

TR-IT-0109

相互情報量を用いた道案内システム

ローケン・キム キュンホ ゴリン アレン 北川 美宏\*

Kyung-ho Loken-Kim Allen Gorin Yoshihiro Kitagawa

1995. 4

概要

本報告書では、わかち書きされた文(単語並び)とその文に関連するひとつのアクション(1単語)から、各単語とアクションとの相互情報量を計算、その相互情報量を用い単語とアクション間をむすぶネットワーク(Connectionist Network, 以下CN)を構築し、そのCNより、道案内に関する文を入力すると、応答するシステムに関して報告する。

## 目次

1	単語とアクションの相互情報量を用いた道案内システム	1
2	使用方法	2
2・1	CNデモシステムの所在	2
2・2	CN稼働環境	2
2・3	CNの実行方法	2
2・4	CNの終了方法	3
3	デモ・システム	4
3・1	CNのデモ概要	4
3・2	相互情報量計算プログラムについて	5
	参考文献	6

## 1 単語とアクションの相互情報量を用いた道案内システム

本デモシステムは、わかち書きされた文(単語並び)とその文に関連するひとつのアクション(1単語)から、各単語とアクションとの相互情報量を計算、その相互情報量を用い単語とアクション間をむすぶネットワーク(Connectionist Network, 以下CN)を構築し、そのCNより、道案内に関する文を入力すると、応答するシステムである。本システムは相互情報量計算プログラムと、道案内システムをイメージしたユーザインタフェースプログラムから構成されている。

学習用のデータとして、文ファイル(例文:「京都駅はどこですか」)とそれに対応するアクションファイル(例:「京都」(駅))を与えることで、各単語と各アクションとの相互情報量が計算される。ユーザインタフェースプログラムは、京都および高の原周辺の駅、線路図を表示するメインウィンドウと、注目している駅周辺を拡大表示する虫メガネウィンドウを持つ。このプログラムの主な機能は、虫メガネウィンドウがある駅に向けてメインウィンドウ上を移動、駅周辺の拡大表示および当該駅の表示色の強調により、現在システムが注目している駅を表現することである。

ユーザーは、メインウィンドウ上である駅をポインティング(マウスプレスイベント)することで、システムに対し注目駅を変更させることができる。システムは、ユーザーからのイベントが一定時間なかった場合、注目度合が減少し他の駅に向けて移動する(boredom function)。

また、当ユーザインタフェースプログラム用に疑似的なマウスのボタンプレスイベントを発行するタッチパネル制御プログラムがある。

デモシステムでは、簡単な質問文処理プログラムにより、上記相互情報量計算プログラムとユーザインタフェースプログラムを連結している。

相互情報量計算の後、起動される質問文処理プログラムは、入力された質問文から求まる相互情報量総和が最大のアクション(駅)をユーザインタフェースプログラムへメッセージ送信します。ユーザインタフェースプログラムはメッセージを受け取ると、注目駅を変更し、その駅へ向けて虫メガネウィンドウを移動させる。

## 2 使用方法

以下、本デモシステムの使用方法などを説明する。

### 2・1 CNデモシステムの所在

本デモシステムは、現在 /data/as13/kitagawa/CN にある。

/data/as13/kitagawa/CN/のディレクトリ構成は、次のとおりである。

README : (このファイル)

SRC/ : ユーザインタフェースプログラム関連のプログラムなど

SRC/NETWORK/

: Allen GorinさんのCN関連プログラム

+ 相互情報量計算&質問文処理プログラム

TouchPanel/

: CN用タッチパネル制御プログラム

algor/ : Allen GorinさんのCN関連プログラム(オリジナル)

デモシステムの実行モジュールは、次の3ファイルである。

1) /data/as13/kitagawa/CN/SRC/cn

: ユーザインタフェースプログラム

2) /data/as13/kitagawa/CN/SRC/NETWORK/cn\_mi

: 相互情報量計算&質問文処理プログラム

3) /data/as13/kitagawa/CN/TouchPanel/TouchPanel

: CN用タッチパネル制御プログラム

### 2・2 CN稼働環境

SUN SPARC station

マイクロタッチシステムズ社"ハイパータッチ"

: タッチパネル制御プログラムを使用する場合。

### 2・3 CNの実行方法

kterm等ターミナルウィンドウから各プログラムを起動する。

1) ユーザーインタフェースプログラム(cn)起動

: メインウィンドウと虫メガネウィンドウがオープンする。

% cn

## 2) 相互情報量計算&質問文処理プログラム(cn\_mi)起動

: 学習データ("input\_0.dat"と"input\_1.dat")を読み込んだ後、  
"--- 質問をどうぞ ---"というメッセージを出し、質問文の  
入力を待つ。

```
% cn_mi
```

## 3) CN用タッチパネル制御プログラム(TouchPanel)起動

: タッチパネル(マイクロタッチシステムズ社"ハイパータッチ")  
が使用できる環境で起動する。

起動後、タッチパネルへのタッチは、ボタンプレスイベント  
としてユーザーインタフェースプログラムのメインウィンドウ  
へ送信される。

```
% TouchPanel
```

### <プログラムの起動順について>

まず最初に、ユーザインタフェースプログラム(cn)を起動する必要がある。

これは、質問文処理プログラムおよびタッチパネル制御プログラムからユーザインタフェースプログラムへのイベント送信用に、ユーザインタフェースプログラムのメインウィンドウのウィンドウIDをXルートウィンドウへプロパティ登録している為である。

## 2・4 CNの終了方法

次の方法で各プロセスを終了する。

1) cnを起動したターミナルウィンドウで、Ctrl-cによる割り込み。

2) 質問文を入力せずに、Returnキーのみ入力。

または、Ctrl-cによる割り込み。

3) TouchPanelを起動したターミナルウィンドウで、Ctrl-cによる割り込み。

### 3 デモ・システム

#### 3・1 CNのデモ概要

##### [attention function]

ユーザーが、ある場所をメインウィンドウ上でポインティングすると、虫メガネウィンドウがその場所へ向けて移動する。

ポインティング場所近くに駅があった場合、虫メガネウィンドウは、現在の注目駅としてその駅を赤色表示する。

##### [boredom function]

虫メガネウィンドウが注目駅へ移動後一定時間経過すると、その駅に対する注目度が減少し、虫メガネウィンドウは、注目駅を黄色表示に変更する。

さらに一定時間が経過すると、システムは他の駅へ注目を移す。現在、新たな注目駅は、駅定義データリスト("/data/as13/kitagawa/CN/SRC/cn\_object.h")における直前の注目駅の次としている。

現在、本来のboredom functionとなる関数は存在せず、いき値となる時間を定数で与えている(/data/as13/kitagawa/CN/SRC/threshold.h)。

```
#define TIME_THRESHOLD 10L /* 単位：秒 */
```

##### [mutual information(相互情報量)]

相互情報量計算&質問文処理プログラム(cn\_mi)において、質問文を入力すると、質問文中の各単語とアクション(駅)との相互情報量を求め、アクション毎に相互情報量を合算する。デモでは、その中から最も相互情報量和の大きいアクションを次のように表示した後、ユーザーインタフェースプログラムのメインウィンドウへ通知する。

--- 質問をどうぞ ---

京都駅はどこにありますか  
maxMI = 3.349234, station = 京都

現在、デモシステムに形態素解析等、文を単語に分割する機能はない。したがって、質問文は単語毎にわかり書きする必要がある。

Connectionist networkの特徴として、質問文の構造に依存しない点がある。すなわち、質問文が非文であってもなんらかのアクション(駅)を選択する。

ただし、質問文中のすべての単語が、学習データに現れなかった単語である場合は、次のようなメッセージを表示し、アクション(駅)は選択しない。

--- 質問をどうぞ ---

はひふへほ

>>> 学習不足の為、適当な駅を見つけられませんでした!!<<<

### 3・2 相互情報量計算プログラムについて

相互情報量計算プログラムについての詳細は、A.Gorinのプログラムドキュメント (/data/as13/kitagawa/CN/SRC/NETWORK/readme)および参考文献 [1,2] を参照させていただきだい。

ここでは、相互情報量計算プログラムの学習データと相互情報量計算式について説明する。

相互情報量計算プログラムは、起動されると次の2つのデータファイルを読み込む。

/data/as13/kitagawa/CN/SRC/NETWORK/input\_0.dat

/data/as13/kitagawa/CN/SRC/NETWORK/input\_1.dat

"input\_0.dat"ファイルには、アクション(駅)を1行毎に書いてある。この各行のアクション(駅)に対応して、"input\_1.dat"ファイルには、そのアクション(駅)に関係のある文を書いてある。

現在、学習データとして37の駅に対し、のべ148アクション(駅)と、148文を使用している。

この学習データをもとに、相互情報量は、次の計算式により求められている。

[相互情報量計算式]

アクション(駅)Aと単語Wの相互情報量

$$= \log \left( \frac{A \text{ と } W \text{ の 相互出現率}}{A \text{ の 出現率} \times W \text{ の 出現率}} \right)$$

$$A \text{ と } W \text{ の 相互出現率} = \frac{A \text{ と } W \text{ の 相互出現数}}{\text{学習データ数}}$$

AとWの相互出現数

= 学習データ中のアクションと文の対それぞれについて、

AとWが同時に出現した回数を積算したもの

学習データ数 = 学習データ中のアクションと文の対の数

Aの出現率 = Aの出現数 / 学習データ数

Wの出現率 = Wの出現数 / 学習データ数

□ 参考文献

[1] ANANTH SANKAR and ALLEN GORIN, "Adaptive language acquisition in a multi-sensory device" in Artificial Neural Networks for Speech and Vision, by R.J. Mammone (ed.), Chapman & Hall, London, ISBN 0 412 54850 X, 1993, pp. 325-356

[2] Allen Gorin, "On Automated Language Acquisition," Presented at JASA meeting in 1994



## 付録

- 1) file : input\_0.dat: page 5 参照  
アクション (駅名) リスト
- 2) file : input\_1.dat: page 5 参照  
アクションと関係がある文リスト
- 3) file: net.log (一部)  
アクションと単語の Mutual Information

"input\_0.dat"

---

京都  
京都  
京都  
京都  
京都  
高の原  
高の原  
高の原  
高の原  
高の原  
西大寺  
西大寺  
奈良  
近鉄奈良  
近鉄奈良  
近鉄奈良  
近鉄奈良  
蹴上  
蹴上  
蹴上  
丹波橋  
京都  
高の原  
西大寺  
奈良  
嵯峨  
嵐山  
桂  
北野白梅町  
帷子の辻  
西院  
二条  
丹波口  
四条大宮  
烏丸  
北山  
鞍馬  
宝ヶ池  
八瀬遊園  
出町柳  
京阪三条  
京津三条  
山科  
蹴上  
京津山科

河原町  
四条  
京阪四条  
竹田  
東福寺  
丹波橋  
中書島  
京阪宇治  
宇治  
祝園  
新祝園  
木津  
近鉄奈良  
京都  
高の原  
西大寺  
奈良  
嵯峨  
嵐山  
桂  
北野白梅町  
帷子の辻  
西院  
二条  
丹波口  
四条大宮  
烏丸  
北山  
鞍馬  
宝ヶ池  
八瀬遊園  
出町柳  
京阪三条  
京津三条  
山科  
蹴上  
京津山科  
河原町  
四条  
京阪四条  
竹田  
東福寺  
丹波橋  
中書島  
京阪宇治  
宇治  
祝園

新祝園  
木津  
近鉄奈良  
奈良  
木津  
宇治  
東福寺  
京都  
近鉄奈良  
西大寺  
西大寺  
高の原  
新祝園  
丹波橋  
竹田  
京都  
木津  
祝園  
中書島  
丹波橋  
東福寺  
京阪四條  
京阪三條  
出町柳  
中書島  
京阪宇治  
桂  
西院  
烏丸  
河原町  
桂  
嵐山  
四條大宮  
西院  
帷子の辻  
嵐山  
帷子の辻  
北野白梅町  
京都  
丹波口  
二條  
嵯峨  
北山  
四條  
京都  
竹田  
出町柳

宝ヶ池  
八瀬遊園  
宝ヶ池  
鞍馬  
京津三条  
蹴上  
京津山科  
京都  
山科

"input\_1.dat"

京都駅はどこにありますか  
新幹線ひかりは京都にとまりますか  
水梨さんはどこに住んでいますか  
私は今京都駅に着いたところです  
京都の景色が好きです  
ATRの最寄り駅はどこですか  
高の原サントウンにマクドナルドがあります  
近鉄高の原からATRまで専用直通バスが走っています  
今京都駅にいますのですがけいはんなへはどう行けばいいですか  
高の原からATRまでタクシーで1500円くらいです  
昨日西大寺の奈良ファミリーへ行きました  
難波から高の原まで行くなら西大寺で乗り換えて下さい  
JR奈良線は奈良が始発駅です  
大仏は近鉄奈良からが近いです  
東大寺へ行きたいのですが  
ひがしむき商店街で買い物をしました  
奈良公園には鹿がいっぱいいます  
シンポジウムは国際交流センターで開催されます  
国際会議に参加したいのですが  
蹴上までは路面電車が走っています  
丹波橋で近鉄京都線を京阪線に乗り換えます  
京都駅はどこですか  
高の原駅はどこですか  
西大寺駅はどこですか  
奈良駅はどこですか  
嵯峨駅はどこですか  
嵐山駅はどこですか  
桂駅はどこですか  
北野白梅町駅はどこですか  
帷子の辻駅はどこですか  
西院駅はどこですか  
二条駅はどこですか  
丹波口駅はどこですか  
四条大宮駅はどこですか

烏丸 駅はどこですか  
北山 駅はどこですか  
鞍馬 駅はどこですか  
宝ヶ池 駅はどこですか  
八瀬遊園 駅はどこですか  
出町柳 駅はどこですか  
京阪三条 駅はどこですか  
京津三条 駅はどこですか  
山科 駅はどこですか  
蹴上 駅はどこですか  
京津山科 駅はどこですか  
河原町 駅はどこですか  
四条 駅はどこですか  
京阪四条 駅はどこですか  
竹田 駅はどこですか  
東福寺 駅はどこですか  
丹波橋 駅はどこですか  
中書島 駅はどこですか  
京阪宇治 駅はどこですか  
宇治 駅はどこですか  
祝園 駅はどこですか  
新祝園 駅はどこですか  
木津 駅はどこですか  
近鉄奈良 駅はどこですか  
京都へ行きたいのですが  
高の原へ行きたいのですが  
西大寺へ行きたいのですが  
奈良へ行きたいのですが  
嵯峨へ行きたいのですが  
嵐山へ行きたいのですが  
桂へ行きたいのですが  
北野白梅町へ行きたいのですが  
帷子の辻へ行きたいのですが  
西院へ行きたいのですが  
二条へ行きたいのですが  
丹波口へ行きたいのですが  
四条大宮へ行きたいのですが  
烏丸へ行きたいのですが  
北山へ行きたいのですが  
鞍馬へ行きたいのですが  
宝ヶ池へ行きたいのですが  
八瀬遊園へ行きたいのですが  
出町柳へ行きたいのですが  
京阪三条へ行きたいのですが  
京津三条へ行きたいのですが  
山科へ行きたいのですが  
蹴上へ行きたいのですが

京津山科へ行きたいのですが  
河原町へ行きたいのですが  
四条へ行きたいのですが  
京阪四条へ行きたいのですが  
竹田へ行きたいのですが  
東福寺へ行きたいのですが  
丹波橋へ行きたいのですが  
中書島へ行きたいのですが  
京阪宇治へ行きたいのですが  
宇治へ行きたいのですが  
祝園へ行きたいのですが  
新祝園へ行きたいのですが  
木津へ行きたいのですが  
近鉄奈良へ行きたいのですが  
JR奈良線に奈良駅はありますか  
JR奈良線に木津駅はありますか  
JR奈良線に宇治駅はありますか  
JR奈良線に東福寺駅はありますか  
JR奈良線に京都駅はありますか  
近鉄奈良線に近鉄奈良駅はありますか  
近鉄奈良線に西大寺駅はありますか  
近鉄京都線に西大寺駅はありますか  
近鉄京都線に高の原駅はありますか  
近鉄京都線に新祝園駅はありますか  
近鉄京都線に丹波橋駅はありますか  
近鉄京都線に竹田駅はありますか  
近鉄京都線に京都駅はありますか  
JR学研都市線に木津駅はありますか  
JR学研都市線に祝園駅はありますか  
京阪本線に中書島駅はありますか  
京阪本線に丹波橋駅はありますか  
京阪本線に東福寺駅はありますか  
京阪本線に京阪四条駅はありますか  
京阪本線に京阪三条駅はありますか  
京阪本線に出町柳駅はありますか  
京阪宇治線に中書島駅はありますか  
京阪宇治線に京阪宇治駅はありますか  
阪急京都線に桂駅はありますか  
阪急京都線に西院駅はありますか  
阪急京都線に烏丸駅はありますか  
阪急京都線に河原町駅はありますか  
阪急嵐山線に桂駅はありますか  
阪急嵐山線に嵐山駅はありますか  
京福本線に四条大宮駅はありますか  
京福本線に西院駅はありますか  
京福本線に帷子の辻駅はありますか  
京福本線に嵐山駅はありますか

京福北野線に帷子の辻駅はありますか  
京福北野線に北野白梅町駅はありますか  
JR山陰線に京都駅はありますか  
JR山陰線に丹波口駅はありますか  
JR山陰線に二条駅はありますか  
JR山陰線に嵯峨駅はありますか  
京都市営地下鉄に北山駅はありますか  
京都市営地下鉄に四条駅はありますか  
京都市営地下鉄に京都駅はありますか  
京都市営地下鉄に竹田駅はありますか  
叡山本線に出町柳駅はありますか  
叡山本線に宝ヶ池駅はありますか  
叡山本線に八瀬遊園駅はありますか  
叡山鞍馬線に宝ヶ池駅はありますか  
叡山鞍馬線に鞍馬駅はありますか  
京阪京津線に京津三条駅はありますか  
京阪京津線に蹴上駅はありますか  
京阪京津線に京津山科駅はありますか  
JR東海道線に京都駅はありますか  
JR東海道線に山科駅はありますか



"net.log"

-----  
1994年11月24日(木) 14時49分44秒 JST  
-----

Word-Word Network

Number of ww connections are 620

ENTRY	FREQ	WORD 0	WORD 1	Prob	MI
0	11	京都	京都	0.074	2.425
1	8	京都	駅	0.054	0.038
2	10	京都	は	0.068	0.210
3	3	京都	どこ	0.020	-0.078
4	9	京都	に	0.061	0.599
5	6	京都	あります	0.041	0.297
6	9	京都	か	0.061	0.156
7	1	京都	新幹線	0.007	2.512
8	1	京都	ひかり	0.007	2.512
9	1	京都	とまりま	0.007	2.512
10	1	京都	水梨さん	0.007	2.512
11	1	京都	住んでま	0.007	2.512
12	1	京都	私	0.007	2.512
13	1	京都	今	0.007	1.819
14	1	京都	着いた	0.007	2.512
15	1	京都	ところ	0.007	2.512
16	4	京都	です	0.027	-0.520
17	2	京都	の	0.014	-0.556
18	1	京都	景色	0.007	2.512
19	2	京都	が	0.014	-0.645
20	1	京都	好き	0.007	2.512
21	3	高の原	A T R	0.020	2.918
22	3	高の原	の	0.020	0.255
23	1	高の原	最寄り	0.007	2.918
24	4	高の原	駅	0.027	-0.250
25	4	高の原	は	0.027	-0.301
26	2	高の原	どこ	0.014	-0.078
27	4	高の原	か	0.027	-0.250