Preface

A rough definition of artificial intelligence might be: the enterprise of programming computers to reason. For instance, an intelligent computer interface would be one which reasoned about its user's commands so as to interpret them sensibly and flexibly; and an intelligent robot would be one which could reason about novel ways to carry out its tasks in a changing environment. Indeed, it is hard to see what intelligence without reasoning could be, and hard to see how, in practice, an artificial device could be made to reason without being controlled by some sort of computer.

As such, AI is, and needs to be, a highly technical discipline. It traffics in symbols and formulae, logics and computer languages, theorems and algorithms. Computers brook no vagueness or hand waving: the theory of reasoning required by AI must therefore be a theory more precise and detailed than any that has gone before. Yet it would be a mistake to think that AI is a discipline concerned primarily with computers. For what makes AI so challenging (and at the same time so fascinating) has less to do with the computers in which it hopes one day to produce intelligence, than with the mystery of what, exactly, the intelligence it hopes to produce in them *is*. That is, the problems of AI are not primarily computing problems (hardware or software), but problems concerning what it is for a device to reason. And the most valuable contributions made by AI are not primarily contributions to computing technology, but the precise expression of insights into the very nature of thought itself.

Thus, a textbook on AI should have two main aims: first, to explain the technical details behind the most important advances in the subject, and second, to explain the essential insights into the nature of thought which

they express. These are the aims of this book.

The material presented in the following pages derives from courses given at the University of Manchester (both at undergraduate- and MSc-level) and at the Universität Hildesheim. It is designed primarily for university courses in AI which aim to take the student to a level where a large proportion of articles in the journal *Artificial Intelligence* are accessible. The book is thus suitable for AI courses from the 2nd year onwards (for students specializing in AI) up to master's level (for students encountering AI for the first time). Thus, this book covers material at a greater level of sophistication than is normal for AI textbooks. At the same time, the book is self-contained in that it presupposes only a minimal background in mathematics and computing. In particular, the appendix contains a tutorial introduction to logic.

Each chapter contains suggestions for further reading, and there is a selection of exercises which vary from the routine to the very challenging. No reliance is made on any specific programming language, although some exercises concern computer implementations of algorithms discussed in the text. Chapters 1–6 form the core of the book, and are best tackled in order; chapters 7–10, by contrast, are more independent, and readers may choose to skip one or more of them. Having gone rather deeper into many issues than other currently available AI textbooks, I have had to be more selective in the topics covered. My response has been to ignore natural language processing, computer vision and motor control. The justification for this selection is purely pragmatic. The omitted topics rely on specialized background knowledge in linguistics or applied mathematics, and are, I believe, best treated in specialist books.

Finally, I have a number of acknowledgements to make. My stay in Hildesheim, where I was privileged to teach some of this material, was made possible by a grant from the *Deutscher Akademischer Austauschdienst*, as well as by the efforts of various members of staff there. Much of the writing was done during sabbatical leave from Manchester at the *Max Planck Institut für Informatik* in Saarbrücken. I am indebted to the Department of Computer Science at Manchester University, to the Royal Society for funding this sabbatical, and of course to my erstwhile colleagues in Saarbrücken for their overwhelming hospitality and friendliness. In addition, progress on this book was aided by a research grant from the Joint Councils' Research Initiative in Cognitive Science, no. SPG 8920254. I wish also to express my gratitude to the fol-

lowing individuals for their help: John Stobo, David Brée, Jens Dörpmund, Detlev Fehrer, Chiara Menchini, Luoping Xu, Jeff Paris, Andrzej Glowinski, Ivan Leudar, Vaughan Marks, Richard Scarborough and Chris Kirkham. Finally, I would like to thank my long-suffering students, both in the UK and in Germany, for being so patient while I was trying out a variety of ideas on them, only some of which I have thought it wise to print.