Handbook of Research on Ambient Intelligence and Smart Environments:

Trends and Perspective

Nak-Young Chong
Japan Advanced Institute of Science and Technology, Japan

Fulvio Mastrogiovanni
University of Genova, Italy

Information Science
REFERENCE

Senior Editorial Director: Kristin Klinger
Director of Book Publications: Julia Mosemann
Editorial Director: Lindsay Johnston
Acquisitions Editor: Erika Carter
Development Editor: Christine Bufton
Production Editor: Sean Woznicki
Typesetters: Keith Glazewski, Natalie Pronio, Jennifer Romanchak, Milan Vracarich, Jr.
Print Coordinator: Jamie Snavely
Cover Design: Nick Newcomer

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data
Handbook of research on ambient intelligence and smart environments: trends and perspective / Nak-Young Chong and Fulvio Mastrogiovanni, editors.
p. cm.
Includes bibliographical references and index.
Summary: "This book covers the cutting-edge aspects of AMI applications, specifically those involving the effective design, realization, and implementation of a comprehensive ambient intelligence in smart environments"-- Provided by publisher.
ISBN 978-1-61692-857-5 (hardcover) -- ISBN 978-1-61692-858-2 (ebook) 1. Ambient intelligence. 2. Telematics. I. Chong, Nak-Young, 1965- II. Mastrogiovanni, Fulvio, 1977-
QA76.9.A48H36 2011
004.01'9--dc22
2010041637

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 22
Recognising Human Behaviour in a Spatio-Temporal Context

Hans W. Guesgen
Massey University, New Zealand

Stephen Marsland
Massey University, New Zealand

ABSTRACT

Identifying human behaviours in smart homes from sensor observations is an important research problem. The addition of contextual information about environmental circumstances and prior activities, as well as spatial and temporal data, can assist in both recognising particular behaviours and detecting abnormalities in these behaviours. In this chapter, we describe a novel method of representing this data and discuss a wide variety of possible implementation strategies.

INTRODUCTION

Over recent years, research in ambient intelligence and smart homes has experienced a significant push, with applications in a wide range of areas. One field where smart homes are of particular interest, and therefore reoccur frequently as a research topic, is those homes that aim to improve the quality of life of the inhabitant in one way or another. This could be by providing a safer environment through security measures such as intruder alerts, adjusting environmental conditions such as temperature to the inhabitants' activities, or supporting the inhabitants in their daily activities. In particular, the last of these has gained a significant amount of attention. This does not come as a surprise, since more and more individuals wish to live independently in their own homes into old age, but are not always able to do so, due to diminishing physical or mental capabilities caused by the aging process or diseases like Alzheimer's. An ambient intelligence system can monitor individuals in a non-obtrusive way in their own homes, and provide assistance whenever necessary. This assistance can range from assurance (making sure that the individual is safe and performing routine

DOI: 10.4018/978-1-61692-857-5.ch022

activities) through support (helping individuals to compensate for impairment by providing reminders) to assessment (determining the physical or cognitive status of the inhabitant and alerting a caregiver if necessary).

To offer effective assistance, the ambient intelligence has to analyse the activities performed by the individuals in the home, and to infer the behaviours from those activities, as they are observed through the sensors. We use the term 'behaviour' to mean a particular set of activities performed with a particular aim in mind, such as doing the laundry, cooking dinner, making breakfast, or watching TV. These are sometimes known as Activities of Daily Living (ADLs). One of the main challenges in behaviour recognition is that the exact activities are not directly observed: the only information provided are the sensor observations, which could be that the kitchen light is on, the oven is turned on and the burner is on; the inference that therefore somebody is cooking is left to the intelligent part of the system. This is particularly true where cameras are not used, something that we tend to assume since cameras are intrusive and the images can be difficult to analyse.

AN AI PERSPECTIVE ON THE SMART HOME

The smart home can be separated into the sensory system and the ambient intelligence that works to interpret the sensor observations. In this chapter, we do not explicitly consider what types of sensor are (or could be) available in the home, but assume that the sensory stream is available in the form of a sequence of 'tokens', i.e., there has been some preprocessing of the sensor readings (and possibly the fusion of different sensor data) into a sequence of observations that can be used by some form of ambient intelligence system. For our purposes, we consider that the task of the smart home is to segment this token stream into different behaviours, and to identify whether or not this behaviour is typical of the user and, if not, whether it is sufficiently abnormal to warrant any action (such as calling a carer, or interacting with the inhabitant, for example in the form of reminders: 'did you remember to turn the gas off?').

Our interpretation of the smart home problem from the point of view of artificial intelligence is that it requires solutions to at least some of the following problems:

- **behaviour recognition and segmentation** Assuming that data from the sensors is represented by tokens, the home needs to process that data into individual behaviours and classify them. The challenges in this are that the data is potentially noisy, certainly provide only a partial snapshot of the person's activities, and may well be difficult to segment with complete accuracy. There is lots of work going on in this area around the world, with most of the more successful approaches being based on probabilistic methods such as Hidden Markov Models, e.g., (Nguyen *et al*., 2005; Duong *et al*., 2005).
- **novelty detection** Once the inhabitant's behaviour has been recognised, the system needs to decide whether or not the behaviour is novel (i.e., the inhabitant is doing something that they have not done before, so that the classification failed) or if they are doing it in an odd way (either the activities that they are performing are not quite correct, or the time or place are unusual). We believe that it is not generally possible to do this by standard classification, and so novelty detection systems are the most promising approach (Rivera-Illingworth, Callaghan and Hagras, 2007; Jakkula and Cook, 2008; Marsland, 2003). For a smart home that performs monitoring, this is the part of the system that is likely to produce

warnings for carers, or some form of interaction with the inhabitant.

- **the ability to utilise temporal, spatial, and contextual data** This chapter is primarily focussed in this area. Using this data can improve the results of both behaviour recognition and novelty detection; indeed, it may be essential for detecting novelty in many cases.
- **ontologies and background knowledge** While a purely evidence-based system would obviously be nice, there are things that are not possible using such a system, such as dealing with new appliances. For example, an ontology of household appliances could be used so that when a new kitchen appliance is bought, the system can theorise tentatively about how the inhabitant could expect to use it. Equally, background knowledge about what is expected may speed up the learning required in the system. While our philosophical preference is for a system that does not require prior knowledge to be installed, this may not be practical in many cases.
- **lifelong learning** Things change within houses and in people's behaviours all the time. The system needs to be able to learn about these things so that it does not annoy the inhabitants and carers by providing multiple false alarms. In addition, not all behaviours may appear during a training phase, and so the system has to be able to receive feedback and update itself.
- **the ability to communicate with other smart homes** One way to avoid the problem of not all behaviours being seen during training could be to use data that other smart homes have seen, and update itself when the inhabitant performs those behaviours. Additionally, if one house detects a potential danger, it can warn other houses that they may see the same thing.

A different way to see some of the same features is to use the classification of abnormal behaviour suggested by Russell Dewey (2009):

- **statistical abnormality** Based on deviation from what statistically is considered usual behaviour.
- **violation of socially-accepted standards** Based on judgment made by religious, cultural, or social groups.
- **theoretical approaches** Based on a theory of personality development.
- **subjective abnormality** Based on a personal assessment of normality.
- **biological injury** Based on the impact that abnormal biological processes such as disease or injury have on the individual.

The first of these categories is the typical one that machine learning systems can deal with. It takes only statistical evidence into consideration and refrains from any judgments made by the individual living in the house or people related to that individual. The idea is that an individual starts off with behaviours that are considered normal (regardless of how strange they might appear when imposing cultural, social, or religious values). These behaviours are used to initialise the ambient intelligence without any prior knowledge of what to expect, and then learns to identify the particular individual's behaviours, for example through machine learning techniques or some logic-based approach. Once trained, the ambient intelligence is used to detect behaviours that deviate from normal behaviours, and thus are considered abnormal.

Although, in principle, this seems to be a straightforward process, there is a major problem with this approach. Behaviours might change over time, but only some of the changes indicate abnormal behaviour. To be able to distinguish between the changes that are normal and those that are abnormal, the behaviours have to be interpreted in context. In other words, the ambient intelligence

not only needs observational information, but also has to be aware of the context in which the observations were made.

The second of the five types of abnormality – and to some extend the third – could be approached using an ontology-based method that had codified knowledge of societal norms, and applied them, while the fifth type could be identified through the integration of knowledge from many different smart homes, so that behaviours that caused accidents, or disease diagnoses that changed behaviours, could be identified.

In this chapter, we focus on how to create a mapping between sensor streams and human behaviours, the problems that this task brings with it, and possible ways to overcome these problems. We will do this by considering the properties of the sensor stream with regard to space, time, and context.

USING CONTEXTUAL DATA IN A SMART HOME

Considering just the token level, contextual information can take two different forms: the input of additional sensors that provide background information either about the environment (e.g., temperature, humidity level) or the user (e.g., emotion detection), and previous outputs from the behaviour recognition system, so that the system has some idea of what the person has been doing recently.

Before going further it is worth considering two different ways in which all of this data can be used. The first is as part of the behaviour recognition, while the second is as part of the 'normality' detection unit. To see the difference, consider the following two statements:

1. it is morning and the toaster is being used, so the inhabitant is probably making breakfast
2. the inhabitant is making breakfast and it is morning, so everything is normal

The first uses the temporal data, together with the sensor data, to decide on a behaviour, while the second uses the behaviour and the temporal data to ensure that the behaviour should not cause concern. Both of these interpretations have implicitly used spatial data, since the sensors that identify use of the toaster are presumably in the kitchen. So is one of these two ways to think about the problem more useful, or are both valid? It seems to us that there is no reason to limit the use of any potentially useful data to just one task. Note, however, that the second one allows for the possibility of negation: if it is morning and the person does not make breakfast, then this could be identified as abnormal. This may well be particularly important for a smart home, where things not done could be important either for safety (e.g., gas not turned off) or as a sign of increasing forgetfulness. Detecting that things are not done can be particularly challenging for probabilistic machine learning methods, and this could be one place where logic-based intelligence is required.

The other challenge is that much of the data is not useful for a particular decision. Statements like 'it is 22.5 degrees and the toaster is being used, so they are probably making breakfast' are unlikely to be correct very often. Working out which correlations are useful is a tricky data mining problem, and one that could require a lot of data. If we need to collect a very large dataset before the smart home can be used, then people may not be prepared to use the system. For these reasons, we prefer to separate out the behaviour recognition system from the environmental data, although ideally the system should be able to take advantage of that data if it is available.

THE DIFFERENCE BETWEEN SPACE AND TIME FOR SMART HOMES

Considering the sequence of sensor-based tokens, the first thing to realise is that spatial and temporal properties of the data are to some degree different:

sensors are physically located in one particular location, and are (in the main) unlikely to move. There are exceptions to this, from furniture rearrangement to sensors on objects such as laptop computers that could be used in different rooms, but in general the location of sensors is fixed, and so if the fridge sensor emits the token that says that the fridge has been opened, then it seems reasonable to assume that there is somebody in the kitchen. For this reason, in general there is no need to tag tokens with the spatial location of the sensor. However, the same thing is obviously not true about time – the same sequence of tokens at two different times of day could well be labelled differently, from trivial differences (cooking in the morning is breakfast, while cooking in the evening is dinner) to more important ones: cooking in the middle of the night could well be a sign of confusion. For this reason, it is usual to consider the sequence of tokens as being tagged with time stamps, but not with spatial locations.

There are a couple of corollaries to this difference between space and time: token activations can occur simultaneously in time, and tokens from different sensors can be interleaved, but if there is only one person in the house, then sensors in two different rooms, or even two parts of the same room, cannot be simultaneously triggered (assuming that the areas covered by the sensors are not overlapping). Another is that we have different resolutions of space and time: the pre-defined set of sensors may well miss out several locations, so that we have no idea what the person is doing at various locations – if the fridge door sensor goes on, and then an hour later goes off, this does not necessarily mean that the person stood at the fridge for an hour, more likely they forgot to close the door properly. However, if they did not trigger any other sensors in that hour, then the system will not know what they were doing. Temporally, there is information in the fact that other sensors were not triggered, but not spatially.

This leads us to the problems of temporal and spatial persistence (Guesgen and Hertzberg, 1996), which can roughly be characterised as follows. Assuming that we do not have complete knowledge about what is and is not true at every individual time point, we need to identify those facts whose truth at some time point will persist until some later time point. For example, if the fridge sensor emits the token that says that the fridge has been opened and we then assume that there is somebody in the kitchen, for how long would we maintain this assumption? If we get a token from the living room sensor after a while, we know that the assumption is no longer true. On the other hand we cannot guarantee the persistence of the assumption if no tokens are received, since it may be that the person left the kitchen, but did not trigger any other sensors.

In addition to the temporal persistence problem, we have to deal with the spatial persistence problem. Assuming that we do not have complete knowledge about what is and is not true at every location in the home, we wish to identify those facts known to be true at some location that persist to other locations of the home. For example, if we know that a particular spot on the sofa has enough light to read a book, would we be able to make this assumption for all locations on the sofa or even the whole living room? At which points in space do we drop the assumption of persistence? In answering this question, we face the problem that space does not have an intrinsic direction, as time does. There are many possible trajectories that we can take to move away from a particular location. To get some order into this large set of possible movements through space, we usually employ concepts such as topology, orientation, and distance; something we discuss briefly later on.

Returning to temporal information, there are a number of ways to specify a particular point in time: either absolute (15:34:34 GMT on 24/06/09) or relative (15 minutes after I get home from work, or the first Sunday after the first full moon after the vernal equinox). Implicit in those three different descriptions is another interesting feature,

which is the 'granularity', i.e., the resolution at which the time is represented. This is very different in the three cases, and choosing the appropriate resolution is certainly non-trivial. This becomes even more important when you consider how behaviours could be classified in time. It is common for working people to have pretty much two modes of behaviour: week days and the weekend. Within these two times, the pattern is often fairly constant, but there is a marked difference between the two. Things that would be normal on a weekend (lying in bed reading the newspaper, for example) would be a definite sign of something being wrong on a weekday. A smart home that treated every day the same would annoy people very quickly!

When looking at spatial information, the issue becomes even more complex in that it is not clear which space we are referring to. One could argue that normal three-dimensional space is natural to describe movement within the smart home, but this space does not directly correspond to the space that the sensors cover. For example, the sensor at the entry door has a particular coverage area, and if this sensor is activated, then we can assume that a person has entered this area. If immediately after that another sensor is activated, we might conclude that the person has moved from the entry-door area to the area covered by the other sensor, but only if the areas do not overlap. In summary, the space we are really looking at is a space of more or less abstract three-dimensional regions that are related to each other in some way. The question is how the regions are related to each other.

Three types of spatial information are commonly distinguished: topological information, orientation information, and distance information. Children acquire a sense for these types of spatial information in that order (Piaget and Inhelder, 1948), which might suggest that topological information is the most essential one among the three. It is indeed the case that by using this type of information, we can draw conclusions that can significantly improve the behaviour recognition in smart homes. For example, suppose that a sensor is activated in the kitchen and then another sensor is activated in the living room. If we know that the corridor is between the kitchen and the living room and that the sensor in the corridor has not been activated, then we can conclude that the sensors are activated by different people and therefore belong to different activities. This assumes that the topological relations between the areas are such that they do not allow activation of the sensors in the kitchen and the living room by one person without activating the sensor in the corridor. Note that this does not necessarily mean that the areas covered by the three sensors have to be pairwise disjoint; they can partially overlap. We provide more details on topological relations in a later section of this chapter.

Orientation information is the second type of spatial information that can be used to put sensor data into context. It can be viewed as a ternary relation of the primary object, the reference object, and a frame of reference. For example, the TV (primary object) might be in front of the chair (reference object), where the edge of the seat of the chair is used as frame of reference. The role of the frame of reference here is to define the front side of the object; without it, a given orientation relation between two objects might have several interpretations. Three different types of reference frame (Levinson, 2003) are usually distinguished:

- **intrinsic** The orientation is given by some inherent properties of the reference object.
- **relative** The orientation is imposed by the point of view from which the reference object is seen.
- **absolute** The orientation is given by fixed directions such as the cardinal directions (north, south, east, west) or the direction provided by gravity.

For example, a statement like 'the banana peel is in front of the person' can either put the banana peel close to the person's toes (intrinsic reference frame) or at the person's side (relative reference

frame with the observer watching a person passing by). In an absolute reference frame, we would rather use a statement like 'the banana peel is north of the person'.

In principle, we can map one reference frame into another, and therefore we are not concerned here about which particular reference frame has been chosen to represent orientation in the ambient intelligence of our smart home. What is important to us is the fact that orientation can add an additional quality to the reasoning process. For example, if the chair is not facing the TV and the sensor in the chair is activated, then it is unlikely that the person is watching TV.

The third type of spatial information is distance information. There are numerous studies of how humans make subjective judgements regarding distances, in particular how humans determine how close objects are to each other. According to (Gahegan, 1995), the human perception of closeness or proximity is influenced by the following:

- In the absence of other objects, humans reason about proximity in a geometric fashion. Furthermore, the relationship between distance and proximity can be approximated by a simple linear relationship.
- When other objects of the same type are introduced, proximity is judged in part by relative distance, i.e., the distance between a primary object and a reference object.
- Distance is affected by the size of the area being considered, i.e., the frame of reference.

Although it is tempting to use Euclidian distance to determine proximity, this is not always the most intuitive way. Humans often have a qualitative rather than a quantitative perception of distance (Guesgen, 2002). For example, if the person sits down in front of the fire and we know that there is a 'safe' distance between the two, then we can conclude that there is no danger of the person getting burned.

CONTEXT AWARENESS

There are a great many forms that context awareness can take within a smart home. The way that it is most commonly understood is that individual activities of people within the house are not interpreted in isolation, but that additional data concerning, for example, the temperature and what other activities they have been engaging in are included. This can be useful since behaviours that appear to be perfectly normal in isolation can look odd with further inspection: sitting by a log fire is perfectly normal in winter, but when the temperature is 30 degrees (Celcius) in the house, it might be an indication of mental instability. Likewise, cooking dinner is a perfectly normal activity, unless it is the third one that the person has cooked today. Considered from this point of view, context awareness is concerned with adding information to assist in either the classification of the activity, or the analysis of whether or not this behaviour is 'normal'.

However, there are many other ways in which context awareness can be included in a smart home, which go beyond the individual home. The ability of the house to communicate with other houses can be very useful. From the public health point-of-view the detection of pandemics and other spreads of illness could be greatly improved by each smart home reporting on the health status of its inhabitants based on its normal monitoring of their behaviour. Considering the same thing in reverse, once houses were informed that there was disease prevalent within their locality, they could warn the inhabitants to be vigilant, or to stock up on food, etc. Additionally, there is a large pool of data that can be accessed in this way to enable houses to learn about behaviours that could be demonstrated by their own inhabitants in the future.

There is another aspect of context awareness, which is that individual behaviours can act to prime (make more likely) or inhibit others. If the house inhabitant has just had breakfast, then behaviours

such as leaving the house (to go to work) or have another cup of coffee (at the weekend) may be quite likely, whereas going to bed is not. By increasing or decreasing the probability of other behaviours being seen after the current one based on experience, the system could improve the accuracy of the behaviour recognition. The length of time for which the priming remains active could also be tuned, so that some behaviours have long-term effects, while others are rather shorter. There are a wide variety of mechanisms that could be used to implement this kind of behaviour, from weights to a Hierarchical Hidden Markov Model (Fine, Singer and Tishby, 1998). A schematic of this idea is shown in Figure 1.

A SYSTEM USING SPATIAL, TEMPORAL, AND CONTEXTUAL DATA FOR BEHAVIOUR RECOGNITION

Figure 2 shows a possible way to visualise our proposed system. Tokens are used to identify possible behaviours using some system, such as Hidden Markov Models. The decision about which behaviour is currently seen may be made using just the token stream, as has been considered in a variety of papers in the literature, e.g., (Fine, Singer and Tishby, 1998). However, if there is other data available, then it could be used to weight the decision. In the figure we have separated the three groupings that we have been discussing, considering different representations for spatial, temporal, and contextual information.

Figure 1. The use of priming between behaviours can increase or decrease the probability of a particular behaviour being recognised next, based on the current behaviour

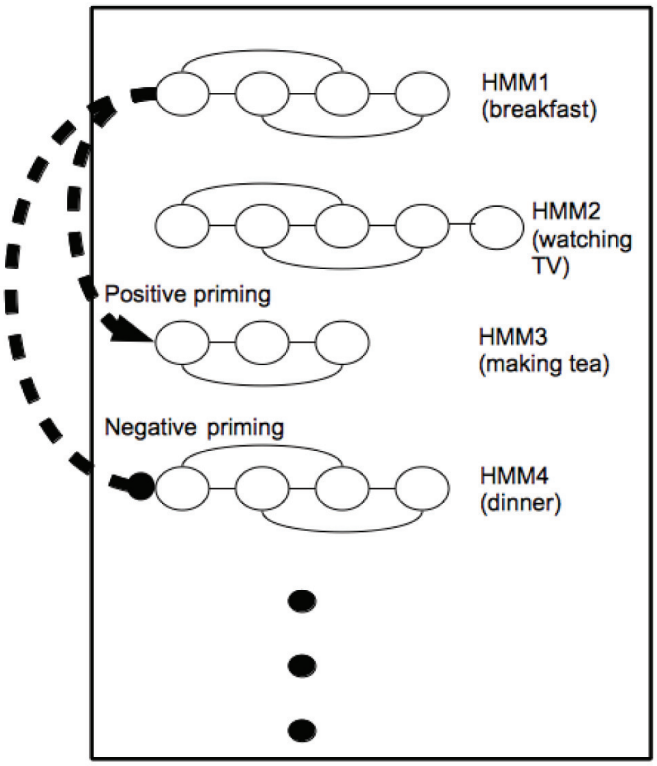
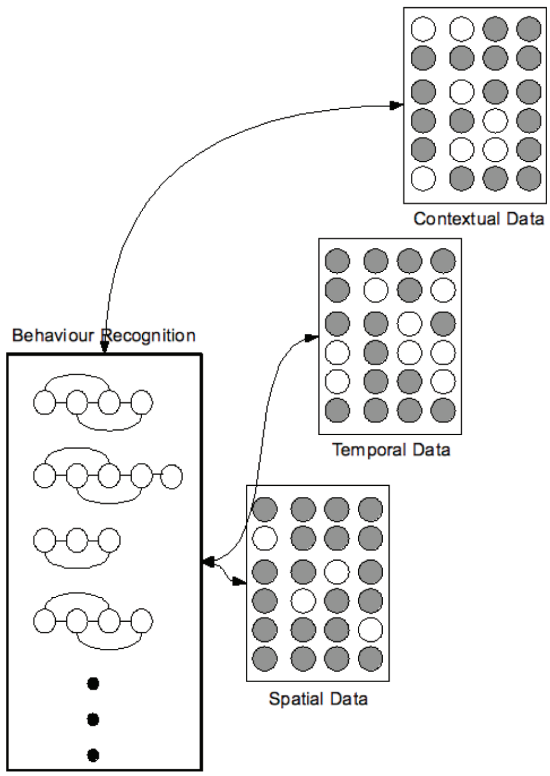


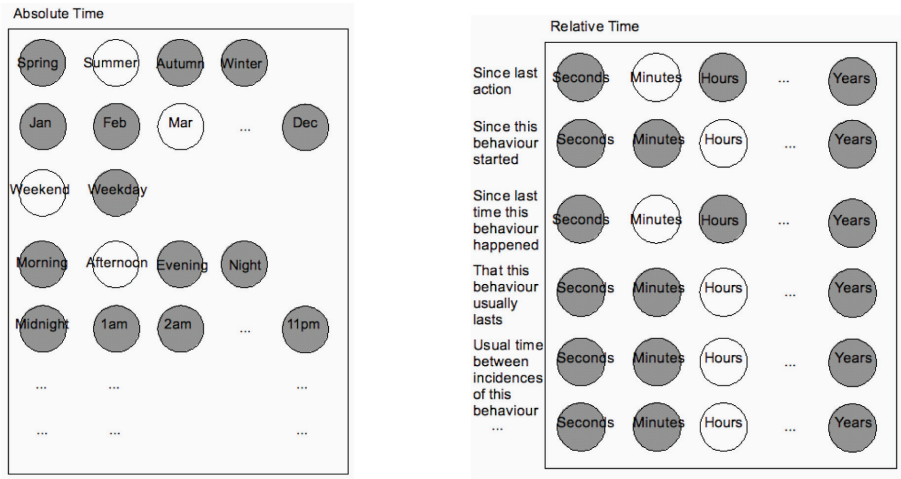
Figure 2. Our conceptual system uses different map layers to represent spatial, temporal, and contextual data



A basic version of the behaviour recognition system could be trained first, to identify behaviours based solely on their token pattern. The spatial, temporal, and contextual data concerning these tokens could then be used as input for each of these different modules. One problem that would have to be solved is to identify suitable granularities for the spatial and temporal representations. Our current thinking is that this should be possible to do by adding many different granularities into the map and allowing the ambient intelligence to work out for itself which are useful. Figure 3 shows two different representations of time, considering it as an absolute measure (left) and as a relative measure (right). Note that there is considerable overlap between different granularities: for example, seasons can be mapped to months without too much difficulty.

If we consider the absolute time map shown on the left of Figure 3 as representing the time that an activity begins (or at least, when the system first detects it), then times at machine precision are unlikely to be useful, since humans do not work at this scale. Each behaviour that occurs will cause a pattern of activation in the various

Figure 3. Two temporal maps based on absolute time (left) and relative time (right). The concepts in white are those that identify the current time, so it is a weekend afternoon in March, which is (Southern hemisphere) summer



'nodes', and as each behaviour is seen repeatedly, so the commonalities of the pattern should be identified and correlated with the particular behaviour. A number of different ways in which this could be done are discussed later.

Similarly, the relative time map should produce patterns of how long behaviours take, and how frequently they occur. This can include historical information, such as the normal length of time that this behaviour takes, and the common time between incidences of this behaviour. As more than one of the timescale options can be switched on simultaneously (to represent, for example, that incidences of a certain behaviour happen between daily and weekly) this is not too restrictive in the variation in time that it allows – this could be considered as something like a fuzzy logic system, where at midday, both the morning and afternoon representations would be on, not a crisp cut-off between the two.

One problem with our representation using this map is that only the time from the one previous behaviour is considered, and there may well be longer-range correlations. However, even using just one previous behaviour can lead to the system requiring massive amounts of data in order to identify patterns, something that will obviously get even worse with more inputs; the so-called curse of dimensionality. A possible way to avoid this problem is to include additional information and reasoning methods, for example by using logical representations of some parts of the data space. For now, we mention just one other thing that might be useful, which is that some actions are more likely than others to be used as a reference in this kind of relative timing – you are more likely to say that you had a cup of tea 10 minutes after you got in from work than that you had the tea 2 minutes after you fed the cat. If the actions that are used in practice for relative timing are possible to identify, then it could partially alleviate this problem.

POSSIBLE IMPLEMENTATION STRATEGIES

The system that we have shown in Figure 2 does not give any indication at all about implementation methods, and during the previous discussion we have highlighted some places where probabilistic methods would be useful, and others where they would not. There are certainly lots of possible methods, and we do not aim to describe a possible implementation that is sure to work; rather, we wish to produce an overview of some technologies that might be useful, particularly in light of the discussion of spatial and temporal data, and context.

We reiterate that the idea behind the different maps of the data is that what is interesting is 'patterns' of data in the maps, not the activation of individual nodes in each of the maps. If the latter were the case, then a set of pure logic rules of the form 'if it is winter and the temperature is less than 20 degrees then the inhabitant should turn the heater on' would be more suitable than our system. There are, of course, ways to learn these types of rules from data, using methods such as inductive logic programming, but that is not our focus here. However, there are plenty of places where symbolic artificial intelligence methods can give us useful data, and we discuss these next, before moving on to probabilistic models.

SPATIO-TEMPORAL REASONING

Spatial, temporal, and contextual data obey a variety of laws. A drop in temperature from 23 degrees to 22 degrees is not possible without an intermediate temperature of 22.5 degrees, meaning that if the temperature is 23 degrees at time t_1 and 22 degrees at time t_2, and if t_1 is before t_2, then we know that there is a time t_3 between t_1 and t_2 at which the temperature is 22.5 degrees. However, this only holds if the temperatures are observed in the same location. If 23 degrees are measured

in the living room at time t_1 and 22 degrees in the kitchen at time t_2, then it is possible that there is no time t_3 at which either the living room or the kitchen has a temperature 22.5 degree.

Researchers have studied universal laws like these for years, particularly in the context of temporal and spatial information. The result is a variety of formalisms that enable us to reason about space and time, which broadly can be put into two categories. One category contains the formalisms used by scientists such as physicists to establish the laws of time and space. We are not primarily interested in this category, but focus on the second category, which contains those formalisms that mimic human reasoning about space and time. The reason for favouring this category is obvious: if a layperson intends to communicate with an ambient intelligence, then it is easier to do so in a language that they are familiar with.

It has been argued in the field of artificial intelligence for many years that the layperson's everyday form of spatio-temporal reasoning is of a qualitative, rather than a quantitative nature; we are not usually interested in precise descriptions of space and time. Coarse and vague qualitative representations frequently suffice to deal with the problems that we want to solve. For example, to know that the meal is prepared in the kitchen (rather than knowing the exact coordinates for this activity) and at lunchtime (rather than 12:03:37) is often enough to decide whether the behaviour is normal or not. For that reason, we have included such vague description in our spatial and temporal layers.

Most qualitative approaches to spatial and temporal reasoning are relational approaches, i.e., approaches that reason about relations between objects (such as regions or time intervals). Of the temporal qualitative approaches, two are dominant and reappear frequently in different scenarios: Allen's temporal logic (Allen, 1983) and Vilain and Kautz's point algebra (1986). The first uses intervals to describe specific events, and a set of 13 basic relations between intervals to describe interdependencies between events. Allen introduced an algorithm that is based on a composition table to reason about the relations. For example, if we know that the kettle is filled with water before it is switched on and if we further know that the water boils while the kettle is switched on, the algorithm can infer that the kettle is filled with water before the water boils. The algorithm caters for uncertainty in the reasoning process by allowing sets of possible relations in addition to the basic relations. For example, it might be possible for the water to continue boiling after the kettle has been switched off (due to the heating element still being hot), in which case the algorithm would use both the 'during' and 'after' relations in the reasoning process.

Unlike Allen's temporal logic, the point algebra uses time points and three possible relations to describe interdependencies among them: $<$ (precedes), $=$ (same as), and $>$ (follows). These relations can be used in a similar way to interval relations. For example, the time point at which we have finished filling the kettle precedes the time point at which we switch on the kettle. In fact, Vilain and Kautz pointed out that many interval relations can be expressed as point relations by using the starting and finishing endpoints of the intervals.

Both Allen's interval logic and Vilain and Kautz's point algebra can be used for spatial reasoning by interpreting intervals as one-dimensional objects and time points as locations in space, respectively. A multi-dimensional representation can then be achieved by using a coordinate system and describing the spatial relations with respect to the coordinate axes. However, this approach suffers from the fact that reasoning can lead to counterintuitive results when the objects are not aligned with the coordinate axes. Researchers have therefore focused on purely topological approaches that do not depend on such an alignment.

We argue that the topological aspects of space are the most relevant ones for modelling human behaviour in an ambient intelligence. For example,

to determine what activity a person might be involved in, it is useful to know which room that person is in (although not necessarily the person's exact position in that room). Or, to determine the possible movements of a person, it is useful to know which rooms are connected to which other rooms. We will therefore focus on the topological aspects of space in our system, but at the same time point out that information about orientation and distance can also play a significant role in the reasoning process.

The topological approach that has gained most attention in the AI community is arguably the Region Connection Calculus (Randell, Cui and Cohn, 1992). The basis of this calculus is a reflexive and symmetric relation, called the connection relation. From this relation, additional relations can be derived, which include the eight jointly exhaustive and pairwise disjoint RCC8 relations. These relations can be used to express, for example, that two regions are disconnected from each other, that they are partially overlapping, or that one is a non-tangential proper part of the other (inside the other region without touching the boundary). Reasoning about space is achieved in the RCC framework in the same way as in Allen's temporal logic: by applying a composition table to pairs of relations.

It should be pointed out that the RCC theory is not the only approach to reasoning about topological relations. Egenhofer and Franzosa (1991) consider the intersections of boundaries and interiors of pairs of regions and derive a formalisation from that, called the 4-intersection calculus (and by adding the complement of the regions to the intersections, the 9-intersection calculus).

BEHAVIOUR RECOGNITION

Until now, we have not discussed the methods by which behaviour recognition and segmentation can be performed. This is because we are considering our temporal and spatial data as separate to the principal behaviour recognition problem, which is based solely on the sensor observations; the categorisation of temporal and spatial data that we have described provides a different way to recognise behaviours, or identify unusual occurrences of behaviours. We do not expect that the exact mechanism used for activity selection will be crucial to this system. Behaviour recognition is one of the more commonly investigated areas of smart home research (Duong *et al*., 2005; Tapia, Intille and Larson, 2004).

The approach that we favour for behaviour recognition and segmentation is a set of competing HMMs, each of which represents a different behaviour. The HMM (Rabiner, 1989) is the simplest dynamic graphical model, and has been used effectively on many problems. In comparison to many graphical models, it has the benefit that there are computationally tractable algorithms for both training and inference (Marslan, 2009). The challenges of using a set of HMMs for this problem are:

- **training a new HMM on some data** Where data is labelled with the behaviour, this is relatively simple, but without the labels, it is considerably more difficult. The LeZi algorithm (Das *et al*., 2002) is one approach to this.
- **segmenting the input stream into behaviours** This part of the problem is concerned with recognising behaviours based on the sensory observations, and then parsing the stream of data into individual behaviours. There are methods to do this using competition between HMMs alone (by monitoring the likelihood values of the HMMs on the data, see Chua, Marsland and Guesgen, 2009), but the system that we have described in this chapter can also assist in this segmentation problem by identifying the common spatio-temporal patterns in them.
- **dealing with behaviours that have not been previously seen** It is important that

sensory observations that do not match any current behaviour are identified, as it is one aspect of novelty detection to see that the home inhabitant is behaving oddly. However, the ability to then add new HMMs that represent additional behaviours provides a mechanism to allow lifelong learning. It may also enable the system to begin building from a *tabula rasa* state, i.e., the smart home starts with sensors being installed, and without any additional programming and knowledge specific to the particular installation.

Much of the other research on activity segmentation in smart homes has focused on more complicated variants of the HMM, such as the Hierarchical Hidden Markov Model (Nguyen *et al*., 2005), or Switching Hidden Semi-Markov Model (Duong *et al*., 2005). In both of these models, a top-level representation of behaviours (e.g., cooking or making coffee) is built up from a set of recognised activities that arise from the individual sensor values. A variant of these methods uses a three level Dynamic Bayesian Network (Liao *et al*., 2007). These models can be seen as adding complexity to the HMM in order to represent a complete model of behaviours arising from sensor activations. The problem is that more complex models require more data for training, and have higher computational cost.

POSSIBLE IMPLEMENTATIONS OF THE LAYERS

We now discuss a number of possible machine learning algorithms that may be useful in order to realise our idea of different spatial, temporal, and contextual layers for behaviour recognition and novelty detection. We will mention several different types of machine learning algorithm here; the interested reader is directed for more details to any of the standard texts in the field, such as (Marsland, 2009; Bishop, 2007).

One place where 'maps' like those we have drawn can be seen is in the Self-Organising Map (SOM) of Kohonen (1993), a neural network that has been very widely used in many different application areas. However, in the SOM, the pattern of activations in the network is created by the inputs through the set of weights, which is exactly the opposite of what we are planning here: it is the patterns of activations in the various maps that should assist with behaviour recognition and possibly with the discovery that this behaviour is abnormal or unusual in the time, place, or circumstances in which it is done.

It would be possible to use the entire set of map activations as inputs for a neural network, which would have outputs for the different behaviours that were identified by the system. This could either be done by considering particular patterns as exemplars in a Radial Basis Function Network, and then using a simple linear predictor (such as a Perceptron) to learn the output classes, or just by using the entire set of map activations as inputs to a standard neural network such as the Multi-Layer Perceptron. However, these methods rather remove the benefit of having the patterns in different maps for spatial and temporal data, something that we believe to be important. They also run the risk that massive amounts of data will be required before any useful output can be introduced, and that even when the data is such that no useful output can be generated, they will still produce an output – all of these are well-documented problems with neural networks.

One way of considering the problem that may be useful is to think of the maps as pictures of the spatio-temporal conditions when a particular behaviour occurred. In this case, it is possible to view the problem as one of taking the current set of spatial or temporal data, which may be different in some small way from those previously seen, and identifying a similar set of conditions that are more common. This can be seen as a clustering

problem, but it is also the kind of representation that Hopfield networks (Hopfield, 1982) can be used to find very effectively. A probabilistic variant on the Hopfield Network, the Boltzmann Machine (Ackley, Hinton and Sejnowski, 1985), may turn out to be useful for this.

The Boltzmann Machine uses a set of elements that are loosely modelled on neurons (like all artificial neural networks) in that they fire or do not fire based on their inputs. However, unlike most artificial neurons, in the Boltzmann Machine, the input is used to make a stochastic decision about whether or not to fire, with the probability of firing being a function of the strength of the inputs. A set of these neurons are created and connected together using weighted connections, and then the learning problem consists of identifying weights so that the target data is reproduced with high probability. There is a relatively simple, although computationally expensive, algorithm for solving this learning problem, which is described in [26]. While the original Boltzmann Machine used only binary nodes, there are a variety of ways to extend it to use non-binary inputs (Peterson and Anderson, 1987; Sejnowski, 1986).

The neurons of the Boltzmann machine can thus be the various layers of nodes in our system, and the learning problem will be to identify common patterns of activity across the different layers and link them to behaviours. One feature that may make this more useful (although at the cost of increased computational time) is the potential to add latent variables, also known as hidden units. These are nodes that do not have a direct meaning in terms of the layers of the system, but that integrate information from a variety of other nodes, so that non-linear structure in the maps can be learnt.

Another method that could be useful is the Conditional Random Field (CRF) (Lafferty, McCallum and Pereira, 2001; Sutton and McCallum, 2006), a probabilistic graphical model with some similarity to the Hidden Markov Model. Two alternative views of the CRF are as a Boltzmann Machine without hidden nodes, but with higher-order input correlations, and as a Hidden Markov Model with the transition probabilities being defined by probability distributions (conditioned on the states) rather than as constant values.

In this discussion we have presented a large number of different algorithms that can be used for implementation. Our aim is not to suggest which will best match the requirements for the problem, but to provoke research in this area. Certainly, it seems to us that a combination of both symbolic (useful for reasoning about time intervals and spatial relations) and probabilistic (to deal with noise and uncertainty) methods are required. We are currently investigating how these methods can be combined and which algorithms from those discussed above are most suited to the task.

SUMMARY

In this chapter we have considered the problem of interpreting the stream of sensory data coming into a smart home intelligence from its various sensors. We have done this by categorising the various properties of the sensor data. An output by a particular sensor can be interpreted as an assertion that the state of whatever that sensor monitors has changed. There are three types of additional data that we can tag to that observation, concerned with the location of the sensor, the time of the observation, and the environmental variables (themselves the output of other sensors) at that time and in that place. Further, since the sensory data arrives in a temporal stream, we also know activities that are occurring at around the same time, and what behaviours the system has judged those sensory observations to be part of.

Taking these things into account, we have suggested a method by which the various pieces of additional information about the sensory data can be used in conjunction with the data in order to more accurately classify the data into behaviours (or Activities of Daily Living) or, possibly more

importantly, to identify abnormal behaviours that could signify illness or otherwise require a carer to be alerted. The system that we propose is based on the representation of the three different forms of additional information in maps, with the pattern of information in the maps being used rather than the individual pieces of data. This enables the system to identify common patterns between different behaviours. We have suggested a variety of artificial intelligence and machine learning techniques that could be used in order to implement such a system.

The maps that represent spatial and temporal information are very different, since the resolution of temporal data can be as fine-grained as you wish (to machine accuracy), while that of spatial data depends on the sensor network. We also separated out relative and absolute time into different maps; the first concerns time from other observations, while the second is based on a system clock; the same separation can be performed with space, bearing in mind the caveats we discussed earlier. We therefore finish by considering some of the properties of data in terms of space, time, and context, and asking questions that need to be considered in order to use our ideas successfully in a smart home.

- **temporal properties**
  - **concurrency:** Do the behaviours correspond to sequences in the sensor stream or are they inter-leaved? If behaviours are interleaved, how can you segment the sensor stream?
  - **multiple scales:** Do the behaviours occur during a short time period or are they extended over longer periods? How does this effect interleaving?
  - **absolute time:** At which time of the day or day of the year does the behaviour occur? Certain activities only occur at certain times or certain days of the year and therefore would not be reflected in the sensor stream if it does not cover these time periods.
  - **relative time:** Do some behaviours always follow one another? Or never follow each other? Are there some things that can only happen after other events?
- **spatial properties**
  - **absolute space:** Which location or type of location does the sensor reading correspond to? Are there some behaviours that only occur in one room, and some that can occur in every one?
  - **discrete space:** Can the observed space be divided into discrete location or do the sensors provide continous space information (such as distance)? If space is not discrete, associating behaviours with particular locations becomes fuzzy. And does the fact that the sensor locations are specific (and fixed) help or hinder the classification?
- **contextual properties**
  - **behaviour dependencies:** How are the behaviours related to each other? Knowledge about these relationships can be used to prime the recogition process. For example, the tea making behaviour is more likely to be observed if previously the breakfast making behaviour has been recognised. This is linked to the relative time properties discussed above.
  - **environmental influences:** What is the state of the environment and how does it influence behaviours? How can you identify which environmental data is useful, and which is coincidental? For example, breakfast never takes place if the temperature is above 20 degrees Celcius, not because it is too warm, but because in some parts

of the world it never gets that warm before mid-morning.

ACKNOWLEDGMENT

We are very grateful for input from the MUSE group at Massey University.

REFERENCES

Ackley, D., Hinton, G., & Sejnowski, T. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, *9*(1), 147–169. doi:10.1207/s15516709cog0901_7

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, *26*, 832–843. doi:10.1145/182.358434

Bishop, C. (2007). *Pattern Recognition and Machine Learning*. Berlin: Springer.

Chua, S.-L., Marsland, S., & Guesgen, H. W. Behaviour recognition from sensory streams in smart environments. In *Proc. Australasian Joint Conference on Artificial Intelligence (AI-09)*, Melbourne, Australia, December 2009.

Das, S. K., Cook, D., Bhattacharya, A., Heierman, E. O. III, & Lin, T.-Y. (2002). The role of prediction algorithms in the MavHome smart home architecture. [Special Issue on Smart Homes]. *IEEE Wireless Communications*, *9*(6), 77–84. doi:10.1109/MWC.2002.1160085

Dewey, R. Psychology: An introduction. http://www.intropsych.com/, last visited on 7 July 2009.

Duong, T. V., Bui, H. H., Phung, D. Q., & Venkatesh, S. Activity recognition and abnormality detection with the switching hidden semi-Markov model. In *Proc. CVPR-05*, pages 838–845. 2005.

Egenhofer, M., & Franzosa, R. (1991). Point-set topological spatial relations. *International Journal of Geographical Information Systems*, *5*(2), 161–174. doi:10.1080/02693799108927841

Fine, S., Singer, Y., & Tishby, N. (1998). The hierarchical hidden Markov model: Analysis and applications. *Machine Learning*, *32*, 41–62. doi:10.1023/A:1007469218079

Gahegan, M. (1995). Proximity operators for qualitative spatial reasoning. In Frank, A. U., & Kuhn, W. (Eds.), *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Science 988*. Berlin, Germany: Springer.

Guesgen, H. W. (2002). Reasoning about distance based on fuzzy sets. [Special Issue on Spatial and Temporal Reasoning]. *Applied Intelligence*, *17*(3), 265–270. doi:10.1023/A:1020087332413

Guesgen, H. W., & Hertzberg, J. (1996). Spatial persistence. [Special Issue on Spatial and Temporal Reasoning]. *Applied Intelligence*, *6*, 11–28. doi:10.1007/BF00117598

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(8), 2554–2558. doi:10.1073/pnas.79.8.2554

V. Jakkula and D. Cook. Anomaly detection using temporal data mining in a smart home environment. *Methods of Information in Medicine*, 2008.

Kohonen, T. (1993). *Self-Organization and Associative Memory* (3rd ed.). Berlin: Springer.

Lafferty, J., McCallum, A., & Pereira, F. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proc. ICML-01*, pages 282–289, San Francisco, CA, 2001. Morgan Kaufmann.

Levinson, S. C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511613609

Liao, L., Patterson, D. J., Fox, D., & Kautz, H. (2007). Learning and inferring transportation routines. *Artificial Intelligence*, *171*(5-6), 311–331. doi:10.1016/j.artint.2007.01.006

Marsland, S. (2003). Novelty detection in learning systems. *Neural Computing Surveys*, *3*, 157–195.

Marsland, S. (2009). *Machine Learning: An Algorithmic Introduction*. New Jersey, USA: CRC Press.

Nguyen, N. T., Phung, D. Q., Venkatesh, S., & Bui, H. Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model. In *Proc. CVPR-05*, pages 955–960, 2005.

Peterson, C., & Anderson, J. R. (1987). A mean field theory learning algorithm for neural networks. *Complex Systems*, *1*(5), 995–1019.

Piaget, J., & Inhelder, B. (1948). *La Représentation de l'Espace chez l'Enfant*. Paris, France: Presses Universitaires de France.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286. doi:10.1109/5.18626

Randell, D. A., Cui, Z., & Cohn, A. G. A spatial logic based on regions and connection. In *Proc. KR-92*, pages 165–176, Cambridge, Massachusetts, 1992.

F. Rivera-Illingworth, V. Callaghan, and H. Hagras. Detection of normal and novel behaviours in ubiquitous domestic environments. *The Computer Journal*, 2007.

Sejnowski, T. J. (1986). Higher-order Boltzmann machines. *AIP Conference Proceedings*, *151*(1), 398–403. doi:10.1063/1.36246

Sutton, C., & McCallum, A. (2006). An introduction to Conditional Random Fields for relational learning. In Getoor, L., & Taskar, B. (Eds.), *Introduction to Statistical Relational Learning* (pp. 93–128). Boston, MA: MIT Press.

Tapia, E. M., Intille, S. S., & Larson, K. Activity recognition in the home using simple and ubiquitous sensors. In *Proc. PERVASIVE-04*, pages 158–175, Linz/Vienna, Austria, 2004.

Vilain, M., & Kautz, H. Constraint propagation algorithms for temporal reasoning. In *Proc. AAAI-86*, pages 377–382, Philadelphia, Pennsylvania, 1986.

KEY TERMS AND DEFINITIONS

Behaviour Recognition: Identifying and classifying activities on the basis of observations such as sensor data streams.

Sensor Observation: Data stream consisting of tokens associated with sensor readings.

Activity of Daily Living: Set of activities performed with a particular aim in mind, often reoccurring on a regular basis.

Novelty Detection: Recognising behaviours that are new or unusual.

Lifelong Learning: The ability of the system to adapt to changes in the environment and in human behaviours.

Hidden Markov Model: Probabilistic graphical model that uses a set of hidden (unknown) states to classify a sequence of observations over time.

Neural Network: Computational model inspired by how neurons work in the brain.

Spatial Calculus: Logic-based approach for reasoning about space.

Temporal Calculus: Logic-based approach for reasoning about time.