

漢字学習支援システム「漢字んカナメ」

猪子徹* 木村将希

(北海道大学情報科学研究科 複合情報学専攻)[†]

概要

既存の漢字学習支援システムは、単調な書き取り学習を行わせるものが多く、非常に退屈である。漢字学習支援システム「漢字んカナメ」は、この問題を解決するために、記憶対象の漢字に纏わる、ストーリー及び、アニメーション+視覚効果を生成することで、漢字学習を支援する。ストーリー及び、アニメーション生成にはGAを用いる。また、アニメーション生成、ストーリー生成実験を行い、結果を示した。

1 はじめに

現在、日本における外国人登録者数は201万1555人^{*1}であり、年々増加している。日本に滞在する外国人の大きな悩みは言葉の壁であり、日本語を学ぶ外国人にとって、漢字学習は困難である。また、携帯電話の普及、パソコンの普及に伴い、手書きの機会が減少するということもあり、日本人の漢字を書く能力の低下が叫ばれている。

そこで、近年、「MOJI BOY (漢字学習支援ソフト)」、「漢字王の王国」等の漢字学習支援システムが開発されている。

しかし、既存の漢字学習支援システムは、漢字を単調な繰り返し書き取り学習を通して身につけることが多く、非常に退屈である。また、麒麟、醤油、鬱、茛菪等、画数の多い漢字は覚えにくいという難点も持ち合わせている。

漢字学習システム「漢字んカナメ」は、既存の漢字学習支援システムの問題であった“飽きやすく”、“退屈で”、“非効率な”繰り返し書き取り漢字学習の問題を解消するために、記憶対象の漢字に纏わる、ストーリー及び、アニメーション+視覚効果を生成する。

2 本研究の位置付け

現在まで、アニメーション生成に関する研究として、これまで、Jeffrey 等が GA を用いたストロークアニメーションの生成手法を研究してきた[1],[2]。これらは有向グラフによって記述された人工生命を GA により進化させ、アニメーションを獲得するという物である。この手法は 3D アニメーションの分野へも応用されており、その成果も確認されている。

本論文では、これらのストロークアニメーション生成部を用いた漢字ストロークアニメーション生成システムを提案する。

次に、ストーリー生成に関する研究として、猪塚等による文生成システムがある[3]。これは、深層意味情報をシステムに与える事により、自動的に文を生成するというものである。また、Dramatica (Screen Play System 社) [4]や、佐久間、小方らによるストーリー生成支援システム[5]。これらのシステムでは、ストーリーの意味情報は予め定義されており、文毎の意味的な対応関係がある物としてストーリーを生成している。

本システムのストーリー生成機能が従来のストーリー生成システムと異なるところは、予め深層意味情報を与えずにストーリーを生成するところである。

3 漢字学習支援システム「漢字んカナメ」

3.1 システム構成

本システムは大きく“Kaname-Plugin”、“Kaname-Main”の2つから成る。又、“Kaname-Main”は漢字アニメーション機能、“ストーリー生成機能”、“ストーリー交換機能”の3つに細分化される。

それぞれの機能概要は以下の様になる。

- Kaname-Plugin :
 - Web ブラウザから漢字テキストの抽出
 - 漢字の部首分割
 - Web ブラウザへのアニメーションの描画
- 文字アニメーション機能 :
 - 漢字アニメーションの生成
- ストーリー生成機能 :
 - 部首に関連する語句の抽出
 - 漢字に纏わるストーリーの生成。
- 視覚効果生成システム :
 - 背景の付加
 - 背景色と漢字の文字色とのコントラストの強調
 - 漢字の文字サイズの変更

図1にシステム全体像を示す。

本稿では、上記の文字アニメーション機能、ストーリー生成機能に焦点を置いて説明する。

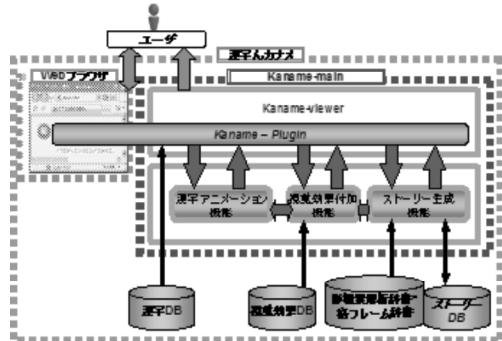


図1：漢字んカナメシステム全体図

4 漢字アニメーション生成システム

漢字アニメーション生成システムの目的は、“ユニークでダイナミックな動きを行う漢字アニメーション”の生成である。

漢字アニメーションは以下の3段階を経て生成される。以下、各詳細を説明する。

4.1 漢字の部首への分割

まず始めに、システムは Kaname-plugin より部首分割された漢字情報を受け取る。

*1. 法務省 (平成17年末現在における外国人登録者統計について)

分割のための漢字 DB は以下の様な構成を取る。

表 1 : 漢字 DB の構成 (一部)

id	text	b_id1	b_id2	b_id3	pattern	s1_start_x	s1_start_y	s1_end_x
0	男	11	15		A1			
1	朝	22	23	28	C3			
...								
11	田						10	30
...								
15	力					10	10	30

s1_end_y	...
10	
10	

各フィールドの説明を以下に示す。

- id : 漢字を一意に示す値
 - text : 漢字のテキスト
 - b*_id : 漢字を構成する部首を一意に示す値。b_id は id と対応する
 - pattern : 漢字がどの様なぶ形構造を持つかを示す値
 - s*_*_* : 部首の各辺の頂点の座標
- 例として、“男”という漢字テキストを受け取った場合、b_id の値として、11 (“田”), 15 (“力”) が漢字 DB より返される。

4.2 漢字アニメーションフィールドでの漢字の組成

次に、システムは b_id を元に、各部首の辺の頂点の座標 (s*_*_*) を取得、この座標を基に、漢字アニメーションフィールド (X-Y 平面) で漢字を構成する。



図 2 : 漢字アニメーションフィールドへの描画

また、辺で接続された各頂点は、それぞれ相互に参照可能な状態で構成する。

4.3 GA による漢字アニメーション生成

最後に、システムは GA により漢字アニメーションの自動生成を行う。Jeffrey は人工生命のアニメーションを獲得するため、各間接の角度を遺伝子として用いた [1]。本論文では、漢字アニメーション生成のため、頂点から辺へ接続された頂点への角速度を遺伝子として用いる。そのため、遺伝子長 (GeneLength) は、

$$\text{GeneLength} = s* \times \text{BitNumber}$$

s* : 漢字の辺の数
BitNumber : 8bit (角速度の Bit 数)

となる。

表 2 : 漢字アニメーションの表現系

s1_anglarSpeed(8bit)	...	s5_anglarSpeed(8bit)
00000000	...	00000000

4.3.1 評価値

本論文では評価値として、漢字の重心の移動量の総和を用いる。重心の移動を評価値として取ることで、漢字全体の動きを評価することが可能になる。

5 ストーリー生成システム

本システムでのストーリーは、記憶対象の漢字を含む部首が含まれる語句の組み合わせにより生成される。

ストーリー作成方針には以下の 2 つがあり、学習対象者によりそれぞれ異なる。

1. 学習対象漢字の各部首の位置関係を用いる
2. 学習対象漢字の各部首の意味を用いる

5.1 漢字の部首への分割

漢字アニメーション生成システムと同様に、システムは Kaname-plugin より部首分割された漢字情報を受け取る。

5.2 ストーリーを構成する句集号の取得

次に、システムは取得した漢字、及び、漢字の部首に対して、それぞれが含まれる名詞句集合、形容詞句集合、動詞句集合を形態素解析辞書 DB より取得する。

5.3 格フレーム辞書を用いた文の構成句の決定

得られた部首の句集号の内、元となる漢字に付随する動詞句の格フレームに一致するものを格フレーム辞書 [6] から探し出し、一致する物が有る場合、文の予備構成句として保持する。

最終的に、全ての部首の句が、元となる漢字に付随する動詞句の格フレームに一致するような、句及び、動詞句が見つかった場合、それらを文の構成句とし、ストーリーとする。

6 実験

本稿で示した手法の内、5 で示した手法の妥当性を確認するための実験を行った。本実験では学習対象漢字の各部首の意味を用いてストーリー生成を行った。

元となる漢字 : 朝

分割後の漢字 : 十, 早, 月

生成されたストーリー : 土人が月曜日の早期に朝日を見る
ストーリーは生成出来ているが、記憶対象の漢字より、難易度の高い漢字が出現しているのが問題である。

7 まとめ・課題

GA による漢字アニメーション自動生成システム、及び格フレーム辞書を用いたストーリー自動生成システムを示した。又、実験により漢字字形構造パターンを用いない場合のストーリー生成が出来たことを示した。

今後の課題は、漢字字形構造パターンを用いた場合のストーリー生成実験、視覚効果生成システムの開発である。

8 謝辞

本研究の一部は、平成 18 年度未踏ソフトウェア創造事業「未踏ユース」に採択され、研究開発致しました。

9 参考文献

- [1] Disney meets Darwin-the evolution of funny animated figures. Ventrella, J. Media Lab., MIT, Cambridge, MA, Computer Animation '95, Proceedings. Publication Date: 19-21 Apr 1995 On page(s): 35-43
- [2] Evolving Virtual Creatures K.Sims, Computer Graphics (Siggraph '94 Proceedings), July 1994, pp.15-22.
- [3] 猪塚 元, 岡部 了也, 柏野 和佳子, 細井 正樹, 村田 賢一, IPAL(SURFACE/DEEP)を用いた文生成システムについて, 1999
- [4] Tsutani, Y., YUZI's INDIE FILMMAKER, <http://www.voltage.co.jp/indiefm/dramatica/dramatical.htm>, 2003
- [5] Sakuma Tomoko, Takeshi Ogata, Story Generation Support System used the Story Theory of Propp, The 19th Annual Conference of the Japanese Society for Artificial Intelligence, 2005
- [6] 笹野遼平, 河原大輔, 黒橋禎夫, 名詞格フレーム辞書の自動構築とそれを用いた名詞句の関係解析, 自然言語処理, Vol.12, No.3, pp129-144 (2005.7)