D5.2

Multi-Parametric Behaviour Interpretation v1

Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support

Dem@Care - FP7-288199

eHealth
Better Healthcare for Europe
SEVENTH FRAMEWORK PROGRAMME
European Commission
Information Society and Media

Deliverable Information

Project Ref. No.		FP7-288199
Project Acronym		Dem@Care
Project Full Title		Dementia Ambient Care: Multi-Sensing Monitoring for Intelligence Remote Management and Decision Support
Dissemination level:		Public
Contractual date of delivery:		Month 18, 30 April 2013
Actual date of delivery:		Month 19, 02 May 2013
Deliverable No.		D5.2
Deliverable Title		Multi-Parametric Behaviour Interpretation v1
Type:		Report
Approval Status:		Final
Version:		1.1
Number of pages:		68
WP:		WP5 Medical Ambient Intelligence
Task:		T5.3 Multi-parametric Patient Behaviour Interpretation
WP/Task responsible:		CERTH
Other contributors:		CS, INRIA, UB1
Authors (Partner)		Stamatia Dasiopoulou (CERTH), Vasiliki Efstathiou (CERTH), Georgios Meditskos (CERTH), Carlos Crispim Junior (INRIA), Anh Tuan Nghiem (INRIA), Vincent Buso (UB1)
Responsible Author	**Name**	Stamatia Dasiopoulou (CERTH)
	Email	dasiop@iti.gr
Internal Reviewer(s)		Alex Sorin (IBM)
EC Project Officer		Gerard Cultot
Abstract (for dissemination)		This deliverable describes the current version of the behavioural interpretation framework of Dem@Care. The aim of the Dem@Care multi-parametric behaviour interpretation framework is to fuse information coming from a multitude of sensors and draw high-level interpretations regarding behavioural aspects of the person with dementia (PwD). The low-level sensory data analysis results are provided to Work Package 5 by Work Package 3 and Work Package 4 using as a common reference point the knowledge structures defined in D5.1 "Semantic Knowledge Structures and Representation" whereas information fusion is implemented in two components: Complex Activity Recognition (CAR) and Semantic Interpretation (SI). The CAR component focuses on extracting information relevant to the position, elementary and more complex states and activities of the PwD, by employing a hierarchical model-based approach that uses a generic constraint-based ontology language to describe the states/activities/events models of interest. The SI component deals with extracting information of higher abstraction such as complex situations, functional problems and summaries of PwD behavioural aspects of clinical interest by espousing a hybrid approach that combines ontology- and rule-based reasoning.

Version Log

Version	Date	Change	Author
0.1	28/02/2013	Deliverable outline	Vasiliki Efstathiou (CERTH)
0.2	06/03/2013	Minor edits	Carlos Crispim Junior (INRIA)
0.3	04/04/2013	Section 4.3	Vincent Buso (UB1)
0.4	09/04/2013	Contributions to state of the art (Section 3) and CAR component description in Sections 4 and 5	Carlos Crispim Junior, Anh Tuan Nghiem (INRIA)
0.5	18/04/2013	Section 1, Section 2, contribution to state of the art (Section 3), SI component description in Section 4, first merging, formatting	Stamatia Dasiopoulou, Vasiliki Efstathtiou, George Meditskos (CERTH)
0.6	18/04/2013	Refinement of Section 4.3	Vincent Buso (UB1)
0.7	22/04/2013	Refinement of CAR description in Sections 4.1 and 5.1, contribution to conclusions	Carlos Crispim Junior (INRIA)
0.8	23/04/2013	Addition of abstract, executive summary and Appendix 1.2; refinement of conclusions and SI description in Sections 4.2 and 5.2; final merged draft	Stamatia Dasiopoulou, Vasiliki Efstathtiou, George Meditskos (CERTH)
	28/04/2013	Internal review feedback	Alex Sorin (IBM)
0.9	30/04/2013	Addressed review comments	Carlos Crispim Junior (INRIA)
1.0	30/04/2013	Addressed review comments - Final version	Dasiopoulou, Vasiliki Efstathtiou, George Meditskos (CERTH)
1.1	02/05/2013	Addressed PMB members comments	Stamatia Dasiopoulou (CERTH)

Executive Summary

This document reports on the methods employed by Work Package (WP) 5 in order to address the medical ambient intelligence interpretation requirements of the first prototype of the Dem@Care system.

Multi-parametric behaviour interpretation involves fusing information coming from heterogeneous resources such as audio and video analysis, and physiological data. The knowledge structures defined in D5.1 "Semantic Knowledge Structures" serve as a common reference point for the aggregation of the descriptions extracted from the various sensory data. WP3 and WP4 provide these pieces of information in the form of observations that constitute the input to WP5 fusion and decision support tasks. The deliverable outlines the interpretation features that are supported in the first prototype for the three settings addressed in Dem@Care, namely lab, home and nursing home. It reviews state-of-the-art approaches relevant to the interpretation objectives of WP5 and proceeds by outlining the methods adapted by the Complex Activity Recognition (CAR) and Semantic Interpretation (SI) components for the realisation of current version of the Dem@Care behavioural interpretation framework. The report provides a description of software implementation aspects of the two components and their interaction within the Dem@Care system, as well as a preliminary evaluation of the components, and closes with a discussion of future directions.

Abbreviations and Acronyms

ADL	Activities of Daily Life
CAR	Complex Activity Recognition
DLs	Description Logics
Gear4	Gear4 Renew Sleep Clock
HAR	Human Action Recognition from static and wearable cameras
IADL	Instrumental Activities of Daily Living
KB	Knowledge Base
KBM	Knowledge Base Manager
MMSE	Mini Mental State Exam
ORWC	Object Recognition from Wearable Camera
OSA	Offline Speech Analyser
OWL	Ontology Web Language
OWL-DL	Ontology Web Language Description Language
PDT-PER	People Detection, Tracking and Primitive Events Recognition
Philips DTI-2	Philips Discrete Tensions Indicator
PwD	Person with Dementia
RDF	Resource Definition Framework
RRWC	Room Recognition from Wearable Camera
SI	Semantic Interpretation
SPARQL	SPARQL Protocol And RDF Query Language
SPIN	SPARQL Inferencing Notation
SVM	Support Vector Machine
SWRL	Semantic Web Rule Language
TURTLE	Terse RDF Triple Language
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
WI	Wearable Inertial Sensor
WIMU-SPS	Wireless Inertial Measurement Unit Signal Processing Software
WP	Work Package
XML	eXtensible Markup Language

Table of Contents

List of Figures

List of Tables

1 Introduction

The goal of multi-parametric behaviour interpretation is to recognise the behaviour of the person with dementia (PwD) and discern traits that have been identified by the clinicians as relevant for diagnostic, status assessment, enablement and safety purposes. In order to support multi-parametric behaviour interpretation, WP5 aggregates a multitude of information accumulated through monitoring of the person with dementia (PwD) herself and the PwD environment. This information is made available to WP5 by WP3 and WP4, which analyse the captured sensor data and provide analysis results, referred thereafter *observations*.

The variety of sensors employed within the Dem@Care system provide multifaceted information that varies from person-specific captured data (such as vocal attributes) to information related to the actions performed by the PwD (such as eating) and the environment of the PwD (such objects the PwD interacts with). Though each one is informative on specific aspects of interest, the individual pieces of information themselves are not capable of delineating the situations in which the PwD may be involved. Combined pieces of information on the other hand can plausibly render the behaviour of the PwD (e.g. by combining that the PwD wakes up, gets out of bed at night and visits the bathroom infers a nocturia incident) and explain incurring situations (e.g. that the PwD had several nocturia incidents because her coffee intake during the preceding day was higher than usual). WP5 aims to aggregate individual pieces of information provided by WP3 and WP4 and meaningfully fuse them in order to derive high-level interpretations of the PwD behaviour.

In order to implement multi-parametric behaviour interpretation, two constituents need to be considered for supporting the underlying fusion tasks, namely representation and reasoning (i.e. decision making) support. In deliverable D5.1 "Semantic Knowledge Structures and Representation" [19] the Dem@Care ontology was presented. The present deliverable reports on the reasoning aspects addressed towards the first version of the multi-parametric behaviour interpretation framework. The activities of WP5 are strongly correlated with those of WP3 and WP4: the knowledge structures serve as a common reference point for the projection and aggregation of the descriptions extracted from the various sensory data, which in turn form inputs for WP5. The decision support techniques employed by WP5 at this stage constitute the first preliminary approach towards multi-parametric behaviour interpretation. As a result, a subset of the interpretation requirements set by WP2 is addressed at this stage with the rest pending for later prototypes.

The deliverable is structured as follows. Section 2 presents the specifications for the first version of multi-parametric behaviour interpretation for the lab, home and nursing home settings that are addressed within the project. Section 3 presents and discusses the relevant literature. Section 4 describes the methodological approaches adopted by the Complex Activity Recognition (CAR) and the Semantic Interpretation (SI) components that realise the current Dem@Care multi-parametric behaviour interpretation. Section 5 reviews architectural considerations of the two components. Section 6 concludes the deliverable, discussing next steps.

2 Multi-parametric Behaviour Interpretation Specifications

The collection of sensors considered in Dem@Care allows capturing a variety of data about the PwD. Dedicated analysis components that are being developed in WP3 and WP4 process these data to extract observations about locomotive, physiological and voice-based attributes (e.g. number of steps, skin conductance, verbal reaction time), about objects within the PwD field of attention (e.g. kettle, watering can) and their location (e.g. near the TV), as well as about elementary activities of the PwD (e.g. eating).

Though already informative, when considered individually, the observations generated by the WP3 and WP4 components provide faceted only views of the PwD's state and conduct. The goal, and challenge, of multi-parametric behaviour interpretation is to collectively analyse the aggregated observations in order to recognise the behaviour of the person with dementia (PwD) and discern clinically relevant traits for the diagnostic, enablement and/or safety purposes addressed within the project.

The clinical objectives maintained at each pilot setting (i.e. lab, home and nursing home) with respect to the aforementioned purposes, designate the desired functional requirements for behaviour interpretation. The type and granularity of the available observations though, set the boundaries of the interpretation pursuits.

In the following, we describe the specifications per pilot setting for the first version of behaviour interpretation. The descriptions reflect the compromise between the clinical requirements (as identified in D2.2 [26] and further refined in D7.1 [50]) and the set of observations that are currently available by the WP3 and WP4 components (D3.1 [47], D4.1 [7] and D4.2 [12]). Revisions are expected upon the completion of the pre-pilot phase that will run before the actual pilots for training[1] and testing purposes. Small if any updates are expected for the lab setting, as real data from the CHUN lab have been recorded and made available to the consortium since M7; for the home and nursing home settings though, testing on real data may lead to a revised set of available observations.

2.1 Behaviour interpretation specifications for the lab setting

The aim of Dem@Care in the lab setting is to support diagnosis/assessment via the implementation of an objective assessment of autonomy and goal oriented cognitive function [18]. To this end, the clinical partners from CHUN defined an experimentation protocol that consists of three steps:

- Step 1 that includes directed activities, where the participant is asked to perform upon request three physical tasks and two vocal tasks,
- Step 2 that includes semi-directed activities, where the participant is given instructions about a set of activities that she needs to perform while respecting certain ordering constraints, and
- Step 3 that includes discussion with the clinician, where the participant is asked to engage in directed and free discussion tasks

[1] Some components (e.g. HAR, ORWC) need to be trained on recorded data from the actual site before the actual deployment.

For each step and task, the experimentation protocol specifies a set of attributes that are relevant to the assessment of certain functions and abilities of the participant (see D2.2 and D7.1). Some of these attributes, such as diadochokinetic regularity (i.e. number of tokens per second), correspond to observations that are directly extracted via analysis of the recorded sensor data. Others, such as recognising the activities executed in the semi-directed step and whether they were performed successfully or not, require the collective interpretation of the aggregated observations.

In the first version of multi-parametric interpretation, the focus is on supporting the elicitation of the attributes specified in the protocol. The derivation of assessments by combining the set of individual attributes in order to support clinicians in the diagnosis is a part of the second and the third versions of multi-parametric interpretation.

Table 2-1 lists the WP3/WP4 components and the observations that are currently available per protocol step for the lab setting[2]. As shown, for Step 1, a number of attributes of interest are already covered; for Step 2, the WP3/WP4 components provide observations that are too low-level to directly serve as attributes; for Step 3, the currently specified vocal attributes are available.

Table 2-1 Input observations for multi-parametric interpretation in the lab setting

WP3/WP4 Component	Step/Task	Available Observations
DTI-2 SW	Step1 (walking & walking and counting backwards); Step 2	moving intensity; resting coefficient; non-sleep passiveness coefficient
ORWC	Step 2	objects present in the lab experimentation room (e.g. kettle, pillbox, teabag)
OSA	Step 1 (sentence repeating)	verbal reaction time
	Step 1 (articulation control)	diadochokinetic regularity; diadochokinetic speed
	Step 3	verbal reaction time; verbal participation
PDT-PER	Step 1; Step 2	person detection; person posture; x-y-z coordinates of person centroid; 3D dimensions of person;
RRWC	Step 2	locations wrt objects present in the lab experimentation room (e.g. TV area, table area)
WIMU-SPS	Step 1; Step 2	moving; resting

[2] The information listed in the table, as in all tables of this section, reflects the state of the components during the writing of the deliverable

Consequently, the requirement for multi-parametric interpretation is to augment the available set of attributes by extracting the pieces of higher-level information that can't be provided by means of WP3/WP4 analysis alone. This translates to recognising the behaviour of the participant with respect to the protocol specifications, i.e. recognising the activities performed, the time required for each activity, the number of repetitions of a certain activity, etc.

Table 2-2 lists the information that interpretation is required to extract, namely the activities and respective properties. The list of activities is not exhaustive, as other, simpler, activities to which the listed ones may be decomposed, need to be recognised first. For example, in order to recognise that the participant is making a phone call, the constituent activities *standing, near office desk*, *near phone* and *use phone* need to be recognised first. The exact list of these simpler activities depends on the conceptualisation adopted by the two interpretation components. As we will see in Section 4, *using the office*, for example, is modelled in the CAR component as positioned within the office zone for more than a predefined time interval, and *bed exit* is modelled in the SI component as waking up and getting out of bed during a night sleep episode.

Table 2-2 Activities extracted in the lab setting and respective properties

Step/Task	Activities	Properties
Step 1	walk; stop; turn	start time; end time
Step 2	turn TV on; establish account balance; pay phone bill; answer the phone; call the psychologist; find bus line on map; prepare the drug box; read article; prepare hot tea; water the plant; leave the room	start time; end time; number of repetitions; successful or not execution
Step 3	n/a	n/a

The definition of the activities depends naturally on the granularity of information of the provided observations too. In the lack of information about the participant picking up the phone or dialling numbers, there is no point for adopting a more elaborate definition such as approaching the telephone, picking up the receiver, then dialling the numbers then speaking, and finishing by placing the receiver down. A much simplified model needs to be used instead. It is envisaged that as components mature they will provide richer observations, which will allow in turn for more sophisticated interpretations.

As illustrated in Table 2-2, interpretation focuses currently primarily on Step 1 and Step 2. In Step 3, the analysis provides already the attributes specified in the protocol as relevant for assessment; since assessment isn't currently addressed by interpretation, these attributes are stored to the knowledge base as all other information, but no further processing is applied for the moment. To present the extracted attributes to the clinicians, the WP6 components need only to query the knowledge base.

2.2 Behaviour interpretation specifications for the home setting

Unlike the lab setting, diagnostic purposes are only implicit in the home setting. The primary aim here is to enable PwD and reduce risks to their well being. This requires recognising the

PwD behaviour and identifying not just what the PwD is doing but also situations that indicate a problematic behaviour and require feedback support.

In the first version of multi-parametric interpretation, the focus is on the recognition of PwD activities and of clinically defined problems. The activities of clinical relevance and interest, as well as the respective problems for the five functional areas (sleep, IADL, exercise, social interaction, mood) addressed by the project have been described in detail in D2.2 and further exemplified in the scenarios and use cases described in D7.1. The assessment of contributing factors for the detected problems will be considered in the future versions of multi-parametric interpretation, along with triangulation.

Table 2-3 lists the WP3/WP4 components and the observations that are (estimated[3] to be) available in the first system prototype for the home setting. As illustrated, the available observations provide various pieces of information. Some observations consider complementary aspects to one another, e.g. PDT-PER provides information about posture and ORWC provides information about objects in the view of attention; others address correlated information, e.g. WIMU-SPS extracts information about moving, while DTI-2 SW extracts information about moving intensity.

Table 2-3 Input observations for multi-parametric interpretation in the home setting

WP3/WP4 Component	Available Observations
DTI-2 SW	moving intensity; resting coefficient; non-sleep passiveness coefficient
Gear4	asleep versus awake
HAR	elementary activities (e.g. person drinking; the exact activities will depend on training at the individual PwD homes)
ORWC	objects present in the PwD home (the exact objects depend on training at the individual PwD homes)
OSA	verbal reaction time; verbal participation; voice rating during conversation
PDT-PER	person detection; person posture; x-y-z coordinates of person centroid; 3D dimensions of person
RRWC	room recognition (e.g. kitchen)
WIMU-SPS	moving versus resting

Table 2-4 lists the types of the inferences currently addressed by interpretation, namely complex activities, clinical problems, and summaries per functional area, as well as properties of interest; it also provides representative examples for each type of expected interpretation type. The primary focus as illustrated is on the areas of sleep, social interaction and IADLs,

[3] As aforementioned, some analysis components require training before being deployed; since the first real data for the home setting will be acquired during the pre-pilot testing phase, only afterwards will the set of available observations be finalised

and in particular eating. Physical exercise and mood will be considered in future versions, as the observations currently available provide too little information to draw meaningful insights. Furthermore, the recognition and assessment of mood requires a more elaborate level of understanding of the PwD behaviour compared for example to that required for the recognition of activities and clinical problems.

Table 2-4 Types of inferences currently supported by the interpretation framework

Interpretation Objectives	Properties	Example inferences
Complex activities/events	start time; end time; date	sleep related activities (sleep episode, bed exit, night bathroom visit, nap, etc.); social interaction related (face to face social interaction, telephone interaction); IADL related activities (having meal, preparing meal, table exit, etc.)
Clinical Problems	date	sleep problems (nocturia problem, fragmented sleep, sleep efficiency, etc.); social interaction problems (insufficient number of social interactions, etc.); meal problems (missed meal, eating in inappropriate places)
Functional Area Summaries	date	sleep summary (bed time, sleep time, wake up time, number of awakenings, number of nocturia incidences, etc.); social interaction summary (number of face to face interactions);

2.3 Behaviour interpretation specifications for the nursing home setting

The nursing home setting can be seen in a way as a controlled version of the home setting. This time however the focus is primary on monitoring the PwD mood, sleep, physical and daily activities in order to inform the staff members about changes in the PwD behaviour and especially of hazardous behaviours that put the PwD safety at risk.

Similar to the home setting, the requirements for the multi-parametric interpretation for the first pilot in the nursing home include the recognition of complex activities, clinical problems and hazardous situations, as well as summaries with respect to the identified areas of interest.

Table 2-5 Input observations for multi-parametric interpretation in the nursing home setting

WP3/WP4 Component	Available Observations
DTI-2 SW	moving intensity; resting coefficient; non-sleep passiveness coefficient
Gear4	asleep versus awake
PDT-PER	person detection; person posture; x-y-z coordinates of person centroid; 3D dimensions of person

Table 2-5 lists the observations that are currently available in the nursing home for the first pilot. As illustrated, a number of observations relevant to understanding sleep behaviour are available. This however is not the case for other areas such as physical activity, or social interactions. As such for the moment the requirements for interpretation are a subset of those presented in Table 2-4 for the home setting.

2.4 Off-line and on-line behaviour interpretation specifications

The behavioural interpretation framework supports two processing modes that satisfy different functional requirements in the Dem@Care system. More specifically:

- **On-line processing mode.** This mode aims to support (near) real-time recognition of the complex activities/situations needed by WP6 to realise immediate interaction feedback services, such as the alert generation service. In the first prototype, PDT-PER is the only of the WP3/WP4 components that is able to deliver real-time analysis results, hence, delineating the possible realm of WP5 real-time interpretations.
- **Off-line processing mode.** The off-line processing mode aims to collectively analyse the aggregated observations generated by the WP3 and WP4 components, so as to recognise the behaviour of a person and identify clinically relevant situations, problems and trends. The off-line processing mode fuses and analyses the results of the WP3/WP4 components that are provided in an off-line mode in combination with the intermediate results of complex activities/situations identified within WP5.

3 Related work

Understanding what a person is doing comprises a key challenge in ambient intelligence application domains, such as health-care monitoring and assisted living. As a result, activity/event[4] recognition has transformed from a typical computer vision challenge, to a key task in ubiquitous and pervasive computing.

Lavee *et al*. [31] categorise computer vision approaches for event recognition in three categories: State models, Pattern Recognition methods, and Semantic models. All three described approaches are generally based on at least one of the following data abstraction levels: pixel-based, feature-based, and event based level.

State models refer to techniques such as Conditional Random Fields, Dynamic Bayesian Networks, and Hidden Markov Models. Pattern Recognition methods are Artificial Neural Networks, Support-Vector Machines (SVM), Nearest Neighbour, etc. In this context, Le *et al*. [32] have presented an extension of the Independent Subspace Analysis algorithm applied at learning invariant spatio-temporal features from unlabelled video data for activity recognition. Wang *et al*. [54] have proposed new descriptors for dense trajectory estimation, which are later used as input for a non-linear SVM. Although these techniques have considerably increased activity recognition performance in benchmark datasets, they extract information from pixel-based and feature-based abstraction, which poses limitations concerning their ability of describing the semantic and hierarchical nature of complex events. Izadinia and Shah [27] have presented a method for learning low-level events from data, and later identifying complex events from the joint relationship among the detected events using a graph representation and a discriminative model.

Alternatively, semantic (or description-based logics) models use a descriptive language and logical operators to build event representations. Its hierarchical nature approach allows explicit modelling of semantic information, besides to the fact they do not require as much data as the Pattern Recognition and State models. These models are also accessible for domain-experts to easy change them, as they generally follow natural terms. Zaidenberg *et al*. [55] have presented a generic framework for activity recognition of group behaviours in an airport, a subway, and shopping centre scenarios. However, a limitation of semantic models is their sensitivity to noise of underlying vision process, like image segmentation and people detection algorithms.

Similar strands of research have been adopted for activity recognition in the ubiquitous and pervasive computing fields, where data from multiple sensors, including inertial sensors (e.g., of accelerometers and gyroscopes) and ambient sensors (e.g. ambient microphones) are fused to monitor daily living activities. Gao *et al*. [23] have demonstrated the fusion of inertial sensors data worn at the waist, chest, thigh, and side of a person body using a Naïve Bayesian Classifiers. See also Rong and Ming, [46]. Fleury *et al*. [21] have presented a multi-modal system using sensors such as Actimeter, Microphones, PIR, and Door contacts, and data fusion is performed using an SVM classifier. Medjahed and Boudy [36] have presented a smart-home setting which performs activity recognition relying on ambient sensors, such as infrared, change state sensors, audio, and physiological sensors fused by a Fuzzy Classifier. Disadvantages of these approaches result from motion noise and the requirement of inter

[4] The terms *activity* and *event* are used interchangeably.

sensor-calibration in case of inertial sensors, or the assumption that the sensors are always placed at the same body position, what could cause noise in large scale studies. A descriptive-based approach has been presented by Cao *et al*. [13] for modelling the context of human status (e.g., body posture) and the environment context (semantic information about the scene). The event models described information provided by a set of cameras for person detection, and accelerometer devices attached to objects of daily living for environment events triggering (*e.g*., TV remote control or doors use). A rule-based reasoning engine is used for processing and combining both models types at activity detection level.

In the literature of semantic approaches to modelling and reasoning about context and activity recognition in pervasive computing, the Semantic Web ontology languages OWL DL [1] and the more recent OWL 2 [2], have motivated a growing body of interest into ontology-based frameworks [15], [9]. In parallel, a number of initiatives have looked into action theories to incorporate notions of cause and effect in activity and context recognition. In the following, we present an overview of seminal approaches.

3.1 OWL-based activity recognition

In [40], a hybrid framework that adapts Allen's temporal operators [1] and ontological modelling for recognizing sequential and multi-tasked activities in a smart home environment is introduced. The suggested framework adopts the common idea that there exist simple and complex ADL activities that take place in the smart home, with the assumption that simple activities consist of a set of primitive actions with no underlying temporal restrictions whereas composite activities are collections of more than one simple activities that occur within a given time interval. In addition, dynamic composite activities represent composite activities that have properties whose values vary in time, incorporating this way the notion of change. OWL is used to formally conceptualise the smart home domain; the 4D fluent approach [1] is adopted to capture temporal knowledge. To support procedural processing of the composite activity model a set of rules is defined based on the Semantic Web Rule Language (SWRL) [25] to overcome OWL's expressiveness limitations and allow inferring complex dependencies among activities and therefore inferring the occurrence of ongoing composite activities.

The potential of the approach is illustrated via a smart home activity recognition example scenario that involves two concurrent activities, where a task of making tea is performed entirely within the duration of a boiling pasta task. The pictorial representation of the example illustrated in Figure 3-1 was taken from [40] where the occurrence of a sequence of activities over a timeline is represented.

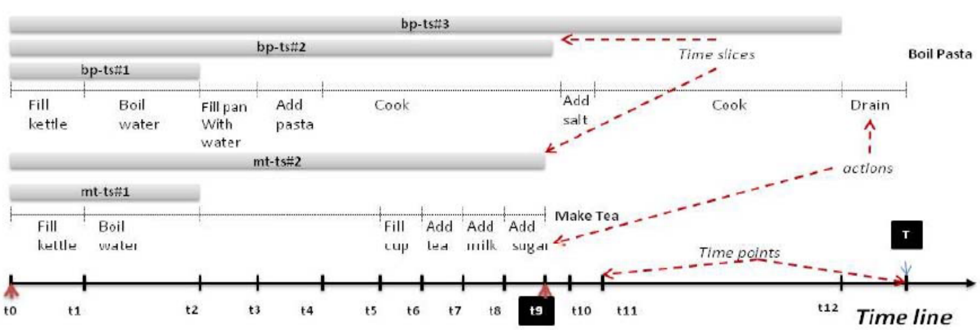


Figure 3-1 A concurrent activity recognition scenario

At time **t2**, the actions *fill kettle, boil water, and fill pan with water* have been activated. The recognition system classifies this as *"make tea"* activity with time slice *mt-ts#1* and corresponding primitive actions{*fill kettle, boil water*} and "*boil pasta"* activity with time slice *bp-ts#1* and corresponding primitive actions {*fill kettle, boil water, fill pan with water*}. The associated intervals are allocated the start time of the first primitive action that initiated the respective composite activity instance. The temporal inference engine concludes that the temporal intervals associated with these time slices are equal so the system reports that most likely the user is concurrently performing "*make tea"* and *"boil pasta"* activities. Further primitive actions occur within the timeline illustrated in Figure 3-1 where the last primitive action that constitutes a partial action of the "*make tea"* activity takes place at **t9** and the last primitive action that constitutes a partial action of the *"boil pasta"* activity takes place at **t12**. After aggregating this information and comparing the respective temporal associations, the intervals associated with time slices mt-ts#2 and bp-ts#3 will be denoted as make-tea (start=t0,end=t9), and boil pasta (start=t0, end=t12), respectively.

SWRL rules are used for generating the necessary composite activity models; however, SWRL rules do not allow for assertion of new individuals, therefore, the assertion of new time slices and the corresponding composite activities has to be done externally. The generation of new named individuals that is, going beyond classifying a given unknown object to the right class but constructing a new object given its constituent parts is not handled within OWL and is an issue that most complex activity recognition proposals surpass and imply that this is handled by some external mechanism.

Another hybrid framework that implements ontological modelling extended with SWRL rules [25] is presented in [28]. In this framework OWL ontologies are used for modelling context, including upper-level information that captures general features of all pervasive computing domains (for representing location, environment, simple events) as well as low-level domain information. The domain presented in this work is a smart classroom environment. The objective is to aggregate low-level simple events detected by sensors and recognise high-level activities. In order to perform complex activity recognition tasks, sensor data from non-overlapping sliding windows are stored in the ontology and SWRL rules are used as a first reasoning step to derive simple events and store them in the ontology; the ontology is then used by an activity recognition system based on Bayesian Networks and case-based reasoning. After the SWRL rules trigger, the *Simple Events* that occurred within a timeframe are passed to the Activity Recognition System building an unsolved case. One or no activity is then recognised using Bayesian Networks and the solved case is added to the case base. This proposal aims to combine and exploit the advantages that ontologies and machine learning techniques offer.

However, the approach requires training data providing an ad-hoc solution to the problem of activity recognition. Moreover, the proposal does not handle temporal characterisations of activities; the information aggregated within a timeframe is handled in a snapshot-like manner where the conjunction of a set of atomic activities at some specific point in time implies the occurrence of a complex activity. So, for instance, rules like the following are used to aggregate low-level information in a smart room and infer higher-level interpretations.

```
Location:SmartClassroom(?c)^Environment(SmartClassroomEnv) ^
Core:hasEnvironment(?c,SmartClassroomEnv)^
Core:noiseLevel(SmartClassroomEnv,?noise)^swrlb:greaterThan(?noise,80) ^
```

```
Core:SimpleEvent(High Level Noise)
→ Core:isActivated(High Level Noise, "true")
```

Using rules like this along with ontological subsumption reasoning may further enhance expressiveness, however neglecting temporal information severely restricts the intricacy of the relations that can be expressed in an activity recognition setting, and hence the scope of the activities to be recognised.

A framework that also combines ontological modelling with hybrids of ontological and statistical reasoning is presented in [44]. In this framework OWL ontologies are used as the underlying models for defining the formal semantics of human activities and specialisations by means of the operators of the ontological language. At run time, context information coming from distributed sources in the intelligent environment is retrieved and aggregated by an aggregation middleware (CARE [8]). CARE interacts with context-aware activity recognition system (COSAR [43]), which includes the ontology of human activities and supports hybrid statistical and ontological reasoning for recognizing simple human activities. Context data are mapped to ontological classes and properties by CARE, and added as instances to the ABox. Ontological reasoning is performed subsequently, by an existing OWL reasoner in order to recognise complex human activities. Unlike the activity recognition system presented in [40], in [44] no concurrent activities are considered and the user is considered to either be performing one activity at a time or to be idle. The framework was tested on a real-world dataset[5] and the results were compared to a purely data driven approach based on Hidden Markov Networks. Activity recognition was taking place at fixed predetermined times every minute and for time slices lasting 60 seconds.

In an attempt to endow partial support for including temporal considerations in complex activity modelling, notions such as *recently used*, *last activity* and *second last activity*, were introduced. Activity recognition tasks using 3 types of ontological representations for activities have been performed. Each of the respective ontologies modelled activities in the following three different ways: *i)* in a snapshot ontology activities were modelled taking into account the user's current one-minute contextual information in a static manner, *ii)* in a "one step" ontology the `recentlyUsed` property was introduced to relate the actor to the objects they used during the last three minutes' time slices, providing this way a temporal characterisation in the activity modelled and *iii)* in a "multi-step" ontology which provides further temporal characterisations by introducing the `secondLastActivity` property to relate the actor to the second last activity they performed before the current one. Since OWL does not support temporal reasoning, an external Java application was used to keep track of recently used objects, and to add assertions to the ABox about the `recentlyUsed` and `secondLastActivity` properties. Even though the properties introduced provide a simplistic approach to temporal characterisation of complex activities the experimental results were encouraging in terms of the accuracy of the activities recognised. The results show that the effectiveness of ontological approaches extended with simple forms of temporal reasoning is comparable to state of the art techniques based entirely on statistical methods (such as Hidden Markov Models for instance).

This study pointed out the effectiveness of recognition based on ontological reasoning, however, 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

[5] http://staff.science.uva.nl/_tlmkaste/research/software.php

habits; they are not general definitions of ADLs. So, the models do not offer the potential of being encapsulated in other frameworks, thus not benefiting of one of the main advantages of ontology-based frameworks, namely interoperability. Besides activity modelling, the activity recognition framework relies on the use of some ad-hoc Java external application for the generation of new named individuals, which further impedes interoperability.

3.2 Activities incorporating event calculus

Allen's interval calculus provides the means for modelling temporal relations between intervals during which events take place, allowing this way to draw associations over temporal relations of the occurring events. Action theories on the other hand allow for drawing cause and effect relationships between events. An approach that reconciles both these ideas and has drawn attention in activity interpretation applications is that of event calculus [29]. Event calculus intends to connect events with the relationships they may initiate or terminate, allowing this way to describe complex, dynamically changing worlds incorporating a temporal dimension. In this formalism, all changes to the world are the results of named events while any property in the world that can change over time is known as a *fluent*. Predicates define relations between entities that specify what happens when, which fluents hold at which times, and describe the initial situation and the effects of events. For instance, `Initiates(a,b,t)` specifies the effects of an event, i.e. the fluent `b` starts to hold after action `a` at time `t`.

A prominent example where event calculus is used for the representation and reasoning of events and their effects is the work in [16]. Activity recognition in the proposed framework is demonstrated through a use case of a making tea task in a smart home environment. The cognitive model consists of:

- *primitive actions* such as `takeTo(thing, location)`, `add(thing,container)`, `remove(thing, container)`, `turnOn(y)`

- a set of *fluents* (i.e. state variables) such as: `at_pos(thing, location)`, `inside(thing, container)`, `available(thing)`, `on(y)`, `off(y)`

- *effect formulae* such as:
```
HoldAt(inside(cold water, kettle), t)
→ Initiates(boilWater,available(boiledWater), t) and

HoldAt(at_pos(teabag,kitchen table), t) ^
HoldAt(at_pos(mug,kitchen table),t)
→ Initiates(add(teabag, mug), inside(teabag, mug), t)
```

- *Heuristics* and *user profile:* user-specific information (e.g. that the user normally takes tea twice a day, around 10.30am and around 4.30pm and the preparation lasts 30mins on average) which along with the background knowledge (e.g. the list of actions that are assumed to comprise the make tea composite activity) can be defined to produce user-specific compound action descriptions as goals to look for.

- *Compound actions and their effects:*
The compound action `makeACupOfTea` for example is described as the sequence of actions:
```
Happen(takeTo(kettle, basin), t0 ) …
```

`Happen(add(coldwater, kettle),` t_2 , t_3)…

`Happen(boilWater,` t_5 , t_6) …

`Happen(add(milk, mug),`t_{10})) …

`Happen(stirAround,` t_{11} ,t_{12}) …

And the effect formula of the compound action is

`Initiates(makingACupOfTea,available(teaReady),`t_{14})

- *Initial situations:*

  `InitiallyP(at_pos(inhabitant, kitchen)),`

  `InitiallyP(at_pos(inhabitant, basin)),`

  `InitiallyP(at_pos(milk, fridge)),`

  `InitiallyP(at_pos(sugar, cupboard)),...`

- *Assistance provisioning*

  The above cognitive formulae can be mapped into individual components of the architecture. As the assistive system has the `makeACupOfTea` compound action and knows two desired goals `HoldAt( available(teaReady), 10:30am )` and `HoldAt( available(teaReady), 4:30pm )`, it can then infer that two compound actions, i.e., `Happen(makeACupOfTea, 10am, 10:30am)` and `Happen(makeACupOfTea, 4pm, 4:30pm)`, should take place at the specified time points in terms of the effect axioms and the 30-minute requirements from the inhabitant's profile. In this case, if the system does not detect `any takeTo(kettle, basin)` action taking place at round 10am or 4 pm it will issue a reminder.

This framework goes beyond Allen's calculus by modelling cause and effect relations of activities besides temporal aspects. However, in order to adapt notions from the event calculus formalism, the compound activity models are uniquely defined sequences of actions, not allowing for deviations and variations of the predetermined sequence. Although according to the authors the formalisation is based on ontologies, no formal ontologies are introduced in [16] for modelling activities or for capturing concepts and relations relative to event calculus; moreover no discussion is made with respect to any ontology language employed. Likewise, no referral to implementation aspects of the rules introduced is made, implying that no standardised rule language is used and indicating that the implemented framework provides an ad-hoc solution to the complex activity recognition problem.

An approach that also aims to demonstrate the usefulness of action theories in an ambient intelligence environment is presented in [41]. This proposal integrates Semantic Web technologies for representing contextual knowledge with rule-based and causality-based reasoning methodologies for supporting a multitude of general-purpose and domain-specific reasoning tasks imposed by an ambient intelligence system. The ontologies employed are implemented in OWL and model key ambient intelligence notions that capture the meaning and relations of concepts regarding low-level context acquired from sensors, high-level context inferred through reasoning, user and device profiling information, spatial features and resource characteristics. Rule-based reasoning is used, on top of the context modelling ontologies, to infer complex contexts from raw context data. In parallel, causality reasoning is employed to capture and reason over preconditions and effects of actions and events, based on the event calculus theory. Resources are collected and translated to event calculus axioms. In order to utilise a compound event e_1 for reasoning tasks, its temporal properties are axiomatised in event calculus as in the following example that serves for verification purposes in an application error detection scenario for the ambient intelligence system.

```
Happens(Start(e1), t)≡
Happens(TurnOnLight(l), t1) ^ Happens(StartLocalizer(), t2)^
Happens(StartMapService(), t3) ^ (t1 < t2) ^ (t = min(t1, t3))
```

as well as its causal properties:

```
Initiates(Start(e1), LightOn(l),t)^
Terminates(Stop(e1), TrainingMode(), t)
```

The framework combines rule based reasoning with Jess[6] also combined with DEC Reasoner[7], a SAT-based event calculus reasoner. The application of the framework focuses more on aspects relative to regulation of the overall operation of an ambient intelligence system rather than in behavioural interpretation aims; however the authors claim that the approach, achieves a general-purpose reasoning framework for ambient intelligence, able to address a broad range of aspects that arise in a ubiquitous domain. The axiomatisation of events that serve for system verification purposes though follow more standard patterns than human activities whose sequential definition may vary largely both in terms of the partial actions that comprise an activity and in the alterations of the sequence in which these occur. Therefore, this approach suffers from the same shortcomings as the one previously discussed in [16]. Moreover, in terms of implementation although the framework employs standard OWL ontologies in one hand it uses non-standardised Jess rules on the other hand, while the Jess rule engine is licensed for limited usage only, hindering this way interoperability.

Another proposal that adopts the event calculus theory for complex activity recognition is presented in [1]. The distinction of *short-term* and *long-term* events is used to capture the notion of atomic and complex respectively. The constraints on a set of time stamped short-term activities that, if satisfied, lead to the recognition of a long-term activity, are expressed using a dialect of the event calculus. Short-term activities are represented as events in the Event Calculus in order to use the `initiatedAt` and `terminatedAt` predicates for expressing the conditions in which these activities initiate and terminate a long-term activity. The output of the system is a set of recognised long-term activities, which are predefined temporal combinations of short-term activities. The system uses a logic programming implementation (Prolog) of the event calculus and representation of activities relies solely on their description through axioms built with predicates, function symbols and constants with no established formal ontology language involved. Additional domain-independent axioms capture general temporal and action-effect information such as the following where F denotes a fluent, V denotes the value the fluent takes and T, T_s and T_f denote time instances.

$$terminatedAt(F=V,T_f) \wedge Ts < Tf < T \rightarrow broken(F=V,T_s,T)$$

According this axiom, a period of time for which $F=V$ holds is broken at T_f if $F=V$ is terminated at time T_f.

[6] Jess, http://www.jessrules.com/

[7] DECReasoner, http://decreasoner.sourceforge.net/

In addition to general, domain-independent rules such as the one introduced above, the framework uses domain-dependent cause and effect rules. So, for instance besides the general domain-independent rule `initially(F =V) → initiatedAt(F =V,0)`, the definitions of `initiatedAt` also has the general form

`happensAt(Ev,T)^Conditions[T] →initiatedAt(F=V,T)` where `Conditions[T]` is some set of further domain-specific conditions referring to time `T`.

The authors present a detailed evaluation of the system through experimentation on a benchmark dataset of surveillance videos. Among all the relative works described in this section this is the only study that does not only evaluate the system's accuracy through use cases where perfect information is available, but also demonstrates how incomplete short-term activity narratives, inconsistent annotation of short-term and long-term activities, and a limited dictionary of short-term activities and context variables affect recognition accuracy.

Although the expressiveness of the proposed event calculus dialect is demonstrated by presenting several complex activity definitions from the domain of public space surveillance, the intricacy of the activity definitions described is way simpler than the ADLs that are normally required to be recognised in a smart home environment. So, while for instance "leaving object" is regarded as a target complex activity of interest in the public surveillance domain, in a smart home this would usually constitute part of a much more complex activity which would require semantic interpretation of the relative object (e.g. leaving the kettle down as part of a "making tea" activity). In addition, in order to further enhance activity recognition the authors use the notion of mutually exclusive short term activities. This may be feasible in their domain of interest where the set of short-term activities is limited to state-like notions such as "abrupt move", "inactive", "moving", but it would lead to an explosion in the number of axioms in the smart domain where the number of short term activities considered in much larger. Finally, the implemented system suffers from the major drawback that it provides a custom-made reasoning framework with limited capabilities in terms of model reuse. The low-level information that is used as input to the complex activity recognition system goes down the level of analysis of pixel positions in frames in order to extract spatio-temporal information; this requires empirical analysis of the respective data and does not provide any reusable activity models.

3.3 Summary

Activity recognition has been focal to computer vision research. Advances in ubiquitous and pervasive computing have transformed the recognition task from a traditional vision analysis problem to a problem of fusing different types of information (visual, audio, physiological, etc). Approaches to multi-parametric behaviour interpretation employ modelling and reasoning techniques that allow for aggregating and fusing information coming from heterogeneous resources.

For semantic interpretation frameworks a common assumption is that analysis of sensor data provide *low-level* information, which, when processed by appropriate reasoning mechanisms, produces *high-level* information such as aspects of behaviour interpretation. Ontologies have been extensively used as the means for representing information. Particular attention has been given to the ontology languages developed for the Semantic Web, thanks to their representational and reasoning power afforded. The OWL family of ontology languages provide a number of features that fit the requirements of the ambient intelligence domain, such as modelling complex logical relations and sharing information coming from heterogeneous sources while OWL ontologies benefit from the availability of reasoning engines that allow deriving higher-level context abstractions.

However, OWL suffers from two shortcomings that are apparent even for interpretation tasks that seem relatively simple. The first shortcoming refers to OWL's lack of support for temporal reasoning; the second is that, within OWL, it is not possible to infer and assert new named individuals. The implications are particularly evident in domains that require the recognition of complex context elements, such as human activities that are generally characterised by intricate temporal associations, and where it is often the case that the aggregation of individual activities entails the existence of a new (composite) activity. The majority of OWL-based frameworks use some OWL-compliant rule language along with some external mechanism in order to overcome these limitations; often resulting this way in non-interoperable solutions. Further attempts have been made in order to develop hybrids of ontological and statistical reasoning to further enhance interpretation by exploiting the advantages of the two methods; however these require training data, further hindering interoperability and model reuse.

4 The Dem@Care Behaviour Interpretation Framework

Based on the specifications presented in Section 2, the behaviour interpretation framework serves a two-fold role in the first Dem@Care prototype:

- It analyses collectively the aggregated WP3/WP4 observations and derives a higher-level understanding of the PwD behaviour in terms of the activities and situations the PwD engages in and the identification of clinically defined functional problems
- It aggregates and summarises the results of the interpretation, offering a single point for collection of the PwD's contextual information needed by WP6 to realise its feedback services.

To serve the aforementioned interpretation tasks, two components have been developed: the Complex Activity Recognition (CAR) component and the Semantic Interpretation (SI) component. The developed components support interpretation tasks at different levels of granularity. CAR serves for identifying complex activities whose modelling is grounded on information at the level of person posture and location. SI on the other hand addresses situations that require encapsulating pieces of information of higher abstraction. Table 4-1 outlines the respective tasks and inputs as supported in their current implementations.

Table 4-1 Current role and inputs of the behaviour interpretation components

Component	Interpretation Task	Input Observations
CAR	complex activities (events), including cases that qualify for real-time alert generation	PDT-PER
SI	complex situations, functional problems, summaries	ORWC, RRWC, OSA, Gear4, DTI-2, WIMU-SPS, HAR, CAR

The Complex Activity Recognition (CAR) currently focuses primarily on the recognition of: i) the position of PwD with respect to predefined zones of interest and her moving from one zone to another (e.g. person inside the office desk zone), ii) elementary states and activities using posture and localisation information (e.g. person bending), and iii) complex states and activities (e.g. person using office desk). As described in the following, CAR adopts a hierarchical model-based approach, using a generic constraint-based ontology language to describe the states/activities/events models of interest.

In turn, the primary focus of Semantic Interpretation (SI) is currently on the recognition of: i) complex situations (e.g. night bathroom visit after the person has gone to sleep), ii) functional problems as defined by clinicians (e.g. nocturia problem in case of more than two bathroom visits during the night and after the person has gone to sleep), and iii) summaries of key attributes of PwD behaviour with respect to the functional areas considered (e.g. for sleep, the number of awakening during night sleep and the number of naps the preceding daytime). SI espouses a hybrid approach that combines ontology- and rule-based reasoning. An OWL 2 ontology is used to model the domain concepts (activities, situations, problems, etc.); SPARQL rules are used to enhance typical ontology-based reasoning with complex activity and problem detection, temporal reasoning and incremental knowledge updates.

Similar to the ontology-based approaches introduced in Section 3, ontologies are used by SI to model the Dem@Care domain of interest; in particular, OWL 2 is the language of choice for representing the underlying knowledge structures, exploiting this way the advantages that OWL offers (representational and reasoning power, interoperability, means for sharing heterogeneous information, widely used reasoning engines). In order to overcome OWL's expressiveness limitations, SI follows the paradigm of the interpretation frameworks introduced in Section 3, combining rules with ontologies in order to support complex activity interpretation. Although rule languages have been used to support temporal reasoning tasks in the literature, the assertion of new individuals is an issue that is not supported by the languages employed; therefore the new individual assertion is usually handled by some external mechanism leading this way to ad-hoc solutions. Employing SPARQL rules allows addressing both these issues using a standardised query language, promoting this way interoperability and providing a novel solution that goes beyond the state of the art.

In the following, the methodologies implemented by the two components are presented in detail.

4.1 Complex Activity Recognition

The Complex Activity Recognition (CAR) component follows a hierarchical model-based approach and uses the generic constraint-based ontology language proposed by Vu *et al*. [53] to describe event models. The event models are built using *a priori* knowledge of the experimental scene (e.g. 3D geometric and semantic information) and attributes of objects (herein called Physical Objects) detected and tracked by the multi-sensor monitoring system. The ontology is a declarative and intuitive language based on a natural terminology to allow domain experts to easily change the models. The *a priori* knowledge of the experimental scene consists of the decomposition of a 3D projection of the scene floor plan in a set of spatial zones which represent semantic information of event models (e.g., TV zone, ArmChair zone, OfficeDesk zone, Coffee machine zone); the corresponding objects of interest (e.g. TV, armchair, office desk, coffee machine) are also part of the a priori knowledge.

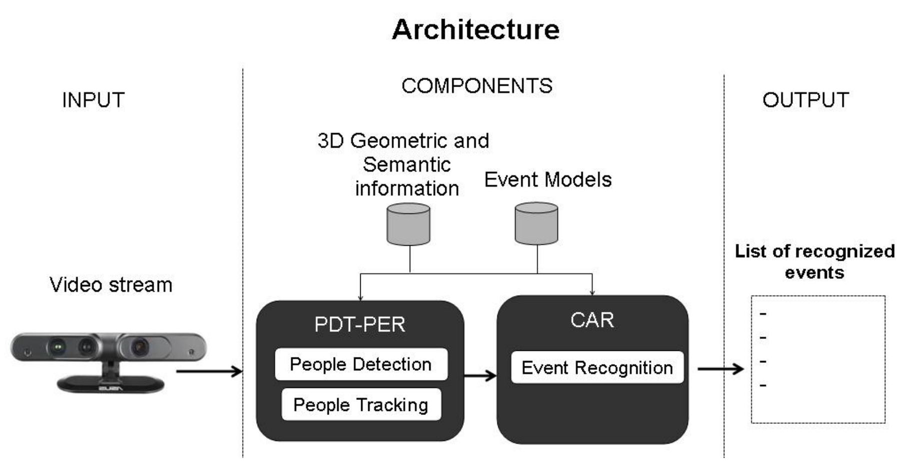


Figure 4-1 Overall architecture of the Complex Activity Recognition framework

Figure 4-1 presents the architecture of the activity recognition. Event and activity terms are herein assumed as the same, although event term has a broader sense than activity in the literature. In the current version of CAR, its inputs are the set of people in the scene and the lower level events (e.g., of primitive state level) detected by PDT-PER component. PDT-PER component takes as input an RGB-D video stream. Both PDT-PER and CAR component uses *a priori* information of the scene (such as contextual objects and zones) and event models provided by domain experts to compute events.

4.1.1 Event model representation

Event models are composed of six components:

- Physical Objects: refers to real objects involved in the recognition of a modelled activity. Examples of physical object types are: mobile objects (e.g. person, vehicle), contextual objects (equipment) and contextual zones (office zone).
- Components: refers to sub-activities that the model is composed of.
- Forbidden Components: refers to activities that should not occur in case the activity model is recognised.
- Constraints: conditions that the physical objects and/or the components should satisfy. These constraints could be logical, spatial and temporal.
- Action: they trigger a specific action which would be performed when an activity of the model is recognised (e.g. send a SMS to a care giver to check a patient over a possible falling down event).

Three types of Physical Object are currently defined: Person, Contextual Zone and Contextual Object. The Person class is an extension of a generic class named Mobile, which captures general information of mobile objects (e.g., x-y-z coordinates, width, height, and depth). The Person class models attributes, like body posture. The Contextual Zone and Object classes refer to zones and objects whose interaction is of particular interest for an activity description, and as aforementioned, constitute part of the prior knowledge.

Constraints define conditions that physical object property(ies) or event components should satisfy for the described event model to be recognised. They could be a-temporal, such as spatial constraints, or temporal, such as the definition of a temporal relationship between two event models: Person_in_zone1 before Person_in_zone2. Temporal constraints are defined using Allen's interval algebra (e.g., BEFORE, MEET, and AND) [3].

The ontology hierarchically categorises activity models according to their complexity (in ascending order): primitive state, primitive event, composite state and composite event.

- Primitive State: models an instantaneous value of a property of a physical object (Person.posture = sitting, or Person.position = inside TV zone area).
- Composite State: refers to a composition of two or more primitive states.
- Primitive Event: models a change in a value of a physical object property (e.g., Person changes from Sitting to standing posture).
- Composite Event: refers to an event composed of two other event models. This model type is generally used to model an existing temporal relationship among the composed ones (the model Person_changes_from_sitting_to_standing should happen before

Person_in_Corridor Zone). The components of a Composite Event model can be of any event model category, even a composite model.

Figure 4-2 presents an example of a primitive state model. This model checks whether the state of the attribute Posture of a Person fits the desired posture type.

```
PrimitiveState( Person_bending,
  PhysicalObjects( (p1:Person) )
  Constraints ( (P1->Posture = Bending) )
)
```

Figure 4-2 Model for Primitive State *Person Bending*

Figure 4-3 presents a composite activity model example, that of "Person using Office Desk". As shown, two primitive states, namely "Person_Inside_Zone_OfficeDesk" and "Person_closeTo_OfficeDesk" are involved. The first is used to verify that the person is inside the semantic zone "OfficeDesk", while the second is used to verify the person's proximity to the "Desk" object. To recognise the composite event, two constraints need to be satisfied: first, the person needs to be inside the respective zone for more than 8 seconds; second, both sub-events (components) should have been detected in the current frame.

```
CompositeEvent(Person_using_OfficeDesk,
PhysicalObjects((p1 : Person), (z1 : Zone), (eq1 : Equipment))
Components( (c1: PrimitiveState Person_Inside_Zone_OfficeDesk(p1, z1))
            (c2: PrimitiveState Person_closeTo_OfficeDesk(p1, eq1)))
Constraints((c1->Interval and c2->Interval) (duration(c1) > 8))
Alarm ((Level : URGENT))
)
PrimitiveState(Person_Inside_Zone_OfficeDesk,
PhysicalObjects((p1 : Person), (z1 : Zone))
Constraints ((p1->Position in z1->Vertices)
            (z1->Name = zoneOffice))
Alarm ((Level : NOTURGENT))
)
PrimitiveState(Person_closeTo_OfficeDesk,
PhysicalObjects((p1 : Person), (eq1 : Equipment))
Constraints ((distance(p1->Position, eq1->Position) <=
MAX_DISTANCE_CLOSE_TABLE)
            (eq1->Name = OfficeTable))
Alarm ((Level : NOTURGENT))
)
```

Figure 4-3 Activity Modelling of Person using Office Desk

4.1.2 Modelling Activities Derived by Different Sensors

Events generated by different sensors are herein modelled using the Primitive State category as it is the basic building block of the ontology. The choice of this model type relies on the fact Primitive State instances can be later filtered using hierarchically higher models (e.g., Primitive Event, Composite Event). This approach is particularly useful for modelling events recognised by heterogeneous sensors, as only the sensor output needs to comply with the ontology vocabulary. For instance, we present the modelling of a Person posture using events recognised by a video-camera and a wearable inertial sensor. The wearable inertial sensor (WI) provides the person current posture (Sitting, Standing), and Actimetry level value. Its posture be used as the major posture estimator for the Person posture description, or as complementary information to the one obtained from the analysis of the video camera by PDT-PER component. Herein the WI posture is added as an attribute to the Person class, in addition to the already existing value provided by the vision system (abbreviated as V, see Figure 4-4 for the Person class description).

```
class Person:Mobile {
    String PostureV;
    String PostureWI;
 }
```

Figure 4-4 Attributes of the class Person

Figure 4-5 presents an example of declaration of Primitive state model which uses the attribute “posture” provided by the WI.

```
PrimitiveState( Person_sitting_WI,
  PhysicalObjects(  (p1 : Person) )
  Constraints( (p1->PostureWI = Sitting) )
)
```

Figure 4-5 Primitive state mapping a wearable sensor value

Figure 4-6 presents an example of a Composite Event which is composed of the event models of the two sensors (WI and video camera).

```
CompositeEvent(Person_Sitting_MS,
 PhysicalObjects(  (p1:Person), (z1:Zone), (eq1:Equipment))
 Components(
      (c1: PrimitiveState Person_sitting_V (p1))
      (c2: PrimitiveState Person_sitting_WI(p1)))
 Constraints( (c1 AND c2) )
)
```

Figure 4-6 Composite event Person Sitting; V: vision-system; WI: wearable inertial sensor

Figure 4-7 presents the Composite Event "Person sitting and using Office Desk", which combines the described multi-sensor events with an event model which checks the person position into a contextual zone. Two constraints are also defined. The first constraint establishes both components need to be detected at the current instant, and the second establishes that the component 1 (Person inside Office Desk zone) has to be already recognised for at least the past 2 seconds.

```
CompositeEvent( Person_sitting_and_using_OfficeDesk,
 PhysicalObjects( (p1:Person), (z1:Zone) )
 Components(
      (c1:CompositeEvent Person_insideOfficeDeskZone(p1,z1))
      (c2:PrimitiveState Person_sitting_MS (p1)))
 Constraints( (duration(c1) > 2 ) (c1 and c2)  )
 )
```

Figure 4-7 Composite Event Person Sitting and using Office Desk

In the above, we have outlined how activities (Events) are modelled in the developed framework. For certain applications, where multiple sensors are available, their fusion can provide more information of a phenomenon. The presented framework enables to handle multiple sensors at decision (event or activity) level by *a priori* modelling multiple sensor activities as sub-components of an activity model (Composite Event). For cases where conflicting evidence arises by the different sensors, disagreement is handled by computing a confidence level for each of the detected activities, as described in the following. Activity fusion at decision level is preferred as it abstracts the sensor (its software and hardware implementation), providing flexibility to the system in case of a change or removal of a sensor.

4.1.3 Solving Conflicts

As aforementioned, the developed Complex Activity Recognition (CAR) framework can also handle disagreement among events recognised by different sensors (in terms of mutually exclusive events). Conflict solving is performed as follows: the Primitive State instantaneous likelihood and its temporal reliability are computed according to the method proposed by Romdhane *et al.* [45]. Once the event probability for each recognised event is computed per sensor, Dempster-Shafer theory is applied to decide which of the conflicting detected events models is performed. An evaluation of the proposed framework for conflict handling is presented in Section 5.1.2.

Primitive State Instantaneous likelihood

A Primitive State is generally associated to a feature value of a physical object. In PDT-PER the height of the person is used to identify if he/she is standing/sitting based on a threshold method (average sitting height). If the person height is equal to or below the average sitting height, the person is considered Sitting, otherwise Standing. We should take into account that failures can happen in the vision algorithms due to illumination changes among other factors, and these failures will affect the estimation of a person height and consequently the Posture identification.

We herein consider that the features used by the Primitive states (e.g., height) follow a Gaussian distribution, therefore a learning step is *a priori* performed to learn the distribution parameters mean (μ) and variance (σ^2) of the feature concerning the event model associated with the sensor (*e.g.*, height of the Person when Sitting and Standing at each sensor).

We compute the instantaneous likelihood of the feature at the current instant based on the learnt distribution parameters (μ, σ^2). See Equation 1.

$$PROB_{k,e,i}^{inst} = e^{-\frac{{Height_{\Omega,k,i} - \mu_{\Omega,i}}^2}{2\sigma_{\Omega,i}^2}} \quad (1)$$

where, k: frame number, Ω: event model, i: sensor.

Since the standard Gaussian distribution likelihood can be considered as a belief level value, it is herein employed as "how strongly it is believed that the event result of the sensor *i* is true at the evaluated time instant (e.g., frame)".

Temporal reliability of a Primitive State

The Primitive State likelihood computation considers whether sitting and standing are recognised in a precise moment (*e.g.*, current frame), but as discussed before, height feature could be misestimated by a failure in the vision algorithms (e.g., image segmentation, people tracking). Therefore, it is interesting to take into account feature probabilities of previous instants to filter possible instantaneous fluctuations of a feature value. Equation 2 and 3 presents an adapted computation of a temporal reliability with a time window of fixed size proposed by Romdhane *et al*. [45]. A cooling function is used to reinforce the information brought by near instants' and lessen those related to farther ones. Generally, a primitive state is a continuous process which lasts for seconds or even minutes depending on the domain of application of an activity recognition system. Therefore, the window size parameter should be tuned accordingly to fit at least the minimum time interval of the respective primitive state.

$$PROB_{\Omega,k,i}^{temp} = \frac{PROB_{\Omega,k,i}^{inst} + M}{\sum_{t=k-w}^{t=k-1} e^{-(k-t)}} \quad (2)$$

$$M = \sum_{t=k-w}^{t=k-1} [e^{-(k-t)}(PROB_{\Omega,k,i,}^{temp} - PROB_{\Omega,k,i,}^{inst})] \quad (3)$$

Where: k: frame number, w: temporal window size, i: sensor number, Ω: target event model

Primitive State Conflict Handling

After obtaining each sensor estimation of Primitive State confidence (Event Temporal reliability), it is necessary to combine the different sensor evidence to infer the Primitive State that is been performed. Dempster-Shafer Theory (DST) is herein proposed for event conflict solving. DST was proposed by Dempster [20] and improved by Shafer [48], and extends the Bayesian inference application by allowing uncertainty reasoning based on incomplete information. The major components of the evidence theory are the frame of discernment (??), and the basic probability assignment (BPA). The frame of discernment contains all possible mutually exclusive hypotheses (e.g. Sitting, Standing, none of those). The BPA is a function m: $2^{\Theta} \rightarrow [0, 1]$ related to a proposition satisfying conditions (X) and (Y), Ali *et al*. [1],

$$m(\emptyset) = 0 \quad (X)$$

$$\sum_{A\in\Theta} m(A) = 1 \qquad \text{(Y)}$$

where A is any subset of the frame of discernment, and ∅ refers to the empty set.

For any A $\in 2^{\Theta}$, m(A) is considered as the subjective confidence level on the event A. Accordingly, the whole body of evidence of one sensor is the set of all the BPAs greater than 0 under one frame of discernment. The combination of multiple evidences defined on the same frame of discernment is the combination of the confidence level values based on BPAs (e.g., pre-defined by experts). Given two sensors (1 and 2), where each sensor has its body of evidence (m1 and m2), these are the corresponding BPA functions of the frame of discernment. We herein adapt the combination rule proposed by Ali *et al*. [1], as it has been demonstrated to be efficient for the combination of evidences from multiple sensors. Equations 4 and 5 present the mass function for computing Sitting and Standing primitive states, respectively:

$$(m_{RGB} \oplus m_{RGBD})(\text{Sitting}) = \frac{1-(1-m_{RGB}(Sitting))\times(1-m_{RGBD}(Sitting))}{1+(1-m_{RGB}(Sitting))\times(1-m_{RGBD}(Sitting))} \qquad (4)$$

$$(m_{RGB} \oplus m_{RGBD})(\text{Standing}) = \frac{1-(1-m_{RGB}(\text{Standing}))\times(1-m_{RGBD}(\text{Standing}))}{1+(1-m_{RGB}(\text{Standing}))\times(1-m_{RGBD}(\text{Standing}))} \qquad (5)$$

An example For instance, activity recognition with data coming from two cameras is used to exemplify this approach (RGB and RGB-D camera) in Section 5.1.2

Note that the events recognised in the current frame are analysed with respect to the events detected in previous frames. If an instance of an event model involving the same physical objects is found in the close past, these instances are assumed to be the same. In the described case the previous instance of an activity has its “end time” attribute updated with the current time. The “close past” distance is defined as in parameter “Const Prolongation” and is the same for all activity models. The output of a determined activity instance (same event ID) could be repeated by several frames before the instance of end finally ends. In case an event model instance is recognised but its “close past” distance to a previously detected instance (of that event) is higher than the threshold established in “Const Prolongation”, it is going to be considered a new instance (different event ID).

4.2 Semantic Interpretation

The abstract architecture of the Semantic Interpretation framework is depicted in Figure 4-8 and consists of the *representation* and *interpretation* layers. The representation layer provides the ontology vocabularies for modelling the Dem@Care application context, such as activities, measurements, problems, summaries, patients, locations, objects and so forth. The interpretation layer encapsulates the inferencing capabilities of the framework by combining the OWL reasoning and SPARQL rule execution processes. In sections 4.2.1 and 4.2.2 we briefly present the role and functionality of each layer, while section 4.2.3 presents in detail the hybrid reasoning architecture and algorithm.

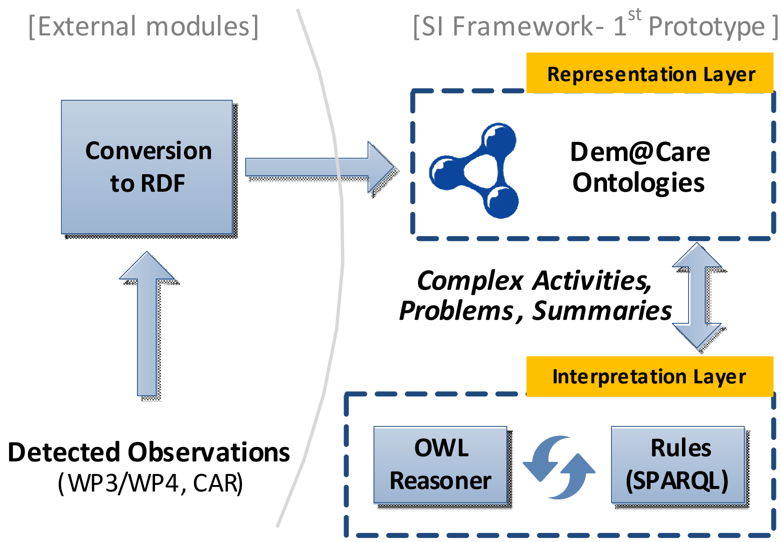


Figure 4-8 The abstract architecture of the Semantic Interpretation framework

4.2.1 Representation Layer

The representation layer provides the vocabulary and infrastructure for capturing and storing information relevant to the lab and home/nursing home environments, such as:

- atomic activities and measurements detected by means of WP3/WP4 monitoring and analysis components (e.g. speak, body temperature, light level), and complex activities inferred by WP5 CAR and SI components (e.g. having meal, sleeping, napping, answering the phone, having a face-to-face conversation, etc.)
- problems and situations that the clinicians need to be informed about (e.g. missed meals, excessive napping, insufficient communication attempts, nocturia, etc.)
- clinically relevant attributes and summaries (e.g. sleep efficiency and duration, number of daily telephone and face-to-face interactions, night sleep summaries, etc.)

Since the first version of the Dem@Care ontology (version 1.1 – 02 November 2012 summarised in deliverable D5.1 "Semantic Knowledge Structures and Representation", several revisions have been made so as to ensure that the representation layer adequately covers the knowledge that it is expected to capture through descriptive, yet lightweight ontology models. The modelling capabilities have been designed with a minimum of semantic commitment to guarantee maximal interoperability. As such, the Dem@Care ontologies can be aligned with relevant foundational ontologies, such as SEM [24] and Ontonym [51], reusing existing vocabularies for modelling different aspects of activities, e.g. entities, places and so forth. The current version of the Dem@Care ontology (version 2.0 – 02 May 2013) includes the following major revisions:

- All the temporal-related object properties, such as `startTime`, `endTime` and `date`, have been converted to datatype properties for simplicity. The OWL Time ontology is no longer part of the Dem@Care ontology model and therefore, instant, interval and date values are represented as plain literals, using the `xsd:dateTime` or `xsd:date` datatypes, and not as individuals of temporal-related classes.

- Several classes have been added to the event hierarchy to support the various reasoning tasks, such as the `Asleep`, `Awake`, `HangUpPhone`, `Conversation`, `DiscardTableObject`, `Visit` atomic activities and the `MovingIntensity`, `TotalTimeAsleep`, `VerbalReactionTime`, `NumberOfSteps` measurements.
- The additional abstraction level on top of the home/nursing home, lab and event ontologies that was formalised by the `Observation`, `InterpretationResult` and `Report` descriptive classes, has been deprecated, simplifying the conceptual model of the Dem@Care ontology application context.
- The definition of the `Event` class has been extended with the `isProvidedBy`, `hasReporting` and `hasPlausability` property assertions of the descriptive vocabulary (see Figure 4-9). Furthermore, the `relatedTo` object property assertion has been added to the `Activity` (subclass of the `Event` class) class for linking activities with relevant scene objects.
- Two object property assertions have been added to the `Problem` class, namely the `isProblemOf` and `date` property assertions, to support the association of a problem with the patient and the date of occurrence, respectively (see Figure 4-10).

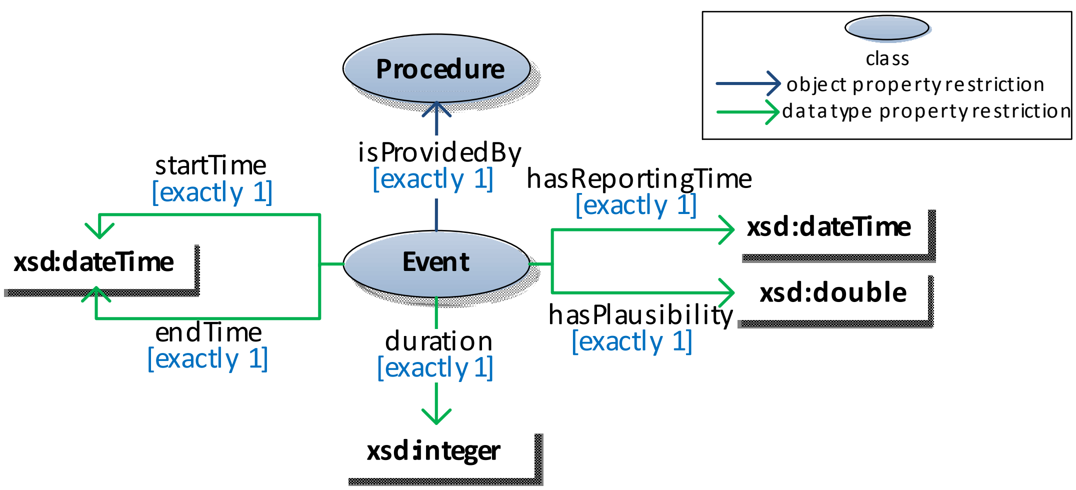


Figure 4-9 The property assertions of the `Event` concept

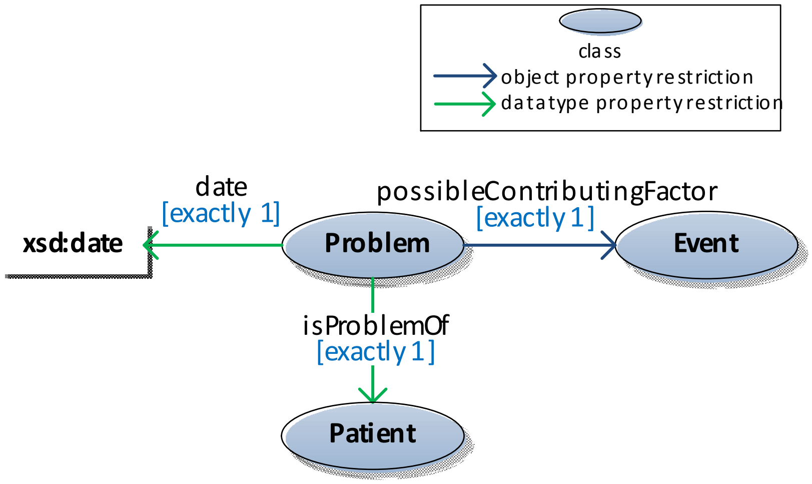


Figure 4-10 The property assertions of the `Problem` concept

4.2.2 Interpretation Layer

The interpretation layer provides a reasoning framework over the representation layer for the integrated interpretation of the PwD's behaviour and recognition of clinically relevant activities, problems and summaries. This is achieved through the combination of the OWL 2 reasoning paradigm and the execution of SPARQL[8] rules in terms of `CONSTRUCT` query patterns over RDF graphs. Essentially, the aim of the hybrid architecture, that is described in section 4.2.3 , is to define a reasoning framework able to deliver key inferencing tasks in the Dem@Care application context that are not supported by the standard semantics of the OWL ontologies, such as:

- **Temporal reasoning:** The ability to reason over the temporal extensions of activities is crucial for the successful correlations of activities and the identification of useful behavioural situations. However, OWL provides no support for temporal reasoning. In order to address this shortcoming, SPARQL rules are used to handle the temporal dependencies among activities, describing the temporal relations (Allen's temporal operators [1]) and the way contextual information can be combined in order to derive the various interpretation elements, such as activities, problems and summaries.
- **Complex activity correlations:** The schema-level axioms in OWL can model only domains where individuals are connected in a tree-like manner [52], [37]. In the activity interpretation domain, however, there is a need to model general relational structures among individuals, i.e. relations among individuals that are not connected. In the interpretation layer, this expressive limitation of OWL is addressed by utilising SPARQL rules for the description of the complex activity correlations that drive the activity recognition procedure.
- **Assertions of individuals:** The successful derivation of composite activities, problems and summaries strongly depends on the ability to update the underlying ontology model by asserting new individuals. With OWL, such assertions are only feasible by external reasoning services, since OWL semantics does not allow the

[8] http://www.w3.org/TR/sparql11-query/

modelling of ABox assertions that refer to named individuals not present in the KB. In the interpretation layer, the derivation of new individuals is based on the native capabilities of SPARQL to update the underlying contextual (RDF) model by asserting composite activity, problem and summary individuals.

SPARQL is a declarative language recommended by the W3C for extracting and updating information in RDF graphs, and it is used in SI to address the aforementioned limitations of the standard OWL semantics relevant to the Dem@Care application context. More specifically:

- Allen's temporal operators are realised as a set of custom SPARQL functions, exploiting the native capabilities of SPARQL for basic `xsd:dateTime` comparisons in `FILTER` clauses.
- SPARQL is an expressive language that allows the description of quite complex relations among activities. The semantics and complexity of the SPARQL query language have been fairly studied theoretically, showing that SPARQL algebra has the same expressive power as relational algebra [42],[4].
- New activity individuals are generated based on the graph pattern matching facilities of SPARQL.

Although SPARQL is mostly known as a query language for RDF, by using the `CONSTRUCT` graph pattern, it is able to define SPARQL rules that can create new RDF data, combining existing RDF graphs into larger ones. Such rules are defined in the interpretation layer in terms of a `CONSTRUCT` and a `WHERE` clause: the former defines the graph patterns, i.e. the set of triple patterns that should be added to the underlying RDF graph upon the successful pattern matching of the graphs in the `WHERE` clause. Figure 4-11 presents two examples of SPARQL rules; the rule in Figure 4-11 (a) can be used to link activity instances to the locations of their associated objects, whereas the rule in Figure 4-11 (b) associates patients with the location of the activities performed (triple variables are marked by the use of "?").

```
# activity location
CONSTRUCT
{
      ?act :location ?l.
}
WHERE
{
      ?act a :Activity ;
            :relatedTo ?obj .
      ?obj :location ?l .
}
```

(a)

```
# patient location
CONSTRUCT
{
      ?p :location ?l .
}
WHERE
{
      ?p a :Patient ;

      :isActorOf ?act .
      ?act a :Activity ;
            :location ?l .
}
```

(b)

Figure 4-11 Examples of SPARQL `CONSTRUCT` graph patterns

4.2.3 Hybrid Reasoning Architecture

The hybrid reasoning architecture of the interpretation layer combines the results of two reasoning modules, as it is depicted in Figure 4-8. More specifically, the O*ntology Reasoning Module* handles the standard OWL semantics[9] of the representation layer, e.g. class subsumption, property domain/range restrictions, instance class memberships, property relationships, e.g. transitive, inverse, and so forth, by incorporating an existing OWL 2 reasoner, such as Pellet [49], Hermit [38] and OWLIM [10]. For example, the inverse relationship between the `hasAgent` and `isAgentOf` properties of the event ontology is handled directly by the ontology reasoner, without implementing from scratch any custom reasoning services. More complex semantic correlations, however, cannot be directly represented in the representation layer and inferred by standard OWL reasoners. The *Rule Reasoning Module* aims to address the limitations of the OWL ontology reasoning paradigm discussed in section 4.2.2 by allowing the definition and execution of SPARQL rules for deriving complex activities, problems and summaries. In the following, we describe the three rule types that are currently supported by the interpretation layer.

Complex Activity Rules (Lab, Home/Nursing Home)

The aim of the complex activity rules is to meaningfully aggregate, correlate and interpret event-related[10] observations originated from the various analysis components of the Dem@Care system, so as to further classify existing activity individuals or assert new ones. More specifically, the complex activity rules can be classified in the following two categories:

- **Classification rules:** The classification rules propagate activity instances in the activity concept hierarchy based on their temporal dependencies with other events. Essentially, the classification procedure can be considered as a temporal-driven rule-based variant of the DL instance realisation procedure [6], allowing temporal information to drive the computation of class extensions, i.e. the set of instances that belong to a concept. Figure 4-12 presents the conceptual structure of a classification rule. The `WHERE` clause contains graph patterns that match event instances in terms of their types, property values and temporal constraints. The `CONSTRUCT` clause defines the classification of an individual matched in the `WHERE` clause (`?actM`) in a new class (`<ActivityClass>`).
- **Composition rules:** The composition rules derive composite activities, that is, activities composed of other activities (sub-activities). The representation of the composite activities requires the generation of new individuals, in contrast to the classification semantics that considers only existing activity individuals. Figure 4-13 presents the conceptual structure of a composition rule that allows the assertion of new named activity individuals (`?new`) in the `CONSTRUCT` clause, when the graph patterns in the `WHERE` clause are successfully matched.

---

[9] http://www.w3.org/TR/owl2-overview/#Semantics

[10] The Dem@Care ontology model defines four subclasses of the Event class, namely Activity, Measurement, ObjectEvent and State.

```
CONSTRUCT
{
      # derivation of additional class type for existing individual
      ?actM a <ActivityClass>.
      ...
}
WHERE
{
      # graph pattern matching on events and their property values
      ?ev1 a <EventClass1> ;
      ...
      ?actM a <ActivityClassM> ;
      ...
      ?evN a <EventClassN> ;

      # temporal filters
      FILTER (before(?ev1, ?actM) && ...))
      ...
}
```

Figure 4-12 The abstract structure of a classification rule

```
CONSTRUCT
{
      # a new composite activity individual
      ?new a <ComplexActivityClass>.
      ...
}
WHERE
{
      # graph pattern matching on events and their property values
      ?ev1 a <EventClass1> ;
      ...
      ?evN a <EventClassN> ;

      # temporal filters
      FILTER (:before(?ev, ?ev) && ...)
      ...

      # generation of unique URI
      BIND (<URI> as ?new)
}
```

Figure 4-13 The abstract structure of a composition rule

Assessment Rules (Home/Nursing Home)

The assessment rules recognise clinically relevant situations of the patient. The first version of the Dem@Care prototype supports assessment rules for the recognition of situations that indicate problems or possibly problematic behaviours that need to be highlighted to the clinician, e.g. a nocturia problem. Figure 4-14 presents the conceptual structure of a problem-related assessment rule that populates the `Problem` class hierarchy with new named individuals (`?new`) by integrating events and checking their temporal extensions.

```
CONSTRUCT
{
      # a new problem-related individual
      ?new a <ProblemClass>.
      ...
}
WHERE
{
      # graph pattern matching on events and their property values
      ?ev1 a <EventClass1> ;
      ...
      ?evN a <EventClassN> ;

      # temporal filters
      FILTER (:before(?ev1, ?ev2) && ...)
      ...

      # generation of unique URI
      BIND (<URI> as ?new)
}
```

Figure 4-14 The abstract structure of a problem-related assessment rule

Summary Rules (Home/Nursing Home)

The summary rules aggregate and summarise the results of the interpretation layer, offering a single point for collection of the PwD's contextual information. The first version of the Dem@Care prototype focuses on daily summaries that contain information about the performance of patients in daily activities, such as sleep and social interactions. Figure 4-15 presents the conceptual structure of a summary rule that correlates and aggregates event-related knowledge, such as activities and measurements, to generate summary individuals.

```
CONSTRUCT
{
      # a new summary-related individual
      ?new a <DaySummaryClass>.
      ...
}
WHERE
{
      # graph pattern matching on event types and property values
      ?ev1 a <EventClass1> ;
      ...
      ?evN a <EventClassN> ;

      # temporal filters
      FILTER (:before(?ev1, ?ev2) && ...)
      ...
```

```
        # aggregation logic
        { SELECT (COUNT(?sum)) ... }

        # generation of unique URI
        BIND (<URI> as ?new)
}
```

Figure 4-15 The abstract structure of a summary-related rule

Hybrid Reasoning Algorithm

Assuming that KB_{atomic} is a set of atomic observation assertions, R_{OWL} is the OWL Reasoning Module and R_{SPARQL} is the Rule Reasoning Module, the algorithm in Figure 4-16 describes the hybrid reasoning procedure that extends the KB_{atomic} set with additional inferences, i.e. complex activities, problems and summaries. More specifically, the architecture follows an iterative combination of the results of the two reasoning modules. Initially, the OWL reasoning module is used over the KB_{atomic} set to derive inferences based on the standard OWL semantics ($R_{OWL}(KB_{atomic})$). These inferences are added back to the KB_{atomic} set (line 3) that is subsequently used as the underlying model of the rule-based reasoning module (line 4). The additional assertions ($KB_{derived}$) are further added to the KB_{atomic} set (line 5), completing a reasoning iteration. Each time a new assertion (individual) is derived, the algorithm assigns to it a URI (Universal Resource Identifier) that encapsulates the assertions used to derive it. Thus, each rule can only be fired once for a certain set of input assertions, ensuring termination [34][35]. If R_{SPARQL} does not produce any inferences, i.e. the $KB_{derived}$ set is empty, the procedure terminates (line 6) with the R_{atomic} set containing both the atomic and the inferred knowledge. Otherwise, a new reasoning iteration begins.

Require: $KB_{atomic} \neq \emptyset$
1: **repeat**
2: $KB_{derived} \leftarrow \emptyset$
3: $KB_{atomic} \leftarrow KB_{atomic} \cup R_{OWL}(KB_{atomic})$
4: $KB_{derived} \leftarrow R_{SPARQL}(KB_{atomic})$
5: $KB_{atomic} \leftarrow KB_{atomic} \cup KB_{derived}$
6: **until** $KB_{derived} \neq \emptyset$

Figure 4-16 The hybrid reasoning algorithm of SI

4.2.4 Implementation and Examples

The first version of the Semantic Interpretation framework is based on the OWLIM [10] semantic repository for the implementation of both the representation and interpretation layers. Although OWLIM is not a complete OWL 2 reasoner comparing, for example, to Pellet and Hermit, it is well-suited to our framework since it provides efficient reasoning and SPARQL-based querying services over OWL 2 ontologies [11]. In practice, however, any OWL 2 reasoner and triple store can be used that support SPARQL queries.

The SPARQL-based reasoning procedure has been realised using the SPARQL Inferencing Notation (SPIN [30]). In SPIN, SPARQL queries can be stored as RDF triples together with any RDF domain model, enabling the linkage of RDF resources with the associated SPARQL queries, as well as sharing and reuse of SPARQL queries. SPIN supports the definition of SPARQL inference rules that can be used to derive new RDF statements from existing ones through iterative rule application. In the following, we illustrate the basic capabilities of the SI framework through a scenario of night sleep monitoring in home. It should be noted that the current version of SI can process observations in an offline mode at the end of a *clinical day*. A clinical day involves the observations that are collected during the respective calendar day and the night sleep-related observations that usually span two calendar dates.

Night Sleep Monitoring Scenario

The scenario involves the aggregation and interpretation of atomic observations collected during the night sleep so as to detect nocturia problems and create night sleep summaries. Our scenario involves the following atomic observations:

- **Night sleep:** It refers to the overall night sleep duration (start/end time) of the person that is detected by the Gear4 Renew Sleep Clock component. A night sleep activity is represented as an individual of the `NightSleep` ontology class.
- **In bed:** It is detected by CAR (instance of the `InBed` ontology class) and represents the fact that the person is inside the bed zone.
- **In bathroom:** It is detected by CAR (instance of the `InBathroom` ontology class) and represents the fact that the person is inside the bathroom zone.

The SI framework is used to semantically interpret and combine the aforementioned atomic observations, so as to derive:

- **Bed exit activities:** Composite activities that are generated when the patient is outside the bed zone during the night sleep.
- **Night bathroom visits:** A bed exit is classified as a night bathroom visit if it also involves a bathroom visit.
- **Nocturia problem:** This problem is detected when there are more than three night bathroom visits during the night sleep.
- **Night sleep summary:** It summarises the night sleep activity of the person, providing information about the sleep duration, the number of bed exits, the number of bathroom visits and so forth.

Figure 4-17 and Figure 4-18 present the SPARQL rules for deriving `BedExit` activities and their corresponding temporal context. More specifically, the rule in Figure 4-17 generates `BedExit` instances by integrating `NightSleep` and `InBed` activities. The start time of the new `BedExit` instance (line 4) is defined by the end time of an `InBed` activity (line 13) that occurs during a `NightSleep` activity (line 14). The rule generates a unique URI using the URI of the `InBed` instance (`?ib`) as a seed, so as to ensure that only a single `BedExit` instance is generated for each `InBed` (lines 15, 16). The rule in Figure 4-18 serves as a helper rule that determines the end time of `BedExit` instances. Essentially, the end time of a `BedExit` activity is defined as the start time of the most recent `InBed` activity after the `BedExit` (line 23).

```
CONSTRUCT {
    ?new a :BedExit; # composition (new individual)
            :hasAgent ?pwd ;
            :startTime ?be_start ;
}
WHERE {
      ?ns a :NightSleep; #detected by Gear4 component
            :hasAgent ?pwd ;
      :startTime ?ns_start ;
      :endTime ?ns_end .
    ?ib a :InBed ; #detected by CAR component
            :hasAgent ?pwd;
            :endTime ?be_start .
      FILTER (:contains(?ns_start, ?ns_end, ?be_start)) .
      BIND (:new(?ib) as ?new) .
      FILTER NOT EXISTS {?new a [] . } .
}
```

Figure 4-17 The composition rule for deriving `BedExit` instances

```
CONSTRUCT {
    ?be :endTime ?ib_start.
}
WHERE {
      ?ns a :NightSleep; #detected by Gear4 component
      :hasAgent ?pwd ;
            :startTime ?ns_start ;
      :endTime ?ns_end .
      ?be a :BedExit ; #detected by SI component
            :hasAgent ?pwd ;
            :startTime ?be_start .
      FILTER NOT EXISTS {?be :endTime []. }.
      FILTER (:contains(?ns_start, ?ns_end, ?be_start)) .
    {
        SELECT ?ib ?ib_start ?pwd
        WHERE {
            ?ib a :InBed ; #detected by CAR component
               :hasAgent ?pwd ;
                          :startTime ?ib_start .
        }
        ORDER BY ASC (?ib_start)
    } .
    FILTER ((?ib_start > ?be_start)
                  && :contains(?ns_start, ?ns_end, ?ib_start)) .
} LIMIT 1
```

Figure 4-18 Helper rule for deriving the end time of `BedExit` activities

Figure 4-19 depicts the classification rule that derives `NightBathroomVisit` instances. The graph pattern in the `WHERE` clause matches `InBathroom` instances that are temporally contained in `BedExit` activities (line 12) and they are further classified in the `NightBathroomVisit` class (line 2). Note that the classification rules do not generate new individuals but classify existing ones. Therefore, the classified instances preserve their temporal extensions, e.g. the `NightBathroomVisit` individuals preserve their `InBathroom` temporal extension (start/end times).

```
 1:  CONSTRUCT{
 2:      ?ib a :NightBathroomVisit . #classification
 3:  }
 4:  WHERE {
 5:      ?be a :BedExit ;
 6:               :hasAgent ?pwd ;
 7:               :startTime ?be_start ;
 8:               :endTime ?be_end .
 9:      ?ib a :InBathroom;
10:               :hasAgent ?pwd ;
11:               :startTime ?ib_start .
12:      FILTER (:contains(?be_start, ?be_end, ?ib_start)) .
13:      FILTER NOT EXISTS {?ib a :NightBathroomVisit . } .
14:  }
```

Figure 4-19 The classification rule for deriving `NightBathroomVisit` instances

The rule in Figure 4-20 derives instances of the `NocturialProblem` class by counting the number of `NightBathroomVisit` instances that are temporally contained in a `NightSleep` activity (line 18). The rule generates an instance of the `NocturiaProblem` (line 2) for each grouped result that contains three or more `NightBathroomVisit` instances (line 21). The `extractClinicalDay` function extracts the `xsd:date` value from the start time (`xsd:dateTime`) of the activity instance that is passed as parameter.

```
 1:  CONSTRUCT{
 2:      ?new a :NocturiaProblem ;
 3:          :isProblemOf ?pwd ;
 4:          :date ?clinicalDay .
 5:  }
 6:  WHERE {
 7:    {
 8:        SELECT ?ns ?pwd (COUNT(?nbv) as ?counter)
 9:        WHERE {
10:              ?ns a :NightSleep;
11:                    :hasAgent ?pwd ;
12:                    :startTime ?ns_start ;
13:                    :endTime ?ns_end .
14:              ?nbv a :NightBathroomVisit ;
15:                    :hasAgent ?pwd ;
```

```
                        :startTime ?nbv_start;
                        :endTime ?nbv_end .
                FILTER(:contains(?ns_start, ns_end, nbv_start, nbv_end)).
        }
        GROUP BY ?ns ?pwd
        HAVING (COUNT(?nbv) >= 3)
    }
    BIND (:new(?ns, ?pwd) as ?new) .
    FILTER NOT EXISTS {?new a [] . } .
    BIND (:extractClinicalDay(?ns) as ?clinicalDay) .
}
```

Figure 4-20 The rule for deriving nocturia problems

Finally, Figure 4-21 and Figure 4-22 present the rules for generating night sleep summaries and populating the `numberOfBedExits` property, respectively. More specifically, the rule in Figure 4-21 is the initialisation rule that generates a `NightSleepSummary` instance (line 2) for each `NightSleep` activity (line 7). Figure 4-22 presents a rule example that populates the `numberOfBedExits` property of night sleep summary instances. The rule counts the number of `BedExit` activities that are detected during a `NightSleep` activity (line 8-18) and populates the corresponding property of the `NightSleepSummary` instance (line 2). Similar rules also exist for the other properties of the `NightSleepSummary` class of the Dem@Care ontology, such as, sleep duration, number of awakenings, and so forth.

```
 1:  CONSTRUCT{
 2:      ?new a :NightSleepSummary;
 3:          :forPatient ?pwd ;
 4:          :date ?clinicalDay ;
 5:  }
 6:  WHERE {
 7:      ?ns a:NightSleep;
 8:          :hasAgent ?pwd .
 9:      BIND (:new(?ns) as ?new) .
10:      FILTER NOT EXISTS {?new a [].} .
11:      BIND (:extractClinicalDay(?ns) as ?clinicalDay) .
12:  }
```

Figure 4-21 The rule for the initialisation of `NighSleepSummary` instances

```
CONSTRUCT{
    ?sleep_sum :numberOfBedExits ?counter .
}
WHERE {
    ?sleep_sum a :SleepSummary .
    FILTER NOT EXISTS {?sleep_sum :numberOfBedExits [] .} .
    {
        SELECT ?ns (COUNT(?be) as ?counter)
        WHERE {
            ?ns a :NightSleep;
                :hasAgent ?pwd ;
                :startTime ?ns_start ;
                :endTime ?ns_end .
            ?be a :BedExit;
                :hasAgent ?pwd ;
                :startTime ?be_start .
             FILTER(:intervalContains(?ns_start, ?ns_end, ?be_start)).
    } GROUP BY ?ns
}
```

Figure 4-22 The rule for populating the `numberOfBedExits` property of `NightSleepSummary` instances

4.3 Towards enhancing situation analysis

In addition to the tasks specified in Section 2, an underlying goal of the interpretation framework is to interact with WP4 and allow for more robust and adaptable visual sensing and activity recognition. To this end, inconsistencies/conflicts detected during the behaviour interpretation will be used to provide feedback to the respective low-level components, so that the latter can automatically adapt their models.

The underlying idea is to assign estimates of confidence and precision to WP4 activity features; these estimates will be used by the respective WP4 components to self-assess their performance in order to learn, adapt and improve over time. Representative features of interest include the PwD location and manipulated objects since they comprise key information for fine tuning the interpretation of activities in which the PwD may be engaged: manipulated objects or a recognised room that are consistent with an on-going detected activity give further confidence in the identification of the respective activity; likewise, the detected objects or rooms that are inconsistent with the ongoing detected activity provide the grounds for tuning the results.

For the moment, only the detection of these features has been addressed. Specifically, two components, the ORWC and RRWC that operate on data coming from a wearable camera, have been developed in WP4 for detecting objects and localisation information respectively (the list of objects and locations currently supported is given in Appendix A.3). Next steps include investigations towards assessing the performance of the object and room recognition algorithms through cross-checking the consistency of the detected objects/rooms with respect to the activity being detected.

5 Implementation of the Dem@Care Behaviour Interpretation

Figure 5-1 illustrates the two behaviour interpretation components, i.e. CAR and SI, within the overall Dem@Care system architecture. CAR uses as input the observations generated by PDT-PER and communicates its interpretation results (i.e. the recognised complex activities) to SI (via the Controller and KB manager components); in case a real-time alert needs to be issued (e.g. in case a fall is recognised), CAR results need also to be communicated to the WP6 GUI Backend component (again via the Controller). SI uses as input the observations of all WP3/WP4 components, besides PDT-PER, as well as the interpretation results of CAR. SI stores its results to the knowledge base (KB); querying the KB, the WP6 GUI Backend component retrieves the information necessary for feedback services. Sections 5.1 and 5.2 present the implemented CAR and SI web service components respectively, as well as preliminary evaluation considerations.

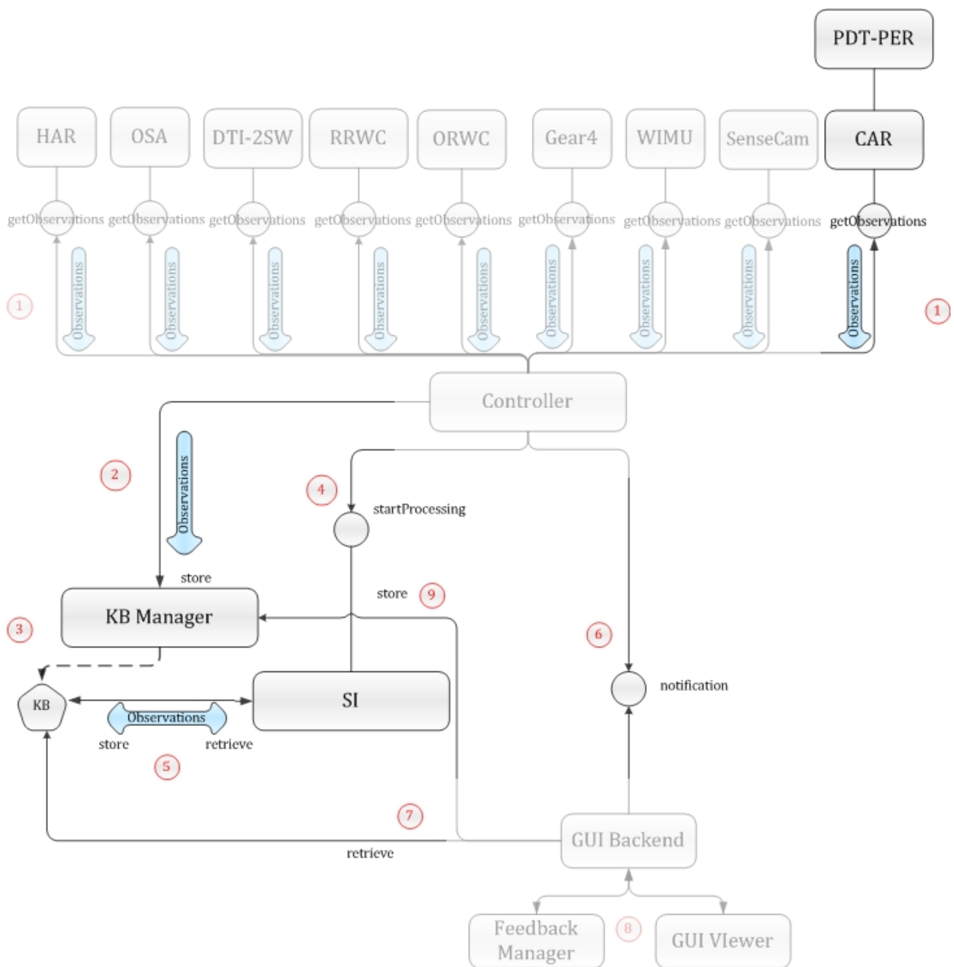


Figure 5-1 CAR and SI in the overall Dem@Care architecture

.

5.1 The Complex Activity Recognition software component

5.1.1 CAR Web Service

The CAR web service performs the semantic interpretation of the observations generated by the PDT-PER component.

The output of the recognition process is provided at frame rate and is stored in XML format and is saved in a PostGreSQL database. Figure 5-2 presents an example of output for a frame. The timestamp information is presented in the form of attributes of the XML tag SUVideoFrame. In this example, two activities have been detected: the primitive state “Person_standing” and the composite event “Person_using_PharmacyBasket”, which as shown is composed of the sub-event "Person_Inside_Zone_Pharmacy".


```xml
<?xml version="1.0"?>
<SUVideoFrame frameID="488" timeYear="2013" timeMonth="4" timeDay="16"
timeHour="11" timeMin="05" timeSec="05" timeMs="937">
  <ListActivities>
    <Activity ID="0" name="Person_standing" confidence="1"
    type="PrimitiveState" AType="URGENT" AText="" status="" hypothesisID="1">
      <TimeStart frameID="17" timeYear="2013" timeMonth="4" timeDay="16"
      timeHour="11" timeMin="05" timeSec="04" timeMs="143" uncertainty="0"/>
      <TimeEnd frameID="488" timeYear="2013" timeMonth="4" timeDay="16"
      timeHour="11" timeMin="05" timeSec="05" timeMs="937" uncertainty="0"/>
      <ListActivityPhysicalObjects>
        <ActivityPhysicalObject type="OBJECT" ID="1" name="p1"
        dynamicObjectID="0"/>
      </ListActivityPhysicalObjects>
      <ListSubActivities/>
      <Properties/>
      <ListActivityCameraCtrl/>
    </Activity>
    <Activity ID="12" name="Person_using_PharmacyBasket" confidence="1"
    type="CompositeEvent" AType="URGENT" AText="" status="" hypothesisID="1">
      <TimeStart frameID="364" timeYear="2013" timeMonth="4" timeDay="16"
      timeHour="11" timeMin="05" timeSec="04" timeMs="860" uncertainty="0"/>
      <TimeEnd frameID="488" timeYear="2013" timeMonth="4" timeDay="16"
      timeHour="11" timeMin="05" timeSec="05" timeMs="937" uncertainty="0"/>
      <ListActivityPhysicalObjects>
        <ActivityPhysicalObject type="OBJECT" ID="1" name="p1"
        dynamicObjectID="0"/>
      <ActivityPhysicalObject type="STATIC_ZONE" ID="11" name="zonePharmacy"
      dynamicObjectID="0"/>
      </ListActivityPhysicalObjects>
      <ListSubActivities>
        <SubActivity ID="11" name="Person_Inside_Zone_Pharmacy"/>
      </ListSubActivities>
      <Properties/>
      <ListActivityCameraCtrl/>
    </Activity>
   </ListActivities>
   <ListMobileObjects/>
</SUVideoFrame>
```

Figure 5-2 Example of XML recognition output

The content of the PostGreSQL database is then exported into observations following the Exchange Model syntax (see D7.1). During this process, the start and end times of the detected activities are pre-processed in order to deliver only one instance per detected activity. As such, assuming that the person with ID *p001* has been detected to be sitting in the sequence of frames captured between 11:05:04 to 11:15:44, the observation shown in Figure 5-3 would result.

```
<Observation xmlns:ns2="http://www.demcare.eu/exchange">
    <ns2:reportingTime>2013-04-16T11:16:30.442</ns2:reportingTime>
    <ns2:observedEvent xmlns:xsi="http://..." xsi:type="ns2:Activity">
        <ns2:startTime>2013-04-16T11:05:04.143</ns2:startTime>
        <ns2:endTime>2013-04-16T11:15:44.325</ns2:endTime>
        <ns2:type>Person_standing</ns2:type>
    </ns2:observedEvent>
    <ns2:provider xmlns:xsi=http://... xsi:type="ns2:ProcessingComponent">
        <ns2:id>CAR</ns2:id>
    </ns2:provider>
    <ns2:subject>
        <ns2:identifier>p001</ns2:identifier>
    </ns2:subject>
    <ns2:plausibility>1.0</ns2:plausibility>
</Observation>
```

Figure 5-3 Example CAR observation in exchange model syntax

CAR Web Service provides two methods: waitingMessage and getObservationsFromCAR. The first method is responsible for treating alert messages coming from the CAR component implementation, and translating them into the observation format defined by the Exchange Model. Its main goal is to make the Controller component aware of the detection of events that require a special treatment (e.g., PwD has fallen down). The alerted activities are exchanged between CAR component and CAR Web Service using Comma Separated Values format (See Figure 5-4 for an example).

The second method, getObservationsFromCAR, provides an interface to retrieve all observations (metadata) detected by CAR during a specific time interval and convert them into the Exchange Model format.

```
0; Person_standing ; 2011-09-12 14:28:18.556+02 ; 2011-09-12 14:29:55.938+02

2; Person_sitting ; 2011-09-12 14:28:20.937+02 ; 2011-09-12 14:28:36.806+02

3;  Person_close_chair_Inside_Zone_UseReadingTable  ;  2011-09-12  14:28:21.437+02;
             2011-09-12 14:28:21.808+02

5; Person_sitting_long ; 2011-09-12 14:28:22.437+02 ; 2011-09-12 14:28:36.806+02
```

Figure 5-4. Data format used by CAR component and CAR Web-Service

The list of activities whose recognition will be supported by CAR in the lab setting is provided in Appendix A1.1. A few activities like writing a check, establishing the account balance, counting backwards, and walking and counting backwards are listed, but their detection will be considered at future versions as they require the fusion of PDT-PER input with other sensor modalities, such as audio and wearable camera events.

5.1.2 CAR evaluation

A preliminary evaluation of CAR has been conducted using a variant of the Dem@Care lab protocol. More specifically, participants aged more than 65 years have been recruited by the Memory Centre (MC) of a collaborating Hospital. Inclusion criteria of the Alzheimer Disease (AD) group are: diagnosis of AD according to NINCDS-ADRDA criteria and a Mini-Mental State Exam (MMSE) [22] score above 15. AD participants which have significant motor disturbances (per the Unified Parkinson's Disease Rating Scale) are excluded. Controls participants are healthy in the sense of behavioural and cognitive disturbances. The participants are asked to undertake a set of physical activities and Instrumental Activities of Daily Living (IADL) in a hospital observation room furnished with home appliances. The monitored IADLs include eight activities, namely:

- Watch TV,
- Make tea/coffee,
- Write the shopping list of the lunch ingredients,
- Write a check to pay the electricity bill,
- Answer/Call someone on the Phone,
- Read newspaper/magazine,
- Water the plant
- Organise the prescribed drugs inside the drug box according to the weekly intake schedule.

Experimental recordings include the use of a 2D video camera (AXIS®, Model P1346, 8 frames per second), a 3D camera (Kinect® sensor) and a wearable inertial sensor (MotionPod®). Figure 5-5 shows the recording viewpoints of the RGB and RGB-D cameras, as well as examples of extracted information, namely activity levels and person trajectory.

All the sensor recordings are time synchronised, and no spatial correspondence is performed among the cameras. After the recording, the MotionPod sensor raw data is pre-processed using proprietary software to extract information about posture, and PDT-PER is used to provide information about people detection and tracking (see D4.2 for exact list of extracted attributes.)

The activity recognition framework performance has been evaluated with respect to three scenarios. First, the overall activity recognition framework is assessed using a single camera; second, we compare a mono and multi-sensor approach using an ambient camera and a inertial sensor; and finally, it is demonstrated the improvement brought by the proposed probabilistic approach using two video cameras. Recognition performance is evaluated using indices of sensitivity, precision, and F-score as defined in Equations 8, 9 and 10, respectively.

$$Sensitivity = \frac{TP}{TP+FN} \tag{8}$$

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

where: TP: True Positive rate, FP: False Positive rate, FN: False Negative rate

$$F-Score = 2 * \frac{Sensitivity * Precision}{Sensitivity + Precision} \tag{10}$$

Figure 5-5 Recording viewpoints and example information: ambient camera view (A); RGB-D camera view (B), showing also in close-up the inertial sensor worn by the participant; actimetry information (bottom of A); trajectory information (C).

Table 5-1 presents the performance of the framework at recognising IADLs. These activities are recognised through using seven composite event models, each composed of a Primitive State for the recognition of the person position inside a contextual zone (a priori defined), and another Primitive State for his/her proximity to a static contextual object located in this zone (also *a priori* defined, e.g., Phone station, Coffee machine). The activities "writing a check" and "writing a shopping list" are not differentiated, and are referred as Person using office desk. The activity "Organise the prescribed drugs…" is shortened as Person using pharmacy basket.

Table 5-1 Framework Performance using a Single Camera[11]

IADL	Sensitivity	Precision
Using Phone	72.83	85.50
Watching TV	80.00	71.42
Using OfficeDesk	92.72	58.62
Preparing Tea/Coffee	90.36	69.44
Using Pharmacy Basket	100.00	88.09
Watering plant	100	64.91
Reading	71.42	69.76
Average	**86.76**	**72.53**

[11] N: 29; 15 min. each, total of 280,400 frames (435 min).

Table 5-2 presents the comparison of a mono-camera and a multi-sensor approach. IADL complex event models used for the single camera are modelled as a sub-component of new complex event models, where the second component refers to a Posture Primitive State. IADL – Sitting/Standing refers to the F-Score rate of the posture when the participant is performing IADL. The Total related to the detection of "only IADL" takes into account the video-camera information only, therefore no change is seen in the results of the evaluation. The Deterministic modelling of Multi-sensor events improve by ~19% the precision index value of Sitting, but the information gain for standing posture is irrelevant. The overall performance of the system drops ("Only IADLs" x "IADL + Posture") as the IADL models now take into account also the posture estimation.

Table 5-2 Comparison of Mono and Multi-sensor approaches[12]

F – SCORE	Mono	Multi-sensor
Sitting during IADLs	52.00	71.00
Standing during IADLs	73.15	71.00
Total (IADL + Posture)	68.00	71.00
Total (Only IADL)	81.22	81.22

Table 5-3 compares the performance of the proposed framework for posture recognition when having input data from: i) the RGBs sensor, ii) the RGB-D sensor, and iii) both sensors. As shown, the proposed probabilistic framework improves the detection of posture-related Primitive states in most of the cases. The fusion approach has higher values of precision for sitting, and sensitivity for standing; in other cases it achieves at least better results than the worst individual performance of one camera. Preliminary results of this approach have been published in [17][33].

Table 5-3 Postures Recognition in Physical Tasks[13]

Posture	Sitting		Standing	
Sensor	**Sensitivity**	**Precision**	**Sensitivity**	**Precision**
RGB	84.29%	69.41%	79.82%	91.58%
RGB-D	100.00%	36.47%	86.92%	97.89%
Fusion	82.35%	91.30%	91.04%	95.31%

Summing up, the framework successfully recognises activities of daily living with a sensitivity average of 86.76 % and a precision of 72.53 % for 435 minutes of recordings (29 participants, 15 min. each). Its extension as a multi-sensor approach has improved by ~19 % the precision of sitting posture recognition during IADLs. But, none information gain is obtained for the recognition of Person standing posture. Future work will extend the evaluation of the probabilistic approach for other primitive activities in order to assess a larger variety of different sensors (heterogeneous and homogeneous) contribution to the activity recognition task.

[12] N: 9; 15 min. each; total of 64800 frames (135 min)

[13] N=10. A 5 second window is used for Temporal Probability.

5.2 The Semantic Interpretation Software Component

The representation and interpretation layers of the SI framework have been implemented as two separate Web service components: the *Knowledge Base Manager (KBM) component* that stores observations in the underlying KB (triple store), and the *SI component* that provides the interface for the invocation of the hybrid reasoning algorithm described in Section 4.2.3. Table 5-4 describes the abstract interfaces that are provided by the Web services.

Table 5-4 The SI and KBM web method interfaces

	Web Method Interface
KBM Web Service	`boolean setObservations(List<Observation) obs)`
SI Web Service	`boolean startDailyAnalysis(String patientID, XMLGregorianCalendar clinicalDay)`

5.2.1 KBM Web Service

The KBM Web service takes as input a list of observations that need to be stored in the KB. The service transforms the observations into RDF statements (triples) following the vocabulary of the Dem@Care ontologies and stores the results in the KB using the HTTP protocol. The execution of the Web services is terminated when all the provided observations have been successfully stored in the KB and the Controller (client) receives the output of the Web service in the form of an acknowledgement message (boolean value). Figure 5-6 depicts an example of an input observation in the Exchange Model format and Figure 5-7 presents the result of the transformation in the TURTLE syntax[14]. The example observation describes an `InBathroom` activity detected by CAR for the patient with `id` *id1*.

```
<Observation xmlns:ns2="http://www.demcare.eu/exchange">
    <ns2:reportingTime>2013-04-16T01:16:30.442+02:00</ns2:reportingTime>
    <ns2:observedEvent xmlns:xsi="http://..." xsi:type="ns2:Activity">
        <ns2:startTime>2013-04-16T01:05:04.143+02:00</ns2:startTime>
        <ns2:endTime>2013-04-16T01:15:44.325+02:00</ns2:endTime>
        <ns2:type>InBathroom</ns2:type>
    </ns2:observedEvent>
    <ns2:provider xmlns:xsi="http://..." xsi:type="ns2:ProcessingComponent">
        <ns2:id>car1</ns2:id>
        <ns2:type>CAR</ns2:type>
    </ns2:provider>
    <ns2:subject>
        <ns2:identifier>id1</ns2:identifier>
    </ns2:subject>
    <ns2:plausibility>1.0</ns2:plausibility>
</Observation>
```

Figure 5-6 Example `InBathroom` observation in the Exchange Model

[14] http://www.w3.org/TeamSubmission/turtle/

```
:InBathroom_1
      a event:InBathroom ;
      desc:hasPlausibility "1.0"^^xsd:double ;
      desc:hasReportingTime "2013-04-16T01:16:30.442Z"^^xsd:dateTime ;
      event:startTime "2013-04-16T01:05:04.143Z"^^xsd:dateTime ;
      event:endTime "2013-04-16T01:15:44.325Z"^^xsd:dateTime ;
      desc:isProvidedBy :car1 ;
      event:hasAgent :p1 ;
      event:startTime "2013-04-16T01:05:04.143Z"^^xsd:dateTime ;
      event:duration "640"^^xsd:long .
```

Figure 5-7 The `InBathroom` observation in the Dem@Care event ontology

5.2.2 SI Web Service

The SI Web service is invoked by the Controller to start the semantic interpretation procedure over the observations that are stored in the KB. The id of the patient and the clinical day of the analysis are provided as parameters to the input message. The execution of the Web service is terminated when all the available observations have been processed and no further inferences are generated by the hybrid reasoning algorithm. For example, assuming that the KB contains the `InBathroom` event in Figure 5-7 and the `BedExit` event in Figure 5-8, the SPARQL rule in Figure 4-19 derives the assertion `:InBathroom_1 a event:NightBathroomVisit`, classifying the `InBathroom_1` individual to the `event:NightBathroomVisit` class.

```
:BedExit_1
      a event:BedExit ;
      desc:hasPlausibility "1.0"^^xsd:double ;
      desc:hasReportingTime "2013-04-16T16:40:30.442"^^xsd:dateTime ;
      desc:isProvidedBy :si1 ;
      event:endTime "2013-04-16T01:22:16.421"^^xsd:dateTime ;
      event:hasAgent :p1 ;
      event:startTime "2013-04-16T01:04:23.541"^^xsd:dateTime ;
      event:duration "1073"^^xsd:long .
```

Figure 5-8 Example `BedExit` event in the Dem@Care event ontology

5.2.3 SI and KBM testing

The first version of the SI framework focuses on the aggregation and semantic correlation of the descriptions extracted from the analysis components of the Dem@Care system, assuming that perfect information is available and without taking into account uncertain, missing information and conflicts. Therefore, the performance of the semantic interpretation procedure strongly depends on the quality of the information (observations) that is provided as input to SI. For example, the rule in Figure 4-19 would always derive the correct consequences (`NightBathroomVisit` instances), provided that the temporal extensions of the `BedExit` and `InBathroom` observations are properly correlated and assuming that there are no conflicts or missing information in the KB. As such, the performance of the SI framework will be meaningfully evaluated in future versions, when the rule-based behaviour interpretation procedure of SI will be able to consider the observations of the lower level Dem@Care analysis components as unreliable or incomplete due to imprecision and errors introduced during their detection.

As far as the KBM component is concerned, Figure 5-9 depicts the scalability in terms of observation loading time, i.e. the time needed by the KBM component to transform the Exchange Model observations into Dem@Care ontology events (see Figure 5-6 and Figure 5-7) and the time the OWLIM repository needs to apply the OWL reasoning procedure over the generated events. Based on the preliminary results obtained by using a synthetic dataset, we have estimated that the representation of an XML observation in the KB requires approximately 30 RDF statements (triples). The dataset of Figure 5-9 involves 46000 observations that were loaded in the KB in approximately two minutes in a PC with Intel Core i7-3770 CPU (3.49GHz) and 8GB RAM, generating more than 1.3 million triples.

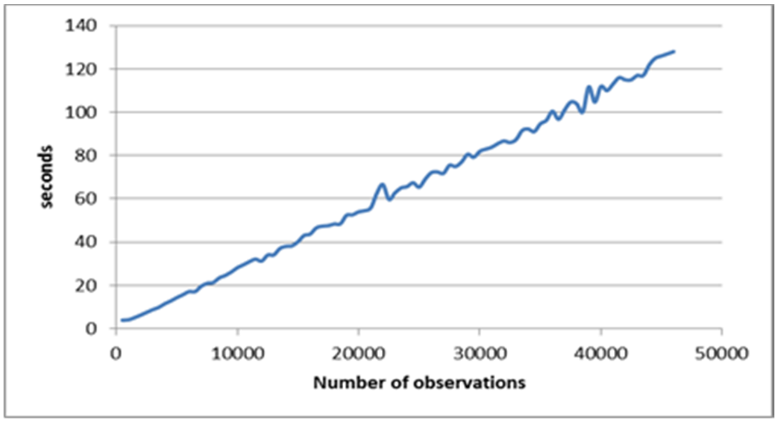


Figure 5-9 KBM scalability in terms of observation loading time.

6 Conclusions

This deliverable introduced the current version of the multi-parametric behaviour interpretation framework of Dem@Care, as realised by the two developed interpretation components, namely CAR (Complex Activity Recognition) and SI (Semantic Interpretation).

In its first version, the developed interpretation framework addresses basic interpretation functionalities that lie in the intersection of clinical requirements and available input observations from Work Package (WP) 3 and WP4 components. More elaborate interpretation tasks will be tackled as the WP3/WP4 analysis components mature and afford richer observations, and as CAR and SI capabilities evolve.

Next steps for CAR include its extension so as to incorporate observations from other WP3/WP4 components, in addition to the already deployed PDT-PER (People Detection, Tracking and Primitive Events Recognition) observations. Utilising, for instance ORWC (Object Recognition from Wearable Camera) observations about objects in the attention field of PwD, it will be possible to distinguish between fine-grained activities such as “paying the phone bill (writing a check)” and “establishing the account balance”.

Next steps for SI focus primarily on extending its inference capabilities so as to handle imperfect information (i.e. missing, uncertain and conflicting observations). Further extensions will allow incorporating additional pieces of information and will address more elaborate interpretation aspects, such as possible contributor factors for the problems detected in the PwD daily life. Within this line of investigation, the use of profile information in the interpretation decision support process is a key consideration.

The experimental evaluation through the pilots planned for months M18-M20 will provide the grounds for evaluating, in real conditions, the WP5 interpretation capabilities, and more important, for identifying the nature of limitations, improvements and extensions that need to be tackled in the future versions.

7 References

[1] T. Ali, P. Dutta, & H. Boruah. A new combination rule for conflict problem of Dempster-Shafer evidence theory. International Journal of Energy, Information and Communications,3(1). (February 2012).

[2] J. F. Allen. Towards a general theory of action and time. *Artificial intelligence* 23.2 (1984): 123-154.

[3] J. F. Allen. Maintaining Knowledge about temporal intervals. Communications of the ACM, 26(11):832-843, (November 1983).

[4] M. Arias, J. D. Fernández, M. A. Martinez-Prieto, and P. de la Fuente, "An empirical study of real-world SPARQL queries," 1st International Workshop on Usage Analysis and the Web of Data, Hyderabad, India, March 28th, 2011

[5] A. Artikis, M. Sergot, & G. Paliouras. A logic programming approach to activity recognition. In *Proceedings of the 2nd ACM international workshop on Events in multimedia* (pp. 3-8). ACM.Reference 2 (October 2010).

[6] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, & P. F. Patel-Schneider, Eds., The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, (2003).

[7] J. Benois-Pineau, R. Megret, H. Boujut, V. Buso, A. Satt, O. Toledo-Ronen, A. Sorin, A. Briassouli, K. Avgerinakis, E. Newman, F. Bremond, C. Crispim. D4.1 Audio & Visual Sensing v1. Dementia Ambient Care: .Multi-Sensing Monitoring of Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[8] C. Bettini, L. Pareschi, & D. Riboni. Efficient profile aggregation and policy evaluation in a middleware for adaptive mobile applications. *Pervasive and Mobile Computing*, *4*(5), 697-718 (2008).

[9] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A.Ranganathan, D. Riboni: A survey of context modelling and reasoning techniques. Pervasive and Mobile Computing 6(2): 161-180 (2010)

[10] B. Bishop, A. Kiryakov, D. Ognyanoff, I. Peikov, Z. Tashev, & R. Velkov, "OWLIM: A family of scalable semantic repositories," Semant. web, vol. 2, no. 1, pp. 33–42, Jan. 2011.

[11] B. Bishop, A. Kiryakov, Z. Tashev, M. Damova, K. Simov, "OWLIM Reasoning over FactForge", Proceedings of OWL Reasoner Evaluation Workshop (ORE 2012), Collocated with IJCAR 2012, Manchester, UK (2012)

[12] V. Buso, R. Mégret, G. Bourmaud, K. Avgerinakis, A. Briasouli, E. Newman, B. Kikhia, F. Bremond, P. Bilinsky, B. Fosty, C. Crispim Junior, A.T. Nghiem. D4.2 Activities Monitoring & Lifelogging v1, Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[13] Y. Cao, L. Tao, & G. Xu. An event-driven context model in elderly health monitoring, Proc. Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, pp. 120-124 (2009).

[14] D. P. Chau, F. Bremond, & M. Thonnat. A multi-feature tracking algorithm enabling adaptation to context variations. In Proceedings of International Conference on Imaging for Crime Detection and Prevention, (2011).

[15] L. Chen, J. Hoey, C. Nugent, D. Cook, and Z. Yu, "Sensorbased activity recognition," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 42, no. 6, pp. 790 –808, (Nov. 2012).

[16] L. Chen, C. Nugent, M. Mulvenna, D. Finlay, X. Hong, & M. Poland. Using event calculus for behaviour reasoning and assistance in a smart home. In *Smart homes and health telematics* (pp. 81-89). Springer Berlin Heidelberg (2008).

[17] C. Crispim, V. Joumier, F. Bremond. A Multi-Sensor Approach for Activity Recognition in Older Patients. In the proceedings of the Second International Conference on Ambient Computing, Applications, Services and Technologies. , AMBIENT 2012. Barcelona, September 23-28, (2012).

[18] C. Crispim, J. McHugh, S. Sävenstedt. D8.2 Evaluation Protocols. Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[19] S. Dasiopoulou, V. Efstathiou, G. Meditskos. D5.1 Semantic Knowledge Structures and Representation, Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[20] A. P. Dempster. Generalization of Bayesian inference. J.Royal Statist. Soc., 30:205–247, (1968).

[21] A. Fleury, N. Noury, & M. Vacher. Introducing knowledge in the process of supervised classification of activities of Daily Living in Health Smart Homes. In Proceedings of 12th IEEE International Conference on e-Health Networking Applications and Services, pages 322-329, (July 2010).

[22] M.F. Folstein, L.M. Robins, and J.E. Helzer; The mini-mental state examination. Arch Gen. Psychiatry, 40(7):812, 1983.

[23] L. Gao, A.K. Bourke, & J. Nelson. A system for activity recognition using multi-sensor fusion. In Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 7869-7872, (August 2011).

[24] W. R. van Hage, V. Malaise, R. H. Segers, L. Hollink, & G. Schreiber. Design and use of the simple event model (sem).Web Semantics: Science, Services and Agents on the World Wide Web, 9(2), (2011).

[25] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, & M. Dean. SWRL: A semantic web rule language combining OWL and RuleML. *W3C Member submission*, *21*, 79 (2004).

[26] K. Irving, D. O'Mahoney, V. Joumier, S. Sävenstedt, D2.2 Functional Requirements and Clinical Scenarios v1, Dementia Ambient Care: Multi-Sensing Monitoring for Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[27] H. Izadinial, & M. Shah. Recognizing complex events using large margin joint low-level event model. In Proceedings of the 12th European conference on Computer Vision, 4:430-444, Firenze, Italy, (October 2012).

[28] M. Koutraki, V. Efthymiou, & G. Antoniou. S-CRETA: Smart Classroom Real-Time Assistance. *Ambient Intelligence-Software and Applications*. Springer Berlin Heidelberg, 2012. 67-74.

[29] R. Kowalski, & M. Sergot. A logic-based calculus of events. In *Foundations of knowledge base management* (pp. 23-55). Springer Berlin Heidelberg (1989).

[30] H. Knublauch, J. A. Hendler, & K. Idehen, "SPIN - overview and motivation," World Wide Web Consortium, W3C Member Submission, Feb. 2011.

[31] G. Lavee, E. Rivlin, & M. Rudzsky. Understanding Video Events: A Survey of Methods for Automatic Interpretation of Semantic Occurrences in Video. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 39:5, pp.489-504, (Sept. 2009).

[32] Q.V. Le, W.Y Zou, S.Y. Yeung, & A.Y. Ng. Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 3361-3368, (June 2011).

[33] Q. MA, B. Fosty, C. Crispim-Junior, & F. Bremond. Fusion Framework for Video Event Recognition. In the proceedings of the 10th IASTED International Conference on Signal Processing, Pattern Recognition and Applications , SPPRA 2013, Innsbruck. (12-14 February, 2013).

[34] G. Meditskos, S. Dasiopoulou, V. Efstathiou, I. Kompatsiaris. SP-ACT: A Hybrid Framework for Complex Activity Recognition Combining OWL and SPARQL Rules, 10th IEEE Workshop on Context Modeling and Reasoning, San Diego, California, March 18-22, 2013

[35] G. Meditskos, S. Dasiopoulou, V. Efstathiou, I. Kompatsiaris. Ontology Patterns for Complex Activity Modelling, 7th International Web Rule Symposium: Research Based and Industry Focused (RuleML 2013), Seattle Metropolitan Area, USA, July 11-13, 2013

[36] H. Medjahed, D. Istrate, J. Boudy, J.-L. Baldinger, & B. Dorizzi. A pervasive multi-sensor data fusion for smart home healthcare monitoring. In Proceedings of IEEE International Conference on Fuzzy Systems, pages 1466-1473, (June 2011).

[37] B. Motik, B. Cuenca Grau, & U. Sattler. "Structured objects in OWL: representation and reasoning," in Proceedings of the 17th international conference on World Wide Web (WWW '08). New York, NY, USA: ACM, pp. 555–564, (2008).

[38] B. Motik, R. Shearer, & I. Horrocks. Hypertableau reasoning for description logics. J. Artif. Int. Res., 36(1):165–228, (September 2009).

[39] A. T. Nghiem, F. Bremond, & M. Thonnat. Controlling background subtraction algorithms for robust object detection. In Proceedings of 3rd International Conference on Imaging for Crime Detection and Prevention, pages 1-6, London, UK, (December 2009).

[40] G. Okeyo, L. Chen, H. Wang, & R. Sterritt. A hybrid ontological and temporal approach for composite activity modelling. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on* (pp. 1763-1770). IEEE (June 2012).

[41] T. Patkos, I. Chrysakis, A. Bikakis, D. Plexousakis, & G. Antoniou, G. A reasoning framework for ambient intelligence. In *Artificial Intelligence: Theories, Models and Applications* (pp. 213-222). Springer Berlin Heidelberg (2010).

[42] J. Perez, M. Arenas, and C. Gutierrez, "Semantics and complexity of SPARQL," ACM Trans. Database Syst., vol. 34, no. 3, pp. 16:1–16:45, Sep. 2009

[43] D. Riboni, & C. Bettini. COSAR: hybrid reasoning for context-aware activity recognition. *Personal and Ubiquitous Computing*, *15*(3), 271-289 (2011).

[44] D. Riboni., L. Pareschi, L. Radaelli., & C. Bettini, (2011, March). Is ontology-based activity recognition really effective?. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on* (pp. 427-431). IEEE, (March 2011).

[45] R. Romdhane, F. Bremond, & M. Thonnat. Complex Event Recognition with Uncertainty Handling. In Proceedings of the 7th IEEE International Conference on Advanced Video and Signal-Based Surveillance, Boston, (August 2010).

[46] L. Rong, & L. Ming. Recognizing Human Activities Based on Multi-Sensors Fusion. In Proceedings of 4th International Conference on Bioinformatics and Biomedical Engineering, pages 1-4, (June 2010).

[47] E. Schuijers, E. Newman, M. Contat, D3.1 Physiological & Lifestyle Monitoring v1. Dementia Ambient Care: Multi-Sensing Monitoring of Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[48] G. Shafer. A mathematical theory of evidence. Princeton University Press, Princeton, NJ, (1976).

[49] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, & Y. Katz. Pellet: A practical OWL-DL reasoner," Web Semantics: Science, Services and Agents on the World Wide Web, vol. 5, no. 2, pp. 51–53, 2011.

[50] D. Stampouli, J. Renoux, , P. Giroux, C. Burak Akgül, J. Bernsen, J., G. Meditskos, K. Akman, D7.1 System Specifications and Architecture v1. Dementia Ambient Care: .Multi-Sensing Monitoring of Intelligent Remote Management and Decision Support, Dem@Care – FP7 288199.

[51] G. Stevenson, J. Ye, & S. Dobson. On the impact of the temporal features of sensed data on the development of pervasive systems. In Proceedings of the International Workshop on Programming Methods for Mobile and Pervasive Systems (PMMPS' 10), Helsinki, Finland, (May 2010).

[52] M. Y. Vardi. "Why is modal logic so robustly decidable?" in Proceedings of a DIMACS Workshop on Descriptive Complexity and Finite Models, ser. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. American Mathematical Society, pp. 149–184, (January 1996).

[53] T. Vu, F. Brémond & M. Thonnat, Automatic Video Interpretation: A Novel Algorithm for Temporal Scenario Recognition. The Eighteenth International Joint

Conference on Artificial Intelligence (IJCAI'03), Acapulco, Mexico, (9-15 August 2003).

[54] H. Wang, A. Klaser, C. Schmid, & C. Liu. Action recognition by dense trajectories. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 3169-3176, (June 2011)

[55] S. Zaidenberg, B. Boulay, F. Bremond, & M. Thonnat. A generic framework for video understanding applied to group behavior recognition. In Proceedings of the IEEE International Conference on Advanced Video and Signal-based Surveillance, pages 136-142, (September 2012).

A Appendix

A.1. Activity Recognition – CAR Component

The list of activities that are currently supported by CAR recognition module refer to the lab setting where recorded data already exist and the algorithms have been tested. The list of activities to be recognised at home/nursing home settings will be consolidated after the pilots are set up, where actual home/nursing training data for the recognition algorithms will be available.

A.1.1 Lab Setting

Directed Activities

Contextual Zones

- zoneExerciseWalking
- zoneUseChair
- zoneStop
- zoneEntrancePath

Elementary Events/Activities

- Person_Inside_ZoneExerciseWalking
- Person_Inside_ZoneUseChair
- Person_Inside_ZoneStop
- Person_Inside_ZoneEntrancePath
- Person_standing
- Person_sitting
- Person_walking
- Person_walking_long
- change_posture_stand_to_sit
- change_posture_sit_to_stand
- Person_standing_long
- Person_sitting_long

Composite/Complex Events

- changeZone_FromExerciseWalkingToStop
- changeZone_FromStopToExerciseWalking
- changeZone_FromExerciseWalkingToUseChair

- changeZone_FromUseChairToExerciseWalking
- Start_WalkingTest
- WalkingTest_FirstAttempt
- WalkingTest_SecondAttempt

Semi-directed Activities

Contextual Objects

- chair
- readingTable
- tv
- plant
- library
- phoneTable
- officeTable
- kettle
- chairCloseReadingTable
- busLinesMap

Contextual Zones

- zoneUseTV
- zoneUseReadingTable
- zoneUseTeaCorner
- zoneUsechair
- zoneUsePlant
- zoneUsePhone
- zoneUsePharmacy
- zoneUseOfficeDesk

Elementary Events

- Person_Inside_Zone_TV
- Person_Inside_Zone_UseReadingTable
- Person_Inside_Zone_UseTeaCorner
- Person_Inside_Zone_UseChair

- Person_Inside_Zone_UsePhone
- Person_Inside_Zone_Pharmacy
- Person_Inside_Zone_OfficeDesk
- Person_Inside_Zone_Plant
- Person_Inside_Zone_BusLinesMap
- Person_closeTo_Phone
- Person_closeTo_TV
- Person_closeTo_OfficeDesk
- Person_closeTo_Kettle
- Person_closeTo_PharmacyBasket
- Person_closeTo_Plant
- Person_closeTo_BusLinesMap
- Person_standing
- Person_sitting
- Person_bending

Composite/Complex Events

- Person_using_Kettle
- Person_reading_inChairReadingTable
- Person_reading
- Person_using_Phone
- Person_watching_TV
- Person_using_OfficeDesk
- Person_using_PharmacyBasket
- Person_using_Plant
- Person_leaving_Room

A.2. Activity Recognition – SI Component

As with the list of activities recognised by CAR, only the list of activities that are currently supported by SI in the lab setting are available at this stage. The list of activities/situations and problems to be recognised by SI at home/nursing home settings will be consolidated after the pilots are set up and relative home/nursing home inputs coming from CAR as well as WP3/WP4 are available.

A.2.1 Lab Setting

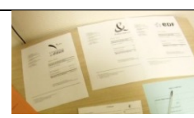
Semi-directed Activities

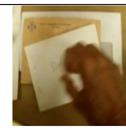
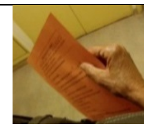
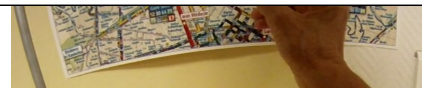
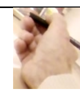
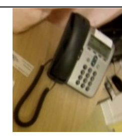
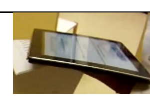
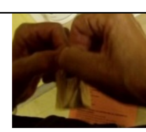
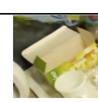
- HavingPhoneInteraction
- EstablishingAccountBalance
- FindingBusLineOnMap
- PayingBill
- PreparingDrugBox
- PreparingTea
- ReadingArticle
- TurningOnTV
- WateringPlant

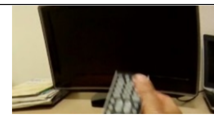
A.3. Object/Location Detection

A.3.1 Lab Setting

Objects

Object category	Example
Basket	
Bills	
Cards	
Checks	
Accounts	

Envelopes	
Instruction	
Kettle	
Map	
MedicalInstruction	
Pen	
Phone	
Pillbox	
PlasticGlass	
Remote	
Tablet	
Teabag	
Teabox	

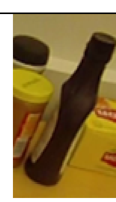
Tv	
WateringCan	

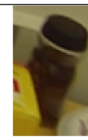
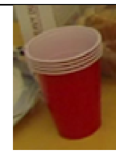
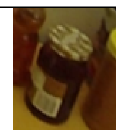
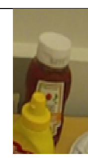
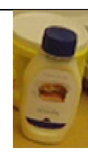
Locations

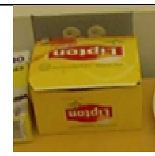
- phone_place
- tea_place
- tv_place
- table_place
- medication_place

A.3.2 Home/Nursing Home setting

The same comment with the list of activities recognised by CAR and SI for home/nursing home settings applies to the list of objects to be recognised at home/nursing home. Below there is an indicative list of object recognition examples in the home setting.

Object category	Example
Bread	
Cheese	
Chocolate	

Coffee	
Cup	
Honey	
Hotdog	
Jam	
Ketchup	
Mayonnaise	
Mustard	
PeanutButter	
Spoon	

Sugar	
Tea	
Water