

高並列推論エンジンPIEにおける

4P-13

ゴールメモリ構成法に関する一考察

平田 圭二 , 相田 仁 , 後藤 厚宏 , 田中 英彦 , 元岡 達

( 東京大学 工学部 )

1. はじめに

知識情報処理の分野、特に自然言語処理などにおいては、大きな構造データ(リスト、構造体)を扱う場合が多い。現在開発中の高並列推論エンジンPIE[1]で構造データをより高速、大量に処理できる処理方式及びアーキテクチャに関する検討、提案を行なう。

2. 構造データと処理方式

構造データはプログラムの性質にも依るが、かなり大きな記憶領域を必要とする。従って各GF(ゴールフレーム)間で構造データを共有することにより、記憶領域の縮小、構造データコピーの省略等が可能となる。また構造データ自身に操作(オペレータ)を送りつけることで演算を行なうこともできる。

構造データの取り扱いから見た種々の処理方式として以下のようなものが考えられる。(図1、表1)

表1. 構造データの取り扱いから見た処理方式

構造データ 書換え	コピー	共有
(MM内) 直接	①	②
(MM外) 間接	③ PIE (第一次モデル)	④

MM ... メモリモジュール(ゴールメモリ)

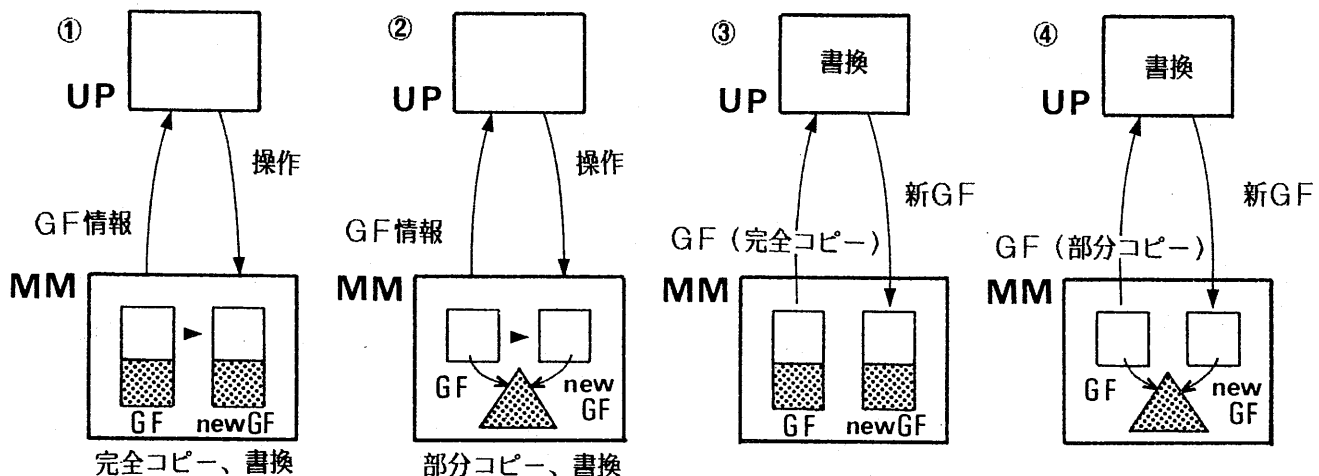


図1. 構造データの取り扱いから見た処理方式

**コピー**とは新GFが親GFの持つ構造データ全体をコピーして持つことである。**共有**とは、OR並列GF間、親子GF間で演算が施されなかった構造データ部分を共有することである。共有した場合、OR関係にあるGFが同一の大きな構造データ全体を個々にかかえる必要がないので、UPに送るデータ量は小さくて済む。従って、  
 ・ネットワーク負荷  
 ・構造データコピーの手間  
 ・必要なメモリ容量は減少する。逆に  
 ・単一化アルゴリズムやガーベジコレクション方式が複雑になり  
 ・メモリアクセス回数の増大を引き起こす。

**間接書換え**とは、MMからUPへGFを送り、UPでGFの書換えを行うことである。**直接書換え**とは、UPからMMに操作を送り、MM内でGFの書換えを行うことである。直接的にMM内で書換えを行えば、UP-MM間ネットワークは単一化に必要な情報のみを転送すればよい。従って、  
 ・ネットワーク負荷は減少するが、  
 ・メモリ側にGF情報抽出、書換えのための高度な処理機能が要求される。

共有及び直接書換えに共通な課題としては、  
 ・ネットワーク形態  
 ・ネットワークの応答性  
 ・負荷分散戦略がある。

3. 構造データ共有方式のシミュレータ

コピー/共有方式と、直接/間接書換え方式により得られる効率を別々に定量的に評価する必要がある。本シミュレータではPIEアーキテクチャにお



この例において、構造データ全体の内、第1レベルまでコピーしたストラクチャ部を含むGFをUPに転送すると、図2のようになる。この例の場合、単一化に必要な構造データはリストの先頭セル（第1レベル）までであるから、これ以上リストをたぐり、データをUPに持ってくる必要はない。

この例において、構造データ全体の内、第1レベルまでコピーしたストラクチャ部を含むGFをUPに転送すると、図2のようになる。この例の場合、単一化に必要な構造データはリストの先頭セル（第1レベル）までであるから、これ以上リストをたぐり、データをUPに持ってくる必要はない。

本シミュレータの評価項目は、構造データの共有率、分割レベル、GFの縮小化率、メモリアクセス回数等で、構造データ共有によってどの程度の効率化が図れるかを把握する。

GF書換え  
ある。従ってGFが  
含まない構造データ  
の未定義変数はUPに  
従って単一化の起動時  
送られる部分は次の三

化する為で  
定義変数  
、すべて  
UPに

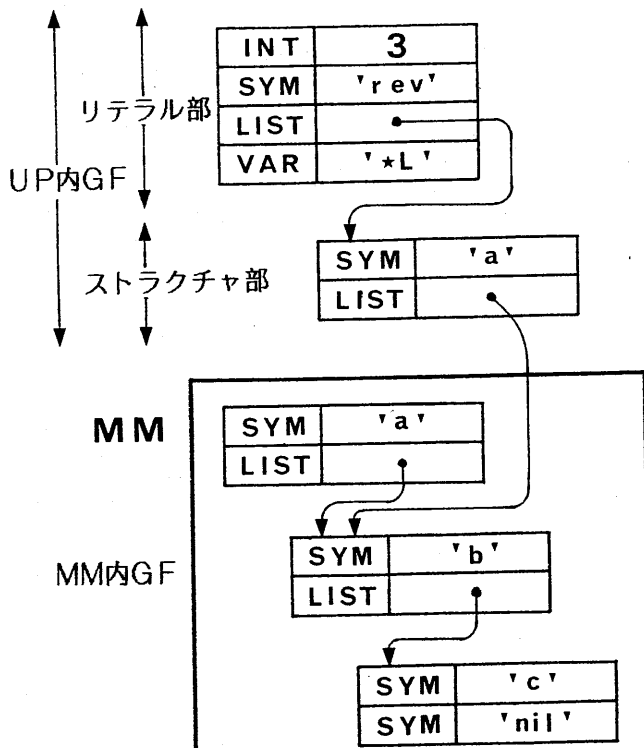
- 1) GF全体に関する情報
- 2) リテラル個々について
- 3) 未定義変数を含む構造データ

レ部  
スト

ただし、未定義変数を一つも含まない構造データは、第1レベル（ポインタを1回たどる深さ）までコピーしてUPに送り込む。多くの場合、構造データは照合操作において第1レベルまで必要とされるからである。

単一化の途中でさらに深いレベルのデータが必要となった時は、MMから1レベルずつ持ってくる。以下例を挙げて説明する。

```
GF: ?-rev ([a, b, c], *L).
定義節: rev ([*H|*T], *L)
        :- rev (.....)
```



#### 4. おわりに

本方式向きのアーキテクチャには図3のようなものが考えられる。UP、MMの結合網を階層化し、それぞれN台づつで1グループを作る。グループ内においてUP-MMは完全結合型ネットワークで接続されており、このグループ単位にセルを回収する。グループ内においてGFは構造データを共有するが、グループ間でGFを転送する時は完全コピーを行う。リングバスは負荷分散のためだけに用いられる。

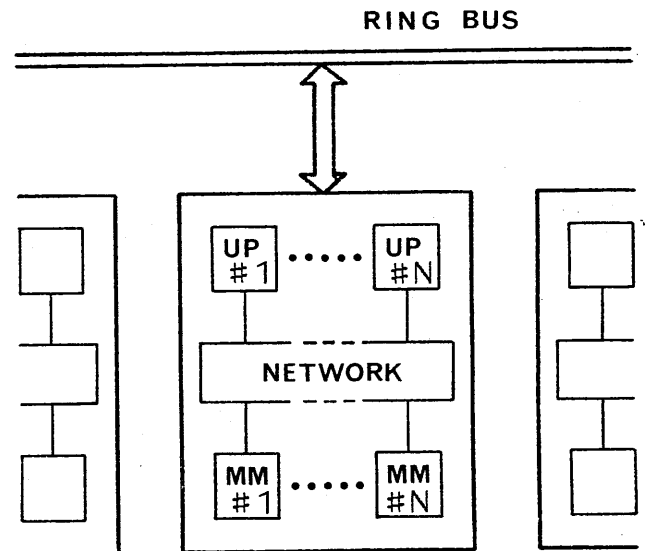


図3. 構造データ共有方式向きのアーキテクチャ

構造データの性質に着目して、PIEにおける構造データの共有方式について検討した。さらに未検討の項目として以下のことがある。

- イ) 未定義変数を含む構造データが共有された場合の単一化アルゴリズム
  - ロ) ガーベジコレクションのアルゴリズム
  - ハ) 1グループあたり最適なUP、MMの台数
- 今後シミュレーションによる定量的な効率の検討を考えている。

#### 《 参考文献 》

[1] 後藤 他：“高並列推論エンジンPIEにつ