What is an intelligent system?

Martin Molina

Dept. of Artificial Intelligence, Universidad Politécnica de Madrid

February 2022

Abstract

The concept of intelligent system has emerged in information technology as a type of system derived from successful applications of artificial intelligence. The goal of this paper is to give a general description of an intelligent system, which integrates previous approaches and takes into account recent advances in artificial intelligence. The paper describes an intelligent system in a generic way, identifying its main properties and functional components, and presents some common categories. The presented description follows a practical approach to be used by system engineers. Its generality and its use is illustrated with real-world system examples and related with artificial intelligence methods.

1 Introduction

Mankind has made significant progress through the development of increasingly powerful and sophisticated tools. In the age of the industrial revolution, a large number of tools were built as machines that automated tasks requiring physical effort. In the digital age, computer-based tools are being created to automate tasks that require mental effort. The capabilities of these tools have been progressively increased to perform tasks that require more and more intelligence. This evolution has generated a type of tool that we call intelligent system.

Intelligent systems help us performing specialized tasks in professional domains such as medical diagnosis (e.g., recognize tumors on x-ray images) or airport management (e.g., generate a new assignment of airport gates in the presence of an incident). They can also perform for us tedious tasks (e.g., autonomous car driving or house cleaning) or dangerous tasks such as exploration of unknown areas (e.g., underwater exploration).

The development of such a type of systems is now an engineering discipline of information technology that requires effective methods and tools. The precise characterization of an intelligent system is non trivial because it is based on terms related to cognition, an area that is not fully understood and admits different interpretations. Some of the used terms can even change with the proposal of new computational models of intelligence and new scientific findings related to our understanding of the mind.

The main purpose of this paper is to present a characterization of an intelligent system that integrates and updates previous conceptions of such type of system. It follows a pragmatic approach to be useful in an engineering context to help system engineers conceive, analyze and build intelligent systems by using concepts and terminology commonly accepted in artificial intelligence. Therefore, the presented description focuses on functional characteristics related to the intelligence that machines can exhibit with the current state of the technology.

The remainder of the paper presents the following contents. Section 2 presents a definition of intelligent system with the description of its properties and its main functional components. Section 3 describes aspects in the development process of an intelligent system such a types of artificial intelligence methods used to develop the different components of intelligent systems. Section 4 describes classes of intelligent systems using the presented characterization, covering both recent achievements and classical approaches of this technology .

Definition 1: Intelligent agents are computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment, and by doing so realize a set of goals or tasks for which they are designed (Maes 1995)
Definition 2: Intelligent agents continuously perform three functions: perception of dynamic conditions in the environment; action to affect conditions in the environment; and reasoning to interpret perceptions, solve problems, draw inferences, and determine actions (Hayes-Roth 1995)
Definition 3: Intelligent agents operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, and create and pursue goals (Russell and Norvig 2014)
Definition 4: Intelligent agents are autonomous or semi-autonomous hardware or software systems that perform tasks in complex, dynamically changing environments. Agents communicate with their environment and effect changes in their environment by executing actions (Muller 1996)

Table 1: Sample of definitions of intelligent agent.

2 Definition of intelligent system

Intelligent systems have been characterized in the literature of AI using the concept of *intelligent agent* (Wooldridge and Jennings 1995) (Franklin and Graesser 1996) (Russell and Norvig 2014). This concept emphasizes the ability to act[1] on an environment. Table 1 shows a sample of definitions of intelligent system considered as an agent. These definitions highlight that the system works in an environment, has a set of capabilities (perception, learning, etc.), and makes decisions about how to act.

The word *system*, instead of agent, is used by some authors (Meystel and Albus 2002) and in academic areas of information technology such as the names of university courses or academic journals (e.g., IEEE Intelligent Systems and International Journal of Intelligent and Robotic Systems). In this case, the word system emphasizes the presence of multiple components that must be adequately combined to create intelligence. Knowing how to do this combination efficiently is one of the key aspects in the development of such type of systems.

The characterization of an intelligent system used in this paper integrates parts of the previous agent-based definitions. For example, the identification of a separate environment where the agent works is important to give an adequate operational context at a certain degree of abstraction. In our definition, we also distinguish from the environment other agents that may interact with the system (e.g., a human user or other artificial systems) in order to consider the social factor of an intelligent system.

This paper also uses a set of capabilities usually enumerated by agent-based definitions (e.g., perception, reasoning, learning, etc.). However, we separate these capabilities into two parts. On the one hand, there are four primary cognitive abilities: perception, action control, deliberative reasoning, and language use (to interact with others). On the other hand, we distinguish the capacity to adapt to changes in the world through learning. This capacity is applicable to each primary cognitive ability.

Finally, our characterization of an intelligent system takes also into account that the system operates according to two behavioral principles: acting rationally and following social norms. The first principle considers that the system behavior is consistent with the individual goals to be achieved (see the following sections). The second principle considers that the system has certain ethical behavior to follow the norms established by the social group with which it cooperates.

According to this, the definition that we follow in this paper a basic scheme to structure the further characterization of an intelligent system is the following:

Definition. An intelligent system (1) operates in an environment with other agents, (2) possesses cognitive abilities such as perception, action control, deliberative reasoning or language use, (3) follows behavioral principles based on rationality and social norms, and (4) has the capacity to adapt through learning.

It is important to note that this definition is not intended to be a rigid characterization to determine whether a system is intelligent or not. Instead, the definition includes usual characteristics related to intelligence that may or may not be present in the analysis or development of specific intelligent systems.

---

[1]The essential meaning of the word *agent* can be expressed as "one who acts", based on its derivation from the Latin word *agens*.

For example, a simple case of intelligent system could be a simple reflex agent, as it is called by Russel and Norvig (Russell and Norvig 2014). This type of system possesses perception and control action abilities and it does not have other characteristics such as learning or deliberative reasoning. A collaborative autonomous robot is a more complex example of intelligent system. In this case, the robot interacts with the user and with the environment in which it operates. The robot has the cognitive abilities of perception, action control, deliberative reasoning (e.g., for motion planning) and language use (to communicate with the user). It can also be assumed that the robot has adaptive capabilities with the help of machine learning techniques.

The following sections describe in more detail the four properties of our definition of intelligent system, presenting additional concepts and categories in which the elements of a system may be classified.

2.1 Property 1: Working in an environment with other agents

As mentioned in the previous section, one of the properties of an intelligent system is that it works in an environment together with other agents with which the system can communicate (e.g., a human user, artificial computer-based agents, etc.). The system observes features from the environment through sensors and performs actions using actuators (see Figure 1). The use of sensors and actuators (real or virtual) separates the body of the intelligent system from the rest of the environment. This characteristic is called *embodiment*.
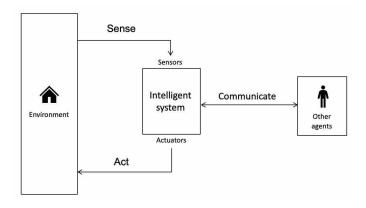

Figure 1: An intelligent system interacts with the environment and with other agents.

Example. A thermostat is a simple example that illustrates the concept of this kind of interaction. The goal of the thermostat is to regulate the temperature in the environment (e.g., house rooms). The thermostat uses a thermometer to sense the temperature of the room, and actuates by starting or stopping a heater. The thermostat also communicates with a human user who starts and stops the thermostat, or establishes the desired temperature.

2.1.1 Environment

Usually, the complexity of the environment is significantly higher than the information processing capacity of the system. The amount of components of the environment, their interactions or the anticipated effect of potential actions in the environment cannot usually be completely represented in the memory of the intelligent system. For example, in the chess game, it is not possible to anticipate the exact effect of all potential movements, because the amount of combinations is too high. The degree of complexity of the environment with respect to the agent can be described with the following features (Wooldridge 2009; Russell and Norvig 2014):

- *Static* (or *dynamic*), the environment does not change (or changes) while an agent is deliberating,

- *Discrete* (or *continuous*), the state of the environment, time, percepts or actions are discrete (or continuous),

- *Fully-observable* (or *partial-observable*), sensors detect (or do not detect) all aspects that are relevant to the choice of action,

- *Deterministic* (or *stochastic*), the next state of the environment is (or it is not) completely determined by the current state and the action,

- *Episodic* (or *sequential*), actions do not influence (or they influence) future actions,

- *Known* (or *unknown*), the outcomes for all actions are known (or they are not known) by the agent in advance.

Example. The environment of a chess player is static, discrete, fully observable, deterministic, sequential and known. In the case of a self-driving car, the environment is continuous, partial observable, stochastic, sequential and known.

2.1.2 Other agents

An intelligent system may interact with other agents as part of an organization. Such an organization can be viewed as a multi-agent system in which individual agents interact using social coordination mechanisms to cooperate to achieve common goals or to compete for limited resources.
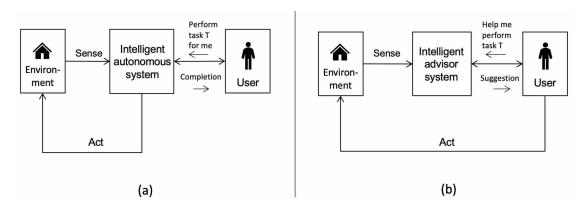

Figure 2: Comparison between (a) autonomous system and (b) advisor system.

An artificial intelligent system interacts at least with a human user that acts as the system's owner. The user starts and stops the execution of the intelligent system and may adjust specific parameters to modify the the system's behavior (for example, the owner of chess player can adjust the level of the game). Depending on who is acting on the environment (the system or the user), an artificial system can play two different roles during its interaction with the user (Figure 2):

- *Autonomous.* An autonomous system *acts in the environment to help the user.* This means that the user delegates to the system a task to be performed and the system makes its own decisions on how to act in the environment to autonomously perform that task. An autonomous system can also reject requested tasks from the user according to certain reasons, based on the current situation of the environment or its own goals (e.g., safety goals or social norms). For this purpose, the system can verify the correctness and feasibility of a requested task before it is performed and explain to the user the reasons that justify why a requested task is rejected.

- *Advisor.* An advisor system *helps the user act in the environment.* In this case, the user is the one who makes the decisions about what actions to do in the environment. The role of the system is to give advice by providing useful information to facilitate such decisions to the user. The system may help the user in prescriptive tasks such as planning, scheduling or resource assignment, but also in analytic tasks such as problem detection, problem diagnosis and temporal projection.

During the interaction, an intelligent system may be *proactive* instead of passive[2]. This means that the system does not have to wait until a user requests a task, but it takes the initiative to perform a task based on its own goals and what it perceives from the environment.
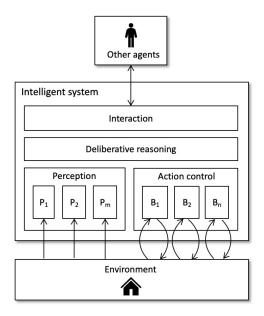

Figure 3: Primary cognitive abilities of an intelligent system.

2.2   Property 2: Having primary cognitive abilities

In this paper we distinguish four primary cognitive abilities that are usually present in intelligent systems:

- *Perception.* Perception may be understood as the interpretation of sensory information to form a mental representation of the world (Schacter et al. 2020). For example, a self-driving car senses the world with the help of multiple sensors (lidar, cameras, accelerometers, acoustic sensors, etc.). Perception is distributed in multiple perceptual components in such a way that each component recognizes a kind of object (e.g., pedestrians, traffic signals, other vehicles, etc.) and updates a dynamic representation of the environment. In general, the perceived world may include parts of the own body that can also be sensed by the system[3].

- *Action control.* Action control refers to the ability of an agent to control the execution of its own actions. Action control sends orders to actuators and uses data from the environment measured with specific sensors operating in a continuous closed loop. For example, the speed control of an autonomous vehicle acts on the accelerator by continuously observing data from a sensor that measures speed. The global control may be divided into a set of *behavior controllers* and each one is in charge of a specific control aspect (e.g., braking, steering, accelerating, etc.)[4].

- *Interaction.* The ability to interact with other agents is based on using a common language. Artificial systems may communicate with users using simplified mechanisms (e.g., graphical user interfaces, prefixed commands, etc.). For example, in a self-driving car, the passenger may indicate the destination to go and the vehicle presents the path to follow together with information about the environment that helps the passenger monitor the correct execution of the trip. Besides these mechanisms, interaction with users can also be based on human language using algorithms on natural language processing.

---

[2]Some authors use the term *reactive* as the opposite of proactive. However, in this paper we will use the term reactive as the opposite of deliberative as it is generally used in agent-based systems.

[3]This type of perception is called *proprioception* (perception of self-movement and body position).

[4]The behavior of certain kinds of intelligent systems may be understood as the combination of multiple behavior controllers as it has been described by the behavior-based paradigm in robotics (Arkin 1998) (Brooks 1991).

- *Deliberative reasoning.* An intelligent agent may be able to reason in order to, for example, plan actions (e.g., find a sequence of steps to travel from one place to another) or diagnose problems (e.g., determine the cause of a failure). In this paper, we use the term deliberation to refer to this type of ability, as it is used by different authors (Wooldridge 2009) (Ingrand and Ghallab 2017) (Murphy 2019).

As Figure 3 shows, perception and action control may be divided in multiple separate components that usually operate in parallel at a high frequency. However, deliberation and communication usually operate sequentially which is necessary for coherent reasoning (Laird, Lebiere, and Rosenbloom 2017) and consistent communication.

2.2.1   Reactive behavior versus deliberative reasoning

The global behavior that emerges from individual behavior controllers is reactive in the sense that it is able to react quickly to events of the environment. The immediate feedback of the state of the world is done through sensors[5]. This reactive behavior is observed, for example, in reflexes of animals with primitive forms of brains. It can also be observed in human intuitive thinking, when decisions are based on intuitions, i.e. based on feelings rather than on facts or evidence.

As an alternative of such behavior, deliberative reasoning uses imaginary information about the world in order to assess alternative hypotheses. For example, chess players imagine the consequences of the moves before deciding on the most appropriate one. Deliberative reasoning uses beliefs, i.e., propositions that the the agent considers true (permanently or provisionally). New beliefs may be inferred by logical deduction or by other forms of reasoning (induction, analogical reasoning, etc.).

The conclusions achieved by deliberative reasoning may be used to activate and inhibit certain behavior controllers. Deliberation can manage conflicting goals and can inhibit reactive behaviors when it is necessary to postpone immediate answers in order to reach long-term goals with more priority. For instance, one person may accept deliberately suffering certain pain during a medical treatment with the expectancy of a future health improvement.

The distinction between deliberative reasoning and reactive behavior has been identified in human thinking by Daniel Kahneman (Kahneman 2011) with the names System 1 and System 2. System 1 (or intuitive thinking) is fast, intuitive, automatic, unconscious and non-linguistic (difficult to verbalize). On the contrary, System 2 (or deliberative reasoning) is slow, logical, effortful, conscious, linguistic (can be described verbally) and makes complex decisions. In the development of computer-based intelligent systems, the combination of deliberative reasoning and reactive behavior is a non trivial problem that requires specific design decisions (e.g., when to favor deliberation over reaction).

2.2.2   Use of symbols

Using symbols has been been pointed out as a relevant feature of intelligence. For example, the philosopher Ernst Cassirer stated that humans should be defined as symbol-making animals. Allen Newell and Herbert Simon proposed the physical symbol hypothesis, i.e., a physical symbol system has the necessary and sufficient means of general intelligent action (Newell and Simon 1976), which was a dominant paradigm in artificial intelligence until the 1990s. Gary Marcus argues that symbolic manipulation seems to be essential for human cognition (Marcus 2020).

Symbols are present in the language used for agent interaction. For example, human language is characterized by the use of symbols in the form of words or other representations (e.g., mathematical notations). Symbols represent classes of physical objects or abstract concepts. Agents agree in advance a shared meaning of the symbols in order to understand the language used in the communication process.

An agent may establish the meaning of symbols through perception (to associate sensory information to symbols) and control action (to translate symbolic commands into actions). This is related to the symbol grounding problem (Harnad 1990), i.e., how to ground the meaning of a symbol in anything different than other symbols. This problem has been addressed in robotics, for example, with the idea of perceptual anchoring (Coradeschi and Saffiotti 2002) under which a symbol (e.g., the concept *tree*) and sensor data (e.g., the image of a particular tree observed

---

[5]It said said that the agent is *situated* in an environment when the agent operates in a close-coupled continuous interaction with environment in a continuous sequence of sense-act. A system conceived in this way may use a limited or nonexistent memory about the world.

through a camera) are related using data structures (called anchors) together with specialized handling procedures.

An agent can represent the beliefs used during reasoning with the help of symbols understandable to other agents. This is useful to explain to others one's own reasoning and the knowledge used to reach the conclusions, showing awareness of own mental states[6] (as humans perform via introspection). This property (sometimes called *explainability*) is a desirable characteristic for an intelligent system to be trusted.

2.3   Property 3: Following principles about rationality and social norms

As presented above, an intelligent system is expected to follow two basic principles of behavior. The first principle is acting rationally and affects the individual behavior of the system with respect to its goals and the environment. The second principle is about following social norms and affects the behavior of the system with respect to other agents. The following sections describe these principles in more detail.

2.3.1   Acting rationally

An agent acts rationally if the decisions it makes about its actions seek to maximize a measure of performance. This measure quantifies the degree of achievement of the goals pursued by the agent. For example, in the game of chess, players try to win their opponents. A financial analyst makes decisions about investments to maximize the economic profit obtained. The decisions made by these agents are oriented to maximize a performance measure (number of opponents beaten or economic benefit obtained). This type of behavior has been described in artificial intelligence as the principle of rationality (Newell 1982) and it has also been analyzed in other disciplines such as philosophy or economy (Edwards 1954).

Acting rational is closely related to the notion of intelligence. Comparing performance measures in achieving goals is one way to compare degrees of intelligence. We say that a chess player A (or investment advisor A) is more intelligent than another player B (or another investment advisor B) if A wins more games (or makes more profit). This goal-based view is one of the approaches followed by different authors to define intelligence (Kurzweil 2000) (Goertzel 2006) (Legg and Hutter 2006). For example Kurzweil defines intelligence as "the ability to use optimally limited resources to achieve goals".

Note that this type of definition provides a method for testing intelligence based on the evaluation of a system in maximizing an objective function. This approach is flexible because it does not create an artificial rigid separation between intelligent and non-intelligent systems. In addition, it takes a practical approach for building intelligent machines, avoiding open scientific questions related to intelligence (e.g., phenomenal consciousness).

2.3.2   Social norms

The behavior of an artificial intelligent system with respect other agents is usually constrained by social norms, especially when it cooperates with human users. This is important to ensure that the system operates in accordance with ethical values such as ensuring fairness (avoiding group biases) or limiting harmful use.

For instance, an intelligent system trained using biased data may be not able to ensure a fair operation. This situation happened with Tay, an experimental conversational chatbot developed by Microsoft in 2016. This chatbot was trained with uncontrolled public data based on the interaction with people which generated a racist bias in the system. The ethical behavior of autonomous vehicles has been analyzed by Bonnefon et al. (Bonnefon, Shariff, and Rahwan 2016) considering extreme situations in which the vehicles have to choose between running over pedestrians or sacrificing their passenger to save the pedestrians.

In this context, Russell argues that artificial intelligence systems should operate as beneficial machines that maximize the realization of human preferences in order to cope with the AI control problem[7] (Russell 2019). According to Russell, the actions of intelligent machines are expected to

---

[6]In psychology, self-awareness is understood as the ability to become the object of one's attention (Duval and Wicklund 1972). Self-awareness can also understood following a view of processing information about the self (Morin 2011) which may be appropriate to simulate this ability with computer-based systems.

[7]The AI control problem is the problem of how to build AI systems that aid rather than harm their creators.

achieve human objectives and, since these objectives are uncertain, machines will defer to humans asking permission, accepting correction and allowing to be switched off. In addition, the behavior of these machines should be constrained by rules and prohibitions just as the actions of humans are constrained by laws and social norms.

This type of behavior has also been analyzed, for example, using the concept of *benevolent agents* in social organizations like multi-agent systems (Castelfranchi 1998). Benevolence may be understood as a kind of goal adoption. An agent is benevolent if the agent adopts other's goals and the adoption of such goals does not help to achieve its own goals.

2.4 Property 4: Capacity to adapt

The capacity for adaptation allows an intelligent system to operate effectively in complex dynamic environments. In general, a system is said to be self-adaptive if it is able to modify the way it performs a certain task in response to changes in the environment so that the system can perform such task more efficiently.

A first approach to achieve self-adaptation is by a dynamic re-combination of behaviors (Hayes-Roth 1995) (Oreizy et al. 1999) (Molina and Santamaria 2021). In this case, the agent uses an adaptation logic with information about the effectiveness of its own behavior methods. In a given situation, the agent uses the adaptation logic to decide which combination of methods is the most appropriate to perform a given task, according to the current state of the environment.

A second approach for adaptation is through learning. The following sections describe in more detail two forms of adaptation (individual and social) that are based respectively on learning from experience and learning from others.

2.4.1 Learning from experience

A system with the capacity of learning is able to improve its performance in the course of multiple interactions with the environment. For example, a chess player can learn more effective strategies after playing multiple games. This approach to learning has been used in the context of machine learning by Tom Mitchell (Mitchell 1997), who formulated the following definition: "a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E".

An intelligent system that is able to learn needs a value judgement method to assess its behavior. The result of such assessment is used by the intelligent system to change its behavior in order to improve the assessment result in the future. The value judgement method can be an *external* agent who works as a instructor (or trainer or supervisor) observing and assessing the behavior of the intelligent system to provide feedback (e.g., with positive or negative rewards). The intelligent system can also use an *internal* method that judges its own actions to generate positive or negative rewards using, for example, simulated feelings (like pain or pleasure) or other more complex mechanisms based on self-observation,

As it is argued by Silver et al. (Silver et al. 2021), systems that learn through trial and error experience to maximize reward could help to acquire multiple abilities related to intelligence. Thus, learning should be understood as a distinguished capacity of intelligent systems that affect primary cognitive abilities (perception, action control, reasoning or interaction). For example, an aerial robot can improve action control by learning to land more precisely on a marked surface after multiple trials.

2.4.2 Learning from other agents

In addition to the approach described above, we can identify another form learning based on knowledge transfer between individuals. With this scheme, an agent A learns new beliefs from another agent B on the basis that agent B is able to describe its own knowledge with a symbolic language whose meaning agent A is able to understand. This form of learning, which allows groups of agents to acquire knowledge directly from other agents, accelerates social adaptation to the environment because some agents can save the effort of having to learn individually from their own experience.

For example, consider the belief expressed as "zebras are horse-shaped and have striped skin". An agent who has never seen a zebra might be able to recognize this animal using the symbolic description. This form of learning has been called *zero shot learning* (Xian et al. 2018), which

consists of learning to recognize new objects without any examples, only having their symbolic description.

3   Building intelligent systems

This section summarizes two important aspects in the construction of an intelligent system: (1) the selection of artificial intelligence methods and (2) the acquisition of the knowledge used by the system to perform the desired tasks (which can be done manually or with the help of automatic methods).

3.1   Artificial intelligence methods

The development of an intelligent system can require combining several artificial intelligence methods. Table 2 shows functionalities related to cognitive abilities together with a sample of common methods used for building intelligent systems.

These methods can be divided according to two distinguished approaches. One approach corresponds to methods using connectionist representations (e.g., convolutional neural networks, recurrent neural networks or deep reinforcement learning). The other approach uses symbolic methods following formal semantics based on logic (e.g., automated logic deduction, rule-based systems and constraint satisfactions algorithms), based on probability (e.g., bayesian networks) or hybrid approaches (e.g., fuzzy logic).

Symbolic and connectionist methods correspond to alternative approaches that have complementary strengths and weaknesses (Franklin 1995). In fact, there is a lack of understanding of how information processing performed by symbolic approaches can be mapped onto computations performed by connectionist methods. The term *computational explanatory gap* (Reggia 2014) has been used to express that it is unclear how the two approaches are related.

Cognitive abilities that are consciously accessible (such as processes related to deliberation or interaction) may be implemented using symbolic methods. These methods can describe the symbols that represent knowledge to justify how to reach a certain conclusion.

On the other hand, perception and action control are usually carried out in an unconscious way and they are better implemented with non-symbolic approaches such as connectionist methods. In practice, it is interesting to note that the implementation of unconscious sensorimotor skills require much more computation compared to deliberative reasoning. This has been stated as the *Moravec's paradox* (Moravec 1988).

3.2   Knowledge acquisition

Intelligent systems may use knowledge that is difficult to be formalized using procedural representations of computer languages such as Java, Python or C++. For example, consider a medical expert system that uses knowledge that relates symptoms and diseases. This knowledge may be modeled using symbolic declarative representations (e.g., rules) that are stored in a *knowledge base*. An inference algorithm processes such rules to diagnose the disease of a given patient. The inference algorithm is general and, therefore, it may be implemented using generic software tools. However, the rules of the knowledge model are specific and must be written for the particular expert system.

The construction of a knowledge base by developers is called *knowledge acquisition*. Many of the knowledge bases of classical expert systems represented heuristic knowledge that was acquired manually from domain experts and formalized using declarative representations (e.g., rules or first order logic). The creation and maintenance of such knowledge bases usually requires significant effort. Therefore, manual acquisition is limited for the development of knowledge bases that are not too large. This problem has been called *knowledge acquisition bottleneck*.

To facilitate this task, the field of ontology engineering in artificial intelligence has proposed technical solutions to formulate common representations, understood as ontologies, that allow sharing and reusing contents of knowledge bases across different intelligent systems.

3.3   Automatic methods for building intelligent systems

In contrast to manual knowledge acquisition, it is possible to apply automatic methods for building intelligent systems, which has gained popularity in the last decades due to the increasing availability

Cognitive ability	Example function	Examples of detailed functions	Examples of computational methods
Perception	Feature extraction	Image recognition, information extraction, signal processing, attention mechanisms	Neural networks (e.g., convolutional neural networks), pattern recognition methods
	Data interpretation	Data abstraction, data fusion, symbol grounding	Statistical data analysis, extended Kalman filters, perceptual anchoring
Deliberation	Descriptive reasoning	Classification, diagnosis, temporal projection (prediction), belief revision	Automated logic deduction, bayesian networks, rule-based systems, semantic networks, truth maintenance systems
	Prescriptive reasoning	Planning, scheduling, configuration, assignment, goal reasoning	Automated planning, constraint satisfaction algorithms, BDI models
Interaction	Language recognition and generation	Natural language understanding, speech recognition, natural language generation	Natural language processing methods (e.g., recurrent neural networks, transformers)
	Dialog management	Dialog coordination, discourse planning	Communicative act theory, interaction protocols, automated planning
Action control	Action selection	Goal-driven or event-driven action selection, action conflict management	Finite-state machines, behavior trees, Petri nets
	Action execution control	Motion control, manipulation control	Control theory, fuzzy control, deep reinforcement learning

Table 2: Functionalities related to cognitive abilities together with examples of computational methods used for building intelligent systems.

of large amounts of data and higher computational power. These methods can use information sources with different formats such as images (photographs or videos), structured data (e.g., data bases with alphanumeric data or temporal series) or non-structured text written in natural language.

Machine learning methods can be used to develop components that implement specific cognitive abilities that will be part of an intelligent system. For instance, a neural network can be trained using thousands of images of tumors labelled as positive and negative examples. The trained neural network can be used as part of an intelligent system that helps physicians recognize the presence of tumors. To build an aerial robot able to land on a moving surface, it would also be possible to apply a method based on deep reinforcement learning and a simulator of the robot operating in a virtual environment to pre-train a neural network able control the robot motion during landing on the surface. In order to build a rule-based expert system for medical diagnosis, a rule induction method could be applied to a data base with examples of diagnoses relating diseases and symptoms to generate candidate rules for the knowledge base.

Natural language processing methods can also be used as an automatic method to assist in the construction of intelligent systems. These methods can be used to extract knowledge from text documents written in natural language. For example, a large corpus of text documents corresponding to a professional domain can be analyzed using natural language processing techniques to generate instances of classes and and relations to be used by a question/answer system.

Note that these two methods, machine learning and natural language processing, can play different roles in the development of intelligent systems. On the one hand, they can be used offline as it is described in this section, playing a role of tools used by a developer to create automatically parts of the system. On the other hand, they can be used online, i.e., the final intelligent system includes components using such methods to implement the capacity to adapt (using machine learning methods) or the ability to interact in natural language (using natural language processing methods).

4   Examples of intelligent systems

This section presents illustrative examples of real-world intelligent systems using the characterization presented in this paper. In order to provide a global view, Figure 4 shows diverse intelligent systems organized in a hierarchy of categories. This hierarchy is not intended to be accurate and exhaustive but illustrative of the kind of existing systems. The following sections describe in more detail some of these categories.
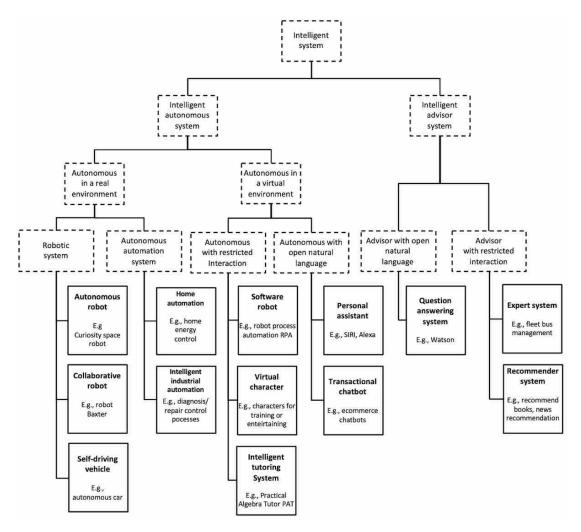
Figure 4: Examples of intelligent systems.

4.1 Autonomous robots

An autonomous robot may be considered as a representative case of intelligent system. An example of this system is an aerial robot for airplane inspection developed as a demonstrative prototype in 2019 for the company Airbus (Bavle 2019). This system performs the following functions:

- *Perception.* The aerial robot is equipped with multiple sensors such as (1) inertial measurement unit (IMU) to provide information about orientation, angular velocities, and linear accelerations, (2) cameras to calculate velocity using visual odometry, (3) a lidar sensor that generates a 3D point cloud, and (4) a high resolution frontal camera. The perception system performs data fusion from multiple sensors for a precise localization. The high resolution camera is used to detect anomalies on the surface of the airplane using computer vision.

- *Deliberation.* In this system, the main deliberative function is path planning. The system may generate an inspection plan to cover the surface of the airplane. In addition, the system stores images to be analyzed later.

- *Action control.* Action control corresponds to plan execution and motion control to complete the navigation plan. Action control also includes a reactive mechanism to avoid unexpected obstacles detected with the help of the lidar sensor.

- *Interaction.* The operator provides the geometry of the airplane and initiates the mission. The operator can stop the mission execution if an emergency happens.

In general, an intelligent robotic system is one of the most complex category of intelligent system because it usually includes the majority of the characteristics considered in this paper and,

in addition, it operates in real-world environments with demanding physical constraints. In this category, there are multiple kinds of systems with different kinds of complexity such as industrial collaborative robots (e.g., Baxter), autonomous vehicles (e.g., self driving cars, planet explorer Curiosity), assistive robots, domestic robots (e.g., house cleaning robot Roomba), entertaining robots (e.g., Sony Aibo, Furby), etc.

4.2   Expert systems

The development of expert systems contributed to the practical success of AI in the 1980s with many commercial systems in multiple domains. They were applied to perform automatically tasks such as the following:

- *Classification*: Find the category of an object using observations (e.g., determine the presence of a mineral, recommend a type of investment, detect abnormal behaviors).

- *Diagnosis:* Find the causes of symptoms (e.g., medical diagnosis, mechanical diagnosis).

- *Temporal projection*: Estimate the future behavior of a dynamic system (e.g., economic market prediction).

- *Assignment/Scheduling*: Assign a set of resources to a set of needs (e.g., assign airplanes to gates in a airport).

- *Configuration*: Create a design that satisfy some requirements (e.g., design the machinery of an elevator).

- *Planning*: Find a set of actions to reach a desired state (e.g., emergency evacuation planner)

4.2.1   First generation expert systems

The first generation of expert systems used mainly heuristic reasoning with symbolic knowledge representations such as rules. For example, Mycin was one of the first expert systems that was used sucessfully as a model in other types of problems. Mycin was developed at the Stanford university in the 1970s (Buchanan and Shortliffe 1984) and it was able to diagnose infectious diseases and propose therapies. This system has the following two main functional components:

- *Interaction.* The information about patients is obtained through the communication with the user. Dialogue with the user is done in a question-answering process with prefixed questions and answers. Mycin asks questions about the patient and cultures. The questions are adapted to the previous answers (questions have context). The user may answer with uncertainty and may ask for explanations. Mycin presents diagnosis and therapy recommendation

- *Deliberation.* Mycin diagnoses the causes of the symptoms finding the possible infectious organisms. Mycin uses a heuristic classification method with a knowledge base with 450 rules and backward chaining. Mycin includes an original method for approximate reasoning (certainty factors as values in [-1, +1]). In addition, Mycin recommends therapies using a method based on a generate and test procedure.

The methods used for building expert systems demonstrated that they are valid in many problems, but they present limitations to be used in certain cases. One of the most important limitation is the problem mentioned above called *knowledge acquisition bottleneck* that restricts this approach for the development of knowledge bases that are not too large.

Expert systems may be able to solve difficult problems, but they operate in narrow specialized domains, without general knowledge about the rest of the world . Therefore, compared to human experts, who have also common sense knowledge, expert systems are brittle and unable to react correctly in unexpected situations.

Another problem is related to the way these systems obtain information about the environment. For example, Mycin gets information about the patient using a prefixed set of questions and answers. This communication mechanism may be too narrow and rigid to be used in situations when the format of available data is more diverse (e.g., unstructured texts). This limitation makes difficult to integrate the expert system into day-to-day operations.

4.2.2 Expert systems for the management of sensorized environments

The behavior of dynamic environments can be monitored with the help of sensor networks and information systems. This is the case, for example, of modern information infrastructures that have been developed in the last decades applied for example to strategic areas such a smart cities (transportation, climate, pollution surveillance) or industry 4.0 (with sensorized industrial installations).

In this context, it emerged the need of using intelligent systems that advise operators to make decisions for the management of such environments. There are multiple examples of expert systems that belong to this category. For instance, there are systems for public transportation management (Molina 2005), scheduling and coordination at an airport (Jo et al. 2000), emergency decision support in floods (Molina and Blasco 2003) or electric power operation (Filho et al. 2012).

The first example corresponds to an expert system that detects incidents and recommends actuation in a urban bus transportation network. This system was included in 2002 as part of the fleet management system of the bus control center of the city of Vitoria in Spain. This system has the following main abilities:

- *Perception.* The system collects data from buses using GPS information and localizes the position of each bus in a map of transportation lines.

- *Deliberation.* The system detects the presence of incidents (e.g., detection of delays based on current position and planned position) and predicts future behavior (e.g., a model is used about historic demand for transportation). The system determines actions to be done to manage incidents (an automated planner based on hierarchical task network (HTN) planning is used to generate action plans).

- *Interaction.* The system presents to the operator detected incidents (e.g., bus delays). The operator can notify other events to the system (e. g., a broken bus, a blocked street, etc.). The system recommends actions (e.g., using additional buses, sending repairing truck, etc.). The system justifies actions (e.g., why some lines are prioritized based on expected demand).

In general, this type of systems include a *perception* component to observe and interpret characteristics of the environment behavior using a sensor network or an information system. In addition, the system includes a *deliberation* component to derive conclusions from data, such as:

- what happens in the environment (detection),

- why it happens (diagnosis),

- what may happen in the future (temporal projection, what-if analysis), and

- what should be done (planning).

The system also includes a *interaction* component to present this information to the human operator. This type of intelligent system does not have an *actuation* component to modify the environment. The operator is responsible of making management decisions and implementing the actions in the environment. Therefore, it is important that the interaction component presents the system suggestions with sufficient justification in such a way that the operator can take responsibility for the decisions.

4.3 Question-answering systems based on cognitive computing

The company IBM developed a system called Watson (Ferrucci et al. 2010) that marked a significant milestone in the history of artificial intelligence when it won the first prize of the quiz TV show "Jeopardy" in 2011 competing against human opponents. The system was able to respond better than humans to open domain questions written in natural language about general culture.

In the quiz game, Watson had to guess an answer based on a given clue provided by the user. For example, Watson receives a clue such as "this drug has been shown to relieve the symptoms of ADD with relatively few side effects". In less than 5 seconds, Watson had to generate the answer expressed as: "what is Ritalin?" (according to the rules of the game, the answer "Ritalin" is presented as a question).

Watson used a selected large corpus of text documents of about 2 million pages as a source of information (e.g., news, encyclopedias and literary works). Watson searches candidate answers in such a corpus that can be possible responses to questions formulated by the end user in natural language. Watson uses a sophisticated structure of algorithms to (1) select relevant search keywords that may be used to find out (partial) candidate partial answers and (2) rank and integrate partial answers to generate the final answer in natural language.

The general characteristics of Watson can be described using to the components presented in this paper. In our characterization, we may consider the corpus of documents as the environment in which the system operates to perform its searching task. According to this, the components of Watson are the following:

- *Interaction.* The interaction component of Watson uses natural language processing to extract the linguistic structure of the given clue. The interaction component also generates the final answer in natural language.

- *Deliberation.* Deliberation in Watson included two parts. A first part generates a number of search keys from the linguistic structure of the given clue. Each search key is generated to be used later (by the perception component) to retrieve information from text documents. The second part of deliberation ranks and combines retrieved candidate answers. For this purpose, Watson was is able to merge answers (Kalyanpur et al. 2012) doing temporal reasoning, geospatial reasoning and taxonomic reasoning (subsumption, disjointness, etc.) using ontologies and databases (e.g., Yago, DBPedia, Freebase). In order to score answers, Watson uses models generated by machine learning.

- *Perception.* To obtain candidate answers from the corpus of documents, Watson used dozens of specialized algorithms for information extraction using natural language processing.

Watson is a successful representative case of a question-answering system that uses a computing approach called *cognitive computing*. The idea of this approach is that the system uses as a source of information large amounts of structured and non-structured data (text, images, etc.). The system performs automatically cognitive tasks that include feature extraction (e.g., using natural language processing or computer vision), answer aggregation (reasoning with ontologies or knowledge bases), natural language communication (using natural language understanding and generation) and evidence assignment of hypotheses using models learned by machine learning.

5   Conclusions

This paper has presented a characterization of an intelligent system that has been formulated with the purpose of being useful for systems engineers to analyze and design this type of systems. The paper defines an intelligent system using certain terminology with commonly accepted meanings in artificial intelligence. The description focuses on functional characteristics related to intelligence that machines are now capable to exhibit with the current state of technology.

In contrast to other work, the description presented in this paper is oriented to provide a global view as a whole, with usual features that may or may not be present in a particular system. This whole view creates a convenient context to describe more precisely concepts and identify their relations. For example, our characterization extends other definitions in order to consider explicitly the social factor of an intelligent system. This includes the interaction with other agents using language and following social norms. Language is related to the use of symbols, which is a significant feature of human intelligence, and social norms are important, especially when the intelligent system should cooperate with human users in accordance to ethical values. In addition, the capacity to adapt is presented in a broad sense, in relation to different forms of learning (from experience and from others).

The characterization provides a unified view of what is meant by an intelligent system that integrates different conceptions of such systems to cover classical systems (e.g., rule-based systems), behavior-based systems, or more recent conceptions for autonomous AI-based systems. The functional components used in the definition of an intelligent system are related to common AI methods (based on both connectionist and symbolic approaches). The paper also demonstrates how the presented characterization may be used to describe the functions of various real systems (e.g., expert systems and autonomous robots).

As future work, we hope to adapt and extend this general characterization to take into account results derived from new research on new computational models of intelligence.

References

Arkin, Ronald C. (1998). *Behavior-based robotics*. MIT press.

Bavle, Hriday (2019). "Positioning and Mapping for Aerial Robots using on-board Perception for Autonomous Missions". PhD thesis. ETSI Industriales, Universidad Politecnica de Madrid.

Bonnefon, Jean-François, Azim Shariff, and Iyad Rahwan (2016). "The social dilemma of autonomous vehicles". In: *Science* 352.6293.

Brooks, Rodney (1991). "Intelligence without representation". In: *Artificial Intelligence* 47.1-3, pp. 139–159.

Buchanan, B.G. and E.H. Shortliffe (1984). *Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.

Castelfranchi, Cristiano (1998). "Modelling social action for AI agents". In: *Artificial intelligence* 103.1-2.

Coradeschi, S. and A. Saffiotti (2002). "Perceptual anchoring: A key concept for plan execution in embedded systems". In: *Advances in Plan-Based Control of Robotic Agents*. Berlin, Heidelberg: Springer, pp. 89–105.

Duval, Shelley and Robert A. Wicklund (1972). "A theory of objective self awareness." In.

Edwards, Ward (1954). "The theory of decision making." In: *Psychological bulletin* 51.4, p. 380.

Ferrucci, David et al. (2010). "Building Watson: An overview of the DeepQA project". In: *AI Magazine* 31.3, pp. 59–79.

Filho, Joao Inacio Da Silva et al. (2012). "Electric Power System Operation Decision Support by Expert System Built with Paraconsistent Annotated Logic". In: *Advances in Expert Systems*. IntechOpen.

Franklin, S. (1995). *Artificial Minds*. MIT Press.

Franklin, S. and A. Graesser (1996). "Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents". In: *International Workshop on Agent Theories, Architectures, and Languages*.

Goertzel, Ben (2006). *The hidden pattern: A patternist philosophy of mind*. Universal-Publishers.

Harnad, Stevan (1990). "The symbol grounding problem". In: *Physica D: Nonlinear Phenomena* 42.1-3.

Hayes-Roth, Barbara (1995). "An Architecture for Adaptive Intelligent Systems". In: *Artificial Intelligence* 72, pp. 329–365.

Ingrand, Felix and Malik Ghallab (2017). "Deliberation for autonomous robots: A survey". In: *Artificial Intelligence* 247, pp. 10–44.

Jo, G. S. et al. (2000). "Ramp activity expert system for scheduling and coordination at an airport". In: *AI Magazine* 21.4, pp. 75–82.

Kahneman, Daniel (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kalyanpur, A. et al. (2012). "Structured data and inference in DeepQA". In: *IBM Journal of Research and Development* 56.3.4.

Kurzweil, Ray (2000). *The age of spiritual machines: When computers exceed human intelligence*. Penguin.

Laird, J. E., C. Lebiere, and P. S. Rosenbloom (2017). "A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics". In: *AI Magazine* 38.4, pp. 13–26.

Legg, Shane and Marcus Hutter (2006). "A formal measure of machine intelligence". In: *Proc. 15th Annual Machine Learning Conference of Belgium and The Netherlands (Benelearn'06)*. Ghent.

Maes, Pattie (1995). "Artificial Life Meets Entertainment: Life like Autonomous Agents". In: *Communications of the ACM* 38.11, pp. 108–114.

Marcus, Gary (2020). *The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence*. arXiv: 2002.06177.

Meystel, Alexander M. and James S. Albus (2002). *Intelligent systems: architecture, design, and control*. New York: Wiley.

Mitchell, Tom (1997). *Machine learning*. McGraw-Hill.

Molina, Martin (2005). "An intelligent assistant for public transport management". In: *International Conference on Intelligent Computing*. Berlin, Heidelberg: Springer, pp. 199–208.

Molina, Martin and Gemma Blasco (2003). "A multi-agent system for emergency decision support". In: *International Conference on Intelligent Data Engineering and Automated Learning*.

Molina, Martin and Pablo Santamaria (2021). "Behavior coordination for self-adaptive robots using constraint-based configuration". In: *arXiv preprint arXiv:2103.13128*.

Moravec, Hans (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.

Morin, Alain (2011). "Self-awareness part 1: Definition, measures, effects, functions, and antecedents". In: *Social and personality psychology compass* 5.10, pp. 807–823.

Muller, Jorg P. (1996). *The design of intelligent agents: A layered approach*. Springer.

Murphy, Robin R. (2019). *Introduction to AI robotics*. MIT press.

Newell, Allen (1982). "The knowledge level". In: *Artificial intelligence* 18.1, pp. 87–127.

Newell, Allen and Herbert A. Simon (1976). "Computer science as empirical inquiry: Symbols and search". In: *ACM Turing award lectures*.

Oreizy, P. et al. (1999). "An architecture-based approach to self-adaptive software". In: *IEEE Intelligent Systems and their Applications* 14.3, pp. 54–62.

Reggia, James A. (2014). "Conscious machines: The AI perspective". In: *AAAI Fall Symposium Series*. Arlington, Virginia, USA.

Russell, Stuart (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.

Russell, Stuart and Peter Norvig (2014). *Artificial intelligence: A modern approach (3rd edition)*. Pearson Education.

Schacter, Daniel L. et al. (2020). *Psychology*. Worth Publishers.

Silver, David et al. (2021). "Reward is enough". In: *Artificial Intelligence* 299, p. 103535.

Wooldridge, M. J. (2009). *An introduction to multiagent systems (Second ed)*. John Wiley and Sons.

Wooldridge, M. J. and N. R. Jennings (1995). "Intelligent agents: Theory and practice". In: *The knowledge engineering review* 10.2, pp. 115–152.

Xian, Yongqin et al. (2018). "Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly". In: *IEEE transactions on pattern analysis and machine intelligence* 41.9, pp. 2251–2265.