CrossMark

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · **John K. Tsotsos[1]**

Abstract

In this paper we present a broad overview of the last 40 years of research on cognitive architectures. To date, the number of existing architectures has reached several hundred, but most of the existing surveys do not reflect this growth and instead focus on a handful of well-established architectures. In this survey we aim to provide a more inclusive and high-level overview of the research on cognitive architectures. Our final set of 84 architectures includes 49 that are still actively developed, and borrow from a diverse set of disciplines, spanning areas from psychoanalysis to neuroscience. To keep the length of this paper within reasonable limits we discuss only the core cognitive abilities, such as perception, attention mechanisms, action selection, memory, learning, reasoning and metareasoning. In order to assess the breadth of practical applications of cognitive architectures we present information on over 900 practical projects implemented using the cognitive architectures in our list. We use various visualization techniques to highlight the overall trends in the development of the field. In addition to summarizing the current state-of-the-art in the cognitive architecture research, this survey describes a variety of methods and ideas that have been tried and their relative success in modeling human cognitive abilities, as well as which aspects of cognitive behavior need more research with respect to their mechanistic counterparts and thus can further inform how cognitive science might progress.

Keywords Survey · Cognitive architectures · Perception · Attention · Cognitive abilities · Practical applications

✉ Iuliia Kotseruba
  yulia_k@eecs.yorku.ca

  John K. Tsotsos
  tsotsos@eecs.yorku.ca

[1] Department of Electrical Engineering and Computer Science, York University, Toronto, ON, Canada

⊴ Springer

1 Introduction

The goal of this paper is to provide a broad overview of the last 40 years of research in cognitive architectures with an emphasis on the core capabilities of perception, attention mechanisms, action selection, learning, memory, reasoning, metareasoning and their practical applications. Although the field of cognitive architectures has been steadily expanding, most of the surveys published in the past 10 years do not reflect this growth and feature essentially the same set comprising a few established architectures.

The latest large-scale study was conducted by Samsonovich (2010) in an attempt to catalog the implemented cognitive architectures. His survey contains descriptions of 26 cognitive architectures submitted by their respective authors. The same information is also presented online in a Comparative Table of Cognitive Architectures.[1] In addition to surveys, there are multiple on-line sources listing cognitive architectures, but they rarely go beyond a short description and a link to the project site or a software repository.

Since there is no exhaustive list of cognitive architectures, their exact number is unknown, but it is estimated to be around three hundred, out of which at least one-third of the projects are currently active. To form the initial list for our study we combined the architectures mentioned in surveys (published within the last 10 years) and several large on-line catalogs. We also included more recent projects not yet mentioned in the survey literature. Figure 1 shows a visualization of 195 cognitive architectures featured in 17 sources (surveys, on-line catalogs and on Google Scholar). It is apparent from this diagram that a small group of architectures such as ACT-R, Soar, CLARION, ICARUS, EPIC, LIDA and a few others are present in most sources, and all other projects are only briefly mentioned in on-line catalogs.

Even though the theoretical and practical contributions of the major architectures are undeniable, they represent only a part of the research in the field. Thus, in this review the focus is shifted away from the deep study of the major architectures or discussion of what could be the best approach to modeling cognition, which has been done elsewhere. For example, in a recent paper by Laird et al. (2017) ACT-R, Soar and Sigma are compared based on their structural organization and approaches to modelling core cognitive abilities. Further, a new Standard Model of the Mind is proposed as a reference model born out of consensus between the three architectures. Our goal, on the other hand, is to present a broad, inclusive, judgment-neutral snapshot of the past 40 years of development of the field. We hope to inform the future cognitive architecture research by introducing the diversity of ideas that have been tried and their relative success.

To make this survey manageable we reduced the original list of architectures to 84 items by considering only implemented architectures with at least one practical application and several peer-reviewed publications. We do not explicitly include some of the philosophical architectures such as CogAff (Sloman 2003), Society of Mind (Minsky 1986), Global Workspace Theory (GWT) (Baars 2005) and Pandemonium Theory (Selfridge 1958), however we examine cognitive architectures heavily influenced by these theories (e.g. LIDA, ARCADIA, CERA-CRANIUM, ASMO, COGNET and Copycat/Metacat, also see discussion in Sect. 5). We also exclude large-scale brain modeling projects, which are low-level and do not easily map onto the breadth of cognitive capabilities modeled by other types of cognitive architectures. Further, many of the existing brain models do not yet have practical applications, and thus do not fit the parameters of the present survey. Figure 2 shows all

---

[1] http://bicasociety.org/cogarch/architectures.htm.

Fig. 1 A diagram showing cognitive architectures found in literature surveys and on-line sources shown with blue and orange colors respectively. The architectures in the diagram are sorted in the descending order by the total number of references in the surveys and on-line sources for each architecture. Titles of the architectures covered in this survey are shown in red. All visualizations in this paper are made using the D3.js library (https://d3js.org). Interactive versions of the figures are available on the project website. The color palettes are generated with ColorBrewer (http://colorbrewer2.org)

architectures featured in this survey with their approximate timelines recovered from the publications. Of these projects 49 are currently active.[2]

As we mentioned earlier, the first step towards creating an inclusive and organized catalog of implemented cognitive architectures was made by Samsonovich (2010). His work contained extended descriptions of 26 projects with the following information: short overview, a schematic diagram of major elements, common components and features (memory types,

[2] Since the exact dates for the start (and end) of the development are not specified for the majority of architectures, we use instead the dates of the earliest and the latest publications we could find. We consider the project under the active development if there was at least one publication, within the last year (2017). In the absense of recent publications we also take into account activity on the websites or online code repositories associated with architectures (i.e. any webpage updates and code commits in 2017).
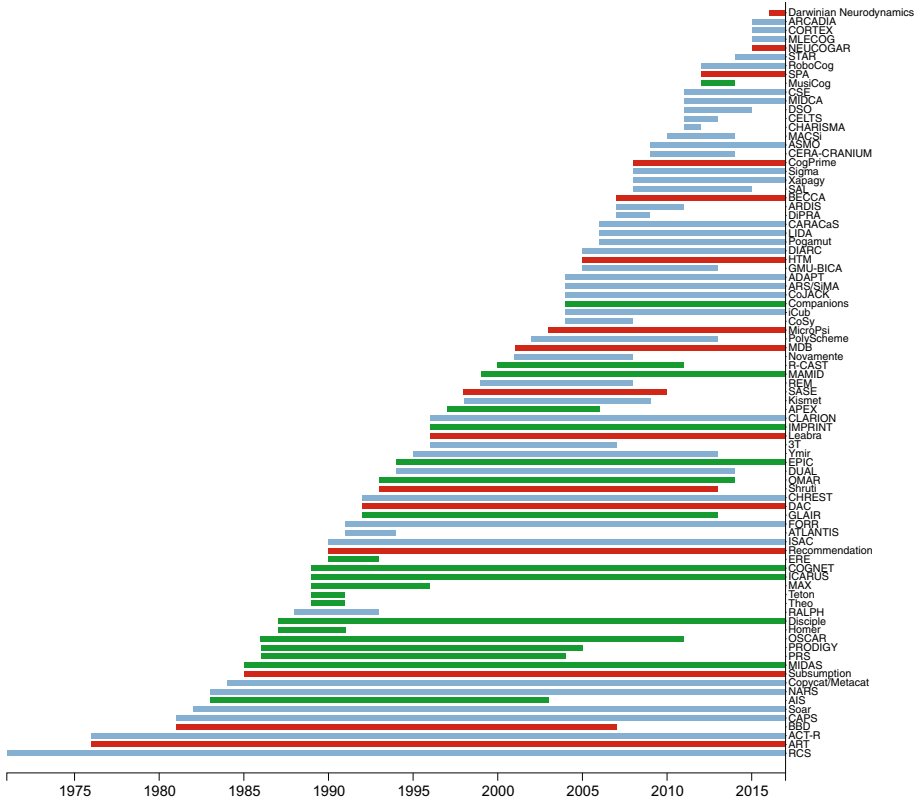
Fig. 2 A timeline of 84 cognitive architectures featured in this survey. Each line corresponds to a single architecture. The architectures are sorted by the starting date, so that the earliest architectures are plotted at the bottom of the figure. Since the explicit beginning and ending dates are known only for a few projects, we recovered the timeline based on the dates of the publications and activity on the project web page or on-line repository. Colors of the lines correspond to different types of architectures: symbolic (green), emergent (red) and hybrid (blue). According to this data there was a particular interest in symbolic architectures since mid-1980s until early 1990s, however after 2000s most of the newly developed architectures are hybrid. Emergent architectures, many of which are biologically-inspired, are more evenly distributed but remain a relatively small group

attention, consciousness, etc.), learning and cognitive development, cognitive modeling and applications, scalability and limitations. A survey of this kind brings together researchers from several disjoint communities and helps to establish a mapping between the different approaches and terminology they use. However, the descriptive or tabular format does not facilitate easy comparisons between architectures. Since our sample of architectures is large, we experimented with alternative visualization strategies.

In the following sections, we will provide an overview of the definitions of cognition and approaches to categorizing cognitive architectures. As one of our contributions, we map cognitive architectures according to their perception modality, implemented mechanisms of attention, memory organization, types of learning, action selection and practical applications.

In the process of preparing this paper, we thoroughly examined the literature and this activity led to an extensive bibliography of more than 2500 relevant publications. We provide

this bibliography, as well as interactive versions of the diagrams in this paper on our project webpage.[3]

2 What are cognitive architectures?

Cognitive architectures are a part of research in general AI, which began in the 1950s with the goal of creating programs that could reason about problems across different domains, develop insights, adapt to new situations and reflect on themselves. Similarly, the ultimate goal of research in cognitive architectures is to model the human mind, eventually enabling us to build human-level artificial intelligence. To this end, cognitive architectures attempt to provide evidence what particular mechanisms succeed in producing intelligent behavior and thus contribute to cognitive science. Moreover, the body of work represented by the cognitive architectures covered in this review, documents what methods or strategies have been tried previously (and what have not), how they have been used, and what level of success has been achieved or lessons learned, all important elements that help guide future research efforts. For AI and engineering, documentation of past mechanistic work has obvious import. But this is just as important for cognitive science, since most experimental work eventually attempts to connect to explanations of how observed human behavior may be generated.

According to Russell and Norvig (1995) artificial intelligence may be realized in four different ways: systems that think like humans, systems that think rationally, systems that act like humans, and systems that act rationally. The existing cognitive architectures have explored all four possibilities. For instance, human-like thought is pursued by the architectures stemming from cognitive modeling. In this case, the errors made by an intelligent system should match the errors typically made by people in similar situations. This is in contrast to rationally thinking systems which are required to produce consistent and correct conclusions for arbitrary tasks. A similar distinction is made for machines that act like humans or act rationally. Machines in either of these groups are not expected to think like humans, only their actions or behavior is taken into account.

Given the multitude of approaches that may lead to human-level AI and in the absense of clear definition and general theory of cognition, each cognitive architecture is built on a particular set of premises and assumptions, making comparison and evaluation of progress across architectures difficult. Several papers were published to resolve these issues, the most prominent being Sun's desiderata for cognitive architectures (Sun 2004) and Newell's functional criteria [first published in Newell (1980) and Newell (1992)], and later restated by Anderson and Lebiere (2003). Newell's criteria include flexible behavior, real-time operation, rationality, large knowledge base, learning, development, linguistic abilities, self-awareness and brain realization. Sun's desiderata are broader and include ecological, cognitive and bio-evolutionary realism, adaptation, modularity, routineness and synergistic interaction. Besides defining these criteria and applying them to a range of cognitive architectures, Sun also pointed out the lack of clearly defined cognitive assumptions and methodological approaches, which hinder progress in studying intelligence. He also noted an uncertainty regarding essential dichotomies (implicit/explicit, procedural/declarative, etc.), modularity of cognition and structure of memory. However, a quick look at the existing cognitive architectures reveals persisting disagreements in terms of their research goals, structure, operation and application.

Given the issues with defining intelligence (Legg and Hutter 2007), a more practical solution is to treat it as a set of competencies and behaviors demonstrated by the system. While

---

[3] http://jtl.lassonde.yorku.ca/project/cognitive_architectures_survey/.

no comprehensive list of capabilities required for intelligence exists, several broad areas have been identified that may serve as guidance for ongoing work in the cognitive architecture domain. For example, Adams et al. (2012) suggest broad areas such as perception, memory, attention, actuation, social interaction, planning, motivation, emotion, etc. These are further split into subareas. Arguably, some of these categories may seem more important than the others and historically attracted more attention (further discussed in Sect. 11.2).

At the same time, implementing even a reduced set of abilities in a single architecture is a substantial undertaking. Unsurprisingly, the goal of achieving Artificial General Intelligence (AGI) is explicitly pursued by a small number of architectures, among which are Soar, ACT-R, NARS (Wang 2013), LIDA (Faghihi and Franklin 2012), and several recent projects, such as SiMA (formerly ARS) (Schaat et al. 2015), Sigma (Pynadath et al. 2014) and CogPrime (Goertzel and Yu 2014). Majority of architectures study particular aspects of cognition, e.g. attention [ARCADIA (Bridewell and Bello 2015), STAR (Tsotsos 2017)], emotions [CELTS (Faghihi et al. 2011b)], perception of symmetry [Cognitive Symmetry Engine (Henderson and Joshi 2013)] or problem solving [FORR (Epstein 2004), PRODIGY (Epstein 2004)]. There are also narrowly specialized architectures designed for particular applications, such as ARDIS (Martin et al. 2009) for visual inspection of surfaces or MusiCog (Maxwell 2014) for music comprehension and generation.

In view of such diversity of existing architectures and their proclaimed goals, naturally, the question then arises as to what system can be considered a cognitive architecture. Different opinions on this can be found in the literature. Laird (2012a) discusses how cognitive architectures differ from other intelligent software systems. While all of them have memory storage, control components, data representation, and input/output devices, the latter provide only a fixed model for general computation. Cognitive architectures, on the other hand, must change through development and efficiently use knowledge to perform new tasks. Furthermore, he suggests that toolkits and frameworks for building intelligent agents (e.g. GOMS, BDI, etc.) cannot themselves be considered cognitive architectures. [Note, however, the authors of Pogamut, a framework for building intelligent agents, consider it a cognitive architecture as it is included in Samsonovich (2010).] Some researchers contrast the engineering approach taken in the field of artificial intelligence with the scientific approach of cognitive architectures (Sun 2007; Jordan and Russell 1999; Lieto et al. 2018b). According to Sun, psychologically based cognitive architectures should facilitate the study of human mind by modeling not only the human behavior but also the underlying cognitive processes. Such models, unlike software engineering-oriented "cognitive" architectures, are explicit representations of the general human cognitive mechanisms, which are essential for understanding of the mind.

In practice the term "cognitive architecture" is not as restrictive, as made evident by the representative surveys of the field. Most of the surveys define cognitive architectures as a blueprint for intelligence, or, more specifically, a proposal about the mental representations and computational procedures that operate on these representations enabling a range of intelligent behaviors (Butt et al. 2013; Duch et al. 2008; Langley et al. 2009; Profanter 2012; Thagard 2012). There is generally no need to justify the inclusion of the established cognitive architectures such as Soar, ACT-R, EPIC, LIDA, CLARION, ICARUS and a few others. The same applies to many biomimetic and neuroscience-inspired cognitive architectures that model cognitive processes on a neuronal level (e.g. CAPS, BBD, BECCA, DAC, SPA). Arguably, engineering-oriented architectures pursue a similar goal as they have a set of structural components for perception, reasoning and action, and model interactions between them, which is very much in the spirit of the Newell's call for "unified theories of cognition" (Hayes-Roth 1995). However, when it comes to less common or new projects, the reasons

for considering them are less clear. As an example, AKIRA, a framework that explicitly does not self-identify as a cognitive architecture (Pezzulo and Calvi 2005), is featured in some surveys anyway (Thórisson and Helgasson 2012). Similarly, a knowledge base Cyc (Foxvog 2010), which does not make any claims about general intelligence, is presented as an AGI architecture in Goertzel et al. (2014).

Recently, claims have been made that deep learning is capable of solving AI by Google (DeepMind[4]). Likewise, Facebook AI Research (FAIR[5]) and other companies are actively working in the same direction. However, the question is where does this work stand with respect to cognitive architectures? Overall, the DeepMind research addresses a number of important issues in AI, such as natural language understanding, perceptual processing, general learning, and strategies for evaluating artificial intelligence. Although particular models already demonstrate cognitive abilities in limited domains, at this point they do not represent a unified model of intelligence.

Differently from DeepMind, Mikolov et al. (2015) from a Facebook research team explicitly discuss their work in a broader context of developing intelligent machines. Their main argument is that AI is too complex to be built all at once and instead its general characteristics should be defined first. Two such characteristics of intelligence are defined, namely, communication and learning, and a concrete roadmap are proposed for developing them incrementally.

Currently, there are no publications about developing such a system, but overall the research topics pursued by FAIR align with their proposal for AI and also the business interests of the company. Common topics include visual processing, especially segmentation and object detection, data mining, natural language processing, human-computer interaction and network security. Since the current deep learning techniques are mainly applied to solving practical problems and do not represent a unified framework we do not include them in this review. However, given their prevalence in other areas of AI, deep learning methods will likely play important role in the cognitive architectures of the future.

In view of the above discussion and to ensure both inclusiveness and consistency, cognitive architectures in this survey are selected based on the following criteria: self-evaluation as cognitive, robotic or agent architecture, existing implementation (not necessarily open-source), and mechanisms for perception, attention, action selection, memory and learning. Furthermore, we considered the architectures with at least several peer-reviewed papers and practical applications beyond simple illustrative examples. For the most recent architectures still under development, some of these conditions were relaxed.

An important point to keep in mind while reading this survey is that cognitive architectures should be distinguished from the models or agents that implement them. For instance, ACT-R, Soar, HTM and many other architectures serve as the basis for multiple software agents that demonstrate only a subset of capabilities declared in theory. On the other hand, some agents may implement extra features that are not available in the cognitive architecture. A good example is the perceptual system of Rosie (Kirk and Laird 2014), one of the robotic agents implemented in Soar, whereas Soar itself does not include a real perceptual system for physical sensors as part of the architecture. Unfortunately, in many cases the level of detail presented in the publications does not allow one to judge whether the particular capability is enabled by the architectural mechanisms (and is common among all models) or is custom-made for a particular application. Therefore, to avoid confusion, we do not make this distinction and list all capabilities demonstrated by the architecture.

---

[4] https://deepmind.com/.

[5] https://research.facebook.com/ai/.

Note also that some architectures underwent structural and conceptual changes throughout their development, notably the long-running projects such as Soar, ACT-R, CLARION, Disciple and few others. For example, Soar 9 and subsequent versions added some non-symbolic elements (Laird 2012b) and CAPS4 better accounts for the time course of cognition and individual differences compared to its predecessors 3CAPS and CAPS.[6] Sometimes these changes are well documented (e.g. CLARION,[7] Soar[8]), but more often they are not. In general we use the most recent variant of the architecture for analysis and assume that the features and abilities of the previous versions are retained in the new version unless there is contradicting evidence.

3 Taxonomies of cognitive architectures

Many papers published within the last decade address the problem of evaluation rather than categorization of cognitive architectures. As mentioned earlier, Newell's criteria (Anderson and Lebiere 2003; Newell 1980, 1992) and Sun's desiderata (Sun 2004) belong in this category. Furthermore, surveys of cognitive architectures define various capabilities, properties and evaluation criteria, which include recognition, decision making, perception, prediction, planning, acting, communication, learning, goal setting, adaptability, generality, autonomy, problem solving, real-time operation, meta-learning, etc. (Asselman et al. 2015; Langley et al. 2009; Thórisson and Helgasson 2012; Vernon et al. 2007).

While these criteria could be used for classification, many of them are too fine-grained to be applied to a generic architecture. Thus, a more general grouping of architectures can be based on the type of representation and information processing they implement. Three major paradigms are currently recognized: symbolic (also referred to as cognitivist), emergent (connectionist) and hybrid. Which of these representations, if any, correctly reflects the human cognitive processes, remains an open question and has been debated for the last 30 years (Kelley 2003; Sun and Bookman 1994).

Symbolic systems represent concepts using symbols that can be manipulated using a predefined instruction set. Such instructions can be implemented as if-then rules applied to the symbols representing the facts known about the world (e.g. ACT-R, Soar and other production rule architectures). Because it is a natural and intuitive representation of knowledge, symbolic manipulation remains very common. Although, by design, symbolic systems excel at planning and reasoning, they are less able to deal with the flexibility and robustness that are required for dealing with a changing environment and for perceptual processing.

The emergent approach resolves the adaptability and learning issues by building massively parallel models, analogous to neural networks, where information flow is represented by a propagation of signals from the input nodes. However, the resulting system also loses its transparency, since knowledge is no longer a set of symbolic entities and instead is distributed throughout the network. For these reasons, logical inference in a traditional sense becomes problematic (although not impossible) in emergent architectures.

Naturally, each paradigm has its strengths and weaknesses. For example, any symbolic architecture requires a lot of work to create an initial knowledge base, but once it is done the architecture is fully functional. On the other hand, emergent architectures are easier to

---

[6] http://www.ccbi.cmu.edu/4CAPS/.

[7] http://www.clarioncognitivearchitecture.com/release-notes.

[8] https://soar.eecs.umich.edu/articles/articles.

Fig. 3 A taxonomy of cognitive architectures based on the representation and processing. The order of the architectures within each group is alphabetical and does not correspond to the proportion of symbolic versus sub-symbolic elements (i.e. the spatial proximity of ACT-R and iCub to nodes representing symbolic and emergent architectures respectively does not imply that ACT-R is closer to symbolic paradigm and iCub is conceptually related to emergent architectures)

design, but must be trained in order to produce useful behavior. Furthermore, their existing knowledge may deteriorate with the subsequent learning of new behaviors.

As neither paradigm is capable of addressing all major aspects of cognition, hybrid architectures attempt to combine elements of both symbolic and emergent approaches.[9] Such systems are the most common in our selection of architectures (and, likely, overall). In general, there are no restrictions on how the hybridization is done and many possibilities have been explored. Multiple taxonomies of the types of hybridization have been proposed (Duch 2007; Hilario 1997; Sun 1996; Wermter 1997; Wermter and Sun 2000). Besides representation, one can classify the systems as single or multi-module, heterogeneous or homogeneous, take into account the granularity of hybridization (coarse-grained or fine-grained), the coupling between the symbolic and sub-symbolic components, and types of memory and learning. However, not all the hybrid architectures explicitly address what is referred to as symbolic and sub-symbolic elements and reasons for combining them. Only a few of them, namely ACT-R, CLARION, DUAL, CogPrime, CAPS, SiMA, GMU-BICA and Sigma, view this integration as essential and discuss it extensively. In general, we found that symbolic and sub-symbolic elements cannot be identified for all reviewed architectures due to lack of such fine-grained detail in many publications, thus we will focus on representation and processing.

Figure 3 shows the architectures grouped according to the new taxonomy. The top level of the hierarchy is represented by the symbolic, emergent and hybrid approaches. The definitions of these terms in the literature are vague, which often leads to the inconsistent assignment of architectures to either group. There seems to be no agreement even regarding the most known architectures, Soar and ACT-R. Although both combine symbolic and sub-symbolic

---

[9] Alternative representations have also been suggested. One example is Conceptual Spaces (?) which advocates the use of geometrical structures to overcome the limitations of symbolic representations and make neural networks more transparent (?). Some preliminary results with this framework have been demonstrated in ACT-R and CLARION (?).

elements, ACT-R explicitly self-identifies as hybrid, while Soar does not (Laird 2012b). The surveys are inconsistent as well, for instance, both Soar and ACT-R are called cognitivist in Vernon et al. (2007) and Goertzel et al. (2010b), while (Asselman et al. 2015) lists them as hybrid.

Further complicating our analysis is the fact that the authors of the architectures rarely mention the types of representations used in their systems. Specifically, out of the 84 architectures we surveyed, only 34 have a self-assigned label. Fewer still provide details on the particular symbolic/subsymbolic processes or elements.

It is universally agreed that symbols (labels, strings of characters, frames), production rules and non-probabilistic logical inference are symbolic and that distributed representations such as neural networks are sub-symbolic. For example, probabilistic action selection is considered as symbolic in CARACaS (Huntsberger and Stoica 2010), CHREST (Schiller and Gobet 2012) and CogPrime (Goertzel 2012), but is described as sub-symbolic in ACT-R (Lebiere et al. 2013), CELTS (Faghihi et al. 2011b), CoJACK (Ritter et al. 2012), Copycat/Metacat (Marshall 2006) and iCub (Sandini et al. 2007). Likewise, numeric data is treated as symbolic in CAPS (Just and Varma 2007), AIS (Hayes-Roth 1995) and EPIC (Kieras 2004), but is regarded as sub-symbolic in SASE (Weng 2002). Similarly, there are disagreements regarding the status of the activations (weights or probabilities) assigned to the symbols, reinforcement learning, image data, etc.[10]

To avoid inconsistent grouping, we did not rely on self-assigned labels or conflicting classification of elements found in the literature. We assume that explicit symbols are atoms of symbolic representation and can be combined to form meaningful expressions and used for inference or syntactical parsing. Sub-symbolic representations are generally associated with the metaphor of a neuron. A typical example of such representation is the neural network, where knowledge is encoded as a numerical pattern distributed among many simple neuron-like objects. Weights associated with units control processing and are acquired by learning. For our classification, we assume that anything that is not an explicit symbol and processing other than syntactic manipulation is sub-symbolic (e.g. numeric data, pixels, probabilities, spreading activations, reinforcement learning, etc.). Hybrid representation may combine any number of elements from both representations.

Given these definitions, we assigned labels to all the architectures and visualized them as shown in Fig. 3. We distinguish between two groups in the emergent category: *neuronal models*, which implement models of biological neurons, and *connectionist logic systems*, which are closer to artificial neural networks. Within the hybrid architectures we separate *symbolic sub-processing* as a type of hybridization where a symbolic architecture is combined with self-contained module performing sub-symbolic computation, e.g. for processing sensory data. Other types of functional hybrids also exist, for example, co-processing, meta-processing and chain-processing, however, these are harder to identify from the limited data available in the publications and thus are not included. The *fully integrated* category combines all other types of hybrids.

The architectures in the *symbolic sub-processing* group include at least one sub-symbolic module for sensory processing, while the rest of the knowledge and processing is symbolic.

---

[10] Activations are considered symbolic only in MusiCog (Maxwell 2014), but majority of the architectures list them as sub-symbolic [e.g. CAPS (Just and Varma 2007), ACT-R (Moon and Anderson 2013; Taatgen 2002), DUAL (Kokinov 1994)]. Reinforcement learning is considered symbolic in CARACaS (Huntsberger and Stoica 2010), CHREST (Schiller and Gobet 2012) and MusiCog (Maxwell 2014), but sub-symbolic in MAMID (Reisenzein et al. 2013) and REM (Murdock and Goel 2001). Image data is sub-symbolic in MAMID (Reisenzein et al. 2013) and SASE (Weng 2002) and non-symbolic (iconic) in Soar (Laird 2012b) and RCS (Albus and Barbera 2005).

For example, 3T, ATLANTIS, RCS, DIARC, CARACaS and CoSy are robotic architectures, where a symbolic planning module determines the behavior of the system and one or more modules process visual and audio sensory data using techniques like neural networks, optical flow calculation, etc. Similarly, ARDIS and STAR combine symbolic knowledge base and case-based symbolic reasoning with image processing algorithms.

The *fully integrated* architectures use a variety of approaches for combining different representations. ACT-R, Soar, CAPS, Copycat/Metacat, CHREST, CHARISMA, CELTS, CoJACK, CLARION, REM, NARS and Xapagy combine symbolic concepts and rules with sub-symbolic elements such as activation values, spreading activation, stochastic selection process, reinforcement learning, etc. On the other hand, DUAL consists of a large number of highly interconnected hybrid agents, each of which has a symbolic and sub-symbolic representation, implementing integration at a micro-level. In the case of SiMA hybridization is done on multiple levels: neural networks are used to build neurosymbols from sensors and actuators (*connectionist logic systems*) and the top layer is a symbol processing system (*symbolic subprocessing*). For the rest of the architectures, it is hard to clearly identify how hybridization was realized. For instance, many architectures are implemented as a set of interconnected competing and cooperating modules, where individual modules are not restricted to a particular representation (Kismet, LIDA, MACSi, ISAC, iCub, GMU-BICA, CORTEX, ASMO, CELTS, PolyScheme, FORR).

Another hybridization strategy is to combine two different stand-alone architectures with complementary features. Even though additional effort is required to build interfaces for communication between them, this approach takes advantage of the strengths of each architecture. A good overview of conceptual and technical challenges involved in creating an interface between cognitive and robotic architectures is given in Scheutz et al. (2013). The proposed framework is demonstrated with two examples: ACT-R/DIARC and ICARUS/DIARC integration. Other attempts found in the literature are CERA-CRANIUM/Pogamut (Arrabales et al. 2009c) and CERA-CRANIUM/Soar (Llargues Asensio et al. 2014) hybrids for playing video games and IMPRINT/ACT-R intended for human error modeling (Lebiere et al. 2002).

Besides hybrids built for specific purposes, few long-standing projects (based on the number of publications) also exist. A canonical example is SAL, which comprises ACT-R and Leabra architectures (Jilk et al. 2008). Here ACT-R is used to guide learning of Leabra models. Another instance is the robotic architecture ADAPT, where Soar is utilized for control and separate modules are responsible for modeling a 3D world from sensor information and for visual processing (Benjamin et al. 2013).

Compared to hybrids, emergent architectures form a more uniform group. As mentioned before, the main difference between *neuronal modeling* and *connectionist logic systems* is in their biological plausibility. All systems in the first group implement particular neuron models and aim to accurately reproduce the low-level brain processes. The architectures in the second group are based on artificial neural networks. Despite being heavily influenced by neuroscience and having the ability to closely model certain elements of human cognition, the biological plausibility of these models is not claimed by their authors.

In conclusion, as can be seen in Fig. 3, hybrid architectures are the most numerous and diverse group, showing the tendency to grow even more, thus confirming a prediction made almost a decade ago (Duch et al. 2008). Hybrid architectures form a continuum between emergent and symbolic systems depending on the proportions and roles played by symbolic and sub-symbolic components. Although quantitative analysis of this space is not feasible, it is possible to crudely subdivide it. For instance, some architectures such as CogPrime and Sigma are conceptually closer to emergent systems as they share many properties with the neural networks. On the other hand, REM, CHREST, and RALPH, as well as the architectures

implementing *symbolic sub-processing*, e.g. 3T and ATLANTIS, are very much within the cognitivist paradigm. These architectures are primarily symbolic but also utilize probabilistic reasoning and learning mechanisms.

4 Perception

Regardless of its design and purpose, an intelligent system cannot exist in isolation and requires input to produce behavior. Although historically all major cognitive architectures focused on higher-level reasoning, it is becoming evident that perception and action also play an important role in human cognition (Anderson et al. 2004).

Perception is a process that transforms raw input into the system's internal representation for carrying out cognitive tasks. Depending on the origin and properties of the incoming data, multiple sensory modalities are distinguished. The five most common ones are vision, hearing, smell, touch and taste. Other recognized human senses include proprioception, thermoception, nocioception, sense of time, etc. Naturally, cognitive architectures implement some of these, as well as other modalities that do not have a correlate among human senses such as symbolic input [e.g. input via keyboard or graphical user interface (GUI)] and various sensors such as LiDAR, laser, compass, etc.

Depending on its cognitive capabilities, an intelligent system can process various amounts and types of data as perceptual input. In this section we will investigate the diversity of the data inputs used in cognitive architectures, what information is extracted from these sources and how it is applied. The visualization in Fig. 4 addresses the first part of this question by mapping the cognitive architectures to the sensory modalities they implement: vision (V), hearing (A), touch (T), smell (S), proprioception (P) as well as the data input (D), other sensors (O) and multi-modal (M).[11] Note that data presented in the diagram is aggregated from many publications and most architectures do not demonstrate all listed sensory modalities in a single model.

Several observations can be made from the visualization in Fig. 4. For instance, vision is the most commonly implemented sensory modality, however, more than half of the architectures use simulation for visual input instead of the physical camera. Modalities such as touch and proprioception are mainly used in the physically embodied architectures. Few senses remain relatively unexplored, e.g. sense of smell is only featured in three architectures [GLAIR (Shapiro and Kandefer 2005), DAC (Mathews et al. 2009) and PRS (Taylor and Padgham 1996)]. As symbolic architectures by design have limited perceptual abilities, they tend to use direct data input as the only source of information (see left side of the diagram). On the other hand, hybrid and emergent architectures (located mainly in the right half of the diagram) implement a broader variety of sensory modalities both with simulated and physical sensors.

In any case, regardless of its source, incoming sensory data is usually not usable in the raw form (except, maybe, for symbolic input) and requires further processing. Below we discuss various approaches to the problem of efficient and adequate perceptual processing in cognitive architectures.

---

[11] The taste modality is omitted as it is featured in a single architecture, Brain-Based Devices (BBD), where it is simulated by measuring the conductivity of the object (Krichmar and Snook 2002).
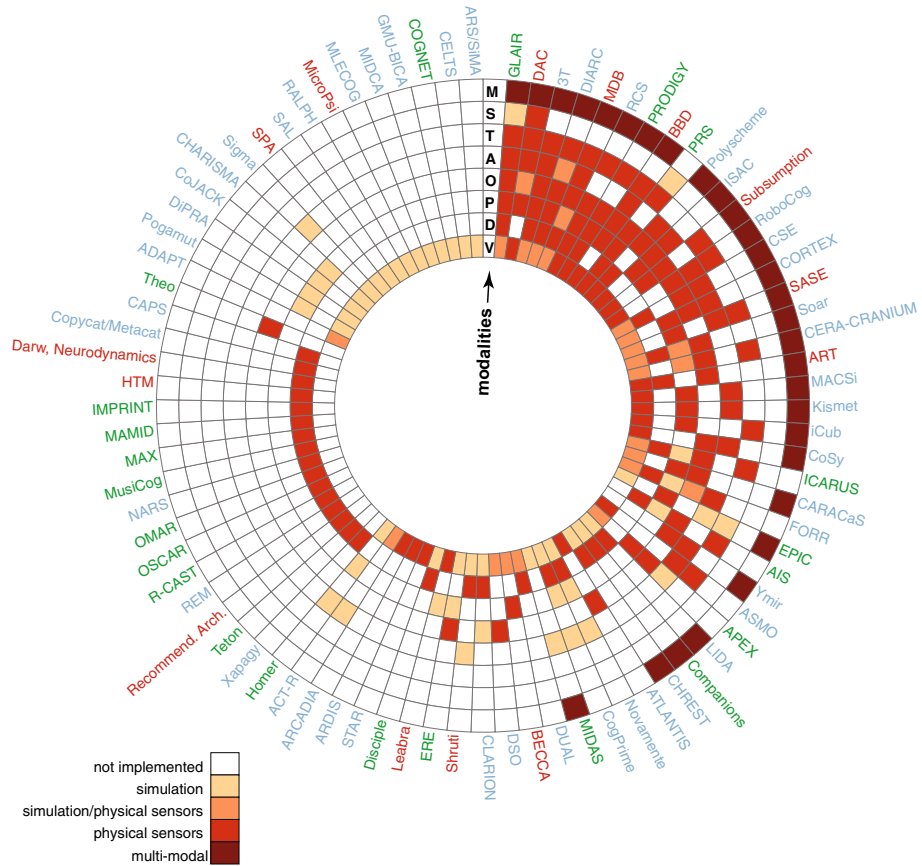
Fig. 4 A diagram showing sensory modalities of cognitive architectures. Radial segments correspond to cognitive architectures. Modalities are arranged in concentric circles and ordered based on how common they are (calculated as a number of architectures that implement this modality): vision (V), symbolic input (D), proprioception (P), other sensors (O), audition (A), touch (T), smell (S) and multi-modal (M). Architecture labels are rendered in different colors depending on the representation: green—symbolic, blue—hybrid, red—emergent. Architectures are ordered clockwise from 12 o'clock depending on how many different sensory modalities they implement and to what extent (scores used for ranking range from 0 when the modality is not implemented to 4 for multi-modal perception)

4.1 Vision

For a long time, vision was viewed as the dominating sensory modality based on the available experimental evidence (Posner et al. 1976). While recent works suggest a more balanced view of human sensory experience (Stokes and Biggs 2014), cognitive architectures research remains fairly vision-centric, as it is the most studied and the most represented sensory modality. Even though in robotics various non-visual sensors (e.g. sonars, ultrasonic distance sensors) and proprioceptive sensors (e.g. gyroscopes, compasses) are used for solving visual tasks such as navigation, obstacle avoidance and search, visual input still accounts for more than half of all implemented input modalities.

According to Marr (2010), visual processing is composed of three different stages: early, intermediate and late. Early vision is data-driven and involves parallel processing of the

visual scene and extraction of simple elements, such as color, luminance, shape, motion, etc. Intermediate vision groups elements into regions, which are then further processed in the late stage to recognize objects and assign meaning to them using available knowledge. Although not mentioned by Marr, visual attention mechanisms, emotion and reward systems also influence all stages of visual processing (Tsotsos 2011). Thus, perception and cognition are tightly interconnected throughout all levels of processing.

We base our analysis of visual processing in the cognitive architectures on image understanding stages described in Tsotsos (1992). These stages include: (1) detection and grouping of intensity-location-time values (results in edges, regions, flow vectors); (2) further grouping of edges, regions, etc. (produces surfaces, volumes, boundaries, depth information); (3) identification of objects and their motions; (4) building object-centered representations for entities; (5) assigning labels to objects based on the task; (6) inference of spatiotemporal relationships among entities.[12] Here only stage 1 represents early vision, and all subsequent stages require additional task or world knowledge. Already at stage 2, grouping of features may be facilitated by the viewpoint information and knowledge of the specific objects being viewed. Finally, the later stages involve spatial reasoning and operate on high-level representations abstracted from the results of early and intermediate processing.

Note that in current computer vision research many of these image understanding stages are implemented implicitly with deep learning methods. Given the success of deep learning in many applications it is surprising to see that very few cognitive architectures employ it. Some applications of deep learning to simple vision tasks can be found in CogPrime (Goertzel et al. 2013), LIDA (Madl et al. 2015), SPA (Stewart and Eliasmith 2013) and BECCA (Rohrer 2013).

Diagrams in Fig. 5 show stages of processing found in cognitive architectures with real and simulated vision. We assume that the real vision systems only receive pixel-level input with no additional information (e.g. camera parameters, locations and features of objects, etc.). The images should be produced by a physical camera, but the architecture does not need to be connected to a physical sensor (i.e. if the input is a dataset of images or a previously recorded video, it is still considered as real vision). The simulated vision systems generally omit early and intermediate vision and receive input in the form that is suitable for the later stages of visual processing (e.g. symbolic descriptions for shape and color, object labels, coordinates, etc.). Technically, any architecture that does not have native support for real vision or other perceptual modalities may be extended with an interface that connects it to sensors and pre-processes raw data into a more suitable format as demonstrated by Soar (Mohan et al. 2012) and ACT-R (Trafton and Harrison 2011).

The illustration in Fig. 5 shows what image interpretation stages are implemented, but it does not reflect the complexity and extent of such processing as there are no quantitative criteria for evaluation. In the remainder of this section we will provide brief descriptions of the visual processing in various architectures.
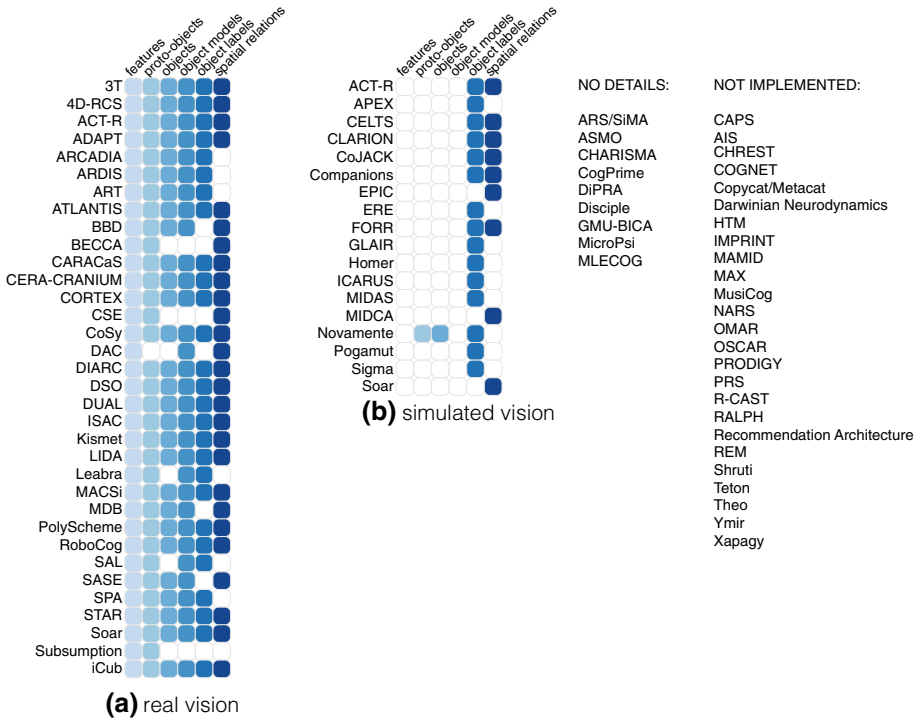
4.1.1 Vision using physical sensors

The majority of the architectures implementing all stages of visual processing are physically embodied and include robot control, biologically inspired and biomimetic architectures.

The architectures in the first group (3T, ATLANTIS and CARACaS) operate in realistic unstructured environments and approach vision as an engineering problem. Early vision

---

[12] The last stage forming consistent internal descriptions—from (Tsotsos 1992) is omitted. In cognitive architectures, such representations would be distributed among various reasoning and memory modules, making proper evaluation infeasible.

(a) real vision

(b) simulated vision

NO DETAILS:

ARS/SiMA
ASMO
CHARISMA
CogPrime
DiPRA
Disciple
GMU-BICA
MicroPsi
MLECOG

NOT IMPLEMENTED:

CAPS
AIS
CHREST
COGNET
Copycat/Metacat
Darwinian Neurodynamics
HTM
IMPRINT
MAMID
MAX
MusiCog
NARS
OMAR
OSCAR
PRODIGY
PRS
R-CAST
RALPH
Recommendation Architecture
REM
Shruti
Teton
Theo
Ymir
Xapagy

Fig. 5 A diagram showing the stages of visual processing implemented by the cognitive architectures. **a** Lists architectures utilizing physical sensors and **b** lists architectures with simulated vision in alphabetical order. Each row contains stages of processing ordered from early to late: (1) features, (2) proto-objects, (3) objects, (4) object models, (5) object labels and (6) spatial relations. Different shades of blue are used to indicate processes that belong to early, intermediate and late vision. In several architectures some type of visual process is implemented but the publications do not provide a sufficient amount of technical detail for our analysis ("No details" list). Architectures in the "Not implemented" list do not implement vision explicitly, although other sensors (e.g. sonars and bumpers) may be used for visual tasks such as recognition and navigation

(stage 1) usually involves edge detection and disparity estimation. These features are then grouped (stage 2) into blobs with similar features (color, depth, etc.), which are resolved into candidate objects with centroid coordinates (stage 3). Object models are often learned off-line using machine learning techniques (stage 4) and can be used to categorize candidate objects (stage 5). For example, in RCS architecture the scene is reconstructed from stereo images, features of the traversable ground are learned from LiDAR data and color histograms (Hong et al. 2002). Similarly, the ATLANTIS robot architecture constructs a height map from stereo images and uses it to identify obstacles (Miller and Slack 1991). The CARACaS architecture also computes a map for hazard avoidance from stereo images and range sensors. The locations and types of obstacles are placed on the map relative to the robot itself. Spatial reasoning (stage 6) is limited to determining the distance to the obstacles or target locations for navigation (Huntsberger et al. 2011).

Biologically inspired architectures also make use of computer vision algorithms and follow similar processing steps. For instance, neural networks are applied for object detection [RCS (Albus and Barbera 2006), DIARC (Scheutz et al. 2004), Kismet (Breazeal and Scassellati 2002)], SIFT features are utilized for object recognition [DIARC (Schermerhorn et al.

2006)], SURF features, AdaBoost learning and a mixture of Gaussians can be applied to hand detection and tracking [iCub (Ciliberto et al. 2011)], Kinect and LBP features in conjunction with SVM classifier can be used to find people and determine their age and gender [RoboCog and CORTEX (Martinez-Gomez et al. 2014; Romero-Garcés et al. 2015b)], etc.

In general, in the biologically inspired architectures vision is more intertwined with memory and control systems and stages of visual processing are explicitly related to the processes found in human visual system. One such example is saliency, which models the ability to prioritize visual stimuli based on their features or relevance to the task. As such, saliency is often used to find regions of interest in the scene [Kismet (Breazeal et al. 2001), ARCADIA (Bridewell and Bello 2015), DIARC (Scheutz et al. 2014), iCub (Leitner et al. 2013), STAR (Kotseruba 2016)]. Another variation on the saliency map, found in robotic architectures, is ego-sphere. It mimics functions of the hippocampus in the integration of sensory information with action, although not in a biologically plausible way (Peters et al. 2001a). Essentially, ego-sphere forms a virtual dome surrounding the robot, onto which salient objects and events are mapped. Implementations of this concept are included in RCS (Albus 1994), ISAC (Kawamura et al. 2008), iCub (Vernon et al. 2010) and MACSi (Ivaldi et al. 2014).

The architectures in the biomimetic subgroup pursue biologically plausible vision. One of the most elaborate examples is the Leabra vision system called LVis (O'Reilly et al. 2013) based on the anatomy of the ventral pathway of the human brain. It models the primary visual cortex (V1), extrastriate areas (V2, V4) and the inferotemporal (IT) cortex. Computation in these areas roughly corresponds to early and intermediate processing steps in the diagram in Fig. 5. LVis possesses many features of the human visual system, such as larger receptive fields of neurons in higher levels of the hierarchy, reciprocal connections between the layers and recurrent inhibitory dynamics that limit activity levels across layers (Wyatte et al. 2012). The visual systems of Darwin VIII (BBD) (Seth et al. 2004), SPA (Spaun) (Eliasmith and Stewart 2012) and ART (Grossberg 2007) are also modeled on the primate ventral visual pathway.

The SASE architecture, also in this group, does not replicate the human visual system as closely. Instead, it uses a hierarchical neural network with localized connections, where each neuron gets input from a restricted region in the previous layer. The sizes of receptive fields within one layer are the same and increase at higher levels (Zhang et al. 2002). This system was tested on a SAIL robot in an indoor navigation scenario (Weng and Zhang 2002). A similar approach to vision is implemented in MDB (Duro et al. 2010), BECCA (Rohrer et al. 2009) and DAC (Mathews et al. 2009).

Although these emergent biomimetic systems do not explicitly assign labels to objects, they are capable of forming representations of spatial relations between the objects in the scene and use these representations for visual tasks like navigation [BBD (Fleischer and Edelman 2009), BECCA (Rohrer et al. 2009), DAC (Mathews et al. 2009), MDB (Duro et al. 2010), SASE (Weng and Hwang 2007)].

4.1.2 Simulated vision

As evident from the diagram in Fig. 5, most of the simulations support only late stages of visual processing. The simplest kind of simulation is a 2D grid of cells populated by objects, such as NASA TileWorld (Philips and Bresina 1991) used by ERE (Drummond and Bresina 1990) and PRS (Kinny et al. 1992), Wumpus World for GLAIR agents (Shapiro and Kandefer 2005), 2D maze for FORR agent Ariadne (Epstein 1995) and tribal simulation designed for CLARION social agents (Sun and Fleischer 2012). Agents in grid-like environments often have a limited view of the surroundings, restricted to few cells in each direction. Blocks

World is another classic domain, where the general task is to build stacks of blocks of various shapes and colors [ACT-R (Kennedy and Trafton 2006), ICARUS (Langley and Allen 1993), MIDCA (Cox 2013)].

Overall, despite their varying complexity and purpose, simulations usually provide the same kinds of data about the environment: objects, their properties (color, shape, label, etc.), locations and properties of the agent itself, spatial relations between objects and environmental factors (e.g. weather and wind direction). As such, simulations mainly serve as a visualization tool and are not too far removed from the direct data input since little to none sensory processing is needed.

More advanced simulations represent the scene as polygons with color and 3D coordinates of corners, which have to be further processed to identify objects [Novamente (Heljakka et al. 2007)]. Otherwise, the visual realism of 3D simulations serves mostly aesthetic purposes, as the information is available directly in symbolic form [e.g. CoJACK (Evertsz et al. 2009), Pogamut (Bida et al. 2012)].

As mentioned earlier, the diagram in Fig. 5 does not reflect the differences in the complexity of the environments or capabilities of the individual architectures. However, there is a lot of variation in terms of scope and realism of the environments for the embodied cognitive architectures. For example, a planetary rover controlled by ATLANTIS performs cross-country navigation in an outdoor rocky terrain (Matthies 1992), the salesman robot Gualzru [CORTEX (Romero-Garcés et al. 2015b)] moves around a large room full of people and iCub [MACSi (Ivaldi et al. 2012)] recognizes and picks up various toys from a table. On the other hand, simple environments with no clutter or obstacles are also used in the cognitive architectures research [BECCA (Rohrer et al. 2009), MDB (Bellas et al. 2006)]. In addition, color-coding objects is a common way of simplifying visual processing. For instance, ADAPT tracks a red ball rolling on the table (Benjamin et al. 2013) and DAC orients itself towards targets marked with different colors (Maffei et al. 2015). Finally, recognition capabilities of most visual systems reviewed in this section are limited to a handful of different object categories, except Leabra, which is able to distinguish dozens of object categories (Wyatte et al. 2012).

Overall, the quality of visual processing has greatly improved with the spread of software toolkits such as OpenCV, Cloud Point Library or Kinect API. But not as much progress has been made within systems that try to model general purpose and biologically plausible visual systems. Currently, their applications are limited to controlled environments.

4.2 Audition

Audition is a fairly common modality in the cognitive architectures as sound or voice commands are typically used to guide an intelligent system or to communicate with it. Since the auditory modality is purely functional, many architectures resort to using available speech-to-text software rather than develop models of audition. Among the few architectures modeling auditory perception are ART, ACT-R, SPA and EPIC. For example, ARTWORD and ARTSTREAM were used to study phonemic integration (Grossberg and Myers 2015) and source segregation (cocktail party problem) (Grossberg et al. 2004) respectively. A model of music interpretation was developed with ACT-R (Chikhaoui et al. 2009).

Using dedicated software for speech processing and communication helps to achieve some degree of complexity and realism, e.g. having scripted interaction with people in a crowded room [CORTEX (Romero-Garcés et al. 2015b)] or a dialog about the objects in the scene in a subset of English [CoSy (Lison and Kruijff 2008)]. A more advanced application involves

using speech recognition for the task of ordering books from the public library by phone [FORR (Epstein et al. 2012)]. Other systems using off-the-shelf speech processing software include the PolyScheme (Trafton et al. 2005) and ISAC (Peters et al. 2001a).

In our sample of architectures, most effort is directed at natural language processing, i.e. linguistic and semantic information carried by the speech (further discussed in Sect. 11.2), and much less attention is paid to the emotional content (e.g. loudness, speech rate, and intonation). Some attempts in this direction are made in social robotics. For example, the social robot Kismet does not understand what is being said to it but can determine approval, prohibition or soothing based on the prosodic contours of the speech (Breazeal and Aryananda 2002). The Ymir architecture also has a prosody analyzer combined with a grammar-based speech recognizer that can understand a limited vocabulary of 100 words (Thorisson 1999). Even the sound itself can be used as a cue, for example, the BBD robots can orient themselves toward the source of a loud sound (Seth et al. 2004).

4.3 Symbolic input

The symbolic input category in Fig. 4 combines several input methods which do not fall under physical sensors or simulations. These include input in the form of text commands and data, and via a GUI. Text input is typical for the architectures performing planning and logical inference tasks [e.g. NARS (Wang 2013), OSCAR (Pollock 1993b), MAX (Kuokka 1989), Homer (Vere and Bickmore 1990)]. Text commands are usually written in terms of primitive predicates used in the architecture, so no additional parsing is required.

Although many architectures have tools for visualization of results and the intermediate stages of computation, interactive GUIs are less common. They are mainly used in human performance research to simplify input of the expert knowledge and to allow multiple runs of the software with different parameters [IMPRINT (Mitchell 2009), MAMID (Hudlicka et al. 2000), OMAR (Deutsch and Cramer 1998a), R-CAST (From et al. 2011)].

Data input can be provided as text or any other format (e.g. binary or floating point arrays) and is primarily used for the categorization and classification applications [e.g. HTM (Lavin and Ahmad 2015), CSE (Henderson and Joshi 2013), ART (Carpenter et al. 1991)].

4.4 Other modalities

Other sensory modalities, such as touch, smell and proprioception, are trivially implemented with a variety of sensors (physical and simulated). Many robotic platforms are equipped with touch-sensitive bumpers for an emergency stop to avoid hitting an obstacle [e.g. RCS (Albus et al. 2006), AIS (Hayes-Roth et al. 1993), CERA-CRANIUM (Arrabales et al. 2009a), DIARC (Trivedi et al. 2011)]. Similarly, proprioception, i.e. sensing of the relative positions of the body parts and effort being exerted, can be either simulated or provided by force sensors [3T (Firby et al. 1995), iCub (Pattacini et al. 2010), MDB (Bellas et al. 2010), Subsumption (Brooks 1990)], joint feedback [SASE (Huang and Weng 2007), RCS (Bostelman et al. 2006)], accelerometers [Kismet (Breazeal and Brooks 2004), 3T (Wong et al. 1995)], etc.

4.5 Multi-modal perception

In the previous sections we have considered all sensory modalities in isolation. In reality, however, human brain receives a constant stream of information from different senses and

integrates it into a coherent world representation. This is also true for cognitive architectures, as nearly half of them have more than two different sensory modalities (see Fig. 4). As mentioned in the beginning of this section, not all of these modalities may be present in a single model and most architectures utilize only two different modalities simultaneously, e.g. vision and audition, vision and symbolic input or vision and range sensors. With few exceptions, these architectures are embodied and essentially perform what is known as feature integration in cognitive science (Zmigrod and Hommel 2013) or sensor data fusion in robotics (Khaleghi et al. 2013). Apparently, it is also possible to use different sensors without explicitly combining their output, as demonstrated by the Subsumption architecture (Flynn et al. 1989), but this approach is rather uncommon.

Multiple sensory modalities improve the robustness of perception through complementarity and redundancy but in practice using many different sensors introduces a number of challenges, such as incomplete/spurious/conflicting data, data with different properties (e.g. mismatched dimensionality or value ranges), the need for data alignment and association, etc. These practical issues are well researched by the robotics community, however, no universal solutions have yet been proposed. Eventually, every solution must be custom-made for a particular application, a common approach taken by the majority of cognitive architectures. There is, unfortunately, little technical information in the literature to determine the exact techniques used and connect them to the established taxonomies [e.g. from a recent survey (Khaleghi et al. 2013)].

Overall, any particular implementation of sensory integration depends on the representations used for reasoning and the task. In typical robotic architectures with symbolic reasoning, data from various sensors is processed independently and mapped onto a 3D map centered on the agent that can be used for navigation [CARACaS (Elkins et al. 2010), CoSy (Christensen and Kruijff 2010)]. In social robotics applications, as we already mentioned, the world representation can take the form of an ego-sphere around the agent that contains ego-centric coordinates and properties of the visually detected objects, which are associated to the location of the sound determined via triangulation [ISAC (Peters et al. 2001b), MACSi (Anzalone et al. 2012)]. In RCS, a model with a hierarchical structure, there is a sensory processing module at every level of the hierarchy with a corresponding world representation (e.g. pixel maps, 3D models, state tables, etc.) (Schlenoff et al. 2005).

Some architectures do not perform data association and alignment explicitly. Instead, sensor data and feature extraction (e.g. coordinates of objects from the camera and distances to the obstacles from laser) are done independently and concurrently. The extracted information is then directly added to the working memory, the blackboard or an equivalent structure (see Sect. 7.2 for more details). Any ambiguities and inconsistencies are resolved by higher-order reasoning processes. This is a common approach in distributed architectures, where independent modules concurrently work towards achieving a common goal [e.g. CERA-CRANIUM (Arrabales et al. 2009a), PolyScheme (Cassimatis et al. 2004), RoboCog (Bustos et al. 2013), Ymir (Thorisson 1997) and LIDA (Madl and Franklin 2015)].

In many biologically inspired architectures the association between the readings of different sensors (and actions) is learned. For example, DAC uses Hebbian learning to establish data alignment for mapping neural representations of different sensory modalities to a common reference frame, mimicking the function of the superior colliculus of the brain (Mathews et al. 2012). ART integrates visual and ultrasonic sensory information via neural fusion (ARTMAP network) for mobile robot navigation (Martens et al. 1998). Likewise, MDB uses neural networks to learn the world model from sensor inputs and a genetic algorithm to adjust the parameters of the networks (Bellas et al. 2010).

All approaches mentioned so far bear some resemblance to human sensory integration as they use the spatial and temporal proximity and/or learning to combine and disambiguate multi-modal data. But overall, only few architectures aim for biological fidelity at the perceptual level. The only elaborate model of biologically-plausible sensory integration is demonstrated using a brain-based device (BBD) architecture. The embodied neural model called Darwin XI is constructed to investigate the integration of multi-sensory information (from touch sensors, laser, camera and magnetic compass) and the formation of place activity in the hippocampus during maze navigation (Fleischer and Krichmar 2007). The neural network of Darwin XI consists of approximately 80,000 neurons with 1.2 million synapses and simulates 50 neural areas. In a lesion study, the robustness of the system was demonstrated by removing one or several sensory inputs and remapping of the sensory neuronal units.

For the most part, cognitive architectures ignore cross-modal interaction. Architectures, including biologically- and cognitively-oriented ones, typically pursue a modular approach when dealing with different sensory modalities. At the same time, many psychological and neuroimaging experiments conducted in the past decades suggest that sensory modalities mutually affect one another. For example, vision alters auditory processing [e.g. McGurk effect (McGurk and MacDonald 1976)] and vice versa (Shimojo and Shams 2001). While some biomimetic architectures, such as BBD mentioned above, may be capable of representing cross-modal effects, to the best of our knowledge, this problem has not been investigated yet in the cognitive architecture domain.

5 Attention

Perceptual attention plays an important role in human cognition, as it mediates the selection of relevant and filters out irrelevant information from the incoming sensory data. However, it would be wrong to think of attention as a monolithic structure that decides what to process next. Rather, the opposite may be true. There is ample evidence in favor of attention as a set of mechanisms affecting both perceptual and cognitive processes (Tsotsos and Kruijne 2014). Currently, visual attention remains the most studied form of attention, as there are no comprehensive attentional frameworks for other sensory modalities. Since only a few architectures have rudimentary mechanisms for modulating auditory data [OMAR (Deutsch et al. 1997), iCub (Ruesch et al. 2008), EPIC (Kieras et al. 2016) and MIDAS (Gore et al. 2009)], this section will be dedicated to visual attention.

For the following analysis, we use the taxonomy of attentional mechanisms proposed by Tsotsos (2011). Elements of attention are grouped into three classes of information reduction mechanisms: selection (choose one from many), restriction (choose some from many) and suppression (suppress some from many). Selection mechanisms include *gaze and viewpoint selection*, *world model* (selection of objects/events to focus on) and *time/region/features/objects/events of interest*. Restriction mechanisms serve to prune the search space by *priming*[13] (preparing the visual system for input based on task demands), *endogenous motivations* (domain knowledge), *exogenous cues* (external stimuli), *exogenous task* (restrict attention to objects relevant for the task), and *visual field* (limited field of view). Suppression mechanisms consist of *feature/spatial surround inhibition* (temporary suppression of the features around the object while attending), *task-irrelevant stimuli suppression*, *negative priming*, and *location/object inhibition of return* (a bias against returning attention

---

[13] Here we consider only priming in vision, other types of priming and their importance for learning are discussed in Sect. 8.6.
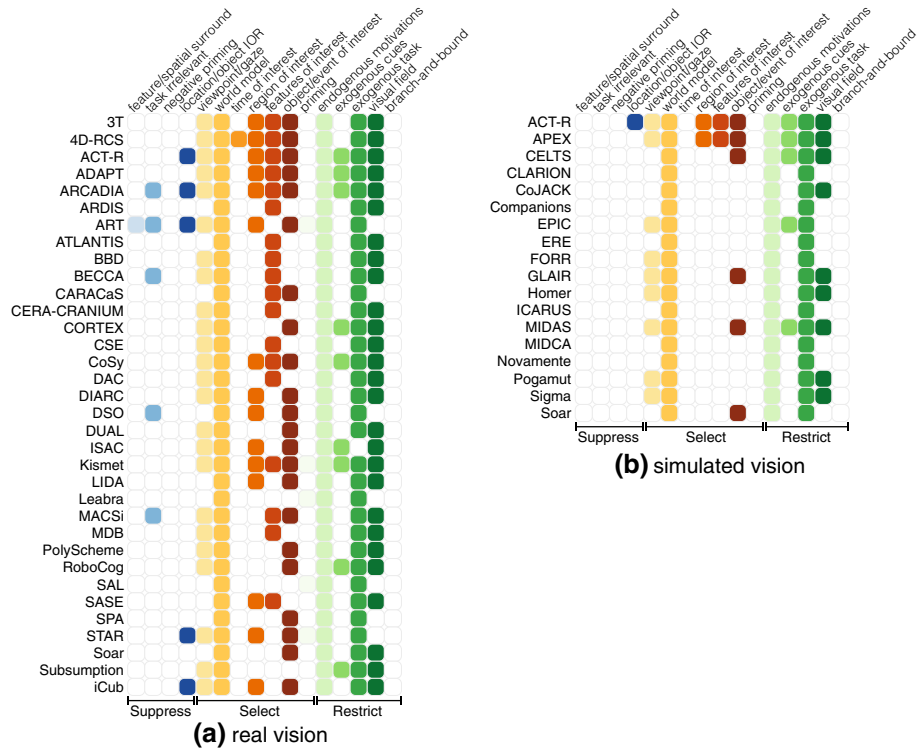
Fig. 6 A diagram showing visual attention mechanisms implemented in cognitive architectures for suppressing, selecting and restricting visual information. Architectures are ordered alphabetically. As in Sect. 4.1, architectures in the left and right columns implement real and simulated vision respectively. Architectures not implementing vision or missing technical details are not shown. Suppression, selection and restriction mechanisms in the diagram are shown in shades of blue, yellow-red and green respectively

to previous attended location or stimuli). Finally, *branch-and-bound* mechanisms combine elements of suppression, selection and restriction. For more detailed explanations of the attentional mechanisms and review of the relevant psychological and neuroscience literature refer to Tsotsos (2011).

The diagram in Fig. 6 shows a summary of the implemented visual attention elements in the cognitive architectures. Here we only consider the architectures with implemented real or simulated vision and omit those with insufficient technical details (as in the previous section). It is evident that most of the implemented mechanisms of attention belong to the selection and restriction group.

Only a handful of architectures implement suppression mechanisms. For instance, *suppression of task-irrelevant* visual information is explicitly done in only four architectures: in BECCA irrelevant features are suppressed through the WTA mechanisms (Rohrer 2011c), in DSO background features are suppressed to speed up processing (Xiao et al. 2015), in MACSi depth is used to ignore regions that are not reachable by the robot (Lyubova et al. 2013) and in ARCADIA non-central regions of the visual field are inhibited at each cycle since the cue always appears in the center (Bridewell and Bello 2015). Another suppression mechanism is *inhibition of return* (IOR), which prevents the visual system from repeatedly attending to the same salient stimuli. In ACT-R activation values and distance are used to

ignore objects for consecutive WHERE requests (Nyamsuren and Taatgen 2013). In ART, iCub and STAR an additional map is used to keep records of the attended locations. The temporal nature of inhibition can be modeled with a time decay function, as it is done in iCub (Ruesch et al. 2008). ARCADIA mentions a covert inhibition mechanism, however, specific implementation details are not provided (Bridewell and Bello 2015).

Selection mechanisms are much more common in the cognitive architectures. For example, a *world model* by default is a part of any visual system. Likewise, *viewpoint/gaze selection* is a necessary component of active vision systems. Specifically, gaze control allows to focus on a region in the environment and viewpoint selection allows to get more information about the region/object by moving away/towards it or by viewing it from different angles. As a result, physically embodied architectures automatically support *viewpoint selection* because a camera installed on a robot can be moved around the environment. Eye movements are usually simulated, but some humanoid robot implement them physically (iCub, Kismet).

Other mechanisms in this group include *selection of various features/objects of interest*. The selection of visual data to attend can be data-driven (bottom-up) or task-driven (top-down). The bottom-up attentional mechanisms identify salient regions whose visual features are distinct from the surrounding image features, usually along a combination of dimensions, such as color channels, edges, motion, etc. Some architectures resort to the classical visual saliency algorithms, such as Guided Search (Wolfe 1994) used in ACT-R (Nyamsuren and Taatgen 2013) and Kismet (Breazeal and Scassellati 1999), the Itti-Koch-Niebur model (Itti et al. 1998) used by ARCADIA (Bridewell and Bello 2015), iCub (Ruesch et al. 2008) and DAC (Mathews et al. 2012) or AIM (Bruce and Tsotsos 2007) used in STAR (Kotseruba 2016). Other approaches include filtering [DSO (Xiao et al. 2015)], finding unusual motion patterns [MACSi (Ivaldi et al. 2014)] or discrepancies between observed and expected data [RCS (Albus et al. 2002)].

Top-down selection can be applied to further limit the sensory data provided by the bottom-up processing. A classic example is visual search, where knowing desired features of the object (e.g., the color red) can narrow down the options provided by the data-driven Fig.-ground segmentation. Many architectures resort to this mechanism to improve search efficiency [ACT-R (Salvucci 2000), APEX (Freed 1998), ARCADIA (Bello et al. 2016), CERA-CRANIUM (Arrabales et al. 2009a), CHARISMA (Conforth and Meng 2011a), DAC (Mathews et al. 2012)]. Another option is to use a hard-coded or learned heuristics. For example, CHREST looks at typical positions on a chess board (Lane et al. 2009) and MIDAS replicates common eye scan patterns of pilots (Gore et al. 2009). The limitation of the current top-down approaches is that they can direct vision for only a limited set of predefined visual tasks, however, ongoing research in STAR attempts to address this problem (Tsotsos 2011; Tsotsos and Kruijne 2014).

Restriction mechanisms allow to further reduce the complexity of vision by limiting the search space to certain features (*priming*), events (*exogenous cues*), space (*visual field*), objects (*exogenous task*) and knowledge (*endogenous motivations*). *Exogenous task* and *endogenous motivations* are implemented in all architectures by default as domain knowledge and task instructions. In addition to that, attention can be modulated by internal signals such as motivation, emotions, which will be discussed in the next section on action selection.

A limited *visual field* is a feature of any physically embodied system since most cameras do not have a 360-degree field of view. This feature is also present in some simulated vision systems. An ability to react to the sudden events (*exogenous cues*), such as fast movement or bright color, is useful in social robotics applications [e.g. ISAC (Kawamura et al. 2004), Kismet (Breazeal et al. 2000)]. On the other hand, *priming* allows to bias the visual system towards particular types of stimuli via task instruction, e.g. by assigning more weight to

skin-colored features during the saliency computation to improve human detection [Kismet (Breazeal et al. 2001)] or by utilizing the knowledge of the probable location of some objects to spatially bias detection [STAR (Kotseruba 2016)]. Neural mechanisms of top-down priming are studied in ART models (Carpenter and Grossberg 1987).

Conclusion Overall, visual attention is largely overlooked in cognitive architectures research with the exception of the biologically plausible visual models (e.g. ART) and the architectures that specifically focus on vision research (ARCADIA, STAR).[14] This is surprising because strong theoretical arguments as to its importance in dealing with the computational complexity of visual processing have been known for decades (Tsotsos 1990). Many of the attention mechanisms found in the cognitive architectures are side-effects of other design decisions. For example, task constraints, world model and domain knowledge are necessary for the functioning of other aspects of intelligent system and are implemented by default. Limited visual field and viewpoint changes often result from physical embodiment. Otherwise, region of interest selection and visual reaction to exogenous cues are the two most common mechanisms explicitly included for optimizing visual processing.

In the cognitive and psychological literature, attention is also used as a broad term for the allocation of limited resources (Rosenbloom et al. 2015b). For instance, in the Global Workspace Theory (GWT) (Baars 2005) attentional mechanisms are central to perception, cognition and action. According to GWT, the nervous system is organized as multiple specialized processes running in parallel. Coalitions of these processes compete for attention in the global workspace and the contents of the winning coalition are broadcast to all other processes. For instance, the LIDA architecture is an implementation of GWT[15] (Franklin et al. 2012). Other architectures influenced by the GWT include ARCADIA (Bridewell and Bello 2015) and CERA-CRANIUM (Arrabales et al. 2009b).

Along the same lines, cognition is simulated as a set of autonomous independent processes in ASMO [inspired by Society of Mind (Minsky 1986)]. Here an attention value for each process is either assigned directly or learned from experience. Attention values vary dynamically and affect action selection and resource allocation (Novianto et al. 2010). A similar idea is implemented in COGNET (Zachary et al. 1993), DUAL (Kiryazov et al. 2007) and Copycat/Metacat (Mitchell and Hofstadter 1990) in which multiple processes also compete for attention. Other computational mechanisms, such as queue [PolyScheme (Kurup et al. 2011)], Hopfield nets [CogPrime (Ikle and Goertzel 2011)] and modulators [MicroPsi (Bach 2015)] have been used to implement a focus of attention. In MLECOG, in addition to saccades that operate on visual data, mental saccades allow to switch between the most activated memory neurons (Starzyk and Graham 2015).

6 Action selection

Informally speaking, action selection determines at any point in time "what to do next". At the highest level, action selection can be split into the "what" part involving the decision making and the "how" part related to motor control (Ozturk 2009). However, this distinction

---

[14] Note that biologically plausible models of vision, such as Selective Tuning (Tsotsos et al. 1995), extensions to HMAX (Riesenhuber 2005; Walther et al. 2002; Walther and Koch 2007) and others (Frintrop et al. 2010), support many of the listed attention mechanisms. However, embedding these models in the cognitive architecture is non-trivial. For instance, STAR represents ongoing work on integrating general purpose vision, as represented by the Selective Tuning model, with other higher-order cognitive processes.

[15] In GWT focus of attention is associated with consciousness, however, this topic is beyond the scope of this survey.

is not always explicitly made in the literature, where action selection may refer to goal, task or motor commands. For example, in the MIDAS architecture action selection involves both the selection of the next goal to pursue as well as actions that implement it (Tyler et al. 1998). Similarly, in MIDCA the next action is normally selected from a planned sequence if one exists. In addition to this, a different mechanism is responsible for the adoption of the new goals based on dynamically determined priorities (Paisner et al. 2013). In COGNET and DIARC, selection of a task/goal triggers an execution of the associated procedural knowledge (Zachary et al. 2016; Brick et al. 2007). In DSO, the selector module chooses between available actions or decisions to reach current goals and sub-goals (Ng et al. 2012). Since the treatment of action selection in various cognitive architectures is inconsistent, in the following discussion the term "action selection mechanisms" may apply both to decision-making and motor actions.

The diagram in Fig. 7 illustrates all implemented action selection mechanisms organized by the type of the corresponding architecture (symbolic, hybrid and emergent). We distinguish between two major approaches to action selection: planning and dynamic action selection. Planning refers to the traditional AI algorithms for determining a sequence of steps to reach a certain goal or to solve a problem in advance of their execution. In dynamic action selection, one best action is chosen among the alternatives based on the knowledge available at the time. For this category, we consider the type of selection (winner-take-all (WTA), probabilistic, predefined) and criteria for selection (relevance, utility, emotion). The default choice is always the best action based on the defined criteria (e.g. action with the highest activation value). Reactive actions are executed bypassing action selection. Finally, learning can also affect action selection but will be discussed in Sect. 8. Note that these action selection mechanisms are not mutually exclusive and most of the architectures have more than one. Even though few architectures implement the same set of action selection mechanisms (as can be seen in the Fig. 7), the whole space of valid combinations is likely much larger. Below we discuss approaches to action selection and what cognitive processes they correspond to in more detail.

6.1 Planning versus reactive actions

Predictably, planning is more prevalent in the symbolic architectures, but can also be found in some hybrid and even emergent [MicroPsi (Bach 2015)] architectures. Task decomposition, where the goal is recursively decomposed into subgoals, is a common form of planning [e.g. Soar (Derbinsky and Laird 2012a), Teton (VanLehn and Ball 1989), PRODIGY (Fink and Blythe 2005)]. Other types of planning are also used: temporal [Homer (Vere 1991)], continual [CoSy (Hawes et al. 2010)], hierarchical task network [PRS (D'Inverno et al. 2004)], generative [REM (Murdock and Goel 2008)], search-based [Theo (Mitchell 1990)], hill-climbing [MicroPsi (Bach 2015)], etc.

Very few systems in our selection rely on classical planning alone, namely OSCAR, used for logical inference, and IMPRINT, which employs task decomposition for modeling human performance. Otherwise, planning is usually augmented with dynamic action selection mechanisms to improve adaptability to the changing environment.

On the other end of the spectrum are reactive actions. These are executed immediately, suspending any ongoing activity and bypassing reasoning, similar to reflexes in humans prompting automatic responses to stimuli. As demonstrated by the Subsumption architecture, combining multiple reactive actions can give rise to complex behaviors (Brooks 1990). However, purely reactive systems are rare and reactive actions are used only under certain conditions, e.g. to protect the robot from the collision [ATLANTIS (Gat 1992), 3T (Bonasso
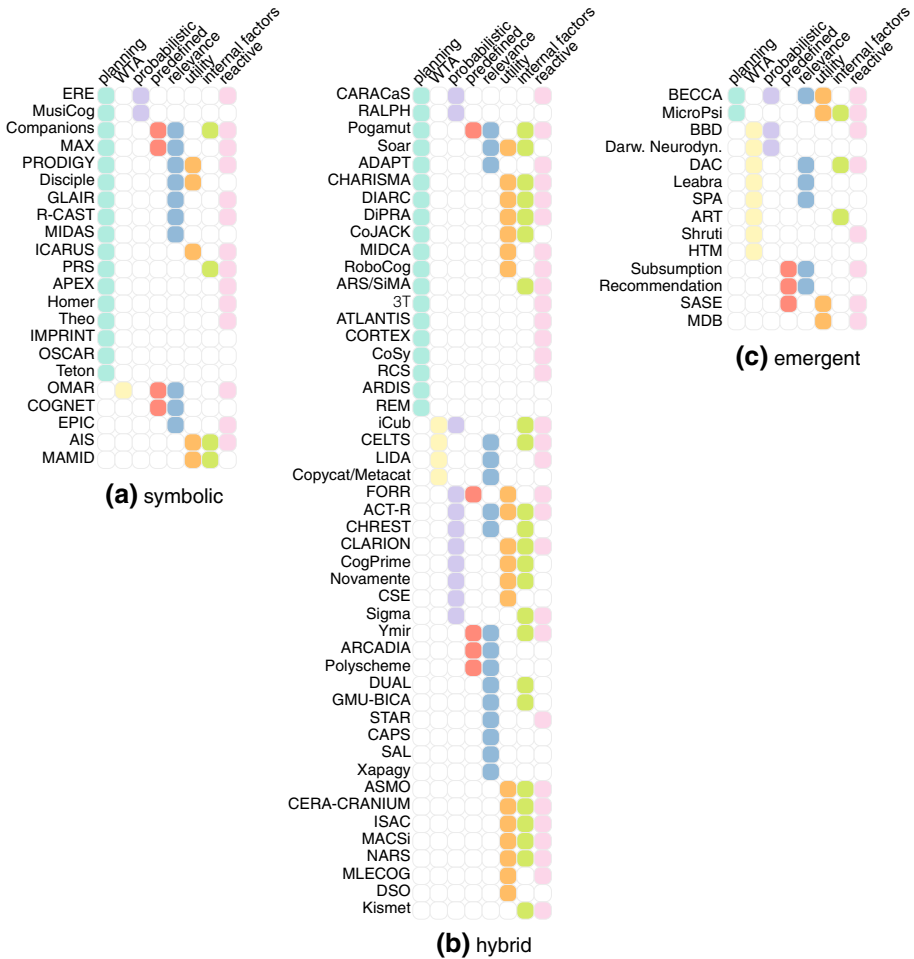
Fig. 7 A diagram of mechanisms involved in action selection. The visualization is organized in three columns grouping symbolic, hybrid and emergent architectures. Note that in this diagram (as well as in the diagrams for the Sects. 7, 8) the sorting order emphasizes clusters of architectures with similar action selection mechanisms (or memory and learning approaches respectively). Since the discussion follows the order in which different mechanisms are shown in the diagram, it can help to identify groups of architectures focusing on a particular mechanism which correspond to clusters within each column. A version of the diagram with an alphabetical ordering of the architectures is also available on the project website

and Kortenkamp 1996)] or to automatically respond to unexpected stimuli, such as fast moving objects or loud sounds [ISAC (Kawamura et al. 2008), Kismet (Breazeal and Fitzpatrick 2000), iCub (Ruesch et al. 2008)].

6.2 Dynamic action selection

Dynamic action selection offers the most flexibility and can be used to model many phenomena typical for human and animal behavior. Below we desribe in more detail typical dynamic action selection mechanisms and action selection criteria.

Winner-take-all (WTA) is a selection process in neuronal networks where the strongest set of inputs is amplified, while the output from others is inhibited. It is believed to be a part of cortical processes and is often used in the computational models of the brain to make a selection from a set of decisions depending on the input. WTA and its variants are common in the emergent architectures such as HTM (Byrne 2015), ART (Carpenter and Grossberg 2017), SPA (Eliasmith and Kolbeck 2015), Leabra (O'Reilly et al. 2012), DAC (Verschure and Althaus 2003), Darwinian Neurodynamics (Fedor et al. 2017) and Shruti (Wendelken and Shastri 2005). Similar mechanisms are also used to find the most appropriate action in the architectures where behavior emerges as a result of cooperation and competition of multiple processes running in parallel [e.g. Copycat/Metacat (Marshall 2002a), CELTS (Faghihi et al. 2013), LIDA (Franklin et al. 2016)].

A *predefined order* of action selection may serve different purposes. For example, in the Subsumption architecture, robot behavior is represented by a hierarchy of sub-behaviors, where higher-level behaviors override (subsume) the output of lower-level behaviors (Brooks 1990). In FORR, decision-making component considers options provided by other modules (advisors) in the order of increasing expertise to achieve robust and human-like learning (Epstein et al. 2002). In Ymir priority is given to processes within the reactive layer first, followed by the content layer and then by the process control layer. Here the purpose is to provide a smooth real-time behavior generation. Each layer has a different upper bound on the perception-action time, thus reactive modules provide automatic feedback to the user (changing facial expressions, automatic utterances) while deliberative modules generate more complex behaviors (Thorisson 1998).

The remaining action selection mechanisms include *finite-state machines* (FSM), which are frequently used to represent sequences of motor actions [ATLANTIS (Gat 1992), CARA-CaS (Huntsberger et al. 2011), Subsumption (Brooks 1990)] and even to encode the entire behavior of the system [ARDIS (Martin et al. 2011), STAR (Kotseruba 2016)]. *Probabilistic action selection* is also common [iCub (Vernon et al. 2010), CSE (Henderson and Joshi 2013), CogPrime (Goertzel et al. 2010a), Sigma (Ustun et al. 2015), Darwinian Neurodynamics (Fedor et al. 2017), ERE (Kedar and McKusick 1992), Novamente (Goertzel et al. 2008a)].

6.2.1 Action selection criteria

There are several criteria that can be taken into account when selecting the next action: relevance, utility and internal factors (which include motivations, affective states, emotions, moods, drives, etc.).

Relevance reflects how well the action corresponds to the current situation. This mainly applies to systems with symbolic reasoning and involves checking pre- and/or post-conditions of the rule before applying it [MAX (Kuokka 1991), Disciple (Tecuci et al. 2010), EPIC (Kieras 2012), GLAIR (Shapiro and Bona 2010), PRODIGY (Veloso et al. 1998), MIDAS (Tyler et al. 1998), R-CAST (Fan et al. 2010a), Disciple (Tecuci et al. 2010), Companions (Hinrichs and Forbus 2007), Ymir (Thorisson 1998), Pogamut (Brom et al. 2007), Soar (Laird 2012b), ACT-R (Lebiere et al. 2013)].

Utility of the action is a measure of its expected contribution to achieving the current goal [CERA-CRANIUM (Arrabales et al. 2009b), CHARISMA (Conforth and Meng 2011a), DIARC (Brick et al. 2007), MACSi (Ivaldi et al. 2014), MAMID (Hudlicka 2006), NARS (Wang and Hammer 2015a), Novamente (Goertzel et al. 2008a)]. Some architectures perform a "dry run" of candidate actions and observe their effect to determine their utility [MLECOG (Starzyk and Graham 2015), RoboCog (Bandera and Bustos 2013)]. Utility can also take

into account the performance of the action in the past and improve the behavior in the future via reinforcement learning [ASMO (Novianto 2014), BECCA (Rohrer 2011a), CLARION (Sun and Zhang 2003), PRODIGY (Carbonell et al. 1992), CSE (Henderson and Joshi 2013), CoJACK (Evertsz et al. 2009), DiPRA (Pezzulo 2009), Disciple (Tecuci et al. 2010), DSO (Ng et al. 2012), FORR (Epstein et al. 2002), ICARUS (Shapiro et al. 2001), ISAC (Kawamura et al. 2008), MDB (Bellas et al. 2005), MicroPsi (Bach 2011), Soar (Laird 2012b)]. Various machine learning techniques can be utilized to associate goals with successful behaviors in the past [MIDCA (Paisner et al. 2013), SASE (Weng and Hwang 2006), CogPrime (Goertzel et al. 2010a)].

Finally, *internal factors* do not determine the next behavior directly, but rather bias the selection. For simplicity, we will consider short-term, long-term and lifelong factors loosely corresponding to emotions, drives and personality traits in humans. Given the impact of these factors on human decision-making and other cognitive abilities it is important to model emotion and affect in cognitive architectures, especially in the areas of human-computer interaction, social robotics and virtual agents. Not surprisingly, this is where most of the effort has been spent so far (see Sect. 11.2.9 on applications of intelligent agents with realistic personalities and emotions).

Artificial emotions in cognitive architectures are typically modeled as transient states (associated with anger, fear, joy, etc.) that influence cognitive abilities (Hudlicka 2004; Goertzel et al. 2008b; Wilson et al. 2013; Bach 2015). For instance, in CoJACK morale and fear emotions can modify plan selection. As a result, plans that confront a threat have higher utility when morale is high, but lower utility when fear is high (Evertsz et al. 2009). Other examples include models of stress affecting decision-making [MAMID (Hudlicka and Matthews 2009)], emotions of joy/sadness affecting blackjack strategy [CHREST (Schiller and Gobet 2012)], analogical reasoning in the state of anxiety [DUAL (Feldman and Kokinov 2009)], effect of arousal on memory [ACT-R (Cochran et al. 2006)], positive and negative affect depending on goal satisfaction [DIARC (Scheutz and Schermerhorn 2009)] and emotional learning in HCI scenario [CELTS (Faghihi et al. 2013)]. The largest attempt at implementing a theory of appraisal led to the creation of Soar-Emote, a model based on the ideas of Damasio and Gratch & Marsella (Marinier and Laird 2004).

Drives are another source of internal motivations. Broadly speaking, they represent basic physiological needs such as food and safety, but can also include high-level or social drives. Typically, any drive has strength that diminishes when the drive is satisfied. As the agent works towards its goals, several drives may affect its behavior simultaneously. For example, in ASMO three relatively simple drives, liking color red, praise and happiness of the user and the robot, bias action selection by modifying the weights of the corresponding modules (Novianto 2014). In CHARISMA preservation drives (avoiding harm and starvation), combined with curiosity and desire for self-improvement, guide behavior generation (Conforth and Meng 2011b). In MACSi, curiosity drives exploration towards areas where the agent learns the fastest (Nguyen et al. 2013). Likewise, in CERA-CRANIUM curiosity, fear and anger affect exploration of the environment by the mobile robot (Arrabales Moreno and Sanchis de Miguel 2006). The behavior of social robot Kismet is organized around satiating three homeostatic drives: to engage people, to engage toys, and to rest. These drives and external events, contribute to the affective state or mood of the robot and its expression as anger, disgust, fear, joy, sorrow and surprise via facial gestures, stance and changes in the tone of its voice. Although its motivational mechanisms are hard-coded, Kismet demonstrates one of the widest repertoire of emotional responses of all the cognitive architectures we reviewed (Breazeal 2003a).

Unlike emotions, which have a transient nature, *personality traits* are unique long-term patterns of behavior that manifest themselves as consistent preferences in terms of internal motivations, emotions, decision-making, etc. Most of the identified personality traits can be reduced to a small set of dimensions/factors necessary and sufficient for broadly describing human personality [e.g. the Five-Factor Model (FFM) (McCrae and John 1992) and Mayers-Briggs Model (Myers et al. 1998)]. Likewise, personality in cognitive architectures is often represented by several factors/dimensions, not necessarily based on a particular psychological model. These traits, in turn, determine the emotions and/or drives that the system is likely to experience.

In the simplest case, a single parameter suffices to create a systematic bias in the behavior of the intelligent agent. Both NARS (Wang 2006) and CogPrime (Goertzel et al. 2013) use so called "personality parameter" to define how much evidence is required to evaluate the truth of the logical statement or to plan the next action respectively. The larger the value of the parameter, the more "conservative" the behavior appears to be. In Novamente personality traits of a virtual animal (e.g. aggressiveness, curiosity, playfullness) are linked to emotional states and actions via probabilistic rules (Goertzel et al. 2008b). Similarly, in AIS traits (e.g. nasty, phlegmatic, shy, confident, lazy) are assigned an integer value that defines the degree to which they are exhibited. Abstract rules define what behaviors are more likely given the personality (Rousseau and Hayes-roth 1997). In CLARION personality type determines the baseline strength and initial deficit (i.e. proclivity towards) of the many predefined drives. The mapping is embodied in a pretrained neural network (Sun 2016).

Remarkably, even these simple models can generate a spectrum of personalities. For instance, the Pogamut agent has 9 possible states and 5 personality factors (based on FFM) resulting in 45 distinct mappings. Each of these mappings can produce any of the 12 predefined intentions to a varying degree and generate a wide range of behaviors (Rosenthal and Congdon 2012).

It should be noted that most of the architectures described above offer a very simplified rendering of human emotions. However, two architectures, CLARION and MAMID, go further and deserve a special mention. CLARION provides a cognitively-plausible framework that is capable of addressing emotions, drives and personality traits and relating them to other cognitive processes including decision-making. Three aspects of emotion are modeled: reactive affect (or unconscious experience of emotion), deliberative appraisal (potentially conscious) and coping/action (following appraisal). Thus emotion emerges as an interaction between the explicit and implicit processes and involves (and affects) perception, action and cognition (Sun 2016; Sun et al. 2016). Several CLARION models have been validated using psychological data, computational model of the FFM personalities (Sun and Wilson 2011), coping with bullying at school (Wilson and Sun 2014), performance degradation under pressure (Wilson et al. 2009) and model of stereotype bias induced by social anxiety (Wilson et al. 2010).

MAMID models both generation and effect of emotions resulting from external events, internal interpretations, goals and personality traits. Internally, belief nets relate the task- and individual-specific criteria, e.g. whether a goal failure will result in anxiety in a particular agent (Hudlicka 2008). MAMID has been instantiated in two domains: peacekeeping mission training (Hudlicka 2001) and search-and-rescue tasks (Hudlicka 2005). Of great value are also contributions towards developing an interdisciplinary theory of emotions and design guidelines (Hudlicka 2016; Reisenzein et al. 2013).

7 Memory

Memory is an essential part of any systems-level cognitive model, regardless of whether the model is being used for studying the human mind or for solving engineering problems. Thus, nearly all the architectures featured in this review have memory systems that store intermediate results of computations, enabling learning and adaptation to the changing environment. However, despite their functional similarity, particular implementations of memory systems differ significantly and depend on the research goals and conceptual limitations, such as biological plausibility and engineering factors (e.g. programming language, software architecture, use of frameworks, etc.).

In the cognitive architecture literature, memory is described in terms of its duration (short-and long-term) and type (procedural, declarative, semantic, etc.), although it is not necessarily implemented as separate knowledge stores. This multi-store memory concept is influenced by the Atkinson-Shiffrin model (1968), later modified by Baddeley (Baddeley and Hitch 1974). Such view of memory is dominant in psychology, but its utility for engineering is questioned by some because it does not provide a functional description of various memory mechanisms (Perner and Zeilinger 2011). Nevertheless, most architectures do distinguish between various memory types, although naming conventions may differ. For instance, architectures designed for planning and problem solving have short- and long-term memory storage systems but do not use terminology from cognitive psychology. Long-term knowledge in planners is usually referred to as a knowledge base for facts and problem-solving rules, which correspond to semantic and procedural long-term memory [e.g. Disciple (Tecuci et al. 2000), MACSi (Ivaldi et al. 2014), PRS (Georgeff and Lansky 1986), ARDIS (Martin et al. 2011), ATLANTIS (Gat 1992), IMPRINT (Brett et al. 2002)]. Some architectures also save previously implemented tasks and solved problems, imitating episodic memory [REM (Murdock and Goel 2001), PRODIGY (Veloso 1993)]. Short-term storage in planners is usually represented by a current world model or the contents of the goal stack.

Figure 8 shows a visualization of various types of memory implemented by the architectures. Here we follow the convention of distinguishing between the long-term and short-term storage. Long-term storage is further subdivided into semantic, procedural and episodic types, which store factual knowledge, information on what actions should be taken under certain conditions and episodes from the personal experience of the system respectively. Short-term storage is split into sensory and working memory following (Cowan 2008). Sensory or perceptual memory is a very short-term buffer that stores several recent percepts. Working memory is a temporary storage for percepts that also contains other items related to the current task and is frequently associated with the current focus of attention.

7.1 Sensory memory

The purpose of sensory memory is to cache the incoming sensory data and preprocess it before transferring it to other memory structures. For example, iconic memory assists in solving continuity and maintenance problems, i.e. identifying separate instances of the objects as being the same and retaining the impression of the object when unattended [ARCADIA (Bridewell and Bello 2015)]. Similarly, echoic memory allows an acoustic stimulus to persist long enough for perceptual binding and feature extraction, such as pitch extraction and grouping [MusiCog (Maxwell 2014)]. The decay rate for items in sensory memory is believed to be in the range of tens of milliseconds [EPIC (Kieras and Hornof 2014), LIDA (Baars et al. 2007)] for visual data and longer for audio data [MusiCog (Maxwell 2014)], although
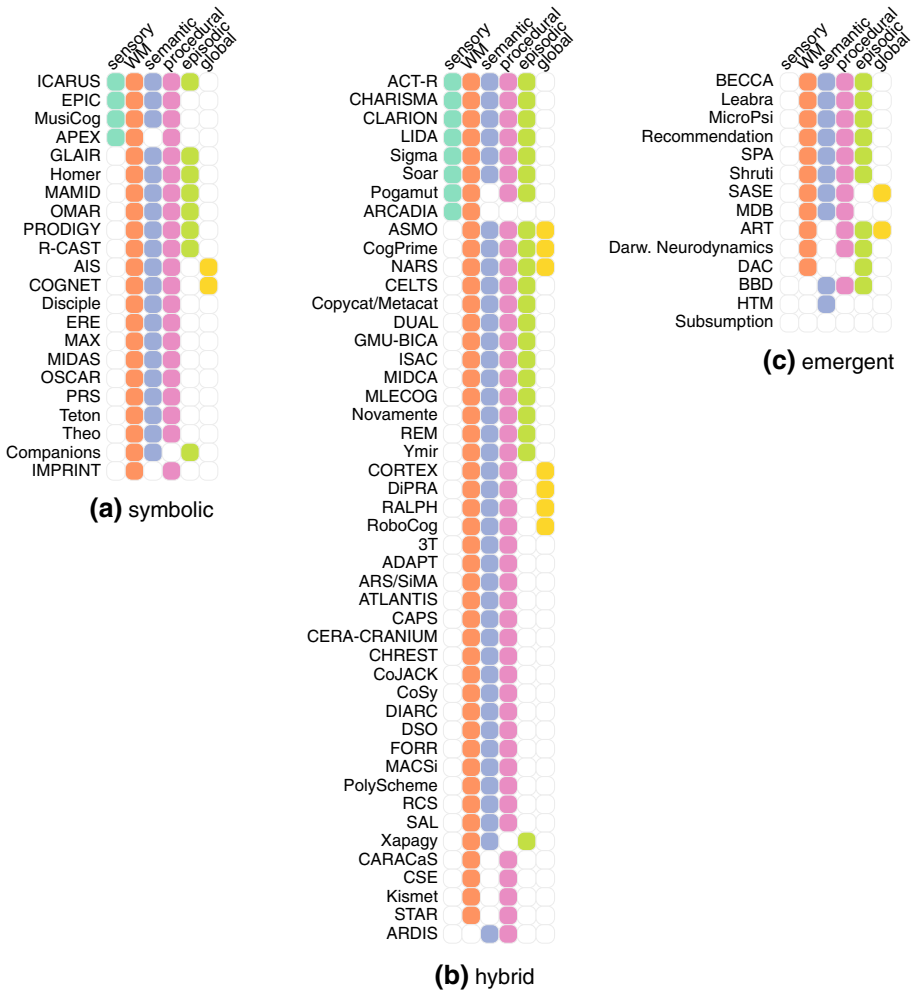
Fig. 8 A diagram showing types of memory implemented in the cognitive architectures. The symbolic, hybrid and emergent architectures are shown in separate columns. The architectures where memory system is unified, i.e. no distinction is made between short-, long-term and other types of memory, are labeled as "global". A version of the diagram with an alphabetical ordering of the architectures is also available on the project website

the time limit is not always specified. Other architectures implementing this memory type include Soar (Wintermute 2009), Sigma (Rosenbloom et al. 2015a), ACT-R (Nyamsuren and Taatgen 2013), CHARISMA (Conforth and Meng 2012), CLARION (Sun 2012), ICARUS (Langley et al. 2005) and Pogamut (Brom et al. 2007).

7.2 Working memory

Working memory (WM) is a mechanism for temporary storage of information related to the current task. It is critical for cognitive capacities such as attention, reasoning and learning, thus every cognitive architecture in our list implements it in some form. Particular realizations

of working memory differ mainly in what information is being stored, how it is represented, accessed and maintained. Furthermore, some cognitive architectures contribute to ongoing research about the processes involved in encoding, manipulation and maintenance of information in human working memory and its relationship to other processes in the human brain.

Despite the importance of working memory for operation, relatively few cognitive architectures provide sufficient detail about its internal organization and connections to other modules. Often WM is vaguely described as "current world state" or "data from the sensors". Based on these statements we infer that in many architectures working memory or an equivalent structure serves as a cache for the current world model, the state of the system and/or current goals. Although usually there are no apparent limitations on the capacity of working memory, new goals or new sensory data often overwrite the existing content. Such simplified account of working memory can be found in many symbolic architectures [3T (Bonasso et al. 1997), ATLANTIS (Gat 1992), Homer (Vere and Bickmore 1990), IMPRINT (Brett et al. 2002), MIDCA (Cox et al. 2012), PRODIGY (Veloso and Blythe 1994), R-CAST (Yen et al. 2010), RALPH (Ogasawara 1991), REM (Murdock and Goel 2008), Theo (Mitchell 1990), Disciple (Tecuci et al. 2007b), OSCAR (Pollock 1993a), ARS/SiMA (Schaat et al. 2014), CoJACK (Ritter et al. 2012), ERE (Bresina and Drummond 1990), MAX (Kuokka 1989), Companions (Forbus et al. 2009), RCS (Albus 2002)] as well as hybrid ones [CSE (Henderson and Joshi 2013), MusiCog (Maxwell 2014), MicroPsi (Bach 2007), Pogamut (Brom et al. 2007)].

More cognitively plausible models of working memory use an activation mechanism. Some of the earliest models of activation of working memory contents were implemented in ACT-R. As before, working memory holds the most relevant knowledge, which is retrieved from the long-term memory as determined by bias. This bias is referred to as activation and consists of base-level activation which may decrease or increase with every access and may also include the spreading activation from neighboring elements. The higher the activation of the element, the more likely it is to enter working memory and directly affect the behavior of the system (Anderson et al. 1996). Mechanisms of this type can be found in the graph-based knowledge representations where nodes refer to concepts and weights assigned to edges correspond to the associations between the concepts [Soar (Laird and Mohan 2014), CAPS (Sanner 1999), ADAPT (Benjamin et al. 2004), DUAL (Kokinov et al. 1996), Sigma (Pynadath et al. 2014), CELTS (Faghihi et al. 2013), NARS (Wang 2007), Novamente (Goertzel and Pennachin 2007), MAMID (Hudlicka 2010)]. Naturally, activation is used in neural network representations as well [Shruti (Wendelken and Shastri 2005), CogPrime (Ikle and Goertzel 2011), Recommendation Architecture (Coward 2011), SASE (Weng and Luciw 2010), Darwinian Neurodynamics (Fedor et al. 2017)]. Activation mechanisms contribute to modeling many properties of working memory, such as limited capacity, temporal decay, rapid updating as the circumstances change, connection to other memory components and decision-making.

Another common paradigm, the blackboard architecture, represents memory as a shared repository of goals, problems and partial results which can be accessed and modified by the modules running in parallel [AIS (Hayes-Roth 1996), CERA-CRANIUM (Arrabales et al. 2011), CoSy (Sjöö et al. 2010), FORR (Epstein 1992), Ymir (Thorisson 1999), LIDA (Franklin et al. 2016), ARCADIA (Bello et al. 2016), Copycat/Metacat (Marshall 2006), CHARISMA (Conforth and Meng 2012), PolyScheme (Cassimatis 2007), PRS (Georgeff et al. 1998)]. The solution to the problem is obtained by continuously updating the shared short-term storage with information from specialized heterogeneous modules analogous to a group of people completing a jigsaw puzzle. Although not directly biologically plausible,

this paradigm works well in situations where many disparate sources of information must be combined to solve a complex real-time task.

Neuronal models of working memory based on the biology of the prefrontal cortex are realized within SPA (Stewart and Eliasmith 2014), ART (Grossberg 2007) and Leabra (O'Reilly and Frank 2006). They demonstrate a range of phenomena (validated on human data) such as adaptive switching between rapid updating of the memory and maintaining previously stored information (O'Reilly 2006) and reduced accuracy with time (e.g. in a list memory task) (Stewart and Eliasmith 2014). In some robotics architectures, ego-sphere (discussed in Sect. 4) besides enabling spatiotemporal integration also provides representation for the current location and the orientation of the robot within the environment [iCub (Ruesch et al. 2008), ISAC (Kawamura et al. 2008), MACSi (Ivaldi et al. 2014)]. Along the same lines, Kismet and RoboCog (Bachiller et al. 2008) use a map built on the camera reference frame to maintain information about recently perceived regions of interest.

Since, by definition, working memory is a relatively small temporary storage, for biological realism, its capacity should be limited. However, there is no consensus on of how this should be done in a cognitive architecture. For instance, in GLAIR the contents of working memory is discarded when the agent switches to a new problem (Shapiro and Bona 2010). A less radical approach is to gradually remove items from the memory based on their recency or relevance in the changing context. The CELTS architecture implements this principle by assigning an activation level to percepts that is proportional to the emotional valence of a perceived situation. This activation level changes over time and as soon as it falls below a set threshold, the percept is discarded (Faghihi et al. 2013). The Novamente Cognitive Engine has a similar mechanism, where atoms stay in memory as long as they build links to other memory elements and increase their utility (Goertzel and Pennachin 2007). It is, however, not guaranteed that the size of working memory will not exceed its capacity without any additional restrictions.

To prevent unlimited growth, a hard limit can be defined for the number of items in memory, for example, 3–6 objects in ARCADIA (Bello et al. 2016), 4 chunks in CHREST (Lloyd-Kelly et al. 2014) or up to 20 items in MDB (Bellas and Duro 2004). Then as the new information arrives, the oldest or the most irrelevant items would be deleted to avoid overflow. Items can also be discarded if they have not been used for some time. The exact amount of time can vary from 4–9 s [EPIC (Kieras 2010)] to 5 s [MIDAS (Hooey et al. 2010), CERA-CRANIUM (Arrabales et al. 2011)] to tens of seconds [LIDA (Franklin 2007)]. In the Recommendation Architecture, a different approach is taken so that the limit of 3–4 items in working memory emerges naturally from the structure of the memory system, and not the external parameter setting (Coward 2011).

7.3 Long-term memory

Long-term memory (LTM) preserves a large amount of information for a very long time. Typically, it is divided into the procedural memory of implicit knowledge (e.g. motor skills and routine behaviors) and declarative memory, which contains explicit knowledge. The latter is further subdivided into semantic (factual) and episodic (autobiographical) memory.

The dichotomies between the explicit/implicit and the declarative/procedural dimensions of long-term memories are rarely implemented directly in cognitive architectures. One of the few exceptions is CLARION, where procedural and declarative memories are separate and both subdivided into an implicit and explicit component. This distinction is preserved on the level of knowledge representation: implicit knowledge is captured by distributed sub-

symbolic structures like neural networks, while explicit knowledge has a transparent symbolic representation (Sun 2012).

Long-term memory is a storage for innate knowledge that enables operation of the system, therefore almost all architectures implement procedural and/or semantic memory. Procedural memory contains knowledge about how to get things done in the task domain. In symbolic production systems, procedural knowledge is represented by a set of if-then rules preprogrammed or learned for a particular domain [3T (Firby 1989), 4CAPS (Varma 2006), ACT-R (Lebiere et al. 2013), ARDIS (Martin et al. 2011), EPIC (Kieras 2004), SAL (Herd et al. 2014), Soar (Lindes and Laird 2016), APEX (Freed and Remington 2000)]. Other variations include sensory-motor schemas [ADAPT (Benjamin et al. 2004)], task schemas [ATLANTIS (Gat 1992)] and behavioral scripts [FORR (Epstein 1994)]. In emergent systems, procedural memory may contain sequences of state-action pairs [BECCA (Rohrer et al. 2009)] or ANNs representing perceptual-motor associations [MDB (Salgado et al. 2012)].

Semantic memory stores facts about the objects and relationships between them. In the architectures that support symbolic reasoning, semantic knowledge is typically implemented as a graph-like ontology, where nodes correspond to concepts and links represent relationships between them [Disciple (Boicu et al. 2003), MIDAS (Corker et al. 1997), Soar (Lindes and Laird 2016), CHREST (Lloyd-Kelly et al. 2015)]. In emergent architectures factual knowledge is represented as patterns of activity within the neural network [BBD (Krichmar and Edelman 2005), SHRUTI (Shastri 2007), HTM (Lavin et al. 2016), ART (Carpenter 2001)].

Episodic memory stores specific instances of past experiences. These can later be reused if a similar situation arises [MAX (Kuokka 1991), OMAR (Deutsch 2006), iCub (Vernon et al. 2010), Ymir (Thorisson 1999)] or exploited for learning new semantic/procedural knowledge. For example, CLARION saves action-oriented experiences as "input, output, result" and uses them to bias future behavior (Sun et al. 1998). Similarly, BECCA stores sequences of state-action pairs to make predictions and guide the selection of system actions (Rohrer et al. 2009), and MLECOG gradually builds 3D scene representation from perceived situations (Jaszuk and Starzyk 2016). MAMID (Hudlicka 2002) saves the past experience together with the specific affective connotations (positive or negative), which affect the likelihood of selecting similar actions in the future. Other examples include R-CAST (Barnes and Hammell 2008), Soar (Nuxoll and Laird 2007), Novamente (Lian et al. 2010) and Theo (Mitchell et al. 1989).

7.4 Global memory

Despite the evidence for the distinct memory systems, some architectures do not have separate representations for different kinds of knowledge or short- versus long-term memory, and instead, use a unified structure to store all information in the system. For example, CORTEX and RoboCog use an integrated, dynamic multi-graph object which can represent both sensory data and high-level symbols describing the state of the robot and the environment (Romero-Garcés et al. 2015b; Bustos et al. 2013). Similarly, AIS implements a global memory, which combines a knowledge database, intermediate reasoning results and the cognitive state of the system (Hayes-Roth et al. 1992). DiPRA uses Fuzzy Cognitive Maps to represent goals and plans (Pezzulo et al. 2007). NARS represents all empirical knowledge, regardless of whether it is declarative, episodic or procedural, as formal sentences in Narcese (Wang and Hammer 2015b). Similarly, in some emergent architectures, such as SASE (Weng and Luciw 2010) and ART (Carpenter 2001), the role of neurons as working or long-term memory is dynamic and depends on whether the neuron is firing.

Conclusion Overall, research on memory in the cognitive architectures mainly concerns its structure, representation and retrieval. Relatively little attention has been paid to the challenges associated with maintaining a large-scale memory store since both the domains and the time spans of the intelligent agents are typically limited. In comparison, early estimates of the capacity of the human long-term memory are within 1.5 gigabits or on the order of 100K concepts (Landauer 1986), while more recent findings suggest that human brain capacity may be orders of magnitude higher (Bartol et al. 2015). However, scaling naïve implementations even to the lowest estimated size of human memory likely will not be possible despite the increase in available computational power. Alternative solutions include tapping into existing methods for large-scale data management and improving retrieval algorithms. Both avenues have been explored, the former by Soar and ACT-R, which used PostgreSQL relational database to load concepts and relations from WordNet (Derbinsky and Laird 2012b; Douglass et al. 2009), and the latter by the Companions architecture (Forbus et al. 2016). Alternatively, SPA supports a biologically plausible model of associative memory capable of representing over 100K concepts in WordNet using a network of spiking neurons (Crawford et al. 2015).

8 Learning

Learning is the capability of a system to improve its performance over time. Ultimately, any kind of learning is based on experience. For example, a system may be able to infer facts and behaviors from the observed events or from results of its own actions. The type of learning and its realization depend on many factors, such as design paradigm (e.g. biological, psychological), application scenario, data structures and the algorithms used for implementing the architecture, etc. However, we will not attempt to analyze all these aspects given the diversity and number of the cognitive architectures surveyed. Besides, not all of these pieces of information can be easily found in the publications.

Thus, a more general summary is preferable, where types of learning are defined following taxonomy by Squire (1992). Learning is divided into declarative or explicit knowledge acquisition and non-declarative, which includes perceptual, procedural, associative and non-associative types of learning. Figure 9 shows a visualization of these types of learning for all cognitive architectures.

We also identified 19 architectures (mostly symbolic and hybrid) that do not implement any learning. In some areas of research, learning is not even necessary, for example, in human performance modeling, where accurate replication of human performance data is required instead (e.g. APEX, EPIC, IMPRINT, MAMID, MIDAS, etc.). Some of the newer architectures are still in the early development stage and may add learning in the future (e.g. ARCADIA, SiMA, Sigma).

8.1 Perceptual learning

Although many systems use pre-learned components for processing perceptual data, such as object and face detectors or classifiers, we do not consider these here. Perceptual learning applies to the architectures that actively change the way sensory information is handled or how patterns are learned on-line. This kind of learning is frequently performed to obtain implicit knowledge about the environment, such as spatial maps [RCS (Schlenoff et al. 2005), AIS (Hayes-Roth et al. 1993), MicroPsi (Bach et al. 2007)], clustering visual features [HTM (Kostavelis et al. 2012), BECCA (Rohrer 2011b), Leabra (O'Reilly et al. 2013)] or finding
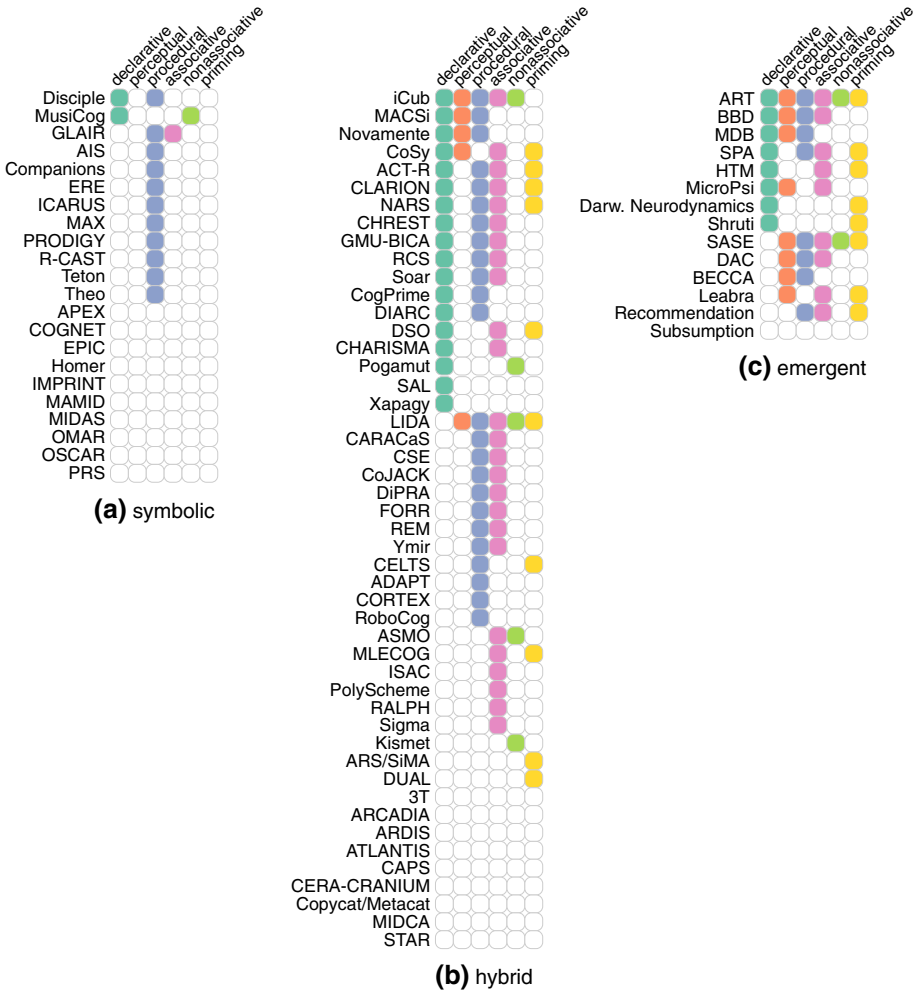
Fig. 9 A diagram summarizing types of learning implemented in the cognitive architectures. Symbolic, hybrid and emergent architectures are grouped and presented in different columns. The architectures are arranged according to the similarity between the implemented learning types. A version of the diagram with an alphabetical ordering of the architectures is also available on the project website

associations between percepts. The latter can be used within the same sensory modality as in the case of the agent controlled by Novamente engine, which selects a picture of the object it wants to get from a teacher (Goertzel 2008). Learning can also occur between different modalities, for instance, the robot based on the SASE architecture learns the association between spoken command and action (Weng and Zhang 2002) and Darwin VII (BBD) learns to associate taste value of the blocks with their visual properties (Edelman 2007).

8.2 Declarative learning

Declarative knowledge is a collection of facts about the world and various relationships defined between them. In many production systems such as ACT-R and others, which imple-

ment chunking mechanisms [(SAL (Jilk et al. 2008), CHREST (Schiller and Gobet 2012), CLARION (Sun et al. 1999)], new declarative knowledge is learned when a new chunk is added to declarative memory (e.g. when a goal is completed). Similar acquisition of knowledge has also been demonstrated in systems with distributed representations. For example, in DSO knowledge can be directly input by human experts or learned as contextual information extracted from labeled training data (Ng et al. 2012). New symbolic knowledge can also be acquired by applying logical inference rules to already known facts [GMU-BICA (Samsonovich et al. 2008), Disciple (Boicu et al. 2005), NARS (Wang 2010)]. In many biologically inspired systems learning new concepts usually takes the form of learning the correspondence between the visual features of the object and its name [iCub (Di Nuovo et al. 2014), Leabra (O'Reilly et al. 2013), MACSi (Ivaldi et al. 2014), Novamente (Goertzel 2008), CoSy (Hawes et al. 2010), DIARC (Yu et al. 2010)].

8.3 Procedural learning

Procedural learning refers to learning skills, which happens gradually through repetition until the skill becomes automatic. The simplest way of doing so is by accumulating examples of successfully solved problems to be reused later [e.g. AIS (Hayes-Roth et al. 1993), R-CAST (Fan et al. 2010a), RoboCog (Manso et al. 2014)]. For instance, in a navigation task, a traversed path could be saved and used again to go between the same locations later [AIS (Hayes-Roth et al. 1993)]. Obviously, this type of learning is very limited and further processing of accumulated experience is needed to improve efficiency and flexibility.

Explanation-based learning (EBL) (Minton et al. 1989) is a common technique for learning from experience found in many architectures with symbolic representation for procedural knowledge [PRODIGY (Etzioni 1993), Teton (VanLehn 1989), Theo (VanLehn 1989), Disciple (Boicu et al. 2005), MAX (Kuokka 1991), Soar (Laird et al. 2004), Companions (Friedman and Forbus 2010), ADAPT (Benjamin and Lyons 2010), ERE (Kedar and McKusick 1992), REM (Murdock and Goel 2008), CELTS (Faghihi et al. 2011a), RCS (Albus et al. 1995)]. In short, it allows to turn an explanation of a single observed instance into a general rule. However, EBL does not extend the problem domain, but rather makes solving problems more efficient in similar situations. A known drawback of this technique is that it can result in too many rules (also known as the "utility problem"), which may slow down the inference. To avoid explosion in the model knowledge, various heuristics can be applied, e.g. adding constraints on events that a rule can contain [CELTS (Faghihi et al. 2011a)] or eliminating low-use chunks [Soar (Kennedy and De Jong 2003)]. Although EBL is not biologically inspired, it has been shown that in some cases human learning may exhibit EBL-like behavior (Vanlehn et al. 1990).

Symbolic procedural knowledge can also be obtained by inductive inference [Theo (Mitchell et al. 1989), NARS (Slam et al. 2015)], learning by analogy [Disciple (Tecuci and Kodratoff 1990), NARS (Slam et al. 2015)], behavior debugging [MAX (Kuokka 1991)], probabilistic reasoning [CARACaS (Huntsberger 2011)], correlation and abduction [RCS (Albus et al. 1995)] and explicit rule extraction [CLARION (Sun et al. 2011)].

8.4 Associative learning

Associative learning is a broad term for decision-making processes influenced by reward and punishment. In behavioral psychology, it is studied within two major paradigms: classical (Pavlovian) and instrumental (operant) conditioning. Reinforcement learning (RL) and

its variants, such as temporal difference learning, Q-learning, Hebbian learning, etc., are commonly used in computational models of associative learning. Furthermore, there is substantial evidence that error-based learning is fundamental for decision-making and motor skill acquisition (Holroyd and Coles 2002; Niv 2009; Seidler et al. 2013).

The simplicity and efficiency of reinforcement learning make it one of the most common techniques with nearly half of all the cognitive architectures using it to implement associative learning. The advantage of RL is that it does not depend on the representation and can be sucessfully applied in the symbolic, emergent or hybrid architectures.

One of the main uses of RL is developing adaptive behavior. In systems with symbolic components it can be accomplished by changing the importance of actions and beliefs based on their success/failure [e.g. RCS (Albus and Barbera 2005), NARS (Slam et al. 2015), REM (Murdock and Goel 2008), RALPH (Ogasawara and Russell 1993), CHREST (Schiller and Gobet 2012), FORR (Gordon and Epstein 2011), ACT-R (Cao et al. 2015), CoJACK (Ritter 2009)]. In the hybrid and emergent systems, reinforcement learning can establish associations between states and actions. One application is sensorimotor reconstruction. The associations are often established in two stages: a "motor babbling" stage, during which the system performs random actions to accumulate data, followed by the learning stage where a model is built using the accrued experiences [ISAC (Kawamura et al. 2008), BECCA (Rohrer et al. 2009), iCub (Metta et al. 2010), DiPRA (Pezzulo 2009), CSE (Henderson et al. 2011)]. Associative learning can be applied when analytical solutions are hard to obtain as in the case of soft-arm control [ISAC (Kawamura et al. 2008)] or when smooth and natural looking behavior is desired [e.g. Segway platform control using BBD (Krichmar 2012)].

8.5 Non-associative learning

Non-associative learning, as the name suggests, does not require associations to link stimuli and responses together. Habituation and sensitization are commonly identified as two types of non-associative learning. Habituation describes gradual reduction in the strength of response to repeated stimuli. The opposite process occurs during sensitization, i.e. when repeated exposure to stimuli causes increased response. Because of their simplicity, these types of learning are considered a prerequisite for other forms of learning. For instance, habituation filters out irrelevant stimuli and helps to focus on important stimuli (Rankin et al. 2009), especially in the situations when positive or negative rewards are absent (Weng and Hwang 2006).

Most of the work to date in this area has been dedicated to the habituation in the context of social robotics and humancomputer interaction to achieve adaptive and realistic behavior. For instance, the ASMO architecture enables the robot to ignore irrelevant (but salient) stimuli during the tracking task, when the robot may be easily distracted by fast motion in the background. Habituation learning (implemented as a boost value attached to motion module) allows it to focus on the stimuli relevant for the task (Novianto et al. 2013). In Kismet (Breazeal and Scassellati 1999) and iCub (Ruesch et al. 2008) habituation (realized as a dedicated habituation map) causes the robot to look at non-preferred or novel stimuli. In LIDA and Pogamut habituation is employed to reduce the emotional response to repeated stimuli in human-robot interaction (Franklin 2000b) and virtual agent (Gemrot et al. 2009) scenarios respectively. MusiCog includes habituation effects to introduce novelty in music generation by aggressively suppressing the salience of elements in working memory that have been stored for a long time (Maxwell 2014).

In comparison, little attention has been paid to sensitization. In ASMO, in addition to habituation, sensitization learning allows the social robot to focus on the motion near the tracked object even though it may be slow and not salient (Novianto et al. 2013). The SASE architecture also describes both mechanisms of non-associative learning (Huang and Weng 2007).

8.6 Priming

Priming occurs when prior exposure to stimulus affects its subsequent identification and classification. Numerous experiments demonstrated the existence of priming effects with respect to various stimuli, perceptual, semantic, auditory, etc., as well as behavioral priming (Higgins and Eitam 2014). In Sect. 4.1 we mentioned some examples of priming in vision, namely spatial [STAR (Kotseruba 2016)] and feature priming [Kismet (Breazeal et al. 2001), ART (Grossberg 2003)], which allow for more efficient processing of stimuli by biasing the visual system.

Priming in cognitive architectures has been investigated both within applications and theoretically. For example, in HTM priming is used in the context of analyzing streams of data such as spoken language. Priming is essentially a prediction of what is more likely to happen next and helps to resolve ambiguities based on those expectations (Hawkins and George 2006). Experiments with CoSy show that priming speech recognition results in a statistically significant improvement compared to a baseline system (Lison and Kruijff 2008). In other instances, application of priming lead to improving perceptual categorization [ARS/SiMA (Schaat et al. 2013b)], skill transfer [SASE (Zhang and Weng 2007)], problem solving [DUAL (Kokinov 1990)] and question answering [Shruti (Shastri 1998)].

Several models of priming validated by human data were developed in the lexical domain and problem solving. For instance, the CLARION model of positive and negative priming in lexical decision tasks models the fact that human participants are faster at identifying sequences of related concepts (e.g. the word "butter" preceded by the word "bread") (Helie and Sun 2014). Darwinian Neurodynamics confirms priming in problem solving by showing that people who were primed for a more efficient solution perform better in solving puzzles (Fedor et al. 2017).

Priming is well studied in psychology and neuroscience (Wentura and Rothermund 2014) and two major theoretical frameworks for modeling priming are found in cognitive architectures: spreading activation [ACT-R (Thomson et al. 2014), Recommendation Architecture (Coward 2011), Shruti (Shastri 1998), CELTS (Faghihi 2011), LIDA (Franklin 2000a), ARS/SiMA (Schaat et al. 2013b), DUAL (Petkov et al. 2006), NARS (Wang 2006)] and attractor networks [CLARION (Helie and Sun 2014), Darwinian Neurodynamics (Fedor et al. 2017)]. The current consensus in the literature is that spreading activation has greater explanation power, but attractor networks are considered more biologically plausible (Lerner et al. 2012). The choice of the particular paradigm may also be influenced by the representation. For instance, spreading activation is found in the localist architectures, where units correspond to concepts. When a concept is invoked, a corresponding unit is activated, the activation is spread to adjacent related units, which facilitates their further use. Alternatively, in attractor networks, a concept is represented by a pattern involving multiple units. Depending on the correlation (relatedness) between the patterns, activating one leads to increase in activation of others. An alternative explanation of priming is offered by SPA based on parallel constraint satisfaction (Schroder and Thagard 2014).

9 Reasoning

Reasoning, originally a major research topic in philosophy and epistemology, in the past decades has become one of the focal points in psychology and cognitive sciences as well. As an ability to logically and systematically process knowledge, reasoning can affect or structure virtually any form of human activity. As a result, aside from the classic triad of logical inference (deduction, induction and abduction), other kinds of reasoning are now being considered, such as heuristic, defeasible, analogical, narrative, moral, etc.

Predictably, all cognitive architectures are concerned with practical reasoning, whose end goal is to find the next best action and perform it, as opposed to the theoretical reasoning that aims at establishing or evaluating beliefs. There is also a third option, exemplified by the Subsumption architecture, which views reasoning about actions as an unnecessary step (Brooks 1987) and instead pursues physically grounded action (Brooks 1990). Granted, one could still argue that a significant amount of reasoning and planning is required from a designer to construct a grounded system with non-trivial capabilities. Otherwise, in the context of cognitive architectures, reasoning is primarily mentioned with regards to planning, decision-making and learning, as well as perception, language understanding and problem-solving.

One of the main challenges a human, and consequently any human-level intelligence, faces regularly is acting based on insufficient knowledge or "making rational decisions against a background of pervasive ignorance" (Pollock 2008). In addition to sparse domain knowledge, there are limitations in terms of available internal resources, such as information processing capacity. Furthermore, external constraints, such as real-time execution, may also be introduced by the task. It should be noted that even in the absence of these restrictions, reasoning is by no means trivial. Consider, for instance, the complex machinery used by Copycat and Metacat to model analogical reasoning in a micro-domain of character strings (Hofstadter 1993).

So far, the most consolidated effort has been spent on overcoming the insufficient knowledge and/or resources problem in continuously changing environments. In fact, this is the goal of general-purpose reasoning systems such as the Procedural Reasoning System (PRS), the Non-Axiomatic Reasoning System (NARS), OSCAR and the Rational Agent with Limited-Performance Hardware (RALPH). PRS is one of the earliest instances of the BDI (belief-desire-intention) model. In each reasoning cycle, it selects a plan matching the current beliefs, adds it to the intention stack and executes it. If new goals are generated during the execution, new intentions are created (Georgeff and Ingrand 1989). NARS approaches the issue of insufficient knowledge and resources by iteratively reevaluating existing evidence and adjusting its solutions accordingly. This is possible with Non-Axiomatic Logic, which associates truth values with each statement thus allowing agents to express their confidence in the belief, a feature not available in PRS (Hammer et al. 2016). RALPH tackles complex goal-driven behavior in complex domain using decision-theoretic approach. Thus, computation is synonymous with action and action with the highest utility value (based on its expected effect) is always selected (Russel and Wefald 1988). The OSCAR architecture explores defeasible reasoning, i.e. reasoning which is rationally compelling but not deductively valid (Koons 2017). This is a more accurate representation of everyday reasoning, where knowledge is sparse, tasks are complex and there are no certain criteria for measuring success.

All the architectures mentioned above aim at creating rational agents with human-level intelligence but they do not necessarily try to model human reasoning processes. This is the

goal of architectures such as ACT-R, Soar, DUAL and CLARION. In particular, Human Reasoning Module implemented in ACT-R works under the assumption that human reasoning is probabilistic and inductive (although deductive reasoning may still occur in the context of some tasks). This is demonstrated by combining a deterministic rule-based inference mechanism with long-term declarative memory with properties similar to human memory, i.e. incomplete and inconsistent knowledge. Together, the uncertainty inherent in the knowledge and production rules enable human-like reasoning behavior, which has been confirmed experimentally (Nyamsuren and Taatgen 2014). Similarly, symbolic and sub-symbolic mechanisms in CLARION (Helie and Sun 2014) and DUAL (Kokinov 1994) architectures are used to model and explain various psychological phenomena associated with deductive, inductive, analogical and heuristic reasoning.

As discussed in Sect. 3, the question of whether higher human cognition is inherently symbolic or not is still unresolved. The symbolic and hybrid architectures considered so far, view reasoning mainly as a symbolic manipulation. Given that in emergent architectures the information is represented by the weights between individual units, what kind of reasoning, if any, can they support? It appears, based on the cognitive architecture literature, that many of the emergent architectures, e.g. ART, HTM, DAC, BBD, BECCA, simply do not address reasoning, although they certainly exhibit complex intelligent behavior. On the other hand, since it is possible to establish a correspondence between the neural networks and logical reasoning systems (Martins and Mendes 2001), then it should also be possible to simulate symbolic reasoning via neural mechanisms. Indeed, several architectures demonstrate symbolic reasoning and planning (SASE, SHRUTI, MicroPsi). To date, one of the most successful implementations of symbolic reasoning (and other low- and high-level cognitive phenomena) in a neural architecture is represented by SPA (Rasmussen and Eliasmith 2013). Architectures like this also raise interesting questions whether it makes sense to try and segregate symbolic parts of cognition from sub-symbolic. As most existing cognitive architectures represent a continuum from purely symbolic to connectionist, the same may be true for the human cognition as well.

10 Metacognition

Metacognition (Flavell 1979), intuitively defined as "thinking about thinking", is a set of abilities that introspectively monitor internal processes and reason about them.[16] There has been a growing interest in developing metacognition for artificial agents, both due to its essential role in human experience and practical necessity for identifying, explaining and correcting erroneous decisions. Approximately one-third of the surveyed architectures, mainly symbolic or hybrid ones with a significant symbolic component, support metacognition with respect to decision-making and learning. Here we will focus on three most common metacognitive mechanisms, namely self-observation, self-analysis and self-regulation.

Self-observation allows gathering data pertaining to the internal operation and status of the system. The data commonly includes the availability/requirements of internal resources [AIS (Hayes-Roth 1995), COGNET (Zachary et al. 2000), Soar (Laird 2012a)] and current task knowledge with associated confidence values [CLARION (Sun 2016), CoJACK (Evertsz et al. 2007), Companions (Friedman et al. 2011), Soar (Laird 2012a)]. In addi-

---

[16] Due to scope limitations we did not examine other concepts related to metacognition. Discussions of the relationship between self-awareness and metacognition, the notion of 'self', consciousness and the appropriateness of using these psychological terms in computational models can be found in Cox (2005, 2007).

tion, some architectures support a temporal representation (trace) of the current and/or past solutions [Companions (Friedman et al. 2011), Metacat (Jensen and Veloso 1998), MIDCA (Dannenhauer et al. 2014)].

Overall, the amount and granularity of the collected data depend on the further analysis and application. In decision-making applications, after establishing the amount of available resources, conflicting requests for resources, discrepancies in the conditions, etc., the system may change the priorities of different tasks/plans [AIS (Hayes-Roth 1995), CLARION (Sun 2016), COGNET (Zachary et al. 2000), CoJACK (Evertsz et al. 2007), MAMID (Hudlicka 2009), PRODIGY (Veloso et al. 1998), RALPH (Ogasawara and Russell 1993)]. The trace of the execution and/or recorded solutions for the past problems can improve learning. For example, repeating patterns found in past solutions may be exploited to reduce computation for similar problems in the future [Metacat (Marshall 2002b), FORR (Epstein and Petrovic 2008), GLAIR (Shapiro and Bona 2010), Soar (Laird 2012a)]. Traces are also useful for detecting and breaking out of repetitive behavior [Metacat (Marshall 2002b)], as well as stopping learning if the accuracy does not improve [FORR (Epstein and Petrovic 2008)].

Although in theory many architectures support metacognition, its usefullness has only been demonstrated in limited domains. For instance, Metacat applies metacognition to analogical reasoning in a micro-domain of strings of characters (e.g. if abc → abd; mrrjjj → ?). Self-watching allows the system to remember past solutions, compare different answers and justify its decisions (Marshall 2002b). In PRS, metacognition enables efficient problem solving in playing games such as Othello (Russell and Wefald 1989), better control of the simulated vehicle (Ogasawara and Russell 1993) and real-time adaptability to continuously changing environments (Georgeff and Ingrand 1989). Internal error monitoring, i.e. the comparison between the actual and expected perceptual inputs, enables the SAL architecture to determine the success or failure of its actions in a simple task, such as stacking blocks, and to learn better actions in the future (Vinokurov et al. 2013). The metacognitive abilities integrated into the dialog system make the CoSy architecture more robust to variable dialogue patterns and increase its autonomy (Christensen and Kruijff 2010). The Companions architecture demonstrates the importance of self-reflection by replicating data from the psychological experiments on how people organize, represent and combine incomplete domain knowledge into explanations (Friedman et al. 2011). Likewise, CLARION models human data for the Metcalfe (1986) task and replicates the Gentner and Collins (1981) experiment, both of which involve metacognitive monitoring (Sun et al. 2006).

Metacognition is also necessary for social cognition, particularly for the skill known in the psychological literature as a Theory of Mind (ToM). ToM refers to being able to acknowledge and understand mental states of other people, use the judgment of their mental state to predict their behavior and inform one's own decision-making. Very few architectures support this ability. For instance, recently Sigma demonstrated two distinct mechanisms for ToM in several single-stage simultaneous-move games, such as the well-known Prisoner's dilemma. The first mechanism is automatic as it involves probabilistic reasoning over the trellis graph and the second is a combinatorial search across the problem space (Pynadath et al. 2013).

PolyScheme applies ToM to perspective taking in a human-robot interaction scenario. The robot and human in this scenario are together in a room with two traffic cones and multiple occluding elements. The human gives a command to move towards a cone, without specifying which one. If only one cone is visible to the human, the robot can model the scene from the human's perspective and use this information to disambiguate the command (Trafton et al. 2005). Another example is reasoning about beliefs in a false belief task. In this scenario, two agents A and B observe a cookie placed in a jar. After B leaves, the cookie is moved to

another jar. When B comes back, A is able to reason that B still believes that cookie is still in the first jar (Scally et al. 2012).

ACT-R was used to build several models of false belief and second-order false belief tasks (answering questions of the kind "Where does Ayla think Murat will look for chocolate?") which are typically used to assess whether children developed a Theory of Mind (Triona et al. 2001; Barslanrugnl et al. 2004). However, of particular interest is the recent developmental model of ToM based on ACT-R, which was subsequently implemented on a mobile robot. Several scenarios were set up to show how ToM ability can improve the quality of interaction between the robot and humans. For instance, in a patrol scenario both the robot and human receive the task of patrolling the south area of the building, but as they start, the instructions are changed to head west. When the human starts walking towards south, the robot infers that she might have forgotten about the new task and reminds her of the new changes (Trafton et al. 2013).

11 Practical applications

Cognitive architectures reviewed in this paper are mainly used as research tools and very few are developed outside of academia. However, it is still appropriate to talk about their practical applications, since useful and non-trivial behavior in arbitrary domains is perceived as intelligent and in a sense validates the underlying theories. We seek answers to the following questions: what cognitive abilities have been demonstrated by the cognitive architectures and what particular practical tasks have they been applied to?

After a thorough search through the publications, we identified more than 900 projects implemented using 84 cognitive architectures. Our findings are summarized in Fig. 10. It shows the cognitive abilities associated with applications, the total number of applications for each architecture (represented by the length of the bars) and what application categories they belong to (the types of categories and the corresponding number of applications in each category are shown by the color and length of the bar segments respectively).

11.1 Competency areas

In order to determine the scope of research in cognitive architectures with respect to modeling various human cognitive abilities, we began with the list of human competencies proposed by Adams et al. (Adams et al. 2012). This list overlaps with cognitive abilities that have been identified for the purposes of evaluation of cognitive architectures (mentioned in the beginning of Sect. 3), and adds several new areas such as social interaction, emotion, building/creation, quantitative skills and modeling self/other.

In the following analysis, we depart from the work in Adams et al. (2012) in several aspects. First, these competency areas are meant for evaluating AI and it is expected that each area will have a set of associated scenarios. Since we are unable to evaluate these concepts directly, we instead will discuss which of these competency areas are represented in existing practical applications based on the published papers. Second, the authors of (Adams et al. 2012) define subareas for each competency area (e.g. tactical, strategic, physical and social planning). Although in the following sections we will review some of these subareas, they are not shown in the diagram in Fig. 10. Perception, memory, attention, learning, reasoning, planning and motivation (corresponding to internal factors in Sect. 6 on action selection) are covered in the previous sections and will not be repeated here. Actuation (which we sim-
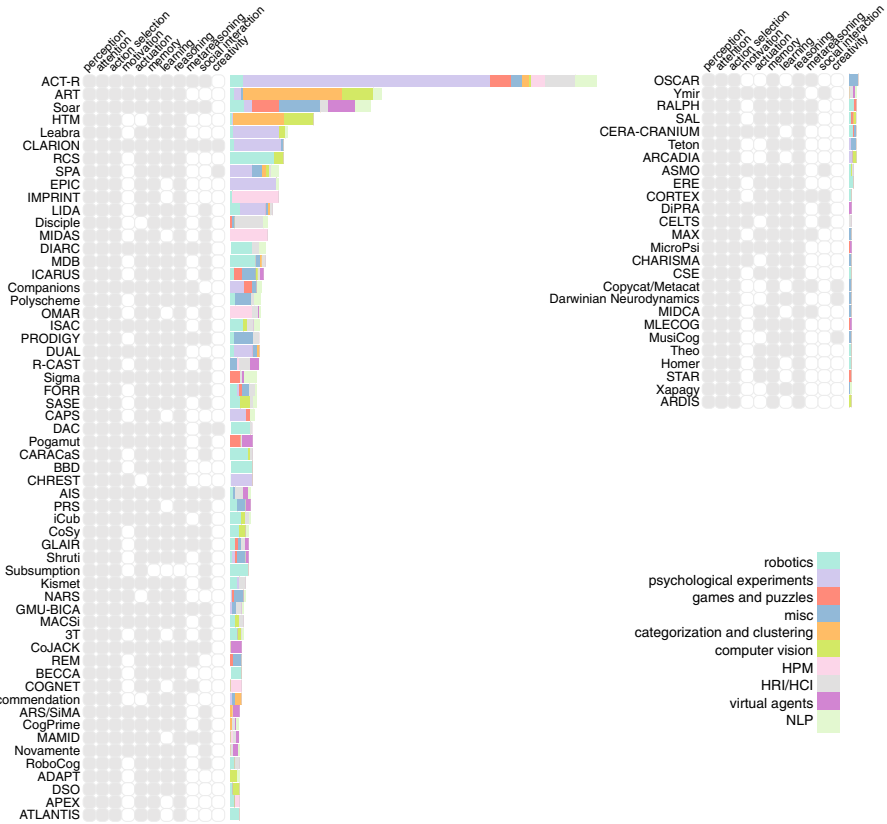
Fig. 10 A graphical representation of the practical applications of the cognitive architectures and corresponding competency areas defined in Adams et al. (2012). The plot is split into two columns for better readability (the right column being the continuation of the left one). Within each column the table shows the implemented competency areas (indicated by gray colored cells). The stacked bar plots show the practical applications: the length of the bar represents the total number of the practical applications found in the publications and colored segments within each bar represent different categories (see legend) and their relative importance (calculated as a proportion of the total number of applications implementing these categories). The architectures are sorted by the total number of their practical applications. For short descriptions of the projects, citations and statistics for each architecture refer to the interactive version of this diagram on the project website

ply interpret as an ability to send motor commands both in real and simulated environments) and social interaction will be discussed in the following subsections on practical applications in robotics (Sect. 11.2.2), natural language processing (Sect. 11.2.5) and human-robot interaction (Sect. 11.2.4). Below we briefly examine the remaining competency areas titled "Building/creation" (which we changed to "Creativity").[17]

---

[17] Quantitative skills, i.e. the ability to use or manipulate quantitative information, are underrepresented in cognitive architectures and therefore are excluded from our analysis. For completeness, we mention few implemented examples: counting moving and stationary objects in a simulated environment [GLAIR (Santore and Shapiro 2003)], computing a sum of two values [SPA (Eliasmith and Stewart 2012)], identifying and counting the number of surrounding faces in a room [DIARC (Scheutz et al. 2007)] and finger counting [iCub (Di Nuovo et al. 2014)]. Psychological experiments involving quantitative skills inlcude: counting tasks investigated by ACT-R (Pape and Urbas 2008) and CLARION (Sun and Zhang 2003), cognitive arithmetic model for addition and subtraction [Soar (Wang and Laird 2006)], models of multi-column subtraction used

The *creativity* competency area in Adams et al. (2012) also includes physical construction. Since demonstrated skills in physical construction are presently limited to toy problems, such as stacking blocks in a defined work area using a robotic arm [Soar (Laird et al. 1991)] or in simulated environments [SAL (Vinokurov et al. 2013)], below we will focus on computational creativity.

Adams et al. (2012) list forming new concepts and learning declarative knowledge (covered in Sect. 8) within the "creation" competency area. While it can be argued that any kind of learning leads to creation of new knowledge or behavior, in this subsection we will focus more specifically on computational creativity understood as computationally generated novel and appropriate results that would be regarded as creative by humans(Duch 2006; Still and D'Inverno 2016). Examples of such results are artifacts with artistic value or innovations in scientific theories, mathematical proofs, etc. Some computational models of creativity also attempt to capture and explain the psychological and neural processes underlying human creativity.

In the domain of cognitive architectures, creativity remains a relatively rare feature as only 7 out of 84 architectures on our list support creative abilities. They represent several directions in the field of computational creativity such as behavioral improvisation, music generation, creative problem solving and insight. An in-depth discussion of these problems is beyond the scope of this survey, but we will briefly discuss the contributions in these areas.

AIS and MusiCog explore improvisation in story-telling and music generation respectively. AIS enables improvisational story-telling by controlling puppets that collaborate with humans in Virtual Theater. The core plot of the story is predefined, as well as a set of responses and emotions for each puppet. These pre-recorded responses are combined with individual perceptions, knowledge base, state information and personality of the puppets. Since puppets react to both the human player and each other's actions, they demonstrate a wide range of variations on the basic plots. Such directed improvisation transforms a small set of abstract directions (e.g. be playful, curious, friendly) into novel and engaging experiences (Hayes-Roth and Gent 1995). A different strategy is followed by MusiCog. It does not start with predefined settings, but instead listens to music samples and attempts to capture their structure. After learning is completed, the acquired structures can be used to generate novel musical themes as demonstrated on Bach BWC 1013 (Maxwell et al. 2012). Another example of synthetic music generation is Roboser, a system comprised of a mobile robot controlled by the DAC architecture and a composition engine which converts the behavioral and control states of the robot to sounds. Together, the predefined transformations and non-deterministic exploratory behavior of the robot result in unique musical structures (Manzolli and Verschure 2005).

Other architectures attempt to replicate the cognitive processes of human creativity rather than its results. One of the early models, Copycat (extended later to Metacat), examined psychological processes responsible for analogy-making and paradigm shift (also known as the "Aha!" phenomenon) in problem solving (Marshall 2002a). To uncover deeper patterns and find better analogies, Copycat performs a random micro-exploration in many directions. A global parameter controls the degree of randomness in the system and allows it to pursue both obvious and peculiar paths. Multiple simulations in a microdomain of character strings show that such approach can result in rare creative breakthroughs (Hofstadter 1993).

Footnote 17 continued

to investigate the hierarchical problem representation and solving [Teton (VanLehn et al. 1989), ICARUS (Langley et al. 2004) and Soar (Rosenbloom et al. 1991)].

CLARION models insight in a more cognitively plausible way and is validated on human experimental data. The sudden emergence of a potential solution is captured through interaction between the implicit (distributed) and explicit (symbolic) representations (Sun and Helie 2015). For example, explicit memory search provides stereotypical semantic associations, while implicit memory allows connecting remote concepts, hence has more creative potential. During the iterative search, multiple candidates are found through implicit search and are added into explicit knowledge. Simulations confirm that the stochasticity of implicit search improves chances of solving the problem (Helie and Sun 2010).

Likewise, Darwinian Neurodynamics (DN) models insight during problem solving. In order to arrive to a solution, the system switches from analytical problem solving based on past experience to implicit generation of new hypotheses via evolutionary search. At the core of the DN model, a population of attractor networks stores the hypotheses which are selected for reproduction based on the fitness criteria. If the initial search through known answers in memory fails, the evolutionary search generates new candidates via mutations. The model of insight during solving a four-tree problem is compared to human behavioral data, accurately demonstrating the effects of priming and prior experience and the size of working memory on problem solving abilities (Fedor et al. 2017).

Last but not least is the spiking neuron model of Remote Associates Test (RAT) often used in creativity research. However, the focus is not on creativity per se, but rather on a biologically realistic representation for the words in the RAT test and associations between them using the Semantic Pointer Architecture (SPA) (Gosmann et al. 2017).

11.2 Practical application categories

We identified ten major categories of practical applications, namely human performance modeling (HPM), games and puzzles, robotics, psychological experiments, natural language processing (NLP), human-robot and human-computer interaction (HRI/HCI), computer vision, categorization and clustering, virtual agents and miscellaneous, which included projects not related to any major group but not numerous enough to be separated into a group of their own. Such grouping of projects emphasizes the application aspect of each project, although the goal of the researchers may have been different.

Note that some applications are associated with more than one category. For example, Soar has been used to play board games with a robotic arm (Kirk and Laird 2016), which is relevant to both robotics and games and puzzles. Similarly, the ACT-R model of Tower of Hanoi evaluated against the human fMRI data (Anderson et al. 2005) belongs to games and puzzles and to psychological experiments.

11.2.1 Psychological experiments

The psychological experiments category is the largest, comprising more than one-third of all applications. These include replications of numerous psychophysiological, fMRI and EEG experiments in order to demonstrate that cognitive architectures can adequately model human data or give reasonable explanations for existing psychological or physiological phenomena. If the data produced by the simulation matches the human data in some or most aspects, it is taken as an indication that a given cognitive architecture can to some extent imitate human cognitive processes.

Most of the experiments investigate psychological phenomena related to memory, perception, attention and decision making. However, there is very little overlap among the specific

studies that were replicated by the cognitive architectures. Notably, the Tower of Hanoi (and its derivative, Tower of London) is the only task reproduced by a handful of architectures (although not on the same human data). The task itself requires one to move a stack of disks from one rod to another and is frequently used in psychology to study problem solving and skill learning. Similarly, in various cognitive architectures ToH/ToL tasks are used to evaluate strategy acquisition [Soar (Ruiz and Newell 1989), Teton (VanLehn 1989), CAPS (Just and Varma 2007), CLARION (Sun and Zhang 2002), ICARUS (Langley and Rogers 2008), ACT-R (Anderson and Douglass 2001)].

Given the importance of memory for even the most basic tasks, many phenomena related to different types of memory were also examined. To name a few, the following experiments were conducted: test of the effect of working memory capacity on the syntax parsing [CAPS (Just and Carpenter 1992)], N-back task testing spatial working memory [ACT-R (Kottlors et al. 2012)], reproduction of the Morris water maze test on a mobile robot to study the formation of episodic and spatial memory [BBD (Krichmar et al. 2005)], effect of priming on the speed of memory retrieval [CHREST (Lane et al. 2003)], effect of anxiety on recall [DUAL (Feldman and Kokinov 2009)], etc.

Attention has been explored both relative to perception and for general resource allocation. For example, models have been built to explain the well-known phenomena of inattentional blindness [ARCADIA (Bridewell and Bello 2016)] and attentional blink [LIDA (Madl and Franklin 2012)]. In addition, various dual task experiments were replicated to study and model the sources of attention limitation which reduces human ability to perform multiple tasks simultaneously. For example, CAPS proposes a functional account of attention allocation based on a dual-task experiment involving simultaneous sentence comprehension and mental rotation (Just et al. 2001), ACT-R models effects of sleep loss on sustained attention performance, which requires tracking a known location on the monitor and a reaction task (Gunzelmann et al. 2009). Similar experiments have been repeated using EPIC (Kieras 2012; Kieras and Meyer 1998).

Multiple experiments relating perception, attention, memory and learning have been simulated using the CHREST architecture in the context of playing chess. These include investigations of gaze patterns of novice and expert players (Lane et al. 2009), effects of ageing on chess playing related to the reduced capacity of working memory and decreased perceptual abilities (Smith et al. 2007) and effects of expertise, presentation time and working memory capacity on the ability to memorize random chess positions (Gobet 2008).

11.2.2 Robotics

Nearly a quarter of all applications of cognitive architectures are related to robotics. Much effort has been spent on navigation and obstacle avoidance, which are useful on their own and enable more complex behaviors. In particular, navigation in unstructured environments was implemented on an autonomous vehicle [RCS (Coombs et al. 2000), ATLANTIS (Gat 1992)], mobile robot [Subsumption (Brooks 1989), CoSy (Pacchierotti et al. 2005)] and unmanned marine vehicle [CARACaS (Huntsberger and Woodward 2011)].

The fetch and carry tasks used to be very popular in the early days of robotics research as an effective demonstration of robot abilities. Some well-known examples include a trash collecting mobile robot [3T (Firby et al. 1995)] and a soda can collecting robot [Subsumption (Brooks and Flynn 1989)]. Through a combination of simple vision techniques, such as edge detection and template matching, and sensors for navigation, these robots were able to find the objects of interest in unknown environments.

More recent cognitive architectures solve search and object manipulation tasks separately. Typically, experiments involving visual search are done in very controlled environments and preference is given to objects with bright colors or recognizable shapes to minimize visual processing, for example, a red ball [SASE (Weng and Zhang 2002)] or a soda can [ISAC (Kawamura et al. 1993)]. Sometimes markers, such as printed barcodes attached to the object, are used to simplify recognition [Soar (Mininger and Laird 2016)]. It is important to note that visual search in these cases is usually a part of a more involved task, such as learning by instruction. When visual search and localization are the end goal, the environments are more realistic [e.g. the robot controlled by CoSy finds a book on a cluttered shelf using a combination of sensors and SIFT features (López et al. 2008)].

Object manipulation involves arm control to reach and grasp an object. While reaching is a relatively easy problem and many architectures implement some form of arm control, gripping is more challenging even in a simulated environment. The complexity of grasping depends on many factors including the type of gripper and the properties of the object. One workaround is to experiment with grasping on soft objects, such as plush toys [ISAC (Kawamura et al. 2004)]. More recent work involves objects with different grasping types (objects with handles located on the top or on a side) demonstrated on a robot controlled by DIARC (Wilson et al. 2016). Another example is iCub adapting its grasp to cans of different sizes, boxes and a ruler (Sauser et al. 2012).

A few architectures implement multiple skills for complex scenarios such as a robotic salesman [CORTEX (Romero-Garcés et al. 2015b), RoboCog (Romero-Garcés et al. 2015a)], tutoring [DAC (Vouloutsi et al. 2015)], medical assessment [LIDA (Madl and Franklin 2015), RoboCog (Bandera et al. 2016)], etc. Industrial applications are represented by a single architecture—RCS, which has been used for teleoperated robotic crane operation (Lytle and Saidi 2007), bridge construction (Bostelman et al. 1999), autonomous cleaning and deburring workstation (Murphy et al. 1988), and the automated stamp distribution center for the US Postal Service (Albus 1997).

The biologically motivated architectures focus on the developmental aspect of physical skills and sensorimotor reconstruction. For example, a childlike iCub robot platform explores acquisition of skills for locomotion, grasping and manipulation (Albus 1994), robots Dav and SAIL learn vision-guided navigation [SASE (Weng and Zhang 2002)] and ISAC learns grasping affordances (Ulutas et al. 2008).

11.2.3 Human performance modeling (HPM)

Human performance modeling is an area of research concerned with building quantitative models of human performance in a specific task environment. The need for such models comes from engineering domains where the space of design possibilities is too large so that empirical assessment is infeasible or too costly.

This type of modeling has been used extensively for military applications, for example, workload analysis of Apache helicopter crew (Allender 2000), modeling the impact of communication tasks on the battlefield awareness (Mitchell et al. 2009), decision making in the AAW domain (Zachary et al. 1998), etc. Common civil applications include models of air traffic control task (Seamster et al. 1993), aircraft taxi errors (Wickens et al. 2008), 911 dispatch operator (Hart et al. 2001), etc.

Overall, HPM is dominated by a handful of specialized architectures, including OMAR, APEX, COGNET, MIDAS and IMPRINT. In addition, Soar was used to implement a pilot model for large-scale distributed military simulations [TacAir-Soar (Jones et al. 1999; Laird et al. 1998)].

11.2.4 Human-robot and human/computer interaction (HRI/HCI)

HRI is a multidisciplinary field studying various aspects of communication between people and robots. Many of these interactions are being studied in the context of social, assistive or developmental robotics. Depending on the level of autonomy demonstrated by the robot, interactions extend from direct control (teleoperation) to full autonomy of the robot enabling peer-to-peer collaboration. Although none of the systems presented in this survey are yet capable of full autonomy, they allow for some level of supervisory control ranging from single vowels signifying direction of movement for a robot [SASE (Weng et al. 1999)] to natural language instruction [Soar (Laird et al. 2004), HOMER (Vere and Bickmore 1990), iCub (Tikhanoff et al. 2011)]. It is usually assumed that a command is of particular form and uses a limited vocabulary.

Some architectures also target non-verbal aspects of HRI, for example, natural turn-taking in a dialogue [Ymir (Thorisson et al. 2010), Kismet (Breazeal 2003b)], changing facial expression [Kismet (Breazeal and Brooks 2004)] or turning towards the caregiver [MACSi (Anzalone et al. 2012)].

An important practical application which also involves HCI is in building decision support systems, i.e. intelligent assistants which can learn from and cooperate with experts to solve problems in complex domains. One such domain is intelligence analysis, which requires mining large amounts of textual information, proposing a hypothesis in search of evidence and reevaluating hypotheses in view of new evidence [Disciple (Tecuci et al. 2013)]. Other examples include time- and resource-constrained domains such as emergency response planning [Disciple (Tecuci et al. 2007a), NARS (Slam et al. 2015)], medical diagnosis [OSCAR (Pollock and Hosea 1995)], military operations in urban areas [R-CAST (Fan et al. 2010b)] and air traffic control [OMAR (Deutsch and Cramer 1998b)].

11.2.5 Natural language processing (NLP)

Natural language processing (NLP) is a broad multi-disciplinary area which studies understanding of written or spoken language. In the context of cognitive architectures, many aspects of NLP have been considered, from low-level auditory perception, syntactic parsing and semantics to conversation in limited domains.

As has been noted in the section on perception, there are only a few models of low-level auditory perception. For instance, models based on Adaptive Resonance Theory (ART), namely ARTPHONE (Grossberg 2003), ARTSTREAM (Grossberg 1999), ARTWORD (Grossberg and Myers 2015) and others, have been used to model perceptual processes involved in speech categorization, auditory streaming, source segregation (also known as the "cocktail party problem") and phonemic integration.

Otherwise, most research in NLP is related to investigating aspects of the syntactic and semantic processing of textual data. Some examples include anaphora resolution [PolyScheme (Kurup et al. 2011), NARS (Kilic 2015), DIARC (Williams and Scheutz 2016)], learning English passive voice [NARS (Kilic 2015)], models of syntactic and semantic parsing [SPA (Stewart et al. 2015), CAPS (Just and Carpenter 1992)] and word sense disambiguation [SemSoar and WordNet (Jones et al. 2016)].

One of the early models for real-time natural language processing, NL-Soar, combined syntactic knowledge and semantics of simple instructions for the immediate reasoning and tasks in blocks world (Lewis 1992). A more recent model (currently used in Soar) is capable of understanding commands, questions, syntax and semantics in English and Spanish (Lindes and Laird 2016). A complete model of human reading which simulated gaze patterns,

sequential and chronometric characteristics of human readers, semantic and syntactic analysis of sentences, recall and forgetting was built using the CAPS architecture (Thibadeau et al. 1982).

More commonly, off-the-shelf software is used for speech recognition and parsing, which helps to achieve a high degree of complexity and realism. For instance, a salesman robot [CORTEX (Bustos et al. 2016)] can understand and answer questions about itself using a Microsoft Kinect Speech SDK. The Playmate system based on CoSy uses dedicated software for speech processing (Lison and Kruijff 2008) and can have a meaningful conversation in a subset of English about colors and shapes of objects on the table. The FORR architecture uses speech recognition for the task of ordering books from the public library by phone. The architecture, in this case, increases the robustness of the automated speech recognition system based on an Olympus/RavenClaw pipeline (CMU) (Epstein et al. 2012).

In general, most existing NLP systems are limited both in their domain of application and in terms of the syntactic structures they can understand. Recent architectures, such as DIARC, aim at supporting more naturally sounding requests like Can you bring me something to cut a tomato?, however, they are still in the early stages of development (Sarathy et al. 2016).

11.2.6 Categorization and clustering

Categorization, classification, pattern recognition and clustering are common ways of extracting general information from large datasets. In the context of cognitive architectures, these methods are useful for processing noisy sensory data. Applications in this group are almost entirely implemented by the emergent architectures, such as ART and HTM, which are used as sophisticated neural networks. The ART networks, in particular, have been applied to classification problems in a wide range of domains: movie recommendations [Netflix dataset (Carpenter and Gaddam 2010)], medical diagnosis [Pima-Indian diabetes dataset (Kaylani et al. 2009)], fault diagnostics [pneumatic system analysis (Demetgul et al. 2009)], vowel recognition [Peterson and Barney dataset (Ames and Grossberg 2008)], odor identification (Distante et al. 2000), etc. The HTM architecture is geared more towards the analysis of time series data, such as predicting IT failures,[18] monitoring stocks,[19] predicting taxi passenger demand (Cui et al. 2015) and recognition of cell phone usage type (email, call, etc.) based on the pressed key pattern (Melis et al. 2009).

A few other examples from the non-emergent architectures include gesture recognition from tracking suits [Ymir (Cassell and Thorisson 1999)], diagnosis of the failures in a telecommunications network [PRS (Rao and George 1991)] and document categorization based on the information about authors and citations [CogPrime (Harrigan et al. 2014)].

11.2.7 Computer vision

Emergent cognitive architectures are widely applied to solving typical computer vision problems. However, these are mainly standalone examples, such as hand-written character recognition [HTM (Stolc and Bajla 2010; Thornton et al. 2008)], image classification benchmarks [HTM (Mai et al. 2013; Zhuo et al. 2012)], view-invariant letter recognition [ART (Fazl et al. 2009)], texture classification benchmarks [ART (Wang et al. 1997)], invariant object recognition [Leabra (O'Reilly et al. 2014)], etc.

---

[18] http://grokstream.com.

[19] http://numenta.com/htm-for-stocks.

Computer vision applications that are part of more involved tasks, such as navigation in robotics, are discussed in the corresponding sections.

11.2.8 Games and puzzles

Playing games has been an active area of research in cognitive architectures for decades. Some of the earliest models for playing tic-tac-toe and Eight Puzzle as a demo of reasoning and learning abilities were created in the 1980s [Soar (Laird et al. 1984)]. The goal is typically not to master the game but rather to use it as a step towards solving similar but more complex problems. For example, Liar's Dice, a multi-player game of chance, is used to assess the feasibility of reinforcement learning in large domains [Soar (Derbinsky and Laird 2012a)]. Similarly, playing Backgammon was used to model cognitively plausible learning in [ACT-R (Sanner et al. 2000)] and tic-tac-toe to demonstrate ability to learn from instruction [Companions (Hinrichs and Forbus 2014)]. Multiple two-player board games with conceptual overlap like tic-tac-toe, the Eight Puzzle and the Five Puzzle can also be used as an effective demonstration of knowledge transfer [e.g. Soar (Kirk and Laird 2016), FORR (Epstein 2001)].

Compared to the classic board games used in AI research since its inception, video games provide a much more varied and challenging domain. Like board games, most video games require certain cognitive skills from a player, thus allowing the researchers to break down the problem of solving general intelligence into smaller chunks and work on them separately. In the past years, graphics, complexity and response times of the video games have greatly improved, hence many games can be used as sensible approximations of the real world environments. In addition to that, a simulated intelligent entity is easier to develop and less prone to damage than the one embodied in a physical robot. A combination of these factors makes video games a very valuable test platform for models of human cognition.

The only drawback of using video games for research is that embedding a cognitive architecture within it requires software engineering work. Naturally, games with open source engines and readily available middleware are preferred. One such example is the Unreal Tournament 2004 (UT2004) game, for which the Pogamut architecture (Kadlec et al. 2009) serves as a middleware, making it easier to create intelligent virtual characters. Although Pogamut itself implements many cognitive functions, it is also used with modifications by other groups to implement artificial entities for UT2004 (Cuadrado and Saez 2009; Mora et al. 2015; Small and Congdon 2009; Van Hoorn et al. 2009; Wang et al. 2009). Other video games used in cognitive architectures research are Freeciv [REM (Ulam et al. 2004)], Atari Frogger II [Soar (Wintermute 2012)], Infinite Mario [Soar (Shiwali and Laird 2009)], browser games [STAR (Kotseruba 2016)] and custom made games [Soar (Marinier et al. 2009)]. It should be noted that aside from the playing efficiency and achieved scores the intelligent agents are also evaluated based on their believability (e.g. 2K BotPrize Contest[20]).

11.2.9 Virtual agents

Although related to human performance modeling, the category of virtual agents is broader. While HPM requires models which can closely model human behavior in precisely defined conditions, the goals of creating virtual agents are more varied. Simulations and virtual reality are frequently used as an alternative to the physical embodiment. For instance, in the military domain, simulations model behavior of soldiers in dangerous situations without

---

[20] http://botprize.org.

risking their lives. Some examples include modeling agents in a suicide bomber scenario [CoJACK (Evertsz et al. 2009)], peacekeeping mission training [MAMID (Hudlicka 2002)], command and control in complex and urban terrain [R-CAST (Fan et al. 2006)] and tank battle simulation [CoJACK (Ritter et al. 2012)].

Simulations are also common for modeling behaviors of intelligent agents in civil applications. One of the advantages of virtual environments is that they can provide information about the state of the agent at any point in time. This is useful for studying the effect of emotions on actions, for example, in the social interaction context [ARS/SiMA (Schaat et al. 2013a)], or in learning scenarios, such as playing fetch with a virtual dog [Novamente (Heljakka et al. 2007)].

Intelligent characters with engaging personalities can also enhance user experience and increase their level of engagement in video games and virtual reality applications. One of the examples is virtual reality drama "Human Trials" which lets human actors participate in an immersive performance together with multiple synthetic characters controlled by the GLAIR architecture [Anstey et al. 2007; Shapiro et al. 2005)]. A similar project called Virtual Theater allowed users to interact with virtual characters on the screen to improvise new stories in real time. The story unfolds as users periodically select among directional options appearing on the screen [AIS (Hayes-Roth and Gent 1995)].

12 Discussion

The main contribution of this survey is in gathering and summarizing information on a large number of cognitive architectures influenced by various disciplines (computer science, cognitive psychology, philosophy and neuroscience). In particular, we discuss common approaches to modeling important elements of human cognition, such as perception, attention, action selection, learning, memory and reasoning. In our analysis, we also point out what approaches are more successful in modeling human cognitive processes or exhibiting useful behavior. In order to evaluate practical aspects of cognitive architectures we categorize their existing practical applications into several broad categories. Furthermore, we map these practical applications to a number of competency areas required for human-level intelligence to assess the current progress in that direction. This map may also serve as an approximate indication of what major areas of human cognition have received more attention than others. Thus, there are two main outcomes of this review. First, we present a broad and inclusive snapshot of the progress made in cognitive architectures research over the past four decades. Second, by documenting the wide variety of tested mechanisms that may help in developing explanations and models for observed human behavior, we inform future research in cognitive science and in the component disciplines that feed it. In the remainder of this section we will summarize the main themes discussed in the present survey and pertaining issues which can guide directions for future research.

Approaches to modeling human cognition Historically, psychology and computer science were inspirations for the first cognitive architectures, i.e. the theoretical models of human cognitive processes and corresponding software artifacts that allowed demonstrating and evaluating the underlying Theory of Mind. Agent architectures, on the other hand, focus on reproducing a desired behavior and particular application demands without the constraints of cognitive plausibility.

Even though they are often described as opposing, in practice the the two paradigms are not as separated in practice (which causes inconsistencies in the survey literature as discussed

in Sect. 2). Despite the differences in the theory and terminology both cognitive and agent architectures have functional modules corresponding to human cognitive abilities and tackle the issues of action selection, adaptive behavior, efficient data processing and storage. For example, methods of action selection widely used in robotics and classic AI, ranging from priority queue to reinforcement learning (Pirjanian 1999), are also found in many cognitively-oriented architectures. Likewise, some agent architectures incorporate cognitively plausible structures and mechanisms.

Biologically and neurally plausible models of human mind primarily aim to explain how known cognitive and behavioral phenomena arise from low-level brain processes. It has been argued that their support for inference and general reasoning is inadequate for modeling all aspects of human cognition, although some of the latest neuronal models demonstrate that it is far from being proven.

The newest paradigm in AI is represented by machine learning methods, particularly data-driven deep learning, which have found enormous practical success in limited domains and even claim some biological plausibility. The bulk of work in this area is on perception although there have been attempts at implementing more general inference and memory mechanisms. The results of these efforts are stand-alone models that have yet to be incorporated within a unified framework. At present these new techniques are not widely incorporated into existing cognitive architectures.

Range of supported cognitive abilities and remaining gaps The list of cognitive abilities and phenomena covered in this survey is by no means exhaustive. However, none of the systems we reviewed is close to supporting in theory or demonstrating in practice even this restricted subset, let alone a set of identified cognitive abilities (a comprehensive survey by Carroll (1993) lists nearly 3000 of them).

Most effort thus far has been dedicated to studying high-level abilities such as action selection (reactive, deliberative and, to some extent, emotional), memory (both short and long-term), learning (declarative and non-declarative) and reasoning (logical inference, probabilistic and meta-reasoning). Decades of work in these areas, both in traditional AI research and cognitive architectures, resulted in creation of multiple algorithms and data structures. Furthermore, various combinations of these approaches have been theoretically justified and practically validated in different architectures, albeit under restricted conditions.

We only briefly cover abilities that rely on the core set of perception, action-selection, memory and learning. At present, only a fraction of the architectures have the theoretical and software/hardware foundation to support creative problem solving, communication and social interaction (particularly in the fields of social robotics and HCI/HRI), natural language understanding and complex motor control (e.g. grasping).

With respect to the range of abilities and realism of the scenarios there remains a significant gap between the state-of-the-art in specialized areas of AI and the research in these areas within the cognitive architectures domain. Arguably some of the AI algorithms demonstrate achievements that are already on par with or even exceed human abilities [e.g. in image classification (He et al. 2014), face recognition (Taigman et al. 2014), playing simple video games (Mnih et al. 2015)], which raises the bar for cognitive architectures.

Due to the complexity of modeling the human mind, groups of researchers have been pursuing complementary lines of investigation. Owing to tradition, work on high-level cognitive architectures is still centered mainly on reasoning and planning in simulated environments, downplaying issues of cognitive control of perception, symbol grounding and realistic attention mechanisms. On the other hand, embodied architectures must deal with the real-world constraints and thus pay more attention to low-level perception-motor coupling, while most of the deliberative planning and reasoning effort is spent on navigation and motor control.

There are, however, trends towards overcoming these differences as classical cognitive architectures experiment with embodiment (Mohan et al. 2012; Trafton et al. 2013) or are being interfaced with robotic architectures (Scheutz et al. 2013) to utilize their respective strengths.

There also remains a gap between the biologically inspired architectures and neural simulations and the rest of the architectures. Even though low-level models can demonstrate how neural processes give rise to some higher-level cognitive functions, they do not have the same range and efficiency in practical applications compared to the less theoretically restricted systems. Some exceptions exist, for example Grok—a commercial application for IT analytics based on the biologically inspired HTM architecture or Neural Information Retrieval System implemented as a hierarchy of ART networks for storing 2D and 3D parts designs built for the Boeing company (Smith et al. 1997).

Finally, most of the featured architectures cannot reuse the capabilities or accumulate knowledge as they are applied to new tasks. Instead, every new task or skill is demonstrated using a separate model, specific set of parameters or knowledge base. [21] Examples of architectures capable of switching between various scenarios/environments without resetting or changing parameters are sparse and are explicitly declared as such, e.g. the SPA architecture which can perform eight unrelated tasks (Eliasmith and Stewart 2012).

Outstanding issues and future directions One of the outcomes of this survey, besides summarizing past and present attempts at modeling various human cognitive abilities, is highlighting the problems needing further research via interactive visualization. We also reviewed future directions suggested by scholars in a large number of publications in search of problems deemed important. We present some of the main themes below, excluding architecture- and implementation-specific issues.

Adequate experimental validation A large number of papers end in a call for thorough experimental testing of the given architecture in more diverse, challenging and realistic environments (Firby et al. 1995; Nunez et al. 2016; Rohrer 2012) and real-world situations (Sun and Helie 2015) using more elaborate scenarios (Rousseau and Hayes-Roth 1996) and diverse tasks (Herd et al. 2013). This is by far the most pressing and long-standing issue [found in papers from the early 90s (Matthies 1992) up until 2017 (Fedor et al. 2017)] which remains largely unresolved. The data presented in this survey (e.g. Sect. 11.2 on practical applications) shows that, by and large, the architectures (including those vying for AGI) have practical achievements only in several distinct areas and in highly controlled environments. Likewise, common validation methods have many limitations. For instance, replicating psychological experiments, the most frequent method in our sample of architectures, presents the following challenges: it is usually based on a small set of data points from human participants, the amount of prior knowledge for the experiment is often unknown and modeled in an ad hoc manner, pre-processing of stimuli is required in case the architecture does not posess adequate perception, and, furthermore, the task definition is narrow and tests only a particular aspect of the ability (Jones et al. 2012). Overall, such restricted evaluation tells us little about the actual abilities of the system outside the laboratory setting.

Realistic perception The problem of perception is most acute in robotic architectures operating in noisy unstructured environments. Frequently mentioned issues inlcude the lack of active vision (Huber and Kortenkamp 1995), accurate localization and tracking (Wolf et al. 2010), robust performance under noise and uncertainty (Romero-Garcés et al. 2015a) and the utilization of context information to improve detection and localization (Christensen

---

[21] Note that the proposed visualizations demonstrate aggregated statistics and do not imply that there exists a single model that encapsulates them all.

and Kruijff 2010). More recently, deep learning is being considered as a viable option for bottom-up perception (Bona 2013; Ivaldi et al. 2014).

These concerns are supported by our data as well. Overall, in comparison to higher-level cognitive abilities, the treatment of perception, and vision in particular, is rather superficial, both in terms of capturing the underlying processes and practical applications. For instance, almost half of the reviewed architectures do not implement any vision (Fig. 5). The remaining projects, with few exceptions, either default to simulations or operate in controlled environments.

Other sensory modalities, such as audition, proprioception and touch, often rely on off-the-shelf software solutions or are trivially implemented via physical sensors/simulations. Multi-modal perception in the architectures that support it, is approached mechanistically and is tailored for a particular application. Finally, bottom-up and top-down interactions between perception and higher-level cognition are largely overlooked both in theory and practice. As discussed in Sect. 5, many of the mechanisms are a by-product of other design decisions rather than a result of deliberate efforts.

Human-like learning Even though various types of learning are represented in the cognitive architectures, there is still a need for developing more robust and flexible learning mechanisms (Hinrichs and Forbus 2014; Menager and Choi 2016), knowledge transfer (Herd et al. 2013) and accumulation of new knowledge without affecting prior learning (Jonsdottir and Thórisson 2013). Learning is especially important for developmental approach to modeling cognition, specifically motor skills and new declarative knowledge (as discussed in Sect. 8). However, artificial development is restricted and far less efficient because visual and auditory perception in robots is not yet on par with even the small children (Christensen and Kruijff 2010).

Natural communication Verbal communication is the most common mode of interaction between the artificial agents and humans. Many issues are yet to be resolved as current approaches do not possess a sufficiently large knowledge base for generating dialogues (Faghihi et al. 2013) and generally lack robustness (Hinrichs and Forbus 2014; Williams et al. 2015). Few, if any, architectures are capable of detecting the emotional state and intentions of the interlocutor (Breazeal 2003a) or give personalized responses (Gobet and Lane 2012). Non-verbal aspects of communication, such as performing and detecting gestures and facial expressions, natural turn-taking, etc., are investigated by only a handful of architectures, many of which are no longer in development. As a result, demonstrated interactions are heavily scripted and often serve as voice control interfaces for supplying instructions instead of enabling collaboration between human and the machine (Myers et al. 2002) (see also discussion in Sect. 11.2.5).

Autobiographic memory Even though memory, a necessary component of any computational model, is well represented and well researched in the domain of cognitive architectures, episodic memory is comparatively under-examined (Bölöni 2012; Borrajo et al. 2015; Danker and Anderson 2010; Menager and Choi 2016). The existence of this memory structure has been known for decades (Tulving 1972) and its importance for learning, communication and self-reflection is widely recognized, however episodic memory remains relatively neglected in computational models of cognition (see Sect. 8). This is also true for general models of human mind such as Soar and ICARUS, which included it relatively recently.

Currently the majority of the architectures save past episodes as time-stamped snapshots of the state of the system. This approach is suitable for agents with a short life-span and a limited representational detail but more elaborate solutions will be needed for life-long learning and managing large-scale memory storage (Laird and Derbinsky 2009). The issues

of scale are relevant to other types of long-term memory as well (Bellas et al. 2014; Henderson et al. 2012; Murdock and Goel 2008).

Computational performance The problem of computational efficiency, for obvious reasons, is more pressing in robotic architectures (Gat and Dorais 1994; Wilson and Scheutz 2014) and interactive applications (Arrabales Moreno and Sanchis de Miguel 2006; Ash and Hayes-Roth 1996). However, it is also reported for non-embodied architectures (Jones et al. 2016; Sun 2016), particularly neural simulations (Tripp and Eliasmith 2016; Wendelken 2003). Overall, time and space complexity is glossed over in the literature, frequently described in qualitative terms (e.g. 'real-time' without specifics) or omitted from discussion altogether.

A related point is an issue of scale. It has been shown that scaling up existing algorithms to larger amounts of data introduces additional challenges, such as the need for more efficient data processing and storage. For example, the difficulties associated with maintaining a knowledge base equivalent to human brain memory-storage capacity are discussed in Sect. 7. These problems are currently addressed by only a handful of architectures, even though solving them is crucial for further development of theoretical and applied AI.

In addition to concerns collectively expressed by the researches (and supported by the data in this survey) below we examine two more observations: (1) the need for an objective evaluation of the results and measuring the overall progress towards understanding and replicating human-level intelligence and (2) reproducibility of the research in the cognitive architectures.

Comparative evaluation of cognitive architectures The resolution of the commonly acknowledged issues hinges on the development of objective and extensive evaluation procedures. We have already mentioned as one of the issues that individual architectures are validated using a disjointed and sparse set of experiments, which cannot support claims about the abilities of the system or adequately assess its representational power. However, comparisons between architectures and measuring overall progress in the field entails an additional set of challenges.

Surveys are a first step towards mapping the landscape of approaches to modeling human intelligence, but they can only identify the blank spots needing further exploration and are ill-suited for comparative analysis. For instance, in this paper, visualizations show only the existence of any particular feature, be it memory, learning method or perception modality. However, they cannot represent the extent of the ability nor whether it is enabled by the architectural mechanisms or is an ad hoc implementation.[22] Thus, any comparisons between the architectures are discouraged and any such conclusions should be taken with caveats.

A compounding factor for evaluation is that cognitive architectures comprise both theoretical and engineering components, which leads to a combinatorial explosion of possible implementations. Consider, for example, the theoretical and technical nuances of various instantiations of the global workspace concept, e.g. ARCADIA (Bridewell and Bello 2015), ASMO (Novianto et al. 2010) and LIDA (Franklin et al. 2012) or the number of planning algorithms and approaches to dynamic action selection in Sect. 6. Furthermore, building a cognitive architecture is a major commitment, requiring years of development to put even the basic components together (average age of projects in our sample is $\approx$ 15 years), making any major changes harder with time. Thus, evaluation is crucial for identifying more promising approaches and pruning suboptimal paths for exploration. To some extent, it has already been happening organically via a combination of theoretical considerations and collective experience gathered over the years. For instance, symbolic representation, once dominant, is being

---

[22] Refer to a relevant discussion in Sect. 7 on how working memory size is set in various architectures.

replaced by more flexible emergent approaches (see Fig. 2) or augmented with sub-symbolic elements (e.g. Soar, REM). Clearly, more consistent and directed effort is needed to make this process faster and more efficient.

Ideally, proper evaluation should target every aspect of cognitive architectures using a combination of theoretical analysis, software testing techniques, benchmarking, subjective evaluation and challenges. All of these have already been sparingly applied to individual architectures. For example, a recent theoretical examination of the knowledge representations used in major cognitive architectures reveals their limitations and suggests ways to amend them (Lieto 2016; Lieto et al. 2018a). Over the years, a number of benchmarks and challenges have been developed in other areas of AI and a few have already been used to evaluate cognitive architectures, e.g. 2K BotPrize competition [Pogamut (Gemrot et al. 2010)], various benchmarks, including satellite image classification [ART (Carpenter and Gaddam 2010)], medical data analysis [ART (Carpenter and Milenova 2000)] and face recognition [HTM (Stolc and Bajla 2010)], as well as robotics competitions such as AAAI Robot Competition (DIARC (Schermerhorn et al. 2006) and RoboCup [BBD (Szatmary et al. 2006)].

Tests probing multiple abilities via series of tasks are suggested for evaluating cognitive architectures and general-purpose AI systems. Some examples include, "Cognitive Decathlon" (Mueller and Minnery 2008) and I-Athlon (Adams et al. 2016), both inspired by the Turing test, as well as Catell-Horn-Carroll model examining factors of intelligence (Ichise 2016) or a battery of tests for assessing different cognitive dimensions proposed by Samsonovich et al. (Samsonovich et al. 2006). However, most of the work in this area remains theoretical and has not resulted in solid proposals ready to be implemented or complete frameworks.

For further discussion of ample literature concerning evaluation of AI and cognitive architectures we refer the readers to a very recent and comprehensive review of these methods in Hernandez-Orallo (2017).

Reproducibility of the results Literature on cognitive architectures gives far more importance to the cognitive, psychological or philosophical aspects of modeling human cognition, while technical implementation details are often incomplete or missing. For example, in Sect. 5 on visual processing we list architectures that only briefly mention some perceptual capabilities but are not specific enough for analysis. Likewise, justifications of using particular algorithms/representations along with their advantages and limitations are not always disclosed. Such lack of full technical detail compromises reproducibility of the research.

In part, this issue can be alleviated by providing access to the software/source code, but only one-third of the architectures do so. To a lesser extent this applies to architectures such as BBD, SASE, Kismet, Ymir, Subsumption and a few others that are physically embodied and depend on a particular robotic platform. However, the majority of architectures we reviewed have a substantial software component, releasing which could be of benefit to the community and the researchers themselves. For instance, cognitive architectures such as ART, ACT-R, Soar, HTM and Pogamut, are used by many researchers outside the main developing group.

In conclusion, our hope is that this work will serve as an overview and a guide to the vast field of cognitive architecture research as we strived to objectively and quantitatively assess the state-of-the-art in modeling human cognition.

References

Adams S, Arel I, Bach J, Coop R, Furlan R, Goertzel B, Hall JS, Samsonovich A, Scheutz M, Schlesinger M, Shapiro SC, Sowa J (2012) Mapping the landscape of human-level artificial general intelligence. AI Mag 33(1):25–42

Adams SS, Banavar G, Campbell M (2016) I-athlon: toward a multidimensional turing test. AI Mag 31(1):78–84

Albus J, Barbera A (2006) Intelligent control and tactical behavior development: a long term NIST partnership with the army. In: 1st joint emergency preparedness and response/robotic and remote systems topical meeting

Albus J, Lacaze A, Meystel A (1995) Theory and experimental analysis of cognitive processes in early learning. In: Proceedings of the IEEE international conference on systems, man and cybernetics, pp 4404–4409

Albus J, Bostelman R, Hong T, Chang T, Shackleford W, Shneier M (2006) THE LAGR PROJECT integrating learning into the 4D/RCS control hierarchy. In: International conference in control, automation and robotics

Albus JS (1994) A reference model architecture for intelligent systems design. In: An introduction to intelligent and autonomous control, pp 27–56

Albus JS (1997) The NIST real-time control system (RCS): an approach to intelligent systems research. J Exp Theor Artif Intell 9(2–3):157–174

Albus JS (2002) 4D/RCS a reference model architecture for intelligent unmanned ground vehicles. In: Proceedings of the SPIE 16th annual international symposium on aerospace/defense sensing, simulation and controls

Albus JS, Barbera AJ (2005) RCS: a cognitive architecture for intelligent multi-agent systems. Annu Rev Control 29(1):87–99

Albus JS, Hui-Min Huang, Messina ER, Murphy K, Juberts M, Lacaze A, Balakirsky SB, Shneier MO, Hong TH, Scott Ha, Proctor FM, Shackleford WP, Michaloski JL, Wavering AJ, Kramer TR, Dagalakis NG, Rippey WG, Stouffer Ka, Legowik S (2002) 4D/RCS: a reference model architecture for unmanned vehicle systems version 2.0. In: Proceedings of the SPIE 16th annual international symposium on aerospace defense sensing simulation and controls

Allender L (2000) Modeling human performance: impacting system design, performance, and cost. In: Proceedings of the military, government and aerospace simulation symposium, pp 139–144

Ames H, Grossberg S (2008) Speaker normalization using cortical strip maps: a neural model for steady-state vowel categorization. J Acoust Soc Am 124:3918–3936

Anderson JR, Douglass S (2001) Tower of Hanoi: evidence for the cost of goal retrieval. J Exp Psychol Learn Mem Cogn 27(6):1331–1346

Anderson JR, Lebiere C (2003) The Newell test for a theory of cognition. Behav Brain Sci 26(5):587–601

Anderson JR, Reder LM, Lebiere C (1996) Working memory: activation limitations on retrieval. Cogn Psychol 30:221–256

Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y (2004) An integrated theory of the mind. Psychol Rev 111(4):1036–1060

Anderson JR, Albert MV, Fincham JM (2005) Tracing problem solving in real time: fMRI analysis of the subject-paced Tower of Hanoi. J Cogn Neurosci 17(8):1261–1274

Anstey J, Bay-Cheng S, Pape D, Shapiro SC (2007) Human trials: an experiment in intermedia performance. ACM Comput Entertain 5(3):4

Anzalone SM et al (2012) Multimodal people engagement with iCub. In: Chella A, Pirrone R, Sorbello R, Jóhannsdóttir KR (eds) Biologically inspired cognitive architectures 2012: proceedings of the third annual meeting of the BICA Society. Springer, Berlin

Arrabales R, Ledezma A, Sanchis A (2009a) A cognitive approach to multimodal attention. J Phys Agents 3(1):53–63

Arrabales R, Ledezma A, Sanchis A (2009b) CERA-CRANIUM: a test bed for machine consciousness research. In: International workshop on machine consciousness

Arrabales R, Ledezma A, Sanchis A (2009c) Towards conscious-like behavior in computer game characters. In: 2009 IEEE symposium on computational intelligence and games, pp 217–224. https://doi.org/10.1109/CIG.2009.5286473

Arrabales R, Ledezma A, Sanchis A (2011) Simulating visual qualia in the CERA-CRANIUM cognitive architecture. In: Hernández C, Sanz R, Gómez Ramirez J, Smith LS, Hussain A, Chella A, Aleksander I (eds) From brains to systems. Springer, New York, pp 239–250

Arrabales Moreno R, Sanchis de Miguel A (2006) A machine consciousness approach to autonomous mobile robotics. In: Proceedings of the 5th international cognitive robotics workshop

Ash D, Hayes-Roth B (1996) Using action-based hierarchies for real-time diagnosis. Artif Intell 88:317–347

Asselman A, Aammou S, Nasseh AE (2015) Comparative study of cognitive architectures. Int Res J Comput Sci 2(9):8–13

Atkinson RC, Shiffrin RM (1968) Human memory: a proposed system and its control processes. Psychology of learning and motivation. Adv Res Theory 2(1):89–195

Baars BJ (2005) Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. Prog Brain Res 150:45–53

Baars BJ, Ramamurthy U, Franklin S (2007) How deliberate, spontaneous and unwanted memories emerge in a computational model of consciousness. In: Mace JH (ed) Involuntary memory. Blackwell Publishing Ltd, pp 177–207

Bach J (2007) Principles of synthetic intelligence. PhD Thesis

Bach J (2011) A motivational system for cognitive AI. In: International conference on artificial general intelligence, pp 232–242

Bach J (2015) Modeling motivation in MicroPsi 2. In: International conference on artificial general intelligence, pp 3–13

Bach J, Bauer C, Vuine R (2007) MicroPsi: contributions to a broad architecture of cognition. In: Annual conference on artificial intelligence, pp 7–18

Bachiller P, Bustos P, Manso LJ (2008) Attentional selection for action in mobile robots. In: Advances in robotics, automation and control, pp 111–136

Baddeley AD, Hitch G (1974) Working memory. Psychology of learning and motivation. Adv Res Theory 8(C):47–89

Bandera A, Bustos P (2013) Toward the development of cognitive robots. In: International workshop on brain-inspired computing

Bandera A, Bandera JP, Bustos P, Calderita LV, Fern F, Fuentetaja R, Garc FJ, Iglesias A, Luis J, Marfil R, Pulido C, Reuther C, Romero-Garces A, Suarez C (2016) CLARC: a robotic architecture for comprehensive geriatric assessment. In: Proceedings of the WAF2016

Barnes A, Hammell RJ (2008) Determining information technology project status using recognition-primed decision-making enabled collaborative agents for simulating teamwork (R-CAST). In: Proceedings of the conference on information systems applied research (CONISAR)

Barslanrugnl BA, Nataatgenrugnl NT, Verbrugge R (2004) Modeling developmental transitions in reasoning about false beliefs of others

Bartol TM, Bromer C, Kinney JP, Chirillo MA, Bourne JN, Harris KM, Sejnowski TJ (2015) Nanoconnectomic upper bound on the variability of synaptic plasticity. eLife 4:e10778

Bellas F, Duro RJ (2004) Some thoughts on the use of sampled fitness functions for the multilevel Darwinist brain. Inf Sci 161(3–4):159–179

Bellas F, Becerra JA, Duro RJ (2005) Induced behavior in a real agent using the multilevel Darwinist brain. In: International work-conference on the interplay between natural and artificial computation, pp 425–434

Bellas F, Becerra JA, Duro RJ (2006) Some experimental results with a two level memory management system in the multilevel Darwinist brain. In: Proceedings of the European symposium on artificial neural networks, computational intelligence and machine learning

Bellas F, Duro RJ, Faiña A, Souto D (2010) Multilevel Darwinist brain (MDB): artificial evolution in a cognitive architecture for real robots. IEEE Trans Auton Ment Dev 2(4):340–354

Bellas F, Caamano P, Faina A, Duro RJ (2014) Dynamic learning in cognitive robotics through a procedural long term memory. Evol Syst 5(1):49–63. https://doi.org/10.1007/s12530-013-9079-4

Bello P, Bridewell W, Wasylyshyn C (2016) Attentive and pre-attentive processes in multiple object tracking: a computational investigation modeling object construction and tracking. In: Proceedings of the 38th annual meeting of the cognitive science society

Benjamin DP, Lyons D (2010) A cognitive approach to classifying perceived behaviors. In: Proceedings of the SPIE 7710, multisensor, multisource information fusion: architectures, algorithms, and applications, vol 7710

Benjamin DP, Lyons D, Lonsdale D (2004) ADAPT: a cognitive architecture for robotics. In: Proceedings of the sixth international conference on cognitive modeling (October), pp 337–338

Benjamin DP, Funk C, Lyons D (2013) A cognitive approach to vision for a mobile robot. In: SPIE defense, security, and sensing

Bida M, Cerny M, Gemrot J, Brom C (2012) Evolution of GameBots project. In: International conference on entertainment computing, pp 397–400

Boicu C, Tecuci G, Boicu M (2005) Improving agent learning through rule analysis. In: Proceedings of the international conference on artificial intelligence

Boicu M, Marcu D, Boicu C, Stanescu B (2003) Mixed-initiative control for teaching and learning in disciple. In: Proceedings of the IJCAI-03 workshop on mixed-initiative intelligent systems

Bölöni L (2012) An investigation into the utility of episodic memory for cognitive architectures. In: AAAI fall symposium: advances in cognitive systems

Bona JP (2013) MGLAIR: a multimodal cognitive agent architecture. PhD Thesis

Bonasso RP, Kortenkamp D (1996) Using a layered control architecture to alleviate planning with incomplete information. In: Proceedings of the AAAI spring symposium\planning with incomplete information for robot problems

Bonasso RP, Firby RJ, Gat E, Kortenkamp D, Miller DP, Slack MG (1997) Experiences with an architecture for intelligent, reactive agents. J Exp Theor Artif Intell 2–3:187–202

Borrajo D, Roubíčková A, Serina I (2015) Progress in case-based planning. ACM Comput Surv 47(2):35

Bostelman R, Hong T, Chang T, Shackleford W, Shneier M (2006) Unstructured facility navigation by applying the NIST 4D/RCS architecture. In: Proceedings of international conference on cybernetics and information technologies, systems and applications, pp 328–333

Bostelman RV, Jacoff A, Bunch R (1999) Delivery of an advanced double-hull ship welding. In: Third international ICSC (international computer science conventions) symposia on intelligent industrial automation and soft computing

Breazeal C (2003a) Emotion and sociable humanoid robots. Int J Hum Comput Stud 59(1–2):119–155

Breazeal C (2003b) Toward sociable robots. Robot Auton Syst 42(3–4):167–175

Breazeal C, Aryananda L (2002) Recognition of affective communicative intent in robot-directed speech. Auton Robot 12(1):83–104

Breazeal C, Brooks R (2004) Robot emotion: a functional perspective. In: Fellous JM, Arbib MA (eds) Who needs emotions? The brain meets the robot. Oxford University Press, Oxford, pp 271–310

Breazeal C, Fitzpatrick P (2000) That certain look: social amplification of animate vision. In: Proceedings of AAAI 2000 fall symposium, pp 18–22

Breazeal C, Scassellati B (1999) A context-dependent attention system for a social robot. IJCAI Int Joint Conf Artif Intell 2:1146–1151

Breazeal C, Scassellati B (2002) Challenges in building robots that imitate people. In: Dautenhahn K, Nehaniv C (eds) Imitation in animals and artifacts. MIT Press, Cambridge, pp 363–390

Breazeal C, Edsinger A, Fitzpatrick P, Scassellati B (2000) Social constraints on animate vision. In: Proceedings of the HUMANOIDS

Breazeal C, Edsinger A, Fitzpatrick P, Scassellati B (2001) Active vision for sociable robots. IEEE Trans Syst Man Cybern Part A Syst Hum 31(5):443–453

Bresina JL, Drummond M (1990) Integrating planning and reaction: a preliminary report. In: Proceedings of AAAI spring symposium on planning in uncertain, unpredictable, or changing environments. NASA Ames Research Center

Brett BE, Doyal JA, Malek DA, Martin EA, Hoagland DG, Anesgart MN (2002) The combat automation requirements testbed (CART) task 5 interim report: modeling a strike fighter pilot conducting a time critical target mission. Technical report

Brick T, Schermerhorn P, Scheutz M (2007) Speech and action: integration of action and language for mobile robots. In: Proceedings of IEEE international conference on intelligent robots and systems, pp 1423–1428

Bridewell W, Bello PF (2015) Incremental object perception in an attention-driven cognitive architecture. In: Proceedings of the 37th annual meeting of the cognitive science society, pp 279–284

Bridewell W, Bello PF (2016) Inattentional blindness in a coupled perceptual-cognitive system. In: Proceedings of the 38th annual meeting of the cognitive science society

Brom C, Pešková K, Lukavsky J (2007) What does your actor remember? Towards characters with a full episodic memory. In: International conference on virtual storytelling, pp 89–101

Brooks RA (1987) Planning is just a way of avoiding figuring out what to do next. Technical report working paper 303

Brooks RA (1989) A robot that walks; emergent behaviors from a carefully evolved network. Neural Comput 1(2):253–262

Brooks RA (1990) Elephants don't play chess. Robot Auton Syst 6(1–2):3–15

Brooks RA, Flynn AM (1989) Robot beings. In: Proceedings of the IEEE/RSJ international workshop on intelligent robots and systems, pp 2–10

Bruce N, Tsotsos J (2007) Attention based on information maximization. J Vis 7:950–950

Bustos P, Martinez-Gomez J, Garcia-Varea I, Rodriguez-Ruiz L, Bachiller P, Calderita L, Manso LJ, Sanchez A, Bandera A, Bandera JP (2013) Multimodal interaction with Loki. In: workshop of physical agents

Bustos P, Manso LJ, Bandera JP, Romero-Garcés A, Calderita LV, Marfil R, Bandera A (2016) A unified internal representation of the outer world for social robotics. In: Proceedings of the second Iberian robotics conference

Butt AJ, Butt NA, Mazhar A, Khattak Z, Sheikh JA (2013) The soar of cognitive architectures. In: Proceedings of the international conference on current trends in information technology

Byrne F (2015) Symphony from synapses: neocortex as a universal dynamical systems modeller using hierarchical temporal memory. arXiv preprint arXiv:151205245

Cao S, Qin Y, Zhao L, Shen M (2015) Modeling the development of vehicle lateral control skills in a cognitive architecture. Transp Res Part F Traffic Psychol Behav 32:1–10

Carbonell JG, Blythe J, Etzioni O, Gil Y, Joseph R, Kahn D, Knoblock C, Minton S, Alicia P, Reilly S, Veloso M, Wang X (1992) PRODIGY4.0: the manual and tutorial. Technical report CMU-CS-92-150

Carpenter GA (2001) Neural-network models of learning and memory: leading questions and an emerging framework. TRENDS Cogn Sci 5(3):114–118

Carpenter GA, Gaddam SC (2010) Biased ART: a neural architecture that shifts attention toward previously disregarded features following an incorrect prediction. Neural Netw 23(3):435–451

Carpenter GA, Grossberg S (1987) Neural dynamics of category learning and recognition: attention, memory consolidation, and amnesia. Adv Psychol 42:233–290

Carpenter GA, Grossberg S (2017) Adaptive resonance theory. In: Sammut C, Webb GI (eds) Encyclopedia of machine learning and data mining. Springer, Boston, MA

Carpenter GA, Milenova BL (2000) ART neural networks for medical data analysis and fast distributed learning. In: Artificial neural networks in medicine and biology. Springer, London

Carpenter GA, Grossberg S, Reynolds JH (1991) ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organizing neural network. Neural Netw 4(5):565–588

Carroll JB (1993) Human cognitive abilities: a survey of factor-analytic studies. Cambridge University Press, Cambridge

Cassell J, Thorisson KR (1999) The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. Appl Artif Intell 13(4–5):519–538

Cassimatis NL (2007) Harnessing multiple representations for autonomous full-spectrum political, military, economic, social, information and infrastructure (PMESII) reasoning. Final technical report AFRL-IF-RS-TR-2007-131

Cassimatis NL, Trafton JG, Bugajska MD, Schultz AC (2004) Integrating cognition, perception and action through mental simulation in robots. Robot Auton Syst 49(1–2):13–23. https://doi.org/10.1016/j.robot.2004.07.014

Chikhaoui B, Pigot H, Beaudoin M, Pratte G, Bellefeuille P, Laudares F (2009) Learning a song: an ACT-R model. In: Proceedings of the international conference on computational intelligence, pp 405–410

Christensen HI, Kruijff GJ, Wyatt JL (eds) (2010) Cognitive systems. Springer

Ciliberto C, Smeraldi F, Natale L, Metta G (2011) Online multiple instance learning applied to hand detection in a humanoid robot. In: Proceedings of IEEE international conference on intelligent robots and systems, pp 1526–1532

Cochran RE, Lee FJ, Chown E (2006) Modeling emotion: Arousal's impact on memory. In: Proceedings of the 28th annual conference of the cognitive science society, pp 1133–1138

Conforth M, Meng Y (2011a) CHARISMA: a context hierarchy-based cognitive architecture for self-motivated social agents. In: Proceedings of the international joint conference on neural networks, pp 1894–1901

Conforth M, Meng Y (2011b) Self-reorganizing knowledge representation for autonomous learning in social agents. In: Proceedings of the international joint conference on neural networks, pp 1880–1887

Conforth M, Meng Y (2012) Embodied intelligent agents with cognitive conscious and unconscious reasoning. In: Proceedings of the BMI ICBM, pp 15–20

Coombs D, Murphy K, Lacaze A, Legowik S (2000) Driving autonomously offroad up to 35 km/h. In: Proceedings of the IEEE intelligent vehicles symposium, MI, pp 186–191

Corker K, Pisanich G, Bunzo M (1997) A cognitive system model for human/automation dynamics in airspace management. In: Proceedings of the first European/US symposium on air traffic management

Cowan N (2008) Chapter 20 what are the differences between long-term, short-term, and working memory? In: Progress in brain research, vol 169. Elsevier, pp 323–338

Coward LA (2011) Modelling memory and learning consistently from psychology to physiology. In: Cutsuridis V, Hussain A, Taylor JG (eds) Perception-action cycle. Springer, pp 63–133

Cox MT (2005) Metacognition in computation: a selected research review. Artif Intell 169(2):104–141

Cox MT (2007) Perpetual self-aware cognitive agents. AI Mag 28(1):32. https://doi.org/10.1609/AIMAG.V28I1.2027

Cox MT (2013) MIDCA: a metacognitive, integrated dual-cycle architecture for self-regulated autonomy. Computer science technical report no CS-TR-5025

Cox MT, Oates T, Paisner M, Perlis D (2012) Noting anomalies in streams of symbolic predicates using A-distance. Adv Cogn Syst 2:167–184

Crawford E, Gingerich M, Eliasmith C (2015) Biologically plausible, human-scale knowledge representation. Cogn Sci 40(4):412–417

Cuadrado D, Saez Y (2009) Chuck Norris rocks! In: Proceedings of the IEEE symposium on computational intelligence and games, pp 69–74

Cui Y, Ahmad S, Hawkins J (2015) Continuous online sequence learning with an unsupervised neural network model. arXiv preprint arXiv:151205463

Danker JF, Anderson JR (2010) The ghosts of brain states past : remembering reactivates the brain regions engaged during encoding. Psychol Bull 136(1):87–102. https://doi.org/10.1037/a0017937

Dannenhauer D, Cox MT, Gupta S, Paisner M, Perlis D (2014) Toward meta-level control of autonomous agents. In: Proceedings of the 5th annual international conference on biologically inspired cognitive architectures, vol 41. Elsevier Masson SAS, pp 226–232

Demetgul M, Tansel IN, Taskin S (2009) Fault diagnosis of pneumatic systems with artificial neural network algorithms. Expert Syst Appl 36(7):10,512–10,519

Derbinsky N, Laird JE (2012a) Competence-preserving retention of learned knowledge in Soar's working and procedural memories. In: Proceedings of the 11th international conference on cognitive modeling

Derbinsky N, Laird JE (2012b) Computationally efficient forgetting via base-level activation. In: Proceedings of the 11th international conference on cognitive modeling, pp 109–110

Deutsch S, Cramer N (1998a) Omar human performance modeling in a decision support experiment. In: Proceedings of the human factors and ergonomics society 42nd annual meeting, pp 1232–1236

Deutsch S, Cramer N (1998b) Omar human performance modeling in a decision support experiment. In: Proceedings of the human factors and ergonomics society 42nd annual meeting, pp 1232–1236

Deutsch SE (2006) UAV operator human performance models. Technical report AFRL-RI-RS-TR-2006-0158

Deutsch SE, Macmillan J, Camer NL, Chopra S (1997) Operability model architecture: demonstration final report. Technical report AL/HR-TR-1996-0161

Di Nuovo A, De La Cruz VM, Cangelosi A (2014) Grounding fingers, words and numbers in a cognitive developmental robot. In: Proceedings of the IEEE symposium on computational intelligence, cognitive algorithms, mind, and brain

D'Inverno M, Luck M, Georgeff M, Kinny D, Wooldridge M (2004) The dMARS architecture: a specification of the distributed multi-agent reasoning system. Auton Agent Multi Agent Syst 9(1–2):5–53

Distante C, Siciliano P, Vasanelli L (2000) Odor discrimination using adaptive resonance theory. Sens Actuators 69(3):248–252

Douglass S, Ball J, Rodgers S (2009) Large declarative memories in ACT-R. In: Proceedings of the 9th international conference on cognitive modeling, p 234

Drummond M, Bresina J (1990) Planning for control. In: Proceedings of the 5th IEEE international symposium on intelligent control, pp 657–662

Duch W (2006) Computational creativity. In: IEEE world congress on computational intelligence

Duch W (2007) Towards comprehensive foundations of computational intelligence. Stud Comput Intell 63:261–316

Duch W, Oentaryo RJ, Pasquier M (2008) Cognitive architectures: where do we go from here? In: Wang P, Goertzel B, Franklin S (eds) Frontiers in artificial intelligence and applications, vol 171. IOS Press, Amsterdam, pp 122–136

Duro RJ et al (2010) Evolutionary architecture for lifelong learning and real-time operation in autonomous robots. In: Angelov P, Filev DP, Kasabov N (eds) Evolving intelligent systems: methodology and applications. pp 365–400

Edelman GM (2007) Learning in and from brain-based devices. Science 318(5853):1103–1105

Eliasmith C, Kolbeck C (2015) Marr's attacks: on reductionism and vagueness. Top Cogn Sci 1–13

Eliasmith C, Stewart TC (2012) A large-scale model of the functioning brain. Science 338:1202–1205

Elkins L, Sellers D, Monach WR (2010) The autonomous maritime navigation (AMN) project: field tests, autonomous and cooperative behaviors, data fusion, sensors, and vehicles. J Field Robot 27(6):81–86. https://doi.org/10.1002/rob

Epstein SL (1992) The role of memory and concepts in learning. Mind Mach 2(3):239–265

Epstein SL (1994) For the right reasons: the FORR architecture for learning in a skill domain. Cogn Sci 18(3):479–511

Epstein SL (1995) On heuristic reasoning, reactivity, and search. In: Proceedings of IJCAI-95, pp 454–461

Epstein SL (2001) Learning to play expertly: a tutorial on Hoyle. In: Furnkranz J, Kubat M (eds) Machines that learn to play games. Nova Science Publishers, Hauppauge, pp 153–178

Epstein SL (2004) Metaknowledge for autonomous systems. In: Proceedings of AAAI spring symposium on knowledge representation and ontology for autonomous systems

Epstein SL, Petrovic S (2008) Learning expertise with bounded rationality and self-awareness. In: Metareasoning: thinking about thinking. MIT Press, Cambridge **(scholarship online)**

Epstein SL, Freuder EC, Wallace R, Morozov A, Samuels B (2002) The adaptive constraint engine. In: Proceedings of the international conference on principles and practice of constraint programming, pp 525–540

Epstein SL, Passonneau R, Gordon J, Ligorio T (2012) The role of knowledge and certainty in understanding for dialogue. In: AAAI fall symposium: advances in cognitive systems

Etzioni O (1993) Acquiring search-control knowledge via static analysis. Artif Intell 62(2):255–301

Evertsz R, Ritter FE, Russell S, Shepherdson D (2007) Modeling rules of engagement in computer generated forces. In: Proceedings of the 16th conference on behavior representation in modeling and simulation, pp 123–34

Evertsz R, Pedrotti M, Busetta P, Acar H, Ritter FE (2009) Populating VBS2 with realistic virtual actors. In: Proceedings of the 18th conference on behavior representation in modeling and simulation

Faghihi U (2011) The use of emotions in the implementation of various types of learning in a cognitive agent. PhD thesis

Faghihi U, Franklin S (2012) The LIDA model as a foundational architecture for AGI. In: Wang P, Goertzel B (eds) Theoretical foundations of artificial general intelligence. Atlantis thinking machines, vol 4. Atlantis Press, Paris, pp 103–121

Faghihi U et al (2011a) Implementing an efficient causal learning mechanism in a cognitive tutoring agent. In: Mehrotra KG, Mohan CK, Oh JC, Varshney PK, Ali M (eds) Modern approaches in applied intelligence: Proceedings of the 24th international conference on industrial engineering and other applications of applied intelligent systems. pp 27–36

Faghihi U et al (2011b) Emotional cognitive architectures. In: D'Mello S, Graesser A, Schuller B, Martin JC (eds) Proceedings of the 4th international conference on affective computing and intelligent interaction (ACII). pp 487–496

Faghihi U et al (2013) CELTS: a cognitive tutoring agent with human-like learning capabilities and emotions. In: Jain LC, Howlett RJ, Ditzinger T (eds) Smart innovation, systems and technologies: 2190–3018. Springer

Fan X, Sun B, Sun S, McNeese M, Yen J (2006) RPD-enabled agents teaming with humans for multi-context decision making. In: Proceedings of the international conference on autonomous agents

Fan X, McNeese M, Sun B, Hanratty T, Allender L, Yen J (2010a) Human-agent collaboration for time-stressed multicontext decision making. IEEE Trans Syst Man Cybern Part A Syst Hum 40(2):306–320

Fan X, McNeese M, Yen J (2010b) NDM-based cognitive agents for supporting decision-making teams. Hum Comput Interact 25(3):195–234

Fazl A, Grossberg S, Mingolla E (2009) View-invariant object category learning, recognition, and search: how spatial and object attention are coordinated using surface-based attentional shrouds. Cogn Psychol 58(1):1–48

Fedor A, Zachar I, Szilágyi A, Öllinger M (2017) Cognitive architecture with evolutionary dynamics solves insight problem. Front Psychol 8:1–15

Feldman V, Kokinov B (2009) Anxiety restricts the analogical search in an analogy generation task. In: Kokinov B, Holyoak K, Gentner D (eds) New frontiers in analogy research: proceedings of the second international conference on analogy. New Bulgarian University Press, pp 117–126

Fink E, Blythe J (2005) Prodigy bidirectional planning. J Exp Theor Artif Intell 17(3):161–200

Firby JR, Kahn RE, Prokopowicz PN, Swain MJ (1995) An architecture for vision and action. In: Proceedings of the 14th international joint conference on artificial intelligence

Firby RJ (1989) Adaptive execution in complex dynamic worlds. PhD thesis

Flavell JH (1979) Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. Am Psychol 34(10):906–911

Fleischer JG, Edelman GM (2009) Brain-based devices: an embodied approach to linking nervous system structure and function to behavior. IEEE Robot Autom Mag 16(3):33–41

Fleischer JG, Krichmar JL (2007) Sensory integration and remapping in a model of the medial temporal lobe during maze navigation by a brain-based device. J Integr Neurosci 6(3):403–431

Flynn AM, Brooks RA, Wells WM, Barrett DS (1989) The world's largest one cubic inch robot. In: Proceedings of IEEE conference on microelectromechanical systems, pp 98–101

Forbus KD, Klenk M, Hinrichs T (2009) Companion cognitive systems: design goals and lessons learned. IEEE Intell Syst PP(99):36–46

Forbus KD, Ferguson RW, Lovett A (2016) Extending SME to handle large-scale cognitive modeling. Cogn Sci 1–50

Foxvog D (2010) Cyc. In: Poli R, Healy M, Kameas A (eds) Theory and applications of ontology: computer applications. Springer, pp 259–278

Franklin S (2000a) Learning in "Conscious" software agents. In: Workshop on development and learning

Franklin S (2000b) Modeling consciousness and cognition in software agents. In: Proceedings of the third international conference on cognitive modeling, pp 27–58

Franklin S (2007) A foundational architecture for artificial general intelligence. Adv Artif Gen Intell Concepts Archit Algorithms 6:36–54

Franklin S, Strain S, Snaider J, McCall R, Faghihi U (2012) Global workspace theory, its LIDA model and the underlying neuroscience. Biol Inspired Cogn Archit 1:32–43

Franklin S, Madl T, Strain S, Faghihi U, Dong D, Kugele S, Snaider J, Agrawal P, Chen S (2016) A LIDA cognitive model tutorial. Biol Inspired Cogn Archit 16:105–130

Freed M, Remington R (2000) Making human-machine system simulation a practical engineering tool: an APEX overview. In: Proceedings of the 3rd international conference on cognitive modelling

Freed MA (1998) Simulating human performance in complex, dynamic environments. PhD Thesis (June)

Friedman SE, Forbus KD (2010) An integrated systems approach to explanation-based conceptual change. In: Association for the advancement of artificial intelligence

Friedman SE, Forbus KD, Sherin B (2011) Constructing & revising commonsense science explanations: a metareasoning approach. In: AAAI fall symposium on advances in cognitive systems

Frintrop S, Rome E, Christensen HI (2010) Computational visual attention systems and their cognitive foundations: a survey. ACM Trans Appl Percept 7(1):6

From J, Perrin P, O'Neill D, Yen J (2011) Supporting the Commander's information requirements: automated support for battle drill processes using R-CAST. In: Proceedings of the IEEE military communications conference MILCOM

Gat E (1992) Integrating planning and reacting in a heterogeneous asynchronous architecture for controlling real-world mobile robots. In: AAAI pp 809–815

Gat E, Dorais G (1994) Robot navigation by conditional sequencing. In: Proceedings of the international conference on robotics and automation, pp 1293–1299

Gemrot J et al (2009) Pogamut 3 can assist developers in building AI (not only) for their videogame agents. In: Dignum F, Bradshaw J, Silverman B, van Doesburg W (eds) Agents for games and simulations (AGS): international workshop on agents for games and simulations. pp 1–15

Gemrot J, Brom C, Kadlec R, Bida M, Burkert O, Zemčák M, Píbil R, Plch T (2010) Pogamut 3—virtual humans made simple. Adv Cogn Sci 211–243

Gentner D, Collins A (1981) Studies of inference from lack of knowledge. Mem Cogn 9:434–443

Georgeff M, Pell B, Pollack M, Tambe M, Wooldridge M (1998) The belief-desire-intention model of agency. In: International workshop on agent theories, architectures, and languages

Georgeff MP, Ingrand FF (1989) Decision-making in an embedded reasoning system. In: Proceedings of the eleventh international joint conference on artificial intelligence (IJCAI-89)

Georgeff MP, Lansky AL (1986) Procedural knowledge. Proceedings of the IEEE 74:1383–1398

Gobet F, Lane PC (2012) Chunking mechanisms and learning. In: Seel NM (ed) Encyclopedia of the sciences of learning. Springer, New York, pp 541–544

Gobet FR (2008) Memory for the meaningless: how chunks help. In: Proceedings of the 20th meeting of the cognitive science society, pp 398–403

Goertzel B (2008) A pragmatic path toward endowing virtually-embodied AIs with human-level linguistic capability. In: Proceedings of the international joint conference on neural networks

Goertzel B (2012) Perception processing for general intelligence: bridging the symbolic/subsymbolic gap AI. https://doi.org/10.1007/978-3-642-35506-6_9

Goertzel B, Pennachin C (2007) The Novamente artificial intelligence engine. Artificial general intelligence. Springer, Berlin, pp 63–129

Goertzel B, Yu G (2014) A cognitive API and its application to AGI intelligence assessment. In: Goertzel B, Orseau L, Snaider J (eds) Proceedings of the international conference on artificial general intelligence. pp 242–245

Goertzel B, Pennachin C, Geissweiller N, Looks M, Senna A, Silva W, Heljakka A, Lopes C (2008a) An integrative methodology for teaching embodied non-linguistic agents, applied to virtual animals in second life. Front Artif Intell Appl 171:161–175

Goertzel B, Pennachin C, Souza SD (2008b) An inferential dynamics approach to personality and emotion driven behavior determination for virtual animals. In: AISB 2008 convention on communication, interaction and social intelligence

Goertzel B, Garis HD, Pennachin C, Geisweiller N, Araujo S, Pitt J, Chen S, Lian R, Jiang M, Yang Y, Huang D (2010a) OpenCogBot: achieving generally intelligent virtual agent control and humanoid robotics via cognitive synergy. In: Proceedings of international conference on artificial intelligence

Goertzel B, Lian R, Arel I, de Garis H, Chen S (2010b) A world survey of artificial brain projects, part II: biologically inspired cognitive architectures. Neurocomputing 74(1–3):30–49

Goertzel B, Sanders T, O'Neill J (2013) Integrating deep learning based perception with probabilistic logic via frequent pattern mining. In: International conference on artificial general intelligence

Goertzel B, Pennachin C, Geisweiller N (2014) Brief survey of cognitive architectures. Engineering general intelligence, part 1. Atlantis Press, Paris, pp 101–142

Gordon J, Epstein SL (2011) Learning to balance grounding rationales for dialogue systems. In: Proceedings of SIGDIAL conference, pp 266–271

Gore BF, Hooey BL, Wickens CD, Scott-Nash S (2009) A computational implementation of a human attention guiding mechanism in MIDAS v5. In: International conference on digital human modeling

Gosmann J, Stewart TC, Wennekers T (2017) A spiking neuron model of word associations for the remote associates test. Front Psychol. https://doi.org/10.3389/fpsyg.2017.00099

Grossberg S (1999) The link between brain learning, attention, and consciousness. Conscious Cogn 8:1–44

Grossberg S (2003) Resonant neural dynamics of speech perception. Technical report CAS/CNS-TR-02-008

Grossberg S (2007) Towards a unified theory of neocortex: laminar cortical circuits for vision and cognition. Prog Brain Res 165:79–104

Grossberg S, Myers CW (2015) The resonant dynamics of speech perception: interword integration and duration-dependent backward effects. Psychol Rev 107(4):735

Grossberg S, Govindarajan KK, Wyse LL, Cohen MA (2004) ARTSTREAM: a neural network model of auditory scene analysis and source segregation. Neural Netw 17(4):511–536

Gunzelmann G, Gross JB, Gluck KA, Dinges DF (2009) Sleep deprivation and sustained attention performance: integrating mathematical and cognitive modeling. Cogn Sci 33(5):880–910

Hammer P, Lofthouse T, Wang P (2016) The OpenNARS implementation of the non-axiomatic reasoning system. In: International conference on artificial general intelligence

Harrigan C, Goertzel B, Ikle M, Belayneh A, Yu G (2014) Guiding probabilistic logical inference with nonlinear dynamical attention allocation. In: International conference on artificial general intelligence, pp 238–241

Hart S, Dahn D, Atencio A, Dalal MK (2001) Evaluation and application of MIDAS v2.0. SAE technical paper 2001-01-2648

Hawes N et al (2010) The playmate system. In: Christensen HI, Kruijff GJM, Wyatt JL (eds) Cognitive Systems. Springer, pp 367–393

Hawkins J, George D (2006) Hierarchical temporal memory: concepts, theory and terminology, technical paper, Numenta

Hayes-Roth B (1995) An architecture for adaptive intelligent systems. Artif Intell 72(1–2):329–365. https://doi.org/10.1016/0004-3702(94)00004-K

Hayes-Roth B (1996) A domain-specific software architecture for a class of intelligent patient monitoring agents. J Exp Theor Artif Intell 8(2):149–171

Hayes-Roth B, Gent RV (1995) Story-making with improvisational puppets and actors. In: Proceedings of the first international conference on autonomous agents

Hayes-Roth B, Washington R, Ash D, Hewett R, Collinot A, Vina A, Seiver A (1992) Guardian: a prototype intelligent agent for intensive-care monitoring. Artif Intell Med 4(2):165–185

Hayes-Roth B, Lalanda P, Morignot P, Pfleger K, Balabanovic M (1993) Plans and behavior in intelligent agents. KSL report no 93-43

He K, Zhang X, Ren S, Sun J (2014) Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: ICCV

Helie S, Sun R (2010) Incubation, insight, and creative problem solving: a unified theory and a connectionist model. Psychol Rev 117(3):994

Helie S, Sun R (2014) An integrative account of memory and reasoning phenomena. New Ideas Psychol 35(1):36–52

Heljakka A, Goertzel B, Silva W, Pennachin C, Senna A, Goertzel I (2007) Probabilistic logic based reinforcement learning of simple embodied behaviors in a 3D simulation world. Front Artif Intell Appl 157:253–275

Henderson TC, Joshi A (2013) The cognitive symmetry engine. Technical report UUCS-13-004

Henderson TC, Peng H, Sikorski K, Deshpande N, Grant E (2011) The cognitive symmetry engine: an active approach to knowledge. In: Proceedings of the IROS 2011 workshop on knowledge representation for autonomous robots

Henderson TC, Joshi A, Grant E (2012) From sensorimotor data to concepts: the role of symmetry. Technical report UUCS-12-005

Herd S, Szabados A, Vinokurov Y, Lebiere C, Cline A, O'Reilly RC (2014) Integrating theories of motor sequencing in the SAL hybrid architecture. Biol Inspired Cogn Archit 8:98–106

Herd SA, Krueger KA, Kriete TE, Huang TR, Hazy TE, O'Reilly RC (2013) Strategic cognitive sequencing: a computational cognitive neuroscience approach. Comput Intell Neurosci 2013:4

Hernandez-Orallo J (2017) Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement. Artif Intell Rev 48(3):397–447. https://doi.org/10.1007/s10462-016-9505-7

Higgins ET, Eitam B (2014) Priming…shmiming: it's about knowing when & why stimulated memory representations become active. Soc Cogn 32:1–33

Hilario M (1997) An overview of strategies for neurosymbolic integration. In: Sun R, Alexandre F (eds) Connectionist-symbolic integration: from unified to hybrid approaches. Psychology Press, Hove, pp 13–35

Hinrichs TR, Forbus KD (2007) Analogical learning in a turn-based strategy game. In: Proceedings of international joint conference on artificial intelligence, pp 853–858

Hinrichs TR, Forbus KD (2014) X goes first: teaching simple games through multimodal interaction. Adv Cogn Syst 3:218

Hofstadter D (1993) How could a COPYCAT ever be creative? AAAI technical report SS-93-01, pp 8–21

Holroyd CB, Coles MGH (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 109(4):679–709

Hong TH, Balakirsky SB, Messina E, Chang T, Shneier M (2002) A Hierarchical world model for an autonomous scout vehicle. In: 16th annual international symposium on aerospace/defense sensing, simulation, and controls (SPIE 2002), pp 343–354

Hooey BL, Gore BF, Wickens CD, Scott-Nash S, Socash CM, Salud E, Foyle DC (2010) Human modelling in assisted transportation. In: Proceeding of the human modeling in assisted transportation conference, pp 327–333

Huang X, Weng J (2007) Inherent value systems for autonomous mental development. Int J Humanoid Robot 4(2):407–433

Huber E, Kortenkamp D (1995) Using stereo vision to pursue moving agents with a mobile robot. In: Proceedings of the IEEE international conference on robotics and automation

Hudlicka E (2001) Modeling affect regulation and induction. In: Proceedings of the AAAI fall symposium 2001, "emotional and intelligent II: the tangled knot of social cognition"

Hudlicka E (2002) This time with feeling: integrated model of trait and state effects on cognition and behavior. Appl Artif Intell 16:611–641

Hudlicka E (2004) Beyond cognition: modeling emotion in cognitive architectures. In: Proceedings of the sixth international conference on cognitive modeling, pp 118–123

Hudlicka E (2005) A computational model of emotion and personality: applications to psychotherapy research and practice. In: Proceedings of the 10th annual cybertherapy conference: a decade of virtual reality

Hudlicka E (2006) Modeling effects of emotion and personality on political decision-making. Programming for peace. Springer, Berlin, pp 355–411

Hudlicka E (2008) Modeling the mechanisms of emotion effects on cognition. In: Proceedings of the AAAI fall symposium on biologically inspired cognitive architectures, pp 82–86

Hudlicka E (2009) Challenges in developing computational models of emotion and consciousness. Int J Mach Conscious 1(1):131–153

Hudlicka E (2010) Modeling cultural and personality biases in decision making. In: Proceedings of the 3rd international conference on applied human factors and ergonomics (AHFE)

Hudlicka E (2016) Computational analytical framework for affective modeling: towards guidelines for designing. In: Psychology and mental health: concepts, methodologies, tools, and applications: concepts, methodologies, tools, and applications, pp 1–64. https://doi.org/10.4018/978-1-4666-7278-9.ch001

Hudlicka E, Matthews G (2009) Affect, risk and uncertainty in decision-making. An integrated computational-empirical approach, Final report

Hudlicka E, Zacharias G, Psotka J (2000) Increasing realism of human agents by modeling individual differences: Methodology, architecture, and testbed. In: Simulating human agents, American association for artificial intelligence fall 2000 symposium series, pp 53–59

Huntsberger T (2011) Cognitive architecture for mixed human-machine team interactions for space exploration. In: IEEE aerospace conference proceedings

Huntsberger T, Stoica A (2010) Envisioning cognitive robots for future space exploration. In: SPIE defense, security, and sensing

Huntsberger T, Woodward G (2011) Intelligent autonomy for unmanned surface and underwater vehicles. In: Proceedings of the OCEANS'11, pp 1–10

Huntsberger T, Aghazarian H, Howard A, Trotz DC (2011) Stereo vision-based navigation for autonomous surface vessels. J Field Robot 28(1):3–18

Ichise R (2016) An analysis of the chc model for comparing cognitive architectures. Proc Comput Sci 88:239–244

Ikle M, Goertzel B (2011) Nonlinear-dynamical attention allocation via information geometry. In: International conference on artificial general intelligence

Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 20(11):1254–1259

Ivaldi S, Lyubova N, Gerardeaux-Viret D, Droniou A, Anzalone SM, Chetouani M, Filliat D, Sigaud O (2012) Perception and human interaction for developmental learning of objects and affordances. In: IEEE-RAS international conference on humanoid robots, pp 248–254

Ivaldi S, Nguyen SM, Lyubova N, Droniou A, Padois V, Filliat D, Oudeyer PY, Sigaud O (2014) Object learning through active exploration. IEEE Trans Auton Ment Dev 6(1):56–72

Jaszuk M, Starzyk JA (2016) Building internal scene representation in cognitive agents. In: Skulimowski, AMJ, Kacprzyk J (eds) Knowledge, information and creativity support systems: recent trends, advances and solutions. Springer, pp 479–491

Jensen R, Veloso M (1998) Interleaving deliberative and reactive planning in dynamic multi-agent domains. In: Proceedings of the AAAI fall symposium on on integrated planning for autonomous agent architectures

Jilk DJ, Lebiere C, O'Reily RC, Anderson JR (2008) SAL: An explicitly pluralistic cognitive architecture. J Exp Theor Artif Intell 20(3):197–218

Jones RM, Laird JE, Nielsen PE, Coulter KJ, Kenny P, Koss FV (1999) Automated intelligent pilots for combat flight simulation. AI Mag 20(1):27–42

Jones RM, Wray REI, van Lent M (2012) Practical evaluation of integrated cognitive systems. Adv Cogn Syst 1:83–92

Jones SJ, Wandzel AR, Laird JE (2016) Efficient computation of spreading activation using lazy evaluation. In: Proceedings of the international conference on cognitive modeling

Jonsdottir GR, Thórisson KR (2013) A distributed architecture for real-time dialogue and on-task learning of efficient co-operative turn-taking. In: Campbell N, Matej R (eds) Coverbal synchrony in human-machine interaction. CRC Press, Boca Raton, pp 293–323

Jordan MI, Russell S (1999) Computational intelligence. In: Wilson RA, Keil FC (eds) The MIT encyclopedia of the cognitive sciences. The MIT Press, Cambridge, MA

Just MA, Carpenter PA (1992) A capacity theory of comprehension: Individual differences in working memory. Psychol Rev 99(1):122–149

Just MA, Varma S (2007) The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. Cogn Affect Behav Neurosci 7(3):153–191

Just MA, Carpenter PA, Keller TA, Emery L, Zajac H, Thulborn KR (2001) Interdependence of nonoverlapping cortical systems in dual cognitive tasks. NeuroImage 14:417–426. https://doi.org/10.1006/nimg.2001.0826

Kadlec R, Gemrot J, Bida M, Burkert O, Havlicek J, Zemcak L, Pibil R, Vansa R, Brom C (2009) Extensions and applications of Pogamut 3 platform. In: International workshop on intelligent virtual agents

Kawamura K, Cambron M, Fujiwara K, Barile J (1993) A cooperative robotic aid system. In: Proceedings of the conference on virtual reality systems, teleoperation and beyond speech recognition

Kawamura K, Peters RAI, Bodenheimer RE, Sarkar N, Park J, Clifton CA, Spratley AW (2004) A parallel distributed cognitive control system for a humanoid robot. Int J Humanoid Robot 1(1):65–93

Kawamura K, Gordon SM, Ratanaswasd P, Erdemir E, Hall JF (2008) Implementation of cognitive control for a humanoid robot. Int J Humanoid Robot 5(4):547–586

Kaylani A, Georgiopoulos M, Mollaghasemi M, Anagnostopoulos GC (2009) AG-ART: an adaptive approach to evolvong ART architectures. Neurocomputing 72:2079–2092

Kedar ST, McKusick KB (1992) There is no free lunch: tradeoffs in the utility of learned knowledge. In: Proceedings of the first international conference on artificial intelligence planning systems, pp 281–282

Kelley TD (2003) Symbolic and sub-symbolic representations in computational models of human cognition: what can be learned from biology? Theory Psychol 13(6):847–860

Kennedy W, Trafton JG (2006) Long-term symbolic learning in Soar and ACT-R. In: Proceedings of the seventh international conference on cognitive modelling, pp 166–171

Kennedy WG, De Jong KA (2003) Characteristics of long-term learning in soar and its application to the utility problem. Proceedings of the fifth international conference on machine learning, pp 337–344

Khaleghi B, Khamis A, Karray FO (2013) Multisensor data fusion: a review of the state-of-the-art. Inf Fusion 14(1):28–44

Kieras D (2010) Modeling visual search of displays of many objects: the role of differential acuity and fixation memory. In: Proceedings of the 10th international conference on cognitive modeling

Kieras D (2012) The control of cognition. In: Gray W (ed) Integrated models of cognitive systems. Oxford University Press, Oxford

Kieras DE (2004) EPIC architecture principles of operation

Kieras DE, Hornof AJ (2014) Towards accurate and practical predictive models of active-vision-based visual search. In: Proceedings of the conference on human factors in computing systems, pp 3875–3884

Kieras DE, Meyer DE (1998) The role of cognitive task analysis in the application of predictive models of human performance. EPIC report no 11 (TR-98/ONR-EPIC-11)

Kieras DE, Wakefield GH, Thompson ER, Iyer N, Simpson BD (2016) Modeling two-channel speech processing with the EPIC cognitive architecture. Top Cogn Sci 8(1):291–304

Kilic O (2015) Intelligent reasoning on natural language data: a non-axiomatic reasoning system approach. PhD thesis

Kinny D, Georgeff M, Hendler J (1992) Experiments in optimal sensing for situated agents. In: Proceedings of the second pacific rim international conference on artificial intelligence

Kirk JR, Laird JE (2014) Interactive task learning for simple games. Adv Cogn Syst 3:13–30

Kirk JR, Laird JE (2016) Learning general and efficient representations of novel games through interactive instruction. Adv Cogn Syst 4

Kiryazov K, Petkov G, Grinberg M, Kokinov B, Balkenius C (2007) The interplay of analogy-making with active vision and motor control in anticipatory robots. In: Workshop on anticipatory behavior in adaptive learning systems, pp 233–253

Kokinov B, Nikolov V, Petrov A (1996) Dynamics of emergent computation in DUAL. In: Ramsay A (ed) Artificial intelligence: methodology, systems, applications. IOS Press, Amsterdam, pp 303–311

Kokinov BN (1990) Associative memory-based reasoning: some experimental results. In: Proceedings of the twelfth annual conference of the cognitive science society

Kokinov BN (1994) The DUAL cognitive architecture: a hybrid multi- agent approach. In: Proceedings of the 11th European conference on artificial intelligence (ECAI)

Koons R (2017) Defeasible reasoning. https://plato.stanford.edu/archives/sum2017/entries/reasoning-defeasible/

Kostavelis I, Nalpantidis L, Gasteratos A (2012) Object recognition using saliency maps and HTM learning. In: Proceedings of the IEEE international conference on imaging systems and techniques, pp 528–532

Kotseruba I (2016) Visual attention in dynamic environments and its application to playing online games. MSc thesis

Kottlors J, Brand D, Ragni M (2012) Modeling behavior of attention-deficit-disorder patients in a N-back task. In: Proceedings of 11th international conference on cognitive modeling (ICCM 2012), pp 297–302

Krichmar JL (2012) Design principles for biologically inspired cognitive robotics. Biol Inspired Cogn Archit 1:73–81

Krichmar JL, Edelman GM (2005) Brain-based devices for the study of nervous systems and the development of intelligent machines. Artif Life 11(1–2):63–77

Krichmar JL, Snook JA (2002) A neural approach to adaptive behavior and multi-sensor action selection in a mobile device. In: Proceedings of the IEEE international conference on robotics and automation

Krichmar JL, Nitz DA, Gally JA, Edelman GM (2005) Characterizing functional hippocampal pathways in a brain-based device as it solves a spatial memory task. Proc Nat Acad Sci USA 102(6):2111–2116

Kuokka DR (1989) Integrating planning, execution, and learning. In: Proceedings of the NASA conference on space telerobotics, pp 377–386

Kuokka DR (1991) MAX: a meta-reasoning architecture for "X". SIGART Bull 2(4):93–97

Kurup U, Bignoli PG, Scally JR, Cassimatis NL (2011) An architectural framework for complex cognition. Cogn Syst Res 12(3–4):281–292

Laird JE (2012a) The soar cognitive architecture. MIT Press, Cambridge

Laird JE (2012b) The soar cognitive architecture. AISB Q 171(134):224–235

Laird JE, Derbinsky N (2009) A year of episodic memory. In: Proceedings of the workshop on grand challenges for reasoning from experiences, IJCAI, pp 7–10

Laird JE, Mohan S (2014) A case study of knowledge integration across multiple memories in Soar. Biol Inspired Cogn Archit 8:93–99

Laird JE, Rosenbloom PS, Newell A (1984) Towards chunking as a general learning mechanism. In: AAAI proceedings, pp 188–192

Laird JE, Yager ES, Hucka M, Tuck CM (1991) Robo-Soar: An integration of external interaction, planning, and learning using Soar. Robot Auton Syst 8(1–2):113–129

Laird JE, Coulter KJ, Jones RM, Kenny PG, Koss F, Nielsen PE (1998) Integrating intelligent computer generated forces in distributed simulations: TacAir-Soar in STOW-97. In: Proceedings of the spring simulation interoperability workshop

Laird JE, Kinkade KR, Mohan S, Xu JZ (2004) Cognitive robotics using the soar cognitive architecture. In: Proceedings of the 6th international conference on cognitive modelling, pp 226–330

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

Landauer K (1986) How much do people remember! Some estimates of the quantity of learned information in long-term memory. Cogn Sci 493:477–493

Lane PCR, Sykes A, Gobet F (2003) Combining low-level perception with expectations in CHREST. In: Proceedings of the European cognitive science conference, pp 205–210

Lane PCR, Gobet F, Smith RL (2009) Attention mechanisms in the CHREST cognitive architecture. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 5395 LNAI:183–196

Langley P, Allen JA (1993) A unified framework for planning and learning. In: Minton S (ed) Machine learning methods for planning. Morgan Kaufmann, Burlington

Langley P, Rogers S (2008) An extended theory of human problem solving. In: Proceedings of the 27th annual meeting of cognitive science society, pp 166–186

Langley P, Cummings K, Shapiro D (2004) Hierarchical skills and cognitive architectures. In: Proceedings of the 26th annual conference of the cognitive science society, pp 779–784

Langley P, Choi D, Rogers S (2005) Interleaving learning, problem solving, and execution in the ICARUS architecture. Technical report, Computational Learning Laboratory

Langley P, Laird JE, Rogers S (2009) Cognitive architectures: research issues and challenges. Cogn Syst Res 10(2):141–160

Lavin A, Ahmad S (2015) Evaluating real-time anomaly detection algorithms—the numenta anomaly benchmark. In: Proceedings of the 14th international conference on machine learning and applications (ICMLA)

Lavin A, Ahmad S, Hawkins J (2016) Sparse distributed representations. https://numenta.com/assets/pdf/biological-and-machine-intelligence/BaMI-SDR.pdf

Lebiere C, Pirolli P, Thomson R, Paik J, Rutledge-Taylor M, Staszewski J, Anderson JR (2013) A functional model of sensemaking in a neurocognitive architecture. Comput Intell Neurosci 2013:5

Lebiere CCMU, Biefeld ECMU, Archer RMA&D, Archer SMA&D, Allender LARL, Kelley TDARL (2002) IMPRINT/ACT-R: Integration of a task network modeling architecture with a cognitive architecture and its application to human error modeling. In: Proceedings of the 2002 advanced simulation technologies conference, San Diego, CA, Simulation Series 34, pp 13–19

Legg S, Hutter M (2007) A collection of definitions of intelligence. Front Artif Intell Appl 157. eprint:0706.3639v1

Leitner J, Harding S, Frank M, Forster A, Schmidhuber J (2013) An integrated, modular framework for computer vision and cognitive robotics research (icVision). Adv Intell Syst Comput 205–210

Lerner I, Bentin S, Shriki O (2012) Spreading activation in an attractor network with latching dynamics: automatic semantic priming revisited. Cogn Sci 36(8):1339–1382

Lewis RL (1992) Recent developments in the NL-Soar garden path theory. Technical report CMU-CS-93-141

Lian R, Goertzel B, Liu R, Ross M, Queiroz M, Vepstas L (2010) Sentence generation for artificial brains: a glocal similarity-matching approach. Neurocomputing 74(1–3):95–103

Lieto A (2016) Representational limits in cognitive architectures. In: Proceedings of the EUCognition, vol 1855

Lieto A, Lebiere C, Oltramari A (2018a) The knowledge level in cognitive architectures: current limitations and possible developments. Cogn Syst Res 48:39–55

Lieto A, Bhatt M, Oltramari A, Vernon D (2018b) The role of cognitive architectures in general artificial intelligence. Cogn Sys Res 48:1–3

Lindes P, Laird JE (2016) Toward integrating cognitive linguistics and cognitive language processing. In: Proceedings of international conference on cognitive modeling

Lison P, Kruijff GJ (2008) Salience-driven contextual priming of speech recognition for human-robot interaction. In: Language, pp 636–640

Llargues Asensio JM, Peralta J, Arrabales R, Bedia MG, Cortez P, Peña AL (2014) Artificial Intelligence approaches for the generation and assessment of believable human-like behaviour in virtual characters. Expert Syst Appl 41(16):7281–7290

Lloyd-Kelly M et al (2014) The effects of bounding rationality on the performance and learning of CHREST agents in tileworld. In: Bramer M, Petridis M (eds) Research and development in intelligent systems XXXI. Springer, pp 149–162

Lloyd-Kelly M, Gobet FR, Lane PCR (2015) Piece of mind: long-term memory structure in ACT-R and CHREST. In: Proceedings of the 37th annual meeting of the cognitive science society

López DG, Sjö K, Paul C, Jensfelt P (2008) Hybrid laser and vision based object search and localization. In: Proceedings of the IEEE international conference on robotics and automation (ICRA)

Lytle AM, Saidi KS (2007) NIST research in autonomous construction. Auton Robot 22(3):211–221

Lyubova N, Filliat D, Ivaldi S (2013) Improving object learning through manipulation and robot self-identification. In: Proceeding of the IEEE international conference on robotics and biomimetics (ROBIO)

Madl T, Franklin S (2012) A LIDA-based model of the attentional blink. In: Proceedings of international conference on cognitive modeling (ICCM), pp 283–288

Madl T, Franklin S (2015) Constrained incrementalist moral decision making for a biologically inspired cognitive architecture. In: Trappl R (ed) A construction manual for robots' ethical systems

Madl T, Franklin S, Chen K, Montaldi D, Trappl R (2015) Towards real-world capable spatial memory in the LIDA cognitive architecture. Biol Inspired Cogn Archit 16:87–104

Maffei G, Santos-Pata D, Marcos E, Sánchez-Fibla M, Verschure PFMJ (2015) An embodied biologically constrained model of foraging: from classical and operant conditioning to adaptive real-world behavior in DAC-X. Neural Netw 72:88–108

Mai X, Zhang X, Jin Y, Yang Y, Zhang J (2013) Simple perception-action strategy based on hierarchical temporal memory. In: Proceeding of the IEEE international conference on robotics and biomimetics (ROBIO), pp 1759–1764

Manso LJ, Calderita LV, Bustos P, Garcia J, Martinez M, Fernandez F, Romero-Garces A, Bandera A (2014) A general-purpose architecture to control mobile robots. In: XV workshop of physical agents: book of proceedings (WAF 2014)

Manzolli J, Verschure PF (2005) Roboser: a real-world composition system. Comput Music J 29(3):55–74

Marinier RP, Laird JE (2004) Toward a comprehensive computational model of emotions and feelings. In: Proceedings of sixth international conference on cognitive modeling: ICCM, pp 172–177

Marinier RP, Laird JE, Lewis RL (2009) A computational unification of cognitive behavior and emotion. Cogn Syst Res 10(1):48–69

Marr D (2010) Vision: a computational investigation into the human representation and processing of visual information. MIT Press, Cambridge

Marshall JB (2002a) Metacat: a program that judges creative analogies in a microworld. In: Proceedings to the second workshop on creative systems

Marshall JB (2002b) Metacat: a self-watching cognitive architecture for analogy-making. In: Proceedings of the 24th annual conference of the cognitive science society

Marshall JB (2006) A self-watching model of analogy-making and perception. J Exp Theor Artif Intell 18(3):267–307

Martens S, Carpenter GA, Gaudiano P (1998) Neural sensor fusion for spatial visualization on a mobile robot. https://doi.org/10.1117/12.326991

Martin D, Rincon M, Garcia-Alegre MC, Guinea D (2009) ARDIS: knowledge-based dynamic architecture for real-time surface visual inspection. Lect Notes Comput Sci 5601:395–404

Martin D, Rincon M, Garcia-Alegre MC, Guinea D (2011) ARDIS: knowledge-based architecture for visual system configuration in dynamic surface inspection. Exp Syst 28(4):353–374

Martinez-Gomez J, Marfil R, Calderita LV, Bandera JP, Manso LJ, Bandera A, Romero-Garces A, Bustos P (2014) Toward social cognition in robotics: extracting and internalizing meaning from perception. In: Workshop of physical agents

Martins J, Mendes V (2001) Neural networks and logical reasoning systems. A translation table. Int J Neural Syst 11(2):179–186

Mathews Z, Lechon M, Calvo JMB, Duff ADA, Badia SBI, Verschure PFMJ (2009) Insect-like mapless navigation based on head direction cells and contextual learning using chemo-visual sensors. In: 2009 IEEE/RSJ international conference on intelligent robots and systems, IROS 2009 pp 2243–2250

Mathews Z, i Badia SB, Verschure PFMJ (2012) PASAR: An integrated model of prediction, anticipation, sensation, attention and response for artificial sensorimotor systems. Inf Sci 186(1):1–19

Matthies L (1992) Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation. Int J Comput Vis 8(1):71–91

Maxwell JB (2014) Generative music, cognitive modelling, and computer-assisted composition in musicog and manuscore. PhD Thesis

Maxwell JB, Eigenfeldt A, Pasquier P, Thomas NG (2012) Musicog: a cognitive architecture for music learning and generation. In: Proceedings of the 9th sound and music computing conference, pp 521–528

McCrae RR, John OP (1992) An introduction to the five-factor model and its applications. J Pers 60(2):175–215

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264(5588):746–748

Melis WJC, Chizuwa S, Kameyama M (2009) Evaluation of hierarchical temporal memory for a real world application. In: Proceedings of the 4th international conference on innovative computing, information and control (ICICIC), pp 144–147

Menager DH, Choi D (2016) A robust implementation of episodic memory for a cognitive architecture. In: Proceedings of annual meeting of the cognitive science society, pp 620–625

Metcalfe J (1986) Dynamic metacognitive monitoring during problem solving. J Exp Psychol Learn Mem Cogn 12:623–634

Metta G, Natale L, Nori F, Sandini G, Vernon D, Fadiga L, von Hofsten C, Rosander K, Lopes M, Santos-Victor J, Bernardino A, Montesano L (2010) The iCub humanoid robot: an open-systems platform for research in cognitive development. Neural Netw 23(8–9):1125–1134

Mikolov T, Joulin A, Baroni M (2015) A roadmap towards machine intelligence. arXiv:151108130v1

Miller DP, Slack MG (1991) Global symbolic maps from local navigation. In: Proceedings of the ninth national conference on artificial intelligence AAAI, pp 750–755

Mininger A, Laird J (2016) Interactively learning strategies for handling references to unseen or unknown objects. Adv Cogn Syst 4

Minsky M (1986) The society of mind. Simon & Shuster Inc., New York

Minton S, Carbonell J, Knoblock CA, Kuokka DR, Etzioni O, Gil Y (1989) Explanation-based learning: a problem solving perspective. Artif Intell 40(1–3):63–118

Mitchell DK (2009) Workload analysis of the crew of the abrams V2 SEP: phase I baseline IMPRINT model. Technical report ARL-TR-5028

Mitchell DK, Abounader B, Henry S (2009) A procedure for collecting mental workload data during an experiment that is comparable to IMPRINT workload data. Tehcnical report ARL-TR-5020

Mitchell M, Hofstadter DR (1990) The emergence of understanding in a computer model of concepts and analogy-making. Phys D 42(1–3):322–334

Mitchell T, Allen J, Chalasani P, Cheng J, Etzioni O, Ringuette M, Schlimmer JC (1989) Theo: a framework for self-improving systems. In: VanLehn K (ed) Architectures for intelligence. Erbaum, Hillsdale, pp 323–356

Mitchell TM (1990) Becoming increasingly reactive. In: Proceedings of the eighth national conference on artificial intelligence, pp 1051–1058

Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D (2015) Human-level control through deep reinforcement learning. Nature 518(7540):529–533

Mohan S, Mininger AH, Kirk JR, Laird JE (2012) Acquiring grounded representations of words with situated interactive instruction. Adv Cogn Syst 2:113–130

Moon J, Anderson JR (2013) Timing in multitasking: memory contamination and time pressure bias. Cogn Psychol 67:26–54

Mora AM, Aisa F, García-Sánchez P, Castillo PÁ, Merelo JJ (2015) Modelling a human-like bot in a first person shooter game. Int J Creative Interfaces and Comput Graphics 6(1):21–37

Mueller ST, Minnery BS (2008) Adapting the turing test for embodied neurocognitive evaluation of biologically-inspired cognitive agents. In: Proceedings of th AAAI fall symposium on biologically inspired cognitive architectures

Murdock JW, Goel AK (2008) Meta-case-based reasoning: self-improvement through self-understanding. J Exp Theor Artif Intell 20(1):1–36

Murdock W, Goel A (2001) Meta-case-based reasoning : using functional models to adapt case-based agents. In: Proceedings of the 4th international conference on case-based reasoning

Murphy KN, Norcross RJ, Proctor FM (1988) CAD directed robotic deburring. In: Proceedings of the second international symposium on robotics and manufacturing research, education, and applications

Myers IB, McCaulley MH, Quenk NL, Hammer AL (1998) MBTI manual: a guide to the development and use of the Myers–Briggs type indicator, vol 3. Consulting Psychologists Press, Palo Alto

Myers KL, Martin DL, Morley DN (2002) Taskable reactive agent communities. Final technical report AFRL-IF-RS-TR-2002-208

Newell A (1980) Physical symbol systems. Cogn Sci 4(2):135–183

Newell A (1992) Précis of unified theories of cognition. Behav Brain Sci 15:425–492

Ng GW, Xiao X, Chan RZ, Tan YS (2012) Scene understanding using DSO cognitive architecture. In: Proceedings of the 15th international conference on information fusion (FUSION), pp 2277–2284

Nguyen SM, Ivaldi S, Lyubova N, Droniou A, Gerardeaux-Viret D, Filliat D, Padois V, Sigaud O, Oudeyer PY (2013) Learning to recognize objects through curiosity-driven manipulation with the iCub humanoid robot. In: Proceedings of the 3rd joint international conference on development and learning and epigenetic robotics

Niv Y (2009) Reinforcement learning in the brain. J Math Psychol 53(3):139–154

Novianto R (2014) Flexible attention-based cognitive architecture for robots. PhD thesis

Novianto R, Johnston B, Williams MA (2010) Attention in the ASMO cognitive architecture. Front Artif Intell Appl 221:98–105

Novianto R, Johnston B, Williams MA (2013) Habituation and sensitisation learning in ASMO cognitive architecture. Lect Notes Comput Sci 8239 LNAI:249–259

Nunez P, Manso LJ, Bustos P, Drews-Jr P, Macharet DG (2016) Towards a new semantic social navigation paradigm for autonomous robots using CORTEX. In: IEEE international symposium on robot and human interactive communication (RO-MAN 2016)—BAILAR2016 workshop

Nuxoll AM, Laird JE (2007) Extending cognitive architecture with episodic memory. In: Proceedings of the national conference on artificial intelligence

Nyamsuren E, Taatgen NA (2013) Pre-attentive and attentive vision module. Cogn Syst Res 211–216

Nyamsuren E, Taatgen NA (2014) Human reasoning module. Biol Inspired Cogn Archit 8:1–18

Ogasawara GH (1991) A distributed, decision-theoretic control system for a mobile robot. SIGART Bull 2(4):140–145

Ogasawara GH, Russell SJ (1993) Planning using multiple execution architectures. In: Proceedings of the international joint conference on artificial intelligence

O'Reilly RC (2006) Modeling integration and dissociation in brain and cognitive development. In: Processes of change in brain and cognitive development: attention and performance XXI, pp 375–401

O'Reilly RC, Frank MJ (2006) Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput 18(2):283–328

O'Reilly RC, Hazy TE, Herd SA (2012) The Leabra cognitive architecture: how to play 20 principles with nature and win! In: The Oxford handbook of cognitive science, pp 1–31

O'Reilly RC, Wyatte D, Herd S, Mingus B, Jilk DJ (2013) Recurrent processing during object recognition. Front Psychol 4:124

O'Reilly RC, Hazy TE, Mollick J, Mackie P, Herd S (2014) Goal-driven cognition in the brain: a computational framework. arXiv preprint arXiv:14047591

Ozturk P (2009) Levels and types of action selection: the action selection soup. Adapt Behav 17:537–554

Pacchierotti E, Christensen HI, Jensfelt P (2005) Embodied social interaction for service robots in hallway environments. In: Proceedings of the 5th international conference on field and service robotics

Paisner M, Cox MT, Maynord M, Perlis D (2013) Goal-driven autonomy for cognitive systems. In: Proceedings of the 36th annual conference of the cognitive science society, pp 2085–2090

Pape N, Urbas L (2008) A model of time-estimation considering working memory demands. In: Proceedings of the 30th annual conference of the cognitive science society, pp 1543–1548

Pattacini U, Nori F, Natale L, Metta G, Sandini G (2010) An experimental evaluation of a novel minimum-jerk Cartesian controller for humanoid robots. In: Proceedings of the international conference on intelligent robots and systems (IROS), pp 1668–1674

Perner A, Zeilinger H (2011) Action primitives for bionics inspired action planning system: abstraction layers for action planning based on psychoanalytical concepts. In: IEEE international conference on industrial informatics (INDIN), pp 63–68

Peters RA, Kawamura K, Wilkes DM, Hambuchen KA, Rogers TE, Alford WA (2001a) ISAC humanoid: an architecture for learning and emotion. In: Proceedings of the IEEE-RAS international conference on humanoid robots, 1, p 459

Peters RAI, Hambuchen KA, Kawamura K, Wilkes DM (2001b) The sensory ego-sphere as a short-term memory for humanoids. In: Proceedings of the IEEE-RAS international conference on humanoid robots

Petkov G, Naydenov T, Grinberg M, Kokinov B (2006) Building robots with analogy-based anticipation. In: Annual conference on artificial intelligence

Pezzulo G (2009) DiPRA: a layered agent architecture which integrates practical reasoning and sensorimotor schemas. Connect Sci 21(4):297–326

Pezzulo G, Calvi G (2005) Dynamic computation and context effects in the hybrid architecture AKIRA. In: International and interdisciplinary conference on modeling and using context

Pezzulo G, Calvi G, Castelfranchi C (2007) DiPRA: Distributed practical reasoning architecture. In: Proceedings of international joint conference on artificial intelligence (IJCAI), pp 1458–1463

Philips AB, Bresina JL (1991) NASA Tileworld. NASA technical report TR-FIA-91-04

Pirjanian P (1999) Behavior coordination mechanisms. Technical report IRIS-99-375

Pollock JL (1993a) Oscar—a general-purpose defeasible reasoner. AAAI technical report FS-93-01

Pollock JL (1993b) Planning in OSCAR. Mind Mach 2:113–144

Pollock JL (2008) OSCAR: an agent architecture based on defeasible reasoning. In: AAAI spring symposium: emotion, personality, and social behavior

Pollock JL, Hosea D (1995) OSCAR-MDA: an artificially intelligent advisor for emergency room medicine

Posner M, NIssen MJ, Klein RM (1976) Visual dominance: an information-processing account of its origins and significance. Psychol Rev 83(2):157–171

Profanter S (2012) Cognitive architectures. In: Hauptseminar human robot interaction

Pynadath DV, Rosenbloom PS, Marsella SC, Li L (2013) Modeling two-player games in the sigma graphical cognitive architecture. In: International conference on artificial general intelligence

Pynadath DV, Rosenbloom PS, Marsella SC (2014) Reinforcement learning for adaptive theory of mind in the sigma cognitive architecture. In: Proceedings of the conference on artificial general intelligence

Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton D, Colombo J, Coppola G, Geyer MA, Glanzman DL, Marsland S, Mcsweeney F, Wilson DA, Wu CF, Thompson RF (2009) Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92(2):135–138

Rao AS, George MP (1991) Intelligent real-time network management. In: Proceedings of the tenth international conference on AI, expert systems and natural language

Rasmussen D, Eliasmith C (2013) Modeling brain function current developments and future prospects. JAMA Neurol 70(10):1325–1329

Reisenzein R, Hudlicka E, Dastani M, Gratch J, Hindriks K, Lorini E, Meyer JJC (2013) Computational modeling of emotion: toward improving the inter- and intradisciplinary exchange. IEEE Trans Affect Comput 4(3):246–266

Riesenhuber M (2005) Object recognition in cortex: neural mechanisms, and possible roles for attention. In: Itti L, Rees G, Tsotsos JK (eds) Neurobiology of Attention. Academic Press, pp 279–287

Ritter FE (2009) Two cognitive modeling frontiers. Emotions and usability. Inf Media Technol 4(1):76–84

Ritter FE, Bittner JL, Kase SE, Evertsz R, Pedrotti M, Busetta P (2012) CoJACK: a high-level cognitive architecture with demonstrations of moderators, variability, and implications for situation awareness. Biol Inspired Cogn Archit 1:2–13

Rohrer B (2011a) A developmental agent for learning features, environment models, and general robotics tasks. ICDL/Eprirob

Rohrer B (2011b) An implemented architecture for feature creation and general reinforcement learning. In: Fourth international conference on artificial general intelligence, workshop on self-programming in AGI systems

Rohrer B (2011c) Biologically inspired feature creation for multi-sensory perception. Biol Inspired Cogn Archit 305–313

Rohrer B (2012) BECCA: Reintegrating AI for natural world interaction. In: AAAI spring symposium: designing intelligent robots, AAAI technical report SS-12-02

Rohrer B (2013) BECCA version 0.4.5. User's Guide

Rohrer B, Bernard M, Morrow DJ, Rothganger F, Xavier P (2009) Model-free learning and control in a mobile robot. In: Proceedings of the 5th international conference on natural computation, ICNC 2009, pp 566–572

Romero-Garcés A, Calderita LV, Martínez-Gómez J, Bandera JP, Marfil R, Manso LJ, Bandera A, Bustos P (2015a) Testing a fully autonomous robotic salesman in real scenarios. In: IEEE international conference on autonomous robots systems and competitions

Romero-Garcés A, Calderita LV, Martinez-Gomez J, Bandera JP, Marfil R, Manso LJ, Bustos P, Bandera A (2015b) The cognitive architecture of a robotic salesman. In: Conference of the Spanish association for artificial intelligence, vol 15(6)

Rosenbloom PS, Laird JE, Newell A, McCarl R (1991) A preliminary analysis of the Soar architecture as a basis for general intelligence. Artif Intell 47(1–3):289–325

Rosenbloom PS, Demski A, Ustun V (2015a) Efficient message computation in Sigma's graphical architecture. Biol Inspired Cogn Archit 11:1–9

Rosenbloom PS, Gratch J, Ustun V (2015b) Towards emotion in sigma: from appraisal to attention. In: International conference on artificial general intelligence

Rosenthal C, Congdon CB (2012) Personality profiles for generating believable bot behaviors. In: Proceedings of the IEEE conference on computational intelligence and games, pp 124–131. https://doi.org/10.1109/CIG.2012.6374147

Rousseau D, Hayes-Roth B (1996) Personality in synthetic agents. Report no KSL 96-21

Rousseau D, Hayes-roth B (1997) Interacting with personality-rich characters. Report no KSL 97-06

Ruesch J, Lopes M, Bernardino A, Hornstein J, Santos-Victor J, Pfeifer R (2008) Multimodal saliency-based bottom-up attention a framework for the humanoid robot iCub. In: Proceedings of the IEEE international conference on robotics and automation, pp 962–967

Ruiz D, Newell A (1989) Tower-noticing triggers strategy-change in the Tower of Hanoi: a Soar model. Technical report AIP-66, pp 522–529

Russel SJ, Wefald E (1988) Decision-theoretic control of reasoning: general theory and an application to game-playing. Technical report UCB/CSD 88/435

Russell S, Norvig P (1995) Artificial intelligence: a modern approach. Prentice Hall, Upper Saddle River

Russell S, Wefald E (1989) On optimal game-tree search using rational meta-reasoning. In: Proceedings of the international joint conference on artificial intelligence

Salgado R, Bellas F, Caamano P, Santos-Diez B, Duro RJ (2012) A procedural Long term memory for cognitive robotics. In: Proceedings of the IEEE conference on evolving and adaptive intelligent systems, pp 57–62

Salvucci DD (2000) A model of eye movements and visual attention. In: Proceedings of the third international conference on cognitive modeling, pp 252–259

Samsonovich AV (2010) Toward a unified catalog of implemented cognitive architectures. In: Proceeding of the conference on biologically inspired cognitive architectures, pp 195–244

Samsonovich AV, Ascoli Ga, Jong KaD, Coletti Ma (2006) Integrated hybrid cognitive architecture for a virtual roboscout. In: Beetz M, Rajan K, Thielscher M, Rusu R (eds) Cognitive robotics: papers from the AAAI workshop, vol 6. AAAI technical reports, AAAI Press, pp 129–134

Samsonovich AV, De Jong KA, Kitsantas A, Peters EE, Dabbagh N, Layne Kalbfleisch M (2008) Cognitive constructor: an intelligent tutoring system based on a biologically inspired cognitive architecture (BICA). Front Artif Intell Appl 171:311–325

Sandini G, Metta G, Vernon D (2007) The iCub cognitive humanoid robot: an open-system research platform for enactive cognition. Lect Notes Comput Sci. https://doi.org/10.1007/978-3-540-77296-5_32

Sanner S, Anderson JR, Lebiere C, Lovett MC (2000) Achieving efficient and cognitively plausible learning in backgammon. In: Proceedings of the seventeenth international conference on machine learning (ICML-2000)

Sanner SP (1999) A quick introduction to 4CAPS programming. http://www.ccbi.cmu.edu/4CAPS/4caps-manual-Sanner.pdf

Santore JF, Shapiro SC (2003) Crystal Cassie: use of a 3-D gaming environment for a cognitive agent. Papers of the IJCAI 2003 workshop on cognitive modeling of agents and multi-agent interactions

Sarathy V, Wilson JR, Arnold T, Scheutz M (2016) Enabling basic normative HRI in a cognitive robotic architecture. In: 2nd workshop on cognitive architectures for social human–robot interaction

Sauser EL, Argall BD, Metta G, Billard AG (2012) Iterative learning of grasp adaptation through human corrections. Robot Auton Syst 60:55–71

Scally JR, Cassimatis NL, Uchida H (2012) Worlds as a unifying element of knowledge representation. Biol Inspired Cogn Archit 1:14–22

Schaat S, Doblhammer K, Wendt A, Gelbard F, Herret L, Bruckner D (2013a) A psychoanalytically-inspired motivational and emotional system for autonomous agents. Industrial electronics society, IECON 2013-39th annual conference, pp 6648–6653

Schaat S, Wendt A, Bruckner D (2013b) A multi-criteria exemplar model for holistic categorization in autonomous agents. Industrial electronics society, IECON 2013-39th annual conference of the IEEE, pp 6642–6647

Schaat S, Wendt A, Jakubec M, Gelbard F, Herret L, Dietrich D (2014) ARS: an AGI agent architecture. Lect Notes Comput Sci 8598:155–164

Schaat S, Wendt A, Kollmann S, Gelbard F, Jakubec M (2015) Interdisciplinary development and evaluation of cognitive architectures exemplified with the SiMA approach. In: EuroAsianPacific joint conference on cognitive science

Schermerhorn P, Kramer J, Brick T, Anderson D, Dingler A, Scheutz M (2006) DIARC: a testbed for natural human-robot interactions. In: Proceedings of AAAI 2006 robot workshop, pp 1972–1973

Scheutz M, Schermerhorn P (2009) Affective goal and task selection for social robots. Handbook of research on synthetic emotions and sociable robotics: new applications in affective computing and artificial intelligence, p 74

Scheutz M, McRaven J, Cserey G (2004) Fast, reliable, adaptive, bimodal people tracking for indoor environments. In: Proceedings of IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 1347–1352

Scheutz M, Schermerhorn P, Kramer J, Anderson D (2007) First steps toward natural human-like HRI. Auton Robot 22(4):411–423

Scheutz M, Harris J, Schermerhorn P (2013) Systematic integration of cognitive and robotic architectures. Adv Cogn Syst 2:277–296

Scheutz M, Krause E, Sadeghi S (2014) An embodied real-time model of language-guided incremental visual search. In: Proceedings of the 36th annual meeting of the cognitive science society, pp 1365–1370

Schiller MRG, Gobet FR (2012) A comparison between cognitive and AI models of blackjack strategy learning. Lect Notes Comput Sci, pp 143–155

Schlenoff C, Madhavan R, Albus J, Messina E, Barbera T, Balakirsky S (2005) Fusing disparate information within the 4D/RCS architecture. In: Proceedings of the 7th international conference on information fusion

Schroder T, Thagard P (2014) Priming: constraint satisfaction and interactive competition. Soc Cogn 32:152–167

Seamster TL, Redding RE, Cannon JR, Ryder JM, Purcell JA (1993) Cognitive task analysis of expertise in air traffic control. Int J Aviat Psychol 3(4):257–283

Seidler RD, Kwak Y, Fling BW, Bernard JA (2013) Neurocognitive mechanisms of error-based motor learning. Adv Exp Med Biol 782:39–60

Selfridge O (1958) Pandemonium: a paradigm for learning in mechanisation of thought processes. In: Proceedings of a symposium held at the national physical laboratory

Seth AK, McKinstry JL, Edelman GM, Krichmar JL (2004) Visual binding through reentrant connectivity and dynamic synchronization in a brain-based device. Cereb Cortex 14(11):1185–1199

Shapiro D, Langley P, Shachter R (2001) Using background knowledge to speed reinforcement learning in physical agents. In: Proceedings of the 5th international conference on autonomous agents, pp 254–261

Shapiro SC, Bona JP (2010) The GLAIR cognitive architecture. Int J Mach Conscious 2(2):307–332

Shapiro SC, Kandefer M (2005) A SNePS approach to the wumpus world agent or cassie meets the wumpus. In: IJCAI-05 workshop on nonmonotonic reasoning, action, and change (NRAC'05): working notes

Shapiro SC, Anstey J, Pape DE, Nayak TD, Kandefer M, Telhan O (2005) MGLAIR agents in a virtual reality drama. CSE technical report 2005-08

Shastri L (1998) Types and quantifiers in SHRUTI—a connectionist model of rapid reasoning and relational processing. In: International workshop on hybrid neural systems

Shastri L (2007) SHRUTI: a neurally motivated architecture for rapid, scalable inference. Studies in computational intelligence. Springer, Berlin, pp 183–203

Shimojo S, Shams L (2001) Sensory modalities are not separate modalities: plasticity and interactions. Curr Opin Neurobiol 11(4):505–509

Shiwali M, Laird JE (2009) Learning to play Mario. Technical report CCA-TR-2009-03

Sjöö K et al (2010) The explorer system. In: Christensen HI, Kruijff GJM, Wyatt JL (eds) Cognitive systems. Springer, pp 395–421

Slam N, Wang W, Xue G, Wang P (2015) A framework with reasoning capabilities for crisis response decision-support systems. Eng Appl Artif Intell 46:346–353

Sloman A (2003) The cognition and affect project: architectures, architecture-schemas, and the new science of mind. Technical report

Small R, Congdon CB (2009) Agent Smith: towards an evolutionary rule-based agent for interactive dynamic games. In: 2009 IEEE congress on evolutionary computation, CEC 2009, pp 660–666

Smith RL, Gobet F, Lane PCR (2007) An investigation into the effect of ageing on expert memory with CHREST. In: Proceedings of the United Kingdom workshop on computational intelligence

Smith SDG, Escobedo R, Anderson M, Caudell TP (1997) A deployed engineering design retrieval system using neural networks. IEEE Trans Neural Netw 8(4):847–51

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

Starzyk JA, Graham JT (2015) MLECOG: motivated learning embodied cognitive architecture. IEEE Syst J 11(3):1272–1283

Stewart TC, Eliasmith C (2013) Parsing sequentially presented commands in a large-scale biologically realistic brain model. In: Proceedings of the 35th annual conference of the cognitive science society, pp 3460–3467

Stewart TC, Eliasmith C (2014) Large-scale synthesis of functional spiking neural circuits. Proc IEEE 102(5):881–898

Stewart TC, Blouw P, Eliasmith C (2015) Explorations in distributed recurrent biological parsing. In: International conference on cognitive modelling

Still A, D'Inverno M (2016) A history of creativity for future AI research. In: Proceedings of the 7th compute create configuration, pp 152–159

Stokes D, Biggs S (2014) The dominance of the visual. In: Stokes D, Matthen M, Biggs S (eds) Perception and its modalities. Oxford University Press, Oxford, pp 1–35

Stolc S, Bajla I (2010) Application of the computational intelligence network based on hierarchical temporal memory to face recognition. In: Proceedings of the 10th IASTED international conference on artificial intelligence and applications (AIA), pp 185–192

Sun R (1996) Hybrid connectionist-symbolic modules. AI Mag 17(2):99–103

Sun R (2004) Desiderata for cognitive architectures. Philos Psychol 17(3):341–373

Sun R (2007) The importance of cognitive architectures: an analysis based on CLARION. J Exp Theor Artif Intell 19:159–193

Sun R (2012) Memory systems within a cognitive architecture. New Ideas Psychol 30(2):227–240

Sun R (2016) Anatomy of the mind: exploring psychological mechanisms and processes with the clarion cognitive architecture. Oxford University Press, Oxford

Sun R, Bookman LA (eds) (1994) Computational architectures integrating neural and symbolic processes: a perspective on the state of the art. Springer, Berlin

Sun R, Fleischer P (2012) A cognitive social simulation of tribal survival strategies: the importance of cognitive and motivational factors. J Cogn Cult 12(3–4):287–321

Sun R, Helie S (2015) Accounting for creativity within a psychologically realistic cognitive architecture. Comput Creat Res Towards Creat Mach 7:3–36

Sun R, Wilson N (2011) Motivational processes within the perception-action cycle. Springer, New York

Sun R, Zhang X (2002) Top-Down versus bottom-up learning in skill acquisition. In: Proceedings of the 24th annual conference of the cognitive science society

Sun R, Zhang X (2003) Accessibility versus action-centeredness in the representation of cognitive skills. In: Proceedings of the fifth international conference on cognitive modeling

Sun R, Merrill E, Peterson T (1998) A bottom-up model of skill learning. In: Proceedings of 20th cognitive science society conference, pp 1037–1042

Sun R, Peterson T, Merrill E (1999) A hybrid architecture for situated learning of reactive sequential decision making. Appl Intell 11:109–127

Sun R, Zhang X, Mathews R (2006) Modeling meta-cognition in a cognitive architecture. Cogn Syst Res 7(4):327–338. https://doi.org/10.1016/j.cogsys.2005.09.001

Sun R, Wilson N, Mathews R (2011) Accounting for certain mental disorders within a comprehensive cognitive architecture. In: Proceedings of international joint conference on neural networks

Sun R, Wilson N, Lynch M (2016) Emotion: a unified mechanistic interpretation from a cognitive architecture. Cogn Comput 8(1):1–14

Szatmary B, Fleischer J, Hutson D, Moore D, Snook J, Edelman GM, Krichmar J (2006) A segway-based human-robot soccer team. In: IEEE international conference on robotics and automation

Taatgen NA (2002) A model of individual differences in skill acquisition in the Kanfer–Ackerman air traffic control task. Cogn Syst Res 3(1):103–112. https://doi.org/10.1016/S1389-0417(01)00049-3

Taigman Y, Ranzato MA, Aviv T, Park M (2014) DeepFace: closing the gap to human-level performance in face verification. In: CVPR

Taylor G, Padgham L (1996) An intelligent believable agent environment. AAAI technical report WS-96-03

Tecuci G, Kodratoff Y (1990) Apprenticeship learning in imperfect domain theories. In: Kondratoff Y, Michalski RS (eds) Machine learning, Vol III. Elsevier, pp 514–551

Tecuci G, Boicu M, Bowman M, Marcu D, Shyr P, Cascaval C (2000) An experiment in agent teaching by subject matter experts. Int J Hum Comput Stud 53(4):583–610

Tecuci G, Boicu M, Hajduk T, Marcu D, Barbulescu M, Boicu C, Le V (2007a) A tool for training and assistance in emergency response planning. In: Proceedings of the annual Hawaii international conference on system sciences, pp 1–10

Tecuci G, Marcu D, Boicu M, Le V (2007b) Mixed-initiative assumption-based reasoning for complex decision-making. Stud Inform Control 16(4):459–468

Tecuci G, Schum D, Boicu M, Marcu D, Hamilton B (2010) Intelligence analysis as agent-assisted discovery of evidence, hypotheses and arguments. Smart Innov Syst Technol 4:1–10

Tecuci G, Boicu M, Marcu D, Schum D (2013) How learning enables intelligence analysts to rapidly develop practical cognitive assistants. In: Proceedings of the 12th international conference on machine learning and applications, pp 105–110

Thagard P (2012) Cognitive architectures. In: Frankish W, Ramsay W (eds) The Cambridge handbook of cognitive science. Cambridge University Press, Cambridge, pp 50–70

Thibadeau R, Just MA, Carpenter PA (1982) A model of the time course and content of reading. Cogn Sci 6:157–203

Thomson R, Bennati S, Lebiere C (2014) Extending the influence of contextual information in ACT-R using buffer decay. In: Proceedings of the annual meeting of the cognitive science society

Thórisson K, Helgasson H (2012) Cognitive architectures and autonomy: a comparative review. J Artif Gen Intell 3(2):1–30

Thorisson KR (1997) Layered modular action control for communicative humanoids. In: Conference proceedings of computer animation, pp 134–143

Thorisson KR (1998) Real-time decision making in multimodal face-to-face communication. In: Proceedings of the international conference on autonomous agents, pp 16–23

Thorisson KR (1999) Mind model for multimodal communicative creatures and humanoids. Appl Artif Intell 13(4–5):449–486

Thorisson KR, Gislason O, Jonsdottir GR, Thorisson HT (2010) A multiparty multimodal architecture for realtime turntaking. In: International conference on intelligent virtual agents

Thornton J, Faichney J, Blumenstein M, Hine T (2008) Character recognition using hierarchical vector quantization and temporal pooling. In: Proceedings of the 21st Australasian joint conference on artificial intelligence: advances in artificial intelligence, vol 5360, pp 562–572

Tikhanoff V, Cangelosi A, Metta G (2011) Integration of speech and action in humanoid robots: iCub simulation experiments. IEEE Trans Auton Ment Dev 3(1):17–29

Trafton JG, Harrison AM (2011) Embodied spatial cognition. Top. Cogn Sci 3:686–706

Trafton JG, Cassimatis NL, Bugajska MD, Brock DP, Mintz FE, Schultz AC (2005) Enabling effective human robot interaction using perspective-taking in robots. IEEE Trans Syst Man Cybern Part A Syst Hum 35(4):460–470

Trafton JG, Hiatt LM, Harrison AM, Tamborello P, Khemlani SS, Schultz AC (2013) ACT-R/E: an embodied cognitive architecture for human-robot interaction. J Hum Robot Interact 2(1):30–54

Triona LM, Masnick AM, Morris BJ (2001) What does it take to pass the false belief task? ACT-R Model 72(15):213

Tripp B, Eliasmith C (2016) Function approximation in inhibitory networks. Neural Netw 77:95–106

Trivedi N, Langley P, Schermerhorn P, Scheutz M (2011) Communicating, interpreting, and executing high-level instructions for human-robot interaction. In: Proceedings of AAAI fall symposium: advances in cognitive systems

Tsotsos JK (1990) Analyzing vision at the complexity level. Behav Brain Sci 13:423–469

Tsotsos JK (1992) Image understanding. In: Shapiro S (ed) The encyclopedia of artificial intelligence, 2nd edn. Wiley, pp 641–663

Tsotsos JK (2011) A computational perspective on visual attention. MIT Press, Cambridge

Tsotsos JK (2017) Attention and cognition: principles to guide modeling. In: Zhao Q (ed) computational and cognitive neuroscience of vision. Elsevier, New York City

Tsotsos JK, Kruijne W (2014) Cognitive programs: software for attention's executive. Front Psychol 5:1–16

Tsotsos JK, Culhane SM, Kei Wai WY, Lai Y, Davis N, Nuflo F (1995) Modeling visual attention via selective tuning. Artif Intell 78(1–2):507–545

Tulving E (1972) Episodic and semantic memory. In: Tulving E, Donaldson W (eds) Organization of memory. Academic Press, Cambridge, pp 382–402

Tyler SW, Neukom C, Logan M, Shively J (1998) The MIDAS human performance model. In: Proceedings of the human factors and ergonomics society, pp 320–324

Ulam P, Goel A, Jones J (2004) Reflection in action: model-based self-adaptation in game playing agents. In: challenges in game artificial intelligence: papers from the AAAI workshop

Ulutas B, Erdemir E, Kawamura K (2008) Application of a hybrid controller with non-contact impedance to a humanoid robot. In: Proceedings of the IEEE 10th international workshop on variable structure systems, pp 378–383

Ustun V, Rosenbloom PS, Kim J, Li L (2015) Building high fidelity human behavior models in the sigma cognitive rchitecture. In: Yilmaz L, Chan WKV, Moon I, Roeder TMK, Macal C, Rossetti MD (eds) Proceedings of the 2015 winter simulation conference

Van Hoorn N, Togelius J, Schmidhuber J (2009) Hierarchical controller learning in a first-person shooter. In: 2009 IEEE symposium on computational intelligence and games, pp 294–301

VanLehn K (1989) Discovering problem solving strategies: What humans do and machines don't (yet). In: Proceedings of the sixth international workshop on machine learning, pp 215–217

VanLehn K, Ball W (1989) Goal reconstruction: how teton blends situated action and planned action. Technical report, Department of Computer Science and Psychology, Carnegie Mellon University

VanLehn K, Ball W, Kowalski B (1989) Non-lifo execution of cognitive procedures. Cogn Sci 13(3):415–465

Vanlehn K, Ball W, Kowalski B (1990) Explanation-based learning of correctness: towards a model of the self-explanation effect. In: Proceedings of the 12th annual conference of the cognitive science society

Varma S (2006) A computational model of Tower of Hanoi problem solving. PhD thesis

Veloso M (1993) PRODILOGY/ANALOGY: analogical reasoning in general problem solving. In: Topics in case-based reasoning

Veloso MM, Blythe J (1994) Linkability: examining causal link commitments in partial-order planning. In: Proceedings of the second international conference on artificial intelligence planning systems

Veloso MM, Pollack ME, Cox MT (1998) Rationale-based monitoring for planning in dynamic environments. In: AIPS 1998 proceedings, pp 171–180

Vere S, Bickmore T (1990) A basic agent. Comput Intell 6(1):41–60

Vere SA (1991) Organization of the basic agent. ACM SIGART Bull 2(4):164–168

Vernon D, Metta G, Sandini G (2007) A survey of artificial cognitive systems: implictions for the autonomous development of mental capabilities in computational agents. IEEE Trans Evol Comput 1–30

Vernon D, von Hofsten C, Fadiga L (2010) The iCub cognitive architecture. In: A roadmap for cognitive development in humanoid robots, pp 121–153

Verschure P, Althaus P (2003) A real-world rational agent: unifying old and new AI. Cogn Sci 27(4):561–590

Vinokurov Y, Lebiere C, Szabados A, Herd S, O'Reilly R (2013) Integrating top-down expectations with bottom-up perceptual processing in a hybrid neural-symbolic architecture. Biol Inspired Cogn Archit 6:140–146

Vouloutsi V, Munoz MB, Grechuta K, Lallee S, Duff A, ysard Llobet Puigbo J, Verschure PFMJ (2015) A new biomimetic approach towards educational robotics: the distributed adaptive control of a synthetic tutor assistant. In: 4th international symposium on new frontiers in human–robot interaction

Walther D, Itti L, Riesenhuber M, Poggio T, Koch C (2002) Attentional selection for object recognition a gentle way. In: International workshop on biologically motivated computer vision

Walther DB, Koch C (2007) Attention in hierarchical models of object recognition. Prog Brain Res 165:57–78

Wang D, Subagdja B, Tan Ah, Ng G (2009) Creating human-like autonomous players in real-time first person shooter computer games. In: Proceedings of the 21st annual conference on innovative applications of artificial intelligence, pp 173–178

Wang J, Naghdy G, Ogunbona P (1997) Wavelet-based feature-adaptive adaptive resonance theory neural network for texture identification. J Electron Imaging 6(3):329–336

Wang P (2006) Rigid flexibility: the logic of intelligence, vol 34. Springer, Netherlands

Wang P (2007) Three fundamental misconceptions of artificial intelligence. J Exp Theor Artif Intell 19(3):249–268

Wang P (2010) Non-axiomatic logic (NAL) specification

Wang P (2013) Natural language processing by reasoning and learning. In: Proceedings of the international conference on artificial general intelligence, pp 160–169

Wang P, Hammer P (2015a) Assumptions of decision-making models in AGI. In: International conference on artificial general intelligence, pp 197–207

Wang P, Hammer P (2015b) Issues in temporal and causal inference. In: Proceedings of the international conference on artificial general intelligence

Wang Y, Laird JE (2006) Integrating semantic memory into a cognitive architecture. Technical report CCA-TR-2006-02

Wendelken C, Shastri L (2005) Connectionist mechanisms for cognitive control. Neurocomputing 65–66:663–672

Wendelken JC (2003) SHRUTI-agent: a structured connectionist architecture for reasoning and decision-making. PhD thesis

Weng J (2002) A theory for mentally developing robots. In: Proceedings of the 2nd international conference on development and learning, pp 131–140

Weng J, Hwang WS (2006) From neural networks to the brain: autonomous mental development. IEEE Comput Intell Mag 1(3):15–31

Weng J, Hwang WS (2007) Incremental hierarchical discriminant regression. IEEE Trans Neural Netw 18(2):397–415

Weng J, Luciw M (2010) Online learning for attention, recognition, and tracking by a single developmental framework. In: Proceedings of the conference on computer vision and pattern recognition

Weng J, Zhang Y (2002) Developmental robots–a new paradigm. In: Proceedings of the second international workshop on epigenetic robotics modeling cognitive development in robotic systems, vol 94, pp 163–174

Weng J, Lee YB, Evans CH (1999) The developmental approach to multimedia speech learning. In: Proceedings of the IEEE international conference on acoustics, speech, and signal processing

Wentura D, Rothermund K (2014) Priming is not priming is not priming. Soc Cogn 32:47–67

Wermter S (1997) Hybrid approaches to neural network-based language processing. Technical report TR-97-030

Wermter S, Sun R (eds) (2000) Hybrid neural systems. Springer, Berlin

Wickens CD, Mccarley JS, Alexander AL, Thomas LC, Ambinder M, Zheng S (2008) Attention-situation awareness (A-SA) model of pilot error. In: Human performance modeling in aviation, pp 213–239

Williams T, Scheutz M (2016) A framework for resolving open-world referential expressions in distributed heterogeneous knowledge bases. In: Proceedings of the thirtieth AAAI conference on artificial intelligence

Williams T, Briggs G, Oosterveld B, Scheutz M (2015) Going beyond literal command-based instructions: extending robotic natural language interaction capabilities. AAAI, pp 1387–1393

Wilson JR, Scheutz M (2014) Analogical generalization of activities from single demonstration. In: Proceedings of Ibero-American conference on artificial intelligence, pp 637-648. https://doi.org/10.1007/978-3-319-12027-0

Wilson JR, Forbus KD, McLure MD (2013) Am I really scared? A multi-phase computational model of emotions. In: Proceedings of the second annual conference on advances in cognitive systems, pp 289–304

Wilson JR, Krause E, Rivers M, Scheutz M (2016) Analogical generalization of actions from single exemplars in a robotic architecture. In: Proceedings of the 2016 international conference on autonomous agents & multiagent systems

Wilson NR, Sun R (2014) Coping with bullying: A computational emotion-theoretic account. In: CogSci

Wilson NR, Sun R, Mathews RC (2009) A motivationally-based simulation of performance degradation under pressure. Neural Netw 22(5–6):502–508

Wilson NR, Sun R, Mathews RC (2010) A motivationally based computational interpretation of social anxiety induced stereotype bias. In: Proceedings of the 2010 cognitive science society conference, pp 1750–1755

Wintermute S (2009) An overview of spatial processing in Soar/SVS investigator. Technical report CCA-TR-2009-01

Wintermute S (2012) Imagery in cognitive architecture: representation and control at multiple levels of abstraction. Cogn Syst Res 19–20:1–29

Wolf MT, Assad C, Kuwata Y, Howard A, Aghazarian H, Zhu D, Lu T, Trebi-Ollennu A, Huntsberger T (2010) 360-Degree visual detection and target tracking on an autonomous surface vehicle. J Field Robot 27(6):819–833. https://doi.org/10.1002/rob.20371

Wolfe JM (1994) Guided search 2.0 A revised model of visual search. Psychon Bull Rev 1(2):202–238

Wong C, Kortenkamp D, Speich M (1995) A mobile robot that recognizes people. In: Proceedings of 7th IEEE international conference on tools with artificial intelligence

Wyatte D, Herd S, Mingus B, O'Reilly R (2012) The role of competitive inhibition and top-down feedback in binding during object recognition. Front Psychol 3:182

Xiao X, Ng GW, Tan YS, Chuan YY (2015) Scene parsing and fusion-based continuous traversable region formation. In: Jawahar C et al. (eds) Computer vision—ACCV 2014 workshops

Yen J, McNeese M, Mullen T, Hall D, Fan X, Liu P (2010) RPD-based hypothesis reasoning for cyber situation awareness. In: Cyber situational awareness, pp 39–49

Yu C, Scheutz M, Schermerhorn P (2010) Investigating multimodal real-time patterns of joint attention in an HRI word learning task. In: 5th ACM/IEEE international conference on human–robot interaction (HRI), 2010, pp 309–316

Zachary W, Santarelli T, Ryder J, Stokes J (2000) Developing a multi-tasking cognitive agent using the COGNET/iGEN integrative architecture. Technical report

Zachary WW, Zaklad AL, Hicinbothom JH, Ryder JM, Purcell JA (1993) COGNET reprezentation of tactical decision-making in anti-air warfare. In: Proceedings of the human factors and ergonomics society 37th annual meeting, pp 1112–1116

Zachary WW, Ryder JM, Hicinbothom JH, Cannon-Bowers JA, Salas E (1998) Cognitive task analysis and modeling of decision making in complex environments. In: Cannon-Bowers J, Salas E (eds) Making decisions under stress: implications for individual and team training, Washington, DC, pp 315–344

Zachary WW, Mentec JCL, Ryder JM (2016) Interface agents in complex systems. Hum Interact Complex Syst 14(1):260–264

Zhang N, Weng J, Zhang Z (2002) A developing sensory mapping for robots. In: Proceedings 2nd international conference on development and learning. ICDL 2002, pp 13–20

Zhang Y, Weng J (2007) Task transfer by a developmental robot. IEEE Trans Evol Comput 11(2):226–248

Zhuo W, Cao Z, Qin Y, Yu Z, Xiao Y (2012) Image classification using HTM cortical learning algorithms. In: Proceedings of the 21st international conference on pattern recognition (ICPR), pp 2452–2455

Zmigrod S, Hommel B (2013) Feature integration across multimodal perception and action: a review. Multisens Res 26:143–157