

# Progress in Informatics ABSTRACTS

No.4, March 2007

## Contents

### Special Contribution

- 1 ——— **Unpredictability and chance in scientific progress**  
科学発展の予測不可能性とチャンス  
John Meurig THOMAS

### Research Papers

- 5 ——— **Named Entity Recognition in Vietnamese documents**  
ベトナム語文書の固有名認識  
Tri Tran Q., Thao Pham T. X., Hung Ngo Q., Dien DINH,  
Nigel COLLIER  
文書中の語を所定のターゲットエンティティ・クラスに分類することを目的とする固有表現抽出 (NER) は現在、情報検索や機械翻訳、情報抽出、質問応答など、多くの自然言語処理タスクの基盤とみなされている。本論文は、NER モデルに基づくサポートベクターマシン (SVM) をベトナム語に応用した実験の結果をまとめたものである。この最先端の機械学習法は、研究の進んだ一部の言語では広く NER に応用されているが、ベトナム語への応用はこれが初めてである。条件付確率場 (CRF) との比較では、フィーチャーウィンドウのサイズの最適化で SVM モデルが CRF を上回り、全体的な F スコアは 87.75 である。本論文はベトナム語の特性についても詳細に論じ、このタスクにおけるパフォーマンスに影響する要因の分析も行う。

- 15 ——— **Rewriting XQuery by child-path folding**  
子軸経路式の畳み込みによる XQuery の書き換え  
加藤弘之, 日高宗一郎, 吉川正俊

式の畳み込みを用いた部分評価に基づく書き換えによる XQuery の最適化手法を提案する。本手法は大きく二つの部分から構成されている。一つはメインアルゴリズム部で、XQuery 式の帰納的定義に基づく再帰アルゴリズムである。もう一つはメインアルゴリズムから呼ばれ式固有の畳み込み関数を呼ぶ部分である。本稿では特に子軸経路式固有の畳み込みである「child-path folding」アルゴリズムに焦点を当てる。別の XQuery によって構築されたエレメントに対する子軸に関する問合せは合成問合せとなるが、提案手法によりエレメント構築子を含む冗長な式を削除することが可能となる。この種の合成問合せは、データ統合問題における Global-As-View ア

プローチにおいて典型的な問合せであるので、提案手法は有用である。また、本稿では子軸経路式中に存在する全ての冗長なエレメント構築子を含む式を削除するための補助変換も提案する。子軸経路式以外の FLWOR 式や限量式を含む主な XQuery 式の畳み込みについても述べている。更に、XQuery 式の構文木に二つのアノテーションを追加することで、枝狩りに基づく書き換えアルゴリズムの改良についても述べている。

- 29 ——— **ILL/DD in Japan across the turn of the century—  
Basic findings about NACSIS-ILL from 1994 to 2005**  
日本の図書館間相互貸借/ドキュメント・デリバリー: 1994  
年から世紀をまいで 2005 年までの NACSIS-ILL 基本調  
査結果

土屋 俊, 竹内比呂也, 佐藤義則, 逸村 裕

本論文は、NACSIS-ILL と名付けられた図書館間相互貸借 (ILL) リクエストメッセージ送付システムによって 1990 年代半ばから実現されてきた大学間の図書館協力サービスに関する基本的事実を記述するものである。この研究は、NACSIS-ILL によって記録された 1994 年から 2005 年までのデータに基づいている。本研究の主要な調査結果として、以下の諸点を挙げる事ができる。すなわち、1) 日本の大学における ILL においては「外国雑誌」に掲載された論文のコピーに対する要求が 1990 年代にはきわめて支配的であるという点が特徴的であること、2) 皮肉にも、特に看護学分野の雑誌に顕著に見られるように「国内雑誌」に掲載された論文に対する要求の増加が、2002 年に始まったコンソーシアム体制下のサイトライセンスによってオンラインで利用可能になった「外国雑誌」掲載論文に対する要求の減少を埋め合わせるかのように顕著になったこと、3) 図書に対する現物貸借の要求は、それが要求全体に占める割合は小さいとはいえ、NACSIS-CAT という総合目録データベースの成長に従って増加してきたこと、4) 充足率は現物課貸出、複写提供のいずれにおいても安定的に高く、ターンアラウンドタイムの平均も概ね 1 週間以内であり、システムはきわめて効率的であるということ、5) システム本来の目的は相互に受益者となるような協力システムの構築にあったが、実際には主に要求するだけの図書館と供給するだけの図書館が存在しており、これは部分的には 1970 年代に指定された「分野別外国雑誌センター」館に起因するものであること、6) いくつかの中小規模の図書館が近年供給を始めておりこれが顕著な動きを示していることである。

## R&amp;D Project Reports

## 51 Overview of SINET3—

## Next-generation Science Information Network

次世代学術情報ネットワーク (SINET3) の概要

漆谷重雄, 阿部俊二, 松方 純, 計 宇生, 福田健介,  
鯉淵道紘, 山田茂樹

本論文では, 2007 年 4 月より運用を開始する次世代学術情報ネットワーク (SINET3) について概説する。SINET3 では, 多様なサービスを柔軟に提供できる革新的なネットワーク基盤の確立により, 学術コミュニティとのさらなる連携強化を目指している。プロトコルフリー化のためのマルチレイヤ (IP, Ethernet, 波長等) 機能, 関連研究機関の連携強化のための豊富な VPN (Virtual Private Network) 機能, 高性能アプリケーションのためのオンデマンド品質保証機能等を基本として, サービス機能の充実を図っている。また, 最先端のレイヤ1スイッチとIP/MPLS ルータの最新実装を高度に連携させたネットワークングに特徴がある。本論文では, 現状と背景を説明した後, ネットワークサービス, ネットワークアーキテクチャ, ネットワーク技術の順で解説を行う。

## 63 Quality Evaluation of Embedded Software in Robot Software Design Contest

ロボットソフトウェアデザインコンテストにおける組込みソフトウェアの品質評価

鷲崎弘宜, 小林靖英, 渡辺博之, 中嶋栄次, 萩原裕志,  
平鍋健児, 福田和也

2005 年に日本で, ライントレースロボットを自律制御するソフトウェアを分析・設計する「ET ソフトウェアデザインロボットコンテスト」が開催された。コンテストは, 組込みソフトウェア開発の領域において, 日本の若い技術者に向けて実践的な学習機会を提供することを意図していた。本稿ではソフトウェア品質評価の観点より, コンテストの内容と結果, および, 結果の分析を通じて得られた知見を報告する。我々は, 設計モデルの品質と最終システム (ロボット) の品質を統合して議論するための枠組みを構築し, 枠組みを用いてコンテストの実施結果を分析した。分析の結果, モデルの構造に関する特徴 (クラス数や複雑さ) の定量的測定結果は, 審査員によるモデルの定性的評価と強く関係していることが分かった。これは, 審査員による定性的評価において, それらの構造上の特徴が暗黙に考慮されたためと考えられる。一方, モデルの定性的評価は, 最終システムの信頼性や効率性と強く関係していないことが分かった。組込みソフトウェアの開発として本来は, 開発の早い段階において分析・設計モデルからシステムの信頼性や効率性を予測可能なことが重要である。得られた分析結果に基づき, 提出されたモデルやその評価方法, コンテストそのもののあり方について, システムの信頼性や効率性の予測につなげるための可能な対策を考察した。具体的には, モデル上での確に (効率性を含む) 種々の品質特性を測定可能とするために, モデル駆動開発によるモデルとプログラムの対応関係の維持といったロボットソフトウェア開発上の対策や, 効率性評価

の重視といったモデル評価時の対策が必要と考えられる。

## Case Study

## 79 Evaluation of Japanese universities' research activity based on the number of awards of Grants-in-Aid for Scientific Research from 1998 to 2002 and in 2003

科学研究費補助金採択研究課題数による 1998-2002 年度および 2003 年度の大学の研究活性度の調査研究

西澤正己, 根岸正光, 柴山盛生, 孫 媛, 野村浩康,  
前田正史, 光田好孝

科学研究費補助金制度は, わが国で最も古い大学教員に対する研究費の提案・公募型の研究費配分制度であり, 申請領域においてピアレビューにより割り当てられる。本研究では 1998 年から 2002 年度まで 5 年間の科学研究費補助金の採択件数を研究分野別ごとに, 大学別に分類し各大学の研究活性度の指標にすることを目的としている。さらに, 科学研究費の申請分野細目表が 2003 年度に大改訂されており, 改定後の採択状況と研究分野の移行の状況を分析した。本論文は, 国立情報学研究所のテクニカルレポートに掲載されている 1998 年度から 2002 年度の 5 年間の分析の 6 分冊, および 2003 年度の分析の 5 分冊の概要報告である。