Attentional Top-down Regulations in a Situated Human-Robot Dialogue

Riccardo Caccavale, Alberto Finzi, Lorenzo Lucignano, Silvia Rossi, Mariacarla Staffa
DIETI, Università degli Studi di Napoli “Federico II”
via Claudio, 80125 Napoli, Italy
{ri.caccavale,alberto.finzi,lorenzo.lucignano,silvia.rossi,maricarla.staffa}@unina.it

ABSTRACT

We propose a framework where the human-robot interaction is modeled as a multimodal dialogue which is regulated by an attentional system that guides the robot towards the execution of structured tasks. Specifically, we propose an approach where the dialogue between the human and the robot is represented as a POMDP, while the associated dialogue policy is enhanced by top-down attentional mechanisms that provides contextual and task-related contents. We introduce a simple case study that illustrates the system at work in different conditions considering top-down regulations and dialogue flows in synergic and conflicting situations.

Categories and Subject Descriptors

I.2.9 [**Robotics**]: [Operator interfaces, Commercial robots and applications]; I.2.8 [**Artificial Intelligence**]: [Plan execution, formation, and generation]

Keywords

Attentional System, Cognitive Control, Dialogue Manager

1. INTRODUCTION

Attentional regulation and dialogue management can play a crucial role in social robotics and human-robot interaction [4]. Indeed, a natural and effective interaction between humans and robots can be modeled as a multimodal dialogue flow, involving speech, gaze orientation, gestures, while attentional mechanisms can be used to orient and focus the robotic (and the human) perceptive and cognitive processes during the interaction. Some authors addressed these issues considering visual attention during human-robot conversation [16, 13, 18] used to detect the human to interact with or the task to be executed. Other authors mainly focused on joint attention and perspective taking methods for human-robot interaction [5, 23]. Differently from these approaches, which are mainly concerned with visual attention only, in this paper we focus on executive attentional mechanisms regulating the human-robot dialogical interaction. More specifically, we aim at defining a system that can manage and regulate the multimodal dialogue between the human and the robot by exploiting top-down and bottom-up attentional regulations. Inspired by attention and cognitive control literature in psychology and neuroscience [19, 9], we assume that the attentional influence can be driven by both higher level tasks/templates (top-down) and external/internal stimuli (bottom-up). In this context, the role of attentional mechanisms is to orchestrate multiple processes, at different levels of abstraction, possibly in conflicts. In this paper, we propose a multimodal real-time HRI system integrating a dialogue manager and a hybrid cognitive control architecture. Following the approach of [12], the dialogue between the human and the robot is modeled as a Partially Observable Markov Decision Process (POMDP) that can capture the inherent ambiguity of the situated communication. In the HRI framework proposed in this paper, the generated dialogue policy provides an interaction multimodal template (involving not only speech, but also gestures, gaze directions, etc.) which can be instantiated and continuously adjusted with respect to the environmental and the operative context by the attentional system. The cognitive control cycle modulates and polarizes the robot execution by enhancing the attentional processes which are aligned with the operative (top-down) and environmental (bottom-up) state, while inhibiting the ones which are not coherent. In order to illustrate the system at work, we introduce a case study where the human and the robot have to interact in simple pick, delivery, and place tasks. In this context, we show the system behavior during complementary and conflicting tasks.

HRI 2014 Bielefeld, Germany

2. THE HRI FRAMEWORK

The HRI architecture proposed in this paper is depicted in Figure 1, the main components of the system will be explained below.

2.1 System Architecture

The cognitive control cycle proposed in this paper involves three main modules: a *behavior pool* (BP), a *working memory* (WM) and a *long term memory* (LTM).

The BP contains a set of behaviors which may contribute to the execution of a complex cognitive task. The WM contains a representation of the current executive state and a representation of the tasks which are in the attentional focus of the system. These include all the tasks the system is executing or willing to execute. Finally, the LTM is a repository which contains the definition of all the tasks available

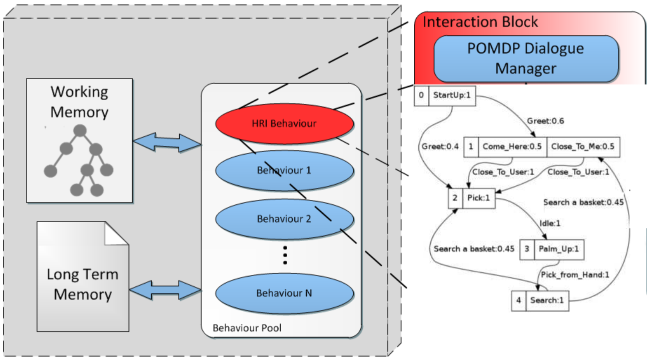


Figure 1: The global Framework

to the system. The cognitive control cycle is managed by a special behavior called *alive* which continuously updates the behavior pool and the working memory exploiting the task definitions provided by the LTM. This process will be better detailed below.

2.1.1 Attentional behaviors.

We assume that each behavior of the BP is structured as in [6]. Specifically, following a schema theory approach, a behavior is composed of a Perceptual Schema (PS) [1], which elaborates sensor data, a Motor Schema (MS), producing the pattern of motor actions, and a control mechanism, based on a combination of a clock and a releaser (see Figure 2). In particular, the releaser enables/disables the activation of the MS, while the clock regulates the behavioral arousal, that is, the sensory sampling rate and the frequency of behavior activations. This regulation represents our bottom-up attentional mechanism, indeed it tunes the resolution at which a behavior is monitored and controlled. For additional details about this model we refer the reader to [6]. Differently from [6], in the system proposed in this work the clock frequency can be modulated either by the perceptual stimuli (bottom-up) or by the executive state of the system exploiting the structures that are represented in the WM (top-down). Analogously, we introduce an additional external releasing mechanism that can enable the behavior activations depending on the executive state of the system.

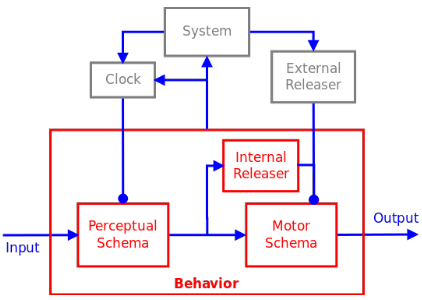


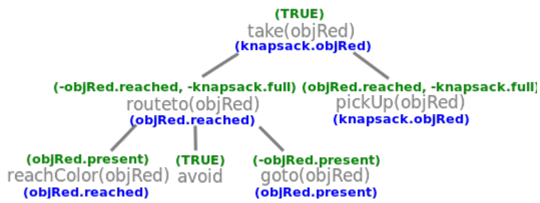
Figure 2: Schema theory representation of an attentional behavior.

2.1.2 Working Memory.

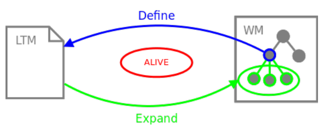
The WM is a critical element of the system because it maintains the executive state and the structure of the tasks in the attentional focus of the system. In our system, the tasks in the WM are represented as a unbalanced tree [17, 7, 8, 2] (see Figure 3) enhanced with additional information about the behavior execution (clock frequency, releaser status, variables, etc.). This herarchical structure follows a typical representation sheared by both artificial and biological models of tasks [20, 17, 7, 8, 2]. Each node of the WM can be classified in two categories:

- *Concrete*, if it represents an instance of an attentional behavior of the system (e.g. *pickUp*(*objRed*) in Figure 3).
- *Abstract*, if it represents only a *chunk* [14] which may be hierarchically decomposed in a different subtasks (e.g. *take*(*objRed*) in Figure 3) .

In this context, our cognitive cycle exploits the WM as follows. Initially, we assume a set of behaviors allocated to manage the basic system activities (e.g. *alive*, *interaction block*, etc.). Each behavior in the BP can affect the WM by inserting new nodes. For example, if the *interaction block* allocates a *take*(*objRed*) as a consequence of a human request, then *alive* (which is periodically activated to check for new nodes at each clock tic) will try to expand *take*(*objRed*) (see Figure 3) allocating other nodes as specified in the LTM. LTM contains production rules (see Figure 3-(b)). Analogously to [11], these represent hierarchical definitions of the available behaviors. When a concrete node is allocated in the WM, the associated behavior is awakened by *alive*. The tree structure of WM, is also endowed with an *external releaser* (ER) for each node. These ERs (green in Figure 3-(a)) are boolean expressions that represent guards to be satisfied in order to enable the execution of a behavior. Therefore, in order to activate a behavior both its ER and the ERs of the ancestors must be satisfied besides the internal releaser. Finally, a node could be also provided with a goal, which is achieved after the completion of the behavior (*teleological* nodes).



(a) Hierarchical task



(b) Subtree expansion

Figure 3: (a) Hierarchical task in the WM; (b) Definition/expansion cycle.

2.1.3 Emphasis.

The control cycle described above connects the execution of multiple behaviors with the hierarchical task structure provided in the WM, however, no explicit mechanism is provided to avoid conflicts or erratic activities. For instance, many behaviors may access to a single resource in WM, generating a *crosstalk interference* [15]. These conflicts can be either prohibited by construction or solved by means of an evaluation function [3]. In our architecture, we follow the latter approach introducing a function that we call *emphasis*. This function provides a modulation mechanism which is obtained as the integration of two types of influences on the concrete behaviors:

- *Frequency* of the behavioral clock induced by the perceptual stimuli.

- *Magnitude* externally induced by heuristics or influenced by other behaviors.

The first one represents a bottom-up and self-induced attentional mechanism, while the second captures top-down and lateral influences. Therefore, the attentional state of each behavioral schema in the tree is defined by the couple (p_i, m_i) representing, respectively, the frequency by the clock period p_i and the magnitude m_i. The emphasis is defined as $e_i = m_i/p_i$. By default, the magnitude is set to 1 for each node in the tree; if a node changes its magnitude, this is inherited by all the child nodes. In our system, we introduce a heuristic that changes the magnitudes according to subgoal achievements. Namely, when a subgoal is accomplished, the parent emphasis of the node is increased by a constant value (which is then propagated to the child nodes). This mechanism induces a soft *teleological* drive towards the completion of the open subtasks. The emphasis affects both the adaptive clock and the output values. More specifically, the clock period is reduced by m_i (with $m_i \geq 1$), hence the updated period is $p'_i = 1/e_i$. As for the output, given a variable v (e.g. the velocities of the motors) in the WM which is affected by the output of a set of behaviors b_i, the emphasis is exploited to weight and combine these multiple contributions as follows: $v = \Sigma_i(e_i \times v_i)/\Sigma_i(e_i)$. These two effects of the emphasis (acceleration of the clock and modulation of the combined outputs) allows us to solve the conflicts in a smooth way: not only the emphasized behaviors provide more frequent updates, but also their contribution is amplified. Since the amplification is associated to a drive towards the goal accomplishment, the goal-oriented behaviors become dominant, in so overcoming behaviors contentions and decisional impasses (see Figure 4).

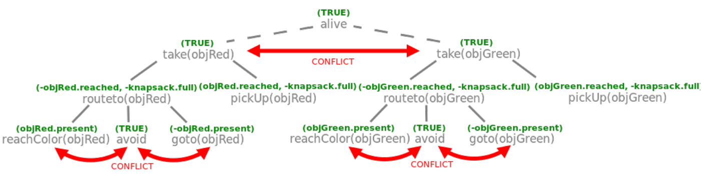


Figure 4: Example of conflicting tasks.

2.2 Dialogue Management System

The multimodal interaction block is appointed to recognize the multiple human commands and actions, such as utterances, gaze directions, gestures or body postures, and to provide an interpretation of user's intentions according to the dialogue context. It is integrated in the overall architecture as a special behavior and it is composed of three layers: the lower layer contains the *classifiers* of the single modalities; the middle layer, the *fusion engine*, performs a Support Vector Machine (SVM)-based late fusion and provides a context-free integration of the multiple inputs [21]; the upper layer, the *dialogue manager* [12], performs the coordination of the dialogue and accomplishes the semantic interpretation of the observations according to the context and the inner knowledge. The main feature of such structure is that the results of each layer are N-best lists of possible interpretations, which are fed to the next layer in order to solve in cascade the ambiguities at the upper layers of the system. The dialogue manager is the upper layer of the interaction block that provides the interaction policy depending on the iteration model. The dialogue models are provided as graph-based specifications. An example of dialogue graph describing a simple interaction scenario is shown in Figure 5. Our approach is to represent the dialogue flow as a POMDP (Partially Observable Markov Decision Process) to cast the inherent *ambiguity* in order to account for noise on the channels, misunderstanding of human actions or commands, multiple interpretations of a particular observation or non-deterministic effects of robot actions. In this context, the solution of the POMDP is a robust dialogue strategy off-line generated for that interaction model.

More specifically, the POMDP model used in our system is a multimodal version of the one proposed by [24, 22]. In this framework, the dialogue is represented as a tuple $(S, A_m, T(\cdot,\cdot,\cdot), r(\cdot,\cdot), O, Z(\cdot,\cdot), b_0)$. $S = S_{flow} \times S_{node} \times A_u$ is the set of states. A state is a triple $s =< s_{flow}, s_{node}, a_u >$, where s_{flow} is the identifier (ID) of a dialogue flow, a possible branch of the conversation; s_{node} is the ID of a situation which may occur in a dialogue flow; a_u is the last observed user action. A_m is the set of machine actions, containing the execution actions, which are the response to the user's commands or activities, and the control actions, which are useful to retrieve confirmation or decision by the user. $T(s', s, a_m)$ is the transition probability $P(s'|s, a_m)$ and R is the reward function $R(s, a_m) \in \mathbb{R}$. O is the set of the observations, which are the N-best lists of results provided by the fusion engine $\bar{o} = [< a_u^1, p_1 > \ldots < a_u^n, p_n >]$, where $p_i = P(a_u^i|\bar{o})$, $a_u^i \in A_u$. Finally, $Z(s, \bar{o})$ is the observation probability $P(\bar{o}|s)$ and b_0 is the initial belief state. Additional details can be found in [12].

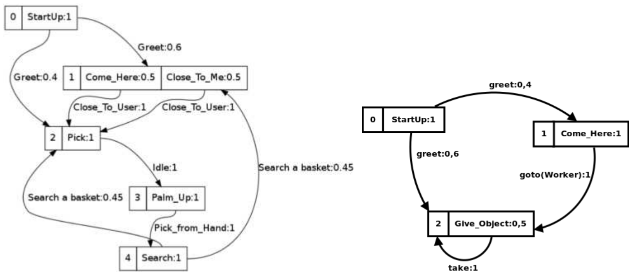


Figure 5: Excerpts extracted from dialogue models: (left) node 1 has the two possible interpretations "Come Here" and "Close to Me". The robot action is to go close to the human from where, in the node 2, the robot expects that user asks to pick something; (right) Simple dialogue flow involving "Come Here" and "Give Object" .

The dialogue policy generated as a solution of the POMDP provides a machine state a_m for each belief state of the dialogue. This machine action is then associated with a task to be allocated in WM whose execution is modulated by top-down and bottom-up attentional mechanisms. In this way, the machine action in the dialogue policy can be instantiated with contextual and task-related subtasks and arguments; moreover, its execution can be regulated by the associated top-down attentional mechanisms.

3. CASE STUDY

In this section, we discuss the system behavior considering as a case study a mobile robot that interacts with humans in a lab scenario. A representation of the environment is illustrated in Figure 6(down-right).

3.1 HRI scenario.

The robot will share the workspace with several users, which can interact with the robot to achieve some tasks such as picking or placing objects like bottles, or carrying paper sheets to other users. The robotic platform setting is the following: Pioneer 3 DX mobile robot provided with ultrasonic sensors and a gripper; RGB-D camera for users and gesture recognition and a High Definition camera for object detection; a microphone and a speech synthesizer; a Cyton 7DOF Pioneer Arm. The users can interact with the robot by speaking or using gestures or body movements, while the robot has a list of user dialogue models describing possible patterns of commands or movements. Each gesture is linked to one or more meanings, so ambiguities are possible. The meaning can be disambiguated according to the dialogue context. On the other hand, some user's acts are not explicit commands, therefore the system should interpret the human's intention supporting the human activity with a proactive behavior. Furthermore, this scenario offers a wide variety of situations for testing the ability of the proposed framework in managing multiple requests and in solving the associated conflicts (pick different objects). In this case, we assume that the robot can pick up an object at a time, but it can carry a maximum of two objects.

3.2 Experimental Results.

In this scenario, we can consider cases where the residual ambiguity in the dialogue policy and the associated decision conflicts should be resolved by the top down and bottom up attentional influences. For instance, if the request interpreted from the dialogue is a generic *take* (without an explicit reference to the object to be taken) and a green and a red object are perceived by the robot during the navigation, the system should decide which object to take. In this case, the perceived affordances associated with the two detected objects can directly elicit two instances of a *take* task to be allocated as schemata in the WM (e.g., $take(objRed)$, $take(objGreen)$). These schemata are then decomposed in two subschemata (see Figure 4) representing the chunks associated with the task: reach the object, pick it up, and give it to the human. In this way, these schemata/subschemata enter into the attentional focus of the robot along with the perceived objects and can be suitably top-down and bottom-up aroused. For instance, in Figure 6 (up) we can observe that, once a first red object is perceived by the robot, the $take(objRed)$ task is bottom-up aroused by the activations of $reachColor(red)$ (from 1 to 30) that instantiates the $routeto(objRed)$ in the WM. After 15 seconds the robot detects also a green object, therefore a decision conflict arises. However, in this case the robot heads towards the read object as an effect of the $reachColor(objRed)$ dominant activations (bottom-up influence) with respect to $reachColor(objGreen)$ (see in Figure 6 (down-left)). Once the red object has been reached, the subtask can be accomplished by the $pickUp(red)$ behavior. At this point the frequency of $take(objRed)$ is relaxed because a new subtask $give(objRed)$ is activated. This behavior receives the emphasis (top-down influence) from the $take(objRed)$ that drives $give(objRed)$ towards the goal accomplishment. In Figure 6 (up), from time 30 to 55 we can see the restriction of the period (frequency enhancement) illustrating the modulation due to the bottom-up influence (dotted red line) and the one that takes into account also the effect of the top-down emphasis (continuous red line).

In Table 1, we illustrate 10 runs where the robot (given a simple dialogue model and its current belief state [12]), interprets as a machine action to be executed an unreferenced *take*. To assess the system behavior in this ambiguous situation, we consider two scenarios: in the first one we have two objects to be taken (red and greed in Table 1, left); in a second scenario we have three objects (red, green, and yellow in Table 1, right). For each scenario we report the executed sequence of tasks and the time needed to accomplish the goal (minutes). The *executed sequence* illustrates the subtasks sequence chosen by system (here *Red*, *Greed*, etc. is an abbreviation for, respectively, reach and pick the object red, reach and pick the object green etc., while *Give* represents the delivery action that ends the task). A maximum of 10 minutes was provided for each run. In order to test the system in the ability of conflict resolution and flexible execution of multiple tasks, we allow the robot to collect two items before the delivery. For instance, in the sequence "*Red Green Give*" $take(objRed)$ and $take(ObjGreen)$ are interleaved, hence the robot first picks the red object, then it picks the green one, and finally it delivers the two objects to the human; in other cases, the task are sequentialized (e.g., in "*Red Give Green Give*"). Notice that the parallel or sequential execution of the task is left to the system decisions and depends on the attentional mechanisms and environmental context.

The results in Table 1 show that the system is always able to accomplish the goal, and when there is an opportunity it can also interleave the execution of the tasks (6 times and 7 times in the first and the second scenario respectively), and, as expected, when this happens the temporal performance is enhanced. In order to better assess the temporal performance, in Table 2 we also report the average and the std of the values collected after the execution of 10 *take* tasks where the referenced object is provided (e.g., $take(green)$). By comparing the average values at the end of Table 1 with the values in Table 2 we can observe that the mean time needed to accomplish the ambiguous requests is comparable with the mean time needed to achieve the tasks where the reference is explicitly defined. This seems to suggest that the conflict resolution mechanism is not time consuming and effective in managing the impasses.

These preliminary results have been provided mainly to show the system behavior in the presence of ambiguities and conflicts during the interaction. Additional and more intensive experimental testing is needed to assess the flexibility and the effectiveness of the platform. The proposed attentional mechanisms are here mainly elicited by the detec-

EXECUTION TIME			
Task Sequence	Time (min)	Task Sequence	Time (min)
TakeRed - TakeGreen		TakeRed - TakeGreen - TakeYellow	
Red Green Give	4.5	Red Green Give Yellow Give	9.19
Green Give Red Give	7.11	Green Give Red Give Yellow Give	8.19
Green Give Red Give	8.04	Red Green Give Yellow Give	7.21
Green Give Red Give	7.14	Yellow Give Green Give Red Give	9.08
Red Green Give	3.53	Yellow Green Give Red Give	7.28
Green Red Give	3.50	Red Green Give Yellow Give	6.41
Red Green Give	4.19	Red Green Give Yellow Give	7.02
Green Give Red Give	6.04	Red Green Give Yellow Give	7.05
Red Green Give	4.48	Yellow Give Green Give Red Give	9.43
Green Red Give	6.26	Red Green Give Yellow Give	8.48
AVG	STD	AVG	STD
5.48	1.64	7.93	1,07

Table 1: Execution time of generic *take* in different contexts.

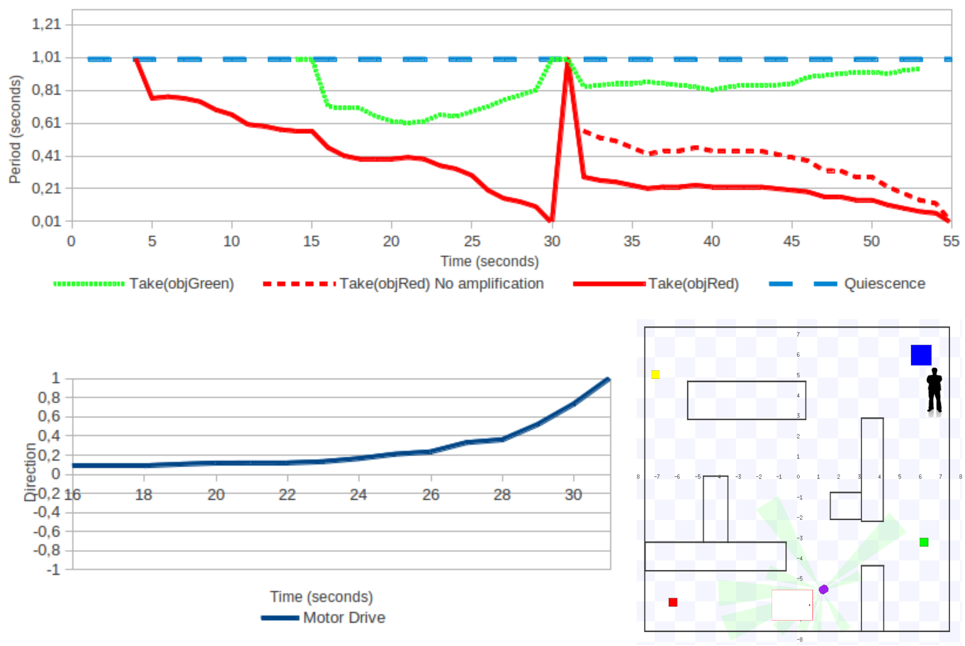


Figure 6: Period modulation (up) and motor drive (down-left) during a conflicting situation in a lab scenario (down-right): $take(objRed)$ is amplified, hence the frequency and the outputs are enhanced (up) driving the robot towards the red target (down-left).

tion of gestures, speech, objects, colors however, additional, and more sophisticated mechanism (e.g. gaze detection and joint attention [23]) can be easily incorporated in this framework. Analogously, the interaction fluency with the human remains to be assessed, our working hypothesis is that the attentional regulation mechanisms presented in this work associated with the implicit communication in a multimodal situated dialogue can enhance the overall interaction flexibility, effectiveness, and naturalness [12, 10]. More extensive tests are left as a future work.

4. CONCLUSIONS

In this paper we presented a novel human-robot interaction system that combines a dialogue system with top-down and bottom-up attentional modulations. The proposed system allows to contextualize the dialogue flow in the operational situation and solve ambiguous communications. We described the proposed HRI system architecture introducing a case study used to illustrate the system behavior in different scenarios. In particular, we have shown how both

EXECUTION TIME (min)					
Take-Red		Take-Green		Take-Yellow	
avg	std	avg	std	avg	std
3.99	0.28	1.48	0.36	2.04	0.27

Table 2: Execution time of the specific *take*.

bottom-up and top-down attentional modulations allow the system to solve decisional impasses driving the system towards the task accomplishment. We presented preliminary empirical results, a deeper validation of the HRI performance is left as a future work.

Acknowledgments

The research leading to these results has been supported by the SAPHARI project (FP7 under grant agreement ICT-287513) and the SHERPA project (FP7 under grant agreement ICT-600958).

5. REFERENCES

[1] M. A. Arbib, E. J. Conklin, and J. A. C. Hill. *From schema theory to language.* Oxford University Press, 1987.

[2] M. Botvinick and D. C. Plaut. Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111:395–429, 2004.

[3] M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen. Conflict monitoring and cognitive control. *Psychological review*, 108(3):624, 2001.

[4] C. Breazeal. *Designing Sociable Robots.* MIT Press, 2002.

[5] C. Breazeal, M. Berlin, A. G. Brooks, J. Gray, and A. L. Thomaz. Using perspective taking to learn from ambiguous demonstrations. *Robotics and Autonomous Systems*, 54(5):385–393, 2006.

[6] E. Burattini, A. Finzi, S. Rossi, and M. C. Staffa. Attentive monitoring and adaptive control in cognitive robotics. In *Cognitive Robotics*, Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2010.

[7] R. Cooper and T. Shallice. Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, 17:297–338, 2000.

[8] R. P. Cooper and T. Shallice. Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, 113(4):887–916, 2006.

[9] A. Gating and E. versus Exploration. A systems-level perspective on attention and cognitive control. *Cognitive neuroscience of attention*, page 71, 2004.

[10] S. Iengo, A. Origlia, M. Staffa, and A. Finzi. Attentional and emotional regulation in human-robot interaction. In *RO-MAN, 2012 IEEE*, pages 1135–1140. IEEE, 2012.

[11] J. E. Laird, A. Newell, and P. S. Rosenbloom. Soar: An architecture for general intelligence. *Artificial intelligence*, 33(1):1–64, 1987.

[12] L. Lucignano, F. Cutugno, S. Rossi, and A. Finzi. A dialogue system for multimodal human-robot interaction. In *Proceedings of the 15th ACM on International conference on multimodal interaction*, pages 197–204, 2013.

[13] P. McGuire, J. Fritsch, J. Steil, F. Rothling, G. Fink, S. Wachsmuth, G. Sagerer, and H. Ritter. Multi-modal human-machine communication for instructing robot grasping tasks. In *Proc. of IROS 2002*, volume 2, pages 1082–1088, 2002.

[14] G. A. Miller. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, 63:81–97, 1956.

[15] M. C. Mozer and M. Sitton. Computational modeling of spatial attention. *Attention*, 9:341–393, 1998.

[16] T. Muller and A. Knoll. Attention driven visual processing for an interactive dialog robot. In *Proc. of the 2009 ACM symp. on Applied Computing*, pages 1151–1155, 2009.

[17] D. A. Norman and T. Shallice. Attention to action: Willed and automatic control of behavior. In *Consciousness and self-regulation: Advances in research and theory*, volume 4, pages 1–18. 1986.

[18] H. G. Okuno, K. Nakadai, K. Hidai, H. Mizoguchi, and H. Kitano. Human-robot interaction through real-time auditory and visual multiple-talker tracking. In *Proc. of IROS 2001*, volume 3, pages 1402–1409. IEEE, 2001.

[19] M. I. Posner and C. R. R. Snyder. Attention and cognitive control. In *Information Processing and Cognition*, pages 55–85. 1975.

[20] D. A. Rosenbaum, R. G. Cohen, S. A. Jax, D. J. Weiss, and R. van der Wel. The problem of serial order in behavior: Lashley's legacy. 26(4):525–554.

[21] S. Rossi, E. Leone, M. Fiore, A. Finzi, and F. Cutugno. An extensible architecture for robust multimodal human-robot communication. In *In Proc. of IROS-2013*, 2013 (to appear).

[22] N. Roy, J. Pineau, and S. Thrun. Spoken dialogue management using probabilistic reasoning. In *Proceedings of ACL '00*, pages 93–100, 2000.

[23] B. Scassellati. Imitation and mechanisms of joint attention: a developmental structure for building social skills on a humanoid robot. In *Computation for metaphors, analogy, and agents*, pages 176–195. 1999.

[24] S. Young, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson, and K. Yu. The hidden information state model: A practical framework for pomdp-based spoken dialogue management. *Computer Speech and Lang.*, 24(2):150 – 174, 2010.