Literature Mining using Bayesian Networks

Péter Antal and András Millinghoffer

Department of Measurement and Information Systems

Budapest University of Technology and Economics

Abstract

In biomedical domains, free text electronic literature is an important resource for knowledge discovery and acquisition, particularly to provide a priori components for evaluating or learning domain models. Aiming at the automated extraction of this prior knowledge we discuss the types of uncertainties in a domain with respect to causal mechanisms, formulate assumptions about their report in scientific papers and derive generative probabilistic models for the occurrences of biomedical concepts in papers. These results allow the discovery and extraction of latent causal dependency relations from the domain literature using minimal linguistic support. Contrary to the currently prevailing methods, which assume that relations are sufficiently formulated for linguistic methods, our approach assumes only the report of causally associated entities without their tentative status or relations, and can discover new relations and prune redundancies by providing a domain-wide model. Therefore the proposed Bayesian network based text mining is an important complement to the linguistic approaches.

1 Introduction

Rapid accumulation of biological data and the corresponding knowledge posed a new challenge of making this voluminous, uncertain and frequently inconsistent knowledge accessible. Despite recent trends to broaden the scope of formal knowledge bases in biomedical domains, free text electronic literature is still the central repository of the domain knowledge. This central role will probably be retained in the near future, because of the rapidly expanding frontiers. The extraction of explicitly stated or the discovery of implicitly present latent knowledge requires various techniques ranging from purely linguistic approaches to machine learning methods. In the paper we investigate a domain-model based approach to statistical inference about dependence and causal relations given the literature using minimal linguistic preprocessing. We use Bayesian Networks (BNs) as causal domain models to introduce generative models of publication, i.e. we examine the relation of domain models and generative models of the corresponding literature.

In a wider sense our work provides support to statistical inference about the structure of the domain model. This is a two-step process, which consists of the reconstruction of the beliefs in mechanisms from the literature by model learning and their usage in a subsequent learning phase. Here, the Bayesian framework is an obvious choice. Earlier applications of text mining provided results for the domain experts or data analysts, whereas our aim is to go one step further and use the results directly in the statistical learning of the domain models.

The paper is organized as follows. Section 2 presents a unified view of the literature, the data and their models. In Section 3 we review the types of uncertainties in biomedical domains from a causal, mechanism-oriented point of view. In Section 4 we summarize recent approaches to information extraction and literature mining based on natural language processing (NLP) and "local" analysis of occurrence patterns. In Section 5 we propose generative probabilistic models for the occurrences of biomedical concepts in scientific papers. Section 6 presents textual aspects of the application

domain, the diagnosis of ovarian cancer. Section 7 reports results on learning BNs given the literature.

2 Fusion of literature and data

The relation of experimental data D_N, probabilistic causal domain models formalized as BNs (G, θ), domain literature $D_{N'}^L$ and models of publication (G^L, θ^L) can be approached at different levels. For the moment, let us assume that probabilistic models are available describing the generation of observations $P(D_N|(G, \theta))$ and literature $P(D_{N'}^L|(G^L, \theta^L))$. This latter may include stochastic grammars for modeling the linguistic aspects of the publication, however, we will assume that the literature has a simplified agrammatical representation and the corresponding generative model can be formalized as a BN (G^L, θ^L) as well.

The main question is the relation of $P(G, \theta)$ and $P(G^L, \theta^L)$. In the most general approach the hypothetical posteriors $P(G, \theta|D_N, \xi_i)$ expressing personal beliefs over the domain models conditional on the experiments and the personal background knowledge ξ_i determine or at least influence the parameters of the model (G^L, θ^L) in $P(D_{N'}^L|(G^L, \theta^L), \xi_i)$.

The construction or the learning of a full-fledged decision theoretic model of publication is currently not feasible regarding the state of quantitative modeling of scientific research and publication policies, not to mention the cognitive and even stylistic aspects of explanation, understanding and learning (Rosenberg, 2000). In a severely restricted approach we will focus only on the effect of the belief in domain models $P(G, \theta)$ on that in publication models $P(G^L, \theta^L)$. We will assume that this transformation is "local", i.e. there is a simple probabilistic link between the model spaces, specifically between the structure of the domain model and the structure and parameters of the publication model $p(G^L, \theta^L|G)$. Probabilistically linked model spaces allow the computation of the posterior over domain models given the literature(!) data as:

$$P(G|D_{N'}^L) = \frac{P(G)}{P(D_{N'}^L)} \sum_{G^L} P(D_{N'}^L|G^L)P(G^L|G).$$

The formalization $(D_N \leftarrow G \rightarrow G^L \rightarrow D_{N'}^L)$ also allows the computation of the posterior over the domain models given both clinical and the literature(!) data as:

$$P(G|D_N, D_{N'}^L) = P(G)\frac{P(D_{N'}^L|G)}{P(D_{N'}^L)}\frac{P(D_N|G)}{P(D_N|D_{N'}^L)}$$
$$\propto P(G)P(D_N|G) \sum_{G^L} P(D_{N'}^L|G^L)P(G^L|G),$$

The order of the factors shows that the prior is first updated by the literature data, then by the clinical data. A considerable advantage of this approach is the integration of literature and clinical data at the lowest level and not through feature posteriors, i.e. by using literature posteriors in feature-based priors for the (clinical) data analysis (Antal et al., 2004).

We will assume that a bijective relation exists between the domain model structures G and the publication model structures G^L ($\mathcal{T}(G) = G^L$), whereas the parameters θ^L may encode additional aspects of publication policies and explanation. We will focus on the logical link between the structures, where the posterior given the literature and possibly the clinical data is:

$$P(G|D_N, D_{N'}^L, \xi) \qquad\qquad (1)$$
$$\propto \quad P(G|\xi)P(D_N|G)P(D_{N'}^L|\mathcal{T}(G)).$$

This shows the equal status of the literature and the clinical data. In integrated learning from heterogeneous sources however, the scaling of the sources is advisable to express our confidence in them.

3 Concepts, associations, causation

Frequently a biomedical domain can be characterized by a dominant type of uncertainty w.r.t the causal mechanisms. Such types of uncertainty show certain sequentiality described below, related to the development of biomedical knowledge, though a strictly sequential view is clearly an oversimplification.

(1) *Conceptual phase*: Uncertainty over the domain ontology, i.e. the relevant entities.

(2) *Associative phase*: Uncertainty over the association of entities, reported in the literature as indirect, associative hypotheses, frequently as clusters of entities. Though we accept the general view of causal relations behind associations, we assume that the exact causal functions and direct relations are unknown.

(3) *Causal relevance phase*: (Existential) uncertainty over causal relations (i.e. over mechanisms). Typically, direct causal relations are theoretized as processes and mechanisms.

(4) *Causal effect phase*: Uncertainty over the strength of the autonomous mechanisms embodying the causal relations.

In this paper we assume that the target domain is already in the Associative or Causal phase, i.e. that the entities are more or less agreed, but their causal relations are mostly in the discovery phase. This holds in many biomedical domains, particularly in those linking biological and clinical levels. There the Associative phase is a crucial but lengthy knowledge accumulation process, where wide range of research methods is used to report associated pairs or clusters of the domain entities. These methods admittedly produce causally oriented associative relations which are partial, biased and noisy.

4 Literature mining

Literature mining methods can be classified into bottom-up (pairwise) and top-down (domain model based) methods. Bottom-up methods attempt to identify individual relationships and the integration is left to the domain expert. Linguistic approaches assume that the individual relations are sufficiently known, formulated and reported for automated detection methods. On the contrary, top-down methods concentrate on identifying consistent domain models by analyzing jointly the domain literature. They assume that mainly causally associated entities are reported with or without tentative relations and direct structural knowledge. Their linguistic formulation is highly variable, not conforming to simple grammatical characterization. Consequently top-down methods typically use agrammatical text representations and minimal linguistic support. They autonomously prune the redundant, inconsistent, indirect relations by evaluating consistent domain models and can deliver results in domains already in the *Associative* phase.

Until recently mainly bottom-up methods have been analyzed in the literature: *linguistic approaches* extract explicitly stated relations, possibly with qualitative ratings (Proux et al., 2000; Hirschman et al., 2002); *co-occurrence analysis* quantifies the pairwise relations of variables by their relative frequency (Stapley and Benoit, 2000; Jenssen et al., 2001); *kernel similarity analysis* uses the textual descriptions or the occurrence patterns of variables in publications to quantify their relation (Shatkay et al., 2002); Swanson and Smalheiser (1997) discover relationships through the heuristic pattern analysis of citations and co-occurrences; in (Cooper, 1997) and (Mani and Cooper, 2000) local constraints were applied to cope with possible hidden confounders, to support the discovery of causal relations; *joint statistical analysis* in (Krauthammer et al., 2002) fits a generative model to the temporal pattern of corroborations, refutations and citations of individual relations to identify "true" statements. The top-down method of the *joint statistical analysis* of de Campos (1998) learns a restricted BN thesaurus from the occurrence patterns of words in the literature. Our approach is closest to this and those of Krauthammer et al. and Mani.

The reconstruction of informative and faithful priors over domain mechanisms or models from research papers is further complicated by the *multiple aspects of uncertainty* about the existence, scope (conditions of validity), strength, causality (direction), robustness for perturbation and relevance of mechanism and the *incompleteness of reported relations*, because they are assumed to be well-known parts of *common sense* knowledge or of the *paradigmatic* already reported knowledge of the community.

5 BN models of publications

Considering (biomedical) abstracts, we adopt the central role of causal understanding and explanation in scientific research and publication (Thagard, 1998). Furthermore, we assume that the contemporary (collective) uncertainty over mechanisms is an important factor influencing the publications. According to this causal stance, we accept the 'causal relevance' interpretation, more specifically the 'explained' (explanandum) and 'explanatory' (explanans), in addition, we allow the 'described' status. This is appealing, because in the assumed causal publications both the name occurrence and the pre-processing kernel similarity method (see Section 6) express the presence or relevance of the concept corresponding to the respective variable. This implicitly means that we assume that publications contain either descriptions of the domain concepts without considering their relations or the occurrences of entities participating in known or latent causal relations. We assume that there is only one causal mechanism for each parental set, so we will equate a given parental set and the mechanism based on it.

Furthermore, we assume that mainly positive statements are reported and we treat negation and refutation as noise, and that exclusive hypotheses are reported, i.e. we treat alternatives as one aggregated hypothesis. Additionally, we presume that the dominant type of publications are causally ("forward") oriented. We attempt to model the transitive nature of causal explanation over mechanisms, e.g. that causal mechanisms with a common cause or with a common effect are surveyed in an article, or that subsequent causal mechanisms are tracked to demonstrate a causal chain. On the other hand, we also have to model the lack of transitivity, i.e. the incompleteness of causal explanations, e.g. that certain variables are assumed as explanatory, others as potentially explained, except for survey articles that describe an overall domain model. Finally, we assume that the reports of the causal mechanisms and the univariate descriptions are independent of each other.

5.1 The intransitive publication model

The first generative model is a two-layer BN. The upper-layer variables represent the pragmatic functions (described or explanandum) of the corresponding concepts, while lower-layer variables represent their observable occurrences (described, explanatory or explained). Upper-layer variables can be interpreted as the intentions of the authors or as the property of the given experimental technique. We assume that lower-layer variables are influenced only by the upper-layer ones denoting the corresponding mechanisms, and not by any other external quantities, e.g. by the number of the reported entities in the paper. A further assumption is that the belief in a compound mechanism is the product of the beliefs in the pairwise dependencies. Consequently we use noisy-OR canonic distributions for the children in the lower layer. In a noisy-OR local dependency (Pearl, 1988), the edges can be labeled with a parameter, inhibiting the OR function, which can be interpreted also structurally as the probability of an implicative edge.

This model extends the atomistic, individual-mechanism oriented information extraction methods by supporting the joint learning of all the mechanisms, i.e. by the search for a domain-wide coherent model. However it still cannot model the dependencies between the reported associations, and the presence of hidden variables considerably increase the computational complexity of parameter and structure learning.

5.2 The transitive publication model

To devise a more advanced model, we relax the assumption of the independence between the variables in the upper layer representing the pragmatic functions, and we adapt the models to the bag-of-word representation of publications (see Section 6). Consequently we analyze the possible pragmatic functions corresponding to the domain variables, which could be represented by hidden variables. We assume here that the explanatory roles of a variable are not differentiated, and that if a variable is explained, then it can be explanatory for any other

variable. We assume also full observability of causal relevance, i.e. that the lack of occurrence of an entity in a paper means causal irrelevance w.r.t. the mechanisms and variables in the paper and not a neutral omission. These assumptions allow the merging of the explanatory, explained and described status with the observable reported status, i.e. we can represent them jointly with a single binary variable. Note that these assumptions remain tenable in case of report of experiments, where the pattern of relevancies has a transitive-causal bias.

These would imply that we can model only full survey papers, but the general, unconstrained multinomial dependency model used in the transitive BNs provides enough freedom to avoid this. A possible semantics of the parameters of a binary, transitive literature BN can be derived from a causal stance that the presence of an entity X_i is influenced only by the presence of its potential explanatory entities, i.e. its parents. Consequently, $P(X_i = 1 | Pa_{X_i} = pa_{x_i})$ can be interpreted as the belief that the present parental variables can explain the entities X_i (Pa_{X_i} denotes the parents of X_i and $Pa_{X_i} \to X_i$ denotes the parental substructure). In that way the parameters of a complete network can represent the priors for parental sets compatible with the implied ordering:

$$P(X_i = 1 | Pa_{X_i} = pa_{X_i}) = P(Pa_{X_i} = pa_{X_i}) \tag{2}$$

where for notational simplicity $pa(X_i)$ denotes both the parental set and a corresponding binary representation.

The multinomial model allows entity specific modifications at each node, combined into the parameters of the conditional probability model that are independent of other variables (i.e. unstructured noise). This permits the modeling of the description of the entities ($P(X_i^{\mathrm{D}})$), the beginning of the transitive scheme of causal explanation ($P(X_i^{\mathrm{B}})$) and the reverse effect of interrupting the transitive scheme ($P(X_i^{\mathrm{I}})$). These auxiliary variables model simplistic interventions, i.e. authors' intentions about publishing an observational model. Note that a "backward" model corresponding to an effect-

to-cause or diagnostic interpretation and explanation method has a different structure with opposite edge directions.

In the Bayesian framework, there is a structural uncertainty also, i.e. uncertainty over the structure of the generative models (literature BNs) themselves. So to compute the probability of a parental set $Pa_{X_i} = pa_{X_i}$ given a literature data set $D_{N'}^L$, we have to average over the structures using the posterior given the literature data:

$$P(Pa_{X_i} = pa_{X_i} | D_{N'}^L) \tag{3}$$
$$= \sum_{(pa_{X_i} \to X_i) \subset G} P(X_i = 1 | pa_{X_i}, G) P(G | D_{N'}^L)$$
$$\approx \sum_G 1((pa_{X_i} \to X_i) \subset G) P(G | D_{N'}^L) \tag{4}$$

Consequently, the result of learning BNs from the literature can be multiple, e.g. using a maximum a posteriori (MAP) structure and the corresponding parameters, or the posterior over the structures (Eq. 3). In the first case, the parameters can be interpreted structurally and converted into a prior for a subsequent learning. In the latter case, we neglect the parametric information focusing on the structural constraints, and transform the posterior over the literature network structures into a prior over the structures of the real-world BNs (see Eq. 1).

6 The literature data sets

For our research we used the same collection of abstracts as that described in (Antal et al., 2004), which was a preliminary work using pairwise methods. The collection contains 2256 abstracts about ovarian cancer, mostly between 1980 and 2002. Also a name, a list of synonyms and a text kernel is available for each domain variable. The presence of the name (and synonyms) of a variable in documents is denoted with a binary value. Another binary representation of the publications is based on the kernel documents:

$$R_{ij}^K = \begin{cases} 1 & \text{if } 0.1 < \text{sim}(k_j, d_i) \\ 0 & \text{else} \end{cases}, \tag{5}$$

which expresses the relevance of kernel document k_j to document d_i using the 'term

frequency-inverse document frequency' (TF-IDF) vector representation and the cosine similarity metric (Baeza-Yates and Ribeiro-Neto, 1999). We use the term *literature data* to denote both binary representations of the relevance of concepts in publications, usually denoted with $D_{N'}^L$ (containing N' publications).

7 Results

The structure learning of the transitive model is achieved by an exhaustive evaluation of parental sets up to 4 variables followed by the K2 greedy heuristics using the BD_{eu} score (Heckerman et al., 1995) and an ordering of the variables from an expert, in order to be compatible with the learning of the intransitive model. The structure learning of the two-layer model has a higher computational cost, because the evaluation of a structure requires the optimization of parameters, which can be performed e.g. by gradient-descent algorithms. Because of the use of the "forward" explanation scheme, only those variables in the upper layer can be the parents of an external variable that succeed it in the causal order. Note that beside the optional parental edges for the external variables, we always force a deterministic edge from the corresponding non-external variable. During the parameter learning of a fixed network structure the non-zero inhibitory parameters of the lower layer variables are adjusted according to a gradient descent method to maximize the likelihood of the data (see (Russell et al., 1995)). After having found the best structure, according to its semantics, it is converted into a flat, real-world structure without hidden variables. This conversion involves the merging of the corresponding pairs of nodes of the two layers, and then reverting the edges (since in the explanatory interpretation effects precede causes).

We compared the trained models to the expert model using a quantitative score based on the comparison of the pairwise relations in the model, which are defined w.r.t. the causal interpretation as follows (Cooper and Yoo, 1999; Wu et al., 2001): *Causal edge (E)* An edge between the nodes. *Causal path (P)* A directed path

linking nodes. *(Pure) Confounded (C)* The two nodes have a common ancestor. The relation is pure, if there is no edge or path between the nodes. *Independent (I)* None of the previous (i.e. there is no causal connection).

The difference between two model structures can be represented in a matrix containing the number of relations of a given type in the expert model and in the trained model (the type of the relation in the expert model is the row index and the type in the trained model is the column index). These matrices (i.e. the comparison of the transitive and the intransitive models to the expert's) are shown in Table 1. Scalar

Table 1: Causal comparison of the intransitive and the transitive domain models (columns with 'i' and 't' in the subscript, respectively) to the expert model (rows).

	I_i	C_i	P_i	E_i	I_t	C_t	P_t	E_t
I	12	0	0	0	0	4	2	6
C	106	20	2	4	4	90	26	12
P	756	72	80	18	188	460	216	62
E	70	6	8	36	6	38	24	52

scores can be derived from this matrix, to evaluate the goodness of the trained model, the standard choice is to sum the elements with different weights (Cooper and Yoo, 1999; Wu et al., 2001). One possibility e.g. if we take the sum of the diagonal elements as a measure of similarity. By this comparison, the intransitive model achieves 148 points, while the transitive 358, so the transitive reconstructs more faithfully the underlying structure. Particularly important is the (E, E) element according to which 52 of the 120 edges of the expert model remains in the transitive model, on the contrary the intransitive model preserves only 36 edges. Similarly the independent relations of the expert model are well respected by both models.

Another score, which penalizes only the incorrect identification of independence (i.e. those and only those weights have a value of 1 which belong to the elements $(I, .)$ or $(., I)$, the others are 0), gives a score 210 and 932 for the transitive model and the intransitive respectively.
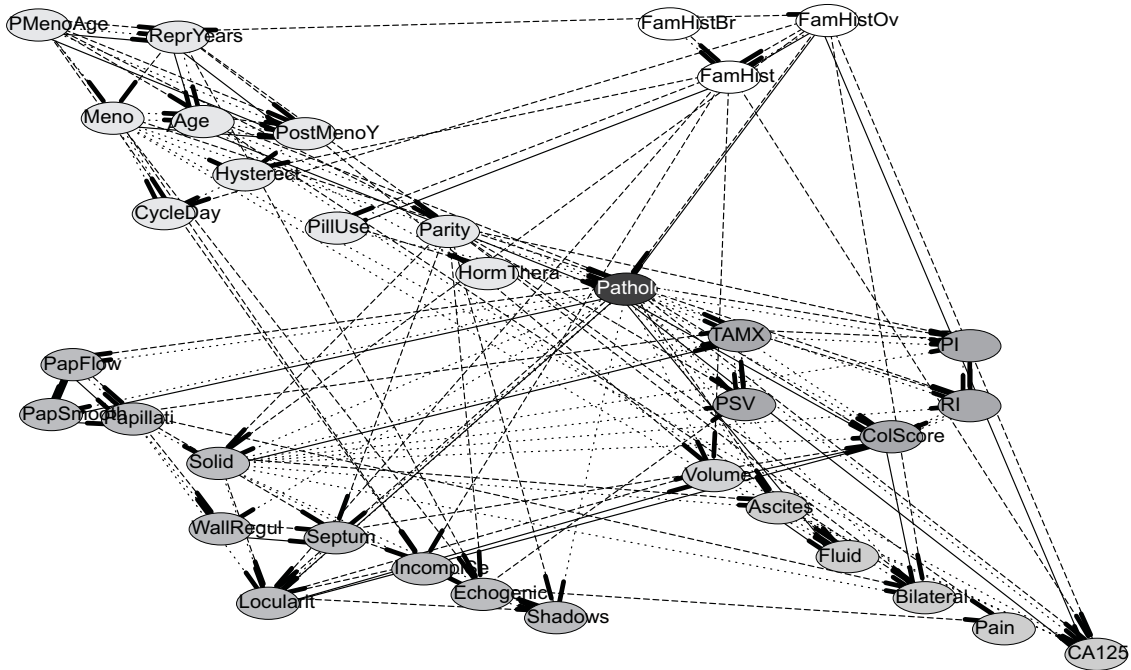
Figure 1: The expert-provided (dotted), the MAP transitive (dashed) and the intransitive (solid) BNs compatible with the expert's total ordering of the thirty-five variables using the literature data set ($PM_{R_3}^{COREL}$), the K2 noninformative parameter priors, and noninformative structure priors.

This demonstrates that the intransitive model is extremely conservative in comparison with both the other learning method and with the knowledge of the expert, it is only capable of detecting the most important edges; note that the proportion of its false positive predictions regarding the edges is only 38% while in the transitive model it is 61%.

Furthermore, we investigated the Bayesian learning of BN features, particularly using the temporal sequence of the literature data sets. An important feature indicating relevance between two variables is the so-called 'Markov Blanket Membership' (Friedman and Koller, 2000). We have examined the temporal characteristics of the posterior of this relation between a target variable 'Pathology' and the other ones using the approximation in Eq. 4. This feature is a good representative for the diagnostic importance of variables according to the community. We have found four types of variables: the posterior of the relevance increasing in time fast or slowly, decreasing slowly or fluctuating. Fig-
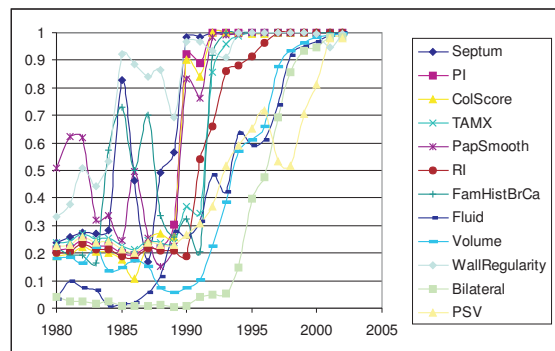

Figure 2: The probability of the relation Markov Blanket Membership between Pathology and the variables with a slow rise.

ure 2 shows examples for variables with a slow rising in time.

8 Conclusion

In the paper we proposed generative BN models of scientific publication to support the construction of real-world models from free-text literature. The advantage of this approach is its

domain model based foundation, hence it is capable of constructing coherent models by autonomously pruning redundant or inconsistent relations. The preliminary results support this expectation. In the future we plan to use the evaluation methodology applied there including rank based performance metrics and to investigate the issue of negation and refutation particularly through time.

References

P. Antal, G. Fannes, Y. Moreau, D. Timmerman, and B. De Moor. 2004. Using literature and data to learn Bayesian networks as clinical models of ovarian tumors. *Artificial Intelligence in Medicine*, 30:257–281. Special issue on Bayesian Models in Medicine.

R. Baeza-Yates and B. Ribeiro-Neto. 1999. *Modern Information Retrieval*. ACM Press, New York.

G. F. Cooper and C. Yoo. 1999. Causal discovery from a mixture of experimental and observational data. In *Proc. of the 15th Conf. on Uncertainty in Artificial Intelligence (UAI-1999)*, pages 116–125. Morgan Kaufmann.

G. Cooper. 1997. A simple constraint-based algorithm for efficiently mining observational databases for causal relationships. *Data Mining and Knowledge Discovery*, 2:203–224.

L. M. de Campos, J. M. Fernández, and J. F. Huete. 1998. Query expansion in information retrieval systems using a Bayesian network-based thesaurus. In Gregory Cooper and Serafin Moral, editors, *Proc. of the 14th Conf. on Uncertainty in Artificial Intelligence (UAI-1998)*, pages 53–60. Morgan Kaufmann.

N. Friedman and D. Koller. 2000. Being Bayesian about network structure. In Craig Boutilier and Moises Goldszmidt, editors, *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence(UAI-2000)*, pages 201–211. Morgan Kaufmann.

D. Geiger and D. Heckerman. 1996. Knowledge representation and inference in similarity networks and Bayesian multinets. *Artificial Intelligence*, 82:45–74.

D. Heckerman, D. Geiger, and D. Chickering. 1995. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20:197–243.

L. Hirschman, J. C. Park, J. Tsujii, L. Wong, and C. H. Wu. 2002. Accomplishments and challenges in literature data mining for biology. *Bioinformatics*, 18:1553–1561.

T. K. Jenssen, A. Laegreid, J. Komorowski, and E. Hovig. 2001. A literature network of human genes for high-throughput analysis of gene expression. *Nature Genetics*, 28:21–28.

M. Krauthammer, P. Kra, I. Iossifov, S. M. Gomez, G. Hripcsak, V. Hatzivassiloglou, C. Friedman, and A. Rzhetstky. 2002. Of truth and pathways: chasing bits of information through myriads of articles. *Bioinformatics*, 18:249257.

S. Mani and G. F. Cooper. 2000. Causal discovery from medical textual data. In *AMIA Annual Symposium*.

J. Pearl. 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Francisco, CA.

D. Proux, F. Rechenmann, and L. Julliard. 2000. A pragmatic information extraction strategy for gathering data on genetic interactions. In *Proc. of the 8th International Conference on Intelligent Systems for Molecular Biology (ISMB'2000), La-Jolla, California*, pages 279–285.

A. Rosenberg. 2000. *Philosophy of Science: A contemporary introduction*. Routledge.

S. J. Russell, J. Binder, D. Koller, and K. Kanazawa. 1995. Local learning in probabilistic networks with hidden variables. In *IJCAI*, pages 1146–1152.

H. Shatkay, S. Edwards, and M. Boguski. 2002. Information retrieval meets gene analysis. *IEEE Intelligent Systems*, 17(2):45–53.

B. Stapley and G. Benoit. 2000. Biobibliometrics: Information retrieval and visualization from co-occurrences of gene names in medline asbtracts. In *Proc. of Pacific Symposium on Biocomputing (PSB00)*, volume 5, pages 529–540.

D. R. Swanson and N. R. Smalheiser. 1997. An interactive system for finding complementary literatures: a stimulus to scientific discovery. *Artificial Intelligence*, 91:183–203.

P. Thagard. 1998. Explaining disease: Correlations, causes, and mechanisms. *Minds and Machines*, 8:61–78.

X. Wu, P. Lucas, S. Kerr, and R. Dijkhuizen. 2001. Learning bayesian-network topologies in realistic medical domains. In *In Medical Data Analysis: Second International Symposium, ISMDA*, pages 302–308. Springer-Verlag, Berlin.