

## Q27a 大量観測データの自動処理：データマイニングによる未同定線の同定

青木 賢太郎 (JST)、大石 雅寿 (国立天文台)、古畑 玲奈、増永 良文 (お茶の水大)

稼動をはじめたすばるや計画が進行している ALMA では、巨大な観測データファイルが次々に生み出される。これらの観測データを従来の方法で処理していくことは時間的にも労力的にも無理があり、高速な計算機の能力を活用した自動処理手法を開発することが必須である。

我々は JST (科学技術振興事業団) の研究資金を導入することにより、その手法の開発を 2 年前から行っており、その初期成果が出たので報告する。本初期成果は、データマイニングという情報技術を活用したものであり、今後の応用範囲は広いと考えられる。

初期成果の対象として用いた観測データは、野辺山 45m 鏡を使って行われたおうし座分子雲 TMC1 での分子線サーベイデータである。この分子線サーベイでは約 700 本の分子線と思われる信号が受かっており、そのうち約 400 本が既知または新発見の分子と同定されている (Ohishi et al. 1999)。残りの約 300 本については同定が待たれている。これら多数の未同定線から、すでに同定された分子線と類似した性質を持った信号をとりだし、より確からしい信号を選ぶことを目標とした。その際にはできるだけ自動化し、作業をおこなう個人には依存しないようにすると同時に、経験を積んだ研究者の知識を利用してより信頼度の高い選別をおこなうことを条件とした。以上の目的のために我々はデータマイニングの手法を用いて未同定線の信号及びノイズへの分類をおこなった。その結果、

1. 研究者の経験と非常によく一致した選別が自動でおこなえた。
  2. 未同定線 300 本中約 1/4 がすでに同定された分子線と類似している。
- ことが得られた。年会では用いたデータマイニングの手法、選別の基準、この選別方法の特徴を中心に発表する。