A-001

変数値域限定による充足可能性判定効率化手法の検討

中山　寛己 †　　千代　英一郎 †

† 成蹊大学大学院理工学研究科

1　はじめに

近年，プログラムの信頼性を高めるため，その不具合を早期に検出するための様々な検査方法が提案されている．プログラム検査を実現するアプローチのひとつとして，検査問題を充足可能性判定問題に帰着させて解く方法が存在する．これは性能向上の著しい SMT ソルバを活用するもので，ループ停止性，等価性，assert 文の検査など様々な種類の検査に応用されている．

このアプローチの問題のひとつに複雑な検査式に対する判定時間の長さがある．たとえば Gulwani らは，ループを含むプログラムの検査で必要となるループ不変条件を人間が記述する手間を省くために，あらかじめ用意したテンプレート（プログラム変数およびテンプレートパラメータを含む不等式を論理和および論理積で組み合わせたもの）の中から，条件を満たすものを探すために SMT ソルバを用いている [1]．しかしながら，これにより生成される検査式は高次の多項式を含むため，テンプレートが大きくなると z3 等の現代の代表的な SMT ソルバであっても判定が数日たっても終了しなくなることがある．

我々は，このような現在の SMT ソルバでは判定の難しい問題に対し，解の範囲（変数の値域）を限定した判定の繰り返しによって確率的に判定を可能とするアプローチから取り組んでいる．本発表では，本アプローチにもとづき考案した 2 つの値域限定方法およびその効果の予備評価結果について報告する．

2　提案手法

本章では変数の値域を限定した判定の繰り返しによる検査手法の考え方について説明する．多くの場合，検査式を充足する値域が整数である変数の値の探索は，各変数の取りうる値域が小さくなればより効率的に行える．この場合 1 度の検査で得られる情報は限定的なものになるが，限定する範囲を変化させながら繰り返し検査を行うことで次第に十分な情報が得られるようになると考えられる．

変数値域の限定にあたっては，限定する変数の選択およびその値域の場所・大きさの検討が必要になる．一般に限定する変数の数を増やし，値域の大きさを小さくするほど，判定時間は短くなるが，検査式を充足する変数の値を見つけるまでに必要となる繰り返しの回数は増えるというトレードオフが存在する．値域を限定する場所に関しては，検査式の簡単な解析により，明らかに解が存在しないことがわかる範囲を除くことが望ましい．たとえば検査式の中に $x > 0$ という制約が存在すれば，x が 0 以下の範囲を含めるのは無駄である．以降ではこのような範囲を除いた値域を実効値域とよぶ．

これらをふまえて設計した方法を以下に述べる．なお以降では限定する値域の大きさのことを値域幅とよぶ．

2.1　値域幅限定法

変数に対して 1 以上の値域幅で値域の限定を行う．その際にどの変数の値域を限定するかはランダムとし，その値域幅と，変数の数は検査ごとに指定するものとする．値域幅を小さくしたり，値域を限定する変数の数を増やすことで検査時間をより短くできると考えられるので，結果が返ってこない場合は，より限定的な検査を行うことで，変化させていきながら検査を行う．注意しなくていけない点としては，値域幅を大きく取りすぎると検査に時間がかかりやすくなり，小さく取りすぎると式が真となるような変数への値の割当を見つけることが難しくなる点がある．しかしながら，検査式によっては，全変数の値域幅を最小の 1 とし，検査を行っても繰り返し検査を行っていくと結果が返ってこないことがある．そのような検査式に対しては，2.2 節で述べる手法を使用していく．

2.2　値代入法

いくつかの変数に実効値域よりランダムに値をひとつ選択して代入する．その際にどの変数に値を指定するかはランダムとし，指定する数は検査ごとに任意とする．その他の変数については値域幅限定法と同様に扱う．これにより値域幅限定法では検査を行えなかった式に対しても，短い時間で検査できることがあると考えられる．デメリットとしては，1 回の判定で式をが真となるような変数への値の割当を見つけられる確率が低いことである．これはランダムに値を指定してもそれが式が真となるような変数の値であることはほとんどないことが原因である．

3　評価

本章ではプログラム検査で用いる検査式を用いて提案手法の予備評価を行った結果について述べる．2 章であげた 2 つの手法について，そのパラメータを変化させたときに判定結果が得られる確率がどう変化するかを測定した．これは提案手法の前提である，値域の限定によって判定結果が短時間で得られるようになることを実証するためのものである．本来であれば，検査式を充足する解が得られるまでの繰り返し回数や検査時間の平均値等を評価すべきであるが，現在のところその段階には至っていない．また，実効値域については，現在その解析が未実装であるため，すべての評価において充足解を含む区間である [-1000,1000] を用いた．

なお，測定時には 1 時間待っても結果が返って来なかったものに関しては，その検査は結果を返さなかったものとした．評価環境は OS が CentOS 6.4, CPU が AMD Opteron 6386SE2.8GHz，メモリが 256GB である．

3.1　評価 1

対象とした検査式はループ停止性検査を目的に作成され，変数を 25 個持ち，そのままでは 1 日待っても結果が得られないものである．最大で 3 次の式を含む．また加算，減算，乗算により構成され，不等式の論理積となっている．

3.1.1　値域幅限定法

全変数の値域を限定し，検査を行ったが，結果は返ってこなかった．そこで，値域幅を変化させながらそれぞれ 100 回検査を行いどのくらい結果が返ってくるかを調べた．値域幅を 500 から少しずつ広げていったときの結果を図 1 に示す．値域幅が 900 以下であるときは，ある程度結果が返ってきたにもかかわらず，1000 以上にした際には結果はひとつも得られなかった．この検査によって得られた結果は全て unsat であった．

3.1.2　値代入法

全変数の値域を限定し，いくつかの変数に実効値域よりランダムに値を選び代入し，代入する変数の数を変化させていったとき，それぞれどのくらい結果が返ってくるかを調べた．値を指定する変数以外の変数については，値域を実効値域である [-1000,1000] とした．結果を図 2 に示す．値を指定する変数の数を最低の 1 にしても，65% の割合で結果を得られた．この検査によって得られた結果は全て unsat であった．

3.2　評価 2

対象とした検査式はループ停止性検査を目的に作成され，変数を 40 個持ち，そのままでは 1 日待っても結果が得られないものである．値域は整数であり，最大で 3 次の式を含む．また加算，減算，乗算により構成され，不等式の論理積となっている．

Efficient satisfiability checking based on range restriction of variables

†Hiroki Nakayama　†Eichiro Chishiro

†Graduate School of Science and Technology, Seikei University

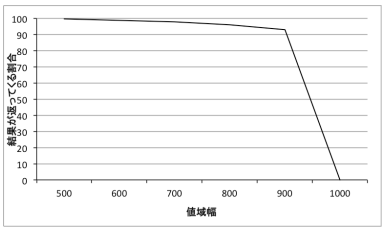

図 1　値域幅に対する結果が返ってくる割合の変化（評価 1）

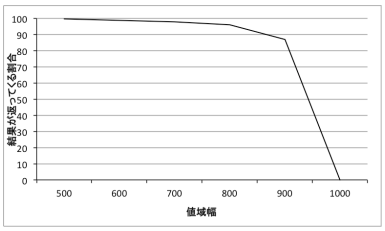

図 3　値域幅に対する結果が返ってくる割合の変化（評価 2）

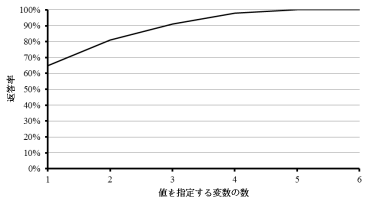

図 2　指定変数の数に対する結果が返ってくる割合の変化（評価 1）

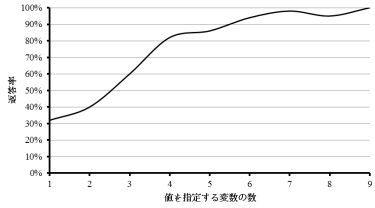

図 4　指定変数の数に対する結果が返ってくる割合の変化（評価 2）

3.2.1　値域幅限定法

始めに全変数の値域を限定し，検査を行ったが結果は返ってこなかった．そこで値域幅を変化させながら 100 回検査を行い，それぞれどのくらい結果が返ってくるかを調べた．値域幅は 500 より徐々に大きくしていった．結果を図 3 に示す．値域幅が 900 以下であるときは，ある程度結果が返ってきたにもかかわらず，1000 以上にした際には結果はひとつも得られなかった．この検査によって得られた結果は全て unsat であった．

3.2.2　値代入法

全変数の値域を限定し，いくつかの変数に実効値域よりランダムに値を選び代入し，代入する変数の数を変化させていったとき，それぞれどのくらい結果が返ってくるかを調べた．値を指定する変数以外の変数については，値域を実効値域である [-1000,1000] とした．結果を図 4 に示す．値を指定する変数の数を最低の 1 にしたとき，結果が返ってきたのは 32% であった．この検査によって得られた結果は全て unsat であった．

3.3　結果の検討

評価 1 では，値代入法において，値を代入する変数の数が最低である 1 のときでも 65%，2 つの変数に値を代入した際には，80% 以上で結果が返ってきた．値を指定する変数の数が少なくても，結果が得られる割合は高かった．すなわち評価 1 では，値代入法の 1 回の検査で結果が見つかる確率は低いというデメリットの影響は小さい．

また値域幅限定法についても値域幅を 900 以下とすると 90% 以上の割合で結果を得られた．この結果より，評価 1 で使用した検査式に関しては，どちらの手法も有効であると考えられるが，1 回の検査での範囲を比べたとき，値域幅限定法の方が大きくなるので，値域幅限定法で検査を行った方が良いと考えられる．

評価 2 では，値域幅限定法において，値域幅 900 のとき，結果が返ってきたのは 80% であり，値域幅が 900 以下とすれば十分結果が返ってくる事が期待できることがわかった．

また値代入法においては，評価 1 のときと違い，値を指定する変数の数がひとつや 2 つのように少ないとき，結果が得られる割合は 30% ほどであまり結果が得られる期待はできないとわかった．もし，評価 2 で使用した式を値代入法で検査を行うとすると，より高い割合で結果を得るためには値を指定する変数の数を増やして検査することが必要である．しかしながら，そのとき得られる結果は値域幅限定法によって得られる結果と比べて，極めて限定的となる．値域幅限定法と値代入法において，十分に結果が期待できる 90% 以上結果が得られときを比べたとき，一度に検査できる範囲を比べると，値域幅限定法の方が広いことより，評価 2 で使用した検査式に関しては，値域幅限定法を使用した方が良いと考えられる．

評価 1 と 2 の結果を比べたとき，値域幅限定法に関して値域幅を変化させていったときの，結果が得られる割合にはあまり差はなかったが，値代入法において，値を指定する変数の数がひとつや 2 つといった少ない場合における結果が返ってくる割合に差がでた．これは評価 1 と 2 の検査式を比べたとき，評価 2 で使用した検査式の方が変数の数は 15 個多く，検査式の長さは 2 倍長いため，より検査式が複雑であるため，効率化の差が出たと考えられる．しかしながら，値域幅限定法では差がでなかったことに理由に関しては，現段階では判明していない．

2 つの検査式に共通して，値域幅限定法において，値域幅を 1000 以上にした際に，1 回も結果を得ることができなかった．これは，検査時間は値域幅に対し指数関数的に増加するためと考えられ，値域幅 900 と 1000 の間で急激に増加してしまったと考えられる．検査を行う検査式や実効値域により，結果が返ってこなくなる値域幅は変わってくると考えられるが，どんな検査式にも結果が返ってこなくなる値域幅の閾値が存在するのではないかと考えることができた．

4　おわりに

本研究では，プログラム検査を充足可能性判定問題に帰着させて解く際に，検査式中の変数の解の範囲を限定することで 1 度の検査にかかる時間を減らし，限定する範囲を変化させながら繰り返し検査を行うことで検査式が真となるような変数への値の割当が存在するかどうか推定する方法について検討を行った．予備評価の結果，いくつかの検査式に対して，解を見つけることはできなかったが，最終的な結果を推定を行う手掛かりとなる限られた範囲に式が真となるような変数への値の割当が存在するかどうかといった限定的な情報を得ることができた．また値域幅限定法において，値域幅を変化させていったとき，結果を得ることができる値域幅の閾値があることがわかった．

今後の課題としては 最終的な結果を推定する上で適切な値域幅の選択方法と変数の選択方法に関する評価などがある．また値域幅限定法と値代入法において，条件を変化させながら検査を行い，より明確に検査式ごとにそれぞれどのくらい結果が返ってくるかを調べることである．特に値代入法において，値を指定する変数以外の値域幅を変化させていったときの結果が得られる割合の変化を調べることである．

参考文献

[1] Gulwani, S., Srivastava, S. and Venkatesan, R.: Program Analysis as Constraint Solving, *PLDI*, pp. 281–292 (2008).