

記憶に基づく推論による天気予測

Weather Forecasting by Memory-Based Reasoning

毛利 隆夫*

Takao Mohri

田中 英彦*

Hidehiko Tanaka

* 東京大学大学院工学系研究科情報工学専攻

Information Engineering Course, Graduate School of Eng., The University
of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113 Japan.

Keywords: weather forecast, memory-based reasoning, case-based reasoning, classification, concept learning.

英文概要

We have implemented a weather forecasting system, WINDOM, which uses memory-based reasoning(MBR). The observation data from Japan Meteorological Agency(JMA)'s networks, AMeDAS and the surface observation centers, are used directly as a case database. WINDOM runs on the parallel computer AP1000 and predicts the weather around Tokyo several hours ahead by matching the past data in the database with the present observational data. In MBR, the weighting methods for features and its metric definition are important to get high accuracy. In the experiments, four weighting methods based on conditional probability and three calculation methods of similarity are tested, and the best pair of methods is selected empirically. The relation between accuracy and parameters such as the area of the observational data, the quantity of data, and prediction hour are shown. By using 9 years of observational data in the middle and the west part of Japan to predict whether the weather of Tokyo 6 hours ahead would be rainy or not, an accuracy of 87.2% is achieved. The average accuracy of prediction in Kanto-area is slightly worse than that of JMA, but in some prefectures our results are comparable with their average results.

1 まえがき

エキスパートシステムを作成する場合、知識ベースの作成に多大なコストが必要であることが、知識獲得のボトルネック問題として広く知られている。専門家の知識をルール形式で書き下すには長時間の専門家との対話が必要であり、得られたルール同士の干渉も問題になる。現在、この知識獲得のボトルネックの解消を目的とした様々な研究が活発に行われている。記憶に基づく推論 [SW86, 北野 92] はルールを用いず、事例をそのままの形式で保持することで知識ベースを構築する。そのため知識獲得ボトルネックの解消方法の一つとして注目されている。

我々は記憶に基づく推論を天気予測に応用したシステム WINDOM を作成した。WINDOM はアメダスなどの気象庁の観測網から得られた 9 年分の観測データを事例として持ち、数時間先の東京近辺の天気を予測する。

記憶に基づく推論では、以下で述べるように類似度の定義が大きな意味を持つ。そのため本論文では、まず条件付き確率を基にした特徴重み付け方法を紹介し、ベンチマークデータによってその評価を行う。次にその結果を利用して天気予測を行い、利用する事例の量や、観測データの領域と正答率の関係を調べ、記憶に基づく推論による天気予測の特性および有効性を明らかにする。

なお、一般に使われている「天気予報」という言葉は、天気を予測して広報する意味であるが、本研究は予測のみを行うため「天気予測」という言葉を用いている。

2 MBR による天気予測

2.1 MBR

従来のエキスパートシステムでは、通常 if-then 型のルールを蓄えた知識ベースを構築し、ルールを何段にも連鎖させて推論を行なう。このようなルールに基づく推論 (Rule-Based Reasoning: RBR) と記憶に基づく推論 (Memory-Based Reasoning: MBR) は、知識ベースの持ち方が大きく異なる。MBR では問題と回答のペアを大量にデータベースに蓄えておき、新たな質問には、データベース内から最も類似した事例を検索し、その類似事例の回答をそのまま質問の回答とするのが基本的な方法である。

このように MBR では、データベースを直接利用して推論を行なうために、エキスパートから知識を引き出す必要がない。そのため知識獲得が容易で、知識の追加・削除が容易、回答の説明が容易、システムの構築が

短期間で行なえるなどの特徴をもつ。MBR はこれまでに英単語の発音問題 [SW86]、機械翻訳 [Sat93, Kit93, SOF+93]、タンパク質の 2 次構造の予測 [ZML92]、米国情勢調査の職業欄の分類 [CMSW92] などに応用されている。

MBR は RBR を大幅に簡略化した枠組みだと考えることができる。RBR が述語論理の枠内でのルールだとすれば、MBR の保持する事例は命題論理でのルールだとみなすことができる。すなわち、RBR ではルールに変数が存在するのに対して、MBR での「ルール」には変数がない。また、ルールが多段に連結されることはなく、1 段しか適用されない。

MBR と類似した枠組みとして、事例に基づく推論 (Case-Based Reasoning: CBR) [Ham86, 小林 91] がある。MBR と CBR とは、事例をもとに推論を行う点では同じであるが、CBR では得られた類似事例の回答を質問に符合するように修正を行うのに対して、MBR では一切修正を行わない。CBR では事例の修正のおかげで典型的な事例のみを事例ベースに保持すればよいが、その反面事例の修正のためのルールが必要になり、やはり知識獲得のボトルネック問題に突き当たる。MBR では事例の修正を最初からあきらめ、事例ベース中の類似事例の回答部分をそのまま質問の回答とする。事例の修正の知識は不要だが、その分事例を大量に揃える必要がある。また、事例ベース中にない回答を生成する能力はなく、回答カテゴリが既知の分類問題のみを対象にしている点が特徴である。

実装レベルでみた場合、MBR はアルゴリズムがデータパラレルな並列性の高い方式であり、事例数だけの並列度が期待できる。そのため並列計算機上に実装した場合にはプロセッサ台数に比例した並列度が期待でき、並列計算機の性能を十分引き出すことが可能である。

2.2 天気予測

現在気象庁で行われている天気予測は、2 つの段階からなっている。第 1 段階は数値予測で、熱力学モデルに基づいてスーパーコンピュータにより差分方程式を数値計算し、定められた時間後の気温、気圧などを求めて、予想気象配置図を作成する。第 2 段階では、第 1 段階で得られた予想気象配置図から「晴れ、曇り、雨」といった天気を推測する。この過程は天気への翻訳と呼ばれており、現在日本の気象庁では、統計的な回帰分析を用いた MOS (Model Output Statistics) と呼ばれる方法が用いられている。しかし、天気予測は完全に自動化されているわけではなく、まだまだ熟練した予報官の経験と勘に頼っている部分が多い。例えば最近では、熟練した予報官が大勢退職した際に、予

測の正答率が停滞したことがあったという [気象 91] .

2.3 MBR による天気予測

このように熟練者の経験と勘が要求される天気予測は AI にとって興味深い応用分野である . 天気予測に AI 手法を導入する研究は現在までに多数行なわれている [RE85, S.Z88] . これらの研究の多くは基本的にルールベースであるが , ニューラル・ネットを用いて雷雨を予測する例も報告されている [McC92] . しかし MBR のように , 大量のデータとのマッチングによる天気予測の試みはなされていない . また , 多くの研究例では雷雨や嵐などが対象であり , WINDOM のように「晴 , 雨」といった天気を予測するシステムは少ない . また , MBR では過去の経験に相当する事例を事例ベースに蓄えることで経験からの学習を行っているが , 予報官も同様に長年の経験を蓄積することで判断を学習しており , 両者の類似性は興味深い .

MBR による天気予測では , データベース中の事例とのマッチングによって予測を行う . これは「気象配置 (降水量や気圧などの分布) が似ているならば , その後の天気も似ていだろう」という , もっともらしい仮説に基づいている . 例えば , 現在から 6 時間先の天気を予測したい場合には , データベースに蓄えている過去の事例の中から , 現在の気象配置に一番似ている事例を探し出してきて , その事例の時点から 6 時間後の天気を , 現在から 6 時間後の天気の予測とするのである .

マッチングの際のデータには , 気象庁の観測網からの気象データを直接用いた . これらの気象観測データは無償で公開されており , 大量のデータが容易に入手・利用できる . そのため , MBR によるシステム作成は容易である .

このマッチングによる天気予測は , 従来の熱力学モデルに基づく方法と比べて , 観測データを直接用いたことによる利点がある . 熱力学モデルによる方法では , モデルの境界地点のデータを定めるのは簡単ではないがマッチングによる方法では , 固定された観測地点の観測データを直接用いるために , その地域の気象の癖や , 山 , 海岸線などの境界条件も最初から観測データに含まれている . そのため新たに境界条件を考慮したり , マッチングのためにデータを補正する必要は全くない .

MBR による天気予測は , 古来人間が周囲の山にかかる雲を見るなどして行なってきた天気予測や , 数値計算が導入される以前の気象庁の天気予測に類似している . しかし , 通信網の発達により広範囲の詳細な観測データの収集が短時間で行なわれるようになった点 , また計算機の高性能化により , 大量の高精度な

データをマッチングに用いることができるようになった点や , マッチング方法自体の定量的な研究に立脚している点などが , 旧来の方式とは根本的に異なる .

3 実装

3.1 並列計算機 AP1000

我々は富士通の並列計算機 AP1000 [SIH92] 上に , MBR による天気予測システム WINDOM を実装した . 我々が使用した AP1000 は 64 プロセッサ構成であり , 一つのセル (プロセッシング・エレメント) は , SPARC IU +FPU (clock 25MHz, 15MIPS) と 16MB の RAM で構成されている . AP1000 はディスクを持っておらず , すべてのデータはホストである Sun Workstation 上のディスクから読み込まれ , 各セルに転送される . WINDOM では各セルで動作するプログラムは同じであるが , それぞれ互いに異なる日時のデータを受け持ち , 並列にマッチングを行なっている .

3.2 観測データ

MBR では問題部と回答部のペアを事例として大量に蓄えるが , 本研究では , 気象要素の観測データを問題部に , 晴れ , 雨といった天気を回答部に割り当てた . 具体的には , 問題部に気象庁の観測網であるアメダスと有人観測所の観測データを用いた . アメダスの観測点は約 17km 間隔で日本全体に敷設されており , 降水量 , 風向 , 風速 , 日照時間 , 気温の 5 要素を 1 時間毎に観測している . WINDOM では , そのうち降水量 , 風向 , 風速の 3 要素を用いた . 有人観測所は約 150km 間隔で設置され , 3 時間毎に 1 3 項目を観測しているが , その中から 海面気圧 , 天気の 2 項目を利用した . 観測要素は , 一旦人間が選択した後でそれぞれ単一要素だけを用いて予測を行い , 高い正答率が得られた観測要素だけを最終的に利用した [毛利 93a] . 事例の回答部には , 予測時間 (例えば 6 時間) 先の東京 (大手町) の天気を割り当てた . この天気には 1 5 区分された天気をを用い , 回答生成時に「晴 (降水なし) , 雨 (降水なし)」の 2 カテゴリにまとめて回答している .

利用する観測点は , 天候は概ね西から変化するとの仮定をもとに , 関東甲信越以西に限定しており , アメダスは約 960 点 , 有人観測所は約 120 点の観測点を利用している (図 1 参照) .

観測データの各気象要素が連続値であった場合には , それぞれ 10 領域程度に離散化した個々の領域を一つの特徴として取り扱った . 結果として観測点毎 , 観測要素毎 , シンボル値または離散値区間毎に異なる特徴として扱われている . そのため , ひとつの事例は



図 1: 利用する観測点
(○: アメダス, ●: 有人観測所)

約 20,000 個の特徴から構成されている。なお 1 年分のデータベースの大きさは約 44 MB であった。

3.3 特徴の重み付けと類似度の計算方法

MBR ではマッチングのみによって推論を行なうため、類似度の決定方法が回答に大きな影響を与える。今回我々は条件付き確率を基にして特徴の重み値を決定する方法を使用した。MBR での類似度の計算では、人間がその領域の専門知識を考慮して、類似度の計算式や計算に用いる特徴の重み値を決める方法 [SW86, ZWM88] もよく用いられるが、条件付き確率を基にした方法は、対象領域の専門知識を必要としない汎用な方法である。

条件付き確率を基にした特徴重み付け方法として、図 2 にある 4 種類を取り上げる。Per-category feature importance (PCF) は、特徴が決まった際の回答カテゴリの確率を重み値とする。天気予測の場合では、例えば静岡の天気が雨だった場合、6 時間後に東京が雨である確率が高ければ、静岡の天気が雨であるという特徴の重みが大きくなる。PCF は特徴の出現頻度には影響を受けないが、頻度の多い回答カテゴリの重み値が大きくなる。

Per-feature category importance (PFC) は、同様に回答カテゴリが決まった際の特徴の確率を重み値とする。PFC は頻度の多い特徴の重み値が大きくなる。

Averaged category feature importance (ACF) は、PCF を回答カテゴリの出現確率 $P(C)$ で正規化した重み値である。これは、PFC を特徴の出現頻度 $P(F)$ で正規化した値と同じになる。

Cross-category feature importance (CCF) は、PCF を全回答カテゴリに対して自乗和をとり平均化

したものであり、回答カテゴリによらず同じ値をとる。

PCF および CCF は [CMSW92] に紹介されていた方法で、PFC および ACF は我々が考案した方法である。

事例間の類似度は、マッチングの際に一致した特徴の重み値を足し合わせるなどして計算されるが、この類似度の計算方法として、SUM, MAX, ERROR の 3 つの方法 [CMSW92] を試した (図 3)。SUM は一致した特徴の重み値の和を求める方法であり、MAX はその中の最大値だけを利用する。ERROR は SUM と MAX の双方の特徴を持っている。つまり、SUM のようにすべての特徴の重み値を考慮するが、重みが 1 であるような特徴がひとつ一致するだけで類似度が 1 になる。

4 実験結果

4.1 実験 1: 特徴重み付け方法のベンチマークテストによる比較

本実験では、PCF, PFC, ACF, CCF の 4 種類の特徴重み付け方法の特性をベンチマークデータを用いて明らかにする。類似度の計算方法としては、最も基本的な SUM を用いている。同時に他の重み付け方法として、IB4 [Aha92], VDM [SW86], および決定木を生成する C4.5 [Qui93] についても試験し、比較対象とした。ベンチマークデータは、UCI 機械学習データベース [MA95] または [Qui93] の付属ディスクから引用した。回答には、1 つの近傍だけとのマッチング結果を用いている。特徴値が連続値である場合には、値の頻度が 5 等分されるように区間を離散化した (C4.5 は独自の方法による離散化を行っている)。また正答率は、データをランダムに 10 等分した cross-validation [WK91] により求めた。結果を表 1 に示す。

正答率の順位は、必ずしも固定されていないが、概ね CCF, PFC, PCF, ACF の順になっている。また CCF は C4.5, IB4, VDM と比較しても遜色のない結果が得られている。

4.2 実験 2: 特徴重み付け方法と類似度計算方法

以下の天気予測に関する実験では、最大 9 年分 (1982 ~ 1990 年) の観測データをデータベース内に訓練データとして蓄えておき、91 年内から予測を行なう日時をランダムに 1000 点選び出し、その日時から 3, 6, 9, 12 時間先の東京の天気を予測した。回答の二項分布を正規分布で近似した場合、グラフで示した正答率の 95% 信頼区間の幅は、87% の正答率に対して $\pm 2.1\%$ である。

表 1: ベンチマークデータによる特徴重み付け方法の比較 (値は正答率 [%])

データ名	iris		vote		soybean		crx		hypo		hepatitis			
データ数	150		300		683		490		2514		155			
連続値特徴数	4		0		0		6		7		6			
シンボル特徴数	0		16		35		9		29		13			
順位	1	CCF 94.7	IB4 94.7	CCF 92.2	IB4 86.7	C4.5 99.6	CCF 81.3	2	C4.5 94.7	VDM 93.0	VDM 91.8	PFC 85.7	PCF 92.3	C4.5 80.0
	3	PFC 93.3	C4.5 92.3	IB4 90.3	CCF 83.7	ACF 92.3	PCF 79.4	4	VDM 94.7	CCF 91.0	C4.5 89.6	VDM 83.3	VDM 90.7	VDM 79.4
	5	PCF 92.7	PCF 88.0	PFC 79.8	C4.5 82.7	CCF 89.2	ACF 79.3	6	ACF 85.1	PFC 88.0	PCF 51.0	PCF 80.0	PFC 70.2	IB4 78.7
	7	IB4 81.3	ACF 82.0	ACF 42.5	ACF 67.5	IB4 67.7	PFC 78.7							

またマッチングに際しては、データベース内のデータのうち質問日と同じ季節(前後を含めて3カ月)とだけマッチングを行なっている。回答を決める際には、予備実験の結果をもとに、高い類似度をもつ上位20事例の回答に対して多数決を行なった。すなわち $k=20$ とした k -nearest neighbor 法を用いている。実行時間は、9年分のデータベースを全セルに読み込むなどの初期化に約1000秒、1回の予測には4.2秒を要した。

本実験では、4種類の条件付き確率による特徴重み付け方法 PCF, PFC, ACF, CCF と、類似度の3種類の計算方法 SUM, MAX, ERROR について、組合せで比較を行ない、天気予測にとって最もよい手法の組合せを求めた。実験結果を表2に示す。訓練データは1年分(1990年のみ)である。なお、ERRORは重み値が0から1までの範囲内であることが前提とされているが、ACFの重み値はこの範囲に留まらないため、ACFとERRORの組み合わせはテストしていない。

類似度の計算方法では、MAXは単一の特徴しか考慮しないため悪く、SUMとERRORでは有意な差は得られていない。特徴の重み付け方法ではCCFが最も良い。なお以下の実験では、この実験の結果を元にCCFとERRORの組合せを用いている。

4.3 実験3: 特徴の重み値の分布

本実験では、9年分の観測データを用いて、CCFにより計算された特徴の重みを観察した。得られた重み値の分布を図4.5,6に示す。図中の点は東京の天気を予測する際、その地点のアメダスの降水量が1mm以上という特徴の重み値であり、半径が重み値の大きさを示している。重み値の大きい地点ほど回答に与える影響が大きい。3時間先の予測では東京近辺に大き

表 2: 重み付け・類似度計算方法の比較

重み付け方法	正答率 [%]		
	SUM	MAX	ERROR
PCF	82.5	82.6	82.5
PFC	83.9	43.4	82.5
ACF	77.1	33.8	—
CCF	84.7	81.3	85.0

い重み値の特徴が集まっているのが、予測時間が長くなるにしたがって西に移動している。これは天気が概ね西から変化することに対応している。また、予測時間が長くなると特徴の大きな点が少なくなるが、これはどの観測点が重要かが決まりにくくなっていくことを示している。なお、図4からは通年の事例による重み値であるが、各季節に限定した事例で重み値を計算し直した場合、春(3~5月)では、予報時間が先になるに従って西側に大きい重み値が現れる傾向が顕著であり、冬(12,1,2月)では北側の観測点が大きい重み値を持つ傾向が観察できた。

4.4 実験4: データ量と正答率との関係

本実験では、データの年数と正答率との関係を調べた。結果を図7に示す。このように、データ量が増加するにつれて、正答率は上昇傾向にあるのがわかる。

マッチングの際は、質問日の前後月を含めて3カ月間のデータを用いているが、本実験では更にデータを間引いて1/2, 1/3とした場合の正答率についても、1/2年, 1/3年と称して図7に記載した。正答率は、

- PCF(Per-Category Feature importance)

$$w(a, c) = P(c|a)$$

- PFC(Per-Feature Category importance)

$$w(a, c) = P(a|c)$$

- ACF(Averaged Category Feature importance)

$$w(a, c) = P(c \wedge a) / (P(c) \cdot P(a))$$

- CCF(Cross-Category Feature importance)

$$w(a, c) = w(a) = \sum_{c=1}^{N_c} P(c|a)^2$$

ただし, c : 訓練事例の属する回答カテゴリ
 a : 事例の特徴
 N_c : 回答カテゴリの数

図 2: 条件付き確率に基づく特徴重み付け手法

SUM:

$$Similarity(u, v) = \sum_i w(a_i, c_u) f(u_i, v_i)$$

MAX:

$$Similarity(u, v) = MAX_i \{w(a_i, c_u) f(u_i, v_j)\}$$

ERROR:

$$Similarity(u, v) = 1 - \prod_i \{1 - w(a_i, c_u) f(u_i, v_j)\}$$

ただし, u : 訓練事例
 v : テスト事例 (質問)
 a_i : i 番目の特徴
 c_u : 訓練事例 u の属するカテゴリ
 $f(x, y) = 1$ (if $x = y$)
 $= 0$ (それ以外)

図 3: 類似度の計算方法

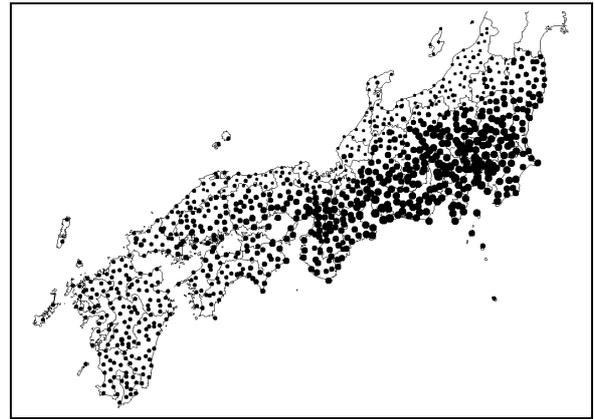


図 4: 3 時間後予測での降水特徴の重み値

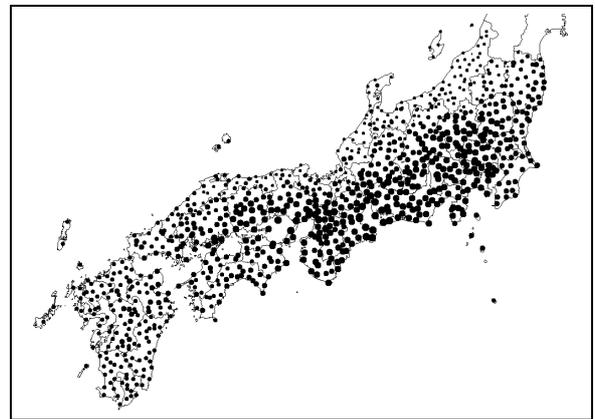


図 5: 6 時間後予測での降水特徴の重み値

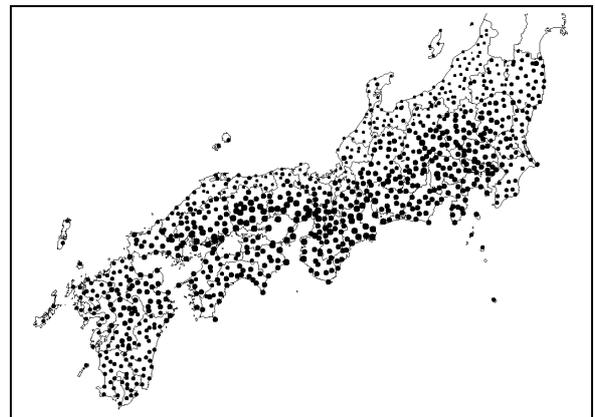


図 6: 9 時間後予測での降水特徴の重み値

データ量が1年の時の正答率に比べて、1/2年と1/3年の場合の方がよい値を示している場合がある。この結果は、単純にデータの量をそろえればよい訳ではなく、データの質も考慮すべきであることを示している。正答率に寄与するようなデータの選択は、今後の興味深い課題である。また、12時間先の予測は、データ量を増やした時の正答率の伸びが顕著であり、事例を追加することでさらに正答率の向上が期待できると考えられる。

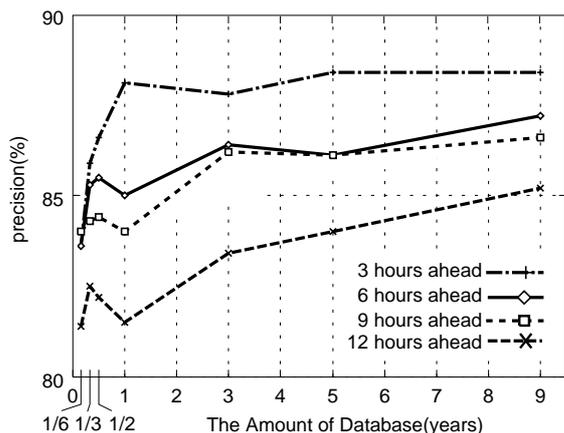


図 7: データ量と正答率との関係

4.5 実験 5: 使用観測地点と正答率の関係

本実験では使用観測地点を変化させた場合の正答率の変化を測定した。結果を表3に示す。利用する観測地点の領域を拡大すると、正答率は上昇する傾向にあることが分かる。この結果は、4.3で示した大きな重みを持つ特徴の分布とも符合する。

表 3: 利用した観測地点の領域と正答率 [%]

観測地点	予測時間	
	6 時間後	12 時間後
大手町 (一点)	82.3	81.8
東京都	84.2	82.3
関東甲信越	84.0	82.9
関東甲信越 ~ 中部	85.2	84.2
関東甲信越 ~ 近畿	86.0	84.6
関東甲信越 ~ 中国・四国	86.2	84.9
関東甲信越 ~ 九州	87.2	85.2

4.6 実験 6: 予測時間と正答率の関係

本実験では、予測時間と正答率の関係を調べるため、3時間前から12時間先までの予測を3時間毎に行ない、それぞれ正答率を求めた。結果を図8に示す。予測時間が短い場合には非常に高い正答率が得られていることがわかる。このことは、マッチングによる天気予測が短時間予測では有望であることを示している。より長時間の予測において高精度を保つのは今後の課題である。

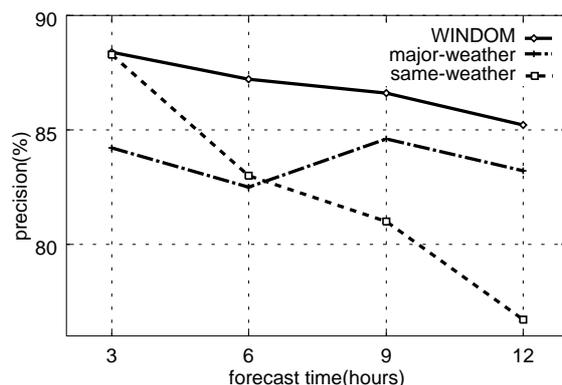


図 8: 予測時間と正答率の関係

図8では、2つの簡単な方法との比較を行っている。ひとつは、必ず多数派カテゴリを回答する方法 (*major-weather*) である。東京の天気の分布は、2カテゴリ (晴または雨) の場合、約 84% が晴に分類される。したがって、常に晴と答える「予測」では、約 84% の正答率が得られる。もう一つは、予測開始時と同じ天気を回答する方法 (*same-weather*) である。多くの場合天気は連続的に変化するため、1,2 時間先であれば現在と同じ天気になる可能性は高い。

図8をみると、*same-weather* の正答率は3時間先の予測では高いものの、その後急速に低下しているのがわかる。一方 WINDOM は予測時間が先になっても、正答率の下降はなだらかである。*major-weather* は、ランダムに選択した1000時点から先の天気は予測時間によって異なるため多少ふらつくが、概ね 84% 程度の正答率である。予測時間が12時間先になると WINDOM と正答率が近づくが、予測の内容は全く異なっている。異なる時刻に対して1000回行ったテストにおいて、12時間後に降水が有る場合は167回あったが、WINDOM はそのうち61回を正答しているのに対して、*major-weather* は1回も当てていない。

4.7 実験 7: 気象庁による予測の正答率との比較

気象庁では日常的に天気予測が行われており、一部の予測の正答率が公表されている [気象 92]。降水有無の予測に関して正答率が公表されている区域で、東京を含む最小のものは、関東甲信地方 (関東 1 都 6 県 + 山梨 + 長野) であり、1 ヶ月毎に正答率が公表されている。本実験では、公表されている予報の中から朝予測を対象にして気象庁の予測と WINDOM とを比較した。朝予測は午前 6 時に何種類か発表されるが、そのうち 6 時から 18 時までの 12 時間の予測が最も短時間であり、これを比較対象にした。

WINDOM は特定の地点の固定された時間先の予測を行う方法であるが、気象庁の予測は地域を対象にし、時間の幅も持っている。そこで比較のために、予測区域内の複数の地点で 1~12 時間後の複数の予測時刻に対して予測を繰り返し行い、その結果を平均した。結果を表 4 に示す。正答率は、平均値では気象庁と比べて、やや劣っている。各都県毎の正答率はバラツキが大きいが、東京都、神奈川県などは気象庁の平均正答率に近い値が得られている。また長野県の正答率は他都県に比べて大きく下回っており、今後の解析が必要である。

表 4: 気象庁による予測の正答率との比較

方法	予測区域	正答率 [%]
WINDOM	東京都	86.3
	神奈川県	87.9
	埼玉県	84.6
	千葉県	84.0
	群馬県	83.2
	栃木県	86.1
	茨城県	82.7
	山梨県	83.7
	長野県	74.4
	関東甲信平均	83.7
気象庁	関東甲信平均	87

5 考察および今後の課題

条件付き確率を基にした 4 種類の特徴重み付け方法では、天気予測を対象とした場合、事例の属する回答カテゴリに依存しない重み値である CCF が最もよい結果が得られた。PCF, PFC, ACF では回答カテ

ゴリ毎に異なる重み値を計算しており、きめ細かな情報を用意しているにもかかわらず、よい正答率が得られなかったのは意外である。一方ベンチマークデータによるテストでは、概ね CCF が良いものの、常に残りの 3 つと比べて良い訳ではない。これは、データの特性とアルゴリズムの特性の相性によるものだと考えられるが、こういった特性を陽に扱う研究は少なく、今後の研究の進展が望まれる。

天気予測では、基盤となる観測データは有限の観測点からしか得られず、また気象現象がカオスであることを考えると 100% の予測はそもそも不可能であろう。ただ、現在の我々のシステムの正答率は、向上の余地が残されていると考えている。

まず、現在の我々のシステムは、気象庁で用いられている入力データのすべてを利用している訳ではない。特に海洋上や高層気象のデータは、予測に重要であると考えられる。これらの新たなデータの追加により、更なる精度向上が期待できる。

また現在は連続値観測データの離散化では人手により区間分割を行ったが、情報量規準により最適な区間分割を行う手法 [F193] などが研究されており、これらの適用も考えている。

6 まとめ

MBR による天気予測システム WINDOM を実装し評価を行なった。特徴の重み付け方法として条件付き確率を基とした 4 つの方法 (PCF, PFC, ACF, CCF) と 3 つの類似度計算方法 (SUM, MAX, ERROR) を試し、実験により最も良い正答率をもたらす組み合わせ (CCF + ERROR) を求めた。ベンチマークデータを用いた実験によって、条件付き確率に基づく重み付け方法が、VDM, IB4 といった他の重み付け方法や、決定木生成手法 C4.5 と比較しても、遜色のない正答率が得られることを示した。天気予測の実験では、正答率の向上に利用する観測データ領域の拡大やデータ量の増加が効果的であることを確認し、6 時間先の東京 (大手町) の降水有無の予測では、87.2% という高い正答率を得た。また、気象庁の正答率との比較を関東甲信地方を対象に行なったところ、平均値では劣るものの、一部の地区では気象庁の平均正答率と同等の正答率が得られており、MBR による天気予測の有効性が実証できた。

謝辞

本システムの初期版の共同作成者である現 NEC の中村真章氏に感謝します。David W. Aha 氏には、彼

のIB4プログラムを使用する機会を与えて頂きました。また本研究を進めるにあたり、気象庁の方々、特に天気相談所および気象庁数値予報課の尾崎友亮氏に大変お世話になりました。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

- [Aha92] David W. Aha. Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms. *Int. J. Man-Machine Studies*, Vol. 36, pp. 267-287, 1992.
- [CMSW92] Robert H. Creedy, Brij M. Masand, Stephen J. Smith, and David L. Waltz. Trading MIPS and memory for knowledge engineering. *Communications of the ACM*, Vol. 35, No. 8, pp. 48-63, August 1992.
- [FI93] Usama M. Fayyad and Keki B. Irani. Multi-interval discretization of continuous-valued attributes for classification learning. In *13th International Joint Conference on Artificial Intelligence(IJCAI-93)*, pp. 1022-1027, 1993.
- [Ham86] Kristian J. Hammond. Chef: A model of case-based planning. In *AAAI-86*, pp. 267-271, 1986.
- [Kit93] Hiroaki Kitano. A comprehensive and practical model of memory-based machine translation. In *13th International Joint Conference on Artificial Intelligence(IJCAI-93)*, pp. 1276-1282, 1993.
- [MA95] P. M. Murphy and D. W. Aha. UCI repository of machine learning databases. Irvine, CA: University of California, <ftp://ics.uci.edu/pub/machine-learning-databases>. 1995.
- [McC92] Donald W. McCann. Forecasting technique: A neural network short-term forecast of significant thunderstorms. *Weather and Forecasting*, Vol. 7, pp. 525-534, Sep 1992.
- [MNT93] Takao Mohri, Masaaki Nakamura, and Hidehiko Tanaka. Weather forecasting using memory-based reasoning. In *Second International Workshop on Parallel Processing for Artificial Intelligence (PPAI-93)*, pp. 40-45, 1993.
- [Qui93] J. Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.
- [RE85] Johannes de Haan Renée Elio. Knowledge representation in an expert storm forecasting system. In *Proceedings of the ninth international joint conference on artificial intelligence (IJCAI-85)*, pp. 400-406, 1985.
- [Sat93] Satoshi Sato. Example-based translation of technical terms. In *The Fifth Int. Conf. on Theoretical and Methodological Issues in Machine Translation*, 1993.
- [SIH92] Toshiyuki Shimizu, Hiroaki Ishihata, and Takeshi Horie. Low-latency message communication support for the AP1000. *19th ICSA*, 1992.
- [SOF⁺93] Eiichiro Sumita, Kozo Oi, Osamu Furuse, Hitoshi Iida, Tetsuya Higuchi, Naoto Takahashi, and Hiroaki Kitano. Example-based machine translation on massively parallel processors. In *13th International Joint Conference on Artificial Intelligence(IJCAI-93)*, pp. 1283-1288, 1993.
- [SW86] Craig Stanfill and David Waltz. Toward memory-based reasoning. *Communications of the ACM*, Vol. 29, No. 12, pp. 1213-1228, December 1986.
- [S.Z88] S.Zubrick. *Validation of a Weather Forecasting Expert System*, Vol. 11 of *Machine Intelligence*, chapter 18, pp. 391-422. Clarendon Press, 1988.
- [WK91] Sholom M. Weiss and Casimir A. Kulikowski. *Computer Systems That Learn*. Morgan Kaufmann, 1991.
- [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.
- [ZWM88] Xiru Zhang, David Waltz, and Jill Mesirov. Protein structure prediction by memory-based reasoning. RI88-3, Thinking Machine Corporation., December 14 1988.
- [気象 91] 気象庁予報部. 天気予報ガイドンスの解説, 1991.
- [気象 92] 気象庁天気相談所. 降水の有無予報及び最高・最低気温予報の評価結果 (1991年), 1992.
- [小林 91] 小林重信. 事例ベース推論の研究課題. 情報処理学会研究会報告, Vol. 91, No. 16, pp. 29-38, 1991.
- [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 93a] 毛利隆夫, 田中英彦. 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 -. 第47回情報処理学会全国大会講演論文集 (2), pp. 41-42, 1993.
- [毛利 93b] 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, pp. 43-46, 1993.