

MRBによる多重参照管理方式

2Q-7 - K L 1 処理系における一括型 GC の特性評価 -

宮内信仁

木村康則

近山隆

久門耕一

(三菱電機)

(ICOT)

(富士通)

1. はじめに

本稿で対象としている K L 1 の評価用処理系では、MRB 方式による実時間 GC [1] を採用している。MRB は、複数の参照状態を確認して ON になると、処理が進んで単一の参照状態に移行しても、OFF にできない。これは、参照のパスが減少しても、参照パスが 1 つになったかどうかは、1 ビットの情報からでは、確認ができないからである。したがって、MRB 方式による実時間 GC では、MRB が一度 ON になると、その参照先のセルの回収は不可能なので、一括型の GC でこれらのセルを回収する必要がある。

本稿では一括型の GC としてコピー方式の GC を採用し、K L 1 処理系に実装した。コピー方式のアルゴリズムとして、今回は 3 種類の方式について検討を行ない、K L 1 処理系全体より眺めた動作特性を測定し、その評価を行なった。

2. 3 種類のコピー方式の GC

コピー方式の GC では、使用可能なデータ領域を 2 つに分けておき、処理系は半分の領域が使用できなくなるまで、処理を行なう。半分の領域中の自由領域を使ってしまうと GC が起動され、生きているデータだけが、旧領域から新領域にコピーされる。GC 時にコピー済みかどうかの判定に、GC ビットとして 1 ビットを使用する。

(1) 単純コピー方式 (Non Dereference GC と呼ぶ。)

生きているデータをデータ構造を変えずにそっくりそのままコピーする。GC の処理時間は生きているデータの量に比例する。

(2) デレファレンス方式 (Dereference GC と呼ぶ。)

コピー時で、不可視参照ポインタ (ref タグのセル) については、その参照先のデータセルを ref セル中にもってきてしまうデレファレンスを行なう。長所として、無駄な ref セルをコピーしないですむ上に、GC 時にデレファレンスを行なうので、GC 後の K L 1 処理過程での同じ作業を省ける。ただし、ref の連鎖をたどっていく途中のセルの MRB が 1 つでも ON になっていれば、デレファレンス結果の MRB を ON にしておく。この方式は、参照パスの消費により単一参照となったデータに関しても、MRB が ON のままであるという欠点があるが、GC のアルゴリズムとしては単純になる。

(3) MRB メンテナンス方式 (MRB Maintenance GC と呼ぶ。)

(2) のデレファレンス方式に加えて正確に MRB のメンテナンスを行なう方式である。旧領域にて ON になっている MRB でも、単一参照であれば OFF にして新領域にコピーする。これにより、MRB 方式の実時間 GC の回収効率が向上することになる。

3. MRB メンテナンス方式の GC

1 回全構造を渡るだけで MRB をメンテナンスすることのできるアルゴリズムについて検討した。

3-1. MRB のメンテナンスの基本的な考え方

基本的には、セルを新領域にコピーする際、その時点ではセルへの参照パスは 1 つしか見つからないので、コピー先の MRB は OFF にできる。コピーが進んで複数の参照パスのあることが認められたならば、参照を行なうセルの MRB を ON にする。

このためには、過去に現われた参照を行なうセルをアクセスできるようにしておく必要がある。普通の GC では、被参照セルが一度コピーされると、旧領域の被参照セルには新領域の被参照セルへのポインタを格納しておくが、(図 1) MRB Maintenance GC では、旧領域の被参照セルには新領域の 1 つ手前の参照側のセルへのポインタを格納することになる。(図 2)

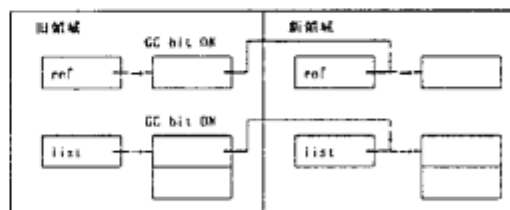


図 1. コピー時の旧領域のポインタの変更 (Non Dereference GC)

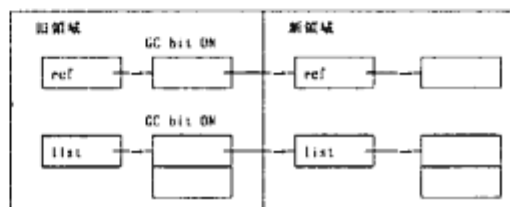


図 2. コピー時の旧領域のポインタの変更 (MRB Maintenance GC)

3-2. MRB の ON, OFF のアルゴリズム

(1) ref セルの MRB

ref セルの連鎖をデレファレンスすることにより、その結果が int, atom, list の各タグをもつセルならば、ref セルは消滅するので、ref セルの MRB のメンテナンスは必要ない。デレファレンス結果が undf, hook セルの場合にだけ、ref セルの MRB のメンテナンスを行なう必要がある。また MRB の定義により、undf, hook セルを指す ref セルの MRB は、ref セルが 2 つまでは OFF であり、3 つ以上ならば 1 つ以外は ON にしなければならない。

旧領域中の ref セルの数の情報として、GC ビットが ON、MRB が OFF であれば既に 1 つ ref セルがあることを示し、GC ビットが ON、MRB が ON であれば、2 つ以上の ref セルがあることを示す。(図 3) ここで不要になった旧領域中の MRB をカウンタビットとして代用している。

(2) 構造体へのポインタの MRB

構造体を指すポインタのセルの MRB は、基本と同様で単一参照ならば OFF であり、複数参照ならば、ON である。(図 4)

4. K L 1 処理系における一括型 GC の特性評価

今回、評価の対象とした試験プログラムは、素数を求める Prime、チェス盤上のクイーンの排他的配置問題を解く 8 Queens、構文解析を行なう Bup の 3 種類である。

GC の起動条件を表 1 のように決定して計測を行なった。

Multiple Reference Management by MRB - Copying GC -

Nobuhito MIYAUCHI, Yasunori KINURA, Takashi CHIKAYAMA and Koichi KUNON

(MELCO)

(ICOT)

(ICOT)

(FUJITSU)

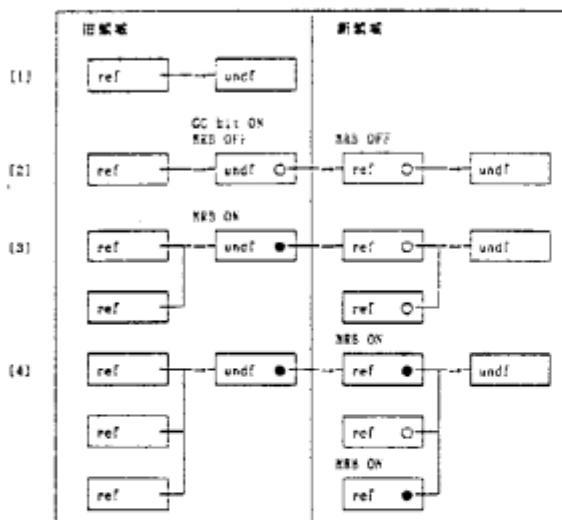


図3. refセルのMRBのメンテナンス

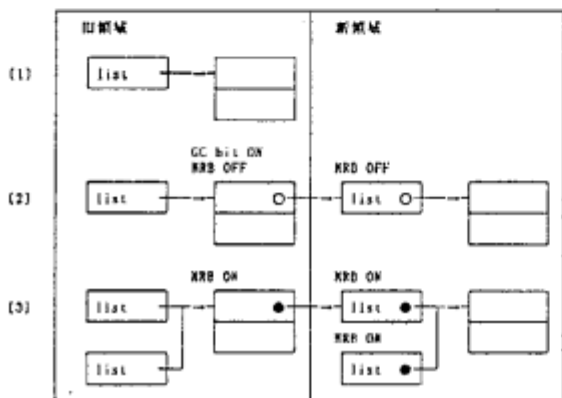


図4. 構造体へのポインタ(list)のMRBのメンテナンス

表1. GCの起動条件の設定

プログラム	value cell 最大消費量	cons cell 最大消費量	GCを起動する セル数の閾値
Prime 500	501	499	value, cons cell: 490
8 Queens	9774	6861	value cell: 5500
Bup	4156	7968	cons cell: 6800

表2. 各プログラムでのGCの回数

プログラム	GCの回数		
	Noe Derefer.	Dereference	MRB Maint.
Prime 500	23	28	20
8 Queens	32	1	1
Bup	3	42	3

デレファレンスの効果を図5の一括型GCの初回におけるrefセルの回収率により示す。8 Queens では、処理過程全般にわたって、refセルが多量につくられ、Dereference GCで大幅に除去される。8 Queensのメモリの消費セルの動特性を図6に示す。

MRBメンテナンスの効果はlistタグのセルのMRBのONからOFFへの変化によくみられる。初回一括型GCでの変化量を図7に示す。Bupにおいて、メンテナンスのされないDereference GCではMRBがONの状態に残されているために、実時間GCにてかなり除去されないリスト構造体セルが存在し、一括型GCの回数が増加している。

5. おわりに

評価を行なった3種類のプログラムについての一括型GCでは、

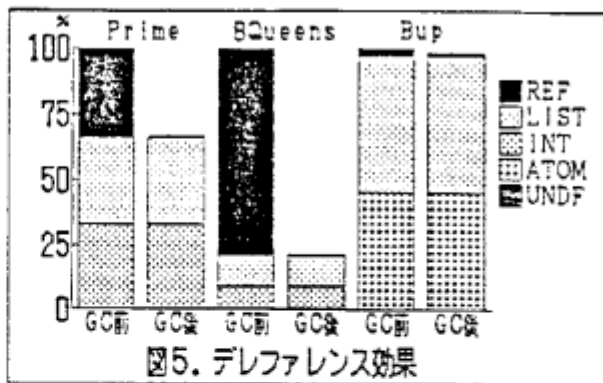


図5. デレファレンス効果

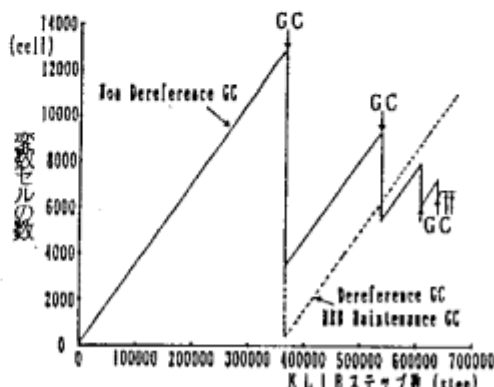


図6. 8 Queensにおける消費されているセル数の推移

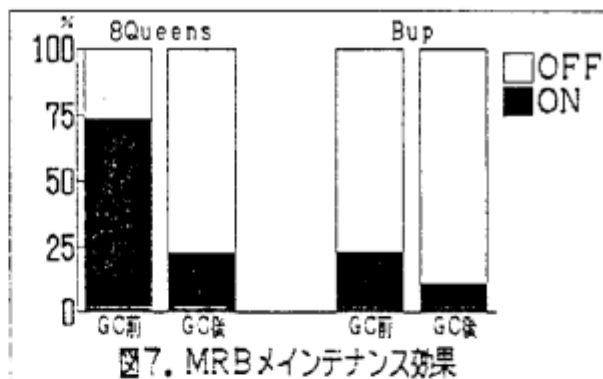


図7. MRBメンテナンス効果

(a) Prime 500 デレファレンス。MRBのメンテナンスは若干の寄与。

(b) 8 Queens デレファレンスの効果が大きい。

(c) Bup MRBのメンテナンスの効果が大きい。

Bupが最も実用的なプログラムであることを考え合わせると、一括型GCのアルゴリズムにデレファレンス、及びMRBのメンテナンスを取り入れることは効果がある。

実により多くのプログラムについてGCの特性の評価を行なう必要がある。今後は、各種GCのコストについても検討する予定である。

謝辞

一括型GC方式の実装、評価に関し、貴重なコメントを下された西田健次氏、後藤厚宏氏をはじめ、ICOTの方々に感謝します。

参考文献

- [1] T.Chikayama and Y.Kimura: "Multiple Reference Management in Flat GC", Proc. of ICLP'87