Humanoid Robots that Behave, Speak, and Think like Humans. A Turing thinking machine.

Alan Rosen
Machine Consciousness Inc.
1917 Clark Lane
Redondo Beach, CA 90278 USA

David B. Rosen
Machine Consciousness Inc.
1917 Clark Lane
Redondo Beach, CA 90278 USA

Abstract An introduction to a radical new approach to programming human-like levels of Artificial Intelligence (AI) into a humanoid robot equipped with a verbal-phoneme sound generator. Behavioral programming techniques may be employed to achieve human-like levels of subjective AI, by use of a Relational Robotic Controller (RRC), Relational Correlation Sequencer (RCS)-module, and a centralized self-location and identification coordinate frame. The existence of the self-coordinate frame and the process of behaviorally programming all the intelligence of the system with respect to the self-coordinate frame, may give rise to a robotic system programmed with subjective experiences that it may retain in its memory system. A multi-tasking RRC-Humanoid Robot equipped with a verbal phoneme sound generator, may be behaviorally programmed to achieve human-like high I.Q. levels of subjective AI for the declarative-verbal words and sentences heard by the robot. This robot meets all the requirements of a Turing 'thinking' machine.

Keywords: human-like artificial intelligence; humanoid robots; thinking machines; behavioral-programming; experiential programming; behavioral speech processing.

Acknowledgment: The authors are grateful to MCon Inc. for allowing us to publish their proprietary data, which was accumulated during the past 20 years. This paper presents an overview/introduction to the design of smart RRC-Humanoid Robots. The reader may obtain technical details relating to the design in the references included in this paper.

1 Introduction

The design of 'thinking computers' has been a goal of the discipline of Artificial Intelligence (AI) since the advent of digital computers. In 1950, Alan Turing, arguably, the founder of AI, posed the question "when is a machine thinking?" His approach to an answer was in terms of the behavior of the machine[1] [2]. He devised a I.Q. 'Turing test' based on the conversational behavior of the machine; and deemed any machine that passed the I.Q.-test to be a thinking machine [2] [3].

We, in this paper, follow Alan Turing and describe a building path for a machine that can reach human-like, high I.Q. levels of AI, defined in terms of the behavior of the machine. But instead of programming the computer (robotic controller) with AI, we first program a 'robotic self' into the system, that identifies the robotic system, and then program, experientially, all the AI that the robot gains with respect to, or into the robotic self coordinate frame of the system. So that it is the robotic self that develops a high IQ-level of intelligence, NOT the objective-mechanical digital computer system.

We have thereby designed a system, called a Relational Robotic Controller (RRC)-system that has a subjective identity and AI-knowledge associated with that identity. It is the 'robotic self,' programmed into the computer, that 'thinks,' not the objective-mechanical digital computer.

An overview of a behavioral programming approach to the design-development of humanoid robots has been described in a previous paper [4]. The auditory RRC-Humanoid Robot is a human-like robotic system, controlled by a proprietary Relational Robotic Controller (RRC) [5] [6] [7] [8].

1.1 But First, a Note About Human-like Levels of AI

All programmable digital computers do not have a "self identity" as a human does, that could absorb and convert all data into subjective knowledge, knowledge absorbed relative to the "self" of the machine. Therefore, the computers do not have human-like intelligence. Computers have machine-like intelligence, NOT human-like intelligence.

Machine-like intelligence may refer to the objective knowledge programmed into all modern day computing devices. Human-like intelligence is obtained relative to the "self" of the machine. Human-like intelligence is called subjective knowledge. The following are six pre-requisites required to achieve human-like levels of AI:

1.1.1 The human brain relates, correlates, prioritizes and remembers sensory input data. Similarly, to achieve human-like intelligence, relating, correlating, prioritizing and remembering input patterns must be the essential analysis tool of the robotic controller. The RRC, a proprietary robotic controller of MCon Inc, was specifically designed to emulate the operation of the human brain. It also was designed to operate with a 'self' circuit that is the central hub of intelligence for the whole robotic system.

1.1.2 Humans have a self-location and identification coordinate frame that is trained from infancy to give the human brain a proprioceptive self-knowledge capability.

Even a baby, with a self-knowledge capability, instinctively knows the location of every surface point on its body, the location of its flailing limbs, and by extension, the location of every coordinate frame point in the near space defined by its flailing limbs. The fundamental design characteristic of the RRC-Humanoid Robot is a centralized hub of intelligence, a proprietary module that is the centralized "self location and identification" coordinate frame of the system. This module gives the RRC-Humanoid Robot a robotic form of proprioceptive knowledge, similar to human proprioceptive intelligence. In the RRC-Robot, the self-knowledge capability is the basis for all knowledge.

1.1.3 Human intelligence is experiential intelligence. Humans learn from, and remember their experiences throughout their lifetime. A behaviorally programmed RRC-Humanoid robot emulates the experiential intelligence of a human.

1.1.4 In order to achieve contextual, or 'self-knowledge' of visual data, auditory data, olfactory data, gustatory data, and vestibular data, all the data obtained from those other human-like sensors must be related and correlated with the self-knowledge, self location and identification coordinate frame.

1.1.5 Human intelligence is gained only from the human-like sensors. In this paper we consider the external sensors: Tactile, visual, auditory, olfactory, gustatory, and vestibular sensors. These sensors provide for the consciousness associated with human 'feeling,' 'seeing,' 'hearing,' 'smelling,' 'tasting,' and 'balancing.' The recording monitors of the RRC-Humanoid Robot emulate the external sensors of humans.

1.1.6 Human-like intelligence may be gained by a human-like RRC-Humanoid Robot.
The mechanical robotic body and associated sensors simulate the human body and the human sensors. The robotic body must be bipedal, standing and walking upright with two arms hands and five fingers per hand free to manipulate objects in the environment. The 6 robotic sensors should be human-like sensors designed to gain the same information as is gained by the human sensors.

2.0 Behavioral Programming and the Development of Human-like AI

Behavioral programming is achieved by training the humanoid robot to control its body, limbs, and verbal-phoneme sound generator on the basis of input data from the six external sensors. It is an experiential supervised programming technique analogous to the autonomous learning of a human. Behavioral programming techniques are employed for all the sensory input signals of the humanoid robot. The most important aspects of behaviorally programming the auditory sensors and the verbal phoneme sound generator are described in the following sections.

2.1 Programming a 'self knowledge ' coordinate frame.
The tactile input signals are used to define the central hub of intelligence, the self-nodal map/coordinate frame, of the humanoid robot. The behavioral programming technique employed for the self location and identification 'self-knowledge' coordinate frame is an itch-scratch methodology, wherein the robot is fully trained and remembers how to a) reach and touch (scratch) all points located on the surface of the robotic body, and all points in the near space surrounding the robotic body, b) to identify and locate all such points, and c) to identify and locate all the "end joint" body parts (ends of fingers, elbow, knee etc.) used to scratch all the itch points. When the level of training reaches the threshold of 'self-knowledge,' the Self Nodal Map Module and associated Task Selector Module (TSM)-memory module will facilitate the robotic identification and recognition of all body parts, and the navigation of all movable parts of the robot towards any and every itch point located on the surface of the robotic body, and all points in the near space surrounding the robotic body.
The totality of the programmed "self location and identification" data, stored in a TSM-memory module, is the basis for the "self-knowledge" level of intelligence. Analogous to the proprioceptive knowledge of a human, a RRC-Robot with a fully programmed 'self-knowledge' capability 'knows,' behaviorally, the location of every surface point of the robotic body, the location of flailing limbs, and by extension, the location of every coordinate frame point in the near space defined by flailing limbs.

2.2 Developing self-knowledge for the visual and auditory sensors.

In the visual and auditory RRC-Humanoid systems, experiential intelligence is obtained by performing behavioral programming on the processed raw data coming from the video visual recording monitor and the auditory recording monitor. The raw data is processed in an Interface Circuit, inputted to the RRC and then behaviorally programmed to reach human-like levels of AI. Behavioral programming reaches the level of experiential human-like intelligence, when the RRC-Humanoid Robot demonstrates behaviorally that it has identified, recognized, visualized and comprehended in the same manner as does a human, the signals coming from the visual sensors, or the auditory sensors. The following sections will describe the processing of the auditory raw data in the Interface Circuit, and the behavioral programming of the processed data within the RRC-Humanoid Robot. On completion of behavioral programming, the RRC-Humanoid Robot demonstrates behaviorally human-like levels of AI for the identification, recognition, visualization or comprehension of the processed raw data. Note: Behavioral-programming of the Auditory RRC-Humanoid Robot generates an operational definition of the 'identification', 'recognition' and ‘comprehension’ levels of AI. Any human need merely verbally ask the RRC-Robot to identify, recognize, comprehend, or visualize any color or 3D-object in the FOV, in order to obtain a response that is indistinguishable from the response of another human.

For example: A person can hold an ‘apple’ in front of the Robot and ask the Robot to “Identify the object I am holding in front of you.” A fully programmed Auditory RRC-Humanoid Robot will respond: “That is an apple.” The robot could also demonstrate comprehension of the image by responding verbally with encyclopedic data relating to the object (e.g. “that is a ‘crab’ apple”).

3.0 The Operation of the auditory RRC-Humanoid Robot.

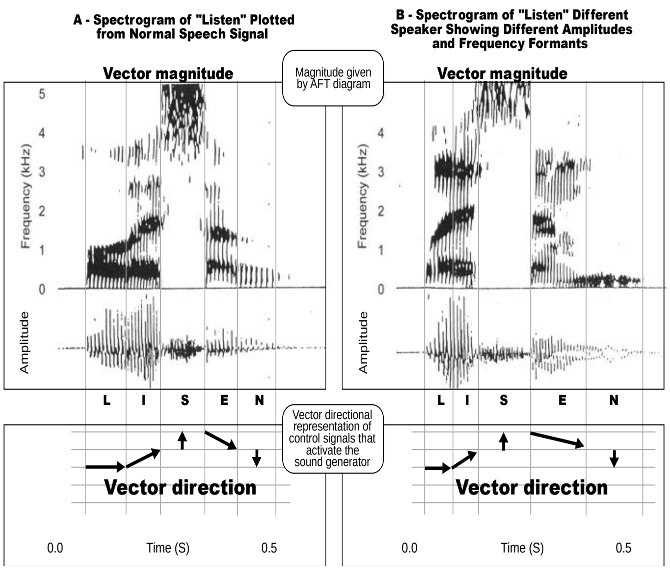

Figure 1 - The magnitudes and directions of a sequence of multi-dimensional p-phoneme vectors representing the word “listen.” The vector direction is shown at the bottom, whereas the functional vector magnitude, the a-f-t-data, is shown at the top. Sections A and B show the differing amplitude and frequency formants for different speakers.

3.1 processing the Auditory Raw Data in the Interface Circuit

The auditory raw data consists of the output of the sound receiving microphones that are sensitive to the auditory frequency range of zero to 20,000 cps (simulating the human ear). When a talking sound is applied to the ear-like receiving microphones, they generate a sequence of frequencies and amplitudes as a function of time. These frequencies and amplitudes may be illustrated in an amplitude-frequency-time (a-f-t)-diagram shown at the top of Figure 1. This (a-f-t) data is inputted to the Interface Circuit.

The Interface Circuit stage performs the following functions: a) It processes the a-f-t data into collective modalities that have been selected to be characteristic of 120 phoneme-sound combinations present in the English language. b) With the aid of a spectrum analyzer, the collective modalities are tuned to the selected phoneme sound combinations in the english language, to musical sounds or to environmental noise. c) It performs behavioral speech processing on sequential sets of phoneme sounds and identifies and recognizes these sequential sets as “ words” or “sentences” formed by the person speaking to the microphones of the auditory recording monitor. (Note: behavioral speech processing starts in the Interface Circuit stage by training the Robot to “repeat the heard words” (training task T-201 in Figure 3). The speech processing described in section 4.0 includes a capability to recognize, identify and correct, incorrect grammatical structures, d) Finally, the (a-f-t) data must be formatted within the Interface Circuit so that it is compatible with the RRC-Phoneme Sound Generator and the input to the RRC-Multi-dimensional Nodal Map Module [5] [6] [7] [8].

The final output of the Interface Circuit stage, the ‘words and sentences,’ music or environmental noise, are inputted to their respective multi-dimensional Nodal Map Modules of the RRC-Humanoid Robot, for behavioral programming. The operation of the RRC-circuits is described in Sections 3.2 and 3.3. The operation of the memory system of the RRC robot is described in Section 3.4. And Section 5.0 describes the behavioral programming of an RRC-robot that ‘knows’ and understands what it ‘hears’, and demonstrates that knowledge verbally.

3.2 The operation of the RRC-circuits that perform identification, recognition, and comprehension of verbal-phoneme words and sentences.

3.2.1 Introduction to the RRC-Robot: The RRC-controller is made up of the array of Relational Correlation Sequencer (RCS) [5], [6], [7], [8]. Control, computation and data storage are performed by relating and correlating each and every input data signal (tactile, visual, or auditory signals) with each other; and relating and correlating each and every input data signal with each and every output-control signal generated by the RRC-controller.

When the input sensory signals are related to the self location and identification Nodal Map Module, the sensory signals are said to be “internalized” to the robotic “self” of the system. When the data from the other sensors is internalized the system gains a contextual form of self location and identification knowledge for those other sensors. Multiple relations and correlations are required in order to achieve higher levels of AI-identification, AI-recognition, AI-visualization and AI-comprehension of all input data signals. It is the internalization of all the input data that allows the RRC-robot to identify, recognize, visualize and comprehend the input signals and patterns.

3.2.2 The Operation of the RRC: The auditory search engine is used to search the multi-dimensional auditory Nodal Map Module of the RRC-robot for a-f-t-sound patterns that will be recognized by the RRC as Task Initiating Triggers (TIT). The TIT-patterns are used to activate any of the tasks listed in the Hierarchical Task

Diagram (HTD) shown in Figure 2. Each of the prime level tasks, shown in Figure 2, has a programmed Task Selector Module (TSM) associated with it. The dominant electronic component of each TSM is a pattern recognition circuit that is programmed to recognize and prioritize the TIT-pattern detected by each of the TSMs as they operate on each of the input Nodal Map Modules. The total collective of TSMs shown in Figure 2, form the declarative memory system of the auditory RRC-Humanoid Robot. The programming/training of the auditory RRC-robot is a process of training the pattern recognition circuits of each TSM associated with each task, to recognize, identify and prioritize input-signal TIT patterns that initiate the lower level tasks shown in the Figure.

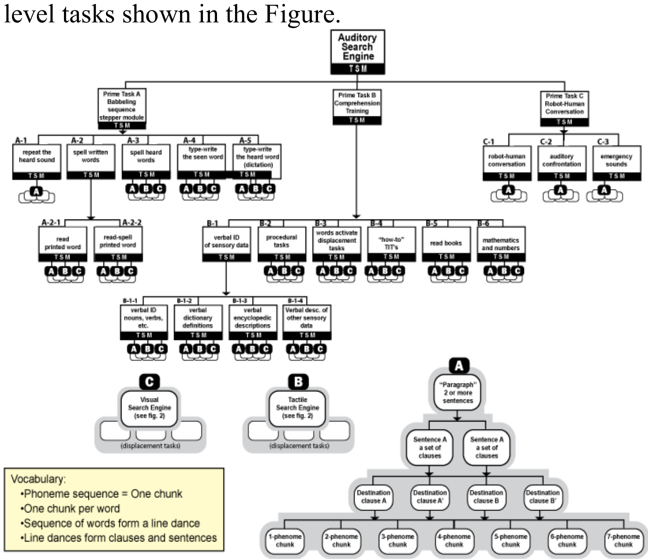

Figure 2 - The declarative HTD: The HTD is the top-level specification of the system. This Figure shows the TSMs of the auditory search engine. Those TSMs form a declarative memory system within the RRC.

3.3 The Search Engine Mode of Operation

The operation of a multi-tasking RRC-robot has been described in the literature [5], [6], [7], [8]. The following presents some aspects of the operation of an RRC-Robot that are applicable to the behavioral programming of the system. An RRC-Humanoid Robot operates by searching the environment for tactile signals or patterns, visual signals or patterns or auditory signals or patterns. The declarative auditory search engine, shown in the Hierarchical Task Diagram (HTD) of Figure 2**,** operates concurrently to guide the robot in the performance of the verbal tasks listed in the Figure. The declarative auditory search engine, shown at the top of the figure, searches the sound environment for phoneme based sounds consisting of words and sentences described by the a-f-t signals shown in Figure 1. The Task Selector Module (TSM), associated with each task and subtask, is shown as a black bar at the bottom of each task or subtask shown in Figure 2. Each TSM consists of pattern recognitions circuits that are programmed to activate only one TIT-task of the number of TIT-tasks that are directly below it. For example, the TSM of the top most Auditory search engine may generate TIT-a-f-t-words or sentence sound patterns that activate prime task A, or TIT-a-f-t-words that activate prime task B, or TIT-a-f-t-words that activate prime task C. For example, if the TSM of the auditory search engine, recognizes the a-f-t-pattern associated with the TIT-words "repeat the heard sound," then the TIT-signal is transmitted to prime task A-1 for further processing. Note that after completion of training, if the robot is confronted with verbal phoneme sounds, but cannot recognize either the speaker or the words, then the signal is transmitted to prime task C-2, (the auditory confrontation mode)). Note that TSMs are associated with each level of the HTD shown in Figure 2. Each TSM, at each level is activated by the TSM at a level above it. Each TSM consists of pattern recognition circuits that recognize TIT-words and sentences, and thereby transmit them to another TSM-level either for further processing or for activation of the verbal phoneme sound generator or robotic body or limbs. The steps for training/programming of the TSMs of the declarative memory system are shown in Figure 3. Note that the programmed/trained TSMs associated with the tasks in Figure 3, form a declarative memory system within the controller that remembers the procedure for performing the various subtasks listed in the HTD.

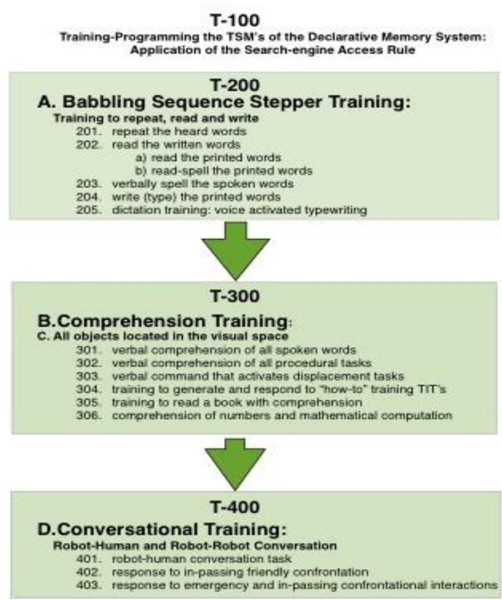

Figure 3 Training-programming the TSMs of the declarative memory system.

3.4 The memory systems within the RRC.

Learning and memory is generally classified as procedural (or reflexive) if the learning or memory involves motor skills, and declarative if the learning or memory involves verbal skills. In the multi-tasking RRC-robot, procedural TITs operate in the motor-joint Nodal Map Module, and procedural memory is the basis for all the control functions of the somatic motor system. In the Auditory RRC-Humanoid robot declarative memory is the basis for all the control functions of the verbal-phoneme sound generator. Figure 2 shows the TSMs associated with the declarative memory system. The programmed/trained TSMs shown in Figure 2, give the robot the capability to "remember how" to perform all the auditory sub-tasks listed in the Figure. The declarative memory system includes a robotic capability to a) repeat, read and write all words and sentences presented to the robot, b) comprehend and identify and describe verbally all nouns, adjectives, verbs and adverbs that are presented to the robotic visual and tactile systems, and c) perform robot-human conversation with comprehension.

4.0 Speech Processing: Recognizing the Acoustic Sequential Set of Phoneme-signals as Phonetic Words and Sentences.

4.1 The Problem:

The problem of converting the perceived acoustic spectrographic (a-f-t) properties of language into an identifiable phonetic structure is an ill posed problem, similar to the 3-dimensional inverse optics problem [9]. There is not a simple one to one mapping between the acoustic properties of speech and the phonetic structure of an utterance. Co-articulation (the segmentation problem) is generally identified as the major source of the problem. Co-articulation gives rise to difficulty in dividing the acoustic signal into discrete "chunks" that correspond to individual phonetic segments. And it also gives rise to a lack of invariance in the acoustic signal associated with any given phonetic segment. Note: The usual methods for solving the problem includes lexical segmentation processing (co-articulation), word recognition processing, context effect processing, syntactic effects on lexical access processing, lexical information and sentence processing, syntactic processing, and intonation-structure processing.

4.2 The Behavioral Speech Processing Methodology for Solving the Inverse Auditory Problem.

Because of the complexity in the mapping between the acoustic signal and phonetic structure, an experiential, behavioral programming methodology was developed for 'unpacking' the highly encoded, context dependent speech signals. 'Unpacking' is performed in the Interface Circuit by programming the RRC to repeat and 'remember' (in the TSM-memory modules) the 'heard' words and sentences of multiple speakers.

4.2.1 Repetition and Babbling the words and sentences taken from a 50,000 word Lexicon (Task 201 in Figure 3). The first step for training the auditory RRC-robot is the requirement for a "babbling" Sequence Stepper Module and an associated TSM that is trained to accurately and quickly repeat the sound of words, strings of words, or sentences heard by the robot. The trained repetition and babbling sub-task A-1 TSM, activates the total vocabulary of the robot. All the words or sentences spoken by the robot and activated by other prime task TSMs must access the sub-task A-1 TSM and form a compound TSM that does not necessarily repeat the sound but accurately enunciates other words and sentences (taken from the sub-task A-1 TSM) and associated with the compound TSM. Most of the design activities of the Task T-201 shown in Figure 3, are aimed at achieving enunciation accuracy in the repetition and babbling sub-task A-1 TSM. In order to achieve repetition accuracy it is necessary to refine the design of the phoneme sound generator, expand the number of phoneme sounds listed in the 120 phoneme sound combinations utilized in the preferred embodiment RRC-Humanoid Robot, and refine the tuning of the spectrum analyzer to the actual collective modalities present in the English language verbal input signal.

4.2.2 Additional speech processing: Further 'unpacking' is performed by behavioral programming techniques that includes the following: First, by relating, correlating, associating and calibrating the heard verbal speech with the corresponding visual and tactile data obtained in the visual and tactile coordinate frames in which the robot is operating. Next, by training the RRC-Robot to be sensitive to such factors as acoustic phonetic context, speaker's 'body language,' speaking rates, loudness and 'emotion laden' intonations. The Auditory RRC-Humanoid Robot takes into account the acoustic consequences of such variations when mapping the acoustic signal onto the phonetic structure. The problems of speaker's 'body language,' 'emotion laden' intonations, acoustic phonetic context, speaking rates, and loudness is solved in the Auditory RRC by coordinating the search engines of the visual and tactile systems with the search engine of the Auditory RRC-Humanoid Robot.

5.0 Behavioral Programming of the (speech) Processed Words and Sentences

Once the words and sentences are recognized as a TIT (in the TSM) and are projected to the Multi-dimensional RRC-Nodal Map Module, the RRC-Humanoid Robot is behaviorally programmed to control body and limbs, and the verbal phoneme sound generator in relation to the words and sentences applied to the Multi-dimensional Nodal Map Module. Words and sentences are identified, recognized, and comprehended by behaviorally programming the RRC-Robot to generate different words by the verbal phoneme sound generator or undertaking distinguishing body or limb actions based on the words or sentences applied to the Multi-dimensional RRC-Nodal Map Module.

5.1 The programming of the auditory RRC-robot

Programming/training the RRC-robot is a process of training the pattern recognition circuits of each TSM associated with each prime level task and all the TSMs associated with the sub-tasks listed under the prime level task (See Figure 2). The pattern recognition circuits must recognize, identify and prioritize input-signal TIT patterns that initiate the prime level task and all the lower priority TIT-sub-tasks that are listed under the prime level task. The programmed TSMs associated with all the tasks in Figures 2 and 3 give rise to a declarative memory system within the controller. Training the declarative memory system of the auditory RRC-robot is presented in the following sections.

5.2 The search engine access rule for training the declarative TSMs

The declarative memory system of the RRC-robot is made up of an array of TSMs with each TSM storing a large number of a-f-t-words phrases and sentences that represent the total vocabulary of the robot. In order to respond verbally with appropriate words and sentences the RRC must analyze the verbal input data, search through the memory TSMs, find the set of TSMs that have parts of the answer stored in them, form a compound TSM that has the total explicit word-answer stored in it, and activate the appropriate word answer that is stored in that compound TSM. The following programming rules have been devised in order to facilitate the search for an appropriate response to any auditory input signal.

1. Search the input signal to determine which TSMs are likely repositories of the appropriate verbal response.
2. Form a compound set of TSMs wherein the response may be stored.
3. Utilize the data present in the auditory input signal and in the compound set of TSMs to home in on an appropriate response.

For example, the application of the auditory search engine access rule when the trainer-supervisor requests the Robot to "identify this visual image," leads to an identification of two TSMs and a compound TSM. The two TSMs are most likely the visual image pattern-TIT presented to the robot (Task T-301 in the B-1 TSM), and the repeat this sound - verbal word or phrase that describes the presented visual image (Task T-201 in the A1 TSM). Note that the sub-task A-1 TSM stores and properly enunciates all the nouns, adjectives, verbs and adverbs taken from the 50,000 word lexicon. The compound TSM is formed in the programmatic development of the access rule and includes the phrase "I see an ----," wherein the training selects that word or phrase from the A-1 TSM that describes the presented visual image.

5.3 Behavioral programming procedures

Behavioral programming procedures are performed on all the TSMs shown in Figure 2. Examples of behavioral programming procedures for 3-TSMs will be described in the following 3-subsections.

5.3.1 Sub-Task A-1: TSM-Training to Repeat Phoneme-Sounds Spoken by the Trainer-Supervisor (sub-task T-201 Figure 3). The trained prime task A-TSM is a memory module that stores all the TITs that identify and properly enunciates all the words listed in the lexicon and the commonly used combinations of words, clauses and sentences selected by the trainer-supervisor. All subsequent verbally generated tasks must access the TITs stored in the prime task-A TSM. The properly enunciated words and phrases taken from the Task A TSM, are then associated with other TITs generated by the visual system, the tactile system, the olfactory system, the gustatory system, or other word TITs generated by the auditory system.

- Repetition and Babbling - The auditory RRC is trained to repeat via the phoneme sound generator, the words and sentences spoken by the trainer-supervisor.
- The high priority TIT that shifts the robot to Sub-Task A-1 is a simultaneous visual recognition image of the trainer, and the command "repeat this sound" spoken by the trainer.
- All the words and sentences repeated by the robot are taken from a 50,000 word lexicon that represents the total vocabulary of the robot.
- The lexicon or vocabulary of the robotic controller consists of the set of words and sentences that the sub-task A-1 TSM has been trained to repeat.
- The trained sub-task A-1 TSM is a memory system that properly enunciates all the words and sentences listed in the lexicon.
- Optimization of the sub-task A-1 TSM to properly enunciate all the words and sentences listed in the lexicon entails a) refining the design of the phoneme sound generator to assure that the lexical segmentation and timing intervals between successive phonemes are optimized. b) expanding the selected number of phoneme sounds to optimize coarticulation problems. And c) refining the tuning of the spectrum analyzer to the actual collective modalities present in the English language verbal input signal.
- The sub-task A-1 TSM memory system is always accessed by other TSMs in order to form compound TITs whenever verbal sounds other than the "repeat this sound" TIT are to be generated by the robot.
- Queries that access the sub-task A-1 TSM generally relate to the verbal enunciation of the words and phrases stored in the A-1 TSM. Therefore, in anticipation of such questions, acceptable and not acceptable grammatical structural forms of verbal enunciation should be programmed into the search engine for each word or phrase in the lexicon; and the specific answer to each anticipated query must be programmed into a compound TSM.

At this point the controller has performed all the speech processing that allows it to recognize and repeat, but not comprehend, all the phoneme constituents of words, sentences and clauses listed in the lexicon. The auditory

RRC-monitor has thereby mapped the acoustic signal onto a linguistic representation that is amenable to declarative comprehension (in prime tasks B and C)

5.3.2 Robotic comprehension: TSM-Programming-learning to comprehend the meaning of a "heard" word (Subtask B-1 in Figure 2). Robotic comprehension is the simultaneous identification and coupling of each word or sentence heard, read, spoken or written by the robot, with the visual image, tactile data, olfactory data, gustatory data and lexical definition that is associated with that word or sentence.

In order to comprehend the meaning of a word or a sequence of word-TITs, it is necessary to use a conditioning learning technique that associates the sequence of words-TITs taken from the lexicon recorded in the prime task A-TSM with the visual TIT, tactile TITs, or other word-sequence TITs that relate to the sequence of words-TITs. For example, a visual image TIT may serve as a visual descriptor by associating or conditioning that image with a q-phoneme sequence TIT that is the "word definition" or "descriptor" of the visually seen image. Thus the visual image in combination with a word TIT may serve as a compound TSM that generates a TIT for the generation of one or more "words" that is descriptive of the image. And hearing the one or more words may serve to generate an "association" with the visual image, which is defined to be comprehension of the meaning of the one or more words.

In order to achieve a high level of verbal comprehension, the training process requires that the trainer-supervisor repeatedly display the sight, smell, feel-touch, verbal definition of the object, and possibly taste of the object that is to be comprehended. For high levels of comprehension of an object such as an apple, the search engine operates on pattern recognition circuits associated with the auditory, visual, tactile, olfactory and taste sensors in order to associate the word apple with the visual image of the apple the verbal definition of an apple, the "feel" of the apple, the "smell" of the apple, or even the "taste" of the apple.

5.3.3 Sub-Task C-1 Conversational Mode (listed as task T-401 in Figure 3). Conversational constraints: Any statement or question posed to the RRC-robot requires that the RRC-controller search through its TSM-declarative memory systems for an appropriate reply. In general the RRC is trained to utilize the search engine access rule described in section 5.2, in order to facilitate the search for an appropriate response to any auditory input signal. Immediately after the RRC-robot is placed in a conversational mode it is necessary that the robot apply the access rule on any verbal input that represents a question or statement made to the RRC-robot. The access rule applied to the query should abstract and point towards a TSM or compound TSM where the reply to the input query may be stored. A sampling of three of the eight most important TSMs available to form compound TSMs for the sub-task C-1 conversational mode are the following: a) The B-1-1 TSM stores sets of TITs of verbal (descriptors) identification of visual images for all nouns, verbs, adjectives and adverbs listed in the 50,000 word lexicon. b) The B-1-2 TSM stores sets of TITs of verbal dictionary definitions of all visual images and the nouns, verbs, adjectives and adverbs listed in the 50,000 word lexicon that are associated with them. c) The B-1-3 TSM stores sets of TITs of verbal encyclopedic descriptions, of all visual images and nouns, verbs, adjectives and adverbs listed in the 50,000 word lexicon that are associated with them.

Programming the access rule into the queries is a process that begins with the programming of the prime task B TSMs themselves. In the design of all the task B TSM (comprehension tasks) all possible queries associated with sub-task C-1 must be anticipated whenever a verbal sentence is programmed into the task B TSM. For each of the possible queries one or more of the array of task B TSMs must have an appropriate and specific answer programmed into it.

The programming technique (and the manpower required) is similar to the programming of some of the word based search engines such as Google or Bing.

- The high priority TIT that shifts the robot to a conversational mode, the sub-task C-1 mode, is a simultaneous visually recognized image of a authorized person, and the spoken words by the person "stop-lets talk" or "lets talk."
- The "stop-lets talk command causes the robot to interrupt the procedural task it is doing and devote itself to conversation, whereas the "lets talk" command allows the robot to continue the procedural task and converse during the ongoing procedural task.
- The sub-task C-1 TSM is trained to utilize the search engine access rule to form a set of compound TSMs and search for the particular TSM that has the proper response programmed into it. (a sampling of verbal queries that point to over 2 dozen task and sub-task TSMs have been identified)
- The generated verbal speech exhibits "comprehension" of the input sound signal by training the response so it is based on the TITs programmed into the task B TSMs that exhibit comprehension (relations to visual-auditory-tactile-olfactory-gustatory data) of the words and sentences stored therein.

Mcon Inc. has developed a detailed overview of the programming requirements for the comprehension and conversational capabilities of the RRC-Robot. The magnitude of the development task and the manpower requirements exceeds the task of developing a global search engine, such as Google. However, the same technology used in the development of a gigantic global search engine, is required in the development of the auditory search engine of the RRC-Humanoid Robot.

Conclusion.

The RRC-Humanoid Robot is a subjective machine. It has a subjective identity and AI-knowledge programmed into that identity. It is the 'robotic self' programmed into the computer that develops high IQ-levels of human-like AI.

The RRC-Humanoid Robot has the same external sensors as a human. The robot can have subjective experiences such as feel, see, hear, smell and taste (but not eat), just like a human, because behaviorally and verbally it responds to tactile stimuli, visual stimuli, auditory stimuli, olfactory stimuli, and gustatory stimuli, just like a human. It can hear and understand verbal speech, and respond intelligently to any query or statement directed at it. The Robot is said to be able to 'think' because it demonstrates behaviorally and verbally a high IQ-level of AI (it passes the 'Turing thinking test'). The robot can be trained-programmed to perform any task that a human may perform.

In addition, if the robotic body of the RRC-Robot has the degrees of freedom, flexibility, and strength of the human body, since the RRC is trained with self knowledge (proprioceptive knowledge) and self awareness knowledge (real time tactile feedback from all joints, all limbs, and all parts of the robotic body), a RRC-Humanoid Robot may be trained-programmed to achieve the manual dexterity, and foot-motion dexterity of humans. That is, they may learn to walk, run or dance gracefully, and depending on the design of the robotic body, they may play piano or violin, do acrobatics, or play basketball, just like a human.

In addition, the physical and intellectual similarity of RRC-Robots to humans, and the capability of robots to converse intelligently with humans, will cause a large number of people to develop a special friendship and empathy towards servant or companion robots. And many would wish to acquire such robots as companions, friends or pets.

This is a highly significant advance in the development of humanoid robots. The authors believe that humanoid robots may be used as the preferred labor force of a modern society: servants in the home, companions and assistance to the sick and elderly, assistance and helpers in the workplace, professionals in business and industry, and assistance and helpers in government projects from DOD to NASA.

The authors believe that the design development and manufacture of human-like RRC Humanoid Robots may become the largest industry of the 21st century.

Therefore, MCon Inc. is launching a program aimed at the commercial design, development and manufacture of highly intelligent RRC-Humanoid Robots. The first phase is a 3-part licensing and joint venture phase, as follows.

1. Publicize and publish the development in the scientific literature, the popular press, newspapers, and the internet (generate publications and art work)
2. Contact, make presentations to potential joint venture partners (i.e. Google, Microsoft, Honda, Sony, Government labs etc.)
3. License and consult with any potential developer for the support of the programming, mechanical body, and sensor development of the RRC-Humanoid Robots

We end this paper with a challenge to industry, (e.g. Google, Microsoft, Honda, Sony, the Japanese government, the US Government, etc.) to commercially develop a large labor force of smart RRC-Humanoid Robots.

We need the support of the RAS and all scientists and engineers who may contribute to this development.

References:

[1] Turing AM (1950), Computing Machinery and Intelligence, Mind 59: 433-460

[2] Turing AM (1953), Solvable and Unsolvable Problems, Science News 31: 7-23

[3] Rosen A. & D.B. Rosen (2007), The Turing Machine Revisited: ASSC-e archive (www.mcon.org)

[4] Rosen A. & D.B. Rosen (2013), A Behavioral Prtogramming Approach to the Design-Development of Humanoid Robots. Presented at ICAI in las Vegas July 22-25, and published in the proceedings. Available for viewing at www.mcon.org.

[5] Rosen A. & D.B. Rosen (2003), The Design of a Volitional, Obstacle avoiding Robot, Machine Consciousness Journal 1: 41-56 (available for viewing at www.mcon.org)

[6] Rosen A. & D.B. Rosen (2006a), The Design of a Sensation Generating Mechanism, ASSC e-archive (available for viewing at www. mcon.org)

[7] Rosen A. & D.B.Rosen (2006b), A Neural Network Model of the Connectivity of the Biological Somatic Sensors, Published in IEEE –Xplore Proceedings of the IEEE-WCCI July 26-21 2006, Canada.

[8] Rosen A. & D.B. Rosen (2006c), An Electromechanical Neural Network Robotic Model of the Human Body and Brain, LNCS Part 1: 105-116; Proceedings of ICONIP-2006; Berlin: Springer-Verlag

[9] Marr D. (1962), 'Vision' Freeman Press