, as a consequence of this understanding, provide more efficient methodologies.

The Editors-in-Chief of the journal currently are G. Ausiello (Rome), D. Sannella (Edinburgh), G. Rozenberg (Leiden), and M.W. Mislove (Tulane).

### ADDITIONAL EATCS INFORMATION

For further information please visit [http://www.eatcs.org](http://www.eatcs.org), or contact the President of EATCS:

Prof. Dr. Burkhard Monien, Department of Computer Science
Universität Paderborn, Via Salaria Fürstenalle 11, 33102 Paderborn, Germany
Email: president@eatcs.org

### EATCS MEMBERSHIP

#### DUES

The dues are €30 for a period of one year. A new membership starts upon registration of the payment. Memberships can always be prolonged for one or more years.

In order to encourage double registration, we are offering a discount for SIGACT members, who can join EATCS for €25 per year. Additional €25 fee is required for ensuring the air mail delivery of the EATCS Bulletin outside Europe.

#### HOW TO JOIN EATCS

You are strongly encouraged to join (or prolong your membership) directly from the EATCS website [www.eatcs.org](http://www.eatcs.org), where you will find an online registration form and the possibility of secure online payment. Alternatively, a subscription form can be downloaded from [www.eatcs.org](http://www.eatcs.org) to be filled and sent together with the annual dues (or a multiple thereof, if membership for multiple years is required) to the Treasurer of EATCS:

Prof. Dr. Dirk Janssens, Dept. of Math. and Computer Science, University of Antwerp
Middelheimlaan 1, B-2020 Antwerpen, Belgium
Email: treasurer@eatcs.org, Tel: +32 3 2653904, Fax: +32 3 2653777

The dues can be paid (in order of preference) by VISA or EUROCARD/MASTERCARD credit card, by cheques, or convertible currency cash. Transfers of larger amounts may be made via the following bank account. Please, add €5 per transfer to cover bank charges, and send the necessary information (reason for the payment, name and address) to the treasurer.

Fortis Bank, Jules Moretuslei 229, B-2610 Wilrijk, Belgium
Account number: 220–0596350–30–01130
IBAN code: BE 15 2200 5963 5030, SWIFT code: GEBABEBB 18A