GLOBAL SOUND EXAMPLE

FUNCTIONALITY:

This library allows the robot Nao to be able to interact with a human, through voice recognition and the robot's own speech, and then to be able to follow, with his head or body, a sound, a red ball or a face.

Therefore, the general operation of the library can be divided into 2 stages:

- Human-robot interaction: recognition and synthesizing of voice by the robot.
- Tracking of a sound, ball or face.

HUMAN-ROBOT INTERACTION:

This stage is done with the help of the module "ALSpeechRecognition" and the module "ALTextToSpeech" present in the NAOqi Audio API of the robot.

On the one hand "ALSpeechRecognition" is based on sophisticated speech recognition technologies provided by: ACAPELA GROUP for Nao Version 3.x and NUANCE for Nao Version 4.

Its principle of operation:

- Before starting, "ALSpeechRecognition" needs to be fed with the list of phrases that must be recognized. This will create Nao's private dictionary and it will only recognize the words or phrases contained in it.
- Once started, "ALSpeechRecognition" places in the key 'SpeechDetected', a boolean that specifies whether or not it is heard by a speaker. When something has been heard, the robot will record it.
- If something is hear by a speaker, the item in the list which best fits the listened sound is placed in the 'WordRecognized' key, which is organized as follows:

$$[Phrase\_1, confidence\_1, sentence2, confidence\_2, \ldots, n-phrase, confidence\_n]$$

where:
$Phase\_i$ is one of the predefined phrases.
$Confidence\_i$ is an estimation of the probability that this sentence is actually what has been pronounced by the human speaker.

The different hypotheses contained in 'WordRecognized' are arranged so that the most likely sentences appear first, so it will suffice to access the first position of this vector to know the word or phrase that Nao has heard.

On the other hand the "ALTextToSpeech" module allows the robot to speak. It works by sending commands to a text-to-speech conversion engine provided by ACAPELA, and also authorizes voice personalization. The result of the synthesis is sent to Nao's speakers

causing him to speak.

TRACKING A SOUND, BALL OR FACE:

For this final stage is needed the help of the "ALTracker" module present in the NAOqi Trackers API. The "ALTracker" module is a generic follower that allows the robot to follow different objectives (red ball, face, reference point, sound ...) using different means (head, whole body, movement, etc). Thus, the main objective of this module is to establish a bridge between the detection of objectives and the movement in order to make the robot keep the target in sight in the center of the camera, as shown in the figure 1.

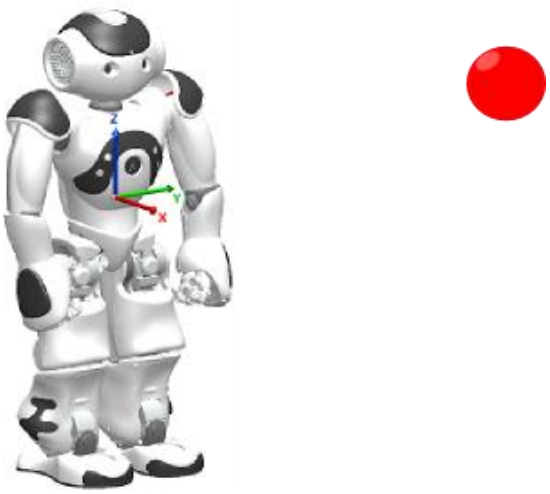

Figure 1. Position of the robot following a colored ball.

"ALTracker" can track the target with any of the modes shown in Table 1 and can in turn follow any of the targets shown in Table 2.

Mode	The robot will track…	Comment
Head	with the head (this is the default mode)	The two head's joint are controlled to track the target.
WholeBody	with whole body and fixed legs	The robot keeps balance autonomously and adapts his posture to track the target.
Move	with move	The robot moves in order to keep a defined distance to the target..

Table 1. Tracking modes

Target	Parameters	Comment
RedBall	diameter of ball (meter)	Used to compute the distance between robot and ball.
Face	width of face (meter)	Used to compute the distance between robot and face..
Sound	[distance, confidence]	distance is used to estimate sound position and confidence to filter sound location.

Table 2. Tracking objectives

OPERATING MODE:

From the modules explained above and combining the two mentioned stages, is possible to obtain a library that allows to make a final example whose operatio is the following one:

1. The configuration of the modules is done by selecting the languages for the recognizer and the speech synthesizer. The vocabulary to be recognized is also added.

2. The robot interacts with the human to ask about the goal to follow (ball, sound or face) and about the tracking mode (head only or body movement towards the target). Nao will ask until it makes sure it has correctly understood what it has to do.

3. With the data obtained from the user in the previous step, the tracker is configured according to the target to be followed and the desired movement mode.

4. The robot will track according to the desired configuration until the program is stopped (Ctrl + C).