AN EVOLUTIONARY AGENT-BASED MODEL OF PROBLEMS AND SOLUTIONS

DAVID BROOKS
MITRE Corporation
dgbrooks@mitre.org

WENDY HAM*
The Wharton School, University of Pennsylvania
wham@wharton.upenn.edu

NATHAN HODAS
Department of Physics, California Institute of Technology
hodas@caltech.edu

BRIAN HOLLAR
Department of Economics, George Mason University
bhollar@gmu.edu

LILIANA SALVADOR
Department of Ecology and Evolutionary Biology, Princeton University
salvador@princeton.edu

**corresponding author*

***Abstract*—We describe an evolutionary agent-based model of problems and solutions, and we explore its potential applications in topics such as: the origins of disciplinary boundaries, the effects of collaboration in problem solving, organizational intelligence, and innovation strategies.**

I. BACKGROUND

Problem solving is often depicted as a cognitive process that begins with the identification of a problem and ends with a solution. While this view is useful for examining isolated, micro instances of problem solving (e.g., search, optimization), it is perhaps unsuitable for examining problem solving at the systems level. In reality, not only do problems look for solutions, solutions also look for problems.[1]

Organization scholars have long recognized the bidirectionality of problem solving. Most prominently, Cohen, March, and Olsen (1972) proposed the Garbage Can Model, which views problem solving and decision-making as the product of semi-accidental collisions involving four relatively independent streams of events: problems, solutions, participants, and choice opportunities.

[1] This topic has been discussed in detail in an unpublished paper (2009).

In the same spirit, we explore the possibility of creating a computational model of problems and solutions in which problems and solutions themselves are the initial units of analysis. As a first step, we introduce an agent-based model in which problems and solutions are the interacting agents. Thus a main difference between the Garbage Can Model and ours is that in our model participants and choice opportunities are allowed to endogenously emerge rather than exogenously given. In developing our model, we try to imagine the origins, emergence, and evolution of a problem-solution landscape: What would a world look like before problems and solutions get distributed into distinct domains? Where do disciplinary boundaries come from?

Our sense is that the distribution of problems and solutions is path-dependent and thereby suboptimal. In other words, the most fitting solution to a problem does not always reside in the same discipline as the problem. As such, an attempt to find matching pairs of problems and solutions must occasionally transcend disciplinary boundaries. (We are agnostic at this point with regards to the possibility of realizing an optimal distribution of problems and solutions.)

II. AGENT-BASED MODEL

A. Representation of problems and solutions.

Our foremost challenge was to find a simple, faithful, and generalizable representation of problems and solutions: What are the fundamental definitions and features of problems, solutions and their relationships to each other? Furthermore, how do related concepts, such as information and knowledge fit in? For example, when does information or knowledge become a problem or a solution? Upon considering these questions, we offer the following insights:

i. *Problems* and *solutions* are made up of *information*
ii. A problem is *an incomplete set of solutions*. (Note that a problem is not the *complete absence of solutions* because knowing that there is a problem is part of the solution set.)
iii. Combining i) and ii) above, all pieces of information can be treated as solutions—whether potential or realized—and these solutions form the basic components of our model.
iv. *Knowledge* is a higher-level concept and therefore is not incorporated as a building block in our model.

Based on these insights, we use NetLogo (Wilensky, 1999) to construct an agent-based model with the following features:

- Solutions are represented as bit strings whose length l_s can be adjusted within a certain range. Each bit represents an attribute of the solution and is randomly generated.

- Problems are represented as bit strings whose length l_p equals $n * l_s$, where n is a non-zero positive integer. In other words, as shown in Figure 1, a problem requires one or more solutions in addition to knowledge of the problem itself as part of the solution set, which is implicit in the problem representation.

- The match between a problem and a solution is determined by Hamming distance (Hamming, 1950) h, where $h = 0$ indicates a perfect match. Some level of matching imperfection *i* is allowed, ranging from 0 to 1, where $i = 0$ indicates a requirement for a perfect match.

- Both problems and solutions can move around in the space to try to find compatible matches. They have a 'radar' that is good within a certain focal radius *r*.

Figure 1: Representation of problems and solutions

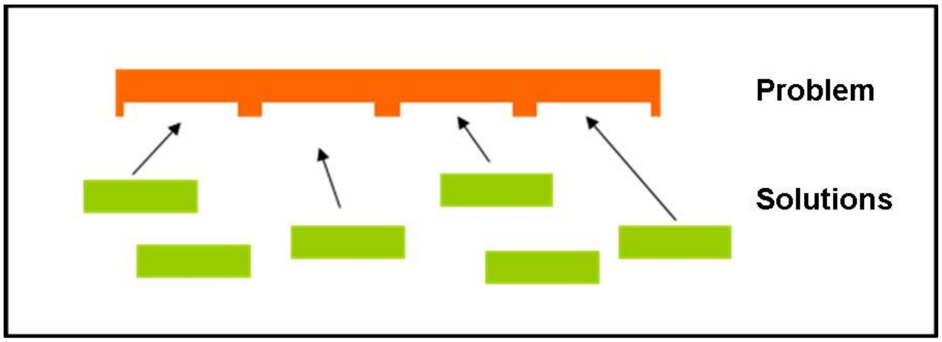
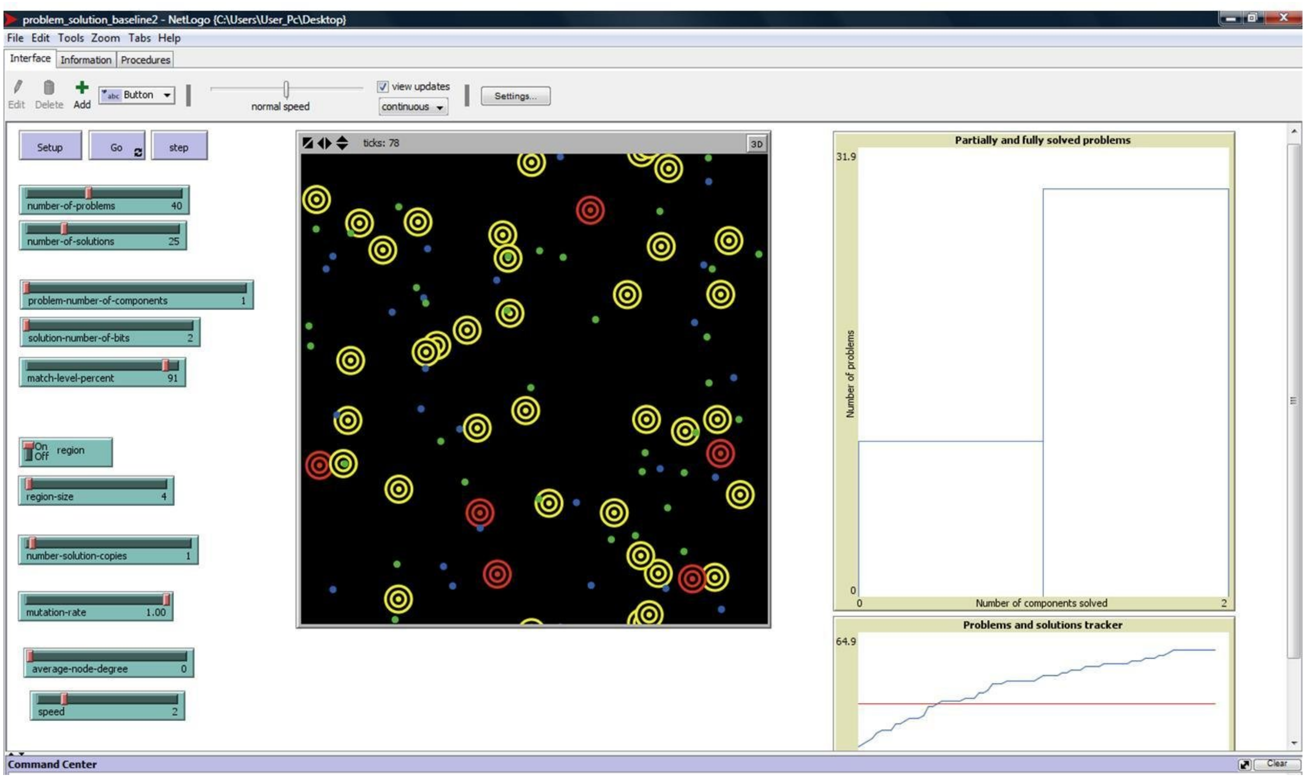


Figure 2: NetLogo screenshot
Red large circle = unsolved problems; yellow large circle = solved problems; blue dots = original solutions, green dots = solution replicas



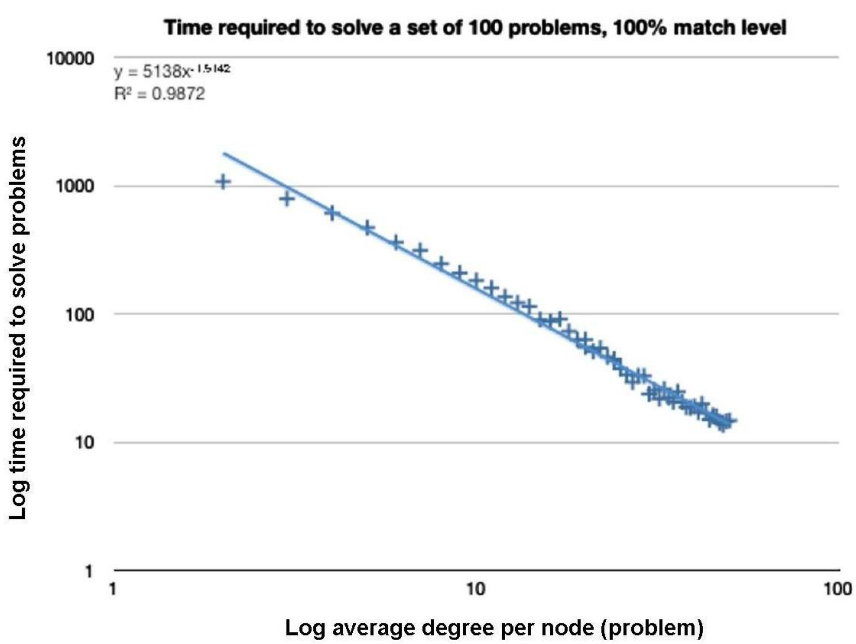
- Problem solving occurs when a problem and all of its solutions have found each other.
- When problem solving occurs, all of the solved problem's solutions replicate. This feature resembles reality in the sense that when a problem is successfully solved, its solution tends to get passed around and made popular.
- The replication process has a mutation probability p_m, ranging from 0 to 1. Additionally, when a mutation occurs, a number of random bits b_m are mutated (from 0 to 1 or 1 to 0). This feature reflects the idea that when solutions are passed around, there is usually some loss of fidelity.
- Some random links can be introduced to connect problems, such that when any one of them gets solved, it shares its solutions with all the other problems in its network. These random links is a mechanism for bypassing the spatial limitation of problems and solutions. This feature models accidental linkages among problems (and consequently their solutions), which may be due to, for example, collaboration activities.

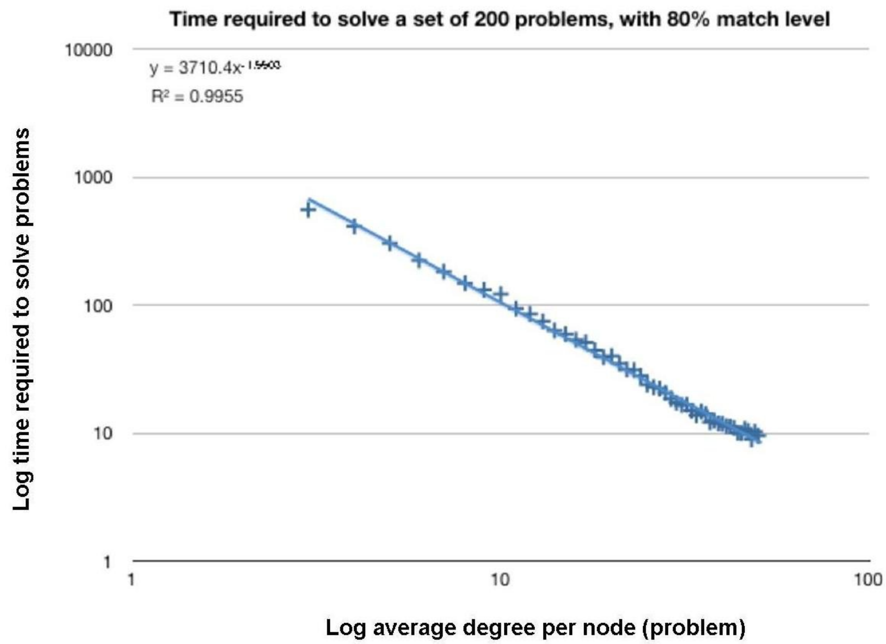
B. Simulations

We explored the above features and ran simulations in which we varied the number of random links. As shown in Figure 3, we found that the time required to solve problems has a power law relationship with problems' average degree. The exponent is ~1.5.

This result is interesting: While it is intuitive that collaboration would speed up problem solving, it is not obvious that the relationship should be a power law.

Figure 3: Log (time required to solve problems) vs. log (degrees)





III. EXTENSIONS

A. Model extensions

- Our simulation results above demand a closer analysis. Is something actually interesting about the power law relationship, or is it simply a network artifact?

- We have not thoroughly explored the features of the current model, so this is an obvious next step. It would be interesting to experiment with various parameters (in addition to the number of random links between problems as described above) to understand better the factors that influence the solvability of problems in a system.

- In addition to studying the effect of throwing in random links into the system, it would be interesting to set a condition according to which links are added between problems and solutions that have previously interacted in a successful problem solving.

- To study the antecedents of cognitive associations, heuristics, disciplinary boundaries, knowledge codification, and technology—among other things—we plan to add to the model a condition according to which if a set of solutions solve problems together often enough, they get bundled and subsequently replicated as a packet (rather than independently).

- At this point, mutations in the model are random, but it might be interesting to add a condition according to which imperfect pairs of problems and solutions mutate to become better matches to each other. In reality, people often modify problems, solutions, or both to make problem solving happen.

- As we understand better the fundamental principles of problem solving, an important next step is to give appropriate representations to distinct types of problems. For example, some problems are better tackled using exhaustive search, whereas others are better tackled using analytical approaches (e.g., the Königsberg bridge problem).

- Similarly, it is important to give appropriate representations to distinct types of solutions. For example, borrowing the theory of public goods (Samuelson, 1954), solutions can be categorized along the dimensions of rivalrousness and excludability.

B. Organizational intelligence

An interesting application of our model is to study ways of conceptualizing and measuring organizational intelligence. Organization scholars have developed several perspectives of organizational intelligence: decision making, organizational learning, and risk taking (March, 1999). For example, according to March (1999), an intelligent organization is "one that adopts procedures that consistently do well (in the organization's own terms) in the face of constraints imposed by such things as scarce resources and competition."

We see an opportunity to contribute to the organizational intelligence literature by isolating the parameters that determine the proportion of unsolved problems and unused solutions in a system (i.e., organization). We propose that an intelligent organization is one that manages to solve as much as possible its problems and utilize as much as possible its solutions. In reality, organizations attempt to do this by buying, developing, and selling expertise. As such, to explore this topic we will likely need to include in the model a mechanism for exchanging problems and solutions across systems.

C. Active search vs. passive uptake in organizational learning

Organizational adaptation relies on the exploration of new possibilities as well as the exploitation of old certainties (March, 1991). In problem solving, exploration can be achieved through at least two distinct mechanisms: active search on the problem-solution landscape and passive uptake of solutions. We see an opportunity to compare and contrast these mechanisms using our model as follows:

- Once a problem-solution landscape (on which problems and solutions presumably form a non-uniform distribution) is established, we can add a third kind of agent (let's call it the 'troll' at least for now) that can move around the landscape.

- The troll's objective is to 'innovate,' which is operationalized as: i) finding an unsolved problem (i.e., a novel combination of solutions), and ii) finding all of the unsolved problem's solutions.

- The troll has two primary features: i) it can move around with certain intensity (in other words, this describes the intensity of active search), and ii) it has a filter that evaluates solutions and accepts or rejects them as a result. (A special filter is not needed for evaluating problems because the criterion for accepting problems is straightforward: they only need to be previously unsolved.)

- The troll can collect problems and solutions in no particular order. In other words, the troll can accumulate solutions before it finds appropriate problems, and vice versa. (Again, this is why the solution filter is useful.)
- The solution filter evaluates solution candidates based on their similarity (i.e., Hamming distance) with the troll's current stock of solutions. (When the stock is zero, a solution is taken up randomly.)
- A stringent filter demands a high degree of similarity. Conversely, a loose filter can result in a somewhat random and unrelated collection of solutions.[2]
- It would be interesting to compare several strategies (as shown in Table 1) and see which one(s) is most useful for achieving the troll's objective to innovate.
- Furthermore, we can study group strategies by examining the different ways in which two or more trolls can share or split strategies.

Table 1: Innovation strategies

Active search / Filter	Low intensity	High intensity
Loose	1	2
Stringent	3	4

IV. ACKNOWLEDGEMENTS

We are extremely grateful to Tom Carter for his patience and generosity in guiding our thinking throughout this project. We also thank Jacopo Tagliabue for helpful discussions that became the basis for this paper. Finally, we thank Santa Fe Institute for providing us with so many wonderful opportunities to grow into adventurous and productive complex systems researchers.

[2] The concept of an unrelated collection of solutions is similar to 'peripheral knowledge,' which has been discussed in unpublished working papers (2008, 2009).

V. REFERENCES

Cohen, M. D., March, J. G., and Olsen, J. P. 1972. A garbage can model of organizational choice. *Administrative Science Quarterly*, 17(1): 1-25.

Hamming, R. W. 1950. Error detecting and Error correcting codes. *Bell System Technical Journal*, 26 (2): 147-160.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2 (1): 71-87.

March, J. G. 1999. *The pursuit of organizational intelligence*. Blackwell: Malden, MA.

Samuelson, P. A. 1954. The pure theory of public expenditure. *Review of Economics and Statistics*, 36 (4): 387–389.

Wilensky, U. 1999. *NetLogo*. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL.