4th AAAI-Sponsored Knowledge Acquisition for Knowledge-Based Systems Workshop
Banff, Canada, October, 1989.

MENTAL MODELS IN THE KNOWLEDGE ACQUISITION PROCESS

Mildred L. G. Shaw and J. Brian Woodward
Knowledge Science Institute
Department of Computer Science
University of Calgary
Calgary, Alberta,
Canada T2N 1N4

email:—mildred@cpsc.ucalgary.ca

ABSTRACT

The main difficulties in knowledge acquisition from domain experts stem from the variety of types of knowledge, the various representations of knowledge, and the intrinsic problems in making these explicit and accessible. In considering these problems it is appropriate to draw on studies in cognitive science and associated disciplines to examine the style of the expert as well as the task demands and goals. The knowledge may be *operative* in that the expert can perform effectively; *tacit* in that the expert cannot express the basis of the performance; *explicable* in a way meaningful to others; essentially *dynamic* in involving continual change; *meta* in that the expert understands the extent of his or her knowledge; and so on. The concepts of *mental model, conceptual model* and *model of conceptual model* have been used to account for the human activities underlying knowledge acquisition and many systems have been developed as a result of the partial application of particular concepts of modeling. This paper reviews several methods of knowledge acquisition, some manual and some of which have been wholly or partially automated, in the context of mental models.

INTRODUCTION

The field of knowledge acquisition lacks a methodological framework which logically supports the process of knowledge acquisition, identifies the characteristics of the products of knowledge acquisition, and offers a set of ordered assumptions about the nature of knowledge and its characteristics. It is imperative that the correct or "operative" knowledge (Johnson, 1986) be identified and made explicit through knowledge acquisition; failing this, the resultant knowledge model may be a poor replica of the original.

Making the knowledge explicit in all its forms challenges the knowledge engineer. Beginning with the domain problem, the initial objectives of knowledge acquisition are to identify both the characteristics of the task situation and the method used by the problem solver to address the task. Without the determination of these two general areas of knowledge, the task of the knowledge engineer becomes an undirected search for whatever knowledge appears to fit the operational formalism chosen for the final computer knowledge base.

If the task demands can be identified and the problem-solving method of the expert ascertained, the types of knowledge required become identifiable. Procedures can then be developed to elicit these types of knowledge, either new procedures or, more likely, those borrowed and modified from other disciplines. The next four sections identify the problems of current approaches and the underlying philosophical, conceptual and methodological issues.

THE "PATCHWORK" APPROACH TO KNOWLEDGE ACQUISITION

Methodologically, the field of knowledge acquisition is characterized by a variety of methods and techniques, mostly imported from other disciplines and adapted for use in knowledge

acquisition. Often a technique has been imported to address a specific issue or problem and then its use is generalized for other purposes. Psychology has offered entity-attribute grids (Shaw & Gaines, 1987; Boose & Bradshaw, 1987) protocol analysis (Ericsson & Simon, 1984) and some forms of interviewing (La France, 1987). Linguistics has offered implicit knowledge structures for text analysis (Woodward, 1988) and forms of discourse analysis, ontological and conceptual analysis (Belkin, Brooks & Daniels, 1986; Hirst & Regoczei, 1989). Epistomology has offered knowledge hierarchies (Hayward, Wielinga & Breuker 1986). The machine learning field has provided induction techniques (Forsyth & Rada, 1986). Artificial intelligence has provided a number of knowledge representation formalisms, inference procedures, and the development of generic structures (McDermott, 1988; Chandrasekaran, 1988), and more procedures are being brought from the area of systems analysis (McGraw & Harbison-Briggs, 1989).

As each method was incorporated into knowledge acquisition, it brought with it assumptions about the nature of knowledge and how knowledge was represented. The assumptions are often implicit, but usually begin to emerge in the knowledge representation language (e.g. entity-attribute grids assume that distinctions are critical primitives underlying knowledge), but also assumptions often remain implicit as in some forms of interviewing. Do the knowledge representations actually represent the knowledge if all assumptions about the knowledge have been left unstated?

The assumptions made about knowledge are reflected in the procedures of the knowledge acquisition method or technique. Methods which emphasize the use of 'task-like' situations (e.g. think-aloud protocols) which provide prompts and cues to draw out or elicit responses are assumed to provide the required knowledge. Methods which emphasize the labeling, selecting, coding or reordering of linguistic material (e.g. discourse analysis, KADS) are assumed to provide an *a priori* structure or schema. The assumptions about knowledge contained in the structure activity also require elucidation. What parts of the knowledge are captured by these methods and what parts are not?

Borrowing methods from other disciplines has helped address the growing knowledge acquisition problem for knowledge based systems and has helped to provide a variety of useable techniques and methods. However, this approach has also resulted in the knowledge acquisition area resembling a patch-work quilt whereby methods and techniques are strung together in a pragmatic rather than systematic fashion. One may ask whether the elicitation, analysis and implementation processes make inconsistent, even conflicting, assumptions about knowledge.

Due to the lack of a systematic framework for knowledge acquisition it is extremely difficult to organize and arrange the proper and appropriate application of the techniques and methods for knowledge acquisition. When and why to apply a particular technique is often left up to the predelictions of the knowledger engineer or the availability of a program for that method.

Some researchers have attempted to solve the problem by extending those methods imported from other areas or by developing a very precise, well-formalized and very powerful method. An example of the first solution is the KADS development. Particular assumptions reflecting the epistomological underpinnings of knowledge representation were used to develop a complex system of knowledge analysis. Examples of the second solution are the knowledge acquisition methods developed by Chandrasekaran (1988) and McDermott (1988). These researchers developed specific methods for identifying repeatable activity during the solution of certain types of problems. These repeatable procedures were then formalized to identify the knowledge required which acts as a basis for knowledge acquisition.

Twine (1989) offers a suggestion for another solution to the "patchwork" problem. He proposes that we begin by making a very clear distinction between "method" and "methodology". A methodology requires the development of a set of guidelines, assumptions and eventually

principles which bring order to a field. In the case of knowledge acquisition, a clearly defined methodology would provide a set of criteria for identifying, ordering and comparing the variety of methods for knowledge acquisition. It would also provide a framework for identifying the "gaps" in knowlege acquisition techniques and methods. Once a gap was found, the methodology would aid in the principled development of new tools which meet methodological requirements. Finally, the methodology provides a framework for logically supporting the knowledge acquisition process, for identifying the characteristics of the products of knowledge acquisition and for making decisions about the nature of knowledge and its characteristics. In this paper we develop a general methodological framework for knowledge acquisition methods that clarifies their underlying assumptions and classifies their structures.

PHILOSOPHICAL ISSUES

The primary philosophical issue in knowledge acquisition may be seen as the contrast between the rational/objectivist approach to knowledge versus the constructivist approach (Woodward 1989). A rational/objectivist approach defines a problem as the observable gap between a given goal or predefined standards of performance and the present state of affairs (Nystedt, 1983). The nature of knowledge is such that knowledge is a body of information acquired by people, something that can be transferred, borrowed, stored, displayed, packaged and bought or sold like any other commodity. Within this framework, the 'task' is an objectively defined set of situational demands which must be met or addressed in order to close the gap. Assumptions are that the gap can be objectified, that the problem-solving process addressing the task can be identified and agreed upon and that the problem can be decomposed into sub-problems. Diagnostic tasks in one domain should be the same as diagnostic tasks in other domains. This philosophical position ignores the presence of the problem-solver; or at the very least, reduces the problem-solver to a container for replicable problem-solving procedures.

In contrast, the constructivist approach defines a problem as an artifact of cognitive processes (Nystedt, 1983) and problem-solving is defined as problem-formulation combined with the steps taken to arrive at cognitive closure (e.g. satisfaction, reduction of dissonance). Problems may be sub-divided but more likely they are reformulated. This approach assumes that the goal may be shared with others but there may be numerous individual methods of reaching the goal. Different experts in the same domain may use different problem-solving methods to solve the same problem. This philosophical position reflects a view which recognizes the problem-solver as central to problem-solving. The problem-solver is the producer of the knowledge. For knowledge acquisition, the focus is on identifying the features of the task which are salient to the problem-solver and making explicit the cognitive procedures which are used. Knowledge about the task is only used as it is seen through the eyes of the problem-solver (Compton & Jansen, 1989).

Both approaches assume that tasks and methods may be classified based on various qualities and characteristics. This is useful for knowledge acquisition. If knowledge engineers can recognize the *task characteristics* and the *problem-solving method* of the expert and match the two, then knowledge base development takes on a more methodical, less idiosyncratic flavor (Welbank, 1987). The salient types of knowledge can be identified and algorithms for the processing of the knowledge constructed or selected. However, there does not yet exist a taxonomy of either task characteristics or problem-solving methods which can assist knowledge engineers throughout the knowledge acquisition process (Woodward, 1989). Also, there is no useful scheme for mapping specific task demands with problem-solving processes.

CONCEPTUAL ISSUES

The conceptual difficulties in knowledge acquisition stem from a lack of consensus on what framework can be used for representing the task demands and problem-solving method (Clancey,

1989). The concept of *mental models* has been developed by a number of researchers in cognitive science. Johnson-Laird (1983) operationalizes this concept as logical propositions represented by quasi-logical propositions. This approach recognizes mental models at the 'micro' level whereas Graesser and Clarke (1985) identify "general knowledge structures" as unique and self-contained mental models. This same level of operationalization is demonstrated by other examples of general knowledge formats or packages such as frames (Minsky, 1975), scripts and/or plans (Schank & Abelson, 1977), productions (Anderson, 1983), memory organization packages (Schank, 1982), and thematic abstract units (Winston, 1984).

At the 'macro' level, the mental models concept can be operationalized as styles of problem-solving. Styles refer to the general approach taken by the problem-solver in addressing a variety of tasks. An example of this level of definition is the formalizing of modes of inquiry. Churchman's (1971) book on inquiry identifies five major schools of western philosophy which provide the basis for different types of problem-solving systems. Each system is characterized by its definition of knowledge, how the knowledge is combined, what constitutes a solution and the differing strengths and weaknesses of the inquiry process. Lockian, Leibnitzian, Kantian, Hegelian and Singerian inquiring systems are outlined. This scheme has been translated into a practical tool which attempts to identify an individual problem-solver's mode or style of inquiry (Harrison & Bramson, 1982).

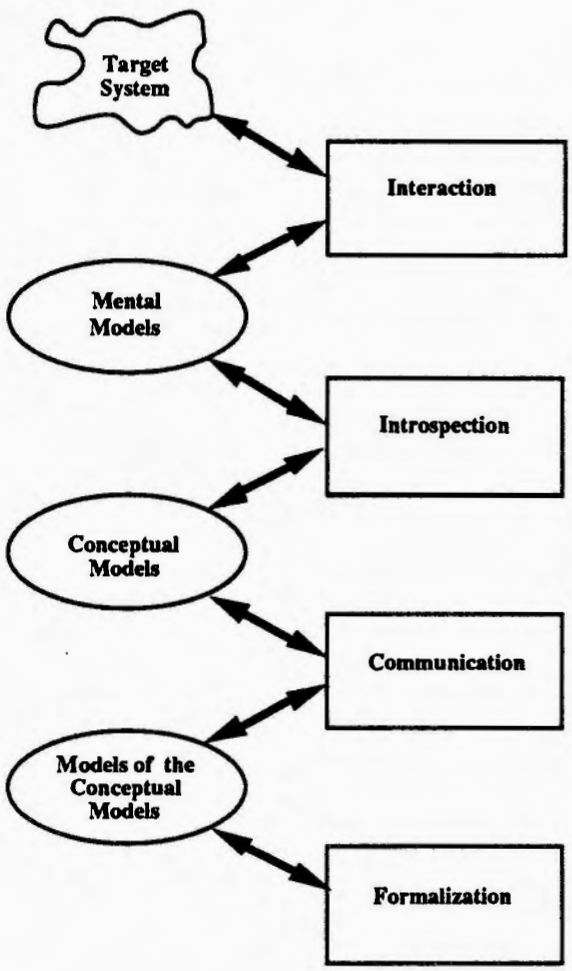


Figure 1. The modeling processes leading to the intermediate knowledge bases.

The conceptualization of mental models at the levels outlined offers useful representation schemes and operational definitions, but these lack an overall framework which would help to organize and arrange the various schemes or models of knowledge. Norman (1983) points out that the first requirement is to define the system that the expert is interacting with, "the system that the person is learning or using, by definition, the *target system.*" Figure 1 shows the relationships between the target system, the mental models of the expert, the conceptual models of the expert and the processes of interaction with the target system, introspection about the interaction, communication of the introspection, and formalism of the communication. No intrinsic model exists in the target system, but rather an ordered or repetitive sequence of activities or events. A variety of potential models are possible for this target system which represents the 'task'.

Using Norman's terminology, the expert interacts with the target system and may produce many *mental models* of the target system. The perceptual system of the expert is capable of isolating and defining 'primitives' by representing them in symbolic form. There is some debate in cognitive science about the nature of these representations and about the level at which the representations are confirmed. Empirical research suggests that it is possible for various models to exist at different levels in perceptual and cognitive processes (Pylyshyn, 1984; Fodor, 1980; Paivio, 1977).

The next modeling process results in a model of the model of the target system. Norman refers to this as the *conceptual model* of the target system because it is the model that is reflected in communicating the mental model to other entities. The conceptual model finds representation in the communication primitives of language and behavior. The mental model affects the form and representation of the conceptual model and is, in turn, constrained by the language and behavioral representational system. It is this conceptual model that is available to others and the one which requires further translation and representation by other modeling activities. As the mental model is unavailable to an observer, the conceptual model is taken to represent the model used to address the task demands of the target system.

The conceptual model becomes the focus of further modeling efforts. Knowledge engineers, or automated knowledge acquisition systems, take these models and enforce an order through the underlying models implicit in their elicitation procedures from the introspection of the expert about his interaction with the target system. The conceptual model provides a set of primitives which are communicated to the analysis processes and recognized by the analytic methods used (which, in turn, are directed by models of their own). The *intermediate knowlege base* is produced from the conceptual model through the elicitation and informal analysis procedures which rely on the introspection and communication of it from the expert. There is no direct link between the mental and conceptual models, between the cognitive processes, or between the elicitation and analysis procedures, but the processes are iterated back and forth, the models providing new data for the procedures which in turn select parts of the essential features from the models until the intermediate knowledge base emerges.

The objective of this modeling activity is to select a set of primitives from the conceptual model which can be used for another model. This activity requires an ordered and systematic procedure of its own. Representing the primitives of the conceptual model necessitates another medium and another representation structure. What is generated is a series of models of the conceptual model of the target system, in various forms of refinement and representation. These models are usually designed to represent different primitives which are considered important to the task. Primitives, which represent facts of the domain, are included with the inferencing procedures. Also, forms of strategic processes which guide the use of domain knowledge primitives and the inferencing processes are included. The result of this activity is a new model of the conceptual model which is the model of the conceptual model of the target system. This model represents the interrelationships between the task demands of the target system and the model for

addressing the demands of the conceptual model. It represents the salient features of the task and its context.

Each successive model requires a representational schema, another medium of representation, and an ordered procedure for translating one representation to another. The model of the conceptual model of the target system should contain the necessary primitives to recreate the operation of the mental model (Norman, 1983). There may be many models of the conceptual model of the target system that can perform the same activity of simulating the mental model. The model of the conceptual model of the target system represents the final working or operational model. By formally analyzing this, a computer knowledge base may be identified and coded in the representation language. As in earlier stages, this process may be iterated and refined by extracting knowledge structures produced until the computer knowledge base is sufficiently well implemented to represent the model of the conceptual model of the target system as shown in figure 2.

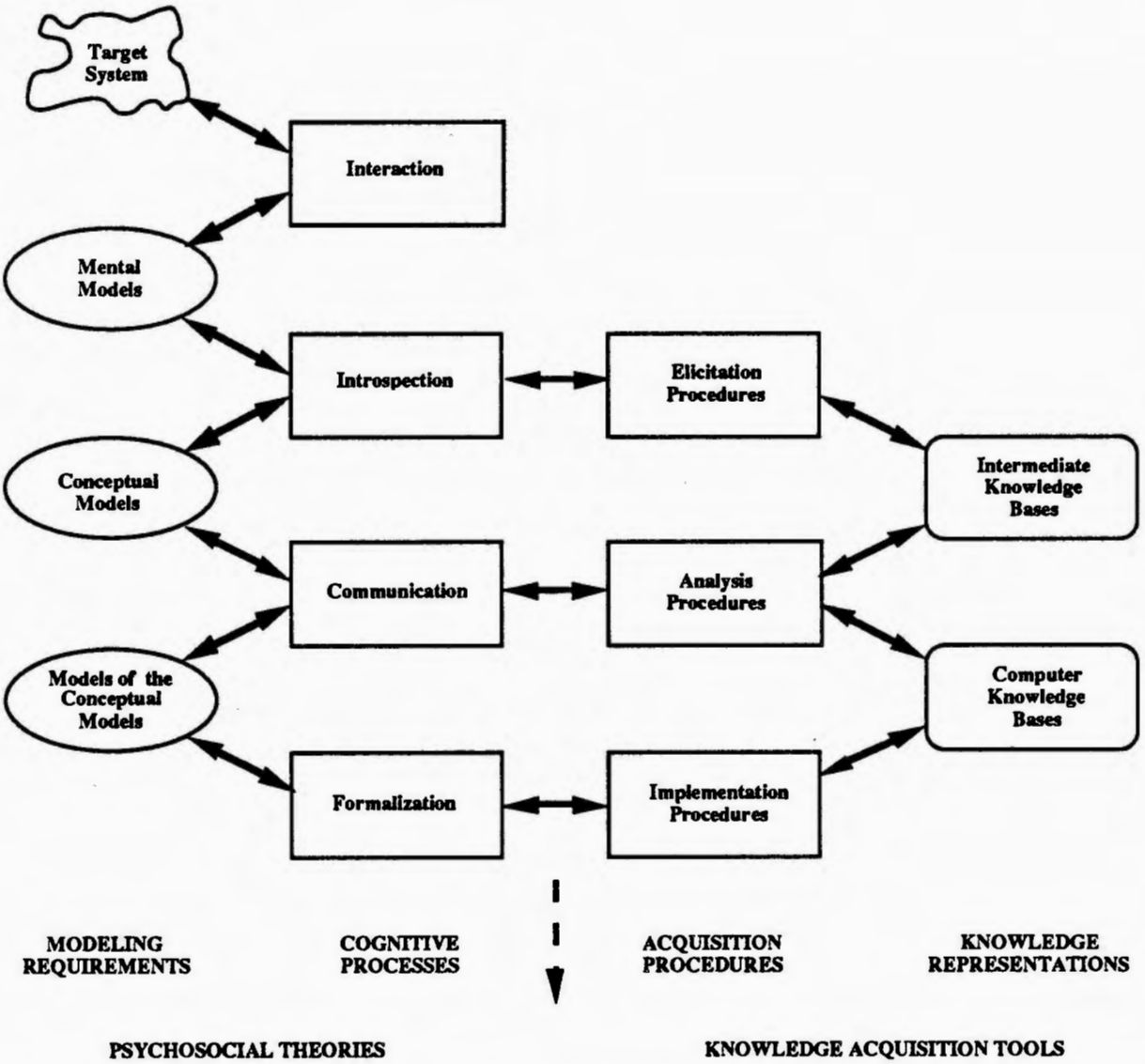


Figure 2. The processes linking the modeling requirements, the cognitive processes, the acquisition procedures and the knowledge bases

This method of development inherently links the task demands and the problem-solving method and models of the conceptual model of the target system which are developed must maintain these connections (Woodward, 1989).

The complete modeling process for the knowledge acquisition phase of development is shown in Figure 2. This shows the links from one part of the model to another and emphasizes the essentially iterative and refining nature of knowledge acquisition. The left side of the figure uses Norman's terminology applied to an expert involved a problem-solving task. It shows the social effect of communication of his problem-solving in a goal-directed situation, where the goal is to produce a computer implementation of his knowledge. The right side is shows the computer-based knowledge acquisition and representation tools.

In summary, the *target system* is the focal point for the problem-solver which has patterns of activity or knowledge inherent to the context or situation. No models exist in the target system, only ordered or repetitive sequences of events. The *mental models* represent unexpressed, uncommunicated knowledge in the expert's mind — fragmentary, partial models generated to account for his or her purposeful interaction with the target system. At this stage primitives in the situation are isolated and defined. The *conceptual models* represent the models of the models of the target system. They represent knowledge expressed in a linguistic form which can be communicated to others. they are not formal and not necessarily consistent and usually requires further refinement. They are the focus of further modeling. The *models of the conceptual models* of the target system are more formally expressed and available for communication, interpretation and discussion with others. They may encompass a variety of media and consist of a series of models which represent the interrelationships of task characteristics as modeled by the problem-solver.

METHODOLOGICAL ISSUES

Methodologically, the field of knowledge acquisition is characterized by a variety of systems, each useful for a selected type of knowledge. Confusion exists between elicitation methods and analysis methods. Norman's modeling framework also offers a way out of this methodological dilemma in knowledge acquisition. Any procedures which aid in the transition, or translation, from the mental model to the conceptual model, may be called *knowledge elicitation* methods. These procedures (interviewing, laddering, brain-storming, repertory grids, MOLE, SALT, etc.) are designed to draw out and represent forms of knowledge. In responding to the elicitation stimuli, the problem-solver arranges his or her knowledge in linguistic form (verbal and behavioral) according to the structure of the prompts. The elicitation methods are particular to specific types of knowledge and must be choosen appropriately. The conceptual model is, however, a rough model of the mental model of the target system and must be further refined into an operational form.

The translation step from the conceptual model to the model of the conceptual model of the target system can be seen as *knowledge analysis* because the end result is a knowledge model or qualitative model (Clancey, 1989) based on the linguistic primitives in the conceptual model. These methods, such as the analysis portion of KSS0 (Gaines & Shaw, 1987), induction programs (Quinlan, 1987; Dietterich & Michalski, 1981; Gaines, 1989), Cognosys (Woodward, 1988), protocol analysis (Ericsson & Simon, 1984), perform analyses on the elicited knowledge and result in another model of knowledge. These models may undergo numerous iterations. The result is a methodological distinction between elicitation and analysis procedures. This suggests the need for operationally connecting the elicitation and analysis procedures through the intermediate knowledge base which is done, for instance, in KSS0 resulting from the earlier work of Shaw (1980).

Integration of a flexible range of tools is of paramount importance (Gaines, 1987) especially with automated knowledge acquisition methods. With current systems, the knowledge engineer is allowed few degrees of freedom. If the current domain does not conform to the tool, the domain is sometimes 'redefined' to fit the tool. Manual methods, on the other hand offer a flexibility and adaptability not found with automated tools but frequently offer help with only a part of the elicitation and analysis tasks required for knowledge acquisition.

PROPERTIES OF MENTAL MODELS

Assumptions made about the mental models are propagated to later models. A simple formulation of the components required for mental representational activity come from the description of verbal reports (Ericsson & Simon, 1984) which are considered as reflecting a subset of the information held in the short-term memory. Cognitive processes are postulated as a sequence of internal states successively transformed by a series of information processes. Also, information is assumed to be stored in several different memories which reflect different capacities and display different capacities and accessing characteristics (p. 11).

The mental model, or the cognitive representations in the human knowledge source, represents that which knowledge engineers attempt to discover and model. These models represent two different processes. The first is the forming of models. Models are formed during information gathering when "some basic level concepts and relationships [are selected] on the basis of perceptual input and prior expectations" (Gilhooly, 1987 p. 23). New situations are initially categorized in terms of an existing model but if this one is insufficient then a new model is developed (Kelly, 1955) from basic concepts or by combining parts of existing models. The second use of models is in prediction and problem solving. The models suggest a set of hypotheses about the system under study. The hypotheses change as new information is perceived and incorporated (Kelly, 1955).

For knowledge acquisition, the nature of these models requires elucidation because elicitation techniques often assume (implicitly or explicitly) particular cognitive structures and processes. For example, the generic task approach (Chandrasekaran, 1988) assumes that problem-solving can be modeled as a series of "modular" processing activities using specific types or forms of knowledge. Elicitation of these processes and the knowledge forms required act as the basis for knowledge acquisition activity.

In order for elicitation procedures to be more useful, the underlying assumptions about knowledge and knowledge processes require elucidation and definition. The structures and descriptions of knowledge types, roles and functions require objectification and definition to be useful to the knowledge engineer. This is accomplished when a well developed theory underlies the elicitation (and hopefully the analysis method as well) and is used to make assumptions about the structure of knowledge. KSS0 (Gaines & Shaw, 1987) is based on Kelly's (1955) method of repertory grids within the framework of personal construct psychology. Cognosys (Woodward, 1988) uses the method of generic knowledge structures (Graesser & Clark, 1985) as a method within the framework of implicit knowledge structures. These methods make very definite assumptions about the nature of the knowledge and the inference processes at play in the mental model. Other researchers (Craik, 1943; Johnson-Laird, 1983; Gentner & Gentner, 1983; Norman, 1983) use the concept of mental models to outline and define the components of knowledge. It has also been postulated that various mental models exist at various levels in the perceptual and cognitive faculties (Pylyshyn, 1984; Fodor, 1980; Paivio, 1977).

PROPERTIES OF INTROSPECTION AND ELICITATION PROCEDURES

Elicitation techniques for knowledge acquisition support introspection of the expert and act as classification grammars for linguistic elements and linguistic levels. After the task situation has

been selected, instructions are normally given to the expert in order to generate verbal and/or behavioral events. These events are collected, then ordered or parsed into useful phrases which form the conceptual model. The specific instructions act as a constraint on the mental model material in that they structure the verbal material from the expert. The method of collection may result in fairly complete collection such as video recording or partial collection such as note-taking.

There appear to be few variations on the main types of elicitation procedures. Each type involves assumptions about the type of knowledge or information and where it resides. Ericsson and Simon (1984) note that information can be elicited retrospectively, inferentially, descriptively, or by direct questioning. In the first case, retrospective elicitation is assumed to tap information stored in the short-term memory if the time between the completion of the task and the elicitation procedure is not too long. It is assumed that the expert can describe the heeded information used for task completion.

When given instructions to report on the steps in solving a task, the expert will make inferences about what he or she 'must' have thought. Inferentially generated material in elicitation procedures constrains the expert's responses by requiring information that is plausible. The information tapped in this case may or may not have anything to do with the heeded material or those salient knowledge cues used in completing the task. Rather, the communication context between the knowledge engineer, or knowledge acquisition tool, and the expert becomes more salient.

Descriptive types of elicitation take the form of thought descriptions at the time of completing the task. These types of elicitation instructions promote the verbalization of information which is currently in short-term memory. It is assumed that the verbalization processes do not interfere with the completion of the task or influence the processes required for task completion (Ericsson & Simon, 1984). The vocalizations are assumed to be a commentary on the problem-solving method used while the expert is completing the task.

Direct probing is assumed to elicit pieces of information directly from the short-term memory with no need for intervening accessing cues. If the information is not in short-term memory, then the direct probing can still be used to elicit information in long-term memory by way of accessing cues. La France's (1986) grid is based on the assumptions that information is stored in many different memories and so can be accessed using a variety of probes.

The information resulting from the above elicitation types may reflect different qualities of verbal material. The simplest verbal material in the conceptual model results from direct verbal descriptions of the information heeded while completing the task. Very little effort is required by the expert for this activity because the information points within the awareness of the expert while completing the task are simply labelled.

More complex material comes from elicitation into the conceptual model when the expert begins to describe the thought content of the mental model in short-term memory. This information assumes an 'observer' role such that the expert is able to describe the thought processes and knowledge roles which are salient and relevant to the task. It is assumed that these verbalizations are descriptions of the mental model in action or descriptions of the mental model in past action.

The most complex level of verbalization occurs when the expert begins to explain the thought processes, thoughts, ideas, hypotheses or motives required when completing the task. The act of explanation requires the linking of cues in short-term memory to earlier information and/or to information in long-term memory. This need to link information not currently in the short-term memory results in an increase in the amount of information and may even after the salience and relevance of the information (Ericsson & Simon, 1984, p.79).

Elicitation procedures implicitly assume characteristics about the knowledge they intend to elicit; they select and order some of the information available in the mental model. The elicitation procedure uses its 'theory' to constrain the types of information from the expert. For example, KSS0 asks the expert for objects or "entities" which the expert believes pertain to a certain purpose or task. These entities may be physical objects or procedures, design components, cases, etc. Then, the expert is asked to describe the distinctions between and among the objects. This form of retrospective elicitation uses its background theory to identify the required elements of the mental model and produces a severely constrained, but highly functional, conceptual model and intermediate knowledge base in the completed entity-attribute grid. The grid can then be used as the basis for a variety of analyses.

PROPERTIES OF CONCEPTUAL MODELS

For knowledge acquisition, this model is the final outcome of elicitation activity. The elicitation procedures structure the information from the expert and it is assumed that this information reflects some of the available knowledge characterizing the expert's mental representation of the particular task as it is interpreted and represented by the elicitation procedures. The conceptual model acts as the first available representation or model of the expert's thinking processes and knowledge concerning the task situation. It contains representations of communicable information about the mental model.

In most cases, the conceptual model is represented in the form of linguistic structures. Because of this representation formalism, a number of assumptions are made about the information available due to the use of linguistic structures. A basic assumption is that linguistic structures are an adequate, reliable and valid means of studying human thought processes. This assumption has been the basis for much of the modern research on linguistics (e.g. Chomsky, 1975) but has not been explicitly stated in knowledge acquisition work. It is possible to extend this main assumption to make it more applicable to knowledge acquisition by outlining a number of assumptions about knowledge and cognitive processes which apply in knowledge acquisition work.

In linguistics it is postulated that the linguistic material produced is a surface-structure representation of a base-grammar, that operations are performed on the base grammar to produce a communicable description of the underlying representation. The view of transformational grammarians (e.g. Chomsky, 1975; Gazdar, 1981; Bresnan, 1982) is that deep structure relations are transformed through a semantic analysis to produce surface, or linguistic, representations. Elicitation methods appear to affect the integration processes of the transformations, and depend upon this transformational process in that such procedures constrain the types of surface-structure components and relations possible.

It is the nature of linguistic material which allows the knowledge engineer the opportunity of making inferences about the material. Graesser and Clark (1985 p. 14) suggest that much of the meaning of the conceptual model is not exclusively in the text. Inferences can be drawn based on contextual cues, on the linguistic material itself, on the goals of the comprehender, on the degree of shared knowledge between the generater and reader of the conceptual model and, finally, about the knowledge structures in long-term memory. The richness of the conceptual model depends somewhat on the elicitation process, and the usefulness of the linguistic representation of the conceptual model lies in the knowledge engineer's ability to generate meaning based on properties of the linguistic structure.

It must also be assumed that the linguistic material somehow reflects abstract structures in the mind which make knowledge possible. The production of linguistic structures are assumed to reflect these mind-based abstract structures and processes. That knowledge engineers and

knowledge acquisition methods are able to aid in the linguistic description of internally represented structures and processes has not been proven experimentally but is assumed to be true. We believe that what we elicit is an approximate representation of knowledge; assumptions are made about the ability of linguistic material to capture the expert's intrinsic competence and internalized knowledge.

Ericsson and Simon (1984, p. 221) make specific assumptions about the cognitive mechanisms underlying verbal descriptions. They state that:

1. "verbalizable cognitions can be described as states that correspond to the contents of short-term memory (i.e. to the information that is in the focus of attention);
2. " the information vocalized is a verbal encoding of the information in short-term memory;
3. " The verbalization processes are initiated as thought is heeded;
4. " the verbalization is a direct encoding of the heeded thought and reflects its structure;
5. " units of articulation will correspond to integrated cognitive structures; and
6. " pauses and hesitations will be good predictors of shifts in the processing of cognitive structures."

PROPERTIES OF INTERMEDIATE KNOWLEDGE BASES

Most knowledge engineers, and knowledge acquisition systems, produces multiple and varied intermediate knowledge bases. Welbank (1987) interviewed a number of knowledge engineers and discovered that there is much variety in how knowledge engineers deal with linguistic material from the conceptual model. Idiosyncratic, ad hoc, methods were reflected in the use of a variety of diagramatic methods and representation procedures. There is more activity in the knowledge acquisition community which has focused on the production of procedures for generating intermediate representations of the conceptual model material (e.g. Johnson & Johnson, 1987).

Intermediate knowledge bases are best characterized in terms of their functions. One main function is as a representation for display to the expert so that validation and confirmation are possible. A diagram is useful to help confirm understanding of the expert's mental model. A second function is to act as a focal point for further elicitation; the initial elicitation acts as the foundation for acquiring more information. An example is the development of a domain dictionary composed of term definitions (McGraw & Harbison-Briggs, 1989) which constitutes a domain vocabulary. Another example is the development of ontological categories based on the syntactic position of words. A noun suggests a static object; a verb suggests action and change; adverbs and adjectives suggests styles and distinctions. Attributes or distinctions can be elicited and combined in groups.

Intermediate knowledge bases are characterized by structures and forms based on qualities already inherent in the natural language structure of the conceptual model rather than on a formalized analysis procedure. Because of their functional properties, intermediate knowledge bases are informally-produced, partially-ordered, linguistically-based, and inherently unstable structures.

PROPERTIES OF COMMUNICATION AND ANALYSIS PROCEDURES

Informal analysis procedures support the expert's first attempts at communication and draw out meaning from the linguistic material of the conceptual model and the partial orderings and structures of the intermediate knowledge bases. Informal analysis procedures are used to identify underlying, implicit or assumed knowledge structures in the elicited material. These procedures are usually theory-laden in that they reflect assumptions made about the structure of knowledge and how the structures can be operationalized. These analysis procedures are possible only because of an underlying theoretical framework which justifies the labelling and ordering of the

material. They are possible also because of their formality and consistency. They reflect a step-by-step, algorithmic nature and they are capable of being applied consistently to elicited material.

Analysis procedures usually begin with some form of encoding of the linguistic material from the conceptual model or from the intermediate knowledge bases Encoding of these materials impose an *a priori* framework over the conceptual model and intermediate knowledge base material. This framework identifies the 'required' linguistic primitives and labels or categorizes them. For example, in the KADS method (Breuker & Wielinga, 1987), verbs or groups of verbs in the text are viewed as "knowledge sources" in the sense that verbs indicate action and that action is usually done to something to produce something else. Knowledge sources are considered as producers of structure or meaning. Nouns or groups of nouns are, in turn, considered as objects which function as either inputs to knowledge sources or as the products or outputs of knowledge sources.

Once material has been coded, more formal analysis techniques group the encoded material into various structural configurations which aid communication of the material to other people and processes. Groups of encoded material are distinguished which can then be ordered, compared, matched, or constructed to produce more meaningful knowledge structures. In Cognosys (Woodward, 1988) "knowledge units", noun-verb-noun phrases, are labelled as events, states, goals or styles. These are then connected to other knowledge units on the basis of linking rules (Graesser & Clark, 1985). Only knowledge units of certain types can be linked to others. Once the linking is complete, the categories are used to develop a variety of sub-structures which appear evident in the task. For example, goal-oriented sub-structures identify the ultimate goal and sub-goals. Implication sub-structures identify implicit connections within the task.

PROPERTIES OF MODELS OF THE CONCEPTUAL MODELS

The results of formal analysis procedures produce a well-defined, conceptually ordered, theoretically consistent model of the task and context under consideration. This model can be viewed as a completed theory of knowledge, based on consistent use of concepts and a methodical application of formal rules which is able to trace the origins of the knowledge and document its transformations. The model of the conceptual model stands as the formalized representation of the more expressive conceptual model.

The model viability is ensured by containing all the necessary elements which allow the implementation process. Addis (1989) has identified three necessary components of knowledge which may be applied to the models of the conceptual models: first, each model should contain the required declarative knowledge reflecting tacit and factual knowledge inherent in the problem-solving method and in the task demands; second, the inference knowledge which operates on the declarative forms is necessary to identify the transformation or processing steps of the knowledge; third heuristic knowledge is required which sometimes act as special constraints, sometimes as inferences and sometimes as strategic procedures.

This model of the conceptual model is also characterized by formal interrelationships between task demands, external to but perceived by the problem solver, and the problem-solving method used by the problem-solver (Woodward, 1988). Task cues and their meanings are represented in the model and these cues are directly linked to inference, heuristic and declarative structures.

PROPERTIES OF FORMALISM AND IMPLEMENTATION PROCEDURES

Implementation processes draw upon the formalized structure of the model of the conceptual model. Based on well established principles of artificial intelligence and underlying assumptions about knowledge, these processes select, order, arrange and represent elements of the model in a form conducive to operation in a computer-based knowledge system. Inference, heuristic and

declarative knowledge are all coded into formats which reflect their different natures and functions. A variety of declarative knowledge forms are used to build portions of the knowledge base. Inference knowledge is represented by the ways in which the declarative knowledge is processed and manipulated. Heuristic knowledge is represented as meta- or control knowledge.

Due to the number of possible formal methods of implementation, different aspects of the model of the conceptual model may be selected. Rule-format coding differs from frame-based coding which differs from logic-based coding. The appropriate choice depends on whether the required knowledge structures are represented in an applicable form within the model of the conceptual model. An efficient model would allow many implementation methods.

PROPERTIES OF COMPUTER KNOWLEDGE BASES

This final result reflects the multiple methods used in the process. The computer knowledge base represents a theory of knowledge that was constructed through the elicitation, analysis and implementation processes and based the assumptions underlying the representations in each modeling phase. Declarative knowledge is usually represented separately from the inferential knowledge but once complete, the final knowledge base operates in algorithmic fashion.

The first part of the paper has identified the primary characteristics of and knowledge acquisition process to be the underlying models and presuppositions they make, and the elicitation, analysis and implementation procedures used. The next section analyses some widely used knowledge acquisition techniques in these terms. The table below summarizes the techniques surveyed, how they relate to the presupposed models and the procedures used in the elicitation, analysis and implementation techniques used in each case.

Method	Presupposed Model	Elicitation Procedures	Analysis Procedures	Implementation
Interviewing (La Franco, 1987) (Johnson & Johnson, 1987)	General but weak	Interview: unstructured semi-structured structured	Extraction and organization of domain concepts	Not supported
Protocol Analysis (Ericsson & Simon, 1984) (Johnson, Zualkernan & Garber, 1987)	Possibility of adequate introspection	Recall of introspection	Extraction and organization of domain concepts	Not supported
Empirical Induction (Quinlan, 1987) (Dietterich & Michalsky, 1981) (Gaines, 1989)	Domain structure already identified; large dataset of expert behavior available	Collect large dataset of expert protocols (evaluations and decisions)	Induction of decision rules	Backward chaining rule-based system
Domain Ontologies (Breuker & Wielinga, 1987)	Applicable multi-level ontology for knowledge representation	Structured interview; protocol analysis	Fit ontological structure to domain	Logic-based interpretation of structure
Problem-Solving Ontologies (McDermott, 1988) (Chandrasekaran, 1988)	Existence of applicable generic problem-solving methods	Elicit repetitive process of problem-solving in domain	Fit generic problem-solving methods to observed behaviors	Parameterize implementation of generic problem-solving methods
Implicit Knowledge Structures (Woodward, 1989)	Implicit knowledge structures	Recall expert problem-solving behavior in domain	Identification of noun-verb phrases	Use as a basis for semantic net, frame or rule-based coding formalisms
Entity-Attribute Grids (Gaines & Shaw, 1987) (Boose & Bradshaw, 1987)	Knowledge characterized by distinctions made	Elicit entities in domain and distinctions between them	Conceptual clustering to confirm domain structure; induction to determine constraints	Generate classes, properties, methods, objects and values

Table 1. Some major knowledge acquisition techniques classified according to the presupposed model assumed, and the procedures used for elicitation, analysis and implementation.

SOME MAJOR TECHNIQUES FOR THE ACQUISITION OF KNOWLEDGE

Interviewing (McGraw & Harbison-Briggs, 1989) is used to identify tasks and major concepts, to structure and refine already-acquired information. Unstructured interviews are appropriate when the knowledge engineer wants to explore an issue or the goal is to establish rapport and to get a broad view. It facilitates the description of the domain in a way that is easy for the expert to understand. However, the data acquired is often unrelated and difficult to integrate due to its lack of structure. It may not allow gathering of specific knowledge and takes time and training to do well. Also, similar questions may be asked in future sessions and this may annoy the expert. Alternatively, structured interviews have the advantages of forcing an organization on the interview and being very goal-directed. They attempt to remove distortion from experts subjectively and allow better integration of material after the interview. They force the expert to be systematic, and the knowledge engineer is able to identify gaps in the knowledge which act as a basis for further questions. The purpose of the session is clear to the expert. However, they need more preparation by the knowledge engineer who must study the background material extensively.

During the interview, of any kind, the knowledge engineer must decide on a note taking approach, for example, video- or audio-taping, each of which has advantages and disadvantages. Both open and closed questions can be used. Open questions are broad and place few constraints on the responder. They are used for determining scope of understanding, response certainty, and the models used. These questions like *what, why, how* allow the expert to express information the knowledge engineer does not know about. He can obtain the expert's vocabulary, concepts, frames of reference, and can help with explanations and underlying theory. Closed questions such as *who, where, when, which* tend to set limits on the type, level and amount of information the expert provides since they often provide a choice of alternatives such as a bipolar or multiple choice response. These are used more for clarifying or probing questions or as feedback as they are less time consuming for specific information. Note-taking is much easier but sometimes one can get too little information as they may stop the expert from volunteering information. They requires an excellent command of vocabulary and concepts from the knowledge engineer. On balance a combination of open and closed questions is preferable.

One example of semi-structured interviewing is the Knowledge Acquisition Grid (LaFrance, 1987). This is a matrix of knowledge types and forms where the forms of knowledge are *layouts, stories, scripts, metaphors* and *rules-of-thumb*; and the question types are *grand tour, cataloging categories, ascertaining attributes, determining inter-connections, seeking advice* and *cross-checking*.

Examples of a "grand tour" are distinguishing boundaries of the domain, perspective, goals organization, e.g. Please describe the kinds of things you do; of "cataloging may be an organized taxonomy of expert terms and concepts, e.g. How would you group and subgroup these things?; of "ascertaining attributes" are distinguishing features and range of values, e.g. Describe similarities and differences of concepts; of "determining interconnections" are uncovering relations and causal model(s), e.g. Why?; of "seeking advice" include revealing strategies, e.g. Which conditions warrant which actions?; and of "cross-checking" are validation concerns, e.g. knowledge engineer plays naive, devil's advocate, poses hypothetical situations and asks How sure are you? Are there exceptions?

The teachback technique of Johnson and Johnson (1987) is a semi-structured interviewing method based on Pask's (1975) teachback techniques and Systemic Grammar Networks from Linguistics. This method attempts to formalize a functional approach to grammar. It is an intermediate representation used to teach back to the expert the knowledge acquired in the interview. It results in a mediating representation which is a machine language independent

formalism between verbal data and software development environments. It is used as a tool to help proceed on early elicitation and conceptualization for synthesizing knowledge from talk with experts. The expert is encouraged to use his or her own words to express facts, procedures and abstractions which are recorded. The transcript is segmented, paraphrased, indexed and annotated by the knowledge engineer and the results captured in a systemic grammar network and checked and amended with the expert. It is mainly used as a means to prepare explicit descriptions of knowledge for the use of the development team to construct an associative network, graphs composed of nodes joined by links. It is said to be a generative form which captures all relevant information but excludes absurdities.

Interviewing protocols whether structured or unstructured, specific or general, probing or expansive, usually require a retrospective approach by the expert. This method makes implicit assumptions about the processing of knowledge but little about the nature of knowledge. Interviewing assumes that the knowledge expressed by the expert is based on the question prompts and is the actual knowledge used by the expert during task completion. Further, this knowledge needs to come into short-term memory before the expert can answer and that the expert needed to use some set of accessing cues (La France, 1987) to bring the knowledge from long-term memory. This process is vulnerable to influence by the structures of the communication situation between expert and knowledge engineer.

This method assumes a neutral stance with respect to types of knowledge elicited, although the question content itself makes assumptions about what knowledge is important and required. The encoding procedures developed for the results of interviewing (e.g. discourse analysis) are left to provide some implication of the nature of the knowledge acquired.

The result of interviewing is a very rich, but extremely unstructured and perhaps somewhat invalid (Bainbridge, 1986) conceptual model. On the other hand, the conceptual model allows a great number of intermediate knowledge bases which may be developed, described, demonstrated, redefined, etc. These intermediate knowledge bases act as useful indicators of validity and as stimuli for further elicitation.

There exist few formal procedures for the analysis of unstructured material. Even though these other procedures are being adapted from other disciplines, they do not yet meet the requirements of producing a formal model of the conceptual model. The results of grounded theory and of the systemic grammar method actually result in a better conceptual model. Jumping from a conceptual model to coding processes without the aid of a formal model of the conceptual model may account for much of the difficulty and inaccuracy of the interviewing method of knowledge acquisition.

Protocol Analysis includes both observations of expert behavior and expert verbalizations. Two examples which are widely used, are Verbal Reports as Data (Ericsson & Simon, 1984) and the Specification of Expertise (Johnson, Zaulkernan & Garber, 1987). Ericsson's method uses concurrent and retrospective verbalization for elicitation. "Concurrent report" reveals the sequence of heeded (attended to) information and "retrospective report" reveals the sequence of specific cognitive processes. A think-aloud transcript is then encoded into segments (assertions propositions, reading, doing arithmetic, visual searching, long-term memory recall, generalizing combinations) and relations of code are derived from verbs, propositions and adjectives. "Arguments" are often nouns and nouns phrases and the task analysis defines *a priori* a space of possible encodings representing information relevant to the task.

The method is based on the facts that:

- heeded (attended to) information is identified by analysis of corresponding verbalization through inference and task analysis

- the original encoding should reflect as closely as possible the verbalized information, preferably by automated or semi-automated means
- encoding uses a problem space approach (behavior is a search through a problem space involving the step-by-step accumulation of knowledge)
- the model distinguishes information directly stored in memory (facts, experiences, perceived events and behavior in past situations) from information generated or produced by inference (reactions and behavior in hypothetical situations)
- the criteria for inferring underlying cognitive processes from thinking aloud protocols ar relevance to task analysis and consistency with preceding verbalizations
- information in focal attention is vocalized directly or is encoded, recovery implies decoding
- segments are analyzed using syntax, pauses and intonation
- each segment is independent (the context is ignored for the reason cited above)
- encoding involves finding a category expressing the same information as the segments

The advantages of this method are that it can provide a complete record using both concurrent and retrospective reports then after definition of problem space and segmentation of the protocol the analysis can be automated or the task made much easier for the coder as it focuses the coder's attention on the segments. It requires consistency and explicitness of underlying assumptions, vocabulary and inference rules and the automated encoding lessens the risk of incorrect inference. A comparison of talking aloud protocols shows a very high consistency for practiced skills and successive protocols show stable task strategies and cognitive representations. It is an appropriate method for reporting specific, recent, episodic memories. On the other hand, retrospective reports may leave out important detail and include inferences and generated information (plausible explanations). It is often difficult to tell if the expert pays attention to the knowledge engineers's instructions. The expert cannot report cues allowing recognition of stimuli since these are in long term memory and his short term memory retains goals and intentions longer than heeded information. There are also incompleteness problems. Thought in non-oral form is much faster, for example, the knowledge engineer cannot instruct a juggler to slow down!

The Specification of Expertise (Johnson, Zaulkernan & Garber, 1987) is one way of using the material gained from protocol analysis. It is a psychology based manual technique in which the expert thinks aloud while working. It encompasses a syntactic and semantic analysis of the protocol using operations, episodes, and data cues. Expertise is characterized by semantic categories, actions abilities, goals, conditions, strategies and solutions. It is based on the facts that thinking is not directly observable, especially that of expert with considerable compiled knowledge and practice and that expertise is characterized by generativity and is a kind of knowledge that is used to perform a task. Part of this knowledge is represented by a set of abilities for problem solving. The representation or specification of expertise consists of: possible solutions and components of solutions; relevant information or data needed; goals for solutions; permissible ways to move between intermediate solutions and the abilities needed to make a transition.

This psychologically-based method of knowledge acquisition makes assumptions about knowledge from an information processing point of view. The method assumes that an individual is able to describe the contents of his or her own short-term memory as if the individual had an internal observer watching information come and go from working memory. Ericsson and Simon (1984) demonstrate support for this method and its underlying assumptions about cognitive information processing but little is stated about the nature of knowledge apart from stating that perceptual information about environmental cues and internal cognitive processes are within the awareness of the 'observer'.

Elicitation takes the form of a variety of 'thinking-aloud' activities while completing the task of interest. As the expert thinks aloud, the entire linguistic material is recorded and usually

transcribed. The transcription can then be parsed or ordered for later coding or encoding. The conceptual model is rich and contains generated material as opposed to retrospective material (e.g. interviewing) and can be used to develop domain dictionaries and parsings based on discourse analysis procedures (e.g. Belkin, Brooks & Daniels, 1986).

Analysis procedures vary for protocol analysis because this method was developed as a generic elicitation strategy. Johnson, Zaulkernan and Garber, (1987) used a method of encoding and ordering of the protocol material. Categories such as cues, triggers, goals, activities, etc. were ordered in graphical form to give a well-defined, formally ordered model of the conceptual model. The KADS method also uses protocol analysis but uses a much different theoretical approach to analyzing the conceptual model.

A variety of coding methods are available but the choice depends on the analysis procedures used and the nature of the model of the conceptual model.

Domain Ontologies include the KADS methodology or Knowledge Acquisition Design System (Breuker & Wielinga, 1987) which deals with the whole process from domain organization to developing a complete system. It breaks the domain organization into stages where the problem is decomposed into component parts and the analysis of knowledge is separated from its machine implementation. Data is then collected and tested to modify the framework. The three refinement cycles are applied: orientation which includes the acquisition of vocabulary, the assessment of characteristics, and the selection of the model and a test for problem feasibility; problem identification in which structures of domain concepts are uncovered, the functional analysis of the prospective expert system and task analysis; and problem analysis where the analysis of the user and operational environment results in the basic architecture of the system. At this point knowledge primitives are identified such as objects, knowledge sources, structures, strategies and models. Interpretation models are produced using a top-down model-driven approach which may include notions of causality, time, space, change and task-dependent concepts.

Building interpretation models is done by knowledge identification and knowledge conceptualization where knowledge primitives are ordered using relationships such as is-a, part-of, causes, depends-on, consists-of, defines, and several types of analysis are performed. The system is then coded using guidelines laid down.

Underlying the KADS method of knowledge acquisition is a philosophically-based theory of knowledge. Epistemologically, knowledge is considered a multi-level phenomenon. Each level represents a different level of complexity and functionality (Hayward, Wielinga & Breuker 1986). They integrate the four degrees of depth at which questions about knowledge can be asked, that is, questions at the *individual, conceptual, formalism* and *mechanism* levels suggested by Sloman (1980), with the five types of knowledge *implementational, logical, epistomological, conceptual* and *linguistic* suggested by Brachman (1979). They identify the five levels as being "knowledge identification", "knowledge conceptualisation", "epistomological analysis", "logical analysis" and "implementational analysis". Knowledge cannot be considered uni-dimensional. This assumption indicates knowledge is not complete without taking into account all levels and their relationships.

Elicitation methods in KADS are designed to develop linguistically rich protocols of expert behaviour. Broader than just protocol analysis, the KADS method generates a variety of verbal protocols from specialized interviews, thinking-aloud instructions, introspection and user-dialogues. The resulting conceptual models are rich linguistic models which are parsed partially on the basis of syntactical position of the words to form domain vocabularies, diagrams, etc. and are sometimes used to identify the format of the final model of the conceptual model.

Analysis procedures (Breuker & Wielinga, 1987) follow closely the multi-level theory of knowledge. Objects, knowledge sources, tasks and strategies are all hierarchically ordered levels pulled out of the conceptual model and intermediate knowledge bases. The final model of the conceptual model is a rich and highly organized formal model called an interpretation model.

Coding procedures are represented by logic-based methods and structures. It is not yet clear how much of the richness of the model of the conceptual model survives the transformation into the computer knowledge base. Other coding methods may be possible with the rich model of the conceptual model but have not, as yet, been described.

Problem-Solving Ontologies are exemplified by *role-limiting methods* (McDermott, 1988) and *generic tasks* (Chandrasekaran, 1988), and reflect a somewhat different nature than most other methods in that they are highly structured and highly specific elicitation procedures based on a great deal of effort prior to elicitation. Both are based on the assumption that there exist repetitive processes of problem-solving used by the expert to complete a task, these methods concentrate first on identifying the repetitive roles of knowledge or those processes of knowledge use. The types of knowledge required are defined by the requirements of the processes.

Role-limiting methods strongly guide knowledge collection and encoding having a broad scope of applicability and providing help in specifying what knowledge needs to be collected to perform a particular task. A role-limiting method typically consists of a simple loop over a sequence of five to ten steps. For example, MOLE is a knowledge acquisition tool for generating expert systems that do heuristic classification. The method has the following control knowledge: determine the events that explain the symptoms; differentiate among several candidates by ruling out explanatory connections or candidate explanatory events, or by providing sufficient support for a candidate or reason for preferring some explanatory connections; get this information and apply it; at this stage, if new symptoms are revealed, go back and perform the sequence again.

Chandrasekaran is aiming to raise the level of knowledge structures from rule type systems to generic tasks which are problem-solving procedures for specific tasks. Elicitation is highly structured and specific in that an algorithm is defined for a problem and an automated elicitation mechanism is developed to structure the expert's verbalizations. If the problem-solving method is, for example, heuristic classification within a diagnostic task, then certain types of knowledge such as symptoms, hypotheses, and differentiation information are requested of the expert.

The conceptual model and the model of the conceptual model are identical in this method. Implementation methods are very specific due to the nature of the model of the conceptual model, and the computer knowledge bases which result are automatically built from the elicitation, analysis and implementation procedures.

Implicit Knowledge Structures are used in linguistics, and one method resulting from this approach is Cognosys (Woodward, 1988) which is mainly used in domain organization to structure the domain in order to isolate and organize key areas of content. The expert provides a description statement of overall domain and a list of general objectives which must be met in order to solve problems or make decisions successfully with the domain. He must then produce short paragraphs which are descriptive of the tasks required to meet each objective. These may be articles provided by the expert, such as case studies, text books or operating manuals. The knowledge engineer takes these and analyses the tasks into categories. He also completes the concept description categories and clarifies and elaborates such things as temporal relationships, event and state sequences, goal specification, cause-effect relationships, inferences, hierarchical relationships. It is important at this stage to clarify the domain boundaries with the expert and discuss specific problems in the domain. He may identify other sources of domain knowledge, before he identifies knowledge acquisition methodologies and tools, report formats and inference processes. A general knowledge structure is then produced by the knowledge engineer from

which goal structures, cause-effect structures, implicational structures and taxonomic structures are identified and isolated.

This linguistic-based method of knowledge acquisition is based on the theory of implicit knowledge structures which underlie general knowledge structures (Woodward, 1988). This method assumes that knowledge structures consist of discrete units which are linked and stored together in memory. When a word is chosen (in linguistic form), it is assumed that there exist a number of semantic meanings attached to the word. The word represents the tip of a meaning iceberg. This representation is not completely static but contains variable links which are partially sensitive to context. When words are chosen for textual material to describe, explain, or label symbolic entities, they can be used to identify the underlying or implicit knowledge structures.

This conceptual model is rich and requires further ordering but it is partially constrained by the objective-task breakdown. A variety of intermediate knowledge bases are possible. Lists of structural concepts (nouns) and functional concepts (verbs) result from the parsing of the textual material. Attributes describing each of the concepts are listed and more may be elicited. Contextual information about each concept can be stored.

The analysis procedures require the application of rules by the knowledge engineer who identifies noun-verb-noun phrases, assigns them a type and then links the phrases with a restricted set of arcs based on the meanings in the paragraphs (Woodward, 1988). These procedures follow directly from the theory of implicit structures and identify generic structures which define set-relations, goal sequences, implications, cause-effect relations and temporal sub-structures.

The set of sub-structures for each task and objective represents the model of the conceptual model. This formalized model then acts as the basis for semantic net, frame, or rule-based coding formalisms to generate a computer knowledge base.

Entity-Attribute Grids are derived from the repertory grids of personal construct theory (Kelly, 1955) which is one example of a modeling system that externalizes human processes. The personal construct theory model views the individual as an anticipatory system anticipating the future as a hypothesized replication of the past; constructs about the world change as a result of feedback about the predictions or actions on it. This is a systemic theory of human cognition based on the single primitive of a construct, or dichotomous distinction. Constructs are:

> "transparent templets which [a person] creates and then attempts to fit over the realities of which the world is composed." (Kelly 1955)

He proposes that all of human activity can be seen as a process of anticipating the future by construing the replication of events:

> "Constructs are used for predictions of things to come, and the world keeps rolling on and revealing these predictions to be either correct or misleading. This fact provides a basis for the revision of constructs and, eventually, of whole construct systems." (Kelly 1955)

Hence his psychological model of man is strongly epistemological and concerned with the way in which man models his experience and uses this model to anticipate the future. The anticipation may be passive, as in prediction, or active, as in action. Kelly developed his theory in the context of clinical psychology and hence was concerned to have techniques which used it to by-pass cognitive defenses and elicit the construct systems underlying behavior. This is precisely the problem of knowledge engineering.

A repertory grid (Shaw 1980) is a two-way classification of data in which events are interlaced with abstractions in such a way as to express part of a person's system of cross-references between his personal observations or experience of the world (elements or entities), and his personal constructs or classifications of that experience (attributes). It is a way of representing

personal constructs as a set of distinctions made about entities relevant to the problem domain. In the development of expert systems the entities might be key elements in the problem domain such as oil-well sites or business transactions, and the attributes express what the particular expert sees as the crucial distinctions between the entities in the context of the problem.

This psychologically-based method of knowledge acquisition assumes that the mental model is partially characterized by distinctions made about the target system by the expert and used to help make predictions about the world. Mental categories are present in which objects act as instances of those categories. These distinctions can be represented by bi-polar relations and objects in the world can be distinguished on the basis of the bi-polar relations.

The entity-attribute grid elicitation procedures follow very directly from the assumptions. Within a given situational or contextual system, objects are elicited which have something to do with the context. Distinctions are elicited by offering triads of objects and asking why two of the objects are the same but different from the third. The distinctions are defined by bipolar descriptions and the remaining objects are ranked upon the bi-polar dimensions depending on the extent to which they are characterized by the distinctions at one or the other end of the dimension. An assumption made by the procedure is that all necessary objects and distinctions are elicited in this manner.

The result of this activity is a very clean conceptual model but one which is limited in linguistic material. The grid does not capture all that occurs or all the knowledge that is used when completing the elicitation exercise. If one were to tape record a session while a grid was elicited while the expert worked under thinking aloud instructions, a fuller conceptual model would be obtained which could include explanations for why choices were made, the decision-making verbalizations and the descriptions of the new understanding acquired by the expert. An intermediate knowledge base becomes available from the elicitation. Any grid which results is part of the conceptual model and may act as one possible intermediate knowledge base in that it reflects a partial ordering of the material in the complete conceptual model.

A variety of analyses for conceptual clustering, and visualization of the domain can be completed using the distance-based ratings in the grid (Shaw, 1980; Shaw & Gaines, 1986). These will aid the expert in reviewing and defining the domain structure in his conceptual model. At a more formal stage of analysis classes, objects etc. can be produced using empirical induction.

Empirical Induction is basically a method that can be applied to a set of rules (or equivalent) from a set of cases (Forsyth & Rada, 1986). It is necessary to have a training set of critical (or other — see Gaines, 1989) cases which encompass crucial and complete infomation. The cases are distinguished by a set of attributes, in a similar way to the repertory grid attributes. The decision of the expert in each of the cases is also needed and an induction algorithm such as ID3 (Quinlan, 1987). ID3, or Interactive Dichotomizer 3, works the best if the data is not corrupted by 'noise'. A subset of a certain size (window) is selected from the training set and an algorithm applied to derive a rule for this window. All exceptions to this rule are noted and the algorithm re-applied with these exceptions until none are found. When the whole set of rules are found they are formed into a decision tree.

Induce (Dietterich & Michalsky, 1981) is another induction algoritmn based on earlier experience with a system called AQ11. A search technique is used to explore the set of examples and generalize very specific rules formed from an initial subset of the training set. There are many other induction systems notable Prism (Cendrowska, 1987) and Induct (Gaines, 1989). In general, the quality of the induced results depends on the quality of the initial training set which must be drawn up by an expert. The resulting rules are usually tested on a quite separate set of examples than the training set, and the final set of rules discussed with the expert (Hart, 1987).

Gaines (1989) reports a study of the tradeoff between the knowledge needed for expertise transfer and the cases required for the induction training set. The results indicate that a mixture of correct rules, adequate rules, a critical set of cases described in terms of a set of relevant (or including some irrelevant) attributes, and even some incorrect decisions can produce the same minimal set of correct decision rules. In other words, the lack of a first class expert who can abstract the essential knowledge from his model, and represent it clearly in an intermediate knowledge base can be offset by the analysis process which produces the computer knowledge base. There is scope for further exploration of this insight into the relationship between empirical induction and expertise transfer.

CONCLUSION

In this paper we have identified one of the major problems in knowledge acquisition as the lack of an overall methodological framework which logically supports the process of knowledge acquisition, identifies the characteristics of the products of knowledge acquisition, and offers a set of ordered assumptions about the nature of knowledge and its characteristics. We have shown the modeling processes involved from the perspective of the expert trying to communicate his view of the target system and transfer it into computer implementable form. We have identified the distinct proccesses of elicitation, analysis and implementation, but have not gone beyond this stage of development of a knowledge-based system. Finally, we have identified the separate knowledge representations of the intermediate knowledge bases which can be used to help the expert review and refine his conceptual model, and the computer knowledge bases which may be unrecognizable by the expert as related to his developing models.

Many manual and automatic elicitation methods exist. Some are well-founded in basic underlying theories from established disciplines, but the majority are based on *ad hoc* methods that seem to work. What has to be established by the advocates of these various methods is an underlying justification, not only from another discipline, but also from a well-established and tested discipline such as psychology, linguistics, philosophy or systems theory. Any less well-established theories need, in turn, to be grounded in a basic discipline.

Shaw and Gaines (1987) list requirements for a computer-based knowledge support system which will support the expert in his or her endeavours to externalize problem-solving behaviors:

1. The knowledge support system should be domain independent;
2. The knowledge support system should be directly applicable by experts without intermediaries;
3. The knowledge support system should be able to access a diversity of knowledge sources including text, interviews with experts, and observations of expert behavior;
4. The knowledge support system system should be able to encompass a diversity of perspectives including partial or contradictory input from different experts;
5. The knowledge support system should be able to encompass a diversity of forms of knowledge and relationships between knowledge;
6. The knowledge support system should be able to present knowledge from a diversity of sources with clarity as to its derivation, consequences and structural relations;
7. Users of the knowledge support system should be able to apply the knowledge in a variety of familiar domains and freely experiment with its implications;
8. The knowledge support system should make provision for validation studies;
9. As much of the operation of the knowledge support system as possible should be founded on well-developed and explicit theories of knowledge acquisition, elicitation and representation;
10. As the overall knowledge support system develops it should converge to an integrated system.

These requirements provide a first checklist for developers, users and supporters of existing systems and techniques. In addition, we would include support of the analysis phase of

knowledge acquisition, in such a way that the conceptual model of the expert is preserved as much as possible, and the analysis results in knowledge representation schemes that can be adequately applied in computerized knowledge bases. The above requirements are a start in aiding people to assess more completely the validity or otherwise of assumptions made about the type of knowledge which is being elicited by the various tools and techniques. Clearly, they must be designed to support the expert as much as possible and to help him or her in preserving an existing series of models that lead to effective interaction with the target system.

ACKNOWLEDGEMENTS

Financial assistance for this work has been made available by the Natural Sciences and Engineering Research Council of Canada. We are grateful to many colleagues for discussions which have highlighted many of the issues raised in this paper, and in particular to Brian Gaines for his critical comments and stimulating discussions.

REFERENCES

Addis, T. (1989). The science of knowledge: a research program for knowledge engineering. **Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW'89)**. Paris, France (June).

Anderson, J.R. (1983) **The Architecture of Cognition.** Harvard University Press: Cambridge, Mass.

Bainbridge, L. (1986). Asking questions and accessing knowledge. **Future Computing Systems** 1(2), 143-149.

Belkin, N.J., Brooks, H.M. & Daniels, P.J. (1986) Knowledge elicitation using discourse analysis. Boose, J.H. & Gaines, B.R. (Eds) **Proceedings of the First AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop**. pp.3-0-3-16. Banff (November).

Boose, J.H. & Bradshaw, J.M. (1987) Expertise transfer and complex problems: using AQUINAS as a knowledge acquisition workbench for knowledge-based systems. **International Journal of Man-Machine Studies 26,** 3-28.

Brachman, R.J. (1979) On the epistomological status of semantic networks. N.V. Findler (Ed) **Associative Networks,** Academic Press: New York..

Bresnan, J. (Ed) (1982) **The Mental Representation of Grammatical Relations.** MIT Press: Cambridge, Mass.

Breuker, J. & Wielinga, B. (1987) Use of models in the interpretation of verbal data. Kidd, A. (Ed) **Knowledge Elicitation for Expert Systems: A Practical Handbook.** Plenum.

Cendrowska, J. (1987) An algorithm for inducing modular rules. **International Journal of Man-Machine Studies 27** (4), 349-370 (October).

Chandrasekaran, B (1988) Generic tasks as building blocks for knowledge-based systems: the diagnosis and routine design examples. **The Knowledge Engineering Review, 3**(3), 183-211.

Chomsky, N. (1975) **Logical Structure of Linguistic Theory.** New York: Plenum.

Churchman, C.W. (1971) **Design of Inquiring Systems.** New York: Basic Books.

Clancey, W. (1989) Viewing knowledge bases as qualitative models. **IEEE Expert,** 4(2), 9-23.

Compton, P. & Jansen, R. (1989). A philosophical basis for knowledge acquisition. **Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW'89)**. Paris, France (June).

Craik, K.J.W. (1943) **The Nature of Explanation.** Cambridge University Press.

Dietterich, T. & Michalsky, R. (1981). Inductive Learning of Structural Descriptions, **Artificial Intelligence, 16.**

Ericsson, K.A. & Simon, H.A. (1984). **Protocol Analysis: Verbal Reports as Data**. Cambridge, Mass. MIT Press.

Fodor, J.A. (1980) **Representations**. Cambridge Mass.: MIT press.

Forsyth, R. & Rada, R. (1986). **Machine Learning Applications in Expert Systems and Information Retrieval,** Ellis Horwood: Chichester.

Gaines, B.R. (1987) Advanced expert system support environments. Boose, J.H. & Gaines, B.R. (Eds) **Proceedings of the Second AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop.** pp.8-0-8-14. Banff (October).

Gaines, B.R. (1989) An Ounce of Knowledge is Worth a Ton of Data: Quantitative Studies of the Trade-Off between Expertise and Data based on Statistically Well-Founded Empirical Induction. **Proceedings of 6th International Workshop on Machine Learning,** pp.156-159. San Mateo, California: Morgan Kaufmann (June).

Gaines, B.R. & Shaw, M.L.G. (1987). Knowledge support systems. **ACM MCC-University Research Symposium.** Austin, Texas: MCC. pp.47-66.

Gazdar G. (1981). On syntactic categories. **Philosophical Transactions of the Royal Society of London,** Series B, **295,** pp 267-83.

Graesser A.C. & Clark, L.F. (1985). **Structures and Procedures of Implicit Knowledge.** New Jersey: Ablex.

Gilhooly K.J. (1987). Mental modeling: a framework for the study of thinking. D.N. Perkins, J. Lochhead & J.C. Bishop (Eds) **Thinking: the Second International Conference.** NJ: Erlbaum.

Harrison, A. & Bramson, (1982) R.M. **Styles of Thinking.** New York: Doubleday.

Hart, A. (1987). Role of induction in knowledge elicitation. A.L. Kidd (Ed) **Knowledge Acquisition for Expert Systems: A practical handbook.** New York: Plenum.

Hayward, S.A., Wielinga, B.J. & Breuker, J.A. (1986) Structured analysis of knowledge. Boose, J.H. & Gaines, B.R. (Eds) **Proceedings of the First AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop.** pp.18-0-18-6. Banff.

Hirst, G. & Regoczei, S. (1989). On 'extracting knowledge from text': modelling the architecture of language users. **Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW'89).** Paris, France (June).

Johnson L. & Johnson, N.E. (1987) Knowledge elicitation involving teachback interviewing. Kidd, A. (Ed) **Knowledge Elicitation for Expert Systems: A Practical Handbook.** Plenum Press.

Johnson, P.E. (1986) Cognitive models of expertise. **Symposium on Expert Systems and Auditor Judgment.** University of Southern California.

Johnson, P.E., Zualkernan & Garber, S. (1987). Specification of expertise: **International Journal of Man-Machine Studies, 27**(2), 161-181 (February).

Johnson-Laird, P.N. (1983) **Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness.** Cambridge University Press.

LaFrance, M. (1987) The knowledge acquisition grid: a method for training knowledge engineers. **International Journal of Man-Machine Studies, 27**(2), 245-255 (February).

Gentner, D., & Gentner, A.L. (1983) Flowing waters or teeming crowds: mental models of electricity. D. Gentner & A. Stevens, (Eds.), **Mental Models.** Hillsdale, N.J.: Erlbaum.

Kelly, G.A. (1955). **The Psychology of Personal Constructs.** Norton: New York.

McDermott, J. (1988) Preliminary steps toward a taxonomy of problem solving methods. Marcus, S. (Ed) **Automating Knowledge Acquisition for Expert Systems,** Kluwer: Boston.

McGraw, K.L., & Harbison-Briggs, K. (1989). **Knowledge Acquisition Principles and Guidelines,** Prentice Hall: New Jersey.

Minsky, M. (1975) A framework for representing knowledge. P.H. Winston (Ed.), **The Psychology of Computer Vision.** New York: McGraw Hill.

Norman, D.A. (1983) Some observations on mental models. D. Gentner & A. Stevens, (Eds.), **Mental Models.** Hillsdale, N.J.: Erlbaum.

Nystedt, L. (1983) The situation: a constructivist approach. J.A. Adams-Webber & J.C. Mancuso (Eds) **Applications of Personal Construct Theory.** New York: Academic Press.

Paivio, A.U. (1977) Images, propositions and knowledge, J.M. Nicholas, (Ed.) **Images, Perception and Knowledge.** Reidel.

Pask, G. (1975). **Conversation, Cognition and Learning,** Elsevier, Amsterdam.
Pylyshyn, Z.W. (1984) **Computation and Cognition.** Cambridge, Mass.: MIT Press.
Quinlan, J.R. (1987) Simplifying decision trees. **International Journal of Man-Machine Studies 27** (3), 221-234 (September).
Schank, R.C. & Abelson, R. (1977) **Scripts, Plans, Goals, and Understanding.** Hillsdale, NJ: Erlbaum.
Schank, R.C. (1982) Reminding and memory organization: an introduction to MOPs. In W.G. Lehnert & M.H. Ringle (Eds), **Strategies of Natural Language Comprehension.** Hillsdale, NJ: Erlbaum.
Shaw, M.L.G. (1980) **On Becoming a Personal Scientist.** London: Academic Press.
Shaw, M.L.G. & Gaines, B.R. (1986) Interactive elicitation of knowledge from experts. **Future Computing Systems, 1**(2) 151-190.
Shaw, M.L.G. & Gaines, B.R. (1987) KITTEN: Knowledge Initiation and Transfer Tools for Experts and Novices. **International Journal of Man-Machine Studies 27**(3), 251-280.
Sloman, A. (1980) **The Computer Revolution in Philosophy,** Harvester.
Twine, S. (1989). A model for the knowledge analysis process. **Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW'89).** Paris, France (June).
Welbank, M. (1987) Knowledge acquisition update: Insight study #5. **Systems Designers,** Pembroke Broadway, Camberley, Surrey, UK.
Winston, P.H. (1984) **Artificial Intelligence.** (2nd ed). Reading. Prentice-Hall.
Woodward, B. (1988) Knowledge engineering at the front-end: defining the domain. Boose, J.H. & Gaines, B.R. (Eds) **Proceedings of the Third AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop.** pp.37-1-37-25. Banff (November).
Woodward, B. (1989) Integrating task demands and problem-solving methods to develop useful taxonomies for knowledge acquisition. **Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW'89).** Paris, France (June).