

シリーズ特集 研究のツールボックス 第6回

<http://www.ai-gakkai.or.jp/jsai/journal/toolbox/>

神畷 敏弘

(産業技術総合研究所)

今日では、研究関連の成果が、論文に加えて、オープンソースのソフトウェアとしても公開されるようになってきている。これらの公開ソフトウェアが利用できることは、研究・開発にとって大きな助けとなるであろう。しかし、利用者が限られるため紹介される機会が少なかったり、利用法がよくわからないといった理由で、すばらしいソフトウェアが有効に利用されない現状がある。

そこで、これらの人工知能研究に有用なソフトウェアを紹介する不定期連載のシリーズ特集を企画し、今まで、以下のようなソフトウェアを紹介した。

- 第1回 19巻3号 (2004年5月) 掲載: Jena, 茶筌と南瓜, BioRuby, Weka
- 第2回 20巻1号 (2005年1月) 掲載: Julius, RoboCup Soccer Simulator, MUSASHI, R
- 第3回 20巻6号 (2005年11月) 掲載: OpenCV, OpenDX, VGJとWalrus, Octave
- 第4回 21巻4号 (2006年7月) 掲載: Protégé, Prolog Cafe, DSpace, GNU Scientific Library
- 第5回 22巻1号 (2007年1月) 掲載: FreeMind, GXPとIbis, Bioconductor, Maxima

第6回となる今回も、すぐに試せるチュートリアル形式で、以下の四つのソフトウェアを紹介する。

最初に、バイオ関連データ処理のフリーウェアをまとめた **EMBOSS** を、多賀谷が解説する。ヒトゲノム配列の解読が終了し、これらのゲノム中の遺伝子の働きを調べることに研究の焦点が移った。これには、機械学習分野などで研究されてきたデータ解析技術が利用されている。こうした解析技術を適用するには、各種データベースからのデータの取出しや、基本的な前処理が必要となる。これらの処理を行うフリーウェア群をパッケージしたもののが **EMBOSS** である。基本的なデータ処理などのほか、第1回で紹介した **BioRuby** との連携についても紹介する。

二つ目は、伊庭が紹介する複雑系シミュレータ **Swarm** である。これは、アリや車のような自律的に動くエージェントの集団が、互いに影響を与えながら、全体としてどのように振る舞うかを調べるためのものであ

る。このシミュレーションにより、形式的な関数の最適化や、より具体的な交通渋滞といった複雑系の問題を扱うことができる。**Swarm** は、こうしたシミュレーションを行うための **Java** のライブラリ群である。これを用いて、系内で行動するエージェントの実装や、それらの可視化ができる。

三つ目の、土肥らが紹介する **JADE** は、ソフトウェアエージェントを実装するためのフレームワークである。人工知能技術を利用したシステムでは、インタフェースと、背後のインテリジェントな処理を個別のエージェントに処理させる。また、一つの大きな処理を、部分処理を担当する複数のエージェントに実行させたりといったことが行われる。こうしたエージェントの相互利用のため、**FIPA** によるプロトコルの標準化が行われている。**JADE** はこの標準プロトコルを実装するためのフレームワークである。

最後に、自然言語テキストなどの非構造データを扱うモジュール群の相互運用のための枠組みである **UIMA** を竹内らが解説する。現在流通しているデータのうち、**RDB** などに構造化されて蓄積されているデータは2割に満たず、そのほとんどはテキストなどの非構造データであるといわれている。こうした自然言語で表されたテキストを処理するには、文抽出、形態素解析、構文解析、談話解析といったステップが必要になる。だが、これらを連結する統一的な枠組みはなく、同じ働きをする違うアルゴリズムと交換するといったことが困難であった。そこで、このような枠組みを提供するのが **UIMA** である。ここでは、各モジュールの仕様記述や、データの受渡しの基本について述べる。

なお、本特集では記事に加えて、サポートページを人工知能学会の **Web** サイトに用意している。

<http://www.ai-gakkai.or.jp/jsai/journal/toolbox/>

本ページには、記事中で紹介された **Web** ページのリンク集や、サンプルなどを掲載したので、本特集とともに利用されたい。