All Else Being Equal Be Empowered

Alexander S. Klyubin, Daniel Polani, and Chrystopher L. Nehaniv

Adaptive Systems Research Group,
School of Computer Science, Faculty of Engineering and Information Sciences
University of Hertfordshire, College Lane, Hatfield Herts AL10 9AB, UK
{A.Kljubin, D.Polani, C.L.Nehaniv}@herts.ac.uk

Abstract. The classical approach to using utility functions suffers from the drawback of having to design and tweak the functions on a case by case basis. Inspired by examples from the animal kingdom, social sciences and games we propose *empowerment*, a rather universal function, defined as the information-theoretic capacity of an agent's actuation channel. The concept applies to any sensorimotoric apparatus. Empowerment as a measure reflects the properties of the apparatus as long as they are observable due to the coupling of sensors and actuators via the environment.

1 Introduction

A common approach to designing adaptive systems is to use utility functions which tell the system which situations to prefer and how to behave in general. Fitness functions used in evolutionary algorithms are similar in spirit. They specify directly or indirectly which genotypes are better.

Most utility functions and fitness functions are quite specific and a priori. They are designed for the particular system and task at hand and are thus not easily applicable in other situations. Each time the task and the properties of the system have to be translated into the "language" of the utility or fitness function. How does Nature address this problem? Is there a more general principle?

One common solution found in living organisms is homeostasis [1]. Organisms may be seen to maintain "essential variables", like body temperature, sugar levels, pH levels. Homeostasis provides organisms with a local gradient telling which actions to make or which states to seek. The mechanism itself is universal and quite simple, however the choice of variables and the methods of regulation is not. They are evolved and are specific to different phyla.

2 Empowerment

2.1 Motivation

Our central hypothesis is that there exist a *local* and *universal* utility function which may help individuals survive and hence speed up evolution by making the fitness landscape smoother. The function is local in the sense that it doesn't

M. Capcarrere et al. (Eds.): ECAL 2005, LNAI 3630, pp. 744–753, 2005.

rely on an infinitely long history of past experience, does not require global knowledge about the world. The utility function is applicable to all species, hence, it should be universal. At the same time it should adapt to morphology and ecological niche. The utility function should be related to other biologically relevant quantities.

In the quest for the function one invariably notices certain traits reappear in different contexts over and over again. In animal kingdom we see the striving for domination and control. Humans and even states strive for money, power and control. In board games such as Reversi or Othello there is a concept of mobility, which is defined as the number of moves a player can make. Everything else being equal players should seek higher mobility.

The unifying theme of these and many other examples is the striving towards situations where *in the long term* one could do many different things *if one wanted to*, where one has more *control* or influence over the world. Predators with better sensors and actuators can hunt better. Having high status in a group of chimpanzees allows one more mating choice. Having a lot of money enables one to engage in more activities. One can choose from an array of options. However, if one doesn't know what to do, a good rule of thumb is to choose actions leading to higher status, more power, money and control. We will now apply this idea to "embodied" agents.

2.2 The Concept of Empowerment

In his work on ecological approach to visual perception [2] Gibson proposed that animals and humans do not normally view the world in terms of geometrical space, independent arrow of time, and Newtonian mechanics. Instead, he argued, the natural description is in terms of what one can perceive and do. Thus, different places in the world are characterized by what they afford one to perceive and do.

This perspective is agent-centric. The concept of "the environment" is a by-product of the interplay between the agent's sensors and actuators. In this spirit we base our utility function solely on the sensors and actuators, without the need to refer to the "outside" of the agent.

We propose *empowerment*, a quite general utility function, which only relies on the properties of "embodiment", the coupling of sensors and actuators via the environment. Empowerment is *the perceived amount of influence or control* the agent has over world. For example, if the agent can make one hundred different actions but the result, *as perceived by the agent*, is always the same, the agent has no control over the world whatsoever. If, on the other hand, the agent can reliably force the world into two states distinguishable by the agent, it has two options and thus two futures to choose from. Empowerment can be seen as the agent's *potential* to change the world, that is, how much the agent could do in principle. This is in general different from the *actual* change the agent inflicts.

In the section 2.4 we will quantify empowerment using Information Theory [3]. Briefly, *empowerment is defined as the capacity of the actuation channel* *of the agent.* The main advantage of using Information Theory for defining empowerment is that the measure is universal in the sense that it does not depend on the task or on the “meaning” of various actions or states.

2.3 The Communication Problem

Here we provide a brief overview of the classical communication problem from Information Theory and define channel capacity for a discrete memoryless channel. For an in depth treatment we refer the reader to [3,4].

There is a sender and a receiver. The sender transmits a signal, denoted by a random variable X, to the receiver, who receives a potentially different signal, denoted by a random variable Y. The communication channel between the sender and the receiver defines how transmitted signals correspond to received signals. In the case of discrete signals the channel can be described by a conditional probability distribution $p(y|x)$.

Given a probability distribution over the transmitted signal, *mutual information* is defined as the amount of information, measured in *bits*, the received signal on the average contains about the transmitted signal. Mutual information can be expressed as a function of the probability distribution over the transmitted signal $p(x)$ and the distribution characterizing the channel $p(y|x)$:

$$I(X;Y) = \sum_{\mathcal{X},\mathcal{Y}} p(y|x)p(x) \log_2 \frac{p(y|x)}{\sum_{\mathcal{X}} p(y|x)p(x)} \ . \tag{1}$$

Channel capacity is defined as the maximum mutual information for the channel over all possible distributions of the transmitted signal:

$$C = \max_{p(x)} I(X;Y) \ . \tag{2}$$

Channel capacity is the maximum amount of information the received signal can contain about the transmitted signal. Thus, mutual information is a function of $p(x)$ and $p(y|x)$, whereas channel capacity is a function of the channel $p(y|x)$ only. Another important difference is that mutual information is symmetric in X and Y and is thus acausal, whereas channel capacity requires complete control over X and is thus asymmetric and causal (cf. [5]).

There exist efficient algorithms to calculate the capacity of an arbitrary discrete channel, for example, the iterative algorithm by Blahut [6].

2.4 Definition of Empowerment

For the sake of simplicity of the argument, let us assume a memoryless agent in a world. Following the information-theoretic approach to modeling perception-action loops described in [7,8] we can split the whole system into the agent's sensor, the agent's actuator and the rest of the system[1] including the environment. The states of sensor, actuator and the rest of the system at different time steps are modeled as random variables (S, A, and R respectively). The perception-action loop connecting these variables is unrolled in time. The pattern of dependencies between these variables can be visualized as a Bayesian network (Fig. 1).

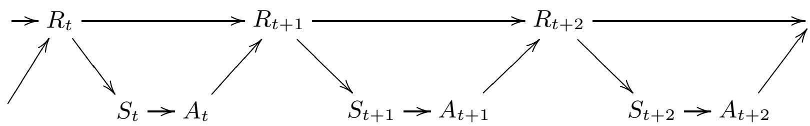


Fig. 1. The perception-action loop as a Bayesian network. S – sensor, A – actuator, R – rest of the system. R is included to formally account for the effects of the actuation on the future sensoric input. R is the state of the actuation channel.

Previously we colloquially defined empowerment as the amount of influence or control the agent has over the world as perceived by the agent. We will now quantify the amount of influence as the amount of Shannon information[2] the agent could "imprint onto" or "inject into" the sensor. Any such information will have to pass through the agent's actuator.

When will the "injected" information reappear in the agent's sensors? In principle, the information could be "smeared" in time. For the sake of simplicity in this paper will be using a special case of empowerment: *n-step sensor empowerment*. Assuming that the agent is allowed to perform any actions for n time steps, what is the *maximum* amount of information it can "inject" into the momentary reading of its sensor after these n time steps (Fig. 2)? The more of the information can be made to appear in the sensor, the more control or influence the agent has over its sensor.

We view the problem as the classical problem of communication from Information Theory [3] as described in Sec. 2.3. We need to measure the maximum amount of information the agent *could* "inject" or transmit into its sensor by performing a sequence of actions of length n. This is precisely the capacity of the channel between the sequence of actions and sensoric input n time steps later.

Let us denote the sequence of n actions taken, starting at step t, as a random variable $A_t^n = (A_t, A_{t+1}, \ldots, A_{t+n-1})$. Let us denote the state of the sensor

---

[1] We include the rest of the system, denoted by R, only to account for the effects of actuation on the future sensoric input. R is the state or memory of the actuation channel. For the problem of channel with side information it is established [4] that knowing the state of the channel may increase its capacity. Thus, in addition to actuator, sensor and the rest of the system it is useful to define *context*, a random variable approximating the state of the actuation channel in a compact form (cf. Information Bottleneck [9], ϵ-machines [10,11]). However, we omit this more general treatment from the present discussion.

[2] The word "information" is always used strictly in the Shannon sense in this paper.

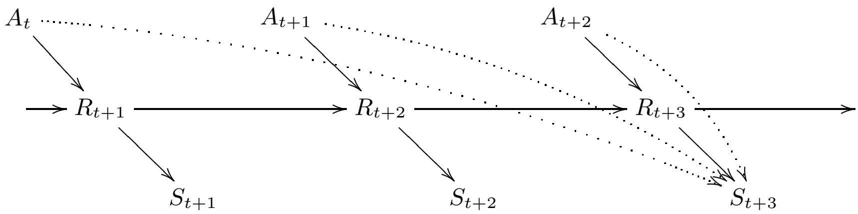


Fig. 2. 3-step sensor empowerment. Actions are independent of system's state (agent with "free will"). The communication channel goes from actions (A_t, A_{t+1}, A_{t+2}) to sensor S_{t+3}.

n time steps later by a random variable S_{t+n}. We now view A_t^n as the transmitted signal and S_{t+n} as the received signal. The system's dynamics induce a conditional probability distribution $p(s_{t+n}|a_t^n)$ between the sequence of actions A_t^n and the state of sensor after n time steps S_{t+n}. This conditional distribution describes the communication channel we need.

We define *empowerment* as the channel capacity of the agent's actuation channel terminating at the sensor (see Eq. 1 and Eq. 2):

$$\mathfrak{E} = C = \max_{p(a_t^n)} \sum_{\mathcal{A}^n,\mathcal{S}} p(s_{t+n}|a_t^n)p(a_t^n)\log_2 \frac{p(s_{t+n}|a_t^n)}{\sum_{\mathcal{A}^n} p(s_{t+n}|a_t^n)p(a_t^n)} . \tag{3}$$

Empowerment is measured in *bits*. It is zero when the agent has no control over what it is sensing, and it is higher the more perceivable control or influence the agent has. Empowerment can also be interpreted as the amount of information the agent could potentially "inject" [8] into the environment via its actuator and later capture via its sensor.

The maximizing distributions over the sequences of actions can be interpreted as distributions of actions the agent should follow in order to inject the maximum amount of information into its sensors after n time steps.

The conditional probability distribution $p(s_{t+n}|a_t^n)$ may induce equivalence classes over the set of sequences of actions. For example, if the various sequences of actions produce only two different outcomes in terms of the resulting probability distribution of sensoric input $p(s_{t+n})$ then the agent may view all the sequences of actions just in terms of two meta-actions corresponding to the two different distributions over the resulting sensoric input.

3 Experiments

In this section we present two experiments to illustrate the concept of empowerment. The first experiment demonstrates how an agent's empowerment looks in a grid world and how it changes when a box is introduced. The second experiment illustrates empowerment of an agent in a maze.

3.1 Box Pushing

Consider a two-dimensional infinite square grid world. An agent can move in the world one step at a time into one of the four adjacent cells. The actuator can perform five actions: go left, right, forward, back, and do nothing. For the sake of simplicity, let's assume that the agent has a sensor which reports the agent's absolute position in the world. What is this agent's n-step empowerment?

For this scenario the n-step empowerment turns out to be the logarithm of the number of different cells the agent can reach in n time steps: $\log_2(2n^2+2n+1)$. This is $\log_2 5$ for 1 step, $\log_2 13$ for 2 steps, and so forth. The empowerment does not depend on where the agent starts with the sequence of actions (Fig. 3, b).

We now add a box occupying a single cell. The agent's sensor, in addition to the agent's position, now also captures the absolute position of the box. Let us assume that the box cannot be moved by the agent and thus remains stationary. If the agent tries to move into the cell occupied by the box the agent remains where it was. In this case the agent's empowerment is lower the closer the agent is to the box (Fig. 3, c). This can be explained by the fact that the box blocks some paths, and as a result it may render unreachable some of the previously reachable cells. Empowerment is high in the box because from there the agent can reach the maximum number of cells including the one occupied by the box.

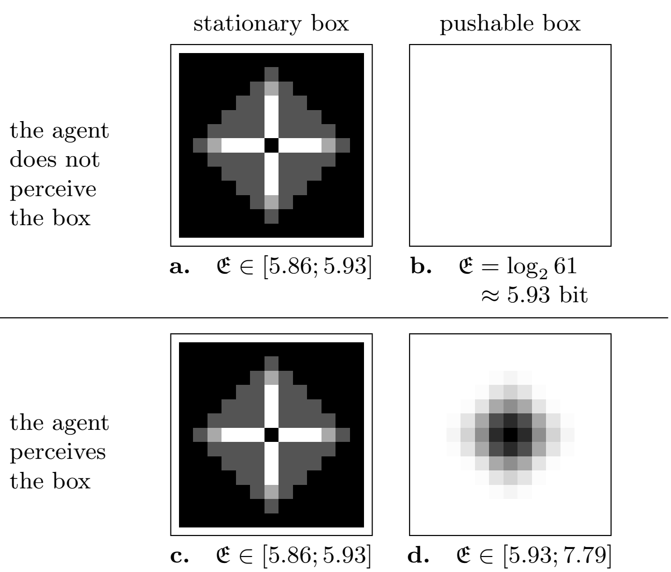


Fig. 3. 5-step empowerment field over the grid. The field is centered at the box. Because empowerment in cells further than 5 cells away from the box is always $\log_2(61) \approx 5.93$ bits, only the 13×13 cells central part of the field is shown. Cells are colored according to scaled empowerment of the agent in the cell. *Darker* color means *higher* empowerment. Maps are scaled independently of each other. Corresponding ranges of empowerment are provided below the maps. Note that the ranges are different in size.

Let us now assume that the box can be pushed by the agent. If the agent tries to move into the cell occupied by the box, it succeeds and the box is pushed in the direction of the agent's move. Empowerment is now more complex than just the number of cells reachable by the agent, because it also includes the position of the box. In this scenario the agent's empowerment in a given cell is the binary logarithm of the number of unique combinations of the agent's and the box's final positions achievable from a given cell. The agent's empowerment is higher the closer the agent is to the box (Fig. 3, d). The number of cells the agent can reach in n time steps is still the same as for the case without the box. However, some paths leading to same cells after n steps can now be differentiated by different positions of the box, because it was pushed differently. Thus, because the position of the box is observable and controllable by the agent, it can be viewed as an extra reservoir for empowerment.

It is also interesting to see what happens if the agent doesn't perceive the box, that is when the sensor captures only the agent's position. In the case of the stationary box, the empowerment field does not change (Fig. 3, a is identical to Fig. 3, c). This is because the position of the box never changes. Excluding it out from the sensor thus cannot decrease the amount of control over the sensor. With a stationary box, a sensor for the box's absolute position is useless. Having no sensor for the box, just by noticing the change in the conditional probability distribution $p(s_{t+n}|a_t^n)$ describing the actuation channel the agent could infer that something changed in the world (no box $\rightarrow$ stationary box).

In the case of the pushable box leaving out the position of the box from the sensor results in the completely flat empowerment field over the grid (Fig. 3, b), exactly as in the initial setup without the box. This is because the movement of the agent and hence its position is not influenced by the box at all. Thus, if the agent doesn't see the box, it cannot perceive it even indirectly.

To summarize, empowerment as a general utility function in this scenario translates to a simple measure of reachability for simple cases (no box, stationary box). Furthermore, it reacts reasonably to changes in the dynamics of the world, which do not need to be explicitly encoded into empowerment. We believe that empowerment discovers intuitively interesting places in the world.

3.2 Maze

Consider a two-dimensional square grid world. Similar to the previous scenario an agent moves in the world one step at a time into one of the four adjacent cells. Some cells have walls between them preventing the agent from moving. A maze is formed using the walls (Fig. 4). The agent has a sensor which captures the agent's global position.

We measure the n-step empowerment of the agent. Similar to the previous scenario, because of deterministic actuation and the nature of the sensor, empowerment is the logarithm of the number of the cells reachable in n moves. Empowerment maps for several time horizons are shown on Fig. 5.

A natural measure for navigation in mazes is the average shortest path from a given cell to any other cell. To navigate through any place in the maze fastest

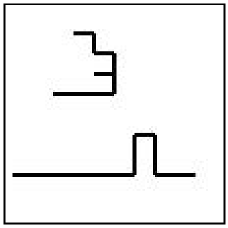
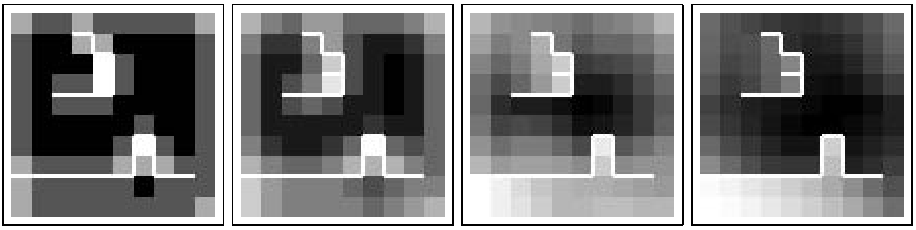


Fig. 4. A 10×10 maze. Walls between cells are shown in black.



$\mathfrak{E} \in [1; 2.32]$ $\mathfrak{E} \in [1.58; 3.70]$ $\mathfrak{E} \in [3.46; 5.52]$ $\mathfrak{E} \in [4.50; 6.41]$

Fig. 5. 1-, 2-, 5-, and 10-step empowerment field over a 10×10 maze (left to right). Walls are shown in white. Cells are colored according to empowerment. *Darker* color corresponds to *higher* empowerment in the cell. Maps are scaled independently of each other. Corresponding empowerment ranges are shown below each map.

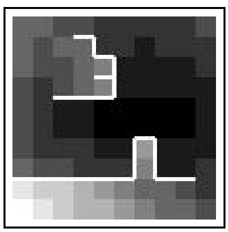


Fig. 6. Map of average shortest distance to other cells. *Darker* color corresponds to *lower* average distance.

one would want to start in a cell with lowest average distance to any other cell. The map of average shortest distances is shown on Fig. 6. It is similar to the map obtained using empowerment with several time steps. In fact, empowerment and average shortest path are roughly anti-correlated (See Fig. 7). However, the two types of maps need not coincide. For instance, if the task were to avoid a predator, the average distance map would not be of much help. However, n-step empowerment with straightforward modifications[3] would implicitly include the effects of the predator into the picture.

[3] A natural way to make the presence of the predator "known" to empowerment is to assume that once the agent is dead, for example, eaten by the predator, all actions have the same effect. As a result, empowerment drops to zero.

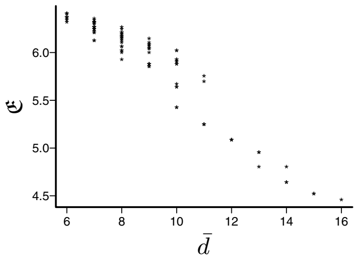


Fig. 7. 10-step empowerment of cells (vertical) vs. the average distance to other cells (horizontal)

4 Discussion and Conclusions

In the search for a general principle for adaptive behavior we have introduced empowerment, a natural and universal quantity derived from an agent's "embodiment", the relation between its sensors and actuators induced by the environment. Empowerment is defined for any agent, regardless of its particular sensorimotor apparatus and the environment, as the information-theoretic capacity of the actuation channel. Empowerment maximization, as a utility or fitness function, can be colloquially summarized as "everything else being equal, keep your options open."

We have shown two simple examples where the empowerment measure captures features of the world which have not and need not be specially encoded. For example, in the box pushing scenario, if the box is pushable the agent is more empowered the closer it is to the box, if the box is not pushable the agent is, vice versa, less empowered the closer it is to the box.

The presence of the box need not be "encoded" into empowerment at all. In both cases empowerment was calculated identically, the sensor and the actuator over which empowerment was measured remained unchanged. It was the dynamics of the world that changed, and empowerment generalized naturally to capture the change. The result was different depending on whether the box was pushable or not.

In the example with walking in a maze, empowerment is anti-correlated with the average shortest distance from a cell to any other cell. However, these two measures will cease to coincide, if, for example, a predator were introduced.

Our central hypothesis is that similar to the two simple examples, where empowerment in most cases was related to the number of reachable cells, empowerment maximization may translate into simpler measures and interpretations, like homeostasis, phototaxis, avoidance, etc.

Empowerment is useful for a number of reasons. Firstly, it is defined universally and independently of a particular agent or its environment. Secondly, it has a simple interpretation – it tells the agent to seek situations where it has control over the world and can perceive the fact. Thirdly, if the agent were to estimate empowerment on-board, it would know what actions lead to what situations in the future – this knowledge could be used for standard planning. Last but not least, empowerment can be calculated on-board in an agent-centric way or externally, as, for example, a fitness function in evolutionary search. In the latter case the agent need not know anything about empowerment – it would just behave as though it maximizes empowerment.

References

1. Ashby, W.R.: An Introduction to Cybernetics. Chapman & Hall Ltd. (1956)
2. Gibson, J.J.: The Ecological Approach to Visual Perception. Houghton Mifflin Company, Boston (1979)
3. Shannon, C.E.: A mathematical theory of communication. The Bell System Technical Journal **27** (1948) 379–423
4. Cover, T.M., Thomas, J.A.: Elements of Information Theory. John Wiley & Sons, Inc. (1991)
5. Pearl, J.: Causality: Models, Reasoning, and Inference. Cambridge University Press (2001)
6. Blahut, R.: Computation of channel capacity and rate distortion functions. IEEE Transactions on Information Theory **18**(4) (1972) 460–473
7. Touchette, H., Lloyd, S.: Information-theoretic approach to the study of control systems. Physica A **331**(1–2) (2004) 140–172
8. Klyubin, A.S., Polani, D., Nehaniv, C.L.: Tracking information flow through the environment: Simple cases of stigmergy. In Pollack, J., Bedau, M., Husbands, P., Ikegami, T., Watson, R.A., eds.: Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems, The MIT Press (2004) 563–568
9. Tishby, N., Pereira, F.C., Bialek, W.: The information bottleneck method. In: Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing. (1999) 368–377
10. Crutchfield, J.P., Young, K.: Inferring statistical complexity. Physical Review Letters **63**(2) (1989) 105–108
11. Shalizi, C.R., Crutchfield, J.P.: Information bottlenecks, causal states, and statistical relavance bases: How to represent relevant information in memoryless transduction. Advances in Complex Systems **5**(1) (2002) 91–95