

2パス限定投機方式向け予測器のエントロピーによる妥当性評価

桐生 恵利佳[†] 金海 和宏^{††} 横田 隆史^{††} 大津 金光^{††} 大川 猛^{††} 馬場 敬信^{††}
[†]宇都宮大学工学部情報工学科 ^{††}宇都宮大学大学院工学研究科情報システム科学専攻

1 はじめに

プログラムのループ中には、条件分岐により複数の実行経路(パス)が存在する。我々は、その予測パスに対応する最適化コードを投機的マルチスレッド実行することでプログラム実行の高速化を実現する、2パス限定投機方式を開発している [1]。この方式によってプログラムの実行を高速化するには、次のイテレーションにおいて実行されるパスの予測機構(以下、パス予測器と呼ぶ)の性能が影響する。これまで様々な方式によるパス予測器が検討されているが [2]、予測成功率が低いプログラムに対して、その性能が限界であるのかについては明らかにされていない。

パス予測器の予測性能は、パスの出現系列の規則性に大きく依存する。その規則性をエントロピーによって測ることで評価対象プログラムで期待可能な予測成功率を算出する。実際の予測成功率と期待可能な予測成功率を比較することで、予測器の性能の妥当性について評価する。

2 パス予測器

1つのループ中では、多くの場合において、高々2つのパスの実行頻度が支配的であるということが分かっており、2パス限定投機方式では、この上位2つのパスに限定した投機実行を行う [1]。実行頻度が上位2本のパスを実行頻度が高い順に #1 パス、#2 パスと呼ぶ。

パス予測器は、それまでのパスの実行履歴を元に次に実行されるパスが #1 パスかそれ以外のパスかを予測する。現在、2パス限定投機方式では2レベル分岐予測器をパス予測に応用した2レベルパス予測方式を用いている。図1に2レベルパス予測器の構成を示す。

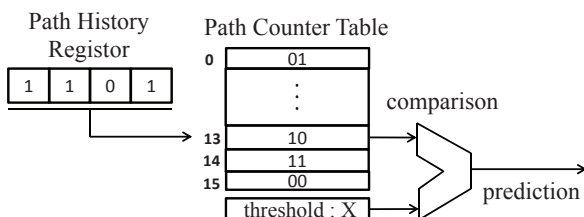


図 1: 2レベルパス予測器

2レベルパス予測方式は履歴レジスタとカウンタテーブルで構成され、履歴レジスタをインデックスとしてカウンタテーブル内の飽和カウンタを参照する。飽和カウンタの値に応じて次に実行されるパスを予測する。

3 情報エントロピー

一般に、ある情報について、その規則性が高いほど情報源のもつエントロピーは小さく予測は容易であり、逆に規則性が低いほど情報源のもつエントロピーは大きく予測は難しいとされている。

予測器の性能を表現するために、この考え方をパスの実行履歴に適用し、パスの出現系列の規則性をエントロピーにより測る。パスの履歴は、#1 パスを通った場合を1、#1以外のパスを0とし、1ビットのシンボルとみなす。

連続した n 個のパス履歴から成るシンボル列による n 次拡大随伴情報源エントロピー $H(\bar{S}^n)$ を、

$$H(\bar{S}^n) = - \sum_i p(S_i^n) \log_2 p(S_i^n) \quad (1)$$

のように定義する。 $p(S_i^n)$ は、シンボル列の出現パターン S_i の出現回数を、総出現回数で除算したものである。これは、2レベルパス予測器に置けるカウンタテーブルの各エントリの参照回数と、出現割合を求めていることと同義である。

$n+1$ 次の拡大随伴情報源エントロピー $H(\bar{S}^{n+1})$ が求めれば、もとの情報源 S のエントロピーの n 次近似 $H^n(S)$ は、

$$H^n(S) = H(\bar{S}^{n+1}) - H(\bar{S}^n) \quad (2)$$

で求めることが可能である。さらに、 $n \rightarrow \infty$ で、元の情報源の真のエントロピー $H(S)$ が求められる [3]。本研究では、このエントロピーを情報源エントロピーと定義する。

4 期待可能予測成功率

ある2値事象が確率 p で発生するとき、その情報量は

$$f(p) = -p \log_2 p - (1-p) \log_2 (1-p) \quad (3)$$

により求められる。逆に、ある2値事象の情報量が与えられた場合、その発生確率 p は式 (3) の逆演算により得られる。本稿では、情報源エントロピーの値を逆算して求めた確率を期待可能予測成功率と定義する。

Appropriateness of Path Predictors for Two-Path Limited Speculation Method

[†]Erika Kiryu, ^{††}Kazuhiro Kinkai, Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa and Takano Baba
 Department of Information Science, Faculty of Engineering, Utsunomiya University ([†])
 Graduate School of Engineering, Utsunomiya University (^{††})

5 SPEC CINT2000 による評価

予測器の性能評価を行うべく、各評価対象プログラムにおける実際の予測成功率 [2], パスの出現系列の規則性を測って算出したエントロピー値, 期待可能予測成功率を評価する. 評価対象は, SPEC CINT2000 の 164.bzip, 175.vpr, 176.gcc, 181.mcf, 186.crafty, 255.vortex, 256.bzip2, 300.twolf とし, データセットは train を用いる. ベンチマークプログラム内で使用するループの選定条件は, 少なくとも 2 本以上のパスが出現すること, およびループ中で関数の呼び出しを含まないものとし, イテレーション数上位 5 つのループを用いる.

上記の評価対象で 2 レベルパス予測方式を用いてパス予測をおこなった際 [2] の履歴レジスタ長と予測成功率との関係を図 2 に示す.

また, 図 3 にシンボル列の長さを 1 ビットから 30 ビットまで変化させたときの n 次拡大随伴情報源エントロピーの値を示す. ベンチマークプログラム名の横に記されている数字は, $n = 28$ の拡大随伴情報源エントロピーの値から算出した情報源エントロピー (単位は bits) である. 図 3 のグラフの n における傾き (情報源エントロピー) の大小と, 図 2 の予測成功率を見比べると, 傾きが小さいプログラムでは予測成功率が高く, 傾きが大きいプログラムでは予測成功率が低くなっている. このことから, 予測成功率と情報源エントロピーの間には相関性があると考えられる.

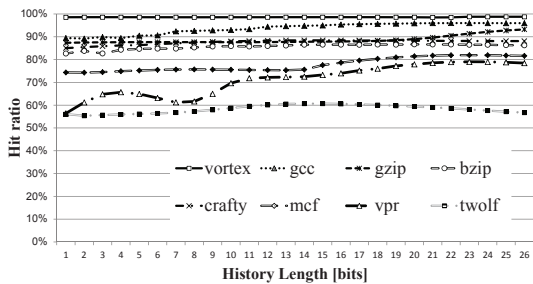


図 2: 2 レベルパス予測方式における履歴レジスタ長と予測成功率の関係

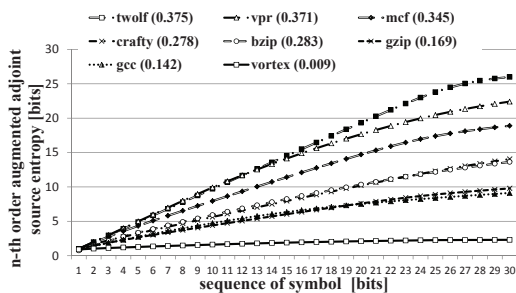


図 3: シンボル列長と n 次拡大随伴情報源エントロピーの関係 $H(\bar{S}^n)$

第 4 節で示したように, 情報源エントロピーから期待可能予測成功率を求める. 図 4 に予測成功率と期待可能予測成功率のベンチマークごとにおける比較結果を示す. 予測成功率は, 2 レベルパス予測方式とパス予測のシミュレーション中に特に高い性能を示した Gshare パス予測方式での結果である [2]. 図 4 から, プログラムによって予測成功率と期待可能予測成功率の差の大小は異なっている. このことから, 予測器は, パスの出現系列の規則性が高いプログラムにもかかわらず規則性を発見できない場合があると考えられる.

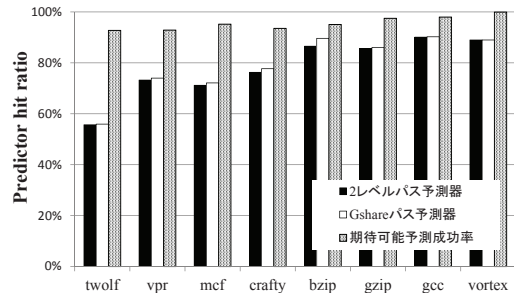


図 4: 予測成功率と期待可能予測成功率の比較

6 おわりに

本稿では, 2 パス限定投機方式向け予測器の性能の妥当性評価を行うため, パスの出現系列の規則性をエントロピーにより評価した. また, 情報源エントロピーから期待可能予測成功率を求め, 実際の予測成功率と比較することで, 各ベンチマークに対する予測器の妥当性について評価した. 評価の結果, 情報源エントロピーと実際の予測成功率には相関性が見受けられた. また, プログラムに応じて, 実際の予測成功率は期待可能予測成功率の差に大小があった. 今後は, 予測成功率が低かったプログラムについてその原因を検証し, プログラムによらず予測性能を発揮する高性能な予測器の開発をする予定である.

謝辞

本研究は, 一部日本学術振興会科学研究費補助金 (基盤研究 (C)24500055, 同 (C)24500054) による.

参考文献

- [1] 横田 隆史ほか: “2 パス限定投機方式の提案”, 情報処理学会論文誌:コンピューティングシステム, Vol.46, No.SIG 16 (ACS-12), pp.1-13, 2005.
- [2] K. Kinkai, T. Baba, H. Jutori, K. Ootsu, T. Ohkawa, and T. Yokota: “Comparative Study of Path Prediction Method for Speculative Loop Execution”, Proc. 3rd International Conference on Networking and Computing (ICNC), pp.283-287, 2012.
- [3] 横田 隆史ほか: “プログラムの実行挙動と分岐予測性能を表現するエントロピー”, 情報処理学会論文誌:コンピューティングシステム, Vol.48, No. SIG 18 (ACS 20), pp.12-31, 2007.