LifeBots I: Building the Software Infrastructure for Supporting Lifelong Technologies

Antonio Bandera[1], Juan P. Bandera[1], Pablo Bustos[2], Fernando Férnandez[3], Angel García-Olaya[3(✉)], Javier García-Polo[3], Ismael García-Varea[4], Luis J. Manso[2], Rebeca Marfil[1], Jesús Martínez-Gómez[4], Pedro Núñez[2], Jose M. Perez-Lorenzo[5], Pedro Reche-Lopez[5], Cristina Romero-González[4], and Raquel Viciana-Abad[5]

[1] Universidad de Málaga, 29071 Málaga, Spain
{ajbandera,jpbandera,rebeca}@uma.es
[2] RoboLab, Universidad de Extremadura, 10003 Cáceres, Spain
{pbustos,ljmanso,pnuntru}@unex.es
[3] Universidad Carlos III de Madrid, 28911 Leganés, Spain
{ffernand,agolaya,fjgpolo}@inf.uc3m.es
[4] Universidad de Castilla-La Mancha, 02071 Albacete, Spain
{ismael.garcia,jesus.martinez,cristina.rgonzalez}@uclm.es
[5] Universidad de Jaén, EPS Linares, 23700 Linares, Spain
{jmperez,pjreche,rviciana}@ujaen.es

Abstract. The goal of the LifeBots project is the study and development of long-life mechanisms that facilitate and improve the integration of robotics platforms in smart homes to support elder and handicapped people. Specifically the system aims to design, build and validate an assistive ecosystem formed by a person living in a smart home with a social robot as her main interface to a gentler habitat. Achieving this goal requires the use and integration of different technologies and research areas, but also the development of the mechanisms in charge of providing an unified, pro-active response to the user's needs. This paper describes some of the mechanisms implemented within the cognitive robotics architecture CORTEX that integrates deliberative and reactive agents through a common understanding and internalizing of the outer reality, which materializes in a shared representation derived from a formal graph grammar.

Keywords: Social robots · Assistive robotics · Smart homes · Software architectures · Lifelong technologies

1 Introduction

While the economic benefits of robotics in industry are already clear, it is expected that their inclusion in everyday life will have a tremendous impact. Personal service constitutes a promising segment of the new market for robotics, which will probably fully boom in the next 10 to 20 years [5]. For that date, it is expected that many, if not all, domestic activities related to personal assistance will be covered and provided by 'robotic' applications. Robot technologies can

A. Ollero et al. (eds.), *ROBOT 2017: Third Iberian Robotics Conference*, Advances in Intelligent Systems and Computing 693, https://doi.org/10.1007/978-3-319-70833-1_32

also take advantage of the growing emergence of ambient intelligence, ubiquitous computing, sensor networks and wireless networking technologies. In South Korea, where practically all households have broadband Internet, the government hopes to put a networked robot in every household by 2020. However, to access these new markets and to be competitive, robots have to be dependable, smarter and able to work in closer collaboration with humans[1]. Although researchers must be cautioned against unrealistic expectations about robot developments, the idea of incorporating them into indoor environments such as (smart) homes, hospitals or public buildings, is a topic encouraged by public and private organisms. Within a global scenario where the average life expectancy at birth on the more developed countries is currently over 80 years, with three quarters of these older people living alone or only with their spouse, the need for assistance platforms that enable individuals with physical limitations and disabilities to continue living independently in their own homes is soaring. The delay or elimination (if possible) of the need for moving an individual to a managed care centre significantly decreases the cost and burden on the individual, family, and health care providers. It also greatly diminishes the likelihood of isolation, depression, and shortened lifespan.

In the context of elder people assisted living, smart homes can be seen as residential houses equipped with sensors and automated devices, whose goal is to deliver care and monitoring. In such settings, socially assistive robotics has arisen not only as a main element to support elder or handicapped people in day-life activities, or as a therapeutic robot, but also as a crucial interface to the person living in the smart home [6]. Assistance must be defined in the long-term, and it must attempt to balance the immediate specific needs of the user with the long-term effects that the robot's and assistance technologies can potentially have on the user's developmental trajectory. To achieve such long-term support, the robotic system must be able to acquire new concepts, adapt to new situations and learn behaviors to solve new tasks by exploiting its past experiences. The term Lifelong refers to the capability of a system to tackle different problems or tasks, throughout its life, with the aim of improving knowledge, skills and competences. The goal of lifelong learning is that the knowledge acquired in the development of past tasks, supports somehow the new ones. Therefore, solving new problems is not afforded from scratch, but supported from previous processes, being such tasks as specific or generic as required. This paper describes the instantiation of the CORTEX architecture [1] within the LifeBots project. The inner representation of the current and previous experiences, as well as the autonomy of the task-dependent software modules within CORTEX, will be the basis for adapting to new goals, user's preferences or environmental issues. Specifically, this paper describes our proposal for merging deliberative and reactive software agents, all of them sharing information through a common representation space.

[1] http://cordis.europa.eu/ictresults.

The paper is structured as follows: The proposed demonstrators of the system are described in Sect. 2. Next, Sect. 3 shows the system architecture and the current state of some of its more significant components. Experimental results are shown in Sect. 4. The paper finishes with the conclusions and future work in Sect. 5.

2 Use Cases

Within the LifeBots project, the system ability to provide long-life support in smart homes and the successful integration of all the components of the CORTEX architecture (see Sect. 3) will be shown by means of two paradigmatic use cases. These use cases will validate the platform as an integrated assistive ecosystem:

- *bringMe (x)*: This use case specifies a situation in which a person asks the robot to bring her some daily object. The boundary conditions here are intentionally loose, since we are looking for a robust, reliable and efficient capability rather than for a one-time, hard to repeat test. Figure 1 provides a schematic storyboard. It should be noted the intense relationship among a large set of functionalities. All partial goals of the project are tested in this use case and a working integration must be running to complete the task. Successive iterations on this activity will validate different lifelong learning technologies, and quantitative measures of improvement will be obtained from the network of sensors and the internal recordings of the robot.
- *letsMove ()*: The robot takes the leading role to stimulate the person into a set of exercises. Exercises will be proposed by medical personnel and will be integrated in the everyday activities that the robot will fulfill in the apartment.

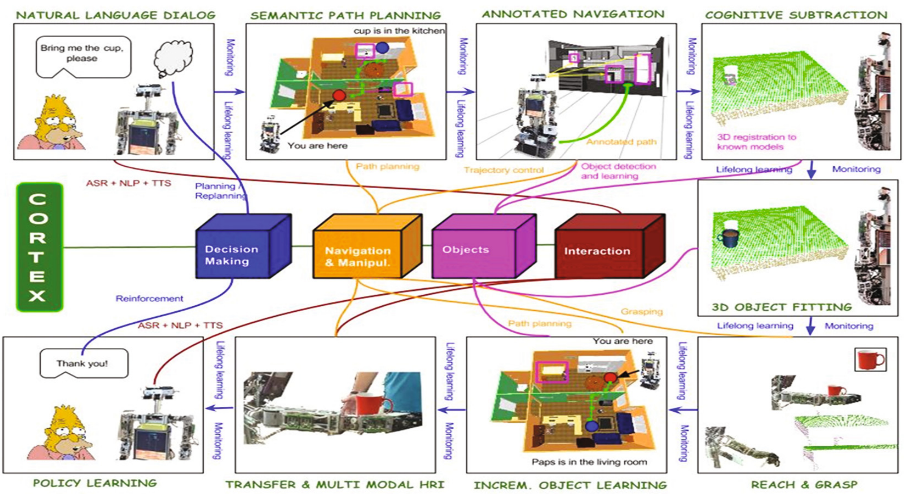

Fig. 1. A schematic view of the BringMe(x) use case

The results of performing these exercises during extended periods of time will be monitored and used to improve the existing exercises routines. This use case will pursue a mixed combination of physical and cognitive exercises in which the role of the robot can change, from leader to follower in imitation games. Health oriented activities like this one are to be interleaved with everyday domestic tasks in social and service robots, creating a pro-active, health seeking environment, for people with limited autonomy.

3 System Architecture

Within the LifeBots project, physical and computational will be intimately linked together. From a physical point of view, the system is composed of the Shelly robot [12] and a series of sensors and actuators making up the smart home infrastructure. The ability of these resources to adapt to the home residents' needs will rest upon embedded computation and communication, and real-time decision making and perception. From a conceptual point of view, these features will be provided by the CORTEX cognitive architecture [1]. The CORTEX architecture is organized around a graph-structured internalization of geometric entities or measures, and symbolic concepts. The world state is perceived by a collection of software agents, which are continuously updating this representation. Simultaneously, these agents are in charge of solving specific functionalities, which can range from reactive tasks to more deliberative ones, but all agents share their individual perception of the outer world by accessing this common representation. The novelty that CORTEX introduces is that the sequence of updated representations are actually sentences of a formal domain graph grammar. All terminal and non-terminal symbols in the grammar are shared by all agents, providing a common communication ground. All modifications of the terminal symbols have to agree with the assigned type and all structural modifications have to come from a valid derivation from the rules of the grammar. Thus, the system as whole evolves writing sentences of its domain grammar. The world is always interpreted as being in a valid sentence of the grammar, and agents do their best to write that specific sentence. See [9] for more detailed description.

Surrounding the inner world there are agents that are able to solve how to navigate to a given goal, perceive the facial expression of a person, update the battery level, and so on. For achieving these tasks, agents need to tie to other software components. Each network of components solving a specific task will be called a *compoNet*. Although there can exist activities that are solved by a single compoNet, it is important to note that the aim is that emergent, complex behaviors result from the interaction of several agents. Thus, it is possible to adapt the dialogue to the facial expression perceived on the user, or to navigate among people by taking into consideration that they are not simple obstacles. The existence of the shared representation avoids to solve the first task by integrating a facial perception module within the agent in charge of the human-robot dialogue, or the second one by integrating a pedestrians' detector and tracker within the

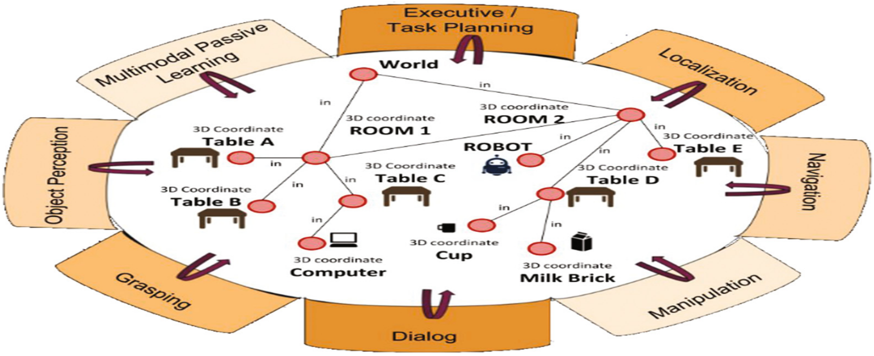

Fig. 2. CORTEX: A hybrid, graphical representation surrounded by a collection of task-related agents

agent responsible of the reactive navigation of the robot. Thus, it allows the parallel execution of tasks and avoids the non-reactivity of those approaches where perception is totally commanded by the deliberative course of action [13]. Within CORTEX, the agents in charge of solving deliberative tasks are also tied to the inner world as the rest of agents. In fact, they constitute a clear example of the aforementioned idea, as they try to approach the short-scale evolution of the inner world towards a 'desirable' state by means of deliberative actions (commands) (Fig. 2).

In the previous paragraph, we have intentionally referred to a short-scale intervention of the agents over the inner world. This is the current situation on CORTEX. However, one of the main objectives of the LifeBots project is to be able to engage users and environment within a large-scale interaction process, which will allow the system to autonomously self-adapt to changes and new situations. The ability of the inner world for augmenting its own capacities for providing an adequate tool for dealing with this new situation is one of the challenges of the LifeBots project. In its current form, the inner world is an efficient working memory. But episodic or semantic memories should be added. On the other hand, the procedural memory should allow the system to be more pro-active to solve situations that were addressed on the past. If the episodic or semantic modules will require to design new graphical representations that extend the current Inner world, the procedural memory - which is currently available within the decision making agent [8] - should be distributed within the agents. They are in charge of solving tasks, and they must have the mechanisms to improve their own activity.

Next section describes the main advances made within the current instantiation of CORTEX for achieving the objectives proposed by the LifeBots project. These objectives were summarized, from a practical point-of-view, on the use cases described at Sect. 2. Hence, we use the needs stated by the bringMe(x) use case, and illustrated on Fig. 1, for unfolding the current results. Future directions and efforts will be discussed at Sect. 5.

4 Experimental Results

The two use cases described in Sect. 2 need the design and development of complete and rich compoNets. Specifically, the brigMe(x) use case implies to endow on the robot the skills needed to:

- Interact with the end-user through verbal and non-verbal channels
- Navigate through an environment that could be populated by other people
- Generate a semantic map of the environment, as all recognized items could be useful for solving future plans
- Recognize and isolate specific targets
- Reach and grasp specific targets

Simultaneously, the software architecture should address additional tasks, related to the understanding and internalizing of the outer environment, and the close integration of deliberative and reactive modules. Next subsections provide details about how some of these mechanisms are currently running for solving this use case.

4.1 Social Navigation

Social and semantic navigation needs the interaction of different specific compoNets within CORTEX. These compoNets endow the robot with the ability for detecting objects in the path and updating the inner world accordingly. Additionally, the skill of detecting humans is also mandatory because robots need to know about people to get commands, avoid collisions and provide feedback. The final, and most important compoNet for social navigation, is the one implementing the navigation algorithms that allow robots to navigate from a point to another in a secure and social manner [12].

An overview of the proposal for social navigation is shown in Fig. 3. The global semantic path planner is a deliberative module whose aim is to choose

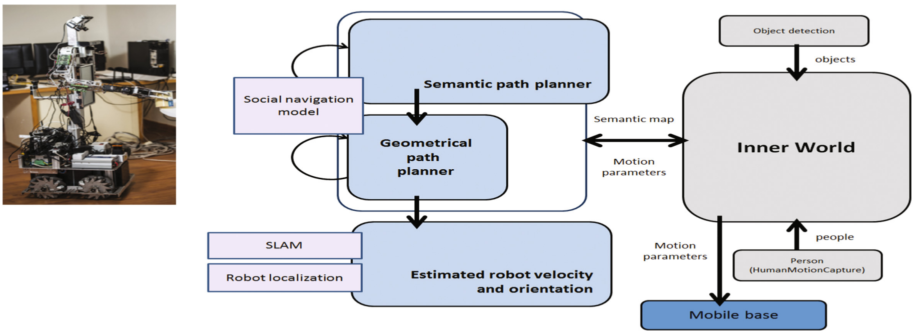

Fig. 3. The Shelly robot and a scheme of the modules involved on the social navigation skill

the optimal route, consisting of a list of way-points. These way-points are characterized by a set of labeled objects in the map that the robot should perceive. The geometrical path planner is in charge of moving the robot from one way-point to the next one. The path traced by both of them is affected by the social navigation model, which takes into account the people and objects detected and internalized on the inner world by other compoNets on the architecture. When it is necessary, the local (or global) route is re-planned [12]. Motion parameters are provided to the mobile base.

4.2 Semantic Mapping

Typically, we can find two types of representation in robotic mapping: metric and semantic. While metric maps refer to the specific positions the robot may reach in the environment, semantic representations incorporate additional information related to high-level descriptions of goals or objects of interest within the map. For example, a semantic map can help to determine the expected behavior of a robot based on its current semantic localization (e.g. a kitchen, a bedroom, a living room). Even though semantic and metric localization can be carried out separately, the explicit association between metric locations and semantic labels, by means of a common map, has proven to be an efficient way to manage the metric semantic gap. The semantic mapping agent can automatically generate semantic maps from sequences of unlabeled RGB-D robot acquisitions. The approach exploits the lexical annotation capabilities provided by previously trained Convolutional Neural Network (CNNs) classification models, given input RGB images. From these annotations, which have also associated a probability distribution over all different lexical labels, we cluster similar RGB images into the same semantic category. A keystone of the proposal is how to compute the similarity between sets of images using the lexical annotations (or more concretely the associated probability distribution) provided by the CNN classification model. In that sense, we use the Kolmogorov-Smirnov distance given that it encodes the maximum dissimilarity between two cumulative distributions. The depth information is used to obtain the global metric position of the robot by means of a RGB-D SLAM procedure. Regarding object recognition and localization, given a RGB-D image, we also propose to exploit pretrained CNN models to identify the objects based on the color information, and assign a 3D position based on the depth information of the image and the robot localized pose.

The proposal is validated using ViDRILO [10], a dataset consisting of sequences of RGB-D images acquired by a mobile robot within an indoor environment. Images in the dataset are annotated with the semantic category of the scene where they were acquired. The evaluation demonstrates that semantic maps can be successfully generated in a completely automatic way. Moreover, we have seen that the use of lexical annotations is not limited to generating semantic categories. These annotations also contribute to the descriptive capabilities of the generated maps. That is, each category can be described by means of its more representative lexical annotations, which improves the interpretability of the environment representation.

4.3 Searching and Grasping Objects

Relying on the semantic map created by the robot while navigating around the apartment, it is possible to answer queries about where an object is located or where is it more probable to be found. This information is used by the task-planner to narrow the search of a requested object. Once the object is in sight, the current action in the plan changes from *approachTheRoom* to *approachThatObject* forcing the *Navigation* agent to perform a servo visual control in cooperation with the ObjectDetector agent.

When the robot is positioned in front of the target object, the *Manipulation* agent gets the corresponding action in the plan and activated to reach and grasp the object. This agent is internally organized in a similar way to the *Navigation* agent [7]. It uses the idea of a trajectory that has to be planned and performed avoiding unexpected events. To avoid planning once and again similar trajectories, a memory of discrete task-space positions is built, holding the angular values needed to move the arm there. Before searching for a trajectory in this memory, it is compared with the point cloud obtained by an RGBD sensor pointing at the zone of interest. Those positions that intersect are tagged as forbidden in this memory, so the search for a trajectory will avoid those areas. The combination of this memory of reachable positions and the quick exclusion of occupied zones detected by the RGBD sensor provide a very fast way to approach objects. Grasping is directly tied to the known or recognizable shape of the object, and proceeds in a preprogrammed way.

4.4 Audio Spatial Cognition

Since hearing is a prominent sense for communication and socialization, it is desirable to strengthen the auditory capabilities of the platform to improve its social behavior. In this sense, in complex acoustical scenarios, where several audio sources can be present (e.g. simultaneous speakers or a speaker voice mixed with noisy sources), it is needed to localize and distinguish the individual sources. These scenarios are known as "the cocktail party problem" [3], and they can be resolved in a first step by the spatial localization of the sources. Besides, the development of audio spatial cognition capabilities not only may be useful for the interaction with human speakers, but also to incorporate acoustic features of the surroundings to the inner world model. Then, the system can work not only with auditory information but also with a multimodal approach by using other kind of information coming from cameras [14] or depth sensors. The acoustic features may contain information such as the localization of the sources or the percentage of time that the sources are active during a period of observation time. By this way, an auditory memory of the surroundings can be developed at different levels using different values for the observation time (e.g. a short-term analysis for more reactive behaviors, and long-term analysis for more deliberative ones). In this sense, some experiments have been carried out to test an unsupervised method for the lateral localization of simultaneous sources, with the aim of incorporating relevant audio spatial features to the robotic platform.

4.5 Planning and Decision Making

This module provides the decision making capabilities of the robot, including creation of high level plans, reasoning about goals, learning and adaptation to user preferences. It relies on Automated Planning (AP) and Reinforcement Learning techniques and uses the planning-learning-execution-monitoring PELEA architecture [13], encapsulated into the Deliberative *compoNet*. From the AP point of view we are currently modeling the use-cases in the Planning Domain Definition Language (PDDL) [11], a standard language used by most modern planners. PDDL uses predicate logic to model the actions the system can perform. Using this model, the information about the current state contained in the inner world, and the tasks the system must perform, the planner can generate a sequence of high-level actions to be executed. Using a declarative language the HighToLow module translates these actions to commands, to be annotated on the inner wold. The rest of agents on the architecture can then read these command and launch the adequate behaviors (e.g. say a sentence, go to a specific goal). As a consequence of the execution of these behaviors the agents update the inner world representation. Obviously, changes can also be the consequence of unexpected events. The monitoring module checks all these changes, which have been previously translated to high-level predicates by the LowToHigh module, and decides whether the plan is being executed correctly and the next action can be annotated on the inner world or whether something unexpected occurred and a new plan must be generated. Tasks to be performed can respond to explicit or implicit user requests (as in the bringMe(x) example) or can be externally generated by caregivers (as in the letsMove() case use). In both cases, preferences can be assigned among tasks, so most useful tasks are performed first.

5 Conclusions and Future Work

Social robots are usually designed thinking on a short-scale (immediate) adaptation to the user [6]. Thus, activities and skills are assumed to be invariable with time, and it is then possible to endow them on the robot in advance. This scenario is opened to non-desirable events (e.g. the person forgot to take medication), which cannot be considered as really 'unexpected' ones. The use of a deliberative planner for dealing with these situations does not differ from the use of a complex state machine. Symbolic and geometric levels can be strictly separated and it is possible to reason using symbolic concepts, emanate the corresponding commands to the task-dependent software modules on the architecture, and wait for a response/change on the world, always from an a priori know set [13].

However, our everyday scenarios are very dynamic, not only from the geometric perspective, but also (and mainly) from the temporal perspective. The collection of possible changes makes difficult to take all of them into account in advance. Within this snapshot, the idea of maintaining a different representation for the deliberative planner, separated from the one employed for the more reactive or situated modules, does not appear to be a good one. If the planner

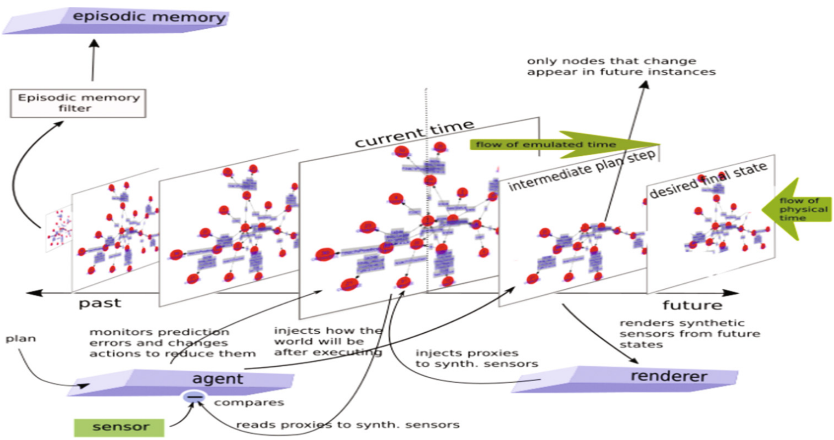

Fig. 4. Schematic view of the dynamic DSR

does not use this common, central representation, it could not monitor the evolution of the outer world. Or it will need a continuous translation from the mid-, low-level representation to the high-level one. Removing the division into deliberative and reactive levels, the goal of the CORTEX architecture is to maintain an integrated, dynamic multi-graph representation that can hold geometric or sensorial data and high-level symbols and predicates. This is the so-called Deep State Representation (DSR) [2]. The DSR allows all agents to share a global state, facilitating the coding of new skills or the adapting of the old ones.

But, in its current implementation, the DSR can only be employed as a short-scale, working memory. The need of storing further information on the representation has distorted this assertion, and the DSR currently merges long-term information, such as the one related with the internal organization of rooms and doors within the apartment, and real short-scale data, such as the current positions of people or objects. The idea of extracting from the short-term DSR those geometric or symbolic items that should be memorized for long periods of time will allow to reduce the size of the DSR, bounding it to an easy-to-use, quasi-constant volume. But this will also need the design of those software modules in charge of synchronizing long-term and working memories (e.g. for deciding that the sofa the robot is currently seeing (working memory) is the same one on the hall (long-term memory)). We are exploring the idea of creating specialized agents to deal with information from the past. These agents will initially capture spatial, episodic and semantic information. Spatial information would be built as a graph of places, with nodes representing locations and edges spatial predicates between them such as *connectsTo*, *insideOf*, etc. Once the robot leaves a place, the DSR information would be extracted by this agent and stored. Objects inside that space would also be stored keeping their spatial positions. Some forgetting algorithm would be used to eliminate some elements and keep the volume of the storage down to a certain size. The second memory agent will be in charge of episodic memory. It will create and store a time sequence of past events made of time frames with references to known objects and spaces. Finally, the semantic information agent would maintain a repository of facts or assertions built over the common language -objects, attributes and predicates- used and understood by all other agents. This agent will have inference capabilities as those provided by semantic reasoners [4]. The interaction of the three memory systems with the present and, more importantly, with the future, is commented below.

Other major problem of the current DSR is how to face the problem of managing previous experiences or emulating future courses of action. Future work within the LifeBots project should focus on evaluating solutions for this problem, as it is on the basis of the long-term adaptation. A plausible solution is to extend the DSR graph on the temporal scale. Figure 4 provides an schematic view of the dynamic DSR. Agents will now be linked to several temporal views of the inner world. Moving backward from the current time, compoNets should be endowed with the software functionality for extracting those subgraphs that allow them to solve/detect a specific situation. This procedure will build a procedural memory, local to each agent. Moving forward, they should be able to emulate and inject, on the DSR+t graph, the future state of the inner world. Conform this DSR+t representation approaches to the present time, the detailed of this future representation will be closer to the one associated to the present time. And the matching of both, predicted and present, representations will provide a cognitive feedback to the agents, allowing to launch corrective actions in advance, before the real problem will appear. The time scale employed for the DSR+t (intermediate plan step, Fig. 4) will not be the same for all concepts stored on the inner world. Moreover, the uncertainty associated to all predictions can be also different. The presence of an expected door on a estimated route can be located with precision in time and location. The need of recharging batteries could be anticipated in time but with a larger uncertainty.

Acknowledgment. This work has been partially funded by the European Union ECHORD++ project (FP7-ICT-601116) and the TIN2015-65686-C5 Spanish Ministerio de Economía y Competitividad projects and FEDER funds. Javier García is partially supported by the Comunidad de Madrid (Spain) funds under the project 2016-T2/TIC-1712. RoboLab is partially supported by the European project POPTEC EUROAGE 4E and by the Extremaduran Government under grant GR15120. We also want to acknowledge the Red de Agentes Físicos TIN2015-71693-REDT.

References

1. Bandera, A., Bandera, J., Bustos, P., García-Varea, I., Manso, L., Martínez-Gómez, J.: CORTEX: a new architecture for social robots. In: 2016 EUCognition Conference on Cognitive Robot Architectures, Vienna, Austria (2016)
2. Calderita Estévez, L.V.: Deep state representation: an unified internal representation for the robotics cognitive architecture CORTEX. Ph.D. thesis (2016)
3. Cherry, E.C.: Some experiments on the recognition of speech, with one and with two ears. J. Acoust. Soc. Am. **25**(5), 975–979 (1953)

4. Dentler, K., Cornet, R., Ten Teije, A., De Keizer, N.: Comparison of reasoners for large ontologies in the OWL 2 EL profile. Semant. Web **2**(2), 71–87 (2011). doi:10.3233/SW-2011-0034
5. EUrobotics: Robotics 2020 multi-annual roadmap for robotics in Europe. Report (2015). https://www.eu-robotics.net/cms/upload/downloads/ppp-documents/Multi-Annual_Roadmap2020_ICT-24_Rev_B_full.pdf
6. Gross, H., Schröter, C., Müller, S., Volkhardt, M., Einhorn, E., Bley, A., Martin, C., Langner, T., Merten, M.: Progress in developing a socially assistive mobile home robot companion for the elderly with mild cognitive impairment. In: IROS, pp. 2430–2437. IEEE (2011)
7. Haut, M., Manso, L., Gallego, D., Paoletti, M., Bustos, P., Bandera, A., Romero-Garcés, A.: A navigation agent for mobile manipulators. In: Advances in Intelligent Systems and Computing, ROBOT2015 Conference, vol. 418, pp. 745–756 (2016). doi:10.1007/978-3-319-27149-1_58
8. Kotseruba, I., Gonzalez, O.J.A., Tsotsos, J.K.: A review of 40 years of cognitive architecture research: focus on perception, attention, learning and applications. CoRR abs/1610.08602 (2016)
9. Manso, L.J.: Perception as Stochastic Grammar-based Sampling on Dynamic Graph Spaces (2013)
10. Martínez-Gómez, J., García-Varea, I., Cazorla, M., Morell, V.: ViDRILO: the visual and depth robot indoor localization with objects information dataset. Int. J. Robot. Res. **34**(14), 1681–1687 (2015). doi:10.1177/0278364915596058
11. McDermott, D.: PDDL - the planning domain definition language. Technical report, CVC TR-98-003/DCS TR-1165, Yale Center for Computational Vision and Control (1998)
12. Nuñez, P., Manso, L., Bustos, P., Drews-Jr., P., Macharet, D.: Towards a new semantic social navigation paradigm for autonomous robots using CORTEX. In: IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2016) - BAILAR 2016 Workshop, New York City, USA (2016)
13. Pulido, J.C., González, J.C., Suárez-Mejías, C., Bandera, A., Bustos, P., Fernández, F.: Evaluating the child–robot interaction of the naotherapist platform in pediatric rehabilitation. Int. J. Soc. Robot., 1–16 (2017). doi:10.1007/s12369-017-0402-2
14. Viciana-Abad, R., Marfil, R., Perez-Lorenzo, J., Bandera, J., Romero-Garces, A., Reche-Lopez, P.: Audio-visual perception system for a humanoid robotic head. Sensors **14**(6), 9522–9545 (2014)