

**ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH**  
**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**  
**KHOA CÔNG NGHỆ THÔNG TIN**

**TRẦN TẤN BẢO – 17C 11002**  
**TRẦN THÚY HIỀN – 17C 11026**  
**NGUYỄN HÀ DUY PHƯƠNG – 17C 11 032**

**ĐỒ ÁN MÔN HỌC**  
**XỬ LÝ NGÔN NGỮ NÓI**

**BÀI TẬP 2: NHẬN DẠNG RỜI RẠC**  
**GIÁO VIÊN**

**PGS. TS. Vũ Hải Quân**

**TP.HCM - 6/2018**

# MỤC LỤC

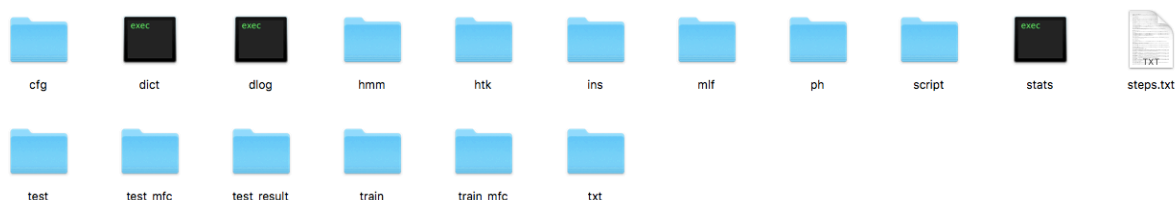
<b>1</b>	<b>Mô tả bài toán .....</b>	<b>2</b>
<b>2</b>	<b>Phương pháp làm .....</b>	<b>2</b>
2.1	Các bước huấn luyện (mở cmd trở đến thư mục htk, lưu ý các file trong project này đều phải lưu dưới dạng ANSI): .....	2
2.2	Các bước test:.....	9
<b>3</b>	<b>Kịch bản thử nghiệm.....</b>	<b>10</b>
<b>4</b>	<b>Kết quả, vẽ biểu đồ.....</b>	<b>10</b>
4.1	Kết quả thực nghiệm .....	10
4.2	Biểu đồ .....	11
<b>5</b>	<b>Đánh giá cá nhân.....</b>	<b>11</b>
<b>6</b>	<b>Tài liệu tham khảo .....</b>	<b>12</b>

## 1 Mô tả bài toán

- Mục tiêu của bài tập là nhận dạng các từ rời rạc, chẳng hạn nhận dạng chữ số từ không đến chín.
- Kiểm tra sự thay đổi chất lượng nhận dạng thông qua thay đổi kịch bản thực nghiệm.

## 2 Phương pháp làm


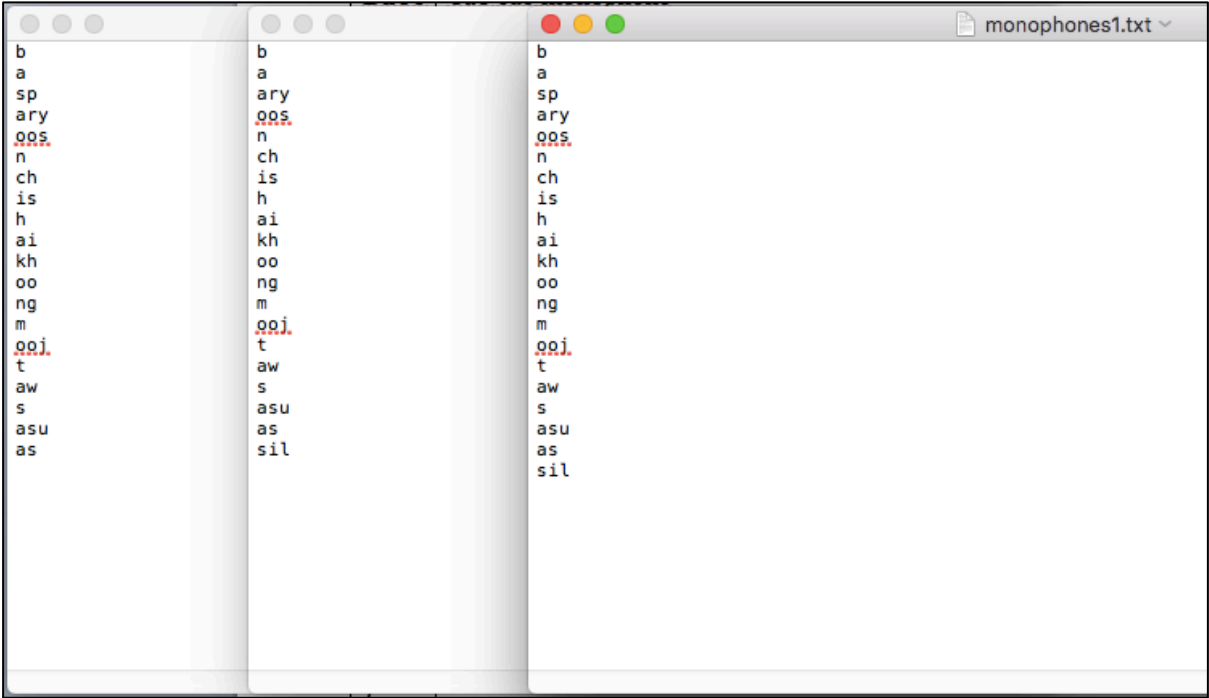
- Sử dụng HTK Toolkit của Đại học Cambridge và mô hình HMM.
- Xây dựng n mô hình tương ứng với n âm vị.
- Tạo thư mục bài tập **htk** (Báo cáo giả định là thư mục chạy ngay trên ổ đĩa C)
- Trong thư mục **htk** có các mục sau: **bin32** chứa các chương trình con của HTK, **cfg** chứa các file config, **scripts** chứa các file Perl (cần cài đặt môi trường cho Perl), **wav** chứa hai thư mục con **train** và **test** để chứa các file \*.wav dùng để train và test tương ứng:
- Gồm có train: 4 file âm thanh đọc liên tục từ 0 đến 9 và test: 60 file âm thanh test rời rạc các chữ số.
- Cài đặt thêm môi trường **perl** và **python** cho một số bước phát sinh tập tên file...



### 2.1 Các bước huấn luyện (mở cmd trở đến thư mục htk, lưu ý các file trong project này đều phải lưu dưới dạng ANSI):

<b>Bước</b> <b>1</b>	<p>Tạo file <b>gram.txt</b> chứa các từ cần nhận dạng (ở bài tập này là từ 0 đến 9)</p> <pre>\$word = ba   bary   boosn   chisen   hai   khoong   mooit   nawm   sasv   tasm; (&lt;\$word&gt;)</pre> <p>(Lưu ý có dấu xuống dòng ở cuối file, nếu không sẽ bị lỗi)</p>
-------------------------	--

<b>Bước</b> 2	Tạo file <b>wdnet</b> bằng lệnh: <b>Htk/HParse.exe txt/gram.txt txt/wdnets</b>
<b>Bước</b> 3	Tạo file dict và sắp xếp theo thứ tự alphabet: <div data-bbox="268 371 919 759" data-label="Text"> <pre> ba          b a sp bary        b ary sp boosn       b oos n sp chisn       ch is n sp hai         h ai sp khoong      kh oo ng sp moojt       m ooj t sp nawm        n aw m sp sasus       s asu sp tasm        t as m sp </pre> </div> <p>(Lưu ý có dấu xuống dòng ở cuối file, nếu không sẽ bị lỗi)</p>
<b>Bước</b> 4	Tạo file thu âm với số lượng train là 4 file, test là 60 file. Ta đã tạo được file <b>train</b> chứa tên và nội dung của các file cần được ghi âm và đưa vào thư mục <b>train</b> . Làm tương tự cho <b>test.s</b> <div data-bbox="268 1086 1482 1341" data-label="Image"> </div>
<b>Bước</b> 5	Chạy lệnh sau (lưu ý khi chạy các file Python cần có từ “Python” đứng đầu câu lệnh trong cmd): <b>Python script/listwavmfc.py train/ train_mfc/ train.scp train_mfc.scp</b> Lệnh này nhằm tạo ra file <b>script/listwavmfc.py</b> , file này nhằm tổng hợp các file cần rút trích đặc trưng mfcc trong <b>train</b> và tạo danh sách các file mfcc xuất ra tương ứng trong thư mục <b>train_mfc</b> .
<b>Bước</b> 6	Rút trích đặc trưng mfcc: <b>htk/HCopy.exe -T 1 -C cfg\HCopy.cfg -S train_mfc.scp</b> Ta sẽ có các file mfcc tương ứng với từng file train trong <b>train</b> . Tạo file <b>train_mfc.scp</b> , lưu danh sách các file mfcc trong <b>train</b> vào thư mục <b>train_mfc</b> . Kết quả:

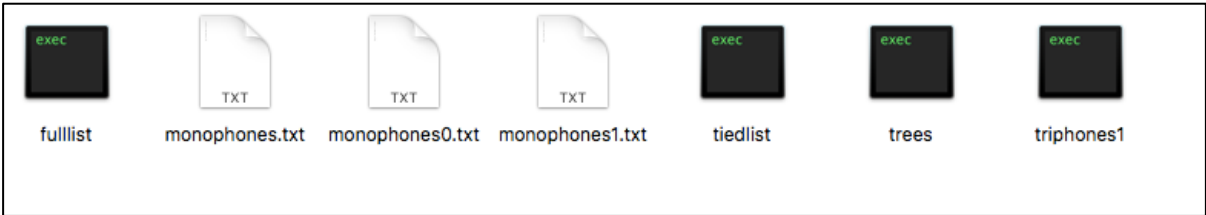
	 <p>TRAIN01_00.mfc   TRAIN02_00.mfc   TRAIN03_00.mfc   TRAIN04_00.mfc</p>
<b>Bước</b> 7	<p>Tạo các monophone</p> <p>Htk/HDMan.exe -m -w txt/wlist.txt -n ph/monophones.txt -l dlog txt/dict.txt txt/srcDict.txt</p> <p>Perl script/mkMonophones.pl ph/monophones.txt ph/monophones0.txt ph/monophones1.txt</p> <p>Sử dụng <i>monophones.txt</i> đã tạo ra để tạo thêm các <i>monophones0.txt</i> và <i>monophones1.txt</i></p> <p>Lần lượt là các monophones: <i>monophones.txt</i>, <i>monophones0.txt</i>, <i>monophones1.txt</i> khác nhau ở các <b>sil</b> và <b>sp</b>.</p> 
<b>Bước</b> 8	<p>Tạo file <b>wlist</b> (word list):</p> <p>Scripts/prompts2wlist.pl train txt/wlist</p>
<b>Bước</b> 9	<p>Tạo MLF <i>words</i></p> <p>perl script/prompts2mlf.pl mlf/words.mlf txt/prompts.txt</p>

<b>Bước</b>  <b>10</b>	Tạo file <b>mkphone0.led</b> : <pre> 1  EX 2  IS sil sil 3  DE sp 4 </pre> (Lưu ý có dấu xuống dòng ở cuối file)
<b>Bước</b>  <b>11</b>	Tạo các MLF <i>phones</i> từ MLF <i>words</i> . <pre> htk/HLEd.exe -l "*" -d txt/srcDict.txt -i mlf/phones0.mlf ins/mkphones0.led mlf/words.mlf  htk/HLEd.exe -l "*" -d txt/srcDict.txt -i mlf/phones1.mlf ins/mkphones1.led mlf/words.mlf </pre>
<b>Bước</b>  <b>12</b>	Tạo thư mục <b>hmm/hmm0</b> , <b>hmm/hmm1</b> , ..., <b>hmm/hmm22</b> Tạo file proto mẫu theo htkbook, copy vào <b>hmm/hmm0</b> , hoặc tạo file proto với nội dụng sau và copy vào <b>hmm/hmm0</b> : <pre> 1  ~o &lt;VecSize&gt; 39 &lt;MFCC_0_D_A&gt; 2  ~h "proto" 3  &lt;BeginHMM&gt; 4  &lt;NumStates&gt; 5 5  &lt;State&gt; 2 6      &lt;Mean&gt; 39 7      0 8      0 0 0 0 0 9      &lt;Variance&gt; 39 10     11     1 1 1 1 1 12 &lt;State&gt; 3 13 &lt;Mean&gt; 39 14 0 15 0 0 0 0 0 16 &lt;Variance&gt; 39 17 18 1 1 1 1 1 19 &lt;State&gt; 4 20 &lt;Mean&gt; 39 21 0 22 0 0 0 0 0 23 &lt;Variance&gt; 39 24 1 25 1 1 1 1 1 26 &lt;TransP&gt; 5 27 0.0    1.0    0.0    0.0    0.0 28 0.0    0.6    0.4    0.0    0.0 29 0.0    0.0    0.6    0.4    0.0 30 0.0    0.0    0.0    0.7    0.3 31 0.0    0.0    0.0    0.0    0.0 32 &lt;EndHMM&gt; </pre>

<b>Bước</b> <b>13</b>	Khởi tạo giá trị train: <code>htk/HCompV.exe -C htk/cfg/HCompV.cfg -f 0.01 -m -S txt/train.scp -M hmm/hmm0 hmm/hmm0/proto</code>
<b>Bước</b> <b>14</b>	Tạo file <b>hmmdefs</b> : <code>Perl script/mkHmmdefsFile.pl hmm/hmm0/proto ph/monophones0.txt hmm/hmm0/hmmdefs</code>
<b>Bước</b> <b>15</b>	Tạo file <b>macros</b> : <code>perl script/mkMacrosFile.pl hmm/hmm0/vFloors hmm/hmm0/macros</code>
<b>Bước</b> <b>16</b>	Xây dựng mô hình <b>hmm1</b> và <b>hmm2</b> <code>htk/HERest.exe -C cfg/HERest.cfg -I mlf/phones0.mlf -t 250.0 150.0 1000.0 -S txt/train_mfc.scp -H hmm/hmm0/macros -H hmm/hmm0/hmmdefs -M hmm/hmm1 ph/monophones0.txt</code>  <code>htk/HERest.exe -C cfg/HERest.cfg -I mlf/phones0.mlf -t 250.0 150.0 1000.0 -S txt/train_mfc.scp -H hmm/hmm1/macros -H hmm/hmm1/hmmdefs -M hmm/hmm2 ph/monophones0.txt</code>
<b>Bước</b> <b>17</b>	Thực hiện tương tự cho huấn luyện <b>hmm4</b> , <b>hmm5</b> .
<b>Bước</b> <b>18</b>	Thêm mô hình <b>sp</b> vào <b>hmmdefs4</b> <code>perl script/makesp.pl hmm/hmm3/hmmdefs hmm/hmm4/hmmdefs hmm/hmm3/macros hmm/hmm4/macros</code>
<b>Bước</b> <b>19</b>	Sử dụng <b>HHed</b> để trói buộc mô hình <b>sil</b> và <b>sp</b> , huấn luyện <b>hmm5</b> <code>htk/HHed.exe -H hmm/hmm4/macros -H hmm/hmm4/hmmdefs -M hmm/hmm5 ins/sil.led ph/monophones1.txt</code>  File sil.led có cấu trúc như sau:


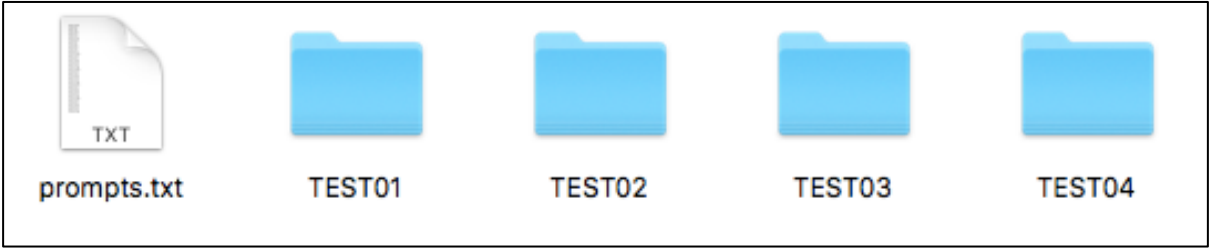
	<pre> 1 AT 2 4 0.2 {sil.transP} 2 AT 4 2 0.2 {sil.transP} 3 AT 1 3 0.3 {sp.transP} 4 TI silst {sil.state[3],sp.state[2]} </pre> <p>Huấn luyện tiếp <b>hmm6</b>, <b>hmm7</b> tương tự bước 14, tuy nhiên ở đây lúc này sử dụng <i>monophones1.txt</i> chứ không dùng <i>monophones0.txt</i> như ở trên.</p>
<b>Bước 20</b>	<p>Tạo ra file <b>aligned.mlf</b> từ <b>phones.mlf</b></p> <pre>htk/HVite.exe -l * -o SWT -b silence -a -H hmm/hmm7/macros -H hmm/hmm7/hmmdefs -i mlf/aligned.mlf -m -t 250.0 -y lab -I mlf/words.mlf -S txt/train_mfc.scp txt/srcDict.txt ph/monophones1.txt</pre>
<b>Bước 21</b>	<p>Huấn luyện tiếp tục <b>hmm8</b>, <b>hmm9</b> từ <b>aligned.mlf</b> có được bằng lệnh</p> <pre>htk/HERest.exe -B -C cfg/HERest.cfg -I mlf/aligned.mlf -t 250.0 150.0 1000.0 -s stats -S txt/train_mfc.scp -H hmm/hmm7/macros -H hmm/hmm7/hmmdefs -M hmm/hmm8 ph/monophones1.txt</pre> <pre>htk/HERest.exe -B -C cfg/HERest.cfg -I mlf/aligned.mlf -t 250.0 150.0 1000.0 -s stats -S txt/train_mfc.scp -H hmm/hmm8/macros -H hmm/hmm8/hmmdefs -M hmm/hmm9 ph/monophones1.txt</pre>
<b>Bước 22</b>	<p>Tạo <b>triphones</b> từ monophones bằng lệnh</p> <pre>htk/HLEd.exe -n ph/triphones1 -l * -i mlf/wintri.mlf ins/mktri.led mlf/aligned.mlf</pre>
<b>Bước 23</b>	<p>Tạo <b>mktri.hed</b> để huấn luyện <b>hmm10</b> là nhái monophones từ <b>hmm9</b></p> <pre>htk/HHEd.exe -B -H hmm/hmm9/macros -H hmm/hmm9/hmmdefs -M hmm/hmm10 ins/mktri.hed ph/monophones1.txt</pre> <p><b>Hmm11</b>: để ước lượng triphones.</p> <pre>htk/HERest.exe -B -C cfg/HERest.cfg -I mlf/wintri.mlf -t 250.0 150.0 1000.0 -s stats -S txt/train_mfc.scp -H hmm/hmm10/macros -H hmm/hmm10/hmmdefs -M hmm/hmm11 ph/triphones1</pre>



	<p>Tiếp tục huấn luyện <b>hmm12</b>.</p> <pre>htk/HERest.exe -B -C cfg/HERest.cfg -I mlf/wintri.mlf -t 250.0 150.0 1000.0 -s stats -S txt/train_mfc.scp -H hmm/hmm11/macros -H hmm/hmm11/hmmdefs -M hmm/hmm12 ph/triphones1</pre>
<b>Bước</b> <b>24</b>	<p>Tạo <b>fulllist</b> từ monophones0</p> <pre>perl script/mkFullList.pl ph/monophones0.txt ph/fulllist</pre>
<b>Bước</b> <b>25</b>	<p>Tạo <b>tree.hed</b></p> <pre>perl script/mkTree.pl 40 ph/monophones0.txt ins/tree.hed</pre> <p>Vào <b>tree.hed</b>, sửa lại đường dẫn muốn lưu <i>fulllist</i>, <i>tiedlist</i> và <i>tree</i></p> <pre>TR 1 AU "D:/Thacsi/NNN/htk1/ph/fulllist" CO "D:/Thacsi/NNN/htk1/ph/tiedlist" ST "D:/Thacsi/NNN/htk1/ph/trees"</pre>
<b>Bước</b> <b>26</b>	<p>Tạo trạng thái ràng buộc cho <b>triphones</b>. Đây là bước quan trọng và tốn thời gian nếu tập dữ liệu quá lớn.</p> <pre>htk/HHed.exe -B -H hmm/hmm12/macros -H hmm/hmm12/hmmdefs -M hmm/hmm13 ins/tree.hed ph/triphones1 &gt; txt/log.txt</pre> <p>Sau bước này, <b>tiedlist</b> và <b>tree</b> sẽ được tạo ra theo đường dẫn lưu đã chọn ở</p> 
<b>Bước</b> <b>27</b>	<p>Cuối cùng, huấn luyện <b>hmm14</b>, <b>hmm15</b> với <b>tiedlist</b> đã được tạo ra ở bước trên.</p> <pre>htk/HERest.exe -B -C cfg/HERest.cfg -I mlf/wintri.mlf -t 250.0 150.0 1000.0 -s stats -S txt/train_mfc.scp -H hmm/hmm13/macros -H hmm/hmm13/hmmdefs -M hmm/hmm14 ph/tiedlist</pre> <pre>htk/HERest.exe -B -C cfg/HERest.cfg -I mlf/wintri.mlf -t 250.0 150.0 1000.0 -s stats -S txt/train_mfc.scp -H hmm/hmm14/macros -H hmm/hmm14/hmmdefs -M</pre>

	<hmm hmm="" hmm15="" ph="" tiedlist<="" tiedlist<br=""></hmm> => Đến đây là xong các bước để train dữ liệu.
--	--

## 2.2 Các bước test:

<b>Bước 1</b>	<ul style="list-style-type: none"> <li>Tạo thư mục <b>test_result</b> để lưu kết quả test</li> </ul> 
<b>Bước 2</b>	<ul style="list-style-type: none"> <li>Thu âm các file test và bỏ vào thư mục <b>test</b>, và tạo tập tin <b>prompts.txt</b> tương ứng các từ trong file âm thanh.</li> </ul>  <ul style="list-style-type: none"> <li>Tạo file <b>test.scp</b> và <b>test_mfc.scp</b>:  python script/listwavmfc.py test/ test_mfc/ txt/test.scp txt/test_mfc.scp </li> <li>Rút trích đặc trưng mfcc của các file test:  htk/HCopy.exe -T 1 -C cfg/HCopy.cfg -S txt/test.scp </li> </ul>
<b>Bước 3</b>	<ul style="list-style-type: none"> <li>Tạo file <b>words.mlf</b> lưu trong thư mục <b>test_result</b>:  perl script/prompts2mlf.pl test_result/words.mlf test_result/prompts.txt </li> </ul>
<b>Bước 4</b>	Thực hiện test dữ liệu: htk/HVite.exe -C cfg/Hvite.cfg -H hmm/hmm15/macros -H hmm/hmm15/hmmdefs -S txt/test_mfc.scp -i test_result/recout.mlf -w txt/wdnet.txt txt/srcDict.txt ph/tiedlist
<b>Bước 5</b>	Tổng kê kết quả, xuất kết quả ra tập tin txt. htk/HResults.exe -f -t -I test_result/words.mlf ph/tiedlist test_result/recout.mlf > test_result/result.txt

Đây là kết quả của lần test dùng hmm15.

```
----- Sentence Scores -----  
===== HTK Results Analysis =====  
Date: Sat Jun 16 20:20:21 2018  
Ref : D:/Thacsi/NNN/htk1/test_result/words.mlf  
Rec : D:/Thacsi/NNN/htk1/test_result/recout15.mlf  
----- File Results -----  
TEST01_00.rec: 0.00( 0.00) [H= 0, D= 0, S= 1, I= 0, N= 1]  
Aligned transcription: D:/Thacsi/NNN/htk1/test_mfc/TEST01_00.lab vs D:/Thacsi/NNN/htk1/  
test_mfc/TEST01_00.rec  
LAB: khoang  
REC: bary  
TEST01_01.rec: 0.00( 0.00) [H= 0, D= 0, S= 1, I= 0, N= 1]  
Aligned transcription: D:/Thacsi/NNN/htk1/test_mfc/TEST01_01.lab vs D:/Thacsi/NNN/htk1/  
test_mfc/TEST01_01.rec  
LAB: mooit  
REC: bary  
TEST01_02.rec: 0.00( 0.00) [H= 0, D= 0, S= 1, I= 0, N= 1]  
Aligned transcription: D:/Thacsi/NNN/htk1/test_mfc/TEST01_02.lab vs D:/Thacsi/NNN/htk1/  
test_mfc/TEST01_02.rec  
LAB: hai  
REC: bary  
TEST01_03.rec: 0.00( 0.00) [H= 0, D= 0, S= 1, I= 0, N= 1]  
Aligned transcription: D:/Thacsi/NNN/htk1/test_mfc/TEST01_03.lab vs D:/Thacsi/NNN/htk1/  
test_mfc/TEST01_03.rec  
LAB: ba  
REC: bary  
  
----- Overall Results -----  
SENT: %Correct=5.00 [H=3, S=57, N=60]  
WORD: %Corr=23.33, Acc=-46.67 [H=14, D=0, S=46, I=42, N=60]  
=====
```

Thực hiện kiểm thử trên hmm14, hmm13 để có các kết quả khác.

### 3 Kịch bản thử nghiệm

- Thử nghiệm với đặc trưng MFCC\_0\_D\_A

Trong đó: MFCC là đặc trưng bố cục khoang miệng khi phát âm, \_0 là độ lớn của vector, \_D là đạo hàm bậc 1 (đặc trưng tương quan theo thứ tự), \_A là đạo hàm bậc 2 (đặc trưng tương quan theo thời gian). \_X nghĩa là kết hợp MFCC với đặc trưng X (X là 0, D, A).

- Tiến hành kiểm thử qua 3 hmm: **hmm13**, **hmm14**, **hmm15**.

### 4 Kết quả, vẽ biểu đồ

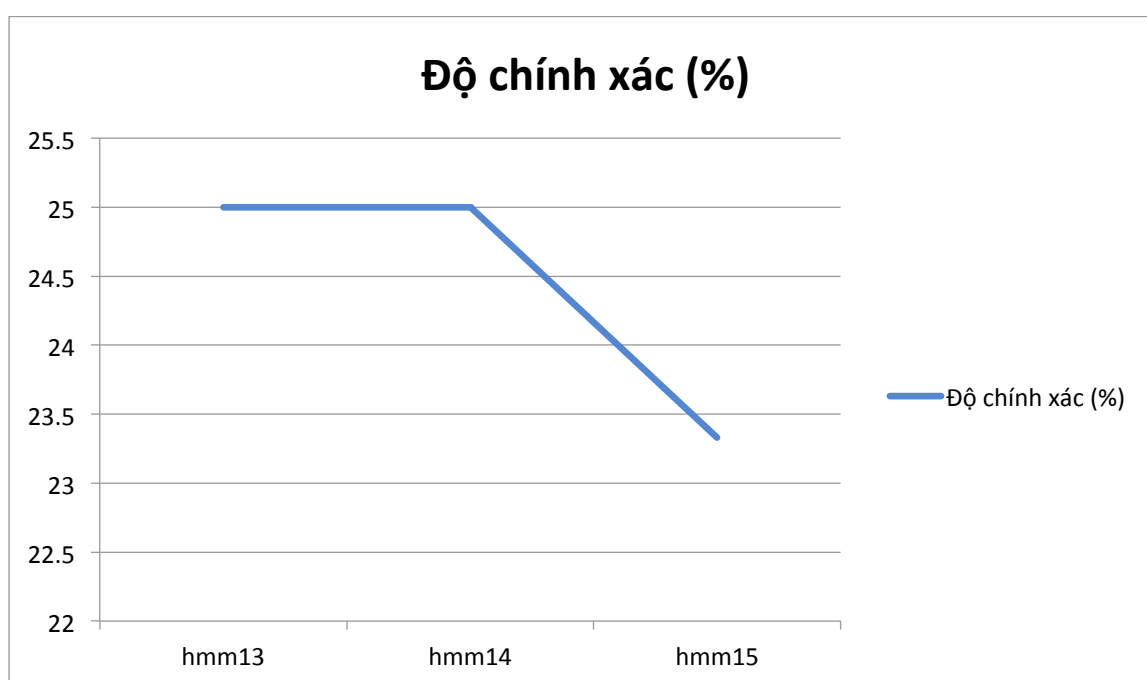
#### 4.1 Kết quả thực nghiệm

Tiến hành kiểm thử qua 3 hmm: **hmm13**, **hmm14**, **hmm15**.

<b>Đặc trưng MFCC_0_D_A Với các hmm</b>	<b>Độ chính xác</b>	<b>Tổng số từ</b>	<b>Số từ đúng</b>	<b>Lỗi chèn</b>
Hmm13	25%	60	15	40
Hmm14	25%	60	15	40
Hmm15	23.33%	60	14	42

## 4.2 Biểu đồ

Biểu đồ đánh giá thay đổi độ chính xác qua các hmm



## 5 Đánh giá cá nhân

- Việc kết hợp MFCC và các đặc trưng khác như đạo hàm bậc 1, đạo hàm bậc 2 làm tăng khả năng nhận dạng của mô hình.
- Chưa thực hiện kiểm thử được cho các kịch bản MFCC\_0\_D, MFCC\_0\_A, ...; thay đổi số vòng lặp huấn luyện, ...
- Việc sử dụng HTK đòi hỏi nhiều thời gian nghiên cứu nên các thông số cũng như các hàm trong các bước trên chưa được mô tả rõ ràng và có thể phát sinh lỗi.

- **Vấn đề:** Chưa thực hiện chạy được toàn bộ tập dữ liệu vivos. Vấn đề xảy ra vì tập dữ liệu quá lớn, khi thực hiện để bước huấn luyện hmm13 thì quá mất thời gian. Nên nhóm dùng 4 tập tin train và 60 tập tin test với mục đích thể hiện hiểu và thực hiện được các bước của quá trình train.

## 6 Tài liệu tham khảo

- Các bước thực hiện huấn luyện và kiểm thử phía trên được tham khảo chủ yếu (có chỉnh sửa, thêm hình ảnh và giải thích) từ file tham khảo ở htk\_training\_vn.pdf và htkbook.pdf