

ENGINEERING OF MIND

— |

| —

┌

└

— |

| —

ENGINEERING OF MIND

An Introduction to the Science of Intelligent Systems

James S. Albus

Senior NIST Fellow

Intelligent Systems Division

Manufacturing Engineering Laboratory

National Institute of Standards and Technology

Alexander M. Meystel

Professor of Electrical and Computer Engineering

Drexel University

and

Guest Researcher

National Institute of Standards and Technology



A Wiley-Interscience Publication

JOHN WILEY & SONS, INC.

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

This book is printed on acid-free paper. ∞

Copyright © 2001 by John Wiley & Sons, Inc. All rights reserved.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM.

For ordering and customer service, call 1-800-CALL-WILEY.

Library of Congress Cataloging-in-Publication Data:

Albus, James Sacra.

Engineering of mind: an introduction to the science of intelligent systems/by James S. Albus and Alexander M. Meystel.

p. cm. — (Wiley series on intelligent systems)

ISBN 0-471-443854-5 (acid-free paper)

1. Intelligent control systems. 2. Robotics. 3. Artificial intelligence. I. Meystel, A. (Alex) II. Title. III. Series.

TJ217.5.A428 2001
629.8'92—dc21

2001024703

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

*To Cheryl and Marina
without whose support
this book would never have been written*

— |

| —

┌

└

— |

| —

CONTENTS

Preface	xi
1 Emergence of a Theory	1
What Is Intelligence?	6
Goals and Success	7
The Origin and Function of Intelligence	8
Learning, Evolution, and Instinct	10
Communication and Language	12
A Computational Theory	15
An Engineering Approach	17
2 Knowledge	21
Epistemology	25
Where Does Knowledge Come From?	29
Representation in the Brain	31
Representation of the Real World	49
Representation in Artificial Intelligence and Robotics	51
3 Perception	57
Sensors	62
The Goal of Perception	70
Entities	72
Entities and Attributes	74
Reasoning Within a Rich World Model	77
Attention	79
What Is Important?	81
4 Goal Seeking and Planning	83
A Shopping Center Scenario	86
Control Theory	92
Feedforward Control	97
Planning	101
The Difficulties of Search	106
Hierarchical Multiresolutional Planning	109
Planning Research	111
Emotional Basis of Behavior	112
Motives	118
Consciousness and Emotion	119
	vii

5 A Reference Model Architecture	123
Background	128
The RCS Reference Model Architecture	135
Managing the Complexity of an Intelligent System In Defense of Hierarchies	150 159
6 Behavior Generation	163
Organizational Units versus Agents	169
Task Commands	171
Task Knowledge	173
Task Decomposition	175
Plans and Planning	176
Commands and Plans	182
Plan Execution	185
Emergency Actions	187
Timing	188
Summary and Conclusions	193
7 World Modeling, Value Judgment, and Knowledge Representation	195
Value Judgment	198
Knowledge Database	200
Attribute–Value Pairs	201
Images	204
The Egosphere	206
Maps	212
Attribute, Entity, Class, and Value Images	216
Entity Frame Representation	218
Images and Frames	221
Geometric Entities	222
Generic Entity Classes	226
Specific Entity Classes	228
Events	228
Situations and Experiences	232
Temporal Persistence of Representation	233
Knowledge of Rules	236
Simulation and Imagination	238
Summary and Conclusions	239
8 Sensory Processing	243
Five Processing Functions	245
Functional Modules and Data Flow	258
Recursive Estimation	263
Initialization	267
Hierarchy of Perception	269

CONTENTS

ix

Example of a SP Hierarchy	271
The Complexity of Perception	290
Summary and Conclusions	292
9 Engineering Unmanned Ground Vehicles	295
4-D/RCS	297
Scout Platoon Example	301
Command Vocabularies	307
Demo III Control Hierarchy	310
Example Data Structures	316
Multilevel Planning	318
Two Kinds of Plans	322
Example Scenarios	324
Sensory Processing and World Modeling	342
Summary and Conclusions	346
10 Future Possibilities	351
Technology Forecasts	353
The Future of Intelligent Systems	354
Implications for Science	355
The Dark Side	360
Fear of Automation	361
Implications for Economic Prosperity	363
The Role of Productivity	364
Impact on Economic Theory and Policy	368
Importance for Military Strength	369
Importance for Human Well-Being	369
The Prospects for Technology Growth	373
Ending the Threat to Employment	374
Widespread Ownership	375
Summary and Conclusions	375
References	379
Index	401

— |

| —

┌

└

— |

| —



The *mind* is not a physical thing. It is a process—just as life is a process. The process of mind occurs within the brain just as the process of life occurs within the body. It is tempting to say that the mind is a process that runs on the brain just as a program is a process that runs on a computer. But the analogy is flawed. First, a program is not a process. A program is simply a list of instructions. A process does not occur until a program is executed. Second, the mind is not a program. The program that runs in the brain is part of the brain itself. The brain's program is stored in the neurons and synapses of the brain. Thus, the brain is both machine and program. Third, the process of mind is not entirely, or even mostly, determined by the brain's program. Input from sensors affects the mind profoundly. The activity of mind is not determined solely by the program that is stored in the brain but also on the input received from the environment. The process of mind results from the brain executing its program while simultaneously being bombarded by signals pouring in from literally millions of sensors reporting on stimulation from the environment. The mind is a closed-loop process that senses, perceives, understands, thinks, plans, and acts in response to stimuli. The mind both acts on the world and reacts to the world.

In this book we hypothesize that the difference between mind and brain is similar to the difference between life and body. Just as life is a process that takes place within the body, so mind is a process that occurs within the brain. If our hypothesis is valid, the process of mind is not beyond the realm of scientific inquiry any more than is the process of life. In our hypothesis, *imagination* is a process of visualization and conceptualization (i.e., generating images from assumptions about state, attributes, and relationships of objects, events, situations, and classes). *Thinking* is a process of imagining what might occur if certain actions were taken and certain conditions were achieved, and analyzing the results. *Reasoning* is a process by which rules of logic are applied to representations of knowledge about the world during the process of thinking. *Emotion* is a mental state or feeling that results from a value judgment process evaluating what is good or bad, attractive or repulsive, important or trivial, loved or hated, hoped for or feared. *Feeling* is a pattern of activity that is perceived as pain, pleasure, joy, grief, hope, fear, love, hate, anxiety, or contentment. *Perception* is a process by which sensory input is transformed into knowledge about the world. *Knowledge* is information that is structured so as to be useful for thinking and reasoning. *Cognition* is a collection of processes by which knowledge is acquired and evaluated, awareness is achieved, reasoning is carried out, and judgment is exercised. *Meaning* is the set of semantic relationships that exist between the knowledge database and the external world. Meaning establishes what is intended or meant by behavioral actions, and defines what entities, events, and situations in the knowledge database refer to in the world.

Understanding occurs when the system's internal representation of external reality is adequate for generating intelligent behavior. *Planning* is a thinking process by which a system imagines the future and selects the best course of

action to achieve a goal state. *Wisdom* is the ability to make decisions that are most likely to achieve high-level long-range goals. *Introspection* is a process by which a system examines its own internal state and capabilities and reasons about its own strengths and weaknesses. *Reflection* is a process by which a system rehearses, analyzes, or thinks about the meaning of situations and events. *Reflexion* is a process whereby a system considers what others think about what it is thinking. *Attention* is a mechanism by which an intelligent system directs sensors and narrows its focus to those sensory inputs and internal representations that are important to its current goals—and ignores what is unimportant. *Awareness* is a condition wherein a system has knowledge of the structure, dynamics, and meaning of the environment in which it exists. *Consciousness* is a state or condition in which an intelligent system is aware of itself, its surroundings, its situation, its intentions, and its feelings.

A scientific theory of mind does not yet exist, but an understanding of mind is developing faster than most people appreciate. Progress is rapid in many different fields. Recent results from a number of disciplines have laid the foundations for a computational theory of intelligence. In this book we outline the mainstreams of research that we believe will eventually converge in a scientific theory of mind. We also propose an avenue of research that we believe could lead to the engineering of mind. We do not presume that ours is the only road to this goal. There may be many others. However, we do argue that any road map leading to the engineering of mind must pass the milestone of highly intelligent systems.

We argue that highly intelligent behavior is the result of goals and plans interacting at many hierarchical levels with knowledge represented in a multiresolutional world model. We argue that high levels of intelligence require a rich dynamic world model that includes both a priori knowledge and information provided by sensors and a sensory processing system. We argue that intelligent decision making requires a value judgment system that can evaluate what is good and bad, important and trivial, and can estimate cost, benefit, and risk of potential future actions. We believe that our model of intelligence can enable the engineering design of intelligent systems that pursue goals, imagine the future, make plans, and react to what they see, feel, hear, smell, and taste. This will enable the development of systems that behave as if they are sentient, knowing, caring, creative individuals motivated by hope, fear, pain, pleasure, love, hate, curiosity, and a sense of priority. We believe that research on highly intelligent systems will yield important insights into elements of mind such as attention, gestalt grouping, filtering, classification, imagination, thinking, communication, intention, motivation, and subjective experience. We believe that as the systems we build grow increasingly intelligent, we will begin to see the outlines of what can only be called mind. We hypothesize that mind is a phenomenon that will emerge when intelligent systems achieve a certain level of sophistication in sensing, perception, cognition, reasoning, planning, and control of behavior.

Clearly, mind is more than intelligence. Intelligence is merely a prerequisite, a substrate in which mind may emerge. The mind goes beyond simply generating appropriate behavior. The mind distinguishes between right and wrong, good and evil. The mind is where resides a sense of justice, honor, duty, reverence, beauty, wonder, and religious experience. The mind can experience hope, fear, happiness, grief, pride, shame, and guilt. The mind can harbor a sense of the mysterious. The mind can wonder, worship, and contemplate its own place in the universe. Many aspects of mind will elude scientific understanding for a long time. Nevertheless, we predict that as more is learned about the fundamental mechanisms of intelligence, some of the deeper mysteries of mind will be revealed.

In this book we present a model of the brain as a hierarchy of massively parallel computational modules and data structures interconnected by information pathways that enable analysis of the past, estimation of the present, and prediction of the future. We hypothesize that algorithms, procedures, and data embedded within this architecture will enable the analysis of situations, the formulation of plans, the choice of behaviors, and the computation of current and expected rewards, punishments, costs, benefits, risks, priorities, and motives.

We believe that a rich internal representation of the world is indispensable to higher levels of intelligent behavior. We assert that goals, motives, and priorities are essential features of intelligent behavior, and that model based planning is the best way to find a behavioral trajectory between a starting state and a goal state. We postulate that the concept of goals, tasks, plans, and control are indispensable to understanding what the brain is for and what the mind does.

We specifically reject the behaviorist model of brain, both as a theory and a guide for intelligent systems research. Although behaviorism yields important insights into simple forms of intelligence such as can be found in insects, fish, and reptiles, it is inadequate as a tool for investigating the higher levels of intelligence where mind is likely to emerge. Behaviorism minimizes or ignores internal knowledge representation and eschews concepts such as understanding, intention, and imagination.

We believe that there is overwhelming evidence that the fundamental architecture of the brain is a hierarchy of laterally interconnected computational modules. We argue that this basic principle will be required in any system designed to approach the performance of the human brain. We claim that a multiresolutional representation of knowledge and hierarchical organization of command and control is the most efficient and effective way for individuals and groups to plan and control intelligent behavior in a complex dynamic world. We believe that the computational power to implement our model can be achieved in practical systems in the foreseeable future through hierarchical and lateral distribution of computational tasks.

We describe a reference model architecture that we believe can support an engineering methodology for the design and construction of intelligent ma-

chine systems. This architecture consists of a multiresolutional hierarchy with layers of interconnected computational nodes each containing elements of sensory processing, world modeling, value judgment, and behavior generation. At the lower levels, these elements generate goal-seeking reactive behavior. At higher levels these elements enable perception, cognition, reasoning, imagination, and planning. Within each level of our proposed architecture, the product of range and resolution in time and space is limited. At low levels, range is short and resolution is high, whereas at high levels, range is long and resolution is low. This enables high precision and quick response to be achieved at low levels over short intervals of time and space, while long-range plans and abstract concepts can be formulated at high levels over broad regions of time and space.

Our proposed reference model architecture accommodates concepts from artificial intelligence, control theory, image understanding, signal processing, and decision theory. The artificial intelligence concept of a plan is expressed as a series of subgoals, waypoints, or desired states and is shown to be equivalent to the control theory concept of a reference trajectory. A plan expressed in the form of a series of actions is equivalent to a feedforward command string. Value judgments can be expressed as cost functions, confidence factors, and measures of importance. Our reference model architecture is expressed in terms of the Real-Time Control System (RCS) that has been developed at the National Institute of Standards and Technology and elsewhere over the last 25 years. RCS provides a design methodology, software development tools, and a library of software that is free and available via the Internet. Application experience with RCS provides examples of how this reference model can be applied to problems of practical importance.

In Chapter 1 we outline evidence for the emergence of a computational theory of mind. In Chapters 2, 3, and 4 we examine the fundamental concepts of knowledge, perception, and goal-seeking behavior. Chapter 5 defines the basic building blocks of a reference model architecture for engineering intelligent systems. Chapter 6 deals with how to engineer behavior generating processes. In Chapter 7 we outline world modeling and value judgment processes and suggest data structures for designing a knowledge database. In Chapter 8 we discuss current issues in sensory processing and propose an engineering approach to designing a perception system. Chapter 9 provides an example of how our proposed reference model architecture is currently being applied to the design of intelligent unmanned military scout vehicles. In Chapter 10 we suggest what impact the ability to engineer mind might have on science, economics, politics, and the well-being of humankind.

JAMES S. ALBUS
ALEXANDER M. MEYSTEL