

(別紙様式)

領域略称名：データマイニング

領域番号： 759

平成14年度科学研究費補助金

特定領域研究に係る研究経過等の報告書

# 情報洪水時代における アクティブマイニングの実現

(平成13年9月 - 平成17年3月)

平成14年 9月

領域代表者 大阪大学・産業科学研究所 教授 元田 浩

連絡先： 06 - 6879 - 8540



# 研究経過等報告書（要約）

## 1. 研究領域名：情報洪水時代におけるアクティブマイニングの実現

## 2. 研究の目的・意義および計画の概要

大量データが氾濫している今日の状況は情報洪水に例えられ、1)膨大な情報空間の的確な監視，効率的な情報収集，2)収集した多様な形態の情報源からの目的にかなった価値ある知識の発掘，3)ユーザの視点の変化や状況変化に即応した知識の洗練，への統合的な対応が急務である．本領域研究では，これら3つの課題に対応し，1)自律的に必要な情報源を探索し前処理を実施するアクティブ情報収集，2)種々の構造を持つデータに適した柔軟なマイニングを実現するユーザ指向アクティブマイニング，3)理解しやすい表示と結果に対するユーザの積極的なフィードバック環境を提供するアクティブユーザリアクションの研究を実施し，3つの連携機能を実現する環境を構築する．実証のためのテストベッドとして医療および化学薬品データのマイニングを取り上げる．前者に関しては千葉大学医学部より提供される慢性肝炎のデータを共通医療データとして各計画研究が取り上げ，上記の3つの機能を統合した“科学発見のスパイラルモデル”による相乗効果を実証する．化学薬品に関しては，新しく開発される薬品の生理活性に対する部分化学構造から，すでに市場に出まわっている薬品に対し必要な警告を与える可能性を検討する．

## 3. 研究の進捗状況

1年間を経過した現在，研究は，上述の3つの課題のそれぞれで順調に進展している．1)の情報収集では，効率的な情報収集のためのメタ情報源の自動学習，マイニング結果に関連のある既存知識の高速探索，前処理の半自動化，伝言ゲーム型の情報収集・前処理結果の交換などが，2)のマイニングでは，構造データからの共起パターン的高速発見，時系列データの抽象化，データの特徴に最適なマイニングアルゴリズムの自動構築，スパイラル的例外性発見，専門家が容易に関与し得る環境の構築などが，3)のユーザリアクションでは，マイニングの対象となる元データやマイニング結果の知識の視覚化，視覚化を通じた専門家の主観的発見プロセスのモデル化などが大きく進展した．本特定領域全体で取り上げている共通医療データである肝炎データの解析では，各要素技術を連携して前処理，マイニング，評価のサイクルが周り始めた．専門医とのインタラクションも活発に行われており，専門医の興味を引く具体的な成果も出はじめた．化学薬品データのマイニングに関しても，部分構造の発見と類似構造の高速探索に成功し，発ガン性予測の国際データマイニングコンテストで，予測精度と専門家による理解の容易さの双方で，参加14グループ中，第1位の評価を得るなど顕著な成果が得られた．



## 1. 領域全体の研究目標

通信技術を含む計算機ハードウェアの急速な進歩により、大量の情報が各種ネットワークを飛び交い個人がアクセスできる情報量が飛躍的に増加している。データマイニングはこのような大量データからの知識発見を目指す新しい技術であり、知識管理業務の強力な武器になるものとその有用性が期待されている。しかし、大量データが氾濫している今日の状況は情報洪水に例えられ、1) 膨大な情報空間の的確な監視、効率的な情報収集が難しい、2) 収集した多様な形態の情報源から目的にかなった価値ある知識を発掘する方法が確率されていない、3) ユーザの視点の変化や状況変化に即応した知識の頻繁な更新に対処できないなどの問題が大きくクローズアップされている。情報収集・データ解析・目的設定変更のサイクルが高速回転し、個人も組織も情報洪水の中で疲弊しているのが現状である。

本領域研究では、このような状況を打破するために、3つの課題に対応し、以下の3つの研究項目を掲げる。

### A01: アクティブ情報収集

不特定・非定常・大規模・分散知識源の中から、ユーザの目的や興味に合致するデータやそれらの関連を効率良く探索し前処理するための情報収集技術を、メタ情報源の活用、ヒューリスティック探索知識の活用、機械学習法の活用など、最新の情報処理技術を駆使して開発する。

### A02: ユーザ指向アクティブマイニング

多様な形式や多種の情報源に対応でき、汎用性と状況の変化にも対応できる柔軟性を持つマイニング手法を開発する。とくに、テキスト情報に代表される半構造化データ、分子化学情報・ネットワーク情報・医療情報に代表される構造化データからのマイニング、これら個別のデータに最適なマイニング手法の自動構築、状況変化検知に強力な例外性の発見技術に注力する。

### A03: アクティブユーザリアクション

具体的な問題領域を対象にマイニングシステムを構築し、発掘した知識を、ユーザにとって有用なものとするための仕組(知識の表示法、評価手法、ユーザからの効果的なフィードバックの手法)を具体化する。

領域全体で連携してこれらの課題を実証するため、具体的な問題として、医療と化学薬品データのマイニングを取り上げる。とくに前者に関しては千葉大学医学部より提供される肝炎のデータを共通医療データとして各計画研究が取り上げ、上記の3つの機能を統合した“科学発見のスパイラルモデル”による相乗効果を実証する。慢性B型、C型肝炎がどのような経過をたどって、肝硬変、肝臓癌に至るかについてはよくわかっていない。肝炎の病理像(繊維化の程度)と血液検査データとの間にある程度の相関があると推測されているが、明

らかな知見に至ってはいない。侵襲度が高く簡単にできる検査ではない肝生検に代わって、(生検)検査以外の血液データから疾患の予後が予測できれば医学的にも大きな貢献になる。共通医療データからのアクティブマイニングは、現在、新しい医療行為として注目されている Evidence Based Medicine (EBM) (「エビデンス(科学的根拠)に基づいた医療」)を実践するための有効な手段を提供するものと期待されており、本プロジェクトで、EBMの効果を実証する。化学薬品に関しては、毎年新しく開発される多数の薬品の生理活性に対する部分化学構造から、すでに市場に出まわっている薬品に対し必要な警告を与える可能性を検討する。

## 2. 領域内における研究組織と研究班の連携状況

本領域の研究は、研究の評価と推進をはかる総括班の下に、1. 領域全体の研究目標の3つの課題に対応して、以下に示すの3つの班から構成され、各班は、それぞれ3, 4, 3の計10個の計画研究からなる。

### A01 班：アクティブ情報収集

計画研究	A01-02	山田誠二	WWWにおけるメタ情報源の獲得
計画研究	A01-03	北村泰彦	分散動的情報源からのアクティブ情報収集
計画研究	A01-04	沼尾正行	多段階学習方式によるデータ収集と前処理の自動化

### A02 班：ユーザ指向アクティブマイニング

計画研究	A02-05	元田 浩	構造データからのアクティブマイニング
計画研究	A02-06	山口 高平	メタ学習機構に基づくアクティブマイニング
計画研究	A02-07	鈴木英之進	例外性発見に基づくスパイラル的アクティブマイニング
計画研究	A02-08	松本裕治	利用者からの要求を考慮したテキストデータからの知識抽出

### A03 班：アクティブユーザリアクション

計画研究	A03-09	津本周作	ラフ集合に基づくアクティブマイニングによる診療情報生成システムの開発
計画研究	A03-10	岡田 孝	アクティブマイニングによる化学物質群からのリスク分子発見
計画研究	A03-11	大澤幸生	ヒューマン・システム・インタラクションに基づく知識評価と選択

本特定領域は、各計画研究間の密接な連携がなければ初期の目標を達成できない。そのため、総括班会議の他に、班会議、計画研究代表者会議を定期的で開催し、各計画研究の進捗状況を全員が把握すると同時に、問題点を徹底的に討議し解決策を全員で議論している。とくに、肝炎データ解析に関しては全ての計画研究が参加しているので、これらの会議には肝炎データ提供者の千葉大学医学部の専門医にも参加してもらい、結果を評価して頂いている。

多くの研究テーマが並行して走っており、それらの成果を交換しながら、領域全体としての

研究成果に集約されていく。共通データ解析の目的は、各要素技術を連携して前処理、マイニング、評価のサイクルを繰り返し、“科学発見のスパイラルモデル”を実践・実証することである。そのため、総体的には各計画研究は所属する班の目標達成のための研究遂行に力を注いでいるが、お互いに少しずつオーバーラップしながらマイニングの全プロセスを経験している。各計画研究がアクティブマイニングのどの部分に寄与しているかを図1に示す。

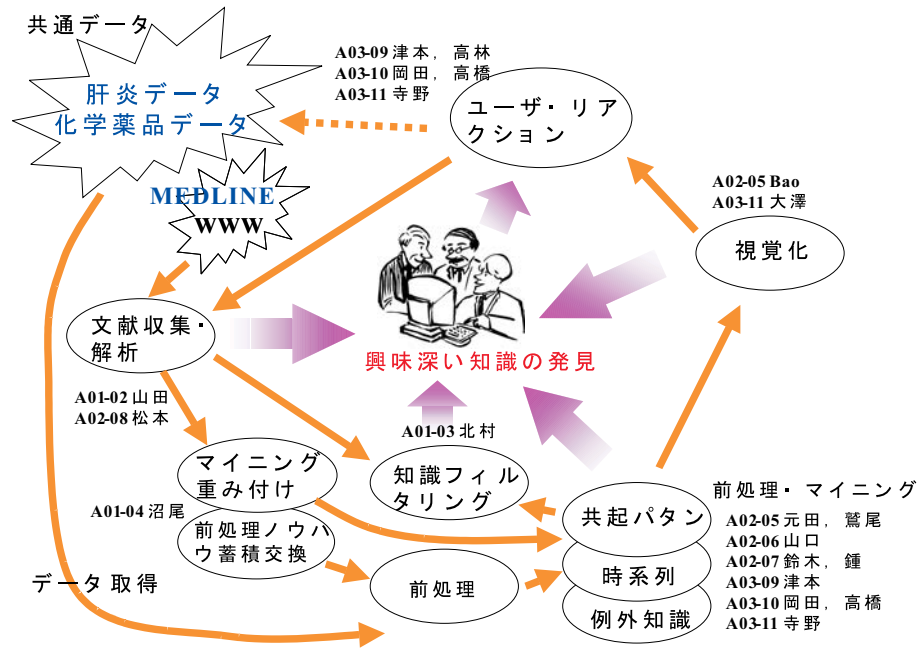


図1: アクティブマイニングプロジェクトの各研究計画の連携状況

### 3. 領域内の研究の進展状況とこれまでの主な研究成果

研究開始後1年間を経過した現在、各研究は、上述の3つの課題のそれぞれで順調に進展している。要約すると、1)アクティブ情報収集では、効率的な情報収集のためのメタ情報源の自動学習、マイニング結果に関連のある既存知識の高速探索、前処理の半自動化、伝言ゲーム型の情報収集・前処理結果の交換などが、2)ユーザ指向アクティブマイニングでは、構造データからの共起パターンの高速発見、時系列データの抽象化、データの特徴に最適なマイニングアルゴリズムの自動構築、スパイラル的例外性発見、専門家が容易に関与し得る環境の構築などが、3)アクティブユーザリアクションでは、マイニングの対象となる元データやマイニング結果の知識の視覚化、視覚化を通じた専門家の主観的発見プロセスのモデル化などが大きく進展した。また、本特定領域全体で取り上げている共通医療データである肝炎データの解析では、各要素技術を連携して前処理、マイニング、評価のサイクルが周り始めた。専門医とのインタラクションも活発に行われており、専門医の興味を引く具体的な成果も出はじめた。化学薬品データのマイニングに関して、部分構造の発見と類似構造の高速探索に成功し、発ガン性予測の国際データマイニングコンテストで、予測精度と専門家に

よる理解の容易さの双方で，参加14グループ中，第1位の評価を得るなど顕著な成果が得られている．

以下，その主要成果を計画研究ごとに箇条書きにして示すが，より詳しくは本報告書の付録1を参照して頂きたい．また，項目別主要成果を別添の研究成果概要（中間ヒアリング参考資料1）に，共通データ解析の観点からまとめた主要成果を別添の共通データ解析報告書（中間ヒアリング参考資料2）にまとめたので，そちらも参照して頂きたい．

#### A01 班：アクティブ情報収集

##### 1. WWWにおけるメタ情報源の獲得

- 1) 適合フィードバックにより，情報源判定のための判別ルールを学習するシステムを開発し，効果を実験で検証した．
- 2) 免疫ネットワークを用いて文書集合間の時系列的関連を考慮した可塑的クラスタリング法を提案し，情報ストリームの抽出・可視化を可能にした．
- 3) メタ情報源などの獲得された情報源の更新を自動的に検知，通知する手法を開発した．

##### 2. 分散動的情報源からのアクティブ情報収集

- 1) 分散動的情報源から効率よく情報収集を行う知的情報収集手法を開発し，発見知識フィルタリング手法への適用可能性を検討した．
- 2) マルチエージェントによる情報収集・統合・監視システムの負荷分散のための動的タスク再割り当て手法を提案した．

##### 3. 多段階学習方式によるデータ収集と前処理の自動化

- 1) 前処理過程に含まれる処理を効率的に行うデータ構造，及び自動化アルゴリズムを提案し，実験システムを構築した．
- 2) 共同作業を支援するための，伝言ゲーム型の情報収集および前処理結果の交換を提案し，システムを実装し実験した．
- 3) 構文的な特徴を共有するWWWページを収集し，HTML構文に関する制約を利用して知識を抽出する手法を提案し，有効性を確かめた．
- 4) 高次元インデックス技術を開発し，空間属性と非空間属性を併せもつ大量の時系列高次元データの前処理について展望を得た．

#### A02 班：ユーザ指向アクティブマイニング

##### 1. 構造データからのアクティブマイニング

- 1) 構造の大きさに線形な処理時間で多頻度連結部分グラフを抽出するアルゴリズムを改良し，肝炎データに適用し，実用的な処理時間で部分構造が抽出できることを確認した．



- 2) 各分岐ごとに同じ操作を再帰的に繰り返すことにより、分類に効果的な属性を逐次的に構築し利用するグラフ構造データ向きの決定木生成法を提案した。
- 3) すべての多頻度誘導部分グラフを高速抽出するアルゴリズムを連結部分グラフのマイニングへ拡張し、処理時間を一桁以上少なくすることに成功した。
- 4) 規則やデータの階層構造を視覚化し、ユーザがマイニングプロセスに積極的に関与し必要なモデルを選択できるマイニング環境 D2MS を構築した。
- 5) ラフ集合モデルを用いて文書を表現し、文書を階層的・非階層的にクラスタリングするアルゴリズムを開発し、情報検索に適用し良好な結果を得た。
- 6) 肝炎時系列データを抽象化し記号データで定性的に表現し、ユーザ指向マイニングシステム D2MS を用いて解析し、B 型肝炎と C 型肝炎の違いや生検データに基づく肝炎の進行状態を判断するルールを発見した。
- 7) 経営データを対象に、従来の相関規則を拡張し、association strength なる概念を導入して、複数の時点における複数のブランド間の購買関連性の分析を可能にした。

## 2. メタ学習機構に基づくアクティブマイニング

- 1) リポジトリに基づく帰納アプリケーション構築支援環境にコミティ学習メソッドを追加し、公開データで性能評価し、既知の結果を上まわる正答率を得た。また、処理の高速化を目指し、6CPU からなる並列マシンを実装し、70 ~ 80 % 程度の台数効果が得られることを確認した。
- 2) 肝炎時系列データより一定長のサブシーケンス群を切り出しクラスタリングした後、決定木学習により、検査項目値から GPT の 1 年間の変化を予見するルールを学習した。乳びと GPT の関連性、GPT の周期性の予測など、専門医より興味深いというコメントを得た。
- 3) 要求レベル - 問題解決レベル - 実装レベルと対応づけた 3 層のメソッドリポジトリを検討し、ビジネスドメインにおいて、プロトタイプを試作し、その有用性を確認した。
- 4) ドメインモデルを構成する概念要素の仕様を与えるドメインオントロジーを構築する手法を開発し、計算機可読型辞書とテキストコーパスを用いてツールを試作した。

## 3. 例外性発見に基づくスパイラル的アクティブマイニング

- 1) 時系列データに関する新しいデータスカッシング手法、集合値属性を許容する決定木学習法、多次元クラスに有効な分類モデルである花つき決定木などを組み込んだ、スパイラル的例外性発見基盤システムを構築した。慢性肝炎データに適用し、異常検査値の類形を発見し、これらの類形の推移に基づき患者をクラスタリングした。
- 2) ルール発見に重要である複数の指標に関する誤差が指定値以下である場合を、指定した確率以上とするために必要な例数を与えるルール発見の新しいサンプリング理論を構築し、状況変化の検出法を提案した。

- 3) ブースティングの結果をフィードバックすることにより，データの前処理であるデータスカッシングのパラメータを自動調節する手法を開発した．
- 4) 対象データを変更し，マルチデータソースから興味深い特異データを抽出することにより，特異ルールを発見する手法を開発し，抗原抗体反応に関するアミノ酸配列データや慢性肝炎データなどに適用し，提案法の有用性を示した．

#### 4. 利用者からの要求を考慮したテキストデータからの知識抽出

- 1) 前後の文脈，あるいは，単語の綴り（特に接頭，接尾表現）を手がかりに，未知語を含む英文文書内の単語の品詞を Support Vector Machine (SVM) を用いて高効率，高精度に推定する手法を開発した．
- 2) 基本句の同定を各単語へのラベル付与問題とみなし，SVMを用いた学習システムを複数混合することにより，従来法を上回る精度で基本句を抽出する手法を開発した．
- 3) 医学/生物学関係の論文概要を提供している Medline アブストラクトを題材とし，テキスト中に現れる名詞句を対象に，その意味クラスを用語の内部情報の情報と前後の文脈情報を用いて推定する実験を行い，約 70%から 88%の正答率を得た．

#### A03 班：アクティブユーザリアクション

##### 1. ラフ集合に基づくアクティブマイニングによる診療情報生成システムの開発

- 1) 肝炎時系列データを平滑化・セグメント化し，多重スケールマッチングを用い，時間スケールのちがう検査値推移パターンを比較可能とした．
- 2) ラフ集合論の識別不能性の概念に基づき，多重スケールマッチングでセグメント化した系列を分類するクラスタリング法を開発し，肝炎データに適用し，GPT の時系列的变化が C 型肝炎のインターフェロンによる投与効果を最も反映しているなど，多くの知見を得た．

##### 2. アクティブマイニングによる化学物質群からのリスク分子発見

- 1) カスケードモデルを用い芳香族ニトロ化合物の変異原性を解析し，「ortho 位置換基の立体障害が変異原性に重要な因子となる」などの有意義な知見を多数得た．同一手法を用い，発ガン性予測の国際データマイニングコンテストで，予測精度と専門家による理解の容易さの双方で，参加 14 グループ中，第 1 位の評価を得た．
- 2) 多数のルールを主ルールと関連ルール群へ組織化し，識別力が大きく低下するリッジ領域を検出する方法を開発し，データスケープ探索（専門家自身が導出されたルール群の助けを借りて，データの特徴を多面的に捉えること）を可能とした．
- 3) フラグメントを質量数で特徴付けた TFS の類似性から活性分子をの同定する手法を開発した．4 種のドーパミン受容体の活性に対する MDDR データベース中の治験薬からの受容体予測に適用し，79%の高予測精度を達成した．

### 3. ヒューマン・システム・インタラクションに基づく知識評価と選択

- 1) 血液中の複数の酵素データを用いて病態モデルを臨床検査の実データから因子分析を用いて構築し，因子スコアを視覚化し，重要な病的異常値の見逃しを防止する効果を得た．
- 2) 血清蛋白分画データの病態パターンを自己組織化マップを用いクラスタリングし，臨床診断につながる付加価値情報を得た．
- 3) メタ学習の結果と背景知識を基に，構造方程式モデリングを用い，EBMの概念に合致する簡潔かつ汎用性のある肝機能検査データの因果モデルを構築した．
- 4) チャンスを評価するメタな指標として，行動の提案可能性，気付かれにくさ，成長性を提案し，人と人の間でチャンスに基づく提案や採択を通してチャンスを選択してゆくプロセスを実現し，視覚化データマイニングの基礎を構築した．
- 5) チャンス発見の二重らせんモデルを提案し，それに基づくヒューマン・マシンインタラクションを促進するシステムを開発し，Interactive Data Mining/Interactive Mining Validation，Data Visualizerの基盤を構築した．この手法を医療データに適用し，肝炎の専門家との対話を二重螺旋モデルによって進める研究を現在進めている．

### 4. 今後の領域の推進方策

共通データ解析が各班間の連携，班内の各計画研究間の連携に大きな役割を果たしている．現在までのところ，研究は順調に進展しており，大きな障害は予想されない．総括班会議，班会議，計画研究代表者会議を定期的開催し，研究の進捗度の評価，問題点の摘出と解決策の提示，各計画研究の個別の研究へのフィードバックを実施してきたことが功を奏している．また，専門医の絶大なる協力を得ていることも順調な進展に大きく寄与している．今後とも，この体制を続行する．

新しいマイニング手法も開発され，個別にもいい成果が出はじめているので，積極的な外部発表を心がける．過去一年間に本特定領域の総括班が中心になって企画した，研究会，国際ワークショップも10件に達した．さらに，今年度内に3件の国際ワークショップを予定している．来年度以降もこの方針を維持して行き，アクティブマイニングの国際発信源としての責を担う．

### 5. 領域を推進するための問題点と対応策（領域内の評価体制と評価状況）

本特定領域では，その研究成果や各種活動を客観的に評価して頂く目的も含めて，総括班メンバーに計画研究代表者以外のデータマイニング，知識発見，機械学習の研究分野の権威の先生方（有川節夫[九州大学]，宮野悟[東京大学]，有村博紀[九州大学]）に，ご参加頂いている．これらの先生方には，日頃から総括班会議や研究会にご参加頂き，研究の推進や成果についてコメントを頂いている．さらに，国際ワークショップなどを開催する機会に，海外の専門家にも本特定領域の成果を紹介し，コメントを頂くように心がけている．今回，中

間ヒアリングを契機に，総括班以外の外部の有識者（長尾真[ 京都大学 ]，大須賀節夫[ 早稲田大学 ]，藤田 稔夫[ 前京都大学 ]，朴 勤植[ 大阪市立大学 ]，野村 文夫[ 千葉大学 ]，Paul Compton [ The University of New South Wales ]，Geoffery Webb [ Monash University ]，Ulrich Reimer [ Swiss Life ]）にも評価していただいた．本報告書の付録 2 として評価結果を添付する．評価は総じて高いものであった．これらの評価コメントを真摯に受けとめ，今後の研究推進に反映させる．

## 6. 研究成果公表の状況

発表件数は多いため，件数のみ記す．招待講演 / 招待論文 21 件，公開ソフトウェア 6 件，受賞 5 件，報道 2 件，分担者によるワークショップ，シンポジウム，パネル開催 15 件，ジャーナル論文 82 件，国際発表 188 件，国内発表 180 件，著書 / 編書 11 件，著書（章分担執筆）48 件，解説 22 件である．詳細は本報告書の付録 3 に添付した．参照して頂きたい．

## 付録 1：主要研究成果概要と今後の計画

### ( A01 班 ) アクティブ情報収集

研究項目 A01 では、必要な情報リソースをシステム自らが能動的に探しに行く仕組みを確立することを目標とし、ネットワーク上に分散したデータベース内から使用者の目的や興味に関連のありそうな情報を、積極的に探索する技術やテキスト情報処理技術、それを後段のマイニング処理に引き渡す前処理を行うことを目的としている。特に、WWW におけるメタ情報源の獲得、分散動的情報からのアクティブ情報収集、多段階学習方式によるデータ収集と前処理の自動化、という三つの課題に焦点をあてて、研究を進めている。以下、各研究課題に関して、現在までの研究成果概要と今後の研究計画についてまとめる。

#### 1. WWW におけるメタ情報源の獲得

本研究課題の目標は、WWW などの膨大な情報空間に存在するメタ情報源を（半）自動的に収集し、それを利用することにより、従来手法よりも質と量、そして効率において飛躍的に向上するような情報収集を実現することである。これまでの WWW における情報収集とは、“Web ページの収集”を意味していた。これに対し本研究では、有用な Web ページの所在や質に関する情報源であるメタ情報源の収集を研究目的とする。メタ情報源には、情報の所在についての情報を提供する「LI 情報源」とコンテンツに関するメタデータを提供する「MD 情報源」がある。最終的には、以上のようなメタ情報源による情報収集を医療データの収集に適用する。また、メタ情報源の獲得のための要素技術として、対話的文書検索、文書集合のクラスタリング、WWW の更新管理などを開発する。

現在までの研究成果

##### 1) 関係学習による対話的文書検索

LI 情報源の収集のためには、インデックスページなどの特定が重要である。LI 情報源は、ハイパーリンクからなるグラフ構造をもち、またその Web ページのコンテンツにも特徴がある。よって、リンク構造とコンテンツからなる制約により、LI 情報源を判定できる。ある Web ページがある LI 情報源に属するか否かの判定は、クエリに対して文書が適合するか否かの判定と同様である。そこで、情報検索で広く使われている手法である適合フィードバックを基盤として、LI 情報源判定のための制約である判別ルールを学習するシステムの開発とそれを用いた実験を行い、良好な結果を得た。

##### 2) WWW 上の情報収集 / 可視化のための免疫ネットワークを用いたクラスタリング

MD 情報源を獲得するには、あるユーザやユーザグループにおいて閲覧されている Web ページの系列に代表される情報ストリームにおいて、情報の流れの抽出、また情報の流れ間の関係を捉えることが必要になる。このような機能実現の要素技術として、免疫ネットワークを用いて文書集合間の時系列的関連を考慮したクラスタリング

手法である可塑的クラスタリングを提案した。

### 3) Web ページの部分更新のモニタリング

収集された WWW における情報は、時々刻々変化する動的なものであるため、一旦獲得された情報源で頻繁に行われる更新を自動的に検知するシステムが必要不可欠になる。そこで、Web ページ上のユーザによって指定された一部分の情報に着目し、その部分に特定の更新があった場合のみ、その更新をユーザに提示する部分更新モニタリングシステムを開発した。このシステムでは、ユーザが監視したい部分を特定するためのルールを分類学習によって獲得する。また、特定された部分の更新がユーザにとって必要なものであるか判定するためのルールも合わせて学習する。これにより、メタ情報源などの獲得された情報源の更新を自動的に検知、通知することが可能になった。

## 今後の研究目標

今後は、これらの成果に基づき、1) LI 情報源の獲得のためのリンク構造制約の学習 2) 対話的文書検索システムへの SVM の適用 3) 医療文献データベース MedLine への対話的文書検索システムの適用 4) 医療情報掲示板における情報ストリームの抽出と追跡 5) マイニングされた知識に基づく情報検索のための知的ユーザインタフェースに取り組む。

## 2. 分散動的情報源からのアクティブ情報収集

本計画研究の目標はインターネット上に存在する動的な情報源から情報収集・統合することでデータマイニングを支援することにある。特にインターネット情報源の多くは頻繁かつ非同期に情報が更新されることがあり、それを考慮した情報収集を行う必要がある。ここでは実時間探索や制約充足手法に基づく効率よい知的情報収集アルゴリズムの開発と肝炎データマイニングにおける発見知識フィルタリングへの応用について研究している。

## 現在までの研究成果

### 1) 分散動的情報源からの知的情報収集と発見知識フィルタリングへの応用

分散動的情報源から効率よく情報収集を行う知的情報収集手法を開発した。本手法では情報収集に関与する情報をメタ情報、静的情報、動的情報に分類している。動的情報は頻繁に更新される情報であり、情報収集の主な対象となる。動的情報を効率よく収集するためには、動的情報の価値を評価するための比較的安定した情報が必要になり、それが静的情報である。メタ情報は動的情報と静的情報を統合する役割を果たす。提案手法の有効性を三つの情報源を統合した航空便空席照会プロトタイプシステムを開発することにより示した。さらに本手法のデータマイニング支援への応用として、発見知識フィルタリング手法の検討を行った。

### 2) マルチエージェントによる情報収集・監視・統合

マルチエージェントによる情報収集・統合・監視システムにおける，特定の情報収集エージェントへのタスクの集中という問題に対し，動的にタスクの再割り当てを行うという手法を提案した．具体的には，タスク割り当ての問題を動的分散制約充足問題として定式化し，既存の分散制約充足アルゴリズムで解くというものである．なおこの手法では，エージェントが安定な値を優先的に選択することにより，動的な問題に対する処理の効率化が図られている．

## 今後の研究計画

知的情報収集アルゴリズムの具体的な応用領域として本特定領域研究の共通分野である肝炎データマイニングを対象とした発見知識フィルタリングに応用する．データマイニングにおいては多量の知識が発見されるが，それをインターネット上の医学情報源からの情報検索と統合することにより，そのフィルタリングを行う．このために，データマイニングを担当する A02 班とより緊密な連携をとる．さらに，肝炎データを扱うためには医学的な知識は不可欠であるので，医療情報専門家に研究協力者としての参加を仰ぎ，その助言を得て研究を進める．

### 3. 多段階学習方式によるデータ収集と前処理の自動化

従来からあるデータ処理技術では，分散した数多くのデータベースの中からあらかじめ必要なデータを選んで収集する作業を行う必要があった．これらの作業は「前処理」と呼ばれ，膨大な人手と時間を要していた．またデータに基づいて分類を行うだけで，対象間の関係を見出す能力も不足していた．以上の背景から，関連する情報を収集し，それらの関係を自動的に発見する手段として，データベース間の通信ネットワークに学習能力を持たせることを提案する．本研究の目的は，トポロジや情報伝達の優先順位を学習により動的に変化できる Global Intelligence Associating Network (GIANT) の構成法を明らかにした上で，現実のデータに基づいて動作するシステムを構築することにある．

## 現在までの研究成果

### 1) 前処理支援システムの構築

蓄積されたデータに対して前処理を施し，解析アルゴリズムが直接扱えるデータに変換する過程を明確に定義し，その過程に含まれる処理を効率的に行えるデータ構造，及び自動化アルゴリズムを提案し，実際のシステムを構築した．具体的には，前処理で扱うデータをすべて XML 形式に統一し，利用者が処理の過程を記述する部分には XML から一意に変換されるデータ構造を用いて効率化を図り，利用者が一度作成したデータ変形フィルタについては自動的に並べ替えて利用者に提案する．

### 2) 伝言ゲーム型の情報収集

情報収集および前処理は，情報提供者，ドメイン専門家，マイニング専門家の共同作業になることがほとんどで，それらの間で大量のデータが交換され，更新が頻繁に行

われることに特徴がある．このような共同作業を支援するため，通常，メールと Web ページが用いられるが，不便な点が多い．データマイニングの作業はルーチンワークではないので，従来のグループウェアの適用も困難である．このような問題を解決する方法として伝言ゲーム型の情報収集および前処理結果の交換を提案し，システムを実装し実験を行った．

### 3) HTML のリンク構造と構文的特徴に基づく知識獲得

現在の検索エンジンのテキスト照合の高速性を利用することで，構文的な特徴を共有する WWW ページを収集する．リンク集についてはバックリンクページがリンク集である頻度が高い事を利用し，データ集については予め XML の雛型を与えておくことによって検索エンジンを利用する．収集した WWW ページから知識を獲得するには，HTML 構文に関する制約を利用して知識を抽出する．これらの特徴を有するシステムを実装し，有効性を確かめた．

### 4) 高次元インデックス技術を用いた検索処理性能向上について

ヒトの思考は空間的な直感に基づいていることが多い．その観点から，データマイニングの対象となるデータとして，位置情報を取り上げ，高次元インデックス技術を開発した．これにより，空間属性と非空間属性を併せもつ大量の時系列高次元データの前処理について，展望を得た．

## 今後の研究計画

伝言ゲーム型の情報収集を実際に各計画研究に使ってもらい，有効性を評価すると共に改良すべき点を洗い出す．組み込まれている推薦機構について検討し，改良を行うと共に，前処理支援システムとの連携を図る．情報収集の結果により，属性を重み付けすることにより，肝炎データのマイニングを制御し，情報収集とマイニングを連携させることを試みる．位置情報の処理を通じて，時系列高次元データの扱いについて研究を進める．A01 班の情報収集および前処理技術を集大成したシステムを設計する．

## ( A02 班 ) ユーザ指向アクティブマイニング

研究項目 A02 では，アクティブ情報収集機構で収集・前処理された情報から，ユーザの目的や興味に照らして重要・有用と思われる知識を発掘することを目標とし，多様な表現形式あるいは多種の情報源に対応できる「汎用性」，並びに，ユーザを含めた状況の変化に対応できる「柔軟性」の両者を兼ね備えた，大規模データからのマイニング手法の開発を目的としている．特に，構造化データからのマイニング，マイニングアプリケーション自動構築，例外性の発見，テキストデータからのマイニング，という 4 つの研究課題に焦点をあてて，研究を進めている．以下，各研究課題に関して，現在までの研究成果概要と今後の研究計画についてまとめる．



## 1. 構造データからのアクティブマイニング

本研究課題の目標は、大規模なグラフ構造データ、空間分布構造データ、時系列構造データ、半構造データ、制約構造データなどの構造データを対象とし、ユーザの価値観を反映した重要なあるいは興味深い部分構造ならびにその特徴を知識として、ユーザの許容時間内に発掘するために必要な基礎技術を開発することである。

現在までの研究成果

### 1) グラフ構造データからの特徴的な部分グラフの発見手法 GBI の性能向上

逐次ペアのチャンキングというアイデアに基づき、構造の大きさに線形な処理時間で多頻度連結部分グラフを抽出するアルゴリズムを改良した。チャンクの指標としては従来通り頻度を用い、それとは別に分類性能に基づく指標を用いて部分構造を並行して抽出することにより、分類問題により柔軟に対応可能とした。また、単純な欲張り探索の簡便さを活かしつつビーム探索へ拡張することにより探索空間を増加させ、重要な部分構造の見落としを減少させた。発癌性化合物や変異源性化合物の分類に適用し良好な結果を得た。また肝炎データに適用し、実用的な処理時間で部分構造が抽出できることを確認し、初期結果を専門医師に評価してもらった。

### 2) GBI 法の属性構築への適用

決定木のノード生成時に分離能力最大のペアを選定し、チャンクして1つのグラフノードとし、各分岐ごとに同じ操作を再帰的に繰り返すことにより、分類に効果的な属性を逐次的に構築し利用するグラフ構造データ向きの決定木生成法を提案し、実装した。小規模の人工データに適用し、正しく機能することを確認した。現在、より規模の大きなデータに適用し問題点を摘出している。

### 3) 多頻度誘導部分グラフ抽出手法 AGM の連結部分グラフマイニングへの拡張

実問題では連結部分グラフパターンを扱えば解ける問題が多いことを考慮し、グラフ構造データからすべての多頻度誘導部分グラフパターンを高速導出する AGM 手法を、すべての多頻度連結部分グラフパターンを高速導出する AcGM 手法に拡張した。これにより計算時間を一桁以上少なくすることに成功し、実規模の問題にも適用可能となった。

### 4) ユーザ指向のマイニングシステムの構築

規則やデータの階層構造を視覚化し、ユーザがマイニングプロセスに積極的に関与し必要なモデルを選択できるマイニング環境 D2MS を構築した。D2MS はデータマイニングのプロセスに基づき、ユーザがマイニングの主導権を持つことができるように配慮したシステムで、現在は決定木構築とルール生成の2つのモジュールならびに多彩な視覚化のツールを内蔵している。得られたルールによってどのデータが説明できるのかなど、元のデータと発掘された知識の関係を容易に把握できる機能があり、マ

イニングの途中経過を専門家が容易に理解出来るよう配慮した使いやすいシステムである。

#### 5) 文書クラスタリング

対称性と反射律のみを許すラフ集合モデル TRSM を用いて文書を表現し，文書を階層的にクラスタリング，非階層的にクラスタリングするアルゴリズムを開発した．それを情報検索に適用し良好な結果を得た．

#### 6) 時系列データの抽象化による肝炎データ解析

肝炎時系列データを抽象化し記号データで表現し，ユーザ指向マイニングシステム D2MS を用いて解析した．エピソードと状態・傾向の2要素を定義し，元のデータをこの2要素で表現し抽象化する．問題に適したエピソードの決定法ならびにエピソード内の状態と傾向を定性的に表現する手法を開発した．B型肝炎とC型肝炎の違いや生検データに基づく肝炎の進行状態を判断するルール，とくに，現在知られている医学的な知見とは異なるルールの発見を試みた．

#### 7) 時系列データからのマイニング手法の開発

経営データを対象に，従来の相関規則を拡張し，association strength なる概念を導入することにより，複数の時点における複数のブランド間の購買関連性の分析を可能にした．その結果，時系列データとして蓄積されている POS データの特徴を活かした，時間の経過にそった状態の遷移を表現することが可能になり，従来，十分に取り扱うことができなかった複雑な社会現象を時系列データとして分析することができる可能性が出てきた．

### 今後の研究計画

各課題に対して，基本的な手法が固まってきた．A02 班の計画研究と連携を取ながら，GBI，D2MS を用いた肝炎データの本格的な解析に着手する．前処理，マイニング，専門医による評価のサイクルを繰り返し，手法を改良しながら，専門医にとって意味のある医学的知識の発見を目指す．また，A03 班の化学物質群からのリスク分子発見の計画研究と協力し，AGcM を生理活性が既知の化学物質の部分構造発見に適用し，活性プロファイル知識ベースの構築を開始する．ビジネスデータに対しては食品、アパレルなどの様々なビジネス領域の様々な競争状態のブランド群に適用し，問題点を明らかにし，時系列データの取り扱いに関する統一的な枠組みを構築する．また，様々なビジネス現場の人々を集め，ビジネスマイニングワークショップを開催し，研究成果を社会に対して積極的に公開する．

### 2. メタ学習機構に基づくアクティブマイニング

本研究課題の目標は，データマイニングシステムの開発プロセスを詳細に分析し，その中で特に重要な「データ加工」と「データマイニング」プロセス，並びに「ユーザの主観的基

準」をリポジトリ化し、メタ学習機構に基づいて、これらのリポジトリを統合しながら、ユーザの使用目的に合致したマイニングアプリケーションを半自動合成するツールを開発することである。

現在までの研究成果

1) リポジトリに基づく帰納アプリケーション構築支援環境

代表的な帰納アルゴリズムを分析し、帰納メソッドリポジトリを構築してきた。あらたにコミティ学習メソッドを追加し、ポルト大学の StatLog プロジェクトから提供されている 8 種類のデータセットを利用して、帰納アプリケーションの自動合成実験を行った。その結果、本ツールで合成された帰納アプリケーションの平均正解率は、StatLog プロジェクトで調査された 24 種類の代表的な帰納アルゴリズムのどの平均正解率よりも高い値を示した。また、合成プロセスを並列処理することを考察し、6CPU からなる並列マシンを実装し、70 ~ 80 % 程度の台数効果が得られることを確認した。さらに、効率的な仕様探索を実現するために、相関ルールに基づいて、仕様書換ルール(メタルール)の学習を試みた。その結果、ランダム探索よりは安定した仕様書換が実現できることが確認されたが、最良の仕様を短時間に見つけ出す点については不十分であり、メタルール自身の構造について課題が残った。

2) 慢性肝炎データセットのデータ前処理とルール発見

本研究領域の共通データである慢性肝炎データセットを分析し、離散値に基づくルール発見を試みた。まず、出現頻度に基づいて、957 種類の検査項目を 41 種類に絞り込んだ。次に、検査周期については、患者に依存して異なっているが、これも出現頻度に基づいて 28 日周期に統一し、空値になるデータ項目については、線形補完によりその値を推定した。また、Das の手法に基づき、時系列データのある一定長で切り出し、EM アルゴリズムに基づき、切り出されたサブシーケンス群のクラスタリングを行った。以上の準備の下に、決定木学習により、検査項目値から GPT の 1 年間の変化を予測するルールの学習を試みた。その結果を専門医に評価してもらった所、ルールによっては理解できないといったコメントもあったが、乳びと GPT の関連性、GPT の周期性の予測など、興味深いというコメントもいくつかもらえた。

3) 多粒度の問題解決メソッドの検討

「慢性肝炎データセットのデータ前処理とルール発見」において、一次の生データからマイニングへの入力データへの変換過程には、いくつもの人手による作業が含まれており、その作業プロセスを自動化することも、アプリケーション開発現場では、大いに有益である。そのような解決に向けて、本テーマでは、要求レベル - 問題解決レベル - 実装レベルと対応づけた 3 層のメソッドリポジトリを検討し、ビジネスドメインにおいて、プロトタイプを試作し、その有用性を確認した。

#### 4) 領域オントロジー構築支援ツール

ユーザにとって重要・有用な知識を発見するには、マイニングだけでなく、モデリング的な機能も必要になってくるものと予想している。そのため、ドメインモデルを構成する概念要素の仕様を与えるドメインオントロジーを構築するツールの開発を手がけ、計算機可読型辞書とテキストコーパスから、ツールの試作を行った。

#### 今後の研究計画

上記の成果を基礎にして、まず、並列マシンのCPU数を6から16に増強し、あわせてメインメモリとHDDも増強し、性能向上を計る。次に、データ前処理では、属性選択法に着目し、フィルターメソッドやラッパーメソッドと比較して、計算時間は中間に位置し、性能は上位となるようなシーズメソッドを新たに提案し評価するとともに、問題解決メソッドの観点から、前処理メソッド群を体系化する。また、前処理と後処理には、やはり専門知識からの評価を支援する機能が不可欠であるため、肝炎におけるオントロジーの構築について、昨年までに開発した領域オントロジー構築支援ツールの適用可能性について検討する。最後に、慢性肝炎データセットに対して、5年以上の中長期的な変化パターンに注目して、専門家の興味ある知識を探る。

#### 3. 例外性発見に基づくスパイラル的アクティブマイニング

例外性は、大多数の傾向とは異なる性質を表し、それ自体が興味深いことに加えて新しい知識発見の糸口となることが多く、データマイニングにおいて有用知識の発見に特につながりやすい発見知識候補として重視されてきた。われわれが提案した従来手法は、発見された知識の興味深さと発見過程の効率性の両面において成功を収めた。本研究課題では、この手法をアクティブマイニング手法に発展させ、データ、知識、および環境を考慮して例外性を連鎖的に発見するスパイラル的例外性発見手法を構築し、医学・商業データなどに適用してその有効性を実証する。

#### 現在までの研究成果

##### 1) スパイラル的例外性発見

既に発見された例外性をデータ、知識、および環境に照らしあわせ、新しい例外知識を発見する基盤システムを構築した。このシステムは、データマイニングの標準問題として提供されている髄膜炎データに適用済みである。スパイラル的発見では、発見過程に応じて種々の学習・発見方法を用いるため、データマイニングにおける重要課題について、有効な学習・発見方法を開発した。具体的には、医療など種々の分野において重視される時系列データに関する新しいデータスカッシング手法、トランザクションデータにおいて重要である集合値属性を許容する決定木学習法、およびクラスが多次元である場合に有効な分類モデルである花つき決定木などである。実際のデータ解析では十分な前処理が必要である上に目標が一意には決められない不確定性が問題となるため、専門家の意見にしたがい重要で現実的と考えられる解析から進めてい

る。薬局の売上に関する商業データに関して、粗利が大きい希少商品を購入することが期待される例外的な顧客を特定する新しい手法を開発し、良好な結果を得た。慢性肝炎データに関しては、異常検査値の類形を発見し、これらの類形の推移に基づき患者をクラスタリングした。さらに、発見問題の設定に関しても医師との共同作業を通して堅実に貢献している。

## 2) 状況変化の検知手法

統一的な規準を確立するために、ルール発見の新しいサンプリング理論を構築した。この理論は従来研究と異なり、ルール発見に重要である複数の指標に関する誤差が指定値以下である場合を、指定した確率以上とするために必要な例数を与える。提案する PAGA (Probably Approximately General and Accurate) 発見は、分類学習において標準的な PAC (Probably Approximately Correct) 学習の自然な拡張ともなっている。さらにサンプリング理論に基づき、状況変化の検出法を構築した。この手法は、期間に分けられたデータからルールを発見し、ルール間の例外性に基づき状況変化を検知する。

## 3) スパイラル発見の自動調節機構

パターン抽出であるブースティングの結果をフィードバックすることにより、データの前処理であるデータスカッシングのパラメータを自動調節する手法を開発した。人工データと実データを用いた実験により、提案手法はデータスカッシングを 1 回だけ行うブースティングよりも正答率が大幅に高く、データスカッシングを行わないブースティングよりも計算時間が大幅に短いのに正答率がほぼ同程度であることが分かった。なお、データスカッシングとブースティングによる異常値発見手法も開発した。

## 4) スパイラル的特異ルール発見手法

対象データを変更し、マルチデータソースから興味深い特異データを抽出することにより、特異ルールを発見する手法を開発した。抗原抗体反応に関するアミノ酸配列データや慢性肝炎データなどに適用し、提案法の有用性を示した。また、特異ルールを相関ルールや例外ルールと形式的に比較・分析し、特異指向マイニングの理論的根拠を確立した。

## 今後の研究計画

主に慢性肝炎データの解析を進め、スパイラル的例外性発見の評価を進める。この評価は、プロジェクト共通データの解析に貢献するだけでなく、他の研究班との連携にも寄与すると予想される。マイニングプロセスのメタレベルの制御メカニズム及び複数のエージェントによるマルチデータソースからのマイニング手法を開発中である。さらに、実験用の複数分野のデータ (fMRI と脳波、ウェブサービスに関するログなど) を収集・データベース化する作業を行っている。

#### 4. 利用者からの要求を考慮したテキストデータからの知識抽出

本研究課題では、医学/生物学分野の論文に含まれる様々な知識を、各自の要求に応じて抽出することを目指している。そのために、技術論文中の自然言語文の自動解析、および、専門用語の識別と意味クラス分けを重要なサブタスクとして設定する。技術論文の解析のためには未知語に対する柔軟な処理が必要である。また、構文の解析においても技術論文特有の表現に対応するため、少数の解析事例から構文解析に関する知識を自動的に学習するシステムが有効であると考えられる。

現在までの研究成果

##### 1) 未知語を含む英文文書内の単語の品詞推定

専門分野の文書には一般の辞書には含まれない語が多く出現し、品詞等の文法情報の特定に支障をきたす。専門性の高い分野には次々に新しい用語が出現するため、それらをすべて辞書に登録することは現実的に不可能である。そこで、前後の文脈、あるいは、単語の綴り(特に接頭、接尾表現)を手がかりとして、未知の単語の品詞を決定し、それによって、専門用語と考えられる名詞句の同定を柔軟に行うことを試み、Support Vector Machine を用いた未知語の品詞推定および文全体への品詞タグ付けを行う手法を開発した。問題が非常に巨大であるため、SVMの単純な適用では実質的な計算時間で動作させることが困難であったため、従来確率に基づく統計モデルの誤りの訂正に特化したSVMを学習することを提案し、従来手法に比べて大きく効率を落とすことなく、より高い解析性能を示すことができた。

##### 2) 文中の基本句の自動抽出および係り受け解析

品詞付与が行われた文書に対して、そこに現れる名詞句や動詞句などの基本句を精度よく同定し、それらの間の文法的な係り受け関係を解析することは、その後の言語処理の性能に大きく影響を与える。特に名詞句と動詞句の間の係り受けのような基本的な関係の同定は、専門用語の抽出にとって必須の処理である。英語の基本句の同定を各単語へのラベル付与問題として見直し、SVMを用いた学習システムを複数混合することにより、従来法を上回る精度での基本句抽出を達成した。また、基本句の間の係り受け関係の同定にもSVMによる学習システムを開発した。

##### 3) テキストからの専門用語の抽出と分類

医学/生物学関係の論文概要を提供している Medline アブストラクトを題材とし、テキスト中に現れる名詞句を対象にして、その意味クラスを推定するタスクに関する実験を行った。既存の専門用語辞書の意味分類とそれに含まれる用語を学習データとし、そこには現れない未知の用語の意味クラスを自動的に推定する実験を行った。MeSH という既存の専門用語辞書の上位5種類の意味クラス(病気、治療法、薬品など)を対象に、未知の名詞句に対し、それがどのクラスに属する用語であるか自動推定可能かどうかを確認することを目的とした。用語の内部情報の情報と前後の文脈情報によっ

てどの程度用語の意味クラス推定が可能かを確認した。その結果、両者の情報を見ることで、約70%から88%の精度で用語の意味クラスを推定することができることがわかった。

#### 今後の研究計画

今回用いた未知語処理を伴う品詞推定、および、基本句へのまとめ上げプログラムは、現在発表されているシステムの中では最も高い精度を示しており、現時点では充分優れたものであると言える。しかし、それぞれのシステムは、現在入手可能な解析付きコーパスである Penn Treebank という主に新聞記事よりなるデータから学習したものであり、今回利用した医学生物学分野の論文とは、内容が大きく異なる。今後、医学分野の解析付きコーパスを蓄積することによって、より精度の高い解析を行える可能性がある。また、現在開発している英語の構文解析システムは文法的な木構造を対象にしているが、これを直接単語あるいは句の間の係り受けを学習できるように改造している。これにより、より高精度な言語処理システムを用いて専門用語推定の実験を再試行することを考えている。また、論文概要における各文の役割(前提, 目的, 手法の説明, 結果の報告など)を自動的に推定し、概要の意味的内容により踏み込んだ解析を行うことを計画している。

#### ( A03 班 ) アクティブユーザリアクション

研究項目 A03 では、具体的な問題領域(医療, 化学薬品)を対象にデータマイニングシステムを構築し、アクティブマイニングの結果得られた知識を、適用領域のユーザにとって有用なものとするための仕組みを具体化することを目標としている。その実現のために、発掘された知識の表示法、評価手法(有用性, 新規性, 意外性など)、ユーザからの効果的なフィードバックの手法など、ユーザのアクティブなリアクションを容易にし、“科学的発見のらせんモデル”を実現するユーザのアクティブな新データ収集、設定目的変更などを容易にする一般的な枠組みの構築に取り組んでいる。具体的には、診療情報生成システムの開発、化学物質群からのリスク分子発見およびヒューマン・システム・インタラクションに基づく知識評価と選択という3つの研究課題に焦点を当て、研究を進めている。以下、各研究課題に関して、現在までの研究成果概要と今後の研究計画についてまとめる。

##### 1. ラフ集合に基づくアクティブマイニングによる診療情報生成システムの開発

本研究課題は、ラフ集合に基づくあいまいな知識の取り扱い方法を利用、専門医のあいまいな知識を定量的に取り扱えるようにして、専門家の知識との総合作用から知識の発掘の促進および医療現場に有用な診療情報の生成を目標としている。現在、共通データが時系列データであることに着目し、時系列データからの診療情報生成に関する基礎的技術の開発を行っている。

## 現在までの研究成果

## 1) 多重スケールマッチングによる時系列データのセグメント化

病院情報システムに蓄積された膨大な時系列データでは、同一患者の検査データを数年から数十年の長期にわたり継続的に収集した時系列検査データが利用可能になりつつある。このような時系列検査データは、数日を単位とする短期間の推移のみならず、年単位の長期にわたる検査値推移パターンと疾患との対応関係を示すものであるため、その解析により慢性疾患を誘発する要因の特定、あるいは発病時期の予測等が可能になると期待される。しかしながら、これらのデータは当初から解析を目的に収集されたものではなく、不均質なものであるため、解析が困難である。本研究では、データの平滑化とともにセグメント化し、セグメント間でのマッチングを容易にする方法として多重スケールマッチングを用いた。

## 2) ラフクラスタリングによる系列データの分類

対象間の類似度が原点を持たない相対的類似度として与えられる場合、クラスタ内の分散、重心等を定義することが困難で、クラスタとしてのまとまりを評価することは容易ではない。ラフクラスタリングは、ラフ集合論の識別不能性の概念に基づくクラスタリング法であり、対象のまとまり具合を識別不能度として表現することで、相対的類似度で表現されたデータにおいても可読性の高いクラスタを生成することができる。本研究では、ラフクラスタリングを用いて、多重スケールマッチングでセグメント化した系列を分類する方法を、共通データを含めた病院情報システムのデータの一部に適用し、良好な結果が得られた。

## 3) 共通データの解析

これまでの肝炎データ解析では、測定回数の多い検査値の内、GPT,ZTTに関して次のようなことがわかった：(1)GPTの時系列的変化がC型肝炎のインターフェロンによる投与効果を最も反映している。(2)GPTの変化の中で、GPTの変動が投与後沈静化しているもの、沈静化せず乱高下しているもの、など、主に9個のクラスタに時系列データが分類された。(3)血液中の免疫グロブリンの総体的変化をとらえるといわれているZTTに関して、インターフェロン投与後にZTTが上昇し、1ヶ月後に正常化するというパターンによるクラスタが存在し、このクラスタを含めて、時系列データが主に7個のクラスタに分類された。他の検査値に関して、同様の解析を行ったが、GPT,ZTTほど顕著にインターフェロンの投与効果を反映した時系列変化は見られていない。



## 今後の研究計画

これらの成果に基づき，1) 共通データにおける他の検査値の時系列データの解析，2) 多重スケールマッチングおよびラフクラスタリングの性能評価と改良，3) 他の医療データへの適用と精度評価を行いつつ，時系列データにおけるユーザリアクションを実現するための新たな手法を開発する．

## 2. アクティブマイニングによる化学物質群からのリスク分子発見

本研究課題は，薬品類とその生理活性のデータベースから，1) それぞれの生理活性に対して特徴的な部分構造や物理化学的性質を見出して知識ベースを作成し，2) 新規の化学物質群中に，その生理活性をこれまでの知識ベースから予測できない例外的な分子があらわれた場合に，類似構造検索により既存物質に対する未知の生理活性を予測して，危険を回避することを目的としている．これまでの研究により，要素技術の開発と有効性の検証をほぼ終えると共に，変異原性と発ガン性およびドーパミン活性について有意義な結果を得ている．

## 現在までの研究成果

### 1) 生理活性に特徴的な分子構造のマイニングと知識ベースの構築

芳香族ニトロ化合物の変異原性と多様な化合物群からの発がん性の解析と予測に，カスケードモデルを適用してマイニングを行った．その結果，「芳香族ニトロ化合物で ortho 位置換基の立体障害が変異原性に重要な因子となる」，「有機塩素化合物でその発ガン性と水素結合受容体の有無および分子の柔軟性との間に高い相関がある」というような化学者が容易に解釈できる有意義な知見を多数得ることができた．なお，発ガン性の解析と予測を対象とする国際ワークショップ Predictive Toxicology Challenge では，モデルの理解容易性と female rat に対する予測精度において，参加 14 グループ中第 1 位の評価を受けることができた．

### 2) データスケープ探索を可能とするルール表現の開発

化学者に理解可能な構造と活性の相関を見出して知識ベース化するためには，専門家自身が導出されたルール群の助けを借りて，データの特徴を多面的に捉えることが不可欠である．この過程をデータスケープの探索と名付け，それを可能とするために，(1) ルール群の主ルールと関連ルール群への組織化，(2) ルールの識別力が大きく低下するリッジ領域の検出，の 2 点についてソフトウェアの開発を行った．これにより，マイニング結果からデータの特徴を素早く把握して，有効な知識に転化することが可能となった．

### 3) 化学構造類似性にもとづくリスクレポート

リスクレポートにおける構造類似性を基礎とした事例ベース推論の有効性を検証するため，3600 化合物を対象に生成フラグメントを質量数で特徴付けた Topological Fragment Spectra (TFS) データベースを作成し，TFS による類似性検索からの活性分子の同定

を試みた。分子グラフの類似性にもとづく事例ベース推論では、化学者が見て納得できる類似構造の分子を選択することができた。さらに、ドーパミンを query とした類似構造検索では 20 分子中で 3 分子が活性分子であることが明らかとなり、リスクレポート問題にも有効であることが示された。さらに 4 種のドーパミン受容体ごとの活性について、MDDR データベース中の治験薬からその受容体を予測したところ、その予測精度は 79% に上り、本方法がこれまでにない有効性を示すことを明らかにできた。また、化合物データベースを TFS 仮想空間上に配置し、化合物の直接探索が可能なデータ可視化ツール MolSpace を開発した。これにより、例外分子のリスクレポートからの専門家による考察の支援が可能となった。

#### 4) タンパク質三次元モチーフ辞書の作成

種々の生体機能発現に重要なタンパク質のモチーフ構造に注目し、PROSITE に登録されている配列モチーフに対応する三次元パターンを集積・分類した三次元モチーフ辞書の作成を行なった。カルシウム結合に関連する EF-hand モチーフの例では三次元パターン分類の結果、ノイズ成分を除去した主要クラスが生成され、その代表幾何パターンを得ることができた。さらに、この代表パターンを質問構造として、三次元構造情報をキーとしたデータベース検索を試みたところ、配列情報をキーとしたこれまでの検索ではヒットしなかったモチーフ部位を検出することができた。

#### 今後の研究計画

これらの成果に基づき、既存薬品の全て(およそ 12 万種)を網羅するデータベース MDDR と毒性のデータベース NCI(およそ 21 万種)を使用して、1) 各種生理活性を対象とする解析と知識ベースの構築とリスクレポートの作成、2) マイニングと類似構造検索および構造特徴解析のための各種ソフトウェアの洗練、3) 最新の PDB 全データセットを対象とする三次元モチーフ辞書の構築と辞書としての精選等の項目について研究を遂行する。また、ワークベンチの構築とオープンラボへの研究者受け入れによる研究も推進する。

#### 3. ヒューマン・システム・インタラクションに基づく知識評価と選択

「知識が有用である」とは、利用者にとって理解ができれば、現場で発生するニーズに応じた的確に使用できること、また、利用者の創造性を刺激しうる機能を備えることを意味する。そのために、本研究課題では、知識の需給関係に注目し、知識の候補を供給するシステムとそれを解釈・選択・利用する専門家とのインタラクションを通じて、知識を評価・選択できるような方式を確立することが目標である。われわれは、利用者個人あるいはグループの主観までを評価尺度に含み、従来研究されてきた客観的な基準での知識評価方法を超える到達点をめざす。特に医学においては、人体を複雑系として捕らえた病態モデルとしての理解が必要である。なかでも特に (1) 既存の医学知識を再評価して複雑な知識構造を確立して得られる、確実な医療を実現する客観的・一般的な知識と (2) 医学の飛躍的進展に寄与する意外性の高い知識の両面の要求がある。特に後者では、知識創造の初期においては、発見

の引き金になる重要な事象への主観的な「気づき」などが重要となるので、人と、(1)の役割をつかさどるデータマイニングシステムが相互作用するためのインタフェースとしての視覚化を研究してゆく。

現時点までの研究成果

1) (1-a) 客観的意志決定の支援: 病態モデルの構築

血液中の複数の酵素データを用いて病態モデルを臨床検査の実データから因子分析を用いて構築し、因子スコアの視覚的表現を試みた。視覚化によって、臨床家の検査データ評価に関わる負担を軽減し、さらに、重要な病的異常値の見逃しを防止する効果が得られた。

2) (1-b) 客観的意志決定の支援: 病態パターンのクラスタリング

血清蛋白分画データの病態パターンのクラスタリングを自己組織化マップを用いて行い、得られたクラスターを病態と対応付けるために、分類クラスと他の臨床検査データを属性として決定木分析を行った。意味付けされた分類クラスには、臨床診断につながる付加価値情報として有用な情報が得られた。

3) (1-c) プロセス時系列のアクティブマイニング

化学プラントの時系列データから、例外性を考慮した制御知識を発見する手法を開発した。これには、データからのプロセスモデル構築、専門家によるモデル評価、機械学習による知識発見のプロセスが含まれ、アクティブマイニングの効果が明確である。手法としては、相互相関分析、ニューラルネット、ならびに学習分類子システムを採用した。学術的な貢献としては、学習分類子システムにおける MDL 規準を適用したルール獲得方法の提案が挙げられる、応用面における貢献としては本手法により、実プラントデータから単純かつ有用なプラント制御知識を獲得したことが挙げられる。

4) (1-d) 肝炎データからの因果モデルの構築

アクティブユーザーリアクションの医療における具体的事例として、肝機能検査データの因果モデルの構築を試みた。そのため、データ自体に内在する知識をより多く発見・獲得するためにメタ学習の枠組みを導入した。そして、メタ学習の結果とユーザの背景知識を基に因果関係を考察し、構造方程式モデリング [1]にて因果モデルを構築した。その結果 EBM の概念に合致するような簡潔かつ汎用性のあるモデルが構築できた。

5) (2-a) 主観的発見: チャンス発見プロセスの実現

チャンスを評価するメタな指標として、P: 行動の提案可能性, U: 気付かれにくさ, G: 成長性を提案した。P と G を具体的に把握するために、一人ではなく数人のグループを構成し、グループディスカッション (GI) における人と人との間でのチャンスに基づ

く提案や採択を通してチャンスを選択してゆくプロセスを実現した。さらにこのプロセスを、意思決定環境についてのデータ（環境データ：販売者における購買データなど）に適用し、視覚的データマイニングによって刺激してチャンスを発見する仮定を促進する結果を得た。

6) (2-b) 主観的発見: チャンス発見二重らせんモデルの実現

新たなチャンスを発見するモデルとして、二重らせんモデルを提案し、これを実現するためのヒューマン・マシンインターアクションを促進するシステムを構築した。具体的には、グループインタビュー（GI）における議論の各参加者の考えを、カードに思い当たるだけ書きこんでもらい、そのテキスト内の単語の相関関係を図示する視覚的なテキストマイニング（KeyGraph）の出力図によって議論におけるチャンス表出化を促進するシステムを構築した。このシステムで顧客像を把握した結果、顧客の購買（POS）データからスーパーマーケットの購入金額増加の鍵となる商品や、その店の経営状態のおおまかな変化を示す予兆を発見することができた。さらに、社会調査を行った社会学者の「関心」を起点として二重螺旋モデルを実行した結果、社会学的に新規性と説明能力の高い仮説が獲得された。この研究を更に進めたアンケート調査によって、人のチャンス発見プロセスそのものをモデル化する研究を行った。この結果、チャンス発見プロセスは確かに上記の二重螺旋モデルのうちで人の螺旋を描いて進んでいることが明らかになった他、モラルやリスク認知といった要素がチャンス発見のプロセスが一段深化するために本質的な役割を果たすことがわかった。これは、この研究の当初からの目標の線に沿った中心的進行といえる、すなわち、

**Interactive Data Mining/Interactive Mining Validation 手法** チャンス発見の場では、何よりもまず主観的な意思決定知識、経験の中に潜んでいるチャンス獲得感性、そして研究あるいは生活において準備されたチャンス志向性を備えた発見主体であるユーザが主人公である。このユーザが自分の内在的な関心をもとにデータを収集し、その内容の構造を視覚化した結果に刺激されて発見を進める。これは、機械による客観的マイニング算法の中から価値ある知識が生まれることに相当する。二重螺旋とは、まさにこのプロセスを実証的にモデル化したものである。

**Data Visualizer の実現** マーケティングあるいは医療などの現場にいる専門家に対して、上記のデータ視覚化ツールを用いて発見を支援する。ここでは、視覚化ツールそのものよりも、それによって支援されるわれわれと医師らの協力体制による研究プロセスそのものが本質的な研究テーマの一つとなる。すなわち、研究コミュニティで価値が発見され伝播し、洗練されることにより、新しくかつ普遍的な価値が生まれるプロセスに関する認知科学的研究ということが出来る。

この手法を医療データに適用し、肝炎の専門家との対話を二重螺旋モデルによって進

める研究を現在進めている。この他、マーケティング分野においてはビジネス現場における成果以外にも、マーケティング学専門の研究者との共同研究において実質的な成果を出している。

#### 今後の研究目標

チャンス発見においては、さらに積極的に人やそのグループの判断をシステムに取り入れ、それにより発見される外界の未知要因の重要性を計る指標とその発見に至るプロセスの支援方法を確立するための研究を進めていく予定である。肝炎データベースの扱いについては、複数の医療機関からのデータを援用して、EBMの観点から有用な医療診断モデルの構築をはかる。これには、計算機科学者と医療専門家の積極的な協力体制が不可欠である。今後の課題は、螺旋モデルにおけるこのような社会的制約を解決することが可能な、すなわち、創造的コミュニケーションを円滑化するような視覚的インタフェースを実現づけることが主眼となる。

## 付録 2：総括班および外部有識者による評価

### (1) 研究目的が妥当か

コメント 1： パソコンを所有しているものは、数十ギガバイトのディスクを個人的なストレージとして利用でき、インターネットに接続すれば、無限ともいえるデータにアクセスが可能である。そして電子的に利用・処理できるもしくはしなければならない情報の量は、個人のレベルでも膨大となっている。この現象は、いまや普通のことであり、大型計算機センターなどに特化されたものではなくなった「情報洪水」という言葉でこの状況を表そうとしているが、洪水ならばいずれ引いていくが、この現象は留まるところを知らない。こうした中で、情報空間のどこを見て、そこから目的に合った知識を取り出し、また知識を更新することは、非常に困難な状況の中にある。実際、ゲノムに関する情報（核酸配列、タンパク質データ、マイクロアレイデータ、パスウェイデータなどのデータベース、MEDLINEなどの文献情報）は、よく整理されているほうであるが、ヒューマン・エキスパートなしではどうしようもない状況である。この研究はこうした状況を打開するために「アクティブマイニング技術」を実現することを目的としており、妥当な研究目的といえる。

コメント 2： 情報化社会における大量情報の氾濫において、組織や個人が迅速な意思決定と効率的な活動を行なうために、これらの大量データを有効に利用するための技術が必要とされている。現在、データマイニングがその鍵となる技術として期待されているが、現在のデータマイニングは個々の要素技術の研究にとどまっており、社会に広く利用される成熟した技術には至っていない。本研究プロジェクトでは、データマイニングを一貫した情報処理プロセスとして研究し、この状況を解決することを目指している。とくに、データマイニング分野では、現在、(i) 多様な情報源への対応、(ii) データの前処理とパターンの後処理の統合、(iii) 自然科学データやインターネットデータに特有な構造データへの適応が、最重点課題であると認識され、欧米を中心に学術と産業の両面で活発に研究が進んでいる。これらの問題の解決は、データマイニングが実際の情報システムに統合されることで、個人と組織が主体的に大量情報の氾濫に対処するための基盤技術となるためにきわめて重要な技術であり、現在、緊急に研究を進めるべきである。本研究プロジェクトは、このような社会的に緊急性の高い問題に、重点的に取り組むものであり、特定領域研究として進めることに意義がある。また、現在わが国は、先行して特定領域研究をもつなど、この分野の研究における競争力をもっており、集中して研究を推進することで学問的な成果が期待できる。

コメント 3： 昨今のインターネットの急速な普及はネットワークを介した大量データの洪水を誘起し、必要なデータを選択し、正しい意思決定をすることが困難になりつつある。社会全体における知識処理の重要性が増してくる状況下で、本特定領域研究は、アクティブマイニングという考えを導入し、情報収集、マイニング、評価・洗練（ユーザリアクション）のサイクルを高度化する方法論の構築を目指すもので、研究目的は時代の要請にかなったものである。

コメント 4: 知識情報処理の分野は、大量データの中からの知識の発見抽出をどこま自動的に  
行えるかに関心が集り、多くの研究が行われて来た。信頼性の高い知識を出来るだけ多く  
確実に取り出すためにはアクティブマイニングという考え方を導入して、より突っ込んだ研  
究をすべきであり、本研究プロジェクトは時宜にかなっている、

コメント 5: 評者は情報技術の専門家ではないが、インターネットの発展や実験室へのロ  
ボット技術の導入を契機として、研究者個人および組織の生み出す研究情報の生産性が飛躍  
的に伸びていることを実感している。評者の専門分野では、オンラインデータベースの重要  
性は ほぼ 30 年前に確立されたのに比して、知識レベルの創出、整理、蓄積、検索に関し  
ては、未だ期待したほどの進展はない。このような状況にあって、特定領域研究「アクティ  
ブマイニング」では、計算機処理が一人歩きするのではなく、あらゆるステップで領域専門  
家との相互作用を重視している。このようなアプローチは評者の期待するところであり、本  
特定研究の目的は妥当なものと言える。

コメント 6: アクティブマイニングという新しいデータマイニング手法を模索しながら医  
療データについてデータマイニングを行うということは、大量なデータが死蔵されている  
といっても過言ではない医療界において、タイムリーな着眼点だと思われる。電子カルテと  
いった IT 化が推進されて、今後医療データはさらに膨大なものとなると思われる。このよ  
うなときにその膨大なデータを解析する手法は従来からのものでは限界があり、データマイ  
ニングは新しい手法として大いに期待ができる。

コメント 7: IT 技術の飛躍的な発展は、ネットワークによる情報の大衆化と情報洪水をも  
たらしているが、これは医療分野においても例外ではない。アクティブマイニング技術の実  
現は、情報洪水に伴う問題の解決を目的としており、研究目的は十分に意義あるものである。

コメント 8, 9: このような目的を明確にしたプロジェクトを特定領域研究として重点的に  
進めるのは大賛成である。

コメント 10: 前処理も含めたトータルなマイニングプロセスを目標にしてることに意義が  
ある。

## (2) 研究方法は妥当か

コメント 1: 研究項目を、A01:アクティブ情報収集、A02:ユーザ指向アクティブマイニン  
グ、A03:アクティブユーザリアクションの 3 つに体系的に分けている。アクティブ情報収  
集 (A01) では、不特定・非定常・大規模・分散知識源の中から、ユーザの目的や興味に合致  
するデータやそれらの関連を効率良く探索し前処理するための情報収集技術を、メタ情報源  
の活用、ヒューリスティック探索知識の活用、機械学習法の活用など、最新の情報処理技術  
を駆使して開発するという研究方法をとっている。これは、非常に妥当なアプローチであり、  
これまでに情報科学で培われた技術を発展させる形となっている。ユーザ指向アクティブ  
マイニング (A02) では、多様な形式や多種の情報源に対応でき、汎用性と状況の変化にも対  
応できる柔軟性を持つマイニング手法を開発することになっている。やや、大風呂敷的に見

えるが、具体的には、データとして、テキスト情報、分子化学情報やネットワーク情報などからのマイニング、及びこれら個別のデータに最適なマイニング手法の自動構築、状況変化検知に強力な例外性の発見技術に力を注ぐ形となっており、フォーカスを絞った研究方法がとられているため、問題はない。どこまで、一般的なパラダイムを構築できるかが研究の達成度を見る時の尺度となるであろう。アクティブユーザリアクション (A03) では、具体的な問題領域を対象にマイニングシステムを構築し、発掘した知識を、ユーザにとって有用なものとするための仕組 (知識の表示法、評価手法、ユーザからの効果的なフィードバックの手法) を具体化するという研究方法をとっている。私のこれまでの経験から考えると、ドメインに特化した知識と、そのドメインの研究カルチャーなど、必要なアドバイスを受けられる環境が必須である。

コメント 2: 一貫した情報処理プロセスとしてのデータマイニング研究という重要かつ大きなテーマに関して、短期間で研究成果をあげるためには、現在課題となっている重要な課題に焦点をあてて、集中して研究を遂行する必要がある。本研究プロジェクトでは、現在、緊急かつ重要な研究課題となっている (i) 多様な情報源への対応、(ii) データの前処理とパターンの後処理の応用システムへの透過的な統合、(iii) 自然科学データやインターネットデータに特有な構造データへの適応に対応した (A01) アクティブ情報収集および、(A02) ユーザ指向アクティブマイニング、(A03) アクティブユーザリアクションの3つの研究項目を立てて、集中的に取り組んでおり、適切な構成である。従来おこなわれてきた要素技術の研究とことなり、データマイニングプロセス全体の研究、とくに (ii) の透過的な統合 (垂直的統合とも呼ばれる) の効果的な研究には、具体的な専門領域の設定と、さらに、専門家の参加、分担研究者間の協調が欠かせない。これらの点についても、共通医療データからの病因予測と化学物質のリスク発見という2つの具体的かつ社会的意義の高い領域を設定し、専門分野との研究者の参加、協同研究体制の整備など、十分な配慮がある点で、妥当であると判断する。

コメント 3: 従来のデータマイニングの研究が、マイニングが中心であったの対し、本特定領域研究では、マイニングの前処理である情報収集、後処理である評価・洗練 (ユーザリアクション) にも力を入れ、データマイニングの全工程をサポートすることを前提に、3つの班に分割し、各班はさらに情報収集が3つの、マイニングが4つの、ユーザリアクションが3つの、計10の計画研究から構成されている。これらの班構成、計画研究の構成は適切である。また、各班ごとに固有の技術を解決するための手法を並行して開発し、それらを、全ての班に共通の医療データ (肝炎データ) に適用し、個別の研究成果を同じ土俵で評価・検証しようとする研究方法は、全体の研究を連携させるのに非常に有効である。領域代表者は計画研究代表者会議を定期的に召集し、専門家 (医師) を交えた技術討論を頻繁に実施しており、また、総括班会議で全体の進捗状況をよく把握しプロジェクト全体を適切に統括している。プロジェクト全員の参画意識も高い。個別の研究はすでに国際的にも高い評価を得ている。肝炎データ解析や化合物 (薬品) データ解析でも、面白い結果が出はじめている。さらなる成果を期待する。



コメント 4: 研究テーマのたて方は妥当であるが、多岐にわたっており、これらの研究成果を有機的に統合して効果的なアクティブマイニングシステムが作れるかどうかは現時点では自明ではない。共通データを取り上げたことは、全員が同じ意識の下で専門家にとって有用な知識を発掘するための方法論と環境を作り上げるのに有効である。共通データ解析を基盤に、踏み込んだ検討を続行し、具体的に検証して欲しい。各研究テーマについての研究を進める方法として、既知の技法の組み合わせや新たな手法の開発に注力しているが、巨大データを扱うという観点からも、どのような新しい技法を創出しようとしているかを明瞭に示してもらいたい。すでに、例外知識の発見、構造データマイニング、マイニングアルゴリズムの自動構築、テキストマイニングなどで国際的に高い評価を得ているが、アクティブマイニングとして、他にはない独創性がどこにあるかも明らかにして頂きたい。

コメント 5: 情報収集、マイニングおよびユーザリアクションの3グループに分けて研究を進めている。評者が情報領域の専門家ではないため、援用する個別情報技術の妥当性についての評価は差し控えるが、全体として正当なアプローチであることは理解できる。なお、統計学を含めて情報処理の専門家が他の専門領域を対象として取り上げる研究プロジェクトの場合、常に研究のベクトルが拡散する危険性をはらんでいる。本特定領域研究は、その点で研究プロジェクト開始時に、肝炎データの解析と化合物生理活性の解析という2つの共通課題を設定して、各班が研究を進めている。また、課題内容を各研究チームへ徹底するための会合も頻繁に開かれ、各チームの参加意識も高いようである。今後の進展に期待できる。

コメント 6: 新しいデータマイニングの手法を作り上げようとして、様々な手法を持ち寄り、総合的な研究成果をあげることを目指している。大きな目標であるので各グループの役割を明確にし、オーガナイズする必要性を感じる。現段階では、前処理・マイニング・Web や文献からの情報収集など様々なセクションにおいて、医療データの利用という共通ミッションを挙げて、研究目的を明確化・具体化しているの、方向性がまとまっていると評価できる。

コメント 7: 本プロジェクトは3つの研究計画班が共通データである肝炎データを用い、知識発見を協力して行おうとするものである。とくに、肝炎データの中から以下のような知見を得ようとしている。1) 病理像と血液検査データとの相関性、2) 肝炎の病理像と発ガンまでの期間、3) 血液データと発ガンまでの期間、4) 時系列に関する血液データ積算の有用性、5) B型肝炎とC型肝炎の経過の違い、6) INF 療法の有用性、7) GOT, GPTが「進行速度」の指標となっているが、実際にどうか(速度×時間=距離、肝炎の進行度が成り立つか)。以上の事項はいずれも大量の医療データをもとにした解析が十分に行われていなかった分野であり、アクティブマイニング手法により新たな知見が得られることは大いに期待させられる。

コメント 8: 医療データと化学データでは性質が大きく違う。医療データはノイズが多すぎ、非常に困難であろう。小人数で取り組んでいる化学データの方もおろそかにすべきではない。成功すればインパクトの大きな成果となる。こちらにも力を入れて欲しい。

コメント 9: マイニングプロセスを分解して各々に集中するのは良い。同時に、各技術を組み合わせた全体をレビューすることも重要である。総括班会議や計画研究代表者会議がその役割を果たす場所であろう。

コメント 10: 具体的な問題に深く取り組むと、その領域にとってよい成果が出る反面、自動化の範囲も問題に依存したものになるなど、非常に限定された領域にしか応用できないものになるという恐れがある。汎用性と領域固有性をどうバランスさせ進めるかが重要である。前処理に関してはヨーロッパの MiningMart プロジェクトが進行中である。外国のプロジェクトの連携は考えられないか？オーバラップしている部分も多いが、目的が違うので難しいとは思いますが、検討する価値はある。

### (3) 研究組織は適切か

コメント 1: 研究目的と研究方法において、日本におけるほぼ最適の人材を集めた研究組織となっている。この研究組織を最もアクティブに活用するために、領域代表者の果たす役目は重要である。総括班会議、計画代表者会議は十分機能しているようであるが、さらなる努力が必要であろう。

コメント 2: 研究目的と方法に関して、現在わが国における最適な人材を集めた機動的な研究組織であり、この分野で国際的に活躍する優秀な若手研究者を中心に、対象分野の専門家を配した組織構成であり、適切である。

コメント 3: プロジェクトとして領域代表者が統括できる範囲で、十分練られた組織である。参加者は我国を代表する若手の優秀な研究者を中心に集められており、これだけの集団が力を合わせれば世界に誇れる成果が期待できる。研究組織は適切である。領域代表者の目が全体に行き届いており、組織としての統一的な見通しに基づいて効率的な研究がなされている。

コメント 4: 30人程度の大きな研究グループではあるが、解決すべきサブテーマもよく検討しており、うまく分担して研究を進めていると判断する。

コメント 5: 若手研究者が多く参加する研究組織となっているため、対象領域への取り組みを領域代表者が主導しやすいと考えられる。また、A03のユーザリアクションでは、データマイニングの専門家でありながら、医学や化学の領域で適切な知識を持つ研究者が、計画研究の代表者となっている。この点では、対象領域固有の知識をプロジェクト全体に反映できる適切な組織といえる。

コメント 6: 医学・工学が互いを理解しながらともに positive に研究を進めている。実験的に進めるとなると、複数のマイニングをしなければ確率的に成功例がでてこないかもしれない。しかしデータを提供するグループと、それをマイニングするグループの比においてあまりに後者が大きいと前者とのやりとりがネックになるであろうし、現状はほぼ最良の関係ではないかと思われる。

コメント 7: 3つの班構成および総括班のプロジェクト管理による研究組織は適切である。

コメント 8: マイニングのすべてをカバーしているので、規模が大きくなっている。もう少し焦点を絞り、小さなグループで進めると、カバーする範囲はせまくなるが、研究がやりやすくなるのではないか？一般に、グループ間の連携は難しいものである。総括班の役割が重要である。

#### (4) 平成 13 年 10 月から平成 14 年 8 月までの研究活動

コメント 1: アクティブ情報収集 (A01) では、「アクティブ」という言葉からわかるように、単に検索によって情報を効率よく収集することを目指しているのではない。そのための研究として、WWW におけるメタ情報源の獲得、分散動的情報からのアクティブ情報収集、多段階学習方式によるデータ収集と前処理の自動化、という 3 つの課題に焦点をあて研究が展開されている。WWW におけるメタ情報の獲得については、関係学習による対話的文書検索の検討、WWW 上の情報収集/可視化のための免疫ネットワークを用いたクラスタリング手法の開発、Web ページの部分更新のモニタリングシステムの開発など、これまでの方法と比べて、質と量、そして効率とともに飛躍的に向上するような情報収集方式の構築に向けて概ね成果が出始めている。また、分散動的情報からのアクティブ情報収集では、知的情報収集アルゴリズムの開発により、「質の高い情報」を収集できることを目指した研究が行われており、MEDLINE データへの応用が計画されている。これは極めて重要な課題であるが、「質の高さ」という曖昧なものを相手にすることになるので、具体的な達成点を明確にして研究を進める必要がある。多段階学習方式によるデータ収集と前処理の自動化に向けては、前処理支援システムの構築、伝言ゲーム型の情報収集と前処理システムの開発とその実験、HTML のリンク構造と構文的特徴に基づく知識獲得システムの実装、検索処理性能向上に向けての高次元インデックス技術の開発など、成果が見えてきている。ユーザ指向アクティブマイニング (A02) は、構造データからのアクティブマイニング、メタ学習機構に基づくアクティブマイニング、例外性発見に基づくスパイラル的アクティブマイニング、利用者からの要求を考慮したテキストデータからの知識抽出など、従来からの先駆的なデータマイニング技術をさらに発展させる研究が展開されており、この分野での独創的な研究が期待される。大規模データという点にどれだけ配慮して研究を進めているかがあまり見えない感があるが、これについても十分な期待をしたい。アクティブユーザリアクション (A03) では、医療、化学薬品を対象にデータマイニングシステムを構築し、アクティブマイニングの結果得られた知識を、適用領域のユーザにとって有用なものとするための仕組みを具体化しようとしている。そして、他の班と連携を取ながら、“科学的発見のらせんモデル”を実現する一般的な枠組みの構築に取り組んでいる。具体的には、ラフ集合に基づくアクティブマイニングによる診療情報生成システムの開発、アクティブマイニングによる化学物質群からのリスク分子発見、ヒューマン・システム・インタラクションに基づく知識評価と選択の研究などの研究が行われている。特にヒューマン・システム・インタラクションに基づく知識評価と選択の研究では、客観的意志決定の支援、客観的意志決定の支援、主観的発見: チャンス発見プロセスの実現、主観的発見: チャンス発見二重らせんモデルの実現など、とてもチャレンジングな研究が展開されており、大きな成果が期待したい。

コメント 2: 論文活動, 受賞等からなど個別の研究成果について, 十分な成果がでている. とくに, (A01) の前処理自動化に関する, メタ情報源の更新検知や分散情報での情報共有, 自動化アルゴリズム, (A02) のユーザ指向化に関する, 高速な構造データマイニング手法の開発, 例外性発見手法, テキストデータからの知識抽出, (A03) のユーザリアクションに関する, ラフ集合手法による医学的知識発見の成功, 化学物質リスク予測の成功などは, 特筆すべき成果であろう. 意欲的なとりくみである 2 つの専門領域への適用に関しては, 専門家をまじえた勉強会や, データ整備の活動, 分担者相互の協同研究など活発に活動を行っており, 着実な進展をみせている. 注目すべき成果として, 当該分野の技術水準を競う国際的データマイニングコンテストにおいて, 本研究プロジェクトが 2 部門で 1 位を獲得している. また, 公開ソフトウェアや報道など, 成果の公開・周知も積極的におこなっている. 現在の活動を継続することで, 今後, 良い成果を出すことが十分に期待できる.

コメント 3: 論文などの通常の研究活動としての成果は十分出ている. まだ, 開始後一年たらずであるが, 肝炎データ解析と化学薬品解析も軌道に乗りつつある. 専門家とのインタラクションも活発に行われており, 前処理, マイニング, 評価のサイクルも着実に周りはじめている. 専門医の興味を引く具体的な成果も出はじめており, 今後の成果が十分期待できる.

コメント 4: この研究の開始前から非公式にグループを作って研究を進めていたこともあって, 研究開始後一年足らずではあるが, 着実な成果をあげており, 今後さらに具体的な良い成果を出すことが十分に期待できる.

コメント 5: 評者の専門領域に関連するマイニング結果について述べれば, 未だ予備的な段階ながら良い結果が出始めていると考えられる. これまでの数理科学的進展のみを重視した研究や, ビジュアル化に過度に頼った枠組みとは異なり, 利用者の理解を触発し, 利用者からのフィードバックを重視したアプローチである. 今後この方向での研究をさらに進展させ, 専門領域の学会へ大きな影響を与え得るような成果を期待したい. なお, 他の課題についても, 十分積極的な研究活動を進めているように見受けられる.

コメント 6: 各グループの研究は, 特徴があって, それぞれ興味深い. データの時系列推移を追ったデータマイニングは医療者からみてもわかりやすく, 有用な知識が得られる期待をもてた. また, 様々な形で形成されるルールは医療者からみてその可読性は必ずしもよくないものがあった. しかしそのルールが正しいのであれば是非, 実際の医療に役立てたいと考える. 1 年の経過の中でデータの前処理とドメイン知識の解説, 実際のマイニングが行なわれ, 何度かのフィードバックがなされてきたと聞いている. これらのトライアルにより今後さらにアクティブマイニングを重ねることで, いよいよ実際の医療に還元できるようなルールの発見が期待できる時期に移りつつあると思われる. 研究成果は研究分担者の医師の意見にもあったが, エキスパートシステムのナレッジベースとして近い将来に使えるであろうし, これを電子カルテに装備することで診断支援への応用が期待できよう.

コメント 7: 研究活動はおおむね順調に進んでいると思われる. 肝炎データマイニングと情報収集に関して医学的な知見を加えておく.

(a) 肝炎データマイニングに関して：A02-06 では「肝炎進行の度合いを示す血液データ (GPT) と検査データとの相関関係を発見し、検体検査データから予後因子を同定する。」ことを目的とした解析を実施している。データのクレンジングおよび時系列データからの属性構築について問題はないが、肝機能の推定としての指標に GPT が用いられている。GPT は、肝細胞に含まれる酵素であり、肝細胞が障害されると逸脱して血中濃度が上昇する。肝炎の活動性、すなわち GPT が高値を示している期間は肝炎ウイルス感染による肝細胞の障害が持続的に存在しており、GPT が正常であれば肝炎ウイルスは感染しているが細胞障害は生じていないことを示している。過去数年にわたる肝細胞障害のパターンと肝炎ウイルスの種類によって、将来の肝細胞障害すなわちウイルスの活動が予測できるならば興味ある結果となる。一方、実際の医療現場では、肝機能異常 (GPT, GOT の上昇) があれば肝庇護剤や安静などの治療が行われるために人為的な影響が多く介在することになる。再現性が確保されなければデータの確度の信頼性は低いと思われるので、この点も配慮されるべきであろう。肝細胞の再生能力は高く、肝炎ウイルスが除去され肝炎が治癒した患者では肝機能の回復がみられる。肝炎の進行度を検証するためには、肝機能の指標としては肝細胞障害の指標である GPT 以外に肝予備能の指標 (アルブミンやコリンエステラゼなど) を用いることも重要ではないかと思われる。膠質反応検査の一つである TTT は、肝炎において免疫グロブリンなどの炎症物質を間接的に評価するパラメータであるが、検査学的には古くからある項目であり、疾患特異性が低いために最近では意義は低いと考えられてきた。今回導出されたルールの中でこの TTT に関する知見が少なからず存在することは興味あることである。

(b) 情報収集に関して：データマイニングにより導出されるルールは多量であり、その中には無用なもの、有用なもの (既知のもの、新規のもの) が含まれる。このような多量のルールを利用者が逐一検証するには現実的に困難である。この作業の効率化にオンラインデータベースを利用した情報収集技術は必要であると考えられる。また、データベースは日々更新されているので最新情報にアクセスする必要がある。これらを実現する方法としてアクティブ情報収集システムの研究方法は妥当と考えられる。ただし、医療データからのルールに対するデータベースとしては、通常の web 情報では信頼性に問題があり、専門の医学データベース (たとえば Medline など) を対象とすることが必要であると考えられる。

コメント 8, 9, 10: 短期間でよく成果を出している。本特定領域研究のホームページの英語版を作成して、外国にも進捗状況が分かるようにして欲しい。

##### (5) 本プロジェクトへの期待

コメント 1: 本プロジェクトが取り組んでいる問題は、このプロジェクトだけで完結するものではなく、直面する情報が巨大化と多様化に応じて、さらなる展開を必要とすると思われる。本プロジェクトの成果が、次の発展のための基礎として位置付けられるようなものであってほしい。

コメント 2: 本研究プロジェクトは、情報化社会における大量情報の氾濫において、個人や組織が有用な情報を獲得するための新しい技術を目指していると考えられる。現在のデー

タマイニング技術もこの方向を目指したもののだが、現状では、データ解析の専門家のための技術となってしまうている。学術研究としての成果ももちろんであるが、本研究プロジェクトが、データマイニングが情報技術に意識されずに組み込まれ、一般の個人が真に有用な情報を主体的に活用できるような、成熟した基盤技術を生み出すことを期待する。

コメント 3: 肝炎データは 20 年にもおよぶ患者の生の病歴データであり、そのままでは解析は非常に難しいと予想する。実世界の生データを避けることなく専門家の協力を得て、真のアクティブマイニングを是非、実証してもらいたい。薬品データ解析によるリスク分子発見も非常に重要なテーマである。両者の解析には違った手法、専門知識を必要とする。本プロジェクトを通し、データマイニングに必要な主要技術を開発し、広く世に公開すると共に、医学、製薬の分野に対しても実質的な貢献が示せることを期待したい。

コメント 4: 共通データとして取り上げている肝炎データや化学薬品関連データを総力を上げて解析し、これらの分野の専門家の気がつかなかった事実、法則性などを発見し、この方法が専門家の信頼を得るに至る研究に発展すること期待したい。

コメント 5: 医療データ、薬品データの解析はともに非常に重要な課題であり、この研究プロジェクトはそれぞれの専門分野に大きく貢献できるような成果を出して欲しい。なお一般にある領域の専門家が情報領域の方法論等を学びながら、プロジェクトを主導する方が成果を出しやすいと考えられている。今回のプロジェクトでは、情報処理の専門家が主導しているが、このような場合に他領域の専門家を含めて、どのように組織化を進めるか、どのようにプロジェクトを運営すれば、実りある成果が得られるかのモデルを提示していただきたい。この特定領域研究は、そのような可能性を十分持った組織と運営を進めているように見受けられる。

コメント 6: 医学においては、今まで様々な分析手法を使用して、データを処理してきた。しかし、その多くは予め対象を設定し、分析の目標や仮説がはっきりした形で存在する、いわゆるデザインされた実験・調査が大半を占めてきた。今回の研究は複数の工学系のグループがそれぞれ今まで研究してきたマイニング手法を惜しげなく投入して、またその過程で得られたノウハウを共有するという、多施設が参加するプロジェクトならではの利点が活かされていると思う。医療側としては、医療関係者では思いも寄らないような全く新しい知識の発見があれば、と期待している。また、このような試みを機に、新しいデータ処理技術が医療に反映されることに大きな期待を持っている。現在、検査の分野では、不要な検査を減らすということが、一つの目標になっている。従来必要だと思われて頻回に測定されていたが、疾患の病態把握の上では他の検査項目でも同様の評価が可能であり、その意義が問われている検査項目がある。このような検査項目を洗い出しその有用性を評価することは医療経済上極めて重要である。医療界ではこの問題について、ROC カーブといったごく限られた評価方法しか浮かんでこないが、自由な発想で取り組み、より有効な手法を提案してほしい。本研究の成果が、限られた医療資源を有効に使うために大きな役割を担うことと期待している。医学の分野における研究は prospective study 全盛の時代である。これは実験科学からすれば

当然の方向ではあるが、実験動物ではなくヒトを対象とする臨床医学においてはできる限り retrospective study から知見が得られることが望ましい。Prospective study における仮説を創生するためにも、また prospective study では行なえないような長期間のデータ解析をするところに retrospective study の意味があろう。今後データマイニングはまさにこうした点において retrospective study における中心的役割を果たすのではないだろうか。

コメント 7: 肝炎データマイニングに関してはいくつかの知識が発見されつつある。ウイルス性肝炎において臨床家をもっとも注目するのは予後である。従来の方法で分析できなかった長期にわたる検査項目の変化トレンドが予後因子として扱うことができるならば重要な発見となるであろう。また、データマイニングに関連する文献を収集するプロジェクトにおいては、発見した知識に関する最新の知見 (evidence) を入手する契機となり、データマイニングと情報収集を統合する本プロジェクトの有用性を示していると考えられる。医療機関には膨大な臨床検査データが蓄積されており、最近では、医療事故に関するインシデント、アクシデントレポートなど従来の統計学的手法では限界のあるデータベースも存在する。このようなデータベースを有効利用する手段としてアクティブマイニングが果たす役割は大きいと考える。本プロジェクトで開発される手法が実用レベルになることを期待している。

コメント 8: プロジェクトの成果として、新しい見識が得られるということも重要だが、学生に対する教育効果も大きい。プロジェクトに参加し多くの経験を学んだ学生、とくに博士課程後期の学生を世に送り出すことの意義は重要である。

#### (6) その他アドバイス

コメント 1: このプロジェクトで目指している技術は、このプロジェクトで取り扱っている WEB 情報や医療情報ばかりでなく、様々の分野で求められているものである。したがって、このプロジェクトでの成果が、このプロジェクトに直接にかかわらなかった分野にも影響を与えることができるようなことができればと思う。たとえば、人工知能学会ばかりでなく、その他の学会誌などに、このプロジェクトの研究成果などをわかりやすい形で紹介することはできないだろうか。

コメント 3: 開発した技術はソフトウェアとしてまとめ、広く公開してほしい。本特定領域研究では医療、薬品データの 2 つを主に扱っているが、多分野への応用可能性も常に念頭において研究を進めて欲しい。

コメント 4: 研究の後半には、少なくとも幾つかのサブテーマは統合してシステム的な研究を得ることを期待したい。そのためには研究分担者間の情報交換、意見交換をさらに密にしながら研究を進めて欲しい。

コメント 5: 科学には統一性を求める流れ以外に、常に例外を求める流れがある。例外性、類似性をソフトウェアに実装して判断するのは結構であるが、ロジックでは計れない個別事象の中にブレークスルーの存在する場合が多い。特に医療や生理活性に関して、この点は重要である。領域専門家との相互作用を考える場合、個別事例を大切に扱うようなソフトウェ

アシシステムを望みたい。また、領域専門家と情報専門家との距離は想像以上に遠い。相手の立場をいくら考えても、考え過ぎることはないと言える。

コメント 6: 医療のデータは本来研究のために集積されたものではなく、診療のために個別に得られた結果の集合体であり、きわめて heterogenous なものである。このため通常の実験データの解析とは異なり、大量の欠損データや統計的に対比し難いデータ、あるいは治療による修飾など解析にあたって考慮すべき点が多いと思われる。これらを正しく処理しないと、重要な知見を見逃してしまう可能性がある。また検査データは当然、誤差を含むなど、分析上の様々な留意点があるので、医療側と協力してより有用な成果を出してほしい。

コメント 8: 医療診断システム Mycin の過去の研究例にあるように、領域の専門家が入っていても、成果が実際に現場で使われないということはよくある。過去にエキスパート・システムが幾つも失敗している。現場で使ってもらえるものを開発するのは非常に難しい。新しい知識の発見も重要だが、標準的な疾患とその治療法の対応をデータから抽出し、名医が実践していることを検証し EBM データとして示すことも重要である。

#### (7) 総合評価

コメント 1: この研究は発展段階であるが、研究のフォーカスが分散しないよう配慮して進められており、(1)の研究目的を達成できるものと推察する。研究は良好な状況であると考える。

コメント 2: 重要な課題に取り組んでいる。このまま計画どおり進めれば、十分な成果が得られるものと期待する。

コメント 3: 難しい課題に挑戦している。研究は順調に進展しており。このまま計画通り進めてよいと判断する。今後の研究の進展に十分期待できる。

コメント 4: 総合的に判断して一年余すぎた現時点としては研究は順調に進んでおり、これからの研究の進展に十分期待できる。

コメント 5: プロジェクトの組織と研究の進め方に配慮し、総体として良好に研究は進展していると考えられる。本質的に非常に難しい課題に取り組んでいるため、領域専門家と情報専門家の相互作用に今後ともより一層の注意を払って、計画通り研究を進めて欲しい。

コメント 6: 今後の展開に大いに期待している。このプロジェクトは工学系・医学系の両方に前進を見いだすものとなると思われる。とくに医学においては retrospective study の基本的ツールとして広く用いられる可能性が大きい。国際的にみても医学分野におけるデータマイニングの集約的研究としては特筆すべきものがある。

コメント 8, 9, 10: 成果に注目している。このまま強力に進めて欲しい。



## 付録 3 : 成果公表一覧

### 招待講演/招待論文

- [1] Y. Kitamura: Interactive Integration of Information Agents on the Web, *5th International Workshop on Cooperative Information Agents* (2002).
- [2] 北村 泰彦: Web とエージェントと教育システム, 人工知能学会第 16 回 AI シンポジウム「e-learning の intelligent 化に向けて」(2001).
- [3] 沼尾 正行: アクティブ情報収集, 第 16 回人工知能学会全国大会論文集, 3A5-01 pp. 1-4, (2002).
- [4] 山口 高平: ユーザ指向アクティブマイニング, 第 16 回人工知能学会全国大会論文集, 3A5-02 pp. 1-4, (2002).
- [5] 津本 周作: アクティブユーザリアクション, 第 16 回人工知能学会全国大会論文集, 3A5-03 pp. 1-4, (2002).
- [6] H. Motoda and H. Liu: Feature Selection, Extraction and Construction, Proc. of Foundation of Data Mining, PAKDD02, pp. 67-72 (2002).
- [7] 元田 浩: AI セミナー「アクティブマイニングの構想と展開」, 第 16 回人工知能学会全国大会論文集, 3A4-01 pp. 1-4, (2002).
- [8] H. Motoda (Panelist): Innovative Projects for Intelligent Systems in the New Century: Active Mining Project, Thirteenth International Symposium on Methodologies for Intelligent Systems (ISMIS) (2002).
- [9] E. Suzuki (Panel Moderator): Innovative Projects for Intelligent Systems in the New Century, Thirteenth International Symposium on Methodologies for Intelligent Systems (ISMIS) (2002).
- [10] 鷲尾 隆: データ収集の新機軸とマイニングシステムの未来, 第 17 回 AI シンポジウム: 人工知能学会研究会資料 SIG-J-A103, pp. 65-70 (2002).
- [11] 鈴木英之進: データマイニングにおけるデータ変換, 第 17 回 AI シンポジウム: 人工知能学会研究会資料 SIG-J-A103, pp. 23-27 (2002).
- [12] T. Washio: Conditions of Law Equations and the Approach of their Discovery, Proc. of International Conference on Advances in Infrastructure for Electronic Business, Education, Science and Medicine on the Internet: SSGRR 2002w, No.83 (2002).
- [13] T.B. Ho: Knowledge Discovery and Data Mining Approach to Computational Biology, International Conference on Differential Equations, Approximations and Applications (invited lecture), Hanoi, pp. 10 (abstract) (2001).
- [14] 山口 高平: 金融データにおける DM アプリケーションの自動合成, 人工知能学会第 17 回 AI シンポジウム, pp. 59-64 (2002).
- [15] 山口 高平, 和泉 憲明: 意味理解のための次世代 Web –Semantic Web– の最新動向, ソ

フトウェア科学会チュートリアル (2002).

- [16] T. Yamaguchi: Implementing Active Mining Based on Meta-Learning, 1st International Conference on Global Research and Education (Inter-Academia 2002), Bratislava, Slovakia. (2002).
- [17] 津本 周作: 医療情報から見た歯科医療界の今後-最新の情報技術と診療支援-, 島根県歯科医師会生涯教育講座, pp. 1-12 (2001).
- [18] 津本 周作: 医療におけるアクティブマイニング, -Medical Data Mining からの新たな展開-, バイオメディカルファジイシステム学会第 14 回年次大会, pp. 1-4 (2001).
- [19] 岡田 孝: データマイニング入門と構造活性相関, 日本環境変異原学会 第 41 回 MMS 研究会 (2002).
- [20] Y. Takahashi, S. Fujishima and K. Yokoe: Similarity Analysis in Chemical Data Space, Asia Pacific Symposium on Information and Communication Technology in Chemical Education, Research & Development 2002, Keynote lecture, KL2, Kuala Lumpur, Malaysia (2002).
- [21] 大澤 幸生: (財) 日本科学技術連盟・多変量解析シンポジウム招待講演「人のチャンス発見プロセスにおける情報視覚化と KeyGraph」(2002).

## 公開ソフトウェア

- [1] ロコミ支援システム WAVE:  
<http://www.nm.cs.titech.ac.jp:12581/wom/>
- [2] 前処理支援システム TransX: <http://active.nm.cs.titech.ac.jp/transx/>
- [3] ユーザ指向モデル選択システム D2MS:  
URL: <http://www.jaist.ac.jp/ks/labs/ho/programs>
- [4] シーケンシャルパターンマイニング 拡張 PrefixSpan:  
<http://cl.taku-ku/software/prefixspan/>
- [5] アプリケーション自動生成ツール CAMLET  
<http://panda.cs.inf.shizuoka.ac.jp/japanese/study/KDD/camlet/>
- [6] カスケードモデル DISCAS version 2.1  
<http://www.clab.kwansei.ac.jp/mining/discas/discas.html>

## 受賞

- [1] 中井 有紀, 山田 誠二: Web ページにおける部分情報の更新モニタリング, 人工知能学会研究奨励賞 (2001).

- [2] 阪本 俊樹, 回り道 康博, 北村 泰彦, 辰巳 昭治: マルチキャラクタを用いた競争型情報推薦システム, 人工知能学会第 15 回全国大会優秀論文賞受賞 (2001).
- [3] T. Washio and H. Motoda: Discovering Admissible Simultaneous Equation Models from Observed Data, Machine Learning: ECML2001, Proc. of the 12th European Conference on Machine Learning, Best Paper Award (2001).
- [4] 工藤 拓, 山本 薫, 坪井 祐太, 松本 裕治: 言語情報を利用したテキストマイニング, 言語処理学会第 8 回年次大会優秀発表賞 (2002) .
- [5] 松村 真宏, 大澤 幸生, 石塚 満: 語の活性度に基づくキーワード抽出法, 第 2 回人工知能学会 MYCOM 優秀プレゼンテーション賞受賞 (2001).

## 報道

- [1] 「 計算機にひらめきを: 天才の思考法を組み込む試み 」(日本経済新聞 2002 年 4 月 7 日)
- [2] 「 アクティブマイニング: 情報洪水時代の救世主( 仮題 )」(読売新聞, 取材中, 掲載決定)

## 分担者によるワークショップ・シンポジウム・パネル開催

- [1] 津本 周作, 大澤 幸生: 特集「アクティブマイニング」および一般, 人工知能学会第 46 基礎論, 人工知能学会第 54 回知識ベースシステム合同研究会, 2001 年 11 月.
- [2] 矢田 勝俊, 羽室 行信, 加藤 直樹: ビジネスマイニングワークショップ( 関西大学 ), 2001 年 12 月 .
- [3] 大澤 幸生, 矢田勝俊, 庄司裕子, 鎌田真由美, 河合深雪: パネルディスカッション「アクティブマイニング時代におけるチャンス発見の役割」第 47 回人工知能基礎論研究会 (SIG-FAI47), 2002 年 1 月.
- [4] 鷲尾 隆: 第 17 回 A I シンポジウム( 人工知能学会 ), 2002 年 3 月.
- [5] 河野浩之: 特集「アクティブマイニング」および一般, 情報処理学会第 128 回知能と複雑系, 人工知能学会第 56 回知識ベースシステム合同研究会, 2002 年 5 月.
- [6] 元田 浩, 沼尾 正行, 山口 高平, 津本 周作: 特集「アクティブマイニング」, 人工知能学会第 16 回全国大会研究会特別セッション, 2002 年 5 月.
- [7] E. Suzuki and H. Motoda: Innovative Projects for Intelligent Systems in the New Century, Panel at Thirteenth International Symposium on Methodologies for Intelligent Systems (ISMIS), 2002 年 6 月
- [8] T. Yamaguchi and H. Motoda: Pacific Rim Knowledge Acquisition Workshop, Workshop at Seventh Pacific Rim International Conference on Artificial Intelligence, 2002 年 8 月
- [9] A. Abe and Y. Ohsawa: The Second International Workshop on Chance Discovery (PRICAI-CDWS02) at Seventh Pacific Rim International Conference on Artificial Intelligence, 2002 年 8 月

- [10] T. Terano: Second International Workshop on Agent-based approach to Economic and Social Complex Systems, 2002年8月.
- [11] T. Terano and M. Myers: Pacific Asian International Conference on Information Systems (PACIS 2002), 2002年9月.
- [12] Y. Ohsawa, S. Parsons, and P. McBurnery: Symposium on Chance Discovery: The Discovery and Management of Chance Events, in The 2002 AAAI Fall Symposium Series (2002年11月開催予定)
- [13] H. Motoda: International Workshop on Active Mining (AM-2002), Workshop at the 2002 IEEE International Conference on Data Mining (2002年12月開催予定)
- [14] K. Yoshida: SAINT2003: Workshop on AI application of the Internet (2003年1月開催予定)
- [15] 鷲尾 隆: AIシンポジウム(人工知能学会)(2003年3月開催予定).
- [16] 北村 泰彦, 津本 周作, 沼尾 正行: 特集「アクティブマイニング」, 電子情報通信学会人工知能と知識処理, 人工知能学会基礎論, 人工知能学会知識ベース, 情報処理学会知能と複雑系合同研究会 (2003年3月開催予定).

## ジャーナル論文

- [1] 山田 誠二, 中井 有紀: 対話的分類学習による Web ページの部分更新モニタリング, 人工知能学会論文誌, Vol.17 No.5 (2002).
- [2] 高間 康史, 廣田 薫: 免疫ネットワーク・メタファに基づく Web 情報可視化手法, 日本ファジィ学会誌, Vol. 14, No. 5, (2002)
- [3] Y. Takama, K. Hirota: Web Information Visualization Method Employing Immune Network Model for Finding Topic Stream from Document-Set Sequence, J. of New Generation Computing, (2002).
- [4] 岡部 正幸, 山田 誠二: 関係学習を用いた対話的文書検索, 人工知能学会誌, Vol. 16, No. 1, F (2001).
- [5] 山田 誠二, 村瀬 文彦: ページ情報エージェントの組織化による Web ブラウザの適応インタフェース, 人工知能学会誌, Vol. 16, No. 1, F (2001).
- [6] 北村 泰彦, 野田 知哉, 辰巳 昭治: 動的情報メディアータのための知的情報収集手法, 電子情報通信学会論文誌 D-I, Vol. J84-D-I, No.8, pp. 1256-1265 (2001).
- [7] Y. Kitamura, T. Noda, and S. Tatsumi: An Intelligent Information Gathering Method for Dynamic Information Mediators, Systems and Computers in Japan, Vol.33, No.9, pp. 11-20 (2002).
- [8] 小久保 卓, 小山 聡, 山田 晃弘, 北村 泰彦, 石田 亨: 検索隠し味を用いた専門検索エンジンの構築, 情報処理学会論文誌, Vol.43, No.6, pp. 1804-1813 (2002).
- [9] C. R. Mofizur and M. Numao: Automated bias shift in a constrained space for logic

- program synthesis, 人工知能学会論文誌, Vol. 16, No. 6, pp. 548-556 (2001).
- [10] T. N. Tran and M. Nuamo: Active Information Gathering by Making Use of Existing Databases, Transactions of the Japanese Society of Artificial Intelligence, 17 [5], pp. 622-629 (2002).
  - [11] 森山 甲一、沼尾 正行: 環境状況に応じて自己の報酬を操作する学習エージェントの構築、人工知能学会論文誌、17 [6], pp. 676-683 (2002).
  - [12] R. Ichise and M. Numao: Learning First-Order Rules to Handle Medical Data, NII Journal, Vol. 2, pp. 9-14 (2001).
  - [13] R. Ichise and M. Numao: Learning Relational Rule from Examples that are Neither Positive nor Negative, Systems and Computers in Japan, Vol. 32, No. 14, pp. 34-40 (2001).
  - [14] 市瀬 龍太郎, 武田 英明, 本位田 真一: 階層的知識間の調整規則の学習, 人工知能学会論文誌, Vol. 17, No. 3, pp. 230-238 (2002).
  - [15] 松田 喬, 元田 浩, 鷲尾 隆: 一般グラフ構造データに対する Graph-Based Induction とその応用, 人工知能学会誌, Vol.16, No. 4, pp. 363-374 (2001).
  - [16] 寺邊 正大, 鷲尾 隆, 元田 浩:  $S^3$ Bagging による高速な分類器生成, 数理モデル化と応用, Vol.42, No.14, pp. 25-38 (2001).
  - [17] 堀 聡, 瀧 寛, 鷲尾 隆, 元田 浩: データマイニングを用いた市場品質監視システム, 電気学会 電子・情報・システム部門誌, Vol.121-c, No.8 (2001).
  - [18] A. Inokuchi, T. Washio, T. Okada and H. Motoda: Applying the Apriori-based Graph Mining Method to Mutagenesis Data Analysis, J. of Computer Aided Chemistry, Vol. 2, pp. 87-92 (2001).
  - [19] T. Matsuda, H. Motoda and T. Washio: Graph-Based Induction and Its Applications, Advanced Engineering Informatics, Vol. 16, No. 2, pp. 135-143 (2002).
  - [20] M. Terabe, T. Washio, H. Motoda, O. Katai and T. Sawaragi: Attribute Generation Based on Association Rules, Knowledge and Information Systems, An International Journal, Vol. 4, No. 3, pp. 329-349 (2002).
  - [21] T. Yoshida, T. Kondo and S. Nishida: Evolving Granules for Classification for Discovering Difference in the Usage of Words, Computational Intelligence, Vol. 17, No.3, 580-592 (2001).
  - [22] 渡辺 正人, 吉田 哲也, 才脇 直樹, 土方 嘉徳, 西田 正吾: イメージに基づいた Web ページデザイン支援システム, ヒューマンインタフェース学会論文誌, Vol.3, No.4, pp. 287-297 (2001).
  - [23] 大西 健介, 吉田 哲也, 西田 正吾: イメージに基づいた Web ページデザイン支援システム, 決定木の相関関係に基づいた概念相違検出手法, 電子情報通信学会論文誌 D I (in press) (2002).
  - [24] T.B. Ho, T.D. Nguyen, D.D. Nguyen and S. Kawasaki: Visualization Support for User-

- Centered Model Selection in Knowledge Discovery and Data Mining, *International Journal of Artificial Intelligence Tools*, pp. 691-713 (2001).
- [25] T.B. Ho and N.B. Nguyen: Nonhierarchical Document Clustering by a Tolerance Rough Set Model, *International Journal of Intelligent Systems*, Vol. 17, No. 2, pp. 199-212 (2002).
- [26] T.B. Ho, S. Kawasaki and N.B. Nguyen: Cluster-based Information Retrieval with a Tolerance Rough Set Model, *International Journal of Fuzzy Logic and Intelligent Systems*, Vol. 2, No. 1, pp. 26-32 (2002).
- [27] V.N. Huynh, T.B. Ho and Y. Nakamori: A Parametric Representation of Linguistic Hedges in Zadeh's Fuzzy Logic, *International Journal of Approximate Reasoning* (in press, <http://www.elsevier.nl/locate/ijar>).
- [28] T.B. Ho and D.D. Nguyen: Chance Discovery and Learning Minority Classes, *Journal of New Generation Computing* (accepted, Special Issue on Chance Discovery). Vol. 21, No. 2 (2003).
- [29] 矢田 勝俊, 羽室 行信, 加藤 直樹: 経営データからの知識発見, *国民経済雑誌*, Vol.184, No.1, pp. 19-33 (2001).
- [30] N. Izumi, T. Yamaguchi: Integration of Heterogeneous Repositories Based on Ontologies for EC Applications Development, *Journal of Electronic Commerce Research And Application*, Vol.1, No.1 pp. 77-91 (2002).
- [31] 阿部 秀尚, 山口 高平: メソッドリポジトリに基づく昨日アプリケーションの並列合成とその洗練化, *人工知能学会論文誌*, 第17巻, 5号, pp. 647-657 (2002).
- [32] 和泉 憲明, 山口 高平: オントロジーに基づくソフトウェアエージェントのパターン指向開発, *電子情報通信学会誌*, Vol.J-84-D-I, No.8, pp. 1181-1190 (2001).
- [33] 和泉 憲明, 山口 高平: 教師支援エージェント構築のためのタスクパターンリポジトリの開発, *教育システム情報学会誌*, Vol.18, No.3・4, pp. 352-363 (2001).
- [34] N. Zhong, C. Liu and S. Ohsuga: Dynamically Organizing KDD Process, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 15, No. 3, World Scientific, pp. 451-473 (2001).
- [35] N. Zhong, J.Z. Dong and S. Ohsuga: Rule Discovery by Soft Induction Techniques, *Neurocomputing, An International Journal*, Vol. 36 (1-4) Elsevier, pp. 171-204 (2001).
- [36] N. Zhong, J.Z. Dong, C. Liu and S. Ohsuga: A Hybrid Model for Rule Discovery in Data, *Knowledge Based Systems, An International Journal*, Vol. 14, No. 7, Elsevier, pp. 397-412 (2001).
- [37] N. Zhong, J.Z. Dong and S. Ohsuga: Using Rough Sets with Heuristics to Feature Selection, *Journal of Intelligent Information Systems*, Vol. 16, No. 3, Kluwer, pp. 199-214 (2001).
- [38] N. Zhong and A. Skowron: A Rough Sets Based Knowledge Discovery Process, *Interna-*

- tional Journal of Applied Mathematics and Computer Science*, Vol. 11, No. 3, Technical University Press, pp. 101-117 (2001).
- [39] N. Zhong: Rough Sets in Knowledge Discovery and Data Mining, *Journal of Japan Society for Fuzzy Theory and Systems*, Vol. 13, No. 6, pp. 581-591 (2001).
- [40] C. Liu and N. Zhong: Rough Problem Settings for ILP Dealing with Imperfect Data, *Computational Intelligence, An International Journal*, Vol. 17, No. 3, Blackwell Publishers, pp. 446-459 (2001).
- [41] 鈴木 英之進 :「一般性と正確性に基づくルール発見の最悪解析」, 人工知能学会論文誌, Vol. 17, No. 5, pp. 630-637 (2002) .
- [42] E. Suzuki: Sampling Theories for Rule Discovery Based on Generality and Accuracy: the Worst Case and a Distribution-Based Case, *Communication of Institute of Information and Computing Machinery, Taiwan*, Vol. 5, No. 2, pp. 83-88 (2002).
- [43] Y. Y. Yao and N. Zhong: Granular Computing using Information Tables, *Data Mining, Rough Sets and Granular Computing*, T. Y. Lin, Y. Y. Yao and L. A. Zadeh (eds.), Physica-Verlag, Heidelberg, pp. 102-124 (2002).
- [44] N. Zhong: Mining Interesting Patterns in Multiple Data Sources, Information Fusion in Data Mining, V. Torra (ed.), *Studies in Fuzziness and Soft Computing Series*, Physica-Verlag, (2002, accepted for publication).
- [45] 山田 悠, 鈴木 英之進 :「動的時間伸縮法に基づく時系列データの高速度クラスタリング」, *Information Technology Letters (ITレターズ)*, (2002, accepted for publication) .
- [46] N. Zhong, J. Z. Dong and S. Ohsuga: Meningitis Data Mining by Cooperatively Using GDT-RS and RSBR, *Pattern Recognition Letters*, Elsevier Science (2002, accepted for publication).
- [47] 山田 寛康, 工藤 拓, 松本 裕治: Support Vector Machine を用いた日本語固有表現抽出, *情報処理学会論文誌*, Vol.43, No.1, pp. 44-53 (2002).
- [48] 松本 裕治: 自然言語処理におけるシステム混合法の利用, *電子情報通信学会論文誌 D-II, 招待論文*, Vol.J85-D-II, No.5, pp. 709-716 (2002).
- [49] 工藤 拓, 松本 裕治: チャンキングの段階適用による日本語係り受け解析, *情報処理学会論文誌*, Vol.43, No.6, pp. 1834-1842 (2002).
- [50] S. Yokoyama, K. Matsuoka, S. Tsumoto, M. Harao, T. Yamakawa, K. Sugahara, C. Nakahama, S. Ichiyama, and K. Watanabe: Study on the Association between the Patient's Clinical Background and the Anaerobes by Data Mining in Infectious Disease Database, *BMFSA*, vol. 7, no. 1, pp. 121-128 (2001).
- [51] 安田 晃, 柳樂 真佐実, 孫 暁光, 津本 周作, 山本 和子: 自主学習における学生の自己評価の変動に関する解析, *医学教育*, vol. 32, pp. 69-75 (2001).
- [52] 津本 周作: 医学における知識発見手法の可能性 (特集: データマイニングコンテスト), *情報処理*, vol. 42, pp. 472-477 (2001).

- [53] 津本 周作: ラフ集合論の現状と課題 (特集: ラフ集合の理論と応用), 日本ファジィ学会誌, vol. 13, pp. 552-561 (2001).
- [54] 柳樂 真佐実, 安田 晃, 孫 暁光, 平野 章二, 津本 周作: 現在の病院情報システムの評価に関する研究, 医療情報学, vol. 22, no. 1, pp. 51-57 (2002).
- [55] S. Tsumoto: Rule Induction, Rough Sets and Matroid Theory. Communications of Institute and Computing Machinery (IICM) vol 5 (2), pp. 89-94 (2002).
- [56] S. Tsumoto, S. Hirano, A. Yasuda, K. Tsumoto: Analysis of Amino Acid Sequences by Statistical Technique, *Information Sciences*, in press (2002).
- [57] S. Hirano, S. Tsumoto, T. Okuzaki, Y. Hata, and K. Tsumoto: Analysis of Biochemical Data Aided by a Rough Sets-Based Clustering Technique, *International Journal of Fuzzy Systems*, in press (2002).
- [58] T. Okada: Discovery of Structure Activity Relationships using the Cascade Model: The Mutagenicity of Aromatic Nitro Compounds, *Journal of Computer Aided Chemistry*, Vol. 2, pp. 79-86 (2001).
- [59] T. Okada: Characteristic Substructures and Properties in Chemical Carcinogens Studied by the Cascade Model, to appear in *Bioinformatics* (2002).
- [60] 加藤 博明, 田所 哲男, 宮田 博之, 近松 信一, 高橋 由雅, 阿部 英次: タンパク質構造データマイニングのための三次元モチーフ辞書の作成, 人工知能学会論文誌, Vol. 17, pp. 608-613 (2002).
- [61] Y. Takahashi, M. Konji and S. Fujishima: MolSpace: A Computer Desktop Tool for Visualization of Massive Molecular Data, *J. Mol. Graph. Model.*, in press.
- [62] 大澤 幸生, 高間 康史: Web に潜む創造的意思決定のチャンス, 人工知能学会誌 Vol.16, No.4 pp. 530-534 (2001).
- [63] Y. Ohsawa: Chance Discoveries for Making Decisions in Complex Real World, *New Generation Computing* (Springer Verlag and Ohmsha), Vol.20 No.2 (2002).
- [64] W. Sunayama, Y. Nomura, Y. Ohsawa and M. Yachida: Support System for User Interests Expression on Searching Web Page, *Systems and Computer in Japan* Vol.32, No.13, pp. 14-22 (2001).
- [65] 松村 真宏, 大澤 幸生, 石塚 満: テキストによるコミュニケーションにおける影響の普及モデル, 人工知能学会論文誌 Vol.17 No.3, pp. 259-267 (2002).
- [66] 平 博司, 福島 伸一, 大澤 幸生, 伊庭 斉志, 石塚 満: AreaView2001:WWWからの構造化した領域総覧提示システム人工知能学会論文誌 Vol.17 No.3, pp. 268-275 (2002).
- [67] N. Matsumura, Y. Matsuo, Y. Ohsawa and M. Ishizuka: Discovering Emerging Topics from WWW, *Journal of Contingencies and Crisis Management* Vol.10, No.2, pp. 73-81 (2002).
- [68] Y. Ohsawa: KeyGraph as Risk Explorer from Earthquake Sequence, *Journal of Contingencies and Crisis Management*, in press (2002).



- [69] Y. Ohsawa and H. Fukuda: Chance Discovery by Stimulated Group of People - An Application to Understanding Rare Consumption of Food - *Journal of Contingencies and Crisis Management*, in press (2002).
- [70] 松尾 豊, 大澤 幸生, 石塚 満: Small World 構造に基づく文書からのキーワード抽出情報処理学会論文誌 Vol.43, No.6, (2002).
- [71] 松村真宏, 大澤幸生, 石塚満: 語の活性度に基づくキーワード抽出法人工知能学会論文誌 Vol.17, No.4, pp. 398-406 (2002).
- [72] N. Matsumura, Y. Ohsawa and M. Ishizuka PAI: Automatic Indexing for Extracting Asserted Keywords from a Document *New Generation Computing*, Vol.21 No.1 (2002).
- [73] Y. Ohsawa and Y. Nara: Understanding Internet Users on Double Helical Model of Chance-Discovery Process Authors, *New Generation Computing* Vol.21 No.2, in press (2002).
- [74] Y. Matsuo, Y. Ohsawa and M. Ishizuka Average-clicks: A New Measure of Distance on the World Wide Web Intelligent Information Systems Journal (Kluwer), Special Issue on Web Intelligence (2002).
- [75] 高田 真好, 寺野 隆雄: 2 段階 CLP 緩和解法によるリソース平準化スケジューリングシステム. 電子情報通信学会論文誌 D-I, Vol.J84-D-I No. 6, pp. 896-905 (2001).
- [76] 倉橋 節也, 寺野 隆雄: エージェントシミュレーションによる共同分配規範モデル. 電子情報通信学会論文誌 D-I, Vol.J84-D-I No.8, pp. 1160-1168 (2001).
- [77] K. Takadama, T. Terano and K. Shimohara: Non-Governance Rather Than Governance in a Multiagent Economic Society, *IEEE Transaction on Evolutionary Computation*, Vol. 5, pp. 535-545 (2001).
- [78] R. Kudo and T. Terano: Automation of Concept Development, *Expert Systems*, Vol.6 No.45, pp1641-1665 (2002).
- [79] 寺野 隆雄: データマイニングの展望. 計測と制御, 第 4 1 巻, 第 5 号, pp. 315-318, 2002.5 月.
- [80] 高橋 大志, 寺野 隆雄: エージェントシミュレーションによる GARCH モデルと Prospect 理論の関連性の分析. シミュレーション, 第 21 巻, 第 2 号, pp. 55-64, June 2002.
- [81] 倉橋 節也, 寺野 隆雄: 学習分類子システムを用いたプロセス時系列のアクティブマイニング. 人工知能学会論文誌, 第 17 巻, 5 号, pp. 638-646, 2002.9
- [82] S. Katsuno, K. Yamazaki, T. Kubo, T. Asami, K. Sugauchi, O. Tsunehiro, H. Enomoto, K. Yoshida and H. Esaki, High-speed IP meter HIM and its Application in LAN/WAN environments, 電子情報通信学会 英文誌 D Special Issue on New Technologies in the Internet and their Applications, in press (2002).

## 国際発表

- [1] S. Yamada and Y. Nakai: Monitoring Partial Updates in Web Pages using Relational Learning, 2nd International Conference on Adaptive Hypermedia and Adaptive Web Based Systems (AH-2002), pp. 452-455 (2002).
- [2] Y. Takama, K. Hirota: Visualization of Topic Distribution from Document Sequence on Web, in P. Sincak et al. Ed., Intelligent Technologies - Theory and Applications - New Trends in Intelligent Technologies, IOS Press, pp. 189-195, (2002).
- [3] Y. Takama: Topic Stream Extraction from Web Online News Articles, 1st Pan-Pacific Symposium on IT, 2002 (to appear).
- [4] Y. Takama: Immune Network Metaphor for Keyword Map Generation, International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'02), 2002. (to appear).
- [5] M. Okabe and S. Yamada: Interactive Web Page Filtering with Relational Learning, The First Asia-Pacific Conference on Web Intelligence (WI-2001), pp. 443-447 (2001).
- [6] S. Yamada and Y. Osawa: Information Gathering of Web pagesto Guide Concept Understanding, The Tenth International World Wide Web Conferenc (WWW10), Poster 1046 (2001).
- [7] N. Nagino and S. Yamada: Assistance of Web browsing by indicating the future Web pages, The 9th International Conference on Human-Computer Interaction, pp. 1140-1144 (2001).
- [8] M. Mase and S. Yamada: Development and Evaluation of an Information Retrieval System for User Groups and The WWW, The 9th International Conference on Human-Computer Interaction, pp. 1125-1129 (2001).
- [9] M. Okabe and S. Yamada: Interactive Document Retrieval with Relational Learning, pp. 27-31, SAC-2001 ACM SYMPOSIUM ON APPLIED COMPUTING (2001).
- [10] Y. Takama and K. Hirota: Consideration of Presentation Timing Problem for Chance Discovery, pp. 429-432, 5th World Multiconference on Systems, Cybernetics and Informatics (SCI2001) (2001).
- [11] Y. Takama, M. Kawabe, K. Hirota: Kansei-keyword Extraction from Japanese Film Scenario Using Sensitivity Information, Joint 9th IFSA World Congress and 20th NAFIPS International Conference (IFSA/NAFIPS2001), 513(CD-ROM Proceedings) (2001).
- [12] Y. Takama and K. Hirota: Employing Immune Network Model for Clustering with Plastic Structure, 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2001), MP-6-1 (CD-ROM Proceedings) (2001).
- [13] Z. Stejic, Y. Takama, K. Hirota: Integrated Querying of Images and Text, pp. 66-70, 2nd Int. Symposium on Advanced Intelligent System Conference (ISIS2001) (2001).
- [14] E. M. Iyoda, Z. Stejic, Y. Takama, K. Hirota: Image Retrieval Using Local Similarity Patterns Inferred by Genetic Algorithm, pp. 101-105, 2nd Int. Symposium on Advanced

- Intelligent System Conference (ISIS2001) (2001).
- [15] Y. Takama, K. Oh, K. Hirota: Finding landmarks in Keyword Map Based on Immune Network, pp. 37-41, 2nd Int. Symposium on Advanced Intelligent System Conference (ISIS2001) (2001).
  - [16] Z. Stejic, Y. Takama, K. Hirota: Improving Retrieval Effectiveness by Integrated Querying of Images and Text, pp. 233-244, 12th Int'l Conf. On Information and Intelligent Systems (IIS2001) (2001).
  - [17] Y. Takama and K. Hirota: Finding Topic Distribution from A Sequence of Document Sets, pp. 132-139, 2nd Vietnam-Japan Symposium on Fuzzy Systems and Applications (VJFUZZY) (2001).
  - [18] Y. Takama and K. Hirota: WWW Information Visualization Based on Plastic Clustering, 10th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE2001), S404 (CD-ROM Proceedings) (2001).
  - [19] Z. Stejic, E. M. Iyoda, Y. Takama, K. Hirota: Automatic Textual Summarization of Image Database Contents Using Combination of Clustering and Neural Network Techniques, pp. 233-239, 2nd International Conference on Intelligent Technologies (InTech'2001) (2001).
  - [20] Y. Takama and K. Hirota: Consideration of Memory Cell for Immune Network-based Plastic Clustering Method, pp. 409-414, 2nd International Conference on Intelligent Technologies (InTech'2001) (2001).
  - [21] Y. Kitamura, T. Yamada, T. Kokubo and Y. Mawarimichi, T. Yamamoto and T. Ishida: Interactive Integration of Information Agents on the Web (Invited Talk), Matthias Klusch, Franco Zambonelli (Eds.), Cooperative Information Agents V, Lecture Notes in Artificial Intelligence 2182, Berlin et al.: Springer-Verlag, pp. 1-13 (2001).
  - [22] S. Oyama, T. Kokubo, T. Yamada, Y. Kitamura and T. Ishida: Keyword Spices: A New Method for Building Domain-Specific Web Search Engines. Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI-01), pp. 1457-1463, (2001).
  - [23] W.E. Walsh, M. Yokoo, K. Hirayama, M. P. Wellman: On Market-Inspired Approaches to Propositional Satisfiability, Proceeding of the 17th International Joint Conference on Artificial Intelligence (IJCAI-01), pp. 1152-1158 (2001).
  - [24] M. Tanaka, S. Nakazono, H. Matsuno, H. Tsujimoto, Y. Kitamura and S. Miyano: Intelligent System for MEDLINE Record Selection by Keyword Recommendation and Learning Text Characteristics, Pacific Symposium on Biocomputing 2001, pp. 130, (2001).
  - [25] Y. Kitamura, H. Tsujimoto, T. Yamada, and Taizo Yamamoto: Multiple Character-Agents Interface: An Information Integration Platform Where Multiple Agents and Human User Collaborate, Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'02), pp. 790-791 (2002).

- [26] Y. Kitamura, T. Sakamoto, and S. Tatsumi: A Competitive Information Recommendation System and Its Behavior, 6th International Workshop on Cooperative Information Agents (CIA-02) (to appear) (2002).
- [27] M. Yokoo, K. Suzuki and K. Hirayama: Secure Distributed Constraint Satisfaction: Reaching Agreement without Revealing Private Information, Proceedings of Eighth International Conference on Principles and Practice of Constraint Programming (CP-2002) (to appear) (2002).
- [28] M. Yokoo, K. Suzuki and K. Hirayama: Secure Distributed Constraint Satisfaction: Reaching Agreement without Revealing Private Information, Proceedings of the Third International Workshop on Distributed Constraint Reasoning, pp. 183-191 (2002).
- [29] K. Hirayama and M. Yokoo: Local Search for Distributed SAT with Complex Local Problems, Proceedings of the First International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS-02), pp. 1199-1206 (2002).
- [30] C. Nattee and M. Numao: Geometric method for document understanding and classification using on-line machine learning, *Sixth International Conference on Document Analysis and Recognition*, pp. 602-606. IEEE Computer Society Press (2001).
- [31] T. N. Tran and M. Numao: Text data mining in biomedical literature by combining with an information retrieval approach, *Proc. the 14th International Conference on Applications of Prolog*, pp. 295-304. INAP Organizing Committee/Prolog Association of Japan, (2001).
- [32] T. N. Tran and M. Numao: Text mining from biomedical literature using term extraction and association rule mining, *Proc. the International Conference on Machine Learning and Applications (ICMLA '02)*, pp. 70-76 (2002).
- [33] M. Numao, D. Kato, and M. Yokoyama: Learning organization in global intelligence. *AAAI Spring Symposia*, AAAI Press, pp. 41-47 (2002).
- [34] M. Numao, S. Takagi and K. Nakamura: Constructive Adaptive User Interfaces — Composing Music Based on Human Feelings, Proc. Eighteenth National Conference on Artificial Intelligence (AAAI-02), pp. 193-198, The AAAI Press / The MIT Press (2002).
- [35] M. Numao, S. Takagi and K. Nakamura: CAUI Demonstration, Proc. Eighteenth National Conference on Artificial Intelligence (AAAI-02), pp. 1010-1011, The AAAI Press / The MIT Press (2002).
- [36] K. Moriyama and M. Numao: Construction of a Learning Agent Handling Its Rewards According to Environmental Situations, Proc. the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2002), pp. 1262-1263, ACM Press (2002).
- [37] H. Kawano: Architecture of Trip Database Systems: Spatial Index and Route Estimation Algorithm, *XIV International Conf. of Systems Science*, Vol. III, pp. 110-117, Poland (2001).

- [38] H. Kawano: On-line Analytical Processing for person trip database, *The sixteenth triennial conference of the International Federation of Operational Research Societies*, p.119, (Edinburgh UK) (2002).
- [39] H. Kawano: On-line Analytical Processing in Person Trip Database *The 4th International Symposium on Operations Research and Its Applications* (2002).
- [40] R. Ichise, H. Takeda and S. Honiden: Automated Alignment of Multiple Internet Directories, Poster Proceedings of the 10th International World Wide Web Conference, pp. 194-195 (2001).
- [41] R. Ichise, H. Takeda and S. Honiden: Rule Induction for Concept Hierarchy Alignment, Proceedings of the IJCAI-01 Workshop on Ontology Learning (OL-2001) (2001)
- [42] R. Ichise, H. Takeda and S. Honiden: An Examination of the Relationships between Internet Directories, Proceedings of the EKAW02 Workshop on Knowledge Management through Corporate Semantic Webs (2002).
- [43] R. Ichise, D. Shapiro and P. Langley: Learning Hierarchical Skills from Observation, Proceedings of the 5th International Conference on Discovery Science (2002).
- [44] T. Wada, H. Motoda and T. Washio: Knowledge Acquisition from Both Human Expert and Data, Proc. of the Fifth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD01), Lecture Notes in Artificial Intelligence, Springer-Verlag, pp. 550-561 (2001).
- [45] T. Washio and H. Motoda: Discovery of Law Equations governing Human Affinity under Trade-off between Cost and Risk Proc. of International Meeting of The Psychometric Society (IMPS-2001), p.74 (2001).
- [46] T. Washio and H. Motoda: Discovering Admissible Simultaneous Equation Models from Observed Data, Machine Learning: ECML2001, Proc. of the 12th European Conference on Machine Learning, Springer, pp. 539-551 (2001).
- [47] M. Terabe, T. Washio and H. Motoda:  $S^3$ Bagging: Fast Classifier Induction Method with Subsampling and Bagging, Advances in Intelligent Data Analysis, Proc. of the Fourth International Symposium on Intelligent Data Analysis, Springer, pp. 177-186 (2001).
- [48] M. Terabe, T. Washio and H. Motoda: The Effect of Subsampling Rate on  $S^3$  Bagging Performance, Proc. of Active Learning, Database Sampling, Experimental Design: Views on Instance Selection, pp. 48-55 (Workshop of ECML/PKDD2001 (2001).
- [49] T. Yoshida, D. Shinkai and S. Nishida: Utilizing the Correlation between Query Keywords for Information, Proc. of the 2nd International Conference on Web-Age Information Management, Springer, pp. 49-59 (2001).
- [50] Y. Hijikata, T. Yoshida and S. Nishida: Adaptive Hypermedia System for Supporting Information Providers in Directing Users through Hyperspace, Proc. of the Third Workshop on Adaptive Hypertext and Hypermedia pp. 147-156 (2001).
- [51] H. Motoda and H. Liu: Feature Selection, Extraction and Construction, Invited, Proc.

- of Foundation of Data Mining, PAKDD02, pp. 67-72 (2002).
- [52] T. Matsuda, H. Motoda, T. Yoshida and T. Washio: Knowledge Discovery from Structured Data by Beam-wise Graph-Based Induction, Proc. of Seventh Pacific Rim International Conference on Artificial Intelligence, Springer, pp. 255-264 (2002).
  - [53] T. Wada, T. Yoshida, H. Motoda and T. Washio: Extension of the RDR method that can Adapt to Environmental Changes and Acquire Knowledge from Both Experts and Data, Proc. of Seventh Pacific Rim International Conference on Artificial Intelligence, Springer, pp. 218-227 (2002).
  - [54] K. Fujiwara, T. Yoshida, H. Motoda and T. Washio: Case Generation Method for Constructing an RDR Knowledge Base, Proc. of Seventh Pacific Rim International Conference on Artificial Intelligence, Springer, pp. 228-237 (2002).
  - [55] T. Matsuda, H. Motoda, T. Yoshida and T. Washio: Beam-wise Graph-Based Induction for Structured Data Mining, Proc. of Pacific Rim Knowledge Acquisition Workshop (PKAW2002), pp. 174-188 (2002).
  - [56] T. Wada, T. Yoshida, H. Motoda and T. Washio: Reorganizing Knowledge Base of RDR for Adaption to Environmental Changes, Proc. of Pacific Rim Knowledge Acquisition Workshop (PKAW2002), pp. 189-202 (2002).
  - [57] T. Matsuda, H. Motoda, T. Yoshida: Mining Patterns from Structured Data by Beam-wise Graph-Based Induction, Workshop Notes on Multi-Relational Data Mining, KDD-2002, pp. 113-129 (2002).
  - [58] T. Matsuda, H. Motoda, T. Yoshida and T. Washio: Preliminary Analysis of Hepatitis Data by Beam-wise Graph-Based Induction, Discovery Challenge Workshop Notes, ECML/PKDD 2002, ISBN 952-10-0639-0 (2002).
  - [59] T.D.Nguyen, T.B. Ho and H. Shimodaira: A Scalable Algorithm for Rule Post-Pruning of Large Decision Trees, 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining PAKDD00, Hongkong. Lecture Notes in Artificial Intelligence 2035, Springer, pp. 467-476 (2001).
  - [60] T.B. Ho, S. Kawasaki and N.B. Nguyen: Cluster-based Information Retrieval with Tolerance Rough Set Model, Second International Symposium on Advanced Intelligent Systems, Daejeon, Korea, pp. 6-11 (2001).
  - [61] T.B. Ho, D.D. Nguyen and S. Kawasaki: Mining Prediction Rules from Minority Classes, 14th International Conference on Applications of Prolog (INAP2001), International Workshop Rule-Based Data Mining RBDM pp. 254-264 (2001).
  - [62] T.B. Ho, T.T. Nguyen, P.C. Nguyen and C.M. Luong: Towards a Practical Framework for Vietnamese Natural Language Processing, Vietnam-Japan Symposium on Fuzzy Systems and Applications, pp. 297-304 (2001).
  - [63] T.B. Ho, S. Kawasaki and D.D. Nguyen: Extracting Predictive Knowledge from Meningitis Data by Integration of Rule Induction and Association Mining, International Work-

- shop Challenge in KDD, Lecture Notes in Artificial Intelligence 2253, Springer, pp. 508-515 (2001).
- [64] T.B. Ho and T.D. Nguyen: Visualization Support for User-Centered Model Selection in Knowledge Discovery in Databases, Proceedings 13th IEEE International Conference on Tools with Artificial Intelligence ICTAI 2001, Texas, pp. 228-235 (2001).
- [65] T.B. Ho: Knowledge Discovery and Data Mining Approach to Computational Biology, International Conference on Differential Equations, Approximations and Applications (invited lecture), Hanoi, pp. 10 (abstract) (2001).
- [66] T. Washio: Conditions of Law Equations and the Approach of their Discovery, Invited, Proc. of International Conference on Advances in Infrastructure for Electronic Business, Education, Science and Medicine on the Internet: SSGRR 2002w, No.83 (2002).
- [67] V.N. Huynh, Y. Nakamori, T.B. Ho and G. Resconi: A Context Model for Constructing Membership Functions of Fuzzy Concepts Based on Modal Logic, Inter. Symp. on Foundations of Information and Knowledge Systems, Kiel, Lecture Note in Computer Science 2284, Springer, pp. 93-104 (2002).
- [68] T.D. Nguyen, T.B. Ho and D.D. Nguyen: Data and Knowledge Visualization in the Knowledge Discovery Process, 5th International Conference Recent Advances in Visual Information Systems, Taiwan, Lecture Note in Computer Science 2314, Springer, pp. 311-321 (2002).
- [69] Y. Fujikawa and T.B. Ho: Cluster-based Algorithms for Filling Missing Values, 6th Pacific-Asia Conf. Knowledge Discovery and Data Mining, Taiwan, Lecture Notes in Artificial Intelligence 2336, Springer, pp. 549-554 (2002).
- [70] T.B. Ho, A. Saito, S. Kawasaki, D.D. Nguyen and T.D. Nguyen: Failure and Success Experience in Mining Stomach Cancer Data, International Workshop Data Mining Lessons Learned, International Conference Machine Learning 2002, Sydney, pp. 40-47 (2002).
- [71] T.B. Ho, T.D. Nguyen and D.D. Nguyen: Visualization Support for a User-Centered KDD Process, ACM International Conference on Knowledge Discovery and Data Mining KDD-02, Edmonton, pp. 519-524 (2002).
- [72] T.B. Ho, D.D. Nguyen, S. Kawasaki and T.D. Nguyen: Extracting Knowledge from Hepatitis Data with Temporal Abstraction, Discovery Challenge Workshop Notes, ECML/PKDD 2002, ISBN952-10-0639-0 (2002).
- [73] T.B. Ho, T.D. Nguyen, D.D. Nguyen: A User-Centered Visual Approach to Data Mining. The System D2MS, IFIP World Computer Congress, Stream 8 Intelligent Information Processing (IIP-2002), Montreal, August 2002 (in press).
- [74] D.D. Nguyen, T.B. Ho and S. Kawasaki: Learning Minority Classes in Unbalanced Datasets, Third International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'02), Kanazawa, 4-6 September 2002 (in press).
- [75] N.A.H. Pham and T.B. Ho: A Distributed Algorithm for Mining Association Rules,

- Third International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'02), Kanazawa, 4-6 September 2002 (in press).
- [76] S. Kawasaki, A. Saitou and T.B. Ho: Mining from Medical Data: Case-Studies in Meningitis and Stomach Cancer Domains, Sixth International Conference on Knowledge-Based Intelligent Information & Engineering Systems, Italy, 16-18 September 2002 (in press).
- [77] T. Nakada, T.B. Ho and S. Kunifuji: Finding and Visualizing Relations between Individuals and between Potential Communities, International Conference on Artificial and Computational Intelligence IASTED'02, Tokyo, 25-27 September 2002 (in press).
- [78] Y. Hamuro, N. Katoh and K. Yada: Discovering Association Strength among Brand Loyalties from Purchase History, Proceeding of 2001 IEEE International Symp. on Industrial Electronics, pp. 114-117 (2001).
- [79] Y. Hamuro, N. Katoh, E.H. Ip, S.L. Cheung, K. Yada: Discovery of Interesting Rules for Purchase Behavior Using String Pattern Analysis, The 2001 World Congress of Mass Customization and Personalization (MCPC 2001) (2001).
- [80] Y. Hamuro, N. Katoh, K. Yada and T. Yano: Discovering Purchase Association among Brands from Purchase History, Proceeding CD of SSGRR2002w (2002).
- [81] K. Terai, M. Sawai, N. Sugiura, N. Izumi and T. Yamaguchi: Reuse of Business Process for Business Application Development, Proc. of Pacific Rim Knowledge Acquisition Workshop (PKAW2002), pp. 73-86 (2002).
- [82] H. Abe and T. Yamaguchi: Evaluating the Automatic Composition of Inductive Applications Using StatLog Repository of Data Set, In Proc. of Pacific Rim Knowledge Acquisition Workshop (PKAW2002) pp. 144-158 (2002).
- [83] N. Izumi, T. Yamaguchi: Semantic Integration of E-Business Models, In Proc. of 7th Pacific-Rim International Conference on Artificial Intelligence (PRICAI2002), pp. 597 (2002).
- [84] M. Komori, H. Abe, T. Yamaguchi: A New Feature Selection Method Based on Dynamic Incremental Extension of Seed Features, In Proc. of Joint Conference on Knowledge-Based Software Engineering (JCKBSE2002) (2002).
- [85] N. Nakaya, M. Kurematsu, T. Yamaguchi: A Domain Ontology Development Environment Using a MRD and Text Corpus, In Proc. of Joint Conference on Knowledge-Based Software Engineering (JCKBSE2002) (2002).
- [86] K. Terai, M. Sawai, N. Sugiura, N. Izumi, T. Yamaguchi: Building Business Application Development Environment With Multi-Grain-Size Repository, In Proc. of Joint Conference on Knowledge-Based Software Engineering (JCKBSE2002) (2002).
- [87] H. Hatazawa, H. Abe, M. Komori, Y. Tachibana and T. Yamaguchi: Knowledge Discovery Support from a Meningoencephalitis Dataset Using an AutomaticComposition Tool for Inductive Applications, JSAI KDD Challenge 2001, pp. 9-17 (2001).
- [88] N. Izumi and T. Yamaguchi: Supporting Development of Software Agents by Integrat-



- ing Heterogeneous Repositories Based on Ontologies, 5th International Conference on Autonomous Agents, WS on Ontologies and Agent Systems (OAS2001)(<http://CEUR-WS.org/Vol-52/>), pp. 53-60,(2001).
- [89] N. Izumi and T. Yamaguchi: Supporting Development of Business Applications Based on Ontologies, International Conference on Enterprise Information Systems, pp. 893-897 (2001).
- [90] T. Yamaguchi: Acquiring Conceptual Relationships from Domain-Specific Texts, 17th International Joint Conference on Artificial Intelligence, WS on Ontology Learning, pp. 14-19 (2001).
- [91] N. Izumi and T. Yamaguchi: Building Business Applications by Integrating Heterogeneous Repositories Based on Ontologies, 17th International Joint Conference on Artificial Intelligence, WS on Ontologies and Information Sharing, pp. 166-173 (2001).
- [92] M. Kuremtsu, N. Nakaya and T. Yamaguchi: Acquiring Conceptual Relationships from a MRD and Text Corpus, ECML/PKDD-2001, WS on Semantic Web Mining, pp. 11-26 (2001).
- [93] N. Izumi and T. Yamaguchi: Development Support of E-Commerce Applications by Integrating Heterogeneous Repositories Based on Ontologies, International Conference on Electric Commerce 2001 (2001).
- [94] M. Ohsaki and T. Sugiyama: A Research on Edge Detector Applications and Definition of Edge Quality, 1st International Symposium on Measurement, Analysis and Modeling of Human Functions (ISHF'01), pp. 322-327 (2001).
- [95] M. Ohsaki, S. Sakamoto and H. Takagi: Development and Evaluation of an IEC Fitting System for Hearing Aids, 17th International Congress on Acoustics (ICA2001), 5A.15.02 (2001).
- [96] E. Suzuki, M. Gotoh and Y. Choki: Bloomy Decision Tree for Multi-Objective Classification, *Principles of Data Mining and Knowledge Discovery, Lecture Notes in Artificial Intelligence 2168 (PKDD)*, Springer-Verlag, pp. 436-447 (2001).
- [97] E. Suzuki: Worst-Case Analysis of Rule Discovery, *Discovery Science, Lecture Notes in Artificial Intelligence 2226 (DS)*, Springer-Verlag, pp. 365-377 (2001). (erratum: <http://www.slab.dnj.ynu.ac.jp/erratumds2001.pdf>)
- [98] N. Zhong, M. Ohshima and S. Ohsuga: Peculiarity Oriented Mining and Its Application for Knowledge Discovery in Amino-acid Data, *Advances in Knowledge Discovery and Data Mining, Lecture Notes in Artificial Intelligence 2035 (PAKDD)*, D. Cheung, G. J. Williams and Q. Li (eds.), Springer-Verlag, pp. 260-269 (2001).
- [99] J. Wu and N. Zhong: An Investigation on Human Multi-Perception Mechanism by Cooperatively Using Psychometrics and Data Mining Techniques, *Proc. 5th World Multi-Conference on Systemics, Cybernetics and Informatics (SCI-01)*, in Invited Session on Multimedia Information: Managing and Processing, Vol. X, pp. 285-290 (2001).

- [100] N. Zhong, Y.Y. Yao, M. Ohshima and S. Ohsuga: Interestingness, Peculiarity and Multi-Database Mining, *Proc. 2001 IEEE International Conference on Data Mining (IEEE ICDM'01)*, IEEE Computer Society Press, pp. 566-573 (2001).
- [101] Y. Yamada and E. Suzuki: Toward Knowledge-Driven Spiral Discovery of Exception Rules, *Proc. 2002 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Vol. 2, pp. 872-877 (2002).
- [102] S. Inatani and E. Suzuki: Data Squashing for Speeding up Boosting-Based Outlier Detection, *Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence 2366 (ISMIS)*, Springer-Verlag, pp. 601-611 (2002).
- [103] F. Takechi and E. Suzuki: Finding an Optimal Gain-Ratio Subset-Split Test for a Set-Valued Attribute in Decision Tree Induction, *Proc. Nineteenth International Conference on Machine Learning (ICML)*, pp. 618-625 (2002).
- [104] Y. Choki and E. Suzuki: Iterative Data Squashing for Boosting Based on a Distribution-Sensitive Distance, *Principles of Data Mining and Knowledge Discovery, Lecture Notes in Artificial Intelligence 2431 (PKDD)*, Springer-Verlag. pp. 86-98 (2002).
- [105] M. Narahashi and E. Suzuki: Subspace Clustering Based on Compressibility, *Discovery Science, Lecture Notes in Artificial Intelligence (DS)*, Springer-Verlag (2002, accepted for publication).
- [106] N. Zhong and J. Z. Dong: Mining Interesting Rules in Meningitis Data by Cooperatively Using GDT-RS and RSBR, *Advances in Knowledge Discovery and Data Mining, Lecture Notes in Artificial Intelligence 2336 (PAKDD)*, M. S. Chen, P. S. Yu and B. Liu (eds.), Springer-Verlag, pp. 405-416 (2002).
- [107] N. Zhong, Y. Matsui, T. Okuno and C. Liu: Framework of A Multi-Agent KDD System, *Proc. Third International Conference on Intelligent Data Engineering and Automated Learning, Lecture Notes in Computer Science 2412 (IDEAL)*, Springer-Verlag (2002, accepted for publication).
- [108] N. Zhong, Y. Y. Yao, J. Z. Dong and S. Ohsuga: Gastric Cancer Data Mining with Ordered Information, *Proc. Third International Conference on Rough Sets and Current Trends in Computing, Lecture Notes in Artificial Intelligence (RSCTC)*, Springer-Verlag (2002, accepted for publication).
- [109] T. Kudo and Y. Matsumoto: Chunking with Support Vector Machines, Proceedings of the Second Meeting of North American Chapter of Association for Computational Linguistics (NAACL), pp. 192-199, Pittsburgh (2001).
- [110] H. Takamura, H. Yamada, T. Kudoh, K. Yamamoto and Y. Matsumoto: Ensemble based on Feature Space Restructuring with Application to WSD, Proceedings of the Sixth Natural Language Processing Pacific Rim Symposium (NLPRS2001), Tokyo, pp. 41-48 (2001).
- [111] T. Nakagawa, T. Kudoh and Y. Matsumoto: Unknown Word Guessing and Part-of-Speech Tagging Using Support Vector Machines, Proceedings of the Sixth Natural Lan-

- guage Processing Pacific Rim Symposium (NLPRS2001), Tokyo, pp. 325-331 (2001).
- [112] T. Nakagawa, T. Kudo and Y. Matsumoto: Revision Learning and its Application to Part-of-Speech Tagging, Proceedings of 40th Annual Meeting of Association for Computational Linguistics (ACL-02), Philadelphia, USA, pp. 497-504 (2002).
- [113] T. Nakagawa and Y. Matsumoto: Detecting Errors in Corpora Using Support Vector Machines, Proceedings of the 19th International Conference on Computational Linguistics (COLING 2002), pp. 709-715 (2002).
- [114] H. Takamura and Y. Matsumoto: Two-dimensional Clustering for Text Categorization, Proceedings of Sixth Conference on Natural Language Learning (CoNLL-2002), pp. 29-35 (2002).
- [115] T. Kudo and Y. Matsumoto: Japanese Dependency Analysis using Cascaded Chunking, Proceedings of Sixth Conference on Natural Language Learning (CoNLL-2002), pp. 63-69 (2002).
- [116] M. Shimbo, H. Yamada and Y. Matsumoto: Using Syntactic Dependency Information for Classification of Technical Terms, PKAW-2002, pp. 131-143 (2002).
- [117] S. Tsumoto: Medical Knowledge Discovery in Hospital Information System, *Proceedings of SPIE vol.4384, Data Mining and Knowledge Discovery: Theory, Tools, and Technology III*, Orlando, pp. 229-237 (2001).
- [118] S. Tsumoto: Statistical Extension of Rough Set Rule Induction, *Proceedings of SPIE vol.4384, Data Mining and Knowledge Discovery: Theory, Tools, and Technology III*, Orlando, pp. 180-188 (2001).
- [119] S. Tsumoto: Discovery of Temporal Knowledge in Medical Time-Series Databases Using Moving Average, Multiscale Matching, and Rule Induction, *5th European Conference on Principles of Knowledge Discovery in Databases (PKDD2001)*, Freiburg, pp. 448-459 (2001).
- [120] S. Tsumoto: Mining Positive and Negative Knowledge in Clinical Databases Based on Rough Set Model, *5th European Conference on Principles of Knowledge Discovery in Databases (PKDD2001)*, Freiburg, pp. 460-471 (2001).
- [121] S. Hirano and S. Tsumoto: A Knowledge-Oriented Clustering Technique Based on Rough Sets, *25th IEEE International Computer Software and Applications Conference (CompSAC2001)*, Chicago, pp. 632-637 (2001).
- [122] S. Hirano and S. Tsumoto: Indiscernibility Degrees of Objects for Evaluating Simplicity of Knowledge in the Clustering Procedure, *2001 IEEE International Conference on Data Mining (ICDM01)*, San Jose, pp. 211-217 (2001).
- [123] S. Tsumoto: Temporal Knowledge Discovery in Time-Series Medical Databases based on Fuzzy and Rough Reasoning, *9th International Fuzzy Systems Association World Congress (IFSA'01)*, Vancouver, (CD-ROM) (2001).
- [124] S. Hirano, S. Tsumoto, T. Okuzaki, and Y. Hata: A Clustering Method Based on

- Rough Set Theory and Its Application to Knowledge Discovery in the Medical Database, *MEDINFO*, London, vol. 10, pp. 206-210 (2001).
- [125] S. Hirano and S. Tsumoto: Analysis of Time-series Medical Databases Using Multiscale Structure Matching and Rough Sets-based Clustering Technique, *2001 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE01)*, Melbourne, pp. 1547-1550 (2001).
- [126] S. Tsumoto, S. Hirano, A. Yasuda, and K. Tsumoto: Analysis of Amino Acid Sequences by Statistical Technique, *4th Conference on Computational Biology and Genome Informatics (CBGI-02)*, Association of Intelligent Machinery, Durham, pp. 1169-1173 (2002).
- [127] S. Tsumoto Comparison of Induced Rules based on Likelihood Estimation. *SPIE AeroSense 2002*, Orlando, vol. 4730, pp. 192-201 (2002).
- [128] S. Tsumoto Statistical Test for Rough Set Approximation. *SPIE AeroSense 2002*, Orlando, vol. 4730, pp. 410-421 (2002).
- [129] X. Sun, S. Hirano, and S. Tsumoto: Evaluation of Similarity Measures for Analysis of Databases on Laboratory Examinations, *SPIE AeroSense 2002*, Orlando, vol. 4730, pp. 324-333 (2002).
- [130] S. Tsumoto: Statistical Evidence for Rough Set Analysis, *2002 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE02)*, Hawaii, pp. 757-762 (2002)
- [131] S. Hirano, X. Sun, and S. Tsumoto: Dealing with Multiple Types of Expert Knowledge in Medical Image Segmentation: A Rough Sets Style Approach, *2002 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE02)*, Hawaii, pp. 884-889 (2002).
- [132] S. Tsumoto Information Refinement and Revision for Medical Expert System *AAAI Spring Symposium on Information Refinement and Revision for Decision Making: Modeling for Diagnostics, Prognostics, Prediction*, AAAI Press (Technical Report SS-02-03), Menlo Park, pp. 115-121 (2002).
- [133] S. Tsumoto: Automated Discovery of Decision Rule Chains Using Rough Sets and Medical Diagnostic Model. *International Symposium on Methods of Information Systems (ISMIS 2002)*, LNCS 2366, Springer Verlag, pp. 321-332 (2002).
- [134] T.Y. Lin and S. Tsumoto: Neighborhood Systems and Qualitative Fuzzy Sets. *Ninth International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU2002)*, Annecy, pp. 259-265 (2002).
- [135] S. Tsumoto Extraction of Hierarchical Decision Rules from Medical Databases using Rough Set Model. *Ninth International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU2002)*, Annecy, pp. 273-280 (2002).
- [136] T. Okada: Characteristic Substructures and Properties in the Chemical Carcinogenicity Studied by the Cascade Model, *International workshop on Predictive Toxicology Challenge 2001*, Freiburg (2001).
- [137] T. Okada: Datascape Survey using the Cascade Model, to appear in *Discovery Science*

- 2002, Springer-Verlag(2002).
- [138] T. Okada: Organizing Rules Yields Datascape, *Pacific Knowledge Acquisition Workshop 2002*, pp .1-14 (2002).
  - [139] T. Niwa, K. Fujikawa, K. Tanaka and M. Oyama: Visual Data Mining Using a Constellation Graph, *International Workshop on Visual Data Mining*, Freiburg (2001).
  - [140] Y. Takahashi and M. Konji: Visualization of Massive Data in Molecular Space and Similar Structure Searching, *6th China-Japan Joint Symposium on Drug Design and Development*, pp. 1-3, Dalian, China (2001).
  - [141] H. Kato, Y. Takahashi and H. Abe: Development of Automated Identification System for Three-dimensional Protein Motifs, *10th German-Japanese Workshop on Chemical Information*, pp. 6, Potsdam, Germany (2001).
  - [142] H. Kato, T. Tadokoro, Y. Takahashi and H. Abe: Construction of a Three-Dimensional Protein Motif Dictionary, *Fifth Australia/Japan Symposium on Drug Design and Development*, pp. 53-55, Nara (2002).
  - [143] S. Yamada and Y. Ohsawa: Information Gathering of Web pages to Guide Concept Understanding *Posters Proc. the World Wide Web Conference (WWW10)* Hong Kong, (2001).
  - [144] Y. Ohsawa, N. Matsumura and M. Ishizuka: Discovering Topics to Enhance Communities' Creation from Links to the Future, *Posters Proc. the World Wide Web Conference (WWW10)* Hong Kong, (2001).
  - [145] N. Matsumura and Y. Ohsawa: Future Deirections of Communities on the Web, *The First Workshop on Chance Discoveries, Japanese Society of Artificial Intelligence*Matsue, Japan (2001).
  - [146] Hiroshi Taira, Yasuhiro Sakata, Y. Ohsawa and M. Ishizuka: AreaView2001: A new WWW organization system with KeyGraph Technology *The First Workshop on Chance Discoveries, Japanese Society of Artificial Intelligence*Matsue, Japan (2001).
  - [147] Y. Matsuo and Y. Ohsawa: A Document as a Small World, *The First Workshop on Chance Discoveries, Japanese Society of Artificial Intelligence*Matsue, Japan (2001).
  - [148] Y. Nara and Y. Ohsawa: A Method for Discovering Seeds of Consensus Applied to Family Risk Perceptions *The First Workshop on Chance Discoveries, Japanese Society of Artificial Intelligence*Matsue, Japan (2001).
  - [149] Y. Matsuo, Y. Ohsawa and M. Ishizuka: Small-World as Asserting Structure of Document, *Proc. The Fifth Multi- Conference on Systems, Cybernetics and Informatics (SCI2001)*, Orland, Florida USA (2001).
  - [150] N. Matsumura, Y. Ohsawa and M. Ishizuka: Future Directions of Communities on the Web, *Proc. The Fifth Multi-Conference on Systems, Cybernetics and Informatics (SCI2001)*, Orland, Florida USA (2001).

- [151] Y. Ohsawa and Y. Nara: Family Perceptions of Risks and Opportunities - Results from Questionnaires to Citizens -, The Fifth Multi-Conference on Systems, Cybernetics and Informatics (SCI2001), Orland, Florida USA (2001).
- [152] Y. Nara and Y. Ohsawa: Family Affection to Children Visualized on the Co-occurring of Questionnaire Answers, *Proc. The Fifth Multi-Conference on Systems, Cybernetics and Informatics (SCI2001)*, Orland, Florida USA (2001).
- [153] Fumiko Yoshikawa and Y. Ohsawa: Text Analysis by Mapping Information Flow, *Proc. The Fifth Multi-Conference on Systems, Cybernetics and Informatics (SCI2001)*, Orland, Florida USA (2001).
- [154] Y. OHSAWA and Y. Nara: Decision Trees as a Model of Chance Perception, *Joint 9th IFSA Congress and 20th NAFIPS International Conference*, Vancouver (2001).
- [155] H. Fukuda and Y. Ohsawa: Discovery of Rare Essential Food by Community Navigation with KeyGraph - An introduction to Data-based Community Marketing - *Proc. KES 2001 (from IOS press)* (2001).
- [156] N. Matsumura, Y. Ohsawa and M. Ishizuka: Discovery of Emerging Topics by Co-citation Graph on the Web *Proc. KES 2001 (from IOS press)* (2001).
- [157] Y. Matsuo, Y. Ohsawa and M. Ishizuka: Discovering Hidden Relation behind a Link *Proc. KES 2001 (from IOS press)* (2001).
- [158] Y. Ohsawa and Y. Nara: Discovery of Virtual Behaviors as Signs of Real Behaviors *Proc. KES 2001 (from IOS press)* (2001).
- [159] Y. Ohsawa and N. Matsumura: Discovering Seeds of New Interest Spread from Premature Pages Cited by Multiple Communities *Proc. Web Intelligence (LNAI2198 from Springer Verlag)* (2001).
- [160] N. Matsumura, Y. Ohsawa and M. Ishizuka: Discovery of Emerging Topics between Communities on WWW *Proc. Web Intelligence (LNAI2198 from Springer Verlag)* (2001).
- [161] Y. Matsuo, Y. Ohsawa and M. Ishizuka: Average-clicks: A New Measure of Distance on the World Wide Web *Proc. Web Intelligence (LNAI2198 from Springer Verlag)* (2001).
- [162] Y. Matsuo, Y. Ohsawa and M. Ishizuka: KeyWorld: Extracting Keywords from a Document as a Small World, *Proc. The Forth International Conference on Discovery Science (An LNAI from Springer Verlag)* (2001).
- [163] N. Matsumura, Y. Ohsawa and M. Ishizuka: Knowledge Navigation on Visualizing Complementary Documents *Proc. The Forth International Conference on Discovery Science (An LNAI from Springer Verlag)* (2001).
- [164] Invited Talk, in the Satellite Session of the 53rd Biennial Session of the International Statistics Institute (ISI2001) "Image and Discover Chance from Rare Events" (2001).
- [165] Y. Matsuo, Y. Ohsawa and M. Ishizuka, A Psychological Metric of Distance on the WWW *Poster, The Eleventh Conf. World Wide Web (WWW11)* (2002).

- [166] N. Matsumura, Y. Ohsawa and M. Ishizuka, Influence Diffusion Model in Text-Based Communiation *Poster, The Eleventh Conf. World Wide Web (WWW11)* (2002).
- [167] N. Matsumura, Y. Ohsawa and M. Ishizuka Mining Mailing List Archives, *Poster, The Eleventh Conf. World Wide Web (WWW11)* (2002).
- [168] Y. Ohsawa, H. Soma, et al, Featuring Web Communities based on Word Co-occurrence Structure of Communications, *Proceedings of the Eleventh Conf. World Wide Web (WWW11)*, ACM, 2002
- [169] N. Matsumura, Y. Ohsawa and M. Ishizuka: Automatic Indexing Based on Term Activities, *Poster, The Seventh Pacific Rim International Conference on Artificial Intelligence (PRICAI-2002)* (to appear)
- [170] N. Matsumura, Y. Ohsawa and M. Ishizuka: Mining and Characterizing Opinion Leaders from Threaded Online Discussions, *Sixth International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies (KES2002)* (2002). (to appear).
- [171] M. Mizuno, H. Shoji, Y. Osawa, Y. Matsuo, N. Matsumura and Y. Miyake: Toward a Chance Discovery-Oriented Recomemender System: A Prototype, *Sixth International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies (KES2002)* (2002). (to appear).
- [172] N. Matsumura, Y. Ohsawa and M. Ishizuka: Profiling of Participants in Online-Community, Chance Discovery Workshop on the Seventh Pacific Rim International Conference on Artificial Intelligence (PRICAI-2002) (2002).
- [173] Y. Ohsawa: Genetic Words - Building Blocks as Context - *Chance Discovery Workshop on the Seventh Pacific Rim International Conference on Artificial Intelligence (PRICAI-2002)* (2002).
- [174] S. Kurahashi and T. Terano: Can We Control Information Free Riders? Analyzing Communal Sharing Norms via Agent-based Simulation. (CASOS 2001), pp. 94-96 (2001).
- [175] H. Matsui, I. Ono, H. Sato, H. Deguchi, T. Terano, H. Kita and Y. Shinozawa: Learning Economics Principles from Bottom. (CASOS 2001), pp. 97-99 (2001).
- [176] S. Kurahashi and T. Terano: Analyzing Norm Emergence in Communal Sharing via Agent-based Simulation. (AESCS-2001), pp. 27-34 (2001).
- [177] H. Deguchi, T. Terano, K. Kurahashi, T. Yuzawa, S. Hashimoto, H. Matsui, A. Sashima and T. Kaneda: Virtual Economy Simulation and Gaming -An Agent Based Approach-. (AESCS-2001), pp. 169-185 (2001).
- [178] Y. Katshumata, S. Kurahashi and T. Terano: Hybridizing Bayesian Optimization and Tabu Search for Multimodel Fancrions. 2001 Genetic and Evolutionary Computation Conference (GECCO-2001) Late Breaking Papers, pp. 227-333 (2001).
- [179] K. Taniguchi, S. Kurahashi and T. Terano: Managing Information Complexity in a Supply Chain Model Agent-Based Genetic Programming. 2001 Genetic and Evolutionary Computation Conference (GECCO-2001) Late Breaking Papers, pp. 413-420 (2001).

- [180] E. Murakami and T. Terano: Collaborative Filtering for a Distributed Smart IC Card System. Intelligent Agents: Specification, Modeling, and Applications, 4th Pacific Rim International Workshop on Multi-Agents, PRIMA 2001, Taipei, Taiwan, July 2001, Proceedings, Springer LNAI 2132, pp. 183-197 (2001).
- [181] T. Terano, Y. Shiozawa, H. Deguchi, H. Kita, H. Matsui, H. Sato and I. Ono: U-Mart: An Artificial Market to Bridge the Studies on Economics and Multiagent Systems. Proc. PRIMA'2001, 4th Pacific Rim International Workshop on Multi-agents, pp. 371-385, Taipei, July 28-29 (2001).
- [182] K. Takadama, T. Terano and K. Shimohara: Learning Classifier Systems Meet Multiagent Environments, Pier Luca Lanzi, Wolfgang Stolzmann, Stephaan W. Wilson (Eds.): Advances in Learning Classifier Systems, Third International Workshop, IWLCS 2000, Paris, France, September 2000. Springer LNAI 1996, pp. 192-212 (2001).
- [183] S. Kuraishi and T. Terano: Inverse Simulation and Genetics-Based Validation for Social Interaction Analysis via Multiagents: UCLA Computational Social Sciences Conf. in Lake Arrowhead No. 100 (2002).
- [184] Y. Katsumata, S. Kurahashi and T. Terano: We Need Multiple Solutions for Electric Equipments Configuration in a Power Plant - Applying Bayesian Optimization Algorithm with Tabu Search - WCCI 2002 (The 2002 IEEE World Congress on Computational Intelligence), pp. 1402-1407 (2002).
- [185] K. Tsuda, M. Kessoku and T. Terano: Collecting Customers' Preference on The WEB by Document Analysis Lecture Notes in Operations Research 4 (ISORA'2002), pp. 294-301 (2002).
- [186] T. Terano, S. Kurahashi and Y. Katsumata: Hybridizing Genetic Algorithms and Tabu Search For Multimodal Function Optimization Lecture Notes in Operations Research 4 (ISORA'2002), pp. 283-293 (2002).
- [187] T. Terano, H. Deguchi, H. Kita, Y. Shiozawa, H. Sato, H. Matsui and I. Ono: UMIE2002: U-Mart International Experiment 2002-What We have Learnt from the Virtual Market- (2002).
- [188] H. Takahashi, T. Terano: Bridging GARCH Model and Prospect Theory in Financial Market Behaviors Via Agent-Based Simulation. Eighth International Conference of The Society for Computational Economics (2002).

## 国内発表

- [1] 小野田 崇, 村田 博士, 山田 誠二: 情報検索における能動学習, 人工知能学会第 56 回「知識ベースシステム」研究会, pp. 93-98 (2002).
- [2] 高間 康史: WWW 文書集合からの時系列的話題情報の抽出・可視化手法の提案, 人工知能学会第 56 回「知識ベースシステム」研究会, pp. 111-116 (2002).



- [3] 高間 康史, 廣田 薫: Web 上の時系列文書集合族を対象とした情報可視化システム, 情報処理学会第 64 回全国大会, pp. 4-643-646, (2002).
- [4] 高間 康史: 情報検索における対象知識獲得支援システムの構築, 平成 14 年電気学会全国大会講演論文集 [4], pp. 505-506, (2002).
- [5] 山田 誠二, 岡部 正幸: 関係学習を用いた対話的 Web ページフィルタリング, 第 54 回人工知能学会「知識ベースシステム」研究会, pp. 67-72 (2001).
- [6] 中井 有紀, 山田 誠二: Web ページにおける部分情報の更新モニタリング, 第 54 回「知識ベースシステム」研究会, pp. 73-78 (2001).
- [7] 岡部 正幸, 山田 誠二: 関係学習を用いたフィルタ生成による対話的 Web ページ検索, 第 61 回情報学基礎研究会, pp. 197-204 (2001).
- [8] 高間 康史, 廣田 薫: 可塑的クラスタリングに基づく WWW 情報可視化システム, 電気学会システム・制御研究会, SC-01-15, pp. 1-6 (2001).
- [9] 高間 康史, 廣田 薫: 免疫ネットワークを用いたキーワード集合抽出の情報可視化システムへの応用, 第 17 回ファジィシステムシンポジウム(日本ファジィ学会), pp. 781-782 (2001).
- [10] 高間 康史, 廣田 薫, WWW 上の情報収集/可視化のための免疫ネットワークを用いたクラスタリング, 第 46 回人工知能基礎論研究会/第 54 回知識ベースシステム研究会(合同研究会)(人工知能学会), pp. 61-66 (2001).
- [11] 北村 泰彦: Web とエージェントと教育システム. 人工知能学会第 16 回 AI シンポジウム「e-learning の intelligent 化に向けて」, 人工知能学会研究会資料, SIG-J-A102-9, pp. 47-48 (2001).
- [12] 北村 泰彦: WWW 情報統合からエージェント統合へ. 情報処理学会連続セミナー 2001「21 世紀のネットサービス社会」第 4 回「サービスプラットフォーム技術」, pp. 1-2 (2001).
- [13] 北村 泰彦: アクティブ情報収集システムに関する検討, 人工知能学会人工知能基礎論・知識ベースシステム研究会, SIG-FAI/KBS-J-14 (2001).
- [14] 辻本 秀樹, 北村 泰彦, 辰巳 昭治: Wizard of Oz 法を使ったマルチキャラクタエージェントインタフェースの評価. 日本ソフトウェア科学会第 10 回マルチエージェントと協調計算ワークショップ (2001).
- [15] 阪本 俊樹, 北村 泰彦, 辰巳 昭治: 競争型推薦システム Recommendation Battlers とその挙動. 人工知能学会人工知能基礎論・知識ベースシステム研究会, SIG-FAI/KBS-J-09 (2001).
- [16] 植村 渉, 辰巳 昭治, 北村 泰彦: 強化学習を用いた 2D メッシュ結合型マルチコンピュータでの耐故障性を持つ適応経路設定, 電子情報通信学会情報・システムソサイエティ大会, D-10-7 (2001).
- [17] 山田 晃弘, 小久保 卓, 北村 泰彦: マルチキャラクタインタフェースを用いた Web 情報統合. 人工知能学会第 15 回全国大会, 1F1-05 (2001).
- [18] 阪本 俊樹, 回り道 康博, 北村 泰彦, 辰巳 昭治: マルチキャラクタを用いた競争型情報

- 推薦システム．人工知能学会第15回全国大会，1F1-01 (2001)．人工知能学会全国大会優秀論文賞受賞
- [19] 阪本 俊樹，回り道 康博，北村 泰彦，辰巳 昭治：マルチエージェントによる競争型情報推薦システム．電子情報通信学会関西支部第6回学生会研究発表講演会，D-1 (2001)．
- [20] 北村 泰彦，鈴東 智，辰巳 昭治：動的迷路における観測と経路探索手法，人工知能学会知識ベースシステム研究会，SIG-KBS-A201-33，情報処理学会知能と複雑系研究会128-33 (2002)．
- [21] 植村 渉，辰巳 昭治，北村 泰彦：強化学習 Profit Sharing への学習確定度の導入，電子情報通信学会総合大会，D-8-14 (2002)．
- [22] 今津 剛行，辰巳 昭治，北村 泰彦，平井 誠：主体別動作イベントに基づく物語文の要約抽出法，電子情報通信学会総合大会，D-5-10 (2002)．
- [23] 平山 勝敏，北村 泰彦：アクティブ情報統合のための動的分散制約充足プロトコル，情報処理学会第128回知能と複雑系研究会および人工知能学会第56回知識ベースシステム研究会の合同研究会資料，pp. 175-180 (2002)．
- [24] 沼尾 正行：Global intelligence による知識流通，情報処理学会知能と複雑系研究会，Vol. 01-ICS-124，pp. 1-8 (2001)．
- [25] ナッティー・チョラウイト，沼尾 正行：知識流通における紙面レイアウトの役割とそれに基づく自動タグ付け，情報処理学会知能と複雑系研究会，Vol. 01-ICS-124，pp. 17-24 (2001)．
- [26] 伊藤 雄介，吉田 匡史，沼尾 正行：口コミ支援システムの実験，情報処理学会知能と複雑系研究会，Vol. 01-ICS-124，pp. 9-16 (2001)．
- [27] 老川 正志，沼尾 正行：帰納論理プログラミングによる嗜好情報の獲得とナビゲーションシステム，人工知能学会全国大会（第15回）論文集 (2001)．
- [28] 高木 将一，中村 啓佑，沼尾 正行：機械学習の手法を用いた感性の抽出と作曲・編曲への応用，人工知能学会全国大会（第15回）論文集 (2001)．
- [29] 五十嵐 建平，大田 佳宏，横山 茂樹，沼尾 正行：データマイニングにおけるXMLを用いたデータ構造の変形，人工知能学会全国大会（第15回）論文集 (2001)．
- [30] 伊藤 雄介，吉田 匡史，沼尾 正行：多くの人の評価を経て情報が吟味される口コミ支援システム，人工知能学会全国大会（第15回）論文集 (2001)．
- [31] C. Nattee and M. Numao: Document analysis and recognition using ILP and the winnow algorithm，人工知能学会全国大会（第15回）論文集 (2001)．
- [32] 沼尾 正行，高木 将一，中村 啓佑：ユーザの感性に合わせた自動編曲及び作曲，情報処理学会音楽情報科学研究会，Vol. 2001-MUS-41，pp. 49-54 (2001)．
- [33] 森山 甲一，沼尾 正行：自己の報酬を操作する学習エージェントの構築，人工知能学会第45回人工知能基礎論研究会，第SIG-FAI-A101巻，pp. 15-20，(2001)．
- [34] 森山 甲一，沼尾 正行：環境状況の変化に応じて自己の報酬を操作する学習エージェントの構築，<http://www-kasm.nii.ac.jp/macc2001-proceedings/MACC2001-15.pdf>．

- MACC2001, (2001).
- [35] 吉田 匡志, 伊藤 雄介, 沼尾 正行: 口コミによる分散型情報収集システム – WAVE を起こそう – Word-of-mouth-Assisting Virtual Environment, <http://www-kasm.nii.ac.jp/macc2001-proceedings/MACC2001-10.pdf>, MACC2001 (2001).
  - [36] 沼尾 正行, 吉田 匡志, 伊藤 雄介: 口コミに基づく情報収集とデータ前処理, 人工知能学会第 46 回人工知能基礎論/第 54 回知識ベースシステム研究会, 第 SIG-FAI/KBS-J 巻, pp. 47-54 (2001).
  - [37] T. N. Tran and M. Numao: Mining biomedical literature by making use of existing databases, 人工知能学会第 46 回人工知能基礎論/第 54 回知識ベースシステム研究会, 第 SIG-FAI/KBS-J 巻, pp. 41-46 (2001).
  - [38] トアンナム・チャン, 沼尾 正行: 生物医学文献データベースを利用する医療データマイニング, 情報処理学会第 128 回知能と複雑系研究会および人工知能学会第 56 回知識ベースシステム研究会の合同研究会資料, p.29-34 (2002) .
  - [39] 沼尾 正行: XML 変形または関係表現の変形による前処理, 情報処理学会第 128 回知能と複雑系研究会および人工知能学会第 56 回知識ベースシステム研究会の合同研究会資料, p.79 (2002) .
  - [40] 山田 有吉, 木村 亮, 矢野 幸司, 沼尾 正行: XML 変形を用いた前処理の事例研究, 情報処理学会第 128 回知能と複雑系研究会および人工知能学会第 56 回知識ベースシステム研究会の合同研究会資料, p.81-86 (2002) .
  - [41] 森山 甲一, 沼尾 正行: 社会ジレンマを解決する内部評価生成法の分析, 情報処理学会第 128 回知能と複雑系研究会および人工知能学会第 56 回知識ベースシステム研究会の合同研究会資料, p.127-132 (2002) .
  - [42] C. Nattee and M. Numao: Weight Updating Algorithm for Multiple-Instance Learning: An Empirical Study, 情報処理学会第 128 回知能と複雑系研究会および人工知能学会第 56 回知識ベースシステム研究会の合同研究会資料, p.133-138 (2002) .
  - [43] 伊藤 雄介, 吉田 匡志, 沼尾 正行: 口コミ支援システム WAVE の新機能, 情報処理学会第 128 回知能と複雑系研究会および人工知能学会第 56 回知識ベースシステム研究会の合同研究会資料, p.151-156 (2002) .
  - [44] 河野 浩之: 位置データ問合せ処理のための空間インデックス手法の検討, 第 45 回システム制御情報学会研究発表講演会 (2001).
  - [45] 中辻 真, 川原 稔, 河野 浩之: ピアツーピアネットワークにおけるトピック主導型検索手法の提案, 知能と複雑系研究会, pp. 47-54 (2001).
  - [46] 南 卓朗, 田名部 淳, 河野 浩之: 空間インデックスを用いた移動オブジェクト管理システムの構成と性能比較, 情処研報 Vol. 2001, No. 70, DBS-125, pp. 225-232 (2001).
  - [47] 河野 浩之: 位置情報システムにおける空間データ利用に関する検討, 人工知能学会 FAI-46&KBS-54 研究会, pp. 159-164 (2001).

- [48] 岸 浩史, 田名部 淳, 河野 浩之: -Tree による時空間 OLAP 技術の交通データへの適用, 第 13 回データ工学ワークショップ, (2002).  
(<http://www.ieice.org/iss/de/DEWS/proc/2002/proceedings.html>)
- [49] 河野 浩之, 田名部 淳, 岸 浩史: 交通管制データに対する OLAP 技術, 電子情報通信学会総合大会, A-17-34, (2002).
- [50] 河野 浩之: 交通データウェアハウスにおける OLAP, 第 128 回 知能と複雑系研究会, (2002).
- [51] 乾 岳史, 櫻井 成一郎: マルチエージェントによる組織形成に関する基礎的研究, 情報処理学会, 予稿 (2001-GI-5), pp. 23-30 (2001).
- [52] 白壁 啓吾, 乾 岳史, 櫻井 成一郎: マルチエージェント系における強化結果に対する解釈について, 情報処理学会, 予稿 (2001-ICS-124), pp. 55-60 (2001).
- [53] 若月 謙太郎, 櫻井 成一郎: リンク解析による WWW ページ群の発見, 人工知能学会, 予稿 (2001-FAI-46, 2001-KBS-54), pp. 109-114 (2001).
- [54] 櫻井 成一郎: ILP による Web ページ作成者の意図抽出について, 人工知能学会知識ベースシステム研究会, SIG-KBS-A201-19, 情報処理学会知能と複雑系研究会 128-19, pp. 99-103 (2002).
- [55] 市瀬 龍太郎, 武田 英明, 本位田 真一: インスタンスに基づく階層的知識源の統合, 人工知能学会 AI チャレンジ研究会, SIG-Challenge-0111, pp. 61-66, (2001).
- [56] R. Ichise, H. Takeda and S. Honiden: An Alignment Algorithm between Concept Hierarchies, NII Technical Report, NII-2001-001 (2001).
- [57] 武田 英明, 市瀬 龍太郎, 村田 剛志, 本位田 真一: 知識共生プロジェクト - ネットワーク情報の自律的生態系を目指して -, 情報処理学会知能と複雑系研究会, SIG-ICS-124, pp. 25-32 (2001).
- [58] 市瀬 龍太郎, 武田 英明, 本位田 真一: WWW における情報源に関する知識の共生, 情報処理学会知能と複雑系研究会, SIG-ICS-124, pp. 33-40 (2001).
- [59] 市瀬 龍太郎, 武田 英明, 本位田 真一: 他者の持つ知識を利用した階層的分類知識の獲得, 第 15 回人工知能学会全国大会, 1D1-06 (2001).
- [60] 武田 英明, 市瀬 龍太郎, 村田 剛志, 本位田 真一: 知識共生 新しい知識流通の基盤を目指して , 第 16 回人工知能学会全国大会論文集 (2002).
- [61] 西村 芳男, 鷺尾 隆, 元田 浩, 猪口 明博: 大量データからの誘導部分グラフデータの検索手法, 2001 年度人工知能学会全国大会資料 (第 15 回 ), 2D1-04 pp. 1-4 (2001).
- [62] 寺邊 正大, 鷺尾 隆, 元田 浩: 相関ルールに基づく属性生成手法 - 連続値属性を含むデータへの適用 -, 2001 年度人工知能学会全国大会資料 (第 15 回 ), 2D1-09, pp. 1-4 (2001).
- [63] 和田 卓也, 元田 浩, 鷺尾 隆: クラス分布の変化に追従するための Ripple Down Rules 法の拡張に関する実験, 2001 年度人工知能学会全国大会資料 (第 15 回 ), 1B2-03, pp. 1-4 (2001).

- [64] 鷺尾 隆, 元田 浩, 丹羽 雄二: 観測データからの第一原理に基づく連立方程式の発見, 2001 年度人工知能学会全国大会資料 (第 15 回), 2D1-03, pp. 1-4 (2001).
- [65] 藤原 啓成, 元田 浩, 鷺尾 隆: 専門家の事例判断のみを利用した RDR 知識ベース構築のための事例生成手法の評価実験, 第 46 回人工知能基礎論研究会, 第 54 回知識ベースシステム研究会 (合同研究会) 資料 (SIG-FAI/KBS-J-25), pp. 153-158 (2001).
- [66] 和田 卓也, 元田 浩, 鷺尾 隆: 環境変化への適応, 異種の知識源からの知識獲得を目的とした Ripple Down Rules 法の拡張に対する評価実験, 第 46 回人工知能基礎論研究会, 第 54 回知識ベースシステム研究会 (合同研究会) 資料 (SIG-FAI/KBS-J-26), pp. 165-170 (2001).
- [67] 松田 喬, 元田 浩, 鷺尾 隆: Graph-Based Induction の部分グラフ抽出能力の改良, 第 46 回人工知能基礎論研究会, 第 54 回知識ベースシステム研究会 (合同研究会) 資料 (SIG-FAI/KBS-J-30), pp. 185-187 (2001).
- [68] 松田 喬, 吉田 哲也, 元田 浩, 鷺尾 隆: グラフ構造に着目した肝炎データからの知識発見, 第 56 回人工知能学会知識ベース研究会資料 (SIG-KBS-A201), pp. 67-72 (2002).
- [69] 松田 喬, 吉田 哲也, 元田 浩, 鷺尾 隆: Graph-Based Induction による分類学習のための構造データからの属性構築, 2002 年度人工知能学会全国大会資料 (第 16 回), 1A4-03, pp. 1-4 (2002).
- [70] 和田 卓也, 吉田 哲也, 元田 浩, 鷺尾 隆: 柔軟な知識ベースシステム構築を目指す Description Length based RDR 法の人工データによるシミュレーション 2002 年度人工知能学会全国大会資料 (第 16 回), 2C1-04, pp. 1-4 (2002).
- [71] 藤原 啓成, 吉田 哲也, 元田 浩, 鷺尾 隆: 近傍事例のクラス情報のみを利用した RDR 知識ベース構築のための事例生成手法と評価, 2002 年度人工知能学会全国大会資料 (第 16 回), 2C1-05, pp. 1-4 (2002).
- [72] ワロドム ジアムサクン, 松田 喬, 元田 浩, 鷺尾 隆, 吉田 哲也: Graph-Based Induction を用いたグラフ構造データに対する分類器の構築, 2002 年度人工知能学会全国大会資料 (第 16 回), 2C1-05, pp. 1-4 (2002).
- [73] 鷺尾 隆: データ収集の新機軸とマイニングシステムの未来, 第 17 回 AI シンポジウム: 人工知能学会研究会資料 SIG-J-A103, pp. 65-70 (招待) (2002).
- [74] 西村 芳男, 鷺尾 隆, 吉田 哲也, 元田 浩: Apriori-based Graph Mining アルゴリズムの高速化, 人工知能学会第 56 回知識ベース研究会資料 SIG-KBS-A201, pp. 11-16 (2002).
- [75] 西村 芳男, 鷺尾 隆, 吉田 哲也, 元田 浩, 猪口 明博: Fast Apriori-based Graph Mining Algorithm (AGM) とその性能評価, 人工知能学会, 第 3 回 MYCOM21 (2002).
- [76] T.B. Ho, S. Kawasaki and D.D. Nguyen: Extracting Predictive Knowledge from Meningitis Data by Integration of Rule Induction and Association Mining, Proc. of International Workshop Challenge in KDD, JSAI Annual Conference, pp. 25-32 (2001).
- [77] T.B. Ho, T.D. Nguyen, D.D. Nguyen and S. Kawasaki: Visualization Support for User-Centered Model Selection in Knowledge Discovery and Data Mining 人工知能基礎論研究会, SIG-FAI/KBS-J-35, pp. 207-212 (2001).

- [78] S. Kawasaki, D.D. Nguyen, T.D. Nguyen and T.B. Ho: Study of Hepatitis Data by Visual Data Mining System D2MS, JSAI SIG-KBS-A201, Workshop Active Data Mining, Pusan, pp. 43-48, May (2002).
- [79] 矢田 勝俊: データマイニング技術とカスタマープロファイリング商業学会関西西部会 (2001).
- [80] 矢田 勝俊, 飯田 洋: ビジネス知識発見の研究展望とアクティブマイニング, 人工知能基礎論研究会, SIG-FAI/KBS-J-35, pp. 213-217 (2001).
- [81] 加藤 直樹, 羽室 行信, 矢田 勝俊: 購買履歴からのデータマイニング, 2001 年情報論的学習理論ワークショップ論文集, pp. 95-104 (2001).
- [82] 小森 麻央, 阿部 秀尚, 山口 高平: シーズ属性拡張に基づく属性選択法の提案と評価, 第 16 回人工知能学会全国大会, 1A4-01 (2002).
- [83] 肥塚 八尋, 和泉 憲明, 山口 高平: オントロジーを利用した学習者コミュニティ形成支援, 第 16 回人工知能学会全国大会, 2C1-01 (2002).
- [84] 上野 真由美, 不知 俊洋, 和泉 憲明, 山口 高平: 個人タスク支援のためのエージェントリポジトリの開発, 第 16 回人工知能学会全国大会, 2C1-03 (2002).
- [85] 中矢 尚美, 樽松 理樹, 山口 高平: 既存の情報リソースを活用するオントロジー構築支援エディタ, 第 16 回人工知能学会全国大会, 3F2-05 (2002).
- [86] 寺井 公一, 杉浦 直樹, 澤井 雅彦, 和泉 憲明, 山口 高平: 多粒度リポジトリに基づくビジネスアプリケーション構築支援環境, 第 16 回人工知能学会全国大会, 3F2-05 (2002).
- [87] 山口 高平: ユーザ指向アクティブマイニング, 第 16 回人工知能学会全国大会, 3A5-02 (2002).
- [88] 斉藤 慎也, 小森 麻央, 杉山 岳弘, 杉山 融, 山口 高平: イメージデータマイニングによる少年剣士コーチルールの発見, 第 1 回情報科学技術フォーラム (2002).
- [89] 小森 麻央, 阿部 秀尚, 畑澤 寛光, 橘 恵昭, 山口 高平: 知識発見システムにおける属性処理と属性値処理に関する一考察, 第 15 回人工知能学会全国大会, 1F1-08 (2001).
- [90] 中矢 尚美, 岩出 剛昌, 樽松 理樹, 山口 高平: 情報リソースを活用した領域オントロジー構築支援環境, 第 15 回人工知能学会全国大会, 2B2-03 (2001).
- [91] 杉浦 直樹, 河口 知幸, 和泉 憲明, 山口 高平: ビジネスリポジトリに基づくビジネスモデル構築支援環境の開発, 第 15 回人工知能学会全国大会, 2F1-05 (2001).
- [92] 和泉 憲明, 杉浦 直樹, 澤井 雅彦, 寺井 公一, 山口 高平: ビジネスアプリケーション開発のための異種リポジトリの統合法, 電子情報通信学会, ソフトウェアインタプライズモデリング研究会, pp. 9-16 (2002).
- [93] 和泉 憲明, 杉浦 直樹, 澤井 雅彦, 寺井 公一, 山口 高平: ビジネスアプリケーション開発のための多粒度リポジトリの開発, 人工知能学会, 第 55 回知識ベースシステム研究会, pp. 91-98 (2002).
- [94] 和泉 憲明, 杉浦 直樹, 澤井 雅彦, 寺井 公一, 山口 高平: ビジネスアプリケーション開発のための異種リポジトリの統合法, 電子情報通信学会, ソフトウェアインタプライズモデリング研究会 (2002)

- [95] 和泉 憲明, 杉浦 直樹, 澤井 雅彦, 寺井 公一, 山口 高平: ビジネスアプリケーション開発のための多粒度リポジトリの構築, 電子情報通信学会・知能ソフトウェア工学と人工知能学会知識ベースシステム研究会合同研究会 (2002).
- [96] 小森 麻央, 阿部 秀尚, 山口 高平: シーズ属性の逐次拡張に基づく属性選択, 情報処理学会第128回知能と複雑系研究会と人工知能学会第56回知識ベースシステム研究会合同研究会, 釜山 (2002).
- [97] 阿部 秀尚, 山口 高平: 共通データによる帰納アプリケーション自動構築支援環境の評価, 人工知能学会, 第46回人工知能基礎論研究会・第54回知識ベースシステム研究会合同研究会, pp. 195-200 (2001).
- [98] 大崎 美穂, 坂本 真一, 高木 英行: IEC フィッティングの実用性の評価, 日本ファジィ学会東海支部主催, 第11回東海ファジィ研究会, pp. 1.1-1.4 (2001).
- [99] 藤井 成清, 林田 憲昌, 高木 英行, 大崎 美穂: PDA 版 Visualized IEC フィッティングシステム, 第3回日本ファジィ学会九州支部学術後援会, pp. 61-64 (2001).
- [100] 鈴木 英之進: 「ルール発見の最悪解析」, 第46回人工知能学会人工知能基礎論研究会&第54回人工知能学会知識ベースシステム研究会 合同研究会, pp. 189-194 (2001).
- [101] 長木 悠太, 鈴木 英之進: 「反復マハラノビスデータ圧縮に基づく高速ブースティング」, 第46回人工知能学会人工知能基礎論研究会 & 第54回人工知能学会知識ベースシステム研究会 合同研究会, pp. 201-206 (2001).
- [102] 長浜 光俊, 山口 直記, 鈴木 英之進: 「粗利と購買履歴に基づく有望顧客の特定」, ビジネスマイニングワークショップ講演論文集, pp. 20-23 (2001).
- [103] N. Zhong and M. Ohshima: Mining Interesting Patterns in Multiple Data Sources, *Research Report of JSAI SIG-FAI/KBS-J-01*, Hakodate, Japan, No. 12-14, pp. 177-184 (2001).
- [104] 長木 悠太, 鈴木 英之進: 「反復データ圧縮型ブースティングの実験的評価」, 第48回人工知能学会人工知能基礎論研究会, pp. 1-8 (2002).
- [105] 中本 和岐, 鈴木 英之進: 「TWS木を用いた例数圧縮による時系列データの高速クラスタリング」, 第48回人工知能学会人工知能基礎論研究会, pp. 9-14 (2002).
- [106] 武智 文雄, 鈴木 英之進: 「集合属性の利得比上限値に基づく決定木の高速学習」, 第48回人工知能学会人工知能基礎論研究会, pp. 15-20 (2002).
- [107] 渡辺 健志, 鈴木 英之進: 「混合多項分布推定を用いた肝炎データにおける異常検査値の類型化」, 情報処理学会第128回知能と複雑系研究会 & 人工知能学会第56回知識ベースシステム研究会 合同研究会, pp. 49-54 (2002).
- [108] 渡辺 健志, 鈴木 英之進: 「決定木とブースティングに基づく異常値発見」, 第16回人工知能学会全国大会論文集, 1A3-04 (2002).
- [109] 山口 直記, 長浜 光俊, 鈴木 英之進: 「少販売数アイテムに着目した顧客 - グループ間類似度に基づく購買予測」, 第16回人工知能学会全国大会論文集, 1A4-01 (2002).
- [110] 松本 裕治, 山田 寛康, 新保 仁: 学習に基づく専門用語分類, 人工知能学会, 人工知能基礎

- 論研究会, 知識ベースシステム研究会合同研究会, SIG-FAI/KBS-J-13, pp. 79-84 (2001).
- [111] 中川 哲治, 工藤 拓, 松本 裕治: 修正学習法による形態素解析, 情報処理学会研究報告, 2001-NL-146, pp. 1-8 (2001).
- [112] 山田 寛康, 松本 裕治: Support Vector Machine の多値分類問題への適用法について, 情報処理学会研究報告, 2001-NL-146, pp. 33-38 (2001).
- [113] 持橋大地, 松本裕治: 意味の確率的表現, 情報処理学会研究報告, 自然言語処理研究会, 2002-NL-147, pp. 77-84 (2002).
- [114] 工藤 拓, 山本 薫, 坪井 祐太, 松本 裕治: 言語情報を利用したテキストマイニング, 情報処理学会研究報告, 自然言語処理研究会, 2002-NL-148, pp. 65-72 (2002).
- [115] 山田 寛康, 新保 仁, 松本 裕治: 文脈情報を用いた医学用語分類, 情報処理学会研究報告, 知能と複雑系研究会, 2002-ICS-128, pp. 23-28 (2002).
- [116] 工藤 拓, 山本 薫, 坪井 祐太, 松本 裕治: テキストデータベースからの構文構造のマイニング, 情報処理学会研究報告, 知能と複雑系研究会, 2002-ICS-128, pp. 139-144 (2002).
- [117] 山田 寛康, 松本 裕治: Support Vector Machine を用いた決定性上昇型構文解析, 情報処理学会研究報告, 自然言語処理研究会, 2002-NL-149, pp. 57-64 (2002).
- [118] 中川 哲治, 松本 裕治: Support Vector Machine を用いたコーパスの誤り検出, 言語処理学会第 8 回年次大会発表論文集, pp. 563-566 (2002).
- [119] 工藤 拓, 山本 薫, 坪井 祐太, 松本 裕治: 言語情報を利用したテキストマイニング, 言語処理学会第 8 回年次大会発表論文集, pp. 643-646 (2002).
- [120] 平野 章二, 津本 周作: ラフ集合論に基づく知識指向型クラスタリング法, バイオメディカルファジイシステム学会第 14 回年次大会, pp. 6-9 (2001).
- [121] 平野 章二, 津本 周作: 多重スケールマッチングとラフクラスタリングによる時系列臨床検査データベースの解析, 人工知能学会第 46 回人工知能基礎論研究会, 第 54 回知識ベースシステム研究会合同研究会, SIG-FAI/KBS-J-42, pp. 257-260 (2001).
- [122] 孫 暁光, 柳樂 真佐実, 平野 章二, 安田 晃, 津本 周作: 臨床データベース解析のための類似性尺度とその評価, 第 21 回医療情報学連合大会, pp. 504-505 (2001).
- [123] 平野 章二, 孫 暁光, 津本 周作: 推移パターンの類似性に基づく時系列臨床検査データの解析, 人工知能学会第 56 回知識ベースシステム研究会, SIG-KBS-A201, pp. 61-66 (2002).
- [124] 平野 章二, 孫 暁光, 津本 周作: ラフ集合の近似に基づく画像分割法, 第 16 回人工知能学会全国大会, 3C4-06 (2002).
- [125] 岡田 孝: Challenge における化学物質による発ガン性の予測, 第 24 回情報化学討論会, 第 29 回構造活性相関シンポジウム, JK06, pp. 13-14 (2001).
- [126] 岡田 孝: アクティブマイニングのためのルール群表現法, 人工知能学会, 第 46 回基礎論研究会 (SIGFAI) 第 54 回知識ベースシステム研究会 (SIG-KBS) 合同研究会, pp. 219-224, 函館 (2001).
- [127] 岡田 孝: カスケードモデルによる高次元データの解析: 化学物質の構造活性相関を対象



- として, 研究集会「高次元データ解析の研究」, 広島 (2002).
- [128] 岡田 孝: ルール群表現の発展による化学発ガン性知識の洗練, 情報処理学会 第 128 回知能と複雑系研究会, 人工知能学会第 56 回知識ベースシステム研究会, 釜山 (2002).
- [129] 岡田 孝: カスケードモデルによる発ガン性化合物からの知識発見, 人工知能学会第 16 回全国大会, 3C1-03 東京 (2002).
- [130] 高橋 由雅, 加藤 博明, 藤島 悟志: 化学物質の構造類似性にもとづくデータマイニング, 人工知能学会第 46 回基礎論研究会 (SIGFAI) 第 54 回知識ベースシステム研究会 (SIG-KBS) 合同研究会, pp. 225-228, 函館 (2001).
- [131] 藤島 悟志, 高橋 由雅: トポロジカルフラグメントスペクトル (TFS) ピーク同定システムの開発, 第 24 回情報化学討論会, J15, pp. 59-60 (2001).
- [132] 加藤 博明, 田所 哲男, 宮田 博之, 高橋 由雅, 阿部 英次: タンパク質三次元構造モチーフデータベースの作成, 第 24 回情報化学討論会, JP14, pp. 125-126 (2001).
- [133] 石原 由一郎, 高橋 由雅: 構造類似性を基礎とした化学物質の毒性予測のシステム化に関する研究, 第 29 回構造活性相関シンポジウム, K14, pp. 213-214 (2001).
- [134] 大野 貴生, 高橋 由雅: Non Terminal Vertex Graph (NTG) を用いた化学構造の特徴解析, 日本化学会第 81 春季年会, 1F8-28, 東京 (2002).
- [135] 松尾 豊, 大澤 幸生, 石塚 満: Small world と Average-clicks, 第 2 回人工知能学会 MYCOM (2001).
- [136] 大澤 幸生, 松村 真宏, 松尾 豊: 自然・社会現象データからの予兆・チャンス発見, 第 43 回人工知能学会 AI チャレンジ研究会 (SIG-Challenge) (2001).
- [137] 大澤 幸生, 相馬 浩隆, 臼井 優樹, 松村 真宏, 松尾 豊: 会話展開キーグラフによる Web コミュニティの特性表現, 人工知能学会第 46 回 SIG-FAI 第 54 回 SIG-KBS 合同研究会 (2001).
- [138] 福田 寿大澤幸生: 消費者の食品メニュー履歴データからの潜在意識発見支援, 人工知能学会・第 47 回人工知能基礎論研究会 (2002).
- [139] 松村 真宏, 大澤 幸生, 石塚 満: 頻出アイテムと希少アイテム間の相関ルールからのチャンス発見, 人工知能学会・第 47 回人工知能基礎論研究会 (2002).
- [140] 松村 真宏, 大澤 幸生, 石塚 満: 語の活性伝播に基づく談話分析, 人工知能学会・第 47 回人工知能基礎論研究会 (2002).
- [141] 大澤 幸生, 奈良 由美子: チャンス発見プロセスの二重らせんモデルに基づくアンケート調査データの解析, 人工知能学会・第 47 回人工知能基礎論研究会 (2002).
- [142] 吉川 史子, 大澤 幸生: 新聞記事に見るチャンスの意味, 人工知能学会・第 47 回人工知能基礎論研究会 (2002).
- [143] 松尾 豊, 大澤 幸生, 石塚 満: 電子掲示板における会話からのハイライト部分の抽出, 人工知能学会・第 47 回人工知能基礎論研究会 (2002).
- [144] 松尾 豊, 大澤 幸生, 石塚 満: Small World のさまざまな拡張についての考察, 人工知能

学会第 47 回人工知能基礎論研究会 (2002).

- [145] 松村 真宏, 大澤 幸生, 石塚 満: Knowledge Navigation through Combining Documents, 人工知能学会全国大会, 1A4-06 (2001).
- [146] 坂田 恭弘, 平 博司, 大澤 幸生, 伊庭 斉志, 石塚 満: AreaBook: WWW エリアビューのブック型情報提示インタフェース, 人工知能学会全国大会, 1A4-06 (2001).
- [147] 松村 真宏, 大澤 幸生, 石塚 満: 質問応答システムのための新しい文書検索手法の提案: 第 40 回情報処理学会全国大会 (2001).
- [148] 松村 真宏, 大澤 幸生, 石塚 満: 文書の組み合わせに基づく知識ナビゲーション: 第 15 回人工知能学会全国大会 (2001).
- [149] 大澤 幸生: 近未来チャレンジのきっかけと展開, 第 43 回人工知能学会 AI チャレンジ研究会 (SIG-Challenge) (2001).
- [150] Y. Ohsawa: Genetic Words - Building Blocks as Context - 行動計量学会全国大会 (2002).
- [151] 大澤 幸生, 相馬 浩隆ほか: 会話展開キーグラフによる Web コミュニティの特性表現 人工知能学会第 46 回基礎論研究会 (2002).
- [152] 松尾 豊, 大澤 幸生, 石塚 満: 電子掲示板における会話からのハイライト部分の抽出, 人工知能学会第 47 回基礎論研究会 (2002).
- [153] 松尾 豊, 大澤 幸生, 石塚 満: Small World のさまざまな拡張についての考察人工知能学会第 47 回基礎論研究会 (2002).
- [154] 福田 寿, 大澤 幸生: 消費者の食品メニュー履歴データからの潜在意識発見支援人工知能学会第 47 回基礎論研究会 (2002).
- [155] 松村 真宏, 大澤 幸生, 石塚 満: 頻出アイテムと希少アイテム間のコレレーションルールからのチャンス発見人工知能学会第 47 回基礎論研究会 (2002).
- [156] 松村 真宏, 大澤 幸生, 石塚 満: 語の活性伝播に基づく談話分析人工知能学会第 47 回基礎論研究会 (2002).
- [157] 吉川史子, 大澤 幸生: 新聞記事に見るチャンスの意味人工知能学会第 47 回基礎論研究会 (2002).
- [158] 大澤 幸生, 奈良 由美子: チャンス発見プロセスの二重らせんモデルに基づくアンケート調査データの解析人工知能学会第 47 回基礎論研究会 (2002).
- [159] 勝又 勇治, 倉橋 節也, 寺野 隆雄: ベイジアンネットワークとタブーリストを利用したハイブリッド GA による多峰性関数の最適化人工知能学会全国大会 (第 15 回) 論文集, p. (2C3-01) (2001).
- [160] 谷口 憲, 倉橋 節也, 寺野 隆雄: エージェント指向サプライ・チェーン・モデルに対する遺伝的プログラミングの適用人工知能学会全国大会 (第 15 回) 論文集, p. (2C3-09) (2001).
- [161] 倉橋 節也, 寺野 隆雄: エージェントモデルによるコミュニティの共同分配規範の成立と崩壊に関する考察人工知能学会全国大会 (第 15 回) 論文集, p. (3C2-03) (2001).
- [162] 村上 英次, 寺野 隆雄: F-CUBE: コミュニティに共有する知識のインターネットフォー

- ラムからの抽出システム人工知能学会全国大会 (第 15 回) 論文集, p. (3B1-02) (2001).
- [163] 出口 弘, 寺野 隆雄, 車谷 浩一, 湯澤 太郎, 橋本 重治, 松井啓之, 幸島 明男, 兼田 敏之: エージェントに基づく仮想経済シミュレータの開発人工知能学会全国大会 (第 15 回) 論文集, p. (2F1-10) (2001).
- [164] 佐藤 浩, 松井 啓之, 小野 功, 喜多 一, 寺野 隆雄, 出口 弘, 塩沢 由典: U-Mart: エージェントシミュレーションで経済を学ぶ人工知能学会全国大会 (第 15 回) 論文集, p. (3F1-11) (2001).
- [165] 山口 高平, 元田 浩, 寺野 隆雄, 鷲尾 隆, 斉藤 和巳, 津本 周作: Discovery of Communicable Knowledge 人工知能学会全国大会 (第 15 回) 論文集, p. (3B4-01) (2001).
- [166] 勝又 勇治, 倉橋 節也, 寺野 隆雄: タブーリストを用いたベイジアン最適化アルゴリズムによる多峰性関数最適化. 情報処理学会 第 8 回 MPS (数理と問題解決) シンポジウム-進化的計算シンポジウム 2001- (2001)
- [167] 寺野 隆雄: 計算組織論とエージェントベースモデル. 社会・経済システム学会 第 20 回大会プログラム「システム論を問いなおす: システム論の新展開-主題と方法-」(2001).
- [168] 寺野 隆雄, 稲田 政則: EBM とデータマイニング: 知識評価の観点から. 人工知能学会研究会資料 SIG-FAI/KBS-J-16(11/13), pp97-102 (2001).
- [169] 高橋 大志, 寺野 隆雄: エージェントシミュレーションによる Prospect 理論と GARCH モデルの関連性の分析. 第 10 回マルチ・エージェントと協調計算ワークショップ (MACC2001) (2001).
- [170] 喜多 一, 出口 弘, 寺野 隆雄: U-Mart: 経済学と工学をエージェントが結ぶ. 第 10 回マルチ・エージェントと協調計算ワークショップ (MACC2001) (2001).
- [171] 國上 真章, 寺野 隆雄: エージェント系としてのレプリケータダイナミクスとその制御. 第 10 回マルチ・エージェントと協調計算ワークショップ (MACC2001) (2001).
- [172] 稲田 政則, 寺野 隆雄: Evidence-Based Medicine とデータマイニング. 計測自動制御学会, システム・情報部門学術講演会 (2001)
- [173] 喜多 一, 出口 弘, 寺野 隆雄: オープン型人工市場 U-Mart: 構想, 成果, 展望. 電子情報通信学会技術研究報告, 人工知能と知識処理 (AI2001 - 58), pp. 17-23 (2002).
- [174] 寺野 隆雄, 出口 弘: 社会科学におけるエージェント研究の動向と課題. 電子情報通信学会技術研究報告, 人工知能と知識処理 (AI2001 - 59), pp. 25-32 (2002).
- [175] 寺野 隆雄, 喜多 一, 出口 弘: U-Mart プロジェクト: 仮想市場とエージェントと経済学. 人工知能学会, 知識ベースシステム研究会 (第 55 回), pp. 99-103 (2002).
- [176] 稲田 政則, 寺野 隆雄: メタ学習による肝機能検査データからの因果モデルの構築知識ベースシステム研究会 (第 56 回), 人工知能学会研究会資料 (SIG-KBS-A201), pp. 73-78 (2002).
- [177] 倉橋 節也, 寺野 隆雄: 学習分類子システムを用いたプロセス時系列からのデータマイニング知識ベースシステム研究会 (第 56 回), 人工知能学会研究会資料 (SIG-KBS-A201), pp. 1-6 (2002).

- [178] 小林 正人, 荒井 篤子, 森川 明美, 寺野 隆雄: ビジネスゲーム開発におけるエージェント技術の適用人工知能学会全国大会 (第 16 回) 論文集, pp. 2-4, (2002) .
- [179] 吉田 健一, 欧米の次世代インターネット・プロジェクトの動向 - トラフィック計測技術と適用プロジェクトの動向, インターネット技術シンポジウム 2001 (2001).
- [180] 勝野 聡, 山崎 克之, 飯作 俊一, 浅見 徹, 菅内 公德, 常 弘敦, 吉田 健一, 榎本 博道: 高速・高精度 IP メータの試作, 信学技報 Vol.101, No.186, pp. 67-73 (2001).

## 著書 / 編書

- [1] 北村 泰彦, 山田 誠二 (編): 情報検索とエージェント, 東京電機大学出版局 (2002).
- [2] 北村 泰彦, 山田 誠二 (編): 情報社会とデジタルコミュニティ, 東京電機大学出版局 (2002).
- [3] 西田 豊明, 木下 哲男, 北村 泰彦, 間瀬 健二: エージェント工学, オーム社, (2002).
- [4] 沼尾 正行 (「知識」編集委員/執筆者): 日本認知科学会 (編) 認知科学辞典、共立出版 (2002).
- [5] H. Liu and H. Motoda: Instance Selection and Construction for Data Mining (Eds.), Kluwer Academic Publishers (2001).
- [6] 鈴木 篤之, 元田 浩: システムの設計・運用・評価, 岩波講座現代工学の基礎 (2002).
- [7] H. Motoda (ed.): Active Mining - New Directions of Data Mining -, Frontiers in Artificial Intelligence and Applications, IOS Press (2002).
- [8] L. Polkowski, S. Tsumoto, and T.Y. Lin (eds.): *Rough Set Methods and Applications : New Developments in Knowledge Discovery in Information Systems*, Physica-Verlag, New York (2001).
- [9] T. Terano, T. Nishida, A. Namatame, S. Tsumoto, Y. Ohsawa, and T. Washio (eds.): *New Frontiers in Artificial Intelligence, Joint JSAI 2001 Workshop Post-Proceedings*, Lecture Notes in Computer Science 2253, Springer-Verlag, Heidelberg (2001).
- [10] N. Callaos, Y. Ohsawa, Y. Zhang, R. Szabo and M. Aveledo, eds.: *World Multiconference on Systemics, Cybernetics, and Informatics Proceedings Volume VIII* (2001).
- [11] A. Namatame, T. Terano and K. Kuruamtani (eds.): Agent-Based Approaches in Economic and Social Complex Systems. IOS Press and Ohmsha (2002).

## 著書 ( 章分担執筆 )

- [1] 北村泰彦: 探索, 教育システム情報ハンドブック, 実教出版, pp. 9-12 (2001).
- [2] Y. Kitamura: Intelligent Tickers: An Information Integration Scheme for Active Information Gathering, H. Motoda (Ed.) Active Mining, IOS Press, pp. 61-69 (2002).

- [3] 北村 泰彦：WWW 情報統合，北村 泰彦，山田 誠二（編）：情報検索とエージェント，東京電機大学出版局，pp. 53-83 (2002).
- [4] 西村 俊和，北村 泰彦：知的インターネット技術を用いた教育，北村 泰彦，山田 誠二（編）情報社会とデジタルコミュニティ，東京電機大学出版局，pp. 97-125 (2002).
- [5] Y. Takama, K. Hirota: Immune Network-based Clustering for WWW Information Gathering/Visualization, *Active Mining: New Directions of Data Mining*, H.Motoda (Ed.), pp.21-30,IOS Press (2002).
- [6] M. Okabe, S. Yamada: Interactive Web Page Retrieval with Relational Learning-based Filtering Rules, *Active Mining: New Directions of Data Mining*, H.Motoda (Ed.), pp.31-40,IOS Press (2002).
- [7] S. Yamada, Y. Nakai: Monitoring Partial Update of Web Pages by Interactive Relational Learning, *Active Mining: New Directions of Data Mining*, H.Motoda (Ed.), pp.41-50,IOS Press (2002).
- [8] 山田 誠二: 情報収集エージェント，北村 泰彦，山田 誠二（編）：情報検索とエージェント，東京電機大学出版局，pp.27-52 (2002).
- [9] M. Numao, M. Yoshida and Y. Ito: Data Mining on the WAVES — Word-of-mouth-Assisting Virtual Environments, *Active Mining, Frontiers in Artificial Intelligence and Applications*, 79, pp. 11-20, IOS Press (2002).
- [10] T. N. Tran and M. Numao: Toward Active Mining from On-line Scientific Text Abstracts Using Pre-existing Sources, *Frontiers in Artificial Intelligence and Applications*, 79, pp. 3-10, IOS Press (2002).
- [11] H. Abe and T. Yamaguchi: Evaluating the Automatic Composition of Inductive Applications Using StatLog repository of Data Set, *Active Mining – New Directions of Data Mining* , IOS press, pp. 139-150 (2002).
- [12] N. Zhong: Knowledge Discovery and Data Mining, *Encyclopedia of Microcomputers*, Vol. 27 (Supplement 6), Marcel Dekker, pp. 235-286 (2001).
- [13] N. Zhong and S. Ohsuga: Automatic Knowledge Discovery in Larger Scale Knowledge-Data Bases, in C. Leondes (ed.) *The Handbook of Expert Systems*, Vol. 4, Academic Press, pp. 1015-1070 (2001).
- [14] E. Suzuki: PAGA Discovery: A Worst-Case Analysis of Rule Discovery for Active Mining, *Active Mining, New Directions of Data Mining*, pp. 127-138, IOS Press, Amsterdam (2002).
- [15] Y. Choki and E. Suzuki: Fast Boosting Based on Iterative Data Squashing, *Active Mining, New Directions of Data Mining*, pp. 151-161, IOS Press, Amsterdam (2002).
- [16] 鈴木 英之進:「データマイニング」,エンサイクロペディア情報処理 改訂 4 版 ,pp. 246-247 ,情報処理学会編 ,オーム社 (2002) .
- [17] T. Wada, T. Yoshida, H. Motoda and T. Washio: Acquiring Knowledge from Both

- Human Experts and Accumulated Data in an Unstable Environment, *Active Mining: New Directions of Data Mining*, H. Motoda (Ed.), pp. 217-228, IOS Press (2002).
- [18] K. Fujiwara, T. Yoshida, H. Motoda and T. Washio: Case Generation Method for Constructing an RDR Knowledge Base, *Active Mining: New Directions of Data Mining*, H. Motoda (Ed.), pp. 205-215, IOS Press (2002).
- [19] T. Matsuda, T. Yoshida, H. Motoda and T. Washio: Knowledge Discovery from Structured Data by Beam-wise Graph-Based Induction, *Active Mining: New Directions of Data Mining*, H. Motoda (Ed.), pp. 115-125, IOS Press (2002).
- [20] T.B. Ho, S. Kawasaki and N.B. Nguyen: Documents Clustering Using Tolerance Rough Set Models and Its Application to Information Retrieval, Chapter in *Intelligent Exploration of the Web*, P.S. Szczepaniak, J. Segovia, J. Kacprzyk, L.A. Zadeh (Eds.), Physica-Verlag (in press, Summer 2002).
- [21] T.B. Ho and T.D. Nguyen: Model Selection in Knowledge Discovery and Data Mining, Book chapter in *Intelligent Systems: Techniques and Applications*, T. Leondes (Ed.), Vol. 4, pp. IV.1-17, CRC Press (2002).
- [22] T.B. Ho, T.D. Nguyen, D.D. Nguyen and S. Kawasaki: Active Participation of Users with Visualization Tools in the Knowledge Discovery Process, *Active Mining: New Directions of Data Mining*, H. Motoda (Ed.), pp. 229-238, IOS Press (2002).
- [23] E. Ip, J. Johnson, K. Yada, Y. Hamuro, N. Katoh and S. Cheung: A Neural Network Application to Identify High-Value Customer for a Large Retail Store in Japan, *Neural Networks in Business: Techniques and Applications*, Idea Group Publishing, pp. 55-69 (2002).
- [24] Y. Hamuro, E. Kawata, N. Katoh and K. Yada: A Machine Learning Algorithm for Analyzing String Patterns Helps to Discover Simple and Interpretable Business Rules from Purchase History, *Progresses in Discovery Science, State-of-the-Art Surveys*, LNCS, Springer-Verlag, pp. 565-575 (2001).
- [25] H. Abe and T. Yamaguchi: Constructing Inductive Applications by Meta-Learning with Method Repositories, *Progress in Discovery Science, Lecture Notes in Artificial Intelligence Vol.2281*, Springer-Verlag, pp. 576-585 (2002).
- [26] E. Suzuki: In Pursuit of Interesting Patterns with Undirected Discovery of Exception Rules, *Progresses in Discovery Science, Lecture Notes in Computer Science 2281, State-of-the-Art Surveys*, Springer-Verlag, pp. 504-517 (2002).
- [27] S. Tsumoto: Discovery of Clinical Knowledge in Databases Extracted from Hospital Information Systems, K.J. Cios (ed.) *Medical Data Mining and Knowledge Discovery*, Physica-Verlag, New York, pp. 433-454 (2001).
- [28] S. Tsumoto: Induction of Rule about Complications with the Use of Rough Sets, W. Pedrycz (ed.) *Granular Computing: an emerging paradigm*, Physica-Verlag, New York, pp. 384-397 (2001).

- [29] S. Tsumoto: Chapter G5: Data Mining in Medicine, W. Kloesgen and J. Zytchow (eds.) *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press (2001).
- [30] S. Tsumoto: Discovery of Positive and Negative Knowledge in Medical Databases Using Rough Sets. *Progress in Discovery Science 2002*: 543-552. Setsuo Arikawa, Ayumi Shinohara (Eds.): *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*. LNCS 2281 Springer Verlag, pp. 543-552 (2002).
- [31] S. Tsumoto: Meta Analysis for Data Mining, H. Motoda (ed.) *ACTIVE MINING*, IOS Press, pp. 279-290 (2002).
- [32] S. Hirano: A Hybrid Approach of Multiscale Matching and Rough Clustering to Knowledge Discovery in Temporal Medical Databases, H. Motoda (ed.) *ACTIVE MINING*, IOS Press, pp. 269-278 (2002).
- [33] T. Okada: Topographical Expression of a Rule for Active Mining, In H. Motoda (eds.) "Active Mining", IOS Press, pp. 247-257 (2002).
- [34] Y. Ohsawa: Discovery of Chances Underlying Real Data, *Progress in Discovery Science*, eds., Arikawa, S., LNAI from Springer Verlag (2001).
- [35] Y. Ohsawa, The Scope of Chance Discovery: *New Frontiers in Artificial Intelligence*, LNAI2253, Takao Terano, et al eds., Springer Verlag, pp. 413 (2001).
- [36] N. Matsumura, Y. Ohsawa and M. Ishizuka: Chapter 59: Future Directions of Communities on the Web *New Frontiers in Artificial Intelligence*, LNAI2253, Takao Terano, et al eds., Springer Verlag, pp. 435-442 (2001).
- [37] Y. Ohsawa and Y. Nara: Chapter 66: Action Proposals as Discovery of Context (An Application to Family Risk Management) *New Frontiers in Artificial Intelligence*, LNAI2253, Takao Terano, et al eds., Springer Verlag, pp. 481-485 (2001).
- [38] Y. Matsuo and Y. Ohsawa: Chapter 60: A Document as a Small World *New Frontiers in Artificial Intelligence*, LNAI2253, Takao Terano, et al eds., Springer Verlag, pp. 444-448 (2001).
- [39] Y. Ohsawa et al, Classifying On-line Communities on the Structure of Word Co-occurrence, ed. Hiroshi Motoda *Active Mining*, IOS Press, pp. 195-204 (2002).
- [40] 大澤 幸生 他著: 情報社会とデジタルコミュニティ, 山田 誠二, 北村 泰彦編東京電機大学出版局 (インターネットの知的情報技術シリーズ) (2002).
- [41] K. Takadama, T. Terano, K. Shimohara, K. Hori and S. Nakasuka: Towards a Multi-agent Design Principle -Analyzing an Organizational-Learning Oriented Classifier System-, Loia, V. and Sessa, S. (Eds.) *Soft Computing Agents: New Trends for Designing Autonomous Systems*" Special Issue of the Series "Studies in Fuzziness and Soft Computing", Physica-Verlag, Springer Publisher (2001).
- [42] S. Kurahashi and T. Terano: Analyzing Norm Emergence in Communal Sharing via Agent-Based Simulation, LNAI 2253, pp. 88-98 (2001).

- [43] H. Sato, H. Matsui, I. Ono, H. Kita, T. Terano, H. Deguchi, and Y. Shiozawa: U-Mart Project: Learning Economic Principles from the Bottom by Both Human and Software Agents. LNAI 2253, pp. 121-131 (2001).
- [44] H. Deguchi, T. Terano, K. Kurumatani, T. Yuzawa, S. Hashimoto, H. Matsui, A. Sashima, and T. Kaneda: Virtual Economy Simulation and Gaming –An Agent Based Approach– LNAI 2253, p. 218-226 (2001).
- [45] S. Kurahashi, T. Terano: Emergence, Maintenance, and Collapse of Norms on Information Communal Sharing: Analysis via Agent-Based Simulation. Agent-Based Approaches in Economic and Social Complex Systems. IOS Press and Ohmsha, pp. 25-34 (2002).
- [46] H. Sato, H. Matsui, I. Ono, H. Kita, T. Terano, H. Deguchi, Y. Shiozawa: Case Report on U-Mart Experimental System: Competition of Software Agents and Gaming Simulation with Human Agents. Agent-Based Approaches in Economic and Social Complex Systems. IOS Press and Ohmsha, pp. 167-178 (2002).
- [47] H. Deguchi, T. Terano, K. Kurahashi, T. Kaneda, H. Matsui, T. Yuzawa, A. Sashima, Y. Koyama, H. Lee, M. Kobayashi: Virtual Economy -Agent Based Modeling and Simulation of National Economy-. Agent-Based Approaches in Economic and Social Complex Systems. IOS Press and Ohmsha, pp. 198-207 (2002).
- [48] K. Yoshida, A. Sakurai: Machine Learning, In Encyclopedia of Information Systems, Academic Press, in press (2002).

## 解説

- [1] S. Yamada, H. Kawano: Information Gathering and Searching Approaches in the Web, New Generation Computing, Vol. 19, No. 2, pp. 195-208 (2001).
- [2] 山田 誠二, 村田 剛志, 北村 泰彦: 知的 Web 情報システム, 人工知能学会誌, Vol. 16, No. 4, pp. 495-502 (2001).
- [3] 大澤 幸生, 高間 康史: WWW に潜む創造的意思決定のチャンス, 人工知能学会誌, Vol. 16, No. 4, pp. 530-534 (2001).
- [4] 沼尾 正行, 宮下 和雄 (インタビュー): 社会へのアンテナ第 4 回 高橋邦芳 氏 - ユーザを納得させられる人工知能技術の実現を, 人工知能学会誌, Vol. 16, No. 5, pp. 732-733 (2001).
- [5] 市瀬 龍太郎, パット ラングレー: 適応的ユーザインタフェースと音声対話, 人工知能学会誌, Vol. 17, No. 3, pp. 291-294 (2002).
- [6] 加藤 直樹, 羽室 行信, 矢田 勝俊: 新規顧客からのロイヤルカスタマーの早期発見, ESTRELA, No. 89, pp. 10-17, (2001).
- [7] 元田 浩, 鷺尾 隆: デ - タマイニング展望, システム制御情報学会誌, Vol. 46, No. 4, pp. 169-176 (2002).
- [8] 元田 浩, 沼尾 正行, 山口 高平, 津本 周作: アクティブマイニングの構想と展開, 人工



- 知能学会誌, Vol.17, No.5, pp. 615-621 (2002).
- [9] 山口 高平: データマイニングアプリケーションの自動構築に向けて, ミニ特集「データマイニングの最前線」計測自動制御学会誌, 第 41 巻, 5 号, pp. 347-351 (2002).
  - [10] 和泉 憲明, 武田 英明, 山口 高平: 意味理解する Web を目指して-次世代 Web の方向性を探る-, 人工知能学会誌, 第 17 巻, 4 号 (2002).
  - [11] 鈴木 英之進, 菅谷 信介, 津本 周作: サポートベクターマシンに基づく医療データからの事例発見, オペレーションズ・リサーチ, vol. 46, no. 5, pp. 243-248 (2001).
  - [12] 鈴木 英之進: 「例外ルールの発見」, システム制御情報学会論文誌, Vol. 13, No. 4, pp. 197-202 (2002).
  - [13] 津本 周作: 医学・医療におけるデータマイニングの現状, システム/制御/情報 vol. 46, no. 4, pp. 184-189 (2001).
  - [14] 津本 周作: ラフ集合論の現状と課題, 日本ファジィ学会誌, vol. 13, no. 6, pp. 552-561 (2001).
  - [15] 津本 周作: 専門家のとの協調と知識発見- データマイニングからアクティブマイニングへ -, 計測と制御, vol. 41, no. 5, pp. 331-336 (2001).
  - [16] 津本 周作: 診療データベースからのデータマイニング, 現代医療, vol. 34, no. 3, pp. 770-776 (2001).
  - [17] 津本 周作: 医療におけるデータマイニングの現状と展望. BME (日本 ME 学会誌), vol. 16, no. 4, pp. 24-29 (2002).
  - [18] 横井 英人, 平野 章二, 高林 克日己, 津本 周作, 里村 洋一: 慢性ウイルス性肝炎データに関するアクティブマイニング-病院情報システムにおける知識発見プロセスの実現に向けて-, 人工知能学会誌, Vo. 17, No. 5, pp. 622-628 (2002).
  - [19] 岡田 孝, 元田 浩: 相関ルールとその周辺, オペレーションズ・リサーチ, No. 9, in press (2002).
  - [20] 加藤 博明: タンパク質三次元共通構造特徴の自動認識システムの開発, SAR News, No. 2, pp. 8-11 (2002).
  - [21] 大澤 幸生: 「予兆発見の科学」(株) 博報堂「未来予兆」(2002).
  - [22] 大澤 幸生, 松尾 豊, 松村 真宏: 予兆発見 - 自然・社会からの意思決定, システム制御情報学会誌 (2002).