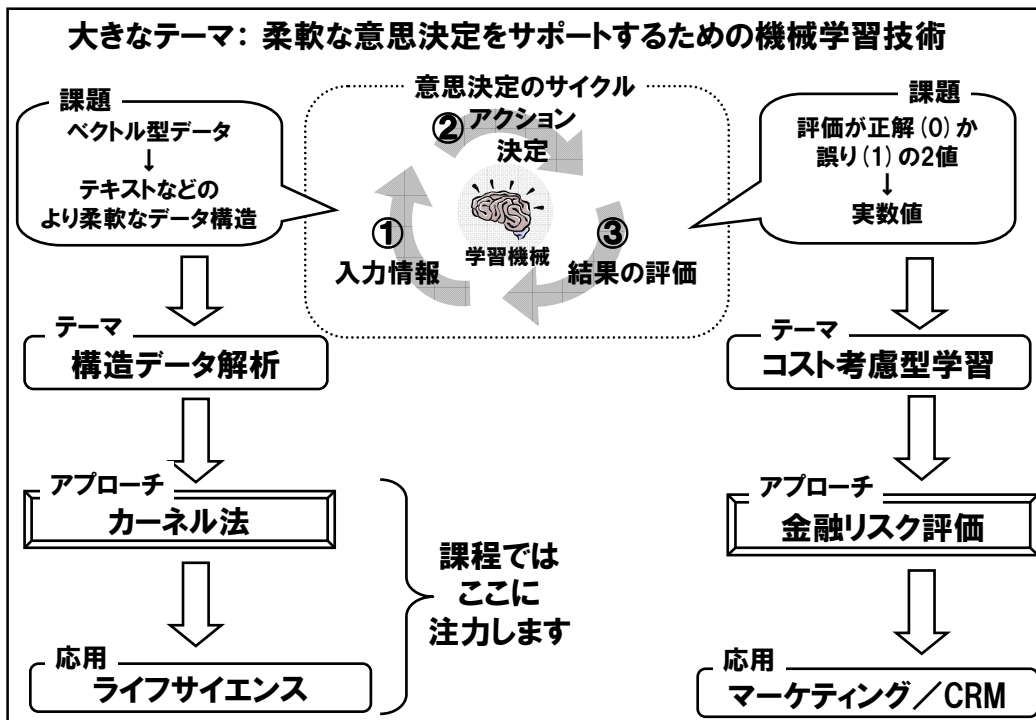


鹿島 久嗣 (かしま ひさし)

日本アイ・ビー・エム株式会社
東京基礎研究所
主任研究員

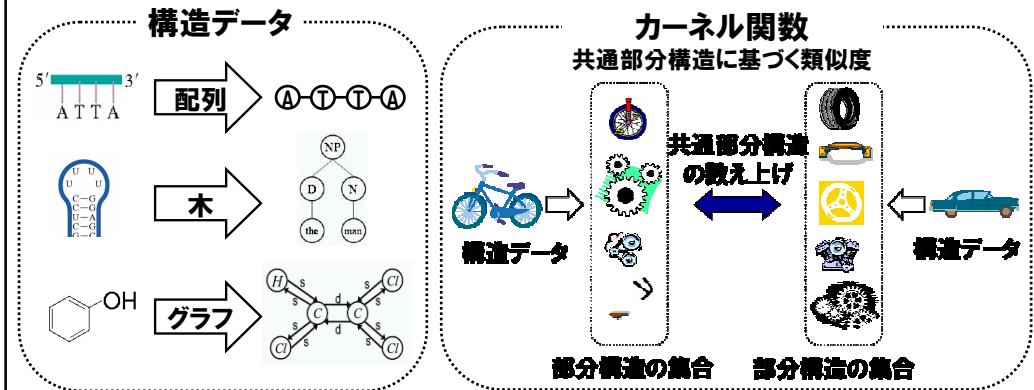
これまでに行った研究の概要と希望する研究

- テーマ
 - ▶ 柔軟な意思決定をサポートするための機械学習技術 ⇒ 構造データの解析法
- これまでの成果
 - ▶ 各種構造データに対するカーネル関数の設計
- 今後の研究
 - ▶ さまざまな構造情報をとりいれ、必要な部分構造だけを用いて、より高速に動く学習法



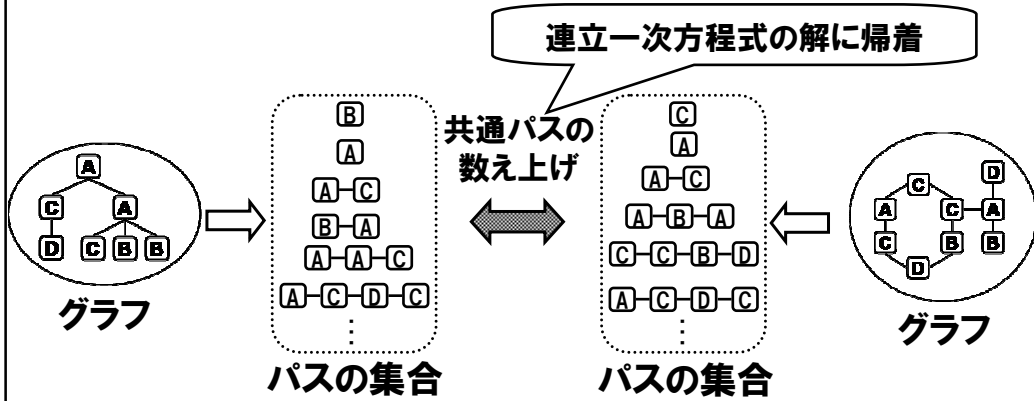
これまでの研究成果：各種構造データに対するカーネル法の適用

- カーネル法：カーネル関数(=2つのデータの類似度)に基づくデータ解析法
- 研究項目：2つの構造データのカーネル関数を、共通にもつ部分構造を用いて定義する(=畳み込みカーネル)
 - ▶ 難しいところ：部分構造の数が多い ⇒ 単純な数え上げが不可能
- これまでにやったこと：適切な部分構造の定義と、効率よいアルゴリズム
 - ▶ ラベルつき順序木カーネル(すべての部分木に基づくカーネル)
 - ▶ グラフカーネル(ランダムウォークで生成されるパスに基づくカーネル)



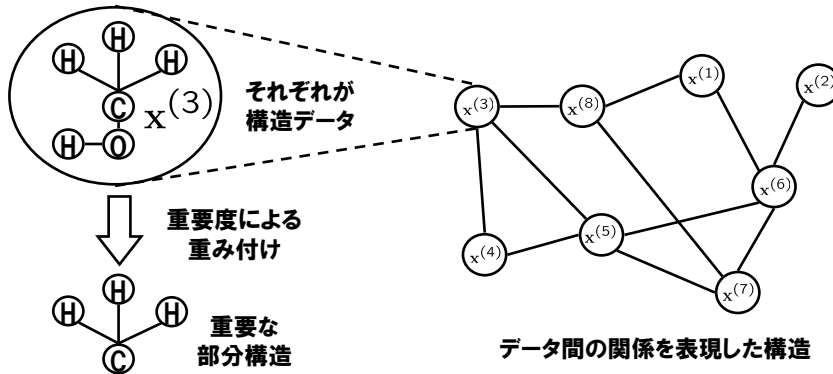
これまでの研究成果：グラフ・カーネル

- 構造データのカーネル：共通部分構造の数え上げ
- グラフカーネルでは、
 - ▶ 部分構造として、グラフ上のランダム・ウォークによって生成されるパスを使う
 - ▶ 2つのグラフから共通のパスが生成される個数の期待値を求める
- 難しいところ：パスは無限個ありうる
 - ← 解決法：数え上げの計算を連立一次方程式に帰着することで計算可能に



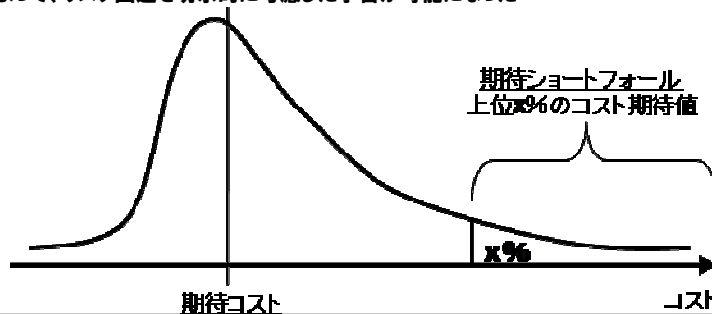
今後の研究: データ間の構造情報を取り入れ、重要な部分構造だけを用い、より高速に動く学習法

- 情報統合
 - ▶ データ内部の構造だけでなく、データ間の構造もとり入れる
- 部分構造選択
 - ▶ すべての部分構造を考慮 → 部分構造の重要度に応じた重みづけをする
- 高速化
 - ▶ 線形時間で計算できるカーネル関数のクラス



これまでの研究成果: リスク回避型学習

- 大きな損失のリスクを回避しながら、平均損失を小さくするような意思決定ルールを学習する
 - ▶ ヒトはリスクに対する嗜好をもっている (リスク嗜好型、リスク回避型、リスク中立型)
 - ▶ ⇔ 従来は平均損失の最小化のみを考慮していた (=リスク中立)
 - 無限の資産を持っているならよいが、数回の大きな損失が致命的な場合には問題
- 機械学習に金融工学のアプローチを導入
 - ▶ 平均損失ではなく「期待ショートフォール」とよばれる金融工学で用いられるリスク指標を目的関数として導入
 - ▶ 既存のリスク中立型学習アルゴリズムを、期待ショートフォールを最小化するようなリスク回避型学習アルゴリズムに変換するメタアルゴリズムを提案
 - ▶ これによって、リスク回避を明示的に考慮した学習が可能になった



これまでの研究成果：主な発表論文等

■ ジャーナル論文

- ▶ (投稿中) H. Kashima: Risk-Sensitive Learning via Expected Shortfall Minimization, submitted to *Journal of Machine Learning Research*, 2006.
- ▶ 鹿島, 坂本, 小柳: 木構造データに対するカーネル関数の設計と解析, *人工知能学会論文誌*, Vol.21, No.1, 2006.
- ▶ 鹿島, 津村, 井手, 野々山, 平出, 江藤, 福田: ネットワークデータを用いた分散システムにおける異常検出, *電子情報通信学会論文誌*, Vol. J89-D, No. 2, 2006.
- ▶ T. Shibuya, H. Kashima and A. Konagaya: Efficient Filtering Methods for Clustering cDNAs with Spliced Sequence Alignment, *Bioinformatics*, 2004.
- ▶ T. Fukao, H. Kashima and N. Adachi: Decentralized Adaptive Control with Improved Transient Performance, *計測自動制御学会論文集*, Vol.35, No.7, pp. 869-878, 1999.

■ 主な国際会議論文

- ▶ H. Kashima: Risk-Sensitive Learning via Expected Shortfall Minimization (Extended Abstract) , To appear in Proc. *2006 SIAM Conference on Data Mining (SDM)* , 2006.
- ▶ H. Kashima, T. Tsumura, T. Ide, T. Nogayama, R. Hirade, H. Etoh and T. Fukuda: Network-Based Problem Detection for Distributed Systems, In Proc. *21st International Conference on Data Engineering (ICDE)* , 2005.
- ▶ T. Ide and H. Kashima: Eigenspace-based Anomaly Detection in Computer Systems, In Proc. *10th ACM SIGKDD Conference (KDD)* ,2004.
- ▶ H. Kashima and Y. Tsuboi: Kernel-Based Discriminative Learning Algorithms for Labeling Sequences, Trees and Graphs, In Proc. *21st International Conference on Machine Learning (ICML)* , 2004.
- ▶ A. Inokuchi and H. Kashima: Mining Significant Pairs of Patterns from Graph Structures with Class Labels, In Proc. *3rd IEEE International Conference on Data Mining (ICDM)* , 2003.
- ▶ H. Kashima , K. Tsuda and A. Inokuchi: Marginalized Kernels Between Labeled Graphs, In Proc. *20th International Conference on Machine Learning (ICML)* , 2003.
- ▶ H. Kashima and T. Koyanagi: Kernels for Semi-Structured Data, In Proc. *19th International Conference on Machine Learning (ICML)* , 2002.