Mixed Depth Representations in the B2 System

Susan W. McRoy[1,2], Susan M. Haller[3], and Syed S. Ali[1]
[1]Department of Mathematical Sciences
[2]Department of Electrical Engineering and Computer Science
University of Wisconsin–Milwaukee
Milwaukee, WI 53201

[3]Computer Science and Engineering Department
University of Wisconsin–Parkside
Kenosha, WI 53141

Introduction

This paper gives an overview of our tutoring system, B2. It describes the natural language and knowledge representation components of B2, and our approach to the representation of questions and requests. The domain that we have developed most thoroughly helps medical students learn a statistical model for medical diagnosis. Many of the examples will be taken from this domain. According to B2's plans for tutoring, the system does this by generating story problems that describe a scenario and then asking the student about conclusions that might be drawn. B2 also supports requests to explain its reasoning and questions about facts.

The B2 Architecture

The B2 system consists of seven components (see Figure 1). In the diagram, solid, directed arrows indicate the direction of information flow between components. The system gets the user's input using a graphical user interface that supports both natural language interaction and mouse inputs. The *Parser* component of the *Parser/Generator* performs the first level of processing on the user input using its grammar and the domain information from the *Knowledge Representation Component.* The Parser interprets the user's inputs to form propositional representations of surface-level utterances for the Discourse Analyzer. The *Generator* produces natural language outputs from the text messages (propositional descriptions of text) that it receives from the *Discourse Planner.*

The system as a whole is controlled by a module called the *Discourse Analyzer.* The Discourse Analyzer determines an appropriate response to the user's actions on the basis of a model of the discourse and a model of the domain, stored in the knowledge representation component. The Analyzer invokes the *Discourse Planner* to select the content of the response and to structure it. The Analyzer relies on a component called the *Mediator* to interact with the Bayesian network processor. This Mediator processes domain level information, such as ranking the effectiveness of alternative diagnostic tests. The Mediator also handles the information interchange between the propositional information that is used by the Analyzer and the probabilistic data that is used by the Bayesian network processor. All phases of this process are recorded in the knowledge representation component, resulting in a complete history of the discourse. Thus, the knowledge representation component serves as a central "blackboard" for all other components.

During the initialization of the system, there is a one-time transfer of information from a file that contains a specification of the Bayesian network both to the Bayesian network processor and to the Knowledge Representation Component. The *Mediator* converts the specification into a propositional representation that captures the connectivity of the original Bayesian network.

The Knowledge Representation Blackboard

B2 represents both domain knowledge and discourse knowledge in a uniform framework as a propositional semantic network. A propositional semantic network is a framework for representing the concepts of a cognitive agent who is capable of using language (hence the term *semantic*). The information is represented as a graph composed of nodes and labeled directed arcs. In a ***propositional*** semantic network, the propositions are represented by the nodes, rather than the arcs; arcs represent only non-conceptual binary relations between nodes. The particular systems that are being used for B2 are SNePS and ANALOG (Ali 1994a; 1994b; Shapiro & Group 1992) which provide facilities for building and finding node as well as for reasoning

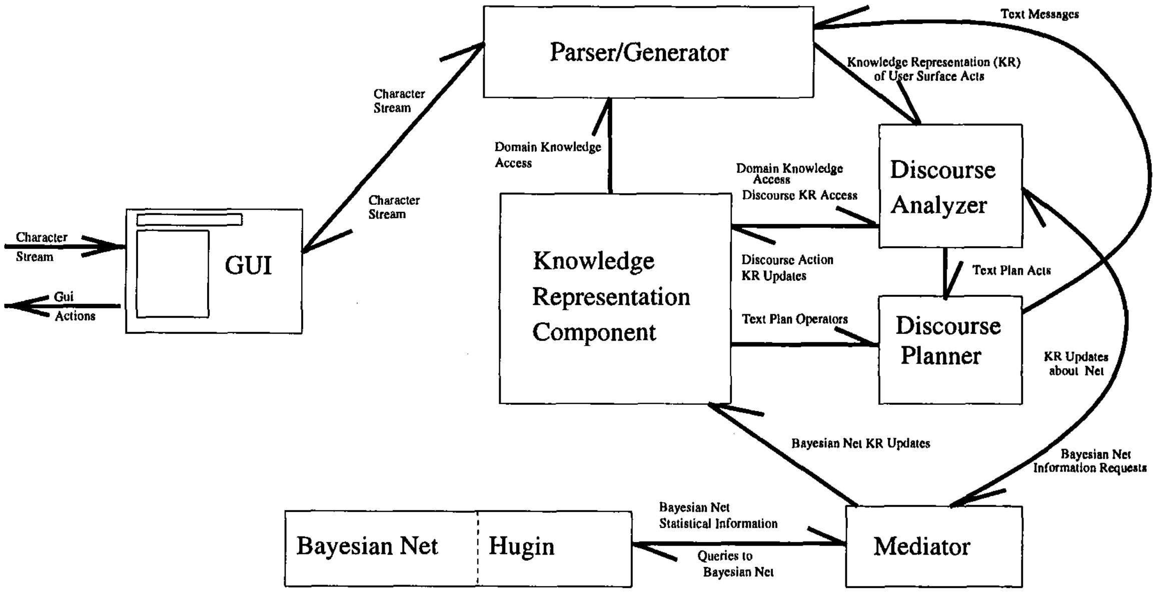

Figure 1: The B2 architecture

and truth-maintenance. These systems satisfy the following additional constraints:

1. Each node represents a unique concept.
2. Each concept represented in the network is represented by a unique node.
3. The knowledge represented about each concept is represented by the structure of the *entire network* connected to the node that represents that concept.

These constraints allow efficient inference when processing natural language. For example, such networks can represent complex descriptions (common in the medical domain), and can support the resolution of ellipsis and anaphora, as well as general reasoning tasks such as subsumption (Ali 1994a; 1994b; Maida & Shapiro 1982; Shapiro & Rapaport 1987; 1992).

We characterize a knowledge representation as *uniform* when it allows the representation of different kinds of knowledge in the same knowledge base using the same inference processes. The knowledge representation component of B2 is uniform because it provides a representation of the discourse knowledge, domain knowledge, and probabilistic knowledge (from the Bayesian net). This supports intertask communication and cooperation for interactive processing of tutorial dialogs.

The rule in Figure 2 is a good example of how the uniform representation of information in the semantic network allows us to relate domain information (a medical case) to discourse planning information (a plan to describe it). This network represents a text plan for describing a medical case to the user. Text plans are represented as rules in the knowledge representation. Rules are general statements about objects in the domain; they are represented by using *case frames*[1] that have `FORALL` or `EXISTS` arcs to nodes that represent variables that are bound by these quantifier arcs. In Figure 2, node `M13` is a rule with three universally quantified variables (at the end of the `FORALL` arcs), an antecedent (at the end of the `ANT` arc), and a consequent (at the end of the `CQ` arc).

This means that if an instance of the antecedent is believed, then a suitably instantiated instance of the consequent is believed. `M13` states that if `V1` (which is at the end of the `CASE-NUMBER` arc) is the case number of a case, and `V2` and `V3` (which are at the end of `CASE-INFO` arcs) are two pieces of case information, then a plan to describe the case will conjoin[2] the two pieces of case information. Node `P1` represents the concept that something is a member of the

[1] Case frames are conventionally agreed upon sets of arcs emanating from a node that are used to express a proposition. For example, to express that A isa B we use the MEMBER-CLASS case frame which is a node with a MEMBER arc and a CLASS. arc (Shapiro *et al.* 1994) provides a dictionary of standard case frames. Additional case frames can be defined as needed.

[2] "Conjoin" is a technical term from Rhetorical Structure Theory (Mann & Thompson 1986); it refers to a co-ordinate conjunction of clauses.

class **case** and P2 represents the concept that the case concept has a case number and case information. For more details about the knowledge representation, see (McRoy, Haller, & Ali 1997).

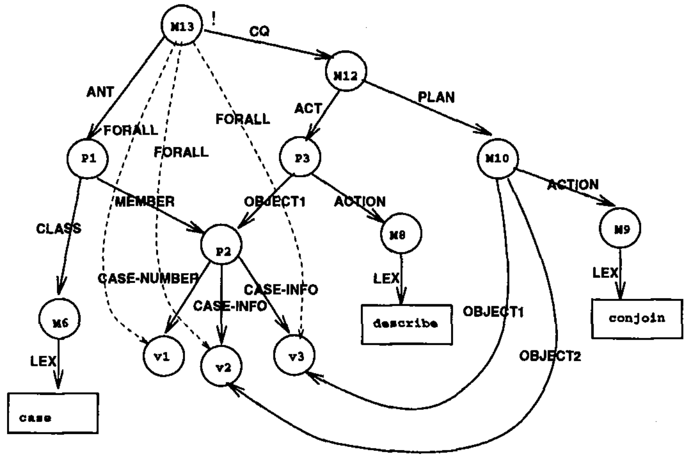

Figure 2: A rule stating that if V1 is the case number of a case, and V2 and V3 are two pieces of case information, then a plan for generating a description of the case will present the two pieces of information in a coordinating conjunction.

The Representation of the Discourse

The discourse model has five levels of representation, shown in Figure 3. These levels capture what the student and the system have each said, as well as how their utterances extend the ongoing discourse. Unlike many systems, B2's model of discourse will include a representation of questions and requests, as well as statements of fact. (Systems that do not represent questions and requests typically give these utterances a procedural semantics, interpreting them as operations to be performed.) Having an explicit representation of questions and requests simplifies the interpretation of context-dependent utterances such as *Why?* or *What about HIDA?* (Haller 1996)[3] It also allows the system to recover from misunderstandings, should they occur (McRoy 1995; McRoy & Hirst 1995).

interpretation of exchanges
exchanges (pairs of interpretations)
system's interpretation of each utterance
sequence of utterances
utterance level

Figure 3: Five Levels of Representation

We will consider each of these levels in turn, starting with the utterance level, shown at the bottom of Figure 3.

The Utterance Level

For all inputs, the parser produces a representation of of its surface content, which the analyzer will assert as part of an occurrence of an event of type SAY. The content of the user's utterance is always represented by what she said literally. In the case of requests, the student may request a story problem directly, as an imperative sentence *Tell me a story* or indirectly, as a declarative sentence that expresses a desire *I want you to tell me a story.* The complete representation of the imperative sentence *Tell me a story* is shown in Figure 4.

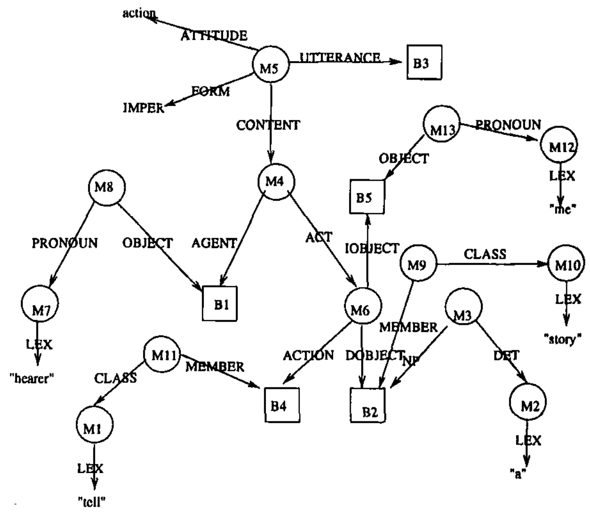

Figure 4: Node B3 represents an utterance whose form is imperative, and whose content (M4) is the proposition that the hearer (B1) will tell a story (B2) to the speaker (B5).

For the system's utterances, the utterance level representation corresponds to a text generation event (this contains much more fine-grained information about the system's utterance, such as mode and tense.) The content of the system's utterance is the text message that is sent to the language generator.

[3]HIDA stands for radio-nuclide hepatobilary imaging, a diagonistic test.

Sequence of Utterances

The second level corresponds to the sequence of utterances. (This level is comparable to the linguistic structure in the tripartite model of (Grosz & Sidner 1986)). In the semantic network, we represent the sequencing of utterances explicitly, with asserted propositions that use the BEFORE-AFTER case frame. The order in which utterances occurred (system and user) can be determined by traversing these structures. This representation is discussed in detail in (McRoy, Haller, & Ali 1997).

The Interpretation Level

In the third level, we represent the system's interpretation of each utterance. Each utterance event (from level 1) will have an associated system interpretation, which is represented using the INTERPRETATION_OF—INTERPRETATION case frame. For example, consider the interpretation of the utterance *Tell me a story* (as well as *I want you to tell me a story.*), shown in Figure 5. (Every utterance has one or more interpretations; at any time, only one is believed and a justification-based truth maintenance system is used to track changes in belief.)

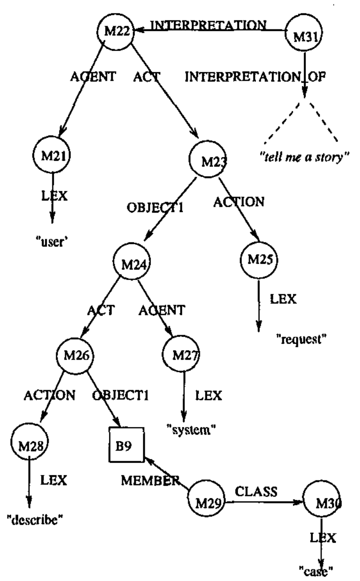

Figure 5: Node `M31` is a proposition that the interpretation of *Tell me a story* (which is glossed in this figure) is `M22`. Node `M22` is the proposition that the user requested that the system describe a case to the user. (Describing a case is a domain-specific action; the pronouns from the utterance level have been interpreted according to the context.)

The Exchange and Exchange Interpretation Levels

The fourth and fifth levels of representation in our discourse model are exchanges and interpretations of exchanges, respectively. A *conversational exchange* is a pair of interpreted events that fit one of the conventional structures for dialog (*e.g.* QUESTION–ANSWER). Figure 6 gives the network representation of a conversational exchange and its interpretation. Node `M113` represents the exchange in which the system has asked a question and the user has answered it. Using the MEMBER-CLASS case frame, propositional node `M115` asserts that the node `M113` is an exchange. Propositional node `M112` represents the system's interpretation of this exchange: that the user has accepted the system's question (*i.e.* that the user has understood the question and requires no further clarification). Finally, propositional node `M116` represents the system's belief that node `M112` is the interpretation of the exchange represented by node `M113`.

Interaction among the Levels

A major advantage of the network representation is the knowledge sharing between these five levels. We term this knowledge sharing ***associativity.*** This occurs because the representation is uniform and every concept is represented by a unique node (see Section). As a result, we can retrieve and make use of information that is represented in the network implicitly, by the arcs that connect propositional nodes. For example, if the system needed to explain why the user had said HIDA, it could follow the arcs from the node representing the utterance that *User said HIDA* to the system's interpretation of that utterance, node `M108`, to determine that

- The user's utterance was understood as the answer within an exchange (node `M113`), and
- The user's answer indicated her acceptance and understanding of the discourse, up to that point `M112`.

This same representation could be used to explain why the system believed that the user had understood the system's question. This associativity in the network is vital if the interaction starts to fail.

Summary

The goal of the B2 project is to give students an opportunity to practice their decision making skills. We give students the opportunity to ask the system to explain what factors were most influential to its decision and why.

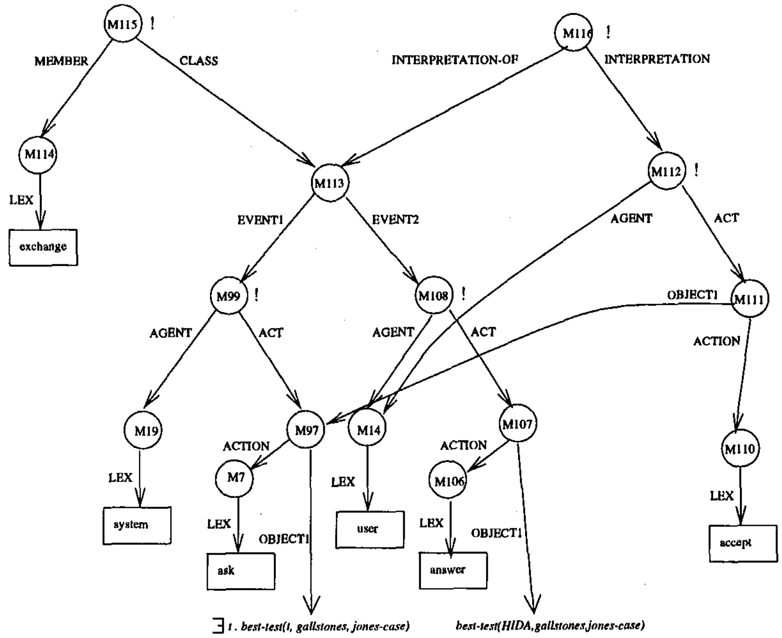

Figure 6: Node M115 represents the proposition that node M113 is an exchange comprised of the events M99 and M108. M108 is the proposition that *The user answered "HIDA is the best test to rule in Gallstones"*. Additionally, node M116 represents the proposition that the interpretation of M113 is event M112. M112 is the proposition that the user has accepted M96. (M96 is the question that the system asked in event M99.)

The natural language processing and knowledge representation components of B2 are general purpose. It builds a five-level model of the discourse, that represents what was literally said, what was meant, and how each utterance and its interpretation relates to previous ones. This is necessary because students' utterances may be short and ambiguous, requiring extensive reasoning about the domain or the discourse model to fully resolve. We have shown how our mixed-depth representations encode syntactic and conceptual information in the same structure. This allows us to defer any extensive reasoning until needed, rather than when parsing. We use the same representation framework to produce a detailed representation of requests and to produce a representation of questions. The representations use the same knowledge representation framework that is used to reason about discourse processing and domain information—so that the system can reason with (and about) the utterances.

Acknowledgements

This work was supported by the National Science Foundation, under grants IRI-9701617 and IRI-9523666 and by a gift from the University of Wisconsin Medical School, Department of Medicine.

References

Ali, S. S. 1994a. A Logical Language for Natural Language Processing. In *Proceedings of the 10th Biennial Canadian Artificial Intelligence Conference*, 187–196.

Ali, S. S. 1994b. *A "Natural Logic" for Natural Language Processing and Knowledge Representation.* Ph.D. Dissertation, State University of New York at Buffalo, Computer Science.

Grosz, B. J., and Sidner, C. L. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12.

Haller, S. 1996. Planning text about plans interactively. *International Journal of Expert Systems* 85—112.

Maida, A. S., and Shapiro, S. C. 1982. Intensional concepts in propositional semantic networks. *Cognitive Science* 6(4):291–330. Reprinted in R. J. Brachman and H. J. Levesque, eds. *Readings in Knowledge Representation*, Morgan Kaufmann, Los Altos, CA, 1985, 170–189.

Mann, W., and Thompson, S. 1986. Rhetorical structure theory: Description and construction of text structures. In Kempen, G., ed., *Natural Lan-*

guage Generation. Boston: Kluwer Academic Publishers. 279–300.

McRoy, S. W., and Hirst, G. 1995. The repair of speech act misunderstandings by abductive inference. *Computational Linguistics* 21(4):435–478.

McRoy, S.; Haller, S.; and Ali, S. 1997. Uniform knowledge representation for nlp in the b2 system. *Journal of Natural Language Engineering* 3(2).

McRoy, S. W. 1995. Misunderstanding and the negotiation of meaning. *Knowledge-based Systems* 8(2–3):126–134.

Shapiro, S. C., and Group, T. S. I. 1992. *SNePS-2.1 User's Manual.* Department of Computer Science, SUNY at Buffalo.

Shapiro, S. C., and Rapaport, W. J. 1987. SNePS considered as a fully intensional propositional semantic network. In Cercone, N., and McCalla, G., eds., *The Knowledge Frontier.* New York: Springer–Verlag. 263–315.

Shapiro, S. C., and Rapaport, W. J. 1992. The SNePS family. *Computers & Mathematics with Applications* 23(2–5).

Shapiro, S.; Rapaport, W.; Cho, S.-H.; Choi, J.; Feit, E.; Haller, S.; Kankiewicz, J.; and Kumar, D. 1994. A dictionary of SNePS case frames.