From Bio-Inspired to Institutional-Inspired Collective Robotics

PTDC/EEA-CRO/104658/2008

Task 3: Towards a unified Bio-Institutional-inspired framework to analyse and synthesize collective systems

Progress Report

Rodrigo Ventura Pedro Lima

26 February 2011

Contents

1 Introduction

This progress report concerns task 3 of the project, aiming at a unified approach to collective systems, building upon the results from two previous tasks: task 1, concerning the simulation of immunological systems, and task 2, establishing the principles of an institutional robotics approach.

The main goal of this task is to create a unified framework upon which the bio-inspired and the institutional-inspired approaches can be cast. Such unification will be pursued at the level of a set of collective properties: stability, robustness, adaptation, and innovation.

Four basic properties of collective systems are here considered: stability, robustness, adaptation, and innovation. Stability concerns the response of a collective to a perturbation on the coupling between the agents. Robustness is assessed by removing or adding individuals with specific roles in the collective, and appraising the consequences. Adaptation is evaluated by performing changes on the environment and evaluating the collective response to those changes. While adaptation assumes small changes, innovation requires radical changes on the environment to be tested, such that rules of interaction between individuals are no longer appropriate, and new rules have to be set.

This report is organized as follows: first, the progress done in tasks 1 and 2 is briefly summarized in section 2, then section 3 introduces the framework under which the four properties, discussed in section 4, are presented; the design of experiments, under the presented framework, is discussed in section 5, and the conclusions are presented in section 6, closing the report.

2 Background work

Tasks 1 and 2 of this project provide the background material for the present task. This section provides a brief overview of these tasks, cast into the framework presented above.

Task 1 consists in a simulation of a collective of T helper (Th) cells, aiming at replicating experimental data taken from biological systems. In this simulation the agents represent Th cells, while the environment comprises the chemical environment surrounding the Th cells (e.g., cytokines) and Antigen Presenting Cells (APCs). The agent sensors are sensitive to the environment state, while its actuators allow the agent to change the environment state chemically. The agent program consists in a logical network with a dynamic internal state (induced by a set of internal feedback loops in the network), which in principle can be modeled by a deterministic automaton. This logical network design is biologically plausible as it was constructed with basis on empirical data. Preliminary results have shown that the simulated Th cells reach the same stable states as in the biological counterparts. Ongoing work aims at obtaining results from a collective of Th cells.

In Task 2 a link between institutional economics and the design of artificial collectives was approached. Two fundamental concepts are presented and discussed: *institutional environments* and *institutionalized individuals*.

Institutional environments comprise several aspects: (1) the existence of objects to which agent assign a function in a deontic fashion, and (2) the existence of rules, of several kinds, that the agents are expected to comply with. In these environments, institutions can be viewed as coordination artifacts among agents. Agents assign a normative meaning to the objects in (1) that result from the institutional environment they belong, rather than any object property or individual interpretation. A simple example of such an object is a traffic sign. The rules condition the behavior of the agents belonging to that institution. In task 2 several levels of these rules are distinguished — operational, collective-choice, and constitutional-choice — distinguishable in terms of at what level they are defined, and the time scale they change.

The concept of institutionalized individuals prescribes a decision-making framework for agents living in these institutional environments. These agents decisions are shaped by four individual variables: expected benefits, expected costs, internal norms, and individual discount rates. Environments are assumed non-deterministic and not necessarily known to the agent. Each decision results from a cost-benefit analysis, based on the agent expectations about outcomes. Internal norms determine the action options available to the agent in each situation, although the agent has the possibility of choosing not to comply with one (or more) norms. This option is modeled by a set of parameters, termed delta parameters, which are values that are added to the expected payoffs. Discount rates regulate the devaluation time imposes to future payoffs. In particular, different discount rates may have a major impact on the course of action of agents. The concept of internal model also plays an important role in this framework: agents have individual mental models of the world, which they use to evaluate expectations. These models can be shared with other agents. Moreover, the concept of shared mental models (called ideologies) is also introduced. This allows for subsets of agents to behave according to a coherent model of the world.

3 Framework

A collective is here understood as collection of entities (agents) embedded in an environment and interacting among them. Collectives can be analysed at two levels of complexity: the *micro-level* and the *macro-level*. The micro-level concerns individuals, as well as how a single individual interacts with others, within the collective. We approach this level modeling individuals as agents. At the macro-level, concerning the collective as a whole, we employ a systems theory perspective. It has been argued, in the field of physics, that systems comprising collections of particles, the individual and the collective levels of analysis are not reducible to a single level [LP00, LPS$^+$00]. We believe the same can be said about artificial systems, with the exception of very simple, uninteresting systems.

3.1 Micro-level

The elementary constituent of a collective is here considered as an *agent*. Following the Russell & Norvig definition, we consider an agent to be "anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators" [RN03]. The sensors and the actuators are the means through which the agent interacts with the environment. Thus, all information the agent receives from the environment is from sensors, and all external manifestations of the agent decisions are performed by the actuators-

In a multi-agent system, from the point of view of one of the agents, all of the other agents are part of the environment. This follows directly from the above definition, that does not distinguish between percepts originated by other agents or by any other entity. However, the environment is often viewed as all aspects of the system not considered as part of the agents.

Although direct communication among agents is in fact always mediated by the environment (e.g., even wireless communication employs electromagnetic waves), when this mediation has no impact on the content, this mediation is often abstracted and the communication is considered as *direct*.

Russell & Norvig further distinguish between agent *function*, the mathematical map between a sequence of percepts and an action, and agent *program*, corresponding to a computer implementation of an agent function. The term *behavior* is often also employed to designate the agent function.

3.2 Macro-level

To map collectives of agents to a systems theory view, a few definitions have to be established first:

- *inputs* are variables whose change is caused by the environment surrounding the collective, including external agents/robots
- *response/output* is the behavior displayed by the collective, measurable by observing several pre-defined features

- *desired value(s)* express what should be the target behavior for the collective, or the target value of the features characterizing the behavior
- *metric* is a well-defined mathematical concept that formalizes and abstracts the notion of distance in a metric space. A trivial example is the Euclidean norm. It is typically used to measure the difference or error between the actual and desired behavior or output of the collective
- *performance* measures how well a collective behaves, according to a function that quantifies behavior features to express the behavior quality numerically
- *parameters* are elements of the collective that can be changed, such as the communication frequency, the robot speed, a gain or similar
- *structure* concerns more fundamental elements of the collective, such as links between collective members, the collective geometry or similar

Given a collection of agents in a given environment, we define as *system state* as a full description of the system, from which a dynamical evolution of this state follows. This description includes both internal aspects of the agent (*i.e.,* the agent internal state) and the state of the environment. If the system is deterministic, given an initial state, the state trajectory is unique. In physical systems, however, such formulation is not feasible, and thus a stochastic model is often used instead.

Therefore, the system state trajectory depends heavily both on each agent behavior and on the environment properties. An interesting property of most collective systems is that the system evolution is often unpredictable, given an individual agent behavior. In these cases, the interaction among agents is determinant to the system evolution.

4 Properties of collectives

The approach taken in this task is based on the four properties of collectives outlined in the project proposal (see section 1). However, these properties have to be defined precisely enough, so that they can be systematically evaluated, given one collective system.

After a sequence of brainstorm discussions, and following Tasks 1 and 2 conclusions, the project team concluded that the original "definitions" in the project proposal are not so much definitions but rather propose operational ways of checking the referred properties, assuming a common-sense concept of what stability, robustness, adaptation and innovation are. This document proposes modifications to the original "definitions" so as to make them general enough, and provides a systematic method of designing an experiment to test them. The instantiation of the method relevant variables and parameters is case-study-dependent.

All of these properties concern the response of the collective, at the macro-level, to a perturbation at the micro-level. Thus, the evaluation of the response is framed into a systems theory view, while the perturbations impact first the micro-level.

For the analysis of collective system we assume that a set of performance measures are established first. For instance, in biological systems these measures may concern ratios of cells of a certain kind, while in robotic systems they may measure rate that the collective reaches their goals in a given task. The system is said to be in a *stationary regime* whenever the statistics of these measures within time window of fixed length do not change when this window is shifted in time.

The experimental methodology to evaluate how a system responds to a perturbation is the following:

1. let the system reach a stationary regime
2. perform the perturbation, and
3. wait until a new stationary regime is reached.

The quantitative evaluation of the response comprises the comparison of the statistics of the stationary regimes, before and after the perturbation.

The definitions proposed in this document are the following:

stability — given a perturbation in the communication links among agents, this property evaluates the change in the performance measures, in stationary regime, after that perturbation. A system is said to be *stable* with respect to a given performance measure, whenever it remains bounded, given a bounded perturbation.

robustness — this property distinguishes from the previous one in terms of the nature of the perturbation: while in the former the perturbation concerns the couplings among agents, in this one the perturbation is structural to the collective. This property evaluates the response of the collective after a change in the distribution of roles among agents. If a performance measure remains bounded after this kind of perturbation, the collective is said to be *robust.*

adaptation — the nature of the perturbations considered in this property concerns the environment. A collective system is said to be *adaptive* when it is able to accomodate for small changes in the environment, *i.e.,* not structural. One way of defining whether these changes are small is to model the environment in function of a set of parameters. While maintaining the same structure, the perturbations considered in this property concern bounded changes on the values of these parameters.

innovation — this property also evaluates the response of the system after a perturbation in the environment, but in this case structural changes are inflicted. Two kinds of changes are considered here: structural changes to the environment (such as the ones that cannot be captures by a change in parameters), and changes in the rules of interaction. By rules of interaction we consider the way an agent interprets actions or communications from other agents in the collective.

5 Designing Experiments

It is proposed that an experiment with collectives to check the four properties defined in section 4, designed according to the principles established in this report, should follow the following steps:

1. list the *inputs* of the collective, which are changed by the external environment
2. list the *response/output* of the collective, i.e., what should one measure to identify univocally what is its behavior at any time step
3. list the *desired value(s)* of the collective
4. list the *parameters* internal to the collective
5. list the *parameters* that characterize the environment
6. define what is considered as the *structure* of the collective
7. choose/define a set of *performance measures* to be used to quantify performance, stability and robustness
8. define the *performance* function for the experiment to be carried out with the collective.

Then, the method describe in section 4 should be carried out.

6 Conclusions

This report is focused on providing the first steps towards a common framework. The approach taken here consisted in fleshing out the four properties considered foundational in terms of crossing the bridge between the biological and the institutional approaches to collective systems. The two levels of analysis — the macro and the micro levels — were also explicitly addressed: the former is addressed using a systems theory approach, while the latter is modeled using the agent paradigm. The four properties were then framed with respect to these levels of analysis. Taken this framework into consideration, recommendations are set concerning the design of experiments.

References

[LP00] R. B. Laughlin and David Pines. The theory of everything. *Proceedings of the National Academy of Sciences of the United States of America*, 97(1):28–31, January 2000.

[LPS+00] R. B. Laughlin, David Pines, Joerg Schmalian, Branko P. Stojkovic, and Peter Wolynes. The middle way. *Proceedings of the National Academy of Sciences of the United States of America*, 97(1):32–37, January 2000.

[RN03] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, second edition, 2003.