CLAIMS

1. A method of performing a dialogue between a humanoid robot (R) and at least one user (U) comprising the following steps, carried out iteratively by said humanoid robot:

i) acquiring a plurality of input signals (s1, s2) from respective sensors (c1, c2), at least one said sensor being a sound sensor and at least one other sensor being a motion or image sensor;

ii) interpreting the acquired signals to recognize a plurality of events (EVI) generated by said user, selected from a group comprising: the utterance of at least a word or sentence, an intonation of voice, a gesture, a body posture, a facial expression;

iii) determining a response of said humanoid robot, comprising at least one event (EVO) selected from a group comprising: the utterance of at least a word or sentence, an intonation of voice, a gesture, a body posture, a facial expression, said determining being performed by applying a set of rules, each said rule associating a set of input events to a response of the robot;

iv) generating, by said humanoid robot, said or each said event;

wherein at least some of said rules applied at said step iii) associate a response to a combination of at least two events jointly generated by said user and recognized at said step ii), of which at least one is not a word or sentence uttered by said user; and

wherein said step ii) comprises searching a match between an acquired signal and an event belonging to a list of expected events stored in a memory of said humanoid robot, or accessible by it, said searching being carried out by successively using a plurality of matching methods (MM1 – MM4) with increasing complexity until an event is recognized with a confidence score greater than a predetermined value, or after the highest recognition method has complexity has been used.

2. A method according to claim 1, wherein the used matching methods are selected depending on a context of dialogue.

3. A method according to any one of the preceding claims, wherein said matching methods include, by order of increasing complexity: the search for an exact match, the search for an approximate match, the search for a phonetic correspondence - only in the case of voice recognition - and the search for a semantic correspondence.

4. A method according to claim 3 wherein said method of searching for a phonetic correspondence comprises:

- a step of phonetic transcription of a set of sounds acquired by a sound sensor;
- a step of simplifying and smoothing the resulting phonetic transcription ;
- calculating an edit distance between said simplified and smoothed phonetic transcription and a plurality of entries, obtained by simplifying and smoothing a predefined set of words in natural language, and
- choosing a natural language word of said predefined set, corresponding to the entry with the lowest edit distance from said simplified and smoothed phonetic transcription.

5. A method according to claim 4 wherein said simplifying and smoothing comprises:

- replacing phonemes prone to confusion by a single phoneme;
- removing vowels other than vowels at the beginning of words and nasal vowels, and
- removing breaks between words.

6. A method according to any one of the preceding claims, wherein said list of expected events is selected, among a plurality of said lists, depending on a dialogue context.

7. A method according to any one of the preceding claims, wherein at least some of said rules applied at said step iii) determine a response comprising at least two events generated jointly by said humanoid robot, of which at least one is not the utterance of a word or sentence.

8. A method according to any one of the preceding claims, wherein, at said step iii, said response of humanoid robot is determined based on at least one parameter selected from: a dialogue context (CTX), the identity of the user, an internal state (RIS) of said humanoid robot.

9. A method according to claim 8, further comprising a step of modifying the value of said or of at least one said parameter according to said at least one event recognized at said step ii) or determined in said step iii).

10. A method according to any one of the preceding claims wherein said step iii) comprises determining a response to a set of events, including the absence of words uttered by said user or identified gestures, by applying rules belonging to a predefined subset (PRO), called proactive rules.

11. A method according to any one of the preceding claims further comprising, if the response determined during step iii) is or comprises at least the utterance of a word or sentence, the execution of a step iii-a) of performing linguistic analysis of the words or sentences to be uttered and determining an animation accompanying said response as a function of said analysis.

12. A method according to claim 11, wherein said step iii-a comprises the substeps of:

α) identifying at least one word of the response to be animated;

β) determining a concept and expressiveness, called one-off expressiveness, associated with said or each said word to be animated;

γ) choosing from a list (ALST) of animations stored in a memory of said humanoid robot, or accessible by it, an animation based on said concept and said one-off expressiveness.

13. A method according to claim 12, wherein said substep α comprises performing a syntactic analysis of a sentence to be uttered to determine each or said word to be animated depending on its function within a structure of said sentence.

14. A method according to any one of claims 12 or 13, wherein, in said substep β, said one-off expressiveness is determined based on at least one parameter selected from: an expressiveness of the word, an expressiveness of one or more other words related to it, and an overall expressiveness of the entire response .

15. A method according to any one of claims 12 to 14, wherein each animation of said list is associated with one or more concepts and has a specific expressiveness, said substep γ including choosing within said list the animation associated with the concept determined in said substep β and having a specific expressiveness closest to said one-off expressiveness.

16. A method according to claim 15 further comprising the following substep:

δ) determining an expressiveness, called final expressiveness, based on said specific expressiveness and said one-off expressiveness.

17. A method according to any one of claims 12 to 16, wherein either said one-off or said final expressiveness determines at least one parameter chosen among a speed and an amplitude of at least one gesture of said animation.

18. A method according to any one of the preceding claims further comprising the following steps, implemented iteratively by said robot simultaneously with said steps i) to iv):

A) determining the position of at least a portion of the body of said user (U) relative to a reference frame fixed to the said robot (R);

B) driving at least one actuator of said robot to maintain the distance between said robot or an element thereof and said at least one or said body part of said user within a predefined range of values .

19. A method according to claim 18 wherein said step B) further comprises driving at least one actuator of said robot to maintain an orientation of the robot with respect to said user in a predetermined angular range .

20. A method according to any one of claims 18 or 19 further comprising the step of:

C) driving said or at least one said actuator to cause said pseudo-random displacements of the robot while maintaining said distance in said predetermined range of values and, where appropriate, said orientation in said predetermined angular range.

21. A method according to any one of claims 18 to 20 further comprising the step of:

D) performing a semantic analysis of an ongoing dialogue between said user and said humanoid robot and, in accordance with said analysis,

changing said predetermined range of distance values and, where appropriate, said predetermined angular range .

22. A method according to one of claims 18 to 21 wherein said step A) comprises determining the position of a lower body of said user relative to said reference frame fixed to the said robot.

23. A computer program product comprising program code instructions for executing a method according to one of the preceding claims when said program is executed by at least one processor embedded on a humanoid robot (R), said robot comprising : a plurality of sensors (c1, c2) operatively connected to said or at least one processor and comprising at least one sound sensor and at least one image or movement sensor, to acquire respective input signals; a speech synthesis module controlled by said or at least one said processor to utter words or sentence; and a set of actuators (A1, A2, A3) driven by said or at least one said processor enabling said robot to perform a plurality of movements or gestures .

24. Humanoid robot (R) comprising:

- at least one embedded processor;
- a sensor assembly (c1, c2) operatively connected to said or at least one said processor and comprising at least one sound sensor and at least one image or movement sensor, to acquire respective input signals;
- a speech synthesis module driven by said or at least one said processor to utter words or sentences, and
- a set of actuators (A1, A2, A3) driven by said or at least one said processor enabling said robot to perform a plurality of movements or gestures;

characterized in that said or at least one said processor is programmed or configured to carry out a method according to any one of claims 1 to 22.

25. Humanoid robot according to claim 24, further comprising a device for connection to at least one remote server, said or at least one said processor being programmed or configured to cooperate with said or at least one said remote server to carry out a method according to any one of claims 1 to 22.