Towards a Theory of Networked Computation[1]

Joan Feigenbaum
http://www.cs.yale.edu/~jf

Michael Mitzenmacher
http://www.eecs.harvard.edu/~michaelm

The increasing prominence of the Internet, the Web, and large data networks in general has profoundly affected social and commercial activity. It has also wrought one of the most profound shifts in Computer Science since the field's inception. Traditionally, Computer-Science research focused primarily on understanding how best to design, build, analyze, and program computers. Research focus has now shifted to the question of how best to design, build, analyze, and operate networks. How can one ensure that a network created and used by many autonomous organizations and individuals functions properly, respects the rights of users, and exploits its vast shared resources fully and fairly?

The Theory of Computation (ToC) community can help address the full spectrum of research questions implicit in this grand challenge by developing a Theory of Networked Computation (ToNC). In our roles as members of the SIGACT Committee on Funding for Theoretical Computer Science [Karp], we have been working with colleagues to chart a research agenda in ToNC and to help secure funding for this research. Two ToNC workshops[2] were held during the Spring of 2006. Our purpose here is to summarize very briefly the findings of those workshops, to point the ToC community to a website [ToNC] with detailed information about the workshops and the evolving ToNC agenda, and to stimulate broad participation in ToNC research and advocacy.

ToC research has already evolved with and influenced the growth of the Web, producing interesting results and techniques in diverse problem domains, including search and information retrieval, network protocols, error correction, Internet-based auctions, and security. A more general Theory of Networked Computation could influence the development of new networked systems, just as formal notions of "efficient solutions" and "hardness" have influenced system development for single machines. To develop a full-fledged Theory of Networked Computation, we in the ToC community should build on past achievements both by striking out in new research directions and by continuing along established directions.

Research Goals

Workshop participants identified three broad, overlapping categories of ToNC-research goals:

---

[1] This article is excerpted and adapted from a forthcoming, longer report with the same title.
[2] These workshops and subsequent ToNC-planning activity were supported by NSF grant CCF-0601893.

- **Realizing better networks:** Numerous theoretical-research questions will arise in the design, analysis, implementation, deployment, operation, and modification of future networks.
- **Computing on networks:** Formal computational models of future networks will enable us both to design services, algorithms, and protocols with provable properties and to demonstrate (by proving hardness results) that some networked-computational goals are unattainable.
- **Solving problems that are created or exacerbated by networks:** Not all of the ToNC-research agenda will involve new computational models. The importance of several established theoretical-research areas has risen dramatically as the use of networked computers has proliferated, and some established methods and techniques within these areas are not general or scalable enough to handle the problems that future networks will create. Examples of these areas include massive-data-set algorithmics, error-correcting codes, and random-graph models.

ToNC-research problems in the first category revolve around finding the right primitives and abstractions with which to study networks; they are exemplified in Goel's breakout-group summary slides from the first ToNC workshops [Goel]. Those in the second category revolve around efficient algorithms for "computing on a network" and hardness results showing that some networked-computational goals are unachievable; they are exemplified in Feigenbaum's talk [Feig] and Impagliazzo's breakout-group summary [Impa] at the second ToNC workshop. Those in the third category revolve around (single-machine) algorithmic questions about network design and network-generated data (as exemplified in Byers's breakout-group summary slides from the second ToNC workshop [Byer]) and mathematical questions about network modeling and network behavior (as exemplified in Kleinberg's breakout-group summary slides from the second ToNC workshop [Klei]). For brevity's sake, we will not elaborate further on the technical agenda here but instead encourage you to peruse the slides from all workshop talks and breakout groups by following the links on [ToNC].

Cross-Cutting Issues

Several cross-cutting, high-level issues are relevant to all three categories and arose repeatedly during plenary and breakout sessions at both workshops

- **Incentive compatibility:** Perhaps the most important distinguishing feature of modern networks is that they are simultaneously built, operated, and used by multiple parties with diverse sets of interests and with constantly changing mixes of cooperation and competition. Formal models of networked computation and notions of hardness and easiness of computation will have to incorporate subnetwork autonomy and user self-interest in an essential way.
- **SPUR:** Achieving the broadest possible vision of "networked computation" will require substantial progress on Patterson's SPUR agenda [Patt]. In his words, "we have taken ideas from the 1970s and 1980s to their logical extreme,

providing remarkably fast and cheap computing and communication (C&C) to hundreds of millions of people. … [F]or our new century, we need a new manifesto for C&C: ... Security, Privacy, Usability, and Reliability (SPUR)."

- **Build on success:** Although today's Internet may leave something to be desired with respect to security, privacy, usability, and reliability, it has far surpassed expectations with respect to several important design goals, *e.g.*, flexibility and scalability. Are the new design criteria compatible with the (manifestly successful) old criteria, and, if not, what are our priorities?
- **"Clean slate":** The phrase "clean-slate design" has become a mantra in networking-research forums and in calls for proposals. Not surprisingly, many people have raised the question of whether anything that requires a "clean slate" could ever be brought to fruition in a world in which networked computation is pervasive and mission-critical. From a research perspective, the crucial point is that clean-slate *design* does not presume clean-slate *deployment*. Part of the ToNC agenda is the evaluation of new technologies, methods, algorithms, *etc.* from the perspective of incremental deployability and paths to adoption.
- **Diversity of "networks":** The scope of the networking research agenda is broader than "next-generation Internet," and thus the ToNC agenda must be broader as well. Interesting theoretical questions arise in the study of special-purpose networks (such as the DoD's Global Information Grid); of moderate-sized but functionally innovative networks; of sensor nets and other technologically constrained networks; of mobile networks; and of P2P and other application-layer networks.

Institutional Support of ToNC

The ToC community can pursue the ToNC-research agenda on many fronts and in many ways. Valuable types of research projects include but are not limited to:

- **Small, single-investigator, purely theoretical projects:** By "small," we mean funded at a level sufficient to pay for one or two months' of PI summer salary per year, one or two PhD students per year, and a few incidentals such as conference travel or commodity computers for the project participants.
- **Medium- and large-sized, multi-investigator projects involving both theory and experimentation:** The distinguishing features of such a project are (1) multiple PIs, at least one of whom is a theorist and at least one of whom is an experimentalist and (2) the inclusion of experimental work on a "real problem" arising in a network that can be built or at least envisioned in the current technological environment. Funding levels for these projects can range from anything that is bigger than "small" up to several million dollars per year.

Program Directors in NSF's Computer and Network Systems Division have explicitly welcomed the type of medium- and large-sized project proposal described here, and the "distinguishing feature" text above comes from them. Careful consideration was given at the workshops to whether small, purely theoretical projects are equally important for success of the ToNC agenda, and participants decided that they are, for two basic reasons: (1) The intellectual scope of ToNC should not be limited by networks that can be built or even envisioned in the current technological environment; technologically untethered but mathematically rigorous investigation of networked computation is also worthwhile. (2) Some of the most eminent and productive members of the ToC community have traditionally worked by themselves or in collaboration with other theorists, and they have established broad and deep research track records in the process. Their potential contribution to the ToNC agenda is immense and should not be conditioned on participation in multi-PI, substantially experimental projects. NSF's Computing and Communication Foundations Division (CCF) should support small, purely theoretical ToNC-research projects, but, ideally, CCF would not be the only source of such support.

Next Steps

For the ToNC-research agenda to be as broad and deep as it promises to be, support will have to be obtained from diverse sources. In particular, funding will have to come from all three divisions in the CISE Directorate at NSF, from Federal agencies other than NSF, and from forward-looking IT companies. This is one of the major challenges ahead for ToNC-community leaders. Vigorous advocacy and outreach will be important in meeting this challenge.

Finally, the ToNC community should coordinate and collaborate with the broader networking community, in both advocacy and in research. ToNC researchers can participate in the Global Environment for Network Innovations [GENI] by formulating testable hypotheses about the inherent power and limitations of networks. The architecture-research community is currently wrestling with fundamental questions about the value, costs, and tradeoffs of various networking primitives and abstractions. Very similar questions must be answered in the pursuit of a rigorous Theory of Networked Computation, and GENI will present a unique opportunity to experiment with new networks that have both innovative functionality and rigorous foundations.

References

[Byer] J. Byers, "ToNC: Summary of Algorithmic Foundations Working Group," http://www.cs.yale.edu/homes/jf/ToNC-Berkeley/Byers-Group4.ppt

[Feig] J. Feigenbaum, "Progress Report on the ToNC Initiative," http://www.cs.yale.edu/homes/jf/ToNC-Berkeley/Feigenbaum-Progress.ppt

[GENI] Global Environment for Network Innovations, http://www.geni.net

[Goel] A. Goel, "Next-Generation Architecure," http://www.cs.yale.edu/homes/jf/ToNC-Princeton/Goel-Group2.ppt

[Impa] R. Impagliazzo, "Complexity-Theoretic Foundations of Networked Computing," http://www.cs.yale.edu/homes/jf/ToNC-Berkeley/Impagliazzo-Group1.ppt

[Karp] R. M. Karp, "Report of the SIGACT Committee on Funding for Theoretical Computer Science" (presented at the FOCS Business Meeting on October 24, 2005), *SIGACT News* **36:4** (December 2005), pp. 99-102.

[Klei] R. Kleinberg, "Analytical Foundations of Networked Computing," http://www.cs.yale.edu/homes/jf/ToNC-Berkeley/Kleinberg-Group5.ppt

[Patt] D. Patterson, "20^{th} century vs. 21^{st} century C&C: the SPUR manifesto," *Communications of the ACM* **48:3** (March 2005), pp. 15-16.

[ToNC] Towards a Theory of Networked Computation, http://www.cs.yale.edu/homes/jf/ToNC.html