

2F-9

太陽フレア予報エキスパートシステムの開発

一 予報知識獲得機能 一

白土 保

郵政省 通信総合研究所

1. はじめに

近い将来、多くの人間が宇宙空間で活動するようになる。ところが宇宙空間においては、太陽フレア(太陽面上で起こる爆発)発生の際地球まで飛んでくるX線、高エネルギー粒子、プラズマのかたまりなどの影響にさらされる。これらの現象は人間に直接有害であったり、機器の異常帯電など機器の故障の原因ともなるため、事前に発生を予測し危険を未然に防ぐことが重要となる。

通信総合研究所平磯宇宙環境センターでは現在、このような現象の発生や影響の程度を予報する「宇宙天気予報システム」の研究開発を進めてきている。このシステムの主要な機能は太陽フレアの予報である。従来においては太陽フレアの予報はほとんど人間の判断に依存してきた。そのため予報のバラツキ、ミス、予報知識の継承、機械化の困難さなどの問題が生じている。本研究のフレア予報システムは、できるだけ専門的なヒューリスティックを用いない、という設計思想にもとづいておりこのような問題は少ない。今回は、過去のデータからフレア予報知識を抽出する機能の実現を中心に太陽フレア予報システムの概要および開発状況について報告する。(以下、太陽フレアを単にフレアと呼ぶ)

2. フレア予報システム

フレアは黒点と呼ばれる太陽面上の領域の近くで生じる。そのためフレア予報は通常黒点ごとにおこなわれる。以下に本研究におけるフレア予報の定義を示す。「ある日の黒点の属性ベクトルをもとに

それが翌日フレアをおこす確率を計算すること」。ここで属性ベクトルとは黒点の状態をあらわす属性値(記号値と数値の2種類ある)の組のことである。

予報システムは前処理部、知識獲得部、予報実行部、核知識の4つの部分から構成される。(図1)

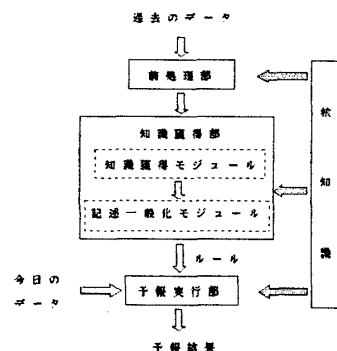


図1 システム全体の構成

前処理部

過去のデータは2種類ありどちらも1日単位のレコードの集まりである。ひとつは各黒点の属性値、もうひとつはフレアに関するデータである。前処理部ではこれらをもとに知識クラスタをつくる(知識クラスタについては後述する)。

知識獲得部

知識獲得部は前処理部から入力された知識クラスタを処理しプロダクションルールの形式でフレア予報知識を抽出する。

予報実行部

ある黒点の今日の属性値の組が入力されるとルールを参照して明日のフレア生起確率を出力する。

核知識

システムの処理プロセスを左右する知識で、次のようなものがある。

- ・システムの機能にかかわるパラメタ的知識。実際のデータ処理の過程で試行錯誤的に最適値が決定される。
- ・記号値をとる属性値間の相違度の定義。本システムではこの部分にのみ専門的ヒューリスティックを用いている。

本システムの最も本質的な部分は知識獲得部の知識獲得モジュールである。以下この部分について述べる。

3. 知識獲得モジュール

知識獲得手法

属性ベクトルと、そのベクトルが翌日起こしたフレアの統計すなわちフレア生起確率、を対にした量を考えこの量を知識クラスタと呼ぶ。図2に知識クラスタの概念を示す。

$$\begin{aligned} \text{知識クラスタ} & \Rightarrow \text{属性ベクトル} \quad \text{フレア生起確率} \\ \text{属性ベクトル} & \Rightarrow \text{属性値}_1 \text{ 属性値}_2 \cdots \text{属性値}_n \\ \text{フレア生起確率} & \Rightarrow \frac{\sum \text{翌日のフレア (有り}=1, \text{なし}=0)}{\text{その属性ベクトルの出現頻度}} \end{aligned}$$

図2 知識クラスタ

1個の知識クラスタは予報知識の断片である。すなわち属性ベクトルがルール条件部、フレア生起確率が結論部にあたる。個々の知識クラスタは以下の2つの理由で、より大きなクラスタにまとめられる(クラスタリング)必要がある。

- ①各知識クラスタのフレア生起確率は、そのほとんどがごく小数のサンプルの代表であり統計的信頼性が小さい。なぜなら属性ベクトルの総数は、ベクトルを構成する各属性の属性値の個数に関し、それらの積の数だけ考えられる。一方、用いる属性をあらかじめ絞ることは強力な専門的ヒューリスティックの使用を意味し、好ましくないからである。
- ②フレア生起確率が互いに近似している知識クラスタに含まれる属性ベクトルはまとめて簡潔な表現に書換えられる可能性がある(書換え処理は記述一般化モジュールでおこなわれる)。

クラスタリングのプロセスは2つの知識クラスタの選択およびそれらのマージによる新しい知識クラスタの生成、の繰り返しである。マージは真のフレア生起確率(その属性ベクトルが本来持っているフレア発生のパテンシャル)の差が小さい知識クラスタの間でのみおこなわれることが必要である。このため、知識クラスタのペアの選択指標として次の2つが用いられる。

①属性ベクトル間の相違度

「みかけが類似した物理対象はふるまいも類似している」というやや普遍的なヒューリスティックにもとづく。サイズが小さい(サンプル数が少ない)知識クラスタに対して用いられる。

②フレア生起確率の差

サイズが大きい知識クラスタはそのフレア生起確率の値をそのまま真の値の近似値としてよい。

サイズが大きくなるにつれ①の指標→②の指標、の乗り換えが少しずつおこなわれる。乗り換えの速さを決めるパラメタは核知識のひとつであり、入力されるデータの数や統計的特性に依存して試行錯誤的に決定される。

クラスタリングは、サイズの小さいたくさん知識クラスタから始まり、それらが成長し最終的に大きいサイズの知識クラスタがいくつか生成されて終了するというプロセスをとる。得られたクラスタに含まれる属性ベクトルの論理和がルールの条件部、フレア生起確率が結論部に相当する。

4. おわりに

太陽フレア予報エキスパートシステムの主要な機能である予報知識獲得機能の実現手法について報告した。記述一般化モジュールを除くシステムの構築はほぼ終了しており、現在この構築および完成部分のテストおこなっている。それらの詳細については別の機会に報告する。

謝辞 本研究の機会を与えて頂いた小川センター長、猪木室長に感謝致します。