

Maximization of Potential Information Flow as a Universal Utility for Collective Behaviour

Philippe Capdepu

Daniel Polani

Chrystopher L. Nehaniv

Adaptive Systems Research Group
School of Computer Science
University of Hertfordshire
College Lane, Hatfield, Herts
AL10 9AB, UK

Email: {P.Capdepu,D.Polani,C.L.Nehaniv}@herts.ac.uk

Abstract— We explore how information theoretic quantities such as potential information flow (empowerment) can be used as a drive toward complex collective behaviour in the context of multi-agent systems. In a first experiment, we investigate the empowerment of two agents interacting in a grid world. We show that some conditions lead to higher empowerment than others, depending on the amount of interaction and the amount of information shared by the agents. We then investigate more deeply the tradeoff between freedom of the agents and the constraints they impose on each other. We show that there exist a trade-off between these where empowerment is maximized. In a third experiment, we show that agents behaving so as to maximize potential information transfer over time generate a wide range of complex collective behaviours. We then discuss how these notions can be compared to what happens in natural systems.

I. INTRODUCTION

Natural systems composed of more or less autonomous agents often exhibit strong collective behaviour. Examples of such systems are multicellular organisms, flocks, ant colonies, animal societies and so on. These collective behaviours are still a very active research topic and some aspects already quite well understood (see for example [2]). For example we know that interaction and communication between the different parts constituting the system are very important conditions for global synchronization to emerge. The question of why such systems should appear has also been studied; one famous example is the emergence of collaboration in the iterated prisoners' dilemma. In this model, it has been shown that given certain conditions, collaborative behaviour gives a better payoff to all the participants in the long run (see [1]). Studies of collective behaviour in animal societies also often take the point of view of a reward maximization. The same approach is used to study the transition from unicellularity to multicellularity (see [9] and [3]).

Here we would like to adopt a different approach to study these phenomena. Recent work suggests that there exist intrinsic drives toward these kind of behaviours that would not be task-dependent (and would thus avoid the problem of reward design). Such kinds of intrinsic drives have been studied mainly in the context of single agent. They have

been applied to evolution of perception-action loops, see for example work on homeokinesis [5], maximization of information structure [12], spatiotemporal coordination [11], autotelic principle [13] and self-motivated learning [10]. If these taskless strategies derive from utilities then we can refer to them as *universal utilities*.

Here we propose to use a universal utility based on the maximization of the control an agent has on its own environment. A measure of such a capacity, referred to as *empowerment*, has been recently introduced in the framework of information theory as a good candidate for a universal utility in the case of single agent systems (see [6] and [7]). An advantage of this measure is that it can be used as a dense feedback for driving behaviour instead of a sparse feedback such as artificial evolution and reinforcement learning. In the present paper we want to investigate empowerment as a universal utility in the case of multi-agent systems. The first question we would like to answer is what are the conditions for which two agents gain something by collaborating. The second question we would like to answer is what would be the collective behaviour of agents trying to maximize their own control on the environment. Our working hypothesis is that empowerment, as a universal (non-zero sum) utility, can provide a useful guide to answer these questions. Validating this hypothesis can help tackling issues such as the transition to multicellularity, or other kinds of self-organization phenomenon in multi-agent systems.

In Sec. 2 we first introduce the mathematical basis of information theory and we describe empowerment and its properties. In Sec. 3 we describe a first experiment in which we measure the empowerment of two agents with various degrees of interaction and communication. Section 4 extends the first results to the case of large number of agents interacting in a grid world with different densities of agents. In Sec. 5 we explore the collective behaviour of systems where behaviour of individual agents is to act toward maximizing their own empowerment. In Sec. 6 we discuss the overall results and how the introduced

notions can be related to what is happening in natural systems.

II. INFORMATION THEORY AND EMPOWERMENT

A. Basis of Information Theory

Shannon's information theory is a mathematical framework that quantifies properties of probability distributions. We refer the reader to [4] for a complete introduction to the field. One of the most important quantity is the *entropy* of a probability distribution. Consider a random variable X for which each event x can take a value in the set \mathcal{X} . The probability of one event x is written as $Pr(X = x)$, to simplify the equations we will use the notation $p(x)$. The Shannon entropy of this random variable is defined as

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log_2 p(x) \quad (1)$$

This value reflects the uncertainty about the outcome of this random variable.

Another important quantity of information theory is the *mutual information* between two random variables. This value measures the mutual dependance of the two variables:

$$I(X; Y) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \quad (2)$$

where X and Y are two random variables, $p(x, y)$ is the joint probability distribution function of X and Y , and $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y respectively.

The last important fundamental value is the *channel capacity* which is the maximum amount of information that can be transmitted through a channel. The channel is defined by the conditional probability $p(y|x)$ where x is the message emitted and y is the received message. It is defined as the maximum of the mutual information between the input X and the output Y of that channel, where the maximization is with respect to the input distribution described by the distribution $p(x)$. This can be written as

$$C = \sup_{p(x)} I(X; Y) \quad (3)$$

B. Empowerment

Empowerment has been recently introduced in [6], it is based on an agent-centric point of view and relates to the perception-action loop of this agent unrolled over time (here we consider a memoryless channel). The idea of empowerment is to measure how much information can be injected by an agent into its environment and then perceived by its sensors. More precisely it is defined as the channel capacity from the sequence of actions $A_t, A_{t+1}, \dots, A_{t+n-1}$ to the perceptions S_{t+n} after an arbitrary number of timesteps

$$\mathfrak{E}(A_t, \dots, A_{t+n-1} \rightarrow S_{t+n}) = \sup_{p(\vec{a})} I(A_t, \dots, A_{t+n-1}; S_{t+n}) \quad (4)$$

with $p(\vec{a})$ being the probability distribution function of the

action sequences (i.e. $\vec{a} = (a_t, \dots, a_{t+n-1})$). Empowerment can be described as the maximum potential information an agent can transfer into its own sensors through the environment. The higher the empowerment the more information it can potentially inject.

We will also consider another measure of empowerment referred to as *context-dependent empowerment* which has been introduced in [8]. Here we mean by context the sensoric state before the action is performed. By opposition we will refer to the first one as *blind empowerment*. Context-dependent empowerment can be computed as

$$\begin{aligned} & \mathfrak{E}(A_t, \dots, A_{t+n-1} \rightarrow S_{t+n} | S_t) \\ &= \sum_{s_t \in \mathcal{S}_t} p(s_t) \sup_{p(\vec{a})} I(A_t, \dots, A_{t+n-1}; S_{t+n} | s_t) \end{aligned} \quad (5)$$

Blind empowerment defines the general ability of an agent to inject information into its own sensors, whereas context-dependent empowerment considers that the agent can make use of different action policies according to what it is sensing. From a global perspective, empowerment can be used as a universal utility to guide the behaviour of an agent, but also to evolve the sensors or actuators of an agent in order to maximize its control abilities on a given environment. In the following sections we will be only considering 1-step empowerment.

III. FIRST EXPERIMENT: WHY BE TOGETHER?

The first question we want to address is "What are the conditions for which two agents gain something by collaborating?" First we have to define what kind of gain we are looking for. As we explained in the introduction, the gain is usually based on some kind of task-dependent reward. Our hypothesis is that the nature of the gain could be independent of any reward mechanism, and would be based on an increase of the control abilities of the agent over its environment. It is then natural to look for a gain in the empowerment of the agents (here we only consider mutual gain).

For the other part of the question, the conditions, we will focus on two aspects. The first one is the amount of interaction or constraint between the two agents, i.e. how much what an agent has done will impact what the other agent can do. The other aspect is the amount of information shared by the two agents, which can also be translated as how much information one agent perceives about the other one. We must here explain that the information shared is not necessarily an intentional communication mechanism, it can be simply perceiving some information about the other agent, whether it is extracted from the other agent's sensors or from an external point of view.

A. Experimental Setup

To address this question, we propose a very simple simulated world in which only two agents are interacting. The world is a 2D grid of finite size whose horizontal and vertical

borders are wrapped around. Each agent has an absolute location in the grid and a rigid body which can occupy one or more tiles and whose orientation never changes. At each timestep both agents are allowed to move rigidly by a distance of one tile either vertically or horizontally. The only constraint of the system is that if a move results in the agents overlapping, then the move is prevented. Agents can sense their absolute position in the world and some information from the other agent's position. The first parameter we study is the shape of the agents which has an impact on the amount of constraint they impose on each other. The second parameter is the amount of information about themselves that they symmetrically share.

The precise settings used are the following: the grid is a 8 by 8 squares (a larger size would have been computationally intractable), so absolute position of an agent is stored in 6 bits. Sensors report the absolute position and a given number of bits (0 to 6) of the other agent's position. The amount of information shared between the two agents is constrained by masking some information of the other agent's position. For a given number of bits, each possible mask is independently evaluated and the average value is taken. For each possible startup position of the two agents, we collect the current sensor values, and for each possible action of the two agents we collect the new sensor values. In order to obtain results that are not too close to the theoretical maximum, we introduce noise in the sensor values (each bit has 20% chances to be randomly set to 0 or 1). Empowerment is computed after all the statistics have been collected $100 * 2^n$ times for each case with n being the number of bits shared. This is done in order to avoid undersampling and also to have the same error for each quantity of information shared. The shape of the agents is chosen from a list of L-shapes of different sizes (see figure 1). For each condition we measure 1-step blind empowerment of the agents.

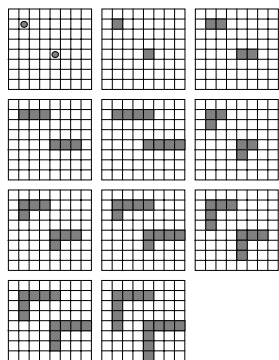


Fig. 1. List of shapes used for the agents. The top left is size 0, i.e. agents cannot collide. For all the other sizes, agents collide if grey tiles are overlapping.

B. Results

Results are shown in figure 2. Horizontal axis shows the size of the agents, vertical axis is the empowerment value. The different lines indicate the differing numbers of bits shared between the two agents. From these results we can see that when there is some interaction between the two agents, i.e. for all sizes except 0 (no collision), the more information that is shared the higher the empowerment is. Also different size of agents leads to different values for empowerment. We can see that empowerment grows when we increase the amount of constraint between agents until a certain point. From other experimental results we have also seen that when size reaches a too high value, e.g. when agents cannot move because they collide most of the time, empowerment decreases again. It seems that empowerment is the highest when there are sufficient interactions between the agents but they still have enough freedom.

We can then partially answer our question. One necessary condition for agents to have a gain in empowerment is that they get information about each other. The more they get, the more they are empowered. The other necessary condition is that they constrain each other's actions. Hence if we assume that empowerment is universal utility driving behaviour, then results show that it is an advantage for agents to share information and interact, hence it would be a natural drive toward collective behaviour. But it is not clear yet how interaction constraints really impact empowerment.

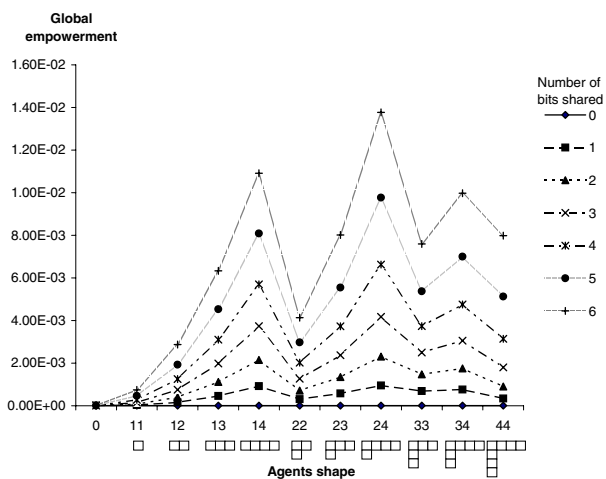


Fig. 2. Plotting of empowerment (vertical axis) for different size of the agents (horizontal axis) and different quantity of information shared (different lines). When agents are asymmetric, i.e. horizontal and vertical size are different, their empowerment is significantly higher than for symmetric agents.

IV. SECOND EXPERIMENT: CONSTRAINT OR FREEDOM?

The aim of the second experiment is to extend the results obtained in the first one to more than two agents. Also we want to explore more precisely the tradeoff between amount of interaction between agents and the freedom of individual agents, i.e how much they constrain each other.

A. Experimental Setup

We use a setup very similar to the first experiment, but we will use a larger grid and several agents. Sensors are not anymore based on the absolute position of the agent but on local sensing of the density of agents around the agent (see figure 3). Therefore the capacity of the sensors is bounded by the number of different values they can take. Agents occupy only one tile and the only constraint to their movement is that they cannot move to a tile where another agent already is. The parameter we will investigate is the global density of agents on the grid. The more agents there are, the higher the probability to interact with another agent, but the lower the freedom of an agent is. According to the results of the first experiment, our hypothesis is that there must exist a maximum in the empowerment for a given value that is the optimal tradeoff between strength of interaction and freedom.

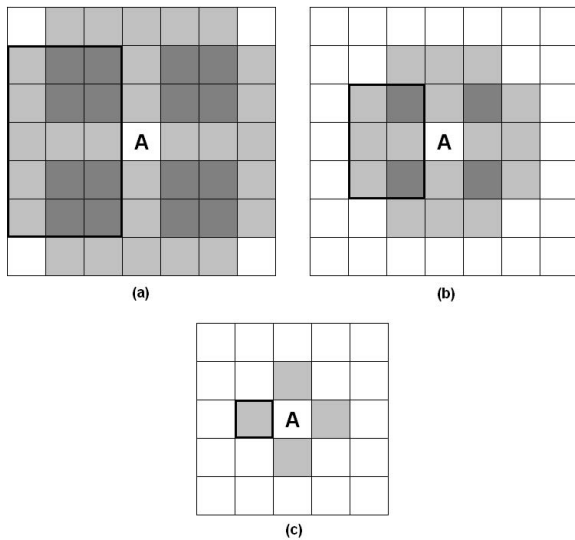


Fig. 3. Sensors used in the experiment. Each sensor is divided into 4 parts (black rectangle) and each part estimates the density of agents in the grey area (dark grey shows overlapping sensory fields). (a) Sensor of range 3. The number of agents is divided by 2 (integer division) in each sensor to fit in 3 bits. The entire sensoric state is then stored on 12 bits. (b) Sensor of range 2. The number of agents in each part is directly stored in 3 bits. (c) Sensor of range 1. The number of agents in each part is directly stored in 1 bit.

As it becomes computationally intractable to evaluate the true empowerment value (because the number of possible states of the system gets very high) we use a Monte Carlo approximation. At each timestep we randomly choose an agent, store its current sensoric state, and for each possible action we compute the new sensoric state. The agent is then allowed to really move according to a purely random policy where each action is equiprobable. We repeat this process for several timesteps (more than 1000000) until statistics converge to stable values. The experiment is performed for the three sensors described in figure 3.

B. Results

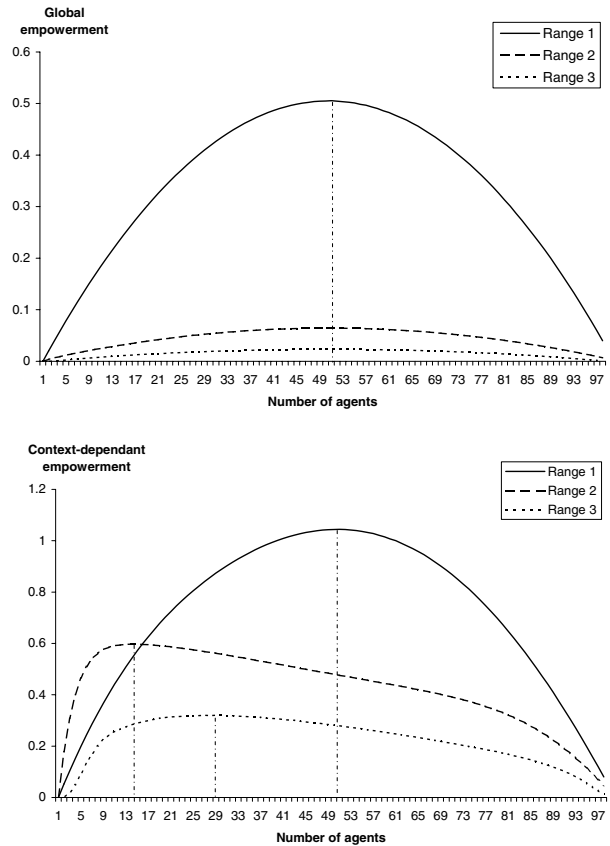


Fig. 4. Plotting of blind and context-dependant empowerment values (vertical axis) for different numbers of agents in the grid (horizontal axis) and different sensoric ranges. Vertical dashed lines indicate maximum values.

Results of this experiment are shown in figure 4. The first thing we can notice is that for each case there exists an optimal density that maximizes empowerment. At the two extreme densities, empowerment reaches minimal values, and it continuously increases when we go toward the optimal density. This observation supports our hypothesis that there exists an optimal tradeoff between the strength of interaction, or the degree of constraint that agents endure, and the freedom they have. Too much freedom leads to a completely unstructured world where there is no information to gather or where the randomness of the environment is too high for the agents to make use of it. On the other side, when there is too much constraint, the agents are so restricted in their actions that they cannot have any control on their environment. At this point a parallel can be made with the behaviour of physical systems such as water. At one extreme the system is so constrained that it results in a very ordered static structure, like an ice crystal. At the other extreme the system is completely free and it ends up in a completely random structure, or barely any structure at all like in a gas. In the middle there exists an area where the system can exhibit complex structures, such as vortices, which can exist

only because there is enough order and disorder at the same time. Nevertheless the parallel has to stop here because in our context agents do have autonomy, which distinguishes them from (passive) water molecules.

Pushing the analysis a bit further, we can observe that in the case of blind empowerment, the maximum values are reached for middle densities. Moreover whichever sensor type is used, the maximum is reached for the same density. Whereas in the case of context-dependent empowerment, changing the sensor has an effect on the position of the optimum density values. If we now consider the different densities as different environments, we can observe that the sensoric apparatus an agent possesses defines a range of environmental conditions for which the agent has optimal control. This observation could speculatively be related to the concept of biological niches. In this context, our information theoretic perspective provides a natural principle for relating sensorimotor capacities to their associated niches. This can be used in two different ways, the first is that an agent with a given sensorimotor apparatus should find the niche where his abilities are optimal. The second way is to use this principle as a criterion for evolving sensors for a given environment (related results are described in [7]).

A last comment we can make on these results concerns the difference between blind and context-dependent empowerment. The fundamental difference between the two measures is that blind empowerment is the maximum potential information an agent can transfer to its sensors if it uses the same action policy all the time, whereas context-dependent empowerment considers that for each sensoric state the agent can use a different action policy. One can see on the plots that in the context of blind empowerment, the higher the capacity of the sensor, the less efficient they turn out to be. This can be explained by the fact that the state of a very simple sensor such as range 1 having very few different states can more easily be correlated with previous actions (empowerment is the maximum possible correlation between actions and future sensoric states). Whereas in the case of the range 2 and range 3 sensors, these have a high number of different states which are then much more difficult to correlate with the previous actions if we do not consider the sensoric context of these actions. Now when we consider context-dependent empowerment, the loss of efficiency is less dramatic. This suggests that when sensors are very simple, the gain of switching from simple to complex behavior is less important than when using more complex sensors. This suggests that the adaptation to different niches crucially depends on making use of complex behaviour, and therefore complex sensors.

V. THIRD EXPERIMENT: COLLECTIVE EFFECTS THROUGH INDIVIDUALISTIC BEHAVIOR?

Concerning the last experiment, one can argue that using a random behaviour for the agents is a very particular case of behaviour. Our idea was to avoid any bias in the structure

contained in the environment, with random behaviour leading to a highly unstructured environment. Agents using specific behaviour rules often generate complex structures in the environment, such as cluster of agents, low density areas. But the question we want to ask now is clearly related to this problem. In fact it is a two part question. The first part is *what kind of behaviour would be obtained if agents act so as to locally maximize their own empowerment?* More precisely, as we can attribute a different empowerment value for each sensoric input, what behaviour will result from the agent trying to reach the sensoric states with maximal empowerment. The other part of the question is *how will empowerment evolve if iteratively computed according to statistics collected for non-random behaviour?* The two questions are interrelated, because there will clearly be some kind of coevolution between the behaviour and the empowerment values. More precisely, empowerment values will define the current behaviour and this behaviour will create some structures in the environment. These structures then have an impact on the action-perception statistics collected, which will then be used to define the new behaviour, and so on.

A. Experimental Setup

As for the last experiment, we will use a Monte-Carlo approximation to compute empowerment. The main difference here is that the action policy will not be random, but instead agents will act so as to increase their own empowerment. The empowerment values that will define the behaviour are computed according to the statistics collected during a given period of time that we will call an epoch. During each epoch, action-perception statistics are collected and merged with the statistics collected during all previous epochs. At the end of the epoch, blind and context-dependent empowerment are computed and the latter is used to define the next behaviour rule. The behaviour rule is the following: when an agent has to move, we virtually execute each action and retrieve the corresponding sensoric state, this information is then used to feed the statistics. According to the statistics collected during the epochs, context-dependent empowerment values for each sensoric state is retrieved. The behaviour rule is then to move toward a sensoric state whose empowerment is maximum. If different accessible sensoric states are associated with the maximum empowerment value, they are considered equiprobable and one of them is chosen randomly.

To sum up, at each epoch a new behaviour rule is applied, whose principle is to maximize empowerment of the agents computed over the statistics collected during all the preceding epochs. In our experimental setup, we used a totalistic sensor that counts the number of agents in a 7x7 square centered around the perceiving agent. We used a toroidal grid world of 100x100 where 1000 agents are randomly placed. One epoch consists in 1000000 movements of randomly chosen agents. The initial behaviour rule is the random policy.

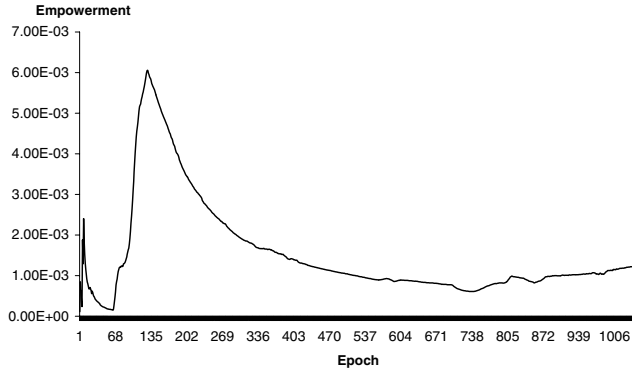


Fig. 5. Plotting of context-dependent empowerment values (vertical axis) during the evolution of the system (horizontal axis) expressed in epochs. The very first peak is an artifact due to undersampling at the beginning of the simulation. A first big transition is visible around epoch 65, stopping at epoch 130, during which empowerment is rapidly increasing. Another transition starts at epoch 700 with a small drop in empowerment followed by a slow durable increase. This transition is detailed in figure 7.

B. Results

The resulting simulation exhibited a wide range of different collective behaviours¹, ranging from highly-structured clusters, world split between stable areas of different densities, oscillating behaviour switching from high to low density clusters. We also measured blind and context-dependent empowerment during the time evolution of the system. The two are very similar but the most striking phenomenon is that in some cases there is a more or less sudden increase in empowerment during which the system gets into a very structured and stable organization. Two examples of this phenomenon are shown on figure 5, the first transition starts at epoch 60 and is illustrated in figure 6, the second one is the transition starting around epoch 700 illustrated in figure 7. One explanation for this increase would be that at some point the behaviour of the system leads to a highly structured organization which then reinforces the behaviour which made it emerge. Therefore there can exist a positive feedback loop between the behaviour and the structure it generates. But all the times that this phenomenon has been encountered so far, the stability of the structure was only temporary, and after some time a new behaviour rule emerged, leading to different structures. Also very similar structures could be observed during many different runs of this simulation.

VI. CONCLUSION

Multi-agent systems can exhibit complex collective behaviour out of simple agent rules. Starting from the hypothesis that quantities such as the potential information flow through the environment from actions to perceptions of an agent could be used as a drive toward these kinds of behaviours. We also expressed the hypothesis that there should exist a trade-off between the freedom of an agent

¹A video of the corresponding simulation is available at <http://homepages.feis.herts.ac.uk/~cp6cl/empowermentbehaviour.avi>.

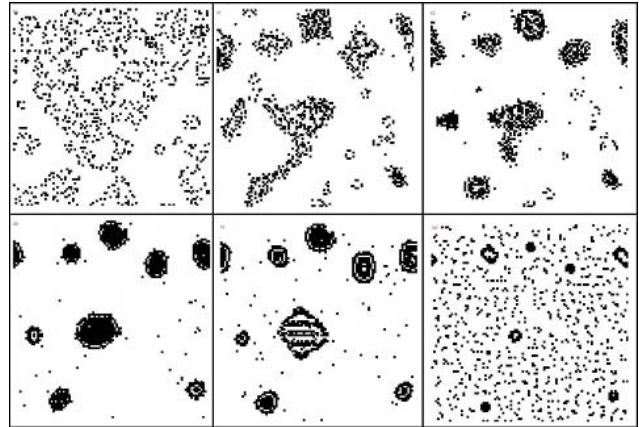


Fig. 6. Snapshots of the system's temporal evolution during the first transition starting around epoch 60 in figure 5. Black tiles are agents, white ones are empty tiles. From left to right and top to bottom corresponding epochs are 60, 61, 62, 69, 75 and 147. The three first snapshots are taken before the transition occurs, agents start forming dense clusters. In the fourth picture, very dense clusters are formed, surrounded by a ring of agents. The fifth picture shows a change in the middle structure (just after the small stop of increase on the plot). The last picture shows the system during the decaying phase. Structures are slowly fading out as agents start moving freely.

and the amount of constraint that the other agent induces, for which this information transfer capacity is maximal. In a first experiment involving two agents we have shown that interaction between agents and sharing of information are two necessary conditions to increase their empowerment. Interacting without sharing information did not increase empowerment, and the same is true for sharing information without interacting. Above a minimal value of interaction, sharing information has been shown to continually increase empowerment. On the other hand, increasing the quantity of interaction did not result all the time in an empowerment increase. From this we concluded that for a given quantity of shared information, there should be an optimal quantity of interaction that maximizes empowerment.

In a second experiment we extended the initial results to the multi-agent case and we investigated the trade-off between freedom of an agent and the amount of constraint due to interaction with other agents. For this we measured empowerment for several agents interacting through collisions in a grid world. The amount of information shared is defined by the sensors used, and the effect of the amount of constraint was investigated by varying the density of agents on the grid. The results obtained show that there is an optimal density of agents where empowerment is maximized. The optimal density is highly dependent on the sensoric apparatus of the agents (and also probably on their actuators). We have also shown that the sensitivity to the sensors is most apparent in the case of context-dependent empowerment, i.e. the use of a complex behavioural rule is necessary to take advantage of the optimal density. Also having complex sensors is an advantage in terms of empowerment only if

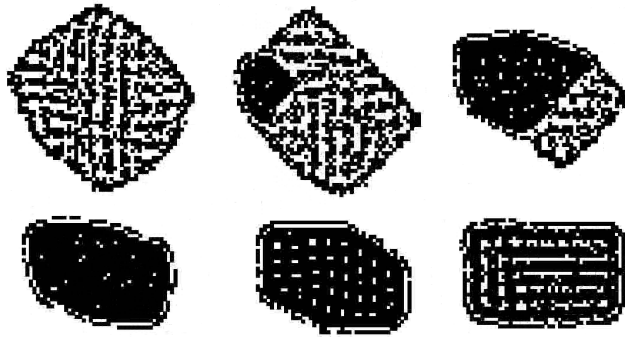


Fig. 7. Snapshots of the system's temporal evolution during the transition starting around epoch 700 in figure 5. By opposition to the first transition, here only the main structure is shown and scaled up. Black tiles are agents, white ones are empty tiles. Top left is the main structure just before the transition (684). On its right are shown later dynamical shifts in the structure after transition has started (699 and 702), we can see that a dense area appears on the left of the structure and then expands toward the right. On bottom line on the left is the structure after the main transition has occurred (704), a very dense area surrounded by a membrane-like ring of agents. The two last pictures are taken later (802 and 985) while empowerment keeps slowly increasing. In the first one we can see a pattern of holes in the structure which are regularly spaced and more or less aligned in the vertical axis. In the last picture, the shape of the structure is nearly rectangular and we can see a pattern of horizontal lines in its right part.

they are used in the context of a more complex behaviour rule.

In the third experiment we investigated collective behaviour resulting from the maximization of empowerment by individual agents. Starting from a random action policy, we iteratively collected statistics of the action-perception of the agents, then computed empowerment, and used this empowerment to define the new behaviour rule. Results have shown that a wide variety of complex collective structures emerged from the behaviour of the agents, such as dense clusters, different density areas, membrane-like structures, regular patterns, pulsing behaviour.

Generally, we have shown that empowerment can efficiently replace a task-dependent rewarding system by a more general utility measure in order to generate collective behaviour. In some conditions, interacting agents become more empowered, hence maximizing their own potential information transfer may be a good intrinsic drive toward collaborative behaviour. Conditions for which agents maximize their empowerment crucially depend on the degree of interaction between the agents, i.e. agents must have enough freedom to act but they also must be constrained by what the other agents are doing. Another condition for them to increase their empowerment by interacting is to be able to get information from or about other agents. Optimal conditions for empowerment to be maximized depend on the embodiment of the agent, i.e. its sensoric apparatus and motoric abilities.

Investigation of collective behaviour where individual agents try to maximize their empowerment by moving toward

the most empowered situations has shown that this principle can lead to a wide range of different behaviours generating complex structures and globally coordinated dynamics. The dynamics of the system are based on a direct and dense feedback from the structures generated and the behaviour that generates them. This is in contrast to the sparse feedback such as that provided by long-term selection in evolutionary processes. We hypothesize that such a dense feedback principle might be exploited in natural or artificial systems to drive them toward generating complex organizations and behaviours.

ACKNOWLEDGMENT

The authors would like to thank Alexander Klyubin for productive discussions about information flow and empowerment, and the anonymous reviewers for their useful comments.

REFERENCES

- [1] Axelrod, R., (1981). The Evolution of Cooperation. *Science*, 211(4489):1390-6 .
- [2] Bonabeau, E., Dorigo, M. and Theraulaz, G., (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press.
- [3] Buck, M. and Nehaniv, C.L., (2006). Effect of multi-level fitnesses on the development of multicellular artificial organisms. In Artmann, S., Dittrich, P., editors, *Proceedings of the Seventh German Workshop on Artificial Life*, pages 51-61. IOS Press.
- [4] Cover, T., and Thomas, J., (1991). *Elements of Information Theory*. New York: John Wiley and Sons.
- [5] Der, R., Steinmetz, U. and Pasemann, F., (1999). Homeokinesis - a new principle to back up evolution with learning. In *Computational Intelligence for Modelling, Control, and Automation*, volume 55 of *Concurrent Systems Engineering Series*, pages 43-47, Amsterdam, 1999. IOS Press.
- [6] Klyubin, A. S., Polani, D., and Nehaniv, C. L., (2005). All Else Being Equal Be Empowered. In Caparrere, M. S., Freitas, A. A., Bentley, P. J., Johnson, C. G., and Timmis, J., editors, *Advances in Artificial Life: Proceedings of the 8th European Conference on Artificial Life*, volume 3630 of *Lecture Notes in Artificial Intelligence*, pages 744-753. Springer, Sep 2005.
- [7] Klyubin, A. S., Polani, D., and Nehaniv, C. L., (2005). Empowerment: A Universal Agent-Centric Measure of Control. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation*, volume 1, pages 128-135. IEEE Press, Sep 2005.
- [8] Klyubin, A. S., Polani, D., and Nehaniv, C. L., (2006). Empowerment, Universal Utilities and the Perception-Action Loop. (in preparation)
- [9] Michod, R. E., (1999). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton University Press.
- [10] Oudeyer, P.-Y., and Kaplan, F., (2004). Intelligent adaptive curiosity: a source of self-development. In Berthouze, L., Kozima, H., Prince, C. G., Giulio Sandini, G., Stojanov, G., Metta, G., and Balkenius, C., editors, *Proceedings of the 4th International Workshop on Epigenetic Robotics*, vol. 117, pages 127-130, 2004. Lund University Cognitive Studies.
- [11] Prokopenko, M., Gerasimov, V., Tanev, I., (2006). Measuring Spatiotemporal Coordination in a Modular Robotic System. In Rocha, L. et al., editors, *Proceedings of the 10th International Conference on Artificial Life*, pages 185-191. MIT Press, Aug 2006.
- [12] Sporns, O. and Lungarella, M., (2006). Evolving coordinated behavior by maximizing information structure. In Rocha, L. et al., editors, *Proceedings of the 10th International Conference on Artificial Life*, pages 323-329. MIT Press, Aug 2006.
- [13] Steels, L., (2004). The Autotelic Principle. In Fumiya, I. et al., editors, *Embodied Artificial Intelligence*, volume 3139 of *Lecture Notes in Artificial Intelligence*, pages 231-242. Springer Verlag, Berlin, 2005.