SWARM INTELLIGENCE: THEORETICAL PROOF THAT EMPIRICAL TECHNIQUES ARE OPTIMAL

DIMA IOURINSKI, SCOTT A. STARKS, and VLADIK KREINOVICH
NASA Pan-American Center for Earth and Environmental Studies
University of Texas at El Paso
El Paso, TX 79968, USA
emails dmitrii@utep.edu, sstarks@utep.edu, vladik@cs.utep.edu

STEPHEN SMITH
Robotics Institute, Carnegie Mellon University
500 Forbes Ave., Pittsburgh, PA 15213, USA
email sfs@cs.cmu.edu

ABSTRACT

A natural way to distribute tasks between autonomous agents is to use *swarm intelligence* techniques, which simulate the way social insects (such as wasps) distribute tasks between themselves. In this paper, we theoretically prove that the corresponding successful biologically inspired formulas are indeed statistically optimal (in some reasonable sense).

KEYWORDS: Swarm intelligence, autonomous agents, optimality proof

INTRODUCTION

What is swarm intelligence

In many real-life situations, we have a large number of tasks, and a large number of autonomous agents which can solve these tasks. The problem is how to best match agents and tasks. This problem is typical:

- in manufacturing, where we have several machines capable of performing multiple tasks;
- in robotics, when we need to coordinate the actions of several autonomous robots;
- in computing, when several parallel computers are available, etc.

In general, if we want an optimal matching, then this problem is difficult to solve. For example, it is known that the problem of optimal manufacturing scheduling is NP-hard; see, e.g., [8]. Since we cannot have an optimal solution, we must look for heuristic solutions to such problems.

One of the natural sources of such heuristics is biology, specifically, the biology of insects. Insects are usually small, so it is difficult for an individual insect to perform complex tasks. Instead, they swarm together and perform tasks in collaboration. Since the existing social insects are the result of billions of years of survival-of-the-fittest evolution, we expect that all the features of their collaboration have been perfected to being almost optimal. Thus, it is reasonable to copy the way social insects interact. The resulting multi-agent systems are called *swarm intelligence* [2,9].

What formulas are used in the existing swarm intelligence systems

The biological observations led researchers to the following model for the insect collaboration: We have several classes of tasks. Each task T of type t is characterized by its degree of relevance $R_t(T)$; in biology, this degree of relevance is called a *stimulus*.

In principle, each agent can perform each task; in this sense, the agents are *universal*. However, different agents have different abilities with respect to different tasks. If an agent is not very skilled in a certain type of tasks, then this agent picks tasks of this type only when they are extremely important, i.e., when the stimulus is very high. If an agent is reasonable skilled in tasks of certain type, then this agent will also pick such tasks when the corresponding stimulus is much lower. This behavior can be characterized by assigning, to each agent A and to each type of tasks t, a *threshold* $\theta_t(A)$:

- if the stimulus $R_t(T)$ corresponding to a task T is much smaller than the threshold, then the agent will not take this task;
- if the stimulus is much larger than the threshold ($R_t(T) \gg \theta_t(A)$), then the agent will take this task.

In other words, whether the agent takes the task or not depends on the ratio $r \stackrel{\text{def}}{=} R_t(T)/\theta_t(A)$: if $r \ll 1$, the agent does not take the task; if $r \gg 1$, the agent takes the task.

When the ratio is close to 1 (i.e., when the stimulus is of the same order of magnitude as the threshold), then the same insect sometimes takes the task, sometimes does not. The frequency (probability) P with which an insect picks the task increases with the ratio r. From the biological observations, it was determined that the dependence of the probability P on the ratio r has the following form:

$$P(r) = \frac{r^2}{1 + r^2}. \tag{1}$$

In other words, the probability P of an agent A to pick the task T of type t is equal to:

$$P = \frac{R_t(T)^2}{R_t(T)^2 + \theta_t(A)^2}. \tag{2}$$

This formula is used in the existing swarm intelligence systems, and it has led to reasonable results [2–7,9].

Formulation of the problem

The idea that a probability P should depend on the ratio r is very convincing. However, the specific dependence of P on r (as described by the formula (1)) is rather ad hoc. Since this formula is successful, it is reasonable to try to find a justification for its use.

In this paper, we provide such a justification.

MAIN IDEA

Since we want to design an *intelligent* system, we should allow agents to learn, i.e., to use their experience to correct their behavior. In the swarm intelligence model, at any given moment of time, the behavior of an agent A towards tasks of all possible types t is characterized by its thresholds $\theta_t(A)$. Thus, learning means changing the agent's thresholds, from the original values $\theta_t(A)$ to new values $\theta'_t(A)$. As a result, the probability

$$P = P(r) = P\left(\frac{R_t(A)}{\theta_t(A)}\right) \tag{3}$$

of an agent A taking the task T changes to a new value

$$P' = P(r') = P\left(\frac{R_t(A)}{\theta'_t(A)}\right). \tag{4}$$

The formula describing the transition from the original probabilities (3) to the new probabilities (4) can be further simplified if we denote the ratio of the old and the new thresholds by

$$\lambda = \frac{\theta_t(A)}{\theta'_t(A)}.$$

In terms of λ, we have $r' = \lambda \cdot r$, hence the new probability is equal to

$$P' = P(\lambda \cdot r). \tag{5}$$

From the statistical viewpoint (see, e.g., [11]), the optimal way of updating probabilities is by using the Bayes formula. Specifically, if we have n incompatible hypotheses $H_1, \ldots, H_n$ with initial probabilities $P_0(H_1), \ldots, P_0(H_n)$, then, after observations E, we update the initial probabilities to the new values:

$$P(H_i \,|\, E) = \frac{P(E \,|\, H_i) \cdot P_0(H_i)}{P(E \,|\, H_1) \cdot P_0(H_1) + \ldots + P(E \,|\, H_n) \cdot P_0(H_n)}. \tag{6}$$

Thus, *an optimal function $P(r)$ can be determined as the one for which the transition from the old probabilities* (3) *to the new probabilities* (4), (5) *can be described by the (fractionally linear) Bayes formula* (6).

FROM THE MAIN IDEA TO THE EXACT FORMULAS

Let us formalize the above condition. In our case, we have two hypotheses: the hypothesis H_1 that it is reasonable for an agent A to take a task of given type t, and the opposite hypothesis H_2 that it is not reasonable for the agent A to take such a task. Initially, the probability of the hypothesis H_1 is equal to P, and the probability of the opposite hypothesis H_2 is equal to $1-P$. According to Bayes formula, after some experience E, the probability P should be updated to the following new value $P' = P(H_1 \,|\, E)$:

$$P' = \frac{P(E \,|\, H_1) \cdot P}{P(E \,|\, H_1) \cdot P + P(E \,|\, H_2) \cdot (1-P)}. \tag{7}$$

If we if we denote $P(E \,|\, H_1)$ by a, $P(E \,|\, H_2)$ by b, and explicitly mention that the probability P depends on the ratio r, then the formula (7) takes the following form:

$$P' = \frac{a \cdot P(r)}{a \cdot P(r) + b \cdot (1 - P(r))}. \tag{8}$$

We want the expression (5) to be representable in this form (8). So, we arrive at the following definition:

MAIN RESULT

Definition. *A monotonic function $P(r) : [0, \infty) \to [0, 1]$ is called optimal if, for every $\lambda > 0$, there exist values $a(\lambda)$ and $b(\lambda)$ for which*

$$P(\lambda \cdot r) = \frac{a(\lambda) \cdot P(r)}{a(\lambda) \cdot P(r) + b(\lambda) \cdot (1 - P(r))}. \tag{9}$$

Theorem. *Every optimal function $P(r)$ has the form*

$$P(r) = \frac{r^\alpha}{r^\alpha + c} \tag{10}$$

for some real numbers α and c.

In other words, for the optimal function $P(r)$, we have

$$P = \frac{R_t(T)^\alpha}{R_t(T)^\alpha + c \cdot \theta_t(A)^\alpha}. \tag{11}$$

If we re-scale the threshold by calling $\theta' = c^{1/\alpha} \cdot \theta$ the new threshold, then the formula (11) simplifies into

$$P = \frac{R_t(T)^\alpha}{R_t(T)^\alpha + \theta_t(A)^\alpha}. \tag{12}$$

Thus, we show that formula (12) – which is a minor generalization of the original formula (2) – is indeed optimal.

PROOF

It is known that many formulas in probability theory can be simplified if instead of the probability P, we consider the corresponding odds $O = P/(1-P)$. (If we know the odds O, then we can reconstruct the probability P as $P = O/(1+O)$.) The right-hand side of the formula (9) can be represented in terms of odds $O(r)$, if we divide both the numerator and the denominators by $1-P(r)$. As a result, we get the following formula:

$$P(\lambda \cdot r) = \frac{a(\lambda) \cdot O(r)}{a(\lambda) \cdot O(r) + b(\lambda)}. \tag{13}$$

Based on this formula, we can compute the corresponding odds $O(\lambda \cdot r)$: first, we compute the value

$$1 - P(\lambda \cdot r) = \frac{b(\lambda)}{a(\lambda) \cdot O(r) + b(\lambda)}, \tag{14}$$

and then divide (13) by (14), resulting in:

$$O(\lambda \cdot r) = c(\lambda) \cdot O(r), \tag{15}$$

where we denoted $c(\lambda) = a(\lambda)/b(\lambda)$. It is known (see, e.g., [1,10]) that all monotonic solutions of the functional equation (15) are of the form $O(r) = C \cdot r^{\alpha}$. Therefore, we can reconstruct the probability $P(r)$ as

$$P(r) = \frac{O(r)}{O(r)+1} = \frac{C \cdot r^{\alpha}}{C \cdot r^{\alpha} + 1}.$$

Dividing both the numerator and the denominator of the right-hand side by C and denoting $c = 1/C$, we get the desired formula (10). Q.E.D.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209 and grant NCC 2-1232, by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grants numbers F49620-95-1-0518 and F49620-00-1-0365, by grant No. W-00016 from the U.S.-Czech Science and Technology Joint Fund, by NSF grants CDA-9522207, ERA-0112968 and 9710940 Mexico/Conacyt, and by IEEE/ACM SC2001 Minority Serving Institutions Participation Grant.

Smith's work was sponsored in part by the Department of Defense Advanced Research Projects Agency and the U.S. Air Force Rome Research Laboratory under contracts F30602-97-2-0066 and F30602-00-2-0503 and by the CMU Robotics Institute.

REFERENCES

1. Aczel, J. *Lectures on functional equations and their applications* (Academic Press, New York, London, 1966).

2. Bonabeau, E., Dorigo, M., and Theraulaz, G. *Swarm Intelligence: From Natural to Artificial Systems* (Oxford University Press, Oxford, 1999).

3. Cesta, A., Oddi, A., and Smith, S. F. "Iterative flattening: a scalable method for solving multi-capacity scheduling problems", *Proc. of the National Conference on Artificial Intelligence AAAI'00*, Austin, TX, July 2000.

4. Cicirello, V. A., and Smith, S. F. "Ant colony control for autonomous decentralized shop floor routing", *Proc. 5th Int'l Symposium on Autonomous Decentralized Systems ISADS'2001*, IEEE Computer Society Press, March 2001.

5. Cicirello, V. A., and Smith, S. F. "Insect societies and manufacturing", *Proc. of IJCAI'01 Workshop on AI and Manufacturing: New AI Paradigms and Manufacturing*, August 2001.

6. Cicirello, V. A., and Smith, S. F. "Wasp nests for self-configurable factories", *Agents'2001, Proc. 5th Int'l Conference on Autonomous Agents*, ACM Press, May-June 2001.

7. Cicirello, V. A., and Smith, S. F., "Improved routing wasps for distributed factory control", *Proc. of IJCAI'01 Workshop on AI and Manufacturing: New AI Paradigms and Manufacturing*, August 2001.

8. Garey, M., and Johnson, D. *Computers and intractability: a guide to the theory of NP-completeness* (Freeman, San Francisco, 1979).

9. Kennedy, J., Eberhart, R., and Shi, Yu. *Swarm Intelligence* (Morgan Kaufmann, San Francisco, CA, 2001).

10. Nguyen, H. T., and Kreinovich, V., *Applications of continuous mathematics in computer science* (Kluwer, Dordrecht, 1997).

11. Wadsworth, H. M. (ed.) *Handbook of statistical methods for engineers and scientists* (McGraw-Hill Publishing Co., N.Y., 1990).