Generating Semantic Graphs through Self-Organization

Marshall R. Mayberry, III and **Matthew W. Crocker**
Saarland University
Germany
martym,crocker@coli.uni-sb.de

Abstract

In this study, a technique called semantic self-organization is used to scale up the subsymbolic approach by allowing a network to optimally allocate frame representations from a semantic dependency graph. The resulting architecture, INSOMNet, was trained on semantic representations of the newly-released LinGO Redwoods HPSG Treebank of annotated sentences from the VerbMobil project. The results show that INSOMNet is able to accurately represent the semantic dependencies while demonstrating expectations and defaults, coactivation of multiple interpretations, and robust processing of noisy input. The cognitive plausibility of the model is underscored by the collective modelling of four experiments from the visual worlds paradigm to show the model's ability to adapt to context.

Introduction

Deep semantic analysis of sentences from real-world dialogues is possible using neural networks: a subsymbolic system can be trained to read a sentence with complex grammatical structure into a holistic representation of the semantic features and dependencies of the sentence. This research breaks new ground in two important respects. First, the model described in this paper, the Incremental Nonmonotonic Self-Organization of Meaning Network (INSOMNet) (Mayberry 2003; Mayberry & Miikkulainen 2003), is the first subsymbolic system to be applied to deep semantic representations derived from the hand-annotated LinGO Redwoods Head-driven Phrase Structure Grammar (HPSG) Treebank (Oepen *et al.* 2002) of real-world sentences from the recently completed VerbMobil project (Wahlster 2000). Second, whereas most previous work has focused on the representation and learning of syntactic tree structures (such as those in the Penn Treebank), the semantic representations taken up in this study are actually dependency graphs represented with flat semantics called Minimal Recursion Semantics (MRS) (Copestake, Lascarides, & Flickinger 2001). The challenge of developing a subsymbolic scheme for handling graph structures led to self-organizing the case-role frame representations that serve as the graph nodes. This semantic self-organization in turn results in a number of interesting cognitive behaviors that will be briefly described in this paper, together with results from modelling more fine-grained psycholinguistic behavior from visual worlds studies.

Copyright © 2004, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

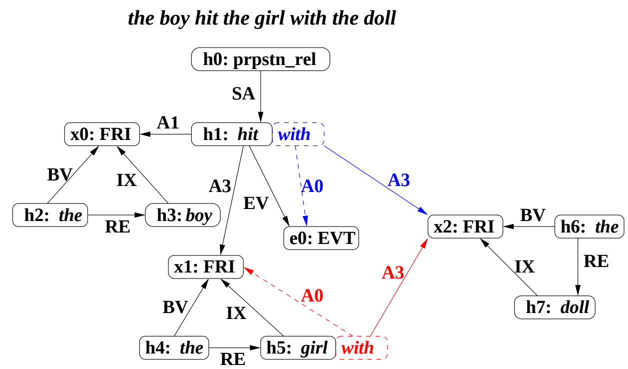

Figure 1: **MRS Dependency Graph.** This graph represents the sentence *the boy hit the girl with the doll*. Nodes in the graph are labeled with a handle and the word in the sentence (or semantic relation if there is no corresponding word). A dashed node connected to another node represents attachment. In this case, both nodes have the same handle, which is how MRS represents modification. The arcs that come out of a node indicate subcategorization and are labeled with the arguments that the word or semantic relation takes. The handles that fill those roles are the labels of the nodes the arcs point to.

Case-role Assignment with SRNs

The INSOMNet system is motivated by earlier work in the semantic task of case-role assignment using neural networks. Based on the theory of thematic case roles (Fillmore 1968), case-role analysis assumes that the syntactic structure of the sentence is specified beforehand, and the goal is to assign the proper roles to the words in the sentence. For example, given a simple sentence with subject, verb, object, and a with-clause, the network's task is to assign those words to the thematic roles **agent**, **act**, **patient**, and **modifier** or **instrument** depending on the selectional preferences of the words in the training corpus (Miikkulainen & Dyer 1991; McClelland & Kawamoto 1986). The sentence is read in one word at a time, and the network is trained to map the sentence into those words that fill the case roles for that sentence. In this manner the network develops expectations and

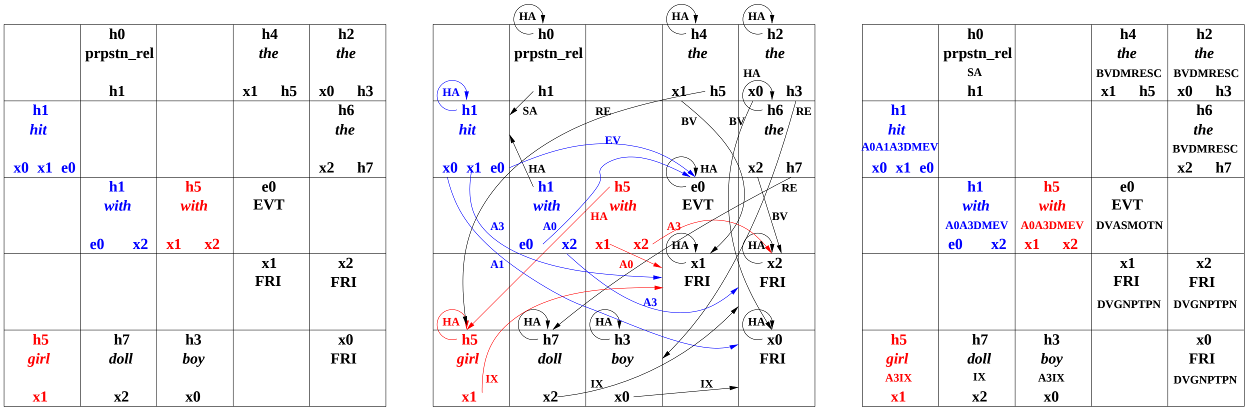

Figure 2: **Stages toward Representing Semantic Graphs in Neural Networks.** The grid on the left represents the basic information in the nodes and values in Figure 1, and the grid in the middle adds the labeled arcs for roles and their fillers. The grid on the right uses subcategorization information to indicate roles and fillers instead of explicit arcs.

defaults in much the same way as people do.

While this approach works well for sentences with fixed structure, it becomes untenable given the full complexity of language. Stopgap solutions, such as fixing the number of any given role, only work on toy grammars and small corpora, and are both linguistically and cognitively undesirable. The building of structures that represent meaning should not impose hard constraints on the complexity of the structure. To be cognitively valid, the ability to build up more complex structures should rather be subject to soft constraints that result in human-like errors in processing.

Sentence Representation

Our solution is motivated by the goal of representing the semantics of a sentence in as explicit and flexible a manner as possible. The semantic formalism used is MRS, which is a flat representation scheme based on predicate logic where nodes represent case-role frames and arcs represent dependencies between them. While space does not permit reviewing MRS in detail, we illustrate how MRS is used in INSOMNet by example. Figure 1 shows the MRS dependency graph for the sentence *the boy hit the girl with the doll*. This representation consists of 14 frames connected with arcs whose labels indicate the type of semantic dependency.

The sentence is declarative, as indicated by the semantic relation **prpstn_rel** in the top node labeled by the handle **h0**. Sentence types subcategorize for a *state-of-affairs*, which is indicated in Figure 1 by the arc labeled **SA**. The filler for this role is the handle **h1**, which is the label of the node for the main predication of the sentence, the verb *hit*. Verbs have an *event* role, and transitive verbs, such as *hit*, have an *arg1* role and an *arg3* role, which correspond to the thematic roles of **agent** and **patient**, respectively. These three roles are represented in the figure by the arcs labeled **EV**, **A1**, and **A3**. The filler for the verb's **EV** argument is the index **e0** for an *event structure* with semantic type **EVT** that specifies features for the verb such as tense, aspect, and mood. These features are not shown in Figure 1 to save space. The semantics of the noun phrases in the sentence are represented by three sets of nodes. Each set represents the determiner, the noun, and an index that indicates the features of the noun, such as gender, number, and person. The determiner is a quantifier which subcategorizes for the *bound variable* (**BV**) and *restriction* (**RE**) arguments (the *scope* role is empty in the current release of the Redwoods Treebank). The **BV** role is filled with the noun's index, and the **RE** role, with the noun's handle. The noun has a single *instance* (**IX**) role, filled with its index. The noun phrase, *the boy*, will make the representation clearer. In Figure 1, the node for the noun *boy* is labeled with the handle **h2**, which fills the **RE** role for the governing determiner *the*. The index of *boy* is the handle **x0**, which labels the node with semantic type **FRI**, indicating that *boy* is a *full referential index*. The index handle **x0** binds the determiner and noun through their **BV** and **IX** roles, respectively, and fills the **A1** role of *hit* to indicate that *boy* is the agent of the verb. Similarly, the index handle **x1** fills the verb's **A3** role to denote the patient. The handle **x2** is the index for the noun *doll* and fills the **A3** role of the preposition *with*. The preposition can either modify the verb for the instrumental sense of the sentence, or the noun for the modifier sense. In MRS, modification is represented by *conjunction* of predicates; for example *big red doll* is denoted by $\bigwedge$[**big**(x), **red**(x),**doll**(x)]. The n-ary connective $\bigwedge$ is replaced by a handle, which is distributed across the operands so that each predicate has the same handle (an operation we call *handle-sharing*). In the case of verb-attachment, the verb *hit* and the preposition *with* both share the handle **h1**, and the preposition's **A0** role is filled with the verb's event structure handle **e0**. For noun-attachment, *with* has the same handle **h5** as *girl*, and its **A0** role points to the index **x1** of *girl*. How the sentence is to be disambiguated is handled in the network by activating one sense more highly than the other.

Before we begin to describe the architecture, its activation and training, it would be useful to motivate its design in terms of the semantic representation in Figure 1. Figure 2 shows how a graph- or frame-based representation like that in Figure 1 could be represented in a grid, where the cells in the grids hold the components of individual MRS frames. This grid representation corresponds to the output of

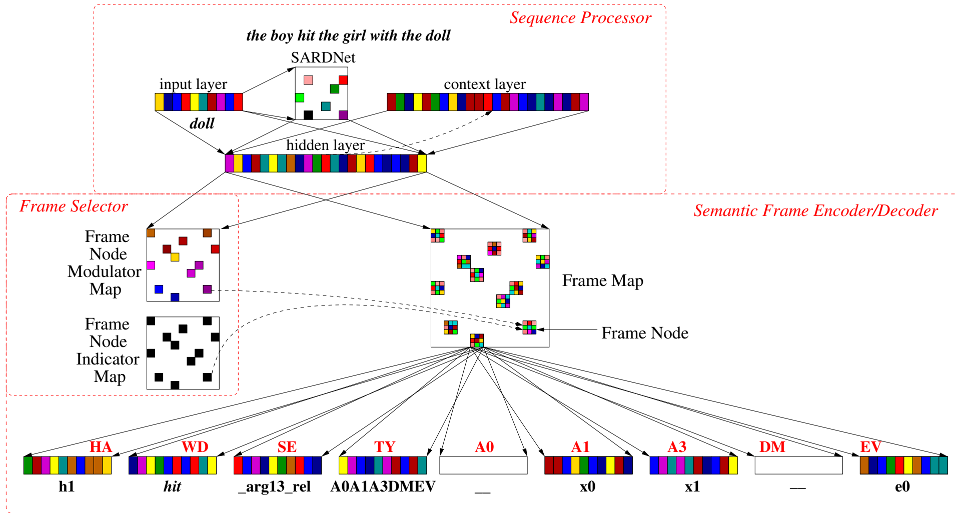

Figure 3: **The INSOMNet Architecture.** The INSOMNet model consists of three operational modules based on how they function together. The Sequence Processor reads the input sentence in one word at a time and activates both the Frame Selector and the Semantic Frame Encoder and Decoder. The Semantic Frame Encoder and Decoder encodes the MRS dependency graph for the semantic interpretation of the sentence as it is incrementally processed. Frame Selector is trained to select frames in a graded manner corresponding to an *a posteriori* probability that those frames belong to the current semantic interpretation.

the SRN. The leftmost grid only shows the graph nodes and values without the labeled arcs. These arcs are added in the middle grid to indicate how the cells are denoted according to their fillers. The grid on the right uses subcategorization information to indicate roles and fillers instead of explicit arcs. Figure 2 reveals how the use of the subcategorization information implicitly represents the labeled arcs. As can be seen, this representation completely describes the graph in Figure 1. Rather than with symbolic components, these frames are encoded through distributed representations, and a *decoder* network is required to pull out their individual components. The grid itself corresponds to the Frame Map in INSOMNet (see Figure 3), described in next.

Network Architecture

The INSOMNet sentence processing architecture (Figure 3) consists of three operational components: the *Sequence Processor*, the *Semantic Frame Encoder/Decoder*, and the *Frame Selector*.

The *Sequence Processor* is based on the SRN and processes a sentence one word at a time. A SARDNet Map (Mayberry & Miikkulainen 1999) retains an exponentially decaying activation of the input sequence to help the network remember long sequences.

The self-organized Frame Map of the *Semantic Frame Encoder/Decoder* is the main innovation of INSOMNet. In the current model, each Frame Node in the map itself consists of 100 units. The Frame Map itself is a 12×12 assembly of these nodes. As a result of processing the input sequence, the *Frame Selector* will activate a number of these nodes to different degrees; that is, a particular pattern of activation appears over the units of these nodes. Through the weights in the Frame Node Decoder, these patterns are decoded into the corresponding MRS case-role frames. The same weights are used for each node in the map. This weight-sharing enforces generalization among common elements across the many frames in any given MRS dependency graph.

The Frame Map is self-organized based on the compression of the frame representations. This process serves to identify which nodes in the Frame Map correspond to which case-role frames in the MRS structure. Because the frame compressions are distributed representations of case-role frames, similar frames will cluster together on the map. Determiners will tend to occupy one section of the map, the various types of verbs another, nouns yet another, and so on. However, although each node becomes tuned to particular kinds of frames, no particular Frame Node is dedicated to any given frame. Rather, through different activation patterns over their units, the nodes are flexible enough to represent different frames, depending on what is needed to represent the input sequence. For example in Figure 3, the Frame Node at the bottom decodes to the **h1** *hit* **_arg_rel A0A1A3DMEV _ x0 x1 _ e0** case-role frame for this particular sentence. In another sentence, it could represent a different verb with a slightly different subcategorization type. This feature of the architecture makes the Frame Map able to represent semantic dependency graphs dynamically, enhancing generalization.

During training of the SRN, the Frame Node serves as a second hidden layer and the case-role frames as the output layer. The appropriate frames are presented as targets for the Frame Node layer, and the resulting error signals are backpropagated through the Frame Node Decoder weights to the Frame Node layer and on up to the first hidden layer.

The Frame Map holds the encoded frame representations so that they can be decoded by the Frame Node Decoder. The Frame Selector system, on the other hand, is used to indicate what frames are active in the current interpretation

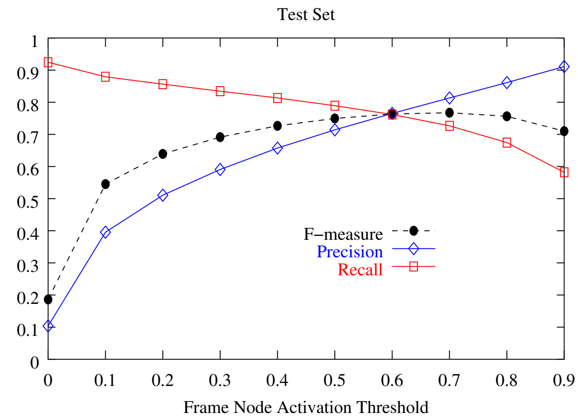

Figure 4: **Sentence Processing Performance.** The nodes on the Modulator Map corresponding to the target frameset are identified, and their activations evaluated according to whether they reach threshold. The true positives are those target frames above threshold, false negatives are the target frames below threshold, and the false positives are non-target frames above threshold. The result is a maximum F-measure of 0.76 at $x = 0.7$.

of a sentence, and the degree to which they are active.

At the same time, a RAAM network is trained to form the compressed frame representations, and the current representations are used to organize the Frame Map. The input word representations are randomly distributed 64-unit vectors on which the SARDNet map is self-organized. Eventually the compressed frame representations converge, and the networks learns to generate the correct MRS dependency graph and the corresponding case-role frames as its output.

Experiments

Ten-fold cross-validation was run on the 4817 sentences from the Redwoods Treebank (Version June 20, 2001) for which at least one analysis was selected as correct by the treebank annotator. The dataset had 1054 tokens and 104 abstract semantic relations. Morphemes such as *-s* and *-ing* were processed in separate steps, but common inflected words such as *am* and *done* were left unchanged.

INSOMNet was trained with an initial learning rate of 0.01 and the Frame Node Indicator Map given an initial learning rate of 0.4. The neighborhood was initialized to half the Frame Map's diameter. The learning rate of the Frame Node Indicator Map was decayed by 0.9 and the neighborhood decremented according to the schedule

$$\text{epoch}_{i+1} = 1.5 * \text{epoch}_i + 1$$

where i indexes each parameter update. Once the Frame Node Indicator Map had stopped self-organizing when its learning rate fell below 0.001, the learning rate for INSOMNet was decayed by 0.5 according to the same schedule.

Results

Figure 4 provides one measure of INSOMNet's average performance on the held-out data. The x-axis denotes the frame activation level, and the y-axis, the precision/recall curves (together with their F-measure). The F-measure is highest at $x = 0.7$, where approximately 72% of target dependencies were selected by INSOMNet, and 81% of selected dependencies were targeted.

Psycholinguistic modelling

Psycholinguists have typically pursued subsymbolic parsing architectures because they show promise not only in being scaled up to real-world natural language, but they do so while retaining the cognitively plausible behavior that sets these systems apart from most other NLP approaches. Here we will present some preliminary results on the integration of multimodal input to INSOMNet to demonstrate the *adaptive* behavior of the network in addition to the more characteristic behaviors of incrementality and anticipation. The cognitive phenomenon that we look at here is people's ability to use context information, such as visual scenes, to more rapidly interpret and disambiguate a sentence. In the four visual worlds experiments modelled in this section, accurate sentence interpretation hinges on proper case-role assignment to sentence participants. All four experiments were conducted in German, a language that allows both subject-verb-object (SVO) and object-verb-subject (OVS) sentence types, so that word order cannot be reliably used to determine role assignments. Rather, case marking in German is used to indicate grammatical function such as subject or object, except in the case of feminine nouns where the article does not carry any distinguishing marking for the nominative and accusative cases. These experiments made it possible to analyse the interplay between scene context and a variety of linguistic factors, as described next.

Morphosyntactic and lexical verb information. Kamide, Scheepers, & Altmann showed how selectional restrictions on verbs involve a compositional linguistic component to predict the semantic class of upcoming arguments. German sentences were presented accoustically with visual scenes showing, for example, a hare, cabbage, a fox, and a distractor. The scene was either combined with a subject-first sentence *Der Hase frisst gleich den Kohl* ("The hare$_{\mathbf{nom}}$ eats just now the cabbage$_{\mathbf{acc}}$") or with the object-first sentence *Den Hasen frisst gleich der Fuchs* ("The hare$_{\mathbf{acc}}$ eats just now the fox$_{\mathbf{nom}}$"). The nominative case marking on *der Hase* made it a typical agent, which, when combined with the semantic information associated with the verb *frisst*, allowed listeners to predict the cabbage as the most plausible forthcoming referent in the scene. On the other hand, accusative case marking on *den Hasen* made it a typical patient, and listeners predicted the fox as the most plausible forthcoming referent. Syntactic case information was apparently combined with semantic verb information to predict post-verbal referring expressions.

Verb type information. To ensure that syntactic case information alone was not sufficient to predict subsequent referents, Scheepers, Kamide, & Altmann used experiencer/theme verbs in which the agent (experiencer) and patient (theme) roles were interchanged. Thus, whereas for verbs like *frisst* ("eats") the subject is typically the agent of a sentence, for verbs like *interessiert* ("interests"), the subject is typically the Patient. Thus, with *interessiert* instead

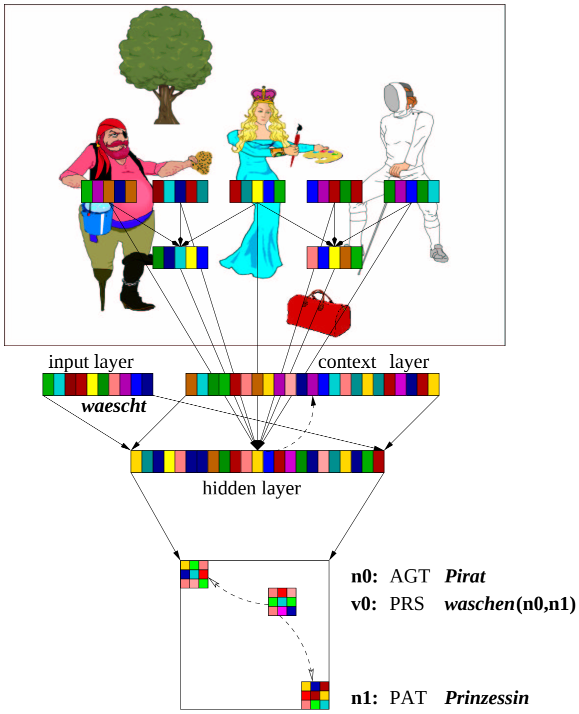

Figure 5: **Scene Integration.** INSOMNet was modified slightly to model the experimental findings in the four visual worlds studies. The decoder was simplified and extra input assemblies for the scene were added. For depicted events, an extra two assemblies were used to represent the events. All extra assemblies were propogated to the hidden layer of the network. Note that some parts of the architecture, such as SARDNet, are not shown in this figure for clarity.

of *frisst*, the role of most suitable referent switches. Accordingly, the pattern of anticipatory eye movements reversed for the two types of verb, confirming that both syntactic case information and semantic verb information are being used to predict subsequent referents.

Ambiguity in word order. Knoeferle *et al.* tested whether depicted events in the absence of selectional restrictions would allow for early disambiguation of argument roles in a visual-world study with sentences in which the initial NP was case-ambiguous and *linguistic* disambiguation took place at the second NP that was clearly case-marked as either accusative (patient) or nominative (agent). An example is the SVO sentence *Die Princessin malt gleich den Fechter* ("The princess$_{\mathbf{nom}}$ paints just now the fencer$_{\mathbf{acc}}$") versus the OVS sentence *Die Princessin wäscht gleich der Pirat* ("The princess$_{\mathbf{acc}}$ washes just now the pirate$_{\mathbf{nom}}$"). Together with the auditorily presented sentence a visual scene was shown in which a princess both paints a fencer and is washed by a pirate (see Figure 5). Because stereotypicality and selectional information were strictly controlled for, they could not be used to disambiguate at the point of the verb; however, the events in the scenes potentially could. As expected, the SVO preference initially triggered eye-movements to the patient for both sentence types. More interesting still, when verb information, combined with depicted events, disambiguated towards an OVS structure, this interpretation was quickly revised, causing anticipatory eye-movements to the agent. The study showed that German listeners initially prefer to interpret case ambiguous sentence-initial NPs as agents, but also that lexical verb information can be rapidly matched with depicted actions to quickly revise this interpretation when appropriate.

Soft temporal adverb constraint. Knoeferle *et al.* also investigated German verb-final active/passive constructions to show that linguistic disambiguation relies on a thematic role-assignment process that did not depend on grammatical function. In both the active future-tense sentence *Die Princessin wird sogleich den Pirat washen* ("The princess$_{\mathbf{nom}}$ will soon wash the pirate$_{\mathbf{acc}}$") and the passive sentence *Die Princessin wird soeben von dem Fechter gemalt* ("The princess$_{\mathbf{acc}}$ is currently painted by the fencer$_{\mathbf{nom}}$"), the initial subject noun phrase is disambiguated as agent in the active sentence and as patient in the passive sentence. To evoke early linguistic disambiguation, temporal adverbs were used to bias the auxiliary *wird* toward either the future ("will") or passive ("is ...ed"). Because the constructions used were verb-final, the interplay of scene and linguistic cues such as those provided by the temporal adverbs were rather more subtle. Thus, when the listener heard a future-biased adverb such as *sogleich*, after the auxiliary *wird*, she would interpret the initial NP as an agent of a future construction, and use the scene in which that NP filled the agent role to anticipate the upcoming patient argument. Conversely, listeners interpreted the passive construction with these roles exchanged.

Experimental Modelling with INSOMNet

INSOMNet was modified to represent just the sentence types in the four experiments in order to model the finer behavior of early disambiguation observed in all cases. The Semantic Frame Decoder was simplified to produce a lexical item and either an agent or a patient role. The greater modification to the network involved adding extra input assemblies for the characters and events from the scenes to the model which fed into the hidden layer (see Figure 5).

The modelling task also differed from the corpus-based approach described earlier in how subsymbolic systems might be scaled up to broad coverage models. In these experiments, we have adopted a grammar-based approach to exhaustively generate a set of sentences based on the experimental materials while holding out the actual materials used for testing. Furthermore, half the sentences were trained with scenes and half were trained without in order to approximate linguistic experience.

INSOMNet is clearly able to perform the early disambiguation task, meaning that it is able to access the scene information and combine it with the incrementally presented sentence to anticipate forthcoming arguments. For the two experiments using non-stereotypical characters and depicted events, accuracy was nearly 100%, and for the other two experiments using selectional restrictions, the performance was somewhat less at 94%. Part of the reason for this discrepancy is that there is no event information provided in these latter experiments, so INSOMNet has more information to work with in the first two experiments.

Discussion

There is clearly room for improvement in the scaling up performance of INSOMNet. Although 50% of the errors can be attributed to annotation errors in the original corpus, the remainder offers opportunities to improve the model's architectural foundation. Apart from this performance issue, INSOMNet does exhibit a number of interesting behaviors that make it a useful cognitive model. First, sentence processing in INSOMNet is incremental and nonmonotonic, with radical interpretation revision possible. Second, INSOMNet represents ambiguities explicitly and in parallel, thereby allowing multiple interpretations to be simultaneously active to the degree warranted by training. Third, experiments have shown that the model is robust to dysfluencies such as hesitations, repairs, and restarts. Fourth, INSOMNet demonstrates the hallmark behavior of expectations and defaults that arise automatically in subsymbolic systems. And fifth, INSOMNet is able to seamlessly integrate multimodal input in an adaptive manner that begins to get at the core of true human language performance, which is not passive and single-channel, but dynamic, adaptive, and characterized by active, anticipatory interpretation. Indeed, the ability to adaptively integrate multimodal input touches directly on the issue of compositionality, which is generally regarded from the viewpoint of a single input mode. The results suggest that subsymbolic systems feature a natural mechanism for handling this extended view of compositionality.

Despite these desirable behavioral characteristics of INSOMNet, there remains the open issue of how best to interpret the network's output. Because each frame is activated by the Frame Selector to denote its inclusion in the sentence interpretation, it often happens that extra frames that are not part of the golden standard are activated in addition (for purposes of rigor, these frames were discounted as "false positives" in our results, bringing down the precision results). Moreover, an effect similar to "dangling pointers" is possible upon decoding frames if components are not activated strongly enough on the Frame Map. The principal concern is that the natural gradience of subsymbolic systems such as INSOMNet makes it difficult to give its output the type of categorical interpretation typical in statistical/symbolic computational linguistic research.

Conclusion

In this paper, we presented a subsymbolic sentence processor, INSOMNet, that is able to analyze a real-world corpus of sentences into semantic representations. A crucial innovation was to use an MRS dependency graph as the sentence representation, encoded in a self-organized Frame Map. As is typical of holistic systems, the interpretation is developed nonmonotonically in the course of incrementally reading in the input words, thereby demonstrating several cognitive behaviors such as coactivation, expectations and defaults, and robustness. Equally crucial is how readily the network is able to adapt to context to anticipate upcoming argument roles. These properties make INSOMNet a promising foundation for understanding situated human sentence processing in the future.

References

[Copestake, Lascarides, & Flickinger] Copestake, A.; Lascarides, A.; and Flickinger, D. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of ACL-2001*. Toulouse, France.

[Fillmore] Fillmore, C. J. 1968. The case for case. In Bach, E., and Harms, R. T., eds., *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston. 0–88.

[Kamide, Scheepers, & Altmann] Kamide, Y.; Scheepers, C.; and Altmann, G. T. M. 2003. Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research* 32(1):37–55.

[Knoeferle *et al.*] Knoeferle, P.; Crocker, M. W.; Scheepers, C.; and Pickering, M. J. accepted. The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*.

[Mayberry & Miikkulainen] Mayberry, M. R., and Miikkulainen, R. 1999. Using a sequential SOM to parse long-term dependencies. In *Proceedings of the 21st Annual Conference of the Cognitive Science Society*, 367–372. Hillsdale, NJ: Erlbaum.

[Mayberry & Miikkulainen] Mayberry, M. R., and Miikkulainen, R. 2003. Incremental nonmonotonic parsing through semantic self-organization. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Mahawah, NJ: Erlbaum. 798–803.

[Mayberry] Mayberry, M. R. 2003. *Incremental Nonmonotonic Parsing through Semantic Self-Organization*. Ph.D. Dissertation, Department of Computer Sciences, The University of Texas at Austin, Austin, TX.

[McClelland & Kawamoto] McClelland, J. L., and Kawamoto, A. H. 1986. Mechanisms of sentence processing: Assigning roles to constituents. In McClelland, J. L., and Rumelhart, D. E., eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2: Psychological and Biological Models*. Cambridge, MA: MIT Press. 272–325.

[Miikkulainen & Dyer] Miikkulainen, R., and Dyer, M. G. 1991. Natural language processing with modular PDP networks and distributed lexicon. *Cognitive Science* 15:343–399.

[Oepen *et al.*] Oepen, S.; Flickinger, D.; Toutanova, K.; and Manning, C. 2002. LinGO redwoods: A rich and dynamic treebank for HPSG. In *Beyond PARSEVAL. Workshop of the Third LREC Conference*. Las Palmas, Spain.

[Scheepers, Kamide, & Altmann] Scheepers, C.; Kamide, Y.; and Altmann, G. T. submitted. The compositional integration of syntactic, semantic, and world knowledge constraints in projecting upcoming arguments in German. submitted to Journal of Experimental Psychology: Learning, Memory, and Cognition.

[Wahlster] Wahlster, W., ed. 2000. *Verbmobil. Foundations of Speech-to-Speech Translation*. Berlin; New York: Springer-Verlag.