

# ベイジアンネットワークを用いた対話相手の知識の推定法 Modeling User's Knowledge with Bayesian Networks

22-5

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

User modeling is an important component of dialog systems. Most previous approaches are rule-based methods. In this paper, we propose to represent user models through Bayesian networks. Some 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 はじめに

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 [3]。

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(ユーザの知識)を予測することが重要となる。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。したがって、システムはユーザが既に知っている場所からの相対位置を説明する必要がある。しかし、初対面のユーザが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。

ユーザの知識を推定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。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。

## 2 ネットワークの構成

ベイジアンネットワークによって推論を行うために、ドメインについての知識をネットワークで表現する必要がある。その詳細は文献 [5] に譲る。ここでは、基本的な考え方について述べるに留める。

形式的には、ベイジアンネットワークは有向非循環グラフ (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、各ノードの確率変数  $x$  の取り得る値  $X$  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  $P_x$  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 [4]。

知識表現のために、まず各ノードがどんな事象に対応するかを決め、次に各事象の間の依存関係を DAG で表現する。どのような事象を考慮するかは、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、ユーザの知識を推定するプロセス一般において重要なのは、ユーザがあるオブジェクトの名前を

\*連絡先: 秋葉友良 東京工業大学 工学部 情報工学科 田中研究室  
〒152 東京都目黒区大岡山 2-12-1 Tel: (03)5734-2831 E-mail: akiba@cs.titech.ac.jp

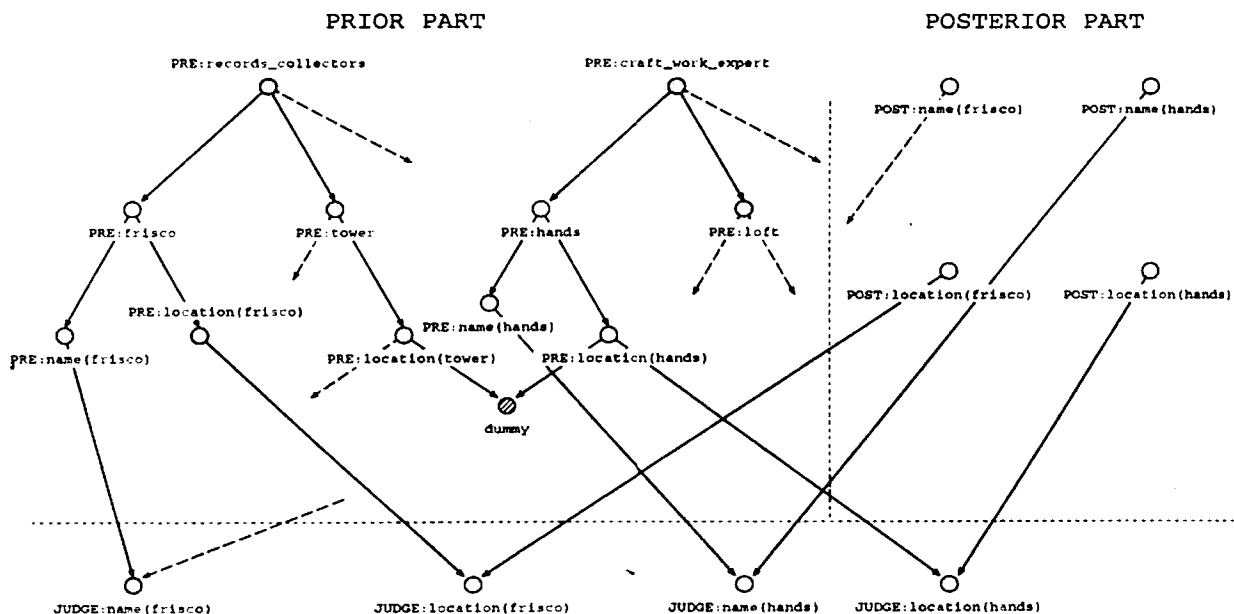


図 1: ネットワークの例 (初期状態)

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「フリスコの場所を教えてください」と質問した場合、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。名前を知っているということは、それと関連する概念を知っていることを推論する手がかりとなる。上の質問の場合、「フリスコ」という専門度の高い記録ショップの名前を知っていることから、他の記録ショップも知っていることが推論できる。

次に、ユーザが対話セッション以前から知っている知識と、対話セッションの間に獲得した知識とを明確に区別する。この区別は重要である。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。このことを表現するために、ネットワークには、対話セッション以前から知っている知識を扱う部分 (prior part) と対話セッション中に獲得する知識を扱う部分 (posterior part) の2つのサブネットワークが含まれる。同じ概念を表す2つのノードが各サブネットワークに存在し、これらには共通の1つの子ノードを持つ。この子ノードはユーザの知識の最終的な判断に使用される。すなわち、ユーザが現在ある概念を知っているのは、対話セッション以前からそれを知っていたか、セッション中にシステムから教えられたかのいずれかである。

我々のネットワークの例を図1に示す。

### 3 証拠の獲得

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  $x$  の後驗確率  $P(x|E)$  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。各ノードの後驗確率は、ネットワーク全体を評価することで求められ、これは一般に、統計シミュレーション [4] によって計算される。

ネットワークへの入力、ユーザの発話あるいはシステムの発話から得られる証拠である。これらは、それぞれ、入力解析コンポーネント、発話生成コンポーネントによって獲得される。しかし、普通、この証拠自体も確実ではない。ベ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。2つ以上の事象に対する曖昧な証拠は、複数の事象ノードの共通の子ノード (証拠ノード) を導入することで表現できる。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。証拠ノードは、 $PRE: name(tower)$ (ユーザがタワーレコードの名前を知っている) と  $PRE: name(power)$ (パワーレコードの名前を知っている) の両方の共通の子ノードとして導入し、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。

もう一つの例として、次のシステムの説明に対する、ユーザの2通りの発話を考える。

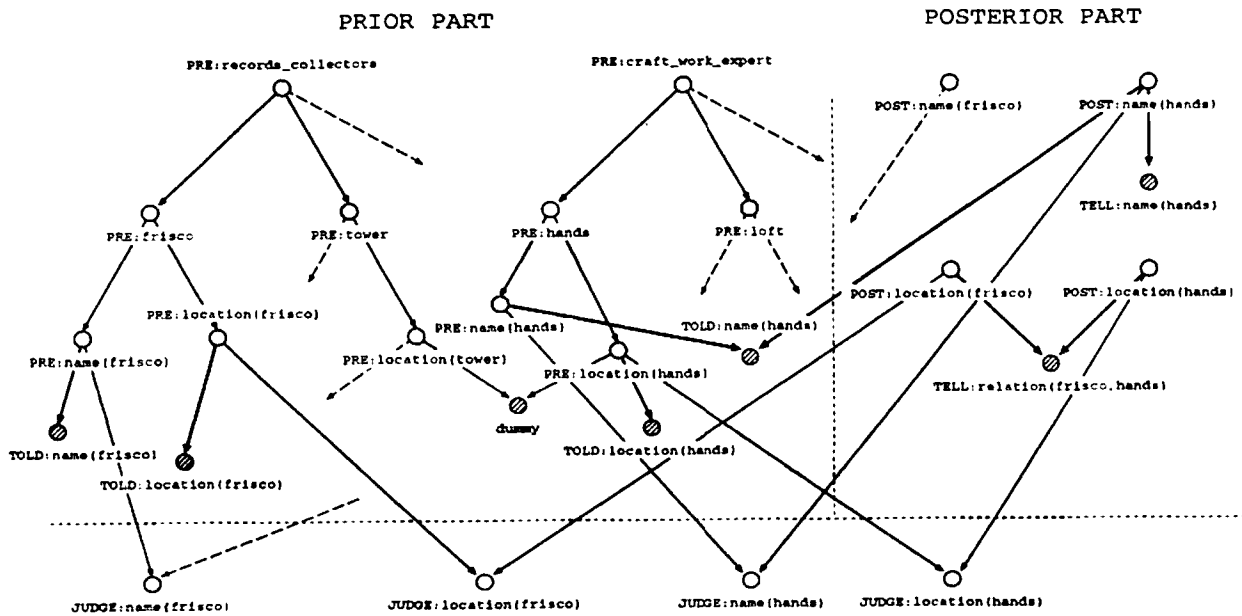


図 2: ネットワークの例 (発話 (3) の後)

(a) system: 東急ハンズはタワーレコードの正面です

(b-1) user: タワーレコードというのはどこにあるのですか

(b-2) user: タワーの場所もわかりません

(a-1) の場合は、ユーザが「タワーレコード」という名前を始めて聞いたという手がかりがある。したがって、「タワーレコード」という名前を対話セッションの前から知っていたという証拠とはならない。一方、(b-2) の場合は、「タワーレコード」という名前を以前から知っていた可能性も残される。すなわち、この名前を知っているのは、以前から知っていたのか、いまシステムから聞いたからか曖昧である。この場合も、先の例と同じように、PRE: name(tower) と POST: name(tower) の共通の子ノードとして証拠ノードを導入することで扱うことができる (図 2 の TOLD: name(hands) 参照)。以降、この 2 つのノードには依存関係が生じる。例えば、ユーザがタワーレコードの名前を知っていることを間接的に指示する証拠 (レコードショップ「フリスコ」の名前を知っている) は、PRE: name(hands) の確率をあげる一方、POST: name(hands) の確率を下げる。このように、我々のネットワークは、ユーザの知識についての発話の微妙な表現の違いを説明する表現力を持ち合わせている。どのような表現が、どのノードに、どの程度の影響を与えるかについての計算言語学的考察は、今後の興味深い課題である。

## 4 例

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 3)。システムは図 1 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

開始直後の発話であるから、入力文を解析することによって、システ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 PRE: name(frisco) の子ノード TOLD: name(frisco)、PRE: location(frisco) の子ノード TOLD: location(frisco) として導入される。これらの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  $P(X|TOLD: name(frisco) = 1, TOLD: location(frisco) = 0)$  として得られる。この場合、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ので、関連する概念の確率が上がる。

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。発話生成コンポーネントは、相手を知っていそうな場所や名前を説明に使用する。ここでは、「東急ハンズ」という語を説明に使うことを決めたと仮定しよう。

(2) system: フリスコは、東急ハンズから南へ 300 メートル行ったところにあります。

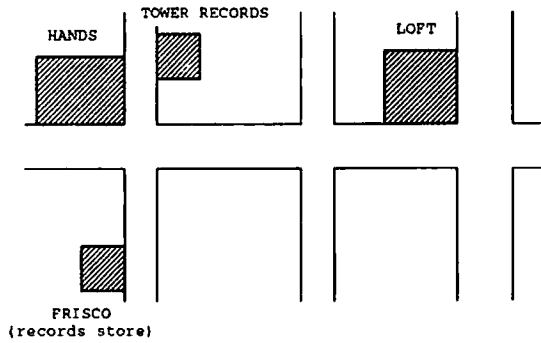


図3 ある街の地図

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: location(frisco) の確率、 $P(\text{POST: location(frisco)}|E)$  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。

(3) user: ハンズのところもわかりません

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

(4) system: ロフトから2ブロック西にあります

この説明は、「ハンズ」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: location(frisco) の確率も、ノード POST: location(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 おわりに

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 [1] と語義の曖昧性解消 [2] 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、音声対話システムを構築中 [6] 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。

表1 評価結果

node	prior	probabilities after the utterance (n)			
		(1)	(2)	(3)	(4)
(JUDGE:)					
location(frisco)	.52	.21	.43	.43	.66
location(tower)	.53	.65	.64	.33	.33
location(hands)	.53	.63	.73	.44	.74
location(loft)	.55	.56	.55	.60	.83
name(frisco)	.49	.86	.86	.84	.84
name(tower)	.49	.78	.78	.72	.72
name(hands)	.55	.58	.88	.95	.95
name(loft)	.55	.57	.57	.68	.68
(PRE:)					
records_collector	.42	.85	.86	.77	.77
craft_work_expert	.51	.54	.54	.70	.69

## 参考文献

- [1] E. Charniak and R.P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, Vol. 64, No. 1, pp. 53-79, 1983.
- [2] 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.
- [3] A. Kobsa and W. Wahlster, editors. *User Models in Dialog Systems*. Springer Verlag, 1989.
- [4] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, 1988.
- [5] 秋葉友良, 田中穂積. 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. 人工知能学会第7回言語・音声理解と対話処理研究会, 1994.
- [6] 伊藤克亘, 秋葉友良, 長谷川修, 速水悟, 田中和世. 音声対話システム構築のためのデータ収集. 日本音響学会全国大会, 1994.