# Defining the Complexity of an Activity

**Yasamin Sahaf, Narayanan C Krishnan, Diane J. Cook**
Center for Advance Studies in Adaptive Systems,
School of Electrical Engineering and Computer Science,
Washington State University
Pullman, WA 99163
{ysahaf, ckn, cook}@eecs.wsu.edu

### Abstract

Activity recognition is a widely researched area with applications in health care, security and other domains. With each recognition system considering its own set of activities and sensors, it is difficult to compare the performance of these different systems and more importantly it makes the task of selecting an appropriate set of technologies and tools for recognizing an activity challenging. In this work-in-progress paper we attempt to characterize activities in terms of a complexity measure. We define activity complexity along three dimensions – sensing, computation and performance and illustrate different parameters that parameterize these dimensions. We look at grammars for representing activities and use grammar complexity as a measurement for activity complexity. Then we describe how these measurements can help evaluate the complexity of activities of daily living that are commonly considered by various researchers.

## Introduction

ADLs (Activities of Daily Living) have been studied in different fields. These are often used in healthcare to refer to daily self-care activities within an individual's place of residence, in outdoor environments, or both. Health professionals routinely refer to the ability or inability to perform ADLs as a measurement of the functional status of a person (Meghan 2002). This measurement is useful for assessing older adults, individuals with cognitive disabilities and those with chronic diseases, in order to evaluate what type of health care services an individual may need. There are many ADL lists published in the Psychology domain; however each research group has concentrated on a subset of this list according to their own needs and requirements (Hindmarch et al. 1998, Garrod et al. 2000).



While sitting, standing, walking etc., appear at one end of the spectrum of activities, the other end consists of complicated activities such as cooking and taking medication, which encompass ambulation, ADLs and instrumental ADLs (iADLs). From a computational standpoint, it is difficult to combine these different activities into a single category for the purpose of designing a recognition system. Having a standard way to classify these activities based on their complexities will help researchers in all fields who want to study activities. This is the primary motivation behind this paper, where we attempt to define a formal complexity measure for activities. The complexity of an activity can be defined in terms of different parameters such as the underlying sensing modality, the computational techniques used for recognition or inherent properties of the activity. We describe each of these parameters in greater detail. Defining such a complexity measure provides a means for selecting activities for conducting benchmarking experiments. Furthermore, it also helps in choosing the correct technology for recognizing a specific set of activities.

## Defining Activity Complexity

In general, the complexity of an activity can be defined in terms of different factors. In this paper we attempt to define it in terms of three components: Sensing complexity, Computational complexity, and Performance complexity.

### Sensing Complexity

Sensing complexity refers to complexity of sensors which are used in collecting data. Research advances in computing have resulted in the development of a wide variety of sensors that can be used for sensing activity.

On one hand there are sensors that have to be worn by individuals (Krishnan and Panchanathan 2008) and on the other hand there are environmental and object sensors that have to be embedded in the environment for gathering activity related information (Singla, Cook and Schmitter-Edgecombe 2009). Each of these sensors provides a rich set of information on a certain set of activities. For example, it is easier to recognize ambulation using wearable sensors over environmental sensors, while iADLs such as cooking and bathing are easier to recognize using environmental sensors. We define the sensing complexity of activities in terms of the following parameters: number of distinct sensors fired, number of sensor types fired, number of objects involved which can put sensor on, sensor size, sensor price, ease of use (Subject, Deployment), type of output data, battery life and type of sensor (wired or wireless). In the following paragraphs, we will discuss each of them in more detail.

The number of sensors used is an important factor that defines this complexity, which in turn can be divided into two groups: number of distinct sensors fired and number of sensor types fired. For example, one particular sensor might be fired many times, but we count it as only one distinct sensor. Based on the technology used in each study, different sensor types can be seen, such as environmental sensors (motion, temperature, light, etc), object sensors (RFID tags, accelerometers, shake sensors, etc) and wearable sensors (accelerometers, RFID, health monitoring sensors, etc). For example if we are using environmental motion sensors, wearable accelerometers and shake sensors on objects, all three sensor types are fired for cooking activity. But for washing hands, only two of them are fired: environmental and wearable (assuming no sensor has been placed on soap). The number of objects involved in an activity that can be sensed through some modality is another factor defining the sensing complexity. For some activities such as brooming, placing sensors on the objects involved (broom) is possible, thus it can be considered simpler than reading books (placing sensor on every book is impractical).

The price and form factor of a sensor is another component of the sensing complexity. An expensive sensor system would be harder to implement, so it can be considered more complex. The same is true with sensor size, especially for wearable and object sensors. Smaller sensors are easy to adopt, while bigger sensors are relatively difficult to deploy. The ease of use of a sensor can be seen from two perspectives: Subject and Deployment. Ease of use with respect to subject refers to ease and level of acceptance with which participants use sensors. For example some wearable sensors could be easier and more comfortable for participants to wear. The deployment aspect of ease of use can be defined in terms of the ease with which experimenters deploy a particular sensor. A sensor might give us helpful data but working with it might be too hard for experimenters that they prefer alternative but less useful ways. This reasoning would be true about type of output of the sensor as well. Some sensor outputs need further complex computations and pre-processing which results in higher sensing complexity.

The battery life of a sensor is an important factor especially in the context of wireless and wearable systems. Choosing wired or wireless sensor depends on the requirements of the system and it has effect on the sensing complexity.

While the values for some of these parameters (e.g., number of sensors, battery life) can be derived empirically, other factors (e.g., form factor and ease of use) require some kind of subjective evaluation. We would expect the measure derived from these parameters to be low for ambulatory activities for wearable sensors such as accelerometers, but will be high for environmental sensors such as motion sensors. In Table 1 we have represented some of the popularly used considered ADLs using these different factors.

## Computational Complexity

Advances in machine learning and pattern recognition domain have resulted in a number of supervised and unsupervised techniques for recognizing activities. Discriminative classifiers such as SVMs (Krishnan and Panchanathan 2008), Logistic regression (Krishnan and Panchanathan 2008), CRFs (Nazerfard et al. 2010) and generative classifiers such as GMMs (Pansiot et al. 2007), HMMs (Singla, Cook and Schmitter-Edgecombe 2009) are very popular for activity recognition. In addition to this, computational complexity also includes the algorithms that transform the raw data stream into a form that is used by some of the recognition algorithms. Examples of these algorithms are FFTs (Huynh and Schiele 2005), wavelets, and other techniques that extract the statistical and spectral properties of the raw data. The main component of the computational complexity is the complexity of the underlying recognition/transformation algorithm. Other factors that affect the computational complexity include memory requirements of the algorithm and real-time performance. The relevance of the computational complexity of an activity depends on the computational resources available. For example, if the goal of the system is to perform recognition on a low power device such as mobile phone, the computational complexity plays an important role in selecting the appropriate set of algorithms.

| Activity | Number of objects involved that cannot put sensors on | Number of distinct sensors fired | Average Time | Duration deviation | Number of people involved | Has a predefin ed time? | Number of distinct location movements |
|---|---|---|---|---|---|---|---|
| **Sweeping** | Low | High | High | Medium | 1 | No | Low |
| **Medication** | Medium | Low | Low | Low | 1 | Yes | Low |
| **Watering plants** | Low | Medium | Medium | Low | 1 | No | Medium |
| **Hand washing** | Medium | Medium | Low | Low | 1 | No | Low |
| **Washing kitchen countertops** | Medium | Low | Medium | Medium | 1 | No | Low |
| **Cooking** | High | Medium | High | Medium | 1 | Yes | Low |

*Table 1: Complexity measurement over activity based on WSU CASAS sensing technology.*

### Performance Complexity

We define the performance complexity to be an abstraction of some of the inherent properties of an activity that is independent of the underlying sensing and computational mechanisms. This complexity term can be defined using different parameters such as: average duration and deviation, duration of non repetitive patterns, predefined time of the activity, number of steps, number of distinct location movements, number of people and objects involved.

The average duration of an activity, even though an important component, does not clearly differentiate the complexity of activities. In other words there is no general rule that can say an activity with higher duration is more complex or vice versa. As an example, cooking is a relatively long and complex activity. At the same time sleeping is also long but not very complex from the perspective of recognition. Thus, this component should be taken into consideration along with other factors.

Perhaps one could look at how much time during the activity the person was active. For example, a person is not active for a large portion of time while sleeping and watching TV. Associated with the average duration of an activity is also the deviation in the duration in the performance of the activity.

The third component is the duration of non-repetitive patterns. Patterns in activities usually give us useful information. Repetitive patterns are easier to recognize. For example, walking or running involve periodic movements of the human body that can be easily recognized, in contrast to movements such as pouring water, or scooping sugar while making a cup of tea. Some activities have a predefined time of occurrence during the daily routine of an individual. Such a unique characteristic of an activity can be effectively utilized by machine learning algorithms for recognition. An example of such an activity is taking medication.

Typically every activity is defined in terms of a number of steps. Some activities have larger number of steps which make them more complex. An activity step can be defined as event that cannot be divided in to sub-events in the current technology. Defining the activity steps in this format facilitates different representations of the steps depending on the underlying technology. The next issue to be considered is the number of distinct location movements; an activity which is performed in different locations can be considered more complex in comparison with an activity that takes place in one location.

Other factors that define the performance complexity of an activity are the number of people and objects involved in that activity. The activities get more complex with an increasing number of people and objects defining the activity.

## Evaluating the Complexity

In Table 1 we have represented 6 common activities and measured some of their complexity measurements discussed before. There are different ways to generate one total value from these measurements. One straight forward approach would be assigning numbers 1, 2, 3 to values low, medium and high respectively, and then summing up all the values for each activity. We can ignore the value of ‘Number of people involved’ in this case, since it is the same for all these activities. Following above rules we will get 8 for ‘cooking’, 7 for ‘sweeping’, 6 for ‘watering plants’, ‘hand washing’ and ‘washing counter tops’ and 5 for 'medication'. Therefore, 'cooking' can be categorized as the most complex activity to recognize with this study's sensing technology and 'taking medication' as the easiest one. For generating these examples we assumed sensing technology of WSU Center for Advanced Studies in Adaptive Systems (CASAS), which consists of three sensor types (environmental, wearable and object).

## Using Grammar Complexity

While the complexity values can be derived from pre-defined measures as described previously, another possible approach is making use of grammars for representing activities. Then, grammar complexity can be used for measuring complexity of the corresponding activity. Using grammar has different benefits. It helps to formally define complex activities based on simple actions or movements. Rules are understandable by human. It can be extended and modified at any time and it can be used by systems with different technologies. In addition, grammar facilitates us with a formal representation of activities which helps researchers in different fields to have a benchmark while trying to choose and compare activities in their studies.

Researchers have used grammar for representing different activities. Ward et al. have used wearable accelerometer and looked at wood workshop activities such as "grinding" and "drilling" (Ward et al. 2005). But most of studies have used camera for gathering data; for example Ryoo and Aggarwal have defined grammar for activities such as "Shake hands", "Hug", "Punch", etc (Ryoo and Aggarwal 2006). Chen et al. have used grammar in gesture recognition (Chen, Georganas and Petriu 2007). There are a few studies on using grammar for representing ADLs, Teixeira et al. has represented ADLs with hierarchical finite state machines (Teixeira et al. 2009).

In other areas such as Human Computer Interaction (HCI) user tasks have been represented by means of task notations. A task defines how the user can reach a goal in a specific application domain. Paterno has defined CTT model which provides a rich set of operators to describe the temporal relationships among tasks and enables designers to describe concurrent tasks (Paternò, Mancini and Meniconi 1997). In addition, for each task, further information can be given; task is described by attributes including Name, Type (abstract, user, application, interaction), Subtask of, Objects, Iterative (a Boolean indicating whether the task is iterative), First action, Last action.

Beyond these descriptive aspects, these notations can also be used to assess the complexity of the tasks. Palanque and Bastide have modeled tasks using the Interactive Cooperative Objects (ICO) formalism, which is based on Petri nets and on the object-oriented approach (Palanque and Bastide 1970). In their quantitative analysis of task complexity they have considered number of nodes (corresponding to the number of states in the task model) the number of actions (corresponding to the number of arcs with different labels) and the length of the path to come back to the initial state which are associated with weights.

To the best of our knowledge, no study has looked at complexity of grammar to derive the activity complexity. Different grammars such as CFG (Teixeira et al. 2009), SCFG (Moore and Essa 2001), DOP (Data Oriented Processing), LFG (Lexical-functional Grammar) can be used for this purpose. In this study we will focus on Context-free Grammar, in which the left-hand side of each production rule consists of only one single non-terminal symbol, and the right-hand side is a string consisting of terminals and/or non-terminals. Human actions and interactions are usually composed of multiple sub-actions which themselves are atomic or composite actions and CFG is able to construct a concrete representation for any composite action (Ryoo and Aggarwal 2006). On the other hand, context-free grammars are simple enough to allow the construction of efficient parsing algorithms (Chen, Georganas and Petriu 2007).

In this study we present a very simple CFG as a baseline for future work which can represent sequential behaviors. In order to define a CFG, we need to define terminals and non-terminals symbols. We can associate the atomic actions with the terminals and complex actions with non-terminal symbols. However, as discussed before, the definition of the atomic action can vary according to the underlying sensing technology. For example, if one is looking at walking patterns, atomic action can be each movement of legs and hands, if one is using accelerometers as the sensing modality. In contrast, in a study that only uses environmental sensors, moving from one part of the room to the other which results in triggering a new sensor is considered atomic. In this paper, we try to define a general definition in a way that any research study will be able to adopt it. Continuing with our previous discussion, we define an atomic action as an event that cannot be divided into smaller sub-events that is recognizable by the underlying sensing modality. If an action contains two or more atomic actions, it is classified as a composite action (Ryoo and Aggarwal 2006). By using CFG, we are able to define a composite action (Non-terminal) based on atomic actions (Terminals).

In order to formally represent an atomic action we follow the linguistic theory of "verb argument structure". Park's operation triplet is <agent-motion-target> (Park and Aggarwal 2004), where agent refers to the body part (i.e. arm, head) directed toward an optional target. Motion set contains action atoms such as "stay", "move right", etc.

But this triplet is too specific to their sensing technology which is using camera and image processing.

As a more generic formal representation we define an atomic action as <agent – motion – location - target> where an agent is the person performing the action, motion represents the event of that atomic action which can be in any form based on the technology, location indicates the location of the event and target is the object or person in interaction. If the action doesn't contain any interaction, target value will remain null. As an example, we chose two common activities and formalized them with this CFG scheme. Following examples show 'Sweeping' and 'Dusting' activities. There is only one person involved in these activities which is represented by 'i'. In order to generate these examples we assumed CASAS sensing technology which we have described before.

**Sweeping:**

RetrieveBroom(i) =
atomicAction(<i, RaiseHand, Near kitchen cupboard, Broom>)
SweepKitchenFloor(i) =
atomicAction(<i, Repetitive pattern & Raise, Kitchen, Broom>)
Sweep(i) -->
RetrieveBroom(i) and SweepKitchenFloor(i)

**Dusting:**

DustLivingRoom(i) =
atomicAction(<i, Repetitive pattern & Raise, Living room, Duster>)
DustDiningRoom(i) =
atomicAction(<i, Repetitive pattern & Raise, Dining room, Duster>)
Dusting(i) -->
DustLivingRoom(i) or DustDiningRoom(i)
DustRooms(i) -->
RetrieveDuster(i) and Dusting(i)

## Summary

In this paper we have defined the complexity of an activity using two approaches. First, we have proposed measurements along three dimensions sensing, computation and performance. We have illustrated some of the parameters that define each of these dimensions, and then categorized some of the popularly used ADLs using these measures. In addition, we propose to use grammars as a formal representation of activities and make use of grammar complexity for categorizing ADLs.

## References


Meghan, M. G. 2002. Activities of Daily Living Evaluation. Encyclopedia of Nursing & Allied Health. Ed. Gale Group, Inc.

Hindmarch, I., Lehfeld, H., Jongh, P. and Erzigkeit, H. 1998. The Bayer Activities of Daily Living Scale (B-ADL). Dement Geriatr Cogn Disord. 9(suppl 2):20-26.

Garrod, R., Bestall, JC., Paul, EA., Wedzicha, JA. and Jones, PW. 2000. Development and Validation of a Standardized Measure of Activity of Daily Living in Patients with Severe COPD: the London Chest Activity of Daily Living Scale(LCADL). Respir Med. 94(6):589-96.

Krishnan, N. C., Panchanathan, S. 2008. Analysis of low resolution accelerometer data for human activity recognition, IEEE conference on Acoustic, Speech and Signal Processing.

Singla, G., Cook, G. and Schmitter-Edgecombe, M. 2009. Recognizing independent and joint activities among multiple residents in smart environments. Journal of Ambient Intelligence and Humanized Computing.

Nazerfard, E., Das, B., Holder, L.B., and Cook, D.J. 2010. Conditional Random Fields for Activity Recognition in Smart Environments. Proceedings of IHI.

Pansiot, J., Stoyanov, D., McIlwraith, D., Lo, B. and Yang, G.Z. 2007. Ambient and Wearable Sensor Fusion for Activity Recognition in Healthcare Monitoring Systems. In Proc. of BSN 07, pp. 208-212.

Huynh, T., Schiele, B. 2005. Analyzing features for activity recognition. In: Proc. Soc-EUSAI 2005. ACM Int. Conf. Proceeding Series, ACM Press: 159–163

Ward, J., Lukowicz, P., Tr ¨oster, G. and Starner, T. 2005. Activity recognition of assembly tasks using body-worn microphones and accelerometers. In PAMI, vol. 28(10).

Ryoo, M. S. and Aggarwal, J. K. 2006. Recognition of Composite Human Activities through Context-Free Grammar based Representation. IEEE Conference on Computer Vision and Pattern Recognition.

Chen, Q., Georganas, N. D. and Petriu, E. M. 2007. Real-time vision-based hand gesture recognition using haar-like features. In Proc. of the IEEE Instrumentation and Measurement Technology Conf., pages 1 – 6.

Teixeira, T., Deokwoo Jung, Dublon, G., Savvides, A. 2009. Recognizing activities from context and arm pose using finite state machines. Third ACM/IEEE International Conference on Distributed Smart Cameras. ICDSC.

Paternò, F., Mancini, C. and Meniconi, S. 1997. ConcurTaskTrees: A Diagrammatic Notation for Specifying Task Models. In: Proc. of Interact'97. Chapman & Hall, 362-369.

Palanque P. & Bastide R. 1970. Synergistic modelling of tasks, system and users using formal specification techniques. Interacting With Computers, Academic Press, 9, 12, pp. 129-153
Wood W.A. Transition network grammars for natural language analysis. Communications of the ACM 13, 10, 591-606

Moore, D. and Essa, I. 2001. Recognizing multitasked activities using stochastic context-free grammar. In *CVPR* Workshop on Models vs Exemplars in Computer Vision.

Park, S., Aggarwal, J.K. 2004. Semantic-level understanding of human actions and interactions using event hierarchy. CVRP Workshop on Articulated and Non-Rigid Motion, Washington DS, USA.