1

Introduction

Webster's Dictionary (2001) defines *intelligence* as the capacity for learning, reasoning, understanding, and similar forms of mental activities; the aptitude in grasping truths, relationships, facts, meanings, etc. The ability to understand the meaning is present in many other definitions of *intelligence*. The *Larousse Dictionary* (1992) adds yet another component to *intelligence*: the ability to adapt. *Defining the meaning of events in a measurable fashion and finding a measure for the capability of a system to adapt are central topics of this book*.

The concept of *artificial intelligence* (AI) is disputed. The definitions of AI have changed to reflect the changes in understanding and scope, which happened as the field of AI evolved. Many attempts have been made to define it, and these definitions could be grouped in several categories. One category is related to the Turing test proposed by Alan Turing (1950). Turing's test was meant to establish criteria to compare an artificial intelligent system with a human with respect to their level of knowledge. Systems that would pass Turing's test would be considered "intelligent." This type of definition focuses on the cognitive aspects of intelligence.

Bellman (1978) took these definitions a step further by adding to cognition typical activities associated with human thinking, e.g., problem solving and decision making. In 1993, Luger and Stubblefield defined AI as a branch of computer science. In their approach, AI is a discipline that studies the automation of "intelligent behavior."

It can be noticed that while the definition of human intelligence does not address an acting or behavior component, the AI definitions do. This is due to the fact that AI is an engineering field, which has to design and build systems required to perform tasks, i.e., to act. It can also be noticed from the literature that many definitions refer to AI systems as computer programs and to AI as a subfield of computer science. While this view has historical roots, in a modern approach AI deals with building complex engineering systems, which perceive, reason, and act. Therefore, for the purposes of this book, the definition proposed by Russell and Norvig (1995) is considered the most adequate: AI deals with the study and construction of artificial agents, which accomplish their tasks with different degrees of success based on knowledge they have built in and on the information they receive through perception. In this view, the actions of an artificial intelligent depend on its perceptions, its built-in knowledge, its ability to act, and, last but not least, on the type of success measure that the agent is using. This definition adds a major component to the intelligent behavior: the idea of maximizing a performance measure, i.e., the idea of a successful behavior.

Intelligent agents have been extensively described in the artificial intelligence literature, for example by Huhns et al. (1998), Jennings and Lesperance (2000), or Brenner et al. (1998). As described by Russell and Norvig (1995), an intelligent agent is one that optimizes its behavior in order to be successful. This description leaves two questions open: How can the agent's success be evaluated? And when should the success evaluation take place? The term *performance measure* is used for the criteria that determine how successful the agent is. Also important is when the performance measure is applied since at different times the agent may be successful at different degrees.

Currently, there does not exist a consistent approach for measuring and predicting a successful agent and how to improve him. Not one single agent-independent performance measure exists. In terms of assessing an agent's performance, the literature is scarce. One relevant approach is related to the concepts of utility and the value of information. In this approach, a utility function is used to capture the agent's preferences among different possible courses of action. The utility function assigns a number to express the desirability of a state or of a course of action. As in the game theory, utilities are combined with outcome probabilities for actions in order to give a weighted utility average value called *expected utility*. In the case of a decision-making agent, the expected utility can be used as a performance measure, such that an intelligent agent chooses a course of action that maximizes the agent's expected utility. Another performance measure is introduced by Russell and Norvig (1995) in the case of an information-gathering agent. By assigning a value and a cost to observed information, the agent uses the value-to-cost ratio to repeatedly select the observation with the highest information value, until the cost of observing is greater than the benefits. In another example found in the literature (Cover, 1991), the information theory is combined with the utility theory for the purpose of optimizing investment decisions.

While it is generally accepted that intelligent agents must receive information – for example, through perception – and must have built-in knowledge in order to be able to react to this information, e.g., by making decision, the concepts of information and knowledge and their relationship have not yet been consistently formalized.

In terms of information, Shannon, in his *Mathematical Theory of Communication* (1948), makes the assumption that the semantic aspects of communication are irrelevant to the engineering problem of communication. Under this assumption, Shannon's concept of information and his definition of entropy as a measure of information, choice, and uncertainty have immensely contributed to the development of communications. At the same time, these concepts have been less instrumental in the engineering of intelligent systems and of those fields, such as artificial intelligence, which are interested in measuring and capturing aspects of reality, like utility and context dependency.

An approach, which this thesis builds on, is represented by the information value theory. According to this theory, the value of information derives from the fact that information may contribute to optimizing one's course of action with respect to the actual situation. The information received by an agent can help the agent discriminate according to the situation, whereas without the information the agent has to do what's best on average, i.e., given a certain level of common knowledge. In general, the value of a piece of information is defined as the difference in expected utility of the agent's actions between actions executed with and without information.

In terms of knowledge, the methods currently in use show a scattered picture when it comes to defining and measuring, i.e., quantifying the amount of knowledge built in or processed by intelligent agents. Typically, the agent's performance is related to the knowledge representation (Ben-Eliyahu-Zohary and Palopoli 1997 or Levy et al. 1997) and is measured in terms of size and time. For example, the number of inference rules or the time required for a graph search or for a matching algorithm. Some qualitative approaches relate to assessing the accuracy of the agent's utilization of knowledge (e.g., Golding and Rosenbloom 1996). An information-theoretical approach has been suggested by Grunwald (1991-a and 1991- b) in order to assess the failure potential and the efficiency of Bayesian inference.

A new performance measure for evaluating an agent's performance and determining the criteria for when to apply it is introduced in this book. The new performance measure is called *efficiency*, and it is established and analyzed within a new mathematical framework introduced herein. The efficiency is deduced herein independent of any particular agent and can be used by particular agent implementations as a criterion for performance improvements. The *cycle of operation* or the diagnostic cycle of an intelligent agent is defined herein. It is shown how, during a cycle of operation, an intelligent agent can receive an amount of information I, related to the amount of U, the agent's built-in knowledge. The amount U is introduced as a function of the *volume* and of the average *expected utility* of the built-in knowledge. It is also described herein how the agent generates an amount of *diagnostic knowledge*, K, related to I and U, and how, due to limited resources, the agent must prepare for the next cycle of operation. The efficiency of an agent's cycle of operation is defined as $\eta = K_{\text{CYCLE}}/I_{\text{RECEIVED}}$, where K_{CYCLE} represents the net amount of knowledge exchanged, and I_{RECEIVED}, the total amount of information received by the agent throughout the cycle. The word *infodynamics* has been selected by the author to describe the mathematical framework introduced herein because of two reasons. On one hand, during the cycle of operation, the information that an intelligent agent processes has its own dynamics. The agent receives information, and based on this information, it generates knowledge. In addition to generating knowledge, the agent executes an appropriate action in response to the received information. As a result of its actions, the agent receives feedback knowledge. The word *infodynamics* captures the idea of the dynamics of information and its transformation into knowledge in intelligent agents. In a certain way, the new mathematical framework called "infodynamics" introduced herein deals with intelligent agents, information, and knowledge similar to how thermodynamics deals with heat engines, heat, and work. In paragraph 9.2, a parallel is drawn between several of the here- introduced concepts and thermodynamic ones. The thermodynamics concepts have been presented in accordance to Sonntag, Borgnakke, and Van Wylen (1998). In paragraph 9.3, a parallel is drawn among concepts available in the information theory. The information-theoretical concepts in use for communications systems are presented in accordance to Shannon (1948) and Cover (1991).