Multi-modal Sensing for Human Activity Recognition

Barbara Bruno, Jasmin Grosinger, Fulvio Mastrogiovanni, Federico Pecora, Alessandro Saffiotti, Subhash Sathyakeerthy and Antonio Sgorbissa

***Abstract*— Robots for the elderly are a particular category of home assistive robots, aiming at assisting the elderly in the execution of daily life tasks to extend their independent life. To this aim, such robots should be able to determine the level of independence of the user and track its evolution over time, to adapt the assistance to the person capabilities and needs. Human Activity Recognition systems employ various sensing strategies, relying on environmental or wearable sensors, to recognize various daily life activities which provide insights on the health status of a person. The main contribution of the article is the design of an heterogeneous information management framework, allowing for the description of a wide variety of human activities in terms of multi-modal environmental and wearable sensing data and providing accurate knowledge about the user activity to any assistive robot.**

I. Introduction

Home assistive robotics addresses the design of robots to be deployed in domestic environments, to assist the residents in the execution of daily life tasks. Robots for the elderly are a particular category of home assistive robots, relying on social interaction with the user and aiming at extending the elderly independent life [1]. The design and development of effective robots for the elderly is a very active topic, which currently gathers the efforts of a large research community.

Among the ample examples provided by Literature it is possible to count socially assistive robots, focusing on training the cognitive capabilities of the elderly [2], systems of assistive robots integrated in smart environments [3] and robots specifically designed for healthcare facilities [4].

To properly and effectively perform the assistive duties, as well as to exhibit a socially acceptable behaviour, robots for the elderly should not only be context-aware, i.e., able to assess the status of the environment they are in, and user-aware, i.e., able to assess the status of the person they are working for, but also *ageing-aware*, i.e., able to perform a long term analysis of the user cognitive and physical evolution, to adapt the assistance to the person capabilities.

Human Activity Recognition (HAR) systems for elderly-care are devoted to the identification, among all actions executed by a person during a day, of specific activities of interest, the Activities of Daily Living (ADL), to be analysed and notified to authorized personnel or devices. ADL require the use of different cognitive and physical abilities and are used by gerontologists to estimate the level of autonomy of a person [5]. As Figure 1 shows, ADL cover a wide variety of human activities, ranging from involuntary movements and postural transfers such as sitting and standing, to complex cognitive activities such as shopping or handling finances, which pose different constraints on the sensory system adopted for their recognition. Consequently, a number of sensing strategies (based on heterogeneous environmental sensors, wearable sensors, etc.) have been developed, each targeting the recognition of a specific subset of ADL.

This paper addresses the problem of endowing robots for the elderly with the ability of monitoring the Activities of Daily Living, by designing a HAR system which allows for a seamless integration with the robot planning system. In particular, for the effective monitoring of different types of ADL, we propose the integration of multiple sensing strategies in a single framework, i.e., the design of a multi-modal HAR system, which would allow for: (i) monitoring of ADL with distinct requirements and features; (ii) an increase in the recognition reliability of all monitored activities, thanks to the combination of the independent information provided by the different sensing systems, in accordance with the multi-modal paradigm; (iii) enhancing the context-awareness and user-awareness capabilities of assistive robots with first-hand information about the user and its ability of autonomously performing daily life tasks.

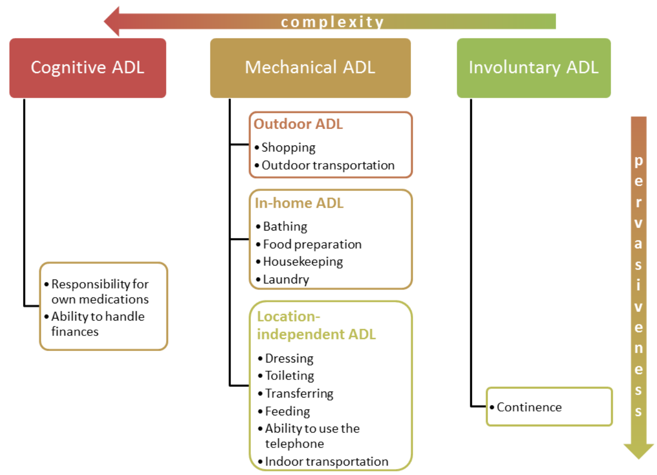

Fig. 1. Commonly considered Activities of Daily Living (ADL) for the assessment of the level of autonomy of a person.

This work was partially supported by a grant of the Italian Ministry of Foreign Affairs and International Cooperation (MAECI) and Ministry of Education, Universities and Research (MIUR) in the context of the Second Executive Programme for Scientific and Technological Cooperation between Italy and Sweden, under Project WearAmI.
B. Bruno, F. Mastrogiovanni and A. Sgorbissa are with the Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Via Opera Pia 13, 16145 Genova, Italy. J. Grosinger, F. Pecora, S. Sathyakeerthy and A. Saffiotti are with the AASS Cognitive Robotic Systems Lab, Örebro University, Fakultetsgatan 1, S-70182 Örebro, Sweden. Corresponding author's email: `barbara.bruno@unige.it`.

The article is organized as follows. Section II outlines the motivations for a multi-modal monitoring system and the knowledge framework supporting it is described in Section III. Section IV details the adopted sensory infrastructure and the proposed system architecture. Preliminary experimental results are analysed in Section V. Conclusions follow.

II. Problem Statement

By introducing the *complexity* of an ADL as a measure of the cognitive abilities required for the execution of the activity and the *pervasiveness* of an ADL as a measure of the requirements posed by the activity on the environment for its execution, it is possible to group ADL in five categories [6], shown with contour boxes in Figure 1. *Cognitive ADL*, such as handling finances, are very complex activities, which we argue to be beyond the capabilities of state-of-the-art HAR systems. The category of *outdoor mechanical ADL* comprises activities such as shopping and using public transports, which occur in a theoretically unbounded area outside of the person house and for which systems based on the sensing and processing capabilities of smartphones appear as the most natural monitoring approach [7], [8]. *In-home mechanical ADL* include a wide set of procedural activities which require the interaction with specific objects and devices, as in the case of bathing, housekeeping or doing the laundry, and which are most commonly and most effectively monitored via smart environments, i.e., systems relying on heterogeneous sensors distributed in the home to infer the status of the person from the context [9]. *Location-independent mechanical ADL* are motions related to toileting, feeding, performing postural transfers or moving indoor, which can virtually take place anywhere and are best and more often monitored with wearable sensing systems, that make use of sensors located on the person body to imply the status of the person from their limb movements [10], [11]. Lastly, involuntary ADL are highly-specific and highly-localized basic body functions, such as continence, that can only be monitored by dedicated systems.

Given the goals of assistive robots for elderly people, specifically designed to help them in the execution of daily life activities mostly occurring inside the person home, it is reasonable to focus on in-home mechanical ADL and location-independent ADL exclusively, monitored with smart environments and wearable sensing systems, respectively.

Unfortunately, wearable sensing systems, due to their limited sensory input, are prone to ambiguity [12], i.e., to misclassifying occurrences of activities which are very close in the sensory space, as it happens, for example, for *climb the stairs* and *walk*. In such cases, merging the activity information with the estimate of the person location, provided by the environmental sensors, would significantly reduce the number of false positive recognitions.

Conversely, a well-known drawback of smart environments is their vulnerability to erroneous conclusions due to incomplete information: for example, a pressure sensor placed under a chair actually detects whether the chair is occupied, but it is usually used to infer whether a person is sitting on it. The implicit underlying assumption that only the user sitting can activate the chair pressure sensor may lead to wrong conclusions; in this case, completing the pressure sensor information with information about the user movements, provided by the wearable sensor, allows improving the accuracy of the inference and, consequently, the correctness of the monitoring system conclusions.

The design of a hierarchical, multi-modal monitoring system able to combine the heterogeneous information provided by sub-systems implementing different sensing strategies, therefore, would: (i) allow for the expansion of the range of recognized activities, in accordance with the outlined taxonomy of ADL, shown in Figure 1; (ii) increase the robustness and reliability of the recognition.

III. The WearAmI Framework

Project WearAmI[1] focuses on the development of a multi-modal monitoring system that integrates the smart environment and wearable sensing approaches, by integrating the information coming from distributed environmental sensors with the information provided by wearable devices.

To this end, a first step is the identification of the necessary sensory infrastructure (i.e., which information is needed and how to acquire it) and the definition of a framework for knowledge representation and management.

An analysis of the ADL listed in Figure 1 allows for extracting general properties of human activities to be used for their definition in terms of sensory data. By introducing a *human motion* as a single, voluntary change in the person status, triggered and executed by the person themselves, it is possible to formulate the following assumptions.

Assumption 1: Human motions are associated with one or more *gestures* voluntarily performed by the user. Therefore, gesture is an intrinsic property of a human motion.

For example, *sitting* and *standing up* are associated with the gestures of, respectively, flexing and extending the legs, while *drinking* requires the sequential execution of the gestures of picking up and putting down the glass.

Assumption 2: Human motions are executed while the person is in a given *posture*. Therefore, posture is an intrinsic property of a human motion.

The ADL *transferring* comprises all human motions for which the posture at the end of the movement is different from the posture at the beginning of the movement. All other motions either require the person to be in a specific posture for their execution (for example, *walking* can only be performed while the person is *standing*), or can be associated with multiple postures (for example, *drinking* can be performed while the person is *sitting* or *standing*).

Denoting as *context* any information that is relevant for the characterization of the status of a person [13], we can categorize context information as either belonging to the spatial domain, i.e., providing insights about the surroundings of the person, or to the temporal domain, i.e., providing insights

[1]WearAmI is the acronym of “Wearable and Ambient Intelligence Make Assistive Robots Smarter”

about the previous status of the person. Spatial information is defined *environment*. An environment is associated with a location, and characterized by the objects within it.

Assumption 3: Human motions are executed while the person is at a given *location*. Locations in the environment put constraints on the actions that can be there executed and therefore are an intrinsic property of a human motion.

The ADL *indoor transportation* comprises all human motions for which the location at the end of the movement is different from the location at the beginning of the movement. Other motions either require the person to be in a specific location for their execution (for example, the ADL *bathing* can only be executed in the *bathroom*) or are location-independent and therefore can be executed anywhere.

An object is a physical entity with a certain position within the environment, that allows for interaction with a person. Each possible interaction of an object with a person is an intrinsic property of that object, called *affordance* [14]. Mobility is the affordance which indicates whether a person can easily move the object from one place to another.

Assumption 4: Objects in the environment provide affordances and constraints to the user and therefore are an intrinsic property of a human motion.

Location-dependent motions, like the ADL *bathing*, are defined as such because of their dependence on the non-movable objects which characterize the environment associated with that location (for example, the bathtub). Conversely, location-independent human motions, such as *drinking*, depend exclusively on movable objects. Finally, motions such as *walking* do not depend on any object.

Assumption 5: The execution of a *Human Motion* allows for the later execution of other activities and therefore is an intrinsic property of a human motion.

As an example, *standing up* from a chair can only be executed after the person has *sit down* on that same chair.

The observations allow for describing human activities in terms of: (i) associated basic gesture; (ii) admissible posture (possibly decomposed into an initial and a final posture); (iii) location-dependency and associated location(s) of interest; (iv) object-dependency and associated object(s) of interest; (v) motion-dependency and associated motion(s) of interest.

This representation, in turn, allows for defining the necessary sensory infrastructure, i.e.: (i) a sensor for the detection of human gestures, whose meaning is given by their context; (ii) a sensor for the detection of the person posture; (iii) a system of homogeneous sensors for the detection of the person location; (iv) a family of heterogeneous sensors for the analysis of objects usage; (v) a framework for the analysis of the temporal constraints relating motions one to the other.

IV. System Architecture

Having identified the information and the sensory infrastructures required by a multi-modal human activity monitoring system, we set up a preliminary test bed in an apartment located in the city of Örebro, in the elderly care facility

Fig. 2. An overview of Ängen apartment, designed to host an elderly person living alone, used as a test bed for the multi-modal activity monitoring system. The stars mark, clockwise from top-left, the bathroom, the bedroom, the kitchen and the living-room.

Ängen[2]. The apartment, shown in Figure 2, is composed of fully furnished living-room, bathroom, bedroom and kitchen, and designed to host an elderly person living alone. A laundry room equipped with washing and drying machines is on the top floor of the building, at the disposal of all tenants.

Accommodations of this type present a number of notable peculiarities: (i) the apartments are built in accordance with the criteria and requirements of elderly care facilities, which already include the installation of simple environmental sensors (such as presence sensors and gas sensors); (ii) the tenants are autonomous or semi-autonomous elderly people who willingly chose such an accommodation to have an easy access to medical care and specific services (such as cleaning, laundry or cooking services), while retaining their independence; (iii) the medical personnel of the elderly care facility is responsible for checking the health status of the tenants and intervening in case of emergencies. This scenario is particularly well suited for the adoption of an automated monitoring system, which provides the caregivers with up-to-date information about the health status of the assisted people without significantly affecting their daily routines.

On the basis of the list of ADL reported in Figure 1, we have decided to focus on: *transferring* (denoting the activities of sitting down, standing up, lying down, getting up); *feeding* (eating, drinking); *food preparation*; *indoor transportation* (climbing stairs, descending stairs, walking).

We propose the multi-modal monitoring system having the architecture shown in Figure 3 for the reliable detection of all activities of interest. In accordance with the outline framework, the system makes use of: (i) a wrist-placed inertial sensor; (ii) a waist-placed inertial sensor; (iii) a network of Passive Infra-Red sensors; (iv) RFID tags, pressure sensors

[2]The apartment is the core of Ängen Research and Innovation program, which aims at developing and showcasing technologies for facilitating ageing in place. Ängen is supported by Örebro University, the city council, the municipality, and Örebro Science Park.
More information are available at: `http://angeninnovation.se/`.

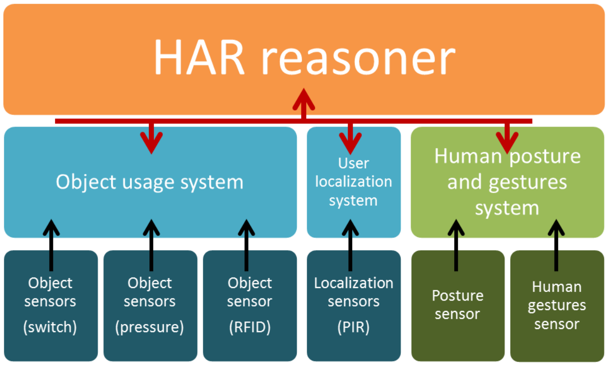

Fig. 3. WearAmI system architecture: dark boxes denote the adopted sensors, sending information to the corresponding analysis system (black arrows). Blue shades represent environmental sensing components, while green shades represent wearable sensing components. The analysis systems, devoted, respectively, to objects usage, user localization and gesture & posture analysis, together with the constraint reasoner responsible for the actual recognition of occurrences of modelled ADL, share a uniform communication model (red arrows).

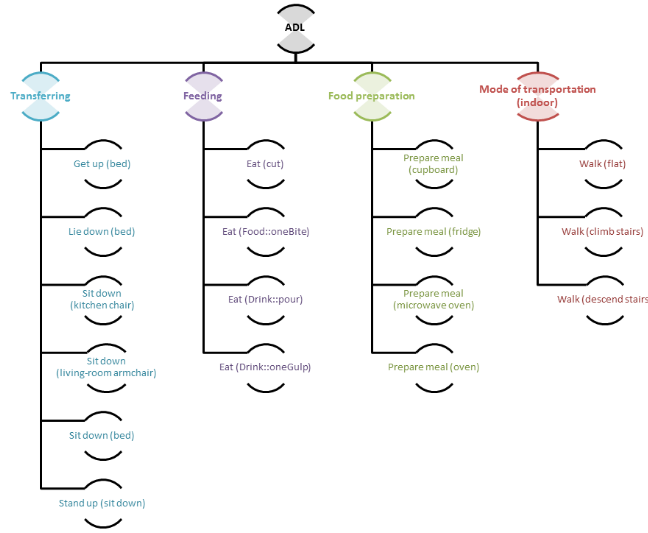

Fig. 4. The four considered ADL (in the coloured patches) and corresponding overloaded Human Motions.

and switches; and (v) a temporal reasoner.

The data extracted by the wrist sensor are used to detect occurrences of basic gestures [15], [10], such as *walking*, *picking up* or *sitting*. The data provided by the waist sensor, instead, are used to estimate the user posture on the basis of the angle between the torso and gravity force. The combined analysis of wrist and waist acceleration data also allows detecting falls with high accuracy [16].

User localization is achieved via a network of Passive Infra-Red (PIR) sensors, distributed throughout the apartment, signalling whether and which rooms are occupied by people. Given the chosen ADL of interest, we identified three categories of objects to monitor: *cutlery and dishes*, which are assumed to be in use when located on the kitchen table and that we detect via an RFID network; *furniture*, such as chairs, armchairs and bed, for which pressure sensors detect whether and which is in use; *household appliances*, such as the fridge and the oven, whose usage can be inferred by checking the status of their doors with switches.

For a proper management of the reported sensory information, it is useful to envision smart environments as ecologies of Physically Embedded Intelligent Systems (PEIS) [17], i.e., devices incorporating computational and communication resources, able to interact with the environment via sensors and/or actuators and connected with each other by a uniform communication model. Concretely, the PEIS-ecology is a middleware based on a shared library, the PEIS-kernel, which is linked by any PEIS and is responsible for communication. In Figure 3 black arrows mark the heterogeneous private communication links between any analysis systems and its related sensors (denoted with dark boxes). The analysis systems (focusing on objects usage, user localization and user posture & gestures, respectively) are implemented as PEIS-components, which can share information among each other and, in particular, with a reasoning system which is responsible for the recognition of all occurrences of activities of interest. The adopted temporal reasoner uses and extends Allen interval algebra to represent the criteria for activity recognition [18]. In particular, the four considered ADL are decomposed in groups of overloaded (i.e., associated with multiple sensory patterns) Human Motions, reported in Figure 4, each modelled as a set of temporal constraints.

The possibility of associating multiple sensory patterns with the same Human Motion is a crucial feature of the proposed monitoring system. Most mechanical ADL, such as *feeding* and *preparing meals*, are composed of a wide variety of activities, which are executed in a procedural form, i.e., in ordered sequences. While people do not necessarily respect the same ordering every time they perform a mechanical ADL, traditional smart environments are forced to a-priori define all admissible orderings. Due to this mismatch, smart environments are forced to choose between generic models (for example, defining *preparing meals* as any execution of opening the fridge), at the expenses of a reduced precision, and specific models (for example, defining *preparing meals* as the sequence: opening and closing the fridge, followed by opening and closing the microwave oven). Overloading allows for establishing a hierarchy of activities with various degrees of complexity and for choosing the way they refer to the ADL of interest, which ultimately increases the robustness and accuracy of the recognition.

V. Experimental Evaluation

Listings 1, 2 and 3 report the models of some of the considered Human Motions, written in a variant of NDDL, which, in turn, is an extension of the Planning Domain Definition Language (PDDL). The goal of PDDL is to provide a common formalism for describing planning domains; adopting the same formalism for the description of human activities allows for the seamless integration between the monitoring system and any actuated device (home assistive robots above all) and leads to a significant improvement in the context assessment capabilities of the assistive devices.

Listing 1. DDL models for the Human Motions *sit down* and *stand up*.

```
(SimpleOperator
 (Head Human::SitDown())
 (RequiredState req1 Gesture::Sit())
 (RequiredState req2 Posture::Sitting())
 (RequiredState req3 Chair::On())
 (Constraint OverlappedBy(Head,req1))
 (Constraint During(Head,req2))
 (Constraint EndEnd(Head,req3))
)

(SimpleOperator
 (Head Human::SitDown())
 (RequiredState req1 Gesture::Sit())
 (RequiredState req2 Posture::Sitting())
 (RequiredState req3 Armchair::On())
 (Constraint OverlappedBy(Head,req1))
 (Constraint During(Head,req2))
 (Constraint EndEnd(Head,req3))
)

(SimpleOperator
 (Head Human::StandUp())
 (RequiredState req1 Gesture::Stand())
 (RequiredState req2 Posture::Standing())
 (RequiredState req3 Human::SitDown())
 (Constraint MetByOrOverlappedBy(Head,req1))
 (Constraint Starts(Head,req2))
 (Constraint MetByOrAfter(Head,req3))
)
```

Listing 2. DDL models referring to the Human Motion *eat*.

```
(SimpleOperator
 (Head Drink::Pour())
 (RequiredState req1 Location::DiningRoom())
 (RequiredState req2 Gesture::PourS())
 (RequiredState req3 Gesture::PourE())
 (RequiredState req4 Glass::On())
 (RequiredState req5 Bottle::On())
 (Constraint During(Head,req1))
 (Constraint StartedBy(Head,req2))
 (Constraint FinishedBy(Head,req3))
 (Constraint During(Head,req4))
 (Constraint During(Head,req5))
)

(SimpleOperator
 (Head Human::Eat())
 (RequiredState req1 Drink::Pour())
 (Constraint Contains(Head,req1))
)

(SimpleOperator
 (Head Drink::OneGulp())
 (RequiredState req1 Location::DiningRoom())
 (RequiredState req2 Gesture::PickUp())
 (RequiredState req3 Gesture::PutDown())
 (RequiredState req4 Glass::On())
 (Constraint During(Head,req1))
 (Constraint StartedBy(Head,req2))
 (Constraint FinishedBy(Head,req3))
 (Constraint During(Head,req4))
)

(SimpleOperator
 (Head Human::Eat())
 (RequiredState req1 Drink::OneGulp())
 (Constraint Contains(Head,req1))
)
```

The field `Head` defines the entity it refers to and the name of the model, separated by a `::`. As an example, `Head Human::SitDown()` indicates that whenever the reported constraints are satisfied, the reasoner should infer that the motion of sitting has been executed by the human user. The field `RequiredState` defines the sensor values which correspond to the execution of the motion, identified in accordance with the knowledge framework reported in Section III. For example, the Human Motion of *sitting on a chair* is defined by the associated basic gesture (`Gesture::Sit()`), admissible posture (and more precisely, the final posture, i.e., `Posture::Sitting()`) and associated object of interest (`Chair::On()`). Since this motion is not location-dependent nor motion-dependent, the requirements do not mention any location or motion of interest. An example of location-dependent Human Motion is *climb the stairs*, while *stand up* is motion-dependent. Lastly, the field `Constraint` defines the temporal relation between each sensor value of interest and the Human Motion.

Listing 3. DDL models of the Human Motions *descend stairs* and *walk*.

```
(SimpleOperator
 (Head Human::ClimbStairs())
 (RequiredState req1 Posture::Standing())
 (RequiredState req2 Gesture::Walk())
 (RequiredState req3 Location::Staircase())
 (RequiredState req4 Location::EntranceHall())
 (Constraint During(Head,req1))
 (Constraint During(Head,req2))
 (Constraint Equals(Head,req3))
 (Constraint MetByOrAfter(Head,req4))
)

(SimpleOperator
 (Head Human::DescendStairs())
 (RequiredState req1 Posture::Standing())
 (RequiredState req2 Gesture::Walk())
 (RequiredState req3 Location::Staircase())
 (RequiredState req4 Location::LaundryRoom())
 (Constraint During(Head,req1))
 (Constraint During(Head,req2))
 (Constraint Equals(Head,req3))
 (Constraint MetByOrAfter(Head,req4))
)

(SimpleOperator
 (Head Human::Walk())
 (RequiredState req1 Posture::Standing())
 (RequiredState req2 Gesture::Walk())
 (Constraint During(Head,req1))
 (Constraint Equals(Head,req2))
)
```

By defining the sequences of values registered by the sensors over time, it is possible to perform a preliminary validation of the Human Motions, i.e., to verify that the proposed temporal rules trigger the recognition of the corresponding activity with the expected behaviour. In the tests, envisioned to help understand the benefits of adopting a multi-modal approach for human activity recognition, we define sequences of sensor values and analyse the reasoner inferences they trigger. Figures 5, 6 and 7 report the outcome of the tests. In all figures, the indexes below the timelines refer to time (in seconds): for the sake of visualization clarity, portions of the recordings with no variation in the sensor values have been shrunk. The timelines of the context variables `Human` and `Drink` are computed by the reasoner and list all corresponding recognized Human Motions, as indicated by the `Head` fields. The other timelines report the sensor values (i.e., gesture, posture, location and objects sensors, as defined by the knowledge management framework). In each figure, the timelines of sensors which are not relevant for the test have been hidden.

At each time instant, the reasoner samples the sensors, keeping track of all Human Motions which are consistent with the sensors readings up to that instant (i.e., all the Human Motions that *could be* the one currently being executed, which we define *possible Human Motions*). As time passes, the number of possible Human Motions progressively reduces, until it converges to the one effectively performed, if it is among the modelled ones, or to none, if the performed motion is not among the modelled ones. A possible Human Motion which at some point has been acknowledged as the one effectively performed is defined as an *inferred Human Motion*: all sensors or context variables statuses supporting an inferred Human Motion are marked with a blue filling. A green filling denotes a sensor or context variable status which, while maybe supporting possible Human Motions, does not support any inferred Human Motion. A grey filling in the context variables timelines denotes a lack of knowledge about the variable status at that time, i.e., the execution of activities which are not modelled in the system.

Figure 5 reports the simulated sensor readings related to a person who drops a heavy bag on the kitchen chair, then walks to the living-room and sits on the armchair. As the timelines show, although the chair pressure sensor is activated by the bag (for $t = [16; 35]$), the wearable gesture and posture sensors do not signal any sitting motion, therefore preventing the reasoner from making an erroneous inference. Later on, when the person sits on the armchair, environmental and wearable sensors agree on indicating that the person sat down, therefore triggering the correct recognition of the *sit down* motion. The example also highlights one advantage deriving from overloading rules: since all three modelled sitting actions are defined as `SitDown`, it is possible to define a single model for the standing up motion, which is constrained by the previous occurrence of *any* sitting action.

Figure 6 represents the sensor readings related to a person who, in the process of having lunch, pours water in a glass and drinks. Feeding is a mechanical ADL composed of a wide variety of activities (such as cutting food with a knife, picking a bite, pouring liquid in a glass, drinking, etc.) and it is not possible to define a priori a preferred order in which they will be executed. Overloading, i.e., associating with `Eat` the occurrence of *any* Human Motion related to feeding, allows not indicating a preferred ordering, which increases the robustness and accuracy of the recognition. In the reported example, in accordance with the models in Listing 2, the recognition of *eat* is first triggered by the execution of *pouring water* and later on by the execution of *drinking* (as shown by the timeline of `Drink`).

Finally, Figure 7 reports the sensor readings related to a person walking upstairs to the laundry room and then returning to her apartment. As the `Gesture` timeline shows, the basic gesture associated to the two motions is the same (since wearable sensing systems can not discriminate between walking and climbing/descending stairs due to their similarities in the acceleration space) and the reasoner relies on the information provided by the environmental location sensors to infer which is the executed one.

VI. Conclusions

In this paper we introduced the idea of a multi-modal monitoring system, i.e., a system for the recognition of a broad set of human activities which integrates information retrieved via different monitoring approaches. Preliminary work has been carried out to set up the necessary hardware and software infrastructure in a test bed apartment located in an elderly care facility in Örebro, Sweden.

Two considerations about the adopted knowledge management solutions are particularly relevant. The ecology of intelligent devices [17] envisions the interconnection of all sentient and acting devices (from sensors, to assistive robots for the elderly): the introduction of wearable sensing systems in this ecology allows enhancing it with first-hand knowledge about the most intelligent actor in the environment, i.e., the human, and is intended as a contribution towards bridging the gap between people and artificial automated systems. Similarly, a key feature of the adopted constraint-based reasoner [18] is its seamless integration between human activity recognition and planning for assistive robots: any refinement in the system understanding of the person status will bring about an improvement in the quality of the provided services and ultimately increase the effectiveness and usefulness of assistive devices.

References

[1] D. Johnson, R. Cuijpers, J. Juola, E. Torta, M. Simonov, A. Frisiello, M. Bazzani, W. Yan, C. Weber, S. Wermter, N. Meins, J. Oberzaucher, P. Panek, G. Edelmayer, P. Mayer, and C. Beck, "Socially assistive robots: A comprehensive approach to extending independent living," *International Journal of Social Robotics*, vol. 6, no. 2, pp. 195–211, 2014.

[2] R. Kachouie, S. Sedighadeli, R. Khosla, and M. Chu, "Socially assistive robots in elderly care: A mixed-method systematic literature review," *International Journal of Human-Computer Interaction*, vol. 30, no. 5, pp. 369–393, 2014.

[3] F. Cavallo, R. Limosani, A. Manzi, M. Bonaccorsi, R. Esposito, M. di Rocco, F. Pecora, G. Teti, A. Saffiotti, and P. Dario, "Development of a socially believable multi-robot solution from town to home," *Cognitive Computation*, vol. 6, no. 4, pp. 954–967, 2014.

[4] F. Mastrogiovanni, A. Sgorbissa, and R. Zaccaria, "Context assessment strategies for ubiquitous robots," in *Proceedings of the 2009 IEEE International Conference on Robotics and Automation (ICRA 2009)*, Kobe, Japan, May 2009, pp. 2717–2722.

[5] S. Katz, A. Chinn, and L. Cordrey, "Multidisciplinary studies of illness in aged persons: a new classification of functional status in activities of daily living," *Journal of Chronic Disease*, vol. 9, no. 1, pp. 55–62, 1959.

[6] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa, "A public domain dataset for adl recognition using wrist-placed accelerometers," in *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2014)*, Edinburgh, UK, August 2014, pp. 738–743.

[7] I. Bisio, F. Lavagetto, M. Marchese, and A. Sciarrone, "A smartphone-centric platform for remote health monitoring of heart failure," *International Journal of Communication Systems*, 2014.

[8] S. Doherty, C. Lemieux, and C. Canally, "Tracking human activity and well-being in natural environments using wearable sensors and experience sampling," *Social Science & Medicine*, vol. 106, pp. 83–92, 2014.

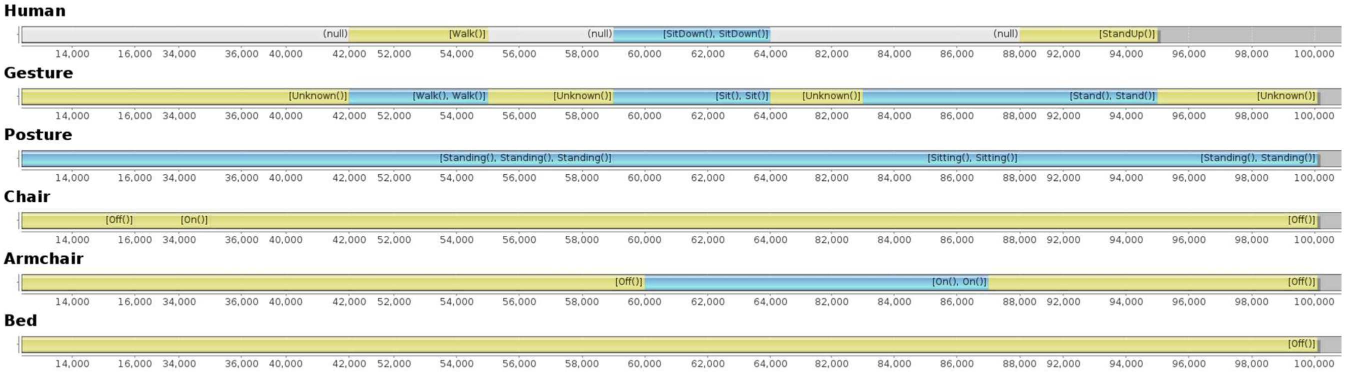

Fig. 5. Validation of the temporal models of the Human Motions *sit down* and *stand up*.

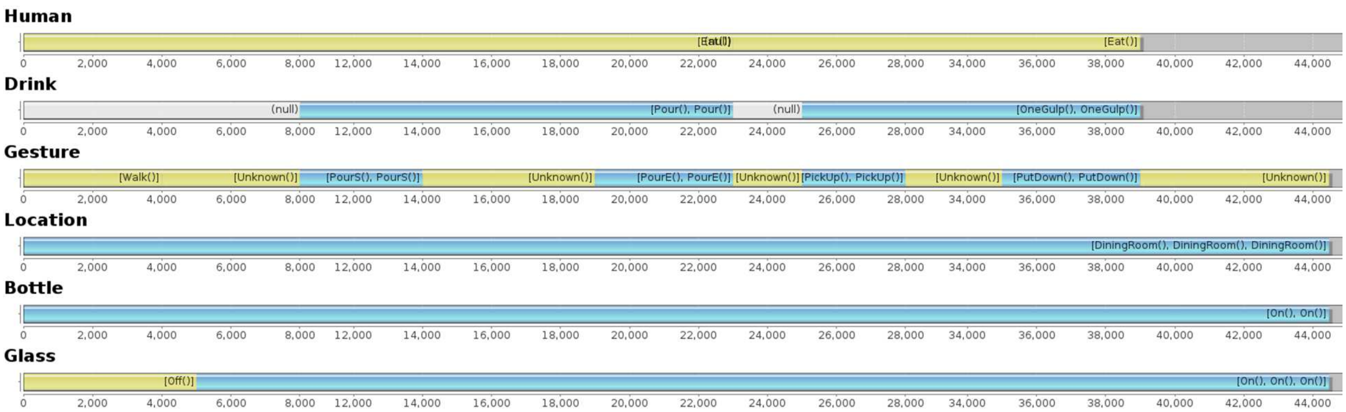

Fig. 6. Validation of the temporal model of the Human Motion *eat*, in relation to pouring water in a glass and drinking.

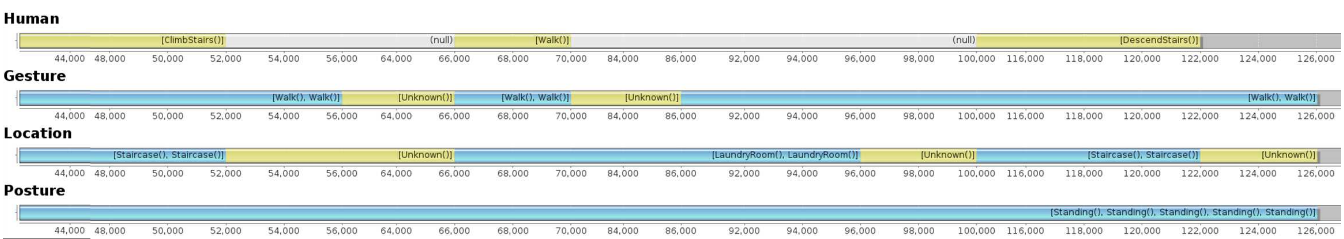

Fig. 7. Validation of the temporal models of the Human Motions *descend stairs* and *walk*.

[9] M. Alam, M. Reaz, and M. Ali, "A review of smart homespast, present, and future," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 42, no. 6, pp. 1190–1203, 2012.

[10] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Analysis of human behavior recognition algorithms based on acceleration data," in *Proceedings of the 2013 IEEE International Conference on Robotics and Automation (ICRA 2013)*, Karlsruhe, Germany, May 2013, pp. 1602–1607.

[11] O. Lara and M. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[12] B. Bruno, F. Mastrogiovanni, A. Saffiotti, and A. Sgorbissa, "Using fuzzy logic to enhance classification of human motion primitives," in *Proceedings of the 15th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2014)*, Montpellier, France, July 2014, pp. 596–605.

[13] G. Abowd, A. Dey, P. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Proceedings of the 1999 International Symposium on Handheld and Ubiquitous Computing (HUC 1999)*, Karlsruhe, Germany, September 1999, pp. 304–307.

[14] J. Gibson, "The theory of affordances," in *R.E. Shaw and J. Bransford (Eds.) Perceiving, Acting, and Knowing*. Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates, 1977.

[15] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Human motion modeling and recognition: a computational approach," in *Proceedings of the 8th IEEE International Conference on Automation Science and Engineering (CASE 2012)*, Seoul, Korea, August 2012, pp. 156–161.

[16] M. Kangas, A. Konttila, P. Lindgren, I. Winblad, and T. Jamsa, "Comparison of low-complexity fall detection algorithms for body attached accelerometers," *Gait & Posture*, vol. 28, pp. 285–291, 2008.

[17] A. Saffiotti, M. Broxvall, M. Gritti, K. LeBlanc, R. Lundh, J. Rashid, B. Seo, and Y. Cho, "The peis-ecology project: vision and results," in *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, Nice, France, September 2008, pp. 2329–2335.

[18] F. Pecora, M. Cirillo, F. Dell'Osa, J. Ullberg, and A. Saffiotti, "A constraint-based approach for proactive, context-aware human support," *Journal of Ambient Intelligence and Smart Environments*, vol. 4, no. 4, pp. 347–367, 2012.