

4H-6

ソースプログラムのタイプミスの  
発見とその訂正

深谷哲司 永田守男  
(慶應義塾大学理工学部)

1. はじめに

コンパイラ言語を利用したプログラミングにおいては、些細な1つのエラーが、致命的になることがある。コンパイルが停止し、プログラミング作業における時間的ロスが増大するうえに、場合によっては、副作用を発生させることもある。このような些細なエラーのうちで、コンピュータによる積極的な発見と訂正が可能となる部分が当然ある。この報告では、タイプミスによるエラーの発見と訂正について取り上げる。

プログラム中においてタイプミスが発生する要素としては、予約語(言語ごとに決定され限定されるもの)、変数、定数、そして手続き名がある。変数等を宣言して使う言語に限定すれば、これらはすべて限られた集合の要素となるはずである[2]。したがって、プログラムの意味を深く理解することなしに、かなりの確率でタイプミスの発見と訂正を自動的に行うことが可能になる。

2. 設計

変数等を宣言する言語のプログラムにおいては、タイプミスした文字列に対応する正しい文字列が、その言語の文法で決められているか、あるいはプログラム中の他の場所にあるかのどちらかである。この特徴を利用して、エラー文字列の発生場所の分析により訂正文字列の種類を限定する。その種類に含まれるべき文字列にタイプミスのパターンを当てはめることと、どの文字列に訂正すべきかの重み付けを行い、ユーザとの対話により訂正を構文解析中に行うシステムを設計した。これによって、エラーがあっても構文解析を中断することなく続行できる。

3. 手法

プログラムの構文解析時に並行して行う。

(1) エラー文字列の種類判定

エラー文字列の前後解析 + 言語の文法

↓

エラー文字列の種類

エラー文字列の前後をsyntax chartとマッチングを取ることによりエラー文字列の取り得る種類(予約語、変数、定数、手続き名)を限定する。

(2) 訂正文字列の集合

限定された種類に関して、さらに構造的に分析する。そこから取り得る訂正文字列の集合を求める。予約語に関しては、言語によっては一意に決定される。変数、定数等に関しては、プログラムのその構造に関して使用可能な集合を決定する。

(3) 重み付け

上で得られた訂正可能な文字列が複数ある場合には、それぞれに重み付けを行い集合からの抽出をおこなう。

I) エラーモデルの適用 [1]

正しい文字列 → エラー文字列

①. 1文字欠落      a b c                  a c

- ②. 1文字余り      a b                  a \* b
- ③. 2文字入れ替え      a b                  b a
- ④. 1文字異なる      a b                  \* b
- ⑤. ①~④の複合

II) 訂正可能文字列の重み付け

1つのエラー文字列に関して、同一エラーモデルを適用できる訂正可能文字列が複数存在する場合、さらに重み付けの方法を取る。エラー文字列、訂正可能文字列の各文字の並び、およびキーボード上の各文字の位置関係から評価関数を作る。この関数の値から、一般的にどれが誤りやすい文字列なのかを評価する。評価関数にはキーボード上の距離を主な要因とした。

エラー文字列



← 変換(エラーモデル①~⑤)

取り得る文字列の集合



← キーボード上の位置関係

← 変換回数

重み付け

(4) 訂正

重み付けによって、最も誤った可能性が高いと考えられる文字列から順にユーザに提示し訂正を進める。ユーザによる入力も認める。

4. 結論

PL/0という、仮想言語を想定してシステムを作成した。結果としてタイプミスはほとんど訂正された。ただしシステムが最初に選んだ文字列が必ずしもユーザの意図するものではないこともある。

タイプミスとはかなり人それぞれに依存する部分が多い。ここでは標準的な利用者を仮定したが、各個人の誤りやすいパターンから積極的に検索する方法を研究する必要がある[3]。

参考文献

- (1) Durham, I., Lamb, D. A., & Saxe, J. B. : "Spelling correction in user interface", Communication of the ACM, 1983, 26, 764-773
- (2) Peterson, J.L. : "Computer programs for detecting and correcting spelling errors", Communication of the ACM, 1980, 23, 676-687
- (3) 永田守男、中島寛 : "個別利用者モデルに基づく対話的ソフトウェアのメッセージ生成", 電子情報通信学会論文誌 1987 Vol. J70-D No. 11 pp. 2077-2082