

天気図理解システムにおける
推論規則の評価と改良

田島幸夫、石塚均
吉田雄二、福村晃夫

名古屋大学工学部

天気図理解システムは、天気図データベースシステムに天気図の内容の理解機能を付加したもので、天気図の等圧線、気圧文字、前線などの気象情報から気象に関するさまざまな推論を行なう。

本報告では天気図理解システムにおける推論方法と推論規則を実際の天気図について実験的に評価し、改良した結果について述べる。また、特に天気の型の推論について、1973年の1年間の天気図についての推論内容、および正答率を調べ、これにもとづいて、推論規則を改良した。推論規則についてCONFUSION MATRIXで誤りの多い推論規則を誤りを少なくするように行なった。この結果、誤推論の多かった例は減少し、正答率も49.3%から52.6%と改良され、またシステムの推論の上位3位までには77.8%の正解が含まれているという結果が得られた。

EVALUATION AND IMPROVEMENT OF INFERENCE RULES
FOR WEATHERMAP UNDERSTANDING SYSTEM

Yukio TAJIMA, Hitoshi ISHIZUKA,
Yuuji YOSHIDA and Teruo FUKUMURA

Department of Information Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464 Japan

Weather map understanding system augments the weather map database system with the facility for understanding retrieved weather maps. The system can infer some kinds of meteorological facts from meteorological information contained in a weather map.

We evaluated experimentally its inference method and inference rules by applying them to real weather maps. Inference rules concerning with the types of weather maps were evaluated in detail and improved based on the results. By the experimental results, the correct inference rate is 52.6% for 13 kinds of weather map types. In cases of 77.8% the correct answers are included among types of the third higher certainties.

1. まえがき

我々は、図形の二次元的特徴に基づいて画像を蓄積し検索するシステムとして、天気図検索システムを構成した。(1)・(2)このシステムでは、気圧文字の位置、前線の位置、日付、気圧配置、及び等圧線図の形状・位相的構造(樹状記述)による天気図の検索が可能である。しかしながら、このシステムでは、検索された天気図の内容に基づいての理解、或いは、逆に天気図の特徴を与えられて、その意味をシステムが理解して天気図を検索することは人手に任されていた。そこで、今回、我々は天気図の内容の理解機能を付与することで天気図データベースシステムに知的インターフェイスを実現し、さらにその中で用いられている種々の推論規則について評価したので報告する。

2. 天気図の理解

2.1 天気図の理解

天気図には、等圧線図、気圧文字(高、低等)とその進行方向・速度、前線、天気、風力・風向など種々の気象情報が含まれている。これらの情報を組み合わせることで、気象に関する推論、判断が行える。ここではその様な推論、判断を行うことを、天気図の理解と呼ぶ。天気図の理解の例として次のものが挙げられる。

- ① 気圧配置、前線の位置、季節、等圧線の形状などから天気の型を認識する。
- ② 前線の位置、気圧配置などから天気概況を推論する。例えば、前線の位置と、低気圧の位置から雨域が推論できる。
- ③ 気圧文字の位置関係から、高気圧、低気圧などの移動方向や進路などが推論でき、このことから数時間後の

天気図の予想が可能となる。

2.2 天気図理解システム

以上の考察を基に、天気図理解機能を天気図のデータベースシステムに付加することで天気図理解システムを実現した。

具体的には、以下の機能を有するシステムを実現した。

① 入力はローマ字日本語文とし、出力はローマ字日本語文、或はカラーグラフィックディスプレイに図形として出力する。

② 天気図に関する簡単な検索機能。

i) 天気の型の認識: 気圧配置・季節などから冬型・梅雨・移動性高気圧型などの天気の型を認識する。

ii) 雨域の推論: 前線や低気圧の近くは天気が悪いことを使って、雨域(悪天域)を推論する。

iii) 数時間後の天気図を推論する。

3. システム構成

3.1 システムの概要

システム構成を図1に示す。ここでは各部分について簡単に触れ、そのうちで本論文で特に重要な部分について次節以降で詳しく述べる。

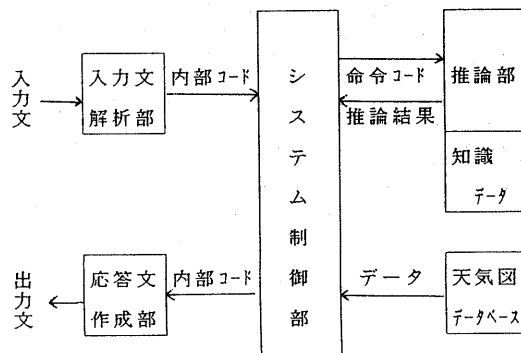


図1 システム構成図

入力文解析部：ローマ字表記の入力文を構文解析し、意味に応じた内部コードをシステム制御部に渡す。

システム制御部：システム全体の動作を制御する。入力文を受け取り、それを入力文解析部に渡してその意味を求める。つぎに、その意味に応じた関数を起動させ、その関数の中で、図形や応答文作成部を呼び出し、表示する。また必要に応じて、その関数は、推論部を呼び出し推論を行う。

推論部：システム制御部からの指示によりデータから推論・認識し、その結果をコード化してシステム制御部に渡す。

知識ベース：文献(4)から天気図に関する種々の規則、知識を収集し、更に、本システムの推論部に適するように変更した天気図に関するルールの集合である。ルールは推論の対象ごとにまとめられルールカテゴリーを形成する。推論は、ルールカテゴリー単位で行われる。

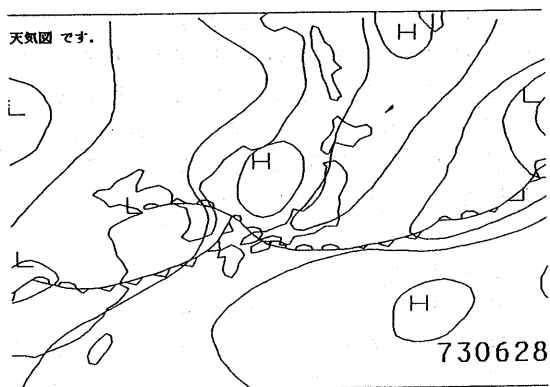


図2 天気図の例

天気図データベース：図2に示すような天気図が1973年一年分365枚格納されている。等圧線・前線などは座標列で表現され、更に前線に関しては、その種類、即ち寒冷、温暖、停滞、閉塞の区別がなされている。気圧文字はその位置が座標で、その種類、即ち高気

圧、低気圧、台風の区別が格納されている。図3には示されていないが各地の天気・風力・風向のデータも付加されている。

3.2 システム制御部

システム制御部は、入力文解析部で解析された入力文の意味を受け取ると、その意味の先頭に書かれた関数を起動する。関数は動作の種類によって分類される。例えば、天気の種類や雨域の推論など推論に関する質問は、推論を行うある1つの関数が起動される。起動された関数の中では、動作データをデータベースから取り出し、それに応じて種々の動作を行う。動作データには、使われるファイル、起動される推論カテゴリ名、新たに呼ばれる関数などが、フレームの形式で書かれている。これにより、システム機能を登録・変更する場合にはデータベース中の対応する動作データの更新のみでよいことになる。

動作データに応じた動作を終えると、システム制御部はメッセージの表示を行って、新たなコマンドを待つ。

3.3 推論部

推論部で用いられている推論方法はプロダクションシステム(P S)の考え方に従っている。このシステムでは、ルールの選択に確信値を導入し、その確信値の比較により結果を選び出すことにした。以下に、ここで用いた推論機構についてやや詳しく述べる。

(1) 推論規則

推論部は先に述べたルールカテゴリー単位で推論を行う。ルールカテゴリー内の各々のルールは図3に示す構造を持つ。図4にルールの一例を示す。このルールはカテゴリー名がL P R E S S - M O V Eで、低気圧の移動に関

するルールであり、ルール名がSP-2である。IF部は、推論対象の現在の天気図の日付が3月から5月までの間であり、ルール変数LPRESSに束縛された低気圧の位置が、(200 400)と(512 512)を対角に持つ長方形の領域に入っていることを示す。THEN部は、このルールの確信値がルールカテゴリLPRESS-MOVEの中で最大値をとるときに解釈される。そのとき、この低気圧の速度は40km/hで、進行方向は30度であることを示す。

```
(カテゴリー名   ルール名
  ( IF      ( P 1   W 1 )
            ( P 2   W 2 )
            |
            ( P n   W n ) )
  ( THEN    ( R 1   V 1 )
            ( R 2   V 2 )
            |
            ( R n   V n ) ) )
```

図2 ルール構造

```
(LPRESS-MOVE
 SP-2
 (IF ((R:SEASON (3 5)) 1)
      ((F:INRESION (L LPRESS)
              ((200 400)(512 512))) 1)
 (THEN ((LPRESS-MOVE (40 30)) 1)))
```

図3 ルールの例

(2) 推論アルゴリズム

以下に推論アルゴリズムを示す。

[推論アルゴリズム]

① 推論部の初期化

システム制御部が推論部の初期化を行う。具体的には、データメモリの中に、パターンマッチングで使われる推論の内容に適したパターンとそのパターン値を格納すること、およびルールに使われる変数を作成することであり、

この変数はパターンマッチングで参照される。

② ルールの取り出しと確信値の計算

推論部はシステム制御部から指示されたルールカテゴリを持つルールを、知識ベースから全て取り出し、各々のルールのIF部に対して確信値CFを以下に示す式で求める。

$$CF = \frac{\sum P_i * W_i}{\sum W_i}$$

ここでP_iはP_iのパターン値で、後のパターンマッチングで述べる方法で求められる。値は、原則として0から1までの実数である。W_iは重みで、P_iがそのルールの中でルールが成立するための重要さを示している。

③ ルール選択

ルールカテゴリ内の全てのルールに対して確信値が全て求められると、推論部はその中から最大の確信値を持つルールを選び、そのルールのTHEN以下に続くリストにもとずき、パターンR_n、パターン値V_nの組をデータメモリーに格納する。但し、最大の確信値が0.5以下の場合には、ルール選択を行わず、推論不可能であったとして、推論結果はNILとなる。

④ 結果選択

データメモリー内を走査して、現在推論しているルールカテゴリを持ったパターンを捜す。 □

(3) データメモリー

いわゆるワーキングメモリに相当し、パターンマッチングで使われるパターンとパターン値の対が格納されている。推論部はこのデータメモリーを走査してパターン値を求めたり、或は、パターンの削除・追加を行う。推論の対象と

なる天気図データも、このデータメモリの一部として位置づけられている。

(4) パターンマッチング

パターンの構造は基本的には以下のようになっている。

(パターン名 (パラメータ))

実際のパターンはこれらの論理式として記述される。パターンに AND, OR, NOT が含まれるときにはそのパターン値は、AND を最小値演算、OR を最大値演算、NOT を 1 からの減算で置き換えることで、論理式の値を求める。

パターンマッチングの方法は以下に示す 3 種類あり、システムは、どの方法に該当するか調べて、パターン値を求める。

i) パターンの直接的マッチング

データメモリーの中から、パターンが一致するものを捜し出し、もしあれば、そのパターン値を取り出し、それをこのパターンの値とする。ここで言うパターンには、AND, OR, NOT を用いて記述されたものも含む。

ii) 関数引用

パターン名が関数引用を表わす "R:" 或は "F:" で始まっていれば、推論部は対応する関数を呼び出し、関数値からパターン値を求める。もし、関数値が数字ならばそのまま、NIL 以外の S 式ならば 1、NIL ならば 0 とする。値が決定されたら、そのパターンとパターン値をデータメモリーの中に加える。"R:", "F:" などの関数については、次項(4)で述べる。

iii) ルールカテゴリー名の引用

そのルールカテゴリー名で推論部を起動し、その後でワーキングメモリーを走査してパターン値を求める。これによりルールの階層化が可能となる。このとき推論部で使われたルールカテゴリー名は記憶しておき、常にそのル

ールカテゴリーが以前に使われたかどうかチェックし、無限ループに陥らないようにする。

パターンが上述のどれにも該当しないときはそのパターン値を 0 とする。

(4) 関数呼び出し

簡単な形で表わせないパターンを LISP 関数として定義し、これを引用して求められるよう作成したものである。これにより、推論部の初期化の段階で数多くのパターンを作る必要がなくなり、自由にルールが記述できるようになった。

パターン名が "R:" で始まるときは、そのパターン名が推論専用の関数と解釈する。パターン名に続くパラメータを評価せずに、そのパターン名の関数に渡して関数値を求め、それをパターン値とする。

パターン名が "F:" で始まるときは、パターン名から "F:" を取り除いた関数を、"R:" の時と同様にパラメータを評価しないで呼び出す。このことはルールの中から任意の関数を呼び出すことが出来ることを示す。

4. システムの実現と推論の評価

システムは、本学の大型計算機センター FACOM M-382 システムの TSS 上に実現されている。システムは、UTILISP で記述され、プログラムサイズは約 2000 行である。また、グラフィックス用パッケージ (TGSP) と UTILISP のリンクによりカラーディスプレイ (F9432A) に種々の形式で天気図や推論結果を表示する。また同時に、カラープロッター、カラーハードコピー装置に画面のハードコピーを取ることも出来る。

4.1 実行例

図5にシステムの応答例を示す。プロンプトの数字は、システムが推論の対象としている天気図の日付である。

推論の対象となった元の天気図は図2の天気図である。

図6に雨域の推論結果を示す。マークが大きいほど雨域である確率が高い。ここで使われている推論方法は、各々の前線に対して前線に近いほど雨の降る確率が高いとしてマーキングをし、更に、低気圧の周りに雨域を設定する方法である。

```
>(START)
HELLO
101>> 6GATU 28NITINO TENKIZUWO HYOUJISEYO.
        6 GATU 28 NICH I NO TENKIZU DESU.
628>> TENKINO KATAWA NANDESUKA.
        TENKI NO KATA HA TUYU TO OMOWAREMASU.
628>> TENKINO WARUITOKOROHA DOKODESUKA.
        KOKOGA WARUITO OMOWAREMASU.
628>> TOKYONO TENKIHA NANDESUKA.
        TOKYO NO TENKI HA KIRI DESU.
628>> OOSAKANO TENKIHA NANDESUKA.
        OSAKA NO TENKI HA KUMORI DESU.
628>> 4JIKANGONO TENKIZUWO HYOUJISEYO.
        4JIKAN GO NO TENKIZU DESU.
628>> BYE.
GOOD-BYE
>
```

図5 応答例

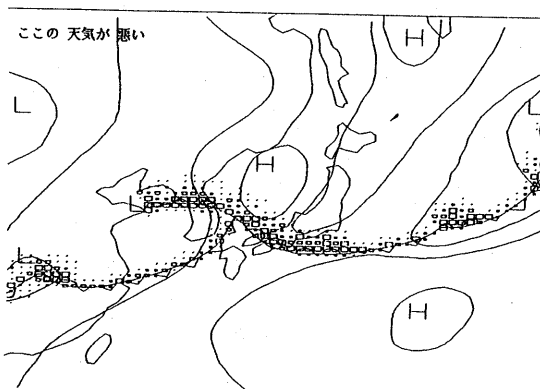


図6 雨域の推論

さらに、図7は数時間後の天気図を推論した結果である。これは、1時間毎の気圧文字の移動量をプロダクションシステムを用いて求め、指示された時間だけこれを繰り返して、数時間後の気圧文字の位置を得る。そして、前線の位置は、その前線が属する低気圧

の移動量に従って移動させる。なお、図の意味は、点線が気圧文字・前線の元の位置、実線が推論された位置である。

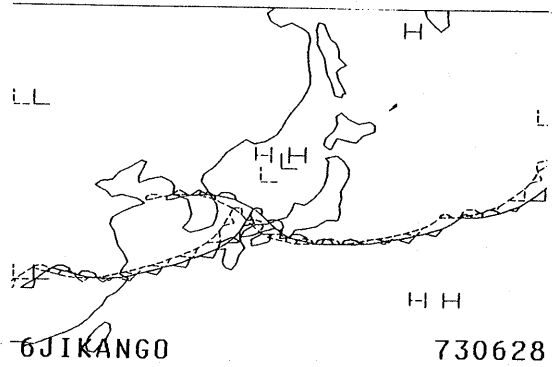


図7 数時間後の天気図の予想例

4.2 推論の評価

ここでは、天気の型の推論について推論機能が正常な推論を行っているかを評価した。正解、不正解の判断は、天気図を見て人が判定した型と比較して行なった。尚、天気の型をはっきりと決定できないものについてはいくつかの候補をあげ、推論によって得られた結果のうちの確信値の3位までにその候補と同じものがあればそれを正解とした。

4.2.1 推論結果

1年間の天気図に対する天気の型の推論結果をCONFUSION MATRIXの形にし、その問題点を次項に示す。

また確信値の大きい順にルールを上から三つ取り出し、その中に正解があるか否かを調べ、その正答率を調べた。結果は次のようであった。

- | | |
|---------------|-------|
| 1位が正解であるもの: | 49.3% |
| 2位までに正解を含むもの: | 62.5% |
| 3位までに正解を含むもの: | 75.3% |

4.2.2 結果の問題点

CONFUSIN MATRIXにおいて、特に誤りの多い場合を表1に示す。

正解	推論	度数
夏型	日本海 低気圧型	10
北高型	冬型	10
夏型	冬型	9
南岸		
低気圧型	冬型	7

表1 誤推論の例

また、この他に、次の傾向がみられる。

i) 冬型の選択日数が実際の出現回数より多い

ii) 夏型の選択日数が実際の出現回数より少ない

これらに見られるように、誤推論の多くが夏型あるいは、冬型に関連していることが知られる。

4.3 推論の改善

推論の結果の評価をもとにルールを変更しその結果を評価した。

4.3.1 ルールの変更の例

誤推論の目だつ場合について、ルールの内容を検討し、それを修正した。修正されたルールの数は12個である。変更はパターンの定義、およびそのウエイトについて行なった。

ルールの変更について、冬型を例にとって述べる。

冬型から、北高型、夏型、または南岸低気圧型への誤推論を減らすために、以下のようにパターンまたはウエイトの変更を行なった。

i) 変更前のパターンとそのウエイト

P1: オホーツクに低気圧がある

$$W1 = 1.5$$

P2: シベリア又は満州に低気圧がある

$$W2 = 1.5$$

P3: (200.150) (300.512) の領域で等圧線が垂直になっている

$$W3 = 1.0$$

P4: P3と同じ領域で、等圧線の密度

が高い $W4 = 1.0$

P5: (200.150) (350.512) の領域に気圧文字無し $W5 = 1.0$

P6: 季節 11月～3月

$$W6 = 2.0$$

ii) 変更点とその理由

P1': オホーツク又は樺太の低気圧 (低気圧が、発達して樺太に達した場合を考慮)

P2': (200.150) (320.250) の領域で等圧線が垂直。

P3': (200.150) (320.250) の領域で等圧線の密度が高い。

P5': 南岸、東シナ、日本海に低気圧なし、及び南岸に前線なし。

$$W5' = 2.0$$

(南岸低気圧型などと区別するため)

4.3.2 改善の結果

変更されたルールによる推論結果のCONFUSION MATRIXを付図に示す。

正解日数が増えたものとしては、夏型14日から27日、北高型12日から20日、などがあるが逆に減ったものもある。また、変更前の夏型を日本海低気圧型に、北高型を冬型になどの誤推論は改善されたが、冬型を南岸低気圧型と誤った例10日、北高型を冬型と誤った例10日、など新たな問題も起きている。

② 正答率は以下のものであった。

1位 : 52.6%

2位まで : 69.6%

3位まで : 77.8%

それぞれにおいて正答率は良くなっており、天気型の型に関するルールは改善されたといえる。3位までの正答率は77.8%とかなり高い。これらの結果から推論を多段階にし、1回目の推論で数個の候補を選び出し、2回目に別のルールで推論を行なうような2段階の推論を行なえば全体の推論結果の正答率が上がるであろう。

5. あとがき

システムの今後の改良課題として、次の項目があげられる。

i) 推論能力の強化

天気型の型に関する推論は、現在、主に気圧文字の位置や季節に基づいて推論を行っているが、これでは高気圧や低気圧の勢力範囲を考慮していないので正確な推論が出来ない。そこで、それらを使えるようにルールを変更あるいは作成すること、及びそれに応じた推論部の変更が必要になる。また、確信値の3位までの正答率は80%近いので、2つのルールカテゴリーを使って2段階推論を行って、推論の正答率を挙げるように推論部の機能の変更も必要である。

ii) ルール記述の簡略化

例えば、天気図の位置領域を座標ではなく、シンボルで表現するようにすることで、ルールの記述を容易にする。

iii) 入力文解析部の改良

もしシステムが理解できない文章が入力された時には、ユーザーに問い合わせて、理解できるようにする。

謝 辞

本研究を進めるに当たり、日々御尽力を賜った本学の稲垣康善教授、鳥脇純一郎教授をはじめ研究室の皆様へ感謝します。

本研究の一部は文部省科学研究費補助金(特定研究『多元知識情報』A班)による。

参考文献

- (1) 山守、吉田、福村: 天気図の符号化とその天気図データベースシステム実現への応用、情報処理学会論文誌、26巻6号、1049-1056(1985-11)
- (2) 中林、吉田、福村: 天気図検索システムと等圧線検索システムの統合、電子通信学会パターン認識と学習研究会資料、PRL85-81(1986-3)
- (3) 吉川、吉田、福村: LISPプログラミングのための知的エディタの実現、電子通信学会オートマトンと言語研究会資料、AL85-25(1985-7)
- (4) 気象庁予報課予報技術研究会: 天気図の書き方と見方、恒星社厚生閣

推論 実際	冬型	夏型	東高 西低	北高 南低	梅雨	流れ 出し	移動 性H	帯状 H	南岸 L	日本 海L	2つ 玉L	台風	台湾 坊主	計
冬型	52	2		1		8		2	10	4				80
夏型	4	27	1	5		5		3		7				52
東高西低		3	3		1	1	1			2				11
北高南低	4	5		20	3	5	5		5	1			2	50
梅雨		2		7	11						2			22
流れ出し			1	1		10				1				13
移動性H	2	1		4	1	1	24		2					35
帯状H	2	2		2		1	2	3						12
南岸L						4	3		14		1		1	23
日本海L		4	1	6		2	4			13				30
2つ玉L	2	4		3		2	2			5	6			24
台風		3	1						1			7		12
台湾坊主													1	1
計	66	53	7	50	16	39	41	8	32	33	9	7	4	365

付図 CONFUSION MATRIX