A Human-Robot Sub-Dialogues Structure Using XML Document Object Model

Majid A. Al-Taee Suhail N. Abood Nada Y. Philip
School of Computing and Information Systems
Kingston University
London, United Kingdom
e-mail: altaeem@kingston.ac.uk, suhail.abood@gmail.com, n.philip@kingston.ac.uk

Abstract—**Improving remote interactivity between patients and their health carers is a key requirement for emerging eHealth systems. This paper presents a new human-robot sub-dialogues structure based on the XML Document Object Model (DOM). The dialogue structure which aims at improving patient-carer interactivity enables both the patients and their carers to control behavior of humanoid robot locally through verbal instructions or tactile sensors as well as remotely through textual Web-based instructions. The textual instructions which are prepared in the form of interactive dialogues are converted to XML DOM and sent to the in-home robot via the Internet. The robot will then verbalize these dialogues and present them to the patient along with appropriate gesture/posture attributes. The XML DOM standard is adopted in this study to facilitate the structure and navigation flexibility between the internal nodes of the dialogue. Numerous human-robot interactions based on the proposed dialogue structure have been designed, built and successfully performed by the robot using both local and remote instructions. These dialogues are tailored to support diabetes self-management in children. The prototype evaluation results and observations showed (i) a seamless and accurate data transfer between the robot and a remote Web-based health portal, and (ii) a high level of navigation flexibility between different nodes of the experimented dialogues.**

Keywords-human-robot interaction; remote control; remote data collection; sub-dialogues structure; XML DOM.

I. INTRODUCTION

Human-Robot Interaction (HRI) is defined as the interdisciplinary field that addresses understanding, designing and evaluating robotic systems for use by or with humans [1]. HRI differs from Human-Computer Interaction (HCI) in several dimensions such as levels of interaction, the environment in which the interaction takes place, and physical and dynamic nature of robots [2]. Communication methods between robots and their users can take numerous forms; of these, verbal, tactile, and sensory are the most common. These communications are translated into actions that aid the robot to accomplish its tasks. To design a human-robot communication interface, the user must be provided with two things; a way to instruct the robot and feedback about what is currently happening on the robot [3]. Feedback can be considered a vital part of human-robot interaction; it reduces disorientation and confusion of users [4].

Human-robot interactions can fall into different categories [5] such as supervisor, operator, mechanic, bystander, and peer. Of these, the latter category, which aims to build and maintain a long-term relationship between the robot and the human, is of a particular interest in this work. A recent review [6] showed increased usage of robots in variety of healthcare applications. Another study showed that emotion expressions yield increased enjoyment of the HRI [7]. A comparison conducted between physical and virtual robots showed that people prefer to work with physical robots over virtual ones [8]. A research on attitudes toward the use of robots in healthcare [9] showed that many people accepted the use of robots in healthcare but some of them raised concerns about reliability, safety, and loss of personal care. Of these, the latter concern is the most challenging due to the fact that medical conditions vary from one patient to another and patients may require personalised treatment/care delivery. This questions the use of robots in behaviour change support from unhealthy to healthy lifestyles for improved self-management of chronic illnesses due to the variations between patients' medical conditions, attitudes, and adherence.

In the case of children, previous studies [10], [11] reported that children exhibited a variety of social and care-taking behaviours toward the robot; they treated it more as a peer than as a toy after spending enough time with the robots. Furthermore, it was reported that human-robot interactivity can also be much improved by using flexible rather than fixed structured dialogues [12]. The main challenge for such relationship/interaction is the patient's gradual loss of interest in the interaction. A rich robot-child communication is reported to be the result of robot evaluation of children's work [13] but not the converse. Benefits of using robot dialogues to help people achieve better lifestyles and physical activity management are reported in [14].

In the context of work reported in this paper, the robot needs to be able to encourage diabetic children toward treatment adherence and express some emotions such as dissatisfaction for bad disease management outcome (e.g. poor blood glucose control), unhealthy lifestyle in terms of diet and physical activity. This can only be achieved through building an effective dialogue structure to support navigation flexibility of various robot-patient dialogues. The work presented in this paper focuses on the development of a new sub-dialogue structure for robot interactions using XML Document Object Model (DOM) [15]. The proposed model exploits a speech recognition engine and a text-to-speech engine existing in a humanoid robot called NAO [16]. Unlike previously reported robot

This research was supported by Kingston University London and a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme, under grant No. 275571.

978-1-4799-5264-9/14 $31.00 © 2014 IEEE
DOI 10.1109/DeSE.2013.29

dialogues, the proposed dialogue structure is based on a well-known standard which allows for building various types of dialogues that can be programmatically read, manipulated, modified, and exchanged over a distance through the XML DOM.

The remainder of this paper is structured as follows. In Section II, the e-health system and types of patient-carer interaction are overviewed. Architecture of the proposed system and its key components are presented in Section III. The sub-dialogues structure model is given in Section IV. Capabilities of the developed dialogue system are demonstrated and in Section V. The paper is then concluded Section VI.

II. SYSTEM OVERVIEW

The proposed human-robot interaction module plays a key role in the e-health system which supports virtual connectivity between diabetic patients and their health care providers. This interaction, which is performed by a humanoid robot on behalf and under control of the health care team, focuses on two important pillars; patients' empowerment and behavioural change support. Applications of this module are distributed between an existing health portal reported in [17] - [19] and a local patient's robot. The remote server applications are linked to the robot through a secured HTTP protocol, as shown in Fig. 1. This e-health system which was proposed to maintain asynchronous interactivity between diabetic patients and their health carers was developed using rapid prototyping methodology [20].

In active-care systems, health carers require continuous data collection and monitoring of the patient's health problems, tests, and treatments over time. In contrast, patients require empowerment and feedback on disease development at each stage in the process of self-management for the health problem at hand. This type of patient-carer interaction can take different ways:

a) **Carer to patient interactions** – built by the carer using a flexible graphical interface developed by the authors. These interactions include treatment plan setup/adjustment, feedback and alerts, and patient's educations.

b) **Patient to carer interactions** – represent the patient's response to the dialogues assigned by his/her carers. These interaction*s* are performed by the patient's robot along with appropriate emotions. This kind of interactions collects patient's measurements, medications, diet, physical activity, depression level and views.

In robot communication interfaces, users are mainly concerned with two things; robot instruction method and robot feedback [3]. The robot instruction method is the mean by which the users instructs the robot or modify its behaviour to complete the interaction. For the purpose of work reported in this paper, the instruction method is twofold: Web-based textual instructions created remotely by health carers and local verbal instructions performed by the patient.

III. ARCHITECTURE OF THE PROPSED SYSTEM

The proposed system is composed of a behaviour manager, network manager, database manager, data collection manager, interaction composer, instruction interpreter, and interaction performer, as shown in Fig. 2. The roles performed by each of these applications are outlined as follows:

Behaviour manager – controls and coordinates communications and dataflow between all sub-modules of the system, as illustrated.

Network manager – handles the synchronization process with the server-side database.

Database manager – manages a local SQLite database, including data retrieval, insertion, update, and deletion functionality on different tables.

Data collection manager – handles wireless collection of measurements from medical sensors.

Interaction composer – generates an XML tree that represents the interaction tree. It uses the data stored in the local database to create an XML tree using the XML DOM.

Instruction interpreter – interprets event instructions and dispatches them for execution.

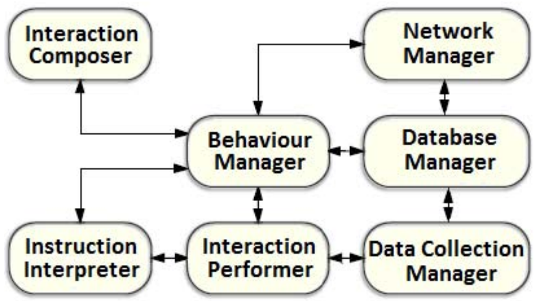


Figure 2. Architecture of the proposed system.

Figure 1. Overview of the e-health system.

Interaction performer – performs XML interaction nodes of the XML document created by the interaction composer. It consists of four main sub-modules namely; interaction dispatcher, text-to-speech, verbal data collection, and gestures/postures presentations. The internal structure of the performer is composed of an interaction dispatcher, verbal data collector, text-to-speech synthesiser, and gestures/postures presenter, shown in Fig. 3. The functions performed by the interaction performer elements are outlined as follows.

a) **Interaction dispatcher** – manages execution of the current interaction node, depending on its type (e.g., information, decision, or data collection), and events through the instruction interpreter (see Fig. 2). It also dispatches the node under execution to one of the performer sub-modules (verbal data collector, text-to-speech synthesiser or gesture/posture presenter).
b) **Verbal data collector** – utilizes the text-to-speech synthesiser and speech recognition engine of robot to handle various types of verbal data collection dialogues on diet, exercise, medications, depression, etc. The robot performs these dialogues in terms of adaptable/extendable multiple-choice questions and collects patient's responses in a timely manner.
c) **Text-to-speech synthesiser** – converts the textual format of the robot-patient dialogues into speech, exploiting the underlying text-to-speech engine of the robot. These dialogues are initially constructed by health care professionals using a user-friendly graphical interface at the remote health portal.
d) **Gesture/posture presenter** – enables the robot to perform numerous emotional responses pre-specified gestures, postures and using head and hands movements and colour patterns on the robot's chest, eyes and ears. These emotional responses, which are performed simultaneously with the verbal dialogues, would enrich the patient-robot interactivity by making it more natural.

IV. SUB-DIALOGUES STRUCTURE

A new platform is proposed to handle all kinds of dialogues to be carried out between the patient and his/her robot. The proposed platform utilises a tree model to represent all kinds of user-robot interaction dialogues in terms of nodes, as shown in the example of Fig. 4. In this model, each of the following nodes, which represent one step in the interaction dialogue, performs one of the following tasks.

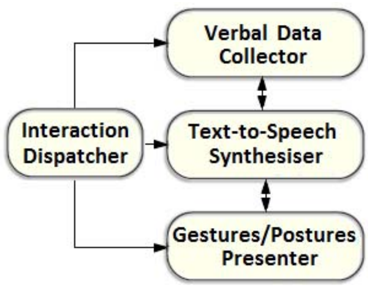


Figure 3. Structure of the interaction performer.

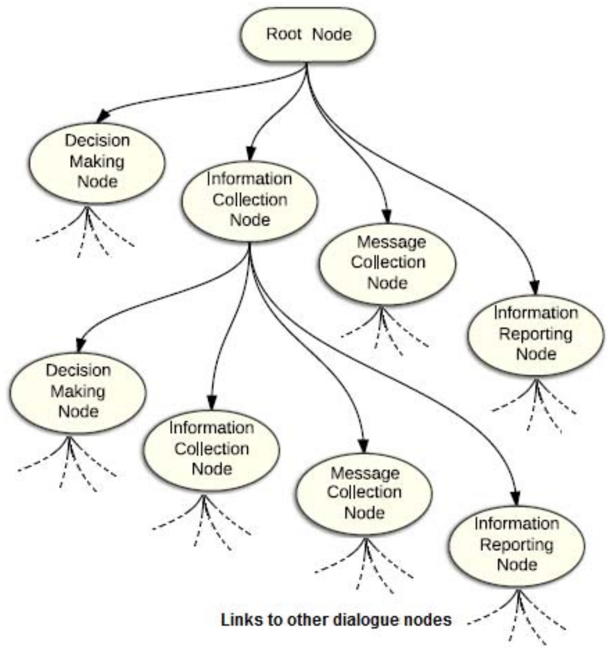


Figure 4. Structure of dialogue nodes.

Decision Making (DM): Makes appropriated navigation decision based on the user response.

Information Reporting (IR): Provides the patient with disease management support information (e.g. feedback on current health status, advices, educational info, etc.).

Information Collection (IC): Collects patient's information (e.g. diet, exercise, wellbeing, etc.).

Message Collection (MC): Collects patient's views in terms of short audio messages.

Execution of these nodes can be repeated by navigating back to the previous node. The node execution can also be accompanied with appropriate robot posture and/or gesture to improve dialogue interactivity with the user. The various attributes of interaction nodes are summarised in Table I.

A. *Instruction set*

In order to use the triggered events to perform an operation on the robot system, the proposed system implements a minified instruction set that provides an interface to either access the main internal features of the robot or to manage the interaction. The instructions can be attached to the events of each node such that when the event is triggered by the performer, the instruction will be executed. Multiple instructions can be grouped together to be executed upon event trigger either in parallel or sequentially. Each instruction consists of a 4-letter keyword followed by a white space and the instruction parameter(s). Parameters can be optional or compulsory depending on the nature of the instruction. For example, the "GOTO" instruction has a compulsory parameter because it affects the interaction execution path. When this parameter is not provided, the instruction will be skipped.

TABLE I ATTRIBUTES OF THE INTERACTION MODEL.

Abbreviations: DM (decision making node), IR (info reporting node), IC (info collection node), MC (message collection node).

Attribute	**Applicability**	**Description**
Category	DM, IR, IC, MC	A category ID to specify which data type this node belongs to. (e.g., 0: physical activity & fitness, 1: diet & food, 2: medications, 3: wellbeing, etc.).
Type	DM, IR, IC, MC	An integer defines the type of node: 0: IR, 1: DM, and 2: IC node.
Content	DM, IR, IC, MC	Textual content to be spoken by the robot. This can be information (or question) to be provided (or asked) to the user.
Choice *(i); i = 1 … n*[(*)]	DM, IC	Textual content used by the robot's speech recognition engine as a reference keyword(s) for the choices of a multiple choice data collection node. This can be a single word or multiple words separated by semicolons.
Weight *(i); i = 1 … n*	IC	A weight value given to each answer of a multiple-choice question.
Back	DM, IC, MC	A verbal instruction used when the user decides to return to a previous node.
Next	DM, IR, IC, MC	This attribute specifies the next destination node in the dialogue execution. It uses a "GOTO" instruction to navigate to the next interaction node.
BeforeExecution, AfterExecution	DM, IR, IC, MC	Instructions executed before or after completion of current interaction node. Usually they contain instruction to provide visual feedback to the user.
BeforeExecutionAttr, AfterExecutionAttr	DM, IR, IC, MC	Events' attributes in terms of a dictionary of key-value pairs.
ChoiceAction *(i); i = 1… n*	DM, IC, MC	Instructions executed in response to patient's selected choice.
ChoiceAttr *(i); i = 1 … n*	DM, IC, MC	Events' attributes passed to event handler as parameters of certain actions.
AttachedToChoice	DM, IR, IC, MC	A status flag Indicates whether the node is attached to a parent multiple-choice answer or not. This is particularly important for navigation between sibling nodes.
Visited	DM, IR, IC, MC	A status flag indicates whether the node has already been performed or not.

[(*)] *n = number of choices in multiple-choice nodes*.

B. Interaction events

Lifecycle of interaction nodes passes through three sequential states: initialisation, execution, and termination, as shown in Fig. 5. Throughout these states, several events are triggered, depending on the type of the interaction node. The execution state of any node can be skipped during initialisation by using a verbal instruction "Skip", as illustrated.

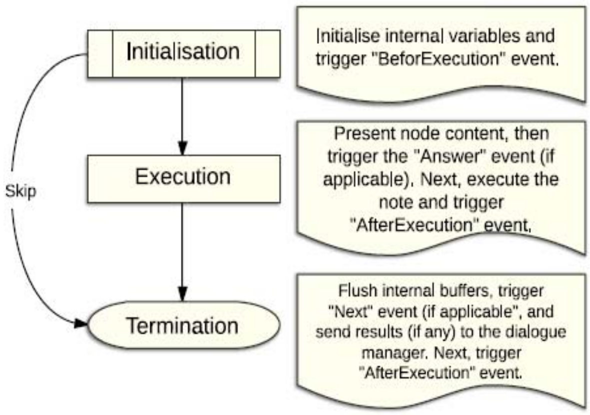


Figure 5. Interaction-node lifecycle.

C. XML DOM

The XML DOM is used to implement the interaction tree model due to its close match to requirements of the proposed interaction system in terms of navigation flexibility and adaptability to user responses. The root interaction element is represented by the following three children elements; internal interaction identity number, timestamp of the interaction assignment, and children nodes of the interaction. Each node has several attributes, including: ID number, type, category, visited status, content, and several choice elements depending on the type of node. Also, a node can have children elements which serve the purpose of either grouping nodes together under a single parent.

D. Navigation mechanism

In order to improve flexibility of the patient-robot interaction dialogues, two mechanisms are suggested to select the next node to be executed. The first mechanism is a traversing algorithm that generates a default execution path. This algorithm is consulted at the beginning of a certain interaction dialogue and after completing execution of each node in order to locate the closest non-visited or non-answered node within the same interaction level (e.g., a child or a sibling node). If none of these node types exists, it navigates back to the parent node in the tree hierarchy where a new execution path will be selected. A pseudo-code implementation of the suggested traversing algorithm is shown in Fig. 6.

Code segment

```
def NextNode(Node):
    if Node is None:
        return None
    if Node.hasUnvisitedChildren():
        for all nodes c in Node.getChildren():
            if not c.isAnswer():
                return c
    if Node.hasUnvisitedSiblings():
        for all nodes n in Node.getSiblings():
            if not n.isAnswer():
                return n
    return NextNode(n.parent())
```

Figure 6. Example of a tree-traversing code.

The second mechanism requires explicit setup of the "next" attribute of an interaction node. After the execution of the current interaction node is completed, the "next" event of the current node is triggered and event instruction(s) are executed. The next event usually contains a single "GOTO" instruction pointing to the next node to be executed within the interaction tree. Another feature of the dialogue system is the users' ability to skip dialogue nodes using the keyword "skip" which skips the current interaction node and moves the execution to the next node.

Example of a possible navigation sequence of an interaction tree is shown in Fig. 7. The dashed arrows represent the traversing steps. As can be seen from the figure, the execution starts with the root Node *A* and then moves to Node *B*, assuming that this node is the first child of the root node. The execution sequence will then move from Node *B* to its first child Node *E*. Since *E* is not attached to a choice of a data collection node and has no child nodes, the execution will move to its first sibling (Node *F*). By applying same conditions of Node *E*, the execution moves on to Node *G* from which, the execution sequence returns back to the parent Node *B*. At node *B*, the algorithm will search for a non-visited sibling of the parent node (e.g., Node *C*) which has no children. Similarly, at node *C*, the execution moves to Node *D* which is a decision node with two children attached to its two choices. Depending on the patient's answer, the execution sequence moves to either Node *H* or Node *I*. As none of these nodes has children-free sibling nodes (i.e. not attached to a choice of a parent node), the execution will return to the parent node (see Fig. 7), and so on until all nodes of the interaction dialogue are completed.

V. Results and Discussion

Numerous experiments were carried out using the proposed dialogue system running on a humanoid robot (called Nao) [16]. This robot incorporates text-to-speech engine and components for gesture/posture presentations that are used to present some emotions such as happiness, excitement, disappointment, etc. These emotions are accompanied with the verbal dialogues to improve patient-robot interactivity. Fig. 8 shows examples of remotely assigned robot-patient interactions using different interfaces; (a) tactile sensors, (b) verbal, or (c) performing

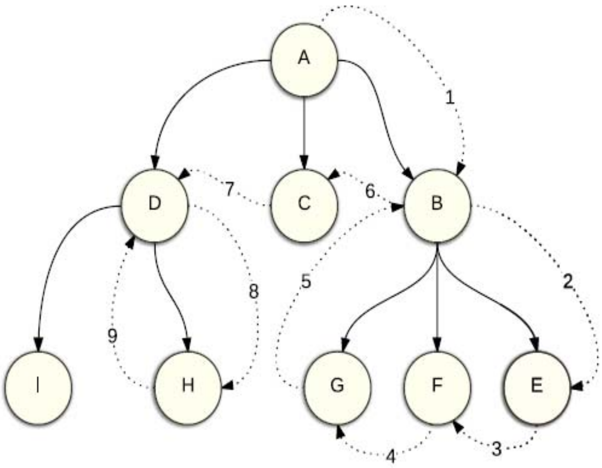


Figure 7. Example of a possible navigation sequence.

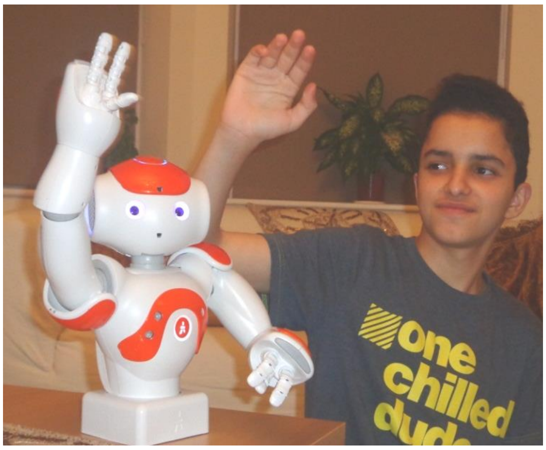
Patient-robot interaction via tactile sensors

(a)

Patient-robot interaction via a verbal dialogue

(b)

Performing joint physical exercise



(c)

Figure 8. Examples of remotely assigned robot-patient interactions.

a joint physical exercise. These and many other experiments which were conducted throughout the evaluation phase of the developed application validated flexibility of the proposed dialogue structure. These interactions are based on textual dialogues prepared remotely at the health portal and sent automatically to the robot's database. The robot in turn verbalizes these dialogues to the user along with appropriate postures/ gestures. The primary goal of the developed dialogue structure is to support diabetes self-management through improving bidirectional information flow between patients and health care providers. This includes but not limited to the following:

a) Remote collection and monitoring of patient's data such as physical activity, diet, medications, and wellbeing as well as patient's views in the form of short audio messages.
b) Automatic recognition of various blood glucose patterns and provision of timely feedback advices to patients.
c) Provision of education modules about diabetes which are tailored to the individual needs of patients as part of their disease self-management programmes.

VI. Conclusions

The design, implementation, and testing of a highly flexible human-robot interaction dialogue structure based on XML DOM have been described in this paper. Performance of the proposed dialogue structure is evaluated by incorporating it into a humanoid robot to perform various types of local and remote interaction between the robot and its users. These dialogues are tailored to support diabetes self-management in children in terms of automatic data collection and remote monitoring, decision support and education about diabetes. The evaluation results and observations showed (i) a seamless and accurate data transfer between the robot and a remote Web-based health portal, and (ii) a high level of navigation flexibility between different nodes of the experimented dialogues. Despite the technical contributions and pilot evaluation of the developed system, the work reported in this paper is still open for further technical improvement and clinical evaluations within the context of diabetes self-management support as well as other chronic diseases. These improvements and others are currently part of the ongoing research of the authors.

References

[1] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: A survey," J. Foundations and Trends in Human-Computer Interaction, vol. 1, no. 3, 2007, pp. 203-275, doi: 10.1561/ 1100000005.

[2] J. C. Scholtz, "Human-robot interactions: creating synergistic cyber forces," Multi-Robot Systems: From Swarms to Intelligent Automata - Proc. from the 2002 NRL Workshop on Multi-Robot Systems, Kluwer Academic Publishers, pp. 177-184, 2002, doi: 10.1007/978-94-017-2376-3_19.

[3] V. Klingspor, J. Demiris, and M. Kaiser, "Human-robot communication and machine learning," J. Applied Artificial Intelligence, vol. 11, no. 7/8, 1997, pp. 719-746.

[4] A. Green, "Designing and evaluating human-robot communication - informing design through analysis of user interaction," PhD thesis, KTH, School of Computer Science and Communication, Human - Computer Interaction, MDI, 2009.

[5] J. Scholtz, "Theory and evaluation of human robot interactions," Proc 36th Annual Hawaii Int. Conf. on System Sciences, 8-9 January 2003, pp. 10, doi: 10.1109/HICSS.2003.1174284.

[6] R. Bogue, "Robots in healthcare," J. Industrial Robot, Emerald Group Publishing Limited, vol. 38, no. 3, pp. 218-223, 2011, doi: 10.1108/ 01439911111122699.

[7] C. Bartneck, "Interacting with an embodied emotional character," Proc. Int. Conf. on Designing Pleasurable Products and Interfaces, New York, NY, USA, 2003, pp. 55-60, doi: 10.1145/782896. 782911.

[8] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric., "The role of physical embodiment in human-robot interaction," Proc. 15th IEEE Int. Symposium on Robot and human interactive communication, September 2006, pp. 117-122, doi: 10.1109/ roman.2006.314404 .

[9] E. Broadbent, H. Kuo, Y. I. Lee, J. Rabindran, N. Kerse, R. Stafford, and B. A. MacDonald, "Attitudes and reactions to a healthcare robot," Telemedicine and e-Health, vol. 16, no. 5, pp. 608-613, June 2010, doi: 10.1089/tmj.2009.0171.

[10] F. Tanaka, A. Cicourel, and J. R Movellan, "Socialization between toddlers and robots at an early childhood education center," Proc. of the National Academy of Sciences of the United States of America, 104(46), 2007, pp. 17954-17958, doi: 10.1073/pnas. 0707769104.

[11] R. Looije, M. A. Neerincx, and V., "Children's responses and opinion on three bots that motivate, educate and play," Journal of Physical Agents, 2(2), 2008, pp. 13 -20.

[12] H. Cuayahuitl, and I. Kruijff-Korbayov´a, "An interactive humanoid robot exhibiting flexible sub-dialogues," Proc. of the NAACL-HLT'12, Montréal, Canada, June 3-8, 2012, pp. 17-20,

[13] M. Goan, H. Fujii, and M. Okada, "Child-robot interaction mediated by building blocks: from field observations in a public space," Artificial Life and Robotics, 10(1), pp. 45-48, 2006, doi: 10.1007/s10015-005-0375-3.

[14] A. Mesquita, "Sociological and philosophical aspects of human interaction with technology: advancing concepts," IGI Global, 2011, doi: DOI: 10.4018/978-1-60960-575-9.

[15] W3C Architecture Domain, "Document object model (DOM) technical reports," available at: http://www.w3.org/DOM, (last accessed on July 27, 2013).

[16] Aldebaran robotics, "NAO humanoid robot platform," available at: http://aldebaran-robotics.com/en/discover-NAO/datasheet.html, (last accessed on July 22, 2013).

[17] M. A. Al-Taee and S. N. Abood, "Mobile acquisition and monitoring system for improved diabetes management using emergent wireless and web technologies," Int. J. of Information Technology and Web Engineering, 7 (1), 2012, pp. 18 – 32, doi: 10.4018/jitwe.2012010102.

[18] M. A. Al-Taee, N. A. Jaradat, and D. M. Abu Ali, "Mobile phone-based health data acquisition system using Bluetooth technology," Proc. IEEE Jordan Conf. on Applied Electrical Engineering and Computing Technologies, Amman-Jordan, 6-8 December, 2011, doi: 10.1109/AEECT.2011.6132524.

[19] M. A. Al-Taee, A. M. Sungoor, S. N. Abood and N. Y. Philip, "Web-of-things inspired platform for integrated diabetes management," Proc. IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT'2013), Amman, Jordan, 3-5 December 2013, doi: 10.1109/AEECT.2013. 6716427.

[20] A. H. El-Mousa, Z. J. Muhsin, and M. A. Al-Taee, "A web-based rapid prototyping workflow management information system for computer repair and maintenance," Int. J. of Computer Science, 4(12), 2008, pp. 991-998, doi:10.3844/jcssp.2008.991.998.