Artificial Swarm Intelligence vs Human Experts

Louis Rosenberg
Unanimous A.I.
San Francisco, CA 94115
unanimous.ai

Abstract— **Artificial Swarm Intelligence (ASI) strives to facilitate the emergence of a super-human intellect by connecting groups of human users in closed-loop systems modeled after biological swarms. Prior studies have shown that "human swarms" can make more accurate predictions than traditional methods for tapping the wisdom of groups, such as votes and polls. To further test the predictive ability of swarms, 75 random sports fans were assembled in the UNU platform for human swarming and tasked with predicting College Bowl football games against the spread. Expert predictions from ESPN were compared. The results are as follows: (i) Individuals – when working alone, test subjects achieved on average, 5 correct predictions out of 10 games (50% accuracy); (ii) Group Poll – aggregating data across all 75 subjects, the group achieved 6 correct predictions out of 10 games (60% accuracy); (iii) Experts - as published by ESPN, the college football experts averaged 5 correct predictions out of 10 games (50% accuracy); and (iv) Swarm – when the 75 subjects worked together as a real-time swarm, they achieved 7 correct predictions out of 10 games (70% accuracy). Thus by forming a real-time swarm intelligence, the group of random sports fans boosted their collective performance and out-performed experts.**

Keywords— Swarm Intelligence, Artificial Intelligence, Human Swarming, Wisdom of Crowds, Collective Intelligence

I. INTRODUCTION

In the field of A.I. research, practitioners have regularly turned to Mother Nature for inspiration and guidance. Not surprisingly, the first path explored was the most familiar – our own brains. Beginning with the Perceptrons of the 1950's and continuing to this day, Neural Networks have emerged as the dominant biologically inspired model for A.I. research. Nature, however, rarely reveals only a single pathway. Billions of years of evolution have produced at least one alternate method for generating high-level intelligence from smaller building blocks and it's not neural – it's collective.

Referred to as Swarm Intelligence (SI), countless species are known to amplify a local group's intellectual ability by forming closed-loop systems among large numbers of independent organisms. These dynamic systems demonstrate that under the right conditions, a collective intelligence can emerge that exceeds the capacity of the individual members in the group. Artificial intelligence researchers have explored swarm-based models for use among groups of networked robots and simulated agents [1], but only recently has swarming been applied to human networks [2, 3, 4, 5].

Known as Artificial Swarm Intelligence (ASI), these computational methods enable human groups to work together in real-time by forming a unified dynamic system that can answer questions, make predictions, reach decisions, or take actions. As a unified system, human swarms collectively explore a decision-space and quickly converge upon preferred solutions. Prior studies have shown that by working in swarms, human groups can outperform their individual members as well as outperform groups taking traditional votes or polls.

In a prior study, a randomly selected human group was tasked with predicting the top awards of the 2015 Oscars, both by taking a poll and by forming a swarm [5]. Across 48 participants, the average poll result achieved 6 of 15 correct predictions (40% success). When taking most popular prediction in the poll, the group achieved 7 of 15 correct predictions (47% success). When working together as a real-time swarm, the group achieved 11 of 15 correct predictions (73% success). This suggests that ASI may be a superior method for tapping the wisdom of crowds than traditional votes, polls, and surveys. The present study aims to explore this further, fielding a larger swarm of users and tasking them with predicting 10 college football games against the spread. In addition, the current study aims to compare the performance of the human swarm, comprised of randomly selected novices, with the performance of individual subject-matter experts.

II. SWARMS AS INTELLIGENT SYSTEMS

Among A.I. researchers, the word "swarm" often refers to groups of robots or simulated agents governed by simple localized rules [1]. These systems are generally inspired by flocks of birds and schools of fish, which navigate complex environments using similar processes. While such systems have many applications, for example enabling robotic drones to navigate in unison, the human swarms discussed herein are modeled less after the motions of flocks and schools, and more after the decision-making processes used by honeybee swarms. This is because the decision-making abilities of honeybees provide a powerful natural proof of the potential for an emergent decentralized parallelized intelligence.

As studied by Seeley et al., the processes that govern decision-making in honeybee swarms and neurological brains are remarkably similar [6]-[9]. Both employ large populations of simple excitable units (i.e., bees and neurons) that work in parallel to integrate noisy evidence, weigh competing alternatives, and converge on decisions in synchrony. In both, decisions are arrived at through a real-time competition among sub-populations of excitable units, each sub-population vying for a different alternative solution. When one sub-population exceeds a threshold level of support, the corresponding alternative is chosen. The threshold in both brains and swarms is not the unanimous support, or even a simple majority, but a sufficient quorum of excitation. This helps to avoid deadlocks and leads swarms to optimal decisions [10].

For example, every spring honeybees face a life-or-death decision to select a new home location for the colony. From hollow trees to abandoned sheds, the colony considers dozens of candidate sites over a 30 square mile area, evaluating each with respect to dozens of competing criteria. Does it have sufficient ventilation? Is it safe from predators? Is it large enough to store honey for winter? It's a complex problem with many tradeoffs and a misstep can mean death to the colony. Using body vibrations known as "waggle dances", hundreds of bees express preferences for competing sites based on numerous quality factors. Through a real-time negotiation, a decision is reached when a sufficient quorum emerges.

Remarkably, the bees arrive at optimal decisions 80% of the time [11]. Thus, although individual bees lack the mental capacity to make a decision this complex and nuanced, when hundreds of scout bees pool their knowledge and experience, they evoke a Collective Intelligence that is not only able to reach a decision, it finds an optimal solution. Thus by working together as a unified dynamic system, the colony amplifies its intelligence beyond the capacity of individual members. It is this emergent amplification of intelligence that human swarming aims to enable among groups of networked people.

III. Enabling Human Swarms

Unlike many social species, human have not evolved the natural ability to form a Swarm Intelligence, for we lack the subtle connections that other organisms use to establish tight-knit feedback-loops among members. Schooling fish detect vibrations in the water around them. Flocking birds detect motions propagating through the group. Swarming bees use complex body vibrations. This suggests that to evoke a real-time Artificial Swarm Intelligence (ASI) among groups of networked humans, technology is required to close the loop among members. To address this need, an online platform called **UNU** was developed to allow distributed groups of users to login from anywhere around the world and participate in a closed loop swarming process.

Modeled after the decision-making of natural swarms, UNU allows groups of independent actors to work in parallel to (a) integrate noisy evidence, (b) weigh competing alternatives, and (c) converge on final decisions in synchrony. Because humans can't waggle dance like honeybees, a novel interface had to be developed to allow participants to convey their individual intent with respect to a set of alternatives. In addition, the interface had to be crafted to allow users to perceive and react to the changing system in real-time, thereby closing a feedback loop around the full population.

As shown in Figure 1, users of UNU answer questions by collectively moving a graphical puck to select among a set of alternatives. The puck is modeled as a physical system with a defined mass, damping and friction. Users provide input by manipulating a graphical magnet with a mouse or touchscreen. By positioning their magnet, users impart their personal intent as a force vector on the puck. The input from each user is not a discrete vote, but a stream of vectors that varies freely over time. Because the full set of users can adjust their intent at every time-step, the puck moves, not based on the input of any individual, but based on the dynamics of the full system. This results in a real-time physical negotiation among the members of the swarm, the group collectively exploring the decision-space and converging on the most agreeable answer.

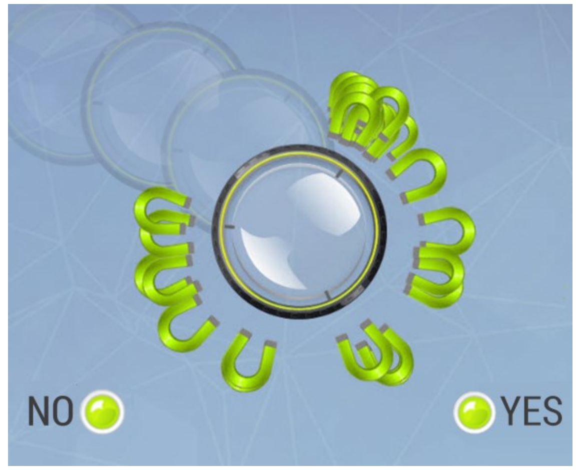

Fig 1. A human swarm comprised of user-controlled magnets.

We must note that users can only see their own magnet during the decision, not the magnets of others users. Thus, although they can view the puck's motion in real time, which represents the emerging will of the swarm, they are not influenced by the specific breakdown of support across the available options. This limits social biasing. For example, if the puck slows due to an emerging deadlock, the participants must evaluate their own willingness to shift support to alternate options without knowing the distribution of support that caused the deadlock. After each decision is over, users can view a replay of all the magnets, allowing them to reflect on how their contribution combined with others to produce the final answer.

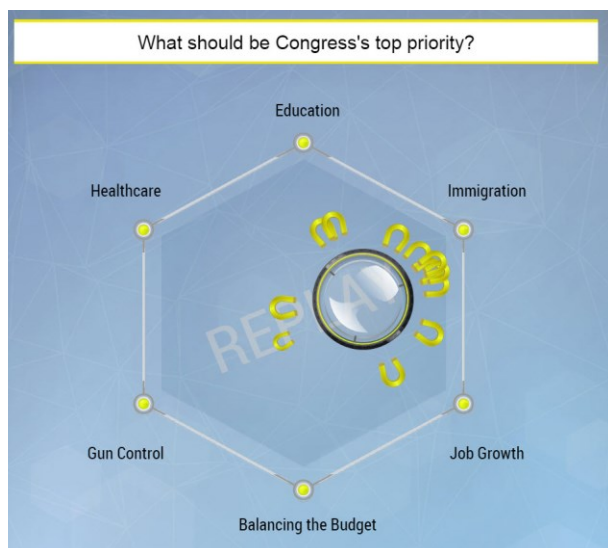

Fig 2. A snapshot of a swarm answering a question.

In Figure 2 above, an example question is shown as it appears simultaneously on the screens of all participants. In this trial, a swarm of 90 users was asked a politically charged question: “What should be Congress’s top priority?” Users are then given a 3,2,1 countdown to coordinate the start of the session. The swarm then springs into action, working in synchrony to guide the puck to a preferred answer.

The decision process is generally a complex negotiation, with individuals shifting their support numerous times to break deadlocks or defend against options they disfavor. When a user pulls towards one option in the answer set, a component of their force also acts to impede the motion of the puck towards competing options. In this way, users don’t only add support a preferred solution when pulling towards it, but also suppress solutions they don’t prefer. This enables the dual process seen in natural swarms and neurological brains wherein individual agents are enabled to both excite and inhibit [8], thereby reducing the chances of a deadlock.

If a group happens to be in substantial agreement at the start of the question, the puck moves smoothly to the preferred answer. But, if two or more competing options have significant support, the swarm negotiates as a unified system. Most users begin by pulling towards the option they prefer most, then shift to alternate choices if the puck starts moving towards an option they dislike. With all users making these changes in parallel, the swarm explores the decision space and converges on an answer that optimizes group satisfaction.

It’s important to note that users don’t just vary the direction of their input, but also the magnitude by adjusting the distance between the magnet and the puck. Because the puck is in motion, to apply full force users need to continually move their magnet so that it stays close to the puck’s rim. This is significant, for it requires all users to be engaged during the decision process. If they stop adjusting their magnet to the changing position of puck, the distance grows and their applied force wanes. Thus, like bees executing a waggle dance or neurons firing activation signals, the users in an artificial swarm must continuously express their changing preferences during the decision process or lose their influence over the outcome.

Post testing interviews with participants suggest that users with high levels of conviction in favor of a particular outcome are more vigilant in maintaining maximum force on the puck. Conversely, users who have lower conviction are less vigilant. In this way, the swarming interface allows the population to convey varying levels of conviction in real-time synchrony. We believe this helps the swarms converge on solutions that optimize the overall satisfaction of the group.

Observations and post-testing interviews also reveal that human swarming yields consistent outcomes across varying spatial placement of answer options. For example, if two highly favored options are placed on opposite sides of the puck’s starting position, the swarm will fall into an early deadlock as it grapples between them. Conversely, if the two highly favored options are placed on the same side of the puck’s starting position, the swarm will not fall into an early deadlock, but instead move the puck towards those two highly favored options. Still, a deadlock will emerge as the puck approaches midpoint between the two favored options. In this way, the decision space can have alternate layouts, but the swarm arrives at the same outcome. A similar robustness has been observed in honeybee swarms, which are known to decide upon optimal nesting locations regardless of the order in which sites are discovered and reported by scout bees [11].

Referring again to Figure 2 the default layout of answers is a set of six options in a hexagon pattern. The hexagonal configuration was chosen because according to social-science research, people are efficient decision-makers when presented with up to six options, but suffer from increasing “choice-overload” inefficiencies when confronted with larger sets [12]. To enable swarms to consider larger sets of answers, the system employs an iterative approach, presenting users with a series of six-option subsets of the full answer pool, then pitting the winner of each subset against each other. The system also allows swarms to select values on a continuous scale. This enables swarms to collectively decide upon quantities, prices, percentages, odds and other numerical values.

As shown in Figure 3 below, a swarm of users was asked to decide upon the fair price of a movie ticket on a scale from $0 to $25. When using scale-based layout, the puck starts at the center of the range and can move smoothly in either direction. The swarm generally overshoots the final answer, then reverses direction, oscillating in narrower and narrower bands. An answer is chosen when the puck settles upon a value for more than a threshold amount of time (e.g., 3 seconds).

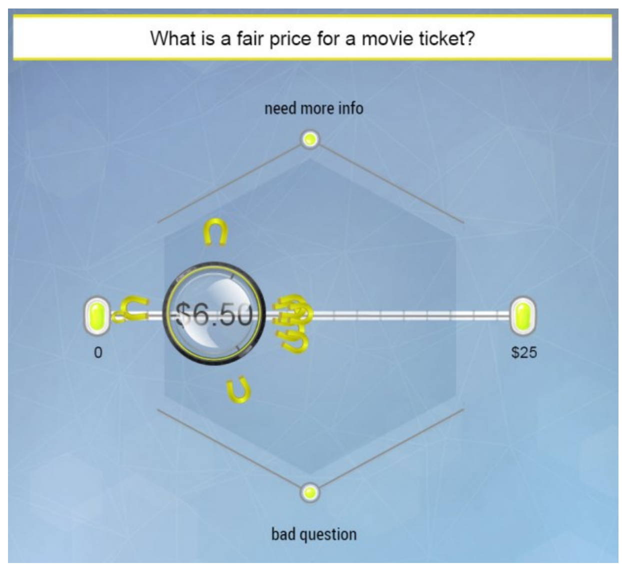

Fig 3. A sample scale-based layout for human swarming

For the current set of tests, range-style questions we asked, allowing users to predict both the winner and the point spread of each bowl game.

IV. Performance Testing

To assess the predictive ability of human swarms, a formal study was conducted with 75 randomly selected subjects. Each participated in the experiment via online access. The only requirement for participation was that each subject was self-identified as a college football fan. Each subject was paid $2.00 for their participation, which required them to make predictions for the outcome of 10 college bowl football games, first by on a blind poll using Survey Monkey, then as part of a real-time human swarm using the UNU platform. In addition, the researchers documented the predictions made by ESPN experts for the same games [13]. Finally, the researchers documented the Las Vegas point-spreads for each of the ten games, which are designed by bookmakers to make each prediction as close to a 50/50 proposition as possible.

When responding on the Survey Monkey poll, each individual gave their own prediction about (a) which team would win each of the 10 games, and (b) by how many points would they win the game (point spread). This allowed for predictions to be made against the spread, without explicitly informing the subjects with what the spread was.

When working as a swarm, the participants were instructed to move the graphical puck along a linear axis labeled with the names of each team competing in a game. As the puck moved along the axis, the closer to a particular team name, the higher the chosen point spread victory for this team. Figure 4 shows a screenshot for the Rose Bowl game, wherein the swarm predicted Stanford University would win by 8 points. It's important to note that although Figure 4 shows all magnets displayed, the subjects were only able to see their own magnet, and thus could not see the pull directions of others.

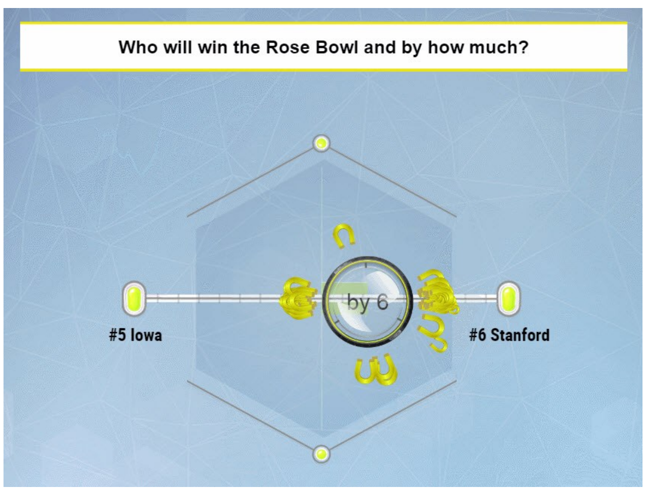

Fig 4. Screenshot of Swarm predicting Rose Bowl

It should be noted that when predicting the 10th game (both by poll and by swarm), a point spread was not used because the 10th game depended upon the outcome of prior games. There were 4 possible winners of the final College Playoff game, the subjects asked to predict which of the four teams would prevail. The ESPN experts did the same.

V. Results

Looking first at the poll results, we find that on average, across 75 participants, the individuals made 5 correct picks for the 10 games. This equates to 50% accuracy against the spread, which is not impressive but confirms that the Las Vegas odds-makers are skilled at what they do, picking spreads that for average individuals, made each bet an approximate toss-up among the competing teams.

Next we computed the most popular predictions in the survey across all 75 subjects. This can be viewed as a collective pick that taps the wisdom of crowds using traditional polling. The group, as a collective, got 6 correct picks out the 10 games in question, yielding 60% accuracy. This supports prior research into collective intelligence which suggest that groups amplify their intelligence when averaging predictions.

Next we compared the swarm results wherein the group of 75 participants worked together as a real-time unified system. The swarm produced 7 correct picks, yielding 70% accuracy. This supports prior studies that show human swarming to be a more effective means of tapping the collective intelligence of groups than votes, polls, and surveys.

Comparing the swarm's 7 correct predictions against the individual picks made by the 75 participants, it was found that the swarm outperformed 95% of participants. Thus, by working together as a real-time swarm, 95% of the subjects would have been better off going with the predictions made by the Artificial Swarm Intelligence than their own picks. This suggests that the ASI achieved a level of intelligence (with respect to this defined task) that was superior to the intelligence of the individual participants who comprised the swarm.

Finally we compared the swarm's predictions to those of the experts at ESPN. Based on publically published picks, ESPN experts made 5 correct predictions against the spread for the 10 games, yielding 50% accuracy. Thus, although their predictions were made with professional expertise, they were unable to beat the Las Vegas odds. The swarm, however, did beat the odds by a good margin. In this way, a human swarm of 75 sports fans, working as a unified system, produced more accurate results than the topic-specific experts at ESPN.

As a final comparison, we computed the payouts that would result if bets were by each of the parties. Because the swarm picked two longshots, and only lost games that were toss-ups (i.e. had nearly even odds), it did very well. The ESPN experts on the other hand, lost both longshot. Had the ESPN experts placed $10 on each of their picks, would have lost $24 of their $100 bet (-24% ROI). The swarm, on the other hand, would have won $34 across the ten games (+34% ROI). This further supports the possibility that swarming can amplify intelligence, allowing groups to behave as topic-specific experts.

VI. Discussion and Conclusions

Can swarms of average people rival the predictive abilities of topic-specific professionals? The results of this study, along with prior studies, suggest this might be the case. Furthermore, swarming appears to be a more effective method of tapping the wisdom of groups than traditional methods, like votes and polls. This may be because unlike polls, which collect data from individuals in isolation, swarms enable groups to negotiate in real-time synchrony, adjusting and adapting as decisions emerge before their eyes. The members of a swarm don't express static views, but continually assess and reassess their own convictions with respect to each of the possible outcomes, weighing their personal confidence and preferences. With all participants doing this in parallel, the swarm converges on solutions that reflect the collective will of the group, tuned by each individual's unique level of confidence.

Because of the potential of human swarming to enable groups to combine their knowledge and intuition in real-time, swarming likely offers the greatest benefit when groups make complex decisions on topics that can be assessed from many unique perspectives. This parallels the benefits of swarming among honeybees, where the decision to pick a new home-site must be evaluated across numerous competing factors. In fact, when honeybee swarms choose a new colony site, they consider dozen of locations, each evaluated with respect to at least six independent attributes. Despite the complexity of the decisions involved, honeybee swarms have been documented as making nearly optimal decisions most of the time [11].

Looking forward, this experiment supports the possibility that artificial swarms of networked humans have the potential to produce an emergent intelligence that exceeds the intellectual abilities of the individual participants for certain tasks. This could lead to the development of a networked super-intelligence that keeps humans in the loop. The fact that human participants are central to the emergent intelligence is promising, for it suggests that our human interests, values, and morals would be integrated into to the process, achieving a safer path to super-intelligence than a purely digital A.I. Further research is needed, exploring how increasing the size of swarms impacts the emergent intelligence produced.

ACKNOWLEDGMENT

Thanks to David Baltaxe and Joe Rosenbaum, both of Unanimous A.I. for their help in making this study possible. Also, this study was made possible by the use of UNU, an online software platform that enables real-time human swarms.

REFERENCES

[1] Beni, G., Wang, J. Swarm Intelligence in Cellular Robotic Systems, Proceed. NATO Advanced Workshop on Robots and Biological Systems, Tuscany, Italy (1989).

[2] Rosenberg, L.B, "Human Swarms, a real-time paradigm for collective intelligence." Collective Intelligence 2015, Santa Clara CA.

[3] Palmer, Daniel W., et al. "Emergent Diagnoses from a Collective of Radiologists: Algorithmic versus Social Consensus Strategies." Swarm Intelligence. Springer International Publishing, 2014. 222-229.

[4] Eberhart, Russell, Daniel Palmer, and Marc Kirschenbaum. "Beyond computational intelligence: blended intelligence." Swarm/Human Blended Intelligence Workshop (SHBI), 2015. IEEE, 2015.

[5] Rosenberg, L.B., "Human Swarms, a real-time method for collective intelligence." Proceedings of the European Conference on Artificial Life 2015, pp. 658-659

[6] Seeley, Thomas D., et al. "Stop signals provide cross inhibition in collective decision-making by honeybee swarms." Science 335.6064 (2012): 108-111.

[7] K.m. Passino, T.F. Seeley, P.K. Visscher, Swarm Cognition in honeybees, Behav. Ecol. Sociobiol. 62, 401 (2008).

[8] J.A.R. Marchall, R. Bogacz, A. Dornhaus, R. Planque, T.Kovacs, N.R. Franks, On optimal decision making in brains and social insect colonies, J.R. Soc Interface 6,1065 (2009).

[9] I.D. Couzin, Collective Cognition in Animal Groups, Trends Cogn. Sci. 13,36 (2008).

[10] Seeley, Thomas D., Visscher, P. Kirk. Choosing a home: How the scouts in a honey bee swarm perceive the completion of their group decision making. Behavioral Ecology and Sociobiology 54 (5) 511-520.

[11] Seeley, Thomas D. Honeybee Democracy. Princeton University Press, 2010.

[12] Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd. "Can there ever be too many options? A meta-analytic review of choice overload." Journal of Consumer Research 37.3 (2010): 409-425.

[13] ESPN College Bowl Picks, published on December 7, 2016. http://espn.go.com/college-football/story/_/id/14306141/forty-games-forty-one-predictions