Artificial Spatial Cognition for Robotics and Mobile Systems: brief survey and current open challenges

P. de la Puente and M.G. Sánchez-Escribano
Universidad Politécnica de Madrid (UPM)
Madrid, Spain

***Abstract*—Remarkable and impressive advancements in the areas of perception, mapping and navigation of artificial mobile systems have been witnessed in the last decades. However, it is clear that important limitations remain regarding the spatial cognition capabilities of existing available implementations and the current practical functionality of high level cognitive models [1, 2]. For enhanced robustness and flexibility in different kinds of real world scenarios, a deeper understanding of the environment, the system, and their interactions -in general terms- is desired. This long abstract aims at outlining connections between recent contributions in the above mentioned areas and research in cognitive architectures and biological systems. We try to summarize, integrate and update previous reviews, highlighting the main open issues and aspects not yet unified or integrated in a common architectural framework.**

***Keywords*—spatial cognition; architecture; surveys; perception; navigation**

I. Brief Survey

A. *Initial models for spatial knowledge representation and main missing elements*

Focusing on the spatial knowledge representation and management, the first contributions inspired by the human cognitive map combined metric local maps, as an Absolute Space Representation (ASR), and topological graphs [3]. This way, they could be merged together into a Memory for Immediate Surroundings(MFIS) [4]. As a related approach, the Spatial Semantic Hierarchy (SSH) [5] was the first fundamental cognitive model for large-scale space. It evolved into the Hybrid SSH [6], which also included knowledge about small-scale space as isolated local maps connected by topological relations. This fundamental work was undoubtedly groundbreaking, but it did not go beyond basic levels of information abstraction and conceptualization [7]. Moreover, the well motivated dependencies among different types of knowledge (both declarative and procedural) were not further considered for general problem solving [8]. The SSH model was considered suitable for the popular schema of a "three layer architecture", without explicitly dealing with processes such as attention or forgetting mechanisms. This lack of principled forgetting mechanisms has been identified by the Simultaneous Localization and Mapping (SLAM) robotics community as a key missing feature of most existing mapping approaches [9, 10].

B. *The role of cognitive architectures and their relation to other works in the robotics community*

Cognitive architectures provide a solid approach for modeling general intelligent agents and their main commitments support the ambitious requirements of high level behavior in arbitrary situations for robotics [11]. A more recent model of spatial knowledge, the Spatial/Visual System (SVS) [12] designed as an extension of the Soar cognitive architecture, proposed a different multiplicity of representations – namely *symbolic*, *quantitative spatial* and *visual depictive*. The spatial scene is a hierarchy tree of objects/entities and their constitutive parts, with intermediate nodes defining the transformation relations between parts and objects. Other works in the robotics field employ similar internal representation ideas [13-15], and other ones included the possibility to hypothesize geometric environment structure in order to build consistent maps [16]. While a complete implementation of this approach not only for geometrical primitives but for all kind of objects requires solving the corresponding segmentation and recognition problems for the given sensor data in a domain independent manner (which is far beyond the state of the art), keeping the perceptual level representations within the architecture enhances functionality. There is a very active research community addressing these difficult perception challenges.

The recognition process should not only use visual, spatial and motion data from the Perceptual LTM but also conceptual context information[8, 17] and episodic memories associated to remembered places [18], from Soar's Symbolic LTM. This should also apply to the control laws and navigation techniques for different situations [19, 20]. The existence of motion models for the objects can significantly improve navigation in changing and dynamic environments, which is one of the main problems in real world robotic applications [21, 22].

A novel cognitive architecture specifically designed for spatial knowledge processing is the Casimir architecture [23], which presents rich modeling capabilities pursuing human-like behavior. Navigation, however, has not been addressed, and this work has scarcely been discussed in the robotics domain.

One of the latest spatial models is the NavModel [24], designed and implemented for the ACT-R cognitive architecture. Besides considering multi-level representations for spatial knowledge, this model presents three navigation strategies with varying cognitive cost. The first developed implementation assumes known topological localization at room level, while a subsequent implementation incorporates a mental rotation model. This work focuses on the cognitive load and does not deal with lower level issues. Many details regarding route generation, route following, dynamic changes, map management etc. are not addressed.

In order to point out how topics are addressed by the Cognitive Architectures and the Robotics communities, we

This work was partially funded by the Spanish Ministry of Economics and Competitivity: DPI 2014-53525-C3-1-R, NAVEGASE. It also received funding from the RoboCity2030-III-CM project (Robótica aplicada a la mejora de la calidad de vida de los ciudadanos. fase III; S2013/MIT-2748), supported by Programas de Actividades I+D en la Comunidad de Madrid and co-founded by Structural Funds of the EU.

compiled Table I as a comparison. The contrast regarding *memory management* capabilities and *uncertainty considerations* seems to be especially relevant. The lack of approaches combining both allocentric and egocentric representations is also remarkable.

TABLE I. COMPARISON OF TOPICS ADDRESSED BY THE COGNITIVE ARCHITECTURES AND ROBOTICS COMMUNITIES

Cognitive Architectures Community	← Topic →	Perception, Robotics, Vehicles Community
ACT-R/S, CLARION	Egocentric spatial models	[25, 26]
LIDA, SOAR-SVS	Allocentric spatial models	[10, 27]
Casimir, LIDA, SOAR-SVS	Object based/ semantic representations	[7, 13-15]
SOAR-SVS	Explicit motion models / dynamic information about the environment	[28, 29]
All	Memory management, forgetting mechanisms	[20]
Extended LIDA [30]	Uncertainty considerations	Most mapping and navigation approaches

To conclude, Table II presents a summary of surveys.

TABLE II. SUMMARY OF SURVEYS

Topic	References
Robotics and Cognitive Mapping	[1]
SLAM and Robust Perception	[9, 10]
Computational cognitive models of spatial memory	[2]
Object recognition	[31, 32]
Cognitive Architectures for Robotics	[11]
Spatial knowledge in brains	[18]

II. CURRENT OPEN CHALLENGES

The previous analysis indicates that the big challenge comes to closing the gap between high level cognitive models and actual implementations for robust perception and navigation competences in artificial mobile systems. To reduce this existing gap, we identify three main goals:

- Combination of allocentric and egocentric models using different levels of features/objects + topology/semantics.
- Acquisition and integration of motion models and dynamic information for the elements/objects.
- Integration of global mapping & loop closure capabilities with extensive declarative knowledge about features relevance and forgetting mechanisms with episodic memory. Management of STM and LTM for localization and navigation.

ACKNOWLEDGMENT

The authors want to thank the EUCog community for fostering interdisciplinary research in Artificial Cognitive Systems and organizing inspiring meetings and events.

REFERENCES

[1] Jefferies and Yeap. Robotics and Cognitive Approaches to Spatial Mapping. Springer, 2008.

[2] Madl et al. Computational cognitive models of spatial memory in navigation space:A review. Neural Networks, 2015.

[3] Yeap. Towards a computational theory of cognitive maps. Journal of Artificial Intelligence, 1988.

[4] Yeap, et al. Using a mobile robot to test a theory of cognitive mapping. Springer, 2008.

[5] Kuipers. The Spatial Semantic Hierarchy. Artificial Intelligence. 2000.

[6] Kuipers et al. Local metrical and global topological maps in the hybrid Spatial Semantic Hierarchy. ICRA, 2004.

[7] Pronobis and Jensfelt. Large-scale Semantic Mapping and Reasoning with Heterogeneous Modalities. ICRA, 2012.

[8] Lathrop. Extending Cognitive Architectures with Spatial and Visual Imagery Mechanisms. PhD Thesis, 2008.

[9] Fernandez-Madrigal and Blanco. Simultaneous Localization and Mapping for Mobile Robots: Introduction and Methods. IGI, 2012.

[10] Cadena et al. Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age. T-RO, 2016.

[11] Kurup and Lebiere. What can cognitive architectures do for robotics? Biologically Inspired Cognitive Architectures, 2012.

[12] Lathrop. Exploring the Functional Advantages of Spatial and Visual Cognition From an Architectural Perspective. TopiCS 2011.

[13] Salas-Moreno et al. SLAM++: Simultaneous Localisation and Mapping at the Level of Objects. CVPR, 2013.

[14] Eslami and Williams. A Generative Model for Parts-based Object Segmentation. Advances Neural Information Processing Systems, 2012.

[15] Uckermann et al. Real Time Hierarchical Scene Segmentation and Classification. Humanoids, 2014.

[16] De la Puente and Rodriguez-Losada. Feature based graph SLAM in structured environments. Autonomous Robots, 2014.

[17] Kunze et al. Combining top-down spatial reasoning and bottom-up object class recognition for scene understanding. IROS, 2014.

[18] M-B Moser, E.I. Moser. The Brain's GPS. Scientific American, 2016.

[19] Gunzelmann and Lyon (2007) Mechanisms for Human Spatial Competence. Spatial Cognition V, LNAI-Springer, 2007.

[20] Dayoub et al. Eight weeks of episodic visual navigation inside a non-stationary environment using adaptive spherical views. FSR, 2013.

[21] Hawes et al. The STRANDS Project: Long-Term Autonomy in Everyday Environments. Robotics and Automation Magazine, 2016.

[22] De la Puente et al. Experiences with RGB-D navigation in real home robotic trials. ARW, 2016.

[23] Schultheis and Barkowsky. Casimir: an architecture for mental spatial knowledge processing. TopiCS, 2011.

[24] Zhao. Understanding Human Spatial Navigation Behaviors: A Cognitive Modeling. PhD Thesis, 2016.

[25] Drouilly et al. Semantic representation for navigation in large-scale environments. ICRA, 2015.

[26] Posada et al. Visual semantic robot navigation in indoor environments . ISR, 2014.

[27] Richardson and Olson. Iterative path optimization for practical robot planning. IROS, 2011.

[28] Ambrus et al. Meta-rooms: Building and maintaining long term spatial models in a dynamic world. IROS, 2014.

[29] Rosen et al. Towards Lifelong Feature-Based Mapping in Semi-Static Environments. ICRA, 2016.

[30] Madl et al. Towards real-world capable spatial memory in the LIDA cognitive architecture. BICA, 2016.

[31] DiCarlo et al. How does the brain solve visual object recognition? Neuron, 2012.

[32] Roth and Winter. Survey of appearance based methods for object recognition. Technical Report, Graz University, 2008.