DOI:10.1145/3122847

Moshe Y. Vardi

Divination by Program Committee

DIVINATION IS THE practice of an occultic ritual as an aid in decision making. It has old historical roots. According to the biblical book of Samuel I, in the 11th century BCE, the Hebrew King Saul sought wisdom from the Witch of Endor, who summoned the dead prophet Samuel, before his impending battle with the Philistines. Alexander the Great, after conquering Egypt in 332 BCE, visited the Oracle of Amun at the Siwa Oasis to learn about his future prospects. Divination can be practiced in many ways, including sortilege (casting of lots), reading tea leaves or animal entrails, random querying of texts, and more. Divination has been dismissed as superstition since antiquity; the Greek scholar Lucian derided divination already in the 2nd century CE. Yet the practice persists.

Developments in mathematics and in computer science in the 20th century shed new light on the power of divination. Unless we believe that divination truly allows us to consult the divine, we can view it simply as a form of randomization, which is recognized as a powerful construct in game theory and algorithm design. The classical game-theoretic example is the game of Rock-Scissors-Paper in which there is no Nash equilibrium of pure strategies, but there is a Nash equilibrium in which both players choose their actions uniformly at random. The classical Dining Philosophers Problem has no symmetric distributed deterministic solution, but, as shown by Michael Rabin, has such a solution if we allow randomization. The essential insight is that randomization is a powerful way to deal with incomplete information. Thus, as realized by the anthropologist Michael Dove in the 1970s, when the Kantu

people of Borneo use birdwatching to decide which sites to farm and which sites to leave fallow, they are simply randomizing in the face of uncertainty about rain, pests, and more, but this randomization comes with a belief in the divine source of the decision. (See essay by Michael Sulson at <https://goo.gl/Ryb264>.)

But what does this have to do with program committees? In 2014, the Neural Information Processing Systems Foundation (NIPS) Conference split the program committee into two independent committees, and then subjected 10% of the submissions—166 papers—to decision making by both committees. The two committees disagreed on 43 papers. Given the NIPS paper acceptance rate of 25%, this means that close to 60% of the papers accepted by the first committee were rejected by the second one and vice versa. (See analysis by Eric Price at <https://goo.gl/fy5jLR>.) This high level of randomness came as a surprise to many people, but I have found it quite expected. My own experience is that in a typical program-committee meeting there is broad agreement for acceptance about the top 10% of the papers, as well as broad agreement rejections about the bottom 25% of the papers. For the other 65% of the submissions, there is no agreement and the final accept/reject decision is fairly random. This is particularly true when the accept/reject decision pivots on issues such as significance and interestingness, which can be quite subjective. Yet, we seem to pretend that this random decision reflects the deep wisdom of the program committee.

I believe the NIPS experiment should not only teach us some humility, but should also suggest that we may want to reconsider the basic

modus operandi of program committees. The standard approach in such committees can be viewed as “guilty until proven innocent.” We expect only 25%–35% of the papers to be accepted, so the default decision is to reject unless there is strong agreement to accept. But the reality is that a different committee may have reached a different decision on the majority of accepted papers. Is it wise to reject papers based essentially on the whim of the program committee? If we switch mode to “innocent until proven guilty,” we would reject only papers on which there is strong agreement to reject, and accept all other papers.

Beyond the increased fairness of “innocent until proven guilty,” this approach would also increase the efficiency of the conference-publication system. A high rejection rate means that papers are submitted, resubmitted, and re-resubmitted, resulting in a very high reviewing burden on the community. It also results in the proliferation of conferences, which fragments research communities. As I argued in an earlier editorial (<https://goo.gl/dUMkwZ>), I believe the proper way to adapt to the growth of the computing research is to grow our conferences rather than proliferate conferences.

NIPS should be lauded for applying the “publication method” to scientific inquiry. It is up to the computing-research community to draw the conclusions and act accordingly!

Follow me on Facebook, Google+, and Twitter.

Moshe Y. Vardi (vardi@cs.rice.edu) is the Karen Ostrum George Distinguished Service Professor in Computational Engineering and Director of the Ken Kennedy Institute for Information Technology at Rice University, Houston, TX. He is the former Editor-in-Chief of *Communications*.

Copyright held by author.