Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mathematics and Computers in Simulation 71 (2006) 1–8
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

On the advantages of non-cooperative behavior in agent populations

Alexander Pudmenzky*

Centre for Mined Land Rehabilitation, The University of Queensland, Brisbane 4072, Australia

Received 19 December 2003; accepted 28 July 2005
Available online 15 November 2005

Abstract

We investigate the amount of cooperation between agents in a population during reward collection that is required to minimize the overall collection time. In our computer simulation agents have the option to broadcast the position of a reward to neighboring agents with a normally distributed certainty. We modify the standard deviation of this certainty to investigate its optimum setting for a varying number of agents and rewards. Results reveal that an optimum exists and that (a) the collection time and the number of agents and (b) the collection time and the number of rewards, follow a power law relationship under optimum conditions. We suggest that the standard deviation can be self-tuned via a feedback loop and list some examples from nature were we believe this self-tuning to take place.

Keywords: Agent population; Co-operation; Reward collection; Armed bandit search; Optimum standard deviation; Exploitation and exploration

1. Introduction

Multi-agent systems include any computational system whose design is fundamentally composed of a collection of interacting parts. Agent based models are simulations based on the global consequences of local interactions of members of a population. Agent based modelling systems are now widely used in many disciplines such as Humans and Artificial Societies [6,8,14], Ecology and Biology [5,15], Economics [9,18,24], Traffic simulations [4,26] and Environmental modelling [7,12,20,21]. We investigate the behavior of a collective systems of agents that are able to communicate locally. The aim is to analyze how communication between agents can be optimized to fulfill a larger common goal such as the minimization of time taken to search for and collect randomly distributed rewards. We begin with a definition of the terminology we will be using in Section 2 which enables us to formulate a generic description of the problem in Section 3. Section 4 then introduces the parameter values we investigate followed by presentation of the results in Section 5. The discussion of the results contained in Section 6 is followed by Section 7 were we suggest the existence of naturally occurring examples. Finally, we conclude by offering some closing remarks in Section 8.

* Tel.: +61 7 33464052; fax: +61 7 33464056.
E-mail address: a.pudmenzky@uq.edu.au

0378-4754/$32.00 © 2005 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.matcom.2005.07.003

2. Definitions

We use the generic term *agent* for an artificial or biological entity playing a part in the behavior of a population. An agent can be a gene or an animal such as an insect or human being, or an artificial entity such as a software-agent, a router in a communications network, a central processor in a multi-CPU cluster or a mechanical robot, to just name a few examples. We also use the word *population* as a generic term for a collection of agents in a defined environment. A population can be represented by terms such as "genome", "group", "swarm", "ant-colony", "collective" or similar. Agents will generally try to collect *rewards* located at *targets* in a certain problem domain (context) which we will call their *world*. Those rewards can consist of food or completed tasks. Total reward collection time is to be always minimized and reverse-proportional to the fitness of the population. Using this terminology we will now attempt a more general formulation of the problem under investigation.

3. Problem description

Located in a d-dimensional world of size A_{world}, at each trial are K targets with a total of R equally distributed rewards so that each individual target consists of R/K rewards. The size of the targets, A_{target}, is chosen so that an arbitrary ratio $a = A_{\text{world}}/A_{\text{target}}$ is achieved. A population of N agents of zero extent (i.e. point-size agents) are uniform-randomly placed into this world at each iteration (cf. Fig. 1 for configuration). If an agent happens to be placed in a target area, the agent removes c (carrying capacity) rewards at this iteration. The agent will remain at this position and continue to remove c rewards at subsequent iterations until all rewards at this target position have been taken. The agent(s) may be joined by other agents that discover the target location at a later iteration. All agents participating in reward collection will again be participating in target location once the reward is exhausted. If all rewards are exhausted the number of

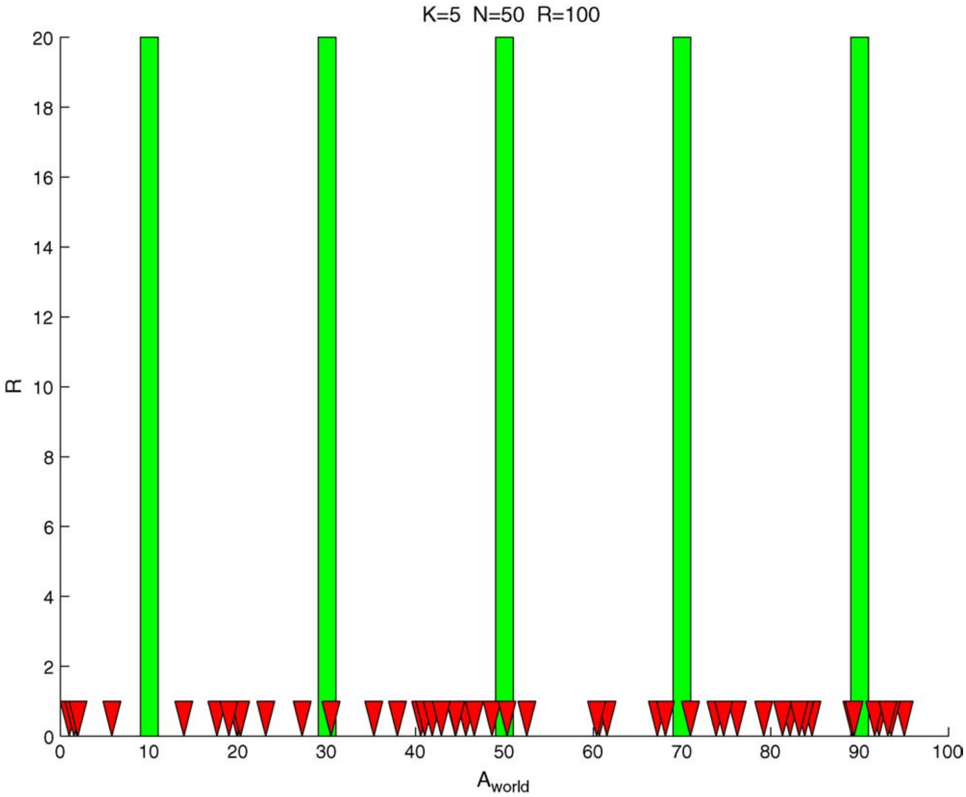

Fig. 1. Example configuration of $K = 5$ targets (bars) containing a total of $R = 100$ rewards and $N = 50$ agents (triangles) in a $d =$ one-dimensional world of size $A_{\text{world}} = 100$. Targets are occupying 10% ($a = 0.1$) of the world size.