

手書き入力された数式の整形

3V-3

野崎信介, 佐藤俊, 中川正樹

(東京農工大学)

1. はじめに

一般の文章については、現在では多くの人ワープロプロセッサ（ワープロ）や文書整形ソフトウェアなど、計算機によるシステムを利用している。ところが、図や数式を含む文書ではこれらのシステムの使い勝手にはかなりの問題が残されている。本稿では特に数式を対象に、手書きで入力することができ、数式整形言語の知識を要せず美しい出力を得る整形方式とその実現について述べる。

2. 従来の数式記述システム

今までの数式を記述できるシステムには、マークアップ言語方式、WYSIWYG方式などがある。

前者では数式の位置や大きさをシステムの方で自動的に決めてくれる。そのため、ユーザは数式のレイアウトについて考えることなく数式を記述することができる。また、テキストで記述するためマシンに依存することがない。その反面、数式を記述するための複雑な命令をテキストに記述しなければならない。また、テキストをシステムで処理しないと出力結果を得ることができない。

後者は、WYSIWYG(What you see is what you get)の言葉通り、見たままを出力できる。このことは裏を返せば、ユーザが自分の考えた通りに数式のレイアウトをしなければならないということである。しかし、この場合においてもユーザは数式の構造を表現するために文字入力以外の多大な時間を要することになる。この手間は、前者のマークアップに対応すると考えられる。

どちらの方法を取っても、頭に浮かんだ数式を即座に表現することができない。これらの方式では数式を作成する際には、どうしても思考の中断がおきてしまうと考えられる。

3. 設計方針

本研究で試作した数式整形システムの設計方針は次の2つである。

(1) ユーザが手書きした数式を自動的に整形

ユーザ（人間）が手で書いた数式は、書いた本人にはその数式の位置関係や意味的なものを理解することができるが、正確に記述されていないことが多い。しかし、それを文書にするなど第三者に提示する時には正確に記述されていなければならない。

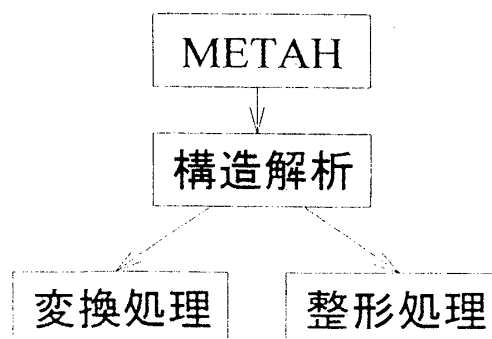
そこで手書きされた数式をシステムが自動的に整形することによって、ユーザの作成した数式を正確に表現することができる。また、整形して表示することによって、ユーザが書いた数式が正しく認識されているかどうかを確認することができる。

(2) 文書整形システム (TeX) への利用

手書きされた数式は、実際に文書中で利用することができなければならない。独自の出力方法を用いることよりもまず既存のシステムで利用することができれば文書作成の支援になる。

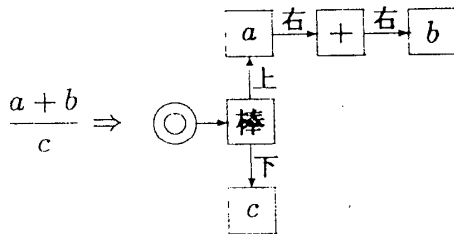
そこで、世界中で広く使われており、数式の記述に優れている文書整形システム TeX のテキスト形式へ変換することを可能にする。

システムの処理の大まかな流れを下図に示す。



4. オンライン数式認識システム METAH の利用

本研究では、我々の研究室で研究中であるオンライン手書き数式認識システム METAH によって出力された認識結果から整形出力を得るシステムについて述べる。METAH では、ユーザが表示一体型タブレットを用いて手書きの数式を認識させることができる。METAH の処理では手書きパターンに対し分割処理・記号認識処理・記号連結処理を行い、認識した最終結果は各記号とその位置関係（「右」、「右上」、「右下」、「左下」、「上」、「下」、「中」の7種）が示されたリストとして出力される。



5. 字句および構造解析

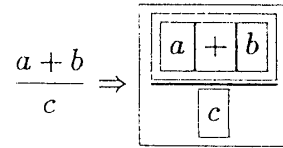
METAH から出力されたリストは各記号の絶対的な位置関係（「右」や「上」など上記の7種）でのみ構成されている。このリストを、整形や TeX への変換といった作業を行うことに対して扱いやすくするために、新たなリストを構築する。この新たなリストでは、sin や cos といった関数名をひとつの記号として扱うようにするための字句解析や、括弧の中身をまとめるといった数式の構造解析を行う。そして、整形や TeX への変換といった作業はこの新たなリストをもとに行う。

6. 整形表示

数式の整形は、字句および構造解析されたリストをもとに行う。整形の際の実際の数式記号列の配置は、「ボックス（箱）」という概念に基づいて行う。「ボックス」とは、それぞれの文字（あるいは文字列）を一つの「箱」とみなしたときの「箱」のことである。数式は複数の記号から成り立っている。言い換えれば複数の「箱」から成り立っている。それぞれの「箱」は最終的にひとつの大きな「箱」になり、「箱」を一つにまとめるときは、常にその箱の中の最長の縦あるいは横幅に合わせる。

また、各記号の他の記号に対する（例えば添字

や肩字の）大きさ・位置・フォントといった情報をファイルに保存することができ、ユーザが自由に変更可能にする。



7. 文書整形システム TeX への変換

TeX への変換方法は、METAH で認識された各記号に対する TeX の命令のテーブルを用意し、それを元に行う。テーブルはファイルにテキストで記述することができるようにして、システムの起動時にこのファイルを読み込んで、メモリ内にテーブルを形成する。そしてこのテーブルを元にして、変換を行う。従って、このテーブルファイルを書き換えることにより、TeX 以外の TeX に似た形式を持つマークアップ言語方式のシステムにも対応することが可能である。

$$\frac{a+b}{c} \Rightarrow \text{\frac{a+b}{c}}$$

8. おわりに

本稿では、手書き数式認識システムによって認識された数式を、整形表示し、文書整形システムのテキスト形式に変換するシステムについて述べた。

参考文献

- [1]村瀬敦史：オンライン手書き数式認識システム「METAH」の試作，情報処理学会 HI 研究会報告 48-4，1993.
- [2]Donald E. Knuth: The TeX book, Addison Wesley Publishing Company, 1984. (邦訳：斎藤信男監修，鷲谷好輝訳：TeX ブック，アスキー出版局，1989.)
- [3]Leslie Lamport: A Document Preparation System LaTeX, User's Guide & Reference Manual, Addison Wesley Publishing Company, 1986. (邦訳：Edgar Cooke, 倉沢良一監訳，大野俊治・小暮博道・藤浦はる美訳：文書処理システム LaTeX, アスキー出版局，1990.)