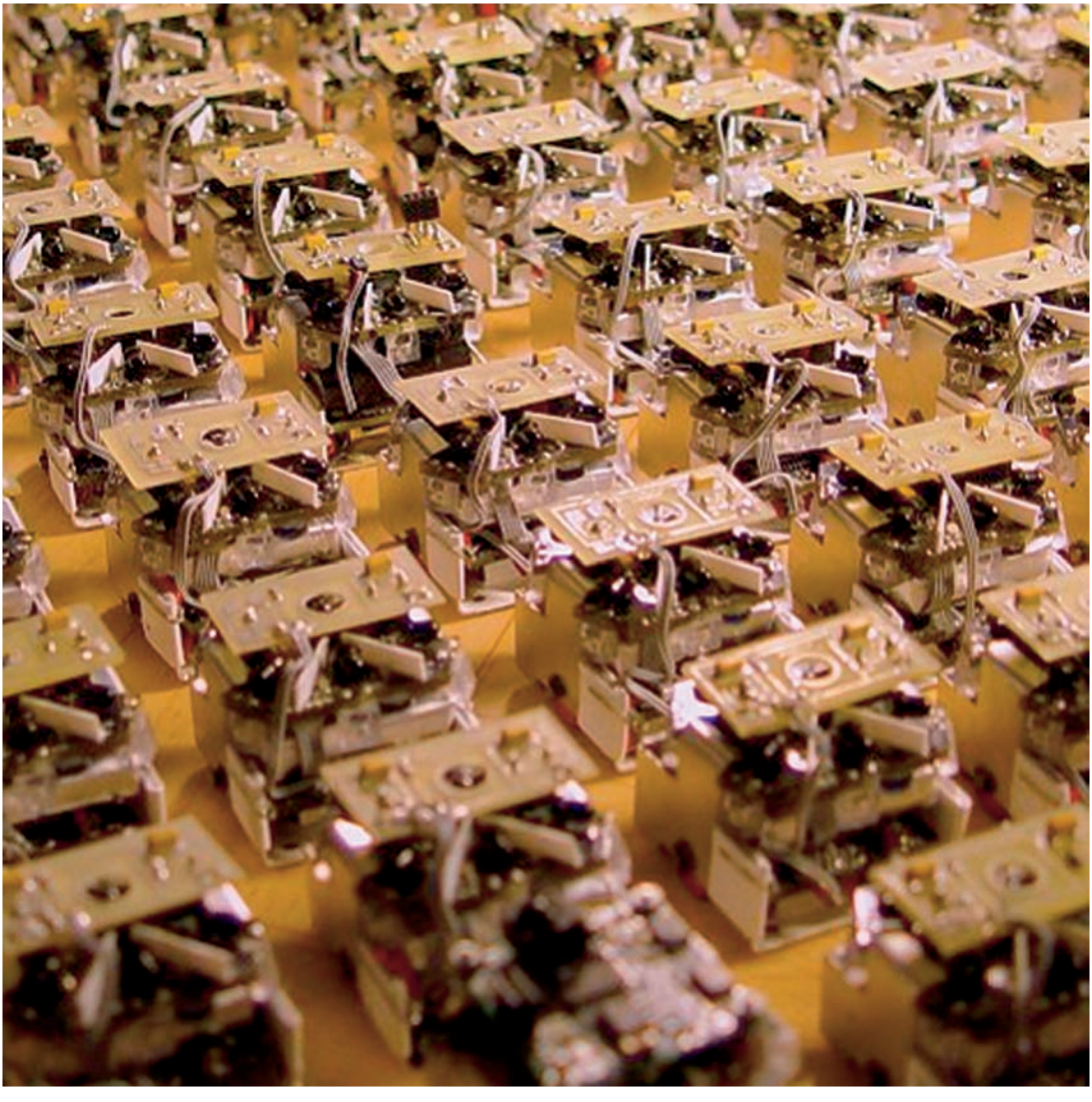
Hermann Haken, Paul Levi

WILEY-VCH

Synergetic Agents

From Multi-Robot Systems to Molecular Robotics



Hermann Haken and Paul Levi

Synergetic Agents

Related Titles

Dunin-Keplicz, B. M., Verbrugge, R.

Teamwork in Multi-Agent Systems

A Formal Approach

ISBN: 978-0-470-69988-1

Wooldridge, M.

An Introduction to MultiAgent Systems

ISBN: 978-0-470-51946-2

Clymer, J. R.

Simulation-Based Engineering of Complex Systems

ISBN: 978-0-470-40129-3

Arendt, W., Schleich, W. P. (eds.)

Mathematical Analysis of Evolution, Information, and Complexity

ISBN: 978-3-527-40830-6

Shamma, J. (ed.)

Cooperative Control of Distributed Multi-Agent Systems

ISBN: 978-0-470-06031-5

Schelter, B., Winterhalder, M., Timmer, J. (eds.)

Handbook of Time Series Analysis

Recent Theoretical Developments and Applications

ISBN: 978-3-527-40623-4

Hermann Haken and Paul Levi

Synergetic Agents

From Multi-Robot Systems to Molecular Robotics

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

The Authors

Prof. Dr. Hermann Haken
Universität Stuttgart
Inst. f. Theoret. Physik
Center of Synergetics
Pfaffenwaldring 57 / IV
70550 Stuttgart
Germany

Prof. Dr. Paul Levi
Universität Stuttgart
Inst. of Parallel & Distributed Systems
Universitätsstr. 38
70569 Stuttgart
Germany

Cover Image:

The cover image shows Jasmine robots prepared by Serge Kernbach, Marc Szymanski, Thomas Schmickl for the large-scale swarm experiment, for more information see Serge Kernbach, Dagmar Häbe, Olga Kernbach, Ronald Thenius, Gerald Radspieler, Toshifumi Kimura and Thomas Schmickl, "Adaptive collective decision-making in limited robot swarms without communication", International Journal of Robotics Research, 32(1), 35–55, 2013 and Serge Kernbach, Ronald Thenius, Olga Kernbach and Thomas Schmickl, "Re-embodiment of Honeybee Aggregation Behavior in an Artificial Micro-Robotic System", Adaptive Behavior, 17(3), 237–259, 2009.

Image reproduced courtesy of Serge Kernbach.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

Print ISBN: 978-3-527-41166-5
ePDF ISBN: 978-3-527-65955-5
ePub ISBN: 978-3-527-65954-8
mobi ISBN: 978-3-527-65953-1
oBook ISBN: 978-3-527-65952-4

Cover Design Adam-Design, Weinheim
Typesetting Thomson Digital, Noida, India

Printed on acid-free paper

Contents

Preface

This book presents entirely new vistas in the following two disciplines:

1) For the first time, it applies basic principles of synergetics – the science of cooperation – to multirobot systems.
2) It applies a modern method developed for active, nonequilibrium quantum systems to *molecular* robots – a rapidly developing, fascinating field within nanoscience and technology.

In both cases (1) and (2), this book deals with *active units*, that is, robots or molecules, capable of forming spatiotemporal structures or collective action based on *cooperation*. In other words, it deals with *synergetic agents*.

In order to reach a broad audience, it is written in a pedagogical style that will allow even nonspecialists to acquaint themselves with our approach. (A few more technical sections are marked by asterisk.)

In fact, both fields, that is, multirobot systems and molecular robots have become highly interdisciplinary endeavors that comprise disciplines such as robotics, mechanical and electrical engineering, physics, informatics, chemistry, biology, medicine, mathematics, and other fields. Our book applies to graduate students, professors, and scientists. Though occasionally we refer to experiments, our emphasis is laid on theoretical approaches. Among our numerous results are

- the Haken–Levi theorem in its classical and quantum mechanical formulation relating robot motion to probability distribution;
- a whole chapter presenting our quantum theory of muscle contraction based on actin–myosin interaction;
- a detailed quantum theoretical model of the motion of molecular robots.

Acknowledgments

We gratefully acknowledge the help provided by Ms. I. Maute who typed (or rather typeset) the manuscript including its complicated formulas quickly and perfectly. We are also indebted to Ms. N. Hepp for her very valuable support to setting up the manuscript. Thanks are also due to Mr. J.-D. Korus for his continuous engagement in preparing the manuscript. We express our gratitude to Dr. U.-Ph. Käppeler for the fast and professional generation of several figures that are included in this book. The authors wish to thank particularly Dr. M. Schanz and Dr. V. Avrutin for many fruitful discussions, calculations, and graphics that they generated untiringly for us.

Stuttgart, April 2012

Hermann Haken
Paul Levi

Prologue I: Synergetic Agents: Classical

Self-Organization in Collective Systems

Collective systems in technique and biology move more and more into the focus of basic research in the fields of natural science such as physics, biology, chemistry, and engineering science such as mechanical engineering, computer science, cybernetics, and robotics. In biology, swarms of bees or ants, flocks of fish or birds, and networks of natural neural networks such as those realized in different kinds of brains demonstrate very impressively the power and the abilities of such collective systems. These advantages are also expected in technical collective systems like cooperative production systems, distributed traffic managements, all facets of internet, and last but not least swarms of mobile robots (on land, in water, and in air). What are the dominant features of natural and artificial collective systems that are so fascinating for all of us?

The two basic scientifically most relevant features of all these aforementioned collaborative systems that immediately catch our eye are the ability to create distributed intelligence (meaning the emergence of intelligence; the whole is more than the sum of all parts) and the competence of self-organization (Levi and Kernbach, 2010). In addition to the dominant property of an “intelligent” collectivity generated by self-organization is the increase of adaptation, of reliability, of flexibility, of self-development, and of self-healing.

In standard artificial collective systems, for example, in traffic management, the intelligence is brought into the system by engineers, but the interplay that creates the self-organization and all the other complementary features like reliability is still very important for inanimate or artificial systems. A very prominent example of a self-organized technical system that is constructed by physicists and engineers is the laser. It demonstrates very clearly the phase transition from noncoherent light of a lamp to coherent light of a laser by self-organization.

The basic concept to define and to implement self-organization is given by the methods of synergetics (Haken, 2004). It is the theory of the cooperation of parts of a system that generate by themselves an “order parameter field” that in turn exerts a strong feedback to its many originators (circular causality). In this book, the parts of a collective system are mainly (but not exclusively) inanimate units.

All features of an intelligent and self-organized system are the result of the cooperative interplay between the global structure (organization principle) of the system, the behavior of the individual units, and the different functionalities "generated" by individual members. As an example of this assertion, consider a swarm of bees. Relevant questions are here: how they organize their foraging, how they perform their navigation and exploration tasks, how they do the foraging, how they distribute the collected nutrition, how they regulate the homeostatic mechanisms, and so on?

The transfer of these features and behaviors to inanimate, artificial swarms of robots, which was mentioned before, is primarily motivated by the approach to get answers to these basic, biological questions by picking up these questions of animated systems and trying to get relevant responses by technical systems. In view of classical artificial intelligence (AI) and its more philosophically based connection to cognitive science that are both characterized by a top-down approach we will present in this book, the new bottom-up approach of collective robotics (Pfeifer and Scheier, 1999) starts from the microscopic parts (e.g., robots) and studies the emergence of intelligence, self-organization, and cognition, for example, in a swarm or even in an organism that is generated out of such a swarm (Floreano and Mattiussi, 2008; Levi, 2009; Siciliano and Khatib, 2008).

Such a distinction between a swarm mode (phase) and an organism mode offers the possibility to analyze the essential features of a part that are inevitable to generate an intelligent swarm (e.g., swarms of house flies never show swarm intelligence). What is different if a swarm is going together in order to build an organism (may be considered a morphogenetic phase transition)? What features of a swarm member are changed if it "mutates" to a "cell" of an organism? How do these new "cells" differentiate themselves to different organs or parts of an organism? Such questions are considered in the so-called symbiotic collective robotics (Levi and Kernbach, 2010). Swarm behavior is also very beneficial in soccer games, for example, in RoboCup, where the robots are no longer small robot cells (about 5 cm^3) but have a bigger volume of about 20 cm^3 (Rajaie *et al.*, 2011).

Besides these basic questions of swarm mode and organism mode, the bottom-up approach in robotics is characterized by the so-called "embodiment," meaning that there can be no intelligence and cognition if there is no body (matter) available; intelligence and cognition require a body.

This statement is augmented by the concept of "situatedness," denoting that each part of such a system can acquire information about the current situation under given environmental conditions, perform an individual interpretation of the existing situation (e.g., by pattern recognition), and finally it makes an individual decision concerning its next activities. The bodies and the situations can be simple or complex. According to the bottom-up approach that is accomplished by our approach, we consider as the first step simple bodies and nonsophisticated situations. In order to complete the two strongly interwoven concepts of "embodiment" and "situatedness," we include in our approach the additional concept of an agent. This is an active part of a whole system (the so-called multiagent system, MAS) that is afflicted with a corpus, is autonomous, and is aware of situations (Weiss, 1999). An agent realizes internally the two concepts of "embodiment" and "situatedness," and it is able to learn. An agent represents a basic concept of robotics and artificial intelligence.

In this book, we create the concept of a synergetic agent. This is an agent that uses internally the methods of synergetics to calculate the sensor-based acquired information and comes to appropriate decisions and actions in response to the calculated information (situation description). The correct information handling is the engine of progress of the interplay between the theory of synergetics – here we mean especially the circular causality of self-organization, the emergence of new qualities by nonequilibrium phase transitions in open systems, and reduction of complexity – and the theory of the emergence of cognition and intelligence of an agent that finally is condensed in intelligent decisions.

The construction and design principles for synergetic agents are based on those of informatics-based agents, but they must be dominantly extended by a paradigm change in physical descriptions of synergetic processes and by a new principle of information.

The paradigm change can very clearly be explained if we compare the classical laser paradigm (a nearly inexhaustible source of inspiration) and a multirobot system (MAS), be it in swarm or in organism mode. The most dominant commonality is, for example, the "circular causality": the participating parts generate one or more "order parameter fields" that operate recurrently and therefore "enslave" the originating parts. Another important commonality is the supplement of the principal coupling to the environment. Here, the following effects have to be considered: damping, fluctuations, and dissipation processes of open systems.

The essential difference is that all atoms that generate the coherent electromagnetic field (order parameter field) are passive and are neither intelligent nor situated, nor able to learn. Robots also obey equations of motion, but their real movement must be generated by controllers for steering (Shen *et al.*, 2004), where an internal force that mimics an external force constrains them to move on a prescribed trajectory. Such controllers have to consider details of the transaction type (type of drive system) and details of the properties of the underground (e.g., land or water) and unforeseen situations (like obstacles or holes). As a result, the "cognitive" decision making of a robot (realized as an agent) generates the appropriate response to unforeseen situations (Levi, 2010). A good example for this latter statement is the kind of response of a soccer robot if it is attacked by one or more robots of the opponent team. This decision is highly influenced by the learned team strategy (how to play the game and if possible to win).

An important supplement for every kind of motion is the coupling to environment. In a classical physical approach, these are the effects of damping, fluctuations (noise), and dissipation. But for mobile robots, we have also to consider new and different types of uncertainties. These are failures in sensor data, aged sensor data, and incorrect steering statements (more generally spoken: degraded information). The correct handling of such degraded information (also including trustworthiness in information source) demands implementation of cognitive processes. In human decision making as part of a cognitive process, the anchoring bias is an example of a dominant focus on a trait of information that is degraded (Kahneman and Tversky, 1996).

This kind of cognitive response has clearly to be distinguished from the elementary stimulus–response cycle that occurs very often on lower levels in biological systems (e.g., consider the Braitenberg vehicle (Braitenberg, 1984)).

In view of information theory, the reaction on unexpected situations is dictated by a minimum of individual information. As bigger the surprise concerning an event the smaller is the probability for this event or for other features like anchoring.

The close connection between the acquired information of an agent and the reaction to this information is formulated by the Haken–Levi information principle.

Each individual synergetic agent minimizes its local information.

$$i_s(\text{agent}) = -\ln p_s\,(\xi_s, \xi_u), \tag{I.1}$$

where p_s is the joint probability of the value ξ_s of the variable of agent "*s*" and of the value ξ_u of the order parameter of the whole system.

This implies that each individual synergetic agent disposes of information that regulates already the feedback caused by the circular causality already mentioned.

This relationship is most clearly expressed when we use the relation

$$p_s(\xi_s, \xi_u) = p_u(\xi_u) p_s(\xi_s|\xi_u), \tag{I.2}$$

where p_u is the probability distribution of the order parameter that is *collectively* generated by all the agents of the system, and $p_s(\xi_s|\xi_u)$ the *conditional* probability that the enslaved variable ξ_s acquires that value provided the value ξ_u is given. By using (I.1) and (I.2), we define the conditional information $i_{s,c}$ of agent *s*:

$$i_{s,c} = -\ln p_s\,(\xi_s|\xi_u). \tag{I.3}$$

If we consider the whole system (e.g., let it be an organism assembled by robot cells), the total system information is maximized. This means the expectation value of all individuals, where information is maximized

$$\text{Information } i\,(\text{system}) = i_u + \sum_s \langle i_s \rangle, \tag{I.4}$$

where

$$i_u = -\sum_{\xi_u} p_u \ln p_u \tag{I.5}$$

is the information of the order parameter and

$$\langle i_s \rangle = \sum_{\xi_u, \xi_s} p_s(\xi_s, \xi_u) i_{s,c} \tag{I.6}$$

the expectation value of $i_{s,c}$(I.3) (Haken, 2006).

By means of the local information (I.1), we may express the equation of motion (or more generally the behavior) of an agent (a robot vehicle):

$$m\,\ddot{\xi}_\delta + \gamma\,\dot{\xi}_\delta = -(Q/\gamma)\nabla i_s + F_s(t). \tag{I.7}$$

(This is a special case of the H-L principle). In (I.7), ξ_s is – in general – the three-dimensional position vector, m the mass of the robot, γ a damping constant, Q the strength of the random force $F_s(t)$ acting on the robot, and ∇ the nabla operator. (For details, cf. Chapter 2.)

There are three main effects attributable to agents. First, we can formulate the circular causality of self-organization by the combined application of the H-L principle to each individual robot and to the system of all robots. Second, we can formulate the equations of motion for one robot or for all robots by the calculation of the gradient of information. Third, we can store the individual and total information gain by the calculation of the Kullback measure. If we perform this calculation by iteration and store each information gain or information loss, then we have implemented a dedicated method to learn.

At first sight, Equation I.7 might look like a simple rewriting of the equation of motion of a robot agent (as part of a multirobot system), namely, instead of using a potential function $V(\xi_s)$ directly, we write it in a somewhat disguised form. In other words, (I.7) seems to rest on some tautology. In mathematics, tautologies are surely not a crime; rather the individual steps used there are just a sequence of tautologies! In the present case, the situation is different, however. First of all, the concepts of the equations of (mechanical) motion and of information in its scientific, mathematical form stemming from information theory originate from two conceptually quite different scientific disciplines. Thus, (I.7) provides us with a qualitatively new insight. As a consequence, we may interpret and use information i under entirely new aspects. Namely, in practice, a robot must *acquire* the appropriate information by its own activities and rather limited preprogramming. Since it does not "know" the positions of the other robots and objects (e.g., obstacles) beforehand, it must measure their relative directions and distances. It then has to attribute to these quantities appropriate artificial potentials. To this end, it has to distinguish between other robots, obstacles, and attractive objects (e.g., energy sources). In specific situations, for example, soccer games, it must distinguish between friend and foe. All these cases require specific preprogrammed potentials (leaving aside aspects of robot learning and evolution).

As we shall see in detail, for instance, when we study docking maneuvers, the robot information may switch from one kind of information to another, depending on the situation. To mention a simple example, the information may switch from the use of one potential function to another one.

Clearly, higher order programs may also be installed in the expression for the information. We will discuss some examples in our book, for example, the self-organized formation of letters by suitable configurations of robots.

Let us discuss how the robot uses the instructions enfolded in the information $i_s(\xi_s, \xi_u)$, (I.1), or, in other words, how it unfolds its information. In principle, it may solve its equation of motion according to (I.7) and use a control mechanism to secure the realization of the wanted motion. In practice, the situation is quite different, at least in general in a multirobot system. First of all, to calculate its future path, the robot must be informed on the future paths of all other robots and vice versa. This requires the action of a "master" computer of very high capacity outside the multirobot system.

In a swarm situation, such a procedure is not possible at all and contradicts the principle of *self-organization*. The practical procedure must be quite different. Based on its measurements of distances and directions to all other objects (including the other robots), the robot under consideration acts on its actuators from moment to moment in such a way that for a *given, measured* value of the r.h.s. of (I.7) the robot accelerates or decelerates, including damping. Because the robot relies on the measured r.h.s., it can even act if slip is present. It may follow its path, though with some *time delay*. In this way, the artificial potential appears as an evaluation function of the quality of reaching the robot's goal.

As we know from the theory of swarms, an essential ingredient of their collective behavior is the requirement that each individual keeps a mean distance to all its nearest neighbors.

There is yet another aspect to our approach: The whole system altogether acts as a *parallel computer* (in contrast to a sequential computer). All its components (the agents!) collect their information in parallel and act in parallel. This information acquisition may be active (e.g., measuring distances to objects) or both active and passive (e.g., communication among agents in collective pattern recognition; see Chapter 3). We believe that our information-based approach opens new vistas to dealing with multirobot or, more generally, multiagent systems. For readers who wish to learn more about the scientific concept of information, we include the following section.

The Tricky Concept of Information (Shannon)

"Information" in ordinary sense is a message (e.g., birth of a child, accident, winning of an election, etc.), an instruction, a set of data, and so on. In more technical terminology, information is essentially a measure of probabilistic uncertainty (not of principal uncertainty, for example, in quantum mechanics (Genz, 1996)). In terms of the discipline of stochastics, the appropriate methodological terms are stochastic events (e.g., unexpected obstacle during an exploration tour of a mobile robot), stochastic variables (e.g., set of sensor data), and stochastic "functions" (e.g., instructions and algorithms). In our book, "information" is a *terminus technicus* that allows a quantitative treatment in terms of the three aforementioned basic definitions. However, the meaning of information is often not very clearly defined, and we will try in the first step to elucidate this meaning before we present it as a useful concept in robotics.

Let us start with the first step by explaining Shannon information ((Shannon, C.E., 1948), (Shannon, C.E., Weaver, W., 1949)) (originally conceived as a measure of the capacity of data transmission channels). We begin our "explanation route" by a set of discrete events labeled by an index λ, where N is a fixed number. Typical examples of such events are tossing of a coin that yields the two events head or number, rolling a die with six outcomes (i.e., events are $\lambda = 1, \ldots, 6$). A more sophisticated example that is for our wanted robot applications more illustrative is the exploration tour of mobile robots in an unknown environment as a task that is a typical part of probabilistic robotics (Thrun, Burgard, and Fox, 2005). This new methodology imposes weaker restrictions on the accuracy (greater uncertainty) of sensor data than the classical deterministic interpretation of measurement data. Typical events during such an exploration tour are the emergence of unexpected obstacles, possibilities of several navigation paths with different lengths (stochastic variable), and the stability (robustness) of the internal power supply (stochastic function in the sense of a homeostasis).

We consider now a very frequent repetition of trials. The probability (frequency) of outcome of event λ is defined by

$$p_\lambda = \frac{\text{number of positive outcomes}}{\text{number of all possible trials}},$$

where we require the normalization of the distribution function $p = (p_1, p_2, \ldots, p_N)$ by

$$\sum_\lambda p_\lambda = 1.$$

The information that a positive occurrence of an individual event delivers is called information of event λ and is defined by

$$i_\lambda = -\ln p_\lambda. \tag{I.8}$$

We can also use $\log_2$ instead ln because $\log_2 = c \ln$, where $c = \log_2 e$, and both logarithmic expressions differ only by a constant factor c.

Shannon defined information as the expectation value of all individual information

$$i_s(p) = -\sum_\lambda p_\lambda \ln p_\lambda = \sum_\lambda p_\lambda \, i_\lambda. \tag{I.9}$$

Formula (I.9) calculates information as a measure of stochastic uncertainty. This term is also called information entropy. The reason for this other naming is the fact that (I.9) is the same mathematical expression as it is used in thermodynamics for entropy. Therefore, von Neumann suggested to Shannon not to use two different names for the identical formula. Today, this argument is no longer fully accepted since the relationship between the information as a measure of (probabilistic) uncertainty and the physical meaning of entropy as the number of microstates, for example, in gases, is clearly distinguished (Penrose, 2006). If we want to point out the equality of the same expression for two different approaches and meanings, (I.9) will be called information entropy in order to accentuate the nonphysical aspect.

Furthermore, in this book we focus on nonequilibrium phase transitions that are characteristics of many dissipative, open systems that not only include living beings but also, for example, robots as artificial "ingredients" of inanimate nature. In open systems, the information can even be increased if a nonequilibrium phase transition occurs and a final system is generated after a bifurcation that has an increased order (Haken, 2006). An example for this declaration is the transition from a lamp (below the bifurcation threshold) to a laser (well above the bifurcation threshold). In closed systems, the opposite effect occurs. In an equilibrium state (constant energy) phase transition, the information entropy decreases if (after the bifurcation) a more ordered system state is achieved.

Maximum and Minimum Principles of Information

We get a feeling on the significance of individual information (I.1) and the total system information of Shannon (I.2) if we treat the following two examples:

1) All but one p_λ are zero: $p_\lambda = 0$; $p_K = 1$, $\ln 1 = 0$, then the individual event information and the total system information are both zero:

   $$i_s = i_\lambda = 0.$$

   This means that there is no uncertainty, no surprise with respect to the outcome of a trial. Or, in other words, there is complete certainty as long as we use the information available to us.
2) All p_λ are equal. The considered probability distribution is given by the uniform distribution $p = (p_1 = (1/N), \ldots, p_N = (1/N))$, $p_\lambda = 1/N$. In this case, the system information i_s maximal: $i_s(p) = \ln N$.

   This is the case if there are no additional constraints besides the standard restriction of normalization of the probability distribution. The uncertainty is maximal since all outcomes are equally likely. Laplace called it the "principle of insufficient reason," which states that all outcomes are equally likely if there is no reason to the contrary (Kapur and Kesavan, 1992). In physics, this result corresponds to the equipartition theorem.

The Kullback measure $K(p, q)$ (Kullback, 1951) calculates the difference between two probability distributions

$$p = (p_1, \ldots, \cdot p_N) \text{ and } q = (q_1, q_2, \ldots, q_N) :$$
$$K(p, q) = \sum_\lambda p_\lambda \ln \frac{p_\lambda}{q_\lambda}, \, (p, q) \geq 0,$$

where each probability distribution (density) is separately normalized to 1, and $K(p, q)$ is nonnegative and vanishes if and only if $p = q$. Usually, the Kullback measure can also be called "information gain" since q is a fixed *a priori* distribution and p is a probability distribution that is searched with the aid of K in order to maximize the divergence of p from q. But this expression can also be used to minimize the difference ("information adjustment"). Closer the distance from p to q, the more the probabilities of the different observed events confirm the *a priori* experiences (knowledge of the experimenter). The application of this method is then directed to find a distribution p that is closest to q and fulfills the same restrictions as q.

However, despite the existing conceptual differences of Shannon measure and Kullback measure there is a central relation between both approaches. There is a tight connection between the maximization of $i_s(p)$ and the minimization of $K(p, q)$ if we assume that q is given by the uniform probability distribution u:

$$K(p, u) = \sum_\lambda p_\lambda \ln \tfrac{p_\lambda}{1/N} = \ln N - i_s(p). \tag{I.10}$$

Maximizing $i_s(p)$ is identical to minimizing $K(p, q)$ if the *a priori* probability distribution is uniform ($q = u$). By maximizing the uncertainty, we minimize the probabilistic distance to a given distribution.

We close this short excursion to two often used expressions for information calculation by the remark that the Shannon approach is not invariant under coordinate transformations, whereas the Kullback approach is invariant under coordinate transformations.

After these remarks, we address ourselves again to the further investigation of the Shannon information in the light of the "maximum information principle" of Jaynes (Jaynes, 1957). This theorem postulates that we are looking for probability distribution p that guarantees $i_s(p) = \text{maximum}$ is fulfilled under all given constraints that also include normalization. In more details, this means that a system tries to realize all "allowed" configurations, that is, configurations that obey the constraints.

For more illustration, we treat another example. Let $N = 2$, $\lambda = 1, 2$ be and $p = (p_1, p_2) = (x, 1-x), 0 \leq x \leq 1$. We introduce again the information of event λ: $i_\lambda = -\ln p_\lambda$.

We discuss two questions that are basic for understanding Jaynes' principle by analyzing the results of Figures I.1–I.3:

1) For which x does becomes $i_s(p)$ a maximum?
   According to Figure I.3, we find $x = 1/2$, $p_1 = p_2, i_s(p) = \ln 2$
2) For which x does p_1 or (p_2) get a maximum?
   According to Figures I.1 and I.2, we find $x = 1, p_1 = 1,\ i_1 = 0$. The probability p_1 is maximum if the information for event 1, i_1, is a minimum.

These two resulting answers lead us to the formulation of two principles that will be important in our book. We start with the general case:

$$i_s(p) = -\sum_\lambda p_\lambda \ln p_\lambda, \quad \sum_\lambda p_\lambda = 1 \tag{I.11}$$

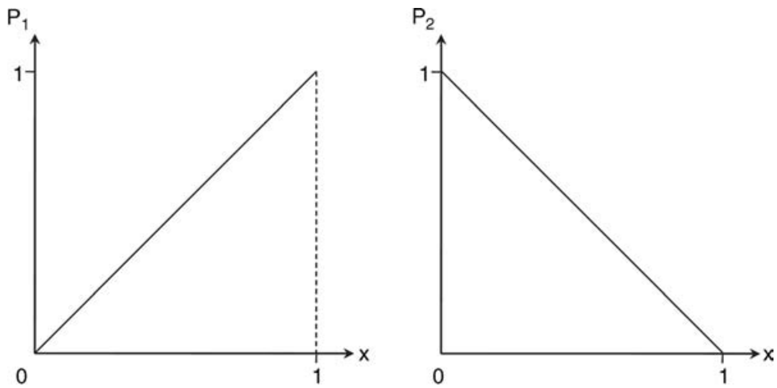


Figure I.1 p_λ versus x; l.h.s.: $\lambda = 1$, r.h.s.: $\lambda = 2$.

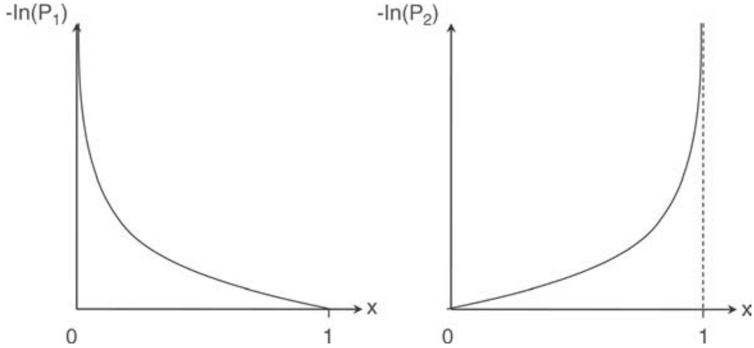


Figure I.2 $-\ln p_\lambda$ versus x; l.h.s.: $\lambda = 1$, r.h.s.: $\lambda = 2$.

and additional constraints that all together will be taken care of by the use of Lagrange multiplicators to maximize $i_s(p)$ (for more details, consult Section 1.4). We consider Figure I.4 and reformulate the two questions put above into one similar, combined question.

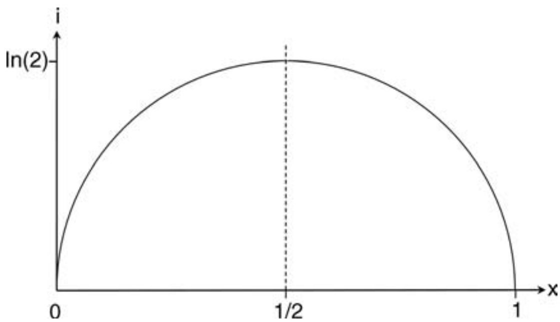


Figure I.3 i versus x.

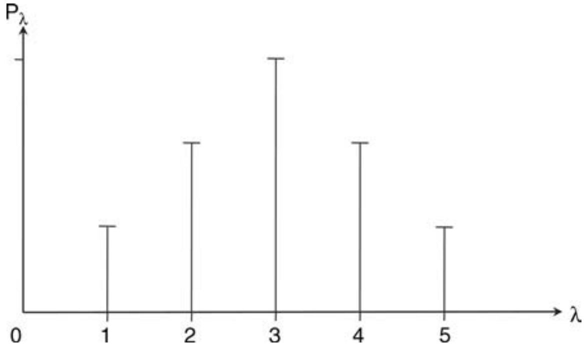


Figure I.4 Discrete probability distribution p_λ as function of λ.

Which p_λ is(are) maximum or which $i_\lambda = -\ln p_\lambda$ is minimum? The answer is: a maximal p_λ minimizes the individual event information i_λ. Or, to turn the argument around, a small p_λ means that the event λ is rare or unexpected, which implies that i_λ must be large. This is the reason why some scientists call the event information i_λ also "surprise," which is another more seldom used *terminus technicus*.

In conclusion, we may state that we have to distinguish between the *maximum information principle* (minimum Kullback measure), according to which we *maximize* the total information $i_s(p)$, and the *minimum information principle*, to which we *minimize* the information i_λ of a specific event λ.

In order to elucidate the relation between the two information principles and, eventually, the role played by synergetics, let us consider some typical cases.

Motion of Multirobot Cells

We first describe the motion of a robot in a multirobot cell example that tries to build an organism. Each robot cell is furnished by adequate sensors for the extraction of environmental features that are relevant for navigation. Here, ξ_s might be the free path length. We denote the probability that in the presence of an obstacle λ the free path length is ξ_s by $p_\lambda(\xi_s)$. In this way, a robot can calculate every time its individual information i_λ and it will navigate by minimizing $i_\lambda(\xi_s)$ – maximum probability $p_\lambda(\xi_s)$. This means that it will try to find such a path where the probability of an unexpected "obstacle event" is small since this event is rare. It favors that route where it knows the positions of obstacles (*a priori*) and it experiences as few as possible surprises. Each individual information constraint is defined by maximum $p_\lambda(\xi_s)$.

The maximum principle of information chooses those probability distributions that maximize the total information $i_S(p(\xi_s)) = \sum_\lambda i_\lambda(\xi_s) p_\lambda(\xi_s)$. Indeed, this formula looks like the simple summation and weighting of the aforementioned local information, but we maximize now by other constraints like

$$p_1\xi_1 + p_2\xi_2 + \cdots + p_N\xi_N = \hat{\xi}, \tag{I.12}$$

where $\hat{\xi}$ is the expected path length (mean path length) of a robot. Under this constraint and the normalization constraint, we obtain as a result for the probability distribution $p_\lambda(\xi_s)$ the famous Boltzmann distribution of statistical mechanics, where we just must replace energy level $\in_\lambda$ by path length ξ_s. The resulting movement of the individual robots will now be dictated by their endeavor to fix the total expected path length $\hat{\xi}$. Here, we know that this kind of maximal system information is not directed toward our original goal to assemble an organism.

Assembly of an Organism

To fulfill this requirement, we have to consider further constraints (moments) like the normalization

$$\sum_{\mu=0}^{1} p_{\lambda\mu} = 1,\ \lambda = 1, 2, \ldots, N \tag{I.13}$$

and

$$\sum_{\lambda=0}^{N}\sum_{\mu=0}^{1}\mu\, \mathrm{p}_{\lambda\mu} = A, \tag{I.14}$$

where each of the N robot cells ($\lambda = 0, \ldots, N$) can occupy only one or no position ($\mu = 0, 1$, like in the Fermi–Dirac distribution), and let $p_{\lambda\mu}$ be the probability that robot λ is on an organism position or not (considered as an event). A defines the expectation value how many robots are in a position that is part of the organism. In addition, we have to claim for a stability constraint, an energy constraint, and so on for the case of a stable and "viable" final organism (Levi and Haken, 2010).

For the calculation of the total information, all participating robots have to exchange the individual results among each other as in a classical distribution system that might be very time consuming. Another approach is a centralized approach where the maximized probability distribution is calculated at one local computer system and then distributed to all robots.

Up to now, a connection between the maximum and the minimum information principle has not been established in our description. This missing link can be defined by the synergetic method of self-organization (slaving principle) that we will deal with later on in this section and in more detail in our book. In a natural parlance such a self-regulating process can be pictured as follows. Assume that the individual robot cells have not enough power to continue their exploration tour, and then they can start to signal to each other that they should begin to lump together in a kind of organism in order to decrease the total power consumption and in this way a change arises that all together can survive. In basic biological systems, such an assembly process increases the concentration of a chemical field (e.g., concentration of cyclic adenosine monophosphate (cAMP) for slime molds) that operates as an order field (organisator) since the increasing gradient of this field (concentration wave) guides the individual cells to a center where the slime mold organism is formed.

Let us go back to our two information principles and transfer this biological example to a higher and more general level. A few of the independent navigating robots stop their tour since the local power supply goes rapidly down and by the support of some fluctuating forces they lump together. Such an unstable small body may launch the generation and calculation of the total maximum information of the stochastic assembled body. The resulting system information $i_{\max}(p)$ chooses then such a probability distribution that not only stabilizes the mean path length $\hat{\xi}$ but also fixes the mean energy consumption, and furthermore the relative distance between two neighboring robot cells will be and the maximum information can be involved to obviate that two cells are on the same position (expression (I.7)).

The circularity in this process is started by the task of the individual robots, the external events are the obstacles (the corresponding stochastic variable ξ_s defines the path length) and an internal event might be the status of the power supply (stochastic variable $\in_\lambda$). So the extended local information can be written $i_\lambda = -\ln p(\xi_s, \in_\lambda) = -\ln\left(p(\in_\lambda)p(\xi_s|\in_\lambda)\right)$, where the energy status will be considered as a local order parameter (see below), and this information will be minimized. If we sum up all these local minimal information $\sum_{\lambda} p_{\lambda} i_{\lambda}$ to minimal mean information, then we must notice that this formula looks, in fact, like the one that is used for maximum information calculation, but the meaning is completely different. The maximum information principle picks up the two stochastic variables ξ_s and $\in_s$ but it considers additional constraints like the requirements of (I.12) and the total energy balancing

$$\sum_{\lambda=0}^{N} \in_{\lambda} \sum_{\mu=0}^{1} \mu p_{\lambda\mu} = B. \tag{I.15}$$

In addition, in the next formula (I.16) where only one ξ_{u} is mentioned as an order parameter, we must also introduce additional order parameters like the power status $\in_{\lambda}$.

The global maximized distribution function has been already shown by expression (I.11) for one variable ξ_{u} (e.g., distance between two roboter cells) and one roboter cell's path length ξ_{s} (or roboter position). As is known from probability theory, the relation between joint probability $p_{\lambda}(\xi_{\mathrm{s}}, \xi_{\mathrm{u}})$, conditional probability $p_{\lambda}(\xi_{\mathrm{s}}|\xi_{\mathrm{u}})$, and probability $p_{\lambda}(\xi_{\mathrm{u}})$ is given by (cf. also (I.2))

$$p_{\lambda}(\xi_s, \xi_u) = p_{\lambda}(\xi_u) p_{\lambda}(\xi_s|\xi_u). \tag{I.16}$$

Some Basic Concepts of Synergetics (Haken, 2004)

In synergetics, formula (I.16) and its variables acquire a specific meaning, and in many cases, the methods of synergetics allow us to calculate the expression on the r.h.s of (I.16) as we will show in our book. The basic idea behind (I.16) is this: in a system composed of (in general) many interacting components (e.g., robots) specific spatial or spatiotemporal configurations (an "organism") can be formed spontaneously by self-organization (i.e., without an external ordering hand). Rather, the system itself establishes one or several collective variables ξ_u that are the *order parameters*.

These order parameters *enslave* the individual components with their variables ξ_s, that is, they determine the behavior of the latter. This is expressed by $p(\xi_{\mathrm{s}}|\xi_{\mathrm{u}})$, (slaving principle). By means of their cooperation, in turn the components determine the dynamics of the order parameters (circular causality). In a number of cases, the variables ξ_s can be eliminated from the fundamental equations of motion so that closed equation for ξ_{u} result. This is reflected by $p(\xi_{\mathrm{u}})$. As we will show in our book, the slaving principle allows us to bridge the gap between the concepts of minimal and maximal information principles. To be sure, this little sketch represents only a small, though characteristic, part of the synergetic methodology. After this interlude, let us continue the previous section.

The mathematical expression (I.16) demonstrates that the individual navigation that can be represented by a maximum probability $p_{\lambda}(\xi_s)$ is replaced by a different maximized total probability $p_{\lambda}(\xi_s, \xi_u)$ that describes the behavior (movement) of a robot cell under the control ("slaving") of the maximal information generated by the participating agents that exercise a strong feedback (causal circularity). In the light of the presented information principles, we consider the maximum information as a superior organizing field that controls the local minimal information of the individual agents.

As an important consequence, this means we do not need a centralized system in order to describe self-organization phenomenon, but after an establishing phase (generation of the maximum information) the agents can immediately move around without any continuous exchange of partial information values of the total maximum information. Like in a gravitation field that is generated only by the participating masses, the individual masses behave coordinated (synchronized) without any explicit information exchange in order to calculate the total information.

The synergetic-based interaction between both optimization principles of information can be extended to objects (organisms) of higher complexity than a pure mass object might represent. The appeal of our approach is the fact that on every hierarchical level (higher semantic level) the same mathematical expression must be calculated; however, the meaning of the information expressed by the probability distribution that minimizes the local information and another one that maximizes the global information is on every hierarchical level entirely different. The scientific challenge is here to formulate the correct restrictions and to find the adequate probability distribution (indeed not a simple task).

Even if we find the correct probability distributions, we must express their dependence on the external or internal event λ by the involvement of the so-called control parameters.

A well-known example for such dependence is the laser distribution function of the order parameter ξ_u

$$p(\xi_\mathrm{u}) = N_\mathrm{u}\exp\{\alpha\xi_\mathrm{u}^2 - \beta\xi_\mathrm{u}^4\}. \tag{I.17}$$

Here, α and β are both control parameters, where the sign of α defines a "bifurcation" of ξ_u. An example for the probability distribution that minimizes the local information is

$$p_s(\xi_s) = N_s\exp\{-\gamma(\xi_s - f(\xi_u))^2/Q_s\}, \tag{I.18}$$

where Q_s is the strength of fluctuating forces (corresponding to a constant diffusion coefficient), γ a damping constant, and f a third-order polynomial in ξ_u. Since the maximum of (I.18) lies at

$$\xi_s = f(\xi_u),$$

this clearly demonstrates the slaving principle: the enslaved variable ξ_s is fixed by the order parameter ξ_u (up to a finite uncertainly due to fluctuations as expressed by the width of the Gaussian).

The synergetic combination of both information principles will be of particular interest to robotics. In order to formulate a motion equation from the minimal or maximal information principle, we consider the probability distribution, in the first step, as the stationary solution of a Fokker–Planck equation (cf. e.g. (Risken, 1989)). This means that we use the following fundamental equation for the local information $i_s(\xi_s)$:

$$F_s(\xi_s) + Q_s \nabla \mathrm{i}_s(\xi_s) = 0, \tag{I.19}$$

with the solution $i_s(\xi_s) = V(\xi_s)/Q_s +$ constant, and with the force $F_s = -\nabla V_{\underline{s}}(\xi_s)$, where ∇ is the nabla operator.

By this approach, we obtain using the principle of minimum information the equation of motion, for example, for an individual robot by the expression

$$\dot{\xi}_s = -Q_s \nabla i_s(\xi_s), \text{ where } i_s(\xi_s) = -\ln p_s(\xi_s). \tag{I.20}$$

For sake of consistencey with the fully stochastic equation of motion, the r.h.s. of (I.20) must be supplemented by a stochastical force of strength Q_s. We formulate this special result in a more general way by the following principle.

Haken–Levi Principle

In a multirobot system (multiagent system), each robot moves (behaves) in such a way that it minimizes its local information, if its motion is overdamped and subject to a fluctuating force (Langevin equation).

This is a special case of the Haken–Levi principle that we formulated above (I.7) for the general motion of a robot with mass, subject to a fluctuating force and damping (H-L II). For massless agents, the acceleration term disappears, or this term can be neglected if the motion is overdamped (H-L I).

The great advantage of expression (I.7) is that robots can immediately move (e.g., in a swarm mode) if their local information is available without any or (a minimal amount) of additional time-consuming message exchanges. In the organism mode, when artificial creatures with higher complexity grow up, we still assume that formula (I.7) is correct and applicable if we consider stochastic variables that do not describe positions or distances between agents but adaptability (fitness) to the change of different environmental factors such as pressure, temperature, humidity, and slippery ground. All these parameters are directly connected to self-regulation problems of homeostasis or foraging. On higher levels, we have also to consider differentiation processes (definition, for example, of "organs" or body parts) or cognitive and decision processes. In the view of H-L principle, we postulate that we describe all these more sophisticated processes finally by information-based activity patterns that can be deduced from adequate potentials, where we do not exclude that our scalar potentials must be replaced by vector potentials or even by gauge potentials $\mathcal{A}$. In the latter case, the forces have to be replaced by corresponding local field strength $\mathcal{F}$ (Naber, 1997).

References

Braitenberg, V. (1984) *Vehicles: Experiments in Synthetic Psychology*, MIT Press, Cambridge.

Floreano, D. and Mattiussi, C. (2008) *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*, The MIT Press, Cambridge.

Genz, H. (1996) *War es ein Gott ?*, Hanser Verlag München, Wien.

Haken, H. (2004) *Synergetics: Introduction and Advanced Topics*, 3rd edn, Springer, Berlin.

Haken, H. (2006) *Information and Self-Organization*, 3rd edn, Springer-Verlag, Berlin Heidelberg.

Jaynes, E.T. (1957) Information Theory and Statistical Mechanics, *Physical Review*, **106** (4), 620–630; Information Theory and Statistical Mechanics, *Physical Review*, **108**, 171–190.

Kahneman, D. and Tversky, A. (1996) On the reality of cognitive illusions. *Psychological Review*, **103** (3), 582–591.

Kapur, J.N., Kesavan, H.K. (1992), *Entropy Optimization Principles with Applications*, Academic Press., Boston.

Kernbach, S. (2010) From robot swarm to artificial organisms: self-organization of structures, adaptivity and self-development, in *Symbiotic Multi-Robot Organisms: Reliability, Adaptability, Evolution* (eds P. Levi and S. Kernbach), Springer-Verlag, Berlin Heidelberg, pp. 5–25.

Kullback, S. (1951) *Information Theory and Statistics*, John Wiley & Sons, Inc., New York.

Levi, P. (2009) Development of evolutionary and self-assembling robot organisms, in Proc. of the 20th Anniversary IEEE MHS-2009 and Micro Global COE, Nagoya, Japan, pp. 1–6.

Levi, P. (2010) Application of embodied cognition to the development of artificial organisms, in *Symbiotic Multi-Robot Organisms: Reliability, Adaptability, Evolution* (eds P. Levi and S. Kernbach), Springer, Berlin, pp. 202–228.

Levi, P. and Haken, H. (2010) Towards a synergetic quantum field theory for evolutionary, symbiotic multi-robotics, in *Symbiotic Multi-Robot Organisms: Reliability, Adaptability, Evolution* (eds P. Levi and S. Kernbach), Springer, Berlin, pp. 25–54.

Levi, P. and Kernbach, S. (eds) (2010) *Symbiotic Multi-Robot Organisms: Reliability, Adaptability, Evolution*, Springer, Berlin.

Naber, G. (1997) *Topology, Geometry, and Gauge Fields*, Springer–Verlag, Berlin Heidelberg.

Penrose, R. (2006) *The Road to Reality*, Alfred A. Knopf, Inc., New York.

Pfeifer, R. and Scheier, Ch. (1999) *Understanding Intelligence*, The MIT Press, Cambridge.

Rajaie, H., Zweigle, O., Haeussermann, K., Käppeler, U.Ph., Tamke, A., and Levi, P. (2011) Hardware design and distributed embedded control architecture of a mobile soccer robot, in *Mechatronics* **21**(2), Elsevier, Amsterdam, 455–468.

Risken, H. (1989) *The Focker–Planck Equation*, 2nd edn, Springer, Berlin.

Shannon, C.E. (1948) A Mathematical Theory of Communication. *Bell System Techn. J.* **27**, 370–423, 623–656.

Shannon, C.E., Weaver, W. (1949) *The Mathematical Theory of Communication* (Univ. of. Illin. Press, Urbana).

Shen, W.-M., Will, P., Galstyan, A., and Chuong, A. (2004) Hormone-inspired self-organization and distributed control of robotic swarms. *Autonomous Robots*, **17** (1), 93–1005.

Siciliano, B. and Khatib, O. (eds) (2008) *Springer Handbook of Robotics*, Springer, Berlin.

Thrun, Burgard, and Fox (2005) Probabilistic Robotics, The MIT Press, Cambrigde, Massachusetts.

Weiss, G. (ed.) (1999) *Multiagent Systems*, The MIT Press, Cambridge.