LA-UR-98-2227

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

TITLE: **COLLECTIVE PROBLEM SOLVING: Functionality beyond the Individual**

AUTHOR(S): N(orman) L. Johnson, T-3
norman@santafe.edu

AVAILABLE AT: *http://CollectiveScience.com*

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2629 5/81

Collective Problem Solving:
Functionality beyond the Individual

Norman L. Johnson

Theoretical Division MS B216
Los Alamos National Laboratory
Los Alamos NM 87545 USA

norman@santafe.edu

Collective Problem Solving

Abstract

Following a non-reductionist approach to the explanation of higher functionality observed in collective problem solvers, a simple agent-based model is used to "solve" a sequential problem - a maze. Larger collectives of the individual agents are observed in the simulations to locate a minimal path, even though the agents are non-interacting, have no global perception of the maze and use rules that do not include logic for finding a shorter path. The convergence to an optimal path is argued to be a demonstration of both an emergent problem formulation and emergent problem solution. Furthermore, many of the dynamics and properties of cooperating collectives are observed: performance of the collective greater than that of the average individual, reduced performance with less diversity, ability to function in the presence of extreme noise and information loss, improved collective performance with established individual problem solvers, path sensitivity to individual contributions but limited sensitivity of group performance, and others. The implications of the results to the formation of self-organizing knowledge and decision making systems are discussed.

Keywords: diversity, collective, self-organizing, emergent problem solving

Introduction

Are there fundamental processes in biological, sociological and economic systems where individuals “solving” their own “problems” result in collective problem definition, knowledge creation or problem solving which is greater than that of the individual? By individuals, we mean organisms, groups, organizations - any entity that is localized in physical or conceptual space. By problems, we mean both problems that are consciously defined and ones that are not explicitly stated but are still solved by the laws or structures governing the global system.

A typical approach to answering this question might be the examination of a system with interacting agents that cooperate and compete with some knowledge of shared resources, be it ecological niche, commodity market, or social position. The individuals are often taken to be agents who can modify their behavior based upon their assessment of their roles and outcomes. As researchers we often attribute significant capability to our idealized agents in order to explain the observed functionality at a global level. But what if there are other mechanisms for functionality that are being overlooked. Two real world examples are given to motivate the possibility of an alternative viewpoint.

The first example is the formation of walkways following a new building development - a modern-day example of path formation in nature. The pre-determination of walkways which best captures path preferences of the users is often an exercise in folly, as judged from the alternative pathways that people quickly develop. These planned solutions probably fail because of the multiplicity of the factors to be considered: different destinations, terrain, security, exposure to the weather and modes of travel. Some planners have learned that often the best solution is to let the “system” determine the paths by first having grass with no paths and then gradually converting emerging paths to formal walkways. These final paths represent the collective action of many individuals solving their own path problem, in a manner that is ultimately useful to the entire population but which is never expressed as a goal at the level of the individual.

A second example is the recommended book lists at Amazon.com, an online bookstore. When a customer finds a possible book of interest, there is also shown a list of books that are related. These lists are constructed by displaying, according to frequency, the books that were purchased by people that also purchased the found book. The lists are exceptionally useful to search a sequence of related books until a desired book is found. Given that the possible choices exceed a million books, staffs of human booksellers would have great difficulty with the success rates of this recommendation method. Yet, this capability is founded on the simple process of capturing the purchase habits of individuals.

These examples illustrate how collective functionality at a global level can occur without intentional "problem solving" on the part of the individual. Arguably both of these examples involve mechanisms that reinforce emerging patterns (selection of an existing path or book), but equally arguably both examples can exhibit collective functionality even if these positive reinforcing mechanisms are eliminated. How then is it possible to reconcile the traditional approaches of collective problem solving involving cooperation and competition of globally “aware” individuals and the above examples of global problems being solved without awareness of the individual? There is a growing body of literature that is addressing this reconciliation in the fields of biology, economics and sociology.

A pivotal analysis of traditional approaches to collective action by evolutionary biologists and cognitive scientists was presented by Hemelrijk (1997). Through a straight-forward simulation of herd formation that includes only aggressive agents (i.e., there is no inherent mechanism for cooperation *embodied* in the individual), the collective behavior of cooperation is observed. She concludes that cooperative behavior from essentially uncooperative individuals is a global structure that emerges from the dynamics and spatial extent of the system. In contrast, traditional approaches to modeling this cooperative behavior arise out of the assumption that the agents themselves embody the choice to cooperate or not, leading to a game theoretical analysis. Said another way, these approaches begin with the potential *in the individual* of the observed global behavior. The explanatory value of these traditional approaches to cooperation in animal and humans is rightly questioned when the same behavior can be observed without assuming it exists at the level of the individual.

The alternative approach taken by Hemelrijk, and the one taken here, is that emergent functionality (a global property that cannot be predicted or observed at a lower level) can be achieved without embodiment of the functionality in the individual. This functionality can include, not only cooperation as in the Hemelrijk's study, but also abstract concepts of problem solving beyond the perception of the individual. The approach of viewing complex global systems as a consequence of the dynamics among relatively simple agents comes out of the field of complexity and complex adaptive systems (e.g., Forrest, 1990; Mitchell, 1996). A review of this field is beyond the scope of this study, but the essential aspects of these systems are illustrated in the following examples.

One of the classic expressions of a self-organizing social/economic system was captured by Adam Smith's description (ca. 1776) of the "invisible hand" in a decentralized capitalistic economy. The understanding is that the individual, be it a consumer or business, pursuing their own interests, contributes to a self-organizing system. At the level of the individual in the economy, global regulation and "decisions" are seemingly being made that are beyond understanding or awareness of the individual. As for the past theoretical treatments of cooperation mentioned above, the reductionist approach of microeconomics to explain macroeconomics is also under criticism (Kochugovindan & Vriend, 1998) as being too quick to attribute sources of the observed global behavior to properties of individual.

Biological evolution (e.g., Fisher, 1930; Hofabauer & Sigmund, 1988), and the associated field of Artificial Life (e.g., Adami, Belew, Kitano & Taylor, 1998) have contributed greatly to the understanding of distributed, self-organizing systems. But to a large extent, the theory and modeling approaches focus on mechanisms based on improvement of functionality by survival of the fittest, in the cycle of mutation, selection and amplification. While changes due to natural selection are driven by global pressures, the source of increased functionality is argued to occur at the level of the individual. So even within the fields that have brought us the greatest understanding of these systems, there still is a significant reductionist approach. Even recently, the source of cooperation at a global level is argued (Sober & Wilson, 1998) from a viewpoint of natural selection, but operating on the competitive survival of collections of similar organisms, instead of individual organisms. This viewpoint still embodies the increased functionality with the entity undergoing selection.

Within evolutionary biology there are developing non-reductionist viewpoints which deemphasize the role of competitive selection at any level. For example, Kauffman (1993) argues that the observed high degree of order on our planet emerges from interactions first, and then is refined by natural selection. Even more ardent is a growing popular work (Margulis & Fester, 1991) that focuses on the role of cooperation and symbiosis in evolution. We conclude that the trend within evolutionary biology is also towards an understanding of increased functionality based on global considerations, with a reduced emphasis on the role of individual functionality.

While the above interpretation of the literature is possibly controversial given the many possible counter examples in the breadth of these fields, this summary sets the stage for a developing major shift in viewpoint by researchers. There is now a major shift from the desire to *understand* distributed, self-organizing systems to begin *creating* these systems. This largely is due to the transformation of a manufacturing-based society to a knowledge-based society (Drucker, 1994). Organizational structuring, management of knowledge and the support of decision makers within the fields of social sciences, economics and library or communication sciences are both responding to and influencing changes in our society.

From the early 1990s, there was an explosion of research and commercialization of online information systems, identified by information filtering, collaborative filtering, computer supported collaborative work, organizational knowledge management, groupware (see, e.g., Maltz &

Ehrlich, 1995; Moukas & Maes, 1998; Smith, 1994; Twidale, Nichols & Paice, 1997). The focus is on the development of online tools that engage the user to collaborate directly with other colleagues or to indirectly assist unknown participants. While the techniques can involve self-organizing approaches, the intent is to enhance existing modes of collaboration and information sharing. The challenges in implementing these systems are extreme due to the breadth of human experience and modes of cooperation in which they operate (Heylighen,1998).

Just as in the examples above, the intent here is to consider mechanisms of self-organization in contrast to the reductionist models of problem solving and knowledge generation. The goal is to study an idealized system where a problem is solved at a global level that is never expressed at the individual level. We wish to address the following questions. What are the dynamics of these systems (stability, instability and chaotic behavior) and from what processes do they originate? What is the global functionality (the emergent property) sensitive to? How do the details of the individual actions affect the collective performance? What are the mechanisms that lead to a collective performing better than the individual? What are the capabilities of the individuals that result in or interfere with the observed dynamics and emergent properties? What is the relationship between the capability of the individual and the level of difficulty of the problems that can be solved at the global level? By answering these questions for a well characterized system, we hope that existing and future systems in society which make use of self-organizing processes will benefit (Johnson et al., 1998).

In the following section an overview of the method is presented. Then, the description of the model system is presented in detail. The next section summarizes the results of the simulations within the narrower context of the virtual simulations. The final section discusses the results in a broader context of social systems and the implications of the present work. These include the discussion of emergent problem solving, its relevance to collective problem solving and the importance of diversity.

Overview of the Model Problem

In the following study, an *individual* agent may represent a single person, group, organization or government within a larger structure of a group, organization or world respectively. The individual is one of many decision-making agents that are localized in either physical space or knowledge space. As used here, a knowledge space is similar to the concept used in the informational sciences (Twidale et al., 1997): a collection of related information within which further knowledge is created or problems solved.

The individuals are identical in the sense that they have the same capabilities and have access to the same information. They differ only in their learned behavior, and their consequential experience and performance. The individual's *goal* is to make a series of sequential decisions, which define a path through the problem domain, ending at a specified terminus. Individuals are taken to be independent - they do not interact or cooperate in any way with one another. We isolate the individuals, because we are interested in the dynamics of collective decision making, in the absence of the complexities of shared learning, cooperation and competition. The problem domain, a *maze*, is a system limited in extent that defines the options for the individual at a particular point in the sequence of decisions. The maze has a starting and end point (goal) for all the individuals. The maze could represent a spatial problem, such as an actual maze, or a conceptual problem, such as collection of possible sequential decisions by individuals in an organization leading to some goal. See, for example, the illustration given by Hong & Page (1998) of a city council allocating public funds to a set of projects.

A population is a large number of individuals (hundreds to thousands). In the first of two sequential phases, the *Learning Phase*, the individuals in the population use *Learning Rules* that specify (1) their movement at each node in the maze and (2) how they modify their own path preference at each node (these are called the Nodal Path Preferences or just *Nodal Preferences*). All individuals use the same set of Learning Rules and only have available local information at any time. *They have no global sense of the maze or their goal and explore the maze until they just* *happen to reach the goal.* This cannot be over-stressed: because the individuals do not have a global sense of the maze, they have no awareness of the correct direction to find the goal. Hence, they have no sense of their path, let alone a short path, through the maze. Another set of rules, the *Application Rules*, then use the *Nodal Preferences* to find the "optimal" path for each individual or, as discussed next, of a collection of individuals. The Application Rules select the preferred path based typically on the largest magnitude of the Nodal Preferences at a node.

Because random choices are made in the rules between paths of equal preference in the Learning Rules, a diversity of areas frequented in the maze and a diversity of final path lengths (*performance*) are created. In general, the individual's optimal path is not necessarily reproducible or unique: if an individual solves for their optimal path many times, the path may differ from solution to solution due to the random selection between choices of equal preference.

To motivate the division of the solution method into two phases, the following metaphor is made. Suppose that an individual is in a town and is looking for a specific café but does not have a map of the town. Her only recourse initially is to search randomly for the café until it is found, optimizing the search based on prior history when paths are repreated. This process corresponds to the Learning Phase. Once the cafe is located, she is able to use the information from the earlier search to optimize her path on a return trip. This process corresponds to the Application Phase.

Another metaphor for the proposed model is an individual's search for a specific site on the distributed database, like the Internet. Assuming the extreme situation that an individual doesn't know what they want until they see it (i.e., all choices appear equally attractive until the goal is reached), she first randomly searches connected sites until the goal is reached (Learning Phase). When the search is repeated (assuming the address of the final site was not saved), she will use the prior information to optimize her path to the goal (Application Phase).

Once the Individual Nodal Preferences are found from the Learning Phase, these can be combined for a group to create *Collective Nodal Preferences*. The identical set of the Application Rules is then used to determine a *collective solution*. This solution represents the decisions of a group of individuals together solving the same problem. Different choices are examined in the creation of the Collective Nodal Preferences from the Individual Nodal Preferences. For example, the simplest construction of a Collective Nodal Preference is an average of the Individual Nodal Preferences at each node for all of the individuals in a collective.

Returning to the earlier café metaphor, what is the analog for the collective solutions? Suppose a group of individuals is heading to the same café. At each corner, they combine their own experiences and collectively chose a preferred path based on the same rules they each used as an individual. This café metaphor is similar to the methods used by an isolated ant to forage for food (Learning Phase) and to optimize the return path (Application Phase) by use of pheromones (Dorrigo & Gambardella, 1997). For a collection of ants, the primary differences are that ants continually optimize the path information (the Learning and Application Phases are concurrent), and they do not move as a group, but continue to move as individuals.

Description of the Model and Analysis Methods

Assumptions

The following assumptions are made in the model and are examined in detail in the discussion section.

1. The Learning and Application Phases are sequentially applied and only once.
2. The problem domain is discrete, i.e., the decision points are finite and localized, and the connectivity fixed.
3. The information available to an individual at a decision point (node) is independent of the path that they took to get there.
4. Individuals "solve" the same problem.

5. Individuals are independent of each other, even in a collective.
6. Individuals have identical capabilities.
7. Individuals have identical assessment of information (values).

Domain

The domain for the simulations is a connected, undirected graph with arbitrary connectivity (edges in graph theory) between nodes or vertices (assumption 2 above). To simplify the solution space, the following restrictions on the graph are assumed: a node cannot be connected to itself and at least one path (sequence of connected edges) must join the start and end nodes. Otherwise, the graph is unrestricted: a node can have any number of connections to other nodes. We note the path "lengths" refer to the number of connections or edges in the path, not to the actual length of a path. This treatment is ideal for the representation of a knowledge space, such as the Internet.

The problem domain or "maze" is created from the graph by defining a start and end node. Fig. 1 shows the example maze with 35 nodes. Note that the maze has three primary minimum solution paths, where a primary path has a common portion of the path that does not change. The maze in Fig. 1 will be used for all later simulations, unless noted. While the maze is simple to solve from the global perspective of Fig. 1, the solution of the maze from a nodal viewpoint (imagine standing at a node and only seeing the options represented by the connections to that node) is challenging. This is the limited perspective of the individual in the simulations.

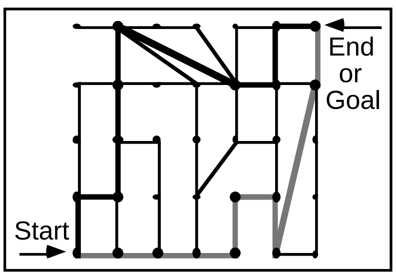


Fig. 1. The example maze. Two of the 14 minimum length paths are highlighted.

Nodal Path Preferences: *P*

At each node in the maze and for each individual, a *Nodal Path Preference*, a scalar number, is stored for the end of each link connected to that node. It is called a *preference*, because the link with a value that is larger compared to the other preferences at that node will be the most likely choice of a path. Together, all the Nodal Path Preferences for one individual can be described as a directed graph overlaying the undirected graph that constitutes the maze. (A directed edge has two values associated with the two ends.) There is a set of *Nodal Path Preferences* for each individual, and because the individuals are independent, each individual has access only to its own Path Preferences.

Let P_{miq} be the *Nodal Path Preference* for individual *m* for link *i* to *j* by link *q*. (Note that we use $\boldsymbol{P}$ for the an individual and Π for the collective; we use bold symbol for these variables when it is being used to describe more than a single value, i.e., when one or more subscripts are absent. The index convention is to use *m* for members, *i* and *j* for nodes, and *p* and *q* for edges.)

Although $\boldsymbol{P}$ is associated with a single individual, it is best thought of as being a property of the maze. An individual only has access to values of $\boldsymbol{P}$ local to its current position. This perspective is consistent with the earlier analogy of an ant solution to a maze using pheromones as a path marker; the pheromones are a property of the spatial location, not of the ant. This approach guarantees the individual has no global sense of the maze.

Learning Phase Rules: Exploring the Domain

The individual begins at the start node with ***P*** initialized to 0.0. The following rules determine the choice of a next node, call it *j*, and the modifications of ***P***, when an individual is at node *i*:

1. If the current node is the end node then stop.
2. If there exists any connecting node with $P_{miq} = 0.0$:
   - Choose a link randomly from the set of links with a value of $P_{miq} = 0.0$,
   - Set P_{miq} for this link to unity and P_{miq} for the reciprocal link from *j* (*j* to *i*) to 0.1*,
   - Set all other links with P_{miq} of unity to zero*.
3. If all links have P_{miq} greater than zero:
   - Choose a link with the maximum value of P_{miq}.

where * means the model results are indifferent to this choice – it can be any value greater than zero and less than unity. The above set of rules is called the *Learning Rules*.

It is important to note that there is nothing special about the above set of rules except that they use only local information and contain no specific test for a shorter path. In terms of the results in this study, any similar set of rules give the same conclusions throughout. In fact, hundreds of rule sets were examined, and the above were chosen because of their simplicity, robustness (no infinite loops) and relative performance. The exception to this statement is a set of rules which produces a random walk; this will be discussed in detail later. Also note that all individuals use the identical set of rules; hence, they are all embodied with the identical capabilities.

Application Phase Rules: Applying Experience of the Domain

The individual begins at the start node. The following rules determine the choice of a next node, call it *j*, when an individual is at node *i*:

1. If *i* is the end node, then stop.
2. From the set of links connected to *i* that have not already been tried:
   - Choose *j* randomly from the links that have a maximum value of P_{miq} (the maximum value can be zero) and if possible, exclude from this choice the last node occupied.
   - Mark the link to *j* as having been tried and exit.
3. If all the links have been tried
   - Pick *j* at random from all links and exit.

The Application Rules select for the maximum value of P_{miq}, excluding the node just vacated or the ones that have already been tried. These, in fact, are just the rules of the Learning Phase reversed: the maximum values of ***P*** are used first, and if they fail to find the path, then a random search is used. Later, we shall see that the choice of selecting a maximum *is* critical to the performance of the collective. Note that only local information is used: information from node *i* and the last node occupied. Also note that the values of ***P*** are not modified in this phase (except as discussed in the next section), i.e., there is no additional learning during the Application Phase.

Novice versus Established Preferences

As will be observed in the simulation results shortly, the ***P*** from the Learning Phase contains information that is never used *by the individual* in the Application Phase. Define a new set of nodal preferences, call it $\boldsymbol{P}^{\dagger}$, using only the values of ***P*** used in the Application Phase by the individual. This new $\boldsymbol{P}^{\dagger}$ will result in the identical path of the individual as the original ***P***, assuming there are no multiple maximum values of P_{miq} (this was observed to be a very rare event).

We associate the unmodified values of ***P*** as the preferences of a novice problem solver that has not experienced which information is useful or not. And we associate $\boldsymbol{P}^{\dagger}$ with an established problem solver that has experienced which information is useful or not. These define the two extremes of the information contained in ***P***. Arguably a novice applying learned information for the first time would possibly retain knowledge of rejected paths. But the established problem solver after many solutions of the same problem, would quickly discard this information. Note that this distinction does not affect the performance of the individual, but is observed to affect the collective solution.

Alternative Application Rules: Probabilistic Selection

In the Application Rules above, the guiding principle was to select the *maximum* value of ***P*** at a node when possible. An alternative would be to select from the Nodal Preferences based on a *Probabilistic* selection: randomly select from the links such that over many selections, the frequency of selection will duplicate the values in ***P***. One view of this option is that, instead of the ***P*** being a preference, it represents a probability of choosing a path. Where we chose a maximum preference typically, in some situations, such as a choice of food, we might select from a variety of options with a Probabilistic selection.

This option is implemented by defining a cumulative distribution function, ***C***, such that:

$$C_{mip} = \frac{\sum_{q=1}^{p} P_{miq}}{\sum_{q=1}^{N_i} P_{miq}} \qquad \text{for } p = 1 \text{ to } N_i \tag{1}$$

where N_i is the number of links connected to node *i*. A link is selected by finding a random number, *r*, between 0 and 1 and then finding a link *q* such that:

$$C_{mi(q-1)} \le r < C_{mi(q)} \quad \text{for } q = 1 \text{ to } N_i \tag{2}$$

where $C_{mi(0)}$ is defined to be zero. In the Application Rules, performance of the method was improved if the node that was just vacated (backstepping) is eliminated from the possible choices. This option can be examined in the Probabilistic approach by setting P_{miq} to zero for this link in Eq. 1.

Forming a Collective Nodal Path Preference

By applying the above Learning Phase to many individuals, a population of ***P*** is found, which differs in exploration of different parts of the maze. We define an *ensemble* for a collective as a set $\{g\}$ of members in a collective as found from a random sampling of individuals from the entire population of individuals without duplication. For example, one ensemble of a collective with 5 individuals is (3,22,6,81,15) from a population of 100. Another ensemble is (24,3,10,5,29). And so on. For a *ensemble* of individuals, ***P*** (or $\boldsymbol{P}^{\dagger}$) can then be combined *at each node* in various ways to create *Collective Nodal Preferences*, Π. Then the same set of the Application Rules, as given above for an individual, is used to determine a path through the maze for the collective. Note that the Individual Application Phase is not required to find the solution for a collective in general. In one sense, the collective is a "super-individual" in that it has access to much more information, but has the same capability as an individual.

For example, a simple Collective Nodal Preference is the average of the Individual Path Preferences at each node for all of the individuals in a sample population.

$$\Pi^{*}_{giq} = \frac{1}{N_g} \sum_{m=\{g\}} P_{miq} \tag{3}$$

where the sum is over the values of *m* in the set $\{g\}$ and N_g is the number of members in $\{g\}$. We attach the superscript * to Π to identify it as the value of Π used for the reference simulation.

Although Eq. 3 is a linear superposition of the individual preferences, it does not follow that the dynamics of the collective system behaves as a linear system when the Application Rules are used. The presence of the selection of the maximum preference in the Application Rules results in a highly non-linear system, with all the associated dynamics as will be demonstrated. If the Probabilistic Selection is used with Eq. 3, then the collective system is linear in the mathematical sense of the usage.

As we shall see in the next section, Eq. 3 results in a collective simulation that typically locates a minimum path of the maze when the Application Rules are used for sufficiently large collectives. Because the formation of Π^* is done without consideration of the contributions of the individuals, this collective represents a group of non-interacting individuals. Other ways of constructing Π are a modification (as described below) of the $\boldsymbol{P}$ of the individuals that make up the collective. Because $\boldsymbol{P}$ represents an individual's preferences or opinions at a juncture in a decision process, a sociological interpretation can be associated with the modification. For example, Π can be created from only the maximum values of $\boldsymbol{P}$ in Eq. 3 at a node from each individual, setting other values of $\boldsymbol{P}$ to zero. The interpretation is a situation where an individual contributes to the collective decision only the opinions that are strongest, omitting the less important ones.

Statistics in the Collective Solution

Even though all individuals are identical in their capabilities and in their access to information, the random choices in the Learning Rules lead to a diversity of paths and performances. Consequently, the simulations of the collectives making up of this diverse population of individuals exhibit paths which have variations even within collectives of a fixed size. This is particularly true when only a few individuals make up the collective. These variations can occurr in both in the path taken and in the path length of the collective.

A collective solution comprised of a fixed number of individuals can differ in two ways. First by the different individuals that contribute to the collective preference as selected from the entire population, and, secondly, by the different random numbers that are used within a fixed set of individuals. The second difference means that if two different sets of random numbers are used in two collective simulations with the identical set of individuals, different results can occur.

To aid in the analysis of the simulation results, we define the following. A *realization* of an ensemble is the collective solution for a given ensemble. For a given ensemble size, there also can be many realizations, either from different member compositions or from different random numbers. A collection of ensembles with the same number of individuals is called an *ensemble set*. The average of the results of a collective simulation over an ensemble set is the *ensemble average*. Statistics for the simulations are collected for ensemble averages for the range of individuals from 1 to the some maximum ensemble size.

Noise and Loss of Information in Communication

In the study of different ways of combining the individual's contribution to form a collective preference, we consider two types of degradation: noise in the contribution and loss of information of a contribution.

Noise is the random *replacement* of information contributed by the individual. Because it replaces *valid* information, noise represents *false* or inconsistent information. The effect of noise is implemented in the simulations by the random replacement of P_{miq} by a small value (0.1 is used but the results are insensitive to this value if it is less than unity) with a specified frequency in the creation of Π. For example, for a frequency of 0.3, any Nodal Preference will, on average, be replaced by 0.1 in Eq. 3 thirty percent of the time.

Loss is the selective *reduction* of information contributed by the individual to the collective decision. For example, loss can be the selective elimination of some individuals out of a set of contributing individuals or by the selective modification of the information contributed by an individual, as in the example in the last subsection of the elimination of weaker opinions. The effect of loss is the selective reduction of information available to the collective, but loss does not add false information to the collective decision.

The following formalism captures many of the options for loss applied later. We define a new path preference ***P'*** with loss of information as follows.

$$\beta' = \frac{1}{N_i}\left(1+\beta\left(N_i - 1\right)\right) \tag{4}$$

$$P'_{miq} = \beta' P_{miq} + \frac{(1-\beta')}{(N_i - 1)} \sum_{q \neq p} P_{miq} \tag{5}$$

$$\Pi_{giq} = \frac{1}{\mathrm{N}_g} \sum_{m=\{g\}} \Theta\left[\alpha \frac{Max[P_{mi}]}{Max[P'_{mi}]} P'_{miq}\right] \tag{6}$$

where

$$\Theta[x] = \begin{cases} x & \text{if } \; x > \tau \\ 0 & \text{if } \; x \leq \tau \end{cases} \tag{7}$$

and where α is the attenuation of the maximum value ($\alpha \geq 0$), β is the reduction of the dynamic range (maximum minus minimum) ($0 \leq \beta \leq 1$) and τ is the of the modified value of ***P*** ' below which the contribution to Π is zero, otherwise it is unchanged. For β=1, the limiting expression for Eq. 6 is

$$\Pi_{gip} = \frac{1}{N_g} \sum_{m=\{g\}} \Theta\left[\alpha P_{mip}\right] \tag{8}$$

Eq. 8 reproduces the reference simulation in Eq. 3 with $\alpha = 1$ and $\tau = 0$. These equations are constructed such that the maximum before and after the "flattening" of $\boldsymbol{P}_{mi}$ using $\beta < 1$ is unchanged. Therefore, in the limit of β=0, only the maximum value of $\boldsymbol{P}_{mi}$ is selected:

$$\Pi_{giq} = \begin{cases} \dfrac{1}{\mathrm{N}_g} \displaystyle\sum_{m=\{g\}} \Theta\left[\alpha \;\; Max[P_{mi}]\right] & \text{for link q with maximum value} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Measures of Diversity

Because one of the main conclusions of the present work is the importance of diversity in distributed self-organizing systems, it is helpful to define quantitative measures of diversity that can be applied to the simulations. In the present model diversity has multifarious aspects, each of which can be evaluated separately. Two primary sources of diversity can be observed: diversity due to difference in capability and diversity due to difference in experience given the same capability, *heuristics* and *perspectives* respectively as used by Hong & Page (1997). In the current study, all individuals have the same capabilities (i.e., they all use the same rules) within a set of simulations, so all the differences in the performance of the individuals arise only from diversity as a result of random choices in the Learning Rules, and *not* from differences in capability.

Diversity that results from experience can be further divided into two aspects: diversity of experience across the domain space (*experiential* or *domain diversity*) and diversity of preferences at one place in the domain space (*preferential* or *nodal diversity*). Both are the extent of the sampling, leading to experience, from the Learning Phase. Experiential diversity of a group is probably closest to the common use of diversity, both in a social (Thomas & Ely, 1996) and biological context (Kauffman, 1993), and is related to the varieties of distinct viewpoints that are used to solve a problem. Preferential diversity reflects the diversity of knowledge at a specific point in a sequence of decisions. For example in the café metaphor used earlier, experiential diversity captures the extent of experience with alternative routes to the café, while preferential diversity is the extent of experience of the choices at one intersection. While one would expect that preferential diversity of a group leads to experiential diversity, they play distinct roles in the current self-organizing system.

Within each of these aspects of diversity, the relative magnitudes (importance) of the sampling can also be examined; i.e., we can distinguish between the *possibility* that a path will be taken and relative probability that it will be taken. The experiential diversity is related to the breadth of the possible sampling of alternative paths through the maze as a whole, but says nothing about the frequency of this sampling. For example, the novice expression of $\boldsymbol{P}$ has more experiential diversity than the established expression of $\boldsymbol{P}$, $\boldsymbol{P}^{\dagger}$, but they both yield the same paths in the Application Phase for an individual and therefore the identical frequency of sampling. Similarly the preferential diversity is the breadth of the possible sampling of different paths at a node in the maze, but says nothing about the relative magnitudes of the preferences of the paths.

While measures for each of the aspects of diversity were developed, only the measures of experiential diversity are presented because relative importance of experiential diversity in the performance of the collective. Note that the viewpoint is taken that experiential diversity is a property of a group and is not meaningful for an isolated individual.

We first define a quantity which is unity if member *m* at node *i* has a non-zero value of P_{miq} and zero otherwise:

$$d_{mi} = \begin{cases} 1 & \text{if } \sum_{q=1}^{N_i} P_{miq} \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

We then define a quantity which is unity if any of the members in group *g* has a non-zero value of d_{mi}

$$d_{gi}^{n} = \begin{cases} 1 & \text{if } \sum_{m=\{g\}} d_{mi} \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

One possible choice of a measure of experiential diversity is simply the relative number of nodes that have non-zero values of P_{miq} for group *g*:

$$D_{g}^{n} = \sum_{i=1}^{N_t} \frac{d_{gi}^{n}}{N_t} \tag{12}$$

where N_t is the number of nodes in the domain. This *cumulative experiential diversity* approaches unity as the members of the group have experience in the entire domain. Because D_g^n was observed to become quickly unity for even small groups and not to correlate with collective performance, an alternative measure is proposed, called the *unique experiential diversity* or simply the *experiential diversity*.

$$D_{g}^{e} = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{d_{gi}^{n}}{N_g - 1} \left(N_g - \sum_{m=\{g\}} d_{mi} \right) \tag{13}$$

This measure gives more weight to the contributions of the individuals that have the least commonality with the group. In other words, individuals with potential experiences that are not shared by others have the most weight. As shared potential experiences increase at a node, the weighting is less, until it is finally zero if all members of the group share the same potential experiences. We use the words “potential experience,” because no consideration is given to the relative magnitude of the preferences at a node; the preferences only need to be non-zero. While Hong and Page (1998) in their work on diversity in heterogeneous individuals do not define an experiential diversity, the experiential diversity in Eq. 13 is analogous to their measure for capability diversity.

Eqs. 12 and 13 have the following properties:

1) D_g^e / D_g^n is unity if the experiences (non-zero path preferences) of all the members in the group have no common nodes,

2) D_g^e is zero if the experiences of the members of the group have all common nodes, and

3) $0 \leq D_g^d \leq 1$.

The first property results from the observation that if no members share a common experience at a node, then the expression in the brackets in Eq. 13 reduces to (N_g-1) and D_g^e is equal to D_g^n. The second property is apparent because the quantity in the brackets in Eq. 13 is zero if all members have experience at a node. Finally the third property results from the observation that the sum in Eq. 13 is always less than or equal to N_t. Because every member in the group shares a common starting and end node, the experiential diversity will never be unity in the current study. Note that D_g^e and D_g^n are global measures of diversity, meaning that depend on the entire context of the problem. For example if the extent of the maze is increased (N_t larger), then for the identical ***P*** of the individuals in the collective, these measures will be smaller. In contrast, D_g^e / D_g^n is a local measure of diversity because it is independent of the extent of the problem domain. In the current study, the limited problem domain and the ability of the individuals to quickly explore the entire domain results in little difference between D_g^e / D_g^n and D_g^e.

In the earlier café metaphor, if all the individuals in a collective know a different route (no common junctions), then D_g^e / D_g^n is unity. If the all know the same routes, then D_g^e is zero, even though they might know every street in the city. This example illustrates the role of unique knowledge in this measure of diversity.

Simulation results

The results of the simulations using the Learning Phase and then the Application Phase defined in the last section are presented. The coding for all the simulations was done within Mathematica, and execution time was about ten minutes for the set of reference simulations and a few hours for the simulations with thousands of members using a Macintosh G3.

Individual Learning Phase

The Learning Phase was applied to a large population using the Learning Rules and a variety of random walk methods (Table 1). The first random walk method is just the random selection of a new node at the current location. The *no-backstep* random walk is the random selection of a new node excluding the node that was just vacated. And the *non-repeating* random walk is the selection, if possible, of only untried links. Also, in the table are the statistics for simulations for a 1000 individuals, indicating the effect of population size on the statistics. Note that the sample range will not converge as the population rises, as this is a measure of the extremes of the distribution. Because the Learning Rules capture both the no-backstep and non-repeating options, in addition to other features, it performs the best of the four methods.

Table 1. Path Lengths for a variety of methods for a population of 100 (except as noted). The simulation marked by * uses a different series of random numbers than the others.

Simulation method	Average	Standard deviation	Sample range
Learning Rules	34.3	24.5	165
Learning Rules (1000)	39.2	30.2	375
Random walk	123.	103.	488
Random walk (1000)	129.	105.	678
Random walk (1000)*	128.	104.	714
No-Backstep Random walk	64.0	66.0	336
No-Backstep Random walk (1000)	57.6	48.7	424
Non-repeating Random walk	50.8	52.3	427
Non-repeating Random walk (1000)	46.8	40.3	294

Individual Application Phase

Simulations using the Application Rules

In Table 2, the results of the simulations using the Application Rules are shown for the four methods listed in Table 1. In order to apply the Application Rules to the simulations that use the random walk methods, $\boldsymbol{P}$ was generated for these methods by initializing $\boldsymbol{P}$ to zero and then incrementing the P_{miq} by 1 each time the link q is used by individual m at node i in the random walk. In all cases the path length and standard deviation are improved in going from the Learning to Application Phase.

Table 2. Path Lengths for the Application Phase for a population of 100 (except as noted) from the different methods listed in Table 1. For comparison, the quantities in brackets are the similar values from the Learning Phase.

Simulation method	Average	Standard deviation	Sample range
Learning Rules	12.8 (*34.3*)	3.06 (*24.5*)	13 (*165*)
Learning Rules (1000)	13.2 (*39.2*)	3.30 (*30.2*)	18 (*375*)
Random walk	48.8 (*123*)	54.9 (*103.*)	330 (*488*)
No-backstep Random walk	38.6 (*64.0*)	40.3 (*66.0*)	211 (*336*)
Non-repeating Random walk	33.7 (*50.8*)	36.2 (*51.5*)	234 (*428*)

Correlation between Performance in the Learning and Application Phases

Because the Learning Phase and Application Phase are separate, the cross-correlation between the performance in the Learning Phase with the performance in the Application Phase can be assessed. The Pearson's correlation coefficient between the number of steps in the phases is 0.12 for a population of 1000, indicating a weak correlation in the two sets of data. This information is shown graphically in Fig. 2. Although 1000 individuals are represented in the figure, fewer symbols appear because much of the data occupy the same coordinate. The absence of a diagonal dominance in the plot supports the weak correlation between the two sets of data. We can conclude for these results that "slow" learners are not necessarily "poor" performers.

One notable aspect in the plot is the absence above the line where performance in the Learning Phase is equal to the performance in the Application Phase. Because the Application Rules preferentially select against options that have not been explored, the Application path will usually not be longer than the Learning path. Hence this region of the cross-plot should be sparsely populated. Because this region is actually void, this indicates that the Application Rules never select paths that have not been already explored (the maximum in Rule 2 is never zero). This is a reasonable result because the Learning Phase continues until the end node is found, and therefore a solution must exist within the values of ***P***. Hence, the Application Rules must at least find the prior solution from the Learning Phase.

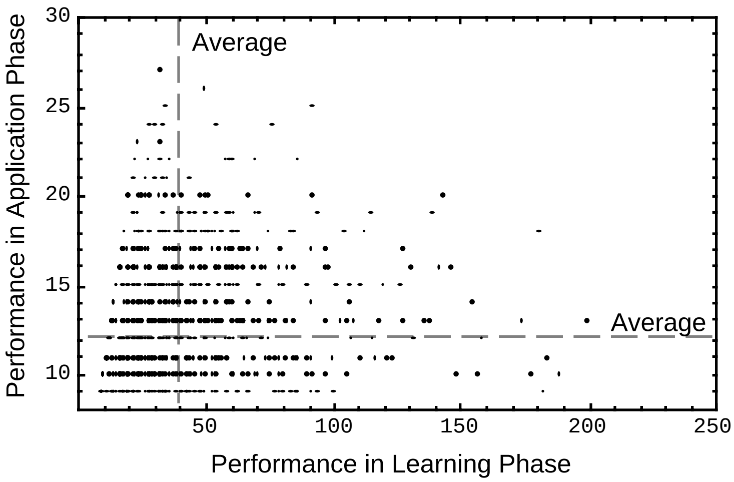


Fig. 2. Cross plot of the performance in the Learning Phase versus the Application Phase for each member of a population of 1000.

Source of the Improvement in the Application Phase

Why is the performance in the Application Phase better than in the Learning Phase? Fig. 3 illustrates how an extraneous loop is eliminated by the Application Phase for a representative individual. The path in the Learning Phase is typically much more complex than shown in Fig. 3, often crossing a prior path many times before the path is terminated, but the mechanism of improvement is the same. From Table 2, the length of rejected loops (about 26 steps) from the Learning Phase is on average about double of the final path length (13.2 steps).

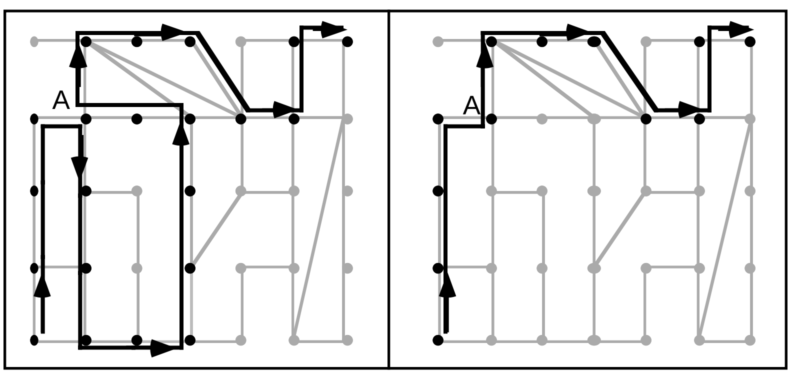


Fig. 3. Plots of the paths in the two phases for a representative individual. The Application Phase removes the extra loop at point A; otherwise the path is unchanged.

The lack of correlation between the phases indicates the predominance of the extraneous loops shown in Fig. 3. If the Learning Phase never crosses its own path, then the Application Rules will duplicate the Learning Phase path and the correlation would be higher. Furthermore, the lack of correlation indicates that the extraneous loops themselves are not correlated with the final path length in the Application Phase but have a random distribution (this can be explicitly shown by examining the correlation between the difference between the path lengths and the final path length). Because the Application Phase duplicates the Learning Phase path except for the extraneous loops, there is no other mechanism for improving the path, for example, by taking a diagonal connection. Mechanisms of this type require a global perspective of the problem, which is not contained in the capability of the individual, but, as we shall see, is a capability of the collective.

Simulations using the Alternative Application Rules: Probabilistic Selection

As an alternative to the Application Rules that select for the maximum of the ***P***, Table 3 shows the results of the Probabilistic selection. We conclude that from the viewpoint of the individual, there is no major advantage between the using the Probabilistic approach over the standard Application Rules (later we will see that this is not true for the collective). The large change in the standard deviation when backstepping is removed from the Probabilistic approach is due to one result with a very high path length (182) when compared to the next highest result (23). This increases the deviation but has little affect on the mean. Excluding this one point results in the two Probabilistic methods being almost identical.

Table 3. Path Lengths using the Probabilistic Selection, compared to the standard Application Phase, for a population of 100, all using the same ***P*** from the Learning Rules.

Approach	Average	Standard deviation	Sample range
Standard Application Phase (Table 2)	12.8	3.06	13
Probabilistic Application Phase			
with backstepping	15.9	5.60	33
without backstepping	15.6	17.2	173
Learning Phase (Table 1)	*34.3*	*24.5*	*165*

Collective Solutions: Emergence, Chaos and Diversity

Collective Performance Using Novice Preferences

The collective solutions, as described in the prior section, are examined using Eq. 3 with the Learning and Application Rules. Note that all the simulations in this subsection use the novice form of ***P***. The effect of established learning and of diversity is examined in the next subsection. In the later sections, noise and loss effects are considered.

Fig. 4 shows the effect of including larger numbers of individuals in the collective solution, with one realization for each ensemble. Although significant variation occurs, particularly for smaller collectives, the collective solution is always better than the average individual when a collective has 20 or more individuals. We define the difference between the path length of the collective solution and the average path length of the individuals making up the collective as the *collective advantage*. The degree of the collective advantage will be more apparent in later figures which include ensemble averages over many realizations. Also shown in Fig. 4 is another selection of the ensembles of different sizes. The differences in the two curves are due to a lesser extent from random choices made in the Application Rules (as we show next), and to a greater extent, from the combination of different sets of individuals that make up the collective. For example, an ensemble with 5 members is shown for two choices: (3,22,6,81,15) and (24, 3,10,5,29).

Another approach to creating an increasing series of ensembles is to add an individual onto an existing ensemble: e.g., (5), (5,66), (5,66,35), (5,66,55,99), etc. This *cumulative* ensemble illustrates the effect of the addition of a single individual to an existing collective solution (Fig. 5). Not surprisingly, the effect of an individual is greater for smaller collectives than larger ones. Even so, the simulations continue to show sensitivity to the addition of a single individual, even in very large collectives. This is essentially a demonstration of the chaotic nature of the system: an infinitesimal change in Π can lead to a large effect in the global path length. In fact, the full chaotic nature of the system is not captured Fig. 5. For a given path length there are many paths with the same length, e.g., the two highlighted paths in Fig. 1. In an animation of the actual paths shown in Fig. 5 (available at the author's website), a major change in the path through the maze can occur by the addition of one individual, without a change in path length. This is a consequence of the multiple primary paths through the maze and the resultant indeterminacy of the selection of one of path over another.

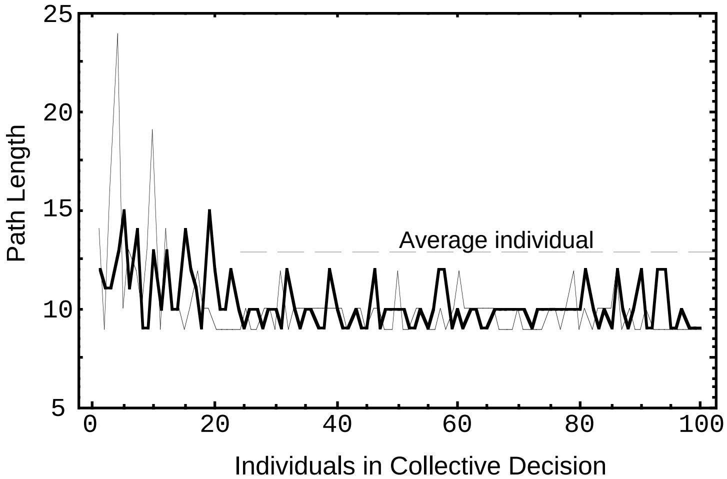


Fig. 4. The path lengths (number of steps) for two different selections for the increasing series of ensembles.

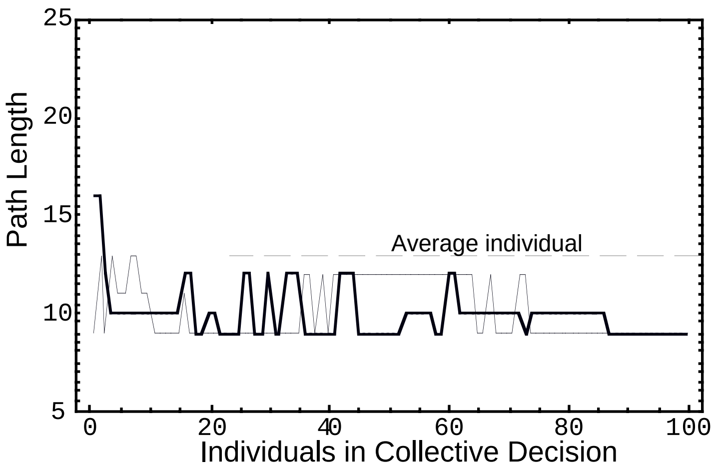


Fig. 5. Path Lengths for a series of collective solutions that adds an individual onto the list of individuals in the previous ensemble, for two different samplings of 100 (same population as in Tables 1 and 2).

To ascertain the effect of the random choices in the Application Phase on the collective solutions, different seeds for the random number generator were used for the Application Phase for 20 identical ensembles. For a given ensemble set, the 20 simulations have the identical members contributing to the collective. The only difference in the simulations were the random numbers used in choices that had equal preference (multiple maxima in Π at a node). In Fig 6, the resulting ensemble average length is shown for the 20 identical ensembles, compared to the solid curve in Fig. 4. Also in Fig. 6, the standard deviation (solid curve) of the 20 simulations is shown, illustrating that the only deviation occurs at points 2-5 and 61 on the abscissa. These results illustrate that sensitivity exists in the solution, but gives no information about the degree of sensitivity. The same study was repeated, but relaxing the test for the maximum of Π at a node in the Application Rules to be within 10 percent of the maximum. These results are also presented in Fig. 6. Two conclusions can be made: 1) the random selection between equally preferred paths is a rare occurrence, 2) the simulations are sensitive to weakening the selection of the maximum preference in Application Rules, but the effect is to increase the randomness, without affecting the overall trends.

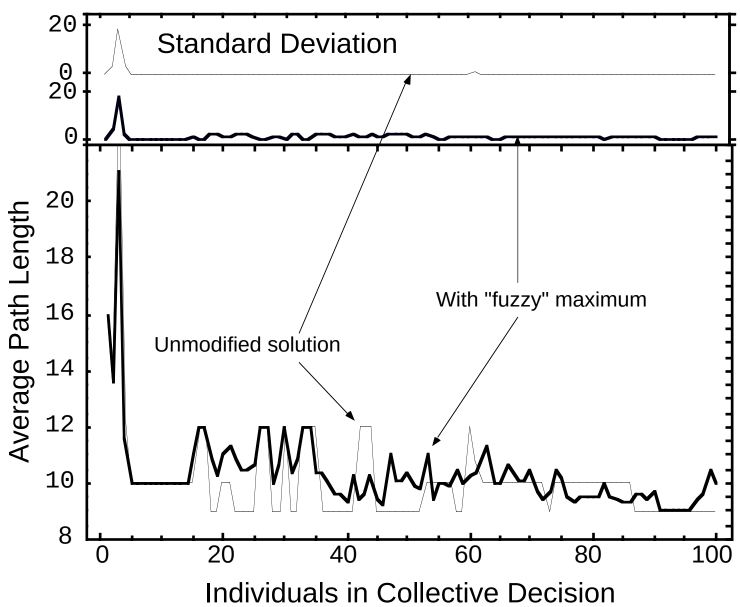


Fig. 6. Average path length and standard deviation for 20 simulations of the ensembles in Fig. 4 with different random numbers. The "fuzzy" curve is with the effect of a 10% relaxation in the maximum value test in the Application Rules.

Collective Performance Using Established Preferences

As illustrated in Fig. 3, the Application Phase of the individual eliminates extraneous loops, and by doing so shortens the path in the Application Phase over the Learning Phase. In an actual problem solving situation, it can be argued that the preference information contained in these rejected loops would be less reinforced, and ultimately lost, as an individual repeatedly solves the same problem. Consequently when an individual contributes to a collective solution this potentially rejected path information might be included for the novice and excluded for the established problem solver.

Fig. 7 shows the simulation results for one of the curves in Fig. 4, along with the simulation results for identical selection of individuals for each ensemble, but using the established preferences, $\boldsymbol{P}^{\dagger}$ in Eq. 3. As seen in Fig. 7, although using the established preferences leads to a general reduction and less variation in the path lengths, the performance is not always better, nor the variation always less. The primary difference is that the solution is more likely to converge to the minimum solution at smaller collective sizes with the established preferences. These observations hold for all studies on different variations of the model that follow. What is remarkable is that the inclusion of the rejected paths in the collective preferences *does not* degrade the performance more. On average, the information contained in the rejected paths is about double the information in the retained paths (see the section on the correlation between the Learning and Application Phase). It appears that the information in the loops is inconsequential to the formation of the collective advantage.

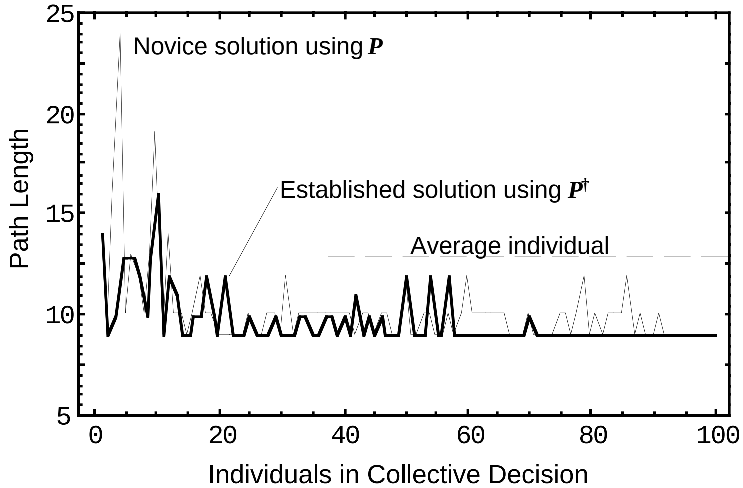


Fig. 7. Comparison path lengths for collectives using novice and established preferences.

Diversity in Novice and Established Collective Solutions

Some insight can be gained on why the simulations using the established preferences outperform the novice preferences by examining the diversity measure presented in Eqs. 12 and 13. In Fig. 8 the cumulative and unique experiential diversity measures are plotted for the simulations in Fig. 7. Most notable is that D_g^n indicates that even small groups (5-15 individuals) quickly gain experience throughout most of the problem domain. Furthermore, D_g^n for the simulations with $\boldsymbol{P}^\dagger$ indicates that experience in some parts of the domain are not represented (for example, dead-ends in Fig. 1), but the absence of this information does not reflect a decline in performance. Based on these observations, D_g^n is concluded to be a poor measure of diversity as influencing performance.

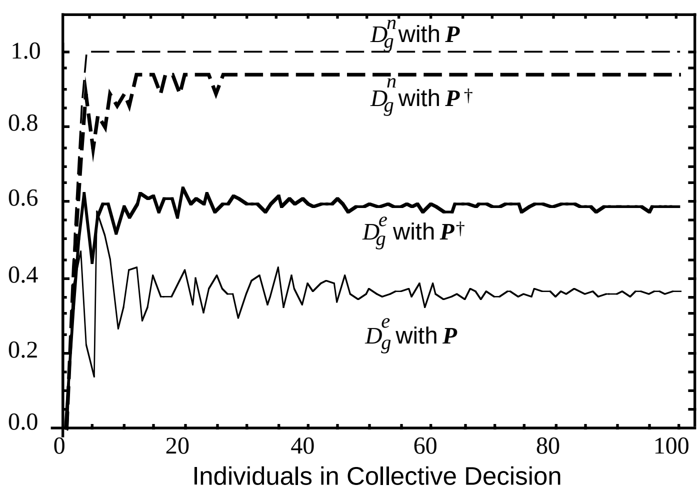


Fig. 8. Measures of diversity for the simulations in Fig. 7.

The other measure of diversity, D_g^e, which accounts for the relative uniqueness of experience among members of a group, shows a different asymptotic value in the two sets of simulations. Despite that fact that the rejected loops contain more information about the domain than the preferred path of an individual, D_g^e is almost 50% larger for the simulations with the established preferences. The difference is due to extraneous information that causes the low diversity groups to be more similar (larger paths overlap more), in comparison to the high diversity groups with limited and unique information. The other notable observation is that D_g^e does not increase with increasing size of a collective. The asymptotic values above a certain size of a collective indicate that the relative uniqueness of an individual is unchanged. Said a different way, increasing the size of a collective does not necessarily increase diversity when the sampling is from a uniformly distributed population. Also, note that the convergence to a minimum solution occurs when the variation in D_g^e declines. Later simulations reinforce these observations.

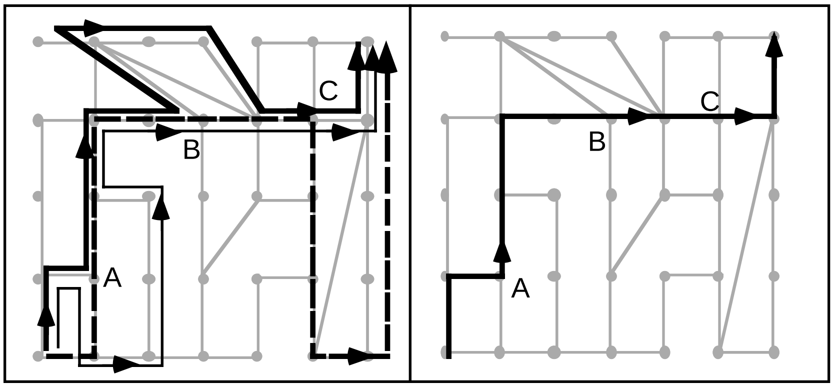
Source of the Collective Advantage

We observe in the above simulations that 1) on average a collective performs better than the average individual making up the collective (called the *collective advantage*) and 2) for collectives of sufficient size, the collective solution converges to one of minimum paths. These results are surprising given that the rules governing the choices of the individual on which the collective decision is formed and the method for forming a collective includes no global perspective (e.g., "The goal is over there, let's go that way"). Nor is there any expression or selection of a minimum path among the collective (e.g., "Let's follow whomever has the shortest path"). Later in the discussion section these global results from local rules are identified as emergent properties of the system and capture a solution to a global problem ("find the shortest path") which is not expressed by the individual. For the purposes here, it means that we cannot look to the model definition to explain the source of the collective advantage. Instead, the collective advantage is a consequence of the interaction of global structures within the collective.

In the following, one mechanism is described that contributes to the collective advantage and is related to the discussion associated with Fig. 3: the elimination of extraneous loops. Because Π in Eq. 3 is a superposition of the contributions of multiple individuals, Π can contain extraneous loops as observed in Fig. 3 which are only partially complete in the individual contributions. Fig. 9 illustrates how this can occur. In the figure each path of an individual contains incomplete extraneous loops that could be eliminated, thereby shortening the path. But because the loops are not closed, each individual is missing essential information to eliminate these paths. By combining the preferences for the paths of multiple individuals, Π now contains complete information and the extraneous loops can be removed from the group path (follow at each juncture the most frequent choice – this typically corresponds to the maximum preference of the collective). In Fig. 9, the information from the paths of the three individuals with path lengths of 12, 14, and 16 are combined to give a collective path length of 10, a result better than any individual in the collective. It is easy to show that by the simple argument made above, the combination of two individuals cannot eliminate unambiguously extraneous paths. This is part of the reason that most couples do no better on average than the individuals that make up the couple. Consequently, the paths of three individuals or more must overlap for this mechanism to work. The above mechanism does not explain why larger collectives result in a minimum solution (although from intuition one might expect this to occur), and a satisfactory explanation has not been found.

Other instances of collective advantage were examined for mechanisms different than described above but trends were difficult to analyze and generalize. These additional mechanisms appear to involve subtle interactions of path preferences and the inherent connectivity of the graph. Further explanations are left to later studies.

Note that in the example given in Fig. 9, the collective solution was better than any of the individual solutions. While this situation was not infrequent, most instances of collective advantage occur when a single "high" performer is combined with "lower" performers. Even though the path of the minimum is not specifically selected ("Who has the shortest path length?") by any of the Rules, the collective solution typically results in a path length that is comparable to the shortest path, but not necessarily the same path as the "best" performer. This results in a collective advantage because the average individual path length is an average of all members in the collective. At first inspection, this may seen to be a variation of a "free rider" problem commonly examined in economics where an individual benefits from the higher performance of a collective without necessarily contributing. But the current dynamics are much more subtle and maybe more profound. All members contribute to the collective in the current model, particularly in regards to reduced sensitivity to noise; the individual with "poorer" performance does make an important contribution as will be seen shortly. And the improved performance of the collective occurs without requiring feedback or cooperation from the individuals.



Paths of three individuals Collective Path

Fig. 9. Elimination of extraneous loops in the collective solution from incomplete loops in the individual contributions.

The Use of Probabilistic Selection on Π

In the subsection on the Individual Application Phase, the Probabilistic selection method was shown to result in path lengths that were only slightly degraded from the average performance of the Application Rules (15.9 versus 12.8 in Table 3). The identical Probabilistic selection method is now applied to Π. As seen in Fig. 10, the Probabilistic method produces significantly degraded results for larger sizes of collectives, particularly for the option where backstepping is allowed.

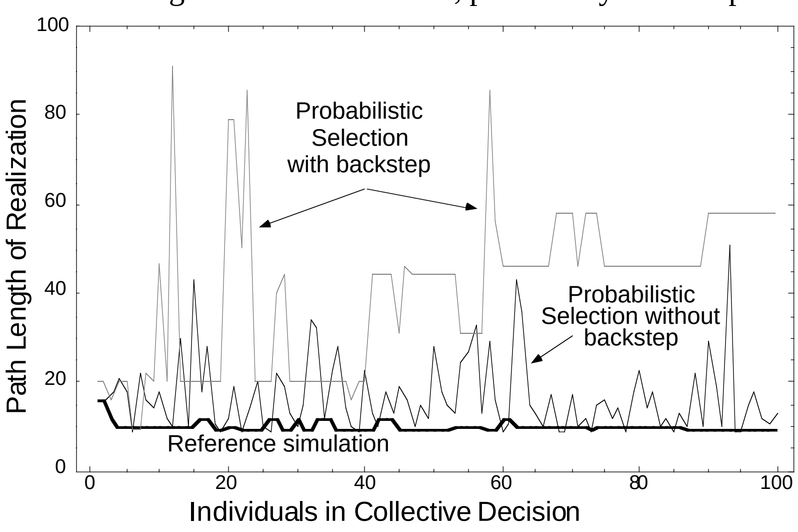


Fig. 10. Comparison of path lengths for the two options for the Probabilistic selection (Table 3) and the Application Rules for a cumulative ensemble. The reference simulation is the same as the solid line in Fig. 4.

In the Probabilistic simulation with backstepping, the cause of the increased path length with larger collectives is a result of the Π becoming more uniform (isotropic) at a node. At the isotropic limit of equal preferences, the Probabilistic solution degenerates to a random walk with a path length around 128 (see Table 1). While Π does not become isotropic for very large collectives, the difference in the minimum and maximum weightings at a node does decrease, causing side paths, which increase the final path length, to become more likely. Both Probabilistic approaches in Fig. 10 were extended to 500 members in the collective; the curve for the Probabilistic method without backstepping behaved similarly, but for the Probabilistic method with backstepping the mean and the variation continued to grow. In the Probabilistic solutions without backstepping, apparently the effect of not returning to the node just vacated is sufficient to prevent the solution from shifting to the random walk solution.

What significance can be given to the sensitivity of the collective solution to the Probabilistic selection relative to the insensitivity of the individual? This question leads to a more general question of how collective systems might require different rules, in comparison to self-organizing systems which do not involve sequential processes. To the author's knowledge, most physical systems use a Probabilistic sampling in either the governing equations or characterization of their dynamics, whether the system is continuous or discrete. Even deterministic systems, such as planetary motion, are being treated as probabilistic systems to account for their chaotic nature (Prigogine, 1998). Exceptions to probabilistic sampling in physical systems are human creations, such as a transistor or laser. Biological systems, by contrast, often use the selection of a maximum state. The common examples are the working of the neurons in the brain or the selection of competing pheromone trails by an ant. We speculate that the ability to select a maximum state in collective biological systems has evolved as a necessary capability to enable the emergent functionality of a distributed self-organizing system.

The Ensemble Average

The prior sections established variations in the realizations of the collective solutions. Henceforth, the results of the collective simulations will be presented as an ensemble average of ensemble sets. For example, the results for the first data point is an ensemble average of fifty simulations with one individual participating, each randomly selected without repeating from a population of 100 members (same populations as in Table 1 and 2). The second point is fifty simulations with two individuals contributing to Π, each pair randomly selected from the population of 100, and so on. To easily identify the collective advantage, the collective path lengths are normalized by the average of the performance of the individuals making up the ensemble set. Because the deviation is relatively low for the Application Rules (Table 2), the actual path lengths are approximately the normalized value times 12.8.

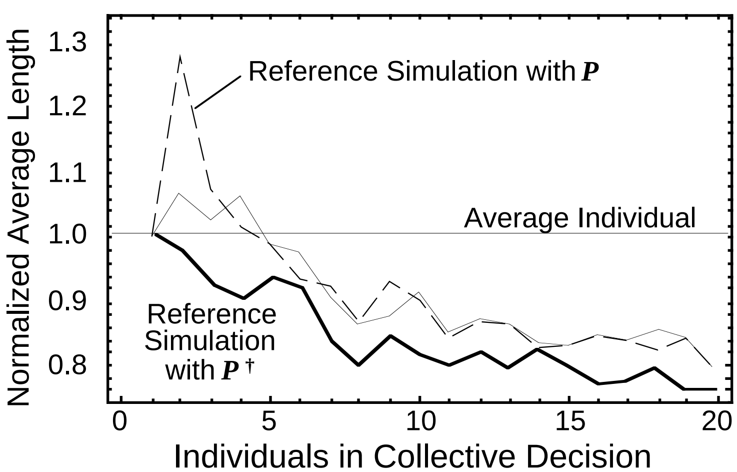


Fig. 11. Normalized path length using $\boldsymbol{P}$ for two selections of ensemble sets (upper curves) and using $\boldsymbol{P}^{\dagger}$ (lower curve). Each point on a curve is an average of 50 simulations.

In Fig. 11, the reference simulation is shown for a collective with the Π^* in Eq. 3 using ***P***. Two of the 50 sets of simulations that make up Fig. 11 are from Fig. 4. Also shown in Fig. 11 is another ensemble set of 50 simulations and with a different members in the ensembles, which are taken from the identical population of 100 individuals. We see the variation of the results is still high at low numbers of individuals but low at higher numbers of individuals. Also shown in Fig. 11 are the simulation results using the established preferences $\boldsymbol{P}^{\dagger}$ (thick line) for the identical ensemble sets as the Reference simulation for ***P***. These ensemble-averaged simulations show a more definite trend than can be observed in the single realizations in Fig. 7. The diversity measures for these ensemble results are similar to Fig. 8 except with less variation; the limiting values for large collectives are unchanged.

Comparison of the Collective Solutions for Different *P*

How is the collective solution, using the same Application Rules, affected by ***P*** from different methods in the Learning Phase? In Table 4 is a summary of the three sets of simulations: the Learning Phase and the Individual and Collective Application Phases. The only difference between the simulation sets is that ***P*** is generated by four different methods, as listed. The number for the Individual Application Phase for each Learning Phase Method is an average over 100 simulations, one for each individual ***P***, as taken from Table 3. The number for the Collective Application Phase, for novice and established preferences, is from a single simulation using 100 individuals contributing to Π. For the collectives, the values of the experiential diversity are given also; while D_g^e broadly correlates to the performance of the collective relative to the individual, the correlation for the results using the Random Walk learning method does not follow the same trend as the other simulations. This suggests that the effect captured by the experiential diversity measure is not the only reason for higher performance of the collective. This is thought to occur because D_g^e measures only the presence of information, but not the importance or quality of information.

Table 4. Path Lengths for the Application Phase for different Learning Phases for a population of 100. All use the Application Rules. The quantities in the parentheses of the Application Phase are the standard deviation for the individuals and the experiential diversity for the collective.

Learning Phase Method	Learning Phase	Application Phase		
		Individual (s.d.)	Collective-***P*** (D_g^e)	Collective-$\boldsymbol{P}^{\dagger}$ (D_g^e)
Learning Rules	34.3	12.8 (3.1)	9 (0.38)	9 (0.60)
Random Walk	123.0	48.8 (54.9)	32 (0.31)	21 (0.46)
No-Backstep Random Walk	64.0	38.6 (40.3)	13 (0.32)	9 (0.42)
Non-Repeating Random Walk	50.8	33.7 (36.2)	10 (0.30)	10 (0.44)

In Figs. 12 and 13 are shown the average path lengths for the four learning methods for novice and established preferences respectively. What is most notable is that the Random Walk solution does not result in an improvement over the individual solution for the novice preferences and, in fact, shows no trend towards converging to a limiting solution for any number of members. The use of the established preferences in Fig. 13 significantly improves the performance of the Random Walk simulation. For very large numbers in the collective, as seen in Table 4, the collective solution using the unmodified random walk still does not converge to a minimum path. Based on this and later studies with increasingly complex mazes, we conclude that the average individual solution can only be improved to a limited degree by increasing the numbers in the collective. The individual performance based on the random walk solution is not sufficient to solve for the minimum path for the maze in Fig. 1. Hence, the collective advantage is limited: the more difficult the problem, the better the individual solution must be.

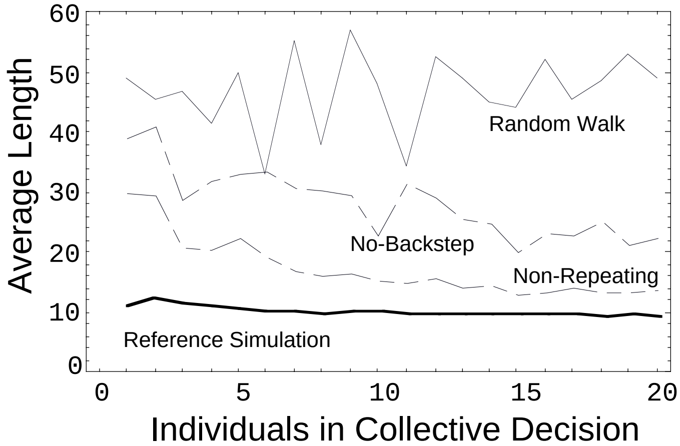


Fig. 12. Path lengths for the four Learning Phases in Table 4 using ***P***.

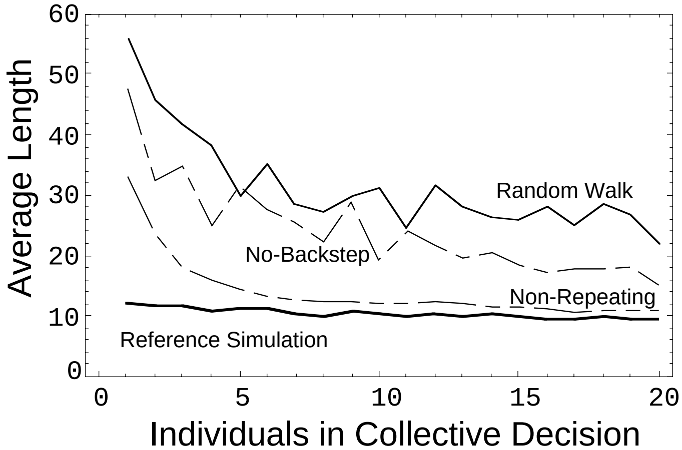


Fig. 13. Path lengths for the four Learning Phases in Table 4 using $\boldsymbol{P}^{\dagger}$.

Simulations of Loss and Noise in the Collective

For the rest of this section, results are presented that show the effect of noise and loss of information on the performance of the collective. Specifically, we wish to explore what failures in communication or even policies of exclusion might degrade the performance of a self-organizing system. Before proceeding, a few comments about the general nature of the simulation that follow can be made. In general, the collective solution is remarkably robust. Degradation of the individual's contribution, however implemented, generally had no effect or postponed the convergence to the minimal solution. In fact, the challenge quickly became trying to find ways to degrade the collective solution. Equally important, and possibly related, was that no alterations were found which generally *improved* the collective solution in comparison to the simple average over ***P***; the robustness and the optimal performance appear to be fundamental properties of the system.

The Effects of Random Noise on a Collective

Noise in this context, as defined earlier, is the random replacement of valid information in the individual’s contribution to the collective, thereby creating *false* information. Fig. 14 shows the effect of the addition of noise in the simulations with different frequencies: 0.0, 0.3, 0.7, and 0.9, where 0.0 represents the reference simulation. These results were insensitive to the magnitude of the noise (0.1), as long as it was less than unity.

Fig. 14 illustrates the remarkable insensitivity of the collective solution to the addition of noise at low frequencies. Even at higher frequencies, the tendency is only to delay the advantage of the collective to larger numbers of individuals. Although not shown, the collective solution degrades to the Non-Repeating Random Walk solution (Table 2) at frequencies of noise above 0.95, a verification of the robustness of the collective solution. Also note that an individual (one on the abscissa) is much more sensitive to noise than the collective. This occurs because noise adds false information that leads an individual to parts of the maze for which they have no experience from the Learning Phase. In unexplored regions, the Application Rules degenerate to a random walk approach and lead to significantly longer path lengths. For collectives, particularly large collectives, experience is available throughout the entire maze and, therefore, the collective cannot be misdirected by false information to unknown parts of the maze. This is a clear demonstration how diversity assists the collective decision.

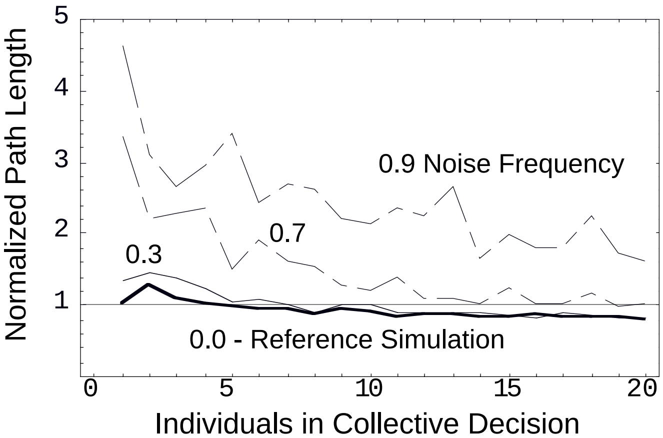


Fig. 14. The effect of the random replacement of the individual's Nodal Preference for different frequencies of replacement.

If the simulations in Fig. 14 are redone using $\boldsymbol{P}^{\dagger}$ instead of $\boldsymbol{P}$, the system becomes even more tolerant of noise. The normalized path length for large collectives is almost identical for noise levels of 0.95, 0.9, and 0.7 as those in Fig. 14 for 0.9, 0.7, and 0.3 respectively. The onset of a random walk solution is shifted to frequencies of 0.99 and above, from 0.95. This suggests that the information that is removed by using $\boldsymbol{P}^{\dagger}$ is more sensitive to noise than the main path preferences of the individual. The interpretation of this is that a collective of novice problem solvers is more sensitive to false information than the seasoned problem solvers, because noise can move the solution to paths that are unproductive. Interestingly small novice collectives do better than small experienced collectives. Until the experienced collective gain more information throughout the domain (the difference of D_g^n in Fig. 8 for small collectives), they are more sensitive to noise. The social analog of this observation is that a seasoned problem solver in a limited problem domain can easily get mislead by false information, where a novice problem solver recovers better from false information because they have experience on how to return to a known path after being lead astray. This suggests that a diversity of novice and established problem solvers would be an optimal collective in the presence of misinformation.

The Effects of Loss on a Collective

As defined earlier, loss is the selective *reduction* of information contributed by the individual to the collective decision. Unlike noise, there are almost unlimited possibilities for loss in the present context. The few examples presented below are one of many that were examined and were chosen either because they have a significant degrading effect or because they have a relevant or interesting interpretation in the context of a decision by a collective.

One less interesting loss mechanism is the effect of α in Eq. 6 with $\beta = 1$ and $\tau = 0$. The resulting collective solutions are identical to the reference simulations (Fig. 11) for any value of α greater than zero. This insensitivity is easily understood by the observation that because the Application Rules contain no parameters that would set a scale, the results must be insensitive to a uniformly applied multiplicative constant. This observation eliminates a whole class of loss mechanisms where the magnitude of the individual preference is reduced for all contributors, either selectively at certain nodes or uniformly across the maze.

Reduction or loss of the extremes of an individual's contribution

What happens if the nodal preferences of each individual are *flattened* so that the difference between the maximum and minimum values is reduced ($\beta < 1$ with $\alpha = 1$ and $\tau = 0$ in Eq. 6)? Fig. 15 shows the results with β=0.1. Except for collectives with two individuals, this loss of information has little effect on the collective solution. The effect of flattening the preferences might be interpreted as reducing the extremes of opinions.

The effect on the couple is subtle but relevant to later results. It was generally found that couples have a lower performance than individuals or larger collectives. The source of the degradation appears to be that with only two contributors, particularly with high performers, there often is direct opposition of a preferred path. As a consequence, the collective solution is degraded by the indeterminacy between two preferred paths. With the addition of more individuals, particularly of different levels of performance, the conflict is moderated and the collective functions better. Here, the same effect is achieved for the couple by the flattening of each individual's preferences.

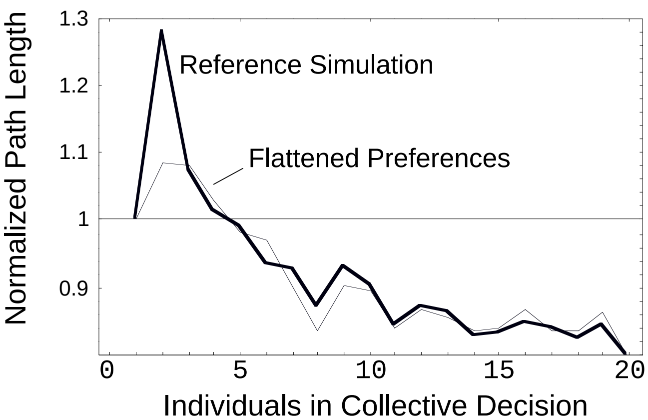


Fig. 15. The effect of reducing the extremes of the preferences of an individual. Note that the same individuals contribute each point in both curves, only their contributions to the Π differ.

Loss of minor preferences or opinions

What happens if only the maximum preferences of an individual at a node are contributed to Π (β=0.0001, τ= 0.9999)? Fig. 16 shows the effect is similar to the one discussed in the last subsection for couples: by emphasizing the dominant preferences, greater variation in performance is introduced and generally the collective has a lower performance. Interestingly, the couple and three person solutions (second and third point) is unchanged; this signifies that the Application Rules are always selecting the maximum preference. For these small groups, this reinforces the comments at the end of the last subsection.

The use of $\boldsymbol{P}^{\dagger}$ in the simulations comprising Fig. 15 and 16 results in no change in the curves for either loss mechanism. This is because the use of $\boldsymbol{P}^{\dagger}$ removes the minor preferences at the node along the path in the individual Application Phase, in addition to removing the preferences for paths not used. This indicates that the preferences in the extraneous paths are destabilizing in the collective solution when not balanced by the minor preferences. This suggests that a collective performs better and is more stable under two conditions: 1) when only the dominant path preferences are communicated to the collective, and 2) if secondary path preferences *are* included, then these need to include the full range of preferences. As the collectives become larger, these degrading effects diminish and the collective becomes stable independent of representation of the minor preferences.

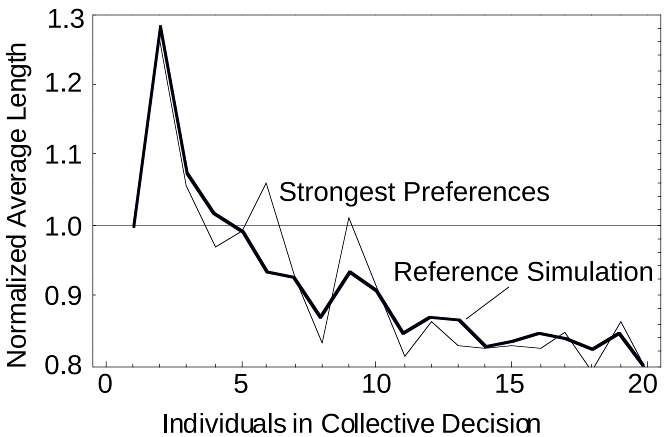


Fig. 16. The effect of contributing only the strongest preferences to the collective.

Complete loss of extremes of the individual

What happens if the individual's contributions at each node are made isotropic ($\beta = 0$), i.e., all nodes (not edges) that have been visited in the Learning Phase have preferences of equal importance? This has the opposite effect as the loss of minor preferences in the last example and is the limit of the loss of extremes in the first example. The social analogy might be an individual that cannot distinguish between important or unimportant contributions.

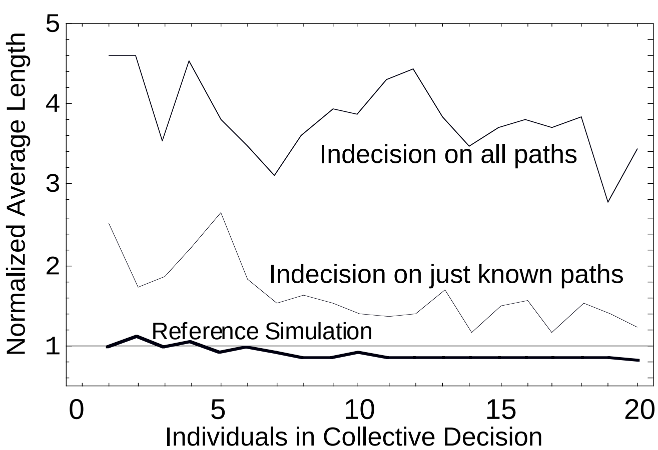


Fig. 17. The effect of an individual with an inability to distinguish between major and minor preferences.

As seen in Fig. 17, the loss of extremes in individual preferences is extremely disruptive to the collective; in fact, no improvement is observed for larger collectives. Had all the nodal preferences been made equal, the Non-Repeating Random Walk would be recovered. But because nodes that have not been visited are unchanged, some information is retained from the Learning Phase, sufficient to give about a 20 percent improvement over the Non-Repeating Random Walk result (Table 1). Also in Fig. 17 are shown the results when just the non-zero preferences are replaced with unity; this simulation limits the collective to portions of the maze for which some prior experience exists. While not as disruptive as the option with $\beta = 0$, this also emphasizes the importance of differentiation of an individual's preferences. When the simulations that comprise Fig. 17 are redone using the established preferences, $\boldsymbol{P}^{\dagger}$, the collective advantage is slightly degraded and has less variation with additional members in the collective.

This concludes the study of modifying uniformly all individual preferences. The general observation can be made that any modification of this type has either no effect or is detrimental. This is contrary to a reasonable expectation that some type of filtering would enhance the collective advantage. As will be further reinforced in the next sections, the collective advantage appears to rely on an unfiltered diversity of experience. The implications of these observations are discussed in detail later.

Random selection of a leader

Another major degradation of the collective's performance was the random selection $\boldsymbol{P}$ of one of the individuals in a collective, with a different individual selected at each node (see Fig. 18). The degradation is severe, resulting in a solution much worse than the average individual and showing no tendency to improve with larger numbers in the collective. This implementation is probably the worst case of leadership in a collective decision, but illustrates how the change of a dominant individual during a sequential solution process can yield results much worse than that of an average individual. Although not examined, a better alternative would be to select one individual for many steps in the maze. But this approach, at best, would only duplicate the performance of an individual and not show a collective advantage.

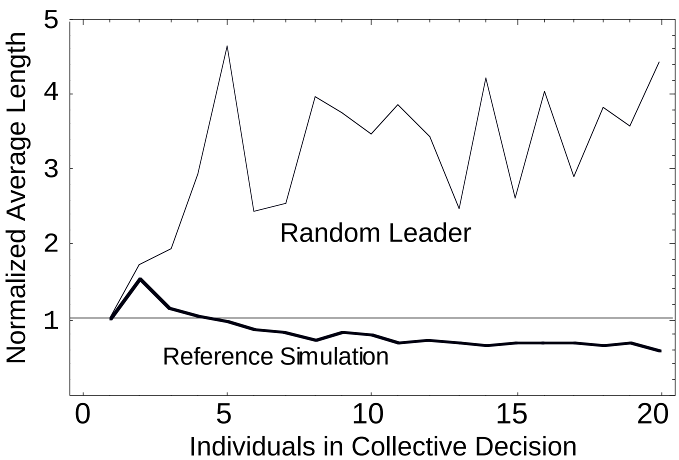


Fig. 18. The effect of randomly selecting a "leader" at each node from the collective.

Limiting Participation by Performance

The following study is the most interesting and initially the most counter-intuitive. Because of the availability of the results from the Application Phase, the effect of eliminating contributors to the collective with different levels of performance can be examined. Unlike the prior loss effects, which were applied on a node-by-node basis, this type of loss is applied to the sample population as a whole. Here the effect of limiting the distribution of individual performance on the collective is examined. In Fig. 19 three sets of simulation results are compared to the reference simulation using $\boldsymbol{P}^{\dagger}$. The results using $\boldsymbol{P}$ are similar except that the collective advantage is reduced. The ensemble set for each of the sets of simulations is comprised of individuals that are randomly selected from the set of individuals with the identical performance in the Application Phase. While they are identical in the individual path length, they have different values of D_g^e of 0.51, 0.53 and 0.53 for the individuals with path lengths of 11, 13 and 15, compared to the reference simulation with random selection of 0.60. Because these values are not near zero, this indicates that even individuals with paths of equal length take different routes through the maze.

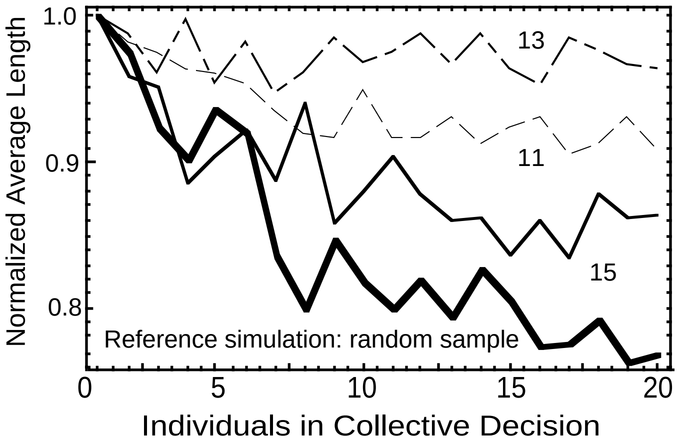


Fig. 19. The effect of homogenous populations of individuals comprising the collective.

Fig. 19 illustrates that narrowing the distribution of performance reduces the collective advantage. Surprisingly the ensemble sets of individuals with a path length of 13, near the average of the sample population as a whole (see Table 2), show almost no collective advantage. One might argue that this degradation is due to the loss of the higher performers with shorter path lengths. And, indeed, a collective with path lengths of 13 *and less* is found to perform better than the homogenous ensemble with a path length of 13. But this observation does not explain the better relative performance of the homogenous ensemble with a path length of 15. Additional simulations that are not reported here suggest that collectives with path lengths around 13 are missing path structures that do not lead to a collective advantage. But collectives made up of individuals with path lengths longer than this critical value, even though starting at a greater disadvantage in individual performance, show a better collective advantage. Because the experiential diversity measure is similar for each of these three groups, more than experiential diversity appears to be important.

What this demonstration suggests is that the self-organizing dynamics in this model system are not simply a linear superposition of information from the individuals. The collective advantage appears to be a complex interaction that requires diversity of performance, when the experiential diversity is relatively constant. Hong and Page (1998) observed that diversity of both individual experience and capability in a group led to better collective performance. But direct comparison of results cannon be made, because they did not report measures of experiential diversity in their study of individuals with identical capability. Unlike the results of Hong and Page, the collective solution is observed to do as good or better when only the “best” performers are included in the collective, assuming no other loss or noise mechanisms. This result is qualified, because the addition of either more complexity to the problem domain (such a noise in the formation of the collective or a hierarchical maze) or consideration of individuals with different capability can lead to similar conclusions. These are separately documented in forthcoming work.

The Effect of Conflicting Goals of Individuals

One of the observed properties of distributed, self-organizing systems is the ability to find solutions in the presence of conflicting information. To model this type of problem in the present system, 33 individuals each were trained in the Learning Phase on three different end nodes in the maze (In Fig. 1, nodes 21, 27 and 35 as counted from the start node going right, then across again going to the right). We then have the population as a whole find each of these end nodes and compare the performance of the collective relative to the average performance of the individual.

Fig. 20 shows the performance of the collective for the three different goals. Clearly the collective does much better than the average individual, but the results are initially less impressive than they appear. The average individual performances for the Application Phase with end nodes {21, 27, 35} is {67, 17, 35}, compared to the minimum path length {9, 7, 9}. The individuals do poorly in the Application Phase compared to the prior simulations. In order to find a node outside their experience from the Learning Phase, the Application Rules degenerate to an optimized Random Walk solution to discover the end node. The collective path length with 100 members is {10, 14, 12} respectively, all longer than the minimum path lengths. This performance, which is worse than was previously observed for the unmodified collective solution, is because the simple method for generating a collective decision assumes that the population is trained on the same goal as will be used in the Application Phase. In the present example, the training on a goal different than the end node leads to conflicting information about the optimal path in the maze. It is not understood how the collective resolves this conflict and achieves a reasonable solution, far better than the average individual's performance. It is believed that the experience of individuals with different goals still contains information useful to the collective, even though the result of a quite different goal. Said another way, while the goals for learning may differ, the connectivity on the problem domain is common. More studies of this type are needed to better understand the dynamics of this system.

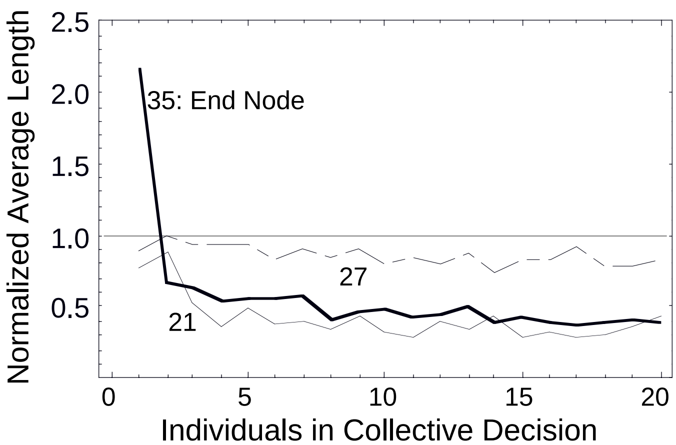


Fig. 20. The performance of a collective solution in the presence of conflicting information.

<u>Study of Maze Complexity with Fixed Individual Ability</u>

All of the above studies examined different individual or collective capabilities while keeping the problem domain fixed. In this subsection, exploitative results are presented for alternative mazes while using the Learning and Application Rules with Π^* in Eq. 3 for a population of 500 individuals.

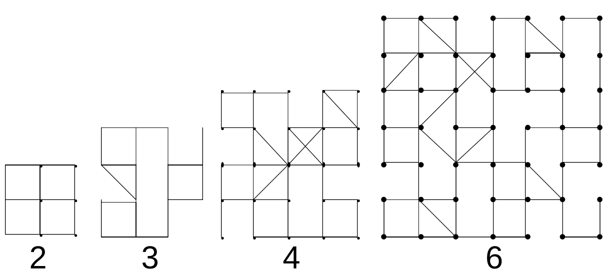


Fig. 21. Four randomly generated mazes. The start and end points are the lower left and upper right in each maze, respectively.

Because the performance of the average individual and the collective were comparable in the reference simulations, a question arises in the prior studies. How does the collective advantage depend on the individual performance? This question is addressed by simulations on increasingly larger mazes. To simplify the study, coding was developed that randomly generates mazes of arbitrary size. Fig. 21 shows four mazes of different sizes used in the study. Table 5 shows the summary of simulations. Fig. 22 shows how performance of the collective solution changes with the different mazes. The simulation for maze number 6 was extended to 2000 members in the collective, and the collective solution asymptotically approaches 12 with 2000 members, with no further improvement expected. Even the collective solution based on $\boldsymbol{P}^{\dagger}$ does not converge to the minimum solution for large collectives.

Table 5. Path lengths for mazes of different sizes (From Figs. 1 and 18) for a population of 500 individuals using the Learning and Application Rules.

Maze:	#2	#3	#4	Fig. 1	#6
Minimum path	4	6	6	9	10
Ave. Learning Phase	8	33	34	37	64
Ave. Application Phase	4	7	10	13	16
Collective	4	6	6	9	13

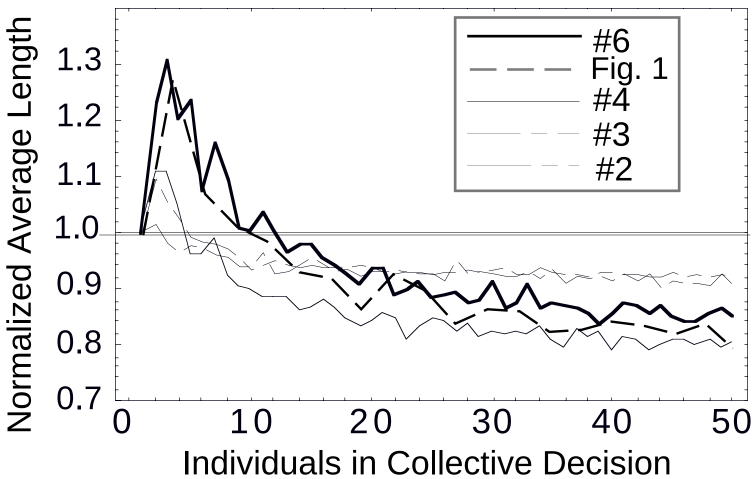


Fig. 22. The normalized path lengths for different mazes using the same Learning and Applications rules with $\prod^*$.

The following conclusions can be drawn from the above study and from the study of different Learning Rules (Table 4) where the global problem difficulty was held constant while the individual capability was varied. 1) A simple maze to a good individual solver is a trivial problem, and no improvement is obtained by a collective solution. 2) The *rate* of improvement of the collective declines as the maze size is increased; larger numbers of individuals are needed to collectively solve harder problems. 3) More difficult global problems require better problem solvers. 4) An extremely difficult problem to an individual with limited capability leads to a random individual solution that shows no collective advantage. The last conclusion is significant; it suggests that harder and harder problems cannot be solved larger and larger collectives of individuals with constant capability. This conclusion suggests that a hierarchical approach or a domain decomposition to a more difficult global problem would allow collectives to solve more difficult problems (Heylighen, 1998).

Discussion and Implications

In the following discussion, an examination of the assumptions is used to tie the present abstracted study in the last section to more realistic applications. The argument for emergent problem solving is made. And, then, the relevance of the work to collective problem solving is addressed. Diversity is discussed separately due to its importance.

Assumptions of the Model

Path independent states

The most significant assumption made in the model is that the state of an individual at a node is not dependent on the path that brought them to that node (except for the caveat below). This assumption is justified in some types of sequential problems, but not all. For example, piling blocks of different sizes to reach a desired height does not depend on the assembly path, but only on the current structure itself. Hong and Page (1998) make the same assumption (although not explicitly stated) and give additional economic examples of this type. Generally, this assumption is violated in a sequential problem that involves complex information. We tend to gain knowledge from excursions and incorporate this knowledge in our current decisions. If we find ourselves back at the same decision point, we likely are not making the decision with the identical information as the last time we were at this point. How much this assumption affects the applicability of the current model is unknown, but a reasonable belief is that the relaxation of this assumption would not change the general conclusions.

The above statements must be qualified somewhat. The model *does* use the prior position in the current decision at a node, by avoiding the choice of the last node occupied. This is a dependence on the most recent path history. The study done of the various random walk approaches suggests that this history information is critical in the self-organization process, as seen by the improvement from the Random Walk to the no-backstep and the no-repeat option. It may be that the current model does capture the most important path dependence, that of the prior node occupied.

Sequential Learning and Application Phases

Another significant assumption is the separation of the two phases in the simulation. The consequence is that an individual does not continue to learn in the Application Phase. This assumption can be appropriate for an individual solving a problem, but probably not for a collective of interacting individuals. The complication of the collective question is treated separately in the next subsection. This assumption is partially relaxed by the option of a novice to become an established problem solver by the elimination of unnecessary information. While this type of learning does not affect the individual's performance, it does influence the collective solution.

Many individuals, solving the same problem

In general, a successful self-organizing system is believed to require many diverse individuals solving problems in a common domain. The collective path formation after a new building given in the introduction is a good example of this belief. If either numbers, diversity or common problem were omitted, the system would not exhibit the same desirable global solution or stability of the solution. In the current simulations, each of these variables was examined. There is a definite need to have many individuals participating in the collective. This was observed both in the need for large collectives in order to achieve a collective advantage and for the need for an ensemble average in order to reduce the inherent noise. The need for individual diversity, other than goals, is discussed in detail below. Here we focus on the assumption of the common problem domain and a common goal.

In almost all of the examples studied, the individuals gained experience based on an identical goal on an identical problem domain. In the study of conflicting goals, the self-organizing collective was observed to make use of learned information from individuals with quite different goals. How this process occurred is not understood; the belief is that the individuals solving for different goals on a common problem domain contribute useful information to the collective. This agrees with our intuition that learning from one problem situation can be useful in another problem if there are commonalities in the solution. But, our intuition also suggests that if the individual experience is too different from what is needed, a collective may be destabilized, as was observed in one study (Huberman & Glance, 1993). While these questions are not resolved, the present system shows the potential of addressing these issues.

Non-interacting individuals

While the relaxation of the assumption that individuals are independent from each other would address many of the potential criticisms of the model, this assumption is essential for the argument of the observance of emergent problem solving in the absence of interaction between the individuals. If the individuals interact with each other during the formation of a collective solution, one could have argued that the improvement in problem solving resulted from the cooperation between the individuals - a property of an individual. By eliminating this possibility, the conclusion of self-organizing problem solving is much stronger. This is similar to the argument made for cooperation from pure aggressive agents by Hemelrijk (1997). The assumption of the independence of the individuals was also made by Hong and Page (1998).

It is believed that by the inclusion of interaction of individuals, a collective advantage would be observed, and it would occur in smaller groups, as early coherence in the collective is reinforced. This is similar to the comment in the introduction that a newly formed path after a building construction forms faster with positive reinforcement, but would still be observed if it were not present.

Individuals with identical capability

The assumption that individuals all use the same rules results in individuals which have the same capability to solve a problem. This assumption was made to simplify the analysis. Different capabilities were presented for the Learning Phase based on the random walk approaches. Because these also gave a self-organizing collective advantage (with the exception of the pure random walk), we conclude that the results do not depend on the details chosen for the capabilities of the individual. Even though the individuals have identical capabilities, the results show that significantly different performance can result from just random choices. This is a reminder that some of the diversity of performance that we observe in our society is not always related to differences in capability, but to less-controllable random circumstances. Hong and Page (1998) examined collectives with different capabilities and observed similar results to their study of populations with the same capability, but different experiences. We expect the same to be true here.

Individuals with identical assessments of value

In the construction of the collective solution, the information of each individual in the reference simulation is treated equally. In other words, the information that an individual contributes is assumed to be understandable and valued the same by the collective and all other individuals. Often when an individual contributes to a collective decision, the value they place on their preferences differs from the perspective of the collective. Similarly the specific information that they contribute is assumed to be compatible with the information from other individuals. These are major assumptions and probably are not generally true. Because the violation of these assumptions is likely associated with noise and loss in communication, the studies related to these effects, therefore, evaluated the sensitivity of the collective advantage and stability to these assumptions. In particular the insensitivity of the collective to noise indicates that significant disagreement can occur in the relative value and content of information. Hence, we conclude that while this assumption is rather poor, the system is not sensitive to violations of it.

Discrete problem domain with fixed connections

The relevant problem domain is assumed to be a graph with fixed connectivity. This assumption has many implications. One is that the problem space has distinct locations where individuals and collectives make decisions, rather than a continuous space where there is a gradation of possible decision locations. This assumption is appropriate for some types of problem domains, such as the Internet or distributed knowledge systems in general, where links between information are discrete. This assumption is not appropriate for problems that are inherently continuous, such as the ant metaphor given earlier. It is believed, but not yet tested, that the general conclusions of the present study can be duplicated on a continuous problem domain. This belief is supported by the comparison done on a lattice (Hemelrijk, 1997) and continuous domain (Hemelrijk, 1998) for the same agent problem. Both types of domains demonstrated the same emergent cooperative behavior, although the continuous solution did capture a clearer social-spatial structure.

The other assumption related to the problem domain is that the connectivity is fixed during the entire problem. This implies that there is no apparent feedback between the activity of the individuals and the connections between the states in the problem domain. One could argue that the connectivity should change according to the knowledge gained about the problem domain (Heylighen, 1998). While the current model does not add or remove connections in the domain, the role of the path preferences accomplished the same intent by having zero and non-zero values. Therefore, the connectivity in the present model should be viewed as a *potential* connectivity in which the potential is realized, or not, through the values of $\boldsymbol{P}$.

Emergent Problem Solving

One of the main goals of the present work is to demonstrate the possibility of emergent problem solving. Admittedly this is a difficult concept, comparable to prior studies of self-organizing systems that exhibit emergent properties. Often agreement on what exactly is an emergent property is hotly debated until its final acceptance, largely due to the disagreement of a universal definition of an emergent property. As argued in the introduction, emergent problem solving occurs when a system solves a problem in which the solution cannot be observed or argued for at the level of the individual. Even more challenging to demonstrate is emergent problem formulation, as is true in the current study, where the formulation of the problem is not even be expressed at the level of the individual but is expressed globally.

The approach that was taken in the present study was to create a system which has a well posed problem to solve, in this case, finding a path between two points. Then, an agent-based system solution method is implemented with three essential restrictions: 1) rules of the agents do not use global information, 2) the agents do not include logic for finding a shorter path and 3) the agents do not interact or cooperate. All three together assure that any global functionality cannot occur from the properties of the individual. When the performance of the system is examined, the collective is not only observed to find a short path, but large collectives on a moderately complex maze locate a minimum path. Therefore, we conclude that finding a minimum path is an example of an emergent problem solving. Furthermore, since the finding of a shorter path is not even formulated at the level of the individual, the definition of the problem is also emergent. This argument is fully parallel to prior definitions of emergent properties: of the cooperation at a global level when the individuals are only competitive (Hemelrijk, 1997) or the regulation and control of an economy when the participants are only competitive (Kochugovindan & Vriend, 1998).

We contrast the above arguments to the work done by Hong and Page (1997). In their formulation they explicitly guide their individual agents to find the best solution possible by iterating until no improvement is found. Hence, the fact that they observed a good or even optimal solution is not emergent property. (It should be noted that this was not the purpose of their study.)

To find emergent problem solving in the current system is not just sophistry, but carries with it implications that are significant. In one of the above studies, the population was divided into homogenous “performance” levels based on their path length in the Individual Application Phase. Because the path length is an emergent property, performance is also an emergent property. This implies that on the level of the individual and collective, it is not possible to assess an unambiguous value of the “performance.” Hence the study can only be done from an omniscient perspective. This has significant implications when self-organizing problem solving is active within an organization. Because the process of solving a problem is outside of the understanding of the individual or group, it is not possible to assess the relative importance of contributions of each individual. This agrees with our understanding that if we don’t know the problem to be solved, then we don’t know who is best to solve it.

Relevance to Collective Problem Solving

The emergent problem formation and problem solving as an abstract expression are easily demonstrated by the results in Figs. 9 and 11, which show the collective outperforming the average individual. What is much more difficult to show is that the problem solving expressed by the present system has any relevance to collective problem solving in social systems. Currently, this can only be demonstrated by a richness of results that appeal to our understanding of actual collective solutions. In the introduction is a series of questions about the behavior self-organizing collective problem solving. As a summary, the insights into these questions as provided by the current study are given.

• What are the dynamics and from what processes do they originate? What are the emergent properties sensitive to? How do the details of the individual actions affect the collective performance?

The dynamics originate from the presence of multiple optimal solutions in the domain and the diversity of individuals in the collective, coupled with the selection of a maximum preference. The system was found to exhibit chaotic dynamics at the level of the individual, but stable solutions at the level of the emergent property: the shortest path. These results appeal to our sense of collective decision making. For example, if it were possible to repeat a problem solving session with a group of people starting from the same initial individual states, we would likely see distinct variations on the specific path taken by the collective to the goal, but the same goal would likely be reached. Furthermore, by the addition of one individual in a small collective, we might observe significantly different dynamics and possibly outcomes. The current model captures these identical features. The collective advantage was observed not to be sensitive to significant amounts of noise in the contribution of the individuals. While this may initially seem contrary to our experience of collective decision making, significant miscommunication and different value judgements do occur, yet collective problem solving is routinely successful.

• What are the mechanisms that lead to a collective performing better than the individual?

The one identifiable mechanism for the collective advantage is given in Fig. 9. What is observed is that a better solution is found in a collective as a result of the "filling in" of a shorter path by missing information from another individual. In an ideal group situation, the more individuals contributing, the more likely critical missing pieces of information will lead to an optimal path. In non-ideal situations, the following can occur to limit the collective advantage or reduce stability: 1) when individuals are limited to only their dominant opinions or when they can only express their opinions uniformly, 2) when a random individual dominates the contribution to the collective at different points in the sequential decision process, or 3) when collectives are comprised of homogeneous performers. Arguably, each of these occurs to some degree in the traditional processes of decision making and are captured by the present model.

• What are the capabilities of the individuals that result in the observed dynamics and emergent properties?

Of primary importance is that the individual must be decisive and a reasonable problem solver. If either the individual's solution is random in nature or the individual indecisive, then the collective fails to be better than the average individual. The need for diversity in the population as a whole is treated separately, because of its significance.

Diversity: An Essential Attribute of Collective Self-Organizing Systems

Diversity in the present work is shown to provide two essential advantages to collective problem solving: better collective performance and more robust solutions. Furthermore, the study distinguishes between the different types of diversity (experiential, preferential and performance) and their relative importance.

One advantage of experiential diversity is clearly captured in the lack of sensitivity of the collective advantage to noise (false information). Because false information can lead to less explored regions of the problem domain, experiential diversity provides the collective with contingencies that are not available to the individual or narrowly focused group (these results were not shown but are easily demonstrated by adding noise to the simulations in Fig. 19). In general, the experiential diversity measure (Eq. 13) was found to correlate with the collective advantage in most examples. We conclude that the experiential diversity is important if not essential, and the system is sensitive, but not excessively so, to the magnitudes of the experiential diversity. The work of Hong & Page (1998) obtained similar conclusions about the importance of experiential diversity, but made no observations about the magnitudes of the experiential diversity. In one example, a difference in the collective advantage was observed in populations with similar experiential diversity when the effect of limiting the contribution by performance was examined. The collective performed better when individuals are included with a distribution of performance, rather than a subset. This suggests that experiential diversity by itself does not capture the full explanation of the collective advantage.

The simulations also provided insight into the effects of preferential diversity and the importance of the relative magnitude of the preferential diversity. These results are unexpected and may not be generally appreciated. The studies on the replacement of nodal preferences with random noise (which alters the preferential diversity) and the studies on the elimination of minor preferences of an individual suggest that the collective solution is not degraded if the preferential diversity is reduced. Only the stability of the solution and, to a lessor degree, the convergence to final solution was affected by the change. The effect of the relative importance of the magnitude of the preferential diversity was examined in detail. Significant degradation was found if all non-zero nodal contributions are made equal. Taken together, the importance of preferential diversity is minimal, but the distribution within this diversity is essential. This is exactly the opposite importance of the experiential diversity. In either case, including the additional information improves the stability or performance of the collective solution.

What are the implications of these results to organizations that include a self-organizing approach to problem solving? Experiential diversity, while essential within the organization, is alone not sufficient. An organization must also include channels by which individuals or groups can express the magnitude of their preferential diversity, particularly avoiding uniform expression of preferences or random expression of one individual or group with the exclusion of others. In addition, there is some justification for an organization valuing a diversity of performance in individuals.

For completeness, we note that alternative diversity studies do show different results. Huberman & Glance (1993) using a free rider problem as a repeated n-person prisoners' dilemma concluded that diversity leads to coalition instability and system failure. There are two possibilities why they observe different results than in the current study. The first was discussed earlier in this section in that identical or similar goals are assumed in this study and in the work of Hong and Page. This allows for a commonality in contributions by diverse individuals. The other possibility, and maybe related to the first, is the major difference in the kind of system examined. The repeated prisoners’ dilemma specifically models whether or not the individual chooses to participate, where in the current study all individuals participate in a collective. There is significant understanding likely to be gained by examining the differences between these approaches to modeling a collective decision.

Conclusions

The main theme of this study is that there may exist an alternative approach to problem solving that operates at a level above our traditional problem solving processes. The system in the current study demonstrates that problem formulation and problem solving can occur at a level above a collection of idealized agents. And a rather simple system can capture a broad range of reasonable behavior of social interactions, even though interactions are *explicitly* omitted. If this capability can be extended to our organizations and society, what are the implications of this alternative approach?

This viewpoint hinges on the idea that human processes are driven by inherently random, but self-organizing dynamics (e.g., Allen, 1994), much like biological evolution. The application of these concepts to the development of future policies or technical developments, or just to our interpretation of history, has been limited by understanding of the basic processes and their consequences. A stochastic theory of self-organizing social systems is being developed (Weidlich, 1971; Weidlich, 1983; Weidlich & Braun, 1992) which may fundamentally answer some of these questions. The simulations in current study can be used to test many of the assumptions in this approach, similar in utility to the use of simulations in the development of kinetic theory in physics (Wu, 1979). An increased understanding of these chaotic but self-organizing processes will enable us to replace what has been an act of faith in the potential of these systems to solve problems with a deep understanding of the advantages and risks.

We have argued the advantages of a self-organizing system as offering problem solving greater than the individual. But, what about the risks posed by this approach to problem solving? For example, a critical concern is how can we tolerate the inherent randomness of these systems. In the current simulations, it was necessary to take averages of many realizations in order to make meaningful comparisons. In real life, this multiple "running" of a system is likely not an option. Can we tolerate the randomness of the solution when a failure of the system may have severe consequences? As in the simulation of the random leader (Fig. 18), would it not be better to have a "poor" leader, but one that is more predictable? The partial answer to this question is the redundancy of the system, such that many realizations are being tested at one time and the emergent property is not random but a stable, well-defined solution.

The ultimate answer to many of the concerns raised about the inappropriateness of a self-organizing approach to social or organizational issues is that self-organizing systems occur because often there are no alternatives (Johnson et al., 1998). Self-organizing systems occur in nature and social systems when the global system is too complex or the centralized problem solver is lacking in capability or control over the system. Because our world is quickly becoming more complex, often changing faster than we can evaluate the changes, let alone respond to them, the process of collective self-organization may be the only option. With greater understanding, we will be able to develop capabilities and policies that will enhance functioning of these system to the benefit of all.

Acknowledgments

The author wishes to thank the members of the Symbiotic Intelligence Project, particularly Steen Rasmussen; Shareen Joshi, Greg Madey and Veronique Longmire for their helpful comments; and Chrissa Constantine for her editorial assistance. The author is grateful to the anonymous referees who made insightful comments and suggested additional sources. This research is supported by the Department of Energy under contract W-7405-ENG-36.

References

Adami, C., Belew, R. K., Kitano, H., & Taylor, C. E. (Eds.). (1998). Artificial Life VI. Cambridge, Mass.: MIT Press.

Dorrigo, M., & Gambardella, L. M. (1997). Ant Colonies for the Traveling Salesman Problem. Accepted for publication in BioSystems, 1997.

Drucker, P. (1994). The Age of Discontinuity. New Brunswick, NJ: Transaction Publishers (originally published in 1969).

Fisher, R. A. (1930). The Genetic Theory of Natural Selection.Oxford: Oxford Univ. Press.

Forrest, S. E. (1990). Emergent Computation. Cambridge, MA: MIT Press (Special issue of Physica D).

Hemelrijk, C. K. (1997). Cooperation without Genes, Games or Cognition. In P. H. a. I. Harvey (Ed.), Fourth European Conference on Artificial Life, (pp. 511-520). Cambridge: MIT Press.

Hemelrijk, C. K. (1998). Spatial centrality of dominants without positional preference. In C. Adami, R. K. Belew, H. Kitano, & C. E. Taylor (Eds.), Artificial Life VI, (pp. 307-315). Cambridge: MIT Press.

Heylighen, F. (1998). Collective Intelligence and its Implications of the Web: algorithms to develop a collecive mental map : Free University of Belgium, mineo.

Hofabauer, J., & Sigmund, K. (1988). The Theory of Evolution and Dynamical Systems. Cambridge: Cambridge University Press.

Hong, L., & Page, S. E. (1997). Problem Solving by Heterogeneous Agents. mimeo .

Hong, L., & Page, S. E. (1998). Diveristy and Optimality. mimeo .

Huberman, B. A., & Glance, N. S. (1993). Diversity and Collective Action. In H. H. a. A. Mikhailov (Ed.), Interdisciplinary Approaches to Non-Linear Systems. Berlin: Springer.

Johnson, N., Rasmussen, S., Joslyn, C., Rocha, L., Smith, S., & Kantor, M. (1998). Symbiotic Intelligence: Self-organizing knowledge on distributed networks driven by human interactions. In C. Adami, R. K. Belew, H. Kitano, & C. E. Taylor (Eds.), Artificial Life VI, . Cambridge, Mass.: MIT Press.

Kauffman, S. (1993). The Origins of Order: Self Organization and Selection in Evolution: Oxford University Press.

Kochugovindan, S., & Vriend, N. J. (1998). Is the Study of Complex Adaptive Systems Going to Solve the Mystery of Adam Smith's "Invisible Hand"? The Independent Review, 3(1), 53-66.

Maltz, D., & Ehrlich, K. (1995). Pointing the way: active collaborative filtering, (Vol. 1995): Association for Computing Machinery, Inc. (ACM).

Margulis, L., & Fester, R. (Eds.). (1991). Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis. Cambridge, MA: MIT Press.

Mitchell, M. (1996). An Introduction to Genetic Algorithms. Cambridge: MIT Press.

Moukas, A., & Maes, P. (1998). Amalthaea: An Evolving Multi-Agent Information Filtering and Discovery System for the WWW. Autonomous Agents and Multi-Agents Systems, 1, 59-88.

Prigogine, I. (1998). The End of Certainty: Time, Chaos, and the New Laws of Nature. N.Y.: The Free Press.

Smith, J. B. (1994). Collective Intelligence in Computer-Based Collaboration. New York: Erlabum.

Sober, E., & Wilson, D. S. (1998). Unto Others: The Evolution and Psychology of Unselfish Behavior: Harvard University Press.

Thomas, D. A., & Ely, R. J. (1996). Making Differences Matter: A new paradigm for managing diversity. *Harvard Business Review*(September–October).

Twidale, M. B., Nichols, D. M., & Paice, C. D. (1997). Browsing is a Collaborative Process. Information Processing & Management, 33(6), 761-783.

Weidlich, W. (1971). The Use of Statistical Models in Sociology. Collective Phenomena, 1, 125-130.

Weidlich, W. (1983). Concepts and Models of a Quantitative Sociology: The dynamics of interacting populations. Berlin: Springer.

Weidlich, W., & Braun, M. (1992). The master equation approach to nonlinear economics. Journal of Evolutionary Economics, 2, 233-265.

Wu, T. (1979). Kinetic Equations of Gases and Plasmas. Reading, MA: Addison-Wesley.