AI magazine
Volume 39 Number 4 Winter 2018

AI Bookie

■ *The AI Bookie column documents highlights from AI Bets, an online forum for the creation of adjudicatable predictions and bets about the future of AI. While it is easy to make a prediction about the future, this forum was created to help researchers craft predictions whose accuracy can be clearly and unambiguously judged when they come due. The bets will be documented on line, and regularly in this publication in The AI Bookie. We encourage bets that are rigorously and scientifically argued. We discourage bets that are too general to be evaluated, or too specific to an institution or individual. The goal is not to continue to feed the media frenzy and pundit predictions about AI, but rather to curate and promote bets whose outcomes will provide useful feedback to the scientific community.*

Place Your Bets: Adversarial Collaboration for Scientific Advancement

Kurt Bollacker, Praveen Paritosh, Chris Welty

It is common these days to hear laments about the loss of rigor in AI (for example, see Lipton and Steinhardt 2018), for researchers to point to the dramatic overspecialization and the tendency of communities to endlessly pursue derivative results well past the point of no return. We like to think that today is more dire than yesterday, of course, when in truth today is only more current, more present. The cycle remains the same. This realization does not, however, avert the need to halt the cycle of derivative results, time and again, helping a field out of one rut and all too often into another.

We have considered the question of whether this cycle can be dampened, and if so, how, and alighted on foster-

Guidelines for Bets

- **The subject of the bet must be relevant to AI or a related field.**
- **We want to encourage bets that promote or advance the field, and discourage vague or nonscience bets such as "when will be the next AI winter" or "company X will be out of business by 2020"**
- **Bets are for reputation, not money.**
- **The period of the bet must be a stated length of time or until a particular event or situation occurs. In all cases, the conditions for time of adjudication must be clear at the onset.**
- **The bettors must provide an argument explaining why the subject of their prediction is important and why they think they will be proved right. All arguments will be in the public domain.**
- **Individual people or small groups make bets using their real names. Large groups, formal organizations, and pseudonyms are not allowed.**
- **Every bet must have an adjudicator. Adjudicators must be agreed upon by the bettors and can be assigned if needed. The adjudicator must approve the operationalization of the bet.**
- **The outcome of predictions is usually decided through mutual agreement by the bettors; the adjudicator serves the role of being a neutral party who can resolve disagreements and the bettors agree to live with that resolution. Thus, while the two bettors have the goal in specifying the bet to make their "side" as clear as possible, the adjudicator has the goal of ensuring the operalization is as clear as possible. There are no appeals.**
- **Bets and predictions are always ultimately win-lose or true-false. There are no partials. This is important. A bet must be formed such that it can be adjudicated in this way.**
- **The data used to adjudicate a bet must be publicly available. There are no secrets.**
- **Bets may not be revoked once final, unless through agreement by all parties.**
- **The bet and outcome will be made public.**

ing adversarial collaboration for scientific advancement. Our current system of reporting almost exclusively positive results is akin to the machine learning problem of training on positive data: we lack solid negatives. Adversarial frameworks, such as the legal system, are important in areas where truth is understood to be unattainable. Adversaries, when properly motivated, can be relied on to find the flaws in one another's arguments or evidence, and would not take a contrary position to vacuous, irrelevant, or merely incremental hypotheses. We believe that science, and especially AI, needs to adopt such an approach, and we propose to start simply — with bets.

Betting has a long history in human society, and scientific bets are an old tradition: legend has it that Newton's *Principia* was motivated by a bet. Hawking

Illustration courtesy James Gary.

"bet against himself" in 1974 to encourage others to find evidence of black holes, in a move known as the Thorne-Hawking bet. Some years earlier and long before such capabilities were a reality, John McCarthy bet that a machine would beat a human at chess. In contrast to debates, the prominent form of adversarial discourse in the early history of natural philosophy and spirituality, bets even the playing field. Winning or losing a bet is less prone to bias, rhetorical device, or the disposition of the audience than is winning or losing a debate. Bets employ discrete criteria to determine the winner, and a well-composed bet should be adjudicatable. You don't win a bet by telling a good joke.

In 1968, Donald Michie, founder of the Department of Machine Intelligence and Perception at the University of Edinburgh, invited David Levy, a master chess player and an expert on computer chess, to play a friendly game of chess against John McCarthy, which Levy won. McCarthy remarked that although Levy was able to beat him now, within 10 years the computer program would exist that could beat Levy. Levy responded by offering the famous bet: within that time frame, no chess program would beat him in a tournament match. The two made a £500 bet, which was later more than doubled when Donald Michie, Seymour Papert from MIT, and Ed Kozdrowicki from the University of California, joined in the wager. David Levy collected on the bet 10 years later, in 1978, after winning a match against Chess 4.7 in Toronto. He won a second five-year bet in 1984, versus Cray Blitz, and then offered a prize for the first computer chess team to beat him. He was finally defeated 4–0 by Deep Thought, the Carnegie Mellon University precursor to Deep Blue, in 1989.

The McCarthy-Levy bet was influential in spurring research in automated chess playing and related areas; however, because this bet wasn't carefully crafted, it resulted in different mutations with different scientific explanations. The initial focus of the entire computer chess community was a single person, David Levy, and the reasons for this limited focus were never articulated. According to Patrick Winston's account (personal communication, August 2017), McCarthy gave up the bet before the 10 years were up, because the only approach to computer chess at the time — which ultimately did prevail — was based on brute-force search. McCarthy had intended, when he made the bet, a computer playing as a human would play. This thinking suggests that the "how" of the intended bet was never articulated.

Other kinds of one-sided bets are common. Newspapers, for example, make informal predictions that are impossible to evaluate ("AI will take away all the jobs"); corporations and venture capitalists make opaque bets on areas of science ("stealth mode" investments in self-driving cars, conversational technologies, or clean energy); government funding agencies such as DARPA, NSF, and the EC make bets on topics and people; editorial boards, hiring committees, award committees (for example, Horvitz and Selman 2012; Marcus, Rossi, and Veloso 2016; and Müller and Bostrom 2016), and even individual researchers (for example, Winston 2012) are making real bets with the use of their limited resources. However, the lack of formality, the lack of adversarial feedback or adjudicatability, the lack of transparency, archivability, and citability of these informal and one-sided bets limits their usability in the scientific process. We would like to change that, and to put a halt to the endless and repeated cycling through of derivative results, by capturing and harnessing scientific disagreement with the use of carefully specified, scientifically rigorous bets.

Supporting Adversarial Collaboration

We address some of the problems we've identified with betting by drawing from the successes in developing the scientific method itself, suitably tuned for "adversarial" contexts, such as hypothesis formulation, experimental design, and analysis. In the traditional scientific method, a testable hypothesis is stated for which an experiment can be performed and the results analyzed. Good experimental design utilizes clear, detailed protocols and mechanisms for observations of phenomena. The analysis of those observations should be straightforward, with the desired outcome being evidence for or against the hypothesis. Science, and AI in particular, has become very biased toward positive evidence — evidence that supports our hypotheses. We've all been there: when we find that evidence, we stop looking.

In a bet, the bettors make mutually exclusive hypotheses. When evidence on one side of the bet is presented, the adversary naturally continues the search. Exactly one of the two hypotheses can be proven true by evidence, and that evidence will disprove the other hypothesis. A final part is played by an adjudicator, an objective party (or committee) who the bettors agree will settle any disputes in resolving the bet. There are no protocols in a bet, but the detail and specificity for terms of the bet ("observations of phenomena") must be high enough to ensure this mutual exclusivity, and thus one and only one of the bettors can win the bet. This process of separation of essential elements of a disagreement from the frequently accompanying "sexy" but tangential issues is the core contribution of a rigorous scientific bet. If created properly, this separation should result in a bet whose outcome can easily and unambiguously be verified by peer-reviewed adjudication.

As compared to traditional hypothesis generation and experimental design, scientific bets may require greater commitment and effort. Bettors must call out what they believe in more detail than in a simple hypothesis, because they have to distinguish it from what they don't believe, what they do not know, and of what they are uncertain. Normally, scientists are required to provide this introspection themselves, but as humans, it is difficult for us to separate the excitement of proving a hypothesis from the rigor of objective analysis. With bets, this introspection is required for practical interaction with the adversary and adjudicator. The parties must cooperate with each other, which in some cases may mean working with someone with whom they have an adversarial relationship.

The AI Bookie

There is no widely used public platform for making scientific bets in AI, so we have decided to create and operate just such a platform, with the AI Bookie. Rather than reinvent wheels, we are inspired by and borrow rules from Long Bets (for example Lowenstein 2016), an "arena for competitive, accountable predictions … [as] a way to foster better long-term thinking." While Long Bets has a focus on long-term bets, the AI Bookie welcomes bets of any duration on the future of AI and will exist to foster clearer thinking about the science around AI in general. The bets will be documented online (at sciencebets.org) and also regularly in this column.

We encourage you to find a partner with whom you have a scientific disagreement and make a bet. Contact us if you would like help finding an adversary with whom you can collaborate. Once you have a collaborator and an idea for a bet, we will help you through the process of crafting a bet to a sufficient level of precision and rigor.

Over time, the history of well-written bets may act as a new type of scientific record that has a distinct role in the scholarly publication process. To this end, we hope that members of the AI community and of the scientific community at large will create bets that have lasting value.

References

Horvitz, E., and Selman, B. 2012. Interim Report from the Panel Chairs: AAAI Presidential Panel on Long-Term AI Futures. In *Singularity Hypotheses: A Scientific and Philosophical Assessment,* edited by A. Eden; J. Moor; J. Søraker; and E. Steinhart, 301–308. Berlin: Springer.

Lipton, Z., and Steinhardt, J. 2018. Troubling Trends in Machine Learning Scholarship. arXiv preprint. arXiv:1807.03341v2 [stat.ML]. Ithaca, NY: Cornell University Library.

Lowenstein, R. 2016. Why Buffett's Million-Dollar Bet Against Hedge Funds Was a Slam Dunk. *Fortune Magazine,* May 10.

Marcus, G.; Rossi, F.; and Veloso, M. 2016. Beyond the Turing Test. *AI Magazine*, 37(1): 3–4. doi.org/10.1609/aimag.v37i1.2650.

Müller, V., and Bostrom, N. 2016. Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In *Fundamental Issues of Artificial Intelligence*, edited by V. Müller, 553–570. Berlin: Springer International Publishing.

Winston, P. 2012. The Next 50 Years: A Personal View. *Biologically Inspired Cognitive Architectures* 1(July): 92–99. doi.org/10.1016/j.bica.2012.03.002.

Kurt Bollacker is the digital research director at The Long Now Foundation.

Praveen Paritosh is a senior research scientist at Google, Inc.

Chris Welty is a senior research scientist at Google, Inc.