

Grado en Informática

Interfaces Inteligentes. 2018-2019

Interfaces multimodales.

Reconocimiento de voz en Unity3D

Objetivo El objetivo de esta práctica es aprender a utilizar las herramientas de reconocimiento de voz que Unity3D ofrece para Windows 10. Para la utilización de estas herramientas se debe importar el paquete UnityEngine.Windows.Speech. Hay que tener en cuenta que este paquete sólo está disponible para el sistema operativo Windows 10.

De este paquete utilizaremos las clases KeywordRecognizer y DictationRecognizer.

KeywordRecognizer La clase KeywordRecognizer intenta casar el audio recibido por el micrófono del ordenador o teléfono móvil con una lista de palabras clave proporcionadas. Es posible mantener varios objetos de esta clase activos al mismo tiempo, siempre que sus conjuntos de palabras no tengan ningún elemento común.

Para iniciar un objeto de este tipo, se utiliza el constructor KeywordRecognizer (string[] lista), donde lista contiene todas las palabras a detectar. Un objeto de esta clase puede activarse y desactivarse con sus funciones Start() y Stop(), respectivamente. Es importante llamar a la función Dispose() al terminar de usarlo para liberar sus recursos. Debe asociarse al objeto una función de escucha del tipo void función (PhraseRecognizedEventArgs datos), donde el parámetro datos contiene:

- La frase recibida datos.text, de tipo string;
- El grado de certeza datos.confidence, de tipo enumerado Windows.Speech.ConfidenceLevel, cuyos valores pueden ser High, Medium, Low y Rejected;
- La duración de la frase datos.phraseDuration, de tipo TimeSpan;
- El momento de comienzo de la frase datos.phraseStartTime, de tipo DateTime;
- El significado semántico de la frase datos.semanticMeanings, de tipo SemanticMeaning[]. Este atributo no tiene utilidad para KeywordRecognizer, pero sí para otras clases de reconocimiento de voz como GrammarRecognizer, que debido a su complejidad no se usará en esta práctica.

Una vez creada la función de escucha, se asociará al objeto KeywordRecognizer mediante la instrucción objeto.OnPhraseRecognized += función.

DictationRecognizer La clase DictationRecognizer intenta transcribir el audio recibido por el micrófono del ordenador o teléfono móvil. Un objeto de esta clase puede activarse y desactivarse con sus funciones Start() y Stop(), respectivamente. Es importante llamar a la función Dispose() al terminar de usarlo para liberar sus recursos. Un objeto DictationRecognizer espera que se le asocien cuatro funciones de escucha:

- objeto.DictationResult, llamada cada vez que se reconozca una frase y que será de tipo void función (string texto, Windows.Speech.ConfidenceLevel confianza)
- objeto.DictationHypothesis, llamada cada vez que se actualice la frase capturada, de tipo void función (string texto)

- objeto.DictationComplete, llamada cada vez que se interrumpa el dictado, de tipo void función (Windows.Speech.DictationCompletionCause causa), donde el parámetro causa es de tipo enumerado y sus valores pueden ser los siguientes: Enlaces de interés
- Uso del micrófono: <https://docs.unity3d.com/ScriptReference/Microphone.html>
- Uso de la cámara: <https://docs.unity3d.com/ScriptReference/WebCamTexture.html>
 - Complete, si el dictado terminó correctamente;
 - AudioQualityFailure, si el dictado se interrumpió por mala calidad del audio;
 - Canceled, si el dictado fue cancelado o la aplicación fue pausada;
 - TimeoutExceeded, si se excede el tiempo estipulado para la captura;
 - NetworkFailure, si el dictado se interrumpió por un fallo de red;
 - MicrophoneUnavailable, si el dictado se interrumpió por no detectarse el micrófono;
 - UnknownError, en cualquier otro caso.
- objeto.DictationError, llamada cuando se detecte un error durante el dictado, de tipo void función (string error, int resultado).

En todos los casos, las funciones de escucha se asociarán al atributo correspondiente del objeto `DictationRecognizer` mediante la instrucción `objeto.atributo += función`.

Qué se pide

Para la realización de esta práctica, se pedirá mostrar el funcionamiento de `KeywordRecognizer`, indicando qué palabras se espera capturar, y de `DictationRecognizer`, mostrando las frases devueltas por el objeto. Ambas clases no deben estar activas al mismo tiempo, sino que el usuario deberá poder escoger cuál quiere utilizar.

Enlaces de interés

Documentación de la clase `KeywordRecognizer`:
<https://docs.unity3d.com/ScriptReference/Windows.Speech.KeywordRecognizer.html>

Documentación de la clase `DictationRecognizer`:

<https://docs.unity3d.com/ScriptReference/Windows.Speech.DictationRecognizer.html>