

Preface

Karl Friston

Active Inference is a way of understanding sentient behavior. The very fact that you are reading these lines means that you are engaging in Active Inference—namely, actively sampling the world—in a particular way—because you believe you will learn something. You are palpating this page with your eyes simply because this is the kind of action that will resolve uncertainty about what you will see next and—indeed—what these words convey. In short, Active Inference puts the action into perception, whereby perception is treated as perceptual inference or hypothesis testing. Active Inference goes even further and considers planning as inference—that is, inferring what you would do next to resolve uncertainty about your lived world.

To illustrate the simplicity of Active Inference—and what we are trying to explain—place your fingertips gently on your leg. Keep them there motionless for a second or two. Now, does your leg feel rough or smooth? If you had to move your fingers to evince a feeling of roughness or smoothness, you have discovered a fundament of Active Inference. To feel is to palpate. To see is to look. To hear is to listen. This palpation does not necessarily have to be overt—we can act covertly by directing our attention to this or that. In short, we are not simply trying to make sense of our sensations; we have to actively create our sensorium. In what follows, we will see why this has to be the case and why everything that we perceive, do, or plan is in the compass of one existential imperative—self-evidencing.

Active Inference is not just about reading or epistemic foraging. It is, on one view, something that all creatures and particles do, in virtue of their existence. This might sound like a strong claim; however, it speaks to the fact that Active Inference inherits from a free energy principle that equates existence with self-evidencing and self-evidencing with an enactive sort of

inference. However, this book is not concerned with the physics of sentient systems. Its focus is on the implications of this physics for understanding how the brain works.

This understanding is not an easy business, as witnessed by millennia of natural philosophy and centuries of neuroscience. Although one can find the roots of Active Inference in first principle accounts of self-organized behavior (i.e., variational principles akin to Hamilton's principle of stationary action), first principles do not help very much when asking how a particular brain works and how it differs from another brain. For example, committing to the theory of evolution by natural selection does not help in the slightest when it comes to understanding why I have two eyes or speak French. This book is about using principles to scaffold key questions in neuroscience and artificial intelligence. To do this, we have to move beyond principles and get to grips with the mechanics to which the principles apply.

As such, Active Inference—and its accompanying Bayesian mechanics—is there to frame questions about how we perceive, plan, and act. Crucially, it does not aim to replace other frameworks, such as behavioral psychology, decision theory, and reinforcement learning. Rather, it hopes to embrace all those approaches that have proven so successful within a unified framework. In what follows, we will pay special attention to linking key constructs from psychology, cognitive neuroscience, enactivism, ethology, and so on to the calculus of belief updating in Active Inference—and its associated process theories.

By *process theories*, we refer to theories about how belief updating is realized by neuronal (and other biophysical) processes in the embodied brain and beyond. Work to date in Active Inference offers a fairly straightforward set of computational architectures and simulation tools to both model various aspects of a functioning brain and enable people to test hypotheses about different computational architectures. However, these tools only solve half the problem. At the heart of Active Inference lies a generative model—namely, a probabilistic representation of how unobservable causes in the world out there generate the observable consequences—our sensations. Getting the generative model right—as an apt explanation for the sentient behavior of any experimental subject or creature—is the big challenge.

This book tries to explain how to meet this challenge. The first part sets up the basic ideas and formalisms that are called on in the second part—to illustrate how they can be applied in practice. In short, this book is for

people who want to use Active Inference to simulate and model sentient behavior, in the service of either scientific inquiry or, possibly, artificial intelligence. Thus it focuses on those ideas and procedures that are necessary to understand and implement an Active Inference scheme without getting distracted by the physics of sentient systems on the one hand or philosophy on the other.

A Note from Karl Friston

I have a confession to make. I did not write much of this book. Or, more precisely, I was not allowed to. This book's agenda calls for a crisp and clear writing style that is beyond me. Although I was allowed to slip in a few of my favorite words, what follows is a testament to Thomas and Giovanni, their deep understanding of the issues at hand, and, importantly, their theory of mind—in all senses.

Acknowledgments

We gratefully acknowledge invaluable input from our friends and colleagues—in particular, past and present members of the Theoretical Neurobiology group at the Wellcome Centre for Human Neuroimaging, University College London; the Cognition in Action (CONAN) Lab at the Institute of Cognitive Sciences and Technologies, National Research Council of Italy; and numerous international collaborators who have been integral to the development of the ideas presented in this book. This young but growing community has been more than generous in providing both intellectual support and motivation. Furthermore, we gratefully acknowledge Robert Prior and Anne-Marie Bono from MIT Press for kindly accompanying and advising us during the preparation of this book and Jakob Hohwy and other thoughtful reviewers for their guidance. Finally, we thank the funding agencies that provided financial support for our research: KJF was funded by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z); GP was funded by the European Research Council under the Grant Agreement No. 820213 (ThinkAhead) and the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3).

