This paper was presented at The First Australian Conference on Neural Networks, ACNN'90, University of Sydney, NSW, Australia, 29-31 January 1990.

Connectionist Chart Parsing

Jacob L. Cybulski

Artificial Intelligence Systems
Telecom Research Laboratories
770 Blackburn Road
Clayton, Vic 3168, Australia
Phone: +613 541 6273, Fax: +613 543 8863

Abstract

This paper describes a method of integrating syntactic and semantic interpretation of natural language utterances via massively parallel constraint satisfaction implemented in a connectionist chart parser - COSIMO.

1.0 Introduction

Connectionism has already found its way into a number of small Natural Language Parsing systems (NLP). The prevalent characteristic of these systems is their ability to propagate weak constraints (lexical, syntactical, semantic and pragmatic) between input-derived terms to form an utterance interpretation. The main mechanism employed in these systems is a neural-like activation spreading between linguistic and cognitive representation elements, activation inhibition and decay (cf. Collins and Loftus 1975, Anderson 1983). Some of the systems focus on the syntactic processing of texts (Slack 1986, Jones 1987, Parisi and Nolfi 1987, Howells 1988), some look at the micro-levels of semantic analysis to resolve lexical or syntactic ambiguity (Cottrell and Small 1983, Pollack and Waltz 1984, Bookman 1987, Cybulski et al 1989), others concentrate on language generation (Kukich 1987, Gasser 1988) or speech understanding (Bengio et al. 1989, Stanfil and Waltz 1986), few address connectionist NLU in an integrated, multi-facet way (Small 1987, also cf. Minsky 1986, Lytinen 1986).

COSIMO, unlike some other systems, combines symbolic and connectionist processing, based on Preference Semantics (Wilks 1975) and chart parsing paradigms (Kay 1980) with connectionist constraint satisfaction of Interactive Relaxation with Decay (McClelland and Rumelhart 1986).

In the spirit of Preference Semantics, COSIMO's long term memory consists of three different types of organisational components :- a dictionary assigning semantic word senses to character strings, formulae giving semantically reducible definitions of domain and universal concepts, and templates specifying a language grammar (refer to Cybulski and Jennings 1989 for more details). The connectionist chart accounts for the system short-term memory. The ultimate objective of the parsing process is to analyse natural language input into a chart structure and then collapse it along the inheritance paths into a number of inter-related application-domain concepts.

The model assumes all chart constituents to have associated with it a momentary energy level, represented by a real number and known as activation. Activations are propagated between the chart elements arranged into a network of excitory and inhibitory connections, from signs up to the phrase semantic interpretations. Excitory connections strengthen the activation of neighbouring nodes and correspond to the representational associations between concepts, i.e. instance-type, frame-slot, formula-constraint, formula-template, sign-sense, or chart constituent concatenations. Inhibitory connections, on the other hand, weaken the activation of neighbouring nodes, and are related to the potentially ambiguous chart clusters, i.e. multiple sign senses, multiple formulae interpretations, multiple template associations, chart constituent overlaps, or near misses. The strength of each excitory and inhibitory link is represented with either positive or negative connection weights respectively. A chart component is said to be active when its activation level is positive. In the absence of input from its neighbours the value of a node activation slowly decays to its resting level, i.e. either zero or negative determined by the relative frequency of the constituent use (recorded with dictionary entries, templates and formulae). The connection weights and formula constants were adapted from VITAL (Howells 1988), thus the weights of the sign/sense connections were assumed to be proportional to the relative element frequency (in its own class), the weight of formula and template chart associations (bottom up activations) is currently a constant inversely proportional to the number of incoming links, the subcomponent associations (top down activation) is inversely proportional to the number of constituents.

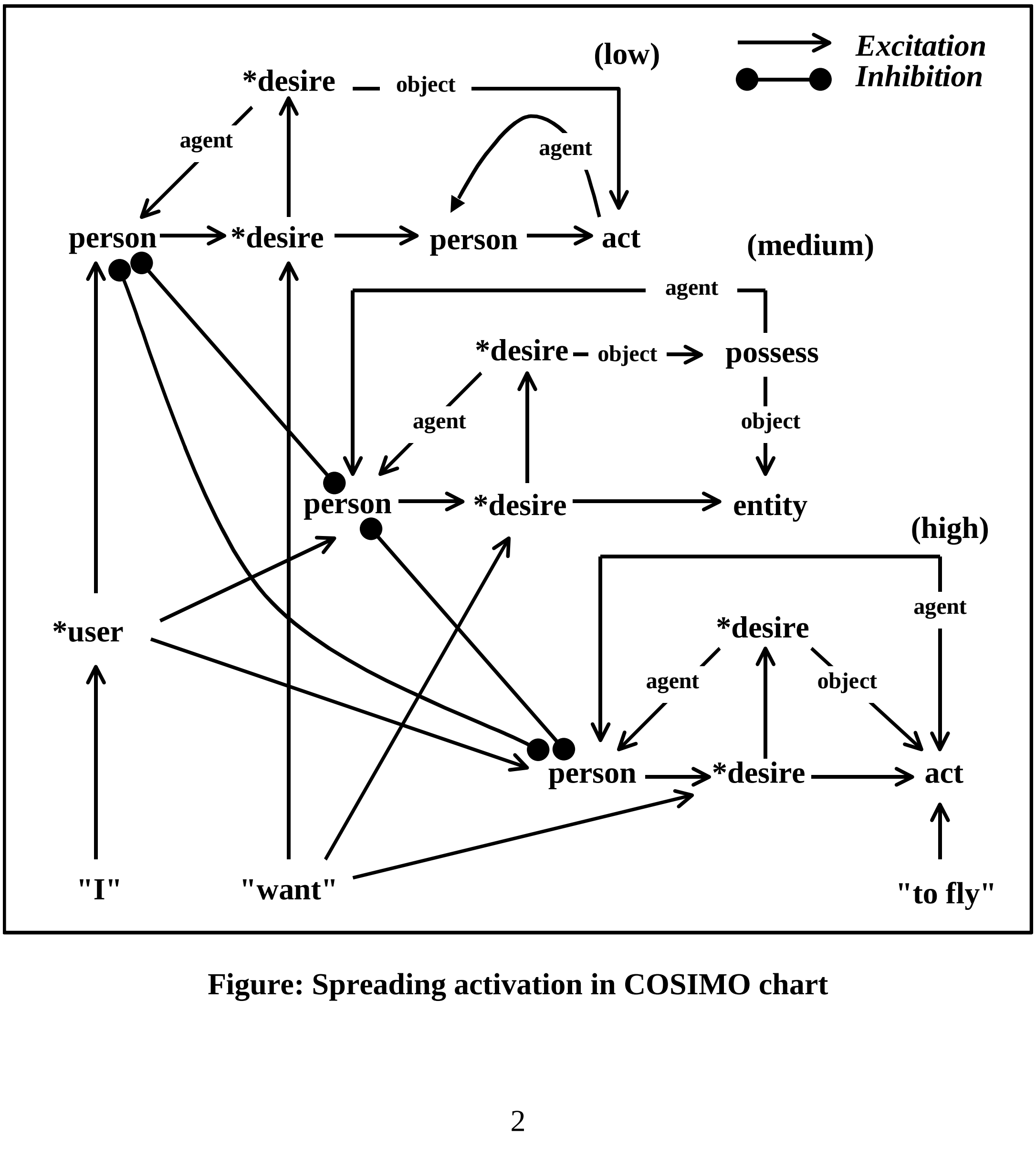

Figure: Spreading activation in COSIMO chart

As an example, let us consider a partial parse of the sentence from the above Figure. The initial processing of string "I want" results in addition of three competing templates (a, b, c), only one matches the expected continuation "to fly" (c), the remaining two are in fact near misses to the input (a omits the person, b is penalised for the use of incorrect semantic category). The spreading activation will prefer the structure with the highest number of shortest connections coming from input signs, thus the correctly matching template gets the highest excitation, inhibits the remaining two interpretations from receiving activation from input, then they slowly decay into their resting energy levels at which they get removed from the chart. The remaining chart constituents achieve equilibrium of up and down activation spreading after a number of chart cycles. If a network relaxes into a number of disjoint interpretations, their selection may use the level of activation as a degree of interpretation correctness.

2.0 Conclusions

We have described the working principles of the COSIMO system. The system prototype was implemented in CommonLisp and C on the Sun 3/60, and is still in its experimental stage. Current research concentrates on increasing the size of COSIMO's dictionary, better coverage of the application domain with formulae, and a significant extension to the pool of available templates. Overall, the results of even our preliminary experiments convince us that the addition of connectionist chart methods to symbolic NLU, and Preference Semantics in particular, greatly enhances the capability of natural language parsers.

3.0 Acknowledgements

We acknowledge the permission of Executive General Manager of Telecom Australia Research Laboratories to publish this paper.

4.0 References

Anderson, J.R. (1983): "A spreading activation theory of memory," Journal of Verbal Learning and Verbal Behavior, V22, 261-295.

Bengio, Y., Cardin, R., DeMori, R., and Merlo, E. (1989): "Programmable execution of multi-layered networks for automatic speech recognition," CACM, V32, N2, 195-199.

Bookman, L.A. (1987): "A microfeature based scheme for modelling semantics," IJCAI 87, 611-614.

Collins, A.M. and Loftus, E.F. (1975): "A spreading activation theory of semantic processing," Psychological Review, V82, 407-428.

Cottrell, G.W. and Small, S.L. (1983): "A connectionist scheme for modelling word sense disambiguation," Cog. and Brain Theory, 6(1), 89-120.

Cybulski, J.L., Ferra, H.L., Kowalczyk, A., and Szymanski, J. (1989): "Determining Word Lexical Categories with a Multi-Layer Binary Perceptron," Proc. Conf. IASTED Expert Systems and Neural Networks, Honolulu, Hawaii, 45-49.

Cybulski, J.L. and Jennings, A. (1989): "COSIMO: A Massively Parallel Integrated Parser," Proc. First Australia-Japan Joint Symposium on Natural Language Processing, Melbourne University, Vic, Australia, 15-27.

Gasser, M.E. (1988): "Connectionist model of sentence generation in a first and second language," University of California, Computer Science Department, Report CSD-880050.

Howells, T. (1988): "VITAL: A connectionist parser," Proc. Conf. Cog. Sci. 88, 18-25.

Jennings, A., Rowles, C.D. and Kowalczyk, A. (1989): "Natural language understanding in the MEDICI project," Proc. Int. Conf. on Comp. and Inf., Toronto, Canada 1989.

Jones, M.A. (1987): "Feedback as a coindexing mechanism in connectionist architectures," IJCAI 87, 602-610.

Kay, M. (1980): "Algorithm schemata and data structures in syntactic processing," CSL-80-12, Xerox Corporation.

Kukich, K. (1987): "Where do phrases come from: Some preliminary experiments in connectionist phrase generation," in Kempen, G. (ed.) Natural Language Generation: New Results in Artificial Intelligence, Psychology and Linguistics, Proc. of the NATO Advanced Research Workshop, 405-421.

Lytinen, S.L. (1986): "Dynamically combining syntax and semantics in natural language processing," AAAI 86, 574-578.

McClelland, J.L., Rumelhart, D.E. and Hinton, G.E. (1986): "The appeal of parallel distributed processing," in Rumelhart, D.E., McClelland, J.L., and the PDP Research Group (eds), Parallel Distributed Processing, V1, 3-40.

Minsky, M. (1986): The Society of Mind, New York: Simon and Shuster.

Parisi, D. and Nolfi, S. (1987): "Connectionist modelling of syntactic constraints on sentence processing, " Proc. IEEE First Int. Conf. on Neural Networks, V2, 505-512.

Pollack, J.B. and Waltz, D.L. (1984): "Parallel Interpretation of Natural Language," Proc. Int. Conf. 5th Gen. Comp. Sys., 686-691.

Slack, J.M. (1986): "Distributed memory: A basis for chart parsing," COLING'86, 476-481.

Small, S.L. (1987): "A distributed word-based approach to parsing," in L. Bolc (ed.) Natural Language Parsing Systems, Springer-Verlag, 161-202.

Stanfill, C. and Waltz, D.L. (1986): "Toward memory-based reasoning," CACM, V29, N12, 1213-1228.

Wilks, Y. (1975): "An intelligent analyser and understander of English," CACM 18, 5,264-274.