DSO Cognitive Architecture: Unified Reasoning with Integrative Memory Using Global Workspace Theory

Khin Hua Ng(✉), Zhiyuan Du, and Gee Wah Ng

Cognition and Fusion Lab 2, DSO National Laboratories, Singapore, Singapore
{nkhinhua,dzhiyuan,ngeewah}@dso.org.sg

Abstract. In this work, we present a design enhancement to the DSO Cognitive Architecture to augment its existing cognitive functions in an attempt to produce more general level of artificial intelligence in computational intelligent systems. Our design is centered on the concept of unified reasoning that indirectly addresses the diversity dilemma in designing cognitive architectures. This is done by implementing an integrative memory with the incorporation of the Global Workspace Theory. We discuss how other cognitive architectures using the Global Workspace Theory have influenced our design and also demonstrate how the new design can be used to solve an image captioning problem.

Keywords: Cognitive architecture · Unified reasoning · Integrative memory · Global workspace theory

1 Introduction

The DSO Cognitive Architecture (DSO-CA) is a top-level cognitive architecture that models the information processing in the human brain using inspirations drawn from Cognitive Science, Neuroscience, and Computational Science [1,2]. It is designed based on the key principles of hierarchical structure, distributed memory and parallelism. These led to an architectural design centered on functional modules (Reasoning, Visual, Association, etc.) that are executed asynchronously and in parallel to one another with each module possessing their own distinct memory system. The DSO-CA has been used to develop useful solutions to problems in applications like scene understanding [3] and mobile surveillance [4]. To further enhance its capability towards producing more human-like general intelligence and dynamic reasoning, we have been researching on advanced design principles and computational algorithms that will permit reasoning across different knowledge domains and representations – a process we termed *unified reasoning*.

Unified reasoning can be implemented either by encoding radically different knowledge domains into a common representation and then implement a single inference engine for the reasoning process, or by using multiple inference engines

T. Everitt et al. (Eds.): AGI 2017, LNAI 10414, pp. 44–53, 2017.
DOI: 10.1007/978-3-319-63703-7_5

and coming up with a way to unify them for different inputs. This leads to a design problem known as the *diversity dilemma*. In cognitive architecture design, the diversity dilemma refers to a need to blend diversity of different cognitive functions with uniformity of structure for efficiency, integrability, extensibility, and maintainability [5]. Diversity refers to the wide range of cognitive functions of different complexity required to operate in a dynamic, complex environment. Uniformity can be interpreted as how different cognitive functions can be realised through interactions between a small set of primitives and functions. This is a dilemma because ensuring uniformity alone may not be adequate in covering most cognitive functions, or accommodating new cognitive functions, but at the same time it is easier to maintain and integrate functions through uniformity. On the other hand, diverse implementations may not gel well with one another and will require a lot of engineering and maintenance effort, but they are easier to extend for new functionalities. In the case of unified reasoning, a single knowledge representation and its inference engine may not cover every domain while a set of different inference engines may require massive engineering to synergise them.

In this paper we present a design enhancement to the DSO-CA that augments its existing cognitive functions to perform unified reasoning. This design is based on the concept of implementing an integrative memory together with the incorporation of the *Global Workspace Theory*. The Global Workspace Theory (GWT) is a neuro-cognitive theory of consciousness developed by Bernard Baars [6,7], where information integration plays an important role. It advances a model of information flow in which multiple, parallel, specialised processes compete and co-operate for access to a global workspace. The global workspace then permits the winning coalition to broadcast to the rest of the specialists. According to the GWT, the mammalian brain instantiates this model of information flow, which enables a distinction to be drawn between conscious and unconscious information processing. To interpret it computationally, consciousness can be described as competition among processors, and outputs for a limited capacity resource that "broadcasts" information for widespread access and use. By grouping these specialists by cognitive functions, their availability to the global workspace causes information in memory to become conscious when the amount of activity representing it crosses a threshold [8].

By making use of an integrative memory system and applying the GWT, the collaboration between vastly different cognitive functions can be achieved and it indirectly provides a resolution to the diversity dilemma. Here, the integrative memory refers to a unified representation for the working memory. It serves as a common language for which the different modules communicate in and the GWT is the protocol of which they use to communicate. In the context of unified reasoning, this means the different reasoners can still use their specialised inference engines on their native representations but when a novel situation arises, which a single reasoner cannot resolve, the global workspace can propagate its content to different reasoners and complement its shortcoming. Hence, the GWT protocol incorporated with an integrative memory can be used as a solution to facilitate unified reasoning. Before we present this new design enhancement to the

DSO-CA, we will first provide a review of related work on cognitive architectures that implement the GWT, which have influenced our design. We will also provide an illustrative example on how the proposed new DSO-CA can be applied to an image captioning problem.

2 Related Works

To investigate how the GWT can be applied to enhance the DSO-CA with the functionality of unified reasoning, we have studied the following cognitive architectures infused with principles inspired by the GWT: MLECOG, CERA-CRANIUM, LIDA, CELTS, and CST. MLECOG (Motivated Learning Embodied Cognitive Architecture) [9] is a cognitive architecture built on top of self-organising artificial neural networks using the idea of pain signals. It is able to dynamically generate goals 'motivated' by desires and needs, and generate actions to fulfil the generated goals. MLECOG has all the semblance of GWT but it does not have a global broadcast mechanism – winner of the competition is routed through a predefined pathway. CERA-CRANIUM [10] uses a layered, hierarchical distributed architecture and it has two main components: CRANIUM, which serves as the workspaces of which the massive parallel, specialised processes operate on, and CERA, which serves as a domain-agnostic control unit to handle higher cognitive functions like selective attention, memory management, etc. LIDA (Learning Intelligent Distribution Agent) [11] aims to build a theoretical framework to unify different theories about human cognition, with a particular focus on the learning aspects of intelligent agents. Using a cognition cycle that starts with perception and ends with action, the GWT serves as a bridge between them by using attention strategies to pick different salient coalitions (a subset of the working memory elements) that are augmented by the various cognitive tasks that occurred before. Similarly, CELTS (Conscious Emotional Learning Tutoring Systems) [12] also implements a cognitive cycle of perception to action. One key difference is the existence of a shorter route for reactive behaviour and the use of 'emotions' to guide selective attention to working memory elements. Lastly, CST (Cognitive Systems Toolkit) [13] serves as a framework for users to create cognitive architecture using codelets as the atomic unit of computation.

To help us compare the design principles of these cognitive architectures we have worked out four features in the context of integrative memory and the GWT, namely, (1) types of memory, (2) competition, (3) purpose of global broadcast, and (4) information flow. Table 1 shows a summary of comparison amongst the different cognitive architectures including the enhanced DSO-CA with GWT using these four features.

Types of memory here refers to the representation used for the memory and the scale for the type of memory ranges from completely integrative to disparate. For example, LIDA, CELTS and MLECOG uses integrated memory whereas CST uses disparate memory due to the content of the memory objects (Table 1). Disparate memory facilitates different forms of representation to populate the working memory, and conversion of memories may be needed between

Table 1. A comparison among the different cognitive architectures along with DSO-CA implemented with GWT.

<table>
<tr><th>CA</th><th>Types of memory</th><th>Global broadcast</th><th>Information flow</th><th colspan="2">Competition</th></tr>
<tr><td>MLECOG</td><td>Integrated</td><td>N/A</td><td>Predefined</td><td>Pain signals and priorities</td><td>Multiple competition</td></tr>
<tr><td>CERA-CRANIUM</td><td>Disparate</td><td>Specific</td><td>Predefined</td><td rowspan="4">Activation based</td><td>Multiple local competition</td></tr>
<tr><td>LIDA</td><td>Integrated</td><td>Specific</td><td>Predefined</td><td>Multiple competition</td></tr>
<tr><td>CELTS</td><td>Integrated</td><td>Specific</td><td>Predefined</td><td>Single competition</td></tr>
<tr><td>CST</td><td>Disparate</td><td>Generic</td><td>Dynamic</td><td>Single competition</td></tr>
<tr><td>DSO-CA</td><td>Integrated but locally disparate</td><td>Generic</td><td>Dynamic</td><td>Activation and speed based</td><td>Multiple competition</td></tr>
</table>

different, incompatible processes. On the other hand, the advantage of using an integrative memory is the uniformity of representation. From a system's perspective, this allows easy maintenance and extension of the architecture. Also, it allows the combination of information contributed from processes, decreasing the complexity of the system in terms of designing mechanisms to handle different representations. In our proposed design enhancement to the DSO-CA, the working memory and content involved is integrated while the different specialised processors can have different memory representations.

The second feature "competition" is defined here as how the coalitions of processes compete for conscious access. Two aspects of competition are laid out in Table 1: the evaluation metric and the level of competition. In general, all cognitive architectures discussed here used similar mechanisms for competition with the most common evaluation metric being activation level. Each element in the working memory is assigned an activation level based on problem-dependent criteria, which indicates how relevant each element is to the goal or state of the agent. Another aspect of competition is the level of competition. In most architectures, there are multiple levels of competition.

The third feature classifies the global broadcast by its designed purpose: specific or generic. For example, the global broadcast in CELTS and LIDA serves mainly to aid learning, and influencing or invoking action selection. With a specific goal in mind, it determines how the global broadcast is implemented including the choice of memory representation that ties in with the integrative memory. Global broadcast with generic purpose simply propagates salient information; the propagated content are not designed according to pre-mapped functions. This is the case for CST and DSO-CA.

Lastly, the feature on information flow defines how information is transferred from one codelet/module to another and its pathway control. The flow can be predefined (top-down) or dynamic (bottom-up). Predefined information flow refers to information flow defined during the design phase where the GWT mechanism is embedded within this flow. This mechanism acts as a gatekeeper and bridge to pass the most salient information to the rest of the modules. As shown In Table 1, all except for CST and DSO-CA have predefined information flow. The main reason we have chosen dynamic information flow for unified reasoning within the DSO-CA is because its current design can readily adopt the dynamism through its pathway control that defines how information flows among modules [1].

3 Design

In this section, we present the proposed enhanced DSO-CA design. The design revolves around certain key aspects. Firstly, the architecture is composed of specialised processors executing asynchronously from one another. Each of the processors is akin to the unconscious processing in the brain, and they can operate on disparate memory representations. However the main working memory and content of interprocess communication must share a common representation with well-defined properties that can be exploited. Additionally, there should exist a mechanism for these specialised processors to compete against one another directly or indirectly. It is through competition that a specialised processor can access the global workspace and broadcast to other processes; this allows the transition of parallel, unconscious processing to serial, conscious processing [6]. Lastly, the global broadcast mechanism must possess an inhibitory function to suppress competing processors while it is broadcasting. With this in mind, we propose the design as illustrated in Fig. 1.

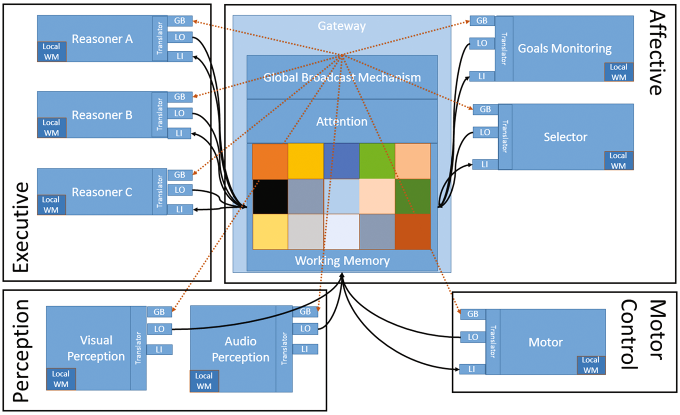

Fig. 1. An overview of the DSO-CA with the GWT. The attention and global broadcast mechanism layer lie on the working memory (integrative memory).

Modules that perform specialised and independent processes in a given domain are referred to as cognitive codelets. Each of these codelet has a translator that converts its local memory representation (as stored in the local working memory) into the integrative memory representation and vice versa. Cognitive codelets communicate with each other through the gateway to other cognitive codelets. Laying on top of the working memory is the attention layer and the Global Broadcast Mechanism; the attention layer seeks out and receives novel, critical, or relevant information embedded within the working memory or the cognitive codelets. Each cognitive codelet has local ports that can receive local inputs, or send local output to other cognitive codelets (routed via the gateway). However, it also has a broadcast port that can receive prioritised input from the Global Broadcast Mechanism.

Figure 2 shows the proposed design of the working memory. It uses a layered architecture. The integrative memory layer holds the actual working memory represented as factor graph. The reference memory layer is made up of cells to hold references to subgraphs of the underlying factor graph. A factor graph is a bipartite graph that represents a complicated global function as a product of local functions [14]. The choice of using factor graph is largely inspired by the SIGMA cognitive architecture [15], which demonstrated the combination of perception, localization, decision-making and learning using only factor graphs and with it, uses sum-product algorithm for cognitive processing [16]. Each reference memory cell contains shared memory among cognitive codelets. A cognitive codelet can send its output to a cell that in turn, can be used as inputs to other cognitive codelets (Fig. 2). This is similar to CST's memory object except that each cell holds a reference to a subgraph of the underlying factor graph.

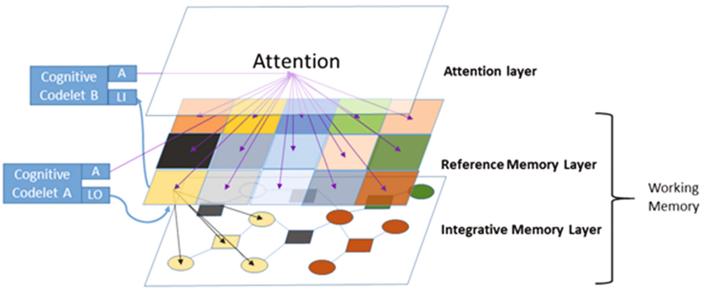

Fig. 2. The working memory has two layers. The integrative memory hosts the factor graph and the reference memory cells contain reference to subgraphs.

Activation level is calculated either on the reference memory cell or on the cognitive codelet (the arrows to and from the attention layer in Fig. 2). Candidates picked from the reference memory cells can be considered as top-down attention as the attention codelets define the criteria of selection. Conversely, candidates from cognitive codelets are bottom-up as the cognitive codelets define the selection criteria instead of the attention codelets. Local competition takes place at this level with the attention codelets, differentiated by their context, picking the most salient candidate – a factor graph which is a subgraph from the reference memory cell or the output of the cognitive codelet. Once the attention codelets have picked their winning candidate, they will be sent to the Global Broadcast Mechanism to compete for access. A temporary, capacity-limited buffer will be created for the candidates upon receiving the first one. The buffer will shut down when either the timer expires or when it fills up to its capacity. Global competition will start here and the winner (with the highest activation level) will be broadcasted to the rest of the cognitive codelets for additional processing. Inhibition starts when the buffer expires and the Global Broadcast Mechanism will reject any candidates sent from attention codelets of which, will also stop selecting or receiving candidates.

While developed independently, some parts of our design bear resemblance to the OpenCog [17] architecture's concept of "cognitive synergy". Both designs are based on the belief that general intelligence requires a synergised way of linking various specialised knowledge (memory) systems, which we called unified reasoning. OpenCog has a memory system known as Atomspace (a hypergraph database) that fits our notion of the integrative memory. Atomspace permits implementation of Probabilistic Logic Networks while our integrative memory is implemented using factor graphs. In essence, while the broad descriptions of both designs share certain level of similarities, the key principles that the designs are based upon are primarily different – one based on the GWT and the other Cognitive Synergy Theory.

4 Example

This section serves to illustrate how the new DSO-CA design with unified reasoning can be applied to an image captioning problem. Figure 3 shows a possible configuration of the cognitive architecture design. The ontology can be encoded as a dependency graph defined by K-parser [19]. Rules can be encoded using a production rule system to represent simple and obvious facts. Ontology and

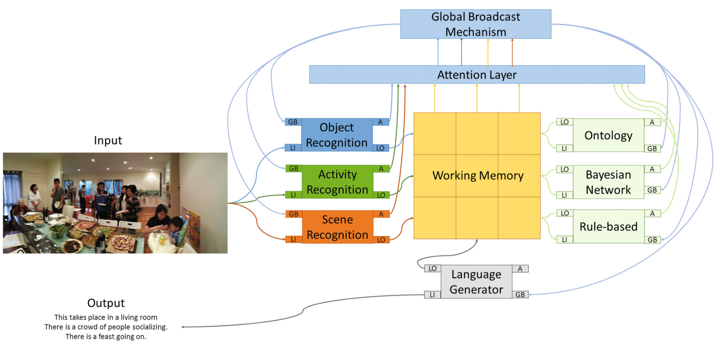

Fig. 3. The proposed design for the application. Image source: MS COCO Captioning Challenge dataset [18].

rules are used to construct the knowledge graph required by the language generator to create the captions. D'Brain [20], which is one of the reasoner modules already implemented in the DSO-CA, can be used to fuse multiple Bayesian networks knowledge fragments [21] to iteratively remove false recognition from the perception through inferencing and providing higher level context.

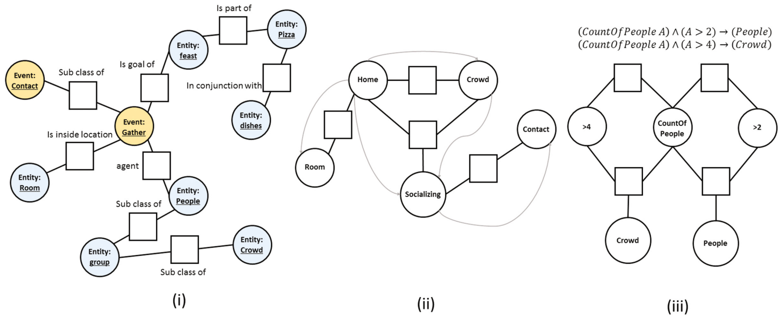

Fig. 4. The different knowledge bases' representations translated into factor graphs: (i) represents an ontology, (ii) is a Bayesian network, the grey arrows represent the original DAG, (iii) is a factor graph representing rules that are activated. These factor graphs are represented in the integrative memory while the various reasoners retain their native representations.

In Fig. 3, potential objects, activities and scenes identified by their respective perception codelets will be put into their respective reference memory cells. Reasoners use these inputs to either generate more candidates using associations and likelihood or remove percepts that are unlikely given its context. Output from the reasoners will return to the reference memory cells again and the attention codelets in the Attention Layer may pick up if the content (in the form of factor graph) is salient; they can also receive them from the other cognitive codelets. Each attention codelet is responsible for a domain and will select the most salient aspect of their respective domain for further competition that takes place in the Global Broadcast Mechanism. There, the candidate with the greatest activation level will be broadcasted to the rest of codelets for further processing, and the process repeats again. Using Fig. 4(iii) and the photo in Fig. 3 as an example, different persons would be initially recognised and tallied, the rule-based reasoner would eventually infer that there are 'people' and a 'crowd' in that photo. If its output are considered the winner of the global competition, it will be broadcasted to the rest of the cognitive codelets. When the ontology reasoner receives the broadcast, 'people' will be associated with 'gather' which in turn is associated with 'feasts' and 'pizza' (Fig. 4(i)) subsequently. This subset of factor graph will be broadcasted to the cognitive codelets again after competition and be picked up by D'Brain, which can calculate the concept association strength through inferencing; concepts, whose posteriors passes a threshold, are considered as candidates for bottom-up attention from the cognitive codelets.

Using Fig. 4(iii), 'home', 'socializing' and 'living room' can be picked out as the most probable associations, and the factor graph is selected for broadcast. Upon receiving the broadcast, ontology and rules-based reasoner expand the knowledge graph by associating more concepts. This example demonstrates the cooperation among the reasoner codelets facilitated by the GWT implementation; rules and ontology can generate associations based on semantics, and Bayesian networks can compute the posterior probability of concepts. Amidst all the competitions and broadcasts, the language generator codelet is concurrently creating sentences from inputs it received through the global broadcast. Sentences are generated by relations between concepts and some rules, and scored by a metric. The top three sentences will then be considered as the caption (Fig. 3). In summary, the dynamically assembled pipeline of unified reasoning process is: Rules → Ontology → Bayesian Network → Ontology → Language Generator. By using the global workspace, this is one of the many different pipelines that can be generated dynamically depending on changing contexts.

5 Conclusion

In this paper we have presented a design enhancement to the DSO-CA to augments its existing cognitive functions in an attempt to produce more general level of artificial intelligence in computational intelligent systems. Our design is centered on the concept of unified reasoning by implementing an integrative memory with the incorporation the GWT. This is done by compartmentalising reasoning functions into different parallelised codelets that contribute their inference results into the integrative memory, which is implemented by factor graphs. The GWT is responsible in picking the most novel and relevant information from the integrative memory and broadcast it to the other codelets, thereby connecting separate information pathways into a unified whole. This design will give us a form of dynamic reasoning that can elegantly adapt to different contexts through the collaboration of different reasoning systems. Through this, we hope it will serve as a building block bringing us a step closer to a cognitive architecture that produces human-like intelligence.

A computational development of the complete enhanced design for the DSO-CA is currently in the works. We aim to test the computational system with other challenging problems that will showcase the usefulness of an intelligent system capable of performing unified reasoning.

References

1. Ng, G.W., Tan, Y.S., Teow, L.N., Ng, K.H., Tan, K.H., Chan, R.Z.: A cognitive architecture for knowledge exploitation. In: 3rd Conference on Artificial General Intelligence (AGI 2010). Atlantis Press (2010)
2. Ng, G.W., Tan, Y.S., Teow, L.N., Ng, K.H., Tan, K.H., Chan, R.Z.: A cognitive architecture for knowledge exploitation. Int. J. Mach. Conscious. **3**(02), 237–253 (2011)

3. Ng, G.W., Xiao, X., Chan, R.Z., Tan, Y.S.: Scene understanding using DSO cognitive architecture. In: 2012 15th International Conference on Information Fusion (FUSION), pp. 2277–2284. IEEE (2012)
4. Ng, G.W., Tan, Y.S., Xiao, X.H., Chan, R.Z.: DSO cognitive architecture in mobile surveillance. In: 2012 Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF), pp. 111–115. IEEE (2012)
5. Rosenbloom, P.S.: Towards uniform implementation of architectural diversity. Artif. Intell. **20**, 197–218 (2009)
6. Baars, B.J.: A Cognitive Theory of Consciousness. Cambridge University Press, Cambridge (1993)
7. Baars, B., Franklin, S., Ramsoy, T.: Global workspace dynamics: cortical "binding and propagation" enables conscious contents. Front. Psychol. **4**, 200 (2013)
8. Reggia, J.A.: The rise of machine consciousness: studying consciousness with computational models. Neural Netw. **44**, 112–131 (2013)
9. Starzyk, J.A., Graham, J.: MLECOG: motivated learning embodied cognitive architecture. IEEE Syst. J. **11**(99), 1–12 (2015)
10. Arrabales, R., Ledezma, A., Sanchis, A.: CERA-CRANIUM: a test bed for machine consciousness research (2009)
11. Franklin, S., Madl, T., D'Mello, S., Snaider, J.: LIDA: a systems-level architecture for cognition, emotion, and learning. IEEE Trans. Autonom. Mental Dev. **6**(1), 19–41 (2014)
12. Faghihi, U., Fournier-Viger, P., Nkambou, R.: A computational model for causal learning in cognitive agents. Knowl. Based Syst. **30**, 48–56 (2012)
13. Paraense, A.L., Raizer, K., de Paula, S.M., Rohmer, E., Gudwin, R.R.: The cognitive systems toolkit and the CST reference cognitive architecture. Biol. Inspired Cogn. Archit. **17**, 32–48 (2016)
14. Loeliger, H.A.: An introduction to factor graphs. IEEE Signal Process. Mag. **21**(1), 28–41 (2004)
15. Rosenbloom, P.S., Demski, A., Ustun, V.: The Sigma cognitive architecture and system: towards functionally elegant grand unification. J. Artif. Gen. Intell. **7**(1), 1–103 (2016)
16. Ustun, V., Rosenbloom, P.S.: Towards adaptive, interactive virtual humans in Sigma. In: Brinkman, W.-P., Broekens, J., Heylen, D. (eds.) IVA 2015. LNCS, vol. 9238, pp. 98–108. Springer, Cham (2015). doi:10.1007/978-3-319-21996-7_10
17. Goertzel, B.: OpenCogPrime: a cognitive synergy based architecture for artificial general intelligence. In: 8th IEEE International Conference on Cognitive Informatics, ICCI 2009, pp. 60–68. IEEE (2009)
18. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft COCO: Common Objects in Context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8693, pp. 740–755. Springer, Cham (2014). doi:10.1007/978-3-319-10602-1_48
19. Sharma, A., Vo, N.H., Aditya, S., Baral, C.: Towards addressing the winograd schema challenge-building and using a semantic parser and a knowledge hunting module. In: IJCAI, pp. 1319–1325 (2015)
20. Ng, G., Ng, K., Tan, K., Goh, C.: The ultimate challenge of commander's decision aids: the cognition based dynamic reasoning machine. In: Proceeding of 25th Army Science Conference (2006)
21. Ng, G.W., Ng, K.H., Tan, K.H., Goh, C.H.K.: Novel methods for fusing Bayesian network knowledge fragments in D'Brain. In: 2007 10th International Conference on Information Fusion, pp. 1–8, July 2007