Check for updates

Cognitive motivations and foundations for building intelligent decision-making systems

Zdzisław Kowalczuk[1] · **Michał Czubenko**[1]

Abstract

Concepts based on psychology fit well with current research trends related to robotics and artificial intelligence. Biology-inspired cognitive architectures are extremely useful in building agents and robots, and this is one of the most important challenges of modern science. Therefore, the widely viewed and far-reaching goal of systems research and engineering is virtual agents and autonomous robots that mimic human behavior in solving known and unknown problems. The article proposes, at a high level of generality, an operational cybernetic model of the human mind, developed with the use of carefully selected ideas taken from psychological knowledge. In particular, the work combines extensive knowledge drawn from both the theory of developmental cognitive psychology and the theory of motivation. The proposed mathematically developed operating blocks create a coherent and functional decision-making system containing all the elements necessary in autonomous robotics. The ISD system is under development. There is still a long way to go to full validation. However, as shown in several articles, the basic subsystems of the ISD system, i.e. motivational and emotional, have already been positively verified in operation. The overall purpose of this article is to show a blueprint of the overall concept of the entire ISD.

Keywords Artificial intelligence · Decision-making systems · Autonomous systems · Mathematical modeling · Cognitive systems · Psychological processes

Zdzisław Kowalczuk and Michał Czubenko have contributed equally to this work.

✉ Zdzisław Kowalczuk
  kova@pg.edu.pl

✉ Michał Czubenko
  micczube@pg.edu.pl

[1] Department of Robotics and Decision Systems, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Pomorskie, Poland

1 Introduction

Today, many autonomous models and systems have been developed for specific or universal robot applications. Some of them are humanoid in nature (Kowalczuk and Czubenko 2013, 2016) with an idea for practical use in areas such as scouting, anti-terrorism, fire fighting, land rovering, etc. On the other hand, there are even more projects related to modeling separate parts of the human mind, such as motivation and emotions, acquiring social skills and developing various technical aspects, for example in the form of behavioral robotics (Brooks 1989; Czubenko et al. 2015; Kowalczuk and Czubenko 2011) or collaboration in a population of autonomous agents (Krolikowski et al. 2016). Among such works are also projects dealing with the coherent mathematical modeling of the totality of the human mind, known as cognitive architectures. The relative complexity of the ISD architecture presented in this review article is primarily due to the fact that more psychological components are implemented here than in other existing cognitive architectures.

One of the most important differences between humans and existing robots lies in emotions. There are some projects which take into account only expressing emotions by a robot, e.g. Kismet, Mexia, iCube, Emys, etc. Several computational emotion models that attempt to mimic human emotions have also been developed (Marsella et al. 2010). However, the crux of the problem is not emotional acting, but the role of motivators to act, control and make decisions (Kowalczuk et al. 2020).

1.1 Motivation & contribution

The concept of the Intelligent System of Decision-making dates back to 2010, when Samsonovich (2010) reviewed cognitive architectures. It turned out that there were many concepts of cognitive architecture, some based in part on human psychology, but none of them using humanistic psychology in conjunction with the theory of emotions. Interestingly, a review of the computational models of emotions by Marsella et al. (2010) appeared in the same year. However, our initial assumption was to create a system dedicated to autonomous agents and mobile robots in order to make decisions in various environmental conditions based on internal states of needs and emotions.

The purpose of this article is to review relevant issues and provide a sketch of an Intelligent System of Decision-making (ISD). Selected fragments of the ISD have already been described in several papers (Kowalczuk and Czubenko 2010, 2021; Czubenko et al. 2015; Czubenko 2017; Czubenko and Kowalczuk 2019; Kowalczuk et al. 2020) where you can find a detailed description of the key issues related to the ISD architecture. The system in general follows from the postulate that a robotic control system based on a coherent model of human psychology can serve as a universal decision-making unit in controlling practical, real or virtual processes. So equipped, the robot can perform the reconnaissance function more effectively in an unfamiliar environment where advanced autonomy is required (Czubenko 2017); it may also appear as a more user-friendly implementation of the concept of human-system interaction than other designs that merely mimic emotions.

Certainly, at this stage of development, it is not possible to model the entire field of psychology. Therefore, in this work we focus exclusively on the theory of humanistic motivation (as the main engine of the decision system) and cognitive psychology (for processing incoming information).

1.2 The structure

The structure of the article is divided into two parts. The first part describes the internal variables of the computerized agent, including the needs system and the emotional system. Then, global cognitive architecture is described primarily as a model based on the structure of human cognitive processes. In particular, the perception section describes the sensing of the environment and the processing of the signals obtained here into sensations/impressions and discoveries. The attention section shows how the environment can influence the above-mentioned states (needs and emotions) of an agent. Then the structure of the memory is presented. And at the end of the cognitive part, the reader will find a thinking engine limited to making decisions about the choice of agent or robot reactions based on internal states. We also briefly describe the basic details of computer-based implementation. Conclusions drawn from the research carried out can be found in the last section.

2 Motivation factors in ISD

The theory of motivation deals with the impulse to act. Thus, the center of the ISD system is built on a synthesized model of human motivation theory, which describes both the external human behavior and the way in which the embodied human mind affects its internal emotions and reasons.

Motivational factors are key ideas in modeling human behavior. Considering them most important to our goals, let us first consider two basic types of motivators: needs and emotions. Need is the main component of a human motivation system. However, in exceptional cases, the thinking subject (agent) must react immediately, although his motivational reaction system may be preoccupied with other motives and driving needs. At such times, the reaction can be quickly derived from the emotions. Such duality (and entanglement) is consistent with currently recognized models of thinking (Pennycook et al. 2015). As we will show later, the ISD architecture allows for the implementation of emotions as a controlling factor (Czubenko et al. 2015) that also influences the needs system.

2.1 Model of needs

The need is an abstract state of an agent experiencing a sense (feeling) of dissatisfaction (Maslow 2012). The more the unmet need, the more this lack (or unfulfillment) needs to be reduced or eliminated. There are a large number of needs. They can be divided into several classes with different hierarchy that can be arranged in a pyramid. The structure of needs in each class may change in the course of development [and become a kind of 'diamond of needs' (Wu 2012; Noltemeyer et al. 2012)].

It should be emphasized that various needs systems [e.g. related to personality (Murray 1938)] and theories of motivation [e.g. about self-determination (Ryan and Deci 2000)] have been developed. One of the more interesting theories of motivation, ERG (Existence, Relatedness and Growth) by Alderfer also uses needs, but in a less hierarchical way. The needs are organized into three clusters here: existence, relatedness and development, but the main idea of human needs is similar. Other categories of needs (hygiene and motivators) were created by Herzberg (1965), who also based on the same concept of needs. We choose the Maslow model because it is flexible and open (from a formal point of view) and easy to implement in a computerized agent management system. The humanistic theory of

Fig. 1 Model of need and its fuzzy-linguistic values (Kowalczuk and Czubenko 2011). The symbol η_i corresponds to the level of this (i-th) unmet need, μ symbolizes the membership functions for satisfaction (μ_s), prealarm (μ_p) and alarm (μ_a). The weight function ω is shown as a dashed blue line, and an exemplary current value of the need is marked with a solid black line segment
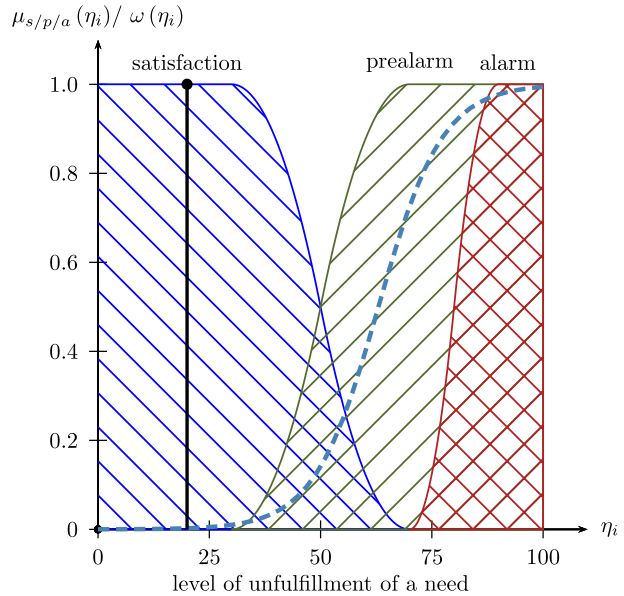
motivation (mentioned above) does not include emotions, which are another part of human psychology. Emotional models have been carefully analyzed in our preliminary research. This article presents psychological theories, while a review of computational emotion models is presented in Kowalczuk and Czubenko (2016), Kowalczuk et al. (2020).

In the ISD architecture (Kowalczuk and Czubenko 2010, 2017; Czubenko et al. 2015; Czubenko 2017) the state of a single need is modeled using fuzzy logic. The level of need-related (dis)satisfaction is represented by a linguistic variable that has three values (states): *s*atisfaction, *p*realarm, and *a*larm. Their membership functions are illustrated in Fig. 1.

Satisfaction is the state of a given need when the agent does not have to take action to eliminate or mitigate it. It has a narrow kernel (at the maximum value of the corresponding membership function) and decreases along the abscissa. Prealarm status tells the agent that it should rather start working on that need. An alarm condition occurs when something is missing that must be immediately compensated (which motivates the agent to act immediately). Coefficients describing the functions of belonging, denominated with appropriate linguistic terms (labels), are subject to modifications depending on the current emotional state of the agent. Generally, negative emotions narrow down the function of the satisfaction state and extend the other functions. In this way, the needs of the ISD agent become unsatisfied (more pressing) more easily. Positive emotions have the opposite effect. Consequently, the corresponding language variable 'unfulfillment of need' ('unmet need') can assume three possible linguistic values which are (separately) described as *s*, *p* and *a*.

Note that all membership functions presented in the figures are expressed in the form of the *z*-, *s*-, and *π*-shaped functions (Kowalczuk and Czubenko 2011; Łęski 2008). The detailed shape of these functions depends on the parameters associated with the parameters and state variables of the ISD system. We recommend using the Gaussian or the generalized bell membership function (Porębski and Straszecka 2016). Another element of the discussed fuzzy model of the needs system is a need weighting function, depicted by a dashed line in Fig. 1, which assumes a simple *s*-shape spanned between 0 and 100. In a natural way, its inflection point is placed in the kernel of the prealarm state. The weighing function

can reach significantly lower values than 1, since additional scaling is taken into account by a factor representing the Maslow class of the analyzed need. This allows ISD to efficiently handle those needs that are currently the most important. It is worth adding that the weight function also modifies inference operations.

Various functions can be used to characterize fuzzy variables. The most popular of these are the trapezoidal, pi-shaped, sigmoidal and Gaussian functions. The types of functions used to describe needs or emotions depend on the agent configuration used, although in our experience this should not be a critical issue.

2.2 Emotions

There are different definitions of emotions. In the physiological approach, it is postulated that emotions evolved from the process of homeostasis. According to the cybernetic theories, emotion is a homeostat itself, which is aimed to "maintain the functional balance of the autonomic system" (Mazur 1976) by a counter action of the information and the energy flow that reduces the possibility of its environmental impact. In other words, emotions adjust the agent's behavior to better respond to life stimuli from the environment.

Apart from the challenge in defining emotions, there are also theories relating to the process of creating or triggering emotions. Psychologists identify two main trends: appraisal and somatic. The evaluation theory suggests that before the appearance of emotions, there are cognitive processes that analyze the stimuli (Frijda 1986; Lazarus 1991). Thus (referring to the history of life entities), emotions follow the process of recognition: first you need to recognize objects, and then determine their relationship with the emotions of the agent (which means the superiority of cognitive processes over emotions). On the other hand, the somatic theory holds that emotions are dominant compared to cognitive processes (Zajonc et al. 1989; Murphy and Zajonc 1994): this means that before analyzing a perceived object (or even registering any sensation), the human brain is able to recall emotions associated with that object. Note that in our concept of cognitive architecture we combine both theories in one ISD system.

2.3 Short review of emotional systems

Note that there are many computational models of emotion created by psychologists or computer science engineers. Many of them are very well described in Marsella et al. (2010), Kowalczuk and Czubenko (2016).

From a psychological point of view, one of the most interesting models of emotions is Russell's (1980). It shows emotions in a two-dimensional space. Although the Cartesian coordinate axes have been called excitation and valence, the model looks much better in polar form. The popular version of the model describes only eight basic emotions, while its extended version, based on experimental research, indicates 28 emotions. Like most theories, this model has evolved over time, and the latest version of this model is called circumplex (Posner et al. 2005).

A similar model was also proposed by Thayer (1989), whose emotional axes were named differently, namely calm-tension and weariness-energy. On the other hand, Bradley's model is based on the assumption that two 'vectors' add up to an emotion (Bradley et al. 1992). According to the arousal value, the emotion is triggered at the point described by these vectors. These kinds of vector models are widely used in laboratory research on emotional triggers (Rubin and Talarico 2009).

A model similar to the one described above is the PANA (Positive Activation–Negative Activation) scheme (Watson and Tellegen 1985). The authors of this model propose two separate systems of (threshold) action to activate emotions, depending on the nature/strength of the emotion. The states of high excitation are used according to their valence, while at a certain low excitation, the valence is considered neutral.

Based on a biological interpretation, Lövheim (2012) presented a 3D model in which the axes are assigned to three neurotransmitters: serotonin, dopamine and norepinephrine. Each of the extreme points of the cube (0 or 1) portrays one of the emotions. Obviously, intermediate states represent the occurrence of a given emotion in a partial range.

From a systemic point of view, the OCC theory (Ortony, Clore & Collins) is the most popular vision used in computational models of emotions for artificial agents (Clore and Ortony 2013; Steunebrink et al. 2009). It describes the evolution of an agent's emotions depending on the occurring events and actions as well as perceived objects. The resulting model uses 22 types of emotions in an appropriate hierarchy.

In computer science, there is a Fuzzy Logic Adaptive Model of Emotions (FLAME) system, also based on the OCC (Clore and Ortony 2013) theory, which takes into account the emotional evaluation of events (El-Nasr et al. 2000). Whenever a new event occurs, FLAME evaluates its value against the agent's agreed targets. In particular, FLAME considers what goals and to what extent are being achieved as a result of the event, and then evaluates them using a set weighting of the goals. Based on this assessment, the utility (purposefulness) of the event is determined using the Mamdani–Assilian fuzzy system. As a result (based on this cognitive assessment of events), fuzzy emotions are created taking into account the desires and rules described by the OCC theory.

EMotion and Adaptation (EMA) is an extensive emotional system with many evaluation variables (Marsella and Gratch 2009). Emotions are generated using a mapping algorithm. They take into account the values of the above-mentioned evaluation variables in the context of a certain perspective. For example, hope comes from believing that something good can happen. Each of the 24 emotions is described by a separate intensity variable. Based on the current emotion, coping strategies are established that work in the opposite direction to evaluation, i.e. identify the causes of emotions and support their improvement.

Affect Simulation Architecture for Believable Interactivity (WASABI) is an example of a computational system of emotions in which emotions have been modeled in a continuous three-dimensional space PAD (Pleasure-Excitement-Domination) (Becker-Asano 2008). The entire system is made up of two parallel emotional and cognitive processes. The emotional process creates a vector of emotions based on evaluative impulses and triggers coming from both the external environment and the cognitive module. At the same time, the cognitive module can generate variables of complex emotion. The module based on the Belief-Desire-Intention (BDI) concept and the Adaptive Control of Thought-Rational (ACT-R) architecture is responsible for transferring a single action or its sequence to agent actors. The system was designed to generate speech taking into account the emotions of the actor.

The hourglass model (Cambria et al. 2012) based on Plutchik's theory of emotions uses the Gaussian function to successfully model realistic behavior of an agent in terms of the evolution of their emotions.

It is worth mentioning here that there are also models based on a broader view of the problem, such as PEM involving Personality, (motivation), Emotions, and Mood (Santos et al. 2011; Shvo et al. 2019). The personality aspect here is about the difference between the individual agents and is modeled by a standard personality trait—the Big Five (also known as FFM, the Five Factor Model). The emotional aspect in the PEM system is based

on the OCC model, while the mood is realized through the aforementioned PAD model. The entire PEM system was used to negotiate between a group of agents.

Certainly, the concepts of the emotional systems presented above do not exhaust the wealth of models available in the literature. On the other hand, we present a computational model of emotions based on different assumptions. First of all, we use the Plutchik model (in the psychological and emotional range) with computationally/systemically significant modifications. We also apply two main theories of creating emotions (somatic and evaluative) for a more complete theoretical completeness and consistency of the ISD system. Note that most emotional models use crisp values to describe emotional states, while in our opinion fuzzy concepts, which can also be used in 3D space, are better.

2.4 Our model

There are a number of parameters that characterize emotional states. A group of similar emotions can be assigned a specific color and appropriately labeled (*joy*, *happiness*, *ecstasy*, etc.). This type of aggregation can be associated with the generalization of human emotions, known in psychology as basic emotions (Plutchik 1980; Ekman and Cordaro 2011). In our system, a color is just a label on some abscissa. However, we have deepened this concept of the color of emotions. Literature research has shown that colors are generally associated with certain emotions, but their meaning may differ from culture to culture (Gilbert et al. 2016; Whitfield and Whiltshire 1990). Our parameterization comes from the concept of Plutchik's emotional configuration. It is instructive that in many computational models of emotions, there are concepts similar to valence/color and arousal/intensity. Therefore, another parameter that describes the emotion more accurately is its intensity (Posner et al. 2005) (meaning that emotions of the same color vary with intensity). Note that emotional intensity determines how much an emotion affects the agent's behavior.

Another parameter concerns the duration of emotions, which can range from a few seconds to several weeks, sometimes even months. Emotional states that last longer than a few months are rather personality traits or emotional disorders. Considering the duration, emotions can be classified as (Biddle et al. 2000; Oatley et al. 2012):

- Autonomous changes (*pre-emotion*): very short (in seconds), spontaneous physical feelings (Ekman 2009), associated with the somatic theory of emotions (and dependent on specific stimuli), without a deeper recognition of the situation, object or event. Therefore, in the ISD model, pre-emotion is associated with a certain stimulus or impression (e.g. detection of sudden movement, which is out of sight, but in close proximity, generates a simplified emotional signal of fear). The agent treats pre-emotion as one of the decision-making factors.
- Expressions (*sub-emotions* and emotional *sub-qualia*[1]): short emotional states (in seconds) which are associated with recognized objects [based on the appraisal theory of emotions (Lazarus 1991)]. Sub-emotions are standard, universal and emotional expressions. By contrast, sub-qualia are subjective, individual, emotional feelings that are qualitative in nature (Hardin 1987). Both expressions are related to perceived and already known objects, situations, or events, and are treated as emotional factors.

---

[1] In this article, for the sake of simplicity, all qualia terms refer to e-qualia—emotional qualia only.

- Classical *emotion* and *emotional quale* (see Footnote 1) are motivational factors (consciously) observed, possibly verbalized and lasting longer. Classical (objective) emotion determines the universal emotional state of the agent (and its basic objectives). It is the result of current sub-emotions, pre-emotions, the level of (un)fulfillment of needs, the previous emotional state and the mute effect. Classic or system emotion is treated as the fundamental emotional state of the agent. Emotional quale is a personal emotion that represents the abstract and individual side of emotions (Kowalczuk et al. 2021). To some extent, as in the case of classical emotion, sub-qualia derived from current perceptions accumulate on system quale.
- Mood (*mood*): long-lasting (days and months), externally observable, with a rather lower intensity than classical emotions, usually slow-changing, positive or negative (Batson 1990). It gives a general indication of whether and how the system emotion affect the agent. The mood allows you to properly modify the functions responsible for satisfying your needs. Technically speaking, the mood can be created through a specific 'differentiation' of the classical emotion, modeled using a *Temporary Amplifier With Saturation* function [TAWS (Kowalczuk and Czubenko 2013)].
- Emotional disorder[2]: a type of long (months and years) depression, phobia, mania, fixation, etc.
- Personality traits[3]: timeless emotions such as temperament, shyness, neuroticism.

Analyzed emotional processes have four functions: information-oriented, activating, meta-cognitive, and modulating. For robotic projects, the main utility is modulation, which is responsible for increasing or decreasing the sense of satisfaction (such modulation is a function of mood). The meta-cognitive function adds more information about the perceived object and thus facilitates its classification (the information function prohibits or allows the agent to look at the perceived object more closely. The simplest emotional factors— pre-emotions, sub-emotions and sub-qualia correspond to meta-cognitive and information functions. The latter function is also associated to the classical emotion and system quale. In addition, some actions can only be unlocked when the classical emotion is in a certain state (e.g. in *rage*, the *impact* response can be unlocked) (Kowalczuk and Czubenko 2016).

Emotions can be grouped according to their similarity or color. For example, grief, sadness and pensiveness have their own common group color (Fig. 2), which varies in intensity. Moreover, for each of the groups of positively interpreted colors, there is an opposite group, associated with 'opposite' emotions. So, as shown in Fig. 3, for the *sadness* family, the opposite group is *joy* (including ecstasy, joy and serenity). Thus, Plutchik identified eight families of 25 emotions (Fig. 2).

Plutchik's theory appears controversial to some researchers. In addition, there are other theories according to which emotions can be located in a two-dimensional space (e.g., relative to arousal and valence). However, these theories describe emotions as points in such a space, while the named variables expressing an emotional state are rather fuzzy in nature. In Plutchik's theory, each emotion is represented by some slice of continuous space. Therefore, this theory is a good reference point for the proposed fuzzy modeling, in which emotion labels and their fuzzy membership functions can be freely shaped by the system designer. As a consequence, the proposed model of emotions seems to be fully functional (Kowalczuk et al. 2020).

---

[2] It has only a very limited representation in the ISD model.

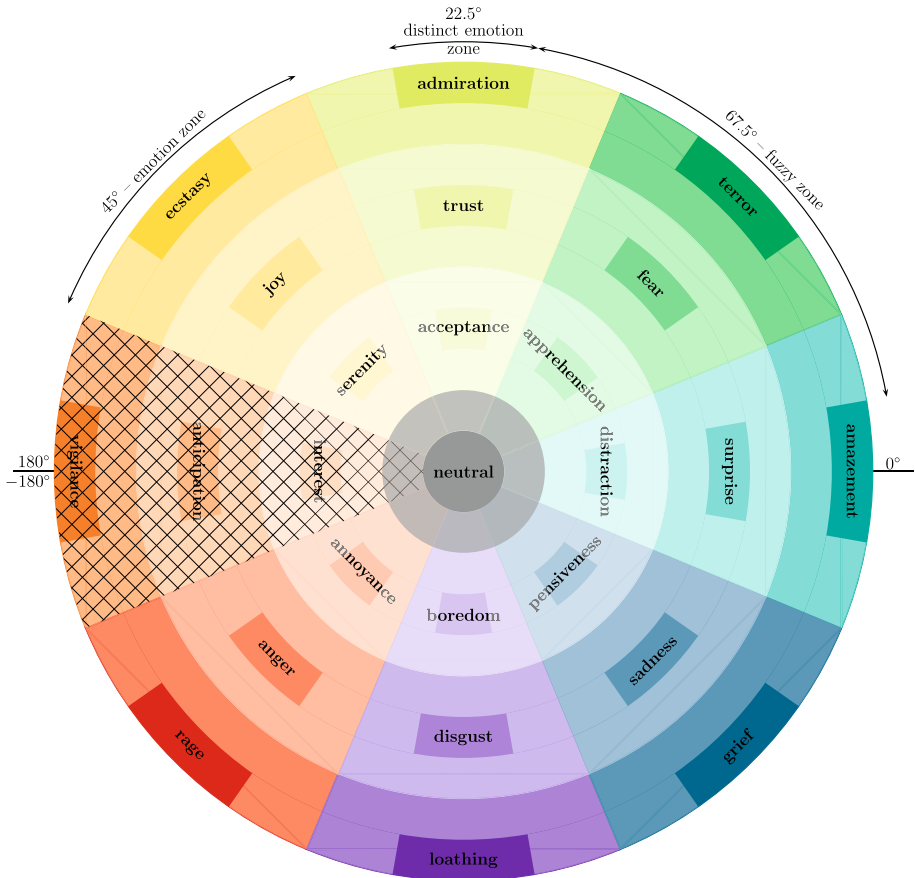[3] They are also not yet used in ISD.

Fig. 2 Rainbow circle of emotions: full color shows where emotions have their full value, while a weaker color shows the slope part of the fuzzy membership functions (the covered parts are not visible) (Kowalczuk and Czubenko 2013, 2016)

In order (as in metric spaces) the norm would grow as you move away from zero, we proposed (Kowalczuk and Czubenko 2013; Czubenko et al. 2015; Kowalczuk et al. 2020) to reverse the emotional scale used in Plutchik's model. Moreover, Plutchik's emotional system is unnaturally centered around zero, which is the most active emotional state, while technical systems require working around a clear neutral state. Despite the dynamics manifested at low levels of natural life and the reactivity of organisms, the existence of neutral states is extremely beneficial in technical systems. Moreover, such a state can represent any emotional state of very low valence (e.g., while sleeping or meditating). In the fuzzy model used, only the smallest circle in Fig. 2 is truly neutral. Moreover, the agent's emotion can reach this state in the absence of any external emotional stimuli. Figure 2 shows the ISD emotion model derived from Plutchik's theory and modified according to the above design guidelines.

The emotion of anticipation creates a non-linear subsystem, which—for a mentally healthy person—is the final state that can be achieved with continuous positive emotional excitement. A person who is not fully emotionally healthy may unexpectedly land on the
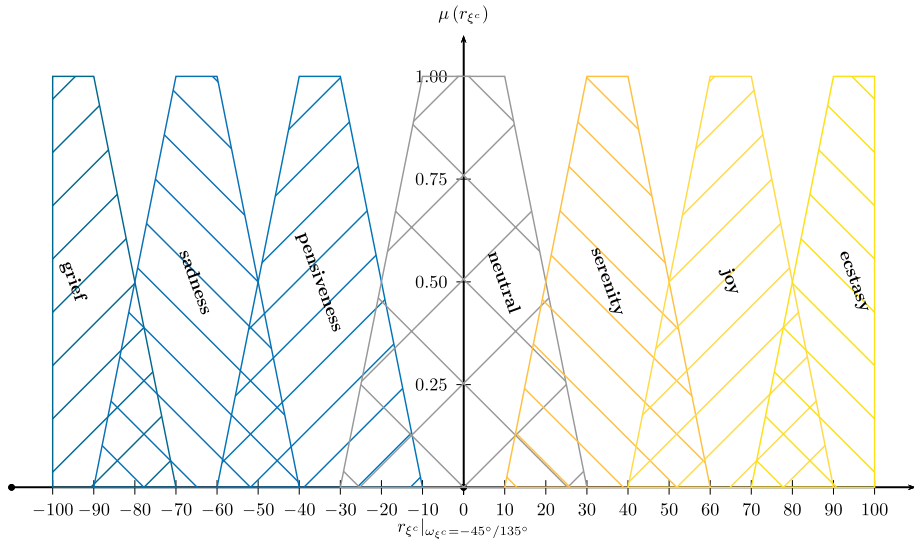
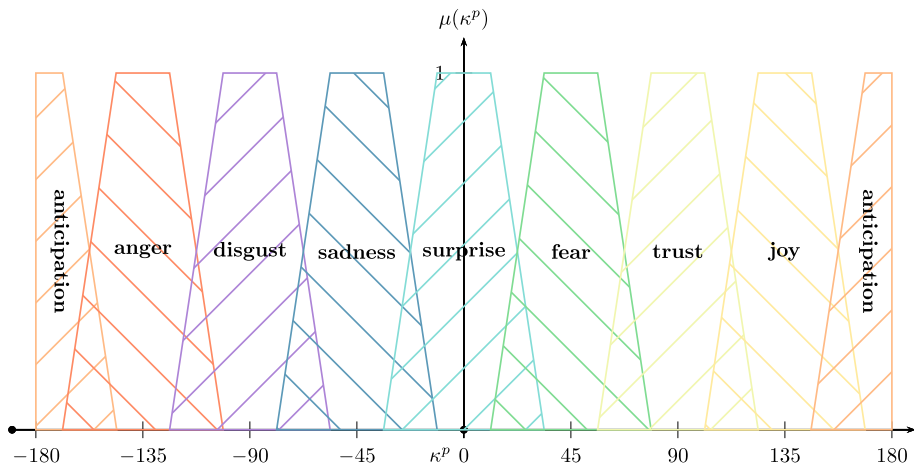**Fig. 3** Cross-section of the emotion circle

Fig. 4 Simplified linear model of emotions (the middle ring on the rainbow of emotions of Fig. 2)

anger emotion/color. Moreover, since a direct (negative) transition from anger to joy (by anticipation) is forbidden (as debatable), the development of emotions from anger to joy can only be achieved naturally through persistent positive excitement.

Following the idea used in our model of needs, in the ISD emotional (sub)system a single emotion is modeled using the developed 2D fuzzy-linguistic system (Figs. 2, 3). The intensity of each emotion can be mapped by four linguistic values; For example, when considering the emotions of *joy*, we have ecstasy, joy, serenity, and a common neutral status.

The middle ring in the emotion rainbow (Fig. 2) makes the classical linear emotions shown in Fig. 4. These emotions are useful in simplified cases—for example, in modeling

pre-emotions whose emotional values are expressed linguistically as one of the eight classic emotions. Other details of the xEmotion model are presented in Kowalczuk et al. (2020), Kowalczuk and Czubenko (2021), where simulations confirming the validity of ISD can be found.

In conclusion, there are many emotional models. Most of them use crisp value to describe an emotional state, while in our opinion it is a fuzzy concept. Our model is based on the Plutchik wheel of emotions and is somewhat reminiscent of the Hourglass concept. Many models use 3D space to model emotions, but only ours shows them as a fuzzy quantity.

The tests of the xEmotion system have already been covered in several articles (Czubenko et al. 2015; Kowalczuk et al. 2020). They prove that apart from the ease of algorithmization, the entire concept of our emotional engine works in the expected, logical way. On the other hand, however, it is not possible to compare different computational emotion models as they are based on different theories and used in different circumstances/scenarios.

2.5 Summary

This section presents a motivational model based on emotions and needs. Despite criticism of the selected theories of motivation (Maslow's needs and Plutchik's emotions), from a cybernetic point of view, both theories show useful motivational factors (subjective impressions) as internal variables that can be represented in terms of fuzzy sets. In the next section, we present the cognitive framework in which the above motivators are implemented.

3 Model of cognitive psychology

In addition to the isolated processing of needs and emotions, the theory of cognitive psychology provides a good basis for supporting the construction of autonomously acting systems. While the theory has different origins and is vast, uncertain, blurred, and inconsistent on many points, it wisely spells out the different types of information processing and the relationship between memory and stimuli. It also recognizes how we make decisions and think, and even what happens during sleep. In particular, the theory describes what happens between the appearance of a stimulus and human reaction. Thus, from an engineering standpoint, cognitive psychology represents a high-level white-box archetype for building autonomous robotic systems.

The practiced cognitive approach to the general problem of decision making is based on the assumption that the knowledge underlying the decision is not created by passive data collection and storage. Instead, active data processing takes place and some sub-processes are parallel/independent. This means that the structures of human cognitive processes provide solid arguments for modeling the desired decision-making process of thinking entities (Lindsay and Norman 1977).

Cognitive processes are responsible for the activities leading to gaining knowledge about reality (Maruszewski 2001; Nęcka et al. 2008), which can be classified as:
(1) Elementary processes:

- Perception,
- Attention,

- Memory,

and (2) Complex processes:

- Thinking,
- Problems solving,
- Decision-making,
- Language.

From the model (functional and structural) point of view, observations can be divided into: sensory, impression and discovery, while attention can be isolated as conscious (intentional, overt, explicit, top-down) and unconscious (unintentional, covert, hidden, bottom-up) (Kowalczuk and Czubenko 2010). At this stage of development, we do not take into account the linguistic aspects as they deserve a separate mathematical treatment. In this way, we combine all complex processes and their relationships into one common thinking process.

During evolution, species took various steps to deal with the environment and its impact. Action can be taken at almost any information processing level. *The unconscious reaction* acts as a reflex or impulse based on a stimulus. This could be a withdrawal reaction to unforeseen movement in sight or to pain development. A *subconscious reaction* can be defined after repeating an action based on the same set of perceived objects multiple times (while driving a car, shifting gears, etc.).

The most developed and relatively slow type of reaction is that which is *consciously* selected according to the knowledge of the environment and the experience of the agent.

Stimuli appear in the receptors. The continuous flow of information passing through the perception system (Nęcka et al. 2008) allows a discovery or an object to be distinguished. Next, such an object is filtered and processed by attention, and then analyzed by thinking to develop an appropriate decision or reaction.

It is worth emphasizing here that in order to better reflect the operation of human psychology, the agent should be able to predict the behavior of other agents present in his environment. Such a mechanism could be based on the theory of mind (Gallagher and Frith 2003), which is related to the ability to explore a multi-agent world (with the presence of other intelligent entities). At the present stage, we focus only on building an agent capable of learning about the world, and we postpone the wider use of the theory of mind to the next stages of the development of the ISD project.

3.1 Perception

The first step of data collection consists in the sensory perception shown in Fig. 5, which receives stimuli from receptors responsible for various senses, such as sight, hearing, taste, smell, touch, balance, temperature, kinesthetics and pain (Nęcka et al. 2008). Sensory perception has two successive phases related to distal (representing real objects) and proximal (images of objects in receptors) stimuli (Maruszewski 2001). Proximal stimuli are stored in sensory memory (ultra-short time), where the stimuli are pre-processed at the primary filtration level. Due to the concept of readiness, receptors can also focus specifically on recognizing selected objects (Bruner 1973).
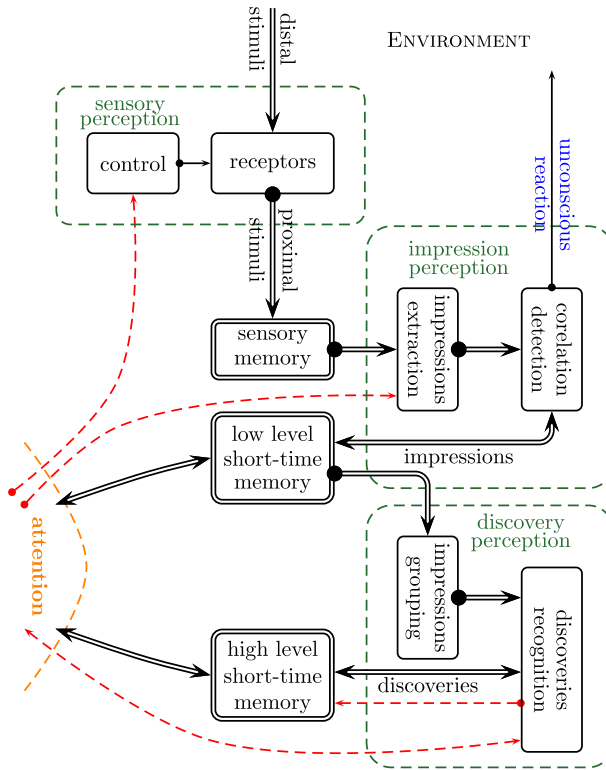
Fig. 5 Perception process

3.1.1 Impression perception

The purpose of perception of impressions, which is the next stage of data processing, is to extract the impressions imprinted in proximal stimuli. Impression is a simple feature of an object (representing color, texture, etc.) that results from the activity of the ascending pathways extending from the receptors (Hebb 1958). Feature detectors (for shapes and even complex patterns, such as the entire face) belong to this part of the perception process. They can be implemented in many ways, for example in the form of a Haar filter (Lienhart and Maydt 2002).

There are four groups of impressions in ISD (Kowalczuk and Czubenko 2014):

(i)   Component (primal)—impressions showing the inherent features of physical objects;
(ii) Intrinsic (complex, abstract, fuzzy)—traits related to complex stand-alone characteristics;
(iii) Functional (system, abstract)—other features related to the functioning of the ISD architecture;
(iv) Recognized (known)—fixed attributes used in system models of learned discoveries.

Single impressions may have a specific physical or geometric meaning (e.g., dots, stripes, face or anthropomorphism) or may represent some abstract class (beauty, redhead, etc.).

These types of functional features (impressions) may refer to various observations of the agent, but also play an important role in the functioning of the system. They can represent feelings (sub-/emotional context), needs (objective motivational context) or other abstract features (such as temporal categories, e.g. monumental, permanent, temporary phenomena) that are systemically important. Certainly, intrinsic (independent) impressions are also ready to constitute functional properties, for example in an evaluative context. Recognized sensations (impressions) refer to features already used in models of discoveries (known, already experienced or learned) stored in the agent's memory, and thus facilitate the identification of the cognized objects.

In this way, once the impression (component or intrinsic) is isolated, it is compared with known features/sensations[4] in order to (further) identify the whole perception. Repeatedly perceived (cognized), although not yet recognized, impressions are stored in low-level short-time memory (L-STM). Therefore, from a systemic point of view, we must take into account the existence of impressions of two kinds:

- Recognized (known);
- New (not yet known).

As mentioned above, intrinsic type impressions, usually describing complex traits, can be functional in the evaluative tasks of the human mind. Meanwhile, all functional impressions relate directly to various motivational factors—such as emotions, representing a subjective context, or needs, relevant from an objective motivation (survival) point of view. Both motivational factors (emotions and needs) are interpreted, applied and stored using a fuzzy representation (Wu and Miao 2013) which seems optimal for making decisions and communicating with people (compared to any strict/crisp mathematical model).

Recognized impressions (from L-STM) are grouped according to their location in the perception space (perceptual scene). In the literature, a meaningful combination of such impressions is referred to as a perception, discovery, observation, or object. "Each sense organ (...) sends connections to a common pool, the non-specific projection system or arousal system, which mixes up these excitations and sends them on to the cortex" (Hebb 1958). As the process of encoding sensory information into impressions is often treated as information re-coding, similarly, the process of discovery synthesis can be interpreted as a form of higher-order recoding. Note that a completely new discovery will be considered unclassified (unknown).

3.1.2 Discovery perception

Recognized (L-STM) impressions are grouped according to their location in the scene of the perceptions made. A discovery is an abstract representation of perceived objects. In many cases, it may be subjective, inaccurate or incomplete. It contains a list of sensations (component, intrinsic and functional) associated with that object, the logogen of that object, and relationships with other discoveries that can be presented as propositions.

In general, there are two types of perceptions: abstract and instance discoveries, a detailed description of which can be found in Sect. 3.3.1. An abstract (classified) discovery represents a category (or a group of objects). The process of perception (recognition) consists in

---

[4] Some of them (such as 'sting' or 'sharpness') are designed to trigger a specific unconscious reaction.

comparing a perceived but unclassified object with abstract discoveries. Each discovery has its own label, describing it with a list of impressions and a list of associations with other discoveries (stored in memory). In addition, it may also have functional impressions with a motivational context (need or sub-emotion). Currently recognized cases of known abstract discoveries are called instance discoveries. Basically, they have the same structure as abstracts but for an additional feature representing their activity level which corresponds to the frequency of their occurrence.

The process of perceiving discoveries involves semantic aspects, such as differentiation, identification, categorization and perceptual orientation, which can be expressed in the form of specific memory networks involved in the discovery recognition process.

Temporary discoveries that have not yet been classified are compared to previously classified discoveries from LTM. A discovery is recognized according to the best match when a certain minimum threshold (e.g. 90% agreement) is exceeded.

If such high-level short-term memory (H-STM) content review is unsuccessful, discovery recognition requires the attention block to retrieve new objects (Fetch New Objects: FNO) from the LTM for further comparison. After a certain number of discovery attempts, the successfully recognized ones are transferred to the memory of the perceptual scene, while the unrecognized discoveries go to the second stage of the search and comparison process.

If there is a match (or some kind of similarity), the number of hits (activity count) on this temporary discovery increases. When no match is found, the perception process generates the Remember New unrecognized Object signal (RNuO) and creates a corresponding new unrecognized discovery. At the end of this stage, when the number of hits for such an unrecognized discovery reaches a certain level, the discovery is transformed into a new i-discovery stored in semantic memory, which is also signaled by an appropriate signal Create New Object (CNO). A temporary name is consciously given to this discovery as a result of a creative thinking process. It is worth noting that an abstract discovery can also be created by generating a Create New Abstract (CNA) signal when multiple instances of a similar nature are in the instance discovery memory.

In case of confusion, when some impressions of the discovery contradict, the discovery is rejected by sending a Drop the Object (DO) signal (Kowalczuk et al. 2016).

Sense perception, controlled by unconscious attention, is responsible for the readiness to perceive the expected object (Bruner 1973). This makes it easier to quickly find specific objects in the nearest vicinity. Such immediate recognition takes place in several cases:

- Narrowing down to a small set of substantive categories (e.g. only food, safety or danger),
- The emergence of a specific motivation for a given unit, person, agent or system,
- Perceived strong relationship between categories.

The situation can also be the opposite: Readiness is weakened, which allows the agent to ignore certain objects (i.e. become immune, not recognize certain objects, ignore obscene words, etc.) (McGinnies 1949). By weakening the focus on most perceived objects, it becomes easier for the agent to recognize new objects.

3.2 Attention processes

Attention is a process that allows us to direct the perception processes to certain objects or scenes. A typical example of such a phenomenon is the *cocktail party effect* (Arons 1992), on which the agent is able to focus its attention on certain participants of a
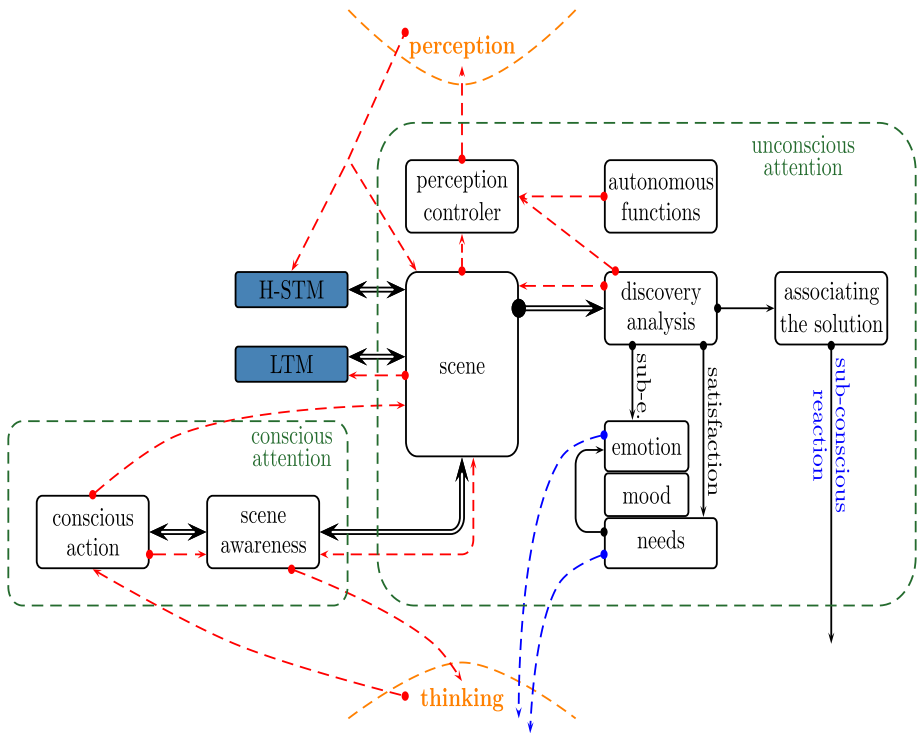
Fig. 6 Attention process

party and their conversation, despite the existing general noise of this party. Attention, founded on the structure of Fig. 6, is responsible for the following functions:

(1) Control of all memory processes,
(2) Allocation of cognitive resources to specific tasks,
(3) Information selection and perceptions analysis,
(4) Shaping (targeting) of cognitive processes.

The first function is completely unconscious, while the other three can be conscious or unconscious.

(**1**) The memory controller performs a set of subordinate tasks that support the transfer and maintenance of information in memory (including LTM and STM). The following subordinate tasks related to memory handling can be mentioned here:

• Writing to LTM,
• Restructuring of LTM,
• Extending the structure of the discoveries,
• Forgetting information,
• Searching for objects,
• Searching for similarities.

The extension of discovery models concerns new properties that are necessary to describe the agent's written knowledge. These activities are discussed in Sect. 3.3.

(**2**) The second function represents a higher level of attention organization to which a cognitive resource group is assigned, also in the subordinate mode and according to the weight assigned to each task. This determines when attention will shift to other tasks, usually in line. It is also about dividing attention, i.e. carrying out two or more activities simultaneously (if it is possible and to what extent).

(**3**) Processing of stimuli in the control of the perception process (agent's attention) works mainly automatically, although in some cases it may be partially controlled by the agent (Ulrich et al. 2015). Information is selected in two stages. Primary selection on the path of information perception concerns the perceived stimuli in the direction of attention (Broadbent 1957). Secondary selection refers to the analysis of information (discoveries) that is done by subordinately controlling the memory and transferring data (read and write) between the short-term STM memory and the long-term LTM memory. This analysis (recoding process) recognizes only those perceived discoveries (i-discoveries and a-discoveries) that are stored in LTM (Maruszewski 2001; Lindsay and Norman 1977). Interestingly, this selection function also has a protective role in the sense that the LTM does not receive all the information from the STM. For psychological reasons, some of the new discoveries may be undesirable in LTM (e.g. brutality).

(**4**) The fourth (superior) function of attention concerns the orientation of cognitive processes, including their intensification and extensification. Thanks to the analysis of discoveries, attention supervises and adapts the processes of perception, in particular, it intensifies/focuses attention (cognition) on certain objects or extensifies/blurs it. This function of shaping the current way of perception is carried out as a process of searching the perceptual field (vigilant scanning). In this way, attention decides whether the sensory perception of sensations is focused on a wide range of objects (objects or phenomena) but with a reduced number of details, or on one object, but with all possible impressions. A good example is the "cocktail party" effect when we focus on one phenomenon or conversation in the presence of many others. As another form of adaptation, superior attention can change the context of perceived discoveries, and can also help change the response currently being performed due to new information (Monsell 1996). The fourth attention function is also responsible for evaluating and providing relative weights to the weighing process of individual tasks performed by the second attention function. This function can also be seen as the next (third) and most powerful level of information selection subject to the thinking process.

Systemic attention includes autonomous (independently operating) unconscious functions (Treisman 1986; Driver 2001):

- Orientation reflex (like intensification) controls the actuators to focus attention on the source of the triggered impulses (which corresponds with the sub-conscious reactions of the subject).
- Defense reflex is the opposite of the orientation reflex, as the agent in a way escapes from the source of the triggered impulses.
- Negative induction mechanism (NIM) allows the blocking of discoveries similar to those already recognized (this way, it can better focus its attention on more important objects).
- Vigilance of attention is a state of readiness for specific stimuli (even a subtle change in the environment) that can technically be easily interpreted in terms of detecting changes.

- Conditioning means adjusting the perception system to search for a specific discovery.

NIM interacts with the scene module, while the remaining blocks of unconscious attention send their control signals to the perception controller.

The overall scheme of attention is shown in Fig. 6. The central parts of the unconscious attention are the blocks representing the scene and the analysis of the discoveries. The current perceptual/context scene remembers (and processes) the recognized and located discoveries based on new (unknown) discoveries stored in high-level short-term memory (H-STM) and knowledge of classified discoveries maintained in the LTM. Additionally, it includes a conscious attention block with indirect access to all agent knowledge that is stored in the LTM.

From a motivational viewpoint, the processing performed by perception analysis begins with defining the context of the needs and emotions associated with the perceived discoveries and their models. It also generates appropriate sub-emotional signals that are the essence of the agent's emotional context and current models of discoveries, all of which are updated. This processing provides the necessary information to trigger subconscious reactions.

The perception controller handles perception channels based on the state of the scene: This plays a key role in the selection process by determining which discoveries visible in the scene are most important. Based on this knowledge, the perception controller decides what the agent should direct its attention to.

3.3 Memory model

As a capacity, human memory can be divided into three categories, due to the storage time, and types of memorized features and elements.

According to the time of storage (Maruszewski 2001) we have

- Ultra-short-time memory of sensors (USTM)—memory used to detect stimuli being temporary and raw form of information,
- Short-term memory (STM)—temporary/current data container, which is divided into:

  – Low-level STM (L-STM) keeping impressions,
  – High-level STM (H-STM) for discoveries,

- Long-term memory (LTM) is the most abstract container that stores objects and reactions.

As *working memory* (Kibbe and Feigenson 2014) or 'liquid intelligence' (Unsworth et al. 2014) served by unconscious and conscious attention, STM contains the currently perceived information (impressions and perceptions) used by attention to interact with the scene and its awareness.

The above cognitive division reflects the structure of the memory processes presented in Tulving (1976) well and is consistent with the classification based on the time of remembering. Edvin Tulving distinguishes *ibid* two types of long-term memory: semantic and episodic. Semantic memory stores commonly used knowledge, word definitions, relationships between them, etc. Episodic memory contains information about events located at specific points in time. Each event can be described as a sequence of basic actions, characters, and relationships between them. Thus, episodic memory can be treated as part of
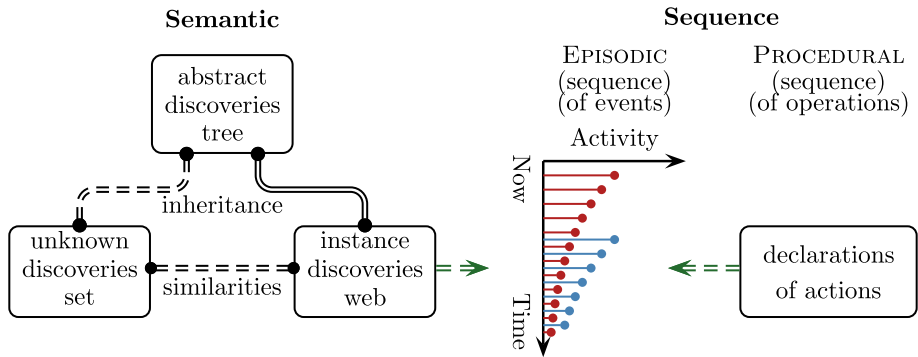
Fig. 7 Functional model of long-time memory (Kowalczuk and Czubenko 2014)

formal sequential memory (Kowalczuk and Czubenko 2014). In this way, remembrances in LTM have two forms: semantic (elementary) and sequential (aggregated) (Fig. 7).

The above categorization also corresponds to other studies (Squire 1992), in which LTM is divided into declarative (facts and events) and non-declarative (skills, non-associative learning and classical conditioning). Declarative memory fits our simple memory structure, while in our opinion, non-declarative memory should rather be integrated with more complex learning mechanisms.

At the implementation stage, STM, which is a working memory containing current discoveries, can be further divided into procedural and semantic parts (Oberauer et al. 2013). Nevertheless, from a functional point of view, the most important is the integration and cooperation of LTM with the *scene* memory, because both memories allow the agent to remember its current environment for consciousness and action (e.g. designing the trajectory of movements). Therefore, the scene module should be seen as the medium-term memory (MTM) being an intermediary between the STM and LTM blocks.

3.3.1 Semantic memory

Semantic memory consists of blocks for declarations (pure knowledge in an encapsulated linguistic form) and semantic networks. Declarative knowledge is crucial for discovery recognition and includes primal impressions, their definitions, and the definitions of shared relationships that are used in defining semantic networks (Milstead 2001) that further describe the discoveries. Note that the discovery may also represent a partially known object.

We distinguish a discovery as an abstract object, called *a-discovery*, which represents a generalized event or finding (e.g., horse, rain, etc.), and instances of discovery, referred to as *i-discoveries* (like the horse called Silver Star or the sunflower that grows in our garden). The terminology used to describe semantic networks is similar to or follows the descriptive logic ($\mathcal{DL}$). However, we must couple the concepts of $\mathcal{DL}$ with the psychological and cognitive point of view (i.e. impressions and discoveries).

All a-discoveries (T-boxes[5]), are arranged using a tree structure of *knowledge*, initiated by a pure abstraction that is an ontological origin (root).

Each abstract discovery has a label, an impressions list, and a relationship list. A-discoveries can also have an emotional context, i.e. a sub-emotion related to that object, and a contextual list of needs that indicates how the possible occurrence of an appropriate i-discovery affects the agent's specific needs.

I-discovery (A-box[6]), represents the specific *experience* of an already known object (a-discovery), which is expressed by a leaf coming out of the corresponding tree node (branch of a given a-discovery). As leaves, i-discoveries are attached to a specific branch (parent), but can also have additional attributes (relative to their parent—a-discovery).

The list of discovery relations can be transformed into a (memory) semantic structure. This means that by using the inheritance relationship (parent–children) and introducing subsequent layers to the tree, the agent's system memory can be formed into a tree hierarchy of a-discoveries and i-discoveries.

The level of activity reflects how often a given i-discovery is detected. It is also related to the necessary forgetting mechanism: if (within a certain time window) a discovery does not occur frequently enough, the agent may have difficulty recognizing it. The index can therefore be used to forget such items as well as to optimize memory space.

3.3.2 Sequential memory

The sequential memory block contains episodic and procedural memories.

Procedural memory is used to store sequences of actions (simple movements, such as moving a hand to a specific position or catching something). Such composed actions can be divided into inner (mental) and outer (with environmental effects). In this memory we place learned solutions (taught or self-developed) representing possible reactions, which are described as a chain of actions that compensate in a desired way for changes noticed by the agent in its world (inner or outer).

The agent can imitate or learn actions effective for its environment (Rizzolatti and Sinigaglia 2008). Own actions may have different origins: The reaction may be conscious (chosen via appropriate thinking processes) or subconscious (usually learned by repeating, without much thinking, just like driving a car by an experienced driver).

As mentioned, episodic memory includes *semantic* events from the past. They represent sequences of (i-)discoveries and related activities located in a specific spatial and temporal context. They also have an activity parameter that allows the agent to control them (filter, forget or recall). This parameter is determined using the forgetting mechanism (Ebbinghaus 2013) and taking into account the agent's current emotion.

3.4 Thinking

Thinking is the most complex of all cognitive phenomena. It can be broken down into a number of sub-processes such as decision making, imagination, inference, problem solving, and more. There are many definitions of thinking worth mentioning (Maruszewski 2001):

---

[5] T-Box is a concept taken from the descriptive logic $\mathcal{DL}$.

[6] As it portrays an assertion field in the $\mathcal{DL}$.

- High activity that fills the gaps in knowledge,
- Creating completely new pieces of information,
- Searching the problem space for a solution,
- Finding a relationship between own experience and the analyzed problem,
- Creating a model of known reality,
- Formal operations concerning syntactic matter,
- Modeling the surrounding environment.

Thinking is a deliberate problem-solving process that involves information processing. New information is the result of thinking. Usually, the main purpose of such a process is to identify relationships between objects, phenomena, events, and concepts that are in the agent's domain of interests. Thinking can also be regarded as a modeling process in which the basic knowledge for constructing models is the *perceived reality* (Nęcka et al. 2008), which is represented both by the perceptual scene and the agent's mental state. Based on the scenic image of the environment, the agent can perform a full simulation of its reaction. The model obtained includes all objects in the scene and their hypothetical reactions. In this way, the agent can mentally consider, predict and evaluate the effects of its environmental impact.

Thinking can be categorized into two types of processes: autistic and realistic (Berlyne 1969). Autistic thinking has no apparent purpose and is characteristic of states of relaxation and deep sleep. It allows the agent to dream (mentally test) of meeting important but unmet needs. It is also responsible for organizing and structuring the LTM memory. In addition, autistic thinking fulfils (Maruszewski 2001):

- Motivational function of sorting goals,
- Compensation function to escape from reality,
- Training (education) function, and
- Self-knowledge development function.

On the other hand, realistic thinking is focused on a specific goal, taking into account the limitations of reality. It is responsible for rational decision making and solving emerging problems.

The process of (realistic) thinking may be productive and reproductive, or creative and recreative (Nęcka et al. 2008). Creative thinking generates new solutions to a given problem, while reconstructive thinking is the search for only a well-established solution, experienced and positively verified. The thinking process can also be evaluative or only critical. Evaluation can refer to the assessment of the various parameters (weight, estimate or grade) that characterize the concept under consideration (object, reaction, relationship, goal, criterion, self aspect, etc.).

In the current version of the Intelligent System of Decision-making (ISD), the thinking mechanisms have been limited to the partial implementation of realistic thinking shown in Fig. 8. ISD provides the most appropriate responses to the current state of interaction with the agent's environment. The appropriate reaction can be chosen in two ways: creative and reconstructive.

Creative thinking can now be derived using the evolutionary computation (Kowalczuk and Białaszewski 2006, 2018) or other exploratory methods (Kowalczuk and Oliński 2012) that are able to create completely new solutions. In such optimisation, the sought reaction, being a sequence of simple movements (see Sect. 3.3), can be computer coded as an appropriate phenotype.
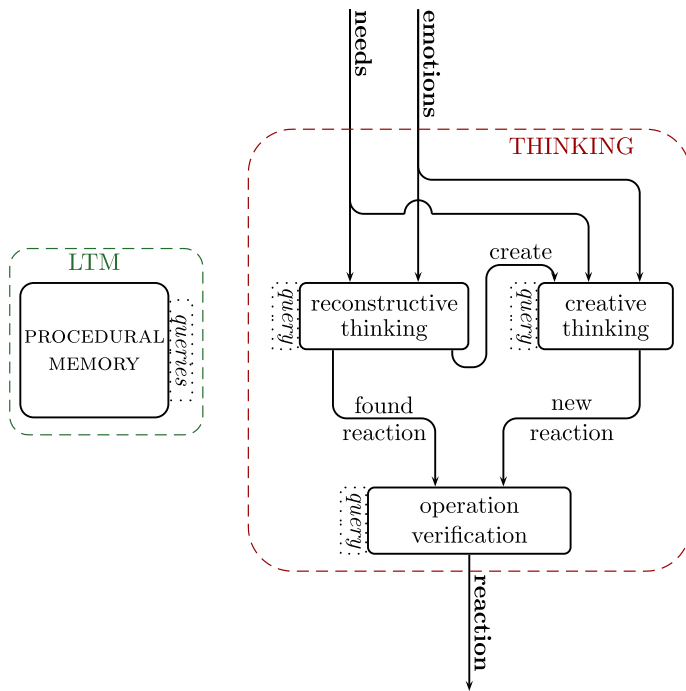
Fig. 8 Cognitive process—realistic thinking built in accordance with the model of cognition (Kowalczuk and Czubenko 2011)

Reconstructive thinking in our implementation consists in choosing the best-suited reaction from a pre-programmed set of reactions (available in procedural memory) depending on the situation of the environment and the internal state of the agent (in terms of needs and emotions).

The effectiveness of the ISD response is assessed on the basis of the anticipated satisfaction of the needs currently 'felt' by the agent that results from that response. This assessment can be done promptly using a fuzzy neural network (FNN) (Kowalczuk and Czubenko 2011). The fuzzy rule implemented by this network can be interpreted as: '*all needs are to be satisfied*' and '*none of the needs should be in a prealarm or alarm state*'. The best available response is indicated by the highest output of this network. Other details of the presented FNN are given in Kowalczuk and Czubenko (2011), including fuzzy models of neurons, and the OR and AND operands (Pedrycz and Rocha 1993).

The performance verification block is responsible for selecting and executing the response, which may also decide to stop the current response in the event of critical changes in the environment.

4 Implementations of ISD

As shown in the previous sections, the ISD architecture, described in part here and partly in our earlier publications (Kowalczuk and Czubenko 2010, 2011, 2013), is the result of in-depth modeling of human psychology. Concepts similar to ISD can be found in the

literature (e.g. Muhlestein 2012). From the point of view of psychology, ISD treated as cognitive architecture can only be compared with the most advanced architecture—Soar. In addition, ISD implements most aspects of the unified standard mind model (Laird et al. 2017). However, ISD places a strong emphasis on the agent's dual motivational system, which is an innovative aspect among cognitive architectures (since most have one motivational aspect).

The ISD system was implemented in Python, with support for several external libraries, such as scikit-fuzzy and numpy. Fuzzy norms and other mathematical operators have been programmed with numpy arrays. For demonstration purposes, we used the PyQt and matplotlib libraries. So far, no computational performance tests have been performed.

The correctness of the applied assumptions and concepts as well as the effectiveness of the ISD project have already been partially proven in several practical applications. The simple implementation of a micro robot (in short *Mu-* or *μ*-robot) proved that ISD can act as an autonomous system (Kowalczuk and Czubenko 2010; Castro-González et al. 2014). Simulation study of a prototype agent *xDriver* showed that ISD can easily act as an autonomous car driver on the road, doing the basic task of speed control (Czubenko et al. 2015; Li and Shao 2015). Also other avatars created in university laboratories are able to demonstrate human-friendly utility (Chybiński 2012; Kowalczuk and Klimczak 2013; Rybka and Janicki 2013). The human-friendly aspect of robots also corresponds to the idea of fuzzy emotions (Pelzl et al. 2020; Karyotis et al. 2018; Bonarini 2016). In this context, another developed system narrowed only to face recognition (Kowalczuk and Chudziak 2018) turns out to be a good candidate for the next part of the ISD architecture.

For the description of needs and emotions (as variables of the internal state) the concept of fuzzy variables/sets seems to be the most natural. Both from the human and computer point of view, the linguistic interpretation of these variables is necessary for their simple practical use (including building human-system interfaces). From the side of a computerized agent, the fuzzy approach allows for an effective interpretation of environmental measurements. The applied fuzzy-neural inference mechanism (so far only in terms of needs) has passed the functional tests of an autonomous driver (Czubenko et al. 2015). Our future goal is to extend this idea to the interpretation of environmental facts (in the form of fuzzy impressions and discoveries) and their respective rules (such as 'only red apples are edible').

The types of functions used to describe needs or emotions depend on the agent configuration used, although in our experience this should not be a critical issue. However, examining the influence of the membership function on the operation of the ISD system can certainly be one of the next research directions.

The main disadvantage of using the fuzzy approach is the difficulty of testing, e.g. when selecting a reaction inferred from a given internal state. Even with only seven needs (as in the xDriver) it is very difficult to successfully develop an agent's behavior. Moreover, the fuzzy system makes the agent nondeterministic, because even in a steady state the selected (sharpened) response may be different.

5 Conclusions

The concepts presented here, based on the achievements of psychology, are in line with the observed current trends in scientific research related to robotics and artificial intelligence. In particular, the development of human behavior models for the construction of virtual agents and autonomous robots is one of the most important challenges of modern science.

The main contribution of this article has been to present the Intelligent System of Decision-making as a new cognitive architecture that attaches great importance to motivational factors (needs and emotions). This system was conceived as a coherent concept containing a large proportion of the results of human psychology. It is to serve the task of building an anthropomorphic control system for an autonomous robot.

In the development of the ISD system, we use various concepts from cognitive, humanistic and motivational psychology, as well as from the theory of emotions. From this point of view, it is both a novel approach and innovative architecture. On the other hand, we also show a systemic, unique and feasible approach to modeling the human brain.

Taking into account the goals of modern autonomous robotics, the work consolidates the broad knowledge known from both developed cognitive psychology and the theory of motivation. The resulting models and algorithms were synthesized into a complete and functional decision-making system containing all the components necessary for robotic purposes.

The ISD architecture, like other cognitive architectures, tries to combine bottom-up and top-down AI approaches and methods. However, it is important to remember that known cognitive architectures vary widely, both in concept and implementation.

The ISD system presented in the article extends the concept of cybernetics in an innovative way, which, apart from modeling various useful subsystems, can also propose a holistic approach to building intelligent systems. On the practical side, ISD is an interdisciplinary project that is (top-down) based on psychology, applies automation methods and systems and software engineering, uses (bottom-up) various computational techniques, combining them all into one system that can be tailored to any agent or robotic tasks.

The developed ISD model is therefore suitable for the construction of coherent autonomous systems in an embodied form with the use of an appropriate engineering platform. Such a system can have various applications, e.g. an intelligent assembly robot, a reconnaissance robot, a human/patient assistant, an internet chatter, a smart toy or a programming agent for solving specific-domain problems based on some form of intelligence, etc. This approach, in turn, will enable the future development of artificial intelligence and the use of fully autonomous systems.

Following the well-known cybernetic engineering approach to system design, instead of using only a rational decision path (classical, with inference and optimization), we introduce to our vision of the autonomous unit (ISD) a form of motivational system that uses an emotional path (fast and reflective), which technically relies on the emotion of the agent, which is here the generalized scheduling variable.

In general, it can be safely stated that the presented concepts (analysis and synthesis/modeling) based on psychology, as well as the results of experimental research, fit into the current trends and their various manifestations in scientific research related to robotics and artificial intelligence.

The proposed approach can also be applied to human–robot interactions and human-system interfaces. It provides functions that make robots or software applications more human: natural, intelligent, and user-friendly.

The system has already proved its functionality and usability in several implemented applications. While key parts of the model have been implemented and tested, there is still a long way to go to fully validate the entire ISD. However, the main purpose of this article was to provide a general overview of the ISD architecture.

Further work expanding the idea of ISD should address several important modeling and design issues, such as building the necessary types of agent memory and solving thinking problems, including creating ontologies, naming newly perceived objects, and conducting autistic thinking, responsible for optimizing and restructuring LTM memory. It is also important to integrate the xEmotion system with the systems of visual recognition as well as the observation and management of the agent's perceptual scene.

References

Arons B (1992) A review of the cocktail party effect. MIT Media Lab, Cambridge

Batson CD (1990) Affect and altruism. In: Moore BS, Isen AM (eds) Affect and social behavior. Cambridge University Press, Cambridge, pp 89–126

Becker-Asano C (2008) WASABI: affect simulation for agents with believable interactivity, vol 319. IOS Press, Amsterdam

Berlyne DE (1969) Structure and direction of thinking process. Wiley, Madison

Biddle S, Fox KR, Boutcher SH (2000) Physical activity and psychological well-being. Psychology Press, London

Bonarini A (2016) Can my robotic home cleaner be happy? Issues about emotional expression in non-bio-inspired robots. Adapt Behav 24(5):335–349

Bradley MM, Greenwald MK, Petry MC, Lang PJ (1992) Remembering pictures: pleasure and arousal in memory. J Exp Psychol Learn Mem Cogn 18(2):379–90

Broadbent DE (1957) A mechanical model for human attention and immediate memory. Psychol Rev 64(3):205–2015

Brooks RA (1989) A robot that walks; emergent behaviors from a carefully evolved network. In: Robotics and automation. International conference. MIT AI Laboratory, Scottsdale, pp 692–696

Bruner JS (1973) Beyond the information given: studies in the psychology of knowing. W.W. Norton & Company Incorporated, New York

Cambria E, Livingstone A, Hussain A (2012) The hourglass of emotions. In: Cognitive behavioural systems. Springer, Berlin, pp 144–157

Castro-González A, Malfaz M, Gorostiza JF, Salichs MA (2014) Learning behaviors by an autonomous social robot with motivations. Cybern Syst 45(7):568–598

Chybiński A (2012) Lingubot—presentation avatar with verbal communication. MSc Thesis (supervisor prof. Z. Kowalczuk), Faculty of ETI, Gdańsk

Clore GL, Ortony A (2013) Psychological construction in the OCC model of emotion. Emot Rev 5(4):335–343

Czubenko M (2017) Anthropoid model of intelligent decision system for autonomous units. (Supervisor prof. Z. Kowalczuk), Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Gdańsk

Czubenko M, Kowalczuk Z (2019) Elements of psychology in the context of robot autonomy. Pomorskie Wydawnictwo Naukowo-Techniczne, Gdańsk

Czubenko M, Kowalczuk Z, Ordys A (2015) Autonomous driver based on intelligent system of decision-making. Cogn Comput 7(5):569–581

Driver J (2001) A selective review of selective attention research from the past century. Br J Psychol 92(1):53–78

Ebbinghaus H (2013) Memory: a contribution to experimental psychology. Ann Neurosci 20(4):155

Ekman P (2009) Telling lies. Clues to deceit in the marketplace, politics and marriage. W.W. Norton & Company, New York

Ekman P, Cordaro D (2011) What is meant by calling emotions basic. Emot Rev 3(4):364–370

El-Nasr MS, Yen J, Ioerger TR (2000) Flame-fuzzy logic adaptive model of emotions. Auton Agents Multi-agent Syst 3(3):219–257

Frijda NH (1986) The emotions. Cambridge University Press, Cambridge

Gallagher HL, Frith CD (2003) Functional imaging of 'theory of mind'. Trends Cogn Sci 7(2):77–83

Gilbert AN, Fridlund AJ, Lucchina LA (2016) The color of emotion: a metric for implicit color associations. Food Qual Prefer 52:203–210

Hardin CL (1987) Qualia and materialism: closing the explanatory gap. Philos Phenomenol Res 48:281–298

Hebb DO (1958) Textbook of psychology. Saunders Company, Philadelphia, London

Herzberg F (1965) The motivation to work among Finnish supervisors. Pers Psychol 18:393–402

Karyotis C, Doctor F, Iqbal R, James A, Chang V (2018) A fuzzy computational model of emotion for cloud based sentiment analysis. Inf Sci 433:448–463

Kibbe MM, Feigenson L (2014) Developmental origins of recoding and decoding in memory. Cogn Psychol 75:55–79

Kowalczuk Z, Białaszewski T (2006) Niching mechanisms in evolutionary computations. Int J Appl Math Comput Sci 16(1):59–84

Kowalczuk Z, Białaszewski T (2018) Gender approaches to evolutionary multi-objective optimization using pre-selection of criteria. Eng Optim 50(1):120–144

Kowalczuk Z, Chudziak P (2018) Identification of emotions based on human facial expressions using a color-space approach. In: Kościelny JM, Syfert M, Sztyber A (eds) Advanced solutions in diagnostics and fault tolerant control. Advances in intelligent systems and computing, vol 635. Springer, Cham, pp 291–303

Kowalczuk Z, Czubenko M (2010) Interactive cognitive-behavioural decision making system. In: Artifical intelligence and soft computing lecture notes in computer science. Lecture notes in artificial intelligence, vol 6114(II). Springer, Berlin, Heidelberg, New York, pp 516–523

Kowalczuk Z, Czubenko M (2011) Intelligent decision-making system for autonomous robots. Int J Appl Math Comput Sci 21(4):621–635

Kowalczuk Z, Czubenko M (2013) Xemotion—a computational model of emotions dedicated to intelligent decision-making systems. Pomiary Automatyka Robotyka 17(2):60–65

Kowalczuk Z, Czubenko M (2014) Cognitive memory for intelligent systems of decision-making, based on human psychology. In: Intelligent systems in technical and medical diagnostics advances in intelligent systems and computing. Advances in intelligent systems and computing, vol 230. Springer, Berlin, Heidelberg, pp 379–389

Kowalczuk Z, Czubenko M (2016) Computational approaches to modeling artificial emotion—an overview of the proposed solutions. Front Robot AI 3(21):1–12. https://doi.org/10.3389/frobt.2016.00021

Kowalczuk Z, Czubenko M (2017) Emotions embodied in the svc of an autonomous driver system. IFAC-PapersOnLine 50(1):3744–3749

Kowalczuk Z, Czubenko M (2021) Intelligent system of decision-making as machine implementation of cognitive and motivational processes. Pomorskie Wydawnictwo Naukowo-Techniczne, Gdańsk

Kowalczuk Z, Klimczak J (2013) A system of intelligent navigation through the internet service, controlled by voice. Pomiary Automatyka Kontrola 59(7):644–647

Kowalczuk Z, Oliński KE (2012) Toolgraph design of optimal and feasible strategies for time-varying dynamical systems. Math Probl Eng 2012(142060):1–16

Kowalczuk Z, Czubenko M, Jędruch W (2016) Learning processes in autonomous agents using an intelligent system of decision-making. In: Kowalczuk Z (ed) Advanced and intelligent computations in diagnosis and control. Advances in intelligent systems and computing, vol 386. Springer, Berlin, pp 301–315

Kowalczuk Z, Czubenko M, Merta T (2020) Interpretation and modeling of emotions in the management of autonomous robots using a control paradigm based on a scheduling variable. Eng Appl Artif Intell 91:103562

Kowalczuk Z, Czubenko M, Gruba M (2021) Qualia: about personal emotions representing temporal form of impressions—implementation hypothesis and application example. Report, Faculty of ETI, Gdańsk

Krolikowski R, Kopys M, Jedruch W (2016) Self-organization in multi-agent systems based on examples of modeling economic relationships between agents. Front Robot AI 3(41):1–16. https://doi.org/10.3389/frobt.2016.00041

Laird JE, Lebiere C, Rosenbloom PS (2017) A standard model of the mind: toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. AI Mag 38(4):13–26

Lazarus RS (1991) Emotion and adaptation. Oxford University Press, New York

Łęski J (2008) Neural fuzzy systems. WNT, Warsaw

Li B, Shao Z (2015) Simultaneous dynamic optimization: a trajectory planning method for nonholonomic car-like robots. Adv Eng Softw 87:30–42

Lienhart R, Maydt J (2002) An extended set of Haar-like features for rapid object detection. In: International conference on image processing. 2002 International conference on image processing, vol 1. p 900

Lindsay PH, Norman DA (1977) Human information processing: introduction to psychology. Academic Press Inc, Warszawa

Lövheim H (2012) A new three-dimensional model for emotions and monoamine neurotransmitters. Med Hypotheses 78(2):341–8

Marsella SC, Gratch J (2009) EMA: a process model of appraisal dynamics. Cogn Syst Res 10(1):70–90

Marsella S, Gratch J, Petta P (2010) Computational models of emotion. In: A blueprint for affective computing—a sourcebook and manual, vol 11, no 1. Oxford University Press, New York, pp 21–46

Maruszewski T (2001) Psychology of cognition. Gdańskie Wydawnictwo Psychologiczne, Gdańsk

Maslow AH (2012) Toward a psychology of being. Start Publishing, New York

Mazur M (1976) Cybernetics and character. PIW, Warszawa

McGinnies E (1949) Emotionality and perceptual defense. Psychol Rev 56(5):244

Milstead JL (2001) Standards for relationships between subject indexing terms. In: Bean CA, Green R (eds) Relationships in the organization of knowledge information science and knowledge management, vol 2. Springer, Dordrecht, pp 53–66

Monsell S (1996) Control of mental processes. In: Bruce V (ed) Tutorial essays in unsolved mysteries of the mind. Erlbaum Taylor & Francis, Hove, pp 93–148

Muhlestein M (2012) Counterfactuals, computation, and consciousness. Cogn Comput 5(1):99–105

Murphy ST, Zajonc RB (1994) Affect, cognition and awareness: the role of affective preceding stimuli at optimal and suboptimal exposures. Przegl Psychol 37:261–299

Murray HA (1938) Explorations in personality. Oxford University Press, Oxford

Nęcka E, Orzechowski J, Szymura B (2008) Cognitive psychology. PWN, Warszawa

Noltemeyer A, Bush K, Patton J, Bergen D (2012) The relationship among deficiency needs and growth needs: an empirical investigation of Maslow's theory. Child Youth Serv Rev 34(9):1862–1867

Oatley K, Keltner D, Jenkins J (2012) Understanding emotions, 2nd edn. Blackwell Publishing, Oxford

Oberauer K, Souza AS, Druey MD, Gade M (2013) Analogous mechanisms of selection and updating in declarative and procedural working memory: experiments and a computational model. Cogn Psychol 66(2):157–211

Pedrycz W, Rocha AF (1993) Fuzzy-set based models of neurons and knowledge-based networks. IEEE Trans Fuzzy Syst 1(4):254–266

Pelzl F, Diepold K, Auernhammer J (2020) Designing a multimodal emotional interface in the context of negotiation. In: International conference on human–computer interaction. Springer, Cham, pp 501–520

Pennycook G, Fugelsang JA, Koehler DJ (2015) What makes us think? A three-stage dual-process model of analytic engagement. Cogn Psychol 80:34–72

Plutchik R (1980) A general psychoevolutionary theory of emotion. In: Plutchik R, Kellerman H (eds) Emotion: theory, research, and experience, vol 1. Academic Press, New York, pp 3–33

Porębski S, Straszecka E (2016) Membership functions for fuzzy focal elements. Arch Control Sci 26(3):395–427

Posner J, Russell JA, Peterson BS (2005) The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev Psychopathol 17(3):715–34

Rizzolatti G, Sinigaglia C (2008) Mirrors in the brain. How we share our actions and emotions. Oxford University Press, Oxford

Rubin DC, Talarico JM (2009) A comparison of dimensional models of emotion: evidence from emotions, prototypical events, autobiographical memories, and words. Memory (Hove, England) 17(8):802–808

Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 39(6):1161–1178

Ryan R, Deci E (2000) Self-determination theory and the facilitation of intrinsic motivation. Am Psychol 55(1):68–78. https://doi.org/10.1037/0003-066X.55.1.68

Rybka J, Janicki A (2013) Comparison of speaker dependent and speaker independent emotion recognition. Int J Appl Math Comput Sci 23(4):797–808

Samsonovich AV (2010) Toward a unified catalog of implemented cognitive architectures. BICA 221(2010):195–244

Santos R, Marreiros G, Ramos C, Neves J, Bulas-Cruz J (2011) Personality, emotion, and mood in agent-based group decision making. IEEE Intell Syst 26(06):58–66

Shvo M, Buhmann J, Kapadia M (2019) An interdependent model of personality, motivation, emotion, and mood for intelligent virtual agents. In: Proceedings of the 19th ACM international conference on intelligent virtual agents. pp 65–72

Squire LR (1992) Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. J Cogn Neurosci 4(3):232–243

Steunebrink BR, Dastani M, Meyer J-JC (2009) The OCC model revisited. In: Proceedings of the 4th workshop on emotion and computing. Association for the Advancement of Artificial Intelligence, Palo Alto, p 62

Thayer R (1989) The biopsychology of mood and arousal. Oxford University Press, Oxford

Treisman A (1986) Features and objects in visual processing. Sci Am 255:114–125

Tulving E (1976) Ecphoric processes in recall and recognition. In: Brown J (ed) Recall Recogn. Wiley, London, pp 37–73

Ulrich R, Schröter H, Leuthold H, Birngruber T (2015) Automatic and controlled stimulus processing in conflict tasks: superimposed diffusion processes and delta functions. Cogn Psychol 78:148–174

Unsworth N, Fukuda K, Awh E, Vogel EK (2014) Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval. Cogn Psychol 71:1–26

Watson D, Tellegen A (1985) Toward a consensual structure of mood. Psychol Bull 98(2):219–235

Whitfield TW, Wiltshire TJ (1990) Color psychology: a critical review. Genetic, social, and general psychology monographs 116(4):385–411

Wu W (2012) The relationship between incentives to learn and Maslow's hierarchy of needs. Phys Procedia 24:1335–1342

Wu Q, Miao C (2013) Modeling curiosity-related emotions for virtual peer learners. IEEE Comput Intell Mag 8(2):50–62

Zajonc RB, Murphy ST, Inglehart M (1989) Feeling and facial efference: implications of the vascular theory of emotion. Psychol Rev 96(3):395–416