

2Q-2

図形処理との協調に基づく
ドリルテキストの解析

遠藤 勉 高岡 和宏
(大分大学工学部)

1. まえがき

一般のテキストは、自然言語で書かれた説明文と図形や表が混在しているのが普通である。通常は、言語表現と図形表現を分離した後、それぞれの解析を独立して行うことになるが、テキスト中の文章の指示に従って、図形を操作させるようなドリルテキストでは、両表現の相互関係の理解、両表現の協調による曖昧性の解消、両表現の相互変換などが要求される。

本稿は、文章表現と図形表現が相補的に機能する形で構成されている小学校1年の算数のドリルテキストを取り上げ、要素的な処理を行う複数のモジュールが相互に通信をしながら、文章解析と図形解析を分散協調に行う方法について述べたものである。

2. 分散協調処理によるテキストの分割と解釈

対象としたドリルテキストの例を図1に示す。テキストは問題番号、問題文、問題図形、解答欄、解答補足文字列などからなるが、以後これらをテキスト要素と呼ぶ。テキスト要素の種類、形状、配置などが、テキストによって大きく異なるだけでなく、問題文によって指示される内容も多様であり、それに伴って要求される図形認識処理も変化するため、テキストに応じて動的に解析方法を決定できることが望ましい。

そこで、非同期並行的に動作する複数個のモジュールを導入し、各モジュールが通信をしながらテキストの解析を行うシステムを考える。図2にシステムの基本構成を示す。解釈モジュール、分割モジュール、テキスト画像データが論理的な処理単位を構成しており、与えられたデータの分割と解釈を行う。その時点でデータの解釈が困難であれば、データを分割し、部分データ毎に新たに解釈および分割モジュールを生成し、同様の処理を行う。

2.1 分割モジュール

分割モジュールは解釈モジュールの指示に従って、画像データをテキスト要素に分割する。まず、画像データを適当なしきい値で2値化する。次に、横あるいは縦方向の周辺分布の空白部分を手がかりにして画像を分割する。周辺分布だけで分割できない画像については、連結成分を抽出し、成分ごとにラベリングした画像から要素図形を取り出す。分割された画像は、各要素の外接長方形になっており、長方形の左上角の点の原画像データにおけるX、Y座標と長方形のサイズをデータ属性として保存する。さらに、これらの属性

とデータ名を、呼び出した解釈モジュールに転送する。

2.2 解釈モジュール

解釈モジュールはシステム起動用モジュール、上位の解釈モジュールあるいは分割モジュールから得られたデータと各種解釈用モジュール(表1参照)を利用して、当該画像データの解釈を試みる。表1において、データタイプとはテキスト、文字列、文字、矩形、連結成分などを、解釈タイプとは文字コード、品詞系列、文節系列、構文木、意味表現、コマンド、具象図形などを示している。以下では解釈モジュールの処理手続きの流れを説明する。

- 1) 親モジュールにプロンプトを送信して、当該画像データのデータタイプを受信。
- 2) 対になる分割モジュール(SM)を起動。
- 3) 低次立案モジュール(LPM)を起動。
 - 3-1) LPMからのプロンプトを確認してデータタイプをLPMへ送信。
 - 3-2) LPMから次に起動すべきモジュール名とパラメータを受信。
 - 3-3) LPMからのプロンプトを確認して終了信号をLPMへ送信。
- 4) 受信したモジュール名が分割(SM)ならば以下を実行、そうでなければ5)へ。
 - 4-1) SMからのプロンプトを確認してパラメータをSMへ送信。
 - 4-2) 分割により得られた部分画像データ名とデータ属性をSMより受信。
 - 4-3) すべての部分画像データに対して判定モジュール(DM)を起動。
 - 4-3-1) DMからのプロンプトを確認してデータ属性をDMへ転送。
 - 4-3-2) DMからデータタイプを受信。
 - 4-3-3) DMからのプロンプトを確認してDMへ終了信号を送信。
 - 4-4) 各部分画像データ毎にデータ名をパラメータにして下位の解釈モジュールを起動。
 - 4-4-1) 下位のSMからのプロンプトを確認してデータタイプを送信し、6)へ。
- 5) 受信したモジュール名を持つモジュールを起動。
- 6) 下位の解釈モジュールあるいは解釈用モジュールからの結果を待つ。
 - 6-1) 解釈タイプと解釈結果を受信。
- 7) 高次立案モジュール(HPM)を起動。
 - 7-1) HPMからのプロンプトを確認して当該画像データのタイプと解釈タイプを送信。
 - 7-2) HPMから次に起動すべきモジュール名を受信。
 - 7-3) HPMからのプロンプトを確認して終了信

Drill Text Understanding in Cooperation with

Picture Processing

Tsutomu Endo and Kazuhiro Takaoka

Oita University

号をHPMへ送信。

- 8) 受信したモジュール名が解釈(IM)ならば、解釈タイプと解釈結果を親のモジュールに送信、そうでなければ、そのモジュールを起動して解釈タイプと解釈結果を送信し、6)へ。

図2では、入力テキストが問題文と、問題図形に分割され、さらに問題文が文字に、問題図形がその要素図形に分割されていく過程を示している。

3. 実験

UNIXワークステーションを用いてシステムのイ

表1 主要な解釈用モジュール

モジュール名	機能
判定(DM)	データ属性から、当該画像データのタイプを求める
低次立案(LPM)	画像データタイプから、起動すべきモジュール名を求める
高次立案(HPM)	解釈およびデータタイプから、起動すべきモジュール名を求める
文字照合(CMM)	画像データと文字構成要素とのテンプレートマッチング
文字認識(CRM)	マッチングの結果から、対応する文字コードを求める
辞書探索(DSM)	文字コード列と単語辞書との照合
形態素解析(PAM)	品詞および意味情報系列から文節系列を求める
構文解析(GAM)	品詞系列より構文解析木を求める
意味解析(SAM)	構文解析木より格構造を求める
意味解釈(SIM)	格構造よりコマンド系列を求める
矩形抽出(REM)	画像データから、長方形あるいは正方形の4頂点の座標を求める
図形正規化(PNM)	平行移動、回転、拡大縮小などによる変形を補正する
図形認識(PRM)	連結成分からなる各種要素図形を認識するモジュール群

ンプリメントを試みており、各モジュールはUNIXのプロセスとして、画像データはディスク上のファイルとして実現されている。また、子プロセスの生成・起動はUNIXのforkとexecシステムコールを、モジュール間の通信は、パイプおよびsocketシステムコールを用いて行う。プログラムは大部分C言語で記述されているが、システムの性格上、プログラミング言語の制約はほとんどない。入力テキストはイメージリーダにより取り込まれ、1024×1024画素(1画素256レベル)の画像データとなる。

解釈モジュールと分割モジュールによる入力テキストの分割実験では良好な結果を得ている。一方、言語解析部と要素図形の認識部は単独に動作することを確認しており、解釈モジュールから起動できるように変更を進めている。

4. むすび

複数の処理モジュールを使って算数のドリルテキストを分散協調的に解析する方式について述べた。本方式の特徴は、(1)上位のモジュールから下位のモジュールへモデルを与えるトップダウン方式、下位のモジュールから上位のモジュールへ解釈を与えるボトムアップ方式のいずれにも対応できる、(2)画素レベルの生データ処理から、理解、問題解決などの記号処理に至るまで統一した環境で実現できることである。

現在、各モジュールの知識表現とアルゴリズムの整備およびモジュール間の通信プロトコルの実験的検討を進めており、システム全体の動作の確認ならびに問題解決モジュールの導入と併せて今後の課題である。

参考文献

- (1) 石田亨: "分散人工知能の技術と応用", 人工知能学会誌, 5, 4 (1990).
- (2) 小1算数5分間トレーニング, 教学研究(1990).
- (3) 遠藤他: "初等数学の問題解決", 昭63電気関係学会九州支部連合大会論文集, 1002, 1003 (1988).

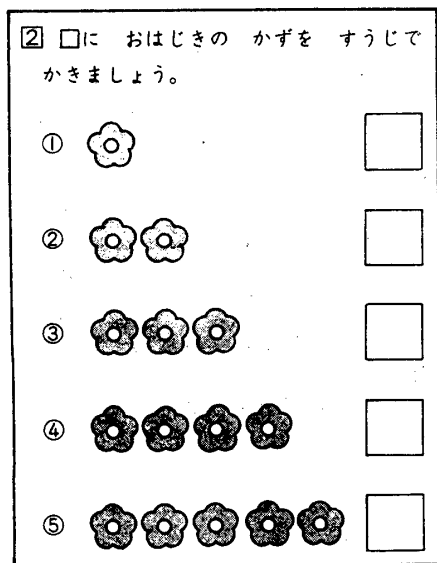


図1 ドリルテキストの例

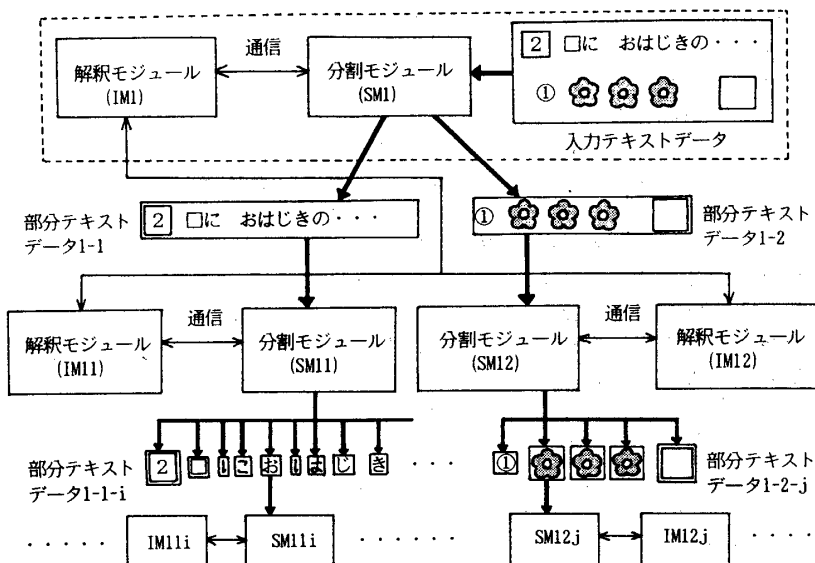


図2 ドリルテキスト解析システムの構成