REPRESENTING CONTEXTUAL INFORMATION

Jacob L. Cybulski

Department of Computing
Royal Melbourne Institute of Technology

GPO Box 2476V
Melbourne, Victoria 3001, Australia

Abstract

The processes of understanding, learning, and communication by humans would not be possible without matching captured information with the knowledge already possessed; without relating together all aspects of the external situation; and without elimination of stimuli irrelevant to the facts and events of interest. Thus the incoming information needs to be analysed in external (situational) and internal (mental) contexts, and in order to perform this analysis it is important to distinguish the attention from the inference mechanism of the mind.

The successful implementation of an automated context sensitive system implies knowledge structures and functions organised to meet the goals of the attention/inference pair. Therefore this paper looks at some available Artificial Intelligence constructs to determine the knowledge representation which best suits the aims of contextual information processing. Different useful features of selected formalisms are synthesised giving a description of a knowledge model - the Context Representation System – CONTEXUS.

Introduction

Before we start talking about mechanisms of contextual information processing, we should elucidate the concept of context itself.

A simple definition is found in the Macquarie Dictionary.

Context is :

> "...The parts of a discourse or writing, which precede or follow, and are directly connected with, a given passage or word,"

or a less linguistic statement

> "...The circumstances or facts that surround a particular situation, event etc."

Also the Latin word "contextus" means a "connection". So in general, context can be viewed as a link between a concept (fact, event, word, etc) and its environment.

Our definitions need further clarification. We can analyse problems of contextual processing from both the linguistic and cognitive points of view, for these two fields deal with the problem, but differ radically.

The Chomskian linguistic theory is based on the assumption that everybody has an inherited, general mechanism for language processing : the "speaker-hearer's intrinsic competence". Any statement belonging to a given language can be synthesized via a sequence of Transformational Generative Rules; amongst these, there are some concerned with context sensing, necessary for grammatical agreement (Chomsky 1957, 1965).

Chomsky can see a necessity to relate the words of a sentence one to another. This aspect of contextual referents in a sentence places any natural language into a category of at least context-sensitive or even unrestricted languages (according to Chomsky's hierarchy).

In spite of the fact that syntax is extremely important in the process of understanding, we must not forget that language does not only consist of linguistic components. The real understanding of discourse cannot be based exclusively an language parsing (as suggested by may); actual knowledge about the world plays a vital part in defining the meaning of words, sentences, or stories. For instance, the same language taught to groups of people in separate contexts (e.g. different countries) results in semantically different language systems, even though the same syntax is used (Lambert, Havelka and Crosby 1958).

We must also remember that the mental processes producing language responses are not necessarily based on verbal associations between known concepts, but can consist of internalized non-verbal responses to various kinds of stimuli. These notions lead us to a completely different perspective an the context problem, that used by Cognitive Science and Cognitive Linguistics, where:

> "... the interest is in the effects of <u>context</u> and <u>knowledge of the world</u> on language behaviours." (Paivio and Begg 1981 p 67).

So we are interested in the meaning of perceived facts, which will be quite different depending an a comprehender's point of view, which is itself determined by his/her past experience, his/her knowledge and the current situation. However, in spite of the variety of individual experience, beliefs and knowledge, we do share certain universal meanings, such as language, mathematics and values, as a result of socialisation and education. To achieve shared meaning we need to restrict our knowledge system to a narrow domain relevant to the current situation, and ignore all unnecessary aspects of the world, which if active would make communication and understanding impossible. This the mechanism responsible for focusing attention has an important role in the process of information apprehension and its analysis. Context and attention are two inseparable elements of human cognition and therefore they both are treated with equal importance in this paper.

Importance of Context

It is evident that humans perceive not what IS but what they SEE (in most general sense). And what we really SEE is susceptible to our misconceptions, misjudgements, miscalculations and other natural "mis-abilities", which cause misperception of the outside world. Certainly we would not like to design an intelligent system which possesses all these miscarriages of Nature, but we must understand that an influence of knowledge of the world on our senses plays an important role in the process of human comprehension; and this ability to comprehend is certainly worth repeating in computer systems.

The first advantage of the sway of knowledge on senses is avoidance of transmission redundancy, for there is no need to transmit (on a higher level of cognition) all the perceived facts that are anticipated to be known; furthermore it is not necessary to construct entire concepts from the incoming stimuli when pre-existing taxonomies can be used; thirdly the sensing peripherals can be reduced to unintelligent scanning devices, because most of the processing is carried out on a higher level. There are many more valid facts to explore but let us concentrate on these three only.

As an example of virtues of the discriminatory capabilities of our minds let us present an observed situation:

> "... twenty girls sitting on pink four legged objects are rhythmically touching with their fingertips small cubes fixed into large plastic boxes, which produce white sheets of paper covered with black symbols ..."

Firstly we don't want to listen to the whole description while we focus our attention on the finger movements of one of these girls (redundancy reduction), secondly it would be much more comprehensible to refer to the picture by a single meaningful term, e.g. "a typing class" instead of reporting typical activities that must take place in such a situation (taxonomy), and lastly isn't it beautiful that we do not need to have special, discrete senses to recognize young girls, chairs and typewriters, and instead we can synthesize the whole image of dots on our retina (peripheral simplicity).

Thus the inherent skill of influencing perception by knowledge must be appreciated even though it carries potential dangers to the quality and correctness of the captured information.

The other side of perception can be associated not with previously acquired knowledge but with information which comes concurrently with the stimulus of interest. This information will be referred to as external context as opposed to the knowledge of the world - internal context. The effect of external contexts on the process of understanding and recognition is quite significant.

Assume that the previous example is set with one of the following three alternatives:

1) It happens in the Department of Secretarial Studies
2) All girls are 5 years old.
3) The four legged objects are real elephants

The first context confirms our guess that girls are participants in "a typing class". The second context excludes this possibility since preschool girls do not learn typing, in this case we would suspect that they are children playing with some sophisticated toys. The third alternative violates our common sense knowledge, and the whole story must be categorised as a fairy tale of a mad man's babble

We can see that both external (situational) and internal (mental) contexts are vital in the process of understanding. The meaning of facts and events can not be universal to all of us, but it is defined by the total information used by an individual to decode the message passed down to him.

Model

First of all we must distinguish three phases of information processing : stimuli acceptance, thinking, and response generation. The first and the last are of a very primitive nature, mapping the external world onto internal data structures (and the other way around), and the remaining one -thinking is of significant complexity and that is why it will be at the focus of this paper.

The sequence of these three phases suggests that thinking is a process which mediates between stimuli and responses, transforming than one into another.

Three mechanisms must then be distinguished, three mechanisms performing the three tasks: the Stimulus Mapper translating the set of in-flowing stimuli onto the internal structures, the Response Mapper generating a physical response from the internal behavioural pattern produced by the Thinking Mechanism.

The Thinking Mechanism is an information processing system capable of linking stimuli with responses depending on the current knowledge activity, this can be achieved by invoking two distinct tasks: focusing attention and reasoning. The former has a role in defining the scope of inference applicability depending on the stimulus and the latter is responsible for inductive and deductive conclusions, leading to the formation of responses.

The attention mechanism controls the flow of information to our memory (selection role of attention), and activates relevant knowledge from our memory (activation role of attention). External stimuli (low level of stimulation) carrying information about an event in context are projected onto the currently active memory creating a perceived context (PC) (Olson 1970). This active field of knowledge about the observation is a base for our reasoning processes which transform it into a new activation field - an inferred context (IC), that is a reflection of generated responses. However, because the attention and reasoning processes work in fact in parallel, the distinction between PC and IC is not physical but conceptual only, and thus both of them are indistinguishable from the currently active memory field - a focus space.

Needless to say, both a stimulus and response can also be mental processes (high level - mental); in addition the lack of external stimuli must be comprehended as a stimulus, the lack of response as a response.

The cooperation between the Attention Mechanism and the Inference Mechanism can be visualised as in the diagram (Figure 1). This simple model for mapping an external situation into a mental context, and then turning it into a response, can account for all the immediate effects of attention (after Norman 69): perception (stimulus selection), distinction (memory activation), conception (new context inference), and remembering (response). It should also be pointed out that attention and inference processes reside in the memory and as such are susceptible to their own activities.

Another significant remark must be made about a .term which has been mentioned and discussed - External Context (EC). It has been said that EC is extra information associated with the stimulus of interest, however all perceived facts are of equal importance for the thinking processes until they are categorised and recognized (at some early stage of knowledge processing) and as such it is not possible to determine which of these stimuli is in the focus of attention. Therefore the External Context is all possible information received from the Stimulus Mapper into the system.

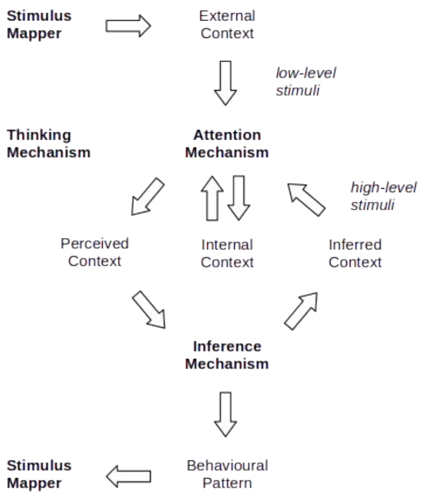

Figure 1. Inference / Attention Mechanism

Representation

We will now further develop the model to construct a functional description of the system representing contextual information. The representation must account for both reasoning and attention. However we will try to design a universal model of the attention mechanism, which could be to some extent independent of reasoning tasks. Here is a list of factors which are the required structural and functional elements of the context sensing mechanisms, together with the effects they have on the qualities of thinking:

- concepts (memory granulation)
- context spaces (physical binding)
- concept relationships (structure organizations)
- activatedness (distinction)
- activation spreading (focus control)
- relevant partitioning (removing concepts)
- relevant clustering (inserting concepts)

The characteristics of all these structures and functions are discusses in the following sections.

Primitives

Any representation system needs some structural and operational primitives which have universal meaning independent of the sensed context, or any other factors. The fact conceptualisations must be expressible in terms of basic units of knowledge. The idea of learning systems is to provide little to get much; that is why self organizing systems evolve their complexity from primitives.

Concepts

One of the structural primitives is of course an atomic memory unit used to represent an incoming stimulus and which is mappable onto a response. This information store will be referred to as a concept (or node). The structure of concepts depends primarily on the interpretation of in-flowing knowledge. And the interpretation is given by the inference mechanism. Ideally, the attention processes do not need to know the nature of concepts kept in the knowledge base. Therefore we can represent facts in an arbitrary form, e.g. as production rules (Newell and Simon 1972), Horn clauses (Kowalski 1979), slots of frames (Minsky 1975, 'Wipers 1975), partitions (Hendrix 1978), etc.

Spaces

It has been observed that human learning contexts are activated together with a recalled fact. Thus we do not memorise a single item at the time but the entire situational context as well (Grosz 1977, Reichman 1978). This knowledge is split into context spaces. The Idea of context spaces is to create subsets of knowledge which contain concepts related one to another in such a way that the inference mechanism does not need to refer to other context spaces to solve a specific problem. The context space's operations must tend to reshuffle information between active spaces to achieve the goal of local processing. This type of a physical binding of concepts is widely known from various knowledge models, such as Frames (Minsky 1975), Scripts (Schank and Abelson 1977), or Vistas (Hendrix 1978, Grosz 1978), which contain local knowledge vital to the process of reasoning, and which can determine when and why a shift of attention (swapping context spaces) should occur.

Relationships

Concepts and contexts themselves would be useless unless we are capable of organizing nodes into bigger conceptual structures, that means we must be able to relate them one to another. This leads to concept relationships (or semantic links) which are used by constructive and analytical reasoning, by search and matching processes, and other inference-related tasks. The meaning of these relationships is also dependent on the nature of the inference mechanism.

The structures described so far indicate that every concept must reside in same context space and therefore the connections between concept nodes can exist either inter-contextually - across two different context spaces, or intra-contextually - within one context space (Schank 1975). Various aspects of semantic links are illustrated and discussed elsewhere (Woods 1975).

Activatedness

The structures such as concepts, semantic links, and context spaces are necessary to represent knowledge in the system, and are related to the inference mechanism rather than the attention mechanism. On the other hand the attention mechanism also needs some memory structures enabling its activities. And because its primary task is to activate knowledge therefore we must incorporate a concept of activatedness into our system (Kantor 1977, after Hirst 1981). The effect of activation is to make a subset of knowledge readily available or retrievable for subsequent processing. Thus inactive information consisting of unactivated structures should elude manipulation, and most of the time should be hidden from the controlling processes.

The knowledge activatedness must account for all available primitive structures: concepts, spaces, and relationships.

Context space activation. The first goal of activatedness is to distinguish a set of context spaces relevant to the problem from the remaining knowledge - to account for perceived and inferred contexts. All highly activated context spaces become a part of the focus space. The focus space is used as a resource of knowledge for reasoning tasks; all the remaining Information kept in the knowledge base is inactive and is ignored by the inference mechanism. The connections between concepts in the focus can be categorised into inter-focus and intra-focus links (analogous to inter- and antra-contextual links).

Concept activation. The activatedness plays a role not only in activation of context spaces to form a focus space, but it is also useful in intra-contextual and inter-contextual manipulation of concepts and their relationships. Inference processes will have a number of alternative paths to follow within the active memory. Clearly, the more active the information link is, the more probability of its being chosen. In this way we can direct inferences and explain various cognitive effects such as having preferences (Wilks 1978) or being interested in something (Schank 1979). Furthermore inferences could use concept activation to determine the placement of various semantic links. Obviously the more active nodes should be considered as worth linking together, and the new relationship between them could be established.

Link activation. What should be the motives to change node or space activation? The answer seems to be easy. The context spaces related to the current problem should be included in the focus space, the ones which are unlikely to be used -removed. Here we have come to the concept of knowledge relevance. Some theorists propose a semantic distance as a ground for determining related items (small semantic distance - high relevance). However semantic distance between two arbitrary concepts is not a simple metric and its computation can involve quite complex inferences (Hirst 1981, a review of different semantic measures are given in Osgood et al 1978). We will define concept relevance in a simpler way: two concepts are said to be relevant (high relevance factor) if they are both in the focus of attention, they are both highly activated and they are semantically related (there exists a semantic link between then). If this happens a relevance link is established between two concepts. When two context spaces are connected by highly relevant inter-contextual links, than if the space restrictions permit they should be viewed, modified and analysed at the same time - they both should be included in the focus of attention.

The relevance between concepts is nothing else but the activatedness of concept to concept relationships, therefore relevance and semantic links overlap, however the relevance link is not associated with the inference mechanism, but with the attention mechanism, and thus it is not concerned with a relationship's interpretations such as the meaning of the link, its direction etc.

So far we have defined some structures necessary to explain human cognitive processes, as they could exist in a mechanical mind. A functional description of the model is also required. Therefore we must describe the operations permissible on the memory structures already defined.

Activation spreading

The activation of nodes, links and spaces is dynamic in its nature, therefore a process of activation spreading must exist.

The process has certain constraints drafted by same cognitive psychologists (see Collins and Loftus 1975):

a) the act of concept processing starts the activation spreading to neighbour-nodes, with intensity decreasing with the distance from the activated node

b) activation goes away gradually when the concept is not used

c) nodes intersecting active paths gain higher activity

d) others

These properties should be helpful in choosing an appropriate model for activatedness, relevance and activation spreading. The activity and relevance can be represented as probabilistic measures. In this case let us define the concept and context space activation as a probability of them being processed, and the relevance as a probability of moving the processing to the connected node in the next cycle (more than one node can be processed at the same time). The relationships between these measures can be given as follows :

t	cycle number (time)
A(n,t)	node "n" activation at time "t"
R(n1,n2)	relevance of "n2" to "n1"
A(n,0)	is an initial activation

$$\forall_n \sum_m R(n,m) = 1$$

$$\forall_n \forall_{t\geq 0}\, A(n,t+1) = \frac{\sum_m A(m,t) \times R(m,n)}{\sum_m R(m,n)}$$

It is not difficult to notice that these formulae describing the activity flow in the network fulfil the Collins and Loftus' requirements (however not all of the original ones like the existence of threshold for activation firing).

a. The activation spreads from any activated node to semantically related concepts. This activation decreases with distance because it can be shown that the activity of the node n_{k+i} received from the node node n_0 via n_1, n_2....,n_k equals to:

$$\mathrm{A}_{n_1,n_2,\ldots,n_k}(n_{k+1}, t+k) =$$

$$\mathrm{A}(\mathrm{n}_0, t) \times \prod_{i=0}^{k} \frac{R(n_i, n_{i+1})}{\sum_m R(m, n_{i+1})}$$

is decreasing (for majority of nodes) because:

$$\forall_{n,i}\ \frac{R(n_i,n_{i+1})}{\sum_m R(m,n_{i+1})} \leq 1$$

b. The activation of a non-activated node goes away gradually because the concept self-activation is equal to:

$$\forall_{n,k,t}\ A(n,t+k) = A(n,t) \times \left(\frac{R(n,n)}{\sum_m R(m,n)}\right)^k$$

and is decreasing with k (for majority of nodes) because:

$$\forall_n\ \frac{R(n,n)}{\sum_m R(m,n)} \leq 1$$

c. Activation of intersecting nodes is a sum of individual activations, so the more incoming links the higher activity of the node is.

Future experiments with the model can prove that some corrections to the above formulae are needed.

The context spaces's activity changes should be directed by the inference mechanism, trying to use inter-contextual links. Every attempt to move to unfocused context space will increase the activity of the space. Also the relevance of different links should Change according to the link use, the more frequent usage the higher relevance.

Partitioning and Clustering

The attention mechanism is responsible mainly for changing knowledge activity, but it should also influence the actual contents of context spaces depending on the relevance of intra-context and inter-contextual links.

It is necessary to have some tools to modify concepts and contexts spaces. Modifications of concepts within one context space are dependent on the inferential abilities of the system, which we try to keep unrestrained so that the Context Representation System could be defined for a variety of representation models. However some changes in the structure of a context space are the responsibility of attention processes. There are two significantly important functions: relevant knowledge partitioning and relevant knowledge clustering. The first is a division of context space into essential and non-essential parts (the latter of which can be removed if necessary); the second is acquiring new facts via inter-contextual links from other context spaces - the facts which are conceptually relevant to the focused context space (even if it creates redundancy within the knowledge base). Let us have an example of these two operations (Figure 2).

Con-text	Relationship	Rlv	Rule
John	Join IS-A Man	0.4	R1 .
	John HAS 1 leg	0.4	R2
	Jahn HAS 2 hands	0.2	R3
Fred	Fred IS-A man	0.6	R4
	Fred HAS 2 hands	0.4	R5
Jack	Jack IS-A man	0.7	R6
	Jack HAS 2 hands	0.3	R7
Man	Man HAS brain	0.1	R8
	Man HAS 2 legs	0.1	R9
	Man INSTANCE John	0.4	R10
	Man INSTANCE Fred	0.2	R11
	Man INSTANCE Jack	0.2	R12

Figure 2. Example

We can see that some of the contexts contain links which connect concepts of relatively little relevance. An example of such a link is x HAS 2 hands in the contexts of "John", "Fred", and "Jack". This fact qualifies for removal from these contexts (partitioning). On the other hand the same fact has a significant relevance to the intersecting node - “Man" in the context “Man" (so that it could be clustered with it). The clustering process requires a duplication of some nodes via high relevant inter-contextual links into the context of interest, then the inference mechanism can transform the copied data into appropriate format.

In this example we can cluster rules R3, R5, and R7 with the context “Man”. The rules will be included into the context space with all the nodes ("John", "Fred", "Jack", "2 hands") and the appropriate links ("HAS”). The inter-contextual links used in the process of clustering will remain intact, however all the relevance weights must be changed. Then the inference mechanism can generalize the fact of having two hands as a property of a man and the unnecessary links and nodes can be removed. Now partitioning of spaces containing irrelevant information can be done at any time.

Inferences

It has been mentioned that the inference mechanism can be designed and analysed independently of the attention mechanism. However the notion of activatedness severely restricts the selection of a formalism fully compatible with the outlined model. Also the physical division of knowledge into context spaces affects reasoning (e.g.: the system should avoid unnecessary shifts of attention, unless the current set of focused context spaces is processed exhaustively).

Fortunately there exist formalism which can account for some of these restrictions.

An obvious choice are models which have structures similar to these of relevance and activatedness. One of them is an inference net, based on probabilistic models (PROSPECTOR - Duda, Gashing, and Hart 1979, INFERNO - Quinlan 1982, MYCIN - Shortliffe 1976, or EMYC1N - Van Melle 1980). These systems utilize tools similar to the ones used in the activation spreading, however their goal is to use and construct inferences dealing with fuzziness, uncertainty, and ambiguity, rather than defining procedures selecting inferences and defining scope of their applicability (our case). A decision to choose a probabilistic model for knowledge representation combined with the attention mechanism defined above leads to a consistent description of the entire system.

Clearly we are not restricted to the PROSPECTOR-like systems, and we can reach for various well defined tools such as predicate logic (and other forms of logic as well). Not all formal logic systems will be appropriate. For instance the use of Horn clauses with Robinson's resolution method (Robinson 1979) is not possible because of the knowledge partitioning into inter-linked context spaces. Here we would have to devise complex strategy for inference backtracking from an inter-contextual link, and retrying it upon the context space inclusion into the focus space. Much better candidates for representing knowledge in logic are those models which use structures similar to the ones we have described.

There were several attempts to use graphical representation of logic: Peirce's Existential Graphs (Roberts 1973), Sowa's Conceptual Structures (Sowa 1984), or Hendrix's Partitioned Semantic Networks (Hendrix 1978). The last proposal is especially interesting since it embeds higher-order logic and uses knowledge partitioning, however the information activatedness is not a part of this formalism.

Other non-logical systems based on either ad hoc inference systems, or linguistic and cognitive models should also be included in this analysis. Some of them have interesting similarities to our proposal.

Most of the cognitive and linguistic models are based on a concept of associative semantic networks, which are a good base for our intra-contextual structure, however only few of these models deal with the problem of knowledge activation. Here we can include Memory Schemata (Bobrow and Norman 1975), neural nets and K-Lines (Minsky 1963, Minsky 1977, Minsky 1981). All these three theories of memory explain various natural phenomena occurring in human minds, and try to give a synthetic description of memory, and the activation spreading in it. K-Lines are the closest model to ours. However, the Collins and Loftus model (which has been used here) does not assume the hierarchical knowledge processing which is forced in the K-Line formalism.

In general any of the models discussed would fit into the attention/inference framework, and the future work an the Context Sensitive System (CONTEXTUS) will attempt to develop a set of tools and methodologies to account for all of these inference systems.

CONTEXTUS

The arguments from the previous sections aim at fully functional description of Context Sensitive Systems.

The basic structures of the system have been listed (concepts, relationships, context spaces), the operations on these structures defined (activation spreading, relevant partitioning, relevant clustering, inferences), the representation of the system discussed (activations, relevance, etc). Now let us describe the model's benefits:

a. The formalism restricts but does not impose the inference model.

b. The system can store knowledge redudantly and ambiguously within the knowledge base. The uniqueness and unambiguity is necessary only within the context space (and it is up to the reasoning processes to achieve it). This enables context switching to similar (redundancy) but different or even contradictory contexts via inter-contextual links.

c. The fuzzy and uncertain descriptions can be kept in the knowledge base, however they are a side effect of the knowledge activatedness. Degrees of activity and relevance correspond to degrees of certainty, and fuzziness, and they depend on the current status of the knowledge base.

d. The available knowledge organizations are not restricted to hierarchical structures. Although the hierarchical processing is possible via concentrically spreading activation. The selection of multi-hierarchical links can be activation and relevance driven.

e. The depth of processing is easy to define. The focus space is an ultimate processing scope, and the activation spreading guarantees that this field cannot expand indefinitely (unprocessed nodes fade).

Many other Important issues can be raised, however some experimental data is necessary to support further claims.

Bibliography

Bobrow, D.G. and Collins, A. (ed 1975): Representation and Understanding: Studies in Cognitive Science, New York, NY: Academic Press.

Bobrow, D.G. and Norman, D.A. (1975): "Some Principles of Memory Schemata," in Bobrow and Collins 1975.

Chomsky, N. (1957): Syntactic Structures, The Hague: Mouton.

Chomsky, N. (1965): Aspects of the Theory of Syntax, Cambridge, MA: The MIT Press.

Collins, A.M. and Loftus, E.F. (1975): "A Spreading Activation Theory of Semantic Processing," Psychological Review 72:6, 407-428.

Duda, R., Gashing, J. and Hart, P. (1979): Model design in the PROSPECTOR consultant system for mineral exploration. In: Mickie (Ed.): Expert Systems in the Microelectronic Age, Edinburgh University Press, 153--167.

Grosz, B.J. (1977): "A representation and use of focus in a system for understanding dialogs," In B Grosz, K Sparck Jones and BL Webber (eds), Readings in Natural Language Processing, Los Altos, CA: Morgan Kaufmann Publishers Inc, 1977.

Grosz, B.J. (1978): "Focus Spaces: A Representation of the Focus of Attention," in Walker 1978, 269-286.

Hendrix, G. (1978): "Semantic Knowledge," in Walker 1978, 121-228.

Hirst, G. (1981): Anaphora in Natural Language Understanding: A Survey, Berlin Heidelberg: Springer-Verlag.

Kantor, R.N. (1977): The Management & Comprehension of Discourse Connection by Pronouns in English, PhD Thesis, Department of Linguistics, Ohio State University.

Kowalski, R. (1979): Logic for Problem Solving, New York NY: North-Holland.

Kuipers, B.J. (1975): "A Frame for Frame: Representing Knowledge for Recognition," in Bobrow and Collins 1975, 151-184.

Lambert, W.E., Havelka, J. and Crosby, C. (1958): "The Influence of Language Acquisition Contexts on Bilingualism," in Journal of Abnormal and Social Psychology 56, 239-244.

Minsky, M. (1963): "Steps towards artificial intelligence," In E.A. Feigenbaum and J. Feldman, editors, Computers and Thought, McGraw-Hill, 406-450.

Minsky, M. (1975): "A Framework for Representing Knowledge," in Winston 1975, 211-280.

Minsky, M. (1977): "Plain talk on neurodevelopmental epistemology" In Proceedings of the Fifth International Joint Conference on Artificial Intelligence, Cambridge, MA, August, 1083-1092.

Minsky, M. (1981): "K-Lines: A Theory of Memory," in Norman 1981, 87-101.

Newell, A. and Simon, H.A. (1963): "GPS: A Program that Simulates Human Thought," in Feigenbaum and Feldman 1963, 279-293.

Norman, D.A. (1969): Memory and Attention: An Introduction to Human Information Processing, New York, NY: Wiley & Sons.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. Psychological Review, 77, 257-273.

Osgood, C., Suci, G. and Tannenbaum, P. (1967): The Measurement of Meaning. University of Illinois Press.

Paivio, A. and Begg, I. (1981): Psychology of Language, Englewood Cliffs, NJ: Prentice-Hall.

Quinlan, J.R. (1983): "Inferno: A cautious approach to uncertain inference," The Computer Journal 26 (3), 255-269.

Reichman, R. (1978):Conversational coherency. Cognitive Science 2 (4), 283-327.

Roberts, D.D. (1973): The Existential Graphs of Charles S. Peirce. Mouton and Co.

Robinson, J.A. (1979): Logic: Form and Function, Edinburgh: Edinburgh University Press.

Schank, R.C. (1975): Conceptual Information Processing, Amsterdam, Holland: North-Holland.

Schank, R.C. (1979): "Interestingness: Controlling Inferences," Artificial Intelligence 12, 273-297.

Schank, R.C. and Abelson, R.P. (1977): Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures, Hillsdale, NJ: Lawrence Erlbaum Associates.

Shortliffe, E.H. (1976): Computer Based Medical Consultations: Mycin, New York, NY: Elsevier North Holland.

Sowa, J.F. (1984): Conceptual Structures, Reading, MA: Addison-Wesley Pub. Co.

Van Melle, W., Scott, A. C., Bennett, J. S. and Peairs (1981): M. The Emycin Manual, Technical Report STAN-CS-81-885, Computer Science Department, Stanford University, October.

Walker, D.E. (ed 1978): Understanding Spoken Language, New York, NY: North-Holland.

Wilks, Y. (1978): "Making preferences more active," Artificial Intelligence, 11(3).

Winston, P.H. (ed 1975): The Psychology of Computer Vision, New York, NY: McGraw-Hill.

Woods, W.A. (1975): "What's in a Link: Foundations for Semantic Networks," in Bobrow and Collins 1975, 35-82.