

7 Q-6

オブジェクト指向知識ベースを用いた漢字フォント生成システム

于冬、広瀬通孝、石井威望

東京大学 工学部

1 始めに

漢字は英字に比べ形が複雑でありかつ字数も膨大であるため、漢字フォントの作成は多大な時間と労力を要する作業である。この問題を解決するために既にいくつかの研究が行われてきた[1][2]。しかしながら、現在までに提案されている方法には一般的に以下のような問題点が含まれている。

1. 普遍性がない。即ち特定の書体にしか適用できず、書体が異なるとフォント作成の方法を変えなければならない。
2. データの入力が大変である。漢字フォント生成に用いられるデータの精度に対する要求が厳しいため、莫大のデータを高精度で入力しなければならない。
3. 既存のデータを違う書体の作成に活用できない。漢字に関するデータが構造化された形で扱われていないため、既存のデータを最大限に活用することが難しい。

本論文では高品質のアウトラインフォントを生成する方法としてエレメント合成法を用いる。(1)の問題を解決するためにすべての書体に適用できるエレメント表現法を用いることを提案する。また(2)と(3)についてはオブジェクト指向による知識表現を採用し、知識に基づく推論機構を導入することで問題の解決を試みる。

2 エレメント合成による漢字フォント生成

2.1 骨格線上相対座標系によるエレメントデータ表現

従来の研究では漢字エレメント輪郭線を表すのに複数種の曲線を用いることが一般的である。この方法には輪郭線の作成と変形が困難であるという欠点があり、漢字フォントの自動生成に適した方法とは言えない。本論文では、輪郭線と骨格線をそれぞれ2次と1次B-Spline曲線を用いて表現している。この方法では輪郭線上の各点は骨格線上の制御点におけるローカル座標系で表現される(図1)。この方法は(1)書体の特徴に依存せず(2)骨格線によるエレメントの変形が容易であるため、フォントの自動生成に適した方法であると考えられる。

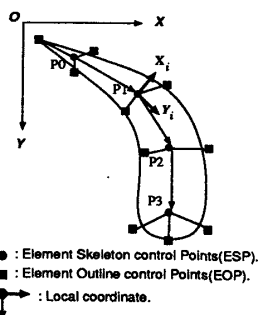


図1: 漢字エレメントデータの表現方法

2.2 エレメント合成による漢字生成

漢字の収まるべき四角い枠内に基本エレメントを適切な位置に並べ、接続すべきエレメント同士を接続させることにより漢字フォントを作成する。エレメントの接続はその部分の接続前後の形状の変化により(1)単純接続、(2)修飾接続、(3)変形接続の三種類に大別できる。

修飾接続の場合は、接続点において曲線による補間が必要である。また変形接続は、あらかじめ変形に必要な形状をエレメントの要素として作成し登録しておくことにより対処できる。

3 CLOSによる漢字データの構造化表現

オブジェクト指向プログラミング言語の発展につれて、オブジェクト指向による知識表現が注目されるようになってきている。知識表現の方法としては既に意味ネットワークとフレーム表現がよく知られているが、オブジェクト指向による知識表現はこれらの方法と同様な表現力と柔軟性を持つ一方、より構造化されたデータ表現が可能であるという長所も有する。本システムでは、CLOS(Common Lisp Object System)を用いて漢字データおよび知識の構造化表現を実現している。

これによって、漢字データの取り扱いが容易になり、さらに、図2のように既存データを他書体フォントの生成に利用することも可能になった。

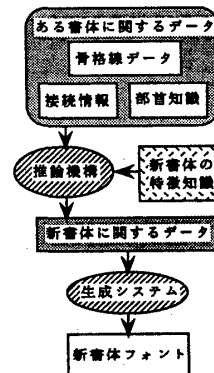


図2: 漢字データの再利用

4 知識を用いた推論

本システムでは、書体の特徴や部首などに関する知識をルール形式で記述している。現在漢字に関する知識として、(1)漢字エレメント接続の知識(2)書体特徴に関する知識(3)部首などの知識を用いている。推論には前向き推論を用いる。これらの知識を採り入れたことにより、柔軟性と拡張性

に優れたシステムを実現している。通常のシステムにおける問題点として、例えば、各書体の各エレメントに対して共通のアルゴリズムで変形を行なうことが困難であるという点が挙げられる。本システムにおいては、変形メソッドを定義し直すことでこの問題を解決することができる。

5 システムインプリメンテーション

本システムは Sun ワークステーション上で CLOS(Allegro CommonLisp) と X-Window を用いて実現している。図3と図4にそれぞれシステムの全体図とグラフィックス・エディタを示す。Lisp と X-Window 間の通信は CTK と呼ばれるプロセス間通信ツールを用いて行なっている。[3]。

6 おわりに

本システムを用い明朝体 100 文字について実験を行なった。図5にその結果をいくつか示す。今後明朝体の骨格線データを再利用し、他の書体の漢字を生成する実験を行なう予定である。システムのパフォーマンスをさらに向上させるために学習機能を探り入れることも考えている。また、生成した漢字のでき具合を評価するための指標についても考える必要がある。

参考文献

- [1] 上原、国西、下位、鍵政、菊池: “ストローク種別に基づく漢字形状生成方式,” 情報処理学会論文誌 Vol.31 No.2(1990)
- [2] 張、真田、手塚: “漢字楷書毛筆体の計算機による生成,” 電子通信学会論文誌 Vol. J67-D, No.5 (1984)
- [3] 葛岡英明、三井博隆、広瀬通孝、石井威望: “プラグブルなネットワーク・アプリケーション・ツールの開発,” 情報処理学会ソフトウェア工学研究会, July 1990

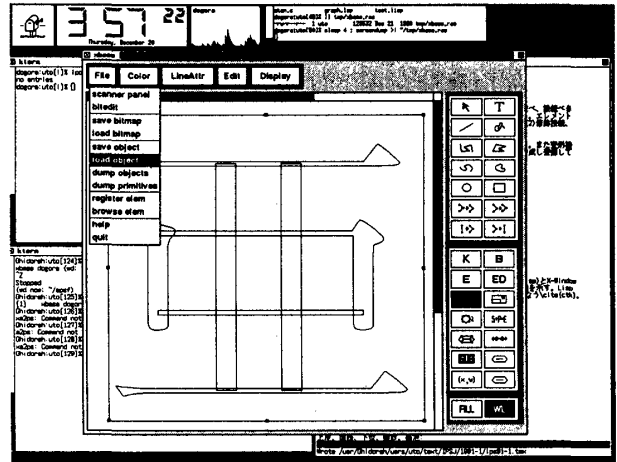


図 4: グラフィックス・エディタ

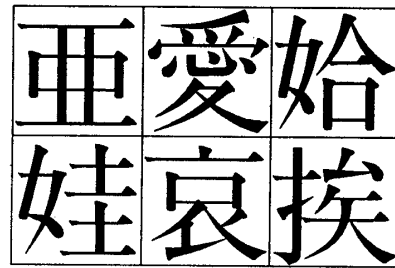


図 5: 漢字生成結果の例

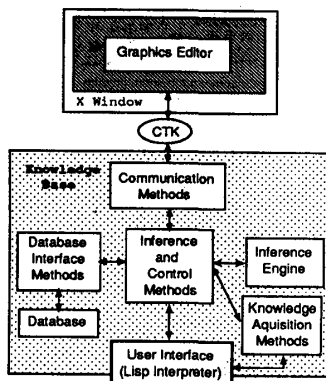


図 3: システム全体図