

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

Robots Móviles

Reporte Práctica 4: “Reconocimiento de voz usando pocketsphinx”

Semestre 2020-2
Grupo: 1 (Plan 2016 Computación)

Durón Luna Oscar



Fecha de entrega: 23 de abril de 2020

Desarrollo

1. Ejemplos de frases que se pueden reconocer de acuerdo con la gramática 'restaurant.gram'

Puede ser desde dos "ramas", una que empieza con "Justina" compuesta de varias palabras mas como por ejemplo "no", "yes", "wait", entre otras. La otra rama es empezando con "I want" seguida de diversas palabras como "cookies" por ejemplo y después puede (o no) ser seguido de la palabra "and" mas otra selección de las mismas palabras anteriores.

Concretamente el código es el siguiente:

```
3 public <command> = (<justina> | <wantCombo> | <wantBeverage>);
4 <justina> = JUSTINA (YES | NO | WAIT | START | TAKE A ORDER | THIS IS THE ORDER);
5 <wantCombo> = I WANT [A] (PRINGLES | CHIPS | PASTA | COOKIES | PASTA | NOODLES | TUNA FISH | PICKLES | CHOCO FLAKES
  | ROBO O'S | MUESLI | M AND M) AND (PRINGLES | CHIPS | PASTA | COOKIES | PASTA | NOODLES | TUNA FISH | PICKLES |
  CHOCO FLAKES | ROBO O'S | MUESLI | M AND M);
6 <wantBeverage> = I WANT A (TEA | BEER | COKE | WATER);
7
```

2. Explicación breve de lo que significan los cuatro parámetros mencionados en el punto 5 de la diapositiva 19.

Se cambian todos los archivos necesarios para que ahora se procesen los comandos de un archivo llamado restaurant que contiene otras palabras que reconoce, por ejemplo la gramática que va a usar (inciso anterior).

3. Capturas de pantalla donde se observen las frases reconocidas.

Ejemplos para la "rama" *I want*:

```
[INFO] [1587696369.632512]: OUTPUT: "I WANT CHOCO FLAKES AND PASTA"
INFO: cmn_live.c(120): Update from < 40.24 12.96 6.44 11.27 5.79 1.31 -0.64 3.48 -0.69 5.53 1.70 0.86 -0.74 >
INFO: cmn_live.c(138): Update to < 38.61 15.03 6.75 9.16 3.71 -0.98 -2.11 3.37 -1.07 4.90 0.90 0.03 -2.11 >
INFO: fsg_search.c(859): 124 frames, 7485 HMMs (60/fr), 16686 senones (134/fr), 1765 history entries (14/fr)
```

```
[INFO] [1587696362.674779]: OUTPUT: "I WANT MUESLI AND M AND M"
INFO: cmn_live.c(88): Update from < 40.79 16.70 6.23 11.29 4.88 -0.54 -0.94 4.30 -0.47 3.76 2.12 1.08 -2.38 >
INFO: cmn_live.c(105): Update to < 40.29 17.25 6.67 11.44 5.20 -0.23 -0.66 4.56 -0.29 3.75 2.05 1.22 -2.27 >
INFO: cmn_live.c(120): Update from < 40.29 17.25 6.67 11.44 5.20 -0.23 -0.66 4.56 -0.29 3.75 2.05 1.22 -2.27 >
INFO: cmn_live.c(138): Update to < 38.85 16.93 7.41 12.07 6.09 0.33 -0.26 4.95 0.25 3.69 1.88 1.50 -1.86 >
INFO: fsg_search.c(859): 62 frames, 2543 HMMs (41/fr), 5957 senones (96/fr), 756 history entries (12/fr)
```

```
INFO: fsg_search.c(869): fsg 0.28 CPU 0.313 xRT
INFO: fsg_search.c(871): fsg 0.99 wall 1.095 xRT
[INFO] [1587696337.521700]: OUTPUT: "I WANT A TEA"
INFO: cmn_live.c(120): Update from < 40.43 14.98 5.90 13.34 6.19 0.52 -3.28 5.39 -1.25 1.28 -0.20 -0.05 -4.29 >
INFO: cmn_live.c(138): Update to < 40.16 12.91 6.58 13.94 6.22 0.93 -2.68 5.52 -1.24 1.42 -0.21 -0.04 -3.93 >
INFO: fsg_search.c(859): 73 frames, 1128 HMMs (15/fr), 3084 senones (42/fr), 350 history entries (4/fr)
```

Ejemplos para la "rama" *Justina*:

```
INFO: fsg_search.c(869): fsg 0.30 CPU 0.323 xRT
INFO: fsg_search.c(871): fsg 1.80 wall 1.935 xRT
[INFO] [1587696184.334289]: OUTPUT: "JUSTINA YES"
INFO: cmn_live.c(88): Update from < 48.58 14.07 -0.91 8.45 3.30 9.21 0.48 -3.00 -3.24 9.52 -0.64 0.91 2.05 >
INFO: cmn_live.c(105): Update to < 49.05 14.29 -1.16 8.11 3.06 10.34 0.66 -3.86 -3.79 9.62 -0.43 0.77 2.46 >
INFO: cmn_live.c(120): Update from < 49.05 14.29 -1.16 8.11 3.06 10.34 0.66 -3.86 -3.79 9.62 -0.43 0.77 2.46 >
INFO: cmn_live.c(138): Update to < 48.85 14.89 -2.09 6.76 2.51 11.44 1.55 -4.58 -4.35 10.13 -0.30 1.02 2.50 >
INFO: fsg_search.c(859): 124 frames, 2488 HMMs (20/fr), 7313 senones (58/fr), 579 history entries (4/fr)
```

```

INFO: fsg_search.c(869): fsg 0.61 CPU 0.224 xRT
INFO: fsg_search.c(871): fsg 3.60 wall 1.324 xRT
[INFO] [1587696139.391336]: OUTPUT: "JUSTINA THIS IS THE ORDER"
INFO: cmn_live.c(88): Update from < 46.07 10.04 7.97 14.75 4.05 1.51 -3.20 2.97 0.16 5.30 -0.29 1.46 2.83 >
INFO: cmn_live.c(105): Update to < 44.28 10.41 8.54 13.91 3.93 0.88 -2.94 2.99 0.32 5.81 0.02 1.19 2.35 >
INFO: cmn_live.c(120): Update from < 44.28 10.41 8.54 13.91 3.93 0.88 -2.94 2.99 0.32 5.81 0.02 1.19 2.35 >
INFO: cmn_live.c(138): Update to < 38.70 10.75 7.42 13.80 7.57 1.45 -2.38 4.82 0.55 5.39 0.45 0.97 0.93 >
INFO: fsg_search.c(859): 301 frames, 9462 HMMs (31/fr), 23238 senones (77/fr), 2415 history entries (8/fr)

```

4. Copia del archivo launch con los parámetros modificados.

Código del archivo pocketsphinx_test.launch modificado:

```

1  <launch>
2    <node name="speech_recog" pkg="pocketsphinx" type="asr_test.py" output="screen">
3      <remap from="jsrf_audio" to="sphnx_audio" />
4      <param name="gram" value="$(find pocketsphinx)/vocab/restaurant"/>
5      <param name="dict" value="$(find pocketsphinx)/vocab/restaurant.dict"/>
6      <param name="grammar" value="restaurant"/>
7      <param name="rule" value="command"/>
8      <param name="lm" value=":default"/>
9      <param name="hmm" value=":default"/>
10   </node>
11
12   <node name="audio_control" pkg="pocketsphinx" type="send_audio.py" output="screen">
13     <param name="input" value=":default"/>
14   </node>
15 </launch>
16

```