Author(s): Naeem, Usman; Bigham, John.
Article title: Activity Recognition using a Hierarchical Framework
Year of publication: 2008
Citation: Naeem, U; Bigham J. (2008) "Activity Recognition using a Hierarchical Framework" in Proceedings of the 2nd International Conference on Pervasive Computing Technologies for Healthcare, Ambient Technologies for Diagnosing and Monitoring Chronic Patients Workshop, Tampere, Finland, 2008, IEEE pp. 24-27.
Link to published version: http://dx.doi.org/10.1109/PCTHEALTH.2008.4571018
DOI: 10.1109/PCTHEALTH.2008.4571018

Activity Recognition using a Hierarchical Framework

Usman Naeem
Department of Electronic Engineering
Queen Mary University of London
London, United Kingdom
usman.naeem@elec.qmul.ac.uk

John Bigham
Department of Electronic Engineering
Queen Mary University of London
London, United Kingdom
john.bigham@elec.qmul.ac.uk

Abstract—**This paper describes an approach for modelling and detecting activities of daily life based on a hierarchy of plans that contain a range of precedence relationships, representations of concurrency and other temporal relationships. Identification of activities of daily life is achieved by episode recovery models supported by using relationships expressed in the plans. The motivation is to allow people with Alzheimer's disease to have additional years of independent living before the Alzheimer's disease reaches the moderate and severe stages.**

Keywords- Alzheimer's Disease, Activities of Daily Life, Episode Recovery, Task Identification

I. Introduction

It is predicted that there will be over one million people with dementia in the UK by 2025 [1]. The cost of Alzheimer's disease is currently estimated at £17 billion each year for the UK. The symptoms are different for each person. However, the main stages are similar for all patients, viz. mild, moderate and severe. The mild stage is when the person starts forgetting daily activities and not being able to carry out straight forward tasks. This on many occasions can be mitigated with the help of a diary and daily activity lists. However as the sufferer experiences memory loss they become anxious [2] in case they lose their independence. As the disease progresses it reaches the moderate stage where mental abilities decline, personality may be subject to change and physical problems may develop. The individual may become increasingly confused, conceive fictional events, find it difficult to find the right words and becomes disorientated [3]. The severe stage is characterised by high levels of disorientation, confusion with hallucinations, and sometimes aggressive behaviour [4].

The work described in this paper targets elderly people who are in the transitional stage between mild and moderate stages of this disease.

Due to forgetfulness patients with Alzheimer's disease often follow a set of activities assigned by carers or health visitors, often in a semi-prescribed order [5]. There can be times when the patient forgets what they are doing. Recognition of activities will provide useful information about their goal and what they are meant to be doing next, or even provide alternative options. It is intended that this support will also assist when goals are interweaved.

II. Related Work

There has been a significant amount of work in the area of recognition of Activities of Daily Life (ADL). Recognition of activities in the home can be split into three subcomponents; feature detection, feature extraction and models for recognition.

A currently popular technique for detecting features of ADLs is known as 'dense sensing' [6], which collects a wide range of sensor data rather than relying on visual based systems. Numerous household objects such as toasters and cupboards are tagged with wireless sensors and transponders that transmit information via a Radio Frequency Identification (RFID) reader when the object is being used or touched.

Another technique for feature detection is the use of wearable sensors such as accelerometers and audio sensors that provide data about body motion and the surroundings where the data has been collected from. Previous work [7] has shown that a variety of activities like climbing stairs and running can be determined using this technique. As well as that Wang et al [8] used this technique to detect fine-grained arm actions like 'drink with glass', 'chop with knife'. These were then combined with object-use data to achieve accurate activity recognition. The accurate recognition was based on a joint probabilistic model of object-use physical actions and activities, which showed that it was possible to combine the data from both for accurate activity recognition.

Different types of Markov models have been used to carry out task identification from a sequence of sensor events. One such approach was by Wilson et al [9], where episode recovery experiments were carried out and analysed by a Hidden Markov Model (HMM) using the Viterbi algorithm which was responsible for determining which task is active from the sequence of sensor events. Although this approach enabled unsupervised task identification it was not as efficient when the tasks were carried out in a random order.

Other approaches that have been developed in order to carry out reliable activity recognition and solve the incomplete sensor problem involve ontologies [10] and data mining techniques [11]. Ontologies have been utilised to construct reliable activity models that are able to match an unknown sensor reading with a word in an ontology which is related to the sensor event. For example, a Mug sensor event could be substituted by a Cup event in the task identification model 'Make Tea' as it uses Cup.

In general the recognition of activities that are represented by sequential models that follow a standard path of execution can be achieved using a range of techniques. However, the approaches are not very reliable when it comes to detecting activities that can be carried out in more than one way or if a particular sensor event is missing, e.g. due to a data transfer problem, or e.g. if a person who normally takes milk or sugar in his or her tea decides not to. (This can sometimes be seen as a missing sensor event.)

III. HIERARCHIES OF ACTIVITIES OF DAILY LIFE

In this paper, ADLs are modelled in a hierarchical structure of plans, which allows us to decompose the ADLs into different subcomponents. With this type of modelling, ADLs can correspond to simple tasks, such as "Switch on Kettle", or more complex activities such as "Make Breakfast" with complex dependencies between the sub-ADLs.

Figure 1 shows that the ADL "Make Breakfast" contains a simple sequence of tasks, Make Tea, Make Toast etc. The sequences of the sensor events at the lowest level (Kettle Sensor, Fridge Sensor, Tea Bag Bowl Sensor, Sugar Bowl Sensor,) correspond to sensors triggered during the task "Make Tea". The lowest nodes of the ADL hierarchy, which have no further decomposition, are defined as tasks. Tasks are identified from the stream of sensor events. A 'dense sensing' approach for gathering information from the sensor events has been used to detect the object usage from activity traces.

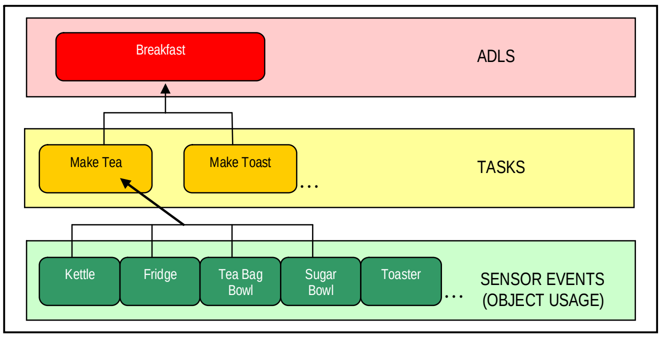

Figure 1. HADL architecture for activity 'Make Breakfast'

We have developed two different approaches to task recognition. One is based on Multiple Behavioural Hidden Markov Models (MBHMM) [12] and the other using a technique inspired from an approach for text segmentation [13]. The validity of these approaches was tested by carrying out episode recovery experiments within a kitchen.

A. *Multiple Behavioural Hidden Markov Models*

The hidden states on this occasion for the MBHMM are the steps (states) which are taken to complete a task. The steps are referred to as states. For example, a simple model of making tea would be to switch the kettle on, followed by putting sugar in cup then adding milk to the cup of tea.

Multiple models are used for each task and the one that fits the sensor events best is chosen as identifying the task. For example, there are separate models for different tasks such as Make Tea or Make Toast. However, in addition, each task may not be represented by just one model but multiple variants. This is because the elderly may carry out the task (e.g. Make Tea) in a different way.

One of the advantages of this approach is that even if the elderly person has not finished completing the task it is still possible for the MBHMM to determine which task is currently active. This is because the probability of being in the final state of the model is computed as each sensor reading is read.

When compared with traditional HMM approaches, the MBHMM out performed a simple Viterbi-based HMM for task recognition when dealing with tasks that are conducted in a different order. This is because the MBHMMs model different feasible orderings of sensor events while the Viterbi-based HMM does not. A HMM implicitly assumes that each task generates, on average the same number of sensor readings regardless of how they are performed, and the probability of returning to the same state is chosen so that the expected number of visits to a state corresponds to this average.

To a certain extent the MBHMM approach was able to solve the problem of missing sensor readings, as the models constructed for each activity modelled the possibility of an unexpected sensor event occurring between expected sensor events. The idea of modelling an unexpected state is based on profile hidden Markov models, where any unexpected sequence data which occurs in a DNA motif is substituted with an insertion.

B. *Task Segmentation*

A simple approach for segmenting tasks can be carried out by simply segmenting sensor events (based on object use data) into segments that corresponds to a task. However, a disadvantage of this approach is that sometimes sensor event segments can be generated that bare no resemblance to the task that is actually being conducted. In order to carry out task segmentation, we assigned a probability $P[a \mid b]$ for each task (a) and sensor event (b).

The entire sensor event stream is segmented into appropriate task segments. The segmented tasks are then used to determine which ADL(s) is/are currently active in the higher tier of the HADL.

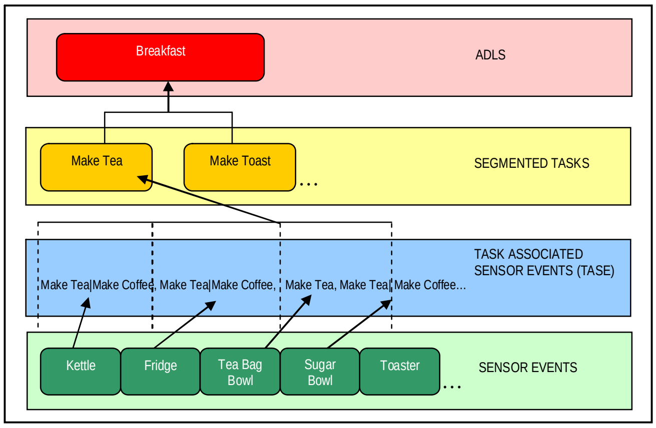

Figure 2. HADL architecture with the TASE tier

In order to accommodate this type of task identification approach, the HADL structure in Figure 1 is slightly modified, as the difference is the introduction of another tier, which is known as the Task Associated Sensor Events (TASE) tier (Figure 2).

The lowest tier of the HADL in Figure 2 deals with the incoming sequence of sensor events that have been detected, these sensor events are then associated with all the tasks that correspond to the sensor event in the TASE level. For example, 'Kettle' sensor event can be associated with 'Make Tea' or 'Make Coffee'. A segmentation algorithm is applied in order to segment tasks efficiently. This algorithm was based on a statistical model which was created for text segmentation by Utiyama et al [14]. This method was used to find the maximum-probability segmentation of text, and does not need any training data, as it estimates probabilities from the stream of text. In the context of segmenting tasks and using the task segmentation algorithm the TASE are converted into letters so that we get a stream of letters, for example;

- Task (Make Tea)= letter (A)
- Task (Make Coffee)= letter (B),
- Task (Make Toast)= letter (C)
- Task (n) = letter (n).

Each of the converted letters in the stream of letters is assigned a probability value, which is based on the number of associations each task has with the total number of sensor events. After this the most likely combinations of segments that occur in the stream of letters, for example, a stream of letters consisting of ABC will have the following combination of segments with different segmentation points: A|B|C, A|BC, AB|C and ABC.

$$\sum_{j=1}^{n_i} \log \frac{n_i + k}{p+1} + \log n_i * 0.2 \tag{1}$$

Equation (1) is then applied to each segment within each stream of letters, which outputs an overall cost for each stream. Correct segmentation of a task is determined by the stream of letters which has the lowest cost. A sample of the 10 lowest cost segmented streams is analyzed, as it gives a good idea of which task is actually being conducted by the person. It is evident that on many occasions that the sample of 10 streams may not be perfect in terms of accuracy, but this is where the higher tier of the HADL is used to refine this interpretation.

C. Higher Tier Activity Recognition

The aim of the higher tier activity recognition is to support recognition of tasks through feedback from beliefs held about ADLs.

The number of levels above the task identification level depends on the complexity of the task. ADLs may occur parallel with other ADLs and have other temporal constraints. Also, not all sub-ADLs need to be executed and there is a range of keywords to express these dependencies. These are represented in Asbru [14], which is a task-specific and intention-oriented plan representation language initially designed to model clinical guidelines. The plans in Asbru have been used to represent ADL and sub-activities within an ADL, e.g. 'Prepare Breakfast' is an ADL, and a sub-activity of this ADL is to 'Enter the Kitchen'.

Asbru has many features which allow each skeletal plan to be flexible and to work with multiple skeletal plans. When a goal is achieved in Asbru the plan is labelled as 'executed'. When the preconditions of an ADL have been met, the ADL is again classified as having been 'executed'. For example, for the goal 'Eat Toast' to start execution a pre-condition could be that the goal 'Make Toast' should be labelled as 'executed'. Additionally, an ADL can be classified as mandatory or optional. If an ADL has sub-goals that are classified as mandatory then these sub-goals must be executed before the ADL is labelled as 'executed'. If optional then the sub-goal need not be executed. Sub-goals can be ordered in many ways. Some of the more common ways include, sequential (in strict order), parallel (executed simultaneously), in any order (activated in any order but where only one sub-goal can be executed at a time) and unordered (executed without synchronisation).

D. Modelled ADL in Asbru

The root ADL plan modelled in Figure 3 is 'Having Breakfast', which is sequential. This means that the child activities within 'Having Breakfast' will be executed in a sequential order, working its way from 'Enter Kitchen' sub-activity to 'Exit Kitchen' sub-activity.

In relation to this root ADL, we suppose that the following actions/tasks are detected in the lower tiers of HADL – Enter Kitchen, Prepare Toast and Clean Dishes - in this order.

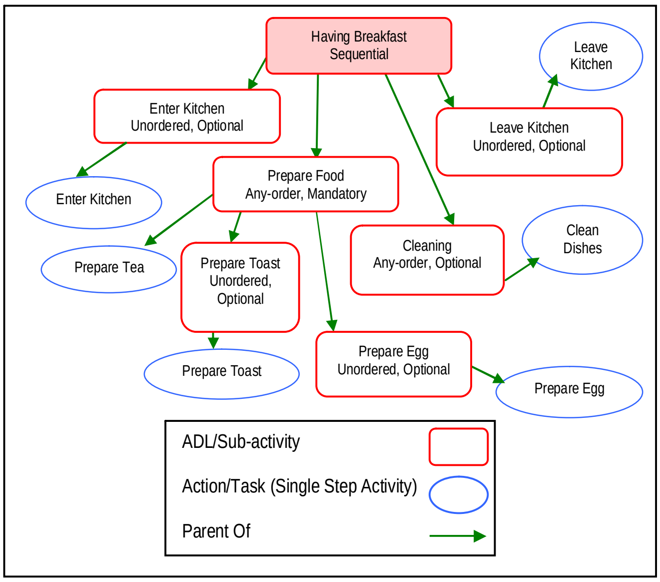

Figure 3. Modelled ADL in Asbru

At the detection of each task in Figure 3, the following processes will take place:

- **Enter Kitchen**: When Enter Kitchen is detected then the sub-activity plan Enter Kitchen is set to complete. The Enter Kitchen, which is a single step activity, allows the system to move onto the sub-activity of the root ADL plan. A single step activity is a plan that cannot be decomposed any further and is called a task, which is what it is called when it is detected in the lower tiers of HADL.
- **Prepare Toast**: As this is also a single step activity then this is also set to complete. However, the system cannot continue to the next sub-activity of the root ADL plan as the sub activity for Prepare Food is mandatory, which means that all the child activity plans and tasks within this plan must be detected before it can proceed to the next sub-activity. Instead of being mandatory, a plan may be optional, which means that a root parent activity does not need its child activities to be set to complete in order for it to move to other sub-activities.
- **Clean Dishes**: The detection of this task indicates the ADL plan in Figure 3 is not the ADL that the person is carrying out, as the mandatory tasks have not been fulfilled in the previous sub activity. This therefore means that the person in question might be having a snack rather than having breakfast.

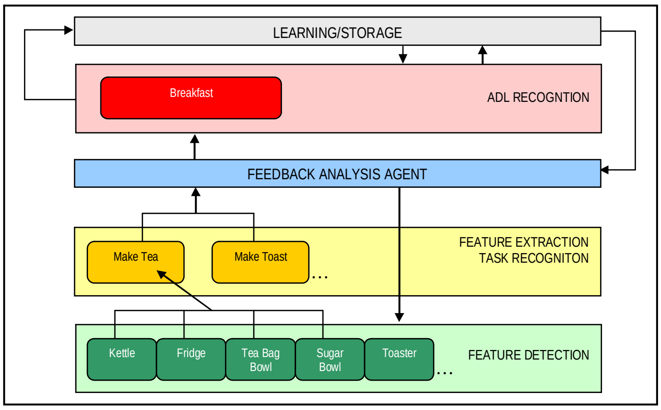

Figure 4. Enhanced Hierarchical ADL architecture for activity 'Make Breakfast'

In terms of communicating information from one tier to another, we are currently developing an approach that enhances the task (and ADL) recognition. This allows information from high level (activities) to provide feedback to the activity recognition level models, so that the task recognition can be re-assessed using the believed context from the models in the tier represented by the planning language (See Figure 4).

IV. Future Work

Currently, ADL models based on typical diary activities from Alzheimer care are being constructed with the intention of evaluating the power of the approach in ADL recognition for patients with Alzheimer's disease. The validity of these models will be tested with actual Alzheimer patients to see if it can assist them in carrying out ADLs.

References

[1] M. Knapp, "Dementia UK". London School of Economics and Institute of Psychiatry at King's College London, 27 February 2007, Published by Alzheimer's Society, UK, 2007

[2] L. Ferretti, S.M. McCurry, R. Logsdon, L.E. Gibbons, L. Teri, "Anxiety and Alzheimer's Disease". Journal of Geriatric Psychiatry and Neurology, Volume 14, Number 1, Page(s): 52-58, 2001

[3] S.J. Tetewsky, C.J. Duffy, "Visual loss and getting lost in Alzheimer's disease". Journal of Neurology, Volume 52, 23 March 1999, Page(s): 958-965, 1999

[4] D. Aarsland, J.L. Cummings, G. Yenner, B. Miller, "Relationships of aggressive behaviour to other neuropsychiatric symptoms in patients with Alzheimer's disease". The American Journal of Psychiatry, Volume 153, February 1996, Page(s): 243-247, 1996

[5] The Alzheimer's Association. "Activities at Home, Planning the day for a person with dementia". Alzheimer's Association National Office 225 N. Michigan Ave., Fl. 17, Chicago, IL 60601, Page(s): 1 -12, 2005

[6] M. Philipose, K.P. Fishkin, M. Perkowitz, D.J. Patterson, H. Kautz, D. Hahnel, "Inferring Activities from Interactions with Objects". IEEE Prevasive Computing Magazine, Volume 3, Issue 4, Page(s): 50-57, 2004

[7] J. Lester, T. Choudhury, N. Kern, G. Borriello, B. Hannaford, "A Hybrid Discriminative/Generative Approach for Modelling Human Activities". Proceddings of the 19th International Joint Conference on Artificial Intelligence, Page(s): 766-772, 2005

[8] S. Wang, W. Pentney, A. Popescu, T. Choudhury, M. Philipose, "Commonsense-based Joint Training of Human Activity Recognizers". Proceedings of the 21st International Joint Conference on Artificial Intelligence, January 2007

[9] D.H. Wilson, D. Wyaat, M. Philipose, "Using Context History for Data Collection in the Home". In Gellersen, H.-W., Want, R., Schmidt, A. (eds.) PERVASIVE 2005. LNCS, Volume. 3468, 2005

[10] E. Munguia-Tapia, T. Choudhury, M. Philipose, "Building Reliable Activity Models using Hierarchical Shrinkage and Mined Ontology". Proceedings of the 4th International Conference on Pervasive Computing, Page(s): 17–32, 2006

[11] D. Wyatt, M. Philipose, T. Choudhury, "Unsupervised Activity Recognition using Automatically Mined Common Sense". Proceedings of the 20th National Conference on Artificial Intelligence, Page(s): 21–27, 2005

[12] U. Naeem, J.Bigham, "A Comparison of Two Hidden Markov Approaches to Task Identification in the Home Environment". Proceedings of the 2nd International Conference on Pervasive Computing and Applications, Page(s): 383-388, 2007

[13] U. Naeem, J. Bigham, J. Wang, "Recognising Activities of Daily Life Using Hierarchical Plans". Proceedings of the 2nd European Conference on Smart Sensing and Context, LNCS 4793, Page(s): 175-189, 2007 in press.

[14] M. Utiyama, H. Isahara,, "A Statistical Model for Domain Independent Text Segmentation". Proceedings of the 39th Annual Meeting on Association for Computational Linguistics, Page(s): 499-506. Toulouse, France, 2001

[15] C. Fuchsberger, J. Hunter, P. McCue, "Testing Asbru Guidelines and Protocols for Neonatal Intensive Care". Proceedings of the 10th European Conference on AI, Page(s):101-110, 2005