Crowdsourced Knowledge: Peril and Promise for Complex Knowledge Systems

Mary Keeler
mkeeler@uw.edu
Josh Johnson
lain0523u.washington.edu
Arun Majumdar
arun@vivomind.com

In recent efforts to create natural-language, question-answering systems for the World Wide Web that can exploit the vast availability of information in its resources, researchers explicitly assume that what is most often repeated in that knowledge base is the truth. This paper considers the implications of that "most often repeated" (MOR) assumption—both hazardous and hopeful—and suggests Peirce's "economy of research" (EOR) in his evolutionary view of logic as a promising alternative to MOR for truth-finding in the increasing complexity of crowdsourced knowledge.

1 Introduction

"Crowdsourcing" on the World Wide Web (WWW) is rapidly becoming a powerful methodology for everything from finding new stars in the galaxy [1] to revolutionizing governments [2]. Employers can even crowdsource a labor force for "human intelligence tasks" [3]. M. Ojala's article, "Everyone's an Expert: The Crowdsourcing of History" in *Data Conservation Laboratory News* explains: "Crowdsourcing is now changing the way we think about knowledge, … [e]xpertise is no longer the exclusive domain of experts" [4]. Social media now make it possible to correct an article's facts and post arguments in a blog. The Flickr Commons project [5] incorporates this power for the identification of items in the photo collections of 30 participating institutions (including the Library of Congress, the Getty Research Institute, and the British Museum) to crowdsource the expertise of ordinary people. These institutions have learned that validating answers to questions on social media as a public process is time consuming, but they conclude that the "greater understanding of our shared past" makes the endeavor worthwhile.

Meanwhile, the vast availability of information on the WWW has inspired the goal of "open domain question answering systems" (QA) that can respond to a natural language question with a short natural language response. The QA technology race to develop more efficient use of "crowdsourced data" challenged researchers to develop standardized evaluation metrics, with the goal "to foster research on systems that retrieve answers rather than documents in response to a question." The chief architect of the most famous QA system, IBM's Watson, explains that the real challenge is to engage humans effectively in the process [6]. Even less ambitious QA system developers explicitly make use of human participation, usually in "voting" processes. Examination of these processes reveals a variety of methods that all rely on a "most often repeated" (MOR) assumption of what is the truth, and developers contend they have no alternative but to rely on this logical fallacy. We argue that Peirce's pragmatic methodology for the "economy of research" (EOR) is an effective alternative to MOR, which could be developed as a "complex adaptive reasoning" game to engage humans.

2 The Evolution of QA Systems

As QA systems have evolved from specialized information retrieval systems with integrated natural language processing into systems that answer questions using the unstructured data of the WWW [6][7][8], their developers have consolidated efforts into work on a few promising systems [9][10][11]. The Fall 2010 issue of *AI Magazine* covered six of the "most interesting" QA systems still under development: Project Halo's Digital Aristotle, Cyc's Semantic Research Assistant, IBM's Watson, Cambridge's True Knowledge, and the University of Washington's TextRunner. The article explains why QA continues to be pursued as a fundamental capability.

> The QA problem extends beyond AI systems to many analytical tasks that involve gathering, correlating, and analyzing information in ways that can naturally be formulated as questions. Ultimately, questions are an interface to systems that provide such analytic capabilities, and the need to provide this interface has increased dramatically over the past decade with the explosion of information available in digital form. … [Users expect] specific answers to specific questions by understanding, synthesizing, and reasoning about the underlying data, knowledge, and text documents. [They] also expect it to be able to provide user-specific explanations or justifications relating the context of the question to what the user currently knows, as well as informing the user of its confidence, especially in cases where the confidence in an answer is low. … recently, with advances in knowledge-based systems, natural language understanding, machine learning, and text understanding across the web, we may be on the threshold of finding combinations of these techniques that achieve a convincing and useful question-answering capability that is radically different from what is available in the market today. [12]

The most ambitious QA system, IBM's Watson, debuted as a contestant on the U.S. television quiz show *Jeopardy*, winning the game against two previously top-scoring humans. Chief Architect D. Ferrucci explains that the real challenge in finding correct answers is to improve the efficiency—not of finding documents, but of gathering knowledge from them, a process in which he says humans must be engaged.

> The human is doing the deep reasoning, but without the technology can't get the information to reason over. Formal reasoning systems can help, but we need to be careful about what role humans really need to and want to play. We need to advance the paradigm from known-item keyword research, because it's a pipe dream that authors are going to tag up the stuff; also, because the users' view is different from authors' (in vocabulary, language, etc.), the annotation problem becomes much worse. The hypothesis here is that open semantics must emerge. The right analysis for the job will likely be a best-of-breed combination integrating across many dimensions. [13]

Ferrucci here refers to the Emerging Component Community at CMU [14]. Watson depends on a system of interoperable components for very large-scale assigning of semantics that weighs evidence and determines the probability that an answer is right. Watson doesn't have the deep understanding or common sense that humans do; it "sees" only words and their relations.

Even less ambitious QA system developers still make use of human participation, usually programmers in a "weighted voting process" [Note 1]. Examination of these processes reveals a variety of methods that cautiously rely on the MOR assumption, which any voting involves as ballot counting. Developers explicitly make this truth assumption, resigning themselves to the position: "What else can we do?" First of all, we can consider the perils of MOR, a "confidence game" which has been identified at least since Aristotle.

3 The Most Often Repeated (MOR) Problem

A query to *Answers.com* finds that the *Oxford Dictionary of Proverbs* attributes this quotation to Charles Haddon Spurgeon: "It is well said in the old proverb, 'a lie will go round the world while truth is pulling its boots on'" [15]. *IWise* ("Wisdom on Demand") quotes Winston Churchill: "A lie gets halfway around the world before the truth has a chance to get its pants on" [16]. Other Web sources mention Shakespeare as the possible source of a similar quotation. Claiming truth by simply repeating an assertion (identified by Aristotle as babbling) is a fallacy in classical logic theory, which has become a master rhetorical strategy in modern politics and advertising. A quick Web search suggests that "A lie, repeated often enough, will end up as truth" is most frequently attributed to Lenin [17].

The WWW is now the medium of choice to propagate the MOR strategy, and with the explosion of "sources for knowledge" on the Web, new sites have emerged that claim to find the truth—or the lies. Truth-finding can be especially complex in political issues. For example, results of a *Fox News* survey, published 19 August 2010, indicate that nearly one in five Americans thinks Obama is Muslim [18]. Searching for "Obama is Muslim" on Google retrieves "About 44,400,000 results (in 0.13 seconds)." And *Politifact.com*'s "Truth-o-Meter" selects a "Lie of the Year," in the U.S., which for 2010 was "A government takeover of health care" [19]. Other sites include *FactCheck.org*, a project of the Annenberg Public Policy Center, and *Truth-out.org*, "progressive journalism and commentary on the web." Of course, these "liberal" sources have been matched by "conservative" truth-finders, such as

Conservapedia, “An encyclopedia with articles written from a conservative viewpoint,” and *ConservativeTruth.org*, “The Antidote to Liberal News Media.”

Journalists now study how political “memes” that are blatantly false get started and circulate on the Web, which entails following an increasing array of political blogs [20]. Reporters are especially aware of the MOR hazard as they negotiate the complexities of finding truth on the Web, and some are self-critical of their industry. A. Huffington says, “the WikiLeaks controversy has found a great deal of the media once again on the wrong side of the secrecy debate. As Harvard's John Perry Barlow tweeted: ‘We have reached a point in our history where lies are protected speech and the truth is criminal’”[21]. Of course, what those documents contained was too complex for most traditional media production methods, and it’s easier to repeat some concise conclusions [22]. As W. Lippmann warned in 1955, “When distant and unfamiliar and complex things are communicated to great masses of people, the truth suffers a considerable and often a radical distortion. The complex is made over into the simple, the hypothetical into the dogmatic, and the relative into an absolute” [23].

To help sort out the complexities there are sites like *RationalWiki.org*, which announces its purpose:

> Analyzing and refuting pseudoscience and the anti-science movement.
> Documenting the full range of crank ideas.
> Explorations of authoritarianism and fundamentalism.
> Analysis and criticism of how these subjects are handled in the media.

This Website features Carl Sagan’s essay “The Fine Art of Baloney Detection,” with a set of warning signs for common fallacies, which Sagan calls a ‘baloney detection kit” listing the types of fallacy, with definitions, and an example found on the current Internet [24]. (It does not specifically list the MOR fallacy, but “begging the question” is a generalization.)

Various social media have taken up the challenge of truth-seeking by crowdsourcing, beyond the “walls” of Wikis. *Quora* on Facebook calls itself, “A continually improving collection of questions and answers created, edited, and organized by everyone who uses it” [25]. And *TruthOrFiction* claims, “Be among the first to know about new eRumors, viruses, Internet hoaxes ... and more!” [26]. This site features “eRumors,” such as emails representing the “lies told by Barack Obama,” analyzes their content and then issues a “Fight the Smears” invitation for public response, through voting. We all know that many “universal truths” are no more than conjectures or simply rumors. “Most often repeated” is a perilous methodology in measuring truth, especially in the evolution of crowdsourced knowledge, which will rely on QA systems as they become common Web utilities.

4 Peirce’s Economy of Research (EOR)

Why we look to Charles Peirce (1839-1914) for help with the MOR problem is delineated by R. Burch in the online Stanford Encyclopedia of Philosophy (SEP) [27]. “Given his lifelong ideas and goals as a scientist-philosopher, he probably would have found the current practical importance of his ideas entirely to be expected.” For example: “He would not be in the least surprised to find that the topic of constructing ‘ontologies’ is in vogue among computer scientists…. He would not find in the least alien many contemporary analytic discussions of the notion of similarity; he would be right at home among them." Burch finishes his entry with a synoptic account of the many contemporary, practical and even crucial uses of Peirce's ideas (for industry, business, intelligence organizations, and the military) in the development of algorithms at the core of what is known as "Social Network Analysis."

Although Peirce scholars have neglected the topic (perhaps due to difficulty of access to his later writings [27] [28]), "the Economy of Research" threads through the evolution of Peirce's work on logic. EOR begins as a part of abduction, which Peirce closely identified with his pragmatism, saying that pragmatism "is nothing else than the question of the logic of abduction" [*CP* 5.196 (1903)]. Then, as logic takes on a more general significance in his development of sign theory, the role of economy extends through his systematic procedure for seeking the truth in science, integrating all stages of inference: abduction, deduction, and induction [*CP* 7.220]. M. Fisch, editor of an edition of Peirce's writings, explains that as a practising scientist Peirce "gradually gave up conceiving science as a mode of apprehension by a single knower, or as systematized knowledge, and came to conceive it as a mode of life common to any community of investigators, and to conceive a particular science as a social group pursuing the same or closely related inquiries" [29: xxv]. We know that he concluded: "the whole service of logic to science, whatever the nature of its services to individuals may be, is of the nature of an economy" [*CP* 7.220 fn (1901)].

Peirce's EOR considers "the relations between the utility and the cost of diminishing the probable error of our knowledge … how, with a given expenditure of money, time, and energy, to obtain the most valuable addition to our knowledge" [*CP* 7.140]. "In effect," says Burch, "the economics of research is a cost/benefit analysis in connection with states of knowledge. Although this idea has been insufficiently explored, Peirce himself regarded it as central to the scientific method and to the idea of rational behavior" [27]. The method begins in *abduction*, to form a hypothesis from conjectures about something uncertain that would, if true, reduce our uncertainty. In the next step of the method, *deduction*, we predict what would be the consequences from the provisionally adopted hypothesis, if it were true. In the third step of *induction*, we experiment to test for evidence of whether those predicted consequences actually occur. This method then enters either of two "feedback loops," as Burch explains. If we do find evidence for the deduced (predicted) consequences occurring, "then we loop back to the deduction stage," to predict further consequences of our hypothesis and experimentally test for them again. If the deduced consequences (our predictions) do not occur, "then we loop back to the abduction stage and come up with some new hypothesis" that explains both our original uncertainty and any new uncertainties uncovered in the course of testing the first, failed, hypothesis. "Then we pass on to the deduction stage, as before …" [27]. In [30-32], we discuss Peirce's theory in more detail [and see Note 2].

Obviously, this "self-corrective" procedure could be made more efficient at each stage of investigation—which is just what technology has done for scientific research. MOR misrepresents the problem of finding truth as naïve induction: simply counting abductions or conjectures, as though these propositions were empirical data. Peirce concludes, "But in induction a habit of probity is needed for success: a trickster is sure to play the confidence game upon himself" [*CP* 1.576 (1902)].

5 EOR as a Crowdsourcing Knowledge Game?

J. McGonigal argues, in her book *Reality is Broken: Why Games Make Us Better and How They Can Change the World*, that we can use games to build stronger communities and to collaborate at vast scales. In a chapter entitled "The Engagement Economy," she reports that it has already been done.

> On June 24, 2009, more than twenty thousand Britons joined forces online to investigate one of the biggest scandals in British history—investigations that led to the resignation of dozens of parliament members and ultimately inspired sweeping political reform. How did these ordinary citizens make such a big difference? They did it by playing a game. [33: 219].

When leaked documents revealed that dozens of members of parliament (MPs) had been filing illegal expense claims, adding up to tens of thousands of pounds, the public was outraged and demanded a full account. The government then "dumped the data" of more than a million expense forms and receipts in the unuseful format of scanned image files. Editors of the Guardian, who knew their reporters could never sort through the data, decided to develop a game to crowdsource the public's help. "Investigate Your MP's Expenses" became "the world's first massively multiplayer investigative journalism project," and a great success.

> Just three days into the game, it was clear that the crowdsourcing effort was an unprecedented success. More than 20,000 players had already analyzed more than 170,000 electronic documents. Michael Andersen, a member of the Nieman Journalism Lab at Harvard University and an expert on Internet journalism, reported at the time: "Journalism has seen crowdsourcing before, but it's the scale of the Guardian's project—170,000 documents reviewed in the first 80 hours, thanks to a visitor participation rate of 56 percent—that's breathtaking. [33: 222]

For comparison, McGonigal reports that "roughly 4.6 percent of visitors to Wikipedia make a contribution to the online encyclopedia," and she assesses the difference in motivation as the sense of moral "self-correctiveness" that a democratic system requires, but usually fails to facilitate. Meanwhile, the Obama administration has recently approached Microsoft to create a game to explain the difficult realities of spending cuts [34], another complex issue for the "Serious Games Initiative," where developers focus on "the growing application of videogames and videogame technologies for purposes outside of commercial entertainment," for use in "education, training, health and public policy" [35].

As McGonigal argues, games can engage collaborators for the purpose of improving a situation by constructing knowledge that would not be possible without their research—and the technology to make that research possible. An EOR game would augment not only researchers' access to evidence but also their constructive reasoning to interpret that evidence; see for example [31] [32]. Peirce's pragmatic theory of research defines "habits of thinking" as beliefs, and distinguishes reasoning from believing as the self-corrective thinking required in learning to improve habits of thought.

> When one's purpose lies in the line of novelty, invention, generalization, theory—in a word, improvement of the situation—instinct and the rule of thumb manifestly cease to be applicable. The best plan, then, on the whole, is to base our conduct as much as

> possible on Instinct, but when we do reason to reason with severely scientific logic. ... Where reasoning of any difficulty is to be done concerning positive facts, that is to say, not mere mathematical deduction, the aid that logic affords is most important. [CP 2 (1902)]

Peirce's pragmatic role for logic is to improve, or economize, reasoning in evolving effective habits of thought, or knowledge. Knowledge evolution conceived in the three stages of research (abduction, deduction, and induction) by no accident corresponds to natural evolutionary theory, in stages of diversity, selection, and adaptation. As Burch says, in Peirce's view, "The world, as it were, evolves by abducing, deducing, and inducing itself," and the evolution of knowledge is an extension of natural evolution [27]. Peirce revised his ideas many times and developed an intricate, mathematical form of EOR [e.g., CP 7.139-157], to maximize the reduction of indeterminacy in evolving knowledge by applying logic in research.

6 Promise and Perspectives

Peirce's motivation for his EOR was to improve the scientific method, which he came to regard as a method for improving reasoning in general. Our natural intuitions and practical reasoning have evolved for our survival in the natural environment; but we now live in an increasingly symbolic (or virtual) environment, as we describe in [36].

In successfully using the scientific method for reasoning, we learn "to play the game of explaining how the world is"—but we remember *this is a game* (of finding out which conjectured representations are probably true). Many of us need to be engaged, each contributing conjectures, to make the game work in improving knowledge. Like the QA system Watson, we can predict but are never sure of the truth, unlike systems that conveniently assume an ideal of truth defined as MOR. Predicting is always subject to risk, but even more so in our highly complex, symbolic world, where we can no longer depend on "common sense" and instincts—which are also vulnerable to other common fallacies that N. Taleb identifies:

> **Narrative Fallacy:** Our need to fit a story or pattern to a series of connected or disconnected facts. The statistical application of this is data mining.
> **The fallacy of silent evidence:** Looking at history, we do not see the full story, only the rosier parts of the process.
> **Confirmation error:** You look for instances that confirm your belief, your construction (or model)—and find them.
> **Round-up fallacy:** Confusing absence of evidence for evidence of absence.
> **Ludic Fallacy:** A manifestation of the Platonic fallacy in the study of uncertainty; basing studies of chance on the narrow world of games and dice … the bell curve (Gaussian) … is the application of the ludic fallacy to randomness.
> **Statistical Regress** (or circularity of statistics): We need the data to tell us what probability distribution to assume; but we need a probability distribution to tell us how much data we need. This causes a severe regress argument (which is somewhat shamelessly circumvented by resorting to the Gaussian and its kin). [37:302-304]

Taleb summarizes the weaknesses of instinctive reasoning:

> We tend to use different mental machinery—so called modules—in different situations: our brains lack a central all-purpose computer that starts with logical rules

> and applies them equally to all possible situations … By the mental mechanism I call naïve empiricism, we have a natural tendency to look for instances that confirm our story and our vision of the world—these instances are always easy to find. You take the past instances that corroborate your theories and you treat them as evidence … Even in testing a hypothesis, we tend to look for instances where the hypothesis proved true. Of course we can easily find confirmation; all we have to do is look, or have a researcher do it for us. I can find confirmation for just about anything, the way a skilled London cabbie can find traffic to increase the fare … Seeing white swans does not confirm the nonexistence of black swans... It is misleading to build a theory from observed facts … Contrary to conventional wisdom, our body of knowledge does not increase from a series of confirmatory observations [37:54-56].

Technology could give us access to what Taleb says "our brains lack," and what QA systems exploit to succeed where humans fail. But, as *Financial Times* reporter R. Waters cautions, in gaining that supplementary cognition, "we should not outsource the brain too much" [38]. Like the human brain, Watson weighs evidence and determines the probability that an answer is correct; but its evidence is entirely symbolic, so the "weighing" is of virtual evidence—symbolically represented evidence (it can't even know if it's playing the MOR "confidence game" on itself). The natural brain often fails us in our symbolic evolution; meanwhile, new evidence-weighing capabilities are emerging. For instance, *Patch* developers claim they "will create a premier global, national, local, and hyper-local content group … [to] mark a seminal moment in the evolution of digital journalism and online engagement" [39]. And *TED* curator C. Anderson claims his global online video lectures will power "crowd accelerated innovation" [40]. *TED* has already spawned *JoVE*, "Journal of Visualized Experiments—a peer reviewed, indexed journal devoted to the publication of biological research in a video format" [41].

While these developments are stunning, they can be misused, even unintentionally. "A fallacy is a pitfall of reasoning that exhibits a general and recurring tendency to deceive and to deceive successfully, to trick even the entirely serious and honest arguer" [42:6]. The MOR assumption is instituting a misleading rule for establishing truth in QA systems; but, as their developers say, what else can we do? Instead of the MOR confidence game, we need an EOR crowdsourcing game for investigating truth in the complex evolution of knowledge on the WWW.

7 Revelator: EOR Complex Adaptive Reasoning Game

A. Burks describes Peirce's pragmatism as a *logical theory of evolution*, in which "inquiry is an intellectual process of adapting to the environment" by learning and discovery [43: 501]. And C. Hausman explains that according to "Peirce's Evolutionary Philsophy," our purposes—such as finding truth and acquiring knowledge—are *emergent* ideals that unpredictably evolve along with our evolving environment. In this complex relationship, Peirce maintained, our *justifications* in any search for truth and knowledge cannot be arguments that depend on any single chain of reasoning, "which is no stronger than its weakest link, but rather a cable whose fibers may be ever so slender, providing they are sufficiently numerous and intimately connected" [44: 33; *CP* 5.265].

Motivated by Peirce's theory, we pursue the Revelator game of *complex adaptive reasoning* (*car)*, a self-correcting methodology for the evolution of truth and knowledge from conjectures. Revelator's *car*, inspired by J. Holland's *complex adaptive systems* (*cas*) [45], is designed to model complex logical relations among conjectures. While "normal form" games have preprogrammed possible strategies and effects, Revelator must be an "extensive form" game to give players the dynamic power to create their own "game world" of strategies: a simulation context for building narrative arguments that may evolve into robust hypotheses worth testing.

In conducting investigation or research, we formulate hypotheses from conjectures about what we anticipate might solve some challenging question. Hypotheses become more robust as they improve our anticipation: when consequences we expect appear to follow from certain conditions, as we guessed they would. In the game of Revelator, players make their conjectures explicit by formulating them in conditional propositions whose antecedents specify a course of action to be performed in obtaining *evidence* and whose consequents describe certain consequences to be expected, as an *inference* from the *evidence*. For example: "If records show that Obama grew up in Indonesia, then he is a Muslim" (even invalid or unwarranted inferences can be legal plays). We refer to this formulation of conjectures as *EII*: "Evidence Implies Inference," which is required for legal plays in the game. These rule-form plays operate as strategies for building robust hypotheses. The imposed linguistic constraints on the form of plays function somewhat like bidding in the game of bridge, to resemble a laboratory experiment in which experts carry out a dialogic, goal-directed, and limited but cognitively complex activity [32].

Revelator's gameplay is designed to engage players in formalizing the strategic process of reasoning, by enforcing the discipline of responsibly formulating conjectures that are logically relatable and explicitly refer to evidence for justification. Rule-form plays explicitly represent the conditional nature of any conjectured fact, by stating what evidence should lead anyone to reach the same conclusion as conjectured; and as rules, they are "strategic contenders" that define argument evolution. Revelator's natural language processing dialogue system tutors players in learning these conventions of play, while also inducing them to clarify ambiguities in natural language expressions that conceal implicit intentions of inference and accuracy of reference to evidence. With explicit conjectures in controlled natural language, players can essentially *program* the deductive process of reducing the variety of conjectures played, by logical selection and inheritance, to construct "surviving hypotheses" for inductive testing. The game encourages egotistically forceful conjecturing and even wild (though well-supported) guessing in a framework where they must contend with all comers, revealing their inevitable flaws along with their fitness in forming robust hypotheses.

Revelator play would improve weaknesses in everyday reasoning [47], especially in abductive and deductive stages, by extending players' sensory and cognitive capabilities and their normally limited commitment to thorough investigation [46]. Players can explore future possibilities and continually bring the state of their evolving conceptual model up to date as new claims are contributed, to improve the faithfulness of the conceptual models they construct. Since Revelator is explicitly a game, players remain aware that "uncertainty lies in the model's interpretation, the mapping between the model and the world" [45: 44-48]. Networkbased Revelator gameplay will increase the inductive efficiency in finding, representing, and checking evidence on the Web. Reasoners can evaluate each play for: what it implies, what other conjectures are consistent with it, what others are inconsistent, and how it stands up to the evidence (that is, what consequences should follow from its truth, to what degree it is confirmed by any consequences that do follow, how it is false if the consequences do not follow). Determining progress in the nonlinear, interconnected character of complex reasoning is like determining the reasonableness of pervasive mutual support among entries in an expanding crossword puzzle. Revelator distinguishes error-related from ignorance-related aspects of fallibility, revealing their pervasive interdependence [46, 32].

8 Knowledge Emergence in *CAR*

Revelator's *car* dynamically models Peirce's three stages in the evolution of reasoning (outlined in section 4): abduction or conjectural reasoning, deduction or necessary reasoning, and induction or experimental reasoning. Abduction suggests an explanatory representation for a mass of facts; deduction proves what must be true as a prediction, if that representation is true; induction indicates what is actually operative, by testing the represented prediction. Deduction draws virtual predictions: experiential consequences deduced from the conjectures and selected from among possible consequences, independently of whether or not they are known or believed to be true. For example: "Indonesia is the most populous Muslim country in the world; records show that Obama grew up in Indonesia; therefore, Obama is a Muslim." Skilled reasoners eliminate false conjectures to improve hypotheses, by virtually predicting the results of possible experiments. These first two stages of reasoning (conjectural and necessary) correlate with Holland's "two major steps" in modeling *cas*: discovery and construction. Just as in *cas,* in Revelator's *car* these stages operate recursively, to reveal emergent hypotheses for inductive testing.

Quite like physical building blocks in the children's game, *car* uses "propositional building blocks." These conditional-propositions establish virtual dimensions, instead of physical dimensions. Geometrical and gravitational (forceful) constraints are replaced by inferential and evidential (factual) constraints. As conditionally-related building blocks, they must adapt to a conceptual environment, in which fallibility serves as gravity does in physical systems. As do regular players of any intellectual game, Revelator players can begin to recognize certain kinds of conceptual patterns that become "building blocks" for longer-term strategies (like "forks," "pins," and "discovered attacks" in chess). Winning involves strategically selecting and combining conjectures to formulate more general propositions to reveal adaptive, higher-order arguments hidden in the complexity of their conceptual environment.

Holland's technique for resolving competition among rules is experience-based: a rule's winning ability depends on its usefulness in the past. Each rule is assigned credit strength that over time comes to reflect the rule's usefulness to the system, which changes the system's performance as it gains experience (for adaptation, by credit assignment). In Revelator, credit strength is represented in player scoring, which measures the survival value of conjectures — not by simply summing each player's plays but progressively, by counting only those conjectures that survive breakthroughs, breakdowns, and wipe-outs [31].

Strategic plays in Revelator are like the "levers" in *cas* that make potential knowledge emergence possible; in Revelator these are robust hypotheses. According to one implementation scenario, these dynamic and evolving arguments could be used to evaluate evidential measures, according to baseline metrics using semantic distances as knowledge capture constraints in [32]. Contradictions and inconsistencies are ultimately resolved by assigning related conjectures to networks that form stable strategies within which the concepts and arguments are consistent and prescribe the evidential tests that support their reliability.

9 Conclusions: No MOR Confidence Games

Game theory tells us that cooperation depends on the stability provided by progressive coordination of outcomes recognized as improvement, and the most significant condition in building stable patterns of cooperation is "hindsight" among participants, who "must be able to observe and respond to each other's prior choices." Although foresight is not necessary, "without foresight, the evolutionary process can take a very long time" [48: 182, 129]. Game-playing has these two key requisites for thriving cooperation, reciprocity and anticipation, which promote the stable evolution of cooperation by making interactions more durable and more frequent [48: 173, 129]. Even without trust, players can build a stable pattern of cooperation if these conditions are fulfilled. But "no form of cooperation is stable when the future is not important relative to the past": time perspective is critical [48: 129]. The emergence, growth, and maintenance of cooperation also rely on some assumptions about recognition and recall, so that players can respond to each other's choices [48: 174]. Any player must recognize other players in the game, from one play to the next, and each player must remember interactions with this player. In addition, "the evolution of any venture requires variety, or new things being tried" [48: 170]. Game theory tournaments capture the variety of possible game strategies, with several rounds to carry out refinements of strategies and to introduce new ideas, and then determine which kind of strategy can flourish [48: 171].

Revelator is a game in which players compete and cooperate in evolving the best strategies—valid conjecture combinations that predict outcomes of testing—toward the goal of justified candidate arguments as testable hypotheses. These strategies *might* emerge from the adapting conjectures, which assert possibilities and describe evidence to be checked that *might* justify these inferences. Overall, the game must help reasoners routinely self-correct, or form habits that minimize error. Revealing and reducing errors is crucial, both in interpreting evidence and in constructing inferences, for evolving productive arguments. Expected benefits (see list, below) of Revelator gameplay encourage us to pursue game building in the context of editorial analysis for blogs or wikis, to clarify and logically relate conjectures identified in free-form conversations, for promoting the emergence of truth and knowledge.

1. The game context encourages competition within a stable pattern of cooperation (as described by game theory), inducing responsible conduct among players as they engage in the conceptual discipline of formulating conjectures to improve collaboratively

developed hypotheses. The game formalizes the strategic process of reasoning, explicitly and engagingly.

2. The rule form of plays enforces the convention of clearly expressing ideas that can be inferentially related, and which explicitly refer to evidence for justification. Revelator plays maintain the distinction between inference and evidence, which is found to be a major difficulty in learning the skills of reasoning [46, 47].

3. The game's technology base supports a mode of interaction that can be augmented by ontological search to increase the efficiency of examining (finding, presenting, and checking) evidence on the WWW.

4. Automated conceptual processing would reduce the cognitive burden of inferring intricate logical relations, and create an automatic credit path (perhaps represented in an ontology) that promotes fair competition among inquiring players. Scores representing all contributions would constitute an objective evaluation of each player's contributions, to establish provenance in the collaborative construction process.

5. An extensive attribution system would make it convenient for players interested in similar aspects of target questions to find and keep track of one another's evolving ideas.

6. Playing the game would create automatic and complete documentation for analysis, implementation, and testing of technology augmentation needs for improvement of particular game contexts and content management. Technologies could even become competing strategies in a Revelator game, for the continuing improvement of game operation itself [8, 9].

Revelator anticipates the growing complexity of the collaborative reasoning, as libraries digitize and research is increasingly conducted by remotely located participants using social network facilities. “Wikipediholics” (people who are addicted to editing Wikipedia articles) explain that “Wikipedia is a massively multiplayer online role-playing game’” [33: 228-29]. Such “MMORPG players” need much better methodology than the “MOR confidence game” in the complex pursuit of truth in crowdsourced knowledge.

Notes

General Note: For all CP references, Collected Papers of Charles Sanders Peirce, 8 vols., edited by Arthur W. Burks, Charles Hartshorne, and Paul Weiss (Cambridge: Harvard University Press, 1931-58). MS references are to Peirce’s manuscripts in the Houghton Library, Harvard University.

[1] Peter Clark, now at Vulcan’s Project Halo, explains: [T]here are a whole variety of methods for weighing up evidence, many of which essentially come down to weighted voting. Basically the evidence is a set of features about the hypothesis answer – the programmer has to decide what those features are, and the typical approach is to throw in as many different features as possible. Then the challenge is to determine the right weights on those features so you give more weight to more informative features. If the programmer included an irrelevant feature, it will end up with zero weight so there’s no penalty for that. To find the weights, people use machine learning: there are algorithmic solutions for finding the “optimal” weights, i.e., the weight values that produce the highest score on the annotated training data. Then you’re done! [personal email 2/10/11]

[2] Burch explains: “Peirce came ever more clearly to see that there are three distinct and mutually incommensurable measures of imperfection of certitude. Only one was probability. The other two he called ‘verisimilitude’ (or ‘likelihood’) and ‘plausibility’. Each of the three

measures was associated with one of his types of argument. Probability he associated with deduction. Verisimilitude he associated with induction. And plausibility he associated with abduction" [27]

Bibliography

[1] *GalaxyZoo*. http://www.galaxyzoo.org/

[2] *WordPress*. "Crowdsourcing Revolution in Egypt." Available: http://siliconcowboy.wordpress.com/2011/02/02/crowdsourcing-revolution-in-egypt/

[3] Mechanical Turk. https://www.mturk.com/mturk/welcome

[4] Ojala, M. *Data Conservation Lab News* (2009). http://www.dclab.com/crowdsourcinghistory.asp

[5] Flickr Commons. Available: http://www.flickr.com/commons

[6] Voorhees, E. "Overview of the TREC 2001 Question Answering Track." National Institute of Standards and Technology, Online (2001). Available: http://www.ai.mit.edu/people/jimmylin/papers/Voorhees01.pdf

[7] Androutsopoulos, I., G. Ritchie, and P. Thanisch. "Natural language interfaces to databases—an introduction." *Natural Language Engineering*, 1(1): 29-81, 1995.

[8] Schank, R. and K. Colby, eds. *Computer Models of Thought and Language*. San Francisco: W.H. Freeman, 1973.

[9] TrueKnowledge. "The Internet Answer Engine." www.trueknowledge.com

[10] Kwok, C., O. Etzioni, and D. Weld. 2001. "Scaling question answering to the Web" *In Proceedings of the Tenth International World Wide Web Conference (WWW10)*. Available: http://www.cs.washington.edu/homes/weld/papers/mulder-www10.pdf

[11] Clark, P. and P. Harrison. "Large-Scale Extraction and Use of Knowledge from Text." *Proceedings of the Fifth International Conference on Knowledge Capture*, ACM: September 2009, pp. 153-160.

[12] Tunstall-Pedoe, W. "Knowledge: Open-Domain Question Answering Using Structured Knowledge and Inference," *AI Magazine*, Fall (2010). http://findarticles.com/p/articles/mi_7446/is_201010/ai_n56441011/

[13] Ferrucci, D. Lecture in Yorktown (recording). Available:

http://www-943.ibm.com/innovation/us/watson/watson-for-a-smarter-planet/building-a-jeopardy-champion/how-watson-works.html

[14] Common Component Metadata Analysis Engine. Available: http://uima.lti.cs.cmu.edu/index.html

[15] Answers.com. www.answers.com/

[16] *IWise*. "Wisdom on Demand." http://www.iwise.com/9lfZT

[17] Thinkexist.com. “Finding Quotations was never this Easy!” http://www.quotationspage.com/, http://www.brainyquote.com/

[18] Fox News Survey. Available: http://www.foxnews.com/politics/2010/08/19/nearly-americans-thinks-obama-muslim-survey-shows/

[19] *PolitiFact*. Available: http://politifact.com/truth-o-meter/article/2010/dec/16/lie-year-government-takeover-health-care/]

[20] Tomasky, M. *Guardian*. Available: http://www.guardian.co.uk/commentisfree/michaeltomasky/2011/feb/01/healthcare-vinson-decision

[21] Huffington, A. Report on "A Symposium on WikiLeaks and Internet Freedom," Personal Democracy Forum (12/11/10), Available: http://www.huffingtonpost.com/2010/12/03/wikileaks-online-presence_n_791699.html

[22] *UTVNews*. www.u.tv/News/

[23] Lippmann, W. Goodquotes.com. http://www.goodquotes.com/quote/walter-lippmann/when-distant-and-unfamiliar-and-complete

[24] Sagan, C. Available: http://www.positiveatheism.org/writ/saganbd.htm#BALONEY http://rationalwiki.org/wiki/The_Fine_Art_of_Baloney_Detection

[25] *Quora*. www.quora.com

[26] *TruthorFiction*. www.truthorfiction.com/ http://www.truthorfiction.com/rumors/o/obama-lies.htm

[27] Burch, R. In Zalta, E. (ed.). *Stanford Encyclopedia of Philosophy*. Available: http://plato.stanford.edu/

[28] Wible, J. *Science Bought and Sold: Essays in the Economics of Science*. University of Chicago Press (2002).

[29] Peirce, C. “Studies in Logic,” in A. Eschbach (ed.) *Foundations of Semiotics Series*, John Benjamins B.V. (1983), pp. 126-181. Originally published as "A Theory of Probable Inference," in *The Johns Hopkins Studies in Logic*, (ed.) Peirce, C. Little Brown and Co. (1883); intended as Essay XIV of the “Search for a Method” (1893).

[30] Keeler, M. and H. Pfeiffer. "Building a Pragmatic Methodology for KR Tool Research and Development." In H. Scharfe, P. Hitzler, P. Oehrstroem (Eds.): *Lecture Notes in Artificial Intelligence*, Vol. 4068. Springer, 314-30 (2006).

[31] Keeler, M. “Revelator Game of Inquiry: A Peircean Challenge for Conceptual Structures in Application and Evolution.” In U. Priss, S. Polovina, and R. Hill (Eds.): *Lecture Notes in Artificial Intelligence*, Vol. 4604. Springer, 443-459 (2007).

[32] Keeler, M. and A. Majumdar. "Revelator's Complex Adaptive Reasoning Methodology for Resource Infrastructure Evolution." In P. Eklund and O. Haemmerle (Eds.): *Lecture Notes in Artificial Intelligence*, Vol. 5113. Springer, 88-103 (2008).

[33] McGonigal, J. *Reality is Broken: Why Games Make us Better and How They Can Change the World*. Penguin Press (2011).

[34] Bailey, K. "Microsoft, Government Working Together on Deficit Reduction Game." Available: http://www.escapistmagazine.com/forums/read/7.188356-Microsoft-to-Make-Budget-Balancing-Game

[35] Serious Games Summit (eighth annual conference). http://www.gdconf.com/conference/sgs.html

[36] Keeler, M. "Learning to Map the Virtual Evolution of Knowledge." In M. Croitoru, S. Ferre, D. Lukose (eds.): *Lecture Notes in Artificial Intelligence*, Vol. 6208. Springer, 108-124 (2010).

[37] Taleb, N. *The Black Swan: the Impact of the Highly Improbable*. Random House (2010).

[38] Waters, R. "Information Technology: Mind Games," *Financial Times*, 16 February (2011).

[39] Bakalis, A. "AOL, Patch's Parent Company, Buys Huffington Post." Available: http://hollywood.patch.com/articles/aol-patchs-parent-company-buys-huffington-post

[40] Anderson, C. "How web video powers global innovation." Available: http://www.ted.com/talks/chris_anderson_how_web_video_powers_global_innovation.html

[41] *JoVE. Journal of Visualized Experiments*. http://www.jove.com/

[42] Woods, J. and D. Walton. *Argument: The Logic of Fallacies*. McGraw-Hill (1982).

[43] Burks, A. "Logic, Learning, and Creativity in Evolution." In N. Houser, D. Roberts, J. Van Evra (eds.), *Studies in the Logic of Charles Sanders Peirce*, Indiana U. Press, pp. 497-534 (1997).

[44] Hausman, C. *Charles S. Peirce's Evolutionary Philosophy.* Cambridge University Press (1993).

[45] Holland, J.H. *Emergence: from Chaos to Order*. Basic Books (1998).

[46] Haack, S. *Evidence and Inquiry*. Blackwell (1993).

[47] Kuhn, D. *Education for Thinking*. Harvard University Press (2005).

[48] Axelrod, R. *The Evolution of Cooperation*, Basic Books (1984).