

Empowerment: A Universal Agent-Centric Measure of Control

Alexander S. Klyubin*, Daniel Polani*[†], and Chrystopher L. Nehaniv*[†]

Adaptive Systems* and Algorithms[†] Research Groups
School of Computer Science, University of Hertfordshire
College Lane, Hatfield, Herts, AL10 9AB, UK

{A.Klyubin, 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 sensorimotor 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. Using two simple experiments we also demonstrate how empowerment influences sensor-actuator evolution.

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. 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, toxin 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

The world is a hostile place. One single wrong action can have disastrous consequences. It can even lead to death. Animals and humans not only survive but also adapt under these circumstances. They all face the classical dilemma of exploration versus exploitation. How do they manage it when errors may be fatal? A standard argument is that evo-

lution biases species towards trying or avoiding certain behaviors in certain situations. However, the population available to evolution is finite and often remarkably small. Evolutionary feedback via death seems to be hardly sufficient.

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

Here are some examples of candidate functions, which, however, are not universal because they are quite specific: sugar concentration around a bacterium, social status of a chimpanzee in a group, money in a bank account of a person. Can these functions be unified?

In his work on ecological approach to visual perception [5] Gibson proposed that animals and humans do not normally view the world in terms of a 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.

In some board games utility may be approximated by what one can afford to do. For example, a breakthrough in strategy occurred in Othello/Reversi when the concept of "mobility" had been discovered. Mobility is the number of moves a player can make in a given situation. The more moves the player can make the better off *on the average* he is. Everything else being equal players should seek higher mobility.

If we look from this perspective on the previously given examples of specific utility functions, they seem to be more related than they appeared at first. To a sugar-feeding bacterium, high sugar concentration means longer survival time and hence more possibilities of moving to different places for reproduction, to a chimpanzee higher social status means more mating choice and interaction, to a person more money means more opportunities and more options.

The common feature of the above examples is the striv-

ing for situations with more options, with more potential for control or influence. To capture this notion quantitatively, as a proper utility function, we need to quantify how much control or influence an animal or human (an agent from now on) has. A fairly straightforward way is to estimate how many actions an agent can perform. However, not all of the actions lead to different results, not to mention the cases when the same action can lead to different results.

Information Theory, originally developed for transmission of information over communication channels by Shannon [11], can be applied to the problem of control and yields quantitative statements [1, 13]. Control can be successfully treated in terms of Shannon information. Recently we demonstrated [6] that information flow optimization can structure systems and, despite the universality of Information Theory, can capture properties of an agent's morphology and the environment.

However, even with Information Theory there is the problem of defining what to look at to determine whether results of actions are different or not. Is turning a page of a book using the left hand different from turning the very same page using the right hand? If one looks only at what page of the book is visible in the end, then the two actions have the same result. If, however, one also takes into the account what one perceives while turning the page, then the results of the two actions are different: when one turns the page using the left hand, one feels and sees that the left hand is moving, similarly, the right hand is perceived to move when turning the page using the right hand. What is the natural way to look at what actions do?

You can move your hand left or right, up or down. Consequently, you may assume you have control over your hand. However, from someone else's perspective you may have other control that you do not suspect. For example, a neurosurgeon may notice that you can make a particular neuron in your brain spike. However, if you never directly or indirectly perceive the result, you will never know that you have control over the neuron.

To make things simpler, similar to Gibson's ecological approach to visual perception [5], we define what an agent does solely in terms of what it perceives. This way it does not matter what other agents perceive about these actions; it does not matter how much control or influence the agent appears to have from a "god's eye view" of the world. The concept of "the environment" becomes a by-product of the interplay between the agent's sensors and actuators. This enables us to base our utility function solely on the sensors and actuators, without the need to refer to the "outside" of the agent.

We build on Gibson's view of perception-action as being central to agents and go even further by: (1) making actuators as important as sensors, and (2) providing *an agent-centric quantification of the amount of control or influence the agent has and perceives*.

In Sec. 2.2 we will provide the quantification using Information Theory. The key concept employed is the channel capacity. We will now briefly provide the necessary background from the Information Theory.

2.1 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 [11, 3].

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 $p(y|x)$:

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

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

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

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 just the conditional probability distribution $p(y|x)$ describing the channel. 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. [8]).

There exist efficient algorithms to calculate the capacity of an arbitrary discrete channel, for example, the iterative algorithm by Blahut [2].

2.2 Definition of Empowerment

Consider an agent in an environment. The agent has a sensor and an actuator. The actions of the agent may reflect the sensoric input, and the sensoric input, in turn, may reflect past actions. The sensor and the actuator are coupled in a perception-action loop [9].

The agent's empowerment is a property of its perception-action loop. Thus, we first model the loop. Its constituent parts are treated as random variables. These are the sensor S , the actuator A , and R – the rest of the system including the environment. We need R to account for the effects of actuation, the agent's morphology and the environment on the sensors. Although we treat the agent as having just one sensor and actuator, S and A can also represent the combined state of some or all of its sensors and actuators.

As in [13, 6, 7] we interpret the perception-action loop in terms of a communication channel-like model. From this perspective actions may “inject” Shannon information into the environment and into the sensors. In order to model the temporal aspects we *unroll* the loop in time by introducing discrete time t . The random variables S_t , A_t , and R_t become the states of the sensor, the actuator, and the rest of the system respectively at different times t . To account for the complete loop we model the dynamics of arbitrary number of time steps, as opposed to [13] where only one time step is modeled.

We model the relations between the variables as a causal Bayesian network [8] which is a directed acyclic graph where any node, given its parents, is conditionally independent from any other node which is not its parent or successor (any node directly or indirectly reachable from the node). In our model this property results in conditional independence from the past.

The pattern of relations between variables at two consecutive time steps is shown in Fig. 1. We assume that the pattern of relations is time-invariant and thus holds for any t . Thus, the graph in Fig. 1 is just a section of the network. The diagram can be read as follows: action A_t is picked given sensor state S_t , sensor state S_t is obtained from the state of the rest of the agent-environment system R_t , R_{t+1} depends only on R_t and A_t .

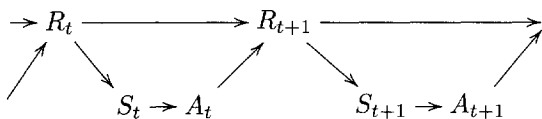


Figure 1: The perception-action loop as a Bayesian network. S – sensor, A – actuator, R – rest of the system. R is included to formally account for the effects of the actuation on the future sensoric input. R is the state of the actuation channel.

So far, we have informally introduced empowerment in terms of the amount of control or influence the agent has over the world and is able to perceive. The only way the agent can impose its influence over the world is via its actuator by performing a sequence of actions.

To illustrate the idea, imagine an agent with “free will”, which at some point t in time can perform an arbitrary sequence of actions of length n . How much information can the agent “inject” via the sequence of action into its sensor at time step $t + n$? The more of the information about the sequence of actions can be made to appear in the sensor, the more control or influence the agent has over its sensor. We redraw the Bayesian network to reflect the fact that the agent is “free” to choose any sequence of actions. The new network is shown in Fig. 2.

We view the situation as the classical problem of communication from Information Theory [11] as described in Sec. 2.1. 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

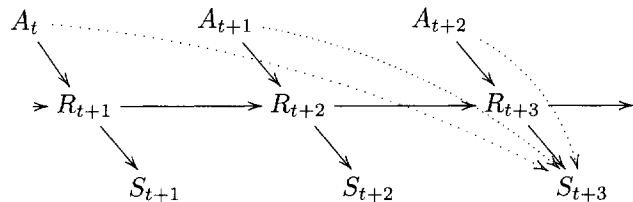


Figure 2: 3-step empowerment. Actions are independent of system’s state (agent with “free will”). The communication channel goes from (A_t, A_{t+1}, A_{t+2}) to S_{t+3} .

actions and sensoric input n time steps later.

Let us denote the sequence of n actions taken starting at time t by a random variable $A_t^n = (A_t, A_{t+1}, \dots, A_{t+n})$, and its instantiation by a_t^n . Let us denote the state of the sensor n time steps later by a random variable S_{t+n} , and its instantiation by 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 can now define *empowerment* as the channel capacity of the agent’s actuation channel terminating at the sensor (see Eq. 1 and Eq. 2):

$$\mathfrak{E}_t = C\left(p(s_{t+n}|a_t^n)\right) = \max_{p(a_t^n)} I(A_t^n; S_{t+n}). \quad (3)$$

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” [7] into the environment via its actuator and later capture via its sensor.

The definition of n -step empowerment above uses only a momentary reading of the sensor at time $t + n$ to see how much of the “injected” information is captured. In general, the momentary reading of the sensor can be replaced by an arbitrary function of the variables downstream of A_t . For example, the function could be the sequence of sensoric states $S_t^n = (S_{t+1}, S_{t+2}, \dots, S_{t+n})$. If the agent has memory, the function could also include the agent’s memory downstream from A_t .

The maximizing distributions over the sequences of actions $p(a_t^n)$ can be interpreted as the distributions of actions the agent should follow in order to inject the maximum amount of information into its sensor after n time steps.

The actuation channel will in general have state due to the specifics of the environment or the agent’s morphology. According to our model (see Fig. 1), the state is R , the rest of the system. For the information-theoretic problem of channel with side information it is established [3] that knowing the state of the channel may increase its capacity. Hence, knowing R could increase empowerment. Accordingly, it is useful to define *context*, a random variable approximating the state of the actuation channel R in a compact form (cf. Information Bottleneck [12], ϵ -

machines [4, 10]). Knowing the context could increase the agent's empowerment. The context could be interpreted as the "environment" relevant to the agent. Moreover, the context can be deduced or constructed from the agent-centric perspective, from the data available "on-board", without the need to resort to global knowledge. However, we omit this more general treatment from the present discussion.

2.3 Maximizing Empowerment

Consider an agent that tries to maximize its empowerment. From the perspective of Information Theory, there are two different ways of doing that: (1) strive to reach the parts of the world where empowerment is highest, that is, where $p(s_{t+n}|a_t^n)$ is less noisy, (2) modify one's sensors and actuators to improve empowerment, that is, modify the sets for which $p(s_{t+n}|a_t^n)$ is used. In this paper we concentrate only on the latter case which is sensor-actuator adaptation or evolution.

If the agent's actuator is fixed (not allowed to adapt or evolve) the sensor may be modified so that the agent "sees" more of the results of its own actions. In other words, the agent's sensor is adapted to better differentiate between outcomes of various actions the agent can take.

If the agent's sensor is fixed, the actuator may be modified so that it can better influence the future sensoric states. In other words, the actuator is adapted to support actions which can be differentiated by the sensor better.

Maximizing empowerment adapts the agent's sensors and actuators to each other. Moreover, it adapts them to the niche in which the agent exists in the world (cf. [5]).

For example, let us assume that there is a cost to sensoric bandwidth. It may be due to energy constraints or due to information-processing constraints. If the agent initially has eyes, but lives in a part of the world where there is no light at all, empowerment, used as a fitness function, will create an evolutionary gradient towards removing the eyes. If, on the other hand, an agent can produce a flash visible to its eyes, empowerment will keep the eyes and the flash-producing action, no matter whether having the capability of producing a flash in the dark is "useful" in the common sense of the word.

The sets of sensoric states and actions available to the agent can be seen as the physical level of its perception-action system. Empowerment builds a logical level on top of that. Actions can be defined through their effects on the sensors (cf. Gibson's approach [5]). For instance, actions with the same effect (as described by $p(s_{t+n}|a_t^n)$) on the sensor may be seen as just one logical action and need not be differentiated, unless other criteria, such as costs of actions, are taken into account. Similarly, sensors may also be seen in terms of how they react to actions. This logical level of description built on top of the physical level may be different in different places in the world. For instance, if one is in a completely dark room, one can disregard one's eyes as a sensor because they capture no effects of one's actions.

We formulate the following hypothesis: evolution adapts the physical level of sensors and actuators to the logical level induced by empowerment for the niche of the species.

Moreover, empowerment maximization will push evolution forward to exploit all existing and newly appearing communication channels.

2.4 Summary

We view the perception-action loop of an agent in terms of random variables describing the state of the agent-environment system. The main parts of the loop are the state of the sensor and the state the actuator. All the variables are linked into a causal Bayesian network. The network allows us to unroll the loop in time.

The agent's actuation channel is a combined probabilistic description of "embodiment", the agent's morphology and environment, in terms of the effects of the agent's actions on its sensors. Empowerment is defined as the information-theoretic capacity of the actuation channel and is measured in bits. Empowerment is zero when the agent has no control over its sensors, and it grows when the control or influence the agent has over its perceptions grows.

The actuation channel in general has state due to the specifics of the environment and the agent's morphology. Knowing certain features of the channel's state may increase the agent's empowerment. These features could be viewed as the context or the "environment" relevant to the agent and can be deduced or constructed from the data available to the agent without the need to resort to global knowledge. Interestingly, these features are projections from the state of the rest of the system and thus capture information about the rest of the system, the "outside" of the agent.

The conditional distribution describing the actuation channel induces a logical structure of the sensor and actuator on top of their physical structure. The logical structure is a potentially more compact description of the agent's actuator in terms of its effects on the sensor, and of the sensor in terms of how it "sees" the actions. For example, actions having the same effect on the sensor may be seen as just one logical action. We hypothesize that natural evolution makes the logical structure and the physical structure adjusted to each other.

3 Sensor and Actuator Evolution Experiments

3.1 The Testbed

Consider an infinite two-dimensional square grid world. A source is located at the center of the grid. The source emits a signal, the strength P of which in any cell of the grid is $P(d) = d^{-2}$, $P(0) = 2$, where d is the Cartesian distance from the source.

An agent moves in the world occupying one cell at a time. Let us assume that the agent has a sensor which samples the strength of the signal in nearby cells and picks the one with highest strength. If there are several such cells, one is picked at random with uniform probability. Let us also assume that at each time step the agent can stay put or jump into one of the nearby cells. We shall now present two scenarios: evolving a sensor for a given actuator, and evolving an actuator for a given sensor.

3.2 Sensor Evolution

Let us assume that the agent's actuator allows it to either stay in the current cell or move into one of the four adjacent cells (von Neumann neighborhood). What is the best layout of the sensor so that the agent's n -step empowerment is maximal?

Let us limit the set of sensors to those finding the cell with highest signal strength near the agent. Each sensor can be thought of as a set of sampling points relative to the agent. For example, a sensor measuring the local gradient using von Neumann neighborhood has four sampling points: $\{(0, -1), (1, 0), (0, 1), (-1, 0)\}$.

3.2.1 Evolutionary Search

To find a good enough sensor we can search in the set of sensors using an evolutionary algorithm. The algorithm treats sensor layout as an individual. The sampling points of any sensor are constrained to lie within a fixed square with side b around the agent. The maximum number of sampling points a sensor can have is fixed. At any point in time, the state of a sensor identifies the sampling point that measures the highest signal strength. Hence, the number of sampling points is also the number of states of the sensor $|\mathcal{S}|$.

The fitness F of a sensor is the 4-step empowerment of the agent equipped with the sensor. The fitness also includes a small penalty for the number of sampling points used $|\mathcal{S}|$:

$$F = \mathfrak{E}_0^4 - \epsilon |\mathcal{S}|. \quad (4)$$

The penalty $\epsilon = 10^{-3}$ is there to prefer more economical sensors with less sampling points. The exact value of the penalty used here is arbitrary, though small.

To evaluate the fitness of a sensor the agent is placed in the world at a predefined position (for instance, in the center which is at $(0, 0)$ in Cartesian coordinates). The conditional probability distribution $p(s_4 | a_0^4)$ is exactly calculated from the Bayesian network [8]. The 4-step empowerment is the capacity of the channel described by the conditional probability distribution. The capacity is found with 10^{-4} bit precision using the iterative algorithm proposed by Blahut in [2].

We initialize the population with five randomly generated sensors. In every generation, five best sensors produce five offspring each. Thus the size of the population is between 5 and 30. Five best individuals from the parents and offspring are selected into the next generation.

An offspring is produced from its parent by mutation. The mutation operator supports two operations: (1) addition of a sampling point, and (2) deletion of an existing sampling point. If the sensor has no sampling points, the mutation operator always adds a point. If the sensor has the maximum number of sampling points, the mutation operator always deletes a point. Otherwise, either a new point is added or an existing one is deleted with equal probability.

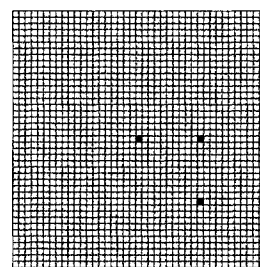
To speed up the search and make it more efficient we have incorporated ideas from simulated annealing and tabu search: (1) the number of mutations performed is uniformly distributed between 1 and $1 + (G \bmod 10)$, where G is the

generation number; and (2) we do not add offspring which have been evaluated before or are already present in the population.

To sample the solution space thoroughly we run the evolutionary algorithm at least 10 times for 1000 generations each. The best individuals are selected across these runs.

3.2.2 Results

We have evolved the sensor for different positions in the world to illustrate the idea that empowerment makes sensors and actuators adapt to the niche in which the agent exists. We have constrained the sensor to a maximum of 20 sampling points which lie inside the square with side 21 around the agent.



Sensors below are evolved for the corresponding places above

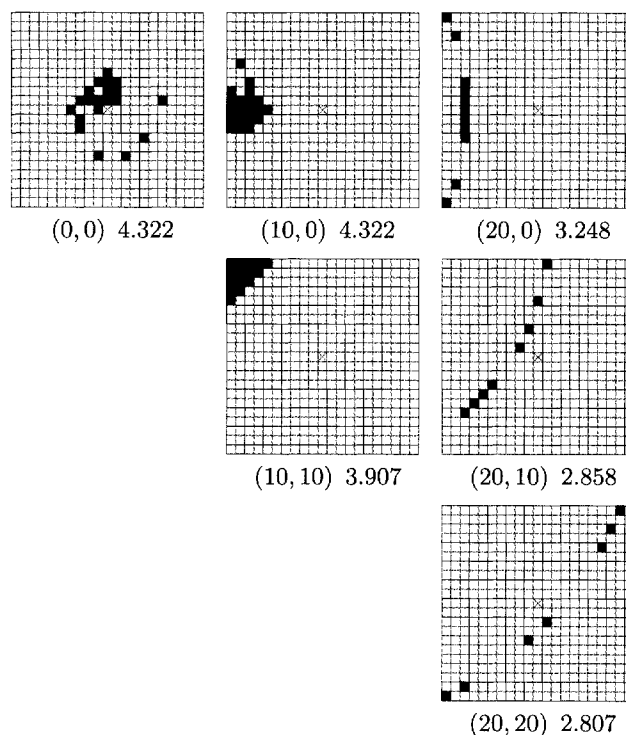


Figure 3: Best sensors evolved for six different places in the world. Maximum number of sampling points $\max |\mathcal{S}| = 20$. All of the points lie inside a square with side $b = 21$ centered the agent. The layout is relative to the agent (center, marked with a cross). Sample points are shown in black. The text under each layout contains the position in the world for which the sensor was evolved and the agent's empowerment with the sensor measured in bits.

Fig. 3 shows the best evolved sensors. Evolution has found two types of sensors. For the areas near the signal

source (Fig. 3 left and middle columns) the sensors capture more or less the absolute displacement from the source. For the areas farther away from the source (Fig. 3 right column) the sensors capture the bearing to the source.

The best sensor layouts for starting positions $(20, 0)$, $(10, 10)$, $(20, 10)$, and $(20, 20)$ are stable in the sense that they stay almost exactly the same if the evolutionary experiment is repeated. The other two layouts, $(0, 0)$ and $(10, 0)$, are much more prone to change while retaining exactly the same empowerment and hence fitness. This suggests that they might belong to a plateau of the fitness landscape. Qualitatively though these layouts still comprise a cluster of sampling points as seen in Fig. 3. The cluster is centered at the agent for the $(0, 0)$ starting position, and on the far left for the $(10, 0)$.

3.3 Actuator Evolution

In this section we show what actuators are optimal in terms of empowerment for a given sensor. The experiment is similar to the sensor evolution experiment described above. The same world is used. The same fitness function is used. The same evolutionary algorithm is used. However, in this experiment the agent has a 16-state sensor shown in Fig. 4. The sensor captures a coarse bearing to the signal source. The actuator is evolved to maximize empowerment.

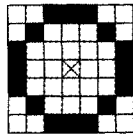
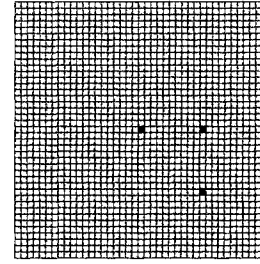


Figure 4: The layout of the fixed sensor used for actuator evolution.

We have constrained the actuator to a maximum of 10 actions which move the agent no further than 4 cells horizontally and 4 cells vertically. The best actuators evolved in 10 runs of 1000 generations each are shown in Fig. 5. When we rerun the experiment evolution settles on different sets of best actuators. However, they retain identical empowerment and fitness. The variety in best evolved actuators is much larger compared to best evolved sensors from the previous section. Apparently, many more actuator layouts are optimal. There is thus less evolutionary pressure on the actuators in this experiment.

Near the center of the world (Fig. 5 left and center columns) it suffices to have just four actions to completely control the sensor after four time steps. When the agent starts farther away it has to cover much longer distances in order to significantly change the bearing to the source and, hence, the sensoric reading. As the step length is limited to 4, there is not enough time to cover the necessary distance. The agent thus has less control over its sensor. For example, it cannot move to the area left from the center. The evolutionary algorithm guided by empowerment “recognizes” this situation (due to nonzero ϵ in Eq. 4) and prunes the set of actions down to just two or three, which are enough to move in the right half of the world.



Actuators below are evolved for the corresponding places above

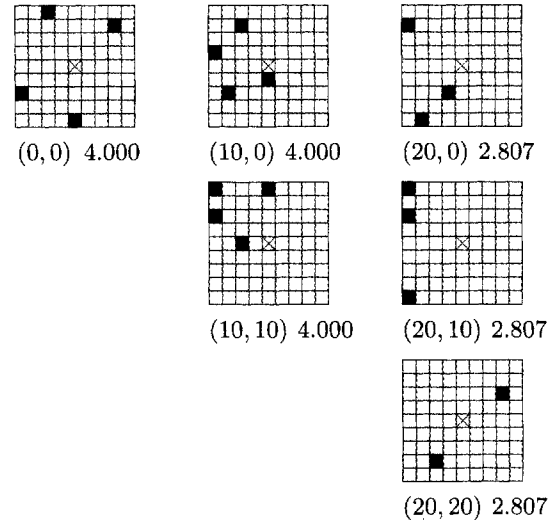


Figure 5: Best actuators evolved for six different places in the world. Maximum number of action $\max |A| = 10$. The displacement generated by any action is constrained to lie inside a square with side $b = 9$ centered at the agent. The layout is relative to the agent (center, marked with a cross). Places the actuator can move the agent to are shown in black. The text under each layout contains the position in the world for which the actuator was evolved and the agent’s empowerment with the actuator measured in bits.

We have also tried allowing the agent to jump farther in one go. In this case, empowerment for all of the tested six places in the world reaches the maximum of 4 bits, which is constrained by the capacity of the sensor. With longer jumps only four actions are sufficient to reach the maximum empowerment in each of the six places tried.

3.4 Summary of the Experiments

We have presented a simple testbed to demonstrate how empowerment could be used for sensor and actuator evolution. We have used an evolutionary algorithm to maximize empowerment by evolving a sensor for a given actuator, and by evolving an actuator for a given sensor. We have demonstrated that: (1) using empowerment as a fitness function is possible, and (2) empowerment may allow evolution to switch from one representation of information to another one (absolute displacement vs. coarse bearing in the sensor evolution experiment).

4 Discussion & Conclusions

In search for a general principle for adaptive behavior we have introduced empowerment, a natural and universal quantity based on an agent's "embodiment", the relation between its sensors and actuators induced by the environment and the agent's morphology.

Empowerment is defined for any agent, regardless of its particular sensorimotor apparatus and the environment, as the information-theoretic capacity of the actuation channel terminating at the sensors. Empowerment 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" into the environment via its actuators and later capture via its sensors. Thus, empowerment maximization can be colloquially summarized as "everything else being equal, keep your options open." Empowerment quantifies what the agent can *potentially* do, not what it *actually* does.

Empowerment is based on the interplay between the agent's sensors and actuators. The concept of the "outside" may arise as a result of maximizing empowerment. This agent-centric view is based on that of Gibson [5] where the agent sees the world in terms of what it can perceive and do. In this work we go further by: (1) treating actions and perceptions using the same formalism of Information Theory, as parts of the agent's perception-action loop, and (2) providing a quantitative formalization of the degree of control or influence the agent has and is able to perceive.

Gibson in [5] was against applying Information Theory to the agent's perceptions, his main argument being that the environment does not intentionally communicate any Shannon information to observers.¹ However, Ashby's work [1] and also more recent work [13] show that Information Theory can be successfully applied to perception-action without the need to assume any intentionality. We believe the concept of empowerment, which is in the spirit of Gibson and at the same time is quantified information-theoretically, also contributes to reconciling perception-action with Information Theory.

To illustrate how empowerment influences evolution we have carried out two simple experiments. In the first experiment we have evolved sensors to maximize empowerment for a given actuator. The evolved sensors were adapted to the niche for which they were evolved. In the second experiment we have evolved actuators to maximize empowerment for a given sensor. Again, the actuators evolved were specific and "meaningful" to the niche of the agent. Empowerment, used as a fitness function for morphology or "embodiment"

evolution, can adapt the agent to its niche. However, as opposed to the majority of other fitness functions, empowerment is agent-centric and universal. It is suitable for any sensorimotor apparatus, it is calculated from the sensor and actuator data available to the agent, and it does not require one to externally assign meaning to the states of sensors and actuators. This, we believe, is its greatest power.

Empowerment is useful for a number of reasons. Firstly, it is defined to be universal, and the definition is independent of a particular agent or its environment. Secondly, empowerment maximization has a simple interpretation – it tells the agent to seek situations where it has perceivable control or influence over the world, where it can change the world most. 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 or homeostasis. 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 – as a result of evolution the agent would just behave as though it maximizes empowerment.

Acknowledgments

We would like to thank the Condor Team from the University of Wisconsin, whose High Throughput Computing system Condor enabled us to conveniently run large numbers of simulations on ordinary workstations.

We are also grateful to the anonymous reviewers of this paper for useful and constructive feedback.

Bibliography

- [1] W. R. Ashby. *An Introduction to Cybernetics*. Chapman & Hall Ltd., 1956.
- [2] R. Blahut. Computation of channel capacity and rate distortion functions. *IEEE Transactions on Information Theory*, 18(4):460–473, Jul 1972.
- [3] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.
- [4] J. P. Crutchfield and K. Young. Inferring statistical complexity. *Physical Review Letters*, 63(2):105–108, 1989.
- [5] J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin Company, Boston, 1979.
- [6] A. S. Klyubin, D. Polani, and C. L. Nehaniv. Organization of the information flow in the perception-action loop of evolved agents. In R. S. Zebulum, D. Gwaltney, G. Hornby, D. Keymeulen, J. Lohn, and A. Stoica, editors, *Proceedings of 2004 NASA/DoD Conference on Evolvable Hardware*, pages 177–180. IEEE Computer Society, 2004.

¹"The term information cannot have its familiar meaning of knowledge communicated to a receiver. . . The only recourse is to ask the reader to remember that picking up information is not to be thought of as a case of communicating. The world does not speak to the observer. . . It [information] is simply there. The assumption that information can be transmitted and the assumption that it can be stored are appropriate for the theory of communication, not for the theory of perception. . . The information for perception, unhappily, cannot be defined and measured as Claude Shannon's information can be." – J. J. Gibson, *The Ecological Approach to Visual Perception*, [5, p. 242]

- [7] A. S. Klyubin, D. Polani, and C. L. Nehaniv. Tracking information flow through the environment: Simple cases of stigmergy. In J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. A. Watson, editors, *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, pages 563–568. The MIT Press, 2004.
- [8] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2001.
- [9] W. T. Powers. *Behavior: The Control of Perception*. Wildwood House Ltd., 1974.
- [10] C. R. Shalizi and J. P. Crutchfield. Information bottlenecks, causal states, and statistical relevance bases: How to represent relevant information in memoryless transduction. *Advances in Complex Systems*, 5(1):91–95, Mar 2002.
- [11] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, Jul 1948.
- [12] N. Tishby, F. C. Pereira, and W. Bialek. The information bottleneck method. In *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, pages 368–377, Sept. 1999.
- [13] H. Touchette and S. Lloyd. Information-theoretic approach to the study of control systems. *Physica A*, 331(1–2):140–172, Jan. 2004.