Testing and Quantifying Collective Intelligence

BEN GREEN, Harvard University

1. INTRODUCTION

Among nature and society's most impressive feats is the ability of organisms and people to collectively accomplish complex goals that none could achieve individually. Ant colonies efficiently find near-optimal paths to food sources, while anonymous people collaboratively produce sites like Wikipedia. Because such systems are decentralized, they are more resilient and adaptive than groups that rely on centralized control to make decisions. Understanding the individual behaviors that generate these desired group outcomes will allow us to design human and engineered systems that utilize cooperative agents to achieve challenging results.

Yet despite countless examples of collective intelligence in nature and society [Edwards and Pratt 2009; Gordon 2014; Katz et al. 2011; Park 1921; Wheeler 1912; Woolley et al. 2010], we have very little formal language or methods to analyze it. Instead, we often operate on a "know it when we see it" principle. This leads to vaguely-defined notions of collective intelligence and makes it difficult to compare instances of collective intelligence.

In this paper, I define tests to identify, measure, and compare collective intelligence. I first define collective intelligence using domain-independent notions that can be formally measured and quantified. The utility of these metrics is then demonstrated using simulated ant colonies.

1.1 Collective Intelligence versus the Wisdom of Crowds

In [2010], Pratt defines collective intelligence as "a group of agents that together act as a single cognitive unit." He describes three key components to collective intelligence:

- —Only local interactions between agents are possible
- —Behavior is decentralized, i.e. no central coordinator knows or controls any agent's actions
- —Positive feedback is crucial to the group's self-organization to achieve their goals

Collective intelligence is often referred to as the "wisdom of crowds" to emphasize the abilities of groups over any individual. But while the two terms are often used interchangeably and there exists overlap between the definitions for collective intelligence and wise crowds, we should take care not to conflate the two.

The most common examples of the wisdom of crowds effect are what James Surowiecki calls cognition problems, in which pooling many people's independent answers to a question yields a better solution than asking only a few people [2005]. Surowiecki's other two classes of problems similarly rely on independent agents to achieve better results than any individual could. So while the wisdom of crowds follows the second criterion above, it clearly defies the first and third points.

A key component to a suitable definition for collective intelligence is its ability to distinguish between these related but separate phenomena. For the sake of this paper I will draw the following distinction: wisdom of crowds occurs when agents solve a problem in isolation and their answers are then aggregated, whereas collective intelligence requires agents to interact with one other (and the environment) while working toward a solution. This distinction will typically be discussed as crowds versus swarms or groups.

2. MEASURING COLLECTIVE INTELLIGENCE

Scientists have long struggled to identify and measure intelligence, especially when studying non-human agents. In his seminal discussion of machine intelligence, Alan Turing sought to answer the question “Can machines think?” [1950]. Recognizing this as an intractable discussion, Turing quickly changes tack and designs an operational definition to identify whether machines can think. In the resulting Imitation Game Turing asks if a machine could, through written communication, convince a human that he or she is conversing with another human.

Defining intelligence becomes even more difficult when dealing with groups rather than individuals and with animals rather than people or human-designed machines. As in Turing’s case, this discussion is immediately thrust into the murky territory of defining thought and intelligence. One is therefore compelled here to make the same dialectical move as Turing, and to design a proxy test for group intelligence rather than attempt to evaluate it directly.

Because collective intelligence can occur among groups of many forms, it is important to define agents broadly. Russell and Norvig provide an appropriate general definition: “An agent is something that perceives and acts in an environment” [2010]. This definition covers what one typically considers as an agent, such as people and animals, as well as others not usually thought to be sentient, such as bacteria and (debatably) cellular automata.

Another challenge in measuring collective intelligence is rejecting pre-conceived notions for how intelligence arises. In the case of computers, for instance, Turing recognized that machines and humans have different strengths and considered the possibility that “machines carry out something which ought to be described as thinking but which is very different from what a man does” [1950]. Ned Block expands on this idea, describing the perils of “crude human chauvinism” when discussing intelligence in agents whose internal processing we scarcely understand [1981].

Humans typically conceive of behavior at the level of individuals rather than groups — insects rather than colonies. Yet William Lyman warns of this “Gestalt blindness:” focusing on individual rather than group behaviors, and thus losing sight of “the forest for the trees” [1981].

The following definition is therefore proposed to act as a Turing Test for Collective Intelligence:

Definition 2.1. Collective intelligence is the ability of a group of agents to improve its ability on a given task by sharing information and responding to cues in the environment while working.

For a group to achieve collective intelligence based on this definition, it must satisfy two criteria:

(1) The group outperforms individuals on the given task. That is, groups with multiple agents achieve better results than each could alone.
(2) The group is more than just a wise crowd. In other words, the group achieves better performance when agents coordinate their behavior than when each agent works alone in parallel.

These points are both quantifiable based on the performance of swarms and crowds. The performance of a swarm is defined as the quality of its solution relative to the optimal solution. This term, also known as the performance above optimal, is calculated as $PAO = P_s/opt$. Similarly, a swarm’s performance relative to a crowd is defined by its performance below a crowd, $PBC = P_s/P_c$.

The two criteria above lay the framework for a Collective Intelligence Index (CII) that could function like an IQ test to compare instances of collective intelligence:

$$CII = (PAO \cdot PBC)^{-1} = \frac{opt \cdot P_c}{P_s^2} \tag{1}$$

Larger values of CII indicate that that a swarm displays more collective intelligence. The most valuable insights from the CII in its current form are from relative rather than absolute scores.

3. EXPERIMENTAL RESULTS: BIOLOGICAL SIMULATION

To provide a case study into the definition presented here, I developed a simulation of ant colony optimization (ACO) to solve the traveling salesman problem (TSP), a well-studied NP-Complete problem that has applications in many areas.

The ACO algorithm was developed by Marco Dorigo, who was inspired by the foraging behavior of ant colonies [1996]. While an optimized version known as the Ant Colony System generally achieves better performance, the Ant System algorithm used here more closely captures how a real ant colony forages. A full specification of the Ant System and its biological inspiration can be found in [Bonabeau et al. 1999].

Simulated swarms were tested on three sample problems taken from the TSPLIB, which contains canonical examples for the TSP [Reinelt 1991; 2008]. These problems contain 42, 150, and 280 cities, and I will refer to them, respectively, as the easy, medium, and hard TSPs. Optimal solutions were also provided. Swarm sizes ranged from 1 to 25, and each swarm was tested 20 times on each problem. The performance of a swarm or crowd is the total length of the tour found.

The performance relative to optimal of swarms varied for each of the three problems. Swarms were most successful on the medium TSP: while individual agents generated solutions 1.33 times the optimal length, swarms of 25 found paths just 1.10 times above optimal. Swarms solving the hard TSP improved from $PAO = 1.41$ down to $PAO = 1.21$. The PAO on the easy TSP improved the least as group size increased, going from 1.30 to 1.22.

Similar results were found for the comparison of swarms and crowds. The PBC for the easy TSP stayed between 0.97 and 1.04, meaning that the swarms and crowds achieved similar performance. The best performance relative to crowds for groups with fewer than 10 agents is on the medium TSP, while groups with more than 10 agents most outperform crowds on the hard TSP.

Based on these tests the simulated ant colony displays collective intelligence on the medium and hard TSPs, but not on the easy TSP. On the medium and hard TSPs the swarm improves its performance with larger groups and achieves better solutions than a crowd of the same size.

Combining the previous two analyses yields a single metric for the collective intelligence of a swarm. This allows for better comparisons between instances of collective intelligence. As expected based on the previous two tests, the CII of swarms solving the easy TSP is low and does not change for larger groups. The highest CII is observed for swarms solving the medium-sized TSP. It increases quickly, indicating the benefits of collaboration, and then remains steady. The CII for the hardest TSP is between that of the easy and medium TSPs. It improves rapidly as swarms increase from 1 to 5 agents, and then steadily grows until it matches the CII of the medium TSP with swarms of 25.

My results indicate that the intelligence of groups depends on the difficulty of their task. The CII increases with group size more for the medium and hard tasks than on the easy task, likely because groups are better able to benefit from coordinating their behavior in more complex environments.

Another interesting result is that although large groups outperform small groups across all tasks, the rate of improvement decreases as group size increases. While groups of 10 significantly outperform an individual agent, groups of 20 perform only slightly better than the groups of 10. This suggests that there may exist a critical number of individuals for each task at which collective intelligence stabilizes.

Based on these results, one promising route for future research is to determine experimentally how collective intelligence changes for different types of groups. Indeed, a recent study suggests that the swarming behavior of insects stabilizes for groups of approximately ten or larger [Puckett and Ouellette 2014]. Analyzing the onset of collective intelligence using the tests and metrics described here may yield similar insights and allow us to further pinpoint specific traits that contribute to collective intelligence.

REFERENCES

Ned Block. 1981. Psychologism and behaviorism. *The Philosophical Review* (1981), 5–43.

Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. 1999. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press.

Marco Dorigo, Vittorio Maniezzo, and Alberto Colorni. 1996. Ant System: optimization by a colony of cooperating agents. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 26, 1 (1996), 29–41.

Susan C Edwards and Stephen C Pratt. 2009. Rationality in collective decision-making by ant colonies. *Proceedings of the Royal Society B: Biological Sciences* 276 (2009), 3655–3661.

Deborah M Gordon. 2014. The ecology of collective behavior. *PLoS Biology* 12, 3 (2014), e1001805.

Yael Katz, Kolbjørn Tunstrøm, Christos C Ioannou, Cristián Huepe, and Iain D Couzin. 2011. Inferring the structure and dynamics of interactions in schooling fish. *Proceedings of the National Academy of Sciences* 108, 46 (2011), 18720–18725.

William G Lycan. 1981. Form, function, and feel. *The Journal of Philosophy* (1981), 24–50.

Robert E Park. 1921. Sociology and the social sciences: The social organism and the collective mind. *Amer. J. Sociology* (1921), 1–21.

Stephen Pratt. 2010. Collective Intelligence. *Encyclopedia of Animal Behavior* 1 (2010), 303–309.

James G Puckett and Nicholas T Ouellette. 2014. Determining asymptotically large population sizes in insect swarms. *Journal of The Royal Society Interface* 11, 99 (2014), 20140710.

Gerhard Reinelt. 1991. TSPLIB - A Traveling Salesman Problem Library. *ORSA Journal on Computing* 3, 4 (1991), 376–384.

Gerhard Reinelt. 2008. TSPLIB. http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/. (2008). [Online; accessed December-2014].

Stuart Russell and Peter Norvig. 2010. *Artificial Intelligence: A Modern Approach*. Prentice Hall.

James Surowiecki. 2005. *The Wisdom of Crowds*. Random House LLC.

Alan M Turing. 1950. Computing machinery and intelligence. *Mind* (1950), 433–460.

William Morton Wheeler. 1912. The Ant-Colony as an Organism. *Journal of Morphology* 22 (1912), 307.

Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science* 330, 6004 (2010), 686–688.