

Original Article

Dennett, Darwin, and Skinner Crows

Frans Blommaert, Capacity Group Human Technology Interaction, Faculty of Technology Management, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

Ruud Janssen, Capacity Group Human Technology Interaction, Faculty of Technology Management, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. Email: R.Janssen@tm.tue.nl

Abstract: The central theme of this paper is the scientific viewpoint taken for understanding behavioral processes. Two classical viewpoints are formulated by Dennett (the intentional stance) and Tinbergen (Tinbergen's four questions). In this paper we argue that the two different viewpoints are linked to the two different processes that underlie complex behavior, namely, the instruction process and the selection process. To zoom in on the similarities and differences between these processes, we model wheel dropping behavior of Northwestern crows as observed by Zach (1978, 1979) from the two different viewpoints: (1) with crows that possess intentional faculties (called Dennett crows), and (2) with crows that possess selectional faculties. The latter type of crows is further divided into a population that is able to adapt over generations only by natural selection (Darwin crows), and a population that, apart from natural selection, is also able to adapt using operant learning (Skinner crows). Salient outcomes are that these two populations need markedly different times to adapt to changes in the environment, and that operant learning needs a value system that is an internal equivalent of the fitness criterion. In conclusion, we propose that understanding behavior should start at a meta-level with identifying whether the nature of the behavioral process under study is intentional or selectional.

Keywords: behavior, prediction, explanation, stance, instruction, selection, intention, rationality, fitness, adaptation, evolution.

Introduction

Marr (1982) claimed that we only fully understand processes in general, and

behavioral processes in particular, if we understand them at different levels of abstraction. For behavioral processes, many proposals for levels differing in degree of abstraction can be found in the literature (see, e.g., Marr, 1982; Dennett, 1981; Newell, 1990; Leont'ev, 1981). The proposals have much in common and can roughly be summarized as consisting of:

1. *A do-level or action level.* At this level it can be observed what the organism is doing at a certain point in time.
2. *An activity level.* Often, actions are only meaningful within a more integrated context of related actions (see, e.g., Norman, 1981).
3. *A semantic level.* What does the activity mean for the organism in terms of its existence.

Behavioral studies can be performed in many different ways. If we take Marr's (1982) claim as a starting point, it is meaningful to distinguish between bottom-up research methods and top-down research methods (Dennett, 1998). Bottom-up research methods start at the action level and try to derive knowledge at higher levels of abstraction. This is, for instance, the accepted research method in ethology (see, e.g., Krebs and Davies, 1993), and behaviorism (Skinner, 1953). Top-down research starts with assumptions at the semantic level (e.g., about goals) and subsequently tries to specify the lower levels of abstraction. Good examples are the computational approach to vision of Marr (1982) and the unified theories of cognition from Newell (1990).

Dennett (1981, 1988) noted that a standard way to interpret human and animal behavior is *as if* it is intentional, and he coined this 'the intentional stance'. It assumes that behavior is goal-directed and optimal with respect to its constraints (rationality). The reason why we use this stance in daily life ('folk psychology') and in the life sciences is that it is an extremely powerful heuristic for predicting complex behavior of biological organisms. There is however one problem, namely that much of the ontogeny and behavior of biological organisms is *not* intentional. As Dennett (1998) himself remarks: "If in retrospect we can identify a goal that has been optimally or sub-optimally achieved by the evolutionary design process, this is something of a misrepresentation of history."

Some scientists have noted that we know of only two alternative processes that result in the sort of complex behavior that we see in animals and humans (Dawkins, 1989, 1990; Simon, 1996; Johnston, 1999). Johnston distinguishes these as (1) the instruction process, and (2) the selection process. In the instruction process there is a specification of the goal, and rational behavior to realize it, just as it is assumed in the intentional stance. In the selection process, on the other hand, there is a generator of behavioral alternatives plus a selection mechanism to choose between them. What is typically lacking in the latter process is foresight into the final result of the behavior.

Results of both processes are often quite similar, which explains why the intentional stance is such a powerful predictor, even if the process considered is selectional in nature. The reason why their results are quite similar is that both processes have a tendency to converge to ‘optimal’ behavior. In the instructional process this is caused by the rationality assumption. In the selection process this is caused by the fact that the selection criterion is often related to fitness in the sense of Darwinian evolution.

Even though the intentional approach is often a good predictor of the outcome of a selection process, it is not a satisfactory means for understanding the selection process *mechanisms*. To this end a selectional stance (interpreting the process at hand as a selection process) would be more suitable. Hence, the aim of this paper is to explore the possibilities of using a generic selectional approach to human and animal behavior, if appropriate. To this end we show, alongside, two interpretations of animal behavior: the instructional (or intentional) interpretation and a selectional interpretation. As an example, we re-interpret experimental results of whelk dropping behavior of Northwestern crows as published by Zach (1978, 1979). In the conclusion of the paper, we try to formulate generic properties of a selectional approach to behavior.

2. Dennett crows

2.1. The intentional stance

We will use a pure form of reverse engineering as described by Dennett (1998). The behavior of crows may be interpreted in two ways:

1. *Foraging behavior.* In this interpretation we focus on the foraging process and assume that the goal of the process is to take in sufficient energy to enable sustained life. Since in the intentional stance it is assumed that goal-directed behavior is ‘rational’, it follows that crows use a ‘rational’ way of foraging. This can be operationalized by assuming that crows try to optimize their foraging behavior in terms of an optimality criterion, for instance, expended time or energy, rate of net energy gain, etc.
2. *Self-regulation.* A broader interpretation is that intentional systems use a self-regulation process to optimize their behavior (see, e.g., Carver and Scheier, 1998; Fischer, Blommaert, and Midden, 2004). The goal of such a process would be to optimize *any* behavior of the organism. In case of foraging behavior this would lead to demands for effectivity and efficiency of the foraging process.

Here we will use the first interpretation and will focus on the foraging process, where the properties of foraging are constrained by the rationality demand. We interpret foraging behavior at three levels of abstraction:

1. *The goal level.* At this level we rigorously define the goal of foraging as taking in sufficient energy to sustain life. We assume that crows have knowledge about food in general and about the particularities of food that is most suited to reach the goal. Furthermore we assume that crows have knowledge about how to access this food (e.g., whelks can be eaten after breaking their shells). Last but not least crows should possess rationality, enabling them to select behavior that is optimal with respect to a set of constraints.
2. *The level of strategy.* At this level choices are made about the type of food (e.g., whelks), how to access this food (e.g., by dropping them from a certain height), which particular whelks are most suited (e.g., large whelks), which particular dropping height is most suited, and so on. The choices are made on the basis of rational comparisons (computations) between the different options.
3. *The level of actions.* Here the crows simply perform the task according to the choices made at the more abstract levels, by searching for a suitable whelk, dropping it repeatedly to break its shell, etc.

2.2. Computations

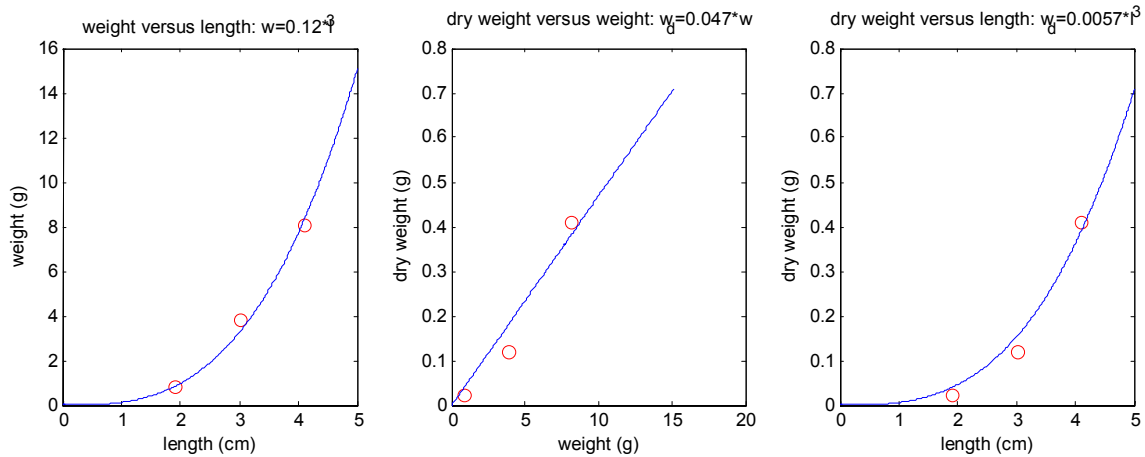
2.2.1. Whelk regarded as a source of energy

We start by assuming that a crow considers a whelk simply as a source of energy. How much energy, then, does a whelk represent to a crow? According to Zach (1979), the caloric conversion rate of a whelk is 4.98 kcal per gram of dry weight. We made several fits to Zach's whelk data (see Fig. 1), and found that dry weight w_d is a fixed fraction of about 4.7% of total whelk weight w . Since the crow's assimilation efficiency is approximately 70% (Zach, 1979), a whelk of weight w would represent to the crow an energy quantity

$$Q = 70\% \times 4.98 \text{ kcal/gram} \times 4.7\% \times w \quad (1)$$

This energy should at least compensate for the energy required to find and open the whelk, otherwise there would be a net energy loss to the whole foraging process.

Fig. 1: Fits (lines) to the whelk data (circles) given by Zach (1978, Table 1; 1979, pages 111-112). In these fits we have assumed a cubic function relating whelk length (l) to whelk weight (w), and a fixed fraction relating whelk weight (w) to dry weight (w_d). From these fits we conclude that dry weight is a fixed fraction of 4.7% of total whelk weight.



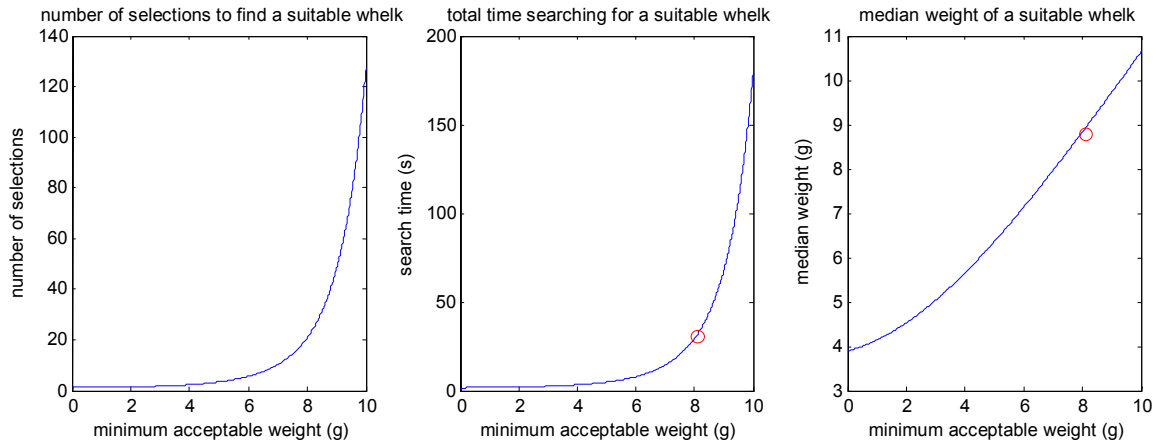
2.2.2. Searching for a suitable whelk

Not all whelks are equally suitable. Small whelks represent less energy to the crow, and they also break less easily when dropped (Zach, 1979). For this reason only whelks with a weight exceeding a certain minimum may be suitable to accept for dropping. We will therefore assume that a crow will select a whelk for dropping only when its weight exceeds some minimum acceptable weight, and that the crow will reject all other whelks. However, depending on the availability of whelks, large whelks may be rare to find. Choosing this minimum acceptable whelk weight too high may then result in extremely long search times.

To model the process of searching for a suitable whelk, we have assumed that a crow will search for a whelk, selecting one whelk at a time and rejecting it when its weight is below a certain minimum acceptable weight. Fig. 2 (left) shows calculations for the expected number of whelks a crow needs to select to find a suitable whelk versus the minimum acceptable weight, Fig. 2 (middle) shows the corresponding time needed to search for a suitable whelk, and Fig. 2 (right) shows the expected weight of a suitable whelk. The figures were obtained by assuming a Normal distribution for the weights of the available whelks; for details see the caption of Fig. 2.

Fig. 2: Model calculations for the process of searching for a suitable whelk. Left: Using Zach's data (1978, Table 2), we assume a Normal distribution for the weights of the available whelks, with mean and standard deviation of 3.58 and 2.66 grams, respectively. The figure shows the expected number of whelks a crow needs to select to find a suitable whelk versus the minimum acceptable weight. Middle: Zach (1978, Table 3) found that crows would only accept 'large' whelks, where large whelks had an average weight of 8.08 grams. If we substitute this value for the minimum acceptable weight, we find that crows need to select 22.1 whelks to find a suitable whelk. Since Zach (1979, Table 1) found that crows spend on average 31.05 seconds looking for a suitable whelk, we find that crows need 1.4 seconds per selection. The

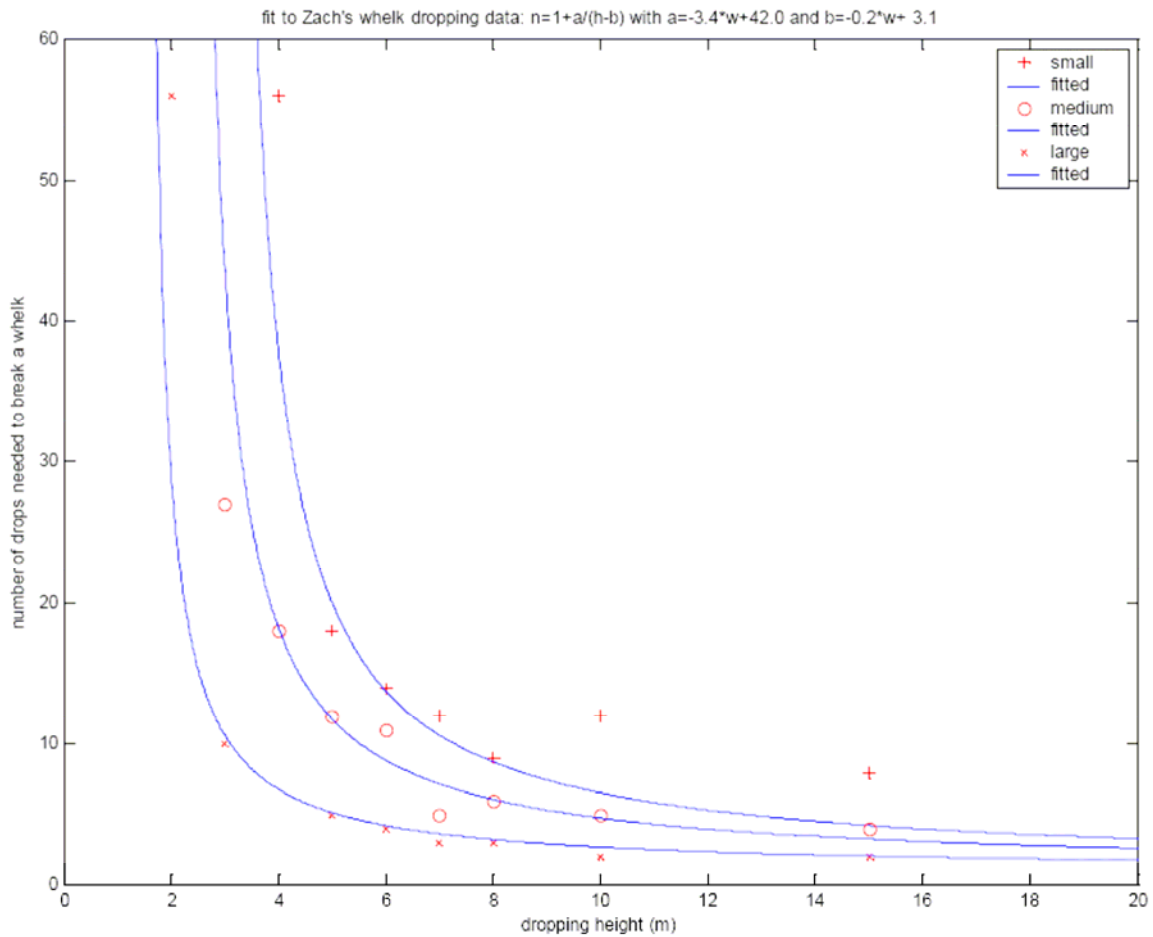
figure shows the resulting fit for total search time versus minimum acceptable weight. Right: This figure shows the median weight of all suitable whelks versus the minimum acceptable weight. For reasons of computational efficiency we will substitute this for the expected weight of a suitable whelk. Zach (1978, Table 2) found that the average weight of a suitable whelk was 8.80 grams; this is shown in the figure by the circle.



2.2.3. Dropping a whelk to break its shell

When a crow has found a suitable whelk, it will drop it repeatedly from a certain height to break its shell and then eat it. Zach (1979) has performed experiments to estimate the average number of drops needed to break a whelk. He did this for small, medium, and large whelks and for several dropping heights between 2 and 15 meters. For our computations we made one fit to the data for all three whelk sizes; it is shown in Fig. 3.

Fig. 3: Our fit to the whelk dropping data given by Zach (1979, Fig. 2). We assume the following relation between dropping height (h) and number of drops needed to break a whelk (N): $N = 1 + a / (h - b)$, where the parameters a and b are given by $a = c_1 \times w + c_2$ and $b = c_3 \times w + c_4$. The resulting fit is shown for small, medium, and large whelks.



2.2.4. Time and energy budgets for the whole process

The whole foraging process consists of seven activities; they are listed in Table 1. Each activity requires time and energy to complete it, and so the time and energy to complete the whole foraging process is simply the sum of all individual activities.

Table 1: Required time and energy for the whole process of searching for a whelk, dropping it until it breaks, and eating it. Required energy is given in units of one BMR (Basic Metabolic Rate, or 0.85 cal/sec). The table is derived from Zach (1979, Table 1), with the following modifications: (a) The time needed to search for a suitable whelk is assumed to be proportional to the number of selections needed to find a suitable whelk. The table therefore gives the normalized value for one selection given in Fig. 2. (b) The time needed for a drop is assumed to be proportional to dropping height. The table therefore gives the normalized value for a dropping height of one meter, obtained by dividing Zach's value by the average dropping height of

5.23 meters. (c) The time needed for handling between drops is also assumed to be proportional to dropping height, since dropping a whelk from a larger height will result in a longer time to search for it on the dropping grounds. The table therefore gives the normalized value for a dropping height of one meter.

Activity	Required time (seconds)	Required energy (BMR)
Flight to beach	7.0	9
Search for whelk	1.4 ^(a)	3
Flight to dropping site	4.0	9
Each drop	0.9 ^(b)	9
Handling between drops	4.8 ^(c)	2
Extraction of whelk	43.2	2
Flight back to perch	7.7	9

Using the fits of the previous sections, Table 1 allows us to estimate the total expended time and energy for the whole foraging process. Given the choices for minimum acceptable whelk weight and dropping height, we can calculate search time and expected suitable whelk weight (Fig. 2), dry weight and the amount of energy gained by eating the whelk (Fig. 1), and the number of drops needed to break the whelk (Fig. 3). Using Table 1 we then calculate total expended time and energy.

2.2.5. Optimality criteria and optimal behavior

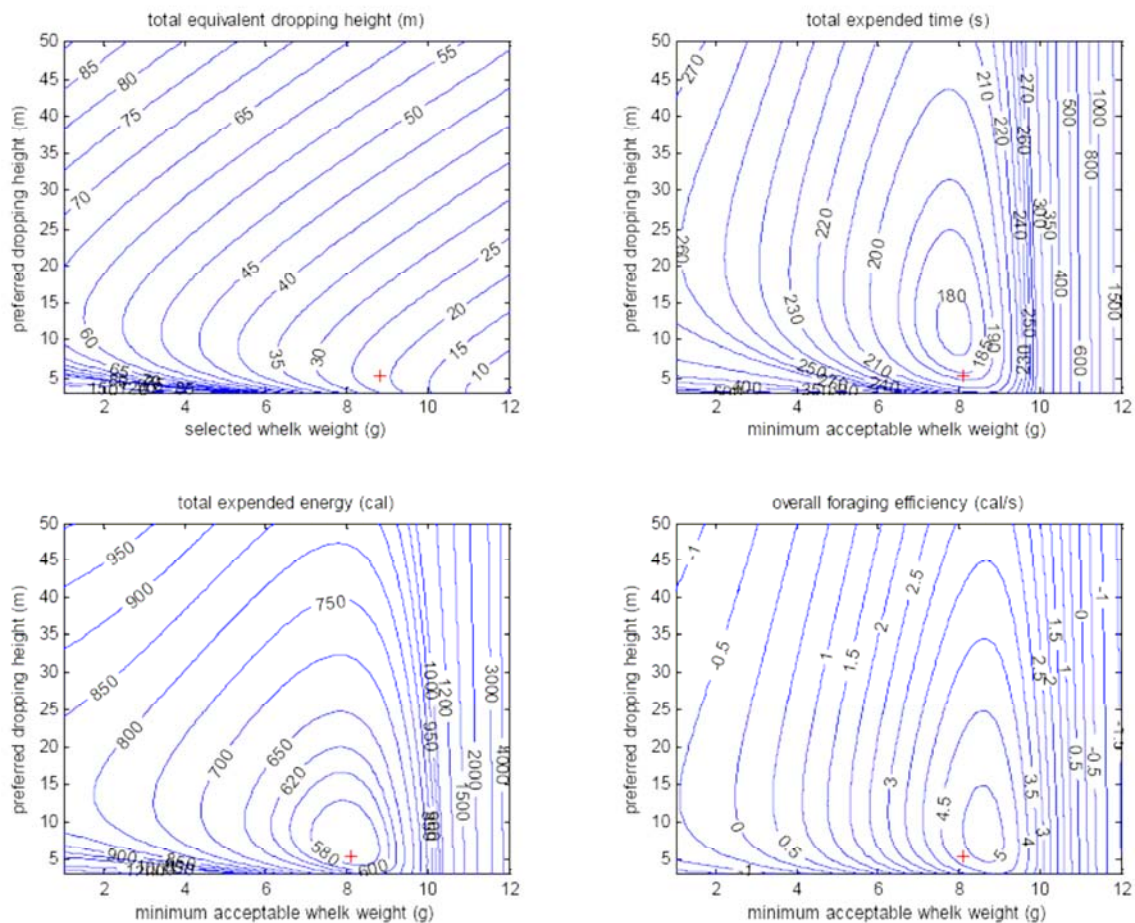
Zach (1978, 1979) found that Northwestern crows selected whelks with an average weight of 8.80 grams and subsequently dropped them from an average height of 5.23 meters. On average, whelks required 4.36 drops to break. Zach (1979) concluded that the average dropping height of 5.23 meters could be explained by assuming that crows minimized the total equivalent dropping height (that is, the average dropping height multiplied by the average number of drops needed to break the whelk) for ‘large’ whelks.

A Dennett crow can choose many different optimality criteria by which to evaluate its foraging behavior, for instance, the number of drops needed to open the whelk, total expended time or energy, net energy gain, etc. Here we consider four ‘reasonable’ criteria: total equivalent dropping height as proposed by Zach (1979), total expended time, total expended energy, and foraging efficiency (rate of net energy gain); see Fig. 4. Note that the last three optimality criteria lead to similar choices for minimum acceptable whelk weight and preferred dropping height: minimum acceptable whelk weights of 7.9, 8.1, and 8.7 grams, respectively, and preferred dropping heights of 11.6, 7.3, and 8.2 meters, respectively.

There is, however, a lot of arbitrariness in the way we modeled whelk dropping behavior and what behavioral and environmental parameters we took into account. From Zach’s observational data alone it is therefore difficult to conclude

which optimality criterion might actually be ‘used’ by the Northwestern crows.

Fig. 4: Four optimality criteria that a Dennett crow might use to evaluate its own whelk dropping behavior. Top left: Total equivalent dropping height, i.e., dropping height multiplied by the expected number of drops needed to break the whelk. This is the criterion used by Zach (1979). Top right: Total expended time (T) for the whole process of searching for a whelk, dropping it until it breaks, and eating it. Bottom left: Total expended energy (E), again for the whole process. Bottom right: Foraging efficiency, i.e., the rate of net energy gain $(Q - E) / T$, also for the whole process. In all four figures the red plus indicates the values observed by Zach (1978, 1979).



2.3. Reflections on results

A major problem that the intentional approach has to deal with is the interpretation of ‘intentions’ and ‘rationality’. First, it is not obvious what the

rationality assumption actually means for a specific behavior, and hence what should be the optimality criterion. Surely, if ‘rational’ would only mean ‘optimize some criterion’, then *any* behavioral strategy can in principle be described under the rationality assumption. However, if the underlying process would be identified as a selection process, the interpretation of rationality might in a more straightforward way be derived from the selection mechanism (survival value, or fitness) of the process.

Second, a strict interpretation of rationality would require that crows ‘know’ all the consequences of their behavior, which would – implausibly – require logical omniscience (Cherniak, 1999). To study real phenomena, less stringent definitions of rationality have been proposed, such as bounded rationality (Simon, 1990).

Third, some scientists will argue that the intentional approach is inadequate for scientific theorizing. For instance, as one of the reviewers of this paper remarked, “[...] we have no idea what an intention is materially or how an intention could cause behavior” and “[...] we have no warrant to suppose that any analog to intention exists in the brain”. This is indeed a very fundamental issue, but as the opinions on this issue vary widely (for an introduction see Wilson, 1999) and as the debate is still continuing, we have chosen to leave this outside of the scope of this paper.

Notwithstanding the above problems, our computational results show that the intentional approach actually works quite well in *predicting* the foraging behavior of crows. Apparently, crows have found a sort of optimal behavior. What the intentional stance does not do is to suggest how crows achieve this. They certainly do not achieve it by using the logic and knowledge of physics as we have done here (and Zach in his papers).

So, as the intentional approach does not use the correct mechanism for predicting behavior, it fails to predict particular aspects of the process, such as, how the behavior of the crow converges to optimal behavior, or the amount of variability in the optimal dropping height as observed by Zach.

3. Darwin crows

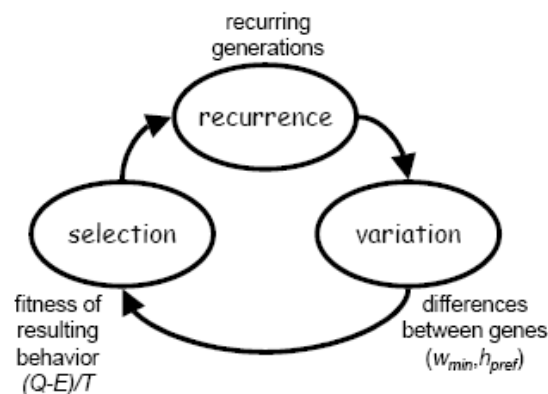
3.1. Darwinian evolution as a selection process

All selection processes (also called evolutionary processes) consist of three basic elements: recurrence, variation, and selection (Baum, 2000). Darwinian evolution is usually chosen as the paradigmatic example of a selection process. Its basic elements are often denoted by replication, mutation, and selection (after Dawkins, 1989, 1990). In this paper we will model Darwinian evolution according to the functional properties of the three basic elements, and neglect the precise implementation in terms of DNA, cross-over, mutations, and so on. We simply treat Darwinian evolution as another selection process.

The basic elements of a selection process have the following properties (see also Fig. 5):

1. *Recurrence*. This is the tendency for something to reappear – repeatedly, and with variations – in a population. The paradigmatic example of recurrence is the reproduction of biological organisms, but other instances (such as the recurrence of a piece of behavior during the lifetime of an organism, or the recurrence of an idea or piece of music within a culture) also exist (Baum, 2000; Dawkins, 1989). For the case of our Darwin crows, recurrence is related to a single crow that is born as a child of two parent crows, and in its turn may become a parent of child crows. Properties of Darwin crows (i.e., the properties that relate to the whelk dropping behavior of a crow) are somehow transferred from parents to offspring.
2. *Variation*. Variation occurs within the population due to the existence of variants. These variants are defined by differences in their environmental effects (Baum, 2000). For the example of Darwin crows, we assume the existence of two ‘genes’ that relate to whelk dropping behavior, one for minimum acceptable whelk weight (w_{min}) and one for preferred dropping height (h_{pref}). The variants differ only in the precise values for these two genes, and hence in the whelk dropping behavior that they will display.
3. *Selection*. Selection occurs when recurrence differs for each of the variants (i.e., when differences in environmental effects cause differences in recurrence) and evolutionary change is the result of changes in the relative frequencies of the variants within the population (Baum, 2000). For instance, in Darwinian evolution selection is related to the reproductive success (or fitness) of the variants. In the example of Darwin crows we assume that foraging efficiency of whelk dropping behavior is determining the fitness of an individual crow. The general argument is that crows that forage efficiently have (1) more time to spend on other activities, and/or (2) have more energy to spend on other activities.

Fig. 5: The basic elements of a (Darwinian) selection process: recurrence, variation, and selection.



Note that in our calculations for the Darwin crows, the genotype of the crow (i.e., the set of ‘genetic’ characteristics of the crow as inherited from its parents) corresponds to (w_{min}, h_{pref}) , and the phenotype (the physical and behavioral characteristics of the crow) corresponds to the actual foraging behavior of the crow. Fitness corresponds to the foraging efficiency of the phenotype, *not* the genotype.

3.2. Further specification of the selection process

We assume that Darwin crows have a minimum acceptable whelk weight w_{min} and a preferred dropping height h_{pref} , which do not change during their lifetime. The minimum acceptable whelk weight and preferred dropping height are independent and vary over the individual crows. In fact, we assume that these two behavioral parameters are represented in the genes of individual crows and are inherited from both parents. Minimum acceptable whelk weight and preferred dropping height of offspring crows will be assigned random values between the respective values of their parents. Hence, the gene pool consists of genes for minimum acceptable whelk weight and preferred dropping height, while the variants are represented by the different values occurring in the gene pool.

Selection is in terms of fitness. Some whelk weights and dropping heights yield a higher foraging efficiency than others. We simply relate reproductive success monotonically to foraging efficiency, based on the following reflection.

The foraging process of a crow may stop for two different reasons:

1. The crow ‘decides’ that it has taken in a sufficient net amount of energy, Q_0 . In this case we assume that crows that have shorter foraging times will have a higher fitness. If it takes a time T and an energy E to open a whelk with an energy equivalent of Q , the total foraging time equals

$$T_0 = Q_0 / (Q - E) \times T = T / (Q - E) \times Q_0, \quad (2)$$
 which is minimal if the foraging efficiency $(Q - E) / T$ is maximal.
2. The alternative is that for some reason an external trigger stops the foraging process. Examples are the appearance of a predator or a human. In this case we assume that crows that have taken in more energy have a higher fitness. If the process is stopped at time T_0 , the crow has, on average, taken in a net amount of energy

$$Q_0 = T_0 / T \times (Q - E) = (Q - E) / T \times T_0, \quad (3)$$
 which is maximal if again the foraging efficiency $(Q - E) / T$ is maximal.

On the basis of these considerations, we use foraging efficiency $(Q - E) / T$ as a quantitative measure that is monotonically related to fitness.

3.3. Computations

We start the computations with a population of 100 crows. Every crow has a

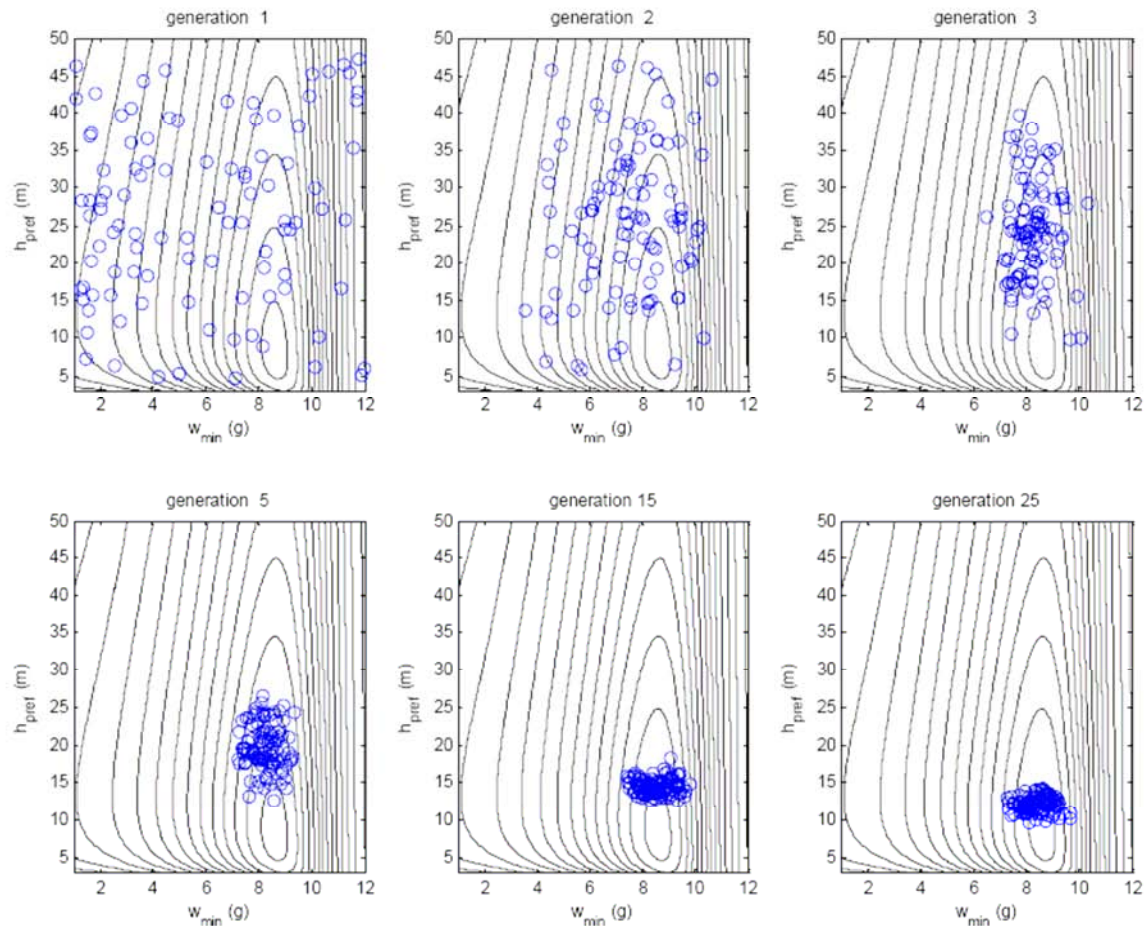
minimum acceptable whelk weight w_{min} somewhere between 1 and 12 grams and a preferred dropping height h_{pref} somewhere between 3 and 50 meters represented in its genes. At the start of the first generation, the distributions of the values w_{min} and h_{pref} are uniform random between these extreme values.

For the simulations we use 25 crow generations. The crows of each generation get 100 foraging trials during which their fitness is measured. In each trial, a small variability of 10% is added to the genetically determined values for w_{min} and h_{pref} to simulate imprecisions in the choices that the crow eventually makes for these parameters. Each trial proceeds as follows:

- Each crow first searches for a suitable whelk, rejecting all whelks until it finds a whelk with a weight equal to or exceeding the crow's minimum acceptable whelk weight. To estimate the total search time and the expected weight of the whelk that the crow finally finds, we use the fits we made for the Dennett crows (Fig. 2), and add small variability of 10% to the results.
- Next, the crow will repeatedly drop the whelk from its preferred dropping height, until the whelk breaks. The expected number of drops needed to break the selected whelk is a function of the whelk's weight and the dropping height. Here, too, we use the fits we made for the Dennett crows (Fig. 3) and add a small variability of 10% to the results.
- Using the time and energy budgets given in Table 1 we calculate total expended time and energy for the whole foraging process, and from these the overall foraging efficiency.

After 100 foraging trials we estimate the crow's fitness using the foraging efficiency averaged over all 100 trials. We allow only the 50% fittest crows to survive before the breeding season starts. Neglecting differences between male and female crows, we randomly assign all surviving crows to couples, and let each couple have an offspring of four baby crows. The genes (w_{min}, h_{pref}) of the baby crows are uniform random distributed between the extremes defined by the values of the two parents. To prevent that variations are being 'swamped out' by this blending inheritance, we add an extra variability of 10% to the resulting values, and let the parent crows die.

Fig. 6: The genotypes for the 1st, 2nd, 3rd, 5th, 15th, and 25th generation, showing the evolution of Darwin crows. Each circle represents the genotype (w_{min}, h_{pref}) of one Darwin crow. The contour lines in the background represent the fitness landscape, i.e., the two-dimensional plot of foraging efficiency versus the parameters w_{min} and h_{pref} (see also Fig. 4, bottom right). As expected, the Darwin crows evolve over generations towards 'optimal' behavior.



3.4. Reflections on results

The selection process works in the sense that it enables to evolve a population of crows that behaves close to the optimum efficiency. Furthermore it demonstrates the evolution of crow behavior over time, as reflected in the gene pool evolution shown in Fig. 6. The particular gene pool convergence we find depends on the specific way we implemented the selection process: the specific properties of recurrence, variation, and selection. In case observational data on the time series of crow behavior were available, the model properties might be adapted in such a way that the evolution of the theoretical gene pool would fit that of the observed gene pool.

Notice that there is a lot of arbitrariness in the initial state of the gene pool we have chosen, as there is in all artificial life research (Langton, 1989). The evolution of crow behavior in principle starts with the introduction of all life on earth, which

eventually evolved into the collection of living creatures that exists today. Somewhere in the evolutionary path crow-like creatures existed that were the forefathers and -mothers of the Northwestern crows that we are modeling in this paper. We have decided to start the evolutionary process with crows that were already exactly the same in appearance, properties, and behavior as the crows that Zach studied. The only thing that was lacking was the optimization of their whelk dropping behavior. We also might have started the evolutionary process at the point where crows did not even have discovered whelk-dropping behavior. Notice also that we do not even know whether the chosen initial state really existed in the evolution of crows.

There seem to be two fundamentally different sub-processes in evolution of behavior. The first sub-process is the ‘invention’ of *new* behavior; in our case this is the invention that whelk dropping is an effective means to open whelks. The second sub-process is the *optimization* of already existing behavior. What we have shown here is that the second sub-process can be modeled well by the selectional stance. It seems much more difficult to model the first sub-process. One way would be to define a gene pool for playful or explorative behaviors. But how can this be done? A possible answer may be given by Darwinian theories on insight (see, e.g., Simonton, 1996).

We have modeled whelk-dropping behavior for a static situation, since we took a whelk population that did not change over time. This is probably not realistic: the availability and properties of whelks probably depends on seasons, and will vary from year to year. Since optimal whelk dropping depends on the properties of the whelk population, the most adaptive solution varies over time. This problem is the subject of the next section.

4. Skinner crows

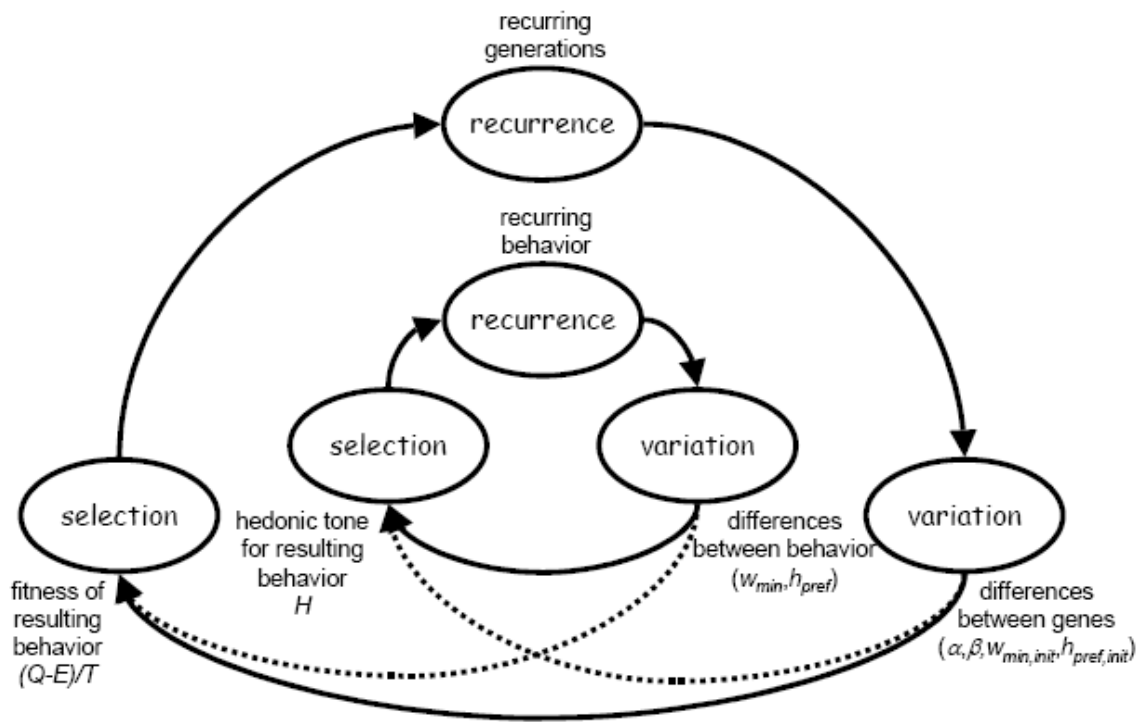
4.1. Two nested selection processes

Adaptation of behavior to the properties of an environment takes time. This is not very important when the properties of the environment are static, but usually they are not. In that case, adaptation is always lagging behind in time: organisms are adapted to an environmental state of the past. Hence, there is fitness increase in speeding up adaptation. In the case of crows it is obvious that, on the one hand, the properties of the available whelks determine which choices for minimum acceptable whelk weight and preferred dropping height are optimal. On the other hand, it is improbable that the properties of whelks, such as available weights and shell hardness, remain the same throughout seasons and throughout years.

In the case of Darwin crows the adaptation of choices for whelk weight and dropping height is very slow since these choices cannot change during the lifetime of a crow. Hence, it is probable that Darwin crows are sub-optimally adapted to their environment. In general there is gain in speeding up adaptation since higher

adaptation velocities succeed better in following environmental change. A way in which adaptation can be speeded up is by using a nesting of selection processes as suggested by Johnston (1999). This nesting enables crows to adapt to environmental changes that occur within their lifetime. The situation is depicted in Fig. 7.

Fig 7: Speeding up adaptation by nesting two selection processes. In the case of our crows, the outer loop ('evolution') acts on generations and the inner loop ('operant learning') acts on foraging trials. The two dotted arrows show the interaction between the outer and inner loops (see text for further details).



The outer loop is comparable to the selection process of the Darwin crows: there is recurrence through the birth of baby crows. Skinner crows differ from Darwin crows by the presence of a faster, nested adaptation process. This inner loop is again a selection process. Recurrence within this loop is the recurrence of a piece of behavior. This piece of behavior may change over the lifetime of a crow; its properties are perceived by the crows and evaluated by their hedonic system. Crows thus select optimal behavior on the basis of the highest anticipated pleasure.

There are two interactions between the inner and outer loops; they are shown by the two dotted arrows in Fig. 7. First, the properties of the hedonic system (the parameters α and β in Fig. 7) determine which behavior will be selected by the crow in the inner loop (i.e., to what values the whelk dropping parameters w_{min} and h_{pref} will converge). Second, the behavior selected in the inner loop in turn determines the crow's fitness in the outer loop and, as a result, the selection of the properties of the crow's hedonic system (i.e., to what values the hedonic system parameters α and β will converge).

Note that in our calculations for the Skinner crows, the genotype of a crow now corresponds to the properties of the crow's hedonic system (α , β) and the *initial* values for the whelk dropping behavior (w_{min}, h_{pref}). The phenotype corresponds to the actual foraging behavior as selected by the crow during its lifetime. Fitness again corresponds to the foraging efficiency of the phenotype.

4.2. Further specification of the nested selection process

During their lifetime, the crows carry out the inner loop of the adaptation process. Hence, the evaluation process of the inner loop resides within the crow's brain. To evaluate its behavior, the crow uses its perceptual systems. For the actual computations, we use the following four perceptions:

1. The perception of whelk weight w ;
2. The perception of dropping height h ;
3. The perception of T , the time spent on the whole foraging process;
4. The perception of E , the energy spent during the foraging process.

The rationale of using these perceptions is that, on the one hand, the perceptions should be rich enough to enable evaluations of behavior with respect to fitness. On the other hand, using more perceptions than necessary causes inefficient use of computational resources.

To compare different costs, a common currency must be used. The literature gives different proposals for this currency, such as pleasure (Cabanac, 1992), somatic marker strength (Damasio, 1994, 1999), or hedonic tone (Johnston, 1999). We follow Johnston's proposal and express costs and benefits in strengths of hedonic tones:

1. The perception of whelk weight leads to a hedonic tone
 $H_w = \alpha_w \cdot (w / w_0) \wedge \beta_w;$ (4)
2. The perception of dropping height leads to a hedonic tone
 $H_h = -\alpha_h \cdot (h / h_0) \wedge \beta_h;$ (5)
3. The perception of expended time leads to a hedonic tone
 $H_T = -\alpha_T \cdot (T / T_0) \wedge \beta_T;$ (6)
4. And the perception of expended energy leads to a hedonic tone
 $H_E = -\alpha_E \cdot (E / E_0) \wedge \beta_E.$ (7)

For the rationale behind using power functions, see Stevens (1957). The constants w_0 , h_0 , T_0 , and E_0 are the same for all crows; we use them only to scale the four perceptions to approximately the same range.

Note that only whelk weight has a positive hedonic tone since it represents a benefit to the crow (energy intake) while the other four terms have a negative hedonic tone since they represent costs. Evaluation of behavior is now simply carried out by summation of all hedonic tones according to:

$$H = H_w + H_h + H_E + H_T. \quad (8)$$

The transformation of the perceptions to a hedonic tone depends on the values of the α 's and β 's and is therefore specific for a certain crow. It eventually determines what choices a crow is going to make for the behavioral parameters minimum acceptable whelk weight (w_{min}) and preferred dropping height (h_{pref}). The four α and β values and the *initial* values for w_{min} and h_{pref} are represented independently in the genes of the Skinner crows. The outer loop eventually decides on the fitness of these values, just as it decides on the fitness of the w_{min} and h_{pref} values in the genes of the Darwin crows.

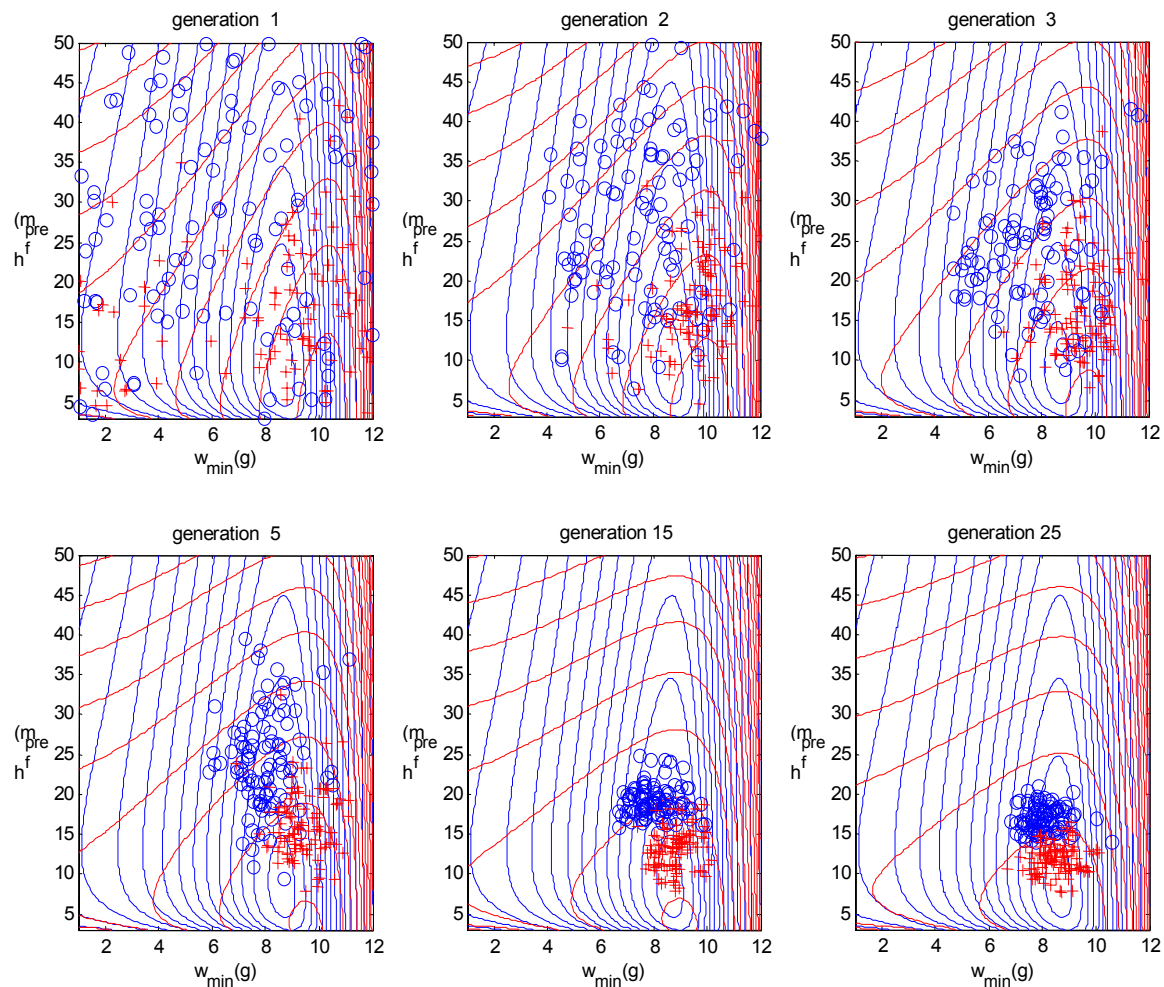
4.3. Computations

We start the computations again with a population of 100 crows. Every crow has four independent values for the α 's somewhere between 0 and 1 and four independent values for the β 's somewhere between 0.3 and 3 represented in its genes. Furthermore, every crow has initial values for minimum acceptable whelk weight w_{min} somewhere between 1 and 12 grams and preferred dropping height h_{pref} somewhere between 3 and 50 meters represented in its genes. At the start of the first generation, the distributions of these α , β , w_{min} and h_{pref} values are uniform random between their respective extreme values.

For the simulations we again use 25 generations. The crows of each generation get 100 foraging trials during which their fitness is measured. The only difference with the Darwin crows is that, during their lifetime, each Skinner crow optimizes minimum acceptable whelk weight and preferred dropping height using its hedonic system. To do this, the crow 'remembers' the best choices for w_{min} and h_{pref} it has made so far (i.e., the choices leading to the highest hedonic tone so far), and uses these values instead of the genetic values as the basis for the next trial.

After 100 foraging trials we estimate the crow's fitness using the foraging efficiency averaged over all 100 trials. We allow only the 50% fittest crows to survive, randomly assign them to couples, and let each couple have an offspring of four baby crows. The genes (α , β , $w_{min,init}$, $h_{pref,init}$) of the baby crows are uniform random distributed between the extremes defined by the values of the two parents. To prevent swamping, we add an extra variability of 10% to the resulting values, and let the parent crows die.

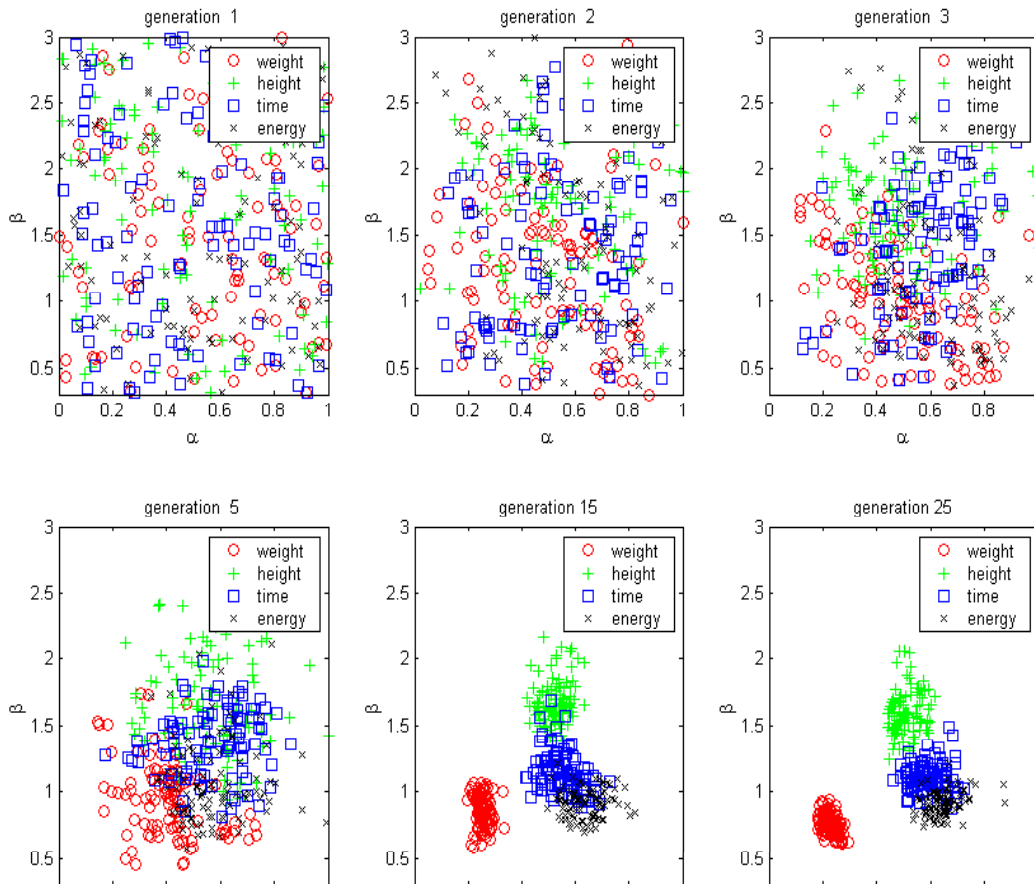
Fig. 8: The genotypes for the 1st, 2nd, 3rd, 5th, 15th, and 25th generation, showing the evolution of Skinner crows. Each blue circle represents the initial (or genotype) values for w_{min} and h_{pref} of a Skinner crow, and each red plus represents the values for w_{min} and h_{pref} that the crow has ‘learned’ after 100 foraging trials. The blue contour lines in the background represent the fitness landscape, and the red contour lines represent the hedonic landscape, i.e., the two-dimensional plot of hedonic tone versus the parameters w_{min} and h_{pref} . The figure shows that the hedonic landscape evolves to a local approximation of the fitness landscape, facilitating an adaptation toward ‘optimal’ behavior within each crow’s lifetime.



4.4. Reflections on results

The hedonic system of the Skinner crows evolves over generations of crows: the α 's and β 's evolve in such a way that the hedonic system becomes a local approximation of the fitness criterion (see Fig. 8). If the properties of the food supply in the environment change, the optimum of the fitness criterion, to which the hedonic system has to adapt, will change as well. In practice this means that more variation of the food supply in the environment generally leads to a less local approximation of the hedonic system to the fitness criterion.

Fig. 9: The (α, β) genotypes of the hedonic system for the 1st, 2nd, 3rd, 5th, 15th, and 25th generation, showing the evolution of the hedonic system of the Skinner crows. This evolution is reflected in the hedonic landscape (the red contours in Fig. 8).



We found that fitting the hedonic system parameters to the fitness criterion is an ill-posed problem: there is no unique solution for the values of the α and β parameters, and hence more than a single combination of these parameters yields a hedonic landscape that is a good local approximation of the fitness landscape. This, however, does not seem to be a problem at all. If the environmental changes put heavier constraints on the adaptation of the hedonic system, its parameter values will probably converge better.

4.5. Computations: speed of adaptation

To test the hypothesis that Skinner crows will adapt faster than Darwin crows to a changing environment, we simulated a stepwise change in the whelk weight distribution. Before the stepwise change we assumed a mean and standard deviation of the weight distribution of the available whelks of 3.58 and 2.66 grams, respectively, and after the stepwise change we used 1.58 and 2.66 grams for these values. The adaptation of the gene pool to the new situation is shown in Fig. 10 for the Darwin crows and in Fig. 11 for the Skinner crows.

Fig. 10: Adaptation of Darwin crows after a stepwise change in the weight distribution of the whelk population. The new fitness landscape is shown by the contours in the background. After five generations the Darwin crows are completely adapted to the new situation.

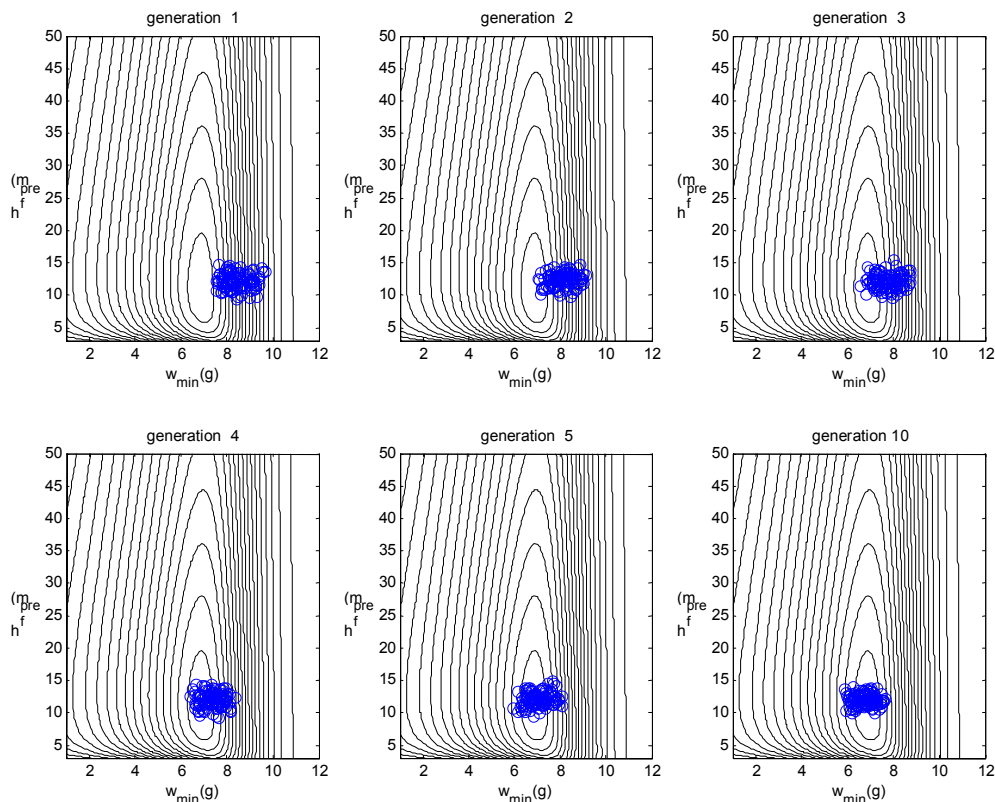
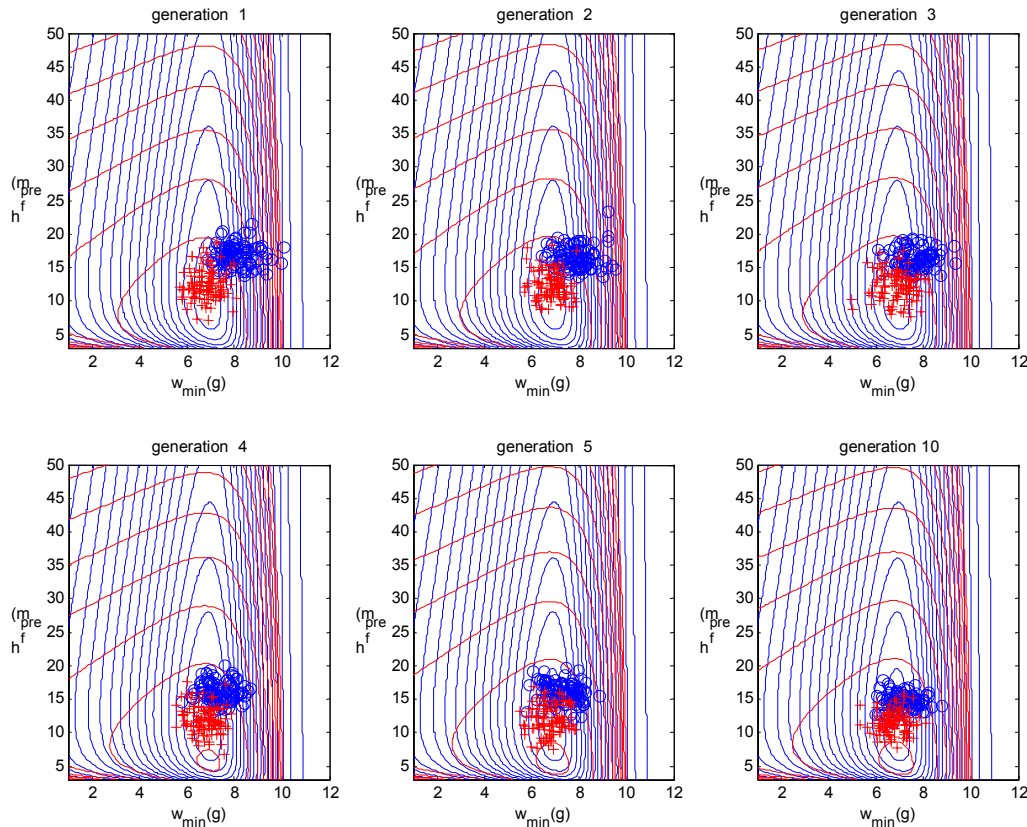


Fig. 11: Adaptation of Skinner crows after a stepwise change in the weight distribution of the whelk population. Skinner crows are able to adapt to the new situation *within* the lifetime of the first generation.

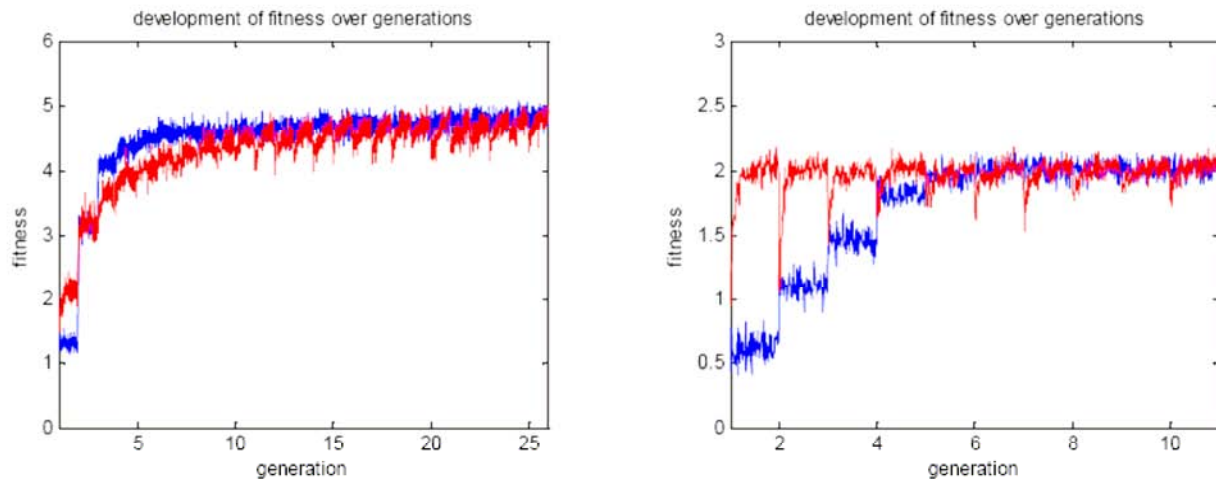


4.6. Reflections on results

Skinner crows indeed adapt faster than Darwin crows to a changing environment, since they are able to adapt *within* their lifetime to changes of the food supply in their environment. Fig. 12 makes this point very explicit by showing the development of fitness of the Darwin and Skinner crows (averaged over the respective populations) over time. To successfully adapt their behavior, Skinner crows need a value system (i.e., their hedonic system) that is an internal version of the fitness criterion. The internal evaluation is carried out by (1) the perception of suitably different aspects of the foraging behavior, and (2) the evaluation of a combination of these perceptions.

Fig. 12: Development of the average fitness of the Darwin and Skinner crow

populations (Darwin crows in blue, Skinner crows in red). Left: adaptation to the initial situation. Right: adaptation after a stepwise change in the weight distribution of the whelk population.



5. Selectional approach to behavior

5.1. Understanding behavioral processes

To start discussing approaches to understanding human and animal behavior, we use a generic state space description of events that encompasses both intentional and selection processes. Even physical processes, such as diffusion, can be well interpreted in state space.

Behavior of an organism can clearly be represented by a begin state, an end state, and the path in state space of behavioral acts which transform the begin state into the end state. In the case of our crows, the begin state would be a crow with an empty stomach, the end state would be a crow with a full stomach, and the path would be the whelk dropping behavior itself. The meta-problem we are discussing here is when, in science, we talk about fully understanding the represented behavior. It is clearly not enough to only know the observational facts (like in a movie). There needs to be more knowledge, but which? We will discuss two answers from different branches of the life sciences (see also Fig. 13 and Fig. 14).

Fig. 13: Representation of behavioral events in state space according to the intentional stance, consisting of a known begin state, a rigorous assumption about the desired end state (the goal or intention), and the rationality assumption. These three elements lead to one optimal path, except in the case of an ill-posed problem.

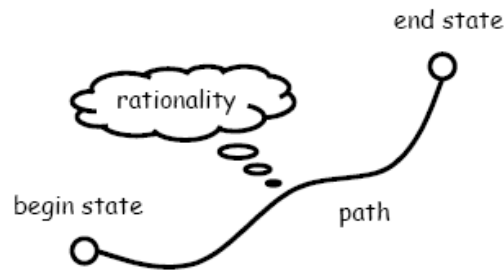
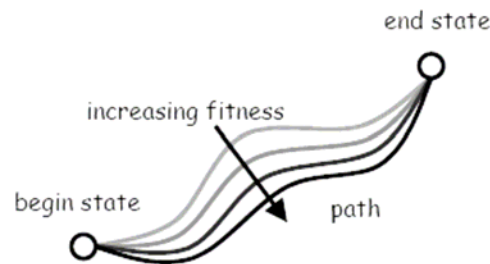


Fig. 14: Representation of behavioral events in state space according to the selectional stance, consisting of a known begin state, a known end state, and an assumed fitness criterion. The selection process will converge the path to an optimal path, even when the problem is ill posed.



From the intentional point of view we take the proposition of Dennett (1998), which is very close to that of Marr (1982) in vision research, and Newell (1990) in cognitive science. Dennett proposes that behavior is only well understood if it is understood from different viewpoints or at different levels of abstraction:

1. *The intentional stance.* What is the goal of the behavior (this specifies the end state in Fig. 13)? How can it be achieved in a rational way ('rationality' in Fig. 13; this assumption constrains the collection of possible paths in Fig. 13)?
2. *The design stance.* What logic can carry out the required processes?
3. *The physical stance.* How can this logic be implemented in hardware (or in wetware)?

From the ethological point of view we have the answer of Tinbergen (1963).

Tinbergen proposes that behavior of animal and man is only well understood if the answers to four questions are known:

1. *Causation*. How does it work?
2. *Survival value*. What is it good for (this specifies the fitness criterion in Fig. 14)?
3. *Ontogeny*. How did the behavior develop in the individual? Is it innate and/or learned (this specifies the development to the optimal path in Fig. 14)?
4. *Evolution*. How did the behavior develop in the species (*idem*)?

The propositions of Dennett and Tinbergen obviously have a lot in common: both contain, besides required knowledge of the observational facts (in Tinbergen's causation question, and in Dennett's physical stance), required knowledge at higher levels of abstraction. There should apparently be knowledge on the benefit of the behavior (in Tinbergen's survival value, and at Dennett's goal level). Additionally, there should be information on the mechanism or logic of how the benefit is achieved (in Tinbergen's ontogeny and evolution questions, and in Dennett's design stance). There is, however, also a striking difference between Tinbergen's and Dennett's proposals: the first requires understanding of how the behavior has *developed* in the individual and in the species (ontogeny and evolution), while in the latter there is no requirement in this direction at all. How can this difference be explained?

Probably this has to do with the fact that Tinbergen and Dennett have different *concepts* for the behavioral processes under study. The concept of Tinbergen is typically that of a process which has developed by evolution in the same way as, for instance, organs have evolved, while the behavioral concept of Dennett is that of intentional behavior, which is intended, logically and rationally derived by using the cognitive system, and after that carried out in adequate motoric actions. The concept of Tinbergen typically relates to behavior of animal and man that is often repeated, more or less automatic, and hence sort of stereotypical. Good examples are the mating behavior of sticklebacks, and the walking behavior of man. The concept of Dennett, on the other hand, is that of intentional problem-solving behavior in man. A prototypical example might be the engineer's design of a new artifact. Hence, it is the concept of a process that in general has not been carried out before and might never be repeated, and that uses the neo-cortex and foresight. In this case the question of how the behavior developed in the individual or in the species is either trivial or unanswerable: trivial in the sense that the behavior always originates in the multifunctional general problem solver that we call the brain, and unanswerable in the sense that there is clearly too little scientific knowledge on brain processes to give a more specific answer. The concept of Tinbergen on the other hand deals with processes that did not originate by willful design but evolved by natural selection. Here it does make sense to ask the question of how the behavior developed in the species and the individual.

The two concepts are of course the concepts for the selection process

(Tinbergen) and the instruction process (Dennett), respectively. Is it possible to observe whether behavior is shaped by selection or by instruction? Certainly not by observing only a single act, but probably by studying the ontogeny of the behavior. It seems obvious that behavior that is perfected by repetition is in fact shaped by a selection process.

5.2. The “Tower of Generate-and-Test”

For understanding the meta-problem better, it probably also helps to use another concept of Dennett, namely, the “Tower of Generate-and-Test”. Dennett (1995) proposed this concept to distinguish between different adaptation processes that evolved in animals and humans. At the base level of the Tower, creatures are only adapted to their environment by genetically imprinted behavior. These are the Darwinian creatures. On the next level of the Tower creatures are also able to use operant conditioning, enabling them to learn adaptive behavior during their lifetime by a ‘trial-and-error’ selection process. These are the Skinner creatures. We explored these two levels of the Tower in our calculations for the Darwin and Skinner crows, however, Dennett proposed two additional levels. On the third level of the Tower, creatures are able to ‘simulate’ the trial-and-error selection process entirely inside their brains (he dubbed these Popperian creatures) while, finally, at the top level of the Tower creatures are also able to pass on learned adaptive behavior to one another (Gregorian creatures). Needless to say, going up in the tower we find increasingly efficient ways to select adaptive behavior. The whole Tower in fact corresponds to a set of four nested selection processes, similar to what is shown in Fig. 7 for the first and second level only.

The creatures inhabiting the first two levels of the Tower typically do not have a neo cortex, and hence are unable to use intentional processes to shape or learn adaptive behavior. All behavior must then be the result of shaping by selection. This holds for the behavior of a lot of species, including all plants and trees, and animals such as reptiles, fish, birds, and small mammals. Popperian creatures may vary in the development of their neo cortex, but they at least possess one. Probably intention and foresight is present in at least some of these creatures. Animals in this category are higher mammals, such as apes and dolphins, and some birds, such as crows. Eventually, only humans are thought to be true Gregorian creatures, although primates may also possess Gregorian qualities (de Waal, 2001).

If we are searching for intentional behavior then, obviously, we only have to consider humans and higher mammals. It should be noted, however, that these organisms not only use the intentional process for shaping their behavior, but also the Darwinian and Skinnerian processes. Therefore, we can easily be misled – by knowing that humans possess intentional capabilities – into thinking that all human behavior is intentional. In particular, by only observing the displayed behavior of a Popperian (or Gregorian) creature at a single point in time we cannot distinguish whether it is carrying out behavior that follows from an intentional process or

whether it has first rapidly ‘simulated’ a selectional process before carrying out the selected behavior. In such cases, the distinction between the selectional and intentional stance becomes blurred (e.g., is the intentional stance implicitly selectional, is rationality or fitness constraining the observed behavior, etc.?), and the greater predictive *efficiency* of the intentional stance may then be a reason to favor it over the selectional stance.

5.3. What did we learn from our computations?

Looking back at the computations for Dennett, Darwin, and Skinner crows, we basically learned three things:

1. If we use the intentional stance, we typically look at behavior using the schema for the intentional process. The particularities of this process are that it has an intended goal, which is achieved in a rational way. Typically, this process is unique (e.g., solving a particular problem), so that the history of the process is not relevant. Using the intentional schema therefore means that we interpret the behavior at hand in terms of its assumed goals, logic and rationality.

If we use the selectional stance, on the other hand, we look at behavior using the schema for the selection process. The particularities of this process are that it has recurrence, variation, and selection. Typically, the process is shaped by a fitness criterion, so that the time series of the process history is a rich source of information. Using the selectional schema therefore means that we interpret behavior in terms of assumed fitness increase, the properties of recurrence, variation, and selection, and history.

2. One of the critical aspects of the selection process is that it may take a lot of time to achieve a substantial fitness increase. This is especially true for Darwinian evolution, where behavioral changes occur only over a number of generations. The use of a nesting of selection processes, as shown by the example of the Skinner crows, can speed up the adaptation process considerably.
3. All selectional adaptation within the lifetime of a creature to changes in the environment needs a selection mechanism or value system or hedonic system that is an internal version of the fitness criterion. This system is shaped over generations by exposure to certain environmental conditions; hence, it can be underdeveloped if the past variation in environmental conditions is small or too fragmentary.

5.4. Understanding behavioral processes better

The proposal that we put forward in this section is that we better understand

behavior if we accept the idea that there are two possible processes that are used by living organisms to perform adaptive behavior: the selection process and the intentional process. For organisms that are able to use intentional processes, hybrid forms are of course also possible. An obvious way to understand complex behavior therefore is to determine in the first place, at a meta-stance, which of the two processes we are dealing with, and second, to interpret the behavior in terms of the corresponding (intentional or selectional) schema. Hence, in order to understand behavior better, we propose that the following five questions be answered:

1. Is the behavior shaped by a selectional or by an intentional process (or by a hybrid process)?
2. What is the benefit of the behavior (the goal or fitness increase)?
3. What is the mechanism or the logic of the behavioral process?
4. How is it implemented?
5. And, in case of a selectional process: what is the history of the behavior?

Received 24 March, 2004, Revision received 11 March, 2005, Accepted 29 March, 2005.

References

- Baum, W. (2000). *Behavior and the General Evolutionary Process*. http://cogprints.org/1051/00/Behavior_and_the_General_Evolutionary_Process.htm.
- Cabanac, M. (1992). Pleasure: the common currency. *Journal of Theoretical Biology*, 155(2), 173-200.
- Carver, C. S. and Scheier, M. F. (1998). *On the Self-Regulation of Behavior*. Cambridge: Cambridge University Press.
- Cherniak, C. (1999). Rational agency. In Wilson R.A. and Keil F.C. (Eds.), *The MIT Encyclopedia of Cognitive Sciences*. Cambridge: MIT Press.
- Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.
- Damasio, A.R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt.
- Dawkins, R. (1989). *The Selfish Gene*. Oxford: Oxford University Press.
- Dawkins, R. (1990). *The Blind Watchmaker*. London: Penguin Books.
- Dennett, D. C. (1981). Intentional systems. In Haugeland, J. (Ed.), *Mind Design: Philosophy, Psychology, Artificial Intelligence*. Cambridge: MIT Press.
- Dennett, D. C. (1988). Précis of the intentional stance. *Behavioral and Brain Sciences*, 11(3), 495-504.
- Dennett, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Schuster.
- Dennett, D. C. (1998). *Cognitive Science as Reverse Engineering: Several Meanings*

- of Top-Down and Bottom-Up*. In D.C. Dennett. *Brainchildren: Essays on Designing Minds*. Bradford Books, Cambridge, MA.
- Fischer, A. R. H., Blommaert, F. J. J., and Midden, C. J. H. (2004). *Evaluation of visual search tasks*. Submitted to International Journal of Human-Computer Studies.
- Johnston, V. S. (1999). *Why We Feel: the Science of Human Emotions*. Cambridge: Perseus Books.
- Krebs, J. R. and Davies, N. B. (1993). *An Introduction to Behavioural Ecology*. Oxford: Blackwell Science.
- Langton, C. G. (1989). Artificial Life. In Langton, C. G. (Ed.), *Artificial Life. SFI Studies in the Sciences of Complexity*, Redwood City: Addison-Wesley.
- Leont'ev, A. N. (1981). *Problems of the Development of the Mind*. Moscow: Progress.
- Marr, D. (1982). *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: Freeman.
- Newell, A. (1990). *Unified Theories of Cognition*. Cambridge: Harvard University Press.
- Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, 88, 1-15.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41: 1-20.
- Simon, H. A. (1996). *The Sciences of the Artificial*. Cambridge: MIT Press.
- Simonton, D. K. (1996). *Foresight in insight? A Darwinian answer*. In Sternberg, R. J. and Davidson, J. E. (Eds.), *The Nature of Insight*. Cambridge: MIT Press.
- Skinner, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64(3), 153-181.
- Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie*, 20, 410-433.
- Waal, F. B. M. de (2001). *The Ape and the Sushi Master: Cultural Reflections by a Primatologist*. New York: Basic Books.
- Wilson, R.A. (1999). Philosophy. In Wilson R.A. and Keil F.C. (Eds.), *The MIT Encyclopedia of Cognitive Sciences*. Cambridge: MIT Press.
- Zach, R. (1978). Selection and dropping of whelks by Northwestern crows. *Behaviour*, 67, 134-147.
- Zach, R. (1979). Shell-dropping, decision-making, and optimal foraging in Northwestern crows. *Behaviour*, 68, 106-117.