

Bayesian Network を用いた意図理解

池田 徹志 荒木 雅弘 堂下 修司

京都大学 工学部 情報工学教室

1 はじめに

本研究では、Bayesian Network を動的に構築することで音声言語を理解する方式を提案する。一般に、音声認識の結果は誤りを含み、対話の履歴からの次発話予測も不確実である。Bayesian Network を用いた確率推論の枠組を用いることにより、認識結果及び構文・意味・対話の知識を確信度の伝播により融合し、最も確からしい発話の解釈・発話の意図を求めることができる。

Bayesian Network は、自然言語処理において必要な不確実な知識に基づく推論を行なうのに有望と考える。Bayesian Network を用いた自然言語処理の研究としては、プラン認識 [1]、語義の曖昧性解消 [2]、ユーザモデリング [3] などが報告されている。

2 Bayesian Network の概説

Bayesian Network [4] は、有向非循環グラフである。ノードには命題が対応し、ノードは命題の確信度を保持する¹。リンクには命題間の直接の依存関係が対応し、リンクは命題間の条件付き確率を保持する。

新しい観測 E は、ノードを加えることなどにより入力する。確信度の更新はベイズ確率の理論にしたがって行なわれ、平衡に達するとノード x の確信度は事後確率 $P(x|E)$ となる。単連結の Bayesian Network ならば、直径 d に対し $O(2d)$ で平衡に達することが保証されている。一般の

¹一般にはノードに多値変数が対応する。[T,F] の 2 値変数とすれば命題となる。

Bayesian Network の平衡状態を求める問題は NP 困難であるが、統計シミュレーションによって近似解を求めることができる。

一般の確率モデルでは、結合分布 $P(X,Y)$ を求めるのに X,Y を除くすべての変数について $P(X_1, \dots, X_n)$ を総和するため、独立と思われる 2 変数間に対しても見せかけの依存関係を考える必要がある。Bayesian Network では、概念的に関係のあるノード間にだけリンクを張り、依存関係を指定すればよい。

3 Bayesian Network を用いた発話理解

ここでは、Bayesian Network を用いて発話理解を行なう手法を示す。具体的には、今回提案するシステムの外部仕様、外部からの入力のネットワークでの表現、ネットワークの構築と最も確からしい解釈の選択について述べる。

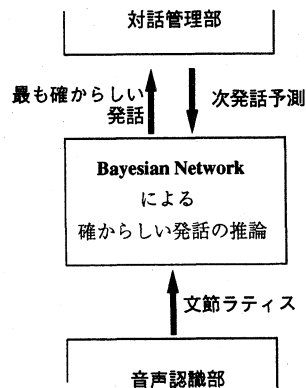


図1: システムの位置付け

3.1 観測事象・事前確率

今回提案するシステムを図1に示す。

音声認識部からの入力としては、確信度付きの文節ラティスを想定する。対話管理部からは次発話の予測が入力される。対話管理部としては2つのものを想定している。

1つは[5]のようにプランを利用して対話構造を認識・次発話予測をおこなうものである。これは対話を理解しながら次発話を予測する方法である。予測として次発話の[意図タイプ, 話題]の組が得られる。しかし、構造の理解に失敗すると正しい予測が得られず頑健性に欠ける。また、確率的な予測情報が得られない。

もう1つは意図タイプのN-gramを用いるものである[6]。発話の意図を分類して意図タイプとして定義し、対話コーパスに意図タイプのラベルを付与することで意図タイプのN-gramを求めることができる。それを次発話の意図タイプ予測に用いる。頑健であり、確率推論と相性がよい。

対話管理部からの予測として、それぞれの片方のみを用いる場合、両方を用いる場合で理解率や頑健性の違いを比較する予定である。

3.2 Bayesian Network による表現

文節ラティスが入力されると、文節仮説に対応してネットワークのノードを作る。文節ノードには事前確率として、音響的な尤度を与える。ノードは表層表現、主辞(文節の中心となる自立語)、文節の占める時間(フレーム単位)などの情報を保っている。

3.3 命題的意味の抽出とユーザ意図の推論

句仮説の推論 日本語では一般には文節がいくつかつながることで、文の中で意味を持つ構成単位になると考えられる[7]。本研究でも文単位

と文節単位の間の中身の大きさを持つ文の構成単位を考え、これを句と呼ぶ。句は発話に現れる基本的な概念に対応しており、多くの文は句を単位にして見ると単純な構造をしていると考えられる。1文節で1つの句となることもある。

1つの句は中心となる単語(主辞)とそれを修飾する文節からなると考える。日本語においては句の右端の自立語が主辞となる。主辞のあとにつく助詞によって句全体が文中で果たす役割(格)が決まる。句タイプを、主辞となり得る単語とそれを修飾する文節の入るスロットで定義する。スロットには助詞の制約と素性の制約をつける。スロットは全部埋まらなくても句として成立し得る。

我々はタスクをスケジュール管理に設定して対話システムの構築を目指しており、そこでの例を示す。「参加イベント」句の主辞としては「会議」などがある。スロットには「~の」(曜日・週)、「~の」(時間)、「~での・~の」(場所)などがある(ここで「」は助詞の、()は素性の制約を表す)。

Bayesian Network 上での実現としては、句に対応するノードを作り句を構成する文節とリンクを結ぶ(図2)。句ノードと文節ノードからのリンクが交わっているノードはフィラー(filler)と呼ばれる。このノードには、文節が句のあるスロットを占める確からしさをあたえる。その確からしさは、その文節と主辞が共起する確率を対話コーパスから求めることなどで得られる。

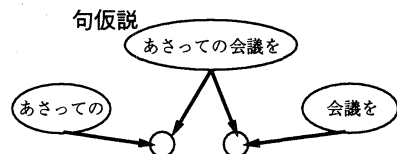


図2: 句仮説の推論

意図を表す部分の解析 1つの発話は命題的意味を表す部分と話者の意図を表す部分であるモダリティにわけられる。モダリティとしては疑問詞や文末の助動詞などがある。これを分類し、そのような表現を持つ文節にモダリティのタイプを付与する [6]。例えば「二時から会議を登録したいのですが。」という発話の「登録したいのですが」という部分はWishというモダリティをもつ。

モダリティが表層表現に対応しているのに対し、前に述べた意図タイプは深層的な意味を反映したものである。モダリティと意図には簡単に対応しない。先ほどの例では「登録したいのですが」という発話で、登録の依頼という意図をもっていると考えられる(意図タイプ request)。

Bayesian Network 上ではモダリティを含む句に対し、意図タイプノードを作りリンクを結ぶ(図3)。意図タイプノードの事前確率として、対話管理部から得られる次発話の意図タイプ予測を用いる。ここで $P(\text{モダリティタイプ} | \text{意図タイプ})$ は、対話コーパスの発話に対し両タイプを付与し、統計的に求めておく。実際に分析して得られたモダリティと意図の対応(一部)を表1に示す。

これにより意図タイプの確からしさを求めることができる。

モダリティ	意図タイプ	システム	ユーザー
Request	questionref	0	41
	request	8	29
	response+	16	17
	acknowledge+	44	1

表1: モダリティ Request の発話の意図タイプの分布(上位4候補、単位: %)

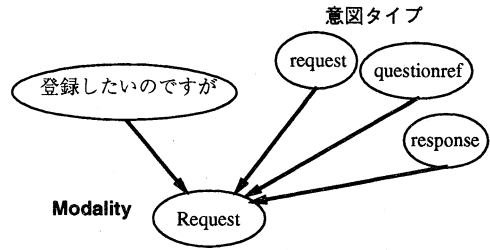


図3: 意図タイプの推論

文仮説の推論 句仮説と意図仮説から、文仮説が作られる。文タイプも同様に素性つきスロットの集合で定義する。格に対応するスロットの他、意図タイプのスロットがある。動詞が文の構造を大きく支配するので、文タイプはほぼ動詞に対応する。

例として「登録する」という動詞に対応する文タイプをみる。スロットには「~の」(曜日・週)、「~から」(時間)、「~まで」(時間)、「~での~で」(場所)、「~を」(イベント)などがある。また、意図タイプは request という制約がある。

Bayesian Network 上では文に対応するノードを作り、文を構成する句とリンクを結ぶ(図4)。句ノードと文節ノードからのリンクが交わるフィルターには、句が文のあるスロットを占める確からしさをあたえる。その確からしさは、文ノードに動詞がある場合、動詞とその句の主辞が共起する確率とする。ない場合または動詞句の場合、0.5とする。この共起確率も対話コーパスから求めることができる。

また、対話管理部としてプラン認識に基づくものを使う場合、次発話予測は[意図タイプ, 話題]の組の集合である。文候補で、同じ組を持つものは強く支持されることになるので、その文

仮説の確信度を上げるようにする。これは確率的情報ではないが、ある一定の支持を与えるようにする。

これにより一発話から得られる命題的情報と意図を表す情報、および対話文脈から得られる予測情報を統合して文仮説の確からしさを求めることができる。

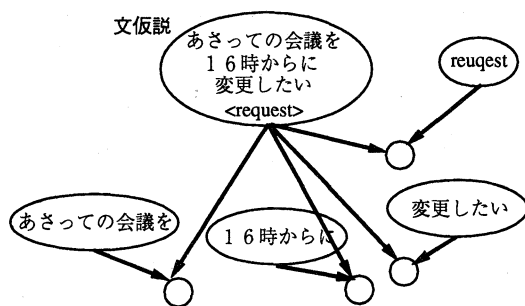


図 4: 文仮説の推論

4 おわりに

音声言語理解を行なうためには、発話の解釈の確からしさを扱う必要があることを論じ、Bayesian Network による確率推論の枠組で実現する方法を提案した。

今後、統計シミュレーションを用いた実装をして複数仮説から最尤仮説の選択の実験をする予定である。

ネットワークの構築をする際には確からしい仮説を優先的に選んで上位の仮説を組み立てていくことが望ましい。このような「焦点機構」を導入することで、ノード数の増加を抑えることができると考えられる。これを実現することは今後の課題である。

参考文献

- [1] E. Charniak and R. P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, Vol. 64, pp. 53-79, 1993.
- [2] L. M. R. Eizirik, V. C. Barbosa, and S. B. T. Mendes. A bayesian-network approach to lexical disambiguation. *Cognitive Science*, Vol. 17, pp. 257-283, 1993.
- [3] 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. 人工知能学会研究会資料, SIG-SLUD-9303-2, pp. 9-16, 1994.
- [4] J. Pearl. *Probabilistic reasoning in intelligent systems*. Morgan Kaufmann, 1988.
- [5] 山岡孝行, 飯田仁. 階層型プラン認識モデルを利用した次発話予測手法. 電子情報通信学会論文誌, Vol. J75-DII, No. 6, pp. 1203-1215, 1993.
- [6] 永田昌明, 鈴木雅実. 日英対話コーパスへの発話行為タイプ付与の試みとその統計的対話モデルへの利用. 人工知能学会研究会資料, SIG-SLUD-9302-7, pp. 49-56, 1993.
- [7] 永井明人, 石川泰, 中島邦男. 話題想起に基づく音声対話文の意味抽出方式. 音響学会全国大会講演論文集, 1-8-13, pp. 25-26, 秋季 1993.