Distributed Computing Column 44
2011 in Review

Idit Keidar
Dept. of Electrical Engineering, Technion
Haifa, 32000, Israel
idish@ee.technion.ac.il

The last column of the year is dedicated, as always, to a review of distributed computing awards and conferences in 2011.

The 2011 Edsger W. Dijkstra Prize in Distributed Computing was awarded to Hagit Attiya, Amotz Bar-Noy, and Danny Dolev, for their fundamental work on "Sharing Memory Robustly in Message-Passing Systems", which appeared in the Journal of the ACM (JACM) 42(1):124-142 (1995). This seminal paper, often dubbed ABD, was the first to establish the equivalence between asynchronous fault-tolerant message passing systems and shared memory systems. Specifically, the paper shows that, unlike consensus, atomic read-write registers can be implemented in fault-prone asynchronous shared memory systems. The implication of this equivalence is that algorithms designed in the more abstract shared memory model can be directly be implemented in message passing systems. The ABD construction lies at the heart of many distributed storage systems built in the last decade, including scalable replicated storage used in today's large data centers and clouds. The Dijkstra Prize is jointly awarded by PODC and DISC; Amotz Bar-Noy and Hagit Attiya received it on behalf of the three authors in DISC this year (see picture). The full award citation appears earlier in this issue of SIGACT News (and on the award's web page[1]), so I do not repeat it here. Instead, I include here Hagit Attiya's account of how the paper came to be, in the spirit of her address at the Edsger W. Dijkstra Prize awarding ceremony.

The Prize for Innovation in Distributed Computing was awarded for the third time this year. The prize was established by SIROCCO to "recognize individuals whose research contributions on the relationships between information and efficiency in decentralized computing have expanded the collective investigative horizon by formulating new problems, or identifying new research areas, that were at the time of their introduction unorthodox and outside the mainstream." The prize

[1] http://www.podc.org/dijkstra/2011.html

Hagit Attiya and Amotz Bar-Noy accept the Dijkstra Prize, which they jointly won with Danny Dolev. Picture by Rotem Oshman.

is awarded in SIROCCO, and its third recipient is David Peleg. David Peleg is recognized for his many important innovative contributions to distributed computing, including in local computing, robot computing, dynamic monopolies, sparse spanners, and compact routing and labeling schemes. The full award citation appears below.

We proceed with reviews of the leading conferences in distributed computing. Continuing with past tradition, I have invited students who have won Best Paper or Best Student Paper Awards in PODC– the ACM Symposium on Principles of Distributed Computing– and DISC– the International Symposium on DIStributed Computing– to review them.

The review of PODC is by Maryam Helmi of the University of Calgary, Canada, who won the PODC 2011 Best Paper Award for her paper "The Space Complexity of Long-Lived and One-Shot Timestamp Implementations", co-authored with Lisa Higham, Eduardo Pacheco, and Philipp Woelfel. This paper studies the problem of implementing timestamps from atomic multi-reader/multi-writer registers in an asynchronous distributed system of n processes with distinct identifiers. The paper establishes a space complexity gap between long-lived and one-shot timestamps. Specifically, it proves an $\Omega(n)$ lower bound (asymptotically matching the known upper bound) for long-lived timestamps, and also provides a one-shot algorithm and lower bound with space complexity $\Theta(\sqrt{n})$. The PODC 2011 Best Student Paper was awarded to Leonid Barenboim, for the paper "Distributed Deterministic Edge Coloring using Bounded Neighborhood Independence", co-authored with his advisor Michael Elkin.

DISC is reviewed by Michael Hakimi and Adam Morrison, who have won the DISC 2011 Best Student Paper Award for their paper "Fast and Scalable Rendezvousing" co-authored with their advisor Yehuda Afek. This paper presents a new highly scalable, high throughput asymmetric rendezvous system, where consumer and producer threads are matched. Their system outperforms prior implementations. The paper further presents a highly scalable and competitive stack implementation. The DISC 2011 Best Paper was awarded to Pierre Fraigniaud, Sergio Rajsbaum, and Corentin Travers, for their paper "Locality and Checkability in Wait-Free Computing".

The final review is of SIROCCO, the venue for "unconventional" ideas in distributed computing. It is reviewed by Adrian Kosowski, Dominik Pająk, and Zuzanna Stamirowska. In all reviews, you will find fun information about the venues, as well as technical content.

Many thanks to Hagit, Maryam, Michael, Adam, Adrian, Dominik, and Zuzanna for their contributions!

Also this year, we saw the inception of a podcast on distributed computing called podccast[2]. And in this column, we have touched in 2011 upon computing over dynamic networks (in March), combining game theory and fault tolerance in distributed computing (in June), and using social networks to overcome Sybil attacks (in September). In 2012, I hope to touch more on dynamic networks, and also on synthesis of distributed and parallel programs, random walks, and perhaps the uses of topology in distributed computing. Best wishes for an exciting and stimulating 2012!

Call for contributions: I welcome suggestions for material to include in this column, including news, reviews, open problems, tutorials and surveys, either exposing the community to new and interesting topics, or providing new insight on well-studied topics by organizing them in new ways.

[2] http://www.podccast.webnode.com

Sharing Memories, Robustly

Hagit Attiya
Department of Computer Science
Technion hagit@cs.technion.ac.il

It is an honor to receive the Dijkstra Prize with Amotz Bar-Noy and Danny Dolev for the ABD simulation of shared memory in message-passing systems.

The award citation does a great job describing the paper's contribution and impact; additional technical description of the result and follow-up work appears in [1]. Here, I would like to offer my personal account of how this paper came about.

The starting point for our interest in relating the shared-memory and message-passing models goes back to the mid-eighties, when Amotz and I were working on our PhDs under the supervision of Danny, at the Hebrew University in Jerusalem. We investigated problems that are solvable in situations where consensus is impossible, e.g., asynchronous failure-prone systems, and in particular, studied the *renaming* problem [4].

At the time, algorithms for the shared-memory model were treated with suspicion, being complicated and un-intuitive; many of these were subtle constructions of various registers from weaker types of registers, and they often had mistakes. For this, and other reasons, we looked for *message-passing* algorithms for renaming, tolerating crash failures. However, rather than design the algorithms from scratch, by using the basic sending and receiving messages operations, we looked for a communication primitive that will abstract away some of the inconsistency inherent in message-passing systems.

We already figured out the partitioning argument showing that in order to solve renaming, the number of failures, f, must be a minority of the number of processors, n (i.e., $n > 2f$). Additionally, since the start of our PhDs, Danny kept coming back to the paper by Thomas [9], using *majority quorums* to solve consensus in partitioned databases. Putting one and one together, we based our renaming algorithms on the concept of a *stable* vector—a vector whose entries are values received from different processors, a majority of which contain exactly the same value. Stable vectors turned out to be a very useful tool in simplifying the renaming algorithms, which, several years later, turned out to achieve optimal name space.

Shortly afterwards, Amotz and I graduated and left for our respective post-docs, at Stanford and MIT, and Danny went for a sabbatical at IBM Almaden.

Danny and Amotz continued to explore the stable vector abstraction. In PODC 1989, they published a paper [5] showing how a communication primitive, similar to stable vectors, can be used to hide the differences between message-passing and shared-memory algorithms. But still the primitive was not "natural" and it was not clear how expressive and easy-to-use it is.

In summer 1989, I visited the Bay Area ("en route" to PODC, which was held in Edmonton, Canada). By that time, the shared-memory model was becoming cleaner, and easier to employ. The fascination with register constructions gave way to solving more sophisticated problems using atomic registers, for example, atomic snapshot algorithms, or various algorithms for randomized consensus.

Inspired by Upfal and Wigderson's simulation of *synchronous* shared-memory (without failures) on fixed-connection machines [10], we realized that we could get a simulation of an atomic register in a message-passing system. The core simulation is rather simple: The value of the register is replicated among all processors, and, like in stable vectors, the value is obtained from a majority of the replicated values; generality is achieved by tagging values with *timestamps* (consecutive sequence numbers) and picking the value with the largest timestamp from the majority. To achieve atomicity of reads (a property not provided by stable vectors), we followed an idea from the construction of atomic registers from regular registers, namely, a *write-back* of the value read, before returning it.

Atomic registers provided just the right abstraction: a good interface to capitalize on algorithms designed for the shared-memory model, at the same time, allowing a simple and efficient implementation. Efficiency was a key issue since it meant that employing the abstraction did not come at a high cost; simplicity was critical to increase the confidence in the correctness of the construction.

It took some more work, a large part of it dedicated to bounding the timestamps, to obtain and prove the final construction. The paper was submitted to STOC 1990, which rejected it, allegedly because the constructions were too simple. The paper was accepted to PODC 1990, where it closed the program of the conference [2].

The simplicity of the ABD construction (especially with unbounded timestamps) meant it was easy to adapt it to different situations, and people felt safe to employ it in practical systems. This lead to follow-up papers, some of which are more cited than our original paper. One key step was phrasing the construction in terms of general quorums, rather than just the majority quorum, and to consider Byzantine failures [8]; another was to extend it to systems with dynamic failures [7]. A validation of our belief that shared-memory algorithms can serve as a blueprint for message-passing algorithms came from *disk Paxos* [6], one of the best ways to understand Paxos state replication, which is employed in many contemporary distributed servers.

In closing, we would like to thank the distributed computing community for taking an interest in our work, in particular, the award committee and the nominators, and I would like to thank Amotz Bar-Noy and Danny Dolev for inviting me to join this research.

References

[1] Hagit Attiya. Robust Simulation of Shared Memory: 20 Years After. *EATCS Distributed Computing Column*, 100:99–113, February 2010.

[2] Hagit Attiya, Amotz Bar-Noy, and Danny Dolev. Sharing memory robustly in message-passing systems. *Journal of the ACM*, 42(1):121–132, January 1995. Previous version in PODC 1990.

[3] Hagit Attiya, Amotz Bar-Noy, Danny Dolev, Daphne Koller, David Peleg, and Rudiger Reischuk. Achievable cases in an asynchronous environment. In *Proceedings of the 28th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 337–346, 1987.

[4] Hagit Attiya, Amotz Bar-Noy, Danny Dolev, David Peleg, and Rudiger Reischuk. Renaming in an asynchronous environment. *Journal of the ACM*, 37(3):524–548, July 1990. Previous version in [3].

[5] Amotz Bar-Noy and Danny Dolev. A partial equivalence between shared-memory and message-passing in an asynchronous fail-stop distributed environment. *Mathematical Systems Theory*, 26(1):21–39, 1993. Previous version in PODC 1989.

[6] Eli Gafni and Leslie Lamport. Disk Paxos. *Distributed Computing*, 16:1–20, February 2003. Previous version in DISC 2000.

[7] Nancy A. Lynch and Alexander A. Shvartsman. Rambo: A Reconfigurable Atomic Memory Service for Dynamic Networks. In *Proceedings of the 16th International Conference on Distributed Computing (DISC)*, pages 173–190, 2002.

[8] Dahlia Malkhi and Michael K. Reiter. Byzantine quorum systems. *Distributed Computing*, 11(4):203–213, 1998. Previous version in STOC 1997.

[9] Robert H. Thomas. A majority consensus approach to concurrency control for multiple copy databases. *ACM Transactions on Database Systems*, 4(2):180–209, 1979.

[10] Eli Upfal and Avi Wigderson. How to share memory in a distributed system. *Journal of the ACM*, 34(1):116–127, 1987. Previous version in FOCS 1984.

Prize for Innovation in Distributed Computing 2011 awarded to David Peleg

The Prize for Innovation In Distributed Computing is awarded by the *Colloquium on Structural Information and Communication Complexity* (SIROCCO). It is established to recognize individuals whose research contributions on the relationships between information and efficiency in decentralized computing have expanded the collective investigative horizon by formulating new problems, or identifying new research areas, that were at the time of their introduction unorthodox and outside the mainstream. The prize recognizes originality, innovation, and creativity. The recipient of the Prize is chosen among the nominated persons for the current year.

The Award Committee has selected **David PELEG** as the recipient of this year's Prize for Innovation In Distributed Computing.

The prize is given in recognition of David Peleg for his many and important innovative contributions to distributed computing, as illustrated by several papers, including some that appeared in the proceedings of past SIROCCO meetings. These papers tackled a wide variety of problems including local computing, robot computing, and the design and analysis of dynamic monopolies, sparse spanners, and compact routing and labeling schemes, as illustrated, in particular, by the following papers:

- B. Awerbuch, O. Goldreich, D. Peleg, R. Vainish. A Trade-Off between Information and Communication in Broadcast Protocols. J. ACM 37(2): 238-256 (1990)
- R. Cohen, D. Peleg. Robot Convergence via Center-of-Gravity Algorithms. SIROCCO 2004: 79-88
- U. Feige, D. Peleg, P. Raghavan, E. Upfal: Randomized Broadcast in Networks. Random Struct. Algorithms 1(4): 447-460 (1990)
- D. Peleg. Size Bounds for Dynamic Monopolies. SIROCCO 1997: 151-161
- D. Peleg. Informative labeling schemes for graphs. MFCS 2000: 579-588
- D. Peleg, E. Upfal. A trade-off between space and efficiency for routing tables. J. ACM 36(3): 510-530 (1989)

The paper *A Trade-Off between Information and Communication in Broadcast Protocols* is a prominent illustration of all studies dealing with compromises between the ability of solving tasks efficiently in a distributed manner, and the amount of global information provided to the individual computing entities about their environment. It is also one of the many seminal papers of David Peleg about information dissemination problems. In this latter framework, the paper *Randomized Broadcast in Networks* is one of the most referenced papers in connection to the analysis of gossip protocols.

The paper *A trade-off between space and efficiency for routing tables* is definitively one of the most influential papers in the theory of compact routing. The richness of the concepts introduced in this paper was the source of many related topics, including the whole framework of compact informative labeling schemes, yet another fundamental research domain initiated by David Peleg (cf. the paper *Informative labeling schemes for graphs*).

David Peleg pioneered so many research domains in the framework of distributed computing that it would be impossible to provide an exhaustive list. In addition to the aforementioned topics, he was interested in robot computing (see, e.g., *Robot Convergence via Center-of-Gravity Algorithms*), in the evolution of dynamic discrete systems (see, e.g., *Size Bounds for Dynamic Monopolies*), and, of course, in graph problems with applications to distributed computing.

By his results and ideas, David Peleg has enriched distributed computing considerably. In particular, he is one of the most significant contributors to the theory of *local* computing, including the study of trade-offs between "local knowledge" and "global performances", which is precisely what SIROCCO is about. His book *Distributed computing: a locality-sensitive approach* (SIAM) is a prominent reference describing the methodology for addressing locality in distributed computing.

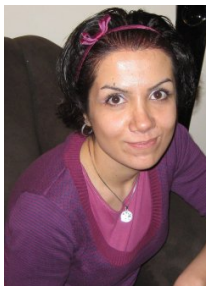
The prize has been officially delivered at the the Business meeting of the 18th International Colloquium on Structural Information and Communication Complexity (SIROCCO), June 26-29, 2011, Gdansk, Poland.

Award Committee 2011:

- Pierre Fraigniaud (CNRS and University Paris Diderot)
- Shay Kutten (Technion)
- Boaz Patt-Shamir (Tel Aviv University)
- Alexander A. Shvartsman (University of Connecticut)
- Paul Spirakis (CTI and University of Patras)

A Review of PODC 2011

Maryam Helmi
Department of Computer Science
University of Calgary, Canada
mhelmikh@ucalgary.ca



1 Introduction

The 30th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC 2011) was held on June 6-8, 2011, in *San Jose, California*, as part of the 5th Federated Computing Research Conference (FCRC). 34 regular papers and 31 brief announcements presented in this conference were selected among 129 regular submissions, and 34 additional brief announcement submissions. The conference took place at the San Jose Convention Center.

Figure 1: San Jose Convention Center.

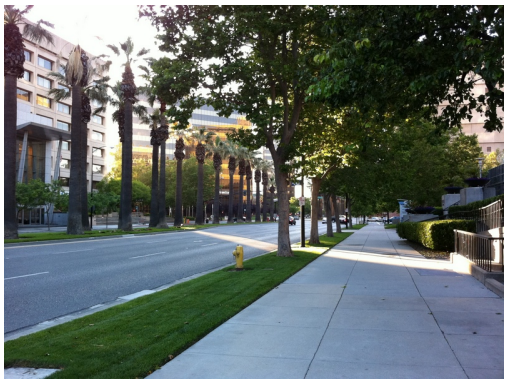


Figure 2: Nice walk to the reception.

2 FCRC Turing Lecture and Plenary Talks

Leslie G. Valiant, who received the 2010 Turing Award, presented the Turing lecture titled *"The Extent and Limitations of Mechanistic Explanation of Nature"* on Sunday evening. He introduced the role of computational learning theory in explaining the organization of physically realizable components of nature. Professor Valiant's talk was one of the highlights of the FCRC conference. His thesis was that the changing and increasing functionality of species is a form of computational

learning. He stated that the Darwinian theory of evolution lacks quantitative explanations. Computer simulations using a variety of algorithms including genetic algorithms have not been able to evolve a mechanism that captures the complexity of existing life forms. He claimed that one can isolate the rate of the change of functionality of biological circuits and formulate it in terms of computational learning. Evidence from other sciences overwhelmingly suggests that humans get here by evolution. He challenged the computer science community to use computational learning theory to explain exactly how this was done.

PODC as a part of FCRC had a plenary talk every day right before lunch. David Ferrucci, an IBM researcher and the principle investigator of the IBM's Watson/Jeopardy project, gave his talk on Monday: *"IBM's Watson/DeepQA"*. He presented a brief explanation of why IBM picked this project and the technical challenges in this project. On Tuesday Ravi Kannan, a Microsoft researcher, gave his talk on *"Algorithms: Recent Highlights and Challenges"*. This talk discussed an interesting vector-based algorithmic approach in information technology. The plenary talk on Wednesday, titled *"Warehouse-Scale Computing: Entering the Teenage Decade"*, was given by Luiz Andre Barroso. He discussed some challenging aspects regarding new technologies such as flash memories and faster processors. He also talked about the challenge to satisfy expectations like real time searching, which is important for warehouse-scale computing.

3 Conference Program

The conference contained eleven sessions of regular papers and three sessions devoted to brief announcements.

Consensus is one of the best studied problems in distributed algorithms; it was the topic of several papers again this year. The first session covered consensus and agreement problem. This was the session that I enjoyed most. Presentations started with Oshman's talk on *"Coordinated Consensus in Dynamic Networks"*, which won the best student presentation award. In this paper, Kuhn, Moses and Oshman studied the problems of consensus, simultaneous consensus, and almost simultaneous consensus where the communication graph changes dynamically under the control of a worst case adversary. The next talk was *"Error-Free Multi-Valued Consensus with Byzantine Failures"*. The authors, Liang and Vaidya incorporated the length of the consensus value into the running time complexity. The session ended with a paper titled *"Byzantine Agreement with Homonyms"*. Delporte-Gallet, Fauconnier, Guerraoui, Kermarrec, Ruppert and Tran-The looked at Byzantine Agreement under the plausible assumption that several processes may be assigned the same identifier. They showed how many identifiers are needed for agreement in a synchronous network, and that this number increases for the partially synchronous case.

I appreciated seeing several papers that tackled problems by applying randomization. The paper *"MIS on Trees"* by Lenzen and Wattenhofer presented an interesting randomized algorithm that finds the maximal set of mutually non-adjacent nodes in a graph in the synchronous message passing model. The running time of this algorithm is approximately the square root of the best previously achieved. Alistarh gave a very good presentation of the paper *"Optimal-Time Adaptive Tight Renaming, with Applications to Counting"*, co-authored with Aspnes, Censor-Hillel, Gilbert and Zadimoghaddam. This paper contained two randomized renaming algorithms against an adaptive adversary. The second algorithm used an elegant idea to transform any sorting network to an adaptive renaming algorithm with step complexity $O(\log k)$, where k is the contention in the current execution.

Figure 3: Security and consistency session.

Message passing algorithms for networks always get some attention. The paper *"Distributed Graph Coloring in Few Rounds"* by Kothapalli and Pemmaraju, presented an optimum distributed algorithm that colors the nodes of an oriented graph with a small number of colors while bounding the number of rounds of computation. Another interesting paper was *"Toward more Localized Local Algorithms: Removing Assumptions concerning Global Knowledge"* by Korman, Sereni and Viennot. They proposed a rather general method to remove the reliance of processors on global information of the network such as the maximum degree or the number of nodes.

Mobile and radio networks was one of this year's hot topics. *"Structuring Unreliable Radio Networks"* by Censor-Hillel, Gilbert, Kuhn, Lynch and Newport studied how to exploit a link detection mechanism to construct a routing backbone in radio networks and how reliable the detection mechanism must be to be effective. In the talk on *"Tight Bounds on Information Dissemination in Sparse Mobile Networks"*, the authors Pettarin, Pietracaprina, Pucci and Upfal studied the dynamics of information dissemination between agents. They showed that the broadcast time for a system below the percolation point does not depend on the transmission radius. *"Order Optimal Information Spreading Using Algebraic Gossip"* by Avin, Borokhovich, Censor-Hillel and Lotker addressed gossip-based information spreading with bounded message size. The authors provided new bounds that are optimal for certain topologies.

Among all the remaining papers, some particularly appealed to me. One was *"A Complexity Separation Between Cache-Coherent and Distributed Shared Memory Models"* by Golab. He gave a nice talk proving that the RMR (remote memory reference) complexity differs between cache-coherent and distributed shared memory models. The paper *"The Complexity of Robust Atomic Storage"* by Dobre, Guerraoui, Majuntke, Suri and Vukolić studied the time complexity of implementing atomic storage from faulty components by providing both algorithms and lower-bounds. *"On The Power of Hardware Transactional Memory to Simplify Memory Management"* by Dragojević, Herlihy, Lev and Moir contained interesting experimental research showing how hardware transactional memory can be used to simplify and replace the memory for concurrent data structures.

The last session on Monday was dedicated to award-winning papers. The paper *"Distributed Deterministic Edge Coloring using Bounded Neighborhood Independence"* by Barenboim and Elkin received the best student paper award. It described an interesting algorithm to solve the edge-coloring problem in the distributed message-passing model. The next talk was *"The Space Complexity of Long-Lived and One-Shot Timestamp Implementations"* by Helmi, Higham, Pacheco and Woelfel, which received the best paper award.

Figure 4: The break time between sessions

There were many other notable and excellent presentations and papers on diverse subjects presented at PODC this year and this review only describes a small subset of those presentations most related to my research interest.

During the banquet, the best papers were announced and there was a warm atmosphere and engaging discussion amongst many of the attendees.

Acknowledgment

Thanks to Maryam Elahi and James King for sharing their photos and to Lisa Higham and Philipp Woelfel for editing assistance.

A Review of DISC 2011

Michael Hakimi Adam Morrison
School of Computer Science
Tel Aviv University
mi5@tau.ac.il adamx@tau.ac.il

The 25th International Symposium on Distributed Computing (DISC) took place on September 19 – 23, 2011 in Rome, Italy. The conference was held at the School of Engineering of Sapienza University, next to the church of *San Pietro in Vincoli*, home of Michelangelo's *Moses* sculpture. The School of Engineering is a gorgeous 15th century building complex that was formerly the church's monastery. In its center is a beautiful courtyard surrounded by covered walkways with open arcades. Affiliated workshops were held at Sapienza University's Anotonio Ruberti Department of Computer and System Sciences building, a short walk from the impressive *Basilica di Santa Maria Maggiore*.

Sapienza University's School of Engineering complex, outside and inside.

The DISC program included 31 paper presentations, 11 brief announcements, five workshops, three invited lectures, two tutorials, and one birthday celebration. On Monday, three workshops were held: on Algorithms and Models for Distributed Event Processing (AlMoDEP), Toward Evolutive Routing Algorithms for scale-free/internet-like Networks (TERANET), and on Theoretical Aspects of Dynamic Distributed Systems (TADDS). The day concluded with a celebration in honor of Nicola Santoro's 60th birthday, followed by the DISC welcome reception. The main symposium opened on Tuesday and concluded on Thursday, followed by Yoram Moses' tutorial "Knowledge Strikes Again" and the workshop on the Theory of Transactional Memory (WTTM) which con-

cluded on Friday. Friday's events also included Christian Cachin's tutorial "From Reliable to Secure Distributed Programming" and DISC's Social Network workshop (SON).

The Program

Each day started with an invited lecture. Presentations of papers and brief announcements followed, with each day ending differently. Tuesday evening featured the Dijkstra Prize award to Hagit Attiya, Amotz Bar-Noy and Danny Dolev. This was followed by the business meeting, where Nicola Santoro reminded everyone that his 2 year term as DISC Steering Committee chair has ended and that Sergio Rajsbaum will be taking his place. Antonio Fernandez was elected as the new SC vice-chair and Shlomi Dolev was elected as a new SC member at-large. On Wednesday afternoon we went on an excursion, after which the DISC Banquet was held and the Best Paper and Best Student Paper awards were presented. Thursday was the final symposium day.

Invited lectures On Tuesday, Andrzej Pelc presented "Deterministic Rendezvous in Networks: Survey of Models and Results". In the problem considered in his talk there are $n \geq 2$ agents starting at distinct initial positions in a network (undirected graph) who must eventually meet using a deterministic algorithm. Andrzej discussed many results about feasibility and cost of rendezvous, which depend on details of the model such as whether agents are anonymous or have labels, execute synchronously or asynchronously, and the amount of memory with which they are equipped. He concluded by outlining other interesting directions to explore. First, what happens when new forms of communication (e.g., wireless) between agents are introduced? Second, some current works assume agents with no memory but who have the ability to take a snapshot of the entire network. What happens if agents are allowed more memory in return for weakening their ability to observe the network? Third, if the network nodes are distinctly labeled, agents can trivially rendezvous at a predetermined node. But what will be the impact if they are supplied with less information?

The following day, Dahlia Malkhi gave a talk titled "Going Beyond Paxos" that described a new storage system which provides strong consistency guarantees without using Paxos. The system is aptly named *Corfu*, after the Greek island just off Paxos. It is a Cluster of Raw/Redundant Flash Units attached to the network that exposes a shared log interface to clients, allowing them to append or read log entries. The sequenced structure of the log offers applications multi-object and transactional consistency guarantees. It also facilitates parallel IO without partitioning data objects by making the log entry the unit partitioned across devices. (As opposed to partitioning by object, which destroys cross-object consistency and may suffer from reduced performance if a single object experiences high write volumes.) Consistency of the log is obtained by exploiting the write-once semantics of flash devices: if two clients try to write the same log entry, only one will succeed. However, such collisions rarely occur because, as an optimization, a Corfu client obtains the next free log entry from a central shared *token server*. A prototype Corfu cluster achieves half a million IO operations per second, close to the theoretical limit of its hardware.

The last invited lecture was by Peter Widmayer on "Polygon Reconstruction with Little Information: An Example for the Power of Simple Micro-Robots". The lecture was a beautiful example of how a limited agent can accomplish an interesting task. The setting is of a robot that starts on some vertex of a polygon. The robot sees all vertices visible from its vertex and can move to any one of them (but cannot sense while moving). The goal is to reconstruct the polygon's shape, and the question is what kinds of robots can achieve this. Peter first showed how a robot that can

measure the inner angle between observed vertices can reconstruct the polygon. Next he showed that measuring the angles is not necessary and it's enough to sense the type of angle, provided the robot can also look back. Surprisingly, he proceeded to show that either looking back or sensing angle type were sufficient by themselves to reconstruct the polygon.

Paper presentations It is impossible to review all the presented papers here. We will try to give a taste of the talks, and hopefully convey the diverse type of work that is presented in DISC: on a wide set of topics, ranging from theoretical to practical. Both theoretical and practical ends of the spectrum were covered by the award papers. The Best Paper award was granted to the paper "Locality and Checkability in Wait-Free Computing" by Pierre Fraigniaud, Sergio Rajsbaum and Corentin Travers. They characterize computational tasks that meet various notions of locality which are inherent to the task's specification and independent of the distributed computing model. One such notion of locality is *monotony*, which says that any output for a partial input is a partial output for the full input. (This property is often implicitly assumed, as we expect a processor that wakes up late to still produce an output.) Here the authors explored other locality properties, proving results both about the relations between tasks with different properties and about their wait-free solvability and checkability. Our paper "Fast and Scalable Rendezvousing", co-authored with Yehuda Afek, received the Best Student Paper award. We gave a practical and efficient algorithm for a fundamental synchronization problem in multi-core machines: threads of two types (e.g., producers and consumers) show up and each thread must be matched with a unique thread of the other type. Our algorithm is simple and significantly outperforms prior algorithms in practice.

Quite a few results were presented about computations in the shared memory model. James Aspnes presented "Randomized Consensus in Expected $O(n^2)$ Total Work using Single-Writer Registers." The only previous consensus protocol to achieve this bound uses a single multi-writer termination bit to (essentially) shut down the protocol once a process terminates. The clever idea in this work is to replace this multi-writer bit with an array of single-writer termination bits that propagate via a gossip protocol. On the lower bound front, Gadi Taubenfeld presented "Tight space bounds for ℓ-exclusion," showing that $2n-2$ single-writer bits are required and sufficient to solve ℓ-exclusion by $n > \ell$ processes. Hence ℓ-exclusion is fundamentally different from mutual exclusion and cannot be solved using one single-writer bit per process. Dmitri Perelman described an interesting software transactional memory (STM) system in "SMV: Selective Multi-Versioning STM." SMV keeps old versions of objects in memory as long as they can help a running reader transaction to obtain a consistent snapshot of memory. The alternative design, of keeping just a constant number of old versions, turns out to be susceptible to exponential growth in memory consumption. SMV's data structures ensure that an object version which isn't useful to any transaction will be garbage collected.

As usual, fault tolerance was a popular topic in DISC. Leslie Lamport talked about "Byzantizing Paxos by Refinement." The paper contains the formalization of the following nice idea: one can view a $3f+1$ process Byzantine Paxos consensus algorithm as $2f+1$ good processes emulating the *ordinary* Paxos algorithm (that tolerates message losses or crash faults) in the presence of f malicious processes whose identity isn't known to the good processes. Rachid Guerraoui then raised an intriguing question, of whether Byzantine versions of other message-passing protocols can be similarly obtained. Elette Boyle presented "Leakage-Resilient Coin Tossing." Coin tossing is an important primitive in randomized distributed computation. Yet known coin tossing protocols break if an attacker can observe some of the local information of the honest parties, as often happens in

practical *side-channel* attacks. This work tackles this problem and obtains an $O(1)$-round common coin protocol which is secure even when a constant fraction of each of the honest parties' local state can leak. The model used is a synchronous network with a broadcast channel and secure communication links between pairs of parties, and an interesting problem for further research is whether leakage-resilient coin tossing can be achieved under weaker assumptions.

Networking papers were not absent as well. Srikanth Sastry presented the paper "Leader Election Using Loneliness Detection." The paper compares the computational power of two models for single-hop wireless networks with different collision detection assumptions. In the strong collision detection (SCD) model every node senses a collision, whereas in a weak collision detection (WCD) model only non-transmitting nodes do. It turns out that the crucial problem in WCD systems is to distinguish the case of a lonely node ($n = 1$) which transmits but no one is listening, from the $n > 1$ case where all nodes simultaneously transmit but the resulting collision is not detected. It is shown that a WCD system can implement an SCD system, provided the WCD is equipped with a *loneliness detector* (LD) — an oracle for determining whether or not the system consists of exactly one node. The paper shows how LD can be implemented in WCD systems and obtains upper and lower bounds for the leader election problem in these models. Rotem Oshman talked about "The Complexity of Data Aggregation in Directed Networks." A data aggregation problem requires collecting data from each node in the network, for example to compute a minimum. It is usually assumed that the communication links are bidirectional, and in such a setting data can be aggregated efficiently. This work shows that in directed networks with restricted bandwidth some data aggregation problems become much harder. Furthermore, for networks with small diameter, it is whether the algorithm *knows* this fact that determines the hardness of computing some functions that depend on all their inputs. Gilles Tredan presented "Misleading Stars: What Cannot Be Measured in the Internet?" which studies the accuracy of Internet topology maps obtained from Traceroute measurements. Such a set of measurements can be consistent with many topologies. Gilles argued that this is not a problem if these topologies have a similar structure, which is what this work seeks to find. The model considered in the paper leads to a negative result: inferrable topologies can differ significantly in many aspects. However, a trace set with certain properties can help to determine global properties such as connectivity.

DISC opened and closed with *extreme* talks, in the sense that the models considered were exceedingly harsh in the limitations placed on the computing processes. Bernhard Haeupler talked about "Beeping a Maximal Independent Set." Here the goal is to compute a maximal independent set in a graph where a node can only *beep* or sense that one of its neighbors beeped. Nodes have no knowledge about the topology or size of the network. It is shown that computing an MIS in sub-polynomial time is impossible in these conditions, but various relaxations (such as having a rough bound on the network size) allow an MIS to be computed in $O(\log^2 n)$ or $O(\log^3 n)$ time, depending on the type of collision detection capabilities assumed. And as if wait-free computation in an asynchronous shared memory system wasn't difficult enough, Eli Gafni described an "Oblivious Collaboration" restriction in which a processor is allowed only to propose an output and receive in return a bit indicating whether it may terminate. Thus, a processor's protocol degenerates to a predetermined sequence of output proposals that it makes until terminating. Amazingly, it turns out that $(2k-1)$-Renaming can still be solved in this model, though the protocol uses sequences whose size grows doubly exponentially in n. However, a (nonconstructive) probabilistic argument establishes the existence of a protocol with linear sized sequences. Finding an explicit construction is just one of the many interesting problems left open in this work.

Social event

On Wednesday afternoon we went on an excursion to the *Necropoli della Banditaccia*, an UNESCO World Heritage Site located in the *comune* of Cerveteri. This necropolis contains thousands of tombs organized in a city-like plan and provides the only surviving evidence of Etruscan residential architecture. After the tour we boarded the buses and headed to the DISC Banquet, with Eli Gafni and Nir Shavit singing along from the front row. (Moods were surely just as high on the other buses.) The banquet was held in the beautiful *Villa dei Principi*, the historic residence of Prince Alessandro Torlonia. In addition to great food and drinks, the evening included the Best Paper award ceremony in which Pierre Fraigniaud, Sergio Rajsbaum and Corentin Travers received the Best Paper award and the review authors received the Best Student Paper award. Two other interesting ceremonies were among the evening's highlights. First, Nicola Santoro crowned Sergio Rajsbaum, his successor SC chair, with the official Captain's Hat (important for steering DISC through the rough conference seas). Sergio then proceeded to perform an enchanting dance with Marcos Aguilera, next year's Program Committee chair.

Left: Nicola Santoro looking forward to bestowing the SC chair Captain's Hat on the new SC chair.
Right: Sergio Rajsbaum and Marcos Aguilera performing the traditional new SC/PC chair dance.

Thanks! We conclude this review with a warm thanks to all the people who took part in organizing this year's DISC: The Program Committee chaired by David Peleg; the Steering Committee chaired by Nicola Santoro; the Organizing Committee chaired by Roberto Baldoni; and the wonderful team that handled local arrangements: Carola Aiello, Gabriella Caramagno, Silvia Bonomi, Leonardo Querzoni, Adriano Cerocchi, Luca Montanari, Alessio Pascucci, Marco Platania, Hani Qusa, Bruno Ambrogio, Giuseppe Bracone, Marta Breno, Giuseppe di Luna, Pasquale Fimiani, Mariano Leva and Amir Soltani Nezhad. Your hard work lead to an excellent event that was both fun and interesting. We also thank Adriano Cerocchi and Luca Montanari for some of the photographs in this article.

Next year DISC will be held in Salvador da Bahia, Brazil, known as Brazil's capital of happiness. This will be the first time in DISC's history that the conference takes place below the Equator!

Review of SIROCCO 2011

Adrian Kosowski
INRIA Bordeaux, France
kosowski@labri.fr

Dominik Pająk
INRIA Bordeaux, France
dominik.pajak@labri.fr

Zuzanna Stamirowska
École Polytechnique, France
zuzanna.stamirowska@polytechnique.edu

The *18th International Colloquium on Structural Information and Communication Complexity*, SIROCCO, took place during June 26-29, 2011. This year's venue was the Hotel Hilton in Gdańsk, Poland, located on the banks of the Motława river, just a minute's walk away from the historic heritage of the city's medieval grain port.

Since its first editions, the areas of interest of SIROCCO have focused around the interplay between the efficiency of algorithms and systems, and the availability of information. The meeting attracts participants working in various areas of distributed computing and network optimization. It is known for its lively and relaxed atmosphere, and for its openness to new, unconventional models and ideas.

During the 2011 edition, we had the opportunity to attend two invited talks, by David Peleg and Colin Cooper, a survey talk by Michel Raynal, and 24 contributed talks selected by a Programme Committee chaired by Masafumi Yamashita. The technical program included most of the "evergreen" topics of SIROCCO: problems of mobile agents and robots, distributed algorithms in graphs, and network optimization with special emphasis on wireless networks. This year many talks were also devoted to issues of fault tolerance and randomization in the contexts of networking and of distributed computing. The high point of the conference was undoubtedly the social event on Tuesday, culminating in the presentation of the SIROCCO *Prize for Innovation in Distributed Computing* to David Peleg.

The talks The SIROCCO program opened on Monday morning with sessions on fault tolerance. In his survey talk, Michel Raynal gave an interesting overview of recent advances in the very broad area of shared memory models in distributed computing, focusing on models with crashing processes. Emmanuel Godard considered the solvability of the consensus and broadcast problems in models where messages can be lost, while Heinrich Moser showed how to simulate a real-time model in a classic fault-tolerant distributed model, and vice versa. Devan Sohier presented a self-stabilizing algorithm for the clustering problem. Finally, Michel Raynal, in his second presentation of the day, brought the audience back to shared memory models, introducing and analyzing a general family of tasks.

After the three-course lunch served in the Hilton's restaurant, many of the participants took advantage of the beautiful weather to walk around the old town area. Still, everyone was back for the afternoon session on routing problems. Chi-Ya Kao talked about the diagnosability of

Participants of SIROCCO 2011 in front of the medieval city wall, next to the conference hotel

multiprocessor systems, while Lakshmi Anantharamu considered the broadcasting problem in a model with dynamic packet arrival and with jamming of channels. The final two presentations of the sessions, given by Matthis Függer and Josef Widder, represented the same line of research on routing algorithms with link reversal.

Monday closed with three talks on mobile agents, an area which unfailingly captures the attention of the SIROCCO audience. Branislav Katreniak described an algorithm which solved the problem of convergence on a plane using very weak mobile devices with a bounded visibility radius and very little common knowledge — all set in the asynchronous CORDA model. During the discussion, the audience tried to relax the assumptions, taking away what little power the robots had left in Branislav's model. Next, Philip Brandes presented an algorithm for forming a line by mobile robots on a plane. This algorithm also works in the CORDA model, and its main competitive advantage is the small total distance traversed by each robot. The last speaker in the session was Anissa Lamani. She revisited the problem of gathering asynchronous robots on a ring, considering scenarios with local multiplicity detection and a large odd number of robots. The audience once again attempted to weaken the assumptions, and were left with more than one challenging open problem to discuss after dinner.

Sessions recommenced on Tuesday morning, starting with a colorful invited talk given by Colin Cooper. Colin talked about random walks in graphs, balancing carefully between mathematical precision and convincing intuitions. He showed us some new ideas of speeding up random walks, and different variants of possible interactions between mobile particles.

The morning continued with four presentations on mobile agents, making this the most studied topic of the conference. The first speaker was Samuel Guilbault, who presented a paper on gathering asynchronous oblivious agents with local vision in regular bipartite graphs. The audience, hoping as always to relax the assumptions of the algorithm even further, found out that this time the authors had done more than enough in this respect — the paper did in fact provide an impossibility result for almost all configurations. Next, Alfredo Navarra presented an algorithm for solving the problem of gathering six robots in a ring in the Look-Compute-Move model with global multiplicity detection. The session closed with two interesting talks on the black hole search problem. Shantanu Das talked about black hole search in the ring using as few agents as possible, in a model with



Preparing for the boat trip (left; old town of Gdańsk in the background). A show of dueling during the conference banquet (center). Shantanu Das vanquishes slavic warriors (right). [Photos by Z. Stamirowska and S. Tixeuil]

token-based communication. Balasingham Balamohan considered the same problem, but using a more classical model of communication among agents, and focussing on a different minimization criterion.

After lunch, the focus of attention switched to randomization issues. The first speaker was Tomek Radzik, who introduced the participants to the notion of random walks on hypergraphs. Stefan Dobrev presented the paper "Routing in Carrier-Based Mobile Networks", while Thomas Nowak talked about synchronization of multiple processes in a model where messages can be lost.

The last day began with the second invited talk of the conference, given by the SIROCCO Prize recipient, David Peleg. David gave a very stimulating talk about maps for the SINR (signal-to-interference & noise ratio) model which describes the availability and quality of connection in wireless networks. After a formal definition and some helpful examples, David pointed out the relationship between SINR Maps and Voronoi Diagrams, differences between uniform and non-uniform models, and discussed complexities of algorithms. After the presentation David answered questions on how SINR could be developed still further, for example, to allow for stations which can use multiple frequencies.

The sessions on Wednesday morning stayed in the area of distributed algorithms and wireless networks. Johannes Schneider talked about efficient algorithms for distributed graph coloring, adding the chromatic number to the spectrum of graph parameters useful for bounding the runtime of a distributed algorithm. Guanfeng Liang then presented perhaps the most unusual result of the conference: a strikingly simple algorithm for computing a three party equality function, improving on the bit complexity of all-to-one communication. Ralf Klasing gave an interesting talk about verifying a network using routing table queries, while Vasco Gallotti presented the only game-theoretic paper of the conference, discussing Social Context Congestion Games. The final session on wireless networks had two speakers: Christian Schindlerhauer showed how to use source-receiver distance in synchronization and localization problems, while Adele Rescigno talked about the data gathering problem. Then it was time for a farewell lunch — with the feeling that we were already looking forward to meeting again at next year's SIROCCO.

The social event and prize ceremony On Tuesday afternoon, we met in the marina right in front of the Hilton hotel and took a water taxi to the coastal resort of Sopot. The water taxies

The award ceremony: the participants (foreground from left to right: Bogdan Chlebus, Leszek Gąsieniec, Nitin Vaidya, Andrzej Proskurowski) watch as steering committee chair Pierre Fraigniaud (center) presents the SIROCCO Prize for Innovation in Distributed Computing to David Peleg (right). [Photos by S. Tixeuil]

turned out, in fact, to be high speed reinforced inflatable boats, reaching the speed of 100 km/h on the open sea. The trip provided a nice dose of adrenalin — leaving many of us asking for more, and some completely speechless! The banquet was held within walking distance from the Sopot pier, in an early medieval settlement which was also a place of archeological research. Wooden huts, fire place and two goats created a nice frame for an informal feast which lasted until night hours. A wide selection of traditional Polish high cuisine was served: duck with apples and cranberries, lamb in vegetables, and zander fish in boletus mushrooms.

After coffee, the steering committee chair, Pierre Fraigniaud, announced that the SIROCCO Prize committee had selected David Peleg as its recipient. In the laudation, Pierre emphasized David's enormous impact in the field of distributed computing, and illustrated his important innovative contributions with several papers, including some that appeared in the proceedings of past SIROCCO meetings.

Later that evening, entertainment was provided by a troupe of medieval warriors and a harpist. The warriors, who had already greeted the guests at the gate at the start of the banquet by striking their shields with swords, provided an impressive proof of their skills in fights. The harpist gave a concert of British, Irish and Polish ballads, which reached the hearts of the more nostalgic and music-sensitive part of audience. For those of us who preferred a more active way of spending time, the chance to practice in the archery range or the javelin throw was a tempting alternative. When everyone was already feeling quite relaxed, a troupe of warriors prepared a surprise, presenting some old slavic games. In one of them, each participant had to run in circles 10 times around a standing pickaxe (a wooden one), with their head touching the pickaxe, before running around the settlement. It was great fun for those of us who could be persuaded to take part — and even greater for the audience!

Perspectives This year we welcomed Shay Kutten as the new SIROCCO steering committee chair, taking the place of Pierre Fraigniaud. Among the challenges for SIROCCO, Pierre mentioned increasing student participation and keeping up the meeting's reputation for touching on new and unconventional topics. We also welcomed the first ever elected member of the steering committee. This role went Ralf Klasing after a popular vote during the business meeting.

We are all looking forward to the future of SIROCCO, and to meeting again at its next editions. SIROCCO 2012 will take place in Iceland, while the 20th edition of SIROCCO in 2013 will be held in a coastal resort on the beaches of Southern Italy.