Event boards as tools for holistic AI

Peter Gärdenfors[1,2], Mary-Anne Williams[2,3], Benjamin Johnston[2], Richard Billingsley[2], Jonathan Vitale[2], Pavlos Peppas[2,4] and Jesse Clark[2]

[1] Lund University Cognitive Science, Lund
[2] University of Technology, Sydney
[3] Stanford University, Stanford
[4] University of Patras

```
Peter.Gardenfors@lucs.lu.se, Mary-Anne@themagiclab.org,
Benjamin.Johnston@uts.edu.au, Richard.Billingsley@uts.edu.au,
Jonathan.Vitale@student.uts.edu.au pavlos.peppas@gmail.com,
                Jesse.Clark@uts.edu.au
```

Abstract. We propose a novel architecture for holistic AI systems that integrate machine learning and knowledge representation. We extend an earlier proposal to divide representations into symbolic, conceptual and subconceptual levels. The key idea is to use event boards representing components of events as an analogy to blackboards found in earlier AI systems. The event components are 'thematic roles' such as agent, patient, recipient, action, and result. They are represented in terms of vectors of conceptual spaces rather than in symbolic form that has been used previously. A control level, including an attention mechanism decides which processes are run

Keywords: Holistic AI, event board, knowledge representation, machine learning, conceptual spaces, attention mechanism.

1 Program: Integrating categorization and reasoning

1.1 Knowledge representation and machine learning system are not enough

AI has two major approaches: Knowledge Representation (KR) and Machine Learning (ML). During the first decades KR, based on symbolic forms such as logic or programming languages, was dominant. Later connectionism, based on neural networks, generated a wave of ML with Deep Learning as a recent development [1,2]. However, ML is criticised for its bias, data greediness, opacity, and brittleness [3,4], while KR suffers from knowledge handling bottlenecks and scalability challenges [5].

In order to address these problems, hybrid systems that combine the benefits of ML and KR have been proposed [6], based on the fact that the two approaches have complementary advantages and disjoint shortcomings. ML's comparative strengths are domains with large amounts of training data and where small changes in the inputs have, generally speaking, small impacts on the outputs. In contrast, KR works best in domains driven by a small number of rules or heuristics or where minor changes can

have large-scale cascading effects.

Attempts to integrate both approaches began in the late 1980s [7] and continue up to the present day [8] with applications including medical diagnosis, text classification, time management, finance, control systems and bioinformatics. These systems are typically designed for specific applications in an ad hoc fashion.

Existing hybrid systems are therefore limited in scope, and there are few, if any, genuinely holistic AI systems. Existing systems often rely on human-intensive interpretation and management to translate ML outcomes into KR systems, and vice versa. Integrating ML and KR is recognised as crucial to attaining general AI.

1.2   A proposal for a new architecture

The main aim of this article is to present an outline of a new architecture for holistic AI systems – an architecture that is cognitively motivated. The key idea is to use *event boards* as mediators between different forms of information and different forms of processing. Event boards can be seen as a development of the blackboards that have been used in some types of integrative AI systems. However, unlike other blackboard models, the event board builds on theories of event representation, which have become a central topic in psychology [9] and in semantics [10]. These theories have so far not been exploited in AI settings.

The event boards contain information about different elements of events – present, past or planned. They take inspiration from cognitive semantics by keeping track of so called thematic roles mainly agents (performing actions) and patients (being affected), but also objects, recipients, instruments. A main motivation for this form of representation comes from the thesis that sentences typically express events [10].

1.3   An outline of a holistic architecture for AI

Gärdenfors [11] argues that there are aspects of cognitive phenomena for which neither symbolism of KR nor the connectionism used in ML offer appropriate modelling tools. He advocates a third conceptual form of representing information that is based on using geometric structures rather than symbols or connections between neurons. The essential aspects of concept formation are best described in such representational structures. Conceptual representations should not be seen as competing with symbolic or connectionist representations. Rather, the three kinds can be seen as three aspects of different granularity of representations. Chella et al. [12] implemented a system for artificial vision that combines and connects the three types of representation. Chella [13] extends the arguments for this type of architecture.

The new holistic AI architecture introduced herein builds on the three forms of processing and their interactions via an event board, but adds an event board as a tool for making information available to all three layers of representation. The event board also forms a basis for overall controlling functions that, among other things, controls the attention of the system. Figure 1 presents an outline of the proposed architecture. In the following sections, the different component of the architecture and their interactions will be presented in greater detail. The description is programmatic, but sufficient to inspire AI researchers to new forms of implementations of holistic AI systems.
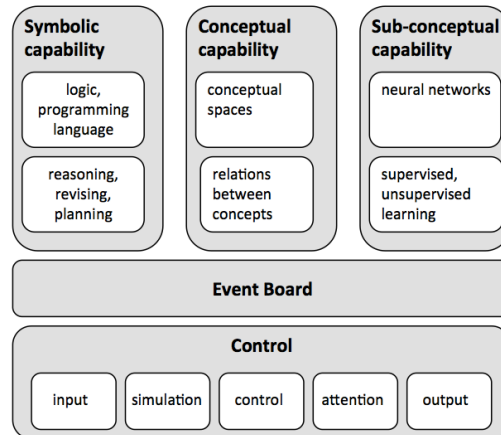
Fig. 1. The components of the holistic AI architecture

2 Symbolic, Conceptual and Sub-conceptual Capabilities

In this section, we briefly describe the three capabilities and their interactions. (For more details and arguments see [11, 12, 13])

Representing and processing symbolic information essentially consists of symbol manipulation, which relies on some logical or programming formalism. The purpose of this capability is to handle complex reasoning and planning tasks.

The information processing involves above all computations of logical consequences. In addition to classical logic, it is natural to include defeasible (non-monotonic) reasoning and belief revision [14, 15, 16] so that some rules function as default principles that can be overridden and changed.

The primary task of the conceptual capability is to provide representations of basic concepts and their relations. We propose that this be done with the aid of conceptual spaces [11]. A conceptual space consists of a number of domains built up from quality dimensions. Examples of basic domains are size, weight, shape, colour and location. However, there are also domains that are more abstract, for example kinship relations. It is assumed that each of the domains is endowed with a certain geometric structure.

An important aspect is that concepts are not independent of each other but can be structured into *domains*: Spatial concepts belong to one domain, concepts for colours to a different domain, kinship relations to a third, concepts for sounds to a fourth, etc.

A feature that clearly distinguishes the conceptual capability from the symbolic is that *similarity* plays a central role for the conceptual processes. Similarities between objects or properties are represented by distances in spaces. Instances of a concept are represented as points in space, and their similarity can be calculated in a natural way in the terms of some suitable distance measure.

The task of the sub-conceptual capability is to transform data into structures that can be processed by the other layers. A central task of ML is to *categorize* data. ML

architectures often involve neural networks that have been trained on large data sets. The output of such a system often involves considerable reduction of the number of dimensions. In this way the multi-dimensional data used by a ML system is filtered into a (small) number of domains that can be subsequently used by the conceptual or the symbolic capabilities. Currently there is a plethora of ML systems that generate information for categorization, not only methods of Deep Learning, but also various sensors, artificial vision, action recognition, visual question and answer systems. When unsupervised semantic approaches [17] are coupled with visualisation techniques [18], such conceptual spaces can be seen through the close clustering of word distributions, leading to forms of automatic discovery.

Since actions are central for the event semantics we use for the event board, ML systems that can generate categorizations of action will be important elements of the sub-conceptual capability. For example, Gu et al. [19] show how latent space representations for actions like opening and closing doors can be learned through robotic manipulation. Currently, there are a number of systems that perform such categorization, but still few that do it in online in real-time [20].

Summing up the comparison between the three capabilities of representation, one can say that the conceptual capabilities function on an intermediary scale between the coarse symbolic and the fine-grained connectionist representations. One can say that the conceptual dimensions 'emerge' from self-organising neural networks. The conceptual layer in turn provides the meanings of the expressions on the symbolic layer.

After this brief description of the three capabilities we next turn to the their interactions. Going from finer to coarser representations, the output from the sub-conceptual capability is often a low-dimensional space that generates a categorization of data. This low-dimensional space can then be used by the conceptual capability to find correlations between categories, for example that blue-eyed parents do not have brown-eyed children or causal relations such as if you eat toadstools you will become ill. These concept relations can then be represented on the symbolic level by the reasoning capability as sentences with varying degrees of defeasibility.

Going from coarser to finer representations, generic symbolic sentences – for example that songbirds build nests – can generate relations between domains for the conceptual capability. Machine learning mechanisms for extracting semantic content from sentences often involve dimensional reduction approaches [17] or recurrent neural networks [21] and output a latent space representation. Next the reasoning and conceptual correlations can function as constraints on the learning processes. Such processes can help reduce biases for ML learning. For example, there is a clear recognition that the now powerful, but too often biased and legally discriminating machine learning algorithms must be carefully managed to avoid undesirable decisions and to ensure compliance with policy, business strategy, law, regulations, ethics and societal values.

3   The structure and mechanisms of an event board

Blackboard architectures that were used, for example, in Hearsay II [22] and in Hofstadter's Copycat [23] functioned as 'libraries' of information that were provided by a number of subsystems. What was added to the blackboard could then in turn be used

by other subsystems. While this service has high utility, a common drawback of the blackboard architectures is that they lacked a unified way of describing what is represented on the blackboard. Related approaches such as publish-subscribe as used in ROS for robotics suffer from the same shortcoming.

3.1   Thematic roles and event representations

The event boards that are proposed here build on theories of event representations as a way of interpreting and connecting the information across the reasoning, conceptual and sub-conceptual capabilities. Our points of departure are accounts of 'thematic roles' in semantics [24] together with the two-vector representation of events developed by Gärdenfors and Warglien [10, 25, 26], where an event is built up from an agent, an action, a patient, a result and possibly other roles such as instrument, recipient, and beneficiary. Agent and patient are objects (animate or inanimate) that have different properties. A related model is the 'reservoir computing' developed by Dominey and his colleagues [27, 28, 29] that also employs thematic roles. However, reservoir computing builds only on neural networks. The event model also reminds of Schank and Abelson's [30] schema theory, although that theory neither systematically relies on thematic roles, nor on vector representations.

The two-vector model states that an event is represented in terms of two components – the force/trigger of an *action* that drives/initiates the event, and the *result* of its application. The result of an event is modelled as a vector representing the change of properties of the patient before and after the event [10]. For example, when somebody (agent) pushes (force vector) a table (patient), the forces exerted make the table move (the result vector). Or when somebody bends a stick, the result may be that the stick breaks. (When the result involves no change, then the event is a *state*.) A central part of an event category is the mapping from actions to results. This mapping is part of the representation of an event category [26] and it contains the central information about *causal* relations. Ideally, the event board will represent persistent background events. The decay of objects, the tendency of objects to fall if unsupported, the fact that liquids will settle may all be attributed to persistent processes.

The event board may be queried for the set of active events or states at any given point in time. Representation capabilities will lead to activities on the event board and they can respond to and contribute new events to the event board. Where concrete observations are unavailable, changes can be extrapolated from previous states by interpolating the change vector of current events and by integrating over the gradient of any continuous events and the background.

The events need not only involve physical forces, but can also be mental (for example commands, threats, insults and persuasive arguments) that can create a change in the emotional state of the addressee. The change is not physical but it can still be represented in terms of changes in a conceptual space (assuming that the concept 'person' has a domain of emotional states). Also social events such as buying, selling, marrying, can be represented in terms of thematic roles and conceptual spaces.

3.2 The role of the event board in the flow of information

By explicitly deconstructing events into constituent parts (agent, action, result, etc.) the event board can translate between semantic representations of events, like descriptions, and empirically witnessed events in a manner that allows statistical observations to be recorded. In this way, we can learn rule based descriptive beliefs that translate to events and actions, which in turn can lead to outcome expectations that can guide action planning.

What distinguishes an event board from a traditional blackboard is that the event structure generates a rich set of *expectations* that can guide the various processes. For example, actions lead to expectations of results. Similarly, if an agent doing something is detected, this can lead the system to search for the patient and what happens to it. Such expectations are also a central factor for the attention process that will be outlined in the following section.

Most blackboard systems use some symbolic form of representation, for example 'object spaces' [31]. In contrast, the event boards are based on *vector* representations: objects are represented as vectors in property spaces and actions and other processes as dynamic vectors. Vector representations allow various forms of similarities to be calculated. This form of representation also allows more realistic simulations of planned actions and their consequences in a dynamic system.

3.3 From events to sentences

A major reason for focussing on event representations is that according to Gärdenfors [10], the meanings of sentences are events (including states). When formulating a sentence from an event representation, a *construal* that picks out a focus of attention on the event must be chosen. For example, if a robot sees that another robot called "Pepper" pushes the box, the construal can select the agent and the action performed. Then the appropriate sentence expressed would be:
    (1) Pepper pushed the box.
If the attentional focus instead is on the patient and the action performed on it, the corresponding sentence would be
    (2) The box was pushed by Pepper.
A similar approach to generating sentences from a board, called a 'reservoir', of facts has been developed by Dominey and his group [27, 28, 29]. The elements in the reservoir are sorted into information about an agent, a patient, an object and a recipient, which covers some of the main thematic roles. In their architecture, all information is generated by neural networks, while in our version the event components can be generated from symbolic and conceptual systems in addition to the sub-conceptual ones.

4 The cognitive control

The control part can use the sub-conceptual capability to learn to map from descriptions (provided by sentences) or sensory data to events in a form that extracts the constituent thematic roles, particularly separating the actions and the results. This

mapping process involves learning an attention model that directs the focus of the system highlighting the facets of information that are most relevant to each component.

4.1   Input and output

All capabilities in the holistic AI architecture can receive input and produce output directly, and share information on the event board. There are two major kinds of input to the proposed architecture: sentences and sensory data. Sentences are either generics ("Spiders have eight legs") that feed directly to relations between regions on the conceptual layer or specific facts that are used by the reasoning capability and for generating data points for the conceptual capability that are made available on the event board. Separating the component roles, particularly the actions and results, along with other thematic constituent parts of a description can be performed by parsing [32] providing either a textual or context rich latent space result. The sub-conceptual capability acquires sensory and symbolic data and provides it to its ML modules for categorization or as data points to the conceptual capability.

Likewise, there are two major kinds of output: sentences and actions. Sentences are communicated as part of cooperative or planning activities. Actions are also performed as part of cooperation by the control part. Output processes can be thought of either as conditional models, extracting information from text, or as generative models, able to create textual output from their latent state inputs [33]. In this way, the mechanism that learns internal representations from descriptions also learns to generate compatible instructions for cooperation and planning purposes.

4.2   The role of the attention mechanism

The role of attention is to direct the system to relevant information. Attention models have been developed for language translation [34], which shows a latent space representation can be formed by learning multi-headed attentions that inform which parts of a sentence correspond in different languages. Attention models have also been used for image captioning and labelling [35] which converts a latent representation of a picture into a latent caption representation and then into a description by learning how to focus attention on each part of the image. Attention has also been used as the basis for the cognitive architecture ASMO [36].

The attention mechanism in the cognitive control determines the entities of an event that are most relevant. In fact, the attention mechanism not only provides a priority – it also determines if an event is posted to the event board. At any time, several ML algorithms can run in parallel and independently process similar information streams and the outcomes can go into direct competition for attention. When writing event parts on the event board, these algorithms can compete for similar resources within an event. For example, assume an agent is pushing an object, thus moving it, while at the same time an apple falls from a tree. If attention is currently biased towards observing the action of the agent, the apple falling from the tree may be considered irrelevant to the current event and not allowed entry to the event board. The

agent action, the object motion and the apple falling can be detected and represented by separate ML algorithms, which demand attention to write the captured entities on the event board.

ASMO [36] is an attention architecture able to resolve conflicts on single resources demanded by multiple processes. This is critical for controlling a robot in real-time. In this architecture, each process is attributed an 'attention value' on the basis of the current available context, which in our proposed model can come from the event board itself. Each time a resource is demanded by multiple processes at the same time, the attention mechanism allows the process with higher attention value to take control of the resource, while inhibiting the competing process(es). The architecture comes with learning mechanisms that can adapt the attention values provided by the demanding processes through experience. Building upon the ASMO architecture, we propose to use a similar attention mechanism to prioritise events on the event board.

4.3 The role of simulations

Learning from simulated environments has gained popularity because simulated events can accelerate learning more than using real-time events. Mnih et al. [37] show how optimum policy functions can be learned through game play, while Gu et al. |38] show how continuous advantage functions can be used to learn skills through simulation. However, in each case, the scope of the model is constrained to simple tasks because of the slow learning regimes and complex calculation overheads of Deep Learning architectures. By transforming these into conceptual spaces, the added constraints can reduce the parameter space allowing more efficient learning to occur and causal relations to be identified.

Johnston's COMIRIT [39] can be used to integrate commonsense reasoning and the geometric inference of conceptual spaces. COMIRIT establishes a novel mechanism for assigning 'semantic attachments' to symbols in KR systems that can be used to automatically construct simulations of the external world and utilise machine learning outputs. Simulations are just one of many representations that use similarity measures and operations for projecting forward and backwards to understand the causes and consequences of actions. Conceptual spaces of actions, such as those proposed Chella et al. [40] and more recently by Gharaee et al. [20] will be similarly used to reason about complex situations. In this way, simulation provides the AI system the imagination to depict a description of events needed to better understand the physical, social and, one day, emotional world we live in.

5   Conclusion

If we want an AI systems to solve the same tasks in similar ways that as humans do, it is natural to take inspiration from the architecture of human cognition. The main idea underlying the new holistic AI architecture is to use semantic event boards to integrate reasoning and learning capabilities and representations. This builds on the fact

that event cognition is characteristic for human deliberation, planning and problem solving.

The architecture we propose allows for representations on three different degrees of coarseness: the sub-conceptual, the conceptual and the symbolic layers. We have proposed that by using event boards, based on vector representations, a new type of holistic architecture can be constructed. Given the limited space, our proposal is by necessity programmatic. It will be evaluated by implementations of some application domains including social robotics.

References

1. Marcus, G.: Deep learning: A critical appraisal, arXiv preprint, arxiv.org/abs/1801.00631 (2018).
2. US Government: Preparing for the future of Artificial Intelligence, Executive Office of the President, National Science and Technology Council Committee on Technology (2016).
3. Pontin, J.: Greedy, brittle, opaque, and shallow: The downsides to deep learning, Wired Magazine, 02/02/2018.
4. Knight, W.: The dark secret at the heart of AI, MIT Technology Review, 120(3), 54-65 (2017).
5. Halevy, A., Norvig, P., Pereira, F.: The unreasonable effectiveness of data, IEEE Intelligent Systems, 24 (2), 8-12 (2009).
6. McCallum, A., Gabrilovich, E., Guha, R., Murphy, K. (eds): Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches: Papers from the 2015 AAAI Spring Symposium, AAAI Press Technical Report SS-15-03 (2009).
7. Towell, G., Shavlik, J.: Knowledge-based artificial neural networks, Artificial Intelligence, 70 (1), 119-165 (2009).
8. Shoham, Y.: Why knowledge representation matters, Communications of the ACM, 59 (1), 47-49 (2016).
9. Radvansky, G.A., Zacks, J.M.: Event Cognition, Oxford University Press, Oxford (2014).
10. Gärdenfors, P.: The geometry of meaning: Semantics based on conceptual spaces. MIT Press, Cambridge, MA (2014).
11. Gärdenfors, P.: Conceptual Spaces: The Geometry of Thought, MIT Press, Cambridge, MA (2000).
12. Chella, A., Frixione, M. Gaglio, S.: A cognitive architecture for artificial vision. Artificiail Intelligence 89, 73-111 (1997).
13. Chella, A.: Grounding ontologies in the external worlds. In: Borgo, S., et al. (eds.), Proceedings of the Joint Ontology Workshops 2017. CEUR-WS (2018).
14. Gärdenfors, P.: Knowledge in Flux: Modeling the Dynamics of Epistemic States. MIT Press Cambridge, MA (1988).
15. Williams, M.-A.: Iterated theory base change: A computational model. Proceeding of IJCAI 1995, 1541-1547 (1995).
16. Peppas, P.: Belief revision. Foundations of Artificial Intelligence 3, 317-359 (2008)
17. Le, Q., Mikolov, T.: Distributed representations of sentences and documents. Proceedings of the 31st International Conference on Machine Learning, (2014).
18. Maaten, L. V. D., Hinton, G.: Visualizing data using t-SNE. Journal of machine learning research, 9, 2579-2605 (2008).

19. Gu, S., Holly, E., Lillicrap, T., Levine, S.: Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates, 2017 IEEE International Conference on Robotics and Automation, Singapore, 3389-3396 (2017).
20. Gharaee, Z., Gärdenfors, P. and Johnsson, M.: Online recognition of actions involving objects, Biologically Inspired Cognitive Architectures, 22, 10-19 (2017).
21. Xu, W., Auli, M., Clark, S.: CCG supertagging with a recurrent neural network. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing Vol. 2, 250-255 (2015).
22. Erman, L. D., Hayes-Roth, F., Lesser, V. R., Reddy, D. R.: The Hearsay-II Speech-understanding system: Integrating knowledge to resolve uncertainty. Computing Surveys, 12(2), 213–253 (1980)
23. Hofstadter, D. R.: Fluid concepts and creative analogies. Basic Books, New York (1995).
24. Dowty, D.: Thematic proto-roles and argument selection. Language 67(3), 547-619 (1991).
25. Warglien, M., Gärdenfors, P., Westera, M.: Event structure, conceptual spaces and the semantics of verbs. Theoretical Linguistics 38, 159-193 (2012).
26. Gärdenfors, P., Warglien, M.: Using conceptual spaces to model actions and events. Journal of Semantics 29, 487-519 (2012).
27. Schank, R. C., Abelson, R. P.: Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. Erlbaum, Hillsdale, NJ (1977).
28. Gelernter, D.: Generative communication in Linda. ACM Transactions on Programming Languages and Systems, vol. 7(1), 80-112 (1985).
29. Lallée, S., Madden, C., Hoen, M., Dominey, P.: Linking language with embodied teleological representations of action for humanoid cognition. Frontiers in Neurobotics, doi:10.3389/fnbot.2010.00008 (2010).
30. Hinaut, X., Dominey, P. F.: Real-time parallel processing of grammatical structure in the fronto-striatal system: a recurrent network simulation study using reservoir computing. PLoS One, 8(2), 1-18 (2013).
31. Mealier, A.-L. Pointeau, G. Gärdenfors, P and Dominey, P. F.: Construals of meaning: The role of attention in robotic language production, Interaction Studies 17(1), 48-76 (2016).
32. Clark, S., Curran, J. R.: The importance of supertagging for wide-coverage CCG parsing. In Proc. of the 20th International Conference on Computational Linguistics,. 282.(2004).
33. Billingsley, R., Curran, J.: Improvements to Training an RNN parser. Proceedings of COLING 2012, 279-294 (2012).
34. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Polosukhin, I.: Attention is all you need. In Advances in Neural Information Processing Systems, 6000-6010(2017).
35. Johnson, J., Karpathy, A., Fei-Fei, L.: Densecap: Fully convolutional localization networks for dense captioning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4565-4574 (2016).
36. Novianto, R. Johnston, B. Williams, M.-A.: Habituation and sensitisation learning in ASMO cognitive architecture. Int. Conference on Social Robotics, 249-259 (2013).
37. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M.: Playing Atari with deep reinforcement learning. arXiv:1312.5602 (2013).
38. Gu, S., Lillicrap, T., Sutskever, I., Levine, S.: Continuous deep q-learning with model-based acceleration. In International Conference on Machine Learning, 2829-2838 (2013).
39. Johnston, B.: Practical Artificial Commonsense, PhD thesis, University of Technology, Sydney (2009).
40. Chella, A., Gaglio, S., Pirrone, R.: Conceptual representations of actions for autonomous robots. Robotics and Autonomous Systems 899, 1-13 (2001).