Development of Robot-enhanced Therapy for
Children with Autism Spectrum Disorders

Project No. 611391

DREAM
Development of Robot-enhanced Therapy for Children with Autism Spectrum Disorders

Grant Agreement Type:	Collaborative Project
Grant Agreement Number:	611391

D6.2 Attention Subsystem

Due date: **01/04/2016**
Submission Date: **30/03/2016**

Start date of project: **01/04/2014** Duration: **54 months**

Organisation name of lead contractor for this deliverable: **Plymouth University**

Responsible Person: **Tony Belpaeme** Revision: **1.3**

Project co-funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	**PU**
PP	Restricted to other programme participants (including the Commission Service)	
RE	Restricted to a group specified by the consortium (including the Commission Service)	
CO	Confidential, only for members of the consortium (including the Commission Service)	

Contents

Executive Summary

Deliverable D6.2 defines the specification, design and implementation of the Attention subsystem within the cognitive architecture in Work Package 6.

Specifically, this report presents the outcome of task T6.2 over the first 24 months of the DREAM project. Motivations and technical details of the Attention Subsystem itself are primarily contained within the annex to this report.

Principal Contributors

The main authors of this deliverable are as follows (in alphabetical order).

Paul Baxter, Plymouth University
Tony Belpaeme, Plymouth University
Hoang-Long Cao, Vrije Universiteit Brussel
Albert de Beir, Vrije Universiteit Brussel
Pablo Gómez, Vrije Universiteit Brussel
Emmanuel Senft, Plymouth University
Bram Vanderborght, Vrije Universiteit Brussel

Revision History

Version 1.0 (P.G. 02-03-2016)
First draft of deliverable structure.

Version 1.1 (P.G. 20-03-2016)
Added state of the art Section in Technical Report.

Version 1.2 (P.G. 25-03-2016)
Added implementation Section in Technical Report.

Version 1.3 (P.G. 30-03-2016)
Minor corrections.

1 Outline of this deliverable

1.1 Description of Task 6.2

[*From the DREAM Description of Work*]: The Attention system is a combination of perceptual attention, in which perceptual stimuli (reported by, for example, face detection or sound localization in work package WP4) that are particularly salient in the current context have to be selected, and attention emulation (from the Deliberative subsystem) to direct the robot's attention and gaze. This provides the robot with a locus of attention that it can use to organize its behavior. The perceptual attention will be based on work of [1] where different weighted visual stimuli are implemented and integrated with the basic intensity and localization of audio signals together with a habituation filter, to obtain a basic real-time and animal/human-like reaction. Attention emulation is introduced so partners of an interaction can orient attention to an object, event or person. Social interaction and joint attention in particular will be guided by the outputs of the Deliberative subsystem. The performances of the system will be gradually improved by adding saliency maps and habituation effects. This subsystem is important since, generally speaking, people with ASD avoid eye contact and have difficulties following the gaze or deictic pointing of others. Several DREAM scenarios will focus on improving this social skill through robot therapy.

1.2 Description of D6.2

The purpose of this deliverable is to report the progress of development of the Attention subsystem. It defines the specification, design and implementation of this component within the cognitive controller architecture as defined in Work Package 6 (see previous section).

2 The Attention Subsystem

As explained in document D6.1, both Reactive and Attention subsystems are combined into a single component since both systems require access to features of the environment and interacting person(s) to respond appropriately (e.g. looking at a face or diverting attention to a loud noise somewhere in the environment). Managing this in a single component therefore seems a sensible choice so that functionality is not replicated. Within this document only the Attention subsystem is described. The functionality of the Reactive system was described in annex 3.2.

The background and technical details of the Attention subsystem may be found in annex 3.1.

References

[1] Cynthia Breazeal and Brian Scassellati. A context-dependent attention system for a social robot. *rn*, 255:3, 1999.

3 Annexes

3.1 Esteban, P.G. et al. (2016), Technical Report: Attention subsystem

Abstract - The purpose of this technical report is to summarize the motivations and constraints underlying the Attention subsystem, and to outline an organisation of it. This is a proposal only; this document is intended to be a working one, to be updated as required during development. This version of the report is based primarily on the discussions that took place in Brussels (23/01/15).

Relation to WP This work outlines the technical implementation of the attention subsystem, and the background thereof. This is relevant to T6.2.

3.2 Esteban, P.G. et al. (2016), A multilayer reactive subsystem for robots interacting with children with autism

Abstract - There is a lack of autonomy on traditional Robot-Assisted Therapy systems interacting with children with autism. To overcome this limitation a supervised autonomous robot controller is being built. In this paper we present a multilayer reactive system within such controller. The goal of this Reactive system is to allow the robot to appropriately react to the child's behavior creating the illusion of being alive.

Relation to WP This work outlines the Reactive subsystem and its relationship with other subsystems. This is relevant to T6.2.

Project No. 611391

DREAM
Development of Robot-enhanced Therapy for
Children with Autism Spectrum Disorders

TECHNICAL REPORT
Attention subsystem

Date: **20/02/2016**

Technical report lead partner: **Vrije Universiteit Brussel**

Primary Author: **P. Gómez Esteban**

Revision: **1.2**

Project co-funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Service)	**PP**
RE	Restricted to a group specified by the consortium (including the Commission Service)	
CO	Confidential, only for members of the consortium (including the Commission Service)	

Contents

Summary

The purpose of this technical report is to summarize the motivations and constraints underlying the Attention subsystem, and to outline an organization of it. This is a proposal only; this document is intended to be a working one, to be updated as required during development. This version of the report is based primarily on the discussions that took place in Brussels (23/01/15) and periodically teleconference calls.

Principal Contributors

The main authors of this document are as follows (in alphabetical order).

Paul Baxter, Plymouth University
Tony Belpaeme, Plymouth University
Hoang-Long Cao, VUB
Albert De Beir, VUB
Pablo Gomez, VUB
Emmanuel Senft, Plymouth University
Bram Vanderborght, VUB

Revision History

Version 1.0 (P.G. 20-02-2016)
Initial outline of ideas for the DREAM attention subsystem.

Version 1.1 (P.G. 18-03-2016)
Added state of the art Section.

Version 1.2 (P.G. 24-03-2016)
Added implementation Section.

1 Overall Organisation

The autonomous controller is informed by four external sources: the child behavior description, sensory information, current intervention script state, and input from a therapist (e.g. emergency stop). Combining these sources, the autonomous controller should trigger as an output the appropriate sequence of action primitives to be performed (as well as some feedback via the WoZ GUI), which then gets executed on the robot.

The autonomous controller is composed of a number of subsystems: Reactive, Attention, Deliberative, Self-Monitor and Expression and Actuation. In the Reactive subsystem, sensory inputs are immediately acted upon with appropriate actuator outputs, see [1] and document D6.1. The Attention subsystem determines the robot's focus of attention. In the Deliberative subsystem, the necessary interventions will be implemented in a general approach so it is not scenario-specific. The Self-Monitoring subsystem acts as an alarm system in two specifications. An internal one when the robot detects that it cannot act because of a technical limitation or an ethical issue. An external alarm can be triggered where the therapist overrules the robot behavior selection. Finally, the Expression and Actuation subsystem is responsible for generating natural motions and sounds that are platform independent. These subsystems interact, and must combine their suggested courses of actions to produce a coherent robot behavior, in the context of constraints provided by the therapist (for example, the script to be followed, types of behavior not permissible for this particular child because of individual sensitivities, etc). In the cognitive controller architecture it is defined the control that the supervising therapist can exert over the behavior of the robot (effectively a limited 'remote control' functionality). This naturally has a number of operational consequences for other subsystems of the cognitive controller, which will be handled through the oversight of the Self-Monitoring subsystem (to prevent conflicting commands for example). As a result, we have formulated the following architecture describing how cognitive control informed by the therapy scripts is to be achieved (Figure 1).

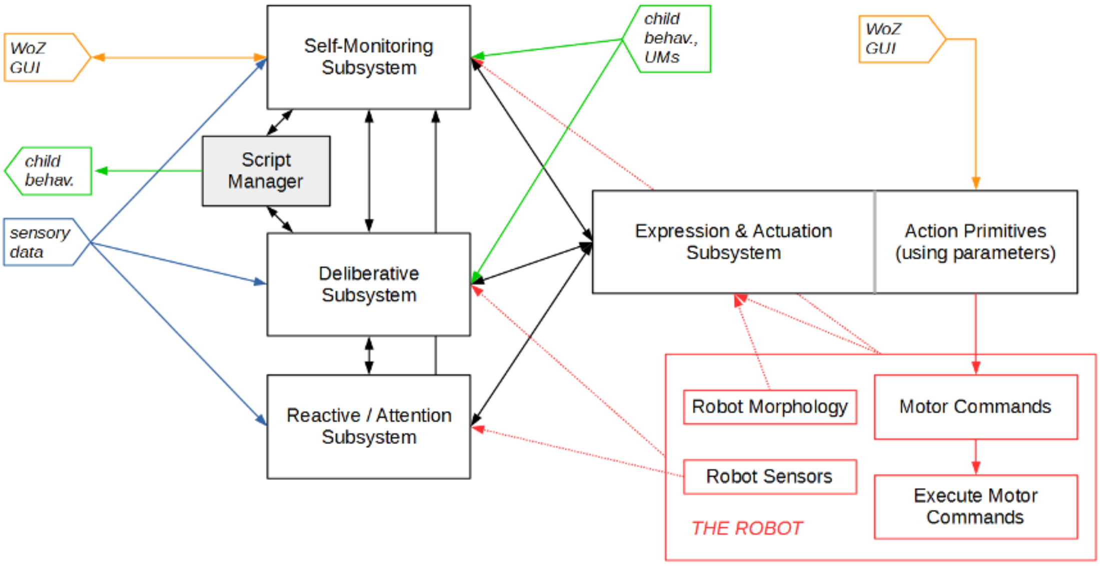


Figure 1: Description of the cognitive controller subsystems. UMs: User Models.

Note that both Reactive and Attention subsystems are combined into a single component. As explained in document D6.1, both systems require access to features of the environment and interacting person(s) to respond appropriately (e.g. looking at a face or diverting attention to a loud noise somewhere in the environment). Managing this in a single component therefore seems a sensible choice so that functionality is not replicated. Within this document only the Attention subsystem is described. The functionality of the Reactive system was described in document D6.1 and [1].

A general high level description of the Attention subsystem is shown in Figure 2. The Attention subsystem is composed of the attention model and the gaze controller. The attention model is a combination of perceptual attention, in which perceptual stimuli (reported by, for example, face detection or sound localization in work package WP4) that are particularly salient in the current context have to be selected, and attention emulation (from the Deliberative subsystem) to direct the robot's attention and gaze. A gaze reaction triggered in social interactions is produced by the Reactive subsystem. These inputs provide the robot with a locus of attention that it can use to organize its behavior. The gaze controller is open for commands from the Self-Monitoring subsystem to overrule this locus of attention whenever the therapist considers it is needed.

The Attention subsystem provides the gaze direction towards the Actuation subsystem and a signal indicating a variation on the gaze of the robot to the eye blinking module of the Reactive subsystem.

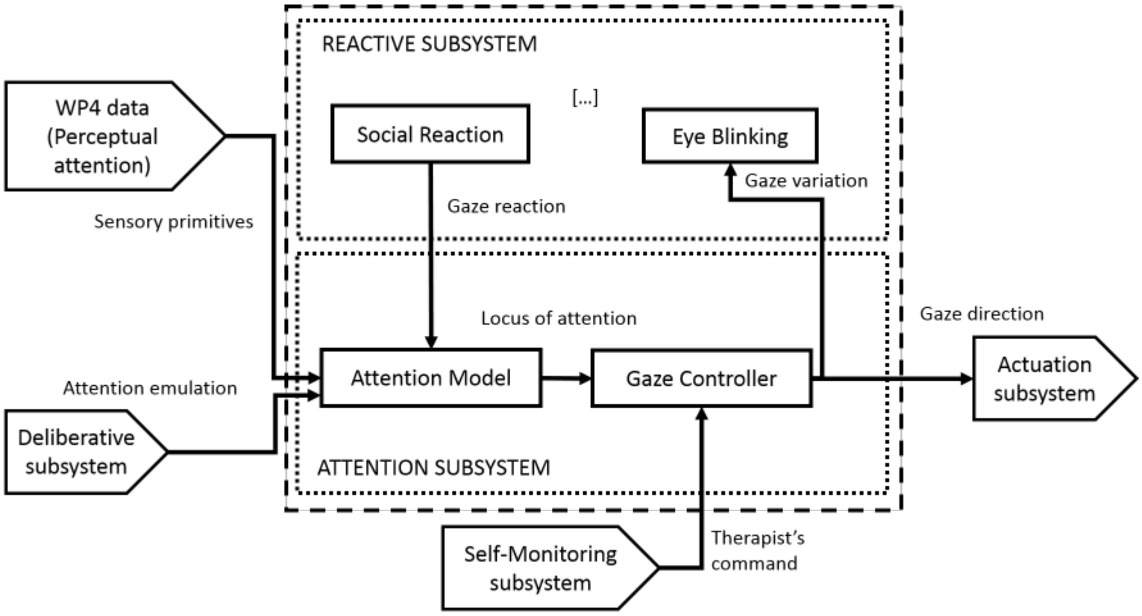


Figure 2: High level description of the Attention subsystem. It receives inputs from several sources and produces outputs for the Actuation subsystem.

A simplified but functional simulator of those components that interact with this one (except the Reactive subsystem which is the actual component) has been developed. For instance, the attention emulation from the Deliberative subsystem or those outputs from WP4, see Section 5 of this Technical Report for further details.

2 Brief state of the art

Social robots require mechanisms to control attention and gaze based on social cues, visual and auditory stimuli. Attention selection is usually described by bottom-up and top-down approaches [2]. The bottom-up approach is driven by external stimuli while the top-down approach is driven by the state of the system itself.

The logic followed by most of bottom-up approaches are as follows. First, algorithms extract visual and/or auditory low-level features of the surroundings. Next, local competition is computed producing feature maps. Finally, such feature maps are combined by weighted sums creating the salience map. Then an algorithm can select attention targets by applying the "winner-takes-all" principle on such salience map.

The top-down approach reflect a processing model comprising high-level feature interpretation such as motivations and goals. These top-down factors modulate the output of the individual feature maps before they are summed to produce the bottom-up contribution.

Within the field of human-robot interaction (HRI), gaze shifts between the speaker and the listener should be replicated by attention models. When designing a social robot attention system high-level social features must be considered, see [3] where limitations of salience-based attention modeling are shown. One of the first works done within HRI is [4] where authors show an attention system that works following bottom-up and top-down approaches. The bottom-up approach considers color, motion and faces as the variables to pay attention to creating salience feature maps. They use habituation effect to modify such maps. The top-down approach depends on the motivations and behaviors of the robot. Each behavior is implicitly associated with one of these features in the salience maps so when selecting a certain behavior, the weight related to such feature will vary and like that the top-down approach is followed. In [5] a cognitive architecture including a visual and auditory attention makes a social robot capable of focusing its attention on the speaker during a conversation. The results are satisfactory, but this approach does not consider some of the fundamental human communicative cues for social attention, like gestures or proxemics. Considering these cues [6] proposed a receptionist robot which drives its attention according to the spatial information of humans interacting with itself. Results show a natural social gaze behavior for these type of situations. Finally, in [7] a modular context-dependent social robot gaze-control system (GCS) has been implemented to direct attention at the appropriate target during interactions with humans. The attention modeling is focused on high-level social features, proxemics and orientation of the speakers. We have based our development on the target selection algorithm of this work due to its simplicity.

3 Attention Model

The goal of the attention model is to provide a point of interest where the robot should look at based on the events occurring within its surroundings. The attention model receives perceptual attention data from DREAM's Work Package 4 and the Reactive subsystem, and attention emulation from the Deliberative subsystem. Attention emulation is introduced so partners of an interaction can orient attention to an object, event or person. Social interaction and joint attention in particular will be guided by the outputs of the Deliberative subsystem.

Given the context in which this subsystem will be implemented, attention behavior will be divided between deliberative (where the attention is determined by the requested scenario) and non-deliberative interactions. Within the first ones, the highest priority should be given to the Deliberative subsystem outputs. Therefore, each time attention emulation is triggered such point of interest is where the robot will look at, unless the therapist decides to overwrite such behavior through the Self-Monitoring subsystem as it will be explained later.

Within non-deliberative interactions the attention model will look for which is the next point of interest to look at. For such purpose we have built a target selection algorithm adapted from [7] where authors present a bottom-up attention model based on social features. Some simplifications of such model have been done to adapt it to our context.

The social features we consider for our model are some of the sensory primitives captured by Work Package 4 and the social reaction output from the Reactive subsystem:

- getFaces ($<$ x,y,z $>$),
- getSoundDirection (threshold, azimuth, elevation),
- getEyeGaze (eye, x, y, z),
- gaze reaction.

Details of these sensory primitives can be found in document D1.3. Each of these inputs provides coordinates (x, y, z) associated with those points of interest within the scene.

Given N points of interest, the Elicited Attention (EA_i) of certain point of interest $i \in N$ is determined by its social value (SV_i) and a value related to its distance (D_i) from the robot position:

$$EA_i = SV_i + D_i.$$

The social value of i is defined as:

$$SV_i = \frac{p_i}{N},$$

being p_i the priority of the point of interest i which is a non-repeating random value between 1 and N. So that from one intervention to another the priority of the points of interest of a robot will vary randomly. Note that this behavior is only running within non-deliberative interactions where unpredictability is not a problem and helps with the purpose of showing aliveness by the robot.

D_i is defined as:

$$D_i = \left(1 - \frac{|dist_i|}{dist_{max}}\right),$$

being $dist_i$ the euclidean distance between the point of interest i and the robot.

The selected target will be the one that satisfies:

$$EA_{winner} = \max(EA_i \times HF_i),$$

where HF_i is the habituation effect of certain i. The habitation effect is a decay factor in response to a stimulus after repeated presentations [8]:

$$HF_i = Peak \times \max\left(0, 1 - \frac{\Delta t}{\tau}\right),$$

being $Peak$ the maximum amplitude of the HF and τ a time constant. According to [7] these parameters have been set to 30 and 10 seconds respectively.

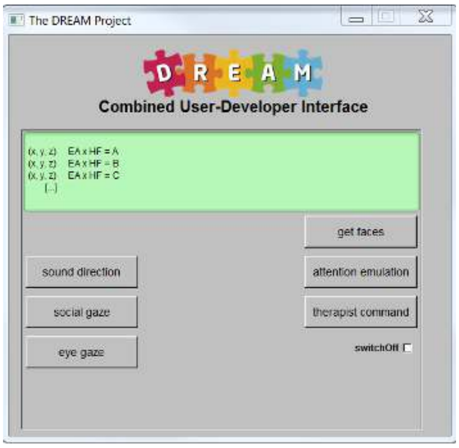
4 Gaze Controller

The gaze behavior of the robot is so far limited to head gaze which does not mean that there will not be eye gaze in the future. We leave that for future developments.

The gaze controller manages the output of the Attention subsystem. It will receive inputs from the attention model and the Self-Monitoring subsystem which will have the highest priority. Therefore, whenever the therapist decides to overrule the attention behavior, the command from the Self-Monitoring subsystem will arrive to the gaze controller which will send the corresponding command to the Actuation subsystem. Otherwise, the resulting output will be the one coming from the attention model.

5 Implementation

The components that interact with this one have been developed as basic simulators that provide the expected output. For some of them, such expected output was provided through a GUI, see Figure 3. That is the case of the Deliberative subsystem (attention emulation) and the Reactive subsystem (social gaze).



```
(x, y, z)   EA x HF = A
(x, y, z)   EA x HF = B
(x, y, z)   EA x HF = C
    [...]
```

Figure 3: Graphical User Interface of the Attention subsystem. The Output panel shows coordinates and computed value of each of the points of interest.

Such GUI has been used to simulate the sensory information of the robot as a first step in the implementation of this system. It should be replaced by a sensory information simulator of an specific robot, i.e. Nao robot. But for the purpose of validating the Attention subsystem the GUI is an equally good option. It also includes other inputs coming from subsystems that are still to be developed.

On the other hand, this system uses actual actuators to show its outputs. This implementation has been done in a Nao Robot using different layers of Yarp, see [9], keeping its platform-independent flavor.

It also includes an Actuation subsystem simulator which receives the outputs of the Attention subsystem and provides the corresponding action primitives.

References

[1] Pablo Gómez Esteban, Hoang-Long Cao, Albert De Beir, Greet Van de Perre, Dirk Lefeber, and Bram Vanderborght. A multilayer reactive system for robots interacting with children with autism. In *Fifth International Symposium on New Frontiers in Human-Robot Interaction*, 2016. To be published.

[2] Timothy J Buschman and Earl K Miller. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *science*, 315(5820):1860–1862, 2007.

[3] Benjamin W Tatler, Mary M Hayhoe, Michael F Land, and Dana H Ballard. Eye guidance in natural vision: Reinterpreting salience. *Journal of vision*, 11(5):5–5, 2011.

[4] Cynthia Breazeal and Brian Scassellati. A context-dependent attention system for a social robot. *rn*, 255:3, 1999.

[5] J Gregory Trafton, Magda D Bugajska, Benjamin R Fransen, and Raj M Ratwani. Integrating vision and audition within a cognitive architecture to track conversations. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 201–208. ACM, 2008.

[6] Patrick Holthaus, Karola Pitsch, and Sven Wachsmuth. How can i help? *International Journal of Social Robotics*, 3(4):383–393, 2011.

[7] Aolfazl Zaraki, Daniele Mazzei, Manuel Giuliani, and Danilo De Rossi. Designing and evaluating a social gaze-control system for a humanoid robot. *Human-Machine Systems, IEEE Transactions on*, 44(2):157–168, 2014.

[8] Richard F Thompson and William A Spencer. Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychological review*, 73(1):16, 1966.

[9] Paul Baxter, Tony Belpaeme, and Emmanuel Senft. Robot Motor Control. Technical report, University of Plymouth, 2014.

A multilayer reactive system for robots interacting with children with autism

Pablo Gómez Esteban [1], **Hoang-Long Cao**[1] , **Albert De Beir**[1] ,
Greet Van de Perre[1] , **Dirk Lefeber**[1] and **Bram Vanderborght**[1]

Abstract. There is a lack of autonomy on traditional Robot-Assisted Therapy systems interacting with children with autism. To overcome this limitation a supervised autonomous robot controller is being built. In this paper we present a multilayer reactive system within such controller. The goal of this Reactive system is to allow the robot to appropriately react to the child's behavior creating the illusion of being alive.

1 INTRODUCTION

Robot-Assisted Therapy (RAT) is widely used, particularly with children with special needs, see [5] and [20] as examples, reducing the workload of the therapy and therefor its cost. While the benefits of using RAT are undisputed, current approaches [10] typically constrain themselves to the Wizard of Oz (WOZ) paradigm [11] [24], where the robot is remotely controlled by a human operator, usually the therapist. According to [19], for a long-term use the WOZ framework is not a sustainable technique. Robots in RAT are required to become more autonomous in order to reduce cost and time within the therapeutic interventions, see [22].

Under such circumstances the DREAM project (Development of Robot-Enhanced therapy for children with AutisM spectrum disorders) was conceived. This project is concerned, among other research challenges, with the development of an autonomous controller. Despise full autonomy is currently unrealistic, a "supervised autonomy", where the operator gives the robot certain goals and the robot autonomously works towards achieving them, is certainly feasible. This controller is composed of a number of systems: Reactive, Attention, Deliberative, Self-monitoring and Actuation, see Figure 1, and complemented by sensory data and a module to assess the performance and motivation of the child. The focus of this paper is on the Reactive system.

The Reactive system is constituted of the lowest-level processes. In natural systems, these processes are genetically determined and not typically sensitive to learning. State information, coming from the sensory inputs, is immediately acted upon with appropriate motor outputs. The Reactive system, while absent in many robot systems, is essential in social robots, see [12]. It creates the illusion of the robot being alive [17], and acts as a catalyst for acceptance and bonding between the young user and the robot. It ensures that the robot can handle the real time challenges of its environment appropriately taking care of small motions, appropriate eye blinking, whole body motion during gesturing and head motion, recovering from falls, and appropriately reacting to affective displays by young users. The behaviors will be configurable by the therapist as it might not be desirable for some children to have the robot display a full gamut of reactive responses (for example, a negative reaction when being pushed).

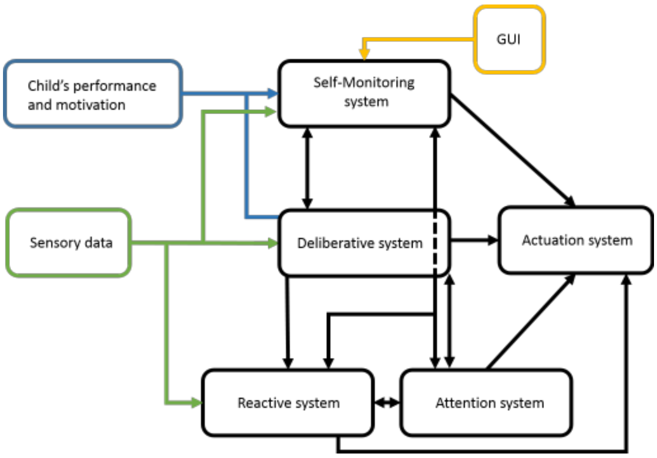


Figure 1. Project DREAM's cognitive architecture is composed of several systems (in black) and complemented by an assessment of the child performance and motivation (in blue) and sensory data (in green). Therapist can control the cognitive architecture through a GUI (in yellow). Arrows show flow of information between the systems.

This paper is structured as follows. In Section 2 a high level description of the system is provided. Through subsections 2.1 to 2.4 the different layers composing the Reactive system are detailed. Finally, some future work is provided in Section 3.

2 A MULTILAYER REACTIVE SYSTEM

A general high level description of the Reactive system is shown in Figure 2. This describes how, given the sensory information the robot reacts to the current situation. Such information is processed by different layers producing each own outputs towards the Actuation system which will combine them all, according to predefined priorities, to produce the final outcome of the cognitive architecture.

The Reactive system is composed of a number of layers:

- The falling reaction triggers a damage avoidance posture when falling. At that moment it interrupts all the running behaviors. Once the robot is back at its feet, it takes care of restoring the intervention behavior.

---

[1] Vrije Universiteit Brussel, Belgium, email: pablo.gomez.esteban@vub.ac.be

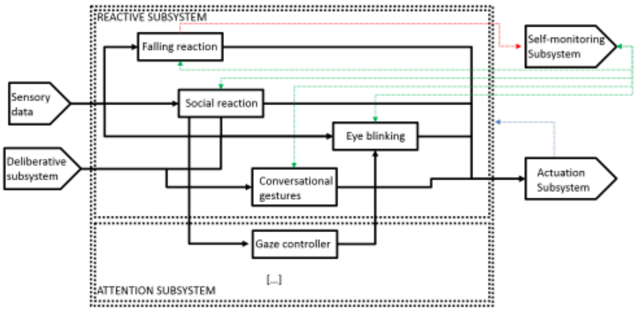


Figure 2. High level description of the Reactive system. It receives inputs from the Deliberative system but mainly from the sensory data, and produces its output towards the Actuation system. The Self-monitoring system allows the therapist, through a GUI (see Figure 1), to switch on/off the functionality of each layer (green arrows). The falling reaction layer might send a signal to interrupt all running behaviors (red arrow). The Actuation system provides feedback about the execution of the motor commands (blue arrow). The remaining arrows show information flow between the layers.

- The social reaction purpose is to appropriately react to social displays of the children and to provide small motions and face/sound tracking features that will give the impression of the robot being alive.
- The eye blinking layer provides a variable blinking rate that complements other gestures and behaviors.
- Conversational gestures complement the speech acts with body gestures.

The therapist might consider that one or more of these layers are not appropriate for being used with a certain child, for such reason, their functionality, which are detailed in the following sections, can be switched on and off when needed through the Self-monitoring system.

One of the main contributions of this system is that it can be easily implemented in different robots due to its platform-independence flavor. The Actuation system is responsible of generating the appropriate motor commands depending on robot morphology. This system, see Figure 3, has access to the degrees of freedom of the robot and generates the corresponding motor commands, see [23] for further details.

2.1 Falling Reaction

Within social interaction with children it may happen that robots lose their balance and have to recover it or even they may fall down. These robotic platforms are expensive so that in case they fall, minimizing the hardware damage would be a priority. According to the intervention protocol we aim to use within DREAM project, the robot will be seated in front of the child, so that a fall is lowly probably to occur. Nevertheless, this module needs to be implemented to face such hypothetical situations.

The Falling Reaction module, see Figure 4, will be periodically checking the balance of the robot using the sensory information available. Changes in the balance may end up in a fall. In such case, a signal will be sent to interrupt any other running behavior, and a damage avoidance behavior that fits the situation will be triggered, see [7] for a case of minimizing damage to a humanoid robot, and [25] for a case of a NAO robot that modifies its falling trajectory to avoid causing injuries in people in front of it. These behaviors might be highly

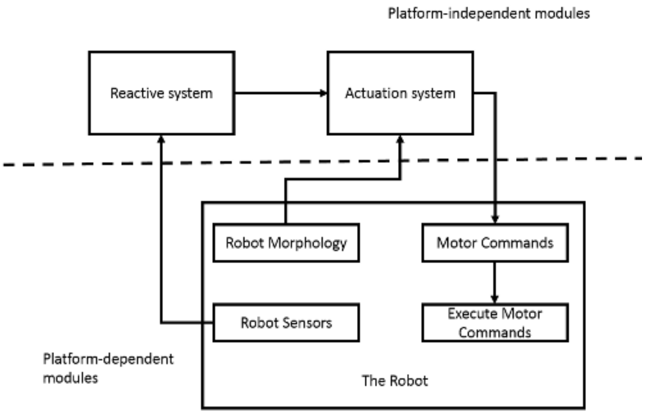


Figure 3. The Reactive system provides different outputs to the Actuation system. This system has access to the morphology and hence the degrees of freedom of the robot.

dependent on the morphology of the robot. Reducing the stiffness of the joints will avoid any mechanical problem independently of its morphology. Additionally, the robot should include some speech acts to reduce the impact of such dramatic situation for the kid as saying that it has been a little bit clumsy or that it is tired today.

Finally, back at its feet, the robot may apologize in order to engage the child back to the intervention and it will send a signal to restore the system functionality.

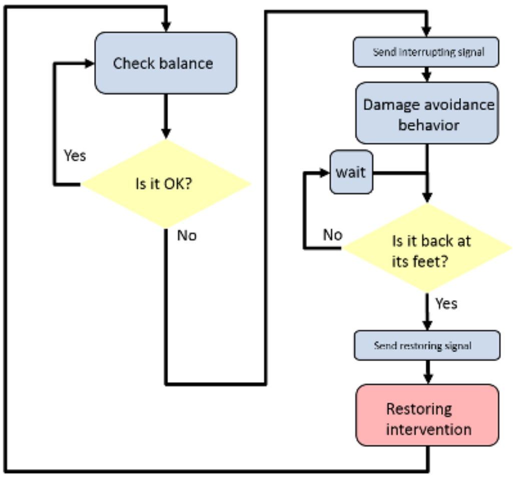


Figure 4. The module is periodically checking the balance of the robot. In case of a fall, a signal will be sent to interrupt any other running behavior, and a damage avoidance behavior will be triggered. Finally, back at its feet, the module will send a signal to restore the intervention.

2.2 Social Reaction

In social situations multiple verbal and non-verbal interactive encounters may occur. The child may behave friendly with the robot affectively touching it or may feel unfavorable to it and eventually hit it. These situations may be very conflicting as a special care must be paid with the potential audience of this system. If it would be the case of a regular social robot, for such both situations the robot may appropriately react, but under these circumstances, the reaction will be simplified to facial expressions and speech acts, always under the supervision of the therapist who might consider that such social reaction is not therapeutically appropriate for a specific child. Moreover, in order to reach an effective social interaction, emulating certain degree of empathy towards the social partner plays a key role in patient-centered therapy [21], i.e. if the child is expressing an emotion, the robot should be aware of that and react accordingly expressing a compatible emotion. In those cases in which there is no social interaction, this module will randomize among a set of small motions to recreate a life-like behavior such as a breathing motion, gaze-shifts or sound and/or face tracking. The purpose of this module is to provide the appropriate social behavior in order to give the impression of the robot being socially alive.

This module receives as input the sensory information where it is specified the child's social and affective state i.e. whether she/he is expressing an emotion or is performing a physical behavior (such as touching the robot unexpectedly). For each of these behaviors there should be a set of facial expressions and speech acts available to choose among them. Ideally it should randomize among them in order to look less predictable.

2.3 Conversational gestures

Exhibiting co-verbal gestures would make the robot appear more expressive and intelligible which will help to build social rapport with their users [15].

Co-verbal gestures are defined as the spontaneous gestures that accompany human speech, and have been shown to be an integral part of human-human interactive communications [14]. There exist evidences that co-verbal gestures have a number of positive effects performed by robots [18][9].

We adopted gestures from [2] where authors use Kendon's Open Hand Supine ("palm up") family of gestures which are related to offering and giving, see Figure 5 where our set of conversational gestures implemented in the Nao robot are shown. As explained in Section 1, we did not consider any negative gesture, as those belonging to the Open Hand Prone ("palm down") family, as it might be not appropriate for this audience.

For the purposes of DREAM project we don't aim at building a highly sophisticated conversational agent as [15] or [1] but to complement speech acts with conversational gestures, that the robot can randomly perform while speaking trying to improve the acceptability of the robot during the social interaction. For that reasons, we include a set of conversational gestures along with the rules to trigger them.

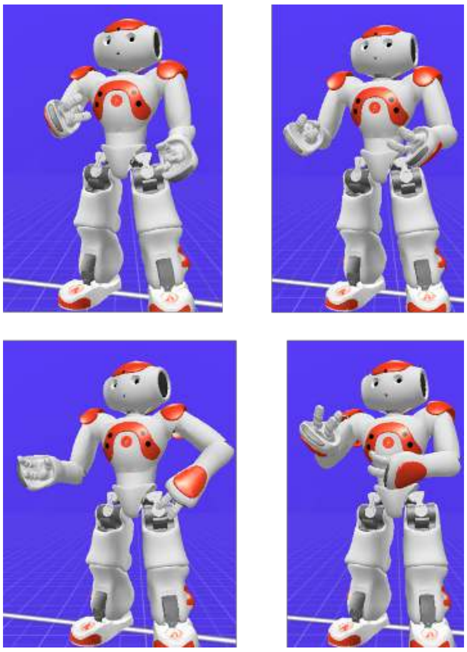


Figure 5. Set of conversational gestures belonging to Kendon's Open Hand Supine family related to offering and giving.

2.4 Eye Blinking

The acceptability of the robot can be further increased if the robot mimics the human blinking behavior. Simulating blinking behavior requires a human-level blinking model that should be derived from real data of human.

Several works have been done concerning the dependencies of human eye blinking behavior on different physiological and psychological factors. Ford et al. [6] proposed the "blink model" for HRI, which integrates blinking as a function of communicative behaviors. Doughty [4] described in his work three distinct blinking patterns during reading, during conversation and while idly looking at nothing specific. Lee et al. [13] proposed a model of animated eye gaze that integrates blinking as depending on eye movements constituting gaze direction.

Given the amount of studies made to model human blinking behavior we don't need to do our own but to use that one that best fits our requirements. Within the context in which DREAM will be applied, we need to recreate a blinking behavior mainly focused on the communicative behaviors and gaze shifts. For such reason, we have simplified and adapted Ford et al.'s model to our needs, see Figure 6, defining a model which considers multiple communicative facial behaviors. For each of them there is a probability to blink. Moreover there is a passive behavior which simulates a natural, or non-interactive, blinking mechanism (for cleaning or humidifying the eye) that can be activated when no other blinking behavior has been triggered. To perform the blinking motion there is a blink morphology module which defines, based on statistics, if the blink is simple or multiple, full or half, its duration, etc.

3 CONCLUSIONS AND FURUTE WORK

In this paper we present a multilayer Reactive system within an autonomous robot controller. The goal of such system is to allow the robot to appropriately react to the child's behavior creating the illusion of being alive. For such purpose it is composed of several layers that can be switched on and off by the therapist depending the needs of the intervention: the falling reaction layer is aimed to prevent and manage falls; the social reaction one to appropriately react to social displays; another one to provide a blinking behavior to complement gestures; and, finally, some conversational gestures to complement speech acts.

This controller has a platform-independent flavor which allows it to be implemented in multiple robotic platforms without spending

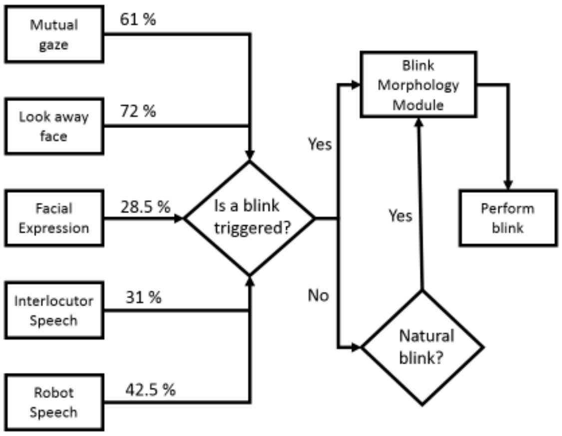


Figure 6. Adapted blinking model. When a communicative facial behavior occurs there is a probability of triggering a blink behavior. Such probabilities come from [6].

too much effort on it. Some test on Nao, Pepper and Romeo are about to be made. Also, studies on the acceptability of this system are under development.

Lip synchronization in robotics looks for matching lip movements with the audio generated by the robot. Several works use synchronization algorithms based directly on the use of audio phonemes to determine the levels of mouth aperture [16] [8]. These approaches require additional information such as dictionaries of phonemes. Currently all the robots available to this research group to implement this system have no mouth. As future work we aim to implement a basic lip synchronization method like [16] in the second version of the huggable robot Probo [3] which is currently under development.

ACKNOWLEDGEMENTS

The work leading to these results has received funding from the European Commission 7th Framework Program as a part of the project DREAM grant no. 611391.

REFERENCES

[1] Paul Bremner, Anthony Pipe, Chris Melhuish, Mike Fraser, and Sriram Subramanian, 'Conversational gestures in human-robot interaction', in *IEEE International Conference on Systems, Man and Cybernetics*, pp. 1645–1649. IEEE, (2009).

[2] Adam Csapo, Emer Gilmartin, Jonathan Grizou, JingGuang Han, Raveesh Meena, Dimitra Anastasiou, Kristiina Jokinen, and Graham Wilcock, 'Multimodal conversational interaction with a humanoid robot', in *IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 667–672. IEEE, (2012).

[3] Albert De Beir and Bram Vanderborght, 'Evolutionary method for robot morphology: Case study of social robot probo', in *11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, (2016).

[4] Michael J Doughty, 'Consideration of three types of spontaneous eyeblink activity in normal humans: during reading and video display terminal use, in primary gaze, and while in conversation', *Optometry & Vision Science*, **78**(10), 712–725, (2001).

[5] David Feil-Seifer and Maja Mataric, 'Robot-assisted therapy for children with autism spectrum disorders', in *Proceedings of the 7th international conference on Interaction design and children*, pp. 49–52. ACM, (2008).

[6] CC Ford, G Bugmann, and P Culverhouse, 'Modeling the human blink: A computational model for use within human–robot interaction', *International Journal of Humanoid Robotics*, **10**(01), (2013).

[7] Kiyoshi Fujiwara, Fumio Kanehiro, Shuuji Kajita, Kenji Kaneko, Kazuhito Yokoi, and Hirohisa Hirukawa, 'Ukemi: falling motion control to minimize damage to biped humanoid robot.', in *IROS*, pp. 2521–2526, (2002).

[8] Fumio Hara, Kouki Endou, and Shingo Shirata, 'Lip-configuration control of a mouth robot for japanese vowels', in *6th IEEE International Workshop on Robot and Human Communication (RO-MAN'97)*, pp. 412–418. IEEE, (1997).

[9] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga, 'Relating conversational expressiveness to social presence and acceptance of an assistive social robot', *Virtual Reality*, **14**(1), 77–84, (2010).

[10] Claire A. G. J. Huijnen, Monique A. S. Lexis, Rianne Jansens, and Luc P. Witte, 'Mapping robots to therapy and educational objectives for children with autism spectrum disorder', *Journal of Autism and Developmental Disorders*, 1–15, (2016).

[11] Thomas K Landauer, 'Psychology as a mother of invention', *ACM SIGCHI Bulletin*, **18**(4), 333–335, (1987).

[12] Nicole Lazzeri, Daniele Mazzei, Abolfazl Zaraki, and Danilo De Rossi, 'Towards a believable social robot', in *Biomimetic and Biohybrid Systems*, volume 8064 of *Lecture Notes in Computer Science*, 393–395, Springer Berlin Heidelberg, (2013).

[13] Sooha Park Lee, Jeremy B Badler, and Norman I Badler, 'Eyes alive', in *ACM Transactions on Graphics (TOG)*, volume 21, pp. 637–644. ACM, (2002).

[14] David McNeill, *Hand and mind: What gestures reveal about thought*, University of Chicago Press, 1992.

[15] Raveesh Meena, Kristiina Jokinen, and Graham Wilcock, 'Integration of gestures and speech in human-robot interaction', in *IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 673–678. IEEE, (2012).

[16] Kyung-Geune Oh, Chan-Yul Jung, Yong-Gyu Lee, and Seung-Jong Kim, 'Real-time lip synchronization between text-to-speech (tts) system and robot mouth', in *RO-MAN, 2010 IEEE*, pp. 620–625. IEEE, (2010).

[17] Jelle Saldien, Bram Vanderborght, Kristof Goris, Michael Van Damme, and Dirk Lefeber, 'A motion system for social and animated robots', *International Journal of Advanced Robotic Systems*, **11**, (2014).

[18] Maha Salem, Friederike Eyssel, Katharina Rohlfing, Stefan Kopp, and Frank Joublin, 'To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability', *International Journal of Social Robotics*, **5**(3), 313–323, (2013).

[19] Brian Scassellati, Henny Admoni, and Maja Mataric, 'Robots for use in autism research', *Annual review of biomedical engineering*, **14**, 275–294, (2012).

[20] R. Simut, C. Pop, j. Vanderfaeillie, D. Lefeber, and B. Vanderborght. Trends and future of social robots for asd therapies: potential and limits in interaction, 2012.

[21] Adriana Tapus, Mataric Maja, and Brian Scassellatti, 'The grand challenges in socially assistive robotics', *IEEE Robotics and Automation Magazine*, **14**(1), N–A, (2007).

[22] S. Thill, C. A. Pop, T. Belpaeme, T. Ziemke, and B.ram Vanderborght, 'Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook', *Paladyn*, **3**(4), 209–217, (2012).

[23] Greet Van de Perre, Michaël Van Damme, Dirk Lefeber, and Bram Vanderborght, 'Development of a generic method to generate upper-body emotional expressions for different social robots', *Advanced Robotics*, **29**(9), 597–609, (2015).

[24] James Wilson and Daniel Rosenberg, 'Rapid prototyping for user interface design', *Handbook of Human-Computer Interaction*, **39**, 859–873., (1988).

[25] Seung-Kook Yun and Ambarish Goswami, 'Hardware experiments of humanoid robot safe fall using aldebaran nao', in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 71–78. IEEE, (2012).