CONTENTS

1

Challenges to the Evolution of Intelligence

- The world offers many resources to organisms, but it is also large and complex.
- Obtaining resources is hard. Exploration takes time, behavior may have delayed consequences, and informative stimuli are often mixed with noninformative ones.
- Animals have evolved several solutions to these challenges, such as learning by trial and error and learning from others.
- In humans, a new solution has emerged that couples cultural information and the capacity to think.
- The transition from animal to human intelligence can be understood by reasoning about *sequences*: sequences of behavior, sequences of information-processing steps, and sequences of stimuli.

1.1 What Happened in Human Evolution?

Until three or four million years ago, our ancestors inhabited a small region in Africa and were probably similar in intelligence to contemporary great apes. Since then, our species has acquired many unique features and has colonized almost all terrestrial habitats. Figure 1.1 shows a coarse summary of our species' history. Characteristics that emerged in human evolution include language, complex societies, material and nonmaterial culture, such as art and science, and a rich inner world of thoughts, hopes, and fears. While it is clear that our species' ability to process and organize information has changed, it is not easy to pinpoint what the changes were and what caused them. We refer to these changes collectively as the *human evolutionary transition*, in analogy with other momentuous evolutionary events, such as multicellularity and sexual reproduction (Maynard Smith and Szathmáry 1995). In this book, we put forward a theory of the *content* of the transition (how human

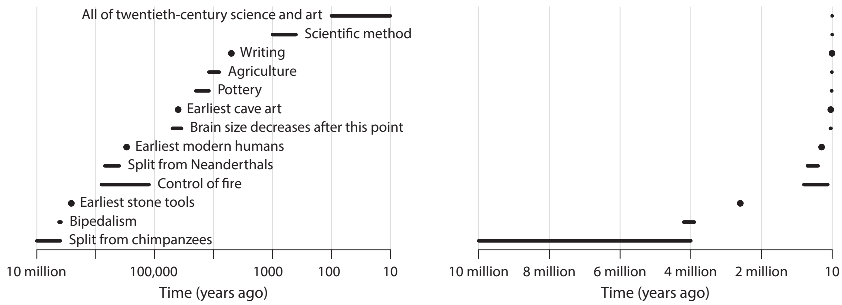

FIGURE 1.1. Some major developments in human history. The two panels include the same events. On the left, a logarithmic time axis enables us to label events clearly. On the right, a linear axis highlights the dramatic speed of cultural evolution: all of twentieth-century science and art has been developed during a mere 0.001% of the time depicted, or about 0.1% of the time since the appearance of modern humans. Data sources: split from chimpanzees: 10–4 million BP (years before present; Dolhinow and Sarich 1971, White et al. 2009); bipedalism: at least since *Australopithecus afarensis*, around 4 million BP ("Lucy"; Ward 2002); earliest stone tools: 3.3 million BP (Harmand et al. 2015); control of fire: between 800k (Berna et al. 2012) and 125k BP (Karkanas et al. 2007); split from Neanderthals: around 600k BP (Schlebusch et al. 2017); earliest modern humans: around 300k BP, based on population genetics (Schlebusch et al. 2017) and fossils (Hublin et al. 2017); brain size decrease: 50–30k BP (Henneberg 1988); cave art: at least 40k BP (Aubert et al. 2014, Brumm et al. 2021); pottery: at least 20k BP (Wu et al. 2012); agriculture: around 10k BP (Vasey 2002); writing: at lea1st 5k1 BP (cuneiform: Walker 1987); scientific method: conventionally, from eleventh-century Arab scholars to seventeenth-century Europeans.

information processing differs from that of other animals) and of its *causes* (the evolutionary events that caused the transition, including genetic and cultural evolution).

In this chapter, as an introduction to a broader argument about the nature of animal and human intelligence, we consider general challenges to the evolution of intelligence. Before beginning, we note that our use of the word *animals* typically excludes humans. Humans are animals, but this usage is convenient to compare humans and other animals.

1.2 The World Is Full of Opportunities

The world is incredibly rich in resources that organisms can exploit, where "resource" is broadly intended as anything that can aid survival and reproduction. Tropical habitats are spectacularly replete with life, but even in hostile environments, such as deserts or the Arctic, organisms can extract enough energy from their surroundings to sustain their species. The behaviors that animals use to obtain resources vary greatly in complexity, depending on what resources they rely on. Filter feeders, for example, rely on resources that are continuously available and floating in water, and therefore their feeding behavior is limited to forcing water through specialized structures that capture nutrients. This simple strategy has evolved many times, such as in molluscs, crabs, shrimp, sharks, and whales. Most animals, however, rely on more complex strategies. Predators may use stealth, speed, venom, or an artifact such as a spider web. Animals that rely on seasonal resources must secure an energy store for hard times, by accumulating fat, hoarding food, or decreasing energy consumption. These are just a few examples of the astonishing diversity of animal life.

Exploiting environmental resources requires organized sequences of actions, in all but the simplest cases. This concept is crucial to the thesis of our book. Weaving a web, hunting prey, or escaping a predator requires executing a sequence of behavior, often with great precision. A mistake results frequently in a lost opportunity, or even in serious harm. Animals have discovered a staggering number of productive sequences of behavior, among many potential alternatives that do not work or that work less well. Humans, however, have developed productive sequences of incredible length. To better appreciate the gap between human and nonhuman sequences, consider a sophisticated nonhuman tool, such as a twig used by chimpanzees to extract termites from their nest, and an outwardly unsophisticated human tool, such as a coat hanger. We are fascinated by chimpanzees' abilities for tool manufacture, but we are typically indifferent to coat hangers. Chimpanzee tools, however, can be built by a single individual with a few actions, such as locating an appropriate twig, detaching it from the branch, and stripping it of its leaves (Nishihara et al. 1995, Sanz et al. 2009). In contrast, the sequence of actions that goes into building a coat hanger is so long as to be untraceable. The coat hanger is made by machines with thousands of parts (figure 1.2), using metal that has ultimately been obtained through an organized mining operation involving thousands of people. The same holds for all but perhaps the simplest objects of daily life in industrial societies (Jordan 2014). To understand why only humans have been able to discover such long sequences, we must first examine why finding productive sequences is difficult at all.

1.3 Sequences and Combinatorial Dilemmas

The design of a behavior system determines what resources it can exploit. For example, an organism whose behavior is completely hardwired is unable to adjust to new food sources or new threats. Likewise, an organism with poor senses might fail to distinguish edible from inedible food. Accordingly,

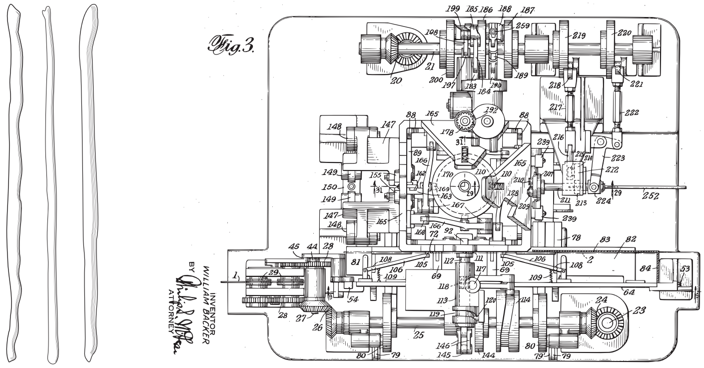

FIGURE 1.2. Left: Twigs used by chimpanzees to perforate termite mounds (redrawn based on Nishihara et al. 1995). Right: One of 32 diagrams describing a machine to manufacture coat hangers (US patent 2,041,805, November 23, 1935).

behavior systems face many challenges. Sense organs must be good enough to perceive relevant information, memory good enough to store such information, and information processing sophisticated enough to drive efficient behavior while also being sufficiently fast. The challenges that are most important for this book are several kinds of combinatorial dilemmas involving sequences. These are difficulties that derive from the exponential increase in possibilities when a task grows in complexity (Bellman 1961, Dall and Cuthill 1997, Keogh and Mueen 2010). In our case, the dilemmas arise when going from shorter sequences to longer sequences. The three combinatorial dilemmas that play a fundamental role in this book involve sequences of behavior, sequences of information-processing steps, and sequences of stimuli.

1.3.1 Behavioral Sequences

In principle, it is possible to find productive sequences of behavior through brute-force exploration, that is, by simply trying out all possible sequences. This, however, is too time consuming to be generally practical. It's like trying to phone someone by dialing all possible phone numbers. More formally, consider an animal exploring sequences of l actions, and assume that each action is chosen randomly out of a repertoire of m. There are then m^l sequences that the animal can try out. If r of these sequences are rewarding, the expected number of attempts before finding a rewarding sequence is

$$\frac{m^l}{r} \tag{1.1}$$

This number increases for longer sequences and more actions. Finding longer sequences is *exponentially* harder than finding shorter ones. Finding a sequence that is just 3 actions long, with a behavioral repertoire of 10 actions, would take $10^3 = 1000$ attempts. Finding a sequence of 10 actions would take $10^{10} = 10$ billion attempts. Furthermore, time spent learning detracts from time spent using what has been learned (the exploration-exploitation dilemma, considered in sections 3.2 and 4.3).

Equation (1.1) also highlights a difficulty with increasing motor flexibility. Most animals have a rather constrained repertoire of possible actions, which is strongly determined genetically (Tinbergen 1951, Eibl Eibesfeldt 1975). Humans and other primates are remarkable for their flexibility, and humans in particular for their ability to learn a diversity of complex skills, such as gymnastics, surgery, or playing the piano. Motor flexibility appears advantageous because it increases the animal's capacity to act on its environment, but it also leads to learning costs that compound with sequence length. For example, an animal that can perform 10 actions can try out $10^2 = 100$ sequences of 2 actions, but this number increases to $50^2 = 2500$ for an animal that can perform 50 actions, yielding a 25-fold increase in the time to try out all possible actions. Thus motor flexibility is not necessarily advantageous, and we can add it to the list of human features whose evolution we would like to understand.

1.3.2 Mental Sequences

Most animals need fast decision making. A foraging bird, for example, is constantly deciding where to hop next, whether to look up to check for predators, whether to switch to a different activity, whether to attack a bug it has spotted, and so on. A more complex information-processing mechanism can be useful if it makes better decisions, but not if this takes too much time. To see how combinatorial dilemmas can arise in information processing, consider planning a sequence of actions. If there are $m = 10$ available actions, planning a single action requires considering 10 alternatives, but planning an action sequence incurs exponentially increasing costs, similar to brute-force exploration. Planning 2 actions may require considering up to $m^2 = 100$ alternatives, and planning 3 actions up to $m^3 = 1000$. Even if imagining actions can be faster (and less risky) than actually performing them, planning sequences can still be prohibitively time consuming. We consider the costs of planning in chapter 11.

1.3.3 Stimulus Sequences

To behave efficiently, organisms typically require information from the environment. Locating food, finding mates, and avoiding predators are some of the activities in which environmental information is essential. Extracting and using information from the stream of sensory experiences poses many challenges. One is stimulus recognition. For example, an organism can only learn to eat a particular fruit if it can recognize sensory states that indicate the presence of that fruit. These sensory states are numerous, arising from differences between individual pieces of fruit as well as differences in background, viewpoint, distance, and lighting conditions. This problem is typically solved through stimulus generalization, that is, by evaluating the current stimulus based on its similarity to familiar stimuli (stimuli to which a response has already been established). The rules of stimulus generalization appear similar regardless of whether knowledge of the familiar stimuli is inborn or learned (Ghirlanda and Enquist 2003).

Stimulus recognition and generalization are very difficult computational problems, as witnessed by the fact that computer systems are just now becoming competent at recognizing objects in the real world (Klette 2014). However, we believe that there are no major differences between humans and animals in this domain, and for this reason we do not devote much space to the topic, with two important exceptions. The first is the recognition of *sequences* of stimuli, for which animals' abilities do appear more limited than humans' (chapter 5). The second is the possibility of learning new *representations* of stimuli, such as symbolic representations or representations that emphasize meaningful features, which we also think is much more limited in animals than in humans (chapter 13). We believe that these differences between animals and humans are rooted in the following combinatorial dilemma.

Organisms experience a continuous stream of stimuli, and have to strike a balance between the potential advantages of remembering more information to use for learning and decision making and the costs that stem from remembering more information. One cost is that more information requires a larger memory. The main cost, however, may be that remembering more might actually make learning and decision making more difficult, rather than easier. To appreciate this fact, suppose that an animal can perceive s different stimuli, and that it decides what to do based on the last n stimuli perceived. If $n = 1$, the animal uses the current stimulus only. It thus needs to know what to do in s different situations. If $n = 2$, the animal remembers the current stimulus and the previous one. This may reveal more information about the environment, but it also means that there are now s^2 situations (s^2 pairs of stimuli) in which the animal must know what to do. With increasing n, the number of potential situations increases exponentially as s^n. Even if not all stimulus sequences actually occur, the animal is still faced with an increased number of possibilities that require a decision. With a realistic number of stimuli, the problem is daunting, and the only solution is to focus on a subset of the incoming

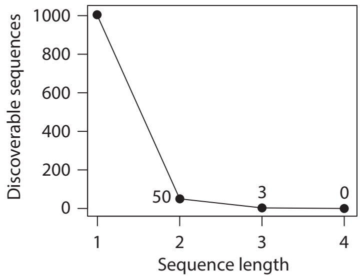

FIGURE 1.3. The number of productive sequences that can be discovered within a given time by brute-force exploration, as a function of sequence length. We assume that each action takes one time unit, so that discovering a sequence of length l takes on average lm^l time units, where m is the number of actions that can be performed. Thus the number of sequences that can be discovered in T time units is $n = \lfloor T/\left(lm^l\right)\rfloor$, with $\lfloor x \rfloor$ indicating the integer part of x. In the figure, $T = 10,000$ and $m = 10$.

information. The dilemma facing the animal is then what to remember and what to discard. We cover this important topic in chapter 5, where we argue that the dilemma lies behind many limitations in animal memory. Additionally, chapter 13 discusses how humans mitigate the problem by learning useful stimulus representations that are transmitted culturally.

1.4 How Can Combinatorial Dilemmas Be Managed?

Given the combinatorial dilemmas just highlighted, finding productive sequences of behavior appears prohibitively hard. With countless possibilities for experienced stimulus sequences and potential behavioral and mental sequences, animals and humans could accomplish very little without sound ways of managing the dilemmas (Zador 2019). Consider, for example, an animal that learns entirely by brute-force exploration, without any genetic or social information to help in the search for productive behavior. Suppose that the animal can perform $m = 10$ different actions and that it has time to try out $T = 10,000$ actions. Such an animal would have time to learn the best action in 1000 stimulus situations; but it could learn the best 2-action sequence in only 50 situations and the best 3-action sequence in only 3 situations, with no guarantee of learning the best 4-action sequence in even a single situation (figure 1.3). These discouraging numbers, moreover, apply under idealized conditions; for example, that the animal learns each correct sequence the first time it performs it.

We think that the evolution of both animal and human intelligence can be seen profitably as the discovery of strategies to manage combinatorial dilemmas (Zador 2019, Quiroga 2020). Animals and humans *can* survive and reproduce, hence they must have found viable strategies. For example, a generalist strategy might focus on learning many different short sequences, while a specialist strategy could focus on a few long ones whose learning could be supported by genetic predispositions. Only humans, it seems, have found a way to discover very many, very long sequences. This difference between humans and animals is the underlying theme of the book. As we will start to see in detail in chapter 2, our conclusion is that human and animal intelligence rely on two very different strategies for managing combinatorial dilemmas.

1.5 Our Approach

In this book, we aim to put forth a strong theory about the differences between human and animal intelligence. By "strong," we mean a theory that is formalized mathematically, from which clear empirical predictions can be derived. We believe that much of the disagreement and fragmentation that exist today in the fields of learning and cognition stem from a lack of formal theory. Human and animal behavior is often described with words that are inherently slippery: intelligence, cognition, understanding, planning, reasoning, insight, mental time travel, theory of mind, and others. These words describe abilities, but not how these abilities are achieved. In this book, we ask what organisms can do, what information their behavior is based upon, and how such information is acquired and used. This approach is not new. Formal models of human and animal behavior have been advanced in psychology, biology, and computer science. As detailed in chapter 2, we are greatly indebted to these traditions, as well as to the deep empirical and conceptual knowledge of behavior accumulated by ethologists and psychologists.

We do not expect all of our ideas to be correct, but we are convinced that our approach can refocus current debate in a direction with a more concrete promise of progress. The rest of this section elaborates our methodology.

1.5.1 What Is in a Mental Mechanism?

In this book, we try to understand how animals and humans arrive at productive behavior by focusing on the information-processing mechanisms that underlie behavior. We call any such mechanism *mental mechanism*. The term *mental* refers simply to the fact that these mechanisms operate within the brain, possibly disjointed from ongoing stimulation, and we do not attach to it any particular significance. We could have used equally well the terms

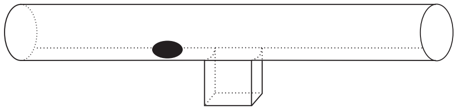

FIGURE 1.4. A trap tube. A food item (black oval) can be pushed out by inserting a stick from one end of the tube. Pushing from the other end causes the food to fall into the trap.

behavior mechanism (common in ethology) or *cognitive mechanism* (common in psychology), but the first may suggest a disregard for internal information processing, which is not our aim, and the second often excludes associative learning, which instead we include as a mental mechanism (see chapter 2).

We base our arguments on explicit models of mental mechanisms, by which we mean formal descriptions of how information is acquired, represented, updated, and used to make behavioral decisions. It is also important to consider which aspects of mental mechanisms develop rigidly (based primarily on genetic information) and which are instead flexible, that is, open to influence from experience. Ideally, all mechanisms should be mathematically defined in enough detail that they can be simulated on computers and implemented in robots. The advantage of this approach is to leave no doubt about how a mechanism is supposed to work, so that it is possible to determine what it can and cannot do.

We focus on behavior without addressing its neural basis. Neuroscience has made tremendous progress in relating behavior to nervous system operation, yet we are still far from understanding how differences in neural processes across species translate into different behavioral abilities. In other words, we start from behavior because we know what animals do much better than what their brains do. At the same time, the models we present in the book can be formulated readily as neural network models (Arbib 2003, Enquist and Ghirlanda 2005, Enquist et al. 2016), which may go some way toward closing the gap between brain and behavior.

1.5.2 What Does It Take to Solve a Problem?

Most conclusions about animal intelligence rest on experiments that challenge animals with various tasks or puzzles, and on inferring underlying mental mechanisms from the results of these experiments. Typically, these tasks can be described and analyzed in different ways. For example, a trap tube experiment (figure 1.4; see Visalberghi and Limongelli 1994, Limongelli et al. 1995) can be described as probing physical cognition, causal inference, or reasoning, or it can be seen simply as a choice between two actions (insert a stick from either end) based on visual information (where the food item lies). What matters, however, is what hypotheses about information processing are required to reproduce how animals behave on the task. To ascertain this conclusively, we see no alternative to formalizing the task and testing different mental mechanisms on it, comparing the behavior of models with that of animals. This often produces surprising results. For example, a task may turn out to be uninformative because it can be solved realistically by many mechanisms (trap tube experiments are one example). Of note, we find that associative learning—often deemed insufficient to reproduce certain aspects of animal intelligence—can behave realistically when modeled formally (see chapter 7).

1.5.3 What Are the Costs and Benefits of Mechanisms?

All mechanisms have costs in terms of time and energy needed to build and operate them, and in terms of the time they require to find and execute productive behavior. A mechanism that is more "intelligent" (can solve more problems) may actually be selected against if it needs too much energy, if it requires information that cannot be obtained easily, or if it takes too long a time to operate. We find that cost-benefit analysis of mechanisms is rare in current debate about animal cognition, yet such analysis is necessary to understand whether a mechanism is a viable evolutionary solution to a problem or set of problems. For this reason, we endeavor to understand the costs and requirements of the mechanisms we consider.

1.5.4 What Data Are Relevant?

Some readers may feel we have omitted studies or experimental paradigms that they deem important, while perhaps focusing on findings that are less central to current debate in animal cognition. Our selection of studies has been guided by our focus on the learning of productive behavioral sequences. For example, we have chosen not to discuss experiments on self-recognition in mirrors. These experiments may be important for a number of questions, but currently it is unclear what mechanisms are involved in self-recognition and what evolutionary advantages they bring.

We also focus on data that can be interpreted with as little ambiguity as possible. It is difficult, for example, to interpret the results of experiments whose subjects have an unknown developmental history, because a given behavior can often result from different mental mechanisms, given appropriate training. For this reason, observing a behavior without knowing the animal's experiences is often uninformative. Therefore, we favor experiments under controlled conditions and avoid, for the most part, observational studies or case studies of single individuals. One should also be aware that, even if we label one part of the experiment as "training" and another as "testing," animals learn from each and every experience. The possibility that behavior is influenced by learning during what we consider a mere test should always be taken into account. We see in chapter 7 that this consideration can be crucial in forming conclusions about mental mechanisms.

1.5.5 Learning Simulator and Online Script Repository

To achieve our goal of rigorous arguments, we have developed a simulation environment that we use throughout the book to specify experimental designs and mental mechanisms. The simulator is described in Jonsson et al. (2021) and is available at https://www.learningsimulator.org. It offers a programming language to specify environments in which learning agents can act in pursuit of goals, such as securing resources and avoiding danger, as well as a library of mental mechanisms that can be tested in any environment. Our simulation scripts can be found at https://doi.org/10.17045/sthlmuni.17082146.

1.6 Animal Rights and Human Responsibility

In this book, we conclude that nonhuman animals are less similar, cognitively, to humans than is often claimed (especially in nonacademic publications). We fear that this claim may be interpreted as detrimental to the humane treatment of animals, but this is not our intent. Rather, we stress that humans are the only species with the cognitive and technical capacity to willingly influence the fate of other organisms. With this power comes a responsibility for stewardship, which, to be effective, requires understanding the needs and capacities of animals (Stamp Dawkins 2008). Indeed, we suggest that species that are often considered less cognitively competent may actually be on the same footing as "more advanced" species considered worthy of ethical concerns. We discuss information-processing capacities only, rather than sentience or capacity to suffer, but these are not irrelevant to ethical arguments. For example, we conclude that, based on available evidence, all vertebrates may be able to assign positive or negative values to experiences (Macphail and Barlow 1985). What to make of this information is a complicated question subject to both ethical and practical considerations (Herzog 2010).

INDEX