

プロジェクト名： 機能と帰納

サブテーマ名： 情報・通信“メタウェア”とその応用

研究代表者

中野純司 [統計数理研究所]

松井知子 [統計数理研究所]

瀧澤由美 [統計数理研究所]

1. 研究目標

“メタウェア”は、現象をモデルとして表現するために必要な概念と構造を意味し、そのモデルをハードウェアやソフトウェアとして具現化できるようにするためのものである。本研究では、次の情報・通信における次の三つの具体的テーマ、

- A) 計算機による帰納的モデリングのための環境
- B) マルチモーダルデータからの不変情報の発見とその方法論の研究
- C) モデル化と高効率データ処理に基づく無線データシステムの研究

に関して、ハードウェア、ソフトウェア、メタウェアを総合的に考察し、情報・通信における有用なメタウェアを明らかにする。

各具体的テーマの目標は次の通りである。

- A) 現代科学の帰納的推論の根幹をなす計算機の機能を、立場の異なる研究者が緊密に連絡を取り合うことによって、有用な方向に発展させることを目指す。具体的に、研究のコアとして統計解析環境 R を用いながら、「物理乱数および疑似乱数の発生とそのシミュレーションでの利用」、「大規模数値計算と統計解析環境における並列化とネットワークおよびグリッドの利用」について研究を行い、科学研究に有用なメタウェアを明らかにする。
- B) マルチモーダルデータを統合的に捕らえ、帰納的アプローチにより不変情報を発見するしくみを見出すことを目指す。具体的に「画像・音声のマルチモーダルデータからの不変事象の学習検索」、「対話における言語・身体動作データからの不変情報の抽出と活用」、「身体性制約下における外界データの不変情報抽出機械としての知覚神経回路の特定」などの課題への取り組むことにより、その方法論の有効性・可能性を検証する。それらの検討を通じて、分野横断的な帰納的アプローチにおけるメタウェアを明らかにする。
- C) 情報化時代の発展を支える高度化公衆情報通信システムの実現を目指す。具体的に「データ伝送の高速化」と「これに伴うチャネルの信頼度低下および消費電力の急激な増大」の課題に取り組み、情報通信におけるメタウェアを明らかにする。

2. 年度研究計画

以下、A)～C)の具体的テーマごとに年度研究計画を示す。

平成16年度（予備研究）

- A) 統計解析環境Rをスーパーコンピュータ上で稼働させるための調査を行う。Rはオープンソースソフトウェアであるため、主としてパーソナルコンピュータレベルでの開発が主であった。これをスーパーコンピュータで安定して利用できるように改良する。
- B) 帰納的アプローチに関する調査を行うとともに、マルチモーダルデータを扱う具体的な課題を設定する。
- C) 移動無線通信におけるチャンネルおよびデータ処理に関する問題点の抽出、解析を行う。

平成17年度

- A) 3種類のスーパーコンピュータ上でRを稼働させる。
- B) 四つの課題、「映像データからの不変事象の学習検索」、「ダイナミカルシステムを用いた身体協調の不変原理の探求」、「身体性制約下における外界データの不変情報抽出機械としての知覚神経回路の特定」、「対話データからの不変情報（コミュニケーション・パターン）を規定する要因の特定」に具体的に取り組む。
- C) 移動無線データ通信における方式的な研究課題を明確化するため、無線システムのモデル化を行う。これに基づきスペクトル拡散変復調部（ベースバンド部）の方式設計を行う。

平成18年度

- A) スーパーコンピュータ上で稼働する並列化されたRを使いやすくするためにWebインタフェースを作成する。さらにRを利用するバイオインフォマティックスのための計算機環境であるBioconductorをスーパーコンピュータ上で稼働させる。そしてそれと遺伝研データベースDDBJのWebを介する結合を行う。また極地科学のための乱数利用のパイロットシステムを極地研で作成し、その上で並列化Rを稼働させ、より小規模の環境への技術移転に備える。また統計解析システムJaspで利用でき、Rを補完することもできるJavaによるデータ可視化のための新たなライブラリも開発する。
- B) 四つの課題、「映像データからの不変事象の学習検索」、「ダイナミカルシステムを用いた身体協調の不変原理の探求」、「身体性制約下における外界データの不変情報抽出機械としての知覚神経回路の特定」、「対話データからの不変情報（コミュニケーション・パターン）を規定する要因の特定」への取り組みをさらに進める。特に最初の課題については、映像検索の国際競争型評価プロジェクトTRECVIDに参加して、本提案法についての客観的な評価を得る。
- C) 移動無線データ通信のチャンネルモデルを研究しこれに基づき方式的な研究課題への対応を行う。ベースバンド部のプロトタイプハードウェアの開発を行う。このプロトタイプモデルにおいてインターネットコンテンツの送受実験を試みる。

平成19年度

- A) 講習会などを通してバイオインフォマティックスの研究者にスーパーコンピュータ上の並列化RとBioconductorの普及をはかる。また、並列化Rのグリッドコンピュータ上での利用

を検討する。また、並列化 R のグリッドコンピュータ上での利用を検討する。さらに Jasp においてはより使いやすい並列計算の仕組み、例えば視覚的な並列計算プログラム入力法などを検討する。

- B) 二つの課題、「ダイナミカルシステムを用いた身体協調の不変原理の探求」と「対話データからの不変情報を規定する要因の特定」を統合し、新たに「対話における言語・身体動作データからの不変情報の抽出と活用」の課題とする。「映像データからの不変事象の学習検索」の課題については昨年度に引き続き、映像検索の国際競争型評価プロジェクト TRECVID に参加して、本提案法についての客観的な評価を得る。各課題について研究を推進するとともに、帰納的アプローチのメタウェアに関する分野横断的な検討を開始する。
- C) 移動無線データ通信の高効率デジタル信号処理として、公衆無線用組み込みシステムに関する研究を行う。具体的解決策として組込み構成論を研究する。高効率データ処理部（組込みシステム）のプロトタイプハードウェアのモデル開発を行う。

平成20年度

- A) 並列化 R のグリッドコンピュータ上での利用を実現する。また、Jasp に関してはデータ視覚化、視覚的プログラミング、並列計算、R との協力的な利用などの機能を設計・実装する。
- B) 二つの課題、「画像・音声のマルチモーダルデータからの不変事象の学習検索」、「対話における言語・身体動作データからの不変情報の抽出と活用」に関して、前者については TRECVID に参加したり、後者についてはモダリティー（発話内容、音韻、身振り等）に着目して対話内での行為の出現パターンとカウンセリングの進行との関係を調査するなどして、さらなる検討を展開する。各課題の検討結果をまとめていき、分野横断的なメタウェア、帰納的手法に関する資料化を行う。
- C) 前年度までに行ったチャンネルモデル（疑似見通しモデル）、マルチパス合成技術、マルチユーザ受信技術、高効率データ処理（組込みシステム）の研究成果の理論的検討を行い、関連領域への発展を試みる。

平成21年度

- A) グリッドコンピュータの利用を含む並列化環境をより使いやすくするために AJAX, Java などの技術を利用する高機能なユーザインタフェースを作成する。そして、R のスーパーコンピュータ化、並列化、高機能だがやさしい利用環境を完成させ、それをすべて公開する。Jasp に関しては特に現在の R に不足している機能である対話的なデータ視覚化、視覚的プログラミングなどの完成度を高める。これらを乱数などの成果も含めて遺伝学および極地科学の研究で利用して、科学研究に有用なメタウェアを明らかにする。
- B) 昨年度に引き続き、各課題の検討結果を踏まえ、分野横断的なメタウェア、帰納的手法に関する資料化を行う。これらにより、映像検索の高度化、人間のコミュニケーションにおける間の問題の解明、知覚神経回路の特定、臨床心理士の支援ツールの開発などを目指す。

- C) 研究成果の論文化，プロトタイプハードウェアのデモンストレーション等，展示会への出展など広報活動を行う。また本方式の開発を実行する過程で得られるソフトウェアおよびプロトタイプハードウェア（具体的メタウェア出力）として，アプリケーションの領域の拡大を図る。

平成22年度以降の展開

本研究で明らかにする情報・通信における有用なメタウェアに基づいて，いろいろなハードウェア，ソフトウェアに関する応用研究が展開できると予想する。例えば，B)の課題では帰納的アプローチによる大規模なマルチモーダルデータを扱う応用研究が展開されると考える。また，C)の課題ではチャンネルモデル（疑似見通しモデル），マルチパス同期技術，マルチユーザ受信技術，高効率データ処理（時変ニコラネットワーク）の理論的展開，応用が期待できる。

3. 研究経費の推移

平成17年度実績： 68,817 千円

平成18年度実績： 55,398 千円

平成19年度見込： 69,300 千円

4. 平成19年度の研究実施体制

[統計数理研究所] 中野純司 田村義保 佐藤整尚 染谷博司 松井知子

福水健次 瀧澤由美 石黒真木夫

[国立情報学研究所] 三浦謙一 速水 謙 佐藤真一 古山宣洋 井上雅史

[国立極地研究所] 岡田雅樹

[国立遺伝学研究所] 五條堀孝 菅原秀明 池尾一穂

[新領域融合研究センター] 深澤敦司 八田野早紀

[早稲田大学] 田邊國士

[京都教育大学] 花田里欧子

[和歌山大学] 入野俊夫

[北海道大学大学院] 宮永喜一

5. 平成19年度研究成果

(1) 成果物（知見・成果物・知的財産権等）

1. A)の成果物：スーパーコンピュータ上で稼働する統計解析システム R，その付加機能として，スレッドを利用した並列化 BLAS と MPI を利用した並列計算機能，疑似乱数と物理乱数利用のための関数，並列化 R のための Web ユーザインタフェース，遺伝研 DNA データバンクとのインタフェース（ソフトウェア）
2. A)の成果物：Java による新しい統計解析環境 Jasp（ソフトウェア）

3. A)の成果物:データビジュアライゼーション用 Java ライブラリ Jaspplot(ソフトウェア)
4. B)の「映像データからの不変事象の学習検索」の知見:高レベル特徴抽出タスクでは、色モーメント、エッジ方向ヒストグラム、Local Binary Patten(LBP)ヒストグラムの3種類の特徴量を用いることが非常に有効である
5. B)の「映像データからの不変事象の学習検索」の知見:Global Alignment(GA)カーネルを導入することにより、可変長の画像セグメント特徴量が有効である可能性がある。
6. B)の「対話データからの不変情報を規定する要因の特定」の知見:クライアントの語句の使用方法がいびつであり、かつそれを言語的に修正することが困難であったときに、セラピストが身振りを用いて非言語的な介入を行うことで、クライアントに自然な概念を持たせうる。
7. B)の「対話データからの不変情報を規定する要因の特定」の知見:セラピストがクライアントの抱える問題を身振りを使って空間的に表現する際に、クライアントの問題把握と食い違うような身振りを使用した場合、混乱が生じて問題の共有が困難になる場合がある。
8. B)の「身体性制約下における外界データの不変情報抽出機械としての知覚神経回路の特定」の知見:一致検出を必要としない両耳信号のコーディングモデル
9. B)の成果物:帰納的学習機械の利用に関する分野横断的なノウハウの報告書「Invariance in Multimodal Data: A Kernel Method Approach」
10. B)の成果物:帰納的学習機械に関するソフトウェア(PLRM, global alignment kernel)
11. B)の「対話データからの不変情報を規定する要因の特定」の成果物:臨床対話における身振りのコーディングスキーマ
12. B)の「対話データからの不変情報を規定する要因の特定」の成果物:コーディングスキーマのチェックツール
13. B)の「対話データからの不変情報を規定する要因の特定」の成果物:編集済み対話ビデオデータ 10 対話分
14. B)の「対話データからの不変情報を規定する要因の特定」の成果物:サンプルトランスクリプト 10 対話分
15. C)の成果物:セル・パス一体化同期方式とそのプロトタイプ装置
16. C)の成果物:高分解能パス同期の方法とそのプロトタイプ装置
17. C)の成果物:低電力同期の方法とそのプロトタイプ装置

(2) 成果発表等

<論文発表>

[学術論文]

- [1] 山本由和, 中野純司, 本多啓介, ” デザインパターンを用いた統計グラフのための Java ライブラリ”, 統計数理, 55, 27-45, 2007

- [2] 本多啓介, 中野純司, ” 3次元平行座標プロット”, 統計数理, 55, 69-83, 2007
- [3] 小林郁典, 山本由和, 中野純司, Lee, Jung Jin, ” 連動する統計グラフィックスによる多変量地理情報データの視覚化”, 統計数理, 55, 113-124, 2007
- [4] Fujiwara, T., Nakano, J., Yamamoto, Y., “Using mathematical expressions in a statistical language”, Computational Statistics & Data Analysis, 52, 650-662, 2007
- [5] Takizawa Y., Yatano S., Fukasawa A., “Advanced Synchronization Scheme for Wideband Mobile Communications,” International Journal of Mathematical Models and Methods in Applied Science, Vol. 2, Issue 1, 2008.
- [6] Takizawa Y, Yatano S., Fukasawa A., “High Performance Digital Signal Processing System for Wideband Mobile Communications,” International Journal of Mathematical Models and Methods in Applied Science, Vol. 2, Issue 1, 2008.

[会議録]

- [1] Ikunori Kobayashi, Junji Nakano, Yoshikazu Yamamoto, Jung Jin Lee, Multivariate geographical data visualization with linked statistical graphics, Bulletin of the International Statistical Institute 56th Session Proceedings, 22-29 August 2007, Lisboa, Portugal
- [2] Duy-Dinh Le and Shin'ichi Satoh, “National Institute of Informatics, Japan at TRECVID 2007: BBC Rushes Summarization,” TRECVID BBC Rushes Summarization Workshop, in conjunction with ACM Multimedia 2007, pp.70-73, 2007.
- [3] D.-D. Le, S. Satoh, and T. Matsui, “NII-ISM, Japan at TRECVID 2007: High Level Feature Extraction,” The TRECVID 2007 Workshop, 2007.
- [4] Marco Cuturi, Jean-Pilippe Vert, Oystein Birkenes, and Tomoko Matsui, “A kernel for time series based on global alignments,” Proc. ICASSP, 2007.
- [5] Oystein Birkenes, Tomoko Matsui, Kunio Tanabe, Tor Andre Myrvoll, “N-best rescoring for speech recognition using penalized logistic regression machines with garbage class,” Proc. ICASSP, 2007.
- [6] Kenji Fukumizu, Arthur Gretton, Xiaohai Sun, and Bernhard Schoelkopf, “Kernel Measures of Conditional Dependence,” Advances in Neural Information Processing Systems 21, 2008 (to appear).
- [7] Takizawa Y., Yatano S., and Fukasawa A., “Digital Signal Processing with Embedded System for Advanced Mobile Communications,” Proceedings of WSEAS International Conference on Circuits, Systems, Signal and Telecommunications, Mexico, Jan. 25, 2008.
- [8] Yatano S., Takizawa Y., and Fukasawa A., “Advanced Synchronization Scheme for Wideband SS Modulation System,” Proceedings of WSEAS International Conference

on Circuits, Systems, Signal and Telecommunications, Mexico, Jan. 25, 2008.

- [9] Takizawa Y., Yatano S., and Fukasawa A., “An Embedded Computer System for Ubiquitous Communications,” Proceedings of the 7th WSEAS International Conference on Applied Informatics and Communications, Athens, Greece, pp.359-362, August 24-26, 2007.
- [10] Yatano S., Takizawa Y., Fukasawa A., “Wideband Spread Spectrum Modulation System for Ubiquitous Communication Services,” Proceedings of the 7th WSEAS International Conference on Applied Informatics and Communications, Athens, Greece, pp.75-80, August 24-26, 2007.

〔解説・総説〕

(該当なし)

〔研究ノート〕

(該当なし)

〔その他〕

- [1] 福水健次, “Kernel Methods for Dependence and Causality,” 9th Machine Learning Summer School 講師 (2007年8月20-31日 ドイツ)。

<会議発表等>

〔招待講演〕

- [1] Ikunori Kobayashi, Junji Nakano, Yoshikazu Yamamoto, Jung Jin Lee, Multivariate geographical data visualization with linked statistical graphics, 56th Session of the ISI, 2007年8月29日
- [2] Toshio Irino, Tom C. Walter, and Roy D. Patterson, “A computational auditory model with a nonlinear cochlea and acoustic scale normalization,” Proc. 19th International Congress on Acoustics (ICA2007), Madrid, 2-7 Sept. (発表日9月3日), 2007.
- [3] Roy D. Patterson, Ralph van Dinther, and Toshio Irino, “The robustness of bio-acoustic communication and the role of normalization,” Proc. 19th International Congress on Acoustics (ICA2007), Madrid, 2-7 Sept. (発表日9月3日), 2007.
- [4] Tomoko Matsui, “Brief overview of TRECVID,” in Norwegian University of Science and Technology (NTNU), 2007年9月3日.
- [5] Tomoko Matsui, “A kernel for time series based on global alignments: toward improvements of isolated-word recognition,” Next-generation statistical models and inference for speech and audio processing, Harvard University, 2007年11月10日.

- [6] 福水健次, “再生核ヒルベルト空間を用いた統計的推論,” 実解析学シンポジウム2007, 2007年10月19日.
- [7] Kenji Fukumizu, “Learning Causal Structure with Kernel-based Dependence Measures,” International Workshop on the Interface between Statistical Causal Inference and Bayesian Networks, 2007年11月4日。

〔一般講演〕

- [1] 本多啓介, 中野純司, “3次元平行座標プロットの格子表示”, 日本計算機統計学会, 2007年5月31日
- [2] Yoshikazu Yamamoto, Junji Nakano, “Parallel and distributed computing in a statistical system Jasp”, The 2007 IASC-ARS Special Conference, Seoul, Korea, 2007年6月8日
- [3] 山本由和, 小林郁典, 中野純司, “カテゴリカルデータの可視化について”, 統計関連学会連合大会, 2007年9月8日
- [4] Ikunori Kobayashi, Yoshikazu Yamamoto, Junji Nakano, “CCmaps and Parallel Coordinate Plot for Visualizing Geographical Data”, The 9th Japan-China Symposium on Statistics, Sapporo, Japan, 2007年9月26日
- [5] Ryoko HANADA, Nobuhiro FURUYAMA and INOUE, Masashi, “Speech-gesture mismatch and how it changes in the problem description by a client,” International Society for Gesture Studies: Third International Conference, Evanston, Illinois, USA, June 18-21, 2007.
- [6] Takizawa Y., “Digital Signal Processing with Embedded System for Advanced Mobile Communications,” WSEAS International Conference on Circuits, Systems, Signal and Telecommunications, Mexico, Jan. 25, 2008.
- [7] Yatano S., “Advanced Synchronization Scheme for Wideband SS Modulation System,” WSEAS International Conference on Circuits, Systems, Signal and Telecommunications, Mexico, Jan. 25, 2008.
- [8] Takizawa Y., “An Embedded Computer System for Ubiquitous Communications,” The 7th WSEAS International Conference on Applied Informatics and Communications, Athens, Greece, August 25, 2007.
- [9] Yatano S., “Wideband Spread Spectrum Modulation System for Ubiquitous Communication Services,” The 7th WSEAS International Conference on Applied Informatics and Communications, Athens, Greece, August 25, 2007.

<著書等>

(該当なし)

<受賞>

(該当なし)

(3) その他の成果発表

1. Super Computing 07

2007年11月9日～17日にアメリカ合衆国レノで開催された Super Computing 07 にブースを出し、テーマ「計算機による帰納的モデリングのための環境」から、スーパーコンピュータ上の R に関するポスターと物理乱数発生装置のデモの展示を行った。その場で多くの研究交流を行うとともに、そのひとりである Michael Mascagni 教授 (Florida State University) を 2008 年 1 月 3 日～11 日に統計数理研究所に招き、並列疑似乱数発生に関する講演およびチュートリアルを行ってもらったとともに、R および乱数に関する討論を行った。

2. 統計数理研究所オープンハウスにおけるデモ展示

2007 年 7 月 19 日に統計数理研究所オープンハウスが開催され、テーマ「モデル化と高効率データ処理に基づく無線データシステムの研究」の成果を展示し、デモを行った。

3. 産学官連携推進会議出展

国立京都国際会館において 2007 年 6 月 16 日 (土)、17 日 (日)、第 6 回産学官連携推進会議が開催され、テーマ「モデル化と高効率データ処理に基づく無線データシステムの研究」の成果を出展した。出展内容は、広帯域無線チャネルの研究と無線通信用組み込みコンピュータの研究である。研究戦略、問題点の明確化、対象のモデル化、対処技術についてパネル説明するとともに、試作プロトタイプハードウェアの動作状況のビデオ映像を放映した。

本研究のブースには大学関係者のみならず、地方公共団体、民間企業などから訪問を受け、大変盛況であった。また、東北大学の無線 LAN 高速化の研究を行うグループとのディスカッション、北海道大学の L S I グループとの研究交流等ができて有益であった。

4. 融合研究シンポジウムでのデモとポスター発表

2007 年 10 月 18 日 (木) にコクヨホールにおいて「融合研究シンポジウム—地球と生命の新パラダイム創造への挑戦—」が開催され、機能と帰納プロジェクトからのポスター発表として、テーマ「モデル化と高効率データ処理に基づく無線データシステムの研究」グループの研究内容のデモ展示を行った。広帯域スペクトル拡散 (W-CDMA) 無線の変復調および同期検出の動態展示、高効率デジタル画像処理装置の動態展示を行った。融合研究プロジェクトの研究者同士のディスカッションのみならず、民間企業からの訪問もあり、有益な研究交流ができた。

従来の研究会型の発表とは異なり、これらデモンストレーションによる目に見える形の成果発表は斬新で面白いとの評を得た。