University of Memphis

University of Memphis Digital Commons

CCRG Papers

Cognitive Computing Research Group

2010

Sensory and perceptual scene representation

R. McCall

J. Snaider

S. Franklin

Follow this and additional works at: https://digitalcommons.memphis.edu/ccrg_papers

Recommended Citation

McCall, R., Snaider, J., & Franklin, S. (2010). Sensory and perceptual scene representation. Retrieved from https://digitalcommons.memphis.edu/ccrg_papers/72

This Document is brought to you for free and open access by the Cognitive Computing Research Group at University of Memphis Digital Commons. It has been accepted for inclusion in CCRG Papers by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

Sensory and Perceptual Scene Representation

Ryan McCall, Javier Snaider, Stan Franklin

Abstract

Intelligent agents (agents) in complex "real-world" environments must generate their own understanding of their current situation from their sensory primitives. For some modalities, especially vision, the process of creating meaning and understanding is exceedingly complex. We present a general framework describing how sensory data can be processed into meaningful information at various levels of abstraction. It proposes ways of representing this perceptual information at different processing states, and details how the more abstract internal representations are grounded and maintain connection to their more detailed sensory underpinnings. The framework's role within a comprehensive cognitive architecture is discussed.

Introduction

Every biological or artificial intelligent agent within a complex environment must continually make sense of its current situations. Without a useful model of the current situation it cannot effectively chose actions to further its agenda. In short, without having a useful "map" of their "territory" it is at the mercy of chance. Such a perceptual map must accommodate a vast, but manageable, variety of objects, categories, actions, feelings, events, etc., as well as their accompanying sensory information. Inputs from the various senses are often, as is the case for humans, continuous, and dynamic while comprising huge amounts of data every tenth of a second or so. To construct a useful model of such a current situation will require suitable representations both at the conceptual and the sensory level. Here we will describe the "Scene Representation" framework, composed of a series of representations that allows for higher-level perception of objects, actions, and events, all of which are grounded in the meticulous details of the sensory stimuli. For example, in vision these representations might include size, location, shape, color, texture, and shading, along with spatial and geometric information. The scene representation framework will be illustrated using a broad, comprehensive cognitive model, LIDA, together with its computational architecture (Franklin and Patterson 2006; Ramamurthy et al. 2006; Baars and Franklin 2009).

We will start by giving brief descriptions of the major concepts related to scene representation. The *Sensory Scene* is a multimodal representation scheme that preserves low-level details of its stimuli, but also allows for processed components as well. In terms of the LIDA model, the Sensory Scene is part of Sensory Memory, which holds the rawest data registered by an agent's sensors as well as more processed information. For the visual modality, the Sensory Scene is composed of layers that correspond to various stages in visual processing. The *Perceptual Scene* consists of portions of the Sensory Scene, and also contains conceptual nodes (instantiated nodes from LIDA's Perceptual Associative Memory, see below). The conceptual part of the Perceptual Scene is termed the *node layer*. As we will describe, the conceptual node layer is grounded in the layers of the Sensory Scene and in Perceptual Associative Memory. The Perceptual Scene, along with what we will define as complex structures in the Current Situational Model of the LIDA Workspace, represent the agent's understanding of the current situation. We divide the Perceptual Scene into a "real" and a "virtual" part. The real part of the Perceptual Scene consists of information originating from the Sensory Scene. Such information is derived from stimuli from the external environment. The virtual part consists of representations of imagination and/or episodic recall. The key distinction is that representations in the virtual part are not directly based on external environmental stimuli while those in the real part are. Some of the sensory information in the virtual scene is hypothesized to be time shared between the virtual and real parts. So the information from these layers can be a part of the real scene or the virtual scene.

The LIDA Model and its Architecture

The LIDA model (Franklin and Patterson 2006; Ramamurthy et al. 2006; Baars and Franklin 2009) is a comprehensive, conceptual and computational model covering a large portion of human cognition[1]. Based primarily on Global Workspace theory (Baars 1988, 2002), the model implements and fleshes out a number of psychological and neuropsychological theories. The LIDA computational architecture is derived from the LIDA cognitive model. The LIDA model and its ensuing architecture are grounded in the LIDA cognitive cycle. Every autonomous agent (Franklin and Graesser 1997), be it human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). More sophisticated agents, such as humans, process (make sense of) the input from such sampling in order to facilitate their decision making. The agent's "life" can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle constitutes a unit of sensing, attending and acting. A cognitive cycle can be thought of as a moment of cognition, a cognitive "moment."

We will now briefly describe what the LIDA model hypothesizes as the rich inner structure of the LIDA cognitive cycle. More detailed descriptions are available elsewhere (Baars and Franklin 2003, Franklin et al. 2005). During each cognitive cycle the LIDA agent first makes sense of its current situation as best as it can *by updating its representation of its current situation, both external and internal*. By a competitive process, as specified by Global Workspace Theory, it then decides what portion of the represented situation is most in need of attention. Broadcasting this portion, the current contents of consciousness[2], enables the agent to chose an appropriate action and execute it, completing the cycle.

Thus, the LIDA cognitive cycle can be subdivided into three phases, the understanding phase, the attention (consciousness) phase, and the action selection phase. Figure 1 should help the reader follow the description. It starts in the upper left corner and proceeds roughly clockwise. Beginning the understanding phase, incoming stimuli activate low-level feature detectors in Sensory Memory. The output is sent to Perceptual Associative Memory where higher-level feature detectors feed in to more abstract entities such as objects, categories, actions, events, etc. The resulting percept moves to the Workspace where it cues both Transient Episodic Memory and Declarative Memory producing local associations. These local associations are combined with the percept to generate a Current Situational Model; the agent's understanding of what is going on right now.

---

[1] "Cognition" is used here in a particularly broad sense, so as to include perception, feelings and emotions.

[2] Here "consciousness" refers to functional consciousness (Franklin 2003). We take no position on the need for, or possibility of, phenomenal consciousness.

Attention Codelets[3] begin the attention phase by forming coalitions of selected portions of the Current Situational Model and moving them to the Global Workspace. A competition in the Global Workspace then selects the most salient, the most relevant, the most important, the most urgent coalition whose contents become the content of consciousness. These conscious contents are then broadcast globally, initiating the action selection phase. The action selection phase of LIDA's cognitive cycle is also a learning phase in which several processes operate in parallel (see Figure 1). New entities and associations, and the reinforcement of old ones, occur as the conscious broadcast reaches Perceptual Associative Memory. Events from the conscious broadcast are encoded as new memories in Transient Episodic Memory. Possible action schemes, together with their contexts and expected results, are learned into Procedural Memory from the conscious broadcast. Older schemes are reinforced. In parallel with all this learning, and using the conscious contents, possible action schemes are recruited from Procedural Memory. A copy of each such is instantiated with its variables bound and sent to Action Selection, where it competes to be the behavior selected for this cognitive cycle. The selected behavior triggers Sensory-Motor Memory to produce a suitable algorithm for the execution of the behavior. Its execution completes the cognitive cycle.

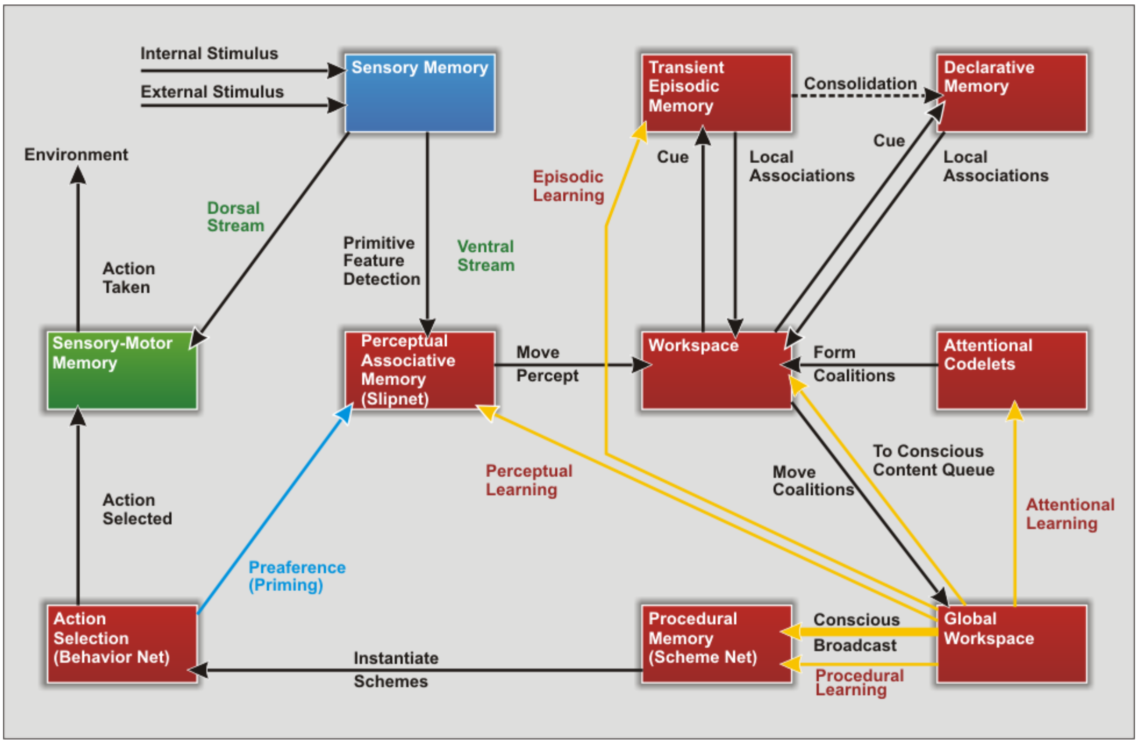


Figure 1. The LIDA Cognitive Cycle Diagram

The Workspace requires further explanation (see Figure 2). Its internal structure is composed of various input buffers, perceptual, and episodic, as well as the Current Situational Model, the Scratchpad and the Conscious Contents Queue. The Current Situational Model, a major subject of this paper, is where the structures representing the actual current internal and external events are stored. Structure building codelets are responsible for the creation of these structures using elements from the various sub-

3 A codelet is a small piece of code that performs a specific task in an independent way. It could be interpreted as a small part of a bigger process, similar to an ant in an ant colony.

modules of the Workspace. The scratchpad is an auxiliary space in the workspace where structure building codelets can construct possible structures prior to moving them to the Current Situational Model. The conscious contents queue holds the contents of the last several broadcasts and permits LIDA, using codelets, to understand and operate upon time related concepts (Snaider, McCall and Franklin 2009).

Sensory Memory

The Sensory Memory is the part of an agent's control structure that contains the rawest representation of data brought in by an agent's sensors. Each specialized set of sensors senses a different aspect of the real world. This is called the modality of the stimulus. Information is produced by the agent internally from this data, as well as from already present information (Oyama 1985). For example, visual sensors sense light from the real world and generate a raw representation of an image. In computational terms, a raw image is a matrix of dots called pixels.

Sensory Memory acts as a buffer for the sensory data, as well as the early information derived from it. The raw data is processed to produce a variety of supplementary representations that we called *layers*. Examples of layers are structures such as matrices containing data about motion or segmentation. In the LIDA model, Sensory Memory provides executing sensory-motor automatisms with detailed information that is used to guide the execution of actions.

Sensory Scene Representation

Previously, node and link representations had been used exclusively in LIDA. However, spatial reasoning and performing actions with only such representations appears to be difficult. Nodes and links are poorly suited for retaining the low-level features of stimuli such as detailed spatial location.

The Sensory Scene is a proposed structure for representing sensory information. We primary think of its use in representing information in the visual modality, but it could be extended to other modalities as well. The Sensory Scene is comprised of a set of layers each containing the results of different types of early visual processing. Both these layers and the processes producing them are part of Sensory Memory. These layers preserve the spatial relations of the related data. For example if one layer represents the lowest pixel representation of an image, another layer can hold the color of these pixels, and yet another would represent their movement. So, for the same pixel, or, in a more generic sense, the same region of the visual field, the different layers represents different low-level information like color or movement. Each layer can obtain its data directly from the rawest layers, from higher layers or from a combination of layers.

.

In order to better understand the Sensory Scene representation it is helpful to know how it is produced. The following paragraphs describe some of the layers that could comprise the Sensory Scene. Other layers can be added to the scene allowing the agent to perceive different aspects (e.g., sensory modalities) of the real world.

It is first assumed that visual input is fed into the system as a 2D matrix of pixels that we have termed a *pixel layer*. In the human visual system such input is constantly incoming from the eyes thus we picture the pixel layer being updated asynchronously. This spatial representation could be a 3D or 2.5D one also. (Marr 1982). However for this work we only address a 2D representation.

Motion can be detected in the visual stimulus from changes in the pixel layer. The result of such processing is represented in what we call the *motion vector field*. In computer vision the motion vector field is known as a flow field (Lishman, 1981). The motion vector field is another representational layer whose geometry corresponds to the agent's visual field. Algorithms in Sensory Memory identify changes in the visual field. *The location of detected motion is stored in the region of motion vector field that corresponds to the region of the pixel layer where the motion occurred*. In this way, the motion has a precise spatial grounding based on where it was sensed. Updates to the motion vector field with new regions of motion occur asynchronously. The more frequently this layer, and representations in general, are updated based upon their lower level groundings, the more up-to-date the representation will be. However, different layers need not be updated at the same rate. As long as the updating between layers occurs frequently enough, the processing in each of these matrices can take place asynchronous.

Another processing step takes the pixel layer as input and produces a color layer. A similar step produces a texture layer. First, an algorithm detects the region(s) of the pixel layer where a particular feature is present. For example a circular region of red may be detected in the upper right corner of the pixel layer. In this case the upper right corner of the color layer is updated to contain a circular region of red. So not only are features identified, their spatial locations are retained as well.

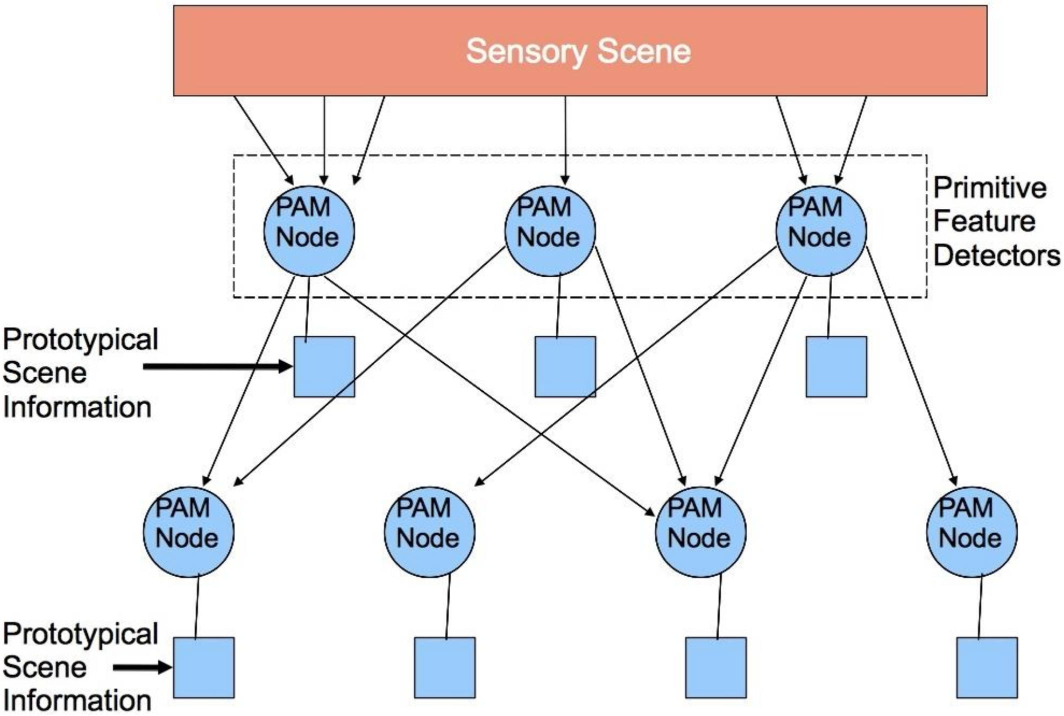
Using a combination of the motion, color, texture, and other such layers, an edge detection algorithm can find the boundaries of possible objects and store them into another layer termed the *segmented layer*. At this point the visual image has been effectively segmented into regions. Being part of Sensory Memory, the information in the motion vector field, the segmented layer, and the feature layer is frequently sent to Sensory-Motor Memory in order to guide action execution. All are part of the Sensory Scene representation.

Using the segmented, motion, and feature matrices, object detection algorithms can be used to identify regions of the image that are believed to be specific objects. Not all such regions may be identified as objects since the region may be part of the background.

In summary, the Sensory Scene is comprised of a set of layers containing different kinds of processed information. Each layers in ultimately grounded in the pixel layer, a short-term store of the raw data from the agent's visual field. Thus, the Sensory Scene contains processed elements including color and texture while keeping the detailed geometry of the elements. The size, shape, and spatial relations of the objects, shapes, and features are directly encoded in the representation.

Feature detectors

Perceptual Associative Memory (PAM) is composed of nodes and links. Each node in PAM represents a concept or feature that the agent is able to recognize. PAM nodes have a base-level activation measuring their general usefulness, as well as a current activation signifying their relevance in the current situation. Feature detectors are represented by nodes in Perceptual Associative Memory (PAM). Each feature detector has a receptive field from which it derives its current activation. Receptive fields can include other PAM nodes, portions of Sensory Memory, or both. Primitive feature detectors are on the periphery of PAM, that is, they are nodes whose receptive fields lie



solely within Sensory Memory (Fig. 2). Thus no primitive feature detector nodes obtain their activation from other nodes. Each non-primitive feature detector node can be viewed as the root node of a tree of nodes whose leaves are primitive feature detector nodes.

Figure 2. Perceptual Associative Memory

We envision a primitive feature detector being implemented by a demon-like process. Each such process would have its own receptive field in the Sensory Scene and a recognition algorithm for the feature it detects. Each process would constantly monitor its receptive field in the Sensory Scene, using its recognition algorithm to determine the extent to which its feature of concern is present. The output of this process is activation of the primitive feature detector's PAM node. As a result the node may have various degrees of activation; however, if the PAM node's activation goes over its threshold the feature is considered to be detected. Thus each primitive feature detector consists of an algorithm to determine what and how it detects, as well as a PAM node whose activation is updated to reflect the degree of detection.

Perceptual Scene

Nodes over their threshold “fire” sending a part of their total activation to their neighbors. Thus activation from primitive feature detector nodes spreads through PAM

(Figure 2). This allows PAM nodes, which may have multiple connections to the primitive feature detector nodes, to gain current activation. If a PAM node's total activation, the sum of its base-level and current activation, is over threshold, an instantiation of it becomes part of the percept. PAM node instantiations are copies of the original PAM nodes. Every instantiated node refers to one or more regions within Sensory Scene layers that were responsible for activating it. Assuming the detection algorithm is accurate, whatever feature or object that node represents must be present in these grounding layers. When PAM nodes are instantiated, these regions are attached to them. They both are part of the percept as well, and become part of the Perceptual Scene. Instantiated PAM nodes also maintain a connection to the original node in PAM. This connection allows structure-building codelets to be able to retrieve additional sensory information.

Within the Perceptual scene, instantiated PAM nodes are said to be a part of the *node layer* representation (see Figure 3). At a high-level, this entire process can be thought of as the layers of the Sensory Scene activating PAM nodes, which represent additional and higher-level understanding of the Sensory Scene.

In summary, the object recognition process is initiated by feature detectors operating on the layers of the Sensory Scene with the result, the instantiated PAM nodes and their Sensory Scene grounding, being instantiated in the Perceptual Scene.

For each region in the segmented layer that might correspond to an object, a *blank node* is added to the node layer (see Figure 3). If a node for a recognized object already exists in PAM, then the corresponding node in PAM is instantiated in the node layer of the Perceptual Scene, and is linked to the blank node of the region corresponding to this object. If the region is not recognized as an existing object in PAM then the blank node may be identified as a new object by a structure building codelet. If this new object, together with its sensory region in the segmented layer, is brought to consciousness by an attention codelet, then the new object will be learned into PAM.

Thus blank nodes allow instantiated nodes to be attached to unrecognized objects, and even to new kinds of objects allowing unrecognized objects and actions in the Sensory Scene to be recognized in PAM and new objects and events to be learned. Previously learned objects and events which have been learned into PAM may be recognized directly.

Blank nodes also support the representation of many similar objects in the Current Situational Model. For example, if the agent perceives a group of cups, all with the same form, color, etc, the agent's Current Situational Model can differentiate them. Each cup corresponds to a region in the segmented layer and a blank node represents each of them. Then instances of nodes that represent or describe this object are linked with each particular blank node.

We hypothesize in the LIDA model that visual representations in the Sensory Scene, including parts of the color, texture and segmented layers almost always have enough activation to enter into the preconscious Workspace as part of the Perceptual Scene.

One way of looking at this activation passing is that the Perceptual Scene continually "picks up" subgraphs from PAM, together with their higher-level sensory information.

The subgraphs contain those nodes in PAM that received sufficient activation from layers of the Sensory Scene.

Instantiated nodes and links maintain a connection with region(s) of the Sensory Scene. If the pattern of information the Sensory Scene that led to the detection of a feature disappears or dies out, the instantiated node for that feature ought to lose activation. Alternatively, if an object changes location, the instantiated node ought to remain activated, albeit, the Sensory Scene grounding will have changed. In this way the high-level representations in the Workspace would remain grounded.

The sensory information that is attached to instantiated nodes comes mainly from the Sensory Scene. In some cases, for example when only partial data is available, *prototypical* sensory information stored with PAM nodes is also used. Moreover, this prototypical sensory information can be used as the default information for a PAM node and can be combined with the real sensory data to complete the representation. This prototypical sensory information can also be used to help construct representations in the virtual window in the perceptual scene. For example, consider the presence of instantiated nodes in the Workspace from a local association. These instantiated nodes may not contain current sensory information. However, they may incorporate the prototypical sensory information retrieved from the original node in PAM by default.

The Perceptual Scene, part of the Current Situational Model inside the Workspace, is accessible to attention codelets and structure-building codelets. The Perceptual Scene is comprised of the node layer together with sensory information from the higher layers of the Sensory Scene (see Figure 3). Thus, some information from the Sensory Scene is shared with the Perceptual Scene. Summing up, the Sensory Scene comprises the rawest layers of data originated by sensors and middle layers with processed data. On the other hand, the Perceptual Scene includes information from the middle layers and the node layer.

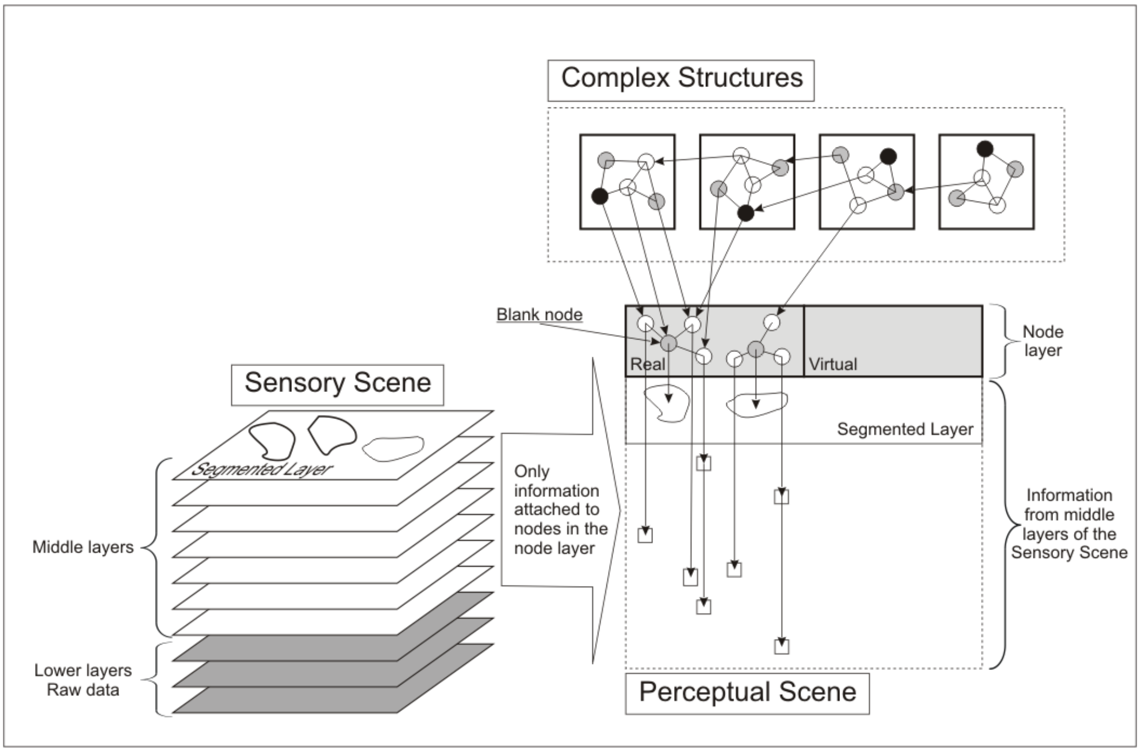


Figure 3. Complex structures

Complex structures

To review, the node layer of the Perceptual Scene contains instantiated nodes and links, which represent the objects, features, and events of the Perceptual Scene. Each node in this layer maintains a reference to its corresponding PAM node. However these nodes and links are only of those objects, events, etc. that feature detectors and PAM could detect from the Sensory Scene. This representation constitutes a somewhat literal and shallow description of the scene.

From the node layer, more abstract, global meanings can be recognized. We have termed representations of this kind of meaning *complex structures.* Complex structures, like representations in the node layer, are also comprised of node and links. They can be thought of as representations sitting atop the node layer of the Perceptual Scene (see Figure 3). In terms of the LIDA architecture, complex structures exist within the Current Situational Model, along with the Perceptual Scene.

Complex structures can represent more abstract entities such as event models (McCall et al. submitted; Zacks & Swallow 2007), plans, or content undergoing deliberation. More than one complex structure can describe the same Perceptual Scene simultaneously. Various complex structures can be at different levels of abstraction (or granularity). For example, event models are complex structures that represent events at different levels of abstraction.

Complex structures can have their roots in other complex structures or in the nodes of the Perceptual Scene. Moreover, a complex structure can have roots in more than one place, for example in the nodes of the Perceptual Scene, real or virtual, or in the Conscious Contents Queue.

Moreover, a complex structure can have nodes with associated sensory information. This information can be “displayed” in the perceptual scene. For example, a complex structure that represents an itinerary can have nodes that denote landmarks in that itinerary. These nodes can have sensory information related to the landmark that can be instantiated in the Perceptual Scene. We hypothesis that there are limits on the number of nodes that can be instantiated concurrently. For example, it is typically not possible to simultaneously imagine (represent) all the landmarks along a route to the airport from your house. So the Perceptual Scene can hold the information of different parts of an itinerary in a temporal order.

The various kinds of complex structures produce different types of meaning, enabling understanding of the current situation in the Current Situational Model. They comprise part of what Bernard Baars calls “the context” (1988). These structures are preconscious most of the time. Hassin et al (2009) describe an implicit working memory, similar but not identical to our workspace. They argue that some of the processes that take place in this implicit working memory produce “context-relevant updating of information, and goal-relevant computations involving active representations.” We take Hassin's assertion to be akin to our proposition that the

Sensory Scene is used to produce (update) our Perceptual Scene with additional input from complex structures.

Attention codelets scout complex structures looking for interesting situations. One example of this is when expectations are not fulfilled. Moreover, attention codelets can form coalitions using elements of complex structures. Thus the agent could broadcast information at various levels of abstraction from cycle to cycle. Suppose the agent's Current Situational Model has various event models pertaining to a soccer match. Then a hierarchy of event models could describe the current situation: an individual kick, a possession, an offensive charge, the entire match, the relation of this match to the playoffs, etc. So here, in one conscious broadcast a structure representing a kick could be sent, while the next one could be a more abstract entity representing the match.

How are these complex structures generated? Some of them can be created by structure building codelets that combine information from many places, such as the Perceptual Scene, the Episodic Memory buffer, the Conscious Contents Queue, etc. Others, like event models or plans can be created or instantiated with other mechanisms, such as selected behaviors or retrieval from memory. The construction and duration of these complex structures can take place over multiple cognitive cycles. For example event models of high level events can describe an event that lasts for many cognitive cycles. Complex structures that represent plans are another example of multi-cyclic complex structures.

Complex structures play an important role in the representation of the current situation. They can describe the situation in various levels of abstraction and granularity. Many of them last for multiple cognitive cycles, giving the agent a more stable representation over time. They combine data from different substructures in the workspace such as Conscious Contents Queue and the Episodic Buffer. Ultimately, as descriptions of the current situation, they can become part of a coalition that enters the competition for consciousness.

Conclusions

Creating and maintaining an internal representation of its current situation is a fundamental feature for sufficiently sophisticated autonomous agents. Such a representation allows the agent to best choose its next action. Such choice is the ultimate goal of any agent control structure.

We have described the Scene Representation, its structure and how it is produced in a comprehensive cognitive architecture. It is comprised of the Sensory Scene, with processed low-level data from the sensors, and the Perceptual Scene, which holds part of this sensory information and incorporates the node layer containing grounded perceptual symbols (Barsalou 1999, 2008).

Complex structures describe the current situation at still higher levels of abstraction. Both complex structures and the Perceptual Scene can take part of a coalition which enters the competition for consciousness.

In this model, various representations exist over different time scales. While the layers of the Sensory Scene are updated frequently, nodes and links in the node layer may persist for longer periods. Event models and other complex structures can remain active even longer, and in the face of conflicting representations at lower levels.

In conclusion, this framework can be used to provide an intelligent agent with a hybrid representation of its current situation, containing both sensory and grounded symbolic information. This representation provides the agent with a description of its current situation at different levels of abstraction and at different time scales. This representation can ultimately provide the agent with a better understanding of its current situation, allowing it to better choose its next action.

References

Baars, Bernard J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baars, Bernard J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Science, 6*, 47–52.

Baars, Bernard J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Science, 7*, 166–172.

Baars, B. J., & Franklin, S. (2009). Consciousness is computational: The LIDA model of Global Workspace Theory. *International Journal of Machine Consciousness, 1*(1), 23-32.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–609.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology, 59*, 617–645.

Franklin, S., Baars, B. J., Ramamurthy, U., & Ventura, M. (2005). The Role of Consciousness in Memory. *Brains, Minds and Media, 1*, 1–38.

Franklin, S., & Graesser, A. C. (1997). Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents *Intelligent Agents III* (pp. 21–35). Berlin: Springer Verlag.

Franklin, S., & Patterson, F. G. J. (2006). The LIDA Architecture: Adding New Modes of Learning to an Intelligent, Autonomous, Software Agent *IDPT-2006 Proceedings (Integrated Design and Process Technology)*: Society for Design and Process Science.

Hassin, R. R., et al. Implicit working memory. *Consciousness and Cognition* (2009), doi:10.1016/j.concog.2009.04.003

Lishman, J. R. (1981). Vision and the optic flow field. *Nature, 293*, 263 - 264.

Marr, D. C. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman

McCall, R., Franklin, S., Friedlander, D., & D'Mello, S. (submitted). Grounded Event-Based and Modal Representations for Objects, Relations, Beliefs, Etc.

Oyama, S. (1985). The Ontogeny of Information. Cambridge: Cambridge University Press.

Ramamurthy, U., Baars, B. J., D'Mello, Sidney K., & Franklin, S. (2006). LIDA: A Working Model of Cognition. In D. Fum, F. Del Missier & A. Stocco (Eds.), *Proceedings of the 7th International Conference on Cognitive Modeling* (pp. 244–249). Trieste: Edizioni Goliardiche.

Snaider, J., McCall, R., Franklin, S. (2009). Time Production and Representation in a Conceptual and Computational Cognitive Model.

Zacks, J. M. and K. M. Swallow. 2007. Event segmentation. Current Directions in Psychological Science, 16, no. 2: 80-84.