1R−4

純リスプに基づいた頭語的評価関数の概要
- Chaitin のアルゴリズムの欠陥とその修正 -

鎮目浩輔
図書館情報大学

1．まえがき

アルゴリズム的情報理論（Algorithmic information theory）は，複雑さに対する定量的な尺度を与え，その性質や応用を研究する理論である。0と1からなる記号列、即ち binary string s が与えられたとき、その複雑さ(complexity)は、"sを出力する最短のアルゴリズム（プログラム）の長さ"として定義される。プログラムの長さを決めるにはそれを実行する計算機、すなわち計算可能な部分評価関数（以下では単に部分評価関数と呼ぶ)が必要であり、complexityの性質はこの部分評価関数の選び方に依存する。そこでChaitinはこの部分評価関数としてその定義域が頭語条件を満たすものを用いれば、complexityがShannonの情報論的エントロピーに類似した、応用上重要な性質を持つことを示した。そして頭語条件を満たす部分評価関数Uが実際に存在することを示すためにまずTuring machineに基づいたUのアルゴリズムを提案し、次に一種の純リスプに基づいたアルゴリズムを与えた[1]。しかし、この純リスプに基づいたアルゴリズムには、Chaitinが与えたままでは欠陥があり，Uの定義域が実際には頭語条件を満たさなくなってしまう。本発表ではそのことを示す。さらに、実際に頭語条件を満たすようアルゴリズムの修正を行なう。

2．ChaitinによるUの定義

Chaitinは次のI,IIの手順でUを定義した。

I. 一種の純リスプを定義し、それを用いて２つのbinary string p, qを引数とする部分的評価関数$V(p,q)$を作る。これは頭語条件を満たさない。

II. Vを元にして頭語条件を満たす関数Uを作る。

ここでは本質的な部分であるIIについてのみ説明を行なう。つまり問題は次の性質を持つ部分関数Uをいかにして定義するかというである：「引き数p,qに対し、$V(p,q)$が値を持ち、しかもpの任意のprefixまたはextension rに対して$V(r,q)$が値を持たなければ$U(p,q)=V(p,q)$、しかし$V(p,q)$と$V(r,q)$ の両方が値を持つ場合には$U(p,q)$と$U(r,q)$ の少なくとも一方は値を持つことが禁止され、両方が同時に値を持つことはない。」

こうするためにもっとも簡単なのは、次でUを定義することある。

$$U(p,q)=\begin{cases} V(p,q) & (p\text{のprefix } r \text{ 全てに対し} V(r,q)\text{が値を持たない場合}) \\ \text{値を持たない（上記以外の場合）} & \end{cases}$$

このアルゴリズムが実行できれば確かにUは頭語条件を満たす。しかし「pのprefix　全てに対しVが値を返すかどうか」を調べるということは、prefix のそれぞれに対し停止性検定を行なうことに他ならず、一般的には不可能である。 そこでChaitin はVの計算において行なわれる再帰呼び出しの深さに注目し、次のアルゴリズムを提案した[1]。

1）再帰呼び出しの深さの限度を指定する変数tを用意し、初期値を0にする。

2）$V(p,q)$の評価に必要な再帰呼び出しの深さdepth(p)がt 以下かどうかを調べる。

3）pのprefixかextension でt 以下の長さをもつものrのそれぞれに対しても$V(r,q)$ の評価に必要な再帰呼び出しの深さdepth(r)がt 以下かどうかを調べる。

4）上の２）と３）の結果に応じて次のように条件分岐を行なう

The Lisp based self-delimiting universal function
Kousuke Shizume
University of Library and Information Science

（イ） depth(p)≤tかつ全てのrに対しdepth(r)＞tの場合：$V(p,q)$の値を$U(p,q)$の値とする。
（ロ） depth(p)＞tかつdepth(r)≤t を満たすrが存在した場合：値を返すことを禁止する。
（ハ） depth(p)≤tかつdepth(r)≤tを満たすrが存在した場合：各rについて$|p|$と$|r|$を比べどのrに対しても　$|p|<|r|$ならば$V(p,q)$の値を$U(p,q)$の値とする。さもなくば値を返すことを禁止する。
（ニ） depth(p)＞tかつdepth(r)＞tの場合：tの値を1だけ増やして２）に戻る。
ここで「値を返すことを禁止する」ということは、例えば意図的に無限ループに入るなどして、停止しなくなることを意味する。

3．Chaitinの定義の欠陥とその修正

上のアルゴリズムで重要な量であるpの長さ$|p|$とdepth(p)に注目するとこのアルゴリズムは次のような意味を持つことがわかる。つまり「pのprefixまたはextension rで$|r|$<depth(p)かつdepth(r)<depth(p)を満たすものが存在するかを調べ、もし存在すれば$U(p,q)$が値をもつことを禁止する。」

ではこうすることにより本当に頭語性が保証されるだろうか。否である。実際、$|p|$＞depth(p)の場合に、pのprefix　rで$|r|$<depth(r)を満たすものがあったとする。　すると$V(r,q)$が値を持つかどうかは$U(p,q)$の評価に影響せず、しかも$V(p,q)$が値を持つかは$U(r,q)$ の評価に影響しないことが示せる。従ってこのアルゴリズムでは$U(p,q)$と$U(r,q)$の両方が値を返すことを妨げることは保証されない。従って頭語条件は保証されなくなる。

この欠陥を修正するには$U(p,q)$の評価の際に、pのprefixまたはextension rで
$|r|\leqq$ max{$|p|$, depth(p)}かつ depth(r)≦ max{$|p|$, depth(p)}
を満たすものがあるかどうかを調べ、あれば値を返すことを禁止し、なければ$V(p,q)$を値として返せばよい。そのようにするためにはChaitinのアルゴリズムにおいて４）の（イ）を次の（イ'）におきかえればすむ：
（イ'） depth(p)≤t かつ全てのrに対しdepth(r)＞tの場合：$|p|\leqq t$であったら$V(p,q)$を$U(p,q)$の値として返し、停止する。さもなくば tを１だけ増やし（２）へ戻る。

この修正したアルゴリズムにより、 pとそのprefixまたはextension rの両方に対して$V(p,q)$と$V(r,q)$の両方が値を返す場合でもmax{$|p|$,depth(p)}<max{$|r|$,depth(r)}　ならば$U(r,q)$ は値を返すことを禁止され、逆ならば$U(p,q)$の方が禁止される。さらに max{$|p|$,depth(p)}　=　max{$|r|$,depth(r)}　ならばpとrのうち長い方（つまり、他方のextensionになっているほう）が、値を返すことを禁止される。従ってUは確実に頭語条件を満たす。

4．まとめ

本研究ではChaitinによる頭語的評価関数の定義の欠陥とその修正法を提示した。

謝辞

日頃ご指導を賜わる図書館情報大の増永教授、および有益な助言をいただいた同大学の阪口助手に感謝いたします。

引用文献

1)Chaitin,G.J. On the length of programs for computing finite binary sequences. J.ACM. Vol.13, p.547-569(1966),
Chaitin,G.J. Algorithmic information theory.Cambridge: Cambridge Univ. Press(1987)