

TM-1244他

情報処理学会第46回全国大会論文集

March, 1993

© 1993, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03)3456-3191 ~ 5
Telex ICOT J32964

Institute for New Generation Computer Technology

TM1244	構造写像エンジンの並列化	山路 孝浩、森下 太郎、 椿 和弘 (シャープ)
TM1245	並列推論マシン上の事例検索方式	椿 和弘、森下 太郎、 山路 孝浩 (シャープ)
TM1247	並列推論マシン PIM/i における KL1 言語処理系の改良	加藤 研児、久野 英治、 六沢 一昭、武田 浩一、 大原 輝彦 (沖)
TM1248	並列推論マシン PIM/i のバスの動 特性の解析	武田 浩一、大原 輝彦 (沖)

構造写像エンジンの並列化

山路孝浩、森下太朗、椿 和弘

シャープ(株) 情報技術開発センター

1. はじめに

我々は構造写像エンジン(Structure Mapping Engine; 以下SME)を利用した事例に基づく回路設計支援システムの開発をPIM上で行って来た([椿91], [椿93])。SMEは、対象の多少の異りを構造の類似により吸収し、かつ類似度を系統的に評価しながら類推を行うことができるという事例ベース推論向きの性質を持つ。しかしながら、SMEには、大規模な構造を持つ問題に対して膨大な処理時間がかかるという問題が存在する([山路91])。処理の負荷の中心は、部分解の組み合わせ処理を行う部分マッチ処理である。この部分に対して幾つかの負荷分散方式を用意し、PIM上で評価実験を行ったのでここに報告する。

2. 部分マッチ処理部の並列化

部分マッチ処理部は、処理の途中段階で求められた部分解(Gmap)を結合の可否を検査しながら組み合わせ、結合しうる最大の部分解の集合(Gmap集合)を求める処理である。この処理は、各部分解同士の2項間の結合可否をビット表現(1が結合可を表す)した転置正方行列から、全要素が1からなる最大の部分正方行列を求める処理と見なすことができる。

この部分マッチ処理部に対して、まず2分岐計算アルゴリズム(ある要素を含む、含まないに分け再帰的に結合集合を計算していく処理)を使い、ある分岐段数になれば、それ以降の処理を割り付けられたプロセッサ上で行う負荷分散方式(段数方式:次節参照)を用意した。しかしながら、この方式では問題によりその段数で生成されるサブタスクの大きさが均一でなく、負荷のバラツキが生じるという問題が生じた。そこで、サブタスクの大きさを反映するパラメータであると考えられる“結合数”(次節参照)に応じてサブタスクを割り振れば、負荷はより均一になるという仮説を立て、負荷分散方式の改良(結合数方式:次節参照)を行った。

更に、分岐数が多いほど並列向きであるという仮説を元にして、多分岐アルゴリズムを用意しそのアルゴリズムに対して、2分岐法と同様な2つの負荷分散方式の適用を試みた。

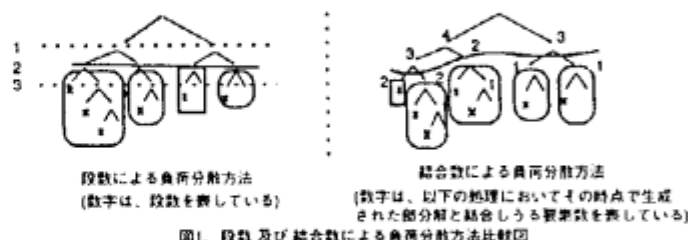


図1. 段数及び結合数による負荷分散方法比較図

3. 段数及び結合数による負荷分散方式

段数方式とは、ある指定する段数になるまで、各プロセッサに生成されるサブタスクを割り振り、指定数になれば割り振りを終了し振られたプロセッサ上で以後の処理を実行する方法である(図1参照)。

一方、結合数方式とは、2分岐法においては結合しうる要素数(その時点において求められた部分解と以下の処理において結合する可能性を持つ要素数)が、多分岐法においては生成される部分行列の次数が、それぞれ指定数以下になるまで各プロセッサにサブタスクを割り振り、それ以下になれば割り振りを終了し割り振られたプロセッサ上で以後の処理を実行する方法である(図1参照)。

尚、両負荷分散方式ともICOTで開発された負荷バランスユーティリティを利用して、基本的な負荷分散の処理を行っている。

4. 2分岐法と多分岐法

多分岐法は、行毎にプロセスを分岐させるアルゴリズムである(図2参照)。

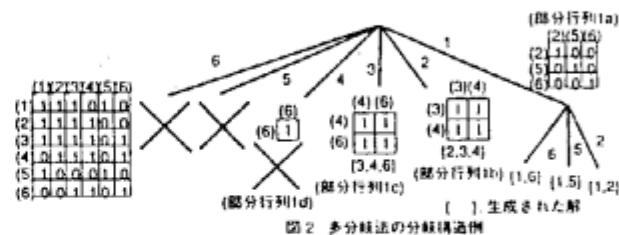
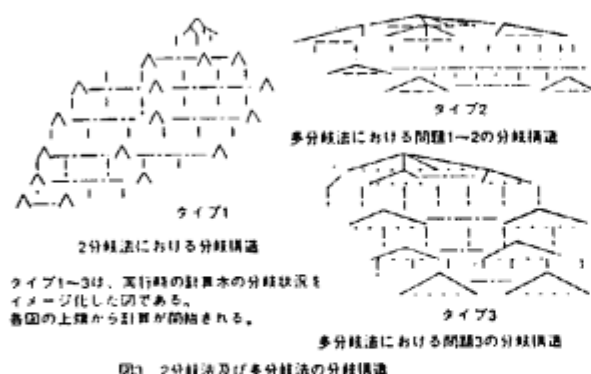


図2. 多分岐法の分岐構造例

2分岐法と多分岐法との大きな違いは、分岐時における枝数である。これにより、実行開始から計算終了までの間に形づくられる計算木の分岐構造は、大きく異なる。図3に、並列実験に用いた問題3題に対する分岐構造のイメージを示した。

尚、並列実験結果より、処理時間は多分岐法の方が短

く、台数効果曲線は2分岐法の方が傾きが良いという結果を得た(表1参照)。



5. 負荷分散方式の評価

5.1 段数方式と結合数方式の比較評価

表1にMulti-PSI 64PEによる各負荷分散方式実行結果の比較表を示す。2分岐法において、段数方式と結合数方式を比較すると、結合数方式の方が、全ての問題において、処理時間が短い。これは、2分岐法の分岐構造が、タイプ1の形をしており、各プロセッサに割り付けるサブタスクの負荷にバラツキが生じることによる。一方、多分岐法における問題1,2はタイプ2の分岐構造をしていることにより、段数による割り付けを行っても、ほぼサブタスクの負荷が均一になる。問題3は、タイプ3の分岐構造をなし、浅い段階で割り付けを行っていることにより、サブタスクの負荷にかなりの差が生じる。深い段階で割り付ければ、結合数による方法の数値にかなり近付くと予想される。しかしながら、全ての問題に妥当である段数を設定することは、一般に困難である。それに比べて、結合数方式は問題に対する依存度が少なく有効であることが確認出来る。

5.2 結合数による負荷分散方式の評価

結合数方式について、PIM/m上において、詳しく調査

	Multi-PSI	段数方式 (64台)	結合数方式 (64台)	PIM/m (結合数方式)	
	逐次時間	固定数: 2分岐 = 30 多分岐 = 4	固定数: 2分岐 = 10 多分岐 = 12	逐次	256台
問題1	1625050 331786	40012(40.0) 7629(43.5)	20007(56.0) 7380(45.0)	615840 118285	5008(120.8) 2756(42.9)
問題2	18073684 4281891	369658(48.9) 72267(50.3)	307613(58.8) 72149(50.3)	7271917 1504760	42221(172.2) 22301(67.5)
問題3	----- 5196469	601757(---) 884859(5.9)	527069(---) 86054(60.4)	12033610 1976592	53025(238.3) 15449(127.9)

表内の数字は、台数効果を表している。また、各問題における上段は、2分岐法による計算時間、下段は、多分岐法による計算時間を表している。単位は、msecである。

表1 各負荷分散方式の比較及びPIM/m上における計算時間

した。その実験結果は、図4及び、表1である。多分岐法に関して、64台までの台数効果曲線はリニアに伸びているが、それ以降伸びていない。2分岐法では、128, 256と台数が多くなるにつれ台数効果曲線は鈍くなっているが、良いもので200倍以上の台数効果を得ている。これはそれぞれの分岐構造によるものと考えられる。タイプ1(2分岐法)の分岐構造は分岐段数が深いので、途中段階においてその都度負荷の重いサブタスクが多数生成される。一方、タイプ2,3(多分岐法)に関して、タイプ2は浅い分岐構造をなしていることより、負荷の軽いサブタスクがほぼ同時に多数生成される。通信コストの問題と絡み合い、64台以降の台数効果が伸びていないと思われる。タイプ3は、タイプ2よりサブタスク生成率は低く、かつ、分岐構造が深いことより負荷の重いサブタスクが多く生成される。このためタイプ2よりも大きな台数効果が得られていると思われる。

6. おわりに

組み合わせ爆発を伴う高負荷な計算問題に対して、サブタスクの負荷を反映するパラメータにより、均等に負荷を分散させる方式の有効性をPIM/m上で確認することが出来た。

並列版SMEは、若干の機能拡張を施した逐次版とともにICOTのフリーソフトウェアとして再登録する予定である。本研究は、ICOT受託研究の一環として行われているものであり、本研究の機会を与えて頂いたICOT第2研究室の新田室長に深く感謝致します。

[参考文献]

- [橋 91] 橋和弘、他、構造画像エンジンを利用した設計支援の考察、情報処理学会第43回全国大会、1991。
- [橋 93] 橋和弘、他、並列推論マシン上の事例検索方式、情報処理学会第46回全国大会、1993。
- [山路 91] 山路孝浩、他、構造画像エンジンの高速化、情報処理学会第43回全国大会、1991。

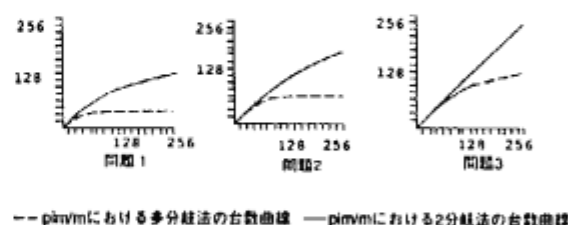


図4 PIM/mによる台数曲線