

CAMLET: 帰納的学習システム構築支援環境 (2)

2M-2

—超並列計算機による仕様の探索—

玉置 俊広

酢山 明弘

山口 高平

静岡大学

1 はじめに

帰納的学習システム構築支援環境 CAMLET[根岸,1998]においては帰納的学習システムを構成するための空間が巨大となるため、有用な探索技法を確立することが課題であった。本稿では探索のための仕様表現、探索オペレータ、探索の制御法を考察した後に、超並列計算機上での実現法について考察する。

2 帰納的学習システムの構成空間

代表的な帰納的学習システムは、generating training/test data sets、generating classifier sets、evaluation sets、modifying training data sets、modifying classifier sets という5種類の基本プロセスから構成される[根岸,1998]。以下、この基本プロセスを基本概念部品と呼び、基本概念部品をさらに詳細化したものを詳細概念部品と呼ぶ。

図1は帰納的学習システムの構成空間であり、CAMLETの探索木となる。この探索木は3レベルから構成され、上位レベルは5種類の基本概念部品から構成される基本設計仕様レベル、中位レベルは詳細概念部品から構成される詳細設計仕様レベル、下位レベルはさらに学習パラメータ情報を特定したパラメータの値が付与された詳細設計仕様レベルを表す。

3 逐次処理型 CAMLET

逐次処理型 CAMLET[根岸,1998]は、学習パラメータの値はデフォルト値を利用して図1の下位レベルを探索することなく中位レベルまでを探索する。

探索に利用する戦略は、以下の局所置換と大域置換である。

局所置換: 同じ基本設計仕様の下で詳細概念部品を置換することにより詳細設計仕様を変更する

大域置換: 基本概念部品の組合せを変えることにより基本概念部品の組合せを基本設計仕様を変更する

逐次処理型 CAMLETは、ランダムに決定された初期詳細設計仕様に対して、まず、局所置換を貪欲的に適用する。すなわち、プロセスオンロジーを利用し

CAMLET: A Computer Aided Inductive Learning Engineering Environment(2) - Searching Best Specifications with Massively Parallel Computer -
Toshihiro TAMAOKI, Graduate School of Science and Engineering, Shizuoka University

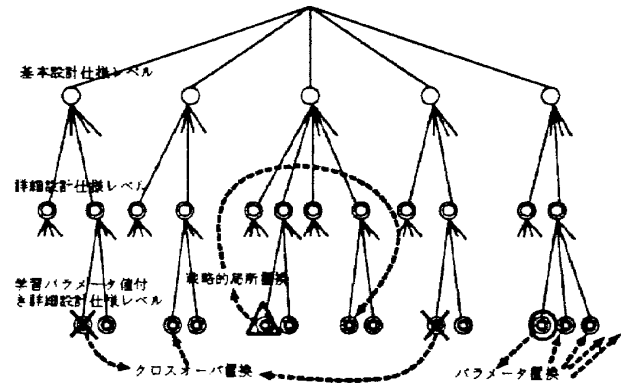


図1: 帰納的学習システムの構成空間

て詳細設計仕様に含まれるプロセスを別のプロセスに置換し、実際にシステムを実行して性能を評価する。このように局所置換を適用した結果得られた詳細設計仕様を実行しユーザの与えた目標精度を満たせば終了するが、満たさない場合、大域置換を適用して別の基本設計仕様を生成した後に貪欲的に局所置換を実行し、目標精度を満たすまでこのプロセスを繰り返し適用する。

4 並列処理型 CAMLET

並列処理型 CAMLETは、学習パラメータ値までも同定することを目標として、図1の探索木全体を探索する。

探索木としては基本設計仕様レベルの節点数は5節点に限定し、一つの基本設計仕様レベルの節点にリンクした詳細設計仕様レベルの節点数は75(=5×5×3)より1575(=5×5×3×3×7)節点となるが、一つの詳細設計仕様レベルの節点にリンクした学習パラメータ情報付き詳細設計仕様レベル節点数は無限個存在するため、無限探索木になる。以上のことから逐次処理型 CAMLETで考察した貪欲的探索法では実時間で解を見出すことができない。

そこで、並列処理型 CAMLETでは、仕様の部分構造に対する評価値を利用したヒューリスティック探索、並びに、多点探索と特徴として大域探索に優れる遺伝的プログラミング(GP)[伊庭,1996]の方法を取り入れた、並列処理向きの探索方法について考察する。

4.1 探索のための仕様表現

帰納的学習システムの仕様は、詳細概念部品の組み合わせにより表現されるため、GP的オペレータの適

用レベルは詳細概念部品とすることが即座に考えられるが、適用レベルの粒度を粗くすれば、部分的な評価を考慮しながらヒューリスティックに探索を進めることが可能になる。そこで本節では、評価可能な部分制御構造を仕様の表現単位とする。すなわち、図2に示した4種類の部分制御構造に対して、データセット全体に対する性能評価、学習されるクラス毎の性能評価、学習されたルールの特徴（条件部の長さなど）を評価値として付与し、探索に利用する。

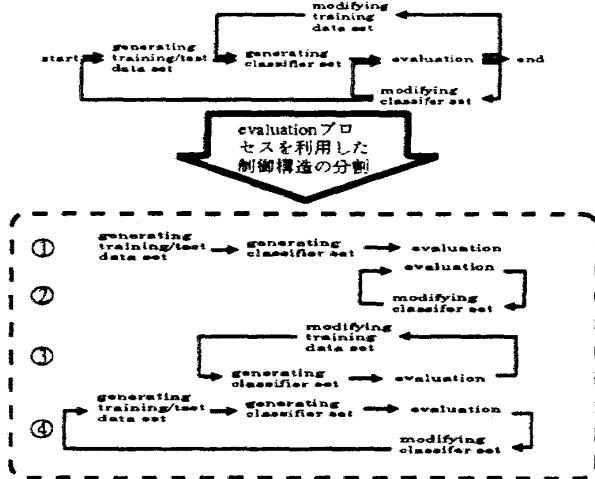


図2: 制御構造の分割

4.2 探索オペレータ

部分制御構造により表現された仕様に対して、現在、下記の3種類の探索（置換）オペレータを考えている。

交差置換: 2つの仕様に対して、各々に含まれる評価値の高い部分制御構造を残す形で、新しい仕様を生成する。

ヒューリスティック置換: 評価値と置換すべき部分制御構造の関連を言及するヒューリスティックを利用して、部分制御構造を変更する。

ランダム置換: 詳細概念部品のもつ学習パラメータについては、変更のためのヒューリスティックが未整理であるため、現時点では、ランダムに変更する。

4.3 探索の制御

基本的に、構成されたシステムの性能が低い時は、システムの制御構造を大きく変化させて対処（大域探索）する必要があり、ある程度性能が高ければ、システムの制御構造は保持したままで細部を変更（局所探索）することにより、対処できる可能性が高い。そこで、以下のように、前節で整理した3種類の置換オペレータを2つの閾値（低閾値と高閾値）により適用制御することを考える。

システム性能 < 低閾値: 交差置換の適用（制御構造の変更）

低閾値 ≤ システム性能 < 高閾値: ヒューリスティック置換の適用（詳細概念部品の変更）

高閾値 < システム性能: ランダム置換とヒューリスティック置換の適用（学習パラメータのみの変更と詳細概念部品の変更）

本制御方法は、大域探索と局所探索の単純な融合（ハイブリッド化）を試みているだけであるが、本実験結果を出発点として発展させていきたいと考えている。

4.4 超並列計算機上での実現法

超並列計算機（AP1000, 最大256PE）には、PE間の通信方法やリソースには制約があるため、それらの制約を考慮した実装方法を考察する必要があるので、3種類の置換オペレータの実装は下記の通りとなる。

交差置換: ホストが2つのPEから仕様を受信し、別のPEに生成された新しい仕様を送信し、実行させる。

ヒューリスティック置換: 単一のPE内で処理する。

ランダム置換: 学習パラメータの変更は、多くの代替案があるので、ホストがその制御にあたる。

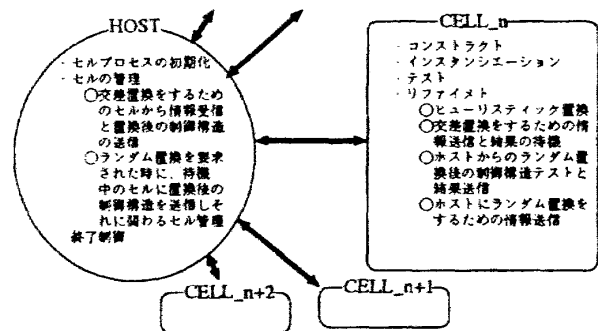


図3: 超並列計算機上での探索制御法の実現

現在、図3のような形で実装を進めている所であるが、リソースの制約から、3種類の置換オペレータの実行に関する制御構造を定める必要がある。この点については、今後の課題としたい。

5 おわりに

本稿では、CAMLETにおける、帰納的学習システムの仕様探索の並列処理化について考察した。通常探索と同様に、大域探索と局所探索の融合法がポイントになると考えられるが、超並列計算機上での実験結果を重視しながら、有効な探索技法を探りたい。実験結果と考察は、発表当日に報告する予定である。

参考文献

[根岸,1998] 根岸 直矢, 酢山 明弘, 山口 高平, “CAMLET: 帰納的学習システム構築支援環境 (1) -仕様と実装の対応付け-”, 情報処理学会第56回全国大会講演論文集, 2M-01, 1998

[伊庭,1996] 伊庭 齊志, “遺伝的プログラミング”, 東京電機大学出版局, 1996