Is Ontology-based Activity Recognition Really Effective?

Daniele Riboni Linda Pareschi Laura Radaelli Claudio Bettini
Universita' degli Studi di Milano, D.I.Co., EveryWare Lab
via Comelico, 39, I-20135 Milano, Italy
Email: {riboni,pareschi,bettini}@dico.unimi.it, laura.radaelli@studenti.unimi.it

***Abstract*—While most activity recognition systems rely on data-driven approaches, the use of knowledge-driven techniques is gaining increasing interest. Research in this field has mainly concentrated on the use of ontologies to specify the semantics of activities, and ontological reasoning to recognize them based on context information. However, at the time of writing, the experimental evaluation of these techniques is limited to computational aspects; their actual effectiveness is still unknown. As a first step to fill this gap, in this paper, we experimentally evaluate the effectiveness of the ontological approach, using an activity dataset collected in a smart-home setting. Preliminary results suggest that existing ontological techniques underperform data-driven ones, mainly because they lack support for reasoning with temporal information. Indeed, we show that, when ontological techniques are extended with even simple forms of temporal reasoning, their effectiveness is comparable to the one of a state-of-the-art technique based on Hidden Markov Models. Then, we indicate possible research directions to further improve the effectiveness of ontology-based activity recognition through temporal reasoning.**

I. INTRODUCTION

There is a growing interest in the use of ontology-based techniques to automatically recognize complex context data such as human activities. In particular, in the area of pervasive computing, the popular ontological language OWL DL [1] has been used to build activity ontologies, and to recognize activities based on context data (e.g., in [2], [3]). The use of ontologies for recognizing human activities has some drawbacks: *a)* it requires good knowledge engineering skills, and significant expertise with the selected knowledge representation language; *b)* OWL DL has serious expressiveness limitations both in terms of the relationships that are needed to represent certain activities, and for the lack of support for temporal reasoning; *c)* ontological reasoning is computationally expensive.

We claim that point *a)* is not so critical, since domain and knowledge engineering experts can be found, and their effort can be shared. With respect to point *c)*, experimental evaluations have shown that the execution of ontology-based techniques is feasible to recognize human activities, at least in small- and medium-scale systems such as smart homes and smart workplaces (see, e.g., the experimental evaluation reported in [4]). However, with respect to point *b)*, at the time of writing, an experimental evaluation of the actual effectiveness of ontology-based activity recognition techniques is missing.

In this paper, we report preliminary experimental results about the effectiveness of ontology-based activity recognition, using a reference OWL-based activity recognition system, and an activity dataset collected in a smart-home setting. Based on a comparison with a data-driven technique relying on Hidden Markov Models (HMM), we indicate the need to extend ontological methods with temporal reasoning in order to improve recognition accuracy. Then, we discuss different possible approaches to incorporate temporal reasoning capabilities into ontology-based activity recognition.

II. ONTOLOGY-BASED ACTIVITY RECOGNITION

The ontological approach to activity modeling consists in a knowledge engineering task to define the formal semantics of human activities by means of the operators of the ontological language. Each activity is defined as a specialization of the abstract ACTIVITY class; for instance, a SOCIALACTIVITY can be defined as an ACTIVITY having more than one actor. Activities are arranged in a hierarchical fashion. Sub-activities are specializations of their parent activity: for instance, TEAPARTY can be defined as *"a specialization of* FRIENDLYMEETING*, in which the actors are sipping tea during the afternoon"*. Ontological reasoning is used to recognize that a user is performing a certain activity starting from some facts (e.g., sensor data, location of persons and objects, properties of actors involved).

The reference OWL-based architecture we adopt for our study is depicted in Figure 1. The design of the OWL ontology is done by means of graphical tools for ontology development that simplify design and testing, such as Protégé[1]. At run time, context information coming from distributed sources in the intelligent environment is retrieved and aggregated by the AGGREGATION middleware (CARE [5]), which is hosted by a possibly non mobile infrastructure. In particular, CARE interacts with the COSAR [3] system to retrieve information about simple human activities, recognized by hybrid ontological/statistical reasoners executed on users' personal mobile devices. Context data are mapped to ontological classes and properties by CARE, and added as instances to the ABox. Ontological reasoning to recognize

[1] http://protege.stanford.edu/

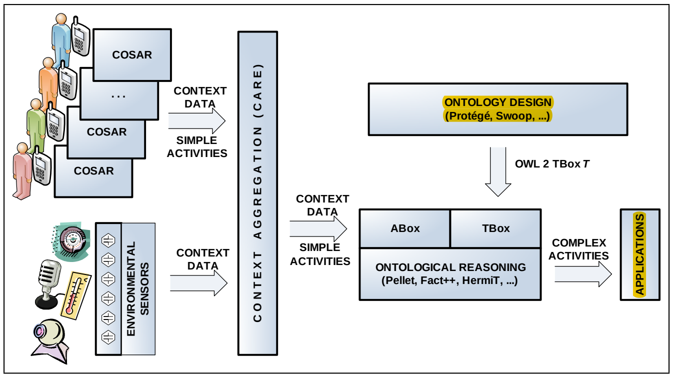

Figure 1. Reference architecture

complex human activities is performed, either periodically or on the occurrence of specific events, by an existing OWL reasoner executed on a dedicated server.

We have defined an OWL DL ontology for the activity recognition domain, which is published on the *PalSPOT* project website[2]. This ontology was derived from the OWL ontology presented in [3], which was used to refine the predictions of statistical activity recognition systems by means of symbolic reasoning. The innovative contribution of this ontology lies in the exploitation of the novel operators of OWL 2 [6] to represent activity axioms that could not be expressed in the former version of OWL.

III. Experimental evaluation

In this section we experimentally compare the ontological approach with a data-driven one that takes into account temporal information.

A. Dataset

In our experiments, we used a real-world dataset[3] of sensor readings, annotated with the activities performed by a person living in a smart-home for 28 days. The smart-home is a three-room apartment instrumented with 14 state-change sensors to detect object use; those sensors are attached to doors and different household appliances. The dataset, as well as the smart-home system, has been presented in [7]. According to the dataset, at each minute the user is either performing a single activity, or he is idle (i.e., concurrent activities are not considered). Eight activities of daily living (ADLs) are monitored, including having dinner, toileting, showering, sleeping.

B. Data-driven technique based on HMMs

In order to compare the ontological approach with a data-driven one, we have implemented the technique described in [7], which is based on Hidden Markov Models (HMMs). Techniques based on HMMs are known to yield state of the art results in recognizing ADLs in smart-home settings. The state duration is 60 seconds; i.e., activity recognition is performed every one minute. Being based on HMMs, with this technique, activity prediction at minute τ depends not only on the observation of the current context, but also on the activity predicted at the previous minute (τ-1), which in turn depends on context information at τ-1 and on the activity predicted at minute τ-2, and so on.

	Timeslice	Class
Data-driven (HMMs)	0.932	0.731
“Snapshot” ontology	0.937	0.530
“One-step” ontology	0.912	0.791
“Multi-step” ontology	0.925	0.803

Table I
Accuracy of the different activity recognition methods

We have repeated the experiments in [7], obtaining substantially equivalent results. Table I reports the accuracy obtained by the different techniques evaluated in the paper. *Timeslice* accuracy is the percentage, over the whole set of one-minute timeslices, of correct activity predictions, while *class* accuracy is the average percentage of correct predictions per activity. For the sake of the experiments about the data-driven technique, we used a “leave-one-day-out” cross-validation method. As it can be observed in Table II (a), activities “leaving house” and “sleeping” have very high recognition rates; since these activities are the most frequent ones, the timeslice accuracy of the data-driven technique is high (0.932). On the contrary, the recognition rate of less frequent activities is lower; hence, class accuracy is only 0.731.

C. “Snapshot” activity ontology

For the sake of the experiments about the ontological approach, we divided the dataset in a 5-days *observation set* and in a 23-days *test set*. We used the observation set to analyze the user’s behavior in order to define the semantics of activities, and the test set to evaluate the effectiveness of recognition based on ontological reasoning. We point out that, in these experiments, the ontological definitions of activities are tightly tailored both to the characteristics of the smart-home system from which data have been collected, and to the specific user’s habits; they are not general definitions of ADLs.

The most common ontology-based approach to activity recognition (e.g., [2], [3]) consists in specifying the semantics of activities based on the observation of a user’s *current* context (current location, current time, objects that the individual is using, etc); we name this approach “snapshot”.

Example 1: Based on the observation of the user’s behavior, we defined “getting drink” (i.e., the generalization of activities “preparing drink” and “drinking”) as an individual activity taking place in a kitchen or in a living room, when

[2]http://everywarelab.dico.unimi.it/palspot
[3]http://staff.science.uva.nl/~tlmkaste/research/software.php

(a) Confusion matrix for the data-driven approach (HMMs)

	1	2	3	4	5	6	7	8
1	**0.61**	0.06	0.01	0.1	0.05	0	0.16	0.01
2	0.01	**0.98**	0	0	0	0	0	0
3	0.13	0.05	**0.72**	0.05	0.04	0.01	0.01	0
4	0.11	0.01	0.06	**0.83**	0	0	0	0
5	0	0	0	0	**0.99**	0	0	0
6	0.28	0	0.01	0	0.01	**0.62**	0.05	0.04
7	0.21	0.01	0	0	0	0.09	**0.67**	0.01
8	0.17	0.05	0.03	0	0	0.12	0.19	**0.44**

(b) Confusion matrix for the "snapshot" ontology approach

	1	2	3	4	5	6	7	8
1	**0.89**	0.07	0	0	0.03	0	0.01	0
2	0.02	**0.98**	0	0	0	0	0	0
3	0.35	0.02	**0.48**	0.02	0.13	0	0	0
4	0.83	0	0.01	**0.17**	0	0	0	0
5	0.07	0	0	0	**0.93**	0	0	0
6	0.53	0	0.04	0.01	0.01	**0.31**	0	0.1
7	0.84	0	0	0	0	0	**0.16**	0
8	0.51	0.03	0.02	0	0	0	0.12	**0.32**

(c) Confusion matrix for the "one-step" ontology approach

	1	2	3	4	5	6	7	8
1	**0.55**	0.06	0.02	0.06	0.03	0.05	0.22	0.01
2	0.01	**0.98**	0	0	0	0	0	0
3	0.10	0.02	**0.73**	0.02	0.11	0.01	0.02	0
4	0.10	0	0.01	**0.88**	0	0.01	0	0
5	0.07	0	0	0	**0.93**	0	0	0
6	0.17	0	0.04	0.02	0.01	**0.73**	0	0.03
7	0.11	0	0.01	0	0	0	**0.87**	0
8	0.09	0.03	0.03	0	0	0	0.20	**0.64**

(d) Confusion matrix for the "multi-step" ontology approach

	1	2	3	4	5	6	7	8
1	**0.67**	0.06	0.02	0.06	0.03	0.01	0.14	0.01
2	0.01	**0.98**	0	0	0	0	0	0
3	0.11	0.02	**0.73**	0.02	0.11	0	0.02	0
4	0.11	0	0.01	**0.88**	0	0	0	0
5	0.07	0	0	0	**0.93**	0	0	0
6	0.17	0	0.04	0.02	0.01	**0.73**	0	0.03
7	0.12	0	0.01	0	0	0	**0.86**	0
8	0.10	0.03	0.03	0	0	0	0.19	**0.64**

Activity labels: 1 = Idle; 2 = Leaving House; 3 = Toileting; 4 = Showering; 5 = Sleeping; 6 = Breakfast; 7 = Dinner; 8 = Getting Drink

Table II
CONFUSION MATRICES

the actor is using the fridge and the cups cupboard:

$$\begin{aligned}\textsc{GettingDrink} \sqsubseteq\ & \textsc{IndividualActivity} \sqcap \\ & \forall\, \textsc{hasActor}.\Big(\textsc{Person} \sqcap \\ \exists\, \textsc{currentLocation}.(&\textsc{Kitchen} \sqcup \textsc{LivingRoom}) \sqcap \\ & \exists\, \textsc{usingArtifact}.\textsc{CupsCupboard} \sqcap \\ & \exists\, \textsc{usingArtifact}.\textsc{Fridge}\Big).\end{aligned}$$

In the snapshot ontology, property USINGARTIFACT relates a person to those objects s/he is using in the current one-minute timeslice.

Note that multiple definitions of each activity can be stated, in order to capture different situations under which the activity is typically executed.

Based on the observation set, we extended the ontology with novel class definitions for the activities considered in the dataset. For each timeslice, we added to the ABox the sensor information extracted from the dataset; we exploited the Java APIs for the Pellet[4] reasoner to perform the realization reasoning task, in order to derive the specific activity performed by the user in the current timeslice.

The snapshot ontology technique slightly outperforms the data-driven one in terms of timeslice accuracy. This result is due to the fact that the snapshot ontology method obtains high recognition rates for activity "idle", which is very frequent in the dataset. On the contrary, in terms of class accuracy, the snapshot ontology method obtains poor results, since the recognition rate of 4 activities out of 8 is less than 0.33 (see Table II (b)). Hence, overall, the snapshot ontology approach underperforms the data-driven one.

D. *"One-step" activity ontology*

A main limitation of the snapshot ontology approach is the lack of support for the temporal characterization of activities. As anticipated in the introduction, handling temporal information within OWL DL ontologies is problematic, due to different limitations of its underlying description logics language. Hence, for the sake of these experiments, we adopted an OWL-compliant solution, based on the simplifying assumption that the user performs a single activity at a time. We stress that this assumption, which is made in most of the activity recognition literature, is unrealistic; however, it is necessary to enable a simple form of reasoning with past activities in OWL DL. Even if this solution –as well as the one proposed in the "multi-step" case– is not adequate to satisfactorily model human activities, it is useful to demonstrate the importance of temporal reasoning.

In the one-step activity ontology, we modified the class definitions to better characterize activities, based on the objects recently used by the actor, and on the activity executed by the user immediately before the current one. In particular, we introduced a RECENTLYUSED property to relate the actor to the objects s/he used during the last three minutes timeslices. Since OWL DL does not support temporal reasoning, we used an external Java application to keep track of recently used objects, and to add assertions

[4]http://clarkparsia.com/pellet/

to the ABox about the RECENTLYUSED property. Property LASTACTIVITY relates the actor to the activity s/he performed before the current one. Once again, assertions regarding the above property were handled by an external Java application.

Example 2: In our one-step activity ontology, we introduced the definition of "drinking" to specialize activity "getting drink": according to the definition below, this activity immediately follows "preparing drink" (another specialization of "getting drink"). Hence, each time "drinking" or "preparing drink" are recognized, the ontological reasoner infers that the current activity is "getting drink". We also substituted property USINGARTIFACT with RECENTLYUSED to take into account the objects used in the near past:

$$
\begin{aligned}
\text{DRINKING} \sqsubseteq\ & \text{INDIVIDUALACTIVITY} \sqcap \\
& \forall\, \text{HASACTOR}.(\text{PERSON} \sqcap\ (\ldots) \sqcap \\
& \exists\, \text{RECENTLYUSED.CUPSCUPBOARD} \sqcap \\
& \exists\, \text{RECENTLYUSED.FRIDGE} \sqcap \\
& \exists\, \text{LASTACTIVITY}.(\text{PREPARINGDRINK})).
\end{aligned}
$$

With respect to timeslice accuracy, this technique obtains slightly worse results than the data-driven one. However, it significantly outperforms it in terms of class accuracy (the confusion matrix is shown in Table II (c)).

E. "Multi-step" activity ontology

In order to incorporate more temporal information, we extended the one-step activity ontology with a property SECONDLASTACTIVITY to relate the actor to the second-last activity s/he performed before the current one. Using this property we were able to better characterize part of the considered activities.

Example 3: Observing the dataset used in our experiments, we noticed that, before "dinner", the user usually performs activities "preparing food" and "preparing drink" (those activities can be recognized by the use of specific artifacts such as the microwave oven, and the cups cupboard). Then, a possible definition of "dinner" is:

$$
\begin{aligned}
\text{DINNER} \sqsubseteq\ & \text{ACTIVITY} \sqcap \\
& \forall\, \text{HASACTOR}.(\text{PERSON} \sqcap\ (\ldots) \\
& \exists\, \text{LASTACTIVITY.PREPARINGDRINK} \sqcap \\
& \exists\, \text{SECONDLASTACTIVITY.PREPARINGFOOD}).
\end{aligned}
$$

Other definitions are possible; for instance, one in which "preparing drink" precedes "preparing food".

We also exploited the information about the duration of the user's activities to refine the predictions of the ontological reasoner, as illustrated below.

Example 4: In the above axiom, activity "breakfast" is characterized by a maximum duration:

$$
\begin{aligned}
\text{BREAKFAST} \sqsubseteq\ & \text{ACTIVITY} \sqcap\ (\ldots) \sqcap \\
& \forall\ (\text{HASDURATION} \leq \text{BFMAXDURATION}),
\end{aligned}
$$

where HASDURATION is the duration of the current activity, and BFMAXDURATION is the maximum duration for activity "breakfast".

For the sake of these experiments, we modeled the duration of a user's activities by a Gaussian probability distribution $N(\mu, \sigma^2)$; mean μ and variance σ^2 were calculated based on the first 5 days of the dataset. We have set the maximum duration as the value τ for which the probability of the activity to last more than τ is 5%.

With respect to timeslice accuracy, this technique obtains substantially equivalent results with respect to the data-driven one; however, it outperforms the other techniques in terms of class accuracy, as it can be observed in Table II (d). However, we claim that this technique is still insufficient to capture the complexity of the possible temporal relations among activities. Possible extensions of the ontological approach to reason with temporal intervals will be discussed in the following section.

IV. DISCUSSION

The experimental evaluation reported in Section III shows the importance of considering temporal information to effectively recognize human activities. Indeed, comparing the confusion matrices regarding the snapshot and the one-step ontology approaches (Tables II (b) and II (c), respectively), the information about the activity previously executed by the user seems to be particularly important to distinguish among interleaved activities, like "dinner" and "getting drink", or "sleeping" and "toileting". In fact, the order in which these activities are performed follows typical patterns; e.g., activity "dinner" is typically interleaved by "getting drink".

Information about the typical duration of activities is also important to refine ontological activity predictions. For instance, observing the dataset used in our experiments, we noticed that, after dinner, the user is idle for a considerable amount of time, before performing a new activity. During this idle period, the individual does not use any artifact; hence, based on the available information, the ontological technique keeps predicting "dinner" as the current activity. By considering the typical duration of activities, we were able to considerably improve the accuracy for "idle", without degrading the accuracy for "breakfast" and "dinner" (see Table II (d)). Results show that accuracy is also improved by considering not only the user's last activity, but also the second-last activity.

However, the OWL-compliant representations we used for our experiments are insufficient to satisfactorily model the temporal characterization of activities. Indeed, with an OWL-compliant solution, it is not possible to perform interval-based temporal reasoning, which seems to be required for capturing complex temporal relations among activities. For instance, Figure 2 shows one of the possible interval-based representations of activities, in which "preparing drink" takes place before "dinner", and overlaps with

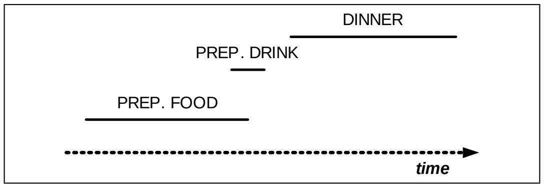

Figure 2. An interval-based representation of activities

"preparing food".

We now discuss the support for temporal reasoning in existing activity recognition techniques. Data-driven techniques usually exploit temporal information; one of the most common solutions is to model the evolution of a user's activities by a Markov model, and to use HMMs to recognize them based on sensor data (like proposed, for instance, in [7]). Many variants of HMMs have been adopted to more extensively exploit temporal information, like activity duration (e.g., in [8]), and variable-length history of past activities (e.g., in [9]).

On the contrary, temporal reasoning is not supported by ontological activity recognition techniques proposed so far, which generally adopt the OWL DL language. Indeed, OWL DL enables the representation of time (e.g., with the OWL-Time[5] ontology), but does not support any form of temporal reasoning. This is due to some restrictions to the language operators of its underlying description logic, necessary for preserving the decidability of reasoning problems.

However, we claim that temporal reasoning must be integrated with ontological techniques to effectively recognize activities. In order to provide description logics with temporal reasoning capabilities, two main approaches exist. The first one consists in the use of temporal description logics [10], in which the temporal and terminological domains are tightly integrated. However, in order to preserve decidability, these languages do not support some expressive operators that are admitted in OWL DL; this makes it difficult to model complex activities. The second approach consists in the use of loosely-coupled techniques, in which temporal reasoning is performed by an external reasoner. With this approach, time is considered a *concrete domain* [11]; individuals in the ontology are related to values of the temporal domain by functional properties. Then, it is possible to use an external reasoner to deal with qualitative and/or quantitative relationships among the time intervals corresponding to activities durations (e.g., "activity x must follow activity y and overlap with activity z"). We believe that the investigation of the latter approach is a promising research direction to improve the results of ontology-based activity recognition.

[5] www.w3.org/TR/owl-time/

V. Conclusions

In this paper, we experimentally evaluated the effectiveness of ontology-based activity recognition, using a dataset collected in a smart-home setting. Results highlight the importance of including temporal reasoning in ontological techniques to effectively recognize activities. We indicated the investigation of loosely-coupled techniques, based on the use of concrete domains, as a promising research direction.

Acknowledgments

The authors would like to thank Tim van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Ben Kröse, for providing the activity dataset used in our experiments.

References

[1] I. Horrocks, P. F. Patel-Schneider, and F. van Harmelen, "From SHIQ and RDF to OWL: The making of a Web ontology language," *J. Web Sem.*, vol. 1, no. 1, pp. 7–26, 2003.

[2] L. Chen and C. D. Nugent, "Ontology-based activity recognition in intelligent pervasive environments," *Int. J. Web Inf. Syst.*, vol. 5, no. 4, pp. 410–430, 2009.

[3] D. Riboni and C. Bettini, "COSAR: hybrid reasoning for context-aware activity recognition," *Pers. Ubiq. Comput.*, to appear. [Online]. Available: http://dx.doi.org/10.1007/s00779-010-0331-7

[4] A. Agostini, C. Bettini, and D. Riboni, "Hybrid reasoning in the CARE middleware for context awareness," *Int. J. Web Eng. Technol.*, vol. 5, no. 1, pp. 3–23, 2009.

[5] C. Bettini, L. Pareschi, and D. Riboni, "Efficient profile aggregation and policy evaluation in a middleware for adaptive mobile applications," *Perv. Mob. Comput.*, vol. 4, no. 5, pp. 697–718, 2008.

[6] B. C. Grau, I. Horrocks, B. Motik, B. Parsia, P. F. Patel-Schneider, and U. Sattler, "OWL 2: The next step for OWL," *J. Web Sem.*, vol. 6, no. 4, pp. 309–322, 2008.

[7] T. van Kasteren, A. K. Noulas, G. Englebienne, and B. J. A. Kröse, "Accurate activity recognition in a home setting," in *Proc. of UbiComp'08*, ser. ICPS, vol. 344. ACM, 2008, pp. 1–9.

[8] T. V. Duong, H. H. Bui, D. Q. Phung, and S. Venkatesh, "Activity recognition and abnormality detection with the switching hidden semi-markov model," in *Proc. of IEEE CVPR'05*. IEEE Comp. Soc., 2005, pp. 838–845.

[9] A. Galata, N. Johnson, and D. Hogg, "Learning variable-length markov models of behavior," *Comp. Vis. Image Under.*, vol. 81, no. 3, pp. 398–413, 2001.

[10] A. Artale and E. Franconi, "A survey of temporal extensions of description logics," *Ann. Math. Artif. Intell.*, vol. 30, no. 1-4, pp. 171–210, 2000.

[11] C. Lutz and M. Milicic, "A tableau algorithm for description logics with concrete domains and general tboxes," *J. Autom. Reasoning*, vol. 38, no. 1-3, pp. 227–259, 2007.