

メディア情報学実験・音声認識 第三週課題レポート

1510151 柳 裕太

2017 年 12 月 1 日

1 認識対象の単語

表 1: 実験で使った数字単語

単語表記	単語名	発音	音素数
3	saN	/saN/	3
4	yoN	/yoN/	3
5	go	/go/	2
7	nana	/nana/	4

2 HMM の学習曲線

[go] に対して行った HMM の学習曲線は、以下の図 1 の通りである。

3 単語認識性能評価

3.1 自分の声

ソースコード 1: countCorr 実行結果

```
1 $ countCorr
2 ( 中 略 )
3 rei , 0
4 ichi , 0
5 ni , 0
6 saN , 10
7 yoN , 10
8 go , 10
9 roku , 0
10 nana , 10
11 hachi , 0
```

12	kyuu, 0
----	---------

3.2 他人の声

ソースコード 2: countXCorr 実行結果

```
1 $ countXCorr m0
2 Evaluation start.  Wait a moment.
3 rei, 0
4 ichi, 0
5 ni, 0
6 saN, 33
7 yoN, 1
8 go, 9
9 roku, 0
10 nana, 44
11 hachi, 0
12 kyuu, 0
13
14 $ countXCorr m1
15 Evaluation start.  Wait a moment.
16 rei, 0
17 ichi, 0
18 ni, 0
19 saN, 22
20 yoN, 1
21 go, 32
22 roku, 0
23 nana, 30
24 hachi, 0
25 kyuu, 0
26
27 $ countXCorr m2
28 Evaluation start.  Wait a moment.
29 rei, 0
30 ichi, 0
31 ni, 0
32 saN, 0
33 yoN, 0
34 go, 0
35 roku, 0
```

```

36 nana, 45
37 hachi, 0
38 kyu, 0
39
40 $ countXCorr f1
41 Evaluation start. Wait a moment.
42 rei, 0
43 ichi, 0
44 ni, 0
45 saN, 4
46 yoN, 0
47 go, 15
48 roku, 0
49 nana, 47
50 hachi, 0
51 kyu, 0

```

4 オンライン単語認識

4.1 自分の声

3, 4, 5, 7 をそれぞれ3回ずつ発声し、結果を得た。

ソースコード 3: recog 実行結果

```

1 $ #saN (spoken by myself)
2 $ recog -20 lib/HMMList
3 Returnキーを押してください
4 Isolated Word Recognition On-The-Fly
5 vad_file= vadwav/10221.wav
6
7 rank word (kana)= log-likelihood
8
9 1. 3 (saN)= -1790.410712
10 2. 7 (nana)= -1838.183389
11 3. 4 (yoN)= -1981.794753
12 4. 5 (go)= -2125.044269
13
14 ハングアップ
15 Returnキーを押してください
16 Isolated Word Recognition On-The-Fly
17 vad_file= vadwav/10454.wav

```

```

18 |-----
19 rank word (kana)= log-likelihood
20 |-----
21 1. 7 (nana)= -1015.582787
22 2. 3 (saN)= -1056.889629
23 3. 4 (yoN)= -1113.365923
24 4. 5 (go)= -1180.399190
25
26 ハングアップ
27 Returnキーを押してください
28 Isolated Word Recognition On-The-Fly
29 vad_file= vadwav/10491.wav
30 |-----
31 rank word (kana)= log-likelihood
32 |-----
33 1. 3 (saN)= -2313.627728
34 2. 7 (nana)= -2391.826858
35 3. 4 (yoN)= -2526.148038
36 4. 5 (go)= -2698.040703
37
38 ハングアップ
39
40 $ #yoN (spoken by myself)
41 $ recog -20 lib/HMMList
42 Returnキーを押してください
43 Isolated Word Recognition On-The-Fly
44 vad_file= vadwav/10277.wav
45 |-----
46 rank word (kana)= log-likelihood
47 |-----
48 1. 4 (yoN)= -1787.122659
49 2. 7 (nana)= -1816.205699
50 3. 3 (saN)= -1825.196272
51 4. 5 (go)= -1979.401992
52
53 ハングアップ
54 Returnキーを押してください
55 Isolated Word Recognition On-The-Fly
56 vad_file= vadwav/11351.wav
57 |-----
58 rank word (kana)= log-likelihood

```

```

59 |-----
60 | 1. 7 (nana)= -1732.293919
61 | 2. 4 (yoN)= -1749.738537
62 | 3. 3 (saN)= -1758.778353
63 | 4. 5 (go)= -1842.883630
64 |
65 | ハングアップ
66 | Returnキーを押してください
67 | Isolated Word Recognition On-The-Fly
68 | vad_file= vadwav/11384.wav
69 |-----
70 | rank word (kana)= log-likelihood
71 |-----
72 | 1. 4 (yoN)= -1623.474475
73 | 2. 3 (saN)= -1675.216760
74 | 3. 7 (nana)= -1679.459238
75 | 4. 5 (go)= -1708.053267
76 |
77 | ハングアップ
78 | $ #go (spoken by myself)
79 | $ recog -20 lib/HMMList
80 | Returnキーを押してください
81 | Isolated Word Recognition On-The-Fly
82 | vad_file= vadwav/10115.wav
83 |-----
84 | rank word (kana)= log-likelihood
85 |-----
86 | 1. 5 (go)= -1445.805241
87 | 2. 7 (nana)= -1490.668070
88 | 3. 3 (saN)= -1556.722549
89 | 4. 4 (yoN)= -1663.274652
90 |
91 | ハングアップ
92 | Returnキーを押してください
93 | Isolated Word Recognition On-The-Fly
94 | vad_file= vadwav/12073.wav
95 |-----
96 | rank word (kana)= log-likelihood
97 |-----
98 | 1. 5 (go)= -1349.897716
99 | 2. 7 (nana)= -1505.549940

```

```

100 3. 3 (saN)= -1550.040932
101 4. 4 (yoN)= -1623.344431
102
103 ハングアップ
104 Returnキーを押してください
105 Isolated Word Recognition On-The-Fly
106 vad_file= vadwav/12121.wav
107 _____
108 rank word (kana)= log-likelihood
109 _____
110 1. 5 (go)= -1386.920504
111 2. 7 (nana)= -1426.955245
112 3. 3 (saN)= -1526.620053
113 4. 4 (yoN)= -1610.494881
114
115 ハングアップ
116
117 $ #nana (spoken by myself)
118 $ recog -20 lib/HMMList
119 Returnキーを押してください
120 Isolated Word Recognition On-The-Fly
121 vad_file= vadwav/10306.wav
122 _____
123 rank word (kana)= log-likelihood
124 _____
125 1. 7 (nana)= -1384.627282
126 2. 3 (saN)= -1536.170493
127 3. 4 (yoN)= -1621.792248
128 4. 5 (go)= -1777.863372
129
130 ハングアップ
131 Returnキーを押してください
132 Isolated Word Recognition On-The-Fly
133 vad_file= vadwav/12208.wav
134 _____
135 rank word (kana)= log-likelihood
136 _____
137 1. 7 (nana)= -1565.372470
138 2. 3 (saN)= -1864.963177
139 3. 4 (yoN)= -1942.985780
140 4. 5 (go)= -2001.167651

```

```

141
142 ハングアップ
143 Returnキーを押してください
144 Isolated Word Recognition On-The-Fly
145 vad_file= vadwav/12236.wav
146 _____
147 rank word (kana)= log-likelihood
148 _____
149 1. 7 (nana)= -1732.933940
150 2. 3 (saN)= -1986.133111
151 3. 4 (yoN)= -2071.052025
152 4. 5 (go)= -2194.124920
153
154 ハングアップ

```

4.2 他人の声

今回の実験では、友人の張氏と浅津氏の2人に協力して頂き、3, 4, 5, 7の順に発声してもらった。

4.2.1 張氏の場合

ソースコード 4: recog 実行結果 (張氏)

```

1 Returnキーを押してください
2 Isolated Word Recognition On-The-Fly
3 vad_file= vadwav/39826.wav
4 _____
5 rank word (kana)= log-likelihood
6 _____
7 1. 3 (saN)= -1505.069939
8 2. 7 (nana)= -1586.825035
9 3. 5 (go)= -1748.309120
10 4. 4 (yoN)= -1753.374475
11
12 ハングアップ
13 Returnキーを押してください
14 Isolated Word Recognition On-The-Fly
15 vad_file= vadwav/39885.wav
16 _____
17 rank word (kana)= log-likelihood
18 _____
19 1. 5 (go)= -1344.840856

```

```

20 2. 4 (yoN)= -1361.261114
21 3. 7 (nana)= -1455.075794
22 4. 3 (saN)= -1500.914987
23
24 ハングアップ
25 Returnキーを押してください
26 Isolated Word Recognition On-The-Fly
27 vad_file= vadwav/39936.wav
28 _____
29 rank word (kana)= log-likelihood
30 _____
31 1. 7 (nana)= -1161.868159
32 2. 4 (yoN)= -1187.774808
33 3. 3 (saN)= -1208.448915
34 4. 5 (go)= -1232.254232
35
36 ハングアップ
37 Returnキーを押してください
38 Isolated Word Recognition On-The-Fly
39 vad_file= vadwav/39971.wav
40 _____
41 rank word (kana)= log-likelihood
42 _____
43 1. 7 (nana)= -877.720638
44 2. 4 (yoN)= -938.691815
45 3. 3 (saN)= -943.252832
46 4. 5 (go)= -970.624385
47
48 ハングアップ

```

4.2.2 浅津氏の場合

ソースコード 5: recog 実行結果 (浅津氏)

```

1 #3, 4, 5, 7の順に発声
2 Returnキーを押してください
3 Isolated Word Recognition On-The-Fly
4 vad_file= vadwav/44822.wav
5 _____
6 rank word (kana)= log-likelihood
7 _____

```



```

8   1.  7  (nana)= -1955.394600
9   2.  3  (saN)= -1960.798787
10  3.  4  (yoN)= -2085.992983
11  4.  5  (go)= -2102.686864
12
13  ハングアップ
14  Returnキーを押してください
15  Isolated Word Recognition On-The-Fly
16  vad_file= vadwav/44962.wav
17  _____
18  rank word (kana)= log-likelihood
19  _____
20  1.  7  (nana)= -951.017733
21  2.  3  (saN)= -979.229846
22  3.  4  (yoN)= -1051.388855
23  4.  5  (go)= -1086.196301
24
25  ハングアップ
26  Returnキーを押してください
27  Isolated Word Recognition On-The-Fly
28  vad_file= vadwav/45027.wav
29  _____
30  rank word (kana)= log-likelihood
31  _____
32  1.  7  (nana)= -951.803934
33  2.  3  (saN)= -983.027531
34  3.  4  (yoN)= -994.104035
35  4.  5  (go)= -1012.153328
36
37  ハングアップ
38  Returnキーを押してください
39  Isolated Word Recognition On-The-Fly
40  vad_file= vadwav/45062.wav
41  _____
42  rank word (kana)= log-likelihood
43  _____
44  1.  7  (nana)= -878.557397
45  2.  3  (saN)= -1032.259370
46  3.  5  (go)= -1033.311989
47  4.  4  (yoN)= -1044.670763
48

```

4.3 考察

今回のオンライン音声認識では、自分の発声ではほぼ安定した結果を得ることができた。しかしながら、友人に発声してもらったところ、判定がかなり不安定な結果となった。

張氏では、中国人留学生で発音が独特だった影響か、特に5の発声では5の最尤順位が最下位となった。また浅津氏の場合では、すべての発声において7が最尤と判定され、更にこちらでも5の発声時に5が最尤順位が最下位となった。これより5においては、自分の発話性が独特であった影響によって判定が不安定になっている可能性が考えられた。

また浅津氏の場合では、3, 4, 5の発話において、最尤順位がすべて7, 3, 4, 5の順序になっていた。これは、浅津氏の発話性による影響を受けたものと考えている。

日本人である浅津氏と中国人留学生である張氏の2者に手伝って頂いたが、最尤順位が正しい数字となる回数では張氏の方が多かった。これは、音声認識の観点では中国語訛りの日本語であろうと、ある程度発話性によっては対応できる可能性を示唆しているのではないかと考えている。同様に日本人の発話においても、発話性によっては判定が不安定になることも同時に示唆していると解釈した。

5 実験のポイント

音声認識において、音声入力から判定までの流れを学ぶことと、過程にて活用する数理モデルの実践

6 よくわかったこと

基本的な処理としては、パターンマッチングによってもっとも近い単語を調べているということ

7 よくわからなかったこと

単語のみならず、文脈を活用した単語の推測・判定

8 要望

システムの信頼性をもう少し改善してほしかった。不具合によって録音不可能になって実験作業に手間取るケースがしばしばあった。

9 感想・その他

特になし

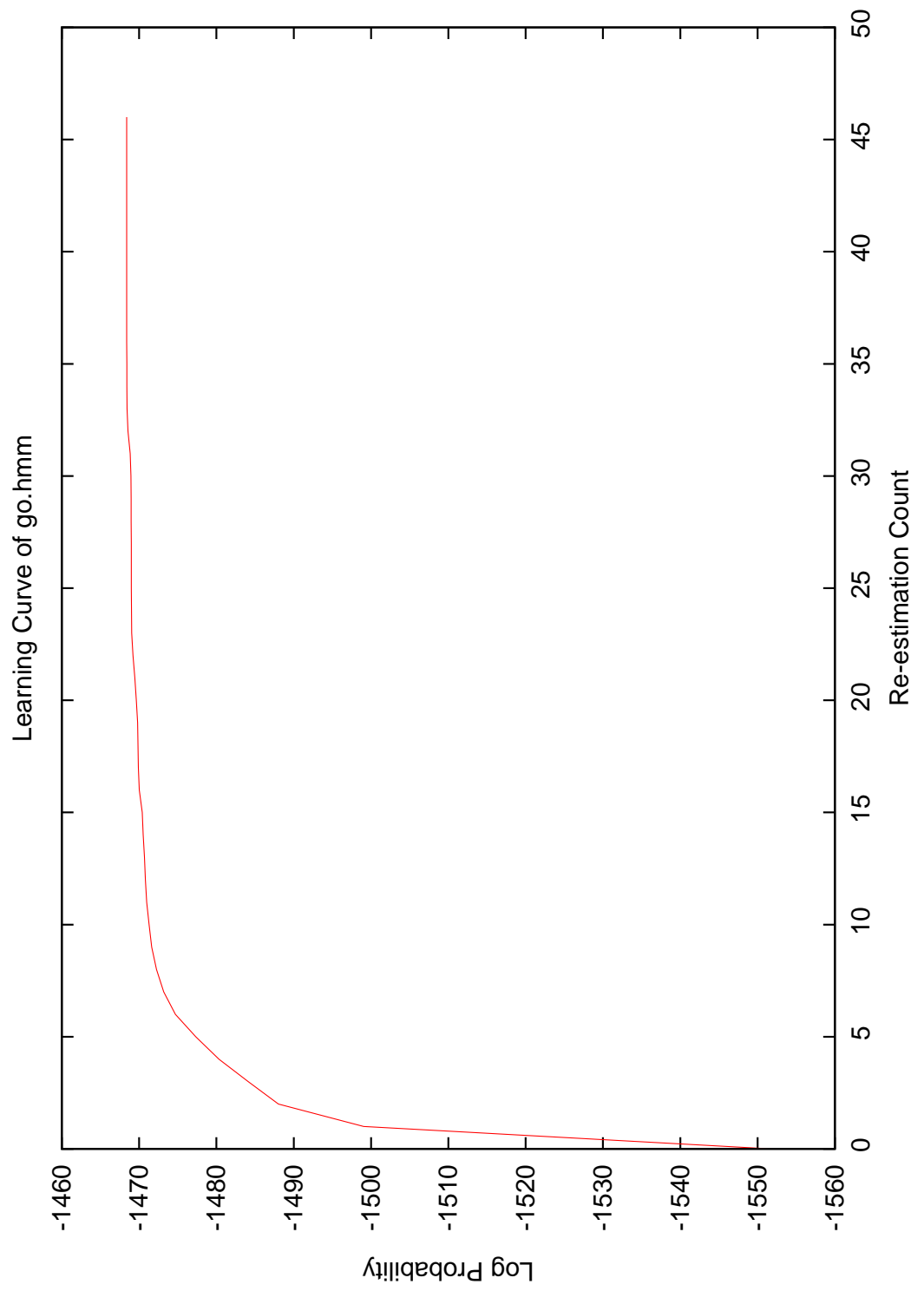


図 1: 実行結果グラフ (go)