

記憶に基づく推論方式による天気予報 —システムの性能評価—

3N-3

毛利 隆夫 田中 英彦

{mohri,tanaka}@MTL.t.u-tokyo.ac.jp

東京大学 工学部

1 はじめに

現在まで我々は、記憶に基づく推論方式 (Memory-Based Reasoning; MBR) を用いた天気予報システム WINDOM を並列計算機 AP1000 上に実装し、その評価を行ってきた。今回新たにマッチング方式を改良することによって、さらに高精度な回答が得られるようになった。

2 MBR による天気予報

MBR[SW86] はルールを用いず、大量のデータとのマッチングにより推論を行う。大量の「問題と回答」のペアがデータベースに記憶されており、新たな問題に対しては、それと最も類似した問題部を持つペアをデータベース中から検索し、そのペアの回答部をそのまま回答する。MBR は数多くの利点をもっている [北野 92]。新しい知識はデータベースに追加するだけで効果を発揮するため、システムの構築は容易であり、ルールベースの推論システムで問題になりがちな、知識獲得のボトルネックは生じない。また、マッチングはデータパラレルに行えるため、並列計算機のメリットを十分に生かせる。最近では機械翻訳 [Sat93]、タンパク質の 2 次構造予測 [ZML92] などに適用され、成果を挙げている。

一方、天気予報は、熱力学モデルを用いてスーパーコンピュータなどによる数値計算により行う方法が主流である。しかし現状では最終的な判断は熟練した予報官に任されている。このような、専門家の経験と技能に依存している天気予報は、今でも AI にとって challenging な分野であると言える。

我々は MBR を用いた天気予報システム WINDOM を並列計算機 AP1000 上に実装し、その評価を行っている [毛利 93]。予測地区は東京 (大手町) 一点のみとし、6 時間後程度の短時間予測を行う。本論文の実験結果は全て「降水有り、降水無し」の 2 カテゴリ予測である。

MBR による天気予報は、「同じような気象状況ならば、その後の天気も同じようになるだろう」という仮説に基づいて天気予報を行う。6 時間後の東京の天気を予測する場合、まず現在の気象状況と最も似ている過去の時点から、データベースから検索する。そして、その時点から

6 時間後の東京の天気を、現在から 6 時間後の東京の天気の予測値として回答するのである。

データベースには、気象庁が無料で公開している気象観測データを直接利用した。具体的にはアメダスのデータから降水量、風向、風速、日照時間、気温の 5 要素を、有人観測所のデータから海面気圧、雲量、天気の 3 項目をそれぞれ使用した。観測点は図 1 のように関東近辺を使用した。データベースには最高 9 年分 (1982 年~90 年) の観測データを蓄え、これを用いて 1991 年の天気を予測する。

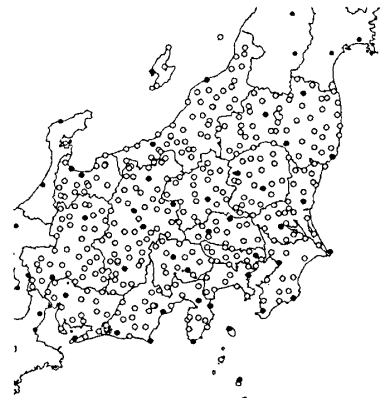


図 1: 利用した気象データの観測点
(○: アメダス, ●: 有人観測所)

3 マッチング方式の改良

本研究では、従来のシステムより高精度の結果を得るために、マッチング方式の改良を行った。

3.1 有効な観測要素の選択

従来は全観測要素をデータとして利用して予測を行っていたが、単独の観測要素のみを用いた予測の回答精度に、大きな差が見られることがわかった。このため、全観測項目から「雲量、日照時間、気温」を除外して 5 要素としたデータ (以下では *select* と呼ぶ) を用いた場合の回答精度を全要素を用いた場合と比較したところ、*select* のほうがよい精度が得られることが判明した (表 1)。

実験では 1991 年中から、重複のないように 1000 点の日時 (1 時間単位) をランダムに選択して予測を行い、回答精度を求めている。回答精度の 95% 信頼区間は、およそ回答精度 $\pm 2.3\%$ である。

⁰Weather Forecast Using Memory-Based Reasoning:
- System Performance Evaluation -
Takao MOHRI and Hidehiko TANAKA
Faculty of Engineering, The University of Tokyo

表 1: 選択観測要素による精度 (データベースは 1 年分)

観測網	観測要素	回答精度 [%]		選択
		3 時間後	6 時間後	
有人観測所	海面気圧	84.2	82.4	○
	雲量	73.7	73.1	×
	天気	87.2	84.6	○
アメダス	降水量	86.8	84.8	○
	風向	85.8	82.8	○
	風速	85.2	83.4	○
	日照時間	82.3	80.8	×
	気温	79.8	79.0	×
全観測要素		87.4	82.7	
select		88.2	85.6	

3.2 数時間分のデータによるマッチング

次に、マッチングの際の単位を、従来の 1 時間分のデータから数時間分に拡大して精度の比較を行った (表 2)。これは、ある時間幅での気象状態の遷移を見る事で、天気为上り坂なのか、下り坂なのかを分かりやすくするのがねらいである。実験の結果、数時間単位でマッチングを行った方がよい精度が得られることが判明した。表中の 2 時間分の得点に同じ重み値をかけて全体の得点とする方法を multi と呼ぶ事にし、以下の実験では select と multi を同時に使った方法を用いる。

表 2: 数時間分のデータによるマッチング

(s_n は予測開始時刻の n 時間前の事例の得点)
(データベースは 9 年分, 6 時間先を予測, †=multi)

match	select	重み付け	精度 [%]
1 時間	無し	$1.0 * s_0$	84.1
1 時間	有り	$1.0 * s_0$	84.9
2 時間	有り	$1.0 * s_0 + 0.1 * s_1$	84.9
2 時間	有り	$1.0 * s_0 + 1.0 * s_1$ †	86.0
3 時間	有り	$1.0 * s_0 + 1.0 * s_1 + 1.0 * s_2$	85.7

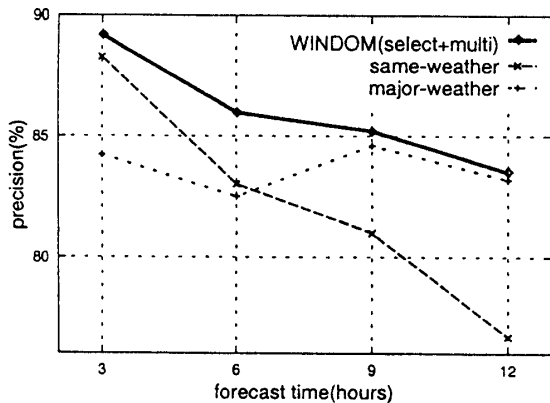


図 2: 予測時間と回答精度の関係 (データベースは 9 年分)

4 実験結果

予測時間と回答精度との関係を図 2 に示す。ここでは、前述の select+multi と、same-weather, major-weather という 2 つのトリビアルな方法とを比較してみる。same-weather は、予測開始時刻と同じ天気をそのまま回答する方法で、短時間予測では高い精度が得られるが、時間

が経つにつれて、ランダムな回答と等価になる。major-weather は、多数を占める天気を常に回答する方法で、本実験の場合は必ず「降水無し」を回答する。なお、major-category は回答から得られる情報がなく、予測としての価値はない。

図 2 において、予測時間が長くなるにつれて精度が低下するのは当然だが、本システム WINDOM の精度低下は、same-weather よりも緩やかである。9, 12 時間後の予測では major-weather と同程度の精度しか得られていないが、WINDOM は「降水有り」の予測も回答している。実際、1000 点の選択した日時の中で降水のあった 169 点のうち、53 点を正しく予測した。

また、9 年分のデータを利用した 6 時間先の予測の場合、従来の方式 (select も multi もなし) よりも 1.9% 高精度な、86.0% という回答精度が得られている (図 3)。

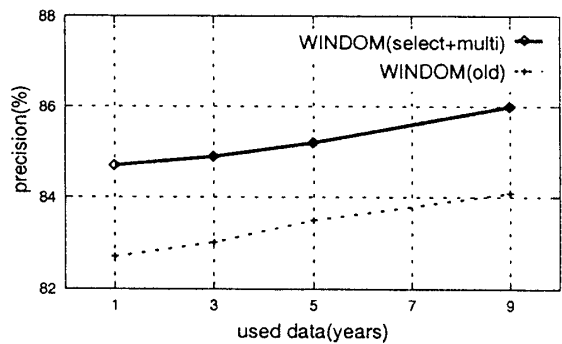


図 3: データ年数と回答精度との関係 (6 時間先を予測)

5 結論および今後の課題

数時間分のデータを用いる場合には各時間の得点を合計する際の重み付けが重要になるが、データベース年数や予測時間などの条件によって良い回答精度を生じる重み値が一定しない場合がしばしばあった。安定して高精度が得られる重み値や合計方法を求めるのは今後の課題である。他にも、有効なデータ観測点の選別、回答の分布の定量的な評価などが今後の課題として挙げられる。

なお本研究は、(財)人工知能研究振興財団の助成金 4AI 第 305 号を受けている。

参考文献

[Sat93] Satoshi Sato. Example-based translation of technical terms. In *The Fifth Int. Conf. on Theoretical and Methodological Issues in Machine Translation*, 1993.

[SW86] Craig Stanfill and David Waltz. Toward memory-based reasoning. *Communications of the ACM*, Vol. 29, No. 12, pp. 1213-1228, December 1986.

[ZML92] Xiru Zhang, Jill P. Mesirov, and David L. Waltz. Hybrid system for protein secondary structure prediction. *Journal of Molecular Biology*, Vol. 225, pp. 1049-1063, 1992.

[北野 92] 北野宏明. 超並列人工知能. 人工知能学会 Vol.7, No.2, pp. 244-262, 3 1992.

[毛利 93] 毛利隆夫, 中村真章, 田中英彦. 記憶に基づく推論方式による天気予報. 人工知能学会全国大会 (第 7 回) 論文集, 1993.