

ベイジアンネットワークを用いた対話システム： ユーザモデルの推定

A Bayesian Approach toward Dialogue Systems: A Method for Expecting User Models

秋葉 友良 田中 穂積

Tomoyosi Akiba Hozumi Tanaka

東京工業大学工学部情報工学科

Department of Computer Science, Tokyo Institute of Technology

Abstract

User modeling is one of the important components of dialog systems. Most of the previous approaches are rule-based methods. We propose to represent user models by Bayesian networks. The advantages of the Bayesian approach over the rule-based approach are as follows: First, rules for updating user models are not necessary because updating is directly performed by the evaluation of the network based on probability theory; this provides us a more formal way of dealing with uncertainties. Second, the Bayesian network provides more detailed information of users' knowledge, because the degree of belief on each concept is provided in terms of probability. We prove these advantages through a preliminary experiment.

1 Introduction

対話システムにおいて、ユーザモデリングの重要性が指摘されている。ユーザモデリングは単に対話をより協調的にするために必要なだけでなく、多くのドメインにおいて適切な対話を達成するために必要不可欠である [8]。ユーザモデルは、システムの他の構成要素と密接に相互作用し、それらを切り離して考えることは簡単ではない。例えば、入力文解析の際の指示表現の曖昧性は、ユーザが何を知っているかによってしばしば解消される。また、出力生成の際、生成文に用いる語彙の選択も、ユーザの知識に左右される。

人に概念を説明するために、我々は相手が既に知っている概念との関係を示す。そのために、対話システムでは相手の知っていることを予測することが重要である。例えば、ユーザにレストランの場所を説明することを考える。レストランの場所を教えるために、地理的な絶対座標を教えるのは無意味である。なぜなら、相手のモデルは絶対座標系によって構成されていそうもないからである。相手がすでに知っている場所と、教える場所との相対的な関係を説明する必要がある。また、他の場所との関係を多く示すほど、ユーザの場所の理解は深まるであろう。しかし、初対面のユーザが確実に知っている場所を知ることは難しい。もちろん質問によって獲得を試みることは出来るが、むなしい質問の連続になる恐れがあり、ユーザをいらだたせ、機械的で非協調的とみなされるであろう。したがって、多少とも、ユーザの発話の中の手がかりから、ユーザの知識を予測するプロセスが必要となる。

この問題に対して、UC[5] で採用されているアプローチは、幾つかの典型的なユーザモデルをあらかじめ用意しておき、ユーザのレベルに合わせてそれらのうちから適切なものを選択するというものである。文献 [10] では、各オブジェクトに専門レベルを与え、ユーザのレベルを推測した後、その専門度以下のオブジェクトについては知っていると判断する。しかし、離散的なユーザモデルを用意するのはあくまでも近似であり、相手の専門度

は連続的で、ユーザモデルは人それぞれ異なるため、あらかじめモデルを用意するようなアプローチには限界がある。また、特に我々の目指すドメイン、地理案内では、単一の専門度でモデルを区別するのでは不十分である。ユーザは、ある特定の地域の地理にだけ詳しい場合もあるし、また特定のカテゴリーの店にだけ詳しい場合もある。

この分野でのもう一つの問題は、対話の進行に合わせてどのようにユーザモデルを更新するかということである。対話セッションの開始時点では、システムは正確にユーザモデルを予測することは難しいが、セッションが進行するにつれて、システムは十分な手がかりを得たり、説明を与えることで、ユーザの知識への理解が深まる。これを実現するために、漸進的な更新手続きが必要になる。

ユーザモデルを更新することの難しさは、不確実性の扱いに存する。ユーザの発話から得られる手がかりは不確実性を伴う。得られた情報は以前に獲得した情報と競合する可能性もあり、その結果、ユーザモデルに対する信念の翻意が起こる。同様に、システムの与える説明の効果も不確実である。さらに、ユーザの知識に関する推論も不確実性のもとで行なわなければならない。Cawsey [2] は、より信頼できる更新規則を先に適用し、相対的に不確かな情報を無視することでこの問題に対処している。ルールをベースとした他のアプローチでは、モデルの無矛盾性を管理するために、真理管理機構 (TMS) や非単調論理 [1] などが利用される。しかし、ルールベースのアプローチは、不確実性を扱うのに本質的に限界があるように思われる [4]。

本稿では、対話システムのユーザモデリングに対する確率的なアプローチを提案する。知識表現・推論にベイジアンネットワークを利用することで、以上で挙げた問題点に対して洗練された解法を与える。ベイジアンアプローチは潜在的に多くの利点を持つにも関わらず、対話システムのユーザモデリングに適用した例はほとんど報告がない。

さらに、ベイジアンアプローチは、ルールベースアプローチに対して、次のような利点がある。第1に、ユーザモデルを更新する規則・プロセスが必要ない。Cawsey [2] は、ユーザモデル更新に必要な4つの情報源を示している。それらは、ユーザのシステム対する発話、システムのユーザに対する発話、ユーザの専門度、ドメイン中の概念の間の関係、である。これらは全て、ベイジアンネットワークの表現の中に組み込むことができ、ネットワークの評価によって自動的にユーザモデルの更新に利用される。第2に、ユーザの知識に関して、より詳細な情報を示すことができる。一般に、知識の二値的な扱い、すなわち知っているか知らないか、は不確実性のもとでモデルを判断するには粗すぎるため、信念の度合を表すため多値あるいは連続的な指標が使われる。しかし、その値をどこから獲得するか、ある値が何を意味するかが明らかではない。ベイジアンアプローチでは、確率理論に基づいた確信度が得られ、その理論的基盤をもとに利用することができる。

以下の節では、まず2節でベイジアンネットワークについて簡単な概観を行なう。3節では、ベイジアンネットワークでの知識表現について論じる。ユーザモデルを表現すれば、更新はネットワークの評価によって自動的に行うことができる。4節では、対話例を示し、実際の評価結果を報告する。最後に、5節で結論をまとめる。

2 Bayesian Networks

確率理論に従って推論を行なうためには、確率モデルを指定しなければならない。一般に、完全な確率モデルは領域上の全ての確率変数の joint probability を指定する必要がある。このためには、非常に多くの確率が必要になる。例えば、全ての確率変数 (総数 n) が

二値 (0 か 1) であると仮定すると、完全な確率モデルの指定には、 $2^n - 1$ の確率が必要である。(n 個の確率変数を x_1, x_2, \dots, x_n とすると、joint probability は $P(x_1 = 0, x_2 = 0, \dots, x_n = 0), P(x_1 = 1, x_2 = 0, \dots, x_n = 0), \dots, P(x_1 = 1, x_2 = 1, \dots, x_n = 1)$ で与えられる。これらは足して 1 となるので、これらのうちの 1 つは自動的に得られる。) さらに、実際問題として、領域全体の joint probability を陽に与えることは難しい。事実、世界の考えうる全ての状態を数えあげ、それら一つ一つに確率を与えることになり、無意味である。

ベイジアンネットワークを用いると、はるかに少ない確率で確率モデルを指定することができる。領域を調べることで得られる定性的な情報が、この差を埋め合わせる。局所的な定性的な知識でネットワークを構成し、定量的に確率推論を行なうことができる。

形式的には、ベイジアンネットワークは有効非循環グラフ (DAG) であり、各ノードは確率変数を、有向弧はそれによって結び付けられる変数の間の直接の依存関係を表す。有効弧がある変数からある変数に達する時、前者を後者の親、後者を前者の子と呼ぶ。確率モデルを与えるために必要な確率は、各ノード x についてその親ノードの集合 $p(x)$ の取り得る値のパターンを条件とする全ての条件付確率 $P(x|p(x))$ である。親のないノード x には、その先験確率 $p(x)$ が与えられる。以上が完全な確率モデルを指定するために必要な全ての情報である [9]。

ベイジアンネットワーク上で推論を行なうことは、証拠の集合 E が与えられた時の各ノード x の後験確率 $P(x|E)$ を求めることに対応する。証拠 E はネットワーク中のノード集合の部分集合に値が与えられたものである (例えば、あるノード y, z について、 $E = \{y = 1, z = 0\}$)。各ノードの後験確率は、ネットワーク全体を評価することで求まり、これは一般に、統計シミュレーション [9] によって計算される。したがって、ベイジアンネットワークによって正しく知識が表現できれば、ネットワークの評価によって自動的に推論結果が得られる。次節では、ベイジアンネットワークでの知識表現について論じる。

3 Knowledge Representation with Bayesian Networks

3.1 Designing The Language

ベイジアンネットワークのノードには確率変数が対応する。ベイジアンネットワークによって知識を表現するために、まず確率変数が実際に何を意味するかを示すための言語を設計する。最初に、各確率変数の値域を、true (1) あるいは false (0) のいずれかを取ると仮定する。この仮定は本質的ではなく、KNOWN, NOT-KNOWN, NO-INFORMATION のように 3 値を取ると仮定しても構わない。

具体的にどのような種類の文を用意するかは、どのようなアプリケーションに使用するかに依存する。しかし、システムがユーザに説明を行なうようなドメイン一般にいえることは、ユーザがあるオブジェクトの名前を知っていることと、その他の属性を知っていることを区別することである。例えば、ユーザがシステムに店の名前を伝えてその位置を尋ねる場合、ユーザはその店の名前という属性は知っているが、その位置という属性は知らないと考えられる。名前を知っていたという事実から、さらなる推論を行なうことができる。そのオブジェクトと関連するオブジェクト、そしてそのオブジェクトの属性 (名前、位置等) を知っているだろうことが推論できる。あるオブジェクトの名前を (恐らく) 知っていることと推論することは、その名前を説明に使うことを (確率的に) 保証されることになる。

本稿で使用する言語を形式的に定義する。すべての文は、 $\langle label \rangle : \langle content \rangle$ という形式を取る。 $\langle label \rangle$ の値は、PRE, POST, JUDGE, TOLD, TELL のいずれかである。

(content) の記述は、通常、その記述が表す言明をユーザが知っているという仮説を表す。ただし、後で述べる専門度を表す記述はこの限りではない。

(content) で表す言明は、表す内容によっていくつかの方法で表現される。まず、オブジェクト自体やクラスを表す文が必要となる。これらはオブジェクト ID あるいはクラス ID をそのまま用いて表す。例えば、shop001, shop, ...。オブジェクトの属性を表す文は、関数記号が属性を表し、引数にオブジェクトをとるような関数表記で表現する。例えば name(shop001), location(shop001), ...。ある分野でのユーザの専門度を表す文は、その分野の名前で表す。例えば、records_collector, craft_work_expert ...。

ラベル PRE を持つ文は、対話セッションに先行してユーザが持っている知識を表し、一方、ラベル POST を持つ文は、その対話セッションの間にユーザが獲得した知識を表す。ラベル JUDGE を持つ文は、PRE, POST とラベル付けされた同じ内容の言明を持つ2つの文に対する判断を統合し、ユーザモデルの利用の際に参照される。ラベル TOLD, TELL を持つ文は、それぞれ、ユーザの発話から得られる証拠、システムが与える説明による証拠を表す。

3.2 Constructing the Networks

ベイジアンネットワークの弧は、連結する変数の間の直接の依存関係を表す。有向性は非遷移的依存関係を表すために本質的である。ベイジアンネットワークによって知識を表現するには、全節で定義した文の間の定性的な関係を、この有向弧、あるいはその組合せに、翻訳しなければならない。

我々のネットワークには、2つのサブネットワークがある。1つは対話セッションに先行してユーザが持っている知識を表し、これを prior part と呼ぶ。もう1つは対話セッションの間にユーザが獲得する知識を表し、posterior part と呼ぶ。prior part サブネットワーク中の文は、PRE あるいは TOLD のラベルを持ち、posterior part サブネットワーク中の文は、POST あるいは TELL のラベルを持つ。これら2つのネットワークを区別することは重要である。システムが教えた概念は、予測に関する概念の間の関係を使って、他のコンセプトを知っていることの根拠になることはない。例えば、ユーザがある店(の位置)を知っていることが分かると、その店に近い店も知っていることが推論できるが、システムがユーザに店(の位置)を教えた場合はその限りではない。

ネットワークの prior part には、領域知識から獲得されたコンセプトの間の関係が表現される。我々のドメインでは、4種類に分類される：(1) オブジェクト自体の間の関係。一般/特定関係、ある専門分野に共通する概念の間の関係などが記される。(2) オブジェクトの属性の間の関係。特に我々のドメインでは、すでに知っていると思われる場所に近い場所は恐らく知っているだろうといった、場所の間の関係。(3) オブジェクトとその属性の間の関係。オブジェクトを知っているならその属性も知っているという一般知識を表現。(4) ある概念とユーザがその概念を知っていることを示す発話(証拠)との間の関係。この証拠ノードには、ラベル TOLD が付けられる。

専門分野に関する関係は、ある分野の専門を表すノードを通して表現される。ネットワークは、専門分野ノードが関連する概念を子ノードとする形式で構成される。例えば、“records_collector” ノードは、概念 “Compact Disk” やレコード店 “Tower Records” などの子ノードとする。専門度のレベルは専門ノードを親ノードとする概念のノードの条件付き確率によって制御される。この枠組によって、任意数の分野の専門度を導入することができる。

上で述べた関係の中で、位置の間の関係は対称的である。ベイジアンネットワークの単

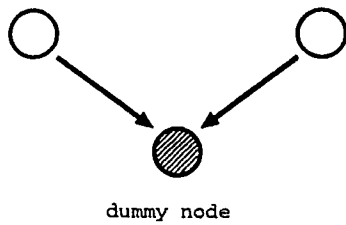


図 1: Symmetric relation

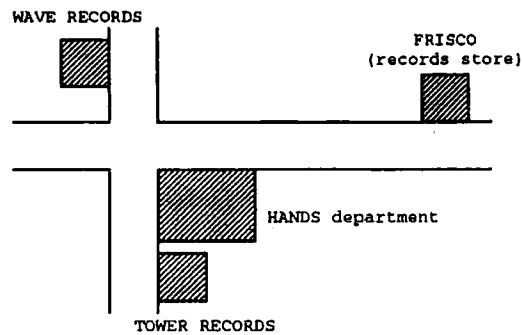


図 2: A map of a certain town

一の有向弧は対称的な関係を表現しない。そこで、対称関係を表すために、ダミーノードを導入しそれを対称関係を持つノードの共通の子ノードとなるようにネットワークを構成する(図1) [9]。

一方、ネットワークの posterior part では、ラベル **POST** を持つ概念ノードは、**TELL** ノードを子ノードとする。説明は、概念間の関係を示すことで行なわれるため **TELL** ノードは2つ以上の **POST** ノードの共通の子ノードとなる場合もある。実際、我々のドメインである街の案内では、ある店の位置の説明は、ユーザがすでに知っている他の店の位置との関係を示すことによって行なわれる。

ペイジアンネットワークでの曖昧な証拠の扱いは、直接的である。証拠ノード、すなわち **TOLD**, **TELL** とラベル付けされたノード、を2つ以上のノードの共通の子ノードとなるようにネットワークを構成することで、曖昧性を表現できる。例えば、音声入力を考えた場合、ユーザの発話が「タワーレコード」か「パワーレコード」で曖昧で、またどちらのレコード店も存在するとする。**TOLD** とラベル付けされた証拠ノードは、**PRE: name(tower)** と **PRE: name(power)** の両方を親ノードとするよう導入でき、証拠ノードに与える条件付き確率によって、認識率から得られる確からしさを反映させることができる。

JUDGE とラベル付けされたノードは、それぞれ同じ言明の内容を持つ、prior part の **PRE** ノードと posterior part の **POST** ノードの2つからサポートされる。すなわち、ユーザがあるコンセプトを知っているのは、それを対話セッションが始まる前からすでに知っていたか、あるいはシステムから教えられたから知っているのかのいずれかである。

結果のネットワークの例を図3に示す。

4 Examples

ある街についてユーザがシステムに質問する場合を考える(図2)。システムは図3のネットワークを利用する。対話はユーザの質問から始まる。

(1) user: フリスコはどこですか

実際には、入力文解析部がユーザの発話からネットワーク中の証拠を獲得する必要があるが、この話題は本稿の扱う範囲を越える。今の場合、入力文を解析することによって、システ

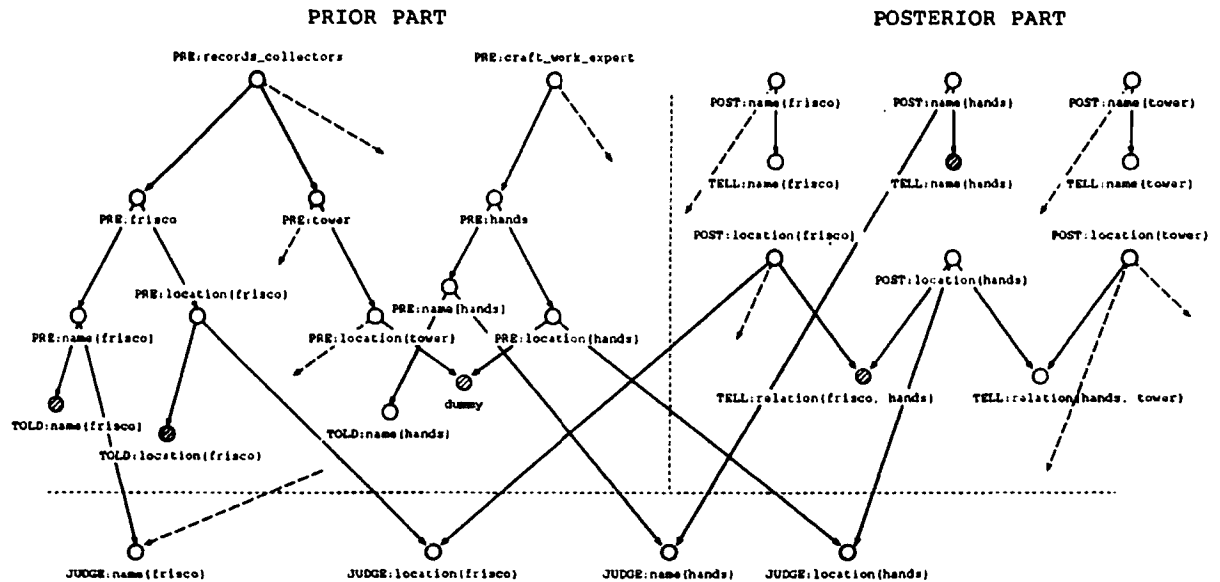


図 3: Example of a network

ムはユーザがある店の名前を知っているが、その位置は知らないということが分かる。ネットワークに与えられる証拠は $E = \{\text{TOLD: name(frisco)} = 1, \text{TOLD: place(frisco)} = 0\}$ である。証拠を考慮した結果、ユーザモデル中のある概念 x の信念は、その後験確率 $P(x|\text{TOLD: name(frisco)} = 1, \text{TOLD: place(frisco)} = 0)$ として得られる。この結果は、ネットワークを評価することによって直接得られるが、ここでは説明のため、その推論過程を簡単に追っていくことにする。実際には、その推論過程は単純ではない。あるノードの確信度を考えるのに、残りの全てのノードの影響を考慮する必要があるからである。ネットワークの評価のため統計シミュレーションが必要になるのはこの理由による。

ユーザは高い専門度を示す“フリスコ”(レコード店)という名前を知っていた。よって、ノード **PRE: records_collectors** の確率を上げ、同時に **records_collectors** が知っていると思われるノード (他のレコード店: **PRE: tower**, **PRE: wave**) の確率を上げる。その結果、これらのノードから、その属性を表すノード (**PRE: location(tower)**, **PRE: name(tower)**, **PRE: location(wave)**, など) に影響を与える。さらに、位置の属性のノードは、その近くにある位置のノードの確率を上げる。ここでは、**PRE: location(tower)** から、結果として **PRE: location(hands)** の確率が上がる。

次に、システムは結果のユーザモデルを利用して、発話を生成する。これは発話生成プランナの仕事である。ここでは、“ハンズ”という語を説明に使うことを決めたと仮定しよう。

(2) system: ハンズから西へ 300 メートルほど進んで下さい

文を発話した結果、システムは証拠 **TELL: name(hands) = 1**, **TELL: rel(hands, frisco) = 1** をネットワークに加える。(場所の説明は他の場所との関係を示すことで行なわれることに注意。) その結果、ノード **POST: place(frisco)** の確率、 $P(\text{POST: place(frisco)}|E)$ が上がる。ここで、 E はこれまでに得られた全ての証拠の集合を表す。次のユーザの発話は以下のものであった。

node (JUDGE:)	prior	probabilities after ...			
		(1)	(2)	(3)	(4)
location(frisco)	.514	.211	.433	.432	.659
location(tower)	.488	.690	.672	.306	.701
location(hands)	.485	.685	.760	.434	.739
name(frisco)	.475	.870	.861	.802	.796
name(tower)	.475	.796	.775	.631	.896
name(hands)	.465	.545	.874	.827	.827
(PRE:)					
records_collector	.396	.871	.841	.643	.620

表 1: the result of evaluation

(3) user: ハンズの様子がわかりません

この発話から証拠 TOLD: place(hands) = 0 が得られる。ここで、信念の翻意が生じる。ノード PRE: place(hands) の確率が落ち、その結果 PRE: place(tower) の確率も減少する。発話生成プランナーは、証拠 TELL: rel(hands,tower)=1 を与え、まず hands の場所を示すことを試みた。

(4) system: ハンズはタワーレコードの向かって左隣りのビルです

この説明は、hands の位置についてのユーザの理解に影響を与えるだけでなく、frisco の位置についても影響を与える。与えられた証拠は、ノード POST: place(frisco) の確率も、ノード POST: place(hands) を介して、影響を与えるからである。

以上の対話を扱ったネットワークの評価結果を表 1 に示す。

5 Conclusion

本稿では、ユーザモデリングへのベイジアンアプローチを提案した。主に、ベイジアンネットワークを用いて知識を表現する方法を論じた。推論はネットワークの評価に従って自動的に行なわれる。

一般には、ベイジアンネットワークを正確に評価するための計算量は NP 困難であるが、計算量的に有利な近似解法 (統計シミュレーション) が知られている。AI における多くの興味深い問題が計算量的には NP 困難であることが知られており、近似解法による解法が現実的であるという事実を考えると、近似解法の知られているベイジアンアプローチは魅力がある。実際には、形式的な計算量は別にしても、現在の計算資源ではシミュレーションのコストは高い。しかし、並列計算による良い結果も報告されている [6]。

ユーザモデルの予測を手に入れた後、それを発話生成に利用する機構が必要となる。発話生成プランナーは、ユーザモデルを参照しつつ、システムの目的を達成するプランを作り出す。プランナーによって提案されたプランは、単純さなどの尺度を統合した評価関数によっ

て評価される。ユーザモデルの確率は、この評価関数に乗じることで、期待値による評価に使用することが出来る。

自然言語処理の分野では、ベイジアンアプローチは、プラン認識 [3] と語義の曖昧性解消 [6] の分野での報告がある。我々は、まず最初にユーザモデリングにベイジアンアプローチを採用した。これは対話システムにとってユーザモデリングが中心的な構成要素となると考えるからである。実際、ユーザモデルは他の構成要素と密接に関係し合う。我々は、音声からの認識を行なう [7] 対話システムを構築中であるが、ユーザモデリングを始めとして、他の多くの構成要素にベイジアンアプローチを採用する予定である。ベイジアンアプローチの利点は、これらの異なる構成要素の間の推論を密に結合することにもある。最終的に、各構成要素が密接に関係し合う統合的な対話システムを目指す予定である。

参考文献

- [1] Douglas E. Appelt and Kurt Konolige. A nonmonotonic logic for reasoning about speech acts and belief revision. In *International Workshop on Nonmonotonic Reasoning*, pp. 164-175, 1988.
- [2] A. Cawsey. *Explanation and Interaction*. MIT Press, 1993.
- [3] E. Charniak and R.P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, Vol. 64, No. 1, pp. 53-79, 1983.
- [4] Peter Cheeseman. In defence of probability. In *the Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1002-1009, 1985.
- [5] David N. Chin. KNOME: Modeling what the user knows in UC. In A. Kobsa and W. Wahlster, editors, *User Models in Dialog Systems*, chapter 4, pp. 74-107. Springer-Verlag, 1989.
- [6] Leila M. R. Eizirik, Valmir C. Babosa, and Sueli B. T. Mendes. A bayesian-network approach to lexical disambiguation. *Cognitive Science*, Vol. 17, pp. 257-283, 1993.
- [7] K. Itou, S. Hayamizu, and H. Tanaka. Continuous speech recognition by context-dependent phonetic HMM and an efficient algorithm for finding n-best sentence hypotheses. In *In Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, 1992.
- [8] A. Kobsa and W. Wahlster, editors. *User Models in Dialog Systems*. Springer Verlag, 1989.
- [9] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, 1988.
- [10] J.W. Wallis and E. H. Shortliffe. Customized explanations using causal knowledge. In B.G. Buchanan and E.H. Shortliffe, editors, *Rule Based Expert Systems: The MYCIN experiments of the Stanford Heuristic Programming Project*, pp. 371-390. Addison Wesley, 1985.