Towards a System Architecture for Recognizing Domestic Activity by Leveraging a Naturalistic Human Activity Model

Michele Dominici
INRIA, Centre Inria Rennes - Bretagne Atlantique
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
Michele.Dominici@inria.fr

Myriam Fréjus and **Julien Guibourdenche**
EDF Recherche & Développement
1 Avenue du Général de Gaulle
92141 CLAMART Cedex, France
firstname.lastname@edf.fr

Bastien Pietropaoli
INRIA, Centre Inria Rennes - Bretagne Atlantique
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
Bastien.Pietropaoli@inria.fr

Frédéric Weis
Université de Rennes 1, IRISA
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
Frederic.Weis@irisa.fr

Abstract

Existing activity recognition approaches in the smart home domain suffer from poor human activity models. Combining expertise from cognitive ergonomics and ubiquitous computing, we discuss the hard technical challenges to address when leveraging a realistic model of human activity. We present the architecture of a prototype smart home system that we are developing and show the gap that exists between our current capabilities in terms of contextual-knowledge extraction and the complexity of the targeted activity recognition. To fill this gap, we propose and discuss the integration of PHATT, an existing algorithm for plan recognition, into our system in order to mine additional information from the dynamics of context.

1 Introduction and Motivation

A *smart home* is a residence equipped with information-and-communication-technology devices conceived to collaborate in order to anticipate and respond to the needs of the occupants, working to promote their comfort, convenience, security and entertainment while preserving their natural interaction with the environment (Aldrich 2003).

When talking about natural interaction, one of the most precious resources to preserve is user attention: during their activities, users should be supported *invisibly*, reducing interruptions and explicit interactions with the system as much as possible. In order to achieve these goals, smart home systems must be able to take into account the *context*, that is the implicit situational information that influences human behavior (Roy et al. 2010), recognizing people activities to provide adapted functionalities. For example, under some conditions, knowing that an inhabitant is executing some long-lasting static activity in a room can suggest that the system should turn on the room's heating and turn off the other rooms' lights.

Existing solutions for human activity recognition often rely on data coming from wearable sensors or video cameras (Chen and Nugent 2009), technologies that are difficult to deploy and get accepted in real-world households. Furthermore, these solutions address the problem of activity pattern discovery directly on raw sensor data or video streams, exploiting data mining techniques to extract recurring patterns in the raw data and to predict or classify future observations, as explained in (Kim, Helal, and Cook 2010). The resulting systems fail to provide adapted services to people in real-world scenarios, as the "gap" between the captured context and the complexity of human behavior is too large. We believe that the main reasons are the poverty (or absence) of the underlying models of human behavior and activities, which don't handle some fundamental aspects of the reality, and/or the lack of computing models taking advantage of these aspects.

To address these issues, we started an interdisciplinary project that brings together researchers from the fields of ubiquitous computing and cognitive ergonomics. Our aim is to develop a smart home system that is able to prevent energy waste and preserve inhabitants' comfort, leveraging on realistic human activity models. Our hypothesis is that human activity models have to be taken into account as challenging implication for informatics, although they shall not be directly integrated into computing models.

In this paper, our contribution is threefold. In Section 2, we present some of the challenging dimensions of human activity, which are not handled by most existing approaches, emerging when considering activity as a situated process, relying on actualization of concerns, and integrated in a network of interactions. In Section 3, we present a functional architecture that is designed to extract high-level situations from low-level raw sensor data. We show the need for an additional activity recognition mechanism and a system architecture that leverages the ubiquitous computing principles and that is at the core of the prototype system that we are developing. In Section 4, we propose to adapt PHATT, an existing algorithm for plan recognition (Goldman, Geib, and

Miller 1999; Geib and Goldman 2009), to be integrated into our architecture, in order to start addressing the challenge of providing adapted functionalities, which are well suited to the complexity of domestic activity. Section 5 illustrates the issues that remain unsolved and that may benefit from exchanging with the GAPRec and ICAPS research communities, while Section 6 concludes the paper.

2 Human Models of Domestic Activity

Recent naturalistic studies (Baillie and Benyon 2008; Crabtree and Rodden 2004; Guibourdenche et al. 2011; Poizat, Fréjus, and Haradji 2009; Salembier et al. 2009) provided some fundamental knowledge for a deeper empirical understanding of human domestic activity. Those studies aimed at orienting the design of ambient systems on the basis of real activity models and definitions of activity contexts, from the inhabitants' points of view. These formal descriptions of real activities and people's contexts are prerequisite for building appropriate applications (Greenberg 2001). These models also raise different issues challenging technical models for activity recognition.

Many existing technical models for activity recognition consider human activities as sequences of targeted actions that are always executed in the same order and which are never concurrently executed or interleaved with actions corresponding to other activities (Gu et al. 2009). Instead, we conducted our work in reference to the course of action empirical research program (Theureau 2003). This theoretical framework, as well as naturalistic studies, demonstrates that (domestic) activity is opportunistic. Inhabitants frequently interrupt a particular task for a while in order to accomplish another one. Individual activity at home is constituted of multiple lines of different concerns which structure a kind of fuzzy involvement in the activity. For example, a mother can be ironing while following a TV-show and looking after children playing at the first floor. Inhabitants manage several activities at the same time with several underlying concerns, which take part in their individual context. Activity is never built according to a pre-established and hierarchical plan but is constantly reoriented according to inter-individual interactions and interactions with the physical environment. This raises design issues relative to the gap between this complex human context and the context of the system based upon an environmental capture.

In addition, the same behavior (e.g. closing shutters) can have several meanings (e.g. reducing the brightness in a room, ensuring some privacy, increasing the sense of safety, reducing the temperature inside the house). This slight gap is due to the asymmetrical relation between environment as raw material, and situation as experienced environment through the individual's activity. Thus, the model has to integrate several layers of inference from a low level (e.g., a shutter is being closed) to a high level (e.g., Julie wants to have more cosiness), the latter the more problematic. Some situations can cause the system to an inability to determine the appropriate action to take; thus, designing a context-aware system implies designing the interaction with the user in order to manage uncertainty (explicit interaction, validation, etc).

Another design limitation rises from the impossibility for a smart home system to act according to deterministic rules only (either manually provided or automatically extracted through machine learning techniques). In our precedent example, a rule such as "closing shutters when night falls" can't be adapted to the several meanings underlying the closing of shutters. However, some solutions are based upon such design principles (Gu et al. 2004; Campo et al. 2006; Bonhomme et al. 2008b; 2008a), but they encounter difficulties to be adapted to inhabitants' practices. Even though some recurrent activities (e.g., cooking, taking children to bed, watching night shows on TV) can be observed at day or week scales, they nevertheless seem to be always accomplished differently, at different times or in a different order. Routines illustrate the recurrence of concerns, not the execution of schemes of action, as some works assume (Chan et al. 2008).

Furthermore, individual and collective scales of activity are intertwined, *mutually* (Crabtree and Rodden 2004; Poizat, Fréjus, and Haradji 2009) and *conflictually* (Baillie and Benyon 2008) giving shape to one another. For example the cleaning can be initiated by an individual and finished by another, the latter doing a different way than what the former had previously thought. Therefore, the system design can't rely on a human activity considered only as individual, a lonely man doing one thing at a time.

Moreover, the activity can't be strictly associated with a specific space (Guibourdenche et al. 2011): families are distributed across multiple scales of physical spaces (floors, rooms, systems of tools, voices, noises). During a local activity (for example, a mother doing the ironing and watching TV in the living room), concerns may refer to different places or people (as supervising children in the example: the mother is also concerned with her daughter alone upstairs). Those characteristics imply that the system can't consider only a local point of view on the activity and must integrate local and global points of views.

Now our aim is to specify an architecture capable of integrating these various constraints.

3 Architecture

In this Section, we present the architecture of a prototype system that we are developing. The aim of our system is to capture physical information from the environment, extract higher-level concepts and combine them to infer human situations and activities, with the ultimate goal of semi-automatically managing household appliances and provide additional functionalities that allow saving energy while preserving comfort. We first present the system architecture, which relies on the principles of the ubiquitous computing paradigm (Weiser 1993) and draws its inspiration from the four-layer model described in (Coutaz et al. 2005), and then highlight the need for an additional activity recognition mechanism.

3.1 Layered architecture

To achieve the goals of our scenario, we decided to adopt the human-computer interaction paradigm called *ubiquitous*

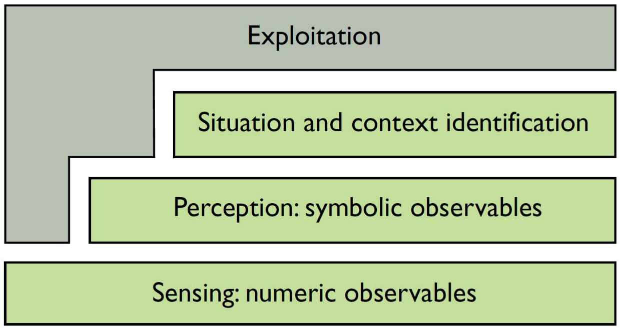

Figure 1: *The four-layer model for context-aware applications proposed by (Coutaz et al. 2005).*

computing. The aim of this paradigm is to seamlessly and invisibly integrate in the physical environment a multitude of digital devices that provide services to the users without asking for their attention (Weiser 1993). To this end, ubiquitous computing applications are typically context-aware, where the word context is used to address any static or dynamic condition that concerns the digital, physical and user-related environments in which a context-aware application is executed. In (Coutaz et al. 2005), a four-layer model is suggested to build context-aware applications, as showed in Fig. 1. The first layer, *sensing*, corresponds to the raw data sensed from the environment. The second layer, called *perception*, can be interpreted as an abstraction of the raw data. *Situation and context identification*, the third layer, concerns the context itself and the situations occurring in the home. The top layer, called *exploitation*, provides contextual information to applications. Our work is partly based on this model.

Considering the aforementioned model, the first layer of our system should be simply composed of sensors, but some constraints have to be fitted. In order to reduce the global system cost and to protect the inhabitants' privacy, the number of sensors dispatched in the environment has to be reduced as much as possible. However, a huge number of different sensors are required to sense context pieces and redundancy can significantly increase the reliability of the sources. With this idea in mind, the sensors will be grouped in nodes, as showed in Fig. 2. These nodes are able to preprocess the data with simple computation such as minimum, maximum and average. They also enable the sensors to communicate, using, for instance, 6LowPAN (IPv6 over LoW Power wireless Area Networks), which is specifically designed for embedded systems (Shelby and Bormann 2009). Another benefit of using nodes is the optimization of energy consumption due to radio communications. This is not negligible as most of the nodes will be running on batteries.

In the second layer of Coutaz' model, the raw data are processed to obtain more abstract data about context and occurring situations. The aggregation of raw data is realized thanks to a data fusion algorithm. The data fusion algorithm that we adopted is called the *belief functions theory* or *theory of evidence*. More specifically, the transferable belief model from (Smets and Kruse 1996) is used to aggregate data from

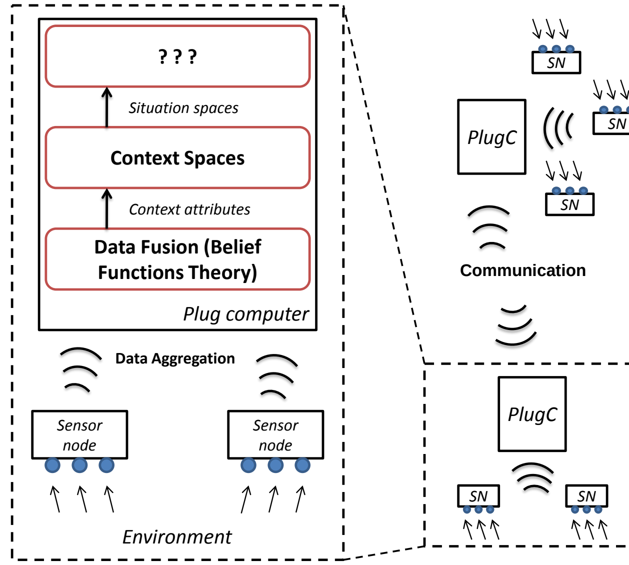

Figure 2: *The system architecture – The sensor nodes send aggregated data to the plug computers, which are in charge of performing sensor data fusion, to produce the context attributes, and context spaces reasoning, to infer the ongoing situation spaces. An additional processing step is needed to perform activity recognition.*

homogeneous as well as heterogeneous sources. The aim in this layer is to extract from raw sensing pieces of higher-level contextual information. Some existing solutions also adopt a similar approach. For instance, in (Ricquebourg et al. 2007), a pressure sensor on a chair, an omnidirectional webcam and a tracking sensor are used to determine the posture of a person. Another example is given in (Chahuara, Vacher, and Portet 2010) with the localization of inhabitants using microphones, presence sensors and contact sensors on house furnishing doors.

The bridge between the second and the third layer is realized integrating the results of sensor data fusion into a context model called *Context Spaces*. This model uses geometrical metaphors to describe context and situations, relying on the following concepts (Padovitz, Zaslavsky, and Loke 2006): the *context attributes*, the *application space*, the *situation spaces* and the *context state*. The *context attributes* are information types that are relevant and obtainable by the system; in our case, the context attribute values are provided by the perception layer, together with a degree of *confidence* on them, needed to cope with the intrinsic uncertainty of sensing systems in real world scenarios. The *application space* is a multi-dimensional space made up of a domain of values for each context attribute. The *situation spaces* are subspaces of the application space defined over regions of acceptable values of selected context attributes; situation spaces model real-life situations, e.g., "the whole family is in the kitchen" or "a person is ironing". A *context state* is the collection of current context attribute values at a given moment (Padovitz, Zaslavsky, and Loke 2006).

In the *situation and context identification* layer, the context state provided by the perception layer is analyzed to infer the ongoing situation spaces (representing real-life situations) and also produce a measure of confidence in their occurrence. As the same context state can correspond to several different situation spaces (and vice versa), reasoning techniques are needed to discern the actual ongoing real-life situations in spite of uncertainty (Padovitz, Zaslavsky, and Loke 2006; Padovitz 2006).

Unfortunately, the computations required by the second and the third layers to obtain abstract data and to analyze context and situations are too heavy for our nodes to be processed on. To remedy to this problem, more powerful nodes acting like sinks are used. These nodes are small "plug and play" computers called *plug computers* (ref. Fig. 2). Their role is to gather data from sensor nodes and to perform data fusion, required to produce the context attributes, and context space reasoning, used to identify ongoing situations.

As explained in (Coutaz et al. 2005), the *exploitation* layer acts as an adapter, allowing applications to address to the infrastructure their requests for context services at a high level of abstraction. In our architecture, this layer will provide information about context to augmented appliances, which will adapt their behavior in a semi-automatic way.

3.2 Need for an additional layer

In our smart home distributed system, applications will directly execute on physical objects and household appliances, adapting their behavior to save energy and preserve inhabitant comfort. To this end, they will exploit context services, provided by the underlying infrastructure, to gain knowledge about the context and to provide the inhabitants with relevant information in a suitable way, following the ubiquitous computing principles (Weiser 1993). In our current implementation, the highest-level contextual information is produced by the situation and context identification layer exploiting the context spaces theory: the result is a set of occurring situations. As we explained in Sect. 2, the same situation can correspond to several different human activities and the same activity can require different forms of assistance depending on the particular situation. Thus, the artifact of situation space currently provided by our context spaces reasoning may not be sufficient to provide the targeted kind of assistance to inhabitants. To fill the gap between the situation spaces and the higher-level contextual information that we target, we need an additional mechanism that extracts a higher level of contextual knowledge from the underlying layer. Since the context spaces reasoning mechanisms only exploit the static contextual information provided by the perception layer, we need to mine additional information from the dynamics of the context. The main idea is that the context spaces theory provides very powerful modeling and reasoning mechanisms, but it can hardly handle the dynamism of context. Some techniques for context verification are developed in the theory, which help solve some ambiguities leveraging on historical context, namely, exploiting the *situation natural flow* (Padovitz et al. 2007). Furthermore, an extension to the context spaces theory has been proposed to perform context prediction (Boytsov, Zaslavsky, and Synnes 2009). Even though these techniques suggest the promising idea that context can be iteratively refined to solve ambiguities, they rely on the assumption that the real-world context obeys the same laws of a point following a trajectory inside a space. This assumption is too restrictive when willing to model the context dealing with the complex behavior of the inhabitants of a house, which often presents quite unpredictable evolutions. A different mechanism has to be adopted, which is able to capture relevant information from the context dynamics and to provide likely explanations for the observed situation sequences. The next Section presents an existing plan recognition algorithm and a way to adapt it to be integrated into our system in order to achieve these goals.

4 Activity Recognition using PHATT

In this section, we present PHATT, an algorithm introduced by Goldman, Geib and Miller in (Goldman, Geib, and Miller 1999) to perform plan recognition, and its application to our architecture. In order to do this, we first present the hierarchical task network planning problem, which is "inverted" by PHATT to perform plan recognition. Then, we show how PHATT can be adapted to be integrated into our system, in order to capture relevant information from the context dynamics.

4.1 Hierarchical Task Network Planning (HTN)

A *Hierarchical Task Network (HTN) planning problem* consists in automatically generating a *plan* starting from a set of *tasks* to execute and some constraints (La Placa, Pigot, and Kabanza 2009; Ghallab, Nau, and Traverso 2004). The problem relies on the specification of a plan library made of two components: the *tasks* to execute, which can be *primitive* if they don't ask for any further planning or *open*, otherwise, and the *methods*, which are prescriptions of how decomposing a task in (partially-) ordered sub-tasks. Note that a same task can be decomposed using different methods, thus resulting in different sub-task sequences. HTN planning proceeds by decomposing non-primitive tasks recursively into smaller and smaller subtasks, until primitive tasks are reached that can be performed directly.

4.2 PHATT

(Goldman, Geib, and Miller 1999; Geib and Goldman 2001; 2009) present PHATT, an algorithm for plan recognition based on a model of plan execution. The principle behind the algorithm is to perform plan recognition relying on three phases: defining the plan library, modeling the plan execution and recognizing the current execution, starting from the observations. The plan library is modeled like in the HTN planning problem presented above. The plan execution is modeled as a stochastic, generative model that selects actions to perform from a set of enabled primitive tasks called *pending set*, which is dynamically defined depending on the previous actions performed by the agent, the agent's goals and the plan library (Geib and Goldman 2009). Assuming this model of plan execution, PHATT takes as input a sequence of observations, which correspond to agent's actions, and generates the set of all possible explanations for the observed sequence of primitive tasks, in terms of executed plans and, thus, goals. It then uses Bayesian inference to calculate the probabilities of the generated explanations and goals.

4.3 Integration of PHATT with the Existing Architecture

In Sect. 3, we saw that the situation and context identification layer of our architecture, implemented exploiting the context spaces theory, lacks effective modeling and recognition of the temporal dimension of the activities. In this Section, we described an algorithm for plan recognition called PHATT; we propose now to adopt that algorithm to provide a way to model complex activities that develop over time. To this end, we show the advantages and a way of adapting PHATT to be integrated into our existing architecture, in order to start addressing the challenge of modeling and recognizing complex activities. We also provide an example that better explains the proposed modifications to the algorithm and the overall approach.

Critical aspects of PHATT and adaptation As showed in Sect. 3, the highest level of abstraction that our existing architecture provides is given by the situation spaces, whose abstraction is realized by the context spaces reasoning mechanism, which provides the set of ongoing high-level situations together with a value of confidence in their actual occurrence. As we said, we propose to integrate PHATT into our system in order to recognize complex activities that develop over time. Existing activity recognition approaches using *hidden Markov models* or *conditional random fields* consider the primitive tasks as elementary actions that can be performed by a person, e.g., use a spoon or a knife (Kim, Helal, and Cook 2010). Even in some of the proposed applications of PHATT, the primitive tasks model elementary actions that can be either directly observed or inferred by their effects (Goldman, Geib, and Miller 1999; Geib and Goldman 2001; 2009). In our case, the existing architecture already includes three layers that abstract high-level situation spaces from the elementary "observations" provided by sensors. For this reason, we propose to replace the primitive tasks of the plan library used by PHATT, which represent the actions that can be observed, with the situation spaces of the context spaces model. In this way, the observations are not elementary actions like those proposed in the original PHATT description (Goldman, Geib, and Miller 1999; Geib and Goldman 2001; 2009), but high-level situations occurring in the smart home. That is, situation spaces become the partially-ordered steps in which methods and root-level taks are decomposed.

Advantages of PHATT We now show the aspects of the complex activity that can be modeled and handled by PHATT and its underlying model of plan execution. In Sect. 2, we presented the domestic activity as opportunistic and constituted of multiple lines of different concerns. This implies that a same person can be preoccupied by multiple concerns at the same time. We also said that activity is never built according to a pre-established and hierarchical plan but is constantly reoriented according to inter-individual interactions and interactions with the physical environment. These aspects may look as being in contrast with the kind of plan library adopted by PHATT, which assumes that the activity can be modeled and fully specified as a hierarchy of alternative partially ordered sequential actions. Instead, even though PHATT relies on this kind of plan library, it adopts a separate model for plan execution, as showed in the previous Section. This model considers the possibility that several root-level tasks are executed in an interleaved fashion and it is thus able to produce explanations that contain multiple high-level goals.

In Sect. 2, we also said that the same behavior of a person can have several motivations, thus making it important to integrate several layers of inference from a low to a high contextual level. PHATT provides an additional contextual inference mechanism to our existing architecture, allowing to logically abducting the possible goals of inhabitants starting from their dynamic behavior. Starting from this knowledge about a person's goal, it may be easier to make decisions about the kind of assistance to provide in a particular situation, since the knowledge about the ongoing situations is in some ways enriched with their "motivation".

Example scenario We presented the advantages of PHATT and a way of integrating it into our existing architecture. We now show an example scenario, illustrating how to model it and recognize the ongoing activities using the previously presented tools. Notice that the scenario we chose is simple enough to allow an effective modeling and recognition of the involved activities using PHATT, while still being enough difficult to handle to represent an improvement to our existing architecture's capabilities (not exploiting PHATT). This scenario does not pretend to cover all the challenging aspects of the activity, which we presented in Sect. 2. Further work will investigate the aspects that we do not consider in this scenario. In Sect. 5, we provide some prospects about possible further adaptations of PHATT in order to model and recognize more complex activities.

Suppose John is involved in the concern of doing the housework. For this, he puts his house in order and does the washing. The washing machine is in a separate room that can be reached walking through a corridor, which also leads to other different rooms. John collects the washing from the whole house, reaches the laundry room with his arms loaded, and then turns the light on in the room. He loads the washing machine, turns it on, and then walks away to continue the housework, turning the light off. After some time, he decides to come back and check whether the washing cycle is over, discovering that it is not. This time, he leaves the light on, since he knows that he'll soon come back to unload the machine, as few minutes of washing are left. When he goes back to the kitchen, he notices that it is time to cook the lunch, so he opens the fridge to check what food is in it and then turns on his laptop to look for a recipe. In the meanwhile, the washing machine cycle is over and the light in the laundry room is still on. He finally decides to go back to the room and unload the machine. Then, he leaves the room, carrying the clean washing, and turns the light off, having

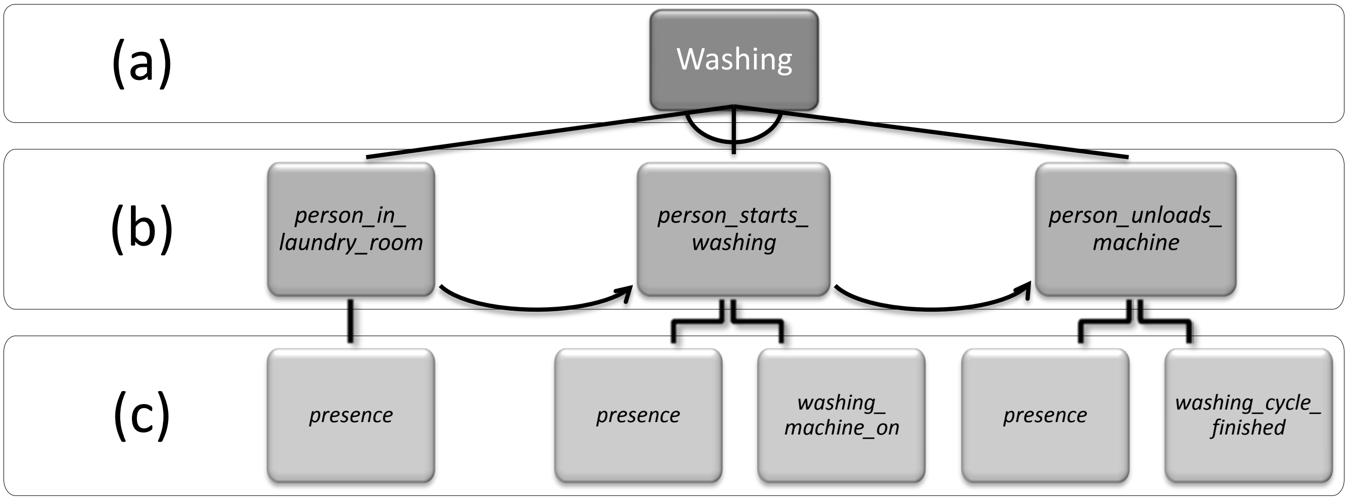

Figure 3: *A cross-layer model of the concern washing: the root-level task, (a), is decomposed in primitive tasks (situation spaces), (b), which in turn are obtained reasoning on context attributes, (c).*

troubles doing that with his arms loaded.

We now show how our current architecture can recognize the situations going on in the house and how PHATT can provide the missing activity recognition functionality. We model the root-level task `Washing` as the ordered sequence of situation spaces depicted in Figure 3.

When John enters the laundry room, the perception layer notifies the upper layer that the context attribute *presence* for the washing room has switched to value *true*. The context spaces reasoning notifies PHATT that the situation space *person-in-laundry-room* is occurring. Thus, PHATT detects that the `Washing` root-level task may be executed by the person. When John turns on the washing machine, the context spaces reasoning infers that the situation space *person-starts-washing* is occurring, which is interpreted by PHATT as the second step in the `Washing` task execution. Now John leaves the laundry room to go to the kitchen and start cooking. PHATT will observe other situation spaces occurring, e.g. *person-using-oven* and *person-using-hotplates*, primitive tasks of a root-level task `Cooking`. As we said, the model of activity execution that underlies PHATT allows multiple root-level tasks to be part of the same explanation. Thus, PHATT will generate an explanation for the observed situation spaces that contains both the goals `Washing` and `Cooking`. Until the situation space *person-unloads-machine* is observed, PHATT will consider the goal `Washing` as still active.

In our architecture, the output of PHATT can be reflected in the exploitation layer, offering contextual information that includes both the occurring situation spaces and the active goals. The augmented appliances can then exploit this information to influence their decisions. For example, if the person is walking in the corridor towards the room and the `Washing` goal is active, the probability that the person wants to enter the laundry room is higher than the probability that the person is going somewhere else. Also, the probability that the person enters the room to use the washing machine is higher than any other activity. The augmented light in the laundry room is notified with this contextual information and prepares to switch itself on as soon as the person actually enters the room. This is useful because other situations and activities may not require turning on the light. For instance, a person may just want to enter the laundry room, grab something and then exit without turning on the light. Knowing that the `Washing` goal is active, instead, helps deciding that the person may need enough light to operate the washing machine. In this way, the decisions taken by the augmented appliances to manage their behavior are helped by a double level of information: the situation spaces, obtained by statically analyzing the sensing data coming from the physical environment, and the root-level goals, produced by PHATT by analyzing the dynamics of situation spaces, generating all the possible explanations and evaluating their probabilities.

5 Open Issues

To conclude this paper, we present some aspects that we will investigate in future work.

In Sect. 4, we provided an example of domestic activity scenario. As we highlighted, that scenario is not representative of the complex model of activity that we presented in this paper.

For instance, the case of multiple people acting and interacting is not taken into account by the proposed scenario. In particular, the person that enters the laundry room may not be the one that is in charge of the laundry. Or, alternatively, the person may want to enter the laundry room for a different reason than doing the laundry. As we explained, taking into account these eventualities is important since it is impossible to model and reproduce what happens in inhabitants' minds. For this reason, future work will have to address the issue of non-interruptive takeover of the system. For instance, if the light automatically turns on in the laundry room, the person should be provided with a proximate interface to turn it off if preferred.

Another aspect of activity that we plan to consider shortly is the recurrence of concerns. Even though the time or order of actualization of concerns cannot be predicted, we can indeed take into consideration the recurrent nature of some of them. For this purpose, the concepts of *prior goal probability* and of influence of the *state of the world* introduced in (Goldman, Geib, and Miller 1999) look very promising.

The recognition of complex ambiguous behaviors of inhabitants could be improved following the principles described in (Coutaz et al. 2005), combining PHATT with the context spaces and with the perception layer using a *holistic* approach. In other words, we may use PHATT to provide feedback to the underlying context spaces reasoning and sensor data fusion layers. The feedback could be positive or negative, depending on the output of PHATT: if the top-ranked explanation has a high probability with respect to the others and as an absolute value, we may return to the lower layers a positive feedback. This feedback allows confirming the results of the context reasoning process and strengthening the belief in the sensor data fusion results, for instance exploiting the *conditioning* function described in (Smets and Kruse 1996).

Concerning the implementation aspects of PHATT, we need to carefully consider some important aspects, described in the rest of this Section.

(Geib and Goldman 2009) presents the assumption that each goal of an agent is known since the beginning of the execution. In Sect. 3, we said that human activity is characterized by inter-individual interactions and interactions with the physical environment, which result in an "on-the-fly" modification of the concerns the person is involved in. The assumption made by PHATT's implementation is thus clearly in contrast with our model of activity. Practically, the consequence of this assumption is that all the explanations and the probabilities have to be recalculated when a new goal is discovered. Future work will investigate the consequences of removing this assumption.

PHATT is implemented with the underlying assumption that all the actions performed by the agent are either directly observable or, in the case of *adversarial plan recognition* (Geib and Goldman 2001), inferable from their effects (changes in the *state of the world* or observation of "disabled" actions (Geib and Goldman 2001)). No concept of degree of uncertainty in the observations is considered, so there is no way to take into consideration the confidence value provided as output of the context spaces reasoning process. Further work will study ways to incorporate the confidence measure into PHATT or effective ways to choose confidence thresholds able to discern the occurrence of situation spaces.

We also need to investigate how the same situation space can be part of different root-level tasks. In (Goldman, Geib, and Miller 1999), the modeling allows the same primitive task to belong to multiple root-level tasks (the authors call *overloaded* these primitive tasks (Goldman, Geib, and Miller 1999)). In the implementation described in (Geib and Goldman 2009), this aspect is left as an open research question. In our system, different human concerns can reflect in the detection of the same situation spaces, and the same concern can be actualized in different situation spaces, leading to the need to model overloaded primitive tasks.

6 Conclusions

In this paper, we showed that the limitations of existing activity recognition approaches in the smart home domain are often due to the poverty of the adopted human-activity models. Combining expertise from the cognitive ergonomics and the ubiquitous computing fields, we discussed the hard technical challenges to address when leveraging on a realistic model of human activity. We presented the architecture of a smart home prototype that we are developing and showed the gap that exists between our current capabilities in terms of contextual-knowledge extraction and the complexity of the targeted activity recognition. To fill this gap, we proposed and discussed the integration of an existing algorithm for plan recognition into our system, in order to mine additional information from the dynamics of context.

Much work is still needed to perform activity recognition when accepting the difficult challenges raised by our complex naturalistic human activity model. Our ultimate aim is to extract a rich basis of contextual information to be used to provide the inhabitants of a smart home with adapted functionalities, targeted to obtain energy saving while preserving inhabitants' comfort.

References

Aldrich, F. 2003. Smart homes: Past, present and future. In Harper, R., ed., *Inside the Smart Home*. Springer London. 17–39.

Baillie, L., and Benyon, D. 2008. Place and technology in the home. *Computer Supported Cooperative Work (CSCW)* 17:227–256. 10.1007/s10606-007-9063-2.

Bonhomme, S.; Campo, E.; Esteve, D.; and Guennec, J. 2008a. Methodology and tools for the design and verification of a smart management system for home comfort. In *Intelligent Systems, 2008. IS '08. 4th International IEEE Conference*, volume 3, 24–2 –24–7.

Bonhomme, S.; Campo, E.; Esteve, D.; and Guennec, J. 2008b. PROSAFE-extended, a telemedicine platform to contribute to medical diagnosis. *J Telemed Telecare* 14(3):116–119.

Boytsov, A.; Zaslavsky, A.; and Synnes, K. 2009. Extending context spaces theory by predicting run-time context. In *NEW2AN '09 and ruSMART '09: Proceedings of the 9th International Conference on Smart Spaces and Next Generation Wired/Wireless Networking and Second Conference on Smart Spaces*, 8–21. Berlin, Heidelberg: Springer-Verlag.

Campo, E.; Bonhomme, S.; Chan, M.; and Esteve, D. 2006. Learning life habits and practices: an issue to the smart home. In C.Nugent, J., ed., *Smart homes and beyond, ICOST'2006 - 4st International Conference On Smart homes and health Telematic*, 355–358. Belfast: IOS Press.

Chahuara, P.; Vacher, M.; and Portet, F. 2010. Localisation d'habitant dans un environnement perceptif non visuel par propagation d'activation multisource. In *MAJECSTIC*, 8pp.

Chan, M.; Estve, D.; Escriba, C.; and Campo, E. 2008. A review of smart homes–present state and future challenges. *Computer Methods and Programs in Biomedicine* 91(1):55 – 81.

Chen, L., and Nugent, C. D. 2009. Ontology-based activity recognition in intelligent pervasive environments. *IJWIS* 5(4):410–430.

Coutaz, J.; Crowley, J. L.; Dobson, S.; and Garlan, D. 2005. Context is key. *Commun. ACM* 48:49–53.

Crabtree, A., and Rodden, T. 2004. Domestic routines and design for the home. *Comput. Supported Coop. Work* 13:191–220.

Geib, C. W., and Goldman, R. P. 2001. Plan recognition in intrusion detection systems. *DARPA Information Survivability Conference and Exposition,* 1:0046.

Geib, C. W., and Goldman, R. P. 2009. A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence* 173(11):1101 – 1132.

Ghallab, M.; Nau, D.; and Traverso, P. 2004. Hierarchical task network planning. 229 – 261.

Goldman, R. P.; Geib, C. W.; and Miller, C. A. 1999. A New Model of Plan Recognition. *Artificial Intelligence* 64:53–79.

Greenberg, S. 2001. Context as a dynamic construct. *Hum.-Comput. Interact.* 16:257–268.

Gu, T.; Wang, X. H.; Pung, H. K.; and Zhang, D. Q. 2004. An ontology-based context model in intelligent environments. 270–275.

Gu, T.; Wu, Z.; Tao, X.; Pung, H. K.; and Lu, J. 2009. epSICAR: An Emerging Patterns based approach to sequential, interleaved and Concurrent Activity Recognition. In *Proc. IEEE International Conference on Pervasive Computing and Communications PerCom 2009*, 1–9.

Guibourdenche, J.; Vacherand-Revel, J.; Grosjean, M.; Frjus, M.; and Haradji, Y. 2011. Using multiple scores for transcribing the distributed activities of a family. In *Proceedings of the 2011 ACM conference on Computer supported cooperative work*, CSCW '11. New York, NY, USA: ACM.

Kim, E.; Helal, S.; and Cook, D. 2010. Human activity recognition and pattern discovery. *IEEE Pervasive Computing* 9:48–53.

La Placa, M.; Pigot, H.; and Kabanza, F. 2009. Assistive planning for people with cognitive impairments. In *Proc. of Workshop on Intelligent Systems for Assisted Cognition hosted by Int'l Joint Conference on Artificial Intelligence (IJCAI)*.

Padovitz, A.; Loke, S.; Zaslavsky, A.; and Burg, B. 2007. Verification of uncertain context based on a theory of context spaces. *International Journal of Pervasive Computing and Communications* 3(1):30–56.

Padovitz, A.; Zaslavsky, A.; and Loke, S. W. 2006. A unifying model for representing and reasoning about context under uncertainty. In *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*.

Padovitz, A. 2006. *Context Management and Reasoning about Situations in Pervasive Computing*. Ph.D. Dissertation, Monash University, Australia.

Poizat, G.; Fréjus, M.; and Haradji, Y. 2009. Analysis of activity in domestic settings for the design ubiquitous technologies. In *European Conference on Cognitive Ergonomics: Designing beyond the Product — Understanding Activity and User Experience in Ubiquitous Environments*, ECCE '09, 14:1–14:2. VTT, Finland, Finland: VTT Technical Research Centre of Finland.

Ricquebourg, V.; Delafosse, M.; Delahoche, L.; Marhic, B.; Jolly-Desodt, A.; and Menga, D. 2007. Fault Detection by Combining Redundant Sensors: a Conflict Approach Within the TBM Framework. In *COGIS 2007, COGnitive systems with Interactive Sensors*. Stanford University.

Roy, P. C.; Bouchard, B.; Bouzouane, A.; and Giroux, S. 2010. *Web Intelligence and Intelligent Agents*. InTech. chapter Combining Pervasive Computing with Activity Recognition and Learning, 447–462. ISBN: 978-953-7619-85-5.

Salembier, P.; Dugdale, J.; Frejus, M.; and Haradji, Y. 2009. A descriptive model of contextual activities for the design of domestic situations. In *European Conference on Cognitive Ergonomics: Designing beyond the Product — Understanding Activity and User Experience in Ubiquitous Environments*, ECCE '09, 13:1–13:7. VTT, Finland, Finland: VTT Technical Research Centre of Finland.

Shelby, Z., and Bormann, C. 2009. *6LoWPAN: The Wireless Embedded Internet*. John Wiley & Sons, Ltd.

Smets, P., and Kruse, R. 1996. The transferable belief model for belief representation. In *Uncertainty Management in Information Systems*. Boston: Kluwer Academic Publishers. 343–368.

Theureau, J. 2003. *Handbook of cognitive task design*. New Jersey: Lawrence Erlbaum Associates. chapter Course-of-action analysis and course-of-action-centered design, 55–81.

Weiser, M. 1993. Some computer science issues in ubiquitous computing. *Commun. ACM* 36:75–84.