

Corpus AIRA

Emulación de Ambientes Acústicos

Objetivos del Curso

- Para el curso, vamos a reducir sus objetivos a:
 - Estimar la dirección de **una** fuente sonora.
 - O *varias*, si me quieren impresionar.
 - Filtrar esa(s) fuente(s) del resto del ambiente.
 - También conocido como “separación de fuentes”.
- Por cierto, este es el proyecto final...

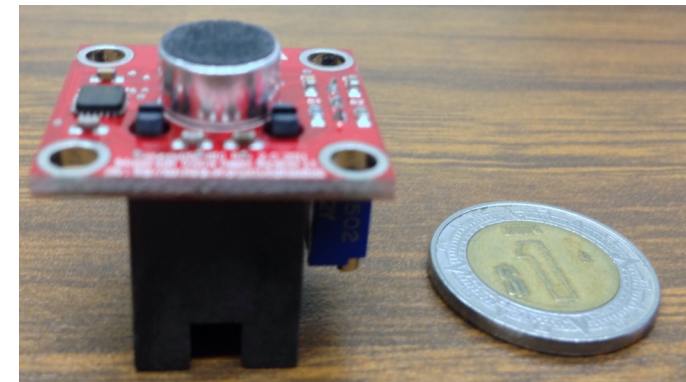
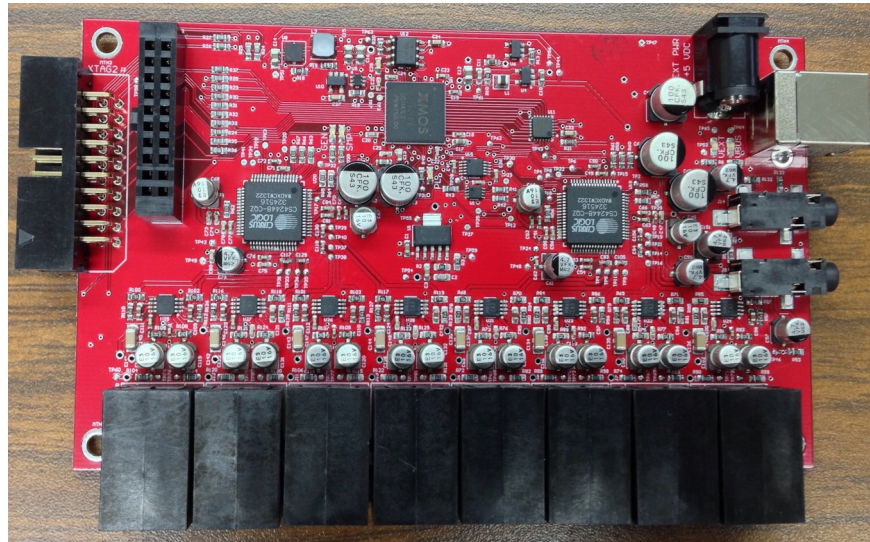
Objetivos del Curso

- Esto involucra el procesamiento de varias señales (micrófonos) *concurrentes*.
 - Hasta 8 micrófonos en algunos casos.
- Probar estas técnicas se vuelve complicado sin el hardware de captura adecuado.
- ¿Que hacemos?

Solución #1

- Les compro a todos ustedes una tarjeta como la 8SoundsUSB del Proyecto ManyEars:

1,000 CAN ~ 15,600 MXN



Solución #1

- O como una interface de M-Audio Fast Track Ultra y micrófonos activos:

500 USD ~ 10,100 MXN por la interface

200 USD ~ 4,100 MXN por cada micrófono



Solucion #2

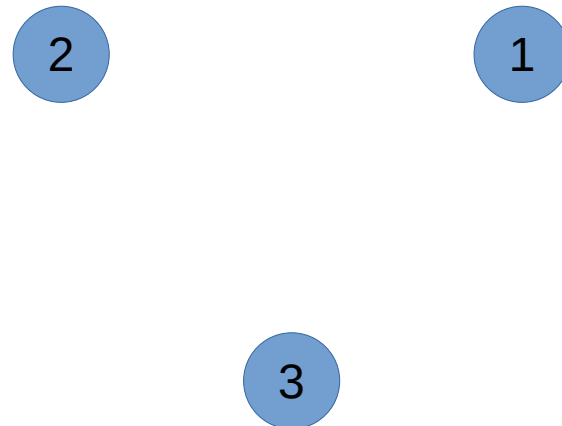
- Hacemos puros ejercicios con sólo un micrófono.
 - Realmente no es solución.

Solucion #3

- Haber grabado anteriormente con hardware ya especializado.
- Y tener un cliente de JACK que emule las entradas de los micrófonos con dichas grabaciones.
- Les presento el corpus: *Acoustic Interactions for Robot Audition* (AIRA).

AIRA Corpus

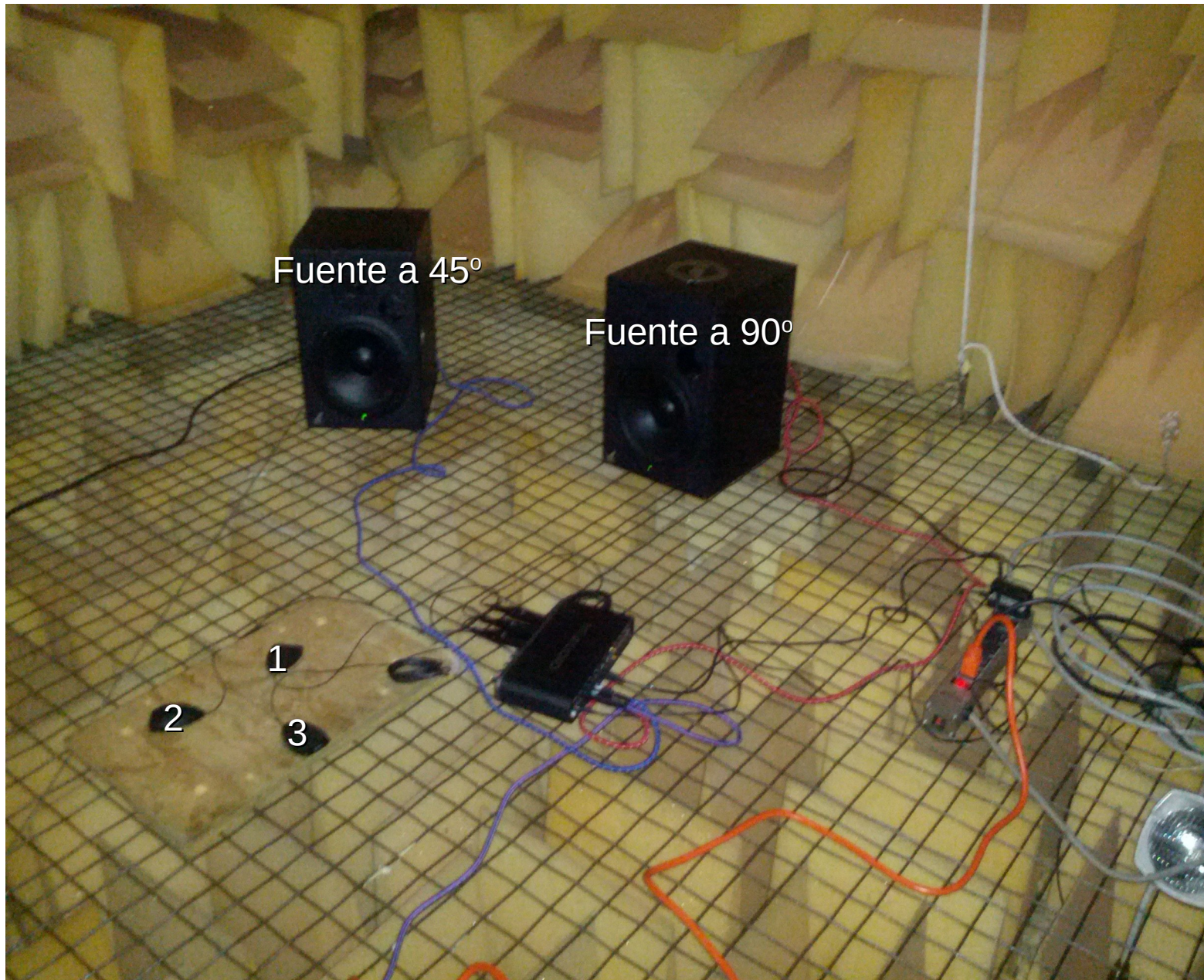
- Hardware utilizado:
 - Fast Track Ultra
 - 3 micrófonos omnidireccionales Shure MX390
 - Triángulo equilátero, con la siguiente configuración:



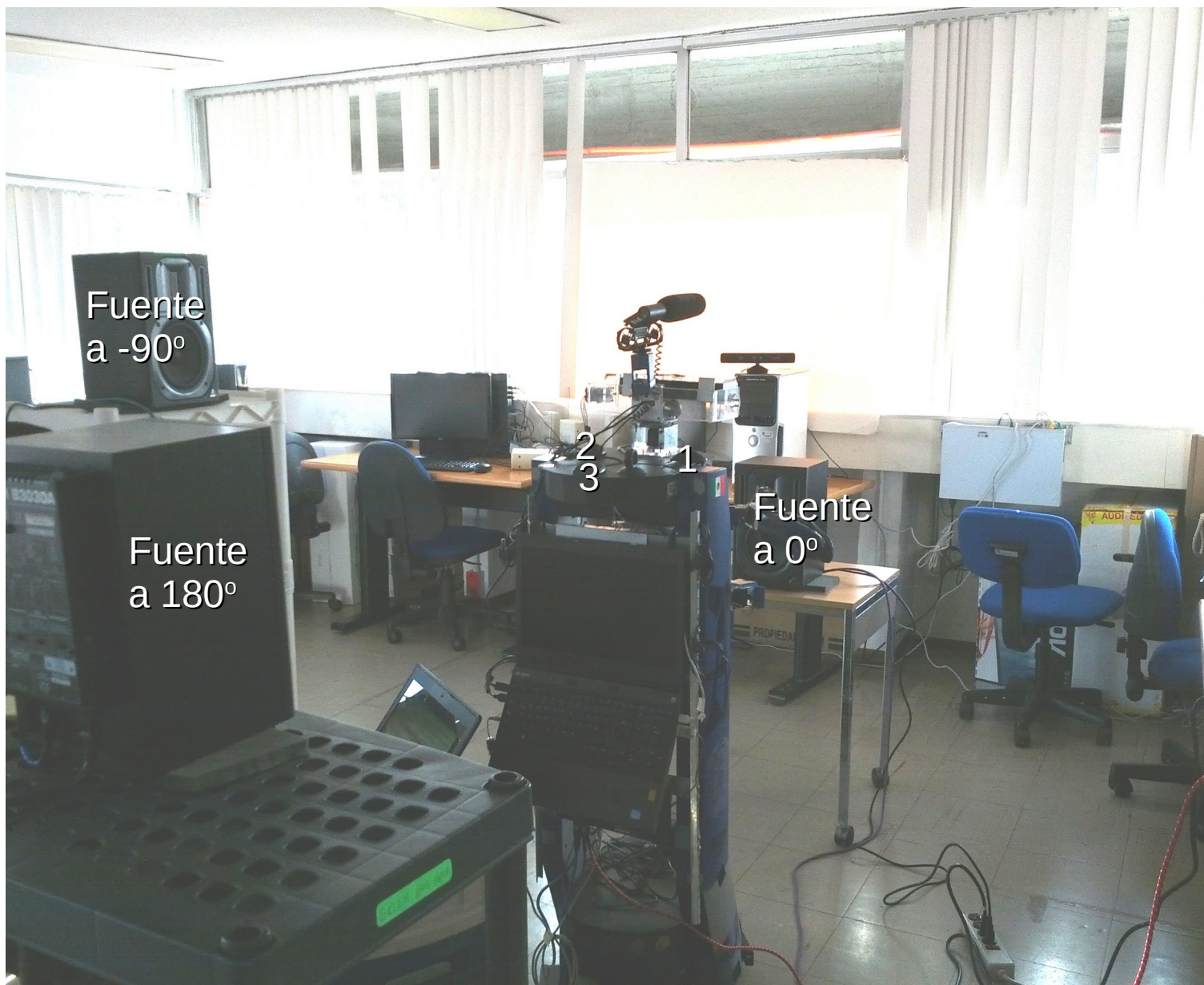
AIRA Corpus

- En dos ambientes diferentes:
 - Cámara anecoica del Laboratorio de Acústica y Vibraciones del ICAT-UNAM.
 - Laboratorio de estudiantes del Departamento de Ciencias de la Computación del IIMAS-UNAM.

Cámara Anecóica



Laboratorio



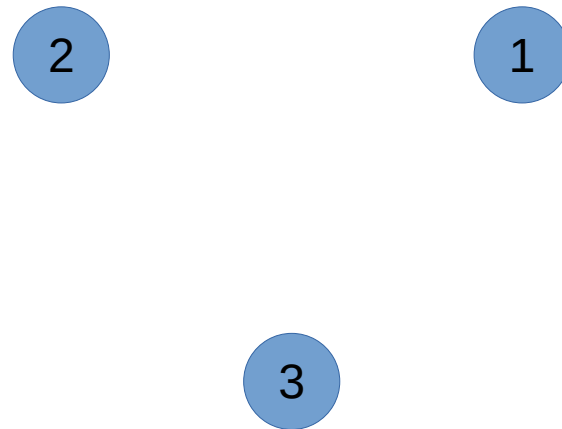
AIRA Corpus

- De 1 a 4 fuentes sonoras estáticas:
 - Utilizando grabaciones al azar del Corpus DIMEx100, en español.
 - Reproducidas en bocinas para estudios de grabación (tipo monitor) Behringer Truth B3030A.



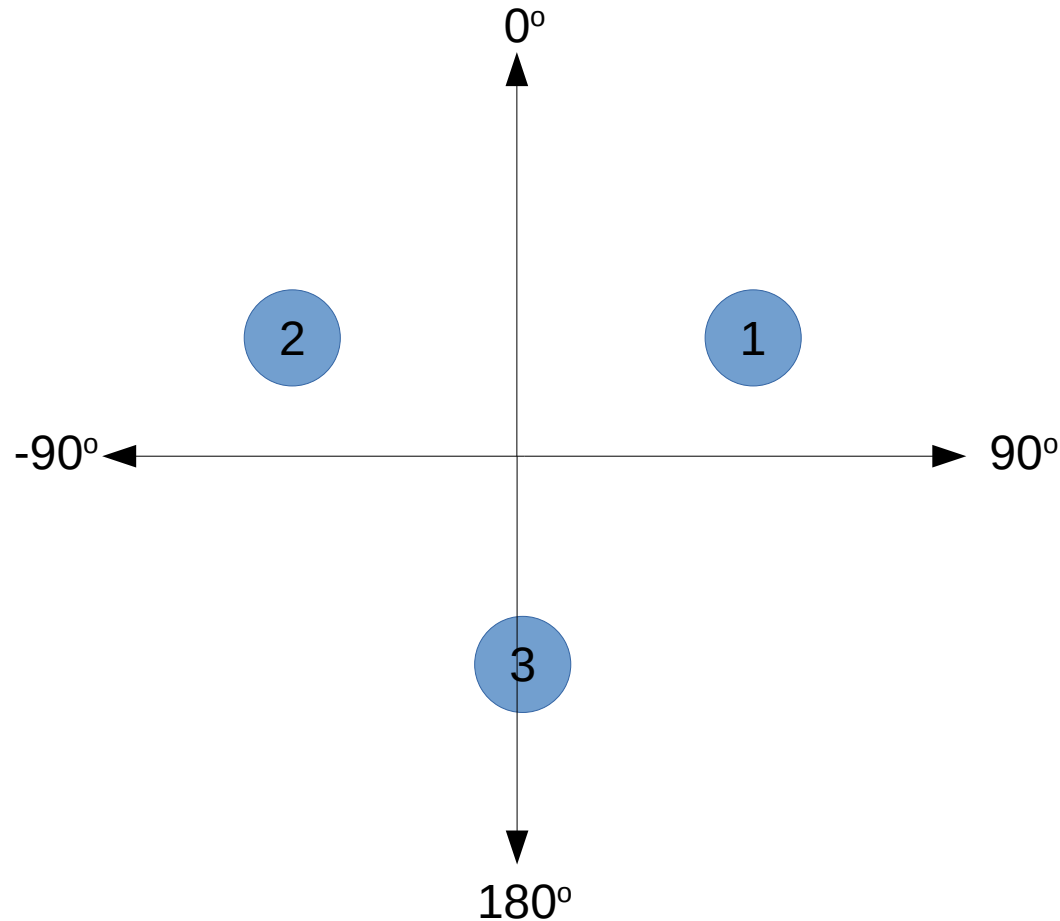
AIRA Corpus

- De 1 a 4 fuentes sonoras estáticas.



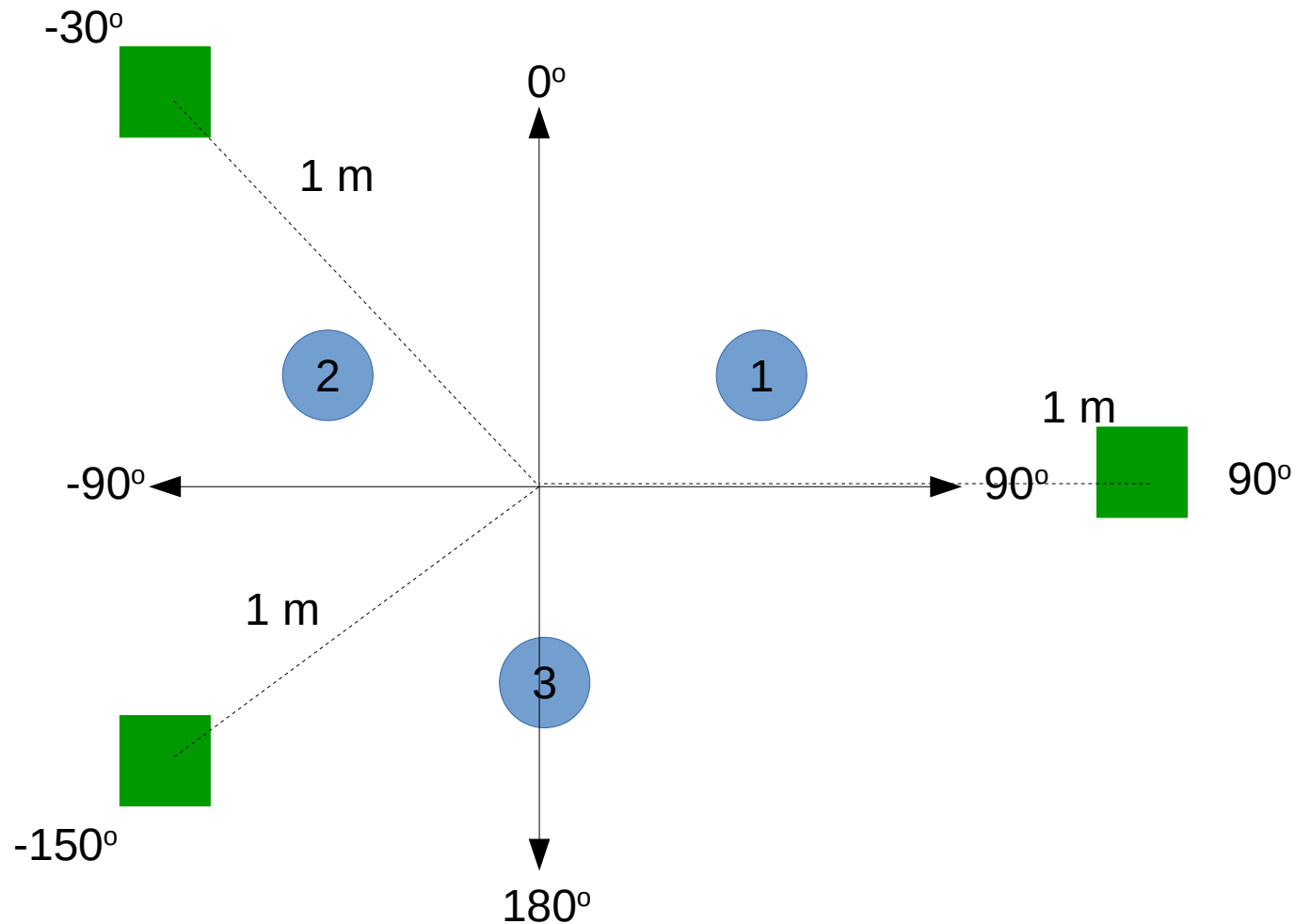
AIRA Corpus

- De 1 a 4 fuentes sonoras estáticas.



AIRA Corpus

- De 1 a 4 fuentes sonoras estáticas.



AIRA Corpus

- Para cada fuente (bocina):
 - Se tiene la grabación antes de ser reproducida.
 - Se tiene lo que se dijo en texto.
 - Se tiene su posición (en ángulo y distancia del centro del arreglo).

Pero...

- Para propósitos de este curso, pueden descargar una parte de este corpus de la página del curso.
 - Para 48 kHz --- `corpus48000.tar.gz`
 - Para 44.1 kHz --- `corpus44100.tar.gz`
- Es una descarga de entre 166 y 180 MB.

corpusxxxxx.tar.gz

- Contiene un folder, con varios subdirectorios:
 - corpus48000
 - clean-1source
 - clean-2source
 - clean-3source
 - clean-4source
 - noisy-1source
 - noisy-2source
 - noisy-3source
 - noisy-4source

Subdirectorios

- Si comienza con “clean”:
 - Grabado en la cámara anecóica, sin ruido.
- Si comienza con “noisy”:
 - Grabado en el laboratorio, con ruido.

Subdirectorios

- Cada subdirectorio contiene los siguientes archivos:
 - info.txt
 - pristine_channel1.wav
 - pristine_channel2.wav ...
 - wav_mic1.wav
 - wav_mic2.wav
 - wav_mic3.wav

wav_micx.wav

- Cada archivo wav_micx.wav es lo grabado en el micrófono del canal x.
- Su frecuencia de muestro es la que indica el nombre del corpus.

wav_mic2.wav

2

1

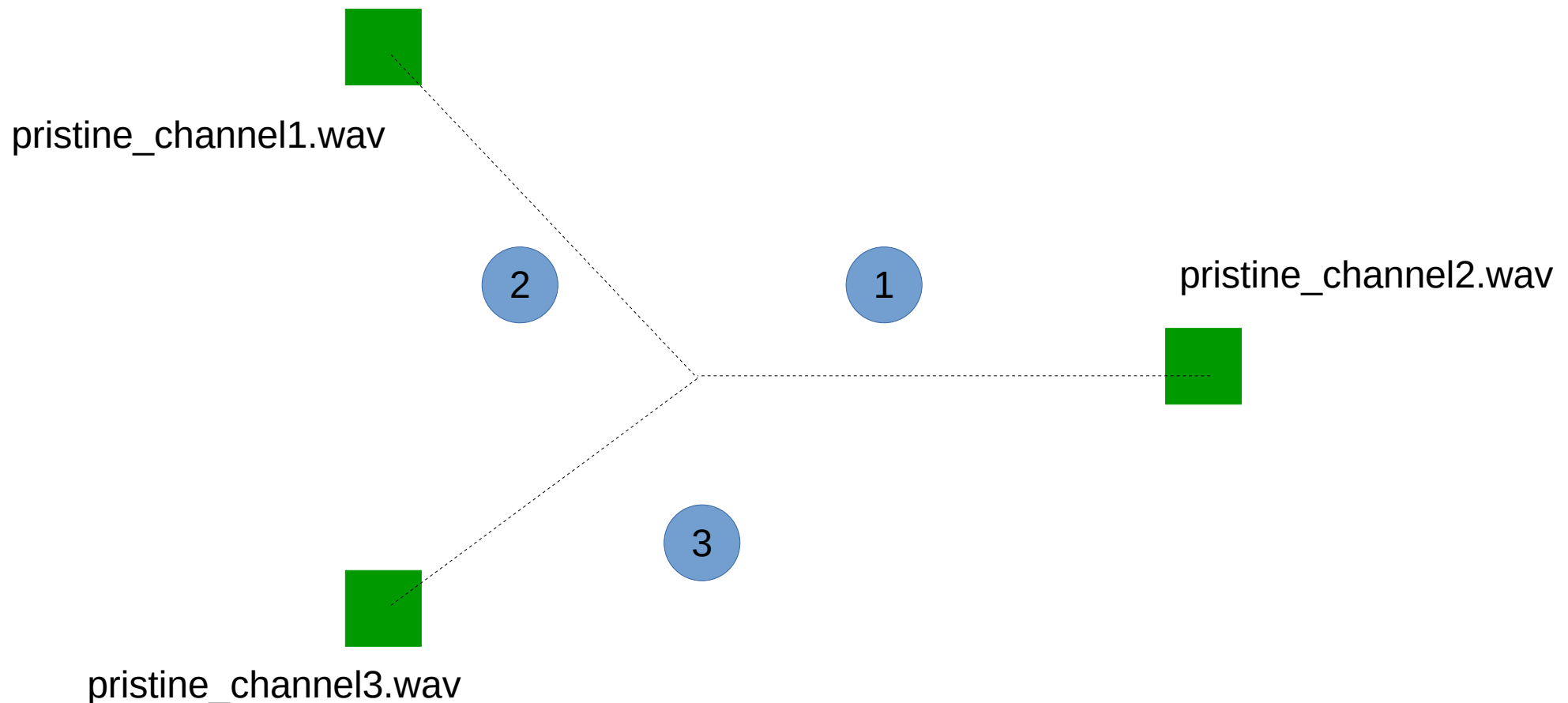
wav_mic1.wav

3

wav_mic3.wav

pristine_channel.y.wav

- Cada archivo `pristine_channel.y.wav` es lo reproducido por la fuente `y`.



info.txt

- El archivo info.txt contiene información adicional del ambiente, específicamente:
 - Primera línea: separación en metros entre los micrófonos.
 - Segunda línea: ubicación en grados de las fuentes sonoras, separadas por comas.

Emular el Ambiente

- Se utilizan los clientes JACK llamado **ReadMicWavs** y **ReadMicWavsMulti**
 - No se conectan a las salidas del sistema, sino a las entradas de otro agente.
 - Y reproducen lo grabado en los diferentes ambientes, un canal por cada archivo wav_micx.wav.
 - Así, emulan los canales como si fueran los micrófonos ya que el agente a cual se conectan, por la forma en que funciona JACK, no percibe la diferencia.
 - Al terminar de reproducir, se mueren automáticamente.

Emular el Ambiente

- Descárguenlos de la página del curso, junto con su Makefile.
- Compílenlos.

ReadMicWavs

- Utiliza la librería libsndfile para leer los archivos WAV, pero no utiliza la librería de libsamplerate para cambiar sus frecuencias de muestreo.
- ¿Por qué?
 - Libsamplerate cambia, de manera pequeña, los datos de los WAV, por lo que ya no sería fiel a lo que realmente fue grabado.
 - Por eso, dos conjuntos de grabaciones, uno por cada frecuencia de muestreo.

ReadMicWavs

- Para correr, se tiene que pasar como argumentos:
 - El nombre del agente de JACK al que va a conectar sus salidas para emular los micrófonos.
 - El nombre de “base” de las entradas del agente de JACK.
 - La localización de los wav_micx.wav.
 - Cuantos micrófonos va a emular (de 1 a 3).

ReadMicWavs

- El nombre del agente es el nombre que ustedes le pusieron a su cliente.

ReadMicWavs

- El nombre “base” de las entradas es el nombre al que le va a agregar un número para identificar el nombre completo de la entrada:
 - input1, input2 --- su nombre base es “input”
 - input_1, input_2 --- su nombre base “input_”

ReadMicWavs

- Digamos que el nombre base es “in_”
- ReadMicWavs asume que:
 - Nombre de primera entrada del agente es:
 - El nombre base más el número 1 (in_1)
 - Por aquí va a emular el primer canal con wav_mic1.wav
 - Nombre de segunda entrada del agente es:
 - El nombre base más el número 2 (in_2)
 - Por aquí va a emular el segundo canal con wav_mic2.wav
 - Nombre de tercera entrada del agente es:
 - El nombre base más el número 3 (in_3)
 - Por aquí va a emular el tercer canal con wav_mic3.wav

ReadMicWavs

- La localización de los archivos wav_micx.wav es algún subdirectorio dentro de corpusxxxxx:
 - corpus48000/clean-1source
 - corpus44100/noisy-3source

ReadMicWavs

- El número de micrófonos a emular debe ser un número entre 1 y 3.
 - 1: emulará canal 1
 - 2: emulará canal 1 y 2
 - 3: emulará canal 1, 2 y 3

ReadMicWavs Ejemplo

- Hemos descargado el corpus de 48 kHz.
- Y lo hemos descomprimido tal que lo tenemos ubicado en:

`/home/caleb/ar/corpus48000`

- Hemos también compilado a ReadMicWavs y está ubicado en:

`/home/caleb/ar/ReadMicWavs`

ReadMicWavs Ejemplo

- Digamos que tenemos:
 - Un agente llamado “in_to_out”
 - Con 2 canales de entrada.
 - Llamados “inp1” y “inp2”
 - Queremos emular el ambiente de:
 - Laboratorio (noisy).
 - Con tres fuentes.

ReadMicWavs Ejemplo

- En este ejemplo se debe:
 - Primero correr el agente in_to_out y desconectarlo de las entradas del sistema.
 - Y luego correr en el folder /home/caleb/ar :
`./ReadMicWavs in_to_out inp corpus48000/noisy-3source 2`

ReadMicWavs Ejemplo

- En este ejemplo se debe:
 - Primero correr el agente in_to_out y desconectarlo de las entradas del sistema.
 - Y luego correr en el folder /home/caleb/ar :

`./ReadMicWavs in_to_out inp corpus48000/noisy-3source 2`

Nombre
de agente

Nombre
base

Localización de
wav_micx.wav

Número de
canales

ReadMicWavs

- ReadMicWavs es un agente de JACK común y corriente, pero no se conecta a las salidas del sistema.
 - Se conecta a las entradas de otro agente.
- Si se quisiera escuchar lo que está reproduciendo, es necesario hacer la conexión de sus salidas a las salidas del sistema manualmente.

ReadMicWavsMulti

- ReadMicWavs conecta las salidas en orden:
 - output1 a input1 del agente
 - output2 a input2 del agente
 - output3 a input3 del agente
- Y sólo trabaja con hasta tres salidas.
- ReadMicWavsMulti permite reordenar y no tiene esa limitación.

ReadMicWavsMulti

- ReadMicWavs es más que suficiente para trabajar con la versión de AIRA que está en la página del curso.
- Pero AIRA tiene otros tipos de arreglos que pueden ser utilizados.

ReadMicWavsMulti

- Para correr, se tiene que pasar como argumentos:
 - El nombre del agente de JACK al que va a conectar sus salidas para emular los micrófonos.
 - El nombre de “base” de las entradas del agente de JACK.
 - La localización de los wav_micx.wav.
 - Una serie de números, cada uno representando una entrada, en el orden de la conexión.

ReadMicWavsMulti Ejemplo

- Con las mismas circunstancias que antes, ReadMicWavsMulti haría lo mismo que ReadMicWavs corriendo:

```
./ReadMicWavsMulti in_to_out inp corpus48000/noisy-3source 2 1 2
```

ReadMicWavsMulti Ejemplo

- Con las mismas circunstancias que antes, ReadMicWavsMulti haría lo mismo que ReadMicWavs corriendo:

`./ReadMicWavsMulti in_to_out inp corpus48000/noisy-3source 2 1 2`

Nombre de agente Nombre base Localización de wav_micx.wav Número de canales Orden de canales

Resultando en las siguientes conexiones:

output1 a inp1

output2 a inp2

ReadMicWavsMulti Ejemplo 2

- Si quisiéramos cambiar el orden de las conexiones, se correría casi lo mismo, pero con el orden invertido:

`./ReadMicWavsMulti in_to_out inp corpus48000/noisy-3source 2 2 1`

Nombre de agente Nombre base Localización de wav_micx.wav Número de canales Orden de canales

Resultando en las siguientes conexiones:

output1 a inp2

output2 a inp1

ReadMicWavsMulti Ejemplo 3

- Si quisiéramos que la primera salida no se conecte, se poner un 0 en la lista que le corresponde en el orden:

./ReadMicWavsMulti in_to_out inp corpus48000/noisy-3source 2 0 2

Nombre de agente	Nombre base	Localización de wav_micx.wav	Número de canales	Orden de canales
------------------	-------------	------------------------------	-------------------	------------------

Resultando en las siguientes conexiones:

output1 no se conecta

output2 a inp2

Ejercicio #1

- Correr ReadMicWavs o ReadMicWavsMulti con alguno de los agentes que hayan ya desarrollado.
- Conectar manualmente alguna de las salidas de ReadMicWavs o ReadMicWavsMulti a alguna del sistema.

Ejercicio #2

- Desarrollar un agente con dos canales y que su salida sea la resta del primer canal menos el segundo.
- Correr con ReadMicWavs o ReadMicWavsMulti y visualizar la salida con Baudline.
 - ¿Que escuchan?

Referencia oficial de AIRA

- Este corpus es una versión pequeña del corpus completo que se puede descargar de:

<https://aira.iimas.unam.mx>

- Su referencia completa es:

C. Rascon, I. Meza, A. Millan-Gonzalez, I. Velez, G. Fuentes. Acoustic interactions for robot audition: A corpus of real auditory scenes. The Journal of the Acoustical Society of America, 144 (5), 2018.

AIRA

- Tiene grabaciones de más ambientes:
 - Tienda UNAM
 - Cafetería de Química
- Tiene grabaciones con más micrófonos: 16 en una base de 3 dimensiones.
- Se incluyeron fuentes móviles.
- Se incluyó que el arreglo de micrófonos sea móvil.

Todavía no lo terminamos...

- Más ambientes.
- Grabaciones con diferentes niveles de volumen de las fuentes sonoras.
- Más grabaciones con fuentes móviles.

wink wink

Siguiente clase:

Estimación de Dirección de Arribo de Fuentes
Sonoras