

MPLUS: A Probabilistic Medical Language Understanding System

Lee M. Christensen, Peter J. Haug, and Marcelo Fiszman

Department of Medical Informatics, LDS Hospital/University of Utah, Salt Lake City, UT

E-mail: ldlchris@ihc.com, ldphaug@ihc.com, ldmfiszm@ihc.com

Abstract

This paper describes the basic philosophy and implementation of MPLUS (M+), a robust medical text analysis tool that uses a semantic model based on Bayesian Networks (BNs). BNs provide a concise and useful formalism for representing semantic patterns in medical text, and for recognizing and reasoning over those patterns. BNs are noise-tolerant, and facilitate the training of M+.

1 Introduction

In the field of medical informatics, computerized tools are being developed that depend on databases of clinical information. These include alerting systems for improved patient care, data mining systems for quality assurance and research, and diagnostic systems for more complex medical decision support. These systems require data that is appropriately structured and coded. Since a large portion of the information stored in patient databases is in the form of free text, manually coding this information in a format accessible to these tools can be time consuming and expensive. In recent years, natural language processing (NLP) methodologies have been studied as a means of automating this task. There have been many projects involving automated medical language analysis, including deciphering pathology reports (Smart and Roux, 1995), physical exam findings (Lin et al., 1991), and radiology reports (Friedman et al., 1994; Ranum, 1989; Koehler, 1998).

M+ is the latest in a line of NLP tools developed at LDS Hospital in Salt Lake City, Utah. Its predecessors include SPRUS (Ranum, 1989) and SymText (Koehler, 1998). These tools have been used in the realm of radiology reports, admitting diagnoses (Haug et al., 1997), radiology utilization review (Fiszman, 2002) and syndromic detection (Chapman et al., 2002). Some of the character of these tools derives from common characteristics of radiology reports, their initial target domain.

Because of the off-the-cuff nature of radiology dictation, a report will frequently contain text that is telegraphic or otherwise not well formed grammatically. Our desire was not only to take advantage of phrasal structure to discover semantic patterns in text, but also to be able to infer those patterns from lexical and contextual cues when necessary.

Most NLP systems capable of semantic analysis employ representational formalisms with ties to classical logic, including semantic grammars (Friedman et al., 1994), unification-based semantics (Moore, 1989), and description logics (Romacker and Hahn, 2000). M+ and its predecessors employ Bayesian Networks (Pearl, 1988), a methodology outside this tradition. This study discusses the philosophy and implementation of M+, and attempts to show how Bayesian Networks can be useful in medical text analysis.

2 The M+ Semantic Model

2.1 Semantic Bayesian Networks

M+ uses Bayesian Networks (BNs) to represent the basic semantic types and relations within a medical domain such as chest radiology reports. M+ BNs are structurally similar to semantic networks, in that they are implemented as directed acyclic graphs, with nodes representing word and concept types, and links representing relations between those types. BNs also have a character as frames or slot-filler representations (Minsky, 1975). Each node is treated as a variable, with an associated list of possible values. For instance a node representing "disease severity" might include the possible values {"severe", "moderate", "mild"}. Each value has a probability, either assigned or inferred, of being the true value of that node.

In addition to providing a framework for representation, a BN is also a probabilistic inference engine. The probability of each possible value of a node is conditioned on the probabilities of the values of neighboring nodes,

through a training process that learns a Bayesian joint probability function from a set of training cases. After a BN is trained, a node can be assigned a value by setting the probability of that value to 1, and the probabilities of the alternate values to 0. This results in a cascading update of the value probabilities in all unassigned nodes, in effect predicting what the values of the unassigned nodes should be, given the initial assignments. The sum of the probabilities for the values of a given node is constrained to equal 1, making the values mutually exclusive, and reflecting uncertainty if more than one value has a nonzero probability. Please note that in this paper, "BN instance" refers to the state of a BN after assignments have been made.

A training case for a BN is a list of node / value assignments. For instance, consider a simple BN for chest anatomy phrases, as shown in Figure 1.

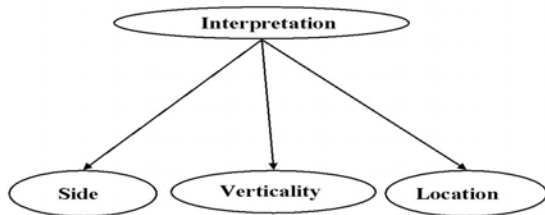


Figure 1. BN for simple chest anatomy phrases.

A training case for this BN applied to the phrase "right upper lobe" could be:

```
side=right
verticality=upper
location=lobe
interpretation= *right-upper-lobe
```

In the context of the Bayesian learning, this case has an effect similar to a production rule which states "If you find the words 'right', 'upper' and 'lobe' together in a phrase, infer the meaning *right-upper-lobe". After training on this case, assigning one or more values from this case would increase the probabilities of the other values; for instance assigning *side*= "right" would increase the probability of the value *interpretation*= *right-upper-lobe.

Interpretive concepts such as *right-upper-lobe are atomic symbols which are either invented by the human trainer, or else obtained from a medical knowledge database such as the

UMLS metathesaurus. By convention, concept names in M+ are preceded with an asterisk.

A medical domain is represented in M+ as a network of BNs, with word-level and lower concept-level BNs providing input to higher concept-level BNs. Figure 2 shows a partial view of the network of BNs used to model the M+ Head CT (Computerized Tomography) domain, instantiated with the phrase "temporal subdural hemorrhage". Each BN instance is shown with a list of nodes and most probable values. Note that input nodes of higher BNs in this model have the same name as, and take input from, the summary nodes of lower BNs. Word level BNs have input nodes named "head", "mod1" and "mod2", corresponding to the syntactic head and modifiers of a phrase. Each node in a BN has a distinguished "null" value, whose meaning is that no information relevant to that node, explicit or inferable, is present in the represented phrase.

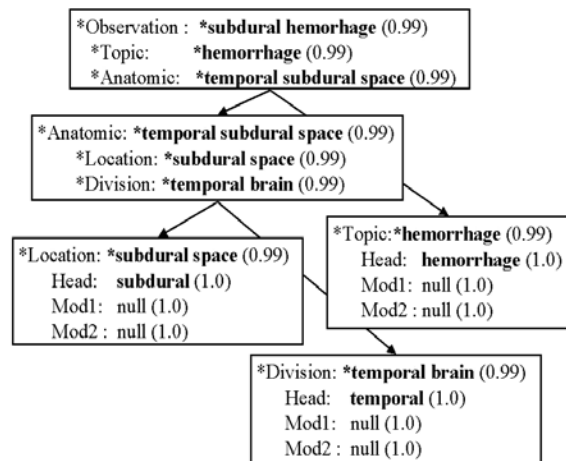


Figure 2. Network of M+ BNs, applied to "temporal subdural hemorrhage".

One way in which M+ differs from its predecessor SymText (Koehler, 1998) is in the size and modularity of its semantic BNs. The SymText BNs group observation and disease concepts together with state ("present", "absent"), change-of-state ("old", "chronic"), anatomic location and other concept types. M+ trades the inferential advantages of such monolithic BNs for the modularity and composability of smaller BNs such as those shown in figure 2. Figure 3 shows a single instance of the SymText Chest Radiology Findings BN, instantiated with the sentence "There is dense infiltrative opacity in the right upper lobe".

*observations : *localized upper lobe infiltrate (0.888)
 *state : *present (0.989)
 state term : null (0.966)
 *topic concept : *poorly-marginated opacity (0.877)
 topic term : opacity (1.0)
 topic modifier : infiltrative (1.0)
 *measurement concept : *null (0.999)
 measurement term : null (0.990)
 first value : null (0.998)
 second value : null (0.999)
 values link : null (0.999)
 size descriptor : null (0.999)
 *tissue concept : *lung parenchyma (0.906)
 tissue term : alveolar (1.0)
 *severity concept : *high severity (0.893)
 severity term : dense (1.0)
 *anatomic concept : *right upper lobe (0.999)
 *anatomic link concept : *involving (1.0)
 anatomic link term : in (1.0)
 anatomic location term : lobe (1.0)
 anatomic location modifier : null (0.999)
 anatomic modifier side : right (1.0)
 anatomic modifier superior/inferior : upper (1.0)
 anatomic modifier lateral/medial : null (0.999)
 anatomic modifier anterior/posterior : null (0.999)
 anatomic modifier central/peripheral : null (0.955)
 *change concept : *null (0.569)
 change with time : null (0.567)
 change degree : null (0.904)
 change quality : null (0.923)

Figure 3. SymText BN instantiation.

2.2 Parse-Driven BN Instantiation

M+ BNs are instantiated as part of the syntactic parse process. M+ syntactic and semantic analyses are interleaved, in contrast with NLP systems that perform semantic analysis after the parse has finished.

M+ uses a bottom-up chart parser, with a context free grammar (CFG). As a word such as "right" is recognized by the parser, a word-level phrase object is created and a BN instance containing the assignment *side*= "right" is attached to that phrase. As larger grammatical patterns are recognized, the BN instances attached to subphrases within those patterns are unified and attached to the new phrases, as described in section 3. The result of this process is a set of completed BN instances, as illustrated in figure 2. Each BN instance is a template containing word and concept-level value assignments, and the interpretive concepts inferred from those assignments. The templates themselves are nested in a symbolic expression, as described in section 2.3, to facilitate composing multiple BN instances in representations of arbitrary complexity.

Each phrase recognized by the parser is assigned a probability, based on a weighted sum of the joint probabilities of its associated BN instances, and adjusted for various syntactic and

semantic constraint violations. Phrases are processed in order of probability; thus the parse involves a semantically-guided best-first search.

Syntactic and semantic analysis in M+ are mutually constraining. If a grammatically possible phrase is uninterpretable, i.e. if its subphrase interpretations cannot be unified, it is rejected. If the interpretation has a low probability, the phrase is less likely to appear in the final parse tree. On the other hand, interpretations are constructed as phrases are recognized. The exception to this rule is when an ungrammatical fragment of text is encountered. M+ then uses a semantically-guided phrase repair procedure not described in this paper.

2.3 The M+ Abstract Semantic Language

The probabilistic reasoning afforded by BNs is superior to classical logic in important ways (Pearl, 1988). However, BNs are limited in expressive power relative to first-order logics (Koller and Pfeffer, 1997), and commercially available implementations lack the flexibility of symbolic languages. Friedman et al have made considerable headway in giving BNs many useful characteristics of first order languages, in what they call probabilistic relational models, or PRMs (e.g. Friedman et al. 1999).

While we are waiting for industry-standard PRMs, we have tried to make our semantic BNs more useful by combining them with a first-order language, called the M+ Abstract Semantic Language (ASL), implemented within M+. Specifically, BNs are treated as object types within the ASL. There is a "chest anatomy" type, for instance, and a "chest radiology findings" type, corresponding to BNs of those same names. The interpretation of a phrase is an expression in the ASL, containing predicates that state the relation of BN instances to one another, and to the phrase they describe. For instance, the interpretation of "hazy right lower lobe opacity" could be the expression

```
(and (head-of #phrase1 #find1)
      (located-at #find1 #loc1))
```

where #phrase1 identifies a syntactic phrase object, and #find1 and #loc1 are tokens representing instances of the findings BN (instanced with the words "hazy" and "opacity") and the anatomic BN (instanced with "right",

"lower" and "lobe"), respectively. The relation 'head-of' denotes that the findings BN is the main or "head" BN for that phrase. Conversely, "hazy right lower lobe opacity" can be thought of as a findings-type phrase, with an anatomic-type modifier.

This expression captures the abstract or "skeletal" structure of the interpretation, while the BN instances contain the details and specific inferences. One can think of the meaning of an expression like (located-at #find1 #loc1) in abstract terms, e.g. "some-finding located-at some-location". Alternatively, the meaning of a BN token might be thought of as the most probable interpretive concept within that BN instance. In this case, (located-at #find1 #loc1) could mean "*localized-infiltrate located-at *left-lower-lobe".

Because the object types in the ASL are the abstract concept types represented by the BNs, semantic rules formulated in this language constitute an "abstract semantic grammar" (ASG). The ASG recognizes patterns of semantic relations among the BNs, and supports analysis and inference based on those patterns. It also permits rule-based control over the creation, instantiation, and use of the BNs, including defining pathways for information sharing among BNs using virtual evidence (Pearl, 1988).

One use of the ASG is in post-parse processing of interpretations. After the M+ parser has constructed an interpretation, post-parse ASG productions may augment or alter this interpretation. One rule instructs *"If two pathological conditions exist in a 'consistent-with' relation, and the first condition has a state modifier (i.e. *present or *absent), and the second condition does not, apply the first condition's state to the second condition"*.

For instance, in the ambiguous sentence "There is no opacity consistent with pneumonia", if the parser doesn't correctly determine the scope of "no", it may produce the an interpretation in which *pneumonia lacks a state modifier, and is therefore inferred (by default) to be present. This rule correctly attaches (state-of *pneumonia *absent) to this interpretation.

One important consequence of the modularity of the M+ BNs, and of the ability to nest them within the ASL, is that M+ can

compose BN instances in expressions of arbitrary complexity. For instance, it is straightforward to represent the multiple anatomic concepts in the phrase "opacity in the inferior segment of the left upper lobe, adjacent to the heart":

```
(and (head-of #phrase1 #find1)
      (located-at #find1 #anat1)
      (qualified-by #anat1 #anat2)
      (adjacent-to #anat1 #anat3))
```

where the interpretive concepts of #anat1, #anat2 and #anat3 are *left-upper-lobe, *inferior-segment, and *heart, respectively.

The set of binary predicates that constitutes a phrase interpretation in M+ forms a directed acyclic graph; thus we can refer to the interpretation as an interpretation graph. The interpretation graph of a new phrase is formed by unifying the graphs of its subphrases, as described in section 3.

2.4 Advantages of Bayesian Networks

As mentioned, a BN training case bears a similarity to a production rule. It would be straightforward to implement the training cases as a set of rules, and apply them to text analysis using a deductive reasoning engine. However, Bayesian reasoning has important advantages over first order logic, including:

1- BNs are able to respond gracefully to input "noise". A semantic BN may produce reasonable inferences from phrasal patterns that only partially match any given training case, or that overlap different cases, or that contain words in an unexpected order. For instance, having trained on multi-word phrases containing "opacity", the single word "opacity" could raise the probabilities of several interpretations such as *localized-infiltrate and *parenchymal-abnormality, both of which are reasonable hypotheses for the underlying cause of opacity on a chest x-ray film.

2- Bayesian inference works bi-directionally; i.e. it is abductive as well as deductive. If instead of assigning word-level nodes, one assigns the value of the summary node, the probability of word values having a high correlation with that summary will increase. For instance, assigning the value *localized-infiltrate will raise the probability that the topic word is "opacity".

Bi-directional inference provides a means for modeling the effects of lexical context. A value assignment made to one word node can alter value probabilities at unassigned word nodes, in a path of inference that passes through the connecting concept nodes. For instance, if a BN were trained on "right upper lobe" and "left upper lobe", but had never seen the term "bilateral", applying the BN to the phrase "bilateral upper lobes" would increase the probabilities of both "left" and "right", suggesting that "bilateral" is semantically similar to "left" and "right". This is one approach to guessing the node assignments of unknown words, a step in the direction of automated learning of new training cases.

Similarly, if the system encounters a phrase with a misspelling such as "rht upper lobe", by noting the orthographic similarity of "rht" to "right" and the fact that "right" is highly predicted from surrounding words, it can determine that "rht" is a misspelling of "right". The spell checker currently used by M+ employs this technique.

3 Generating Interpretation Graphs

As mentioned, in M+ the interpretation graph of a phrase is created by unifying the graphs of its child phrases. High joint probabilities in the resulting BN instances are one source of evidence that the words thus brought together exist in the expected semantic pattern. However, corroborating evidence must be sought in the syntax of the text. Words which appear together in a training phrase may not be in that same relation in a given text. For instance, "no" and "pneumonia" support different conclusions in "no evidence of pneumonia" and "patient has pneumonia with no apparent complicating factors". M+ therefore only attempts to unify sub-interpretations that appear, on syntactic grounds, to be talking about the same things. This is less constraining than production rules that look for words in a specific order, but more constraining than simply pulling key words out of a string of text.

The following are examples of rules used to guide the unification of ASL interpretation graphs. For convenience, several shorthand functional notations are used: If P represents a phrase on the parse chart, root-bn(P) represents the root or head BN instance in

P's interpretation graph, and type-of(root-bn(P)) is the BN type of root-bn(P). If A and B are sibling child phrases of parent phrase C, then C = parent-phrase(A,B). Note that for convenience, BN instances in the interpretation graphs in Figures 4 - 6 are represented alternately as the words slotted in those instances, and as the most probable interpretive concepts inferred by those instances.

3.1 Same-type Unification

If phrase A syntactically modifies phrase B, then M+ assumes that some semantic relation exists between A and B. The nature of that relation is partly determinable from type-of(root-bn(A)) and type-of(root-bn(B)). If type-of(root-bn(A)) = type-of(root-bn(B)), that relation is simply one where root-bn(A) and root-bn(B) are partial descriptions of a single concept. If root-bn(A) and root-bn(B) are unifiable, M+ composes their input to form root-bn(parent-phrase(A,B)).

If in addition there are two unifiable same-type BN instances X and Y linked to root-bn(A) and root-bn(B) respectively, via arcs of the same name, then X and Y also describe a single concept, and the arcs describe a single relationship. For instance, if X and Y describe the anatomic locations of root-bn(A) and root-bn(B), and if root-bn(A) and root-bn(B) are partial descriptions of a single "finding", then X and Y are partial descriptions of a single anatomic location, and ought to be unified.

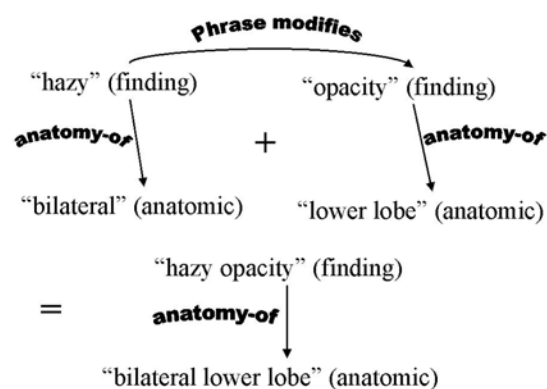


Figure 4: Same-type unification

In figure 4, in the Chest X-ray domain, the phrase "bilateral hazy lower lobe opacity" is interpreted by unifying the interpretations of its subphrases "bilateral hazy" and "lower lobe opacity". Note that without any corresponding syntactic transformation, this rule brings about a

"virtual transformation", whereby words are grouped together within BN instances in a manner that reflects the conceptual structure of the text. In this example "bilateral hazy lower lobe opacity" is treated as ("bilateral lower lobe") ("hazy opacity").

3.2 Different-type Unification

If phrase A syntactically modifies phrase B, and $\text{type-of}(\text{root-bn}(A)) \diamond \text{type-of}(\text{root-bn}(B))$, then $\text{root-bn}(A)$ and $\text{root-bn}(B)$ represent different concepts within some semantic relation. M+ uses the ASG to identify that relation and to add it to the interpretation graph in the form of a path of named arcs connecting $\text{root-bn}(A)$ and $\text{root-bn}(B)$. This path may include implicit connecting BN instances.

For instance, to interpret "subdural hemorrhage" in the Head CT domain, M+ attempts to unify the graphs for the subphrases "subdural" and "hemorrhage", where $\text{type-of}(\text{root-bn}(\text{"subdural"})) = \text{location}$, and $\text{type-of}(\text{root-bn}(\text{"hemorrhage"})) = \text{topic}$. M+ identifies the connecting path for these two types as shown in figure 2, and adds that path to the interpretation as shown in figure 5. Note that this path contains instances of the "observation" and "anatomy" BN types.

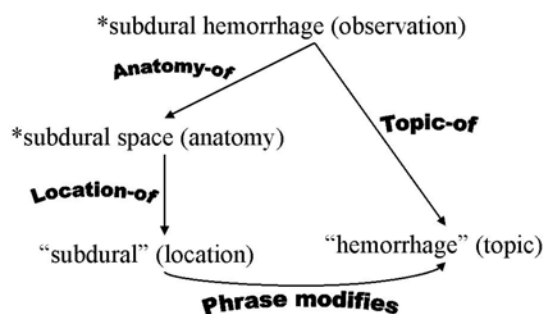


Figure 5. Different-type unification.

3.3 Grammar Rule Based Unification

Individual grammar rules in M+ can recognize semantic relations, and add connecting arcs to the interpretation graph. For instance, M+ has a rule which recognizes findings-type phrases connected with strings of the "suggesting" variety, and connects their graphs with a 'consistent-with' arc. This is used to interpret "opacity suggesting possible infarct" in the Head CT domain, as shown in figure 6.

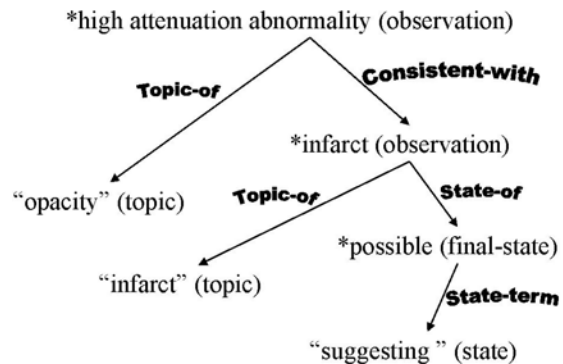


Figure 6: Grammar rule - based unification.

4 M+ Implementation

M+ is written in Common Lisp, with some C routines for BN access. The M+ architecture consists of six basic components: The parser, concept space, rule base, lexicon, ASL inference engine, and Bayesian network component.

As mentioned, the parser is an implementation of a bottom up chart parser with context free grammar.

The concept space is a table of symbols representing types, objects and relations within the ASL. These include BN names, BN node value names, inter-BN relation names, and a small ontology of useful concepts such as those related to time.

The rule base contains rules, which comprise the syntactic grammar and ASG.

The lexicon is a table of Lisp-readable word information entries, obtained in part from the UMLS Specialist Lexicon.

The ASL inference engine combines symbolic unification with backward-chaining inference. It can be used to match an ASG pattern against an interpretation graph, and to perform tests associated with grammar rules.

The Bayesian network component utilizes the Norsys Netica(TM) API, and includes a set of Lisp and C language routines for instantiating and retrieving probabilities from BNs.

5 Training M+

Porting M+ to a new medical domain involves gathering a corpus of training sentences for the domain, using the Netica(TM) graphical interface

to create domain-specific BNs, and generating training cases for the new BNs.

The most time-consuming task is the creation of training cases. We have developed a prototype version of a Web-based tool which largely automates this task. The basic idea is to enable M+ to guess the BN value assignments of unknown words, then use it to parse phrases similar to phrases already seen. For instance, having been trained on the phrase "right upper lobe", the parser is able to produce reasonable parses, with some "guessed" value assignments, for "left upper lobe", "right middle lobe", "bilateral lungs", etc. The BN assignments produced by the parse are output as tentative new cases to be reviewed and corrected by the human trainer.

The training process begins with an initial set of interpreted "seed" phrases. From this set, the tool can apply the parser to phrases similar to this set, and so semi-automatically traverse ever widening semantically contiguous areas within the space of corpus phrases. As the training proceeds, the role of the human trainer increasingly becomes one of providing correction and interpretations for semantic patterns the system is increasingly able to discover on its own.

To parse phrases containing unknown words, M+ uses a technique based on a variation of the vector space model of lexical semantic similarity (Manning and Schutze, 1999). As M+ encounters an unknown word, it gathers a list of training corpus words judged similar to that word, as predicted by the vector space measure. It then identifies BN nodes whose known values significantly overlap with this list, and provisionally assigns the unknown word as a new value for those nodes. The assignment resulting in the best parsetree is selected for the new provisional training case.

6 Evaluation

M+ was evaluated for the extraction of American College of Radiology (ACR) utilization review codes from Head CT reports (Fizman, 2002). The ACR codes compare the outcome in a report with the suspected diagnosis provided by emergency department physicians. If the outcome relates to the suspected diagnosis then the report should be encoded as positive (P). If the outcome is negative and does not

relate to the suspected diagnosis then the report should be encoded as negative (N). In order to extract those ACR codes we trained M+ to extract eleven broad disease concepts, then inferred the ACR codes based on the application of a rule to the M+ output: If any of the concepts was present, the report was considered positive, else the report was considered negative.

Twenty six hundred head CT scan reports were used for this evaluation. Six hundred reports were randomly selected for testing, and the rest were used to train M+ in this domain. The performance of M+ on this task was measured against that of four board certified physicians, using a gold standard based on majority vote, as described in (Fizman, 2002). For each subject we calculated recall, precision and specificity with their respective 95 % confidence intervals for the capture of ACR utilization codes.

From 600 head CT reports, 67 were judged to be positive (P) by the gold standard physicians and 534 were judged to be negative (N). Therefore the positive rate for head CT in this sample was 11%. Recall, precision and specificity for every subject are presented with their respective 95% confidence intervals in Table 1. The physicians had an average recall of 88% (CI, 84% to 92.%), an average precision of 86% (CI, 81% to 90%), and average specificity of 98% (CI, 97% to 99%). M+ had recall of 87% (CI, 78% to 95%), precision of 85% (CI, 77% to 94%) and specificity of 98% (CI, 97% to 99%).

Table 1. Results of ACR utilization code study.

Subject	Recall	Specificity	Precision
Physician1	0.83 (0.74-0.92)	0.99 (0.98-1.00)	0.91 (0.84-0.99)
Physician2	0.88 (0.81-0.97)	0.98 (0.97-0.99)	0.84 (0.75-0.93)
Physician3	0.93 (0.87-1.00)	0.98 (0.97-0.99)	0.86 (0.78-0.95)
Physician4	0.88 (0.96-0.99)	0.97 (0.96-0.99)	0.81 (0.71-0.90)
M+	0.87 (0.78-0.95)	0.98 (0.97-0.99)	0.85 (0.77-0.94)

The results on Head CT reports are encouraging, but there are limitations. We only evaluated 600 reports, because it's very hard to get physicians to produce gold standard data for medical reports. The prevalence of positive

reports is only 11% and reflects the fact that the individual brain conditions have very low prevalence.

7 Conclusions

M+ and its predecessors have demonstrated that BNs provide a useful semantic model for medical text processing. In practice, a medical NLP system will frequently encounter missing and unknown words, unknown and ungrammatical phrase structures, and telegraphic usages. Knowledge databases will be imperfect and incomplete. Using BNs for semantic representation brings a noise-tolerant, partial match-tolerant, context-sensitive character to the recognition of semantic patterns, and to relevant inferences based on those patterns. In addition, BNs can be used to guess the semantic types of unknown words, providing a basis for bootstrapping the system's semantic knowledge.

8 Acknowledgements

Many thanks to Wendy W. Chapman for her advice and input in this paper, and her efforts to make M+ a useful addition to the RODS project at the University of Pittsburgh.

References

- Chapman W., Christensen L. M., Wagner M., Haug P. J., Ivanov O., Dowling J. N., Olszewski R. T. 2002. Syndromic Detection from Free-text Triage Diagnoses: Evaluation of a Medical Language Processing System before Deployment in the Winter Olympics. *Proc AMIA Symp.* (submitted).
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Special technical report (Massachusetts Institute of Technology, Research Laboratory of Electronics); no. 11. Cambridge, MA: MIT Press.
- Fiszman M., Blatter D.D., Christensen L.M., Oderich G., Macedo T., Eidelwein A.P., Haug P.J. 2002. Utilization review of head CT scans: value of a medical language processing system. *American Journal of Roentgenology (AJR)*. (submitted)
- Friedman C, Alderson PO, Austin JH, Cimino JJ, Johnson SB. 1994, A general natural-language text processor for clinical radiology. *J Am Med Inform Assoc.* Mar-Apr;1(2) pp. 161-74.
- Friedman N., Getoor L., Koller D. and Pfeffer A. 1999. Learning Probabilistic Relational Models. *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI)*: pp. 1300-1307.
- Haug P. J., Christensen L., Gundersen M., Clemons B., Koehler S., Bauer K. 1997. A natural language parsing system for encoding admitting diagnoses. *Proc AMIA Symp.* 81: pp. 4-8.
- Koehler, S. B. 1998. SymText: A natural language understanding system for encoding free text medical data. Ph.D. Dissertation, University of Utah.
- Koller D., and Pfeffer A. 1997. Object-Oriented Bayesian Networks. *Proceedings of the 13th Annual Conference on Uncertainty in AI*: pp. 302-313.
- Lin R, Lenert L, Middleton B, Shiffman S. A free-text processing system to capture physical findings: Canonical Phrase Identification System (CAPIS). *Proc Annu Symp Comput Appl Med Care.* pp. 843-7.
- Manning C. D. and Schutze H. 1999. Foundations of Statistical Natural Language Processing. MIT Press.
- Minsky, M. 1975. A framework for representing knowledge. In *The Psychology of Human Vision*, ed. P. H. Winston, pp. 211-277. McGraw Hill.
- Moore, R. C. 1989. Unification-based Semantic Interpretation. *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp33-41.
- Pearl, Judea. 1988. Probabilistic inference in intelligent systems. Networks of plausible inference: Morgan Kaufmann.
- Ranum D.L. 1989. Knowledge-based understanding of radiology text. *Comput Methods Programs Biomed.* Oct-Nov;30(2-3) pp.209-215.
- Romacker, Martin and Hahn, Udo. 2000. An empirical assessment of semantic interpretation. *ANLP/NAACL 2000 -- Proceedings of the 6th Applied Natural Language Processing Conference & the 1st Conference of the North American Chapter of the Association for Computational Linguistics.* pp. 327-334.
- Schank, R.C. and R. Abelson. 1997. *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Lawrence Erlbaum.
- Smart, J. F. and M. Roux. 1995. A model for medical knowledge representation application to the analysis of descriptive pathology reports. *Methods Inf Med.* Sep;34(4) pp. 352-60.