Psychotocical Review
1988, Vol. 95, No. 2, 163-182

0033-:«X;88/$00.75

The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model

Walter Kintsch
University of Colorado

In contrast to expectation-based, predictive views of jibeouue comprehension, a model is developed in which the initial processing is strictly bottom-up. Word meanings are activated, propositions are formed, and inferences and elaborations are produced without regard to the discourse context. However, a network of interrelated items is created in this manner, which can be integrated into a coherent structure through a spreading activation process. Dau concerning the time course of word identification in a discourse context are examined. A simulation of arithmetic word-problem understanding provides a plausible account for some well known phenomena in this area.

Discourse comprehension, from the viewpoint of a computational theory, involves constructing a representation of a discourse upon which various computations can be performed, the outcomes of which are commonly taken as evidence for comprehension. Thus, after comprehending a text, one might reasonably expect to be able to answer questions about it, recall or summarize it, verify statements about it, paraphrase it, and soon.

To achieve these goals, current theories use representations with several mutually constraining layers. Thus, there is typically a linguistic level of representation, conceptual levels to represent both the local and global meaning and structure of a text (e.g., the micro- and macrostructure, constituting the *text base in* van Dijk & Kintsch, 1983), and a level at which the text itself has lost its individuality and its information content has become integrated into some larger structure (e.g., van Dijk & Kintsch's situation model).

Many different processes are involved in constructing these representations. To mention just a few, there is word identification, where, say, a written word like *bank* must somehow provide access to what we know about banks, money, and overdrafts. There is a parser that turns phrases like *the old men and women* into propositions such as AND(OLD[MEN],OLD [WOMEN]]. There is an inference mechanism that concludes from the phrase *The hikers saw the bear* that they were scared. There are macro-operators that extract the gist of a passage. There are processes that generate spatial imagery from a verbal description of a place.

It is one thing for a theorist to provide some formal description (e.g., a simulation model) for how such processes can occur and for what the computational steps were that led to a particular word identification, inference, or situation model. It is quite another to control construction processes in such a way that at each point in the process exactly the right step is taken. Part of the nroblem has to do with the characteristic ambiguity of language· How do we make sure that we access the financial meaning of *bank,* and not the meaning of *riverbank?* Why did we parse the *old men and women* as we did—maybe the women were not old at all. Why did we infer that the hikers were scared rather than that they had their eyes open, or a myriad of other irrelevances? Of all the many ways macro-operators could be applied, how did we get just the right sequence to reach a plausible gist without making the wrong generalizations? The number of possible alternative steps is distressingly large in constructing discourse representations, and without firm guidance, a computatiov-a! model could not function properly for long. That is where knowledge comes in.

Gencrr;! knowledge about words, syntax, the world, spatial relations—in short, general knowledge about anything—constrains the construction of discourse representations at all levels. Indcco. this is what makes it possible to construct these representations. There is a striking unanimity among current theories about how this is done.

Our conceptions about knowledge use in discourse comprehension are dominated by the notions of top-down effects and expectation-driven processing. Knowledge provides part of the context within which a discourse is interpreted. The context is thought of as a kind of filter through which people perceive the world. At the level of word recognition and parsing, it lets through only the appropriate meaning of an ambiguous word or phrase and suppresses the inappropriate one. Through semantic priming, the feature counter of the logogen for *bank as a financial institution* will be incremented and will reach its threshold before that of *riverbank* in the right context (Morton. 1969). Parsing a sentence is often thought of as predicting each successive constituent from those already analyzed on the basis of syntactic rules (Winograd, 1983). Scripts, frames, and schemata constrain the inferences an understander makes (as in Schank & Abelson, 1977), thereby preventing the process from ^ir»g swamped in a flood of irrelevancies and redundancies. Arithmetic strategies generate just the right hypothesis in solving a word problem and preclude the wrong ones (Kintsch &

This research was supported by Grant MH 15872 from the National Institute of Mental Health. The work on word arithmetic problems was supported by Grant BNS 8741 from the National Science Foundation.

Correspondence concerning this article should be addressed to Walter Kintsch, Department of Psychology, Univeristy of Colorado, BouMe;. Colorado 80309-0345.

Greeno, 1985), in a word, knowledge makes understanding processes smart: It keeps them on the right track and avoids exploring blind alleys. People understand correctly because they sort of know what is going to come. This program of research is well expressed by the following quotation from Schank (1978, p. 94), which served as a motto for Sharkey's (1986) model of text comprehension:

> We would claim that in natural language understanding, a simple rule is followed. Analysis proceeds in a top-down predictive manner. Understanding is expectation based. It is only when the expectations are useless or wrong that bottom-up processing begins.

Hmpirically, this position is questionable: Even fluent readers densely sample the words of a text, as indicated by their eye fixations (Just & Carpenter, 1980), making the bottom-up mode appear the rule rather than the exception. Computationally, it is not an easy idea to make work, ft is difficult to make a system smart enough so that it will make the right decisions, yet keep it flexible enough so that it will perform well in a broad range ; of situations. On the one hand, one needs to make sure that exactly the right thing (word meaning, proposition, inference) will be constructed: for that purpose one needs powerful, smart rules that react sensitively to subtle cues. On the other hand, humans comprehend well in ever-changing contexts and adapt easily to new and unforeseen situations; for that purpose one needs robust and general construction rules. Scripts and frames, as they were first conceived, are simply not workable: If they are powerful enough, they are too inflexible, and if they are general enough, they fail in their constraining function. This dilemma has long been recognized (e.g., Schank, 1982; van Dijk & Kintsch, 1983), and efforts have been undertaken to make expectation-driven processes sufficiently flexible (e.g., Schank's memory organization packets, or MOF's). In this article, an alternative solution to this problem will be explored.

Construction of Discourse Representations

The traditional approach to modeling knowledge use in comprehension has been to design powerful rules to ensure that the right elements are generated in the right context. The problem is that it is very difficult to design a production system powerful enough to yield the right results but flexible enough to work in an environment characterized by almost infinite variability. The approach taken here is to design a much weaker production system that generates a whole set of elements. These rules need to be just powerful enough so that the right element is likely to be among those generated, even though others will also be generated that are irrelevant or outright inappropriate. An integration process will then be used to strengthen the contextually appropriate elements and inhibit unrelated and inappropriate ones. Weak productions can operate in many different contexts because they do not have to yield precise outputs; on the other hand, a context-sensitive integration process is then required to select among the outputs generated. The integration phase is the price the model pays for the necessary flexibility in the construction process.

The model proposed here has been termed a *construction-integration* model to emphasize its most salient feature. It combines a construction process in which a text base is constructed from the linguistic input as well as from the comprehender's knowledge base, with an integration phase, in which this text base is integrated into a coherent whole. The knowledge base is conceptualized as an associative network. The construction process is modeled as a production system. Indeed, it is a generalization of the production system used in earlier work, such as the simulation-of-comprehension processes developed by Fletcher (1985) and Dellarosa (1986) after the model of Kintsch and Greeno (1985). The main difference is that instead of precise inference rules, sloppy ones are used, resulting in an incoherent, potentially contradictory output. However, this output structure is itself in the form of an associative net, which can be shaped into a coherent text base via relaxation procedures in the connectionist manner (e.g., Rumelhart & McClelland, 1986). Thus, the model represents a symbiosis of production systems and connectionist approaches.'

Certain limitations of the present article are worth noting at this point, for it does not offer a solution to all the problems in discourse understanding. Thus, it is not primarily concerned with the specific strategies (or rules) for the construction of text propositions or inferencing. Instead, it relies in this respect on what is available in the literature as well as on whatever future researchers will be able to come up with. The only point it makes is that whatever these strategies or rules are, they will be easier to formulate within the present framework, which allows them to be both weaker and more general. Thus, one need not worry about constructing just the right inference, but can be content with a much sloppier rule. Sometimes, of course, even the latter type of rule may oe hard to come by, whereas in other cases (e.g., in the word problems discussed later) promiscuous hypothesis generation is straightforward (while selecting just the right one can be tricky).

Knowledge Representation

The process of constructing a discourse representation relies heavily on knowledge. To understand how it operates, one must first have an idea of how the to-be-used knowledge is organized. Typically, theorists have tried to create knowledge structures to support smart processes: semantic nets, frames, scripts, and schemata. As has been argued elsewhere (Kintsch, in press), such fixed structures are too inflexible and cannot adapt readily enough to the demands imposed by the ever-changing context of the environment. Instead, a minimally organized knowledge system is assumed here in which structure is not prestored, but generated in the context of the task for which it is needed. An associative net with positive as well as negative interconnections serves this purpose.

Knowledge is represented as an associative net, the nodes of

[1] Conceivably, a purer connectionist model might be constructed. In the present mcxlel, an associative knowledge net is used to build a text-base net, which is then integrated. McClelland (1985) has put forth the idea of a connection information distributor, which is a subnetwork in which the units are not dedicated and connections are not hardwired. Instead, this subnetwork is programmable by inputs from the central network where the knowledge that controls processing in the subnetwork is stored. One could say that the production rules in the present model have the function of programming such a subnetwork.

which are concepts or propositions.[2] The nodes in this net are interconnected. Connections among nodes have a strength value, which may be positive, zero, or negative, ranging from 1 to -1. Nodes consist of a head plus a number of slots for arguments. Thus, the nodes ofthe knowledge net are formally equivalent to the propositions used to represent texts (e.g., Kintsch, 1974),[3] The slot specifies the nature ofthe relation between the head and the argument. Slots may represent attributes, parts, cases of verbs, or arguments of functions. They need not be named, but may be named ifthe relation is a common one (such as the cases of verb frames). The arguments of a proposition are concepts or other propositions. The number of arguments in a proposition may vary from one to some small number. Examples of common types of nodes in the knowledge net are (a) MARY, (b) CAKE. (C) SWEETfCAKE], (d) BAKE[agenttMARY.object:CAKE], (e) CONSEQUENCE(condition:NOT[watch] agent: MARY,object:CAKE],effect:BURN[object:CAKE]]. Examples A and B are lexical nodes that have associated with them perceptual procedures that identify certain patterns in the environment—either the objects themselves or the written or spoken words, such as MARY and CAKE, respectively. In the following I shall not deal with these perceptual procedures explicitly The semantic and associative relations into which MARY and CAKE enter, which constitute a part ofthe general knowledge net, are the focus of interest here. MARY and CAKE appear as arguments in Examples C through E in various roles (the agent and object slots, etc.).

There are two ways of looking at the list of propositions in Examples A through E. On the one hand, it could be considered simply as a portion of a general knowledge network, whereas on the other hand, it could be considered the prepositional base of a (brief) discourse, in which a particular Mary bakes and bums a particular cake.[4] Thus, the elements ofwhich knowledge nets and text bases are constructed are the same. Indeed, as will lx? detailed later, text bases are formed by selecting, modifying, and rearranging prepositional elements from the knowledge net. However, text bases are not part ofthe knowledge net, but separate structures with their own properties.

Concepts are not defined in a knowledge net. but their meaning can be constructed from their position in the net. The immediate associates and semantic neighbors of a node constitute its core meaning. Its complete and full meaning, however, can be obtained only by exploring its relations to all the other nodes in the net. Meaning must be created. As a first step one could add all propositions in the net directly related to a node to obtain what Mudersbach (1982) termed the first level of meaning; then all propositions directly related to the propositions at the first level can be added to form a second level, and so on, until the whole knowledge net is involved. Note, however, that such a construction is a theoretical exercise without direct psychological correspondence. It is not possible to deal with the whole, huge knowledge net at once. Instead, at any moment only a tiny fraction ofthe net can be activated, and only those propositions of the net that are actually activated can affect the meaning of a given concept/Thus, the meaning of a concept is always situation specific and context dependent. It is necessarily incomplete and unstable: Additional nodes could always be added to the activated subnet constituting the momentary meaning of a concept, but at the cost of losing some of the already activated nodes.'

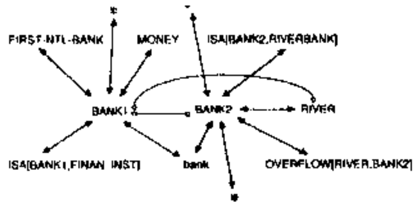


figure I. A fragment of the associative net for BANK. (Positive connections are indicated by arrows, negative ones by circles; asterisks indicate further, unnamed nodes.)

The notion of an associative net is not unfamiliar, but it is usually thought of as relating concepts only, not prepositional nodes. Two extremely simple examples will illustrate the nature of such an associative net. First, consider the representation of the homonym BANK in an associative net. Positive connections are indicated by arrows, negative ones by circles. Asterisks indicate further, unspecified nodes. Of course, each of the concepts arid propositions shown in Figure 1 participate in ihe general knowledge net beyond the single connection shown here. As a second example, consider the proposition BAKE[agent:PERs<>N,object:CAKE] (see Figure 2). Once again, only a fragment of the complete network is shown, just to illustrate certain tyrx;s of connections.

Representing knowledge in a propositiona! network has several advantages. Primarily, it provides a common format for the knowledge base and for the mental representation of discourse. Furthermore, we have by now considerable experience woi xing with propositional structures, whereas other forms of represen-

[2] Formally, concepts and propositions can be treated alike (e.g., Anderson, 1980).

[3] This use of the term *proposition* differs from the standard one in logic. Furthermore, not all authors who use comparable semantic units in iheir analyses use the same term. For instance, Dik (1980) talked about *predicates* and *terms* combining to form *predications.* Wilensky (1986) used *relation* and *aspectuals* In spite ofthis terminological disarray and the possibility of confusion with the meaning of proposition in logic, *proposition* appears to be the most widely accepted term and will be retained here.

[4] The extreme informality ofthis notation is chosen for ease of exposition. Frequently, of course, a more precise formalism is needed. It is fairly straightforward to elaborate the present informal notation whenever that is the case. For example, in the computer simulation of word problem solving by Dellarosa (1986), the LOOPS language provides a ready-made type-token distinction. There seems to br no reason, however, to burden a general discussion like the present one with a cumbersome, formal notation when it is not needed.

[5] As with *proposition,* this is a nonstandard use ofthe term's meaning. Meaning is used here as shorthand for the momentary, subject- and situation-specific activated semantic and experiential context of a concept. Clearly, this is not what many people have in mind when they speak about the meaning of a word—though it is a conception of meaning quite appropriate for a psychological processing model.

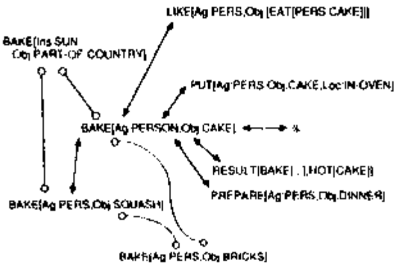


Figure 2 A fragment of the associative net for BAKE.

tation are less well understood (e.g., the spatial-imagery and linear structures of Anderson, 1983; the mental models of Johnson-Laird, 1983; or whatever the appropriate representation in the affective system might be, as in Zajonc, 1980). However, the decision to use a propositionai representation does not imply that all other forms of knowledge are to be considered unimportant or nonexistent. It would be desirable to expand the model to include nonpropositional representations, but one would first have to learn how to operate with such forms of representation.

Construction Processes

The steps in constructing a text base according to the construction-integration model involve, (a) forming the concepts and propositions directly corresponding to the linguistic input; (b) elaborating eacn of these elements by selecting a small number of its rhost closely associated neighbors from the general knowledge net; (c) inferring certain additional propositions; and (d) assigning connection strengths to all pairs of elements that have been created.

The result is an initial, enriched, but incoherent and possibly contradictory text base, which is then subjected to an integration process to form a coherent structure.

In Step A of this process, a propositionai representation of the text is constructed from a parsed linguistic input, such as the words of a text with suitable syntactic annotations, and from a knowledge system as envisioned earlier. Note that the parser itself is not a part of the present model. The basic process of proposition building has been described in van Dijk and Kintsch (1983, chapter 4) and Kmtsch (1985). I will illustrate it here with some simple examples. Consider the sentence *Mary bakes a cake.* The parser output needed is *Mary* (agent of BAKE) *bakes* (predicate) *a cake(*object of BAKE). *Mary, bake,* and *cake* activate their corresponding lexical nodes, and MARY and CAKE are assigned the roles of agent and object in the BAKE proposition. As was suggested in Figure *2,* BAKE requires a PERSON as agent, hence a test is made whether MARY is a person. This may either involve a search through the knowledge net for the proposition ISA[MARY,PERSON] or, should that search prove unsuccessful, an attempt to infer this proposition (e.g., the net may contain only propositions to the effect that MARY is a name and that persons have names; exactly how such problem-solving activity occurs within an associative net will not be considered here).

The present model, however, differs in a significant way from my earlier conceptions: It does not require that the right, and only the right, proposition always be formed. Instead, the construction rules for building propositions can be weakened, allowing for the formation of incomplete or "wrong" propositions. Proposition building is on-line, and frequently, all the relevant information for building just the right one is not available on-line, leading to false starts or incompleted attempts. In the aforementioned example, this has no interesting consequences; for example, if in response to the phrase *Mary bakes* . the proposition BAKEfMARY.Sj—the dollar sign indicates an unfilled slot—is formed, it will simply be replaced by the complete proposition when the rest of the sentence is processed. However, consider an example discussed by Fra/ier and Rayner (1982): *The linguists knew the solution of the problem would not be easy.* Here, the on-line construction of propositions is not so simple. First, the proposition KNOW[LINGUISTS,$] is formed. Then, by the strategy of minimal attachment, the subsequent noun phrase is interpreted as the object of KNOW, yielding KNOW[LINGUISTICS,SOLUTION]. The final verb phrase, however, requires a subject, so |NOT(EASY{SOLUTION]] is constructed. As Frazier and Rayner pointed out, this does not involve a reinterpretation of the sentence. Subjects do not go back, noting in some way that *solution of the problem* had been attached to the wrong proposition, and repair this error. Instead, the incorrectly formed KNOW proposition somehow just disappears; the description of the integration process that follows shows how.

A third example of proposition building, involving pronoun identification, will be discussed here. There exists good psychological evidence that pronouns may activate more than one possible referent (e.g., Fredenksen, 1981). Thus, in *The lawyer discussed the case with (he judge He said "I shall send the defendant to prison"* the following propositions would be formed: DISCUSS [LAWYER . JUDGE .CASE]; SAY I LAWYER, [SEND (LAWYER.DEFENDANT , PRISON] }], and SAY (JUDGE , [SEND | JUDGE, DEFENDANT.PRISON]]). Eventually, of course, the right interpretation comes to dominate the wrong one, as will be shown shortly.

In Step B of the construction process, each concept or proposition that has been formed in Step A serves as a cue for the retrieval of associated nodes in the knowledge net. The retrieval process itself is modeled after well-known theories that have been developed and tested in the memory literature (Raaijmakers & Shiffrin, 1981). Suppose that node / in the knowledge net is positively associated with η other nodes in the net. Let *s(ij)* be the associative strength between nodes *i* and/ Then the probability that the retrieval cue y will retrieve node y is

(1)

Note that each concept or proposition in the text base serves as an independent retrieval cue, hence the particularly simple form of the retrieval process. (An intersection search would be required if the items in the text base acted as a compound cue.)

On each retrieval attempt, an item among the associates of *i* is selected according to Equation 1. A sampling-with-replacement process is assumed so that dominant associates may be retrieved more than once The number of retrieval attempts with item / as the cue is assumed to be fixed and is a parameter of the model, *k* In the examples that follow, *k* was chosen to be 2 or 3, mostly to reduce the complexity of these examples. However, one may speculate that the most realistic value of *k* would not be much higher, perhaps between 5 and 7.

Consider some simple examples.

1. Suppose the word *bank* is presented as part of a text. It will activate the lexical nodes BANKι (financial institution) as well as BANK2 (riverbank), plus some of their associates; for example, the construction process might pick from Figure I: BANK!, MONEY, FIRST-NATIONAL-BANK, BANK2. RIVER, OVERFLOW[R1VER,BANK2J.

2. Suppose the sentence *Lucy persuaded Mary to bake a cake* is presented as part of a text. The parser should provide a phrase structure tree as output, from which the proposition PERSUADE(LUCY.MARY,BAKE[MARY,CAKE]] is constructed. Each text proposition activates propositions closely related to it in the general knowledge net, regardless of the discourse context. For instance, in the case of BAKE(MARY,CAKE]we might thus obtain LIKE(MARY,EATIMARY,CAKE||, PU I (MARY.CAKt ,IN-OVFN), RESULTIBAKF.IMARY,CAKE|,HOTK'AKF.JJ. PREPARf.[MARY,D1NNERj. These propositions are all closely associated with baking a cake (Figure 2). Note, however, that elaborating the text base in this way is not just a question of retrieving associated propositions from the knowledge net. The arguments of these retrieved propositions must be treated as variables that are to be bound to the values specified by the retrieval cue. Thus, because MARY is the agent of the text proposition, MARY is made the agent in the knowledge propositions it brings into the text representation, instead of PERSON in Figure 2. Similarly, although the informality of the present notation hides this, CAKE now is the particular one MARY bakes, not the generic one in Figure 2. These knowledge propositions function as potential inferences. Out of context there is no way of determining which of them are relevant: Maybe Mary really likes to eat cake, but perhaps she is in the process of cooking dinner, in which case PREPARE (MARY,PINNER] might become a macroproposition (what van Dijk, 1980, calls a *construction).* But it is also possible that next she will burn her fingers when she takes the cake out of the oven, making nor. which plays no role at all in the other contexts, the relevant inference. At this point, the construction process lacks guidance and intelligence; it simply produces potential inferences, in the hope that some of them might turn out to be useful.

3. In the third example, if the proposition SENDILAWYER.DEFENDANT,PRISON) has been formed, the knowledge net contributes nothing, because one presumably does not know anything about lawyers sending defendants to prison. (Of course, LAWYER, DEFENDANT, and PRISON would each be associatively elaborated separately.) If, however, JUDGE rather than LAWYER were the agent of SEND, the elaboration process would contribute the information that this implies that the judge is sentencing the defendant and so forth.

Sterj» C in the construction process, the generation of additional inferences, is necessary because not all inferences that are required for comprehension will, in general, be obtained by

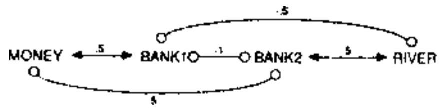


Figure ,i Connections between BANK.ι and BANK2 and their associates.

the random elaboration mechanism described earlier. In some cases more focused problem-solving activity is necessary to generate the desired inferences. Exactly how this is to be done is, however, beyond the scope of this article. I merely wish to point out here that in addition to the undirected elaboration which results from Step B of the construction process, there is still a need for controlled, specific inferences. Two types of such inferences are of particular importance in comprehension. Bridging inferences (Haviland & Clark, 1974; Kinlsch, 1974) are necessary whenever the text base being constructed is incoherent (i.e., whenever either the original text base itself or the elaborated text base remains incoherent by the criteria discussed in van Dijk and Kintsch, 1983, chapter 5). Second, rnacropropositions have to be inferred (as discussed in general terms in chapter 6 of van Dijk &. Kintsch, 1983, and operationalized as a production system by Turner, McCutchen, & Kintsch, 1986) Macropropositions are also elaborated associatively, as described in Step B for micropropositions.

What has been constructed so far is a set of propositions contaming the (micro)propositions directly derived from the text, a randomly selected set of associates for each of these, the macioproposltions generated from the texi, and their associates. The final Stepι p of the construction process involves the sj>№ncatum oFTtie interconnections bt-tween all of these elements. There are two ways in which elements are interconnected. U) The propositions directly derived from the text (hence referred to as "text propositions") are positively interconnected with strength values proportional to their proximity in the text base. Specific realizations of this principle are described in the discussion of Figure 4. (b) If propositions / and *j* are connected in the general knowledge net with the strength value *s(ij),* $-1 <$ $s(i,j) <$ I, and if *ι* and ; become members of a text base, the strength of their connection in the text base is *s(ij).* In other words, propositions in the text base inherit their interconnections from the general knowledge net. Strength values are additive, up to a maximum of 1, in those cases in which an inherited strength value combines with a text-base-determined connection.

Consider, for instance, the portion of a network that is generated when the word *bank* activates both BANKI and BANK2, as well as the associations MONEY and RIVER. A possible pattern of connections is shown in Figure 3, where for simplicity, connection strengths have been limited to ±.5 or 1. Alternatively, the graph shown in Figure 3 can be expressed in matrix form as shown in Table !. BANKI is associated with MONEY, BANK2 with RIVER, but inhibitory connections exist between MONEY and BANK2 and between RIVER and BANKI.

An example of text propositions that are interconnected via their positions in the text base is shown in Figure 4. LUCY is connected most strongly to WEED{LUCY,GARDENJ, and least

Table 1
Connectivity Matrixfor the Graph Shown in Figure 3

Proposition	1	2	3	4
1. MONFY	—	0.5	-0.5	0.0
2. BANKI	0.5		- 1.0	-0.5
3. BANK2	0.5	1.0		0.5
4. RIVER	0.0	-0.5	0,5	—

Table 2
Connectivity Matrixfor the Graph Shown in Figure 4

Proposition	1	2	3	4
1. LUCY	—	0.9	0.7	0.4
2. WF.KD	0.9		0.9	0.7
3. GARDEN	0.7	0,9		0.9
4. VEGETABIE	0.4	0.7	0.9	—

strongly to VF,.GFTABLF[GARDFNJ Although there are many possible ways to assign numerical connection strengths to express this pattern ofconnectivity, the one chosen here results in the matrix shown in Table *2.*

Inferences inherit positive and negative interconnections from the general knowledge net, as seen in Figure 5. The result ofthe construction process is, therefore, a network expressable as a connectivity matrix, consisting ofall the lexical nodes accessed, all the propositions that have been formed, plus all the inferences and elaborations that were made at both the local and global level and their interconnections.

Integration

The network that has been constructed so far is not yet a suitable text representation It was carelessly constructed and is therefore incoherent and inconsistent. At all levels ofthe representation, components associated with the text elements were included without regard to the discourse context, and many of them are inappropriate. An integration process in the connectionist manner can be used to exclude these unwanted elements from the text representation (e.g., see Rumelhart & McClelland, 1986. and Walt? & Pollack, 1985, for discourse).

Text comprehension is assumed to be organized in cycles, roughly corresponding to short sentences or phrases (for further detail, see Kintsch & van Dijk, 1978; Miller Λ Kintsch. 1980). In each cycle a new net is constructed, including whatever is carried over in the short-term buffer from the previous cycle.[6] Once the net is constructed, the integration process takes over: Activation is spread around until the system stabilizes. More specific, an activation vector representing the initial activation values of all nodes in the net is postmultiplied repeatedly with the connectivity matrix. After each multiplication the activation values are renormalized: Negative values are set to zero, and each of the positive activation values is divided by the sum ofall activation values, so that the total activation on each cycles remains at a value of one (e.g., Rumelhart & McClelland, 1986). Usually, the system finds a stable state fairly rapidly; if

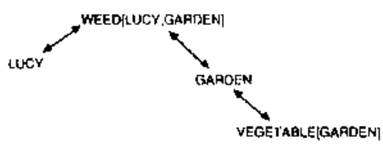


Figure 4. The text base for *Lucy weeded the vegetable garden.*

the integration process fails, however, new constructions are added to the net, and integration is attempted again. Thus, there is a basic, automatic construction-plus-integration process that normally is sufficient for comprehension. This process is more like perception than problem solving, but when it fails, rather extensive problem-solving activity might be required to bring it back on track. These processes will not be considered further here.

The result of the integration process is a new activation vector, indicating high activation values for some ofthe nodes in the net and low or zero values for many others. The highly activated nodes constitute the discourse representation formed on each processing cycle. In principle, it includes information at many levels: lexical nodes, text propositions, knowledge-based elaborations (i.e., various types of inferences), as well as macropropositions

A few simple examples will illustrate what is at issue here. Consider *Lucy persuaded Mary ío hake a cake,* which was discussed earlier. The PERSUADF proposition will pull in related knowledge items, just as was shown for BAKE However, out of context the integration process will not yield any striking results. In the context of *Lucy made tomato soup and .sauleed .some porkchops with herhs. She set the tahle and persuaded Mary lo bake a cake,* the integration process has very different results: PRF,PARF,[LUCY.DINNI;R) emerges as the dominant proposition (macroproposition) because most of the other propositions in the text base contribute to its activation value. That the cake was hot, or that she put it into the oven, disappears from the representation with activation values around zero.

Next, consider the example just discussed, where a perfectly good propositional strategy led to a wrong result. For *The linguists knew the solution of the problem would not be easy,* the text base that was constructed is shown in Figure 6. It corresponds to the connectivity matrix exhibited in Table 3 if connection strengths are assigned as in Table 2. (KNOW[SO1 UTIONJ and NOT[EASY] are connected positively via KNOW[$] but negatively via FASY, which adds up to 0.) The activation vector (.25,

[6] That integration occurs at the end ofeach processing cycle is proposed here merely asa simplifying assumption. Although (here isdcarly something going on at the end of sentences (e.g., Aaronson & Scarborough, 1977), integration does not need to wait for a sentence boundary (see the evidence for the "immediacy assumption"; Just A Carpenter, 1980. Sanford & Garrod, 1981). It would be quite possible to apply the relaxation procedure outlined here repeatedly in each cycle, as propositions are being constructed. This would allow for the disambiguation of word senses before the end of a cycle. Because inferences and macropropositions are usually not available before the end of a processing cycle, end-of-cycle integration plays an especially important role.

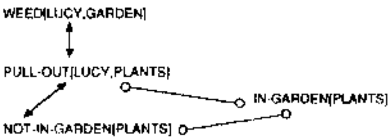


Figure 5. **Inferences generated from WEED[LUCY, GARDEN] and their interconnections.**

Table 3
Connectivity Matrix for the Graph Shown in Figure 6

Proposition	1	2	3	4
1. KNOWjS)		0.9	0.7	0.9
2. KNOWISOI.)	09	—	-1.0	0.0
3. EASY	0.7	-1.0	—	0.9
4. NOT	0.9	0.0	0.91	—

.25, .25, .25) corresponding to the assumption that all text propositions are equally activated initially is repeatedly multiplied with this matrix, renormalizing the obtained activation values after each multiplication as described earlier. To decide when the activation vector has stabilized, the following criterion was established: A stable slate is reached when the average change in the activation values after a multiplication is less than .001. Although this is an arbitrary criterion, even large changes (by one order of magnitude in either direction) make only minor differences in the final activation values obtained in this and many other cases. In the present case, this criterion is reached after 10 operations, yielding the final activation vector (.325, .(XX), .325, .350)—that is, the wrong KNOW[I.INGUISTS,SOLUTIONJ. which does not fit into the text base, has been deactivated. The integration process similarly resolves the problem of multiple pronoun referents. For *The lawyer discussed the case with the judge. He said "I shall send the defendant to prison,"* propositions were constructed for both lawyer and judge as referents of *he.* However, the process of associative elaboration generated some additional information for SEND[JUDGE,DEFFENDANT,PRISON], but not for SEND[L AWYER,OF,FENDANT,PRISON], The resulting text base is shown in Figure 7. To obtain the corresponding connectivity matrix (see Table 4), connection strengths among text base propositions were assigned as in Table 2, and among associates as in Table 3 (other assignments result in different numerical values for the final activation vector, but its pattern remains the same as long as the essential features of the matrix are preserved—for example, which connections are positive, negative, and zero). Assume an initial activation vector of (.25, .25, .25, .25, .25,0,0), reflecting the fact that only the text propositions themselves are activated initially. After 19 multiplications with the connectivity matrix, the two propositions in which *he* had been identified as the *lawyer* have activation values of 0, whereas the corresponding *judge* propositions have activation values of .261 and .283, respectively. Just a little knowledge was enough to choose the correct referent.

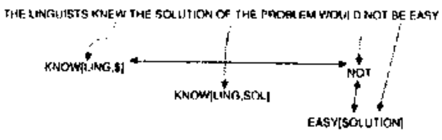


Figure 6 The strategic construction of a text base: soi DTION-OF-THF-PROBI.EM is first assigned to KNOW, then to EASY. (The dollar sign is a placeholder.)

After this general description of the construction-plus-activation model, two specific applications will be discussed in more detail: how words are identified in a discourse context, and how a propositional text base and situation model are constructed when comprehension depends heavily on activating a rich knowledge set. For that purpose, arithmetic word problems were chosen as the example, because the knowledge that needs to be activated is particularly well defined in that domain, and unambiguous criteria of understanding exist—a solution is either right or wrong. The purpose of these examples is twofold; to show how the general framework proposed can be elaborated into specific models in these experimental situations, and to compare the performance of these models with empirical observations and experimental results as a first test of the psychological adequacy of these models.

Word Identification in Discourse

The first problem to be considered in detail is how knowledge is used in understanding the meaning of words in a discourse. The previously sketched model implies that word meanings have to be created anew in each context, that this »s initially strictly a bottom-up process with context having its effects in the integration phase, and that this construction-plus-integration process takes time, with different factors influencing successive phases of the process.

Context effects in word recognition are ubiquitous in the experimental literature, and the explanation of these context effects has been a primary goal of theories of word recognition, lypically, it is taken for granted in these theories that because

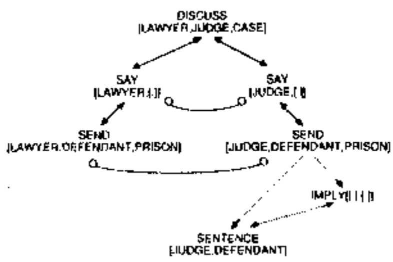


Figure 7. The strategic construction of a text base: The pronoun *he* is identified with two potential, mutually exclusive referents. (Instead of writing out whole propositions, the abbreviation [.] is used for the arguments of a proposition when they can be readily inferred.)

Table 4
Connectivity Matrix for the Graph Shown in Figure 7

Proposition	1	2	3	4	5	6	7
1 DISC	—	0,9	09	0.7	0.7	0.0	0.0
2. SAYfIAWYERI	0.9	—	10	0.9	0.0	0.0	0.0
3. SAY|Jt!OOF.)	09	1 0	⋯_	0.0	0.9	0.0	0.0
4 SIND|I.AWYF.R|	0.7	09	0.0	·_·	- 1.0	0.0	0.0
5. SENDf JUDGE)	0.7	0.0	0.9	1.0	—	0.5	0.5
6. IMPI Y	0.0	0.0	0.0	0.0	0.5	—	0.5
7, SFNT	0.0	0.0	0.0	0.0	0.5	0.5	—

context influences word recognition, contextual factors interact with the perceptual processes. Context effects are said to be top-down and expectation driven and are said to facilitate (or sometimes interfere with) the perceptual analysis. Similar ideas were once current in related fields, such as the "New Look" in perception (Bruner & Postman, 1949) and the filter theory ofattention (Broadbent, 1958). People perceive what they expect or want, attention filters out the irrelevant. Some words are recognized because the context favors them; others are not because the context inhibits them How these top-down effects of context are realized differs in detail among theories, but all the most influential current theories postulate interactive processes in which contextually expected words are favored In the logogen model (Morton, 1969, 1979), context primes semantic features that enter into the feature counter of a logogen and therefore strengthen that logogen. In Forster's search model (Forster, 1976), perceptual analysis defines a candidate set that is then searched by semantic relations or by word frequency. In Becker's verification model (Becker, 1976), both a set ofsensory candidates and a set of semantic candidates are created, with the latter being verified first. In the cohort model of Marslen-Wilson and Welsh (1978), context is used to exclude members of the cohort from the very beginning. Norris (1986) has recently reviewed these models and pointed out that they all involve some sort ofpriming mechanism through which context effects are mediated.

The model of how knowledge is used in discourse suggests a fundamentally different approach. Following earlier work by Kintsch and Mross(1985) and Norris (1986), the present model is neither interactive, nor does it involve priming. As these authors have argued, word identification is not simply a matter of lexical access. Rather, it is a complex process that responds to different influences at different stages. These stages, however, are merely convenient verbal labels. In fact, processing is continuous, and there is significant temporal overlap between the different subprocesses defining these stages. In the first stage (which was termed *seme activation* by Kintsch & Mross, 1985), the number of word candidates consistent with the perceptual input is progressively reduced through perceptual feature analysis. As in Forster or Becker, a set of sensory candidates is created through perceptual analysis, but its size decreases as the analysis progresses. This process rapidly reduces the number of word candidates to some manageable number, but not necessarily to one. At this point (probably at about 50 ms, see Fischler & Goodman, 1978), the semantic context comes into play. Some small number of lexical nodes has now been selected, each one activating a few of its strongest semantic or associative neighbors in the knowledge network. If there is a node whose associates fit into whatever context is present, it will be taken as the meaning of the to-be-identified word. What fits is determined by the integration process sketched earlier. This is the *sense-selection* stage of Kintsch and Mross.

Note that if the perceptual analysis had been allowed to continue for a sufficient period of time, for most words it would have yielded a result eventually by itself, and probably the same one. It is just that the association check helped to shortcut this process. With homonyms, however, the association check plays a crucial role: Perceptual analysis alone cannot decide which meaning *of bank* to select in any given context.

Sense selection by means of an association check is the very first of a possibly very long series of contextual plausibility checks (Norris's term). It comes first because the associative/semantic context of a lexical node can be computed rapidly. *i* As more information about the context becomes available, the sentence and discourse meaning begin to emerge, and more and deeper plausibility checks can be performed as long as there still • is time. This is the *sense-elaboration* phase, in which the meaning of a word is contextually explored and elaborated. However, once a response has been made in a recognition experiment, or once the process moves on in a discourse, elaboration is terminated. Thus, word meanings are usually identified long before complex inferences are made in comprehending a discourse.

At this point, a "meaning" has been constructed for the word in this particular context. It consists of the lexical node that has been activated (the contextually inappropriate nodes that had been activated have by now been deactivated through the various context checks), the associative and semantic neighbors of that node, the sentence and discourse context in which the word participated, and some inferences and elaborations that were produced in the course of the various plausibility checks that explored the role of that word in the given context.

What do we need to make such a model of word identification work? We shall disregard the perceptual analysis and take for granted that a certain number of appropriate lexical nodes has been activated (e.g., multiple semantic nodes for a homonym). We then need to compute the sentences and phrases in which the word in question participates, or more accurately, the propositions in which the corresponding concept token (for which the lexical node serves as the type) plays a role. Finally, we need to construct inferences and elaborations when necessary.

A model of word recognition that thus far is identical with the one favored here has recently been developed by Norris (1986). Norris called it the "checking model" and compares and contrasts it with the other extant models of word recognition in the literature. In Norris's model, the plausibility of word candidates in any given context is evaluated. The recognition criterion for contextually plausible words is lowered and that for implausible words is increased. By manipulating criterion bias in this way, Norris accounted for a wide range of observations from threshold and other types of recognition experiments.

Instead of equating plausibility with criterion bias, a different mechanism—integration—is used here. This mechanism has the great advantage of being applicable not only at the word-recognition level (which is what Norris was concerned with).

but it is equally suited to modeling knowledge integration at higher levels.

When a word is perceived, one or more lexical nodes are accessed, and some of their neighboring nodes that are closely related associatively or semantic-ally are also activated. Similarly, when a proposition is constructed, a number of associatively and semantically related propositions are also constructed, Both related concepts and related propositions serve to determine the plausibility of the core words and propositions. A richly interconnected structure is thus formed, through which activation can spread, so that positively interconnected items strengthen each other, while unrelated items drop out and inconsistent items become inhibited. Or, said differently, implausible items will be suppressed, whereas plausible ones support each other—at the level of word recognition as well as of textual integration.

Time Course of Activation of Words in a Discourse Context

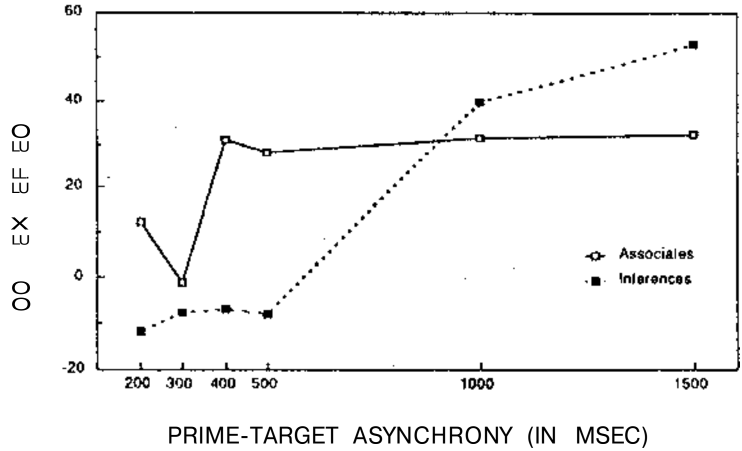
The model of word recognition just outlined is consistent with a great deal of experimental data. Norris (1986) has reviewed the word recognition literature in great detail and shown that his checking model accounts for the rich empirical findings in that area better than any of its competitors. The construction-integration model is closely related to Norris's model. On the one hand, it is more specific in that it proposes computational procedures by means of which Norris's "plausibility check" could actually be achieved, whereas on the other hand it replaces Norris's shift in criterion bias with the computationally more feasible integration mechanism. It appears likely that the present model can handle all the data the checking model accounts for, in just the same way and for just the same reasons as the checking model. There is, however, another part of the literature on word recognition that is not discussed in Norris (1986): the work on word identification in discourse. The empirical findings in this area are also in good agreement with the construction-integration model.

In a lexical decision task, the subject sees a string of letters and must decide as quickly as possible whether it forms an English word. If a target word is preceded by a closely related word, the response to the target word is speeded up (on the order of 20 to 40 ms) in comparison with unrelated control words. This priming effect has been well documented for some time and is obtained in list contexts (e.g., Meyer & Schvaneveldt, 1971) as well as in discourse contexts (e.g., Swinney, 1979). However, the discourse context is actually irrelevant to the priming effect. What matters is merely the associative relation between the prime word and the target word. As has been shown repeatedly (Kintsch & Mross, 1985; Swinney, 1979; Till, Mross, & Kintsch, in press; also equivalent results obtained with a naming task by Seidenberg, Tanenhaus, Leiman, & Bienkowsky, 1982), homonyms will prime strong associates of both their meanings, irrespective of the discourse context and in spite of the fact that the context-inappropriate meaning of the homonym never enters consciousness. Furthermore, context appropriate inferences that are not associatively related to a priming word are not responded to any faster than unrelated control words. However, all of this depends on the amount of time allowed for the processing of the priming word. If the target word closely follows the priming word, so that the processing of the prime is still in its initial stages, the results are as already described. However, if there is enough time for mote complete processing of the priming word in its discourse context, quite different results are observed. In this case, context-appropriate associates are still primed, but inappropriate associates no longer are, whereas context-appropriate inferences now become strongly primed. This time course of knowledge activation can be described in more detail by some illustrative experimental results.

In the study by Till et al. (in press), subjects read sentences like *The townspeople were amazed to find that all the buildings had collapsed except the mint* After the priming word *mint* they were given a lexical decision task, with the target word being either *money, candy,* or *earthquake.* That is, the target was a context-appropriate associate of the prime *(money),* a context-inappropriate associate *(candy),* or a topical inference word *(earthquake),* respectively. In addition, the interval between the presentation of the prime and the target word (stimulus-onset asynchrony, or SOA) was varied from 200 ms to 1500 ms. In the first case, the prime could only be incompletely processed; with an SOA of 500 ins, a somewhat deeper processing of the prime was possible before a response had to be given to the target word; and with 1,000 ms, extensive processing of both the prime word and its discourse context was possible. The data are shown in Figure 8. To keep this presentation simple, Figure 8 shows the average priming effects observed in the three experiments of Till et al. for SOAs of 200, 300, 400, 500, 1,000, and 5,500 ms. The value shown for associates at 200 ms, for instance, is the difference between the mean reaction time for context-inappropriate and context-appropriate associates a* that prime target asynchrony. !t is the average of two such values obtained in two different experiments—showing the data separately for each experiment merely complicates the picture without changing its essential features. The value for inferences, similarly, is based on the difference between topic words and unrelated control words. The purpose of Figure 8 is, therefore, merely to give an impression of the over-all shape of the results of this study (for more detailed analyses, the original article must be consulted).

Targets that are contextually appropriate associates of the priming word are primed at all four SOAs. Contextually inappropriate targets, however, are primed only when the priming word is still in its initial processing stages; by 400 ms inappropriate associates are no longer activated. Topical inferences are primed only if there is ample time, more than 500 ms, for the processing of the prime and its discourse context. This observation implies that the topic was not inferred immediately as soon as the relevant information became available, but was left for the sentence wrap-up period. Till et al.'s sentences were written in such a way that the topic could have been inferred before the last word in the sentence. This, however, is not what happened: Topics were inferred only after the whole sentence was read, requiring more than 500 ms of processing time. Thus, the full contextual meaning of the prime required about 1 s to emerge.

Data like these suggest that the initial activation of lexical knowledge is independent of the discourse context. What matters is only the (relatively fixed and stable) associative/semantic



figureS. Context effects as indexed by the reaction time difference to context inappropriate and appropriate associates or inferences as a function of processing time, after Till, Mross, and Kintsch (1988).

context of each word by itself. This stage of sense activation, however, is quickly followed by a process of sense selection in which the discourse context becomes effective: By 500 ms, context-inappropriate associates are deactivated (see also Seidenberg et al., 1982, and Swinney, 1979). Ifgiven more time, context effects grow even stronger: By 1,000 ms, contextually appropriate inference words are strongly and reliably primed even in the absence ofassociative connections (similarly for recognition, see McKoon & Ratcliff, 1986)

| Clearly, this pattern ofresults is in excellent agreement quali-/ tatively with the model ofknowledge use in discourse presented f earlier.' Right after a word is perceived, it activates its whole associative neighborhood in a context-independent way, with the consequence that strong associates of a word are likely to be represented in working memory and hence will be primed in a lexical decision task, whether they are context appropriate or not The knowledge-integration process then results in the deactivation of material that does not fit into the overall discourse context (such as context-inappropriate associates). Note that in order to disambiguate words on-line, the integration phase cannot be delayed until the end of a processing cycle; word senses are disambiguated before that. In the model, therefore, as soon as a text proposition is constructed and its associates have been generated, they will be integrated into whatever context exists at that time in working memory. Thus, each processing cycle involves many integrations, and the single integration operation performed at the end of each cycle in many of the examples discussed here is merely a simplification, adopted whenever one is not concerned with the on-line generation of word meanings. Finally, contextual inferences should require the most time to become activated on the average because although they sometimes result from the initial knowledge sampling, in other cases repeated sampling or, further, strategic elaboration might be required.

Earlier, an example was given of one of the texts used in the Till et al. (in press) study. The predictions of the model will be illustrated by means of this example. The aforementioned text (*The townspeople were amazed to find that all the buildings had collapsed except the mini)* has the following propositional representation:

1. TOWNSPEOPLE
2. AMAZED(TOWNSPEOPIE,P3]
3. COLLAPSF(P4)
4. ALL-BUT(BUILDING,MINT]
5. BUILDING
6. MINT

Connection strengths of .9, .7, .4, and 0 were assigned to text propositions one, two, three, or more steps apart in the text base (e.g., Pi is two steps away from P3, connected via P2). Next. each text proposition was allowed to access at random two of its neighbors in the long-term associative net. This process was simulated by having an informant provide free associations to phrases based on each of these six propositions. For instance, the phrase *all buildlings but the mini* elicited the associations *many buildings* and *mint is a building.* Of course, MONKY and CANDY were chosen as the associates of MINT Each text proposition was connected by a value of .5 to its associates, yielding an 18 X 18 connectivity matrix. Activation was then allowed to spread from the text propositions to the knowledge elaborations. Specifically, an initial activation vector with 1/6's corresponding to the text propositions and zeros otherwise was multiplied with the connectivity matrix until the pattern of activation stabilized. As a result, text propositions achieved activation values between .0987 and .1612, depending on how closely they were tied into the text base, and the knowledge elaborations had much lower activation values, between .0142 and .0239. with both MONEY and CANDY having a value of .0186. Thus, at this stage of processing, MONEY and CANDY are equally activated.

Activation continues to spread, however, and differences begin to emerge among the activation values for the various knowledge elaborations that have been added to the text base. The reason for this is that the knowledge elaborations are connected not only to the text propositions that had pulled them into the net but also to other text propositions as well as to each other. To approximate these interrelations, a connection value of .5 was assigned to any two propositions sharing a common argument. Because the homophone *mini* contributed associations to the subnet that refers to both of its senses, an inhibiting connection of-.5 was assigned to MINT/CANDY and BUH..DINO. whereas CANDY and MONEY themselves were connected by a -1. Continued multiplication of the activation vector with this connectivity matrix yielded a stable pattern (average change < .001) after 11 operations. At this point text propositions had activation values ranging between .1091 and .0584. Several of the knowledge elaborations reached values in this range, for example, .0742 for both ISA(MINT.BUILDING) and MONEY and .0708 for KILLJBUILDING.TOWNSPEOPLE}, whereas others had faded away by this time; for example, MAN. which entered the subnet as an associate of TOWNSPEOPLE, had an activation value of .0070 and, most significantly, .0000 for CANDY. This stage of processing corresponds to the 400- and 500-ms points in Figure 8: MINI is now clearly embedded in its context as a kind of building, and the inappropriate association CANDY is no longer activated.

The next processing stage involves the construction of a topical inference—what is the sentence about? While the exact operations involved in the construction of such inferences are beyond the scope of this article, van Dijk and Kintsch (1983, chapter 6) have discussed some of the mechamsms involved, such as a strategy of looking for causal explanations, which is what actual subjects appear to use predominantly in the following case. If given enough time, the modal response of human readers is that the sentence is about an earthquake that destroyed a town. Thus, the (empirically determined) propositions EARTHQUAKE and CAUSE[EARTHQUAKE,P3i were added to the text base and connected with the text-base propositions from which they were derived by a value of .5. The two new propositions were given initial activation values of zero, and the integration process was resumed; that is, activation now spread from the previously stabilized subnet into the newly constructed part of the net. Nine more integration cycles were required before the expanded net stabilized. As one would expect, the two new inferences did not alter the pattern of activation much, but both of them became fairly strongly activated (thereby diminishing activation values in the already existing portion of the net). The topical inferences EARTHQUAKE and CAUSE{EARTHQUAKE.P3) ended up with activation values of .0463 and .0546, respectively, among the most strongly activated inferences in the net. At this point, the process appears to coincide with the time interval between 1,000 and 1,500 ms shown in Figure 8.

The construction-integration model thus accounts for the data in Figure 8 by means of an intricate interplay between construction and integration phases: the construction of the text base and the context-free, associative knowledge elaboration during the first 350 ms of processing; the establishment of a coherent text base, which appears to be complete by 400 ms; and finally, an inference phase, involving new construction and new integration and requiring more than 500 ms of processing under the conditions of the Till el al. study. The model does not account for the time values cited here, but ii describes a processing sequence in accordance with the empirically determined time sequence.

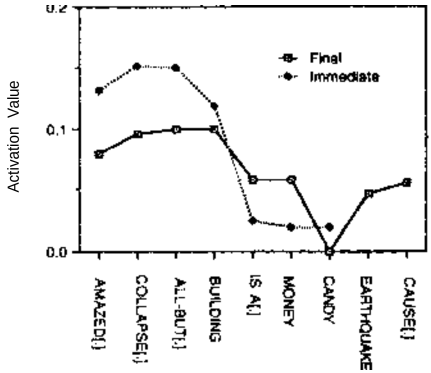


Figure 9 The changing meaning of MINT (The activation values of all propositions directly connected to MINT at the beginning and at the end of the process. The (.] notation is used as an abbreviation for the arguments of a proposition.)

In many models of word identification, the problem is thought to be "How do we get from a certain (acoustic or visual) stimulus pattern to the place in the mental lexicon where the meaning of this word is stored?" In the present model, word identification is much more deeply embedded into the process of discourse understanding. The lexical node itself provides just one entry point into the comprehender's long-term memory store of knowledge and experiences, and what eventually becomes activated from that store depends on the discourse context. In conceptions of the lexicon like that of Mudersbach's (1982), the meaning of a word is given by its "neighborhood" in the associative network into which it is embedded. Neighborhoods may be denned narrowly or broadly (nodes one link away vs. nodes several links away). In the present model, the meaning of a word is also given by its neighborhood—narrowly or broadly defined—not in the long-term memory net as a whole, but in the subnet that has been constructed as the mental representation of the discourse of which the word is a part. Because that representation changes as processing proceeds, word meanings change with it.

Figure 9 depicts the changing meaning of MINT in our example. MINT is directly linked to nine propositions in the network; indirectly it is linked to the whole net, of course. If one takes as its contextual meaning only its immediate neighbors, one finds at the beginning of processing mostly closely related proposilions from the text base plus three weakly activated knowledge elaborations that in part do not fit into the context at all ((ANDY). At the end of the process, however, the context-inappropriate association has dropped out, other inferences have been added, and the activation is more evenly distributed among text propositions and knowledge elaborations. Thus, textual information becomes part of the contextual meaning of a word, in contrast to most traditional conceptions of "meaning."

This example is, of course, no more than an illustration. Parameters in our calculations could be changed. For example, more than just two associates could *be* sampled initially in the process of knowledge elaboration In this case the neighborhood of MINT would contain many more knowledge elaborations than are shown in Figure 9, where there is a strong predominance of text propositions. Not enough is known at present to set some of these parameters with confidence. But Figure 9 does reflect certain aspects of the data correctly: the equal initial activation of MONFY and CANDY, the later emergence of the topical inference EARTHQUAKE. Although much more research is needed to produce a more adequate picture of how the contextual meaning of words is constructed during discourse comprehension, here is a technique that at least may help us to do so.

Arithmetic Word Problems

How children understand and solve simple word arithmetic problems provides an excellent domain to try out the construction-plus-integration model. Unlike with many other types of discourse, there are clear-cut criteria for when a problem is solved correctly, and the formal knowledge of arithmetic that is necessary for its solution is easily denned. However, word problems, like all other texts, share the ambiguity and fu//iness of all natural language. Not only formal, arithmetic knowledge is involved in understanding these problems, but all kinds of linguistic and situational knowledge. What makes word problems hard—and interesting—are often not their formal properties, but the way a problem is expressed linguistically and the way formal arithmetic relations map into the situations being described. Thus, word problems are ideal from the standpoint of knowledge integration because it is precisely the integration of formal arithmetic knowledge and linguistic and situationai understanding that is at issue here.

Another reason for choosing the domain of word problems is that there already exist alternative formal models of how children solve simple word arithmetic problems (Briars & Larkin, 1984; Kintsch & Greeno, 1985). Specihcally, the work of Kintsch and Greeno will be taken as a starting point here. Their model represents a union of the work on problem solving in arithmetic by Riley, Greeno, and Heller (1983) on the one hand, and that on discourse understanding by van Dijk and Kintsch (1983) on the other. Kintsch and Greeno (1985) added to the discourse-comprehension strategies of the van Dijk and Kintsch model some special purpose strategies for solving word arithmetic problems, which they named the *arithmetic strategies.* For instance, if the model encounters a quantity proposition, such as "six marbles," it forms a set and tries to fill in the various slots of the set schema: what the objects are, the cardinality of the set, a specification of the objects (e.g., that the marbles are owned by Fred), and the relation between the present set and other sets in the problem (the six marbles were given to Fred by Tom, which might identify them as a "transfer set"). Thus, the Kintsch and Greeno model for word problems builds a text base in quite the same way as in the van Dijk and Kintsch general theory of text comprehension, but it then forms a very specialized situation or problem model in terms of sets of objects and their interrelations. It solves a problem by recognizing a particular pattern of relations among sets (such as TRANSFER-IN or SUPERSET) and then using a stored-solution procedure appropriate to that case.[7] Thus, in terms of the foregoing discussion about knowledge use in discourse, the Kintsch and Greeno model is a "smart" model: Production rules are formulated in such a way that in each situation exactly the right arithmetic strategy is tired.

The Kintsch and Greeno model of solving arithmetic word problems is useful in several ways. The model identifies different classes of errors, such as errors caused by a lack of arithmetic knowledge, errors caused by linguistic misunderstandings, and errors that do not reflect a lack of knowledge at all but result from resource limitations. Certain formulations of word problems overload the resources of the comprehender, especially short-term memory, leading to a breakdown in processing. As Kintsch and Greeno have shown, within each arithmetic problem type there exists a strong correlation between the frequency of errors made in solving the problem and the memory load imposed by it, even though there are no differences within problem types in either the arithmetic or linguistic knowledge required for solution.

The model distinguishes between linguistic and arithmetic errors and helps us to investigate to what extent errors made by second- and third-grade pupils are caused by a failure to understand properly the text of the word problem, rather than by a faulty knowledge of arithmetic (e.g., Dellarosa, 1986; Dellarosa, Kintsch, Reusser, & Weimer, in press; Kintsch, 1987). If certain linguistic misunderstandings about the meanings of such key words as *have more than, have altogether,* or *some* are built into the knowledge base of the model, the model produces a pattern of wrong answers and misrecaJl of the problem statements that strikingly parallels some of the main types of errors that experimental subjects make. This is a good example of how much can be achieved even with the use of knowledge-poor representations in studies of discourse processing. The Kintsch and Greeno model knows about arithmetic (its arithmetic strategies), and it knows about the meaning of words (its lexicon; a *semantic net* in Dellarosa, 1986). However, it has no general world knowledge that would allow it to understand the situation described in a word problem. It merely picks out the crucial arithmetic information from the discourse and builds a propositional text base for it. This is good enough for some purposes (e.g., the investigation of resource limitations or linguistic factors in understanding as mentioned earlier, or to predict recall, summarization, or readability as in Kintsch & van Dijk, 1978, and related work), but it is not good enough for other purposes.

[7] Computer simulations of this model have been developed by Fletcher (1985) and Dellarosa (1986) and are available from the author.

The limits of this approach are illustrated by a well-known observation: If a word problem is embedded into a concrete, familiar situation or action context, it is much easier to solve than when the same problem is expressed abstractly (e.g., Dellarosa et al., in press; Hudson, 1983). Thus, *five birds saw three worms on the ground, and each hird tried to get a worm, flow many birds didn 'I get a worm*[9] is easy for first graders, but *There are five red marbles and three green marbles. How many more red marbles are there than green marbles?* is very hard, even though the two problems are equivalent in form.

The Kintsch and Greeno model does not account for this difference. What is needed is a model in which all knowledge relevant to the understanding of a word problem becomes integrated into a representation that is sensitive to arithmetic as well as to situational information. In the model to be described shortly, this is achieved by forming many different hypotheses about the arithmetic relations in the problem, instead of only a single one, and then by looking for information in the text in support of each hypothesis. Thus, situational and arithmetic information can combine in forming the problem interpretation.

Arithmetic Strategies

Arithmetic knowledge forms a special subset of a person's general knowledge network. Sets of objects can be represented by a prepositional schema with the slots object, specification, quantity, and role (i.e., their relation to other sets)—equivalent to the set schema of Kintsch and Greeno (1985). Superordinate schemata can be similarly defined. Thus, a TRANSFER-IN schema can be set up with slots for a START, TRANSFER-IN, and RESULT SET. With each such superordinate schema, various arithmetic procedures (such as the counting strategies of Kintsch & Greeno, 1985) can be associated.

Arithmetic knowledge is used in the same way as general world knowledge. That is, propositions that represent various hypotheses about the arithmetic structure of a word problem are constructed as the text of the word problem is read and become part of the subnet. Several possible arithmetic hypotheses are constructed at each point, and whatever evidence in the text favors one or the other of these hypotheses is connected positively to it.

The strategies required for solving arithmetic word problems have been described in Kintsch and Greeno (1985) and Kintsch (1984) and have been incorporated into the computer simulation of Dellarosa (1986). However, they are used here in a different way. In the aforementioned works, the intent was to circumscribe the conditions under which each strategy is applied so accurately that only the correct one is fired in each word problem. Here, strategies fire promiscuously whenever they are supported, however weakly, by the text, and it is left for the integration process to weed out what is not wanted, just as all sorts of general knowledge propositions are activated that later turn out to be useless. A problem is solved when the right superordinate schema is more strongly activated than its alternatives, which then triggers the desired arithmetic procedures.

Three forms of arithmetic strategies need to be considered. There are strategies that form arithmetic hypotheses about sets, strategies that determine the nature of the connections between various text propositions and these hypotheses, and strategies that form superordinate arithmetic schemata on the basis of which arithmetic calculations can be performed.

1. I lypotheses about sets are propositions of the form SET-[object:X,specification:Y,quantity:Z,role:W], where X refers to an object, such as TULIP, Y is one or more other text propositions, specifying X further—for example, PASHLOCATioNHULUMN-GARUENi]; Z is a quantity proposition wiih X as argument—for example, FOURiHENnuLiP); and W indicates the role of the set in some superschema, such as WHOLE or PART.

2. Whenever a quantity proposition is encountered in the text base, possible arithmetic hypotheses derivable from it are constructed (e.g., two otherwise identical propositions with the roles WHOLE and PART). Propositions in the text base that provide evidence for any of these alternatives are then connected positively to it. Key words can be used for this purpose, as in Kmlsch and Greeno (1985): Collection terms such as *altogether* indicate WHOLE sets; *give/take, oj these,* and *have more/less than* indicate PART sets. In addition, general world knowledge about situations and actions is used to determine what is a WHOLE and what are its PARTS. The strategies involved have been described in Kintsch (1984): restricted subsets, conjunction, and time-ordered possession/location.

Restricted subsets. If the specification of one set is more general than that of another, the former is assigned the role of WHOLE and the latter that of PART Examples are LARC3E-WINDOW, SMALL-WINDOW versus WINDOW, OΓ ON-UPPER-SHELF, ON-L.OWF.R-SHFI.F versus ON-SHFLF.

Conjunction. If the object or specification of one set consists of the conjunction of the objects or specification of two other sets, the former is the WHOLE and the others the PARTS. This conjunction may be explicit a? in YESTERDAY. TODAY, and **YESTERDAYATODAY, OΓ implicit, as in TEDDYBEAR, DOLL,** and TOY

Time-ordered possession/local ion. If the specification slots of three sets contain either HAVF[agent,object] or LOCATiON[object,place], or the negations of these propositions, as well as information to establish a temporal order, WHOLE and PART roles can be assigned to the three sets according to the resulting patterns. For instance, if the specifications of three sets are **I IMF 1 | HAVE , JOE , P (MARBLES j 1; TIME2 (GIVE , JOE , TOM . Q (MARB! ESI], which implies HME21NOTIHAVE,JOE,Q(MARBLESJ];** and TiME.i[HAVE(JOF:,ziMAR»LES]j, SETi is indicated as the WHOLE set.

3. the PART-WHOLE schema is the only arithmetic superschema to be considered in the examples that follow, though various TRANSFER and C:OMPARISON schemata could have been treated in the same way, as in Kintsch and Greeno (1985). Three hypotheses can be formed about the PART-WHOLE schema, depending on whether the first, second, or third of the sets formed is to be considered the WHOLE set. (Note that the order in which the sets were formed in the word problem, not their true temporal order, is at issue here.) Thus, a proposition with the head PPW, which is simply a mnemonic for PART-PART-WHOLE, expresses the hypothesis that the problem is a PART-WHOLE problem with the third set as the WHOLE: ppw{role[SETi,PART],rolelSET2.PART],role(SET3,WHOLE]]. Associated with this schema is the equation $Q\backslash$ -f $Q2 = Q3,$ where $Q\backslash$ is the quantity of the Mh set, as well as procedures to solve

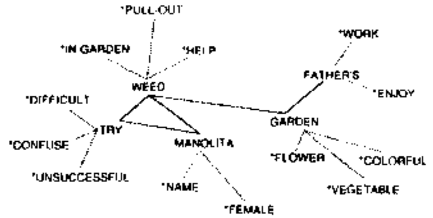


figure 10. The elaborated text base for the first sentence ofthe Manolita problem. (Fach proposition is indicated by a single word. Text propositions are connected by solid lines, their associates by broken lines. Associates are marked with an asterisk.)

this equation, depending on which ofthe quantities happens to he unknown.

Examples and Issues

Three examples will be analyzed here to show how the model understands, or fails to understand, as the case may be, arithmetic word problems. To see how these examples work, it is necessary to present at least the first one in sufficient detail. This problem is intended simply as an illustration ofthe basic mechanisms of the model—nothing much of interest happens with respect to the arithmetic, and textually. the only thing of significance is that a simple inference is formed, which, however, is crucial for the understanding ofthe problem. Two more examples, which will not be presented in as much detail, will serve as illustrations of how the model can account for some well-known facts about word-problem solving that alternative models (Briars & Larkin, 1984; Kintsch & Greeno, 1985) do not handle readily.

Inferences

> Manolita tried to weed her father's garden. "You sure weeded it," said Mr. Mundoza "There were fourteen tulips in the garden and now there are only six." How many tulips did she pull out by mistake?

This problem, modified from the "Thinking Stories" of Willoughby, Bereiter, Hilton, and Rubinstein (1981), requires for its solution the application of one of the LOCATION strategies: There were so many tulips in the garden, then some were pulled out, and now so many are left. A simple, knowledge-based inference becomes necessary: that the tulips that were pulled out are no longer in the garden. The knowledge-activation mechanism of the present model readilv supplies this inference, and the problem will be solved successfully.

The model processes this problem in three cycles, which includes the first sentence, the statement by Mr. Mundoza, and the question sentence. The first sentence simply sets up a context and is not directly relevant to the arithmetic. In Figure 10, the way the model understands this sentence is indicated, albeit in abbreviated form. The propositions constructed from the sentence itself are (PI) MANOLITA. (P2) GARDEN, (P3) TRY(MANOI ΓΓΛ.P4). (P4) **WEHDfMANOUTA.GARDEN],** (P5) FA-THF,R\s{GARDEN|. Only the first terms ofthese prepositions are shown in Figure 10. Also shown in Figure 10 are the propositions that were added to the text base through the process of associative-knowledge elaboration (they are marked with an asterisk and, once again, abbreviated: *NAME stands for ISAJMAN-OLITA.NAMEJ, etc.). Because no simulation of a general knowledge network is available, or even conceivable, the process of knowledge elaboration must be approximated empirically. An informal procedure was adopted for this purpose: Three persons were asked to provide free associations to phrases corresponding to PI through P5 (as well as to corresponding phrases from the remaining two sentences of this word problem), and the responses generated by at least two persons were considered as the top associates of each proposition in the general knowledge net (up to a maximum of three associations per proposition).

The text base shown in Figure 10 serves as a basis for deriving a connectivity matrix, using the principles illustrated earlier in Tables I and 2: Text propositions are connected depending on their proximity in the text base, each text proposition is connected to its associates by a value of .5, and knowledge derived propositions are interconnected by the same value if they share an argument, or by -.5 if different word senses are involved (this does not occur in the present example).

An initial activation vector consisting of .2 s for the five propositions directly derived from the text, followed by 13 zeros for the propositions generated from the knowledge net, was then repeatedly updated by multiplying it with the connectivity matrix until the activation values stabilized, as in the examples discussed previously. In the present case, activation levels stabilize after 10 iterations. The resulting pattern of activation is shown in the first panel of Figure 1 1 . WF.EPfMANOi ITA.GAR-DENJ, whose centrality in the text base is apparent in the graphical representation, has the highest activation value, and the other text-derived propositions also have fairly high activation values. Knowledge-derived propositions are considerably less activated. The four most strongly activated propositions (PI through P4) are retained in the short-term buffer and enter the second processing cycle.

The second processing cycle is shown in Figure 12. The four propositions held over in the short-term memory buffer from Cycle 1 are joined by 9 new text propositions and 11 associated propositions from the knowledge base. (Because of lack of space, the latter are indicated only by asterisks.) The quantity propositions FOURTEENITULIPJ and SIXJTIJLIP] generate four arithmetic hypotheses: that the 14 tulips that were in the garden in the past are, respectively, a PART or WHOLE set, and that the 6 tulips now in the garden are a PART or WHOLE set. What the reader knows about weeding gardens provides the crucial information that discriminates among these hypotheses: The tulips before the garden was weeded are the WHOLE set, and only a PART is left after the weeding. This knowledge is expressed in the connectivity matrix by connecting PAST with WHOLE[14], and PRESENT with PAR r[6].

The last three propositions that enter the subnet are the su~perordinate arithmetic hypotheses PPW, PWP, and WPP. They receive support from their corresponding first-order arithmetic

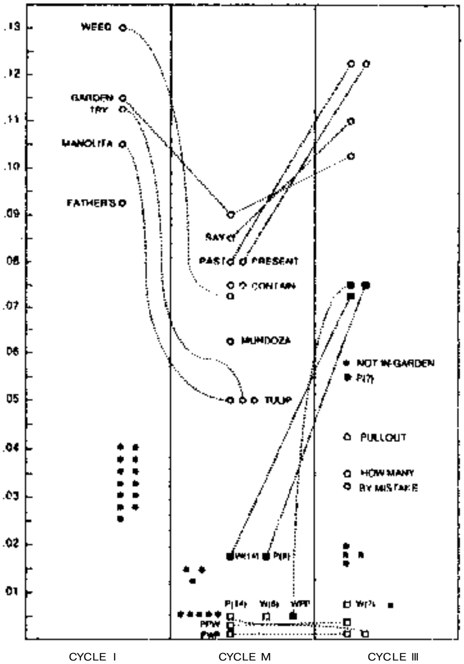


Figure 11 The result ofthe integration process for the three sentences in the Manolita problem. (Propositions are indicated by single words; inferences are marked by an asterisk; their arrangement in the figure is approximate. The ordinate shows the activation values ofeach proposition after the process has stabitoed. Propositions carried over from one processing cycle to the next are connected by arrows.)

hypotheses. Thus, whatever strength each arithmetic hypothesis gathers from the text is fed into the superordinate arithmetic schemata consistent with it. These schemata are mutually exclusive and inhibit each other with connection values of 1. Note that only at this final level is inhibition among arithmetic hypotheses used: The hypotheses that a particular set of objects plays the role of WHOLE or PART set are also mutually exclusive, but they are not allowed to inhibit each other, they merely collect more or less positive evidence, which they then transmit to the superordinate stage where a selection among alternatives is made.

The resulting connectivity matrix then becomes the multiplier ofthe activation-state vector for the 28 propositions participating in this second processing cycle. Initially, these activation values are positive for the text-derived propositions, and zero otherwise, except for the propositions carried over in the buffer, which retain the activation values they reached in the last cycle. In this case, the activation vector stabilizes already after seven operations. The results are shown in the second panel of Figure 11. (If the activation process is extended to twice the number of cycles, the activation values for the arithmetic hypotheses, measured to four decimal places, do not change at all.) All text-derived propositions remain strongly activated, while none of the textual inferences (e.g., MUNDOZA is a NAME of a MALE, TUMPS are FLOWERS. RF,D, and GROW-IN not LAND) reach a high level of activation. This is intuitively quite plausible. As far as the arithmetic is concerned, the problem is at this point understood correctly and practically solved: WHOLEI 14] is more strongly activated than its alternative, PARi[14]. Similarly, FAR τ[6] is stronger than WHOM{6]. The correct hypothesis, WPP, is the most strongly activated of the three alternative superschemata.

Note that the text propositions and inferences are, in general, much more strongly activated than the arithmetic hypotheses. Therefore, the activation values ofthe latter must be considered separately, relative to each other, rather than in relation to the text propositions when it comes to selecting propositions to be maintained in the short-term memory buffer. This imbalance is required for the model to work. If the arithmetic hypotheses are weighted more heavily, they draw the activation away from the text itself, and the system cannot stabilize: It will flip-flop between alternative, mutually contradictory arithmetic schemata. The arithmetic hypotheses have to be anchored in a stable text representation.

For the third and final sentence, the short-term memory buffer needs to carry over both text propositions to establish textual coherence and arithmetic hypotheses to take advantage ofthe, understanding ofthe problem that has been achieved so far. It has been assumed here that the four strongest text propositions as well as the four strongest arithmetic hypotheses are carried over in the buffer, as shown in Figure 13. (There are, of course, other plausible alternatives,) The three tex\ propositions

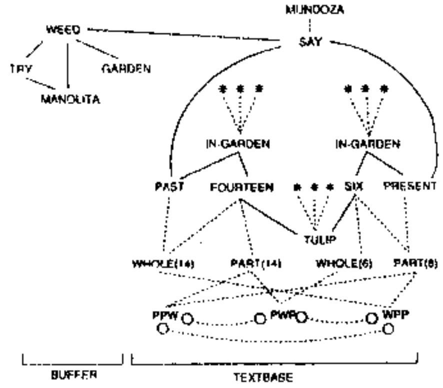


Figure 12 The elaborated text base for the second sentence ofthe Manolita problem. (Pour propositions were carried over from the previous cycle in the short-term memory buffer. Solid lines connect text propositions, broken lines inferences, nonarithmetic inferences are indicated by asterisks only.)

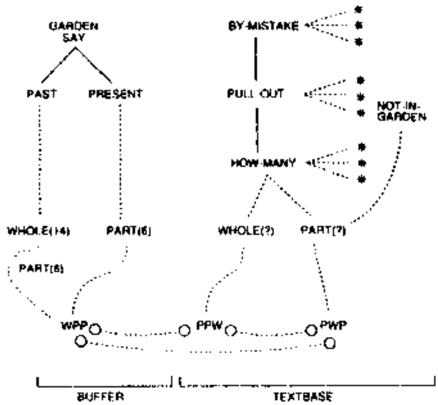


figure 13 The elaborated textbase for the third sentence of the Manolita problem.

generated on the basis of this sentence bring with them into the net six knowledge propositions, one of which is NOTJCON-TAiN(tiAROEN,TtJLipj), which turns out to be crucial for the solution of the problem. In addition, new hypotheses about the question set are formed, and the schemata PPW and PWP, which were lost after the second cycle, are reconstructed. Because the child knows about weeding gardens, the tulips that were pulled out are identified as a part of those that were in the garden in the beginning. Hence, a connection that favors the PART hypothesis over the WHOLE hypothesis is formed between the inference NOT|CONrAiN(GARDEN,TULiP|i and PARTI?}. It completes the pattern that is the condition for the use of a LOCATION strategy: some tulips at one place in the past, then some not there, now some are left.

The new net requires 43 operations to stabilize. The knowledge-based inference NOT|CONTAIN(GARDEN,TULIP]) achieves an activation level above the range of the text propositions (Figure 11, third panel). The picture is completely clear as far as the arithmetic is concerned: All the correct hypotheses are strongly activated, and all incorrect alternatives have low or zero activation values.

The final steps in the solution of the problem are procedural. From information associated with the WPP pattern the equation $14 = \beta + ?$ is generated, which is then used to obtain the correct answer. A lot of mountains had to be moved to achieve a very simple result!

The Manolita problem was solved without problem solving. The basic comprehension operations were sufficient; that is, it produced the inference that the pulled-out tulips are not in the garden, which was required for the application of the LOCATION strategy. However, this is not always the case. In many, not necessarily difficult, problems, more focused problem-solving operations are required because the random-inference generation process described earlier fails to generate the required inference. Consider the following "thinking problem":

> Mrs. Nosho was telling Mark about the two huge aquariums she kept when she was a little girl. "There were 30 fish in one and 40 fish in the other, so you can tell how many fish I had." How many fish did Mrs. Nosho have?

In a simulation run of this problem the model failed because it did not come up with the transitive inference HAVF[X,YI*CT>N-TAINIY./.J implies HAVFJX.ZJ. At this point, the process needs to go into a problem-solving mode in which the information in the text is elaborated in a more focused manner than is possible with the automatic-comprehension mechanisms discussed here.

Context fleets

Problems embedded into a familiar situational context are much easier to solve than problems that must be solved without this situational support (e.g., Hudson, 1983). Thus, birds catching worms present a concrete, understandable situation that makes it clear what is the whole and what are the parts, whereas abstract, ill-constrained problems do not. All depends on whether the right arithmetic strategy is used; the situation is of no help.

In the worm-and-bird problem, the text provides a situational constraint for the interpretation of the problem that has very little to do with arithmetic per se. It is the knowledge about birds eating worms that matters. Ihe birds trying to catch the worm are understood as the WHOLE set, with the birds catching worms as one PART, and the birds unable to get a worm as the other PART This understanding was achieved not because a certain key phrase, like *how many more,* was parsed correctly b*ut on the basis of general world knowledge. If there are birds, some of whom catch and some of whom do not catch a worm, what is the WHOLf: set and what are the PARTS is given by general world knowledge that is not specific to arithmetic. The arithmetic can hardly go wrong here because the well-known situation guarantees the right interpretation of the problem. It is this aspect that the present model deals with most effectively.

Context, however, does not always facilitate problem solution, it may also interfere with it. Consider this typical school problem, with its highly impoverished context:

> Fred has four Chevies and three Fords, (a) How many cars docs he have altogether? (b) How many more Chevies does he have than Fords?

Context is no help with this problem; it must be solved on the basis of specialized arithmetic strategies, on the basis of the key words *have altogether* for Question A and *have more than* for Question B. Of course, children are much more familiar with the former (e.g., Riley et al., 1983), but if the right strategies are available, both problems will be solved. In the model, too, the *altogether* in Question A will be connected with the now-MANY/WHOLE hypothesis, and the *have more than* will be connected with the HOW MANY/PARι hypothesis in Question B, and both questions will be answered equally well. After the first sentence, PART and WHOLE hypotheses are established for both the Chevies and the Fords, but there is not much to distinguish them; the superordinate schemata PPW, PWP, and WPP are only weakly activated and hardly differentiated. Question A, on the other hand, correctly activates the PPW hypothesis, and

Question B yields the WPP result. Thus, if the arithmetic knowledge is available, it makes very little difference which question follows the problem statement.

In contrast, if the problem is only slightly contextualized, the model can be biased in favor of one of the questions, and actually fails when it gets the wrong one. Suppose, the foregoing problem is changed to read

> Fred has a nice collection of antique cars. Four of his cars are Chevies, and three are Fords.

Collection, like some, is constructed as a quantity proposition, and hence PART and WHOLE hypotheses for a set of cars with unspecified quantity are established in the first processing cycle. They are both activated equally, however, at this point. This changes dramatically with the second sentence: The *four Chevies* and *three Fords* are both identified as PART sets because of the phrase *of his.* in consequence, the model begins to favor the WPP hypotheses. When it receives Question A, the WPP hypothesis is decisively strengthened, and the problem is solved correctly. On the other hand, if it is given Question B, the model becomes confused between the WPP and PWP hypotheses, which are both equally activated, and fails to solve the problem.

Thus, we have here an example where the problem context interferes with the solution of a problem. It biases the problem in favor of one particular interpretation, so that when another interpretation is required, the whole process fails. It is important, however, to analyze exactly why the model failed to answer Question B correctly: After processing the second sentence, it was so strongly convinced that the *four Chevies* and *three Fords* were both PART sets that it did not carry over the corresponding WHOLE set hypotheses and therefore had no way of using the information in the *have-more-than* question in support of the CHEVIES/WHOLE hypothesis. Thus, rather special circumstances prevented the model from answering Question B. In slightly different circumstances, it could have done so: (a) if the buffer were large enough, the CHEVY/WHOLE hypothesis would not have been lost, or (b) if the model had been allowed to reread the problem statement.

Question Specificity

The final example illustrates some different aspects of word-problem solving; namely the complex role that redundant specifications of sets may have. On the one hand, overspecifying a set can be helpful because it provides more than one way to refer to it. On the other hand, redundant specifications increase the length of the text and thus the likelihood that some important piece of information is no longer in active memory when it is required. In the following problem, three versions of the question are possible:

> Joe had a collection of nine marbles. He started his collection with some beautiful red marbles. Then Lucy added six pink marbles to his collection as a present, (a) How many beautiful red marbles did he start his collection with? (b) How many marbles did he start his collection with? (c) How many beautiful red marbles did he have?

The first processing cycle results in undifferentiated hypotheses about the nine marbles. The set constructed in the second cycle, on the other hand, is clearly a PART set, as is the one constructed in the third cycle. Indeed, at the end of the third cycle, the model understands the problem essentially correctly, with the WPP schema greatly exceeding alternative hypotheses in activation value. To understand what happens next, it is necessary to know which text propositions were maintained in the buffer at the end of the third cycle: Only propositions from the third sentence are carried over, while the propositions from the second sentence are no longer held in active memory at this point. This has nontrivial consequences when the question is asked. In Versions A and B everything is all right, because the question itself identifies the question set as a PART set—*starting a collection* serves this function, just as it did in Sentence 2. Version C of the question, on the other hand, does not yield a correct solution. The question itself does not indicate the role of the question set, and there is no information from the second sentence still available in active memory that would help to identify its role either; because there are already several strong PART hypotheses around, the model tends toward the hypothesis that the question set has the role of a WHOLE; the PWP schema thus becomes more activated than the correct WPP schema.

However, this is far from an unequivocal prediction of failure for Version C of the question. With a slightly larger buffer, or with a little less irrelevant material intervening *(pink marbles, as a present},* the critical information from the second sentence could have been maintained in the buffer and used to solve the problem. Or even more obviously, the problem solver could reread the problem or perform a reinstatement search (Kintsch A van Dijk, 1978; Miller A Kintsch, 1980) to activate the required information from long-term memory. Rather the prediction is that children, like the model, would have more trouble with Question C, and fail more frequently, than with either A or B.

Thus, the more specific the question the better. 3ut how irrelevant or redundant material will affect the difficulty of a word problem is a more complex story. It may be quite harmless, or may even facilitate problem solving, if the question exploits a redundancy in the specification of a set. But it may be a source of difficulty and even a cause of failure when the question is asked in an unhelpful way. The present model has the flexibility to handle these complex effects of context: Many small effects are allowed to add up and pull the model one way or another. The "smart" models of Kintsch and Greeno (1985) and Briars and l>arkin (1984) have no ready way to cope with these subtle contextual demands: Either the right strategy is used or not.

Discussion

How people recall relevant knowledge when they read a text is reminiscent of another experimental paradigm that has been studied extensively in psychological laboratories: how people recall lists of words. A widely used explanation for the recall of word lists is based on the generation-recognition principle. Some words are recalled directly, perhaps from a short-term memory buffer, and these words are then used to generate other semantic-ally or contexlually related, plausible recall candidates. Words that have actually appeared in the to-be-learned list will be recognized among these candidates and recalled, whereas intrusions will tend to be rejected. Generation-recognition theories have had their detractors, and in their most primitive form they are certainly inadequate to account for the more complex phenomena of list recall. However, sophisticated versions of this theory are widely accepted now. Almost every current model of list recall includes a generation/retrieval as well as a recognition/editing stage.

The model of knowledge use in discourse comprehension proposed here has two analogous stages: First, a prepositional network must he constructed, and then it must be edited or integrated. The way the construction process is thought of here is a straight extension of previous work on discourse processing. The words and phrases that make up a discourse are the raw material from which a mental representation of the meaning of that discourse is constructed. This mental representation takes the form of a propositional text base. Text bases combine two sources of information: the text itself and knowledge—knowledge about language as well as knowledge about the world. To construct even a single proposition, an appropriate frame must be retrieved from one's store of knowledge, and its slots must be filled in the way indicated by the text. The novel aspect of the present model is that the role of knowledge is greatly expanded in this process. Previously, one could think of the text base—to put it crudely—as a translation into "propositionaleese" of the sentences in the text Now, the text base becomes a much richer structure than before. Not only does it contain the propositions directly derivable from the text, but also each of these propositions brings with it a number of other propositions that are closely connected to it in the general knowledge net. Thus, propositions are constructed just as before (e.g., van Dijk A Kintsch, 1983); yet where previously a single proposition was formed, a whole cluster is generated now.

Crucial in the present model is how this cluster of propositions is obtained: by a context-free process of activation of the closest neighbors of the original text-derived proposition in the general knowledge net. Of course, such a process will inevitably activate a lot of material that is irrelevant *for* any given context and. indeed, inconsistent with it. However, the price that has to be paid for promiscuity is not very high: The resulting text base is a connectionist net in which further spreading activation processes rapidly take care of inconsistencies and irrelevancies. What is gained by this dumb and seemingly wasteful process of random knowledge activation is flexibility and context sensitivity. 'The research on knowledge actiyatipn in psychology, as well as the experience with artificial intelligence systems, suggests that it is very difficult to activate knowledge intelligently. Prediction or expectation-based systems that use frames or scripts do not adapt easily to new contexts; prestructured knowledge hardly ever is exactly in the form that is needed. The construction-integration scheme proposed here may turn out to be more successful in this respect.

The general framework sketched earlier could be extended and elaborated in various ways as more experience with it is obtained. It might prove necessary, for instance, to resort to greater formali/ation in the propositional notation used here. However, until it becomes quite clear what the gains of greater formalization would be, a robust, easy-to-use system is to be preferred, even at the cost of some imprecision.

Perhaps more important might be elaborations of the knowledge-sampling mechanism. As presented here, each text-derived proposition activates its own strongest associates. It might be worthwhile to explore schemes whereby pairs or clusters of propositions activate their strongest joint associates.

Similarly, other criteria for stabilizing a network might be explored. For instance, networks might be made to maximi/e some statistic like harmony, as in Smolensky (1986). This might have considerable advantages. For instance, it is not always possible now to compare different networks in terms of how fast they reach equilibrium, because the number of cycles required depends strongly on the number of nodes in the network. In addition, at present there is no really satisfactory way to tell how good an equilibrium a process achieves. In the word arithmetic problems, all one can tell is whether the right hypothesis is more strongly activated than its competitors, but comparisons of the size of that difference across problems are problematic.

Constructive processes other than the ones explored here will need to be considered. For word arithmetic problems, the most important constructions involved the arithmetic hypotheses. The construction of macropropositions could be neglected, mostly because the word problems were short ones and their macrostructure played no role in the problem-solving process.[8] For many other types of text, construction rules to form successive layers of abstractions and generalizations, as described by Turner et al. (1986), would be of primary interest. The macrostructure of a text could thus be made an integral part of a text base rather than a separate component, as it is presently treated.

Thus, there are a great many rules necessary to make the construction-integration model work for proposition building, assigning references and coreferences, bridging inferences, forming macrostructures. elaborating knowledge, and so on. Some of these construction rules are reasonably well worked out at this point, others are available within restricted domains, but many problems remain as yet unsolved. Thus, some of the same problems are encountered here as in conventional expectation-driven, top-down models of comprehension—but with one difference: Weaker, more general rules can be used here because these rules need not be fine-tuned to an ever-changing context. Whatever rules are still needed ought to be easier to work out within the construction-integration framework.

In van Dijk and Kintsch (1983), an imporant distinction was made between text bases and situation models. The former correspond to the propositional representation of a text, both at the level of the micro- and macrostructure. The latter correspond to a representation of the text that is integrated with other knowledge. Thus, in terms of the present model, the integrated text base—after irrelevant and inconsistent information has been deactivated and important knowledge elements have been absorbed—is a kind of situation model. The qualifying phrase "a kind of" is needed because text bases, integrated or not, are always propositional, whereas van Dijk and Kintsch specifically left open the possibility that situation models may *be* nonpropositional (e.g., Perrig & Kintsch, 1985). Situation models, under certain circumstances, may thus be like Johnson-Laird's (198.3) mental models.[9]

* Longer problems in which the macrostructure does play a role have been investigated by Dellarosa et a!. (1988). Depending on whether a word problem establishes a theme of competition or cooperation between two protagonists, *compare* or *combine* problems will he solved most easily.

The theory of knowledge use in discourse comprehension has been presented here at two levels: first, it is presented in terms of a general computational mechanism, at the level of what Pylyshyn (1985) called the "cognitive virtual machine"; and second, as a particular model that specifies how this mechanism is used in word identification in discourse and in understanding and solving word problems. The function of the model is primarily explanatory. Certain phenomena can now be interpreted within the framework of the model; for example, why a particular formulation of a word problem is especially hard or easy. Unlike less complex theories, however, there is no direct link between explanation and prediction in the present case. Unqualified experimental predictions are hard to come by in a model as complex as the present one. At best, one might predict that a particular problem should be a difficult one, but that might mean several different things at the empirical level: that the solution fails, that a particular error occurs, that extra memory resources are required, that a reinstatement search will occur, that the problem must be read twice, and so forth. Even if we knew precisely what the "knowledge-use virtual machine" was like, our ability to make precise experimental predictions that are testable in conventional ways would still be severely limited. That, however, is not to say that such theories are without empirical consequences. Although we cannot predict particular events, predictions concerning classes of events are quite feasible (e.g., the different ways people might have trouble with word problems). Furthermore, our new-found understanding of why and how certain things happen can have important consequences for how certain texts are created in the first place or for instructional practices designed to help people with particular comprehension tasks.

[9] Unlike the representation of the text itself—the text base, which is always prepositional—situation models may have a different representation format, although this possibility was not considered in the present article. Both text bases and situation models are mental models of one kind or another in the sense of Gentner and Stevens (1983), though not necessarily in the more restrictive sense of Johnson-Laird (1983).

References

Aaronson, D., A Scarborough, H. S (1977). Performance theories for sentence coding. Some quantitative models. *Journal of Verbal Learning and Verbal Behavior, 16.* 277-303.

Anderson, J, R. (1980). Concepts, propositions, and schemata: What are the cognitive units'* *Nebraska Symposium on Motivation, 28,* 121-162.

Anderson, J. R. (1983). *The architect lire of cognition.* Cambridge: Harvard University Press.

Becker, C. A. (1976). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Ps*cholvgy: Human Percept ion and Performance, 2,* 556-566.

Briars, D., A Larkin, J. H. (I984). An integrated model of skill in solving elementary word problems. *Cognition and Instruction, I.* 245-296.

Broadbent, D. A. (1958). *Perception and communication.* New York: Pergamon Press.

Bruner, J. S., & Postman, L. (1949). Perception, cognition, and behavior. *Journal of Personality. 18,* 15-31.

Dellarosa, O. (1986). A computer simulation of children's arithmetic word problem solving *Behavior Research Methods, Instruments. <& Computers. 18,* 147-154.

Dellarosa, D., Kintsch, W., Reusser. K., A Weimer. R. (in press). The role of understanding in solving word problems. *Cognitive Psychology.*

Oik, S. C. (1980). *Studies in functional grammar* London: Academic Press.

Fischler, I., A Goodman. G. O. (1978). latency of associative activation in memory. *Journal of Experimental Psychology: Human Perception ami Performance, 4,* 455-470.

Fletcher, C. R. (1985) Understanding and solving arithmetic word problems: Λ computer simulation. *Behavior Research Methods. Instruments, and Computers. 17.* 565-571.

Forster. K. I. (1976). Accessing the mental lexicon In R. J. Wales A E. *C.* T Walker (Eds.), Mm *approaches to language mechanisms* (pp. 257-287). Amsterdam: North-Holland.

Fra?,ier, I,., A Rayner, K. (1982). Making and correcting errors during sentence comprehension: F.ye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology. 14.* 178-210.

Fredcriksen.*).* R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal icferents. *Discourse Processes, 4,* 323-347.

Gentner, D., A Stevens. A. L. (1983). *Mental models.* Hillsdale, NJ: F.rlbaiim

Haviland, S I., A Clark. H. H. (1974). What's new? Acquiring new information as a process in comprehension..*Journal of Verbal Learning and \ rrhal Behavior. 13.* 512-521.

Hudson. I. (1983). Correspondences and numerical differences between disjoint sets. *Child Development. 54,* 84-90.

Johnson-l aird, P. N. (1983). *Mental models.* Cambridge: Harvard University Press.

Just. M. Λ , A Carpenter, P. Λ. (1980). A theory of reading: From eye fixation to comprehension. *Psychological Review, 87.* 329-354.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

Kintsch, W. (I9H4). *Strategies for solving word arithmetic problems: Extensions of the Kintsch A Oreeno Model* (Technical Report No. 133). Boulder: Institute of Cognitive Science, University of Colorado.

Kintsch, W. (1985). Tex» processing: A psychological model In T. A. van Dijk (Ed.). *Handbook of discourse analysis (Vv\.* 2, pp. 231-243). London: Academic Press.

Kintsch, W. (1987). Understanding word problems: Linguistic factors in problem solving. In M. Nagao(Ed.), *language and artificial intelligence (pp.* 197-208). Amsterdam. North Holland.

Kintsch, W. (in press). 'The representation of knowledge and the use of knowledge in discourse comprehension. In C. Graumann & R. Dietrich (Ed.), *Language in the social context* Amsterdam: Flsevier.

Kintsch, W., A van Dijk. T. A. (1978). Towards a model of text comprehension and production *Psychological Review, 85,* 363-394.

Kintseh, W, A Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92.* 109-129.

Kintsch, W,, A Mross, E. F Π 985). Context effects in word identification *Journal of Memory and Language, 24,* 336-349.

MarsJen-Wilson, W f> , A Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognition. 10.* 29-63.

McCleiland, J. 1... (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitnv Science.* ◊', 113 146.

McKoon, G., A Ralcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 12. S2-91.*

Meyer, D. E.. A Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations *Journal of Experimental Pswhology. 90,* 227-234.

Miller, J. R., A Kinisch, W.(1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory. 6,* 335-354.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76.* 165-178.

Morton, J. (1979). Word recognition. In J. Moron & J. C. Marshall (Eds.), *Psycholinguistics, Series 2 'Structures and processes* (pp. 107-156). London: Paul Elek.

Mudersbach, K. (1982). Hol-Atomismus als Vereinheitlichung von Atomismus und Holismus. In *Proceedings of the 7;h International Wittgenstein Symposium* (pp. 347-349). Wien: Hoelder-Pichler-Tempsky.

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition,* 22.93-136.

Perrig, W., & Kintsch, W. (1985). Prepositional and situational representations of text. *Journal of Memory and Language, 24,* 503-518.

Pylyshyn, Z. W. (1985). *Computation and cognition.* Cambridge : MIT Press.

Raaijmakers, J. G., & Shifffin, R. M. (1981). A theory of probabilistic search of associative memory. *Psychological Review, 88,* 93-134.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The Development of mathematical thinking* (pp. 153-196). New York: Academic Press.

Rumelhart, D. E., & McClelland. J. L. (1986). *Parallel distributed processing Explorations in the microstructure of cognition. Vol. 1.'Foundations.* Cambridge: MIT Press.

Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language.* New York: Wiley.

Schank, R. C. (1978). Predictive understanding. In R. N. Campbell & P. T. Smith (Eds.), *Recent advances in the psychology of language—Formal and experimental approaches* (pp. 91-101). New York: Plenum Press.

Schank, R. C. (1982). *Dynamic memory.* New York: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowsky, M. (1982). Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14,* 489-537.

Sharkey, N. E. (1986). A model of knowledge-based expectations in text comprehension. In J. A. Galambos. R. P. Abelson, & J. B. Black (Eds.), *Knowledge structures (pp.* 49-70). Hillsdale, NJ: Erlbaum.

Smoiensky. P. (1986). Information processing in dynamical systems: Foundations of Harmony Theory. In D. E. Rumelhart & J. L. McClelland, *Parallel distributed processing. Explorations in the microstructure of cognition. Vol. 1 Foundations* (pp. 194-281). Cambridge: MIT Press.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Reconsideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18,* 523-534.

Till, R., Mross, E. F, & Kintsch, W. (in press). Time course of priming for associate and inference words in a discourse context. *Memory A Cognition.*

Turner, A. A., McCutchen, D., & Kintsch, W. (1986). *The generation of macropropositions during comprehension.* Paper presented at the meeting of the Psychonomic Society, New Orleans.

van Dijk, T. A. (1980). *Macrosimctures.* Hillsdale, NJ: Erlbaum.

• van Dijk, T. A., & Kintsch, W. (!983). *Strategies of discourse comprehension.* New York: Academic Press.

Waltz, D. L., & Pollack. J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science, 9,* 51-74.

Wilensky, R. (i 986). *Some problems and proposals for knowledge representation* (Tech. Rep. No. UCB/CSD 86/294). Berkeley: Computer Science Division, University of California.

Willoughby, S. S., Bereiter, C., Hilton, P., & Rubinstein, J. H. (1981). *Heal math: Thinking stories book. Level 2.* La Salle, IL: Open Court.

Winograd, T. (1983). *Language as a cognitive process. Volume I: Syntax.* Reading, MA: Addison-Wesley.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35.* 151-175.

Received December 16,1986
Revision received September 14, 1987
Accepted September 23, 1987 ·

Kintsch Appointed Editor of *Psychological Review,* 1989-1994

The Publications and Communications Board of the American Psychological Association announces the appointment of Walter Kintsch, University of Colorado, as editor *of Psychological Review* for a 6-year term beginning in 1989. As of January 1, 1988, manuscripts should be directed to

Walter KJntsch
Department of Psychology
University of Colorado
Campus Box 345
Boulder, Colorado 80309