

LRCによるインクリメンタルGCの評価

宮内 信仁

(三菱電機)

中川 貴之

(ICOT)

1. はじめに

本稿では、LRC方式のインクリメンタルGCを評価用逐次型KL1処理系に実装して動作の確認を行い、メモリ消費特性の評価をMRBインクリメンタルGCと比較して行なったので、その結果について報告する。

2. LRC方式の実行時の処理

LRC方式の処理系がMRB方式の処理系と大きく異なる点は、データオブジェクトが単一参照の状態から多重参照になるとき、データオブジェクトの手前に参照数カウンタを持つRC(Reference Count)セルを挿入し、参照数の増減に応じ参照数カウンタのメンテナンスを行なうことである[3]。この処理により単一参照のオブジェクトについてはMRB方式と同じコストで回収ができるとともに、MRB方式では回収不可能な多重参照のオブジェクトデータも回収が可能になる。次にLRC方式特有の実行時の処理について述べる。

2-1. RCセルの割り付け

KL1の抽象命令KL1-B[2]において、ゴールや構造体の引数のコピー命令には put_value, set_value, write_value がある。コピー時に参照数が増やす場合、MRB方式では図1のようにmark命令を用いて引数に現われたデータオブジェクトへのポイントのMRBをON(多重参照ポイント)にしてから単にポイントのコピーを行なえばよい。

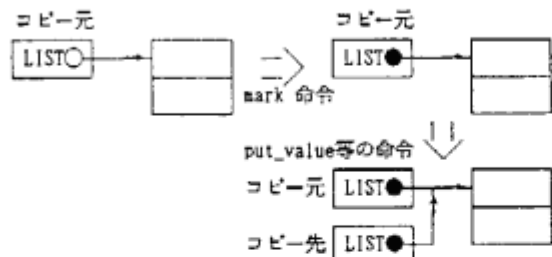


図1. MRB方式の単一参照ポイントの参照数増加のコピー

一方、LRC方式の引数のコピーには、コピー元の参照が消滅する参照数の増えないコピーと、コピー元とコピー先の両方から参照される参照数の増えるコピーとを区別する。参照数の増えるコピーにおいてコピー元が単一参照ポイントの場合にはRCセルを挿入する。このた

め、RCセルの分だけメモリの使用量が増える。

このほかに多重参照されている構造体の受動部でのユニフィケーションのときにも、要素に単一参照ポイントがあれば、要素を取り出し参照パスが増える場合には、RCセルを割り付けることになる。

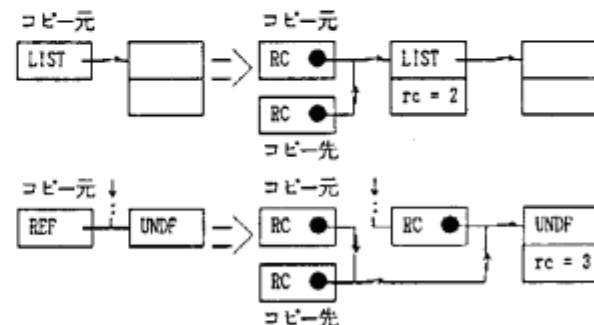


図2. LRC方式の単一参照ポイントの参照数増加のコピー

2-2. デレファレンス

デレファレンスを行なうとき、RCセルが参照パスに挿入されている場合は参照をたどる処理がMRB方式に比べて増える。

2-3. 参照数カウンタの更新とRCセルの回収

MRB方式では多重参照になったデータに対しては、参照数管理を行なわないので、後で単一参照になったとしても回収できない。一方、LRC方式では多重参照のデータの参照数管理のために、多重参照のオブジェクトデータの参照数を増やす処理とユニフィケーションや回収命令により参照数を減らす処理ではRCセルの参照数カウンタの参照、更新を行なうオーバーヘッドがある。

参照数管理の結果、デレファレンスやガーベジ回収のときにRCセルの参照数が1であれば、そのRCセルを回収する。

3. 評価結果

使用した評価プログラムは、Prime, 8Queen, Layered Queen, Bup である。評価プログラムでは次の3種類のデータセルを使用している。本処理系ではMRB処理系と同様にメモリをフリーリスト管理して再利用を行なう。

- (1) Value cell 1 word のセル
- (2) Cons cell 2 word のリスト構造体セル

Evaluation of the Incremental GC by LRC
Nobuhito MIYAUCHI, Takayuki NAKAGAWA
(MELCO) (ICOT)

(3) RC cell 2 word の RC セル
(参照数カウンタ付きセル)

3-1. メモリの使用状況

Prime ではMRB方式においてMRBが常にOFFであり全てのガーベジセルを回収できるので、LRC方式でもRC cell は全く使用されずMRB方式とLRC方式のメモリ使用特性は同じであった。

Prime を除くメモリの使用状況を図3に示す。フリーリストに返却した分を除いた実メモリ最大使用量がpeakに相当する。left は残存したガーベジセルの数であり、LRC方式ではループガーベジがないので当然0になる。

どのテストプログラムにおいてもLRC方式ではメモリ最大使用量がかなり縮小されている。LRC方式ではRC cell を余分に消費するので、そのメモリ消費の影響が大きいと予測されたが、MRB方式で回収できずに蓄積するガーベジセルに比較すれば、RC cell 全体の量はあまり多くはない。ガーベジセルは即時に回収されるため、他のValue cellやCons cellのメモリ消費も減少し、全体からみればメモリ最大使用量は減少している。

8Queen ではメモリ最大使用量が大きく減少した。これはMRB方式ではMRBがONのため回収できないガーベジセルが処理過程で大量に発生するためである。

RC cell のメモリ占有率の高いLayered Queen は、参照数カウンタのメインテナンスの処理の負担が大きいプログラムといえる。

実用的なプログラムに近いとされるBupではRC cellが10数%のメモリ占有率を示しており、単一参照データが大部分であることが予測される。

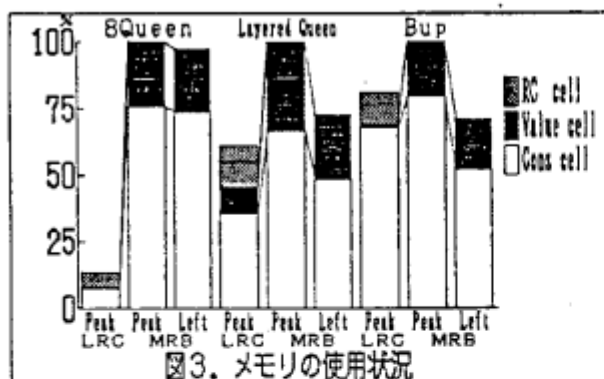


図3. メモリの使用状況

3-2. 動的なメモリ消費の傾向

図4に実際のメモリ消費の推移の一例として特徴的なLayered Queen の場合を示す。縦軸は消費メモリの総ワード数、横軸はKL1B命令ステップ数である。

Layered Queen ではある時期にまとまりのあるデータ回収が行われるが、各種セルが単調に消費される傾向が認められる。データが不要になった時点ですぐに回収処理が行われるLRC方式の特徴を示した非常に典型的な

例といえる。

なお8QueenやBupではLRC方式においても単調増加的なメモリ消費の傾向を示している。これはメモリの消費もメモリの回収も単調に行われるためである。

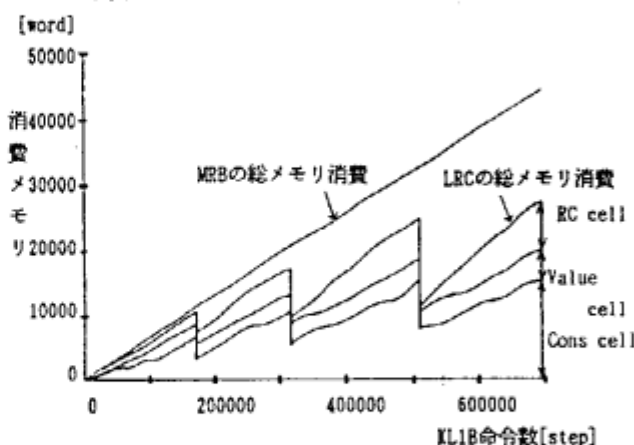


図4. 消費メモリの推移 (Layered Queen)

4. おわりに

LRC方式の長所は単一参照の状況が多いほどMRB方式に近いコストでインクリメンタルGCが可能であり、ループ以外のガーベジセルを全て回収できることである。

今回の評価により、LRC方式はメモリを効率良く使用しており、MRB方式と比較してワーキングセットが小さくできることが判明した。なおLRC方式ではMRB方式よりメモリアクセス回数が増加するが、ループガーベジ以外の全ガーベジセルの回収の代償として不可欠であり、ワーキングセット縮小によりキャッシュヒット率が向上するので他方式のGCよりそのコストは小さいことが予想される。メモリアクセス特性及びキャッシュ特性については別稿[4]を参照されたい。

今後の課題としてはさらにメモリアクセスを減少させる最適化が重要である。

参考文献

- [1] T. Chikayama and Y. Kimura :
"Multiple Reference Management in Flat
GBC". Proc. of ICLP'87, also as ICOT-TR248
- [2] Y. Kimura and T. Chikayama :
"An Abstract KL1 Machine and its
Instruction Set". Proc. of SLP'87, also as
ICOT-TR246
- [3] A. Goto et al. :
"LAZY REFERENCE COUNTING : An Incremental
Garbage Collection Method for Parallel
Inference Machines". Proc. of SLP'88
- [4] 中川 他:
Lazy-Reference-Counting GC 方式の評価
—共有バス結合マルチプロセッサへの適応性
並列処理に関する「火の国」ミニシンポジウム