Towards Activity Recognition Using Probabilistic Description Logics

Rim Helaoui1, Daniele Riboni2, Mathias Niepert1, Claudio Bettini2, Heiner Stuckenschmidt1

1 KR&KM research group, University of Mannheim
2 EveryWare Lab, Universita’ degli Studi di Milano, D.I.Co.

Email: {rim,mathias,heiner}@informatik.uni-mannheim.de, {daniele.riboni,claudio.bettini}@unimi.it

Abstract
A major challenge of pervasive context-aware computing and intelligent environments resides in the acquisition and modelling of rich and heterogeneous context data. Decisive aspects of this information are the ongoing human activities at different degrees of granularity. We conjecture that ontology-based activity models are key to support interoperable multilevel activity representation and recognition. In this paper, we report on an initial investigation about the application of probabilistic description logics (DLs) to a framework for the recognition of multilevel activities in intelligent environments. In particular, being based on Log-linear DLs, our approach leverages the potential of highly expressive description logics with probabilistic reasoning in one unified framework. While we believe that this approach is very promising, our preliminary investigation suggests that challenging research issues remain open, including extensive support for temporal reasoning, and optimizations to reduce the computational cost.

Introduction
The spectacular progress of low-cost and low-power sensing has given rise to appealing computing fields. Designated as pervasive computing, this new paradigm is witnessing an increasing interest and a growing research community. Especially, the concept of emerging intelligent environments is progressively evolving into a commonplace. Such environments necessitate context awareness to support and assist the user with reactive and proactive services. Ongoing human activities are a decisive aspect of such contextual information. Hence, many researchers have been recently concerned with automatically recognizing human activities from lightweight dense sensing. Accordingly, several approaches have been proposed that can be classified as data-driven, knowledge-based, or hybrid approaches. A multitude of works tried to apply diverse machine learning algorithms to recognize human activities. The majority proposed supervised learning methods such as Hidden Markov Models (Patterson et al. 2005) and conditional Random Fields (Buettnner et al. 2009). The bottleneck of obtaining large amounts of training data has motivated some recent works to explore weakly supervised learning (Stikic and Schiele 2009) as well as unsupervised learning (Gu et al. 2010).

Despite being well suited for simple activities, data-driven techniques have a number of problems with the recognition of complex high-level activities. Their poor portability as well as the severe scalability problems they face, make them hardly deployable in other environments. Adding new activities without newly training and possibly designing the given model remains also out of reach. Furthermore, these techniques are generally doomed to a flat representation of human activities which contradicts their hierarchical nature. Finally, the lack of formal semantics prevents data-driven models from encoding the inherent common-sense knowledge underlying human activities.

As an alternative to data-driven approaches, some researchers adopted different logical modelling and reasoning algorithms to address human activity recognition. These include (Cirillo et al. 2009) where general predefined rules are used to recognize activities based on a constraint-based temporal reasoning framework (OMPS). In (Bouchard, Giroux, and Bouzouane 2006), the authors leverage action description logics and lattice theory for plan recognition to predict human behaviour. Despite their ability to cope with some of the limitations of data-driven approaches, rule-based systems suffer from many restrictions, including limited support for uncertainty and temporal reasoning.

Combining both paradigms has been the motivation of our previous works (Helaoui, Niepert, and Stuckenschmidt 2011) and (Riboni and Bettini 2011). In the former, we applied Markov Logic Networks to unite both logical statements as well as probabilistic ones in one single framework. This allows to address the uncertain aspect of human activities and integrate crucial background knowledge. Due to the specific rules and the learning phase, the same system can not be reused under different settings. The capability of thoroughly exploiting the semantic features between the context and the activities leaves much to be desired. In the latter, we proposed a loosely coupled technique to integrate supervised learning methods with DLs. We defined an ontology to formally model the semantics of activities, and exploited ontological reasoning to refine the prediction of the machine learning algorithm based on the current context. While useful for improving the recognition rates of simple actions, that method was not well suited to recognize complex ac-
tivities, since the ontological language lacked support for uncertainty and temporal reasoning.

In this paper, we explore the convergence of our previous works by the use of probabilistic DLs to exploit semantic features of dense sensing for activity recognition. Compared to symbolic activity recognition based on manual specification of a set of rules, our approach aspires to facilitate portability and interoperability of the recognition system. As such, we propose to use ontologies as a commonly shared conceptual knowledge description standard. To leverage the full potential of DLs, we use the same unified language for both modelling multi-level activities and their context and probabilistically reasoning about them. This hybrid framework is achieved through adoption of log-linear DLs as proposed in (Niepert, Noessner, and Stuckenschmidt 2011).

The rest of the paper is structured as follows. In the next section we describe related works. Then, we present our overall recognition framework and techniques. Finally, we conclude the paper with a discussion of our approach and future work.

Related Work

In this section, we describe approaches closely related to our work. Especially, we discuss those that utilize ontologies and semantic information to enable and improve activity recognition. There are only few works that keep both semantic description of the activities and their recognition process tightly-coupled. Such an approach has been adopted in (Chen, Nugent, and Wang 2011) and in (Springer and Turhan 2009). Similarly to our work, the authors load the current contextual information to populate their generated ontology then employ inference reasoner to obtain the most specific equivalent activity or situation. Hence, activity recognition can be mapped to equivalency and subsumption and Turhan 2009). Similarly to our work, the authors load

The rest of the paper is structured as follows. In the next section we describe related works. Then, we present our overall recognition framework and techniques. Finally, we conclude the paper with a discussion of our approach and future work.

Representing and Recognizing Multilevel Activities with Log-Linear DLs

In this section, we describe our technique for multilevel activity recognition.

System overview

As shown by the following example, an activity recognition system should be able to recognize activities at both coarse- and fine-grained level of detail.

Example 1. Consider a healthcare system to remotely monitor the activities of an elderly person. Recognized activities are periodically communicated to a medical center, in order to evaluate the evolution of the patient’s physical and cognitive capabilities. A similar system should be able to recognize not only high-level activities such as “preparing breakfast” and “cleaning house”, but also the more specific actions that compose those activities. For instance, it should be possible, for the medical center, to inspect the simpler activities that were executed during breakfast preparation (e.g., washing fruit, cooking egg, fill a jug with milk, . . .), to better evaluate the patient’s behavior.

Figure 1 shows a multilevel decomposition of complex activities (Level 1) in more specific components; we adopt the structure proposed in (Lukowicz et al. 2010) for the Opportunity dataset. Level 2 represents simple activities, which can be described as temporal sequences of manipulative gestures and modes of locomotion (Level 3). Finally, Level 4
describes the atomic gestures that characterize manipulative gestures.

As shown in Figure 1, atomic gestures may be recognized through the application of supervised learning methods, based on data acquired from body-worn sensors (e.g., accelerometers to track movements of the body), environmental sensors, and RFID tags to detect objects usage. In order to achieve the goal of higher-level activity recognition, in our technique we propose the use of DLs to express background knowledge about the domain, and to model the relationships among high-level activities and their simpler component actions. In order to cope with the uncertainty of both context facts and activity definitions, we adopt log-linear description logics (DLs), a probabilistic DL that has been recently proposed in (Niepert, Noessner, and Stuckenschmidt 2011).

Description Logics and Log-linear Description Logics

Description logics are a commonly used representation for knowledge bases. There are numerous tools and standards for representing and reasoning with knowledge using DLs. The DLs framework allows one to represent both facts about individuals (concept and role assertions) as well as axioms expressing schema information. Log-linear DLs integrate description logics with probabilistic log-linear models (Niepert, Noessner, and Stuckenschmidt 2011). In particular, Log-linear DLs allow to model both probabilistic and deterministic dependencies between DL axioms through extending uncertain axioms with weights. Regarding the expressiveness of the language, Log-linear DL supports the same operators of the well-known OWL 2 language.

The syntax of log-linear DLs is equivalent to that of the underlying DL except that it is possible to assign weights to general concept inclusion axioms (GCIs), role inclusion axioms (RIs), and assertions. We explicitly allow assertions (facts about individuals) in the knowledge base and we will use the terms constraint box (CBox) and knowledge base (KB) interchangeably. Moreover, for ease of presentation, we will use the term axiom to denote GCIs, RIs, and concept and role assertions. A log-linear knowledge base $C = (C^D, C^U)$ is a pair consisting of a deterministic CBox $C^D$ and an uncertain CBox $C^U = \{\{c, w_c\}\}$ with each $c$ being an axiom and $w_c$ a real-valued weight assigned to $c$. The deterministic CBox contains axioms that are known to hold and the uncertain CBox contains axioms with weights. The greater the weight of an axiom the more evidence there is for it to hold. Each axiom can either be part of the deterministic or the uncertain CBox but not both. The deterministic CBox is assumed to be coherent and consistent.

The simple semantics of log-linear DLs is based on probability distributions over coherent and consistent knowledge bases. The weights of the axioms determine the log-linear probability distribution. For a log-linear knowledge base $C = (C^D, C^U)$ and a CBox $C'$ with $C^D \subseteq C' \subseteq C^D \cup \{c : (c, w_c) \in C^U\}$, we have that

$$\Pr_C(C') = \begin{cases} \frac{1}{Z} \exp \left( \sum_{c \in C' \setminus C^D} w_c \right) & \text{if } C' \text{ is coherent and consistent;} \\ 0 & \text{otherwise} \end{cases}$$

where $Z$ is the normalization constant of the log-linear distribution $\Pr_C$. An axiom with weight 0 that is not in conflict with any other axiom has the marginal probability of 0.5. Hence, the semantics of log-linear DLs leads to a distribution compatible with the open-world assumption.

**Example 2.** Let $S$ and $P$ be two concepts and let $C = (C^D, C^U)$, with $C^D = \emptyset$ and $C^U = \{(S \sqsubseteq P, 0.8), (P \sqsubseteq S, -1.0), (S \sqcap P \subseteq \bot, 0.4)\}$. Then,
ties “put in dishwasher” and “clean table”. Hence, we can activity “clean up” in terms of its component simple actions and current context data. Suppose that we want to represent the complex activity “clean up”, which are associated to weighted axioms.

Example 3. Suppose that we want to represent the complex activity “clean up” in terms of its component simple activities “put in dishwasher” and “clean table”. Hence, we can add the following axioms to the knowledge base:

\[
\text{CleanUp} \sqsubseteq \text{ComplexActivity} \quad (1)
\]

\[
\forall \text{hasActor.} \{\text{Person}\} \quad (1)
\]

\[
\exists \text{hasSimpleActivity.} \text{PutInDishwasher} \quad , 1.8
\]

\[
\text{CleanUp} \sqsubseteq \text{ComplexActivity} \quad (2)
\]

\[
\forall \text{hasActor.} \{\text{Person}\} \quad (2)
\]

\[
\exists \text{hasSimpleActivity.} \text{CleanTable} \quad , 1.5
\]

Axiom (1) has weight 1.8, and essentially states that if a person is currently putting things in the dishwasher, she is probably performing the complex activity “clean up”. Axiom (2) is similar, apart that it has a slightly lower weight, and considers activity “clean table” instead of “put in dishwasher”. Note that the axioms weights can be manually defined based on background knowledge, or automatically learned from a training set of performed activities.

A major issue with our approach is that the description logic underlying Log-linear DL (which has essentially the same expressiveness of OWL 2 (Grau et al. 2008)) does not natively support temporal reasoning. This is a serious limitation, since reasoning with temporal intervals is needed to recognize many human activities. Consider, for instance, the simple activity “put in dishwasher” depicted in Figure 1. A possible situation describing that activity is: “the individual takes an object, walks to the dishwasher, opens its door; puts the object inside, and closes the dishwasher door”. Other definitions for the same activity may exist, based on the individual’s habits. For instance, many objects may be put inside the dishwasher while its door is left open. In order to express these temporal relationships among activities and actions, we should be able to reason with temporal intervals. Unfortunately, this kind of reasoning is not naturally supported by the description logic underlying Log-linear DL, due to some restrictions to the language operators that are necessary for preserving the decidability of ontological reasoning problems. This shortcoming is shared by any activity recognition system strictly based on OWL 2. On the contrary, DLs for interval-based temporal languages, like the one presented in (Artale and Franconi 1998), are well suited for reasoning with the temporal characteristics of activities and actions, but lack support for some expressive operators that are admitted in OWL 2, which are needed to model complex context data. An alternative approach we are considering is to exploit an external reasoner to perform temporal reasoning, while maintaining the whole expressiveness of our ontological language. We further discuss this issue in the last section of this paper.

Our Ontology for the Opportunity Dataset

As a testbed for our technique, we chose to use the dataset acquired within the Opportunity EU research project (Activity and Context Recognition with Opportunistic Sensor Configurations)\(^1\). The dataset has been acquired in a smart home

\(^1\)http://www.opportunity-project.eu/
with dense sensing equipments (72 sensors of 10 modalities, including accelerometers, RFID tags attached to objects, and environmental sensors for audio and video). A number of multilevel activities in an “early morning” setting has been acquired by 12 subjects. Gestures and activities are structured similarly as shown in Figure 1. The dataset is presented in detail in (Lukowicz et al. 2010).

The Opportunity dataset provides labels for atomic gestures (Level 4). The goal of our experiments is to evaluate the feasibility of our technique to recognize gestures and activities at Levels 3, 2, and 1 starting from atomic gestures. For the sake of these experiments, we will assume that a state-of-the-art method is used to recognize activities at Level 4 based on raw context data. For our experiments, we have developed an ontology for representing the multilevel activities considered in the Opportunity dataset. The ontology has been developed using the Protégé OWL editor (Knublauch et al. 2004); axioms weights have been added as annotations to axioms definitions. Figure 2 shows a snapshot of our ontology.

**Determining Axioms Weights**

According to our log-linear model, the marginal probability of an axiom is the sum of the probabilities of each consistent and coherent ontology in which that axiom holds. We adopt the association rules learning principle to obtain a-priori probabilities of individual axioms: We first expand our training data with the corresponding multi-level activity labels. This is achieved based on the concept definitions provided in the *TBox*. Each of these axioms is considered as a rule. For example, considering the lowest activity level, the *Atomic Gesture* would be the antecedent and the *Manipulative Gesture* the consequent. After deleting the consecutive duplicate events, we determine the support value for the considered rules which actually coincide with the probability values of the corresponding axioms. In the example above, the obtained support value corresponds to the conditional probability of *Manipulative Gesture given an Atomic Gesture*. Since the semantics of log-linear DLs is based on log-linear probability distributions, we compute the a-priori weights by taking the log of the a-priori probabilities.

**The Recognition Framework**

At the time of writing, we are developing a prototype system in Java to evaluate our activity recognition technique. At first, the Java program adds to the assertional part of our ontology an instance of PERSON, which represents the current individual. Then, the program parses the Opportunity dataset to acquire the atomic gestures executed by the individual. Each gesture is added as an instance of its subclass of ATOMICGESTURE, and a role HASATOMICGESTURE is added, to relate the gesture with the individual. At each new gesture, we use the ELOG reasoner to calculate the most probable consistent ontology, given the current gestures. Then, we use the Pellet reasoner for realizing the ontology calculated by ELOG, in order to derive the specific manipulative gestures, simple, and complex activities performed by the individual.

**Discussion and Future Work**

In this paper, we propose a tightly-coupled hybrid system for human activity recognition. Our framework unites both symbolic and probabilistic reasoning. This is achieved through the adoption of highly expressive log-linear DLs to represent and reason about the current activity at different granularities and complexity levels simultaneously. This paper outlines an initial implementation of the proposed framework. It describes the recognition algorithm and illustrates the concepts of our work through several examples. While the full evaluation of the proposed approach awaits further steps and experiments, our ontology-based approach combines promising features to address human activity recognition. The benefits of the proposed approach are manifold. Unlike the majority of related works, it supports the inherent uncertain nature of human activities without sacrificing the advantages of ontological reasoning. These advantages include consistency checking, the ability of integrating rich background knowledge, and the simultaneous recognition of coarse and fine-grained activities. The use of a standard description formalism enhances the portability and reusability of the proposed system, and supports the representation of heterogeneous and uncertain context data. Moreover, the declarative nature of DLs reinforces the flexibility and intelligibility of the system.

As pointed out earlier, however, there are some challenging open issues that need to be addressed. Currently, our approach does not support temporal reasoning, which is a key requirement for human activity recognition. There are two alternative approaches for enabling temporal reasoning in...
References


