Assisted Cognition:

Compensatory Activity Assistance Technology

Donald J. Patterson

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

2005

Program Authorized to Offer Degree: Computer Science and Engineering

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Donald J. Patterson

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Co-Chairs of Supervisory Committee:

Henry Kautz

Dieter Fox

Reading Committee:

Henry Kautz

Dieter Fox

Gaetano Borriello

Date: 10/6/2005

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature

Date 10/5/2005

University of Washington

Abstract

Assisted Cognition: Compensatory Activity Assistance Technology

Donald J. Patterson

Co-Chairs of Supervisory Committee:
Professor Henry Kautz
Department of Computer Science and Engineering
Professor Dieter Fox
Department of Computer Science and Engineering

The predicted increase in the number of elderly members of the industrialized world suggests an associated increase in the number of people who are going to be diagnosed with various forms of cognitively disabling dementias, including Alzheimer's disease.

A possible solution to the cost of caring for these members of society is to augment our care network with sophisticated cognitive aids which can compensate for simple cognitive errors.

In this thesis I present the design and implementation of two cognitive aids. The first is an outdoor navigation assistant and the second is an indoor household activity monitor.

Both systems are characterized by a fusion of sensor data with background knowledge and are interpreted in a probabilistic framework.

The result of this work is a demonstration of the feasibility of developing cognitive aids based on real-time streaming sensor data.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

GLOSSARY

2TBN: A two-slice temporal Bayesian Network: A 2TBN is a type of DBN.

ACCESS: Assisted Cognition in Community, Employment and Support Settings: ACCESS is a project to create novel technologies that will enhance the quality of life of people with cognitive disabilities by supporting personal navigation and use of public transportation.

ACTIVITY COMPASS: A class of portable devices that are collocated with an individual and that assists her in carrying out navigation tasks, especially in the presence of cognitive errors.

AC: Assisted Cognition

ADL: Activity of Daily Living

AD: Alzheimer's Disease

BABY-BOOMERS: This term refers to those people born between 1946 and 1964.

BARISTA: Bayesian Recognition Assistant

BLUETOOTH: This is an industrial specification for short-range wireless communications of data and voice between both mobile and stationary devices.

CPT: Conditional Probability Table

DBN: Dynamic Bayesian Network

EM: Expectation Maximization

FAST: Functional Assessment Stages: FAST is a way to quantify cognitive decline.

FMMSE: Folstein Mini-Mental Status Examination: FMMSE is a way to quantify cognitive decline.

GDS: General Dementia Scale: GDS is a way to quantify cognitive decline.

GIS: Geographic Information System

GPRS: General Packet Radio Service: GPRS is a radio technology for GSM networks that makes efficient use of available radio spectrum for higher data rate communications on mobile devices.

GPS: Global Positioning System: GPS is a constellation of non-geosynchronous satellites which provide an infrastructure for passive location computation.

GSM: Global System for Mobile Communication: GSM is a digital mobile phone standard.

GUIDE: This is a project at Intel Research Seattle to recognize activities based on object usage.

HMM: Hidden Markov Model

IADL: Instrumental Activity of Daily Living

J2ME: Java 2 Platform, Micro Edition: J2ME is an application development platform for ubiquitous computing platforms

MCI: Mild Cognitive Impairment, MCI is a precursor to Alzheimer's Disease

MIDP: Mobile Information Device Profile: MIDP is a set of Java APIs that is designed for mobile phones.

MMSE: Mini-Mental Status Examination: MMSE is a way to quantify cognitive decline.

OK: Opportunity Knocks, see below.

OPPORTUNITY KNOCKS: "Opportunity Knocks" is a specific implementation of an Activity Compass which assists users in making and completing transportation plans based on GPS sensor readings.

PARTICLE FILTERS: These are a Monte-Carlo based inference technique.

PDA: Personal Digital Assistant

RFID: Radio Frequency Identification

SISR: Sequential Importance Sampling with Replacement

UI: User Interface

WAAS: Wide Area Augmentation System: WAAS is a method of augmenting the existing GPS infrastructure to provide more accurate positioning by adding geostationary satellites to the GPS constellation.

ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to the faculty and staff of the University of Washington Department of Computer Science and Intel Research Seattle who helped me in ways small and great to get a Ph.D. Special thanks to:

- Professor Henry Kautz for meeting with me for many years, providing valuable guidance, finding funding, calming my panic attacks and assisting me with my writing.

- Associate Professor Dieter Fox for helping to frame the technical aspects of the work so that they would be compelling and for not letting me get away with saying something was too hard.

- Professor Gaetano Borriello for tactical advice over numerous conference dinners and for helping me to achieve the right balance among competing technical approaches.

- Kay Beck-Benton, Jeremy Buhler, Seth Bridges, Krzysztof Gajos, Lin Liao, Luke McDowell, Gerome Miklau, Kristi Morgansen, Matthai Philipose, Larry Ruzzo, Shaiokai Wang Ken Yasuhara,

- and my wife, Sandy, who was working hard by my side to help me finish this thesis.

DEDICATION

We have not known Thee as we ought,
Nor learned Thy wisdom, grace and power;
The things of earth have filled our thought,
And trifles of the passing hour.
Lord, give us light Thy truth to see,
And make us wise in knowing Thee.

...

We have not served Thee as we ought,
Alas, the duties left undone,
The work with little fervor wrought,
The battles lost or scarcely won!
Lord, give the zeal, and give the might,
For Thee to toil, for Thee to fight.

Thomas B. Pollock,
from Supplemental Hymns to Hymns Ancient and Modern, 1889.

Chapter 1

INTRODUCTION

1.1 The Aging of the Baby-Boomers

In the next 50 years, the industrialized world is going to see a dramatic change in the demographics of its population. A huge number of people, the children who were the baby-boom, are going to begin to reach retirement age and experience the predictable effects of aging. Not only is the absolute size of this demographic group a first for the world, but the relative numbers of elderly compared to working age individuals is also shifting.

The economic and social cost of providing a good quality of life for the baby-boomers is a real and present challenge that we are facing. One of the most effective ways to alleviate the costs associated with aging is to increase the time during which people are able to live independently.

One of the major risks to independence, and a reason for the decline in the accomplishment of typical Activities of Daily Living (ADLs) by the elderly is the onset mild cognitive impairment (MCI), a precursor to Alzheimer's Disease. Regardless of the presence of cognitive disabilities, research suggests that one of the best ways to prolong independence is to encourage the successful completion of ADLs [13]. A side benefit of such an improvement is an increase in the quality of life of caregivers, and improved socialization of our seniors – both of which correlate to increased quality of life.

1.2 Finding Solutions

Independently of these aging trends, great strides have recently been made in the development of mobile and ubiquitous sensors and software technology for reasoning about them. The term Assisted Cognition (AC) has been coined at the University of Washington to refer to methods from artificial intelligence that are applied toward understanding, generalizing, and extending the data created by these sensors.

The number of such sensors, both in the marketplace and deployed in the environment, is quickly growing. They include familiar consumer grade technologies like the Global Positioning System (GPS) and Wi-Fi, as well as more specialized technologies like the Intel Personal Server[1], Radio-Frequency Identification (RFID) tags and MICA motes[2]. In many ways this phenomenon parallels the emergence of Internet information sources, which are also heterogeneous with varying degrees of trustworthiness. However, these embedded and mobile sensors are unique in that they are typically placed in the context of a location or a user, are continuously streaming data, and are observing actual physical processes.

1.3 The Hypothesis

This growth suggests an exciting vision of computer systems research that combines ideas from machine learning, artificial intelligence, and mobile and embedded computing.

This thesis seeks to determine if it is possible to use next generation sensors to recognize activities of individuals with sufficient detail that cognitive errors can be recognized, and corrected.

I will explore two facets of this space under the umbrella of *compensatory activity assistance technology.* The first is an outdoor recognition and assistance system called "Opportunity Knocks" (OK), which relies on GPS data, and the second is an indoor human activity recognition system, "BARISTA", which uses RFID tags. Creating these systems required extending current machine learning techniques, a principled integration of external knowledge bases, and the development of novel modes of interaction with ubiquitous computing systems.

---

[1] A portable CPU without a human-computer interface, http://www.intel.com/labs/features/rs08031.htm

[2] Small low power CPUs optimized for sensing,http://computer.howstuffworks.com/mote4.htm

1.4 Contributions of Assisted Cognition

1.4.1 Opportunity Knocks

Opportunity Knocks is an outdoor urban transportation assistance system designed for individuals with mild cognitive decline. It is accessed with a cell-phone and requires the user to carry a GPS receiver. By monitoring a user's position over time, OK is able to develop a probabilistic model of a user's transportation behavior. Behaviors include activities such as walking to a bus stop, boarding a bus, and riding in a bus or a car. It uses this model to monitor his or her progress, even in the face of signal loss, and to predict future movements, locations and destinations. It can also provide real-time, customized, proactive assistance by detecting explicit user errors, such as missing a bus stop. We implemented this system and conducted experiments using over 60 days of continuous GPS data from two (fully functioning) individuals. These experiments showed that our techniques could correctly infer a user's current transportation mode 84% of the time, an improvement of 30% over other machine learning techniques. In addition it could correctly predict where a user would be 17 city blocks in the future 50% of the time, even if the user changed modes of transportation. Other research contributions include:

- **An Inference Engine:** OK models user transportation decisions as a hierarchical Dynamic Bayesian Network (DBN). At the lowest level of the hierarchy are raw GPS sensor signals and at the highest level is abstract information about the user's transportation decisions. Inference is performed using a sample based inference algorithm and individual user behavior is learned in an unsupervised manner using Expectation-Maximization (EM). The system automatically discovers unique aspects of user behavior, such as frequently used bus stops and parking lots, common destinations, and routes between them.

- **A Novel User Interface:** OK utilizes a simple user interface in which candidate destinations are displayed as images. The most likely destinations are chosen by the inference engine according to the user's history and context. When the user needs assistance, she selects one of four digital images and receives personalized routing information and real-time transportation status. If the system detects that the user is at a new destination, it asks her to take a representative picture using the camera-phone. That image is used in the future to refer to that location as a possible destination.

- **An Error Detection Algorithm:** OK reasons with three DBN model variants. The first uses parameters that are consistent with physical constraints (*e.g.*, people get on and off buses at bus stops, people travel slower by foot than by car, etc.). The second is trained to a user's personal behavior. The third is the same, but is clamped to presume only one particular destination. We use model selection techniques to determine the probability that a user's actions are compatible with his or her goal.

1.4.2 **BARISTA**: **B***ayesian* **A***ctivity* **R***ecognition ass***ISTA***nt*

BARISTA is an indoor household activity monitoring system, motivated by the desire to keep a diary of daily activities for the long-term health monitoring of elderly individuals. This system assumes a world in which consumer goods are tagged with RFID tags, and users wear a bracelet sensor that can detect an object being touched. By observing the order and timing of touches, we demonstrated the ability to recognize activities at a variety of different granularities. This work made the following research contributions:

- **Broad Activity Recognition:** Using a general inference engine, we were able to show for the first time the ability to identify a broad range of ADLs with high accuracy using a single sensor suite. In the past such systems were highly engineered to a particular activity and were rarely transferable to new activities. By contrast, we were able to use one system to recognize over 14 classes of activities, ranging from doing laundry to preparing a simple meal, with greater than 88% accuracy.

- **Fine-Grained Activity Recognition:** We systematically explored what aspects of activity recognition helped our models recognize activities at a fine granularity. We were able to show that timing distributions and knowledge of object classes improved accuracy. Such detailed models are a prerequisite for performance evaluation, trending, guidance, and helpful intervention.

- **Scalable Activity Recognition through Data-Mining:** We demonstrated the potential for scaling activity recognition through web data-mining. Using automatic model creation techniques from natural language texts, we were able to recognize which of 120 recipes a user was making, with accuracy rates 48 times better than chance, despite the fact that many objects and ingredients were shared between recipes.

1.5 Summary

Using sensors and probabilistic reasoning to assist people with cognitive disabilities is a challenging and compelling goal. Not only does it hold the promise of addressing imminent needs for the disabled, but it also has the promise of assisting normally functioning individuals as well, promoting their health, safety and quality of life.

This thesis looks at two ways of accomplishing this. Both ways fall into the broad category of compensatory activity assistance technology.

Chapter 2

THE ALZHEIMER'S CHALLENGE

The number of elderly in the industrialized world is growing quickly. With age come associated health care challenges. My concern in this chapter is to look at the impact of a particular form of dementia called Alzheimer's disease (AD) on future society. I will show how it is defined and measured so that we will have an understanding of where to look for technological opportunities for intervention.

Such interventions are warranted because the rising numbers of elderly with AD are going to extract significant economic and social costs. On the economic side we will see that the costs of aging are largely due to long-term health care and that dementia and reduced functionality are the leading causes for admission to long-term health care facilities. On the social side, mental health for both caregivers and people with AD depends on functional independence.

Finally, I will conclude with an analysis of the best places to direct technological interventions to have the most impact on people with AD. Clinical research will also suggest some of the forms that such interventions should take.

Let's begin with a look at how demographics in America (and the rest of the world) are changing.

2.1 America's Changing Demographics

The industrialized world is growing older rapidly. Not only is the absolute number of older people growing, but their population is also growing in relative terms to the rest of society. In this section we will look at both trends starting with the absolute numbers.

The U.S. Census Bureau tracks older Americans according to two age brackets, which will serve as a good reference point, those who are over 65 years of age and those who are over 85. These two groups are also referred to as "older Americans" and the "oldest-old

Americans," respectively.

2.1.1 The Aging Baby Boomers

Two separate effects are combining to create a surge in the size of these populations: a post WWII population boom and increased longevity. First, there was an increase in population associated with the end of WWII in many countries. This increase is particularly noticeable against a prior decrease in population growth attributed to the Great Depression. The post WWII generation is frequently referred to as the *baby boomers* and their birth coincides with the return of soldiers from the war front. Demographers typically identify this generation as those people born in the years 1946–1964 [30]. In 2011 the first of the baby boomers will turn 65 and they will continue entering the ranks of older Americans until 2029. In 2031 the first of the baby boomers will enter the age range of the oldest-old Americans. This increase in population is shown visually in figure 2.1 and is described in the following quote from the U.S. Administration on Aging:

> "The older population–persons 65 years or older–numbered 35.6 million in 2002 (the most recent year for which data are available). They represented 12.3% of the U.S. population, about one in every eight Americans. The number of older Americans increased by 3.3 million or 10.2% since 1992, compared to an increase of 13.5% for the under-65 population. However, the number of Americans aged 45-64 – who will reach 65 over the next two decades – increased by 38% during this period." [5]

Secondly, in America, and by extension, in the rest of the industrialized world, an increase in the quality of health care and in healthy lifestyles is causing the average life expectancy to increase. The average life expectancy of a child born in 2001 was 30 years longer than a child born in 1900. Even among older Americans, life expectancy is increasing. The life expectancy of a 65 year old increased 3.8 years in the 36 year period between 1965 and 2001. This is in comparison to only a 2.5 year increase in the 60 year period from 1900 to 1960 [5].

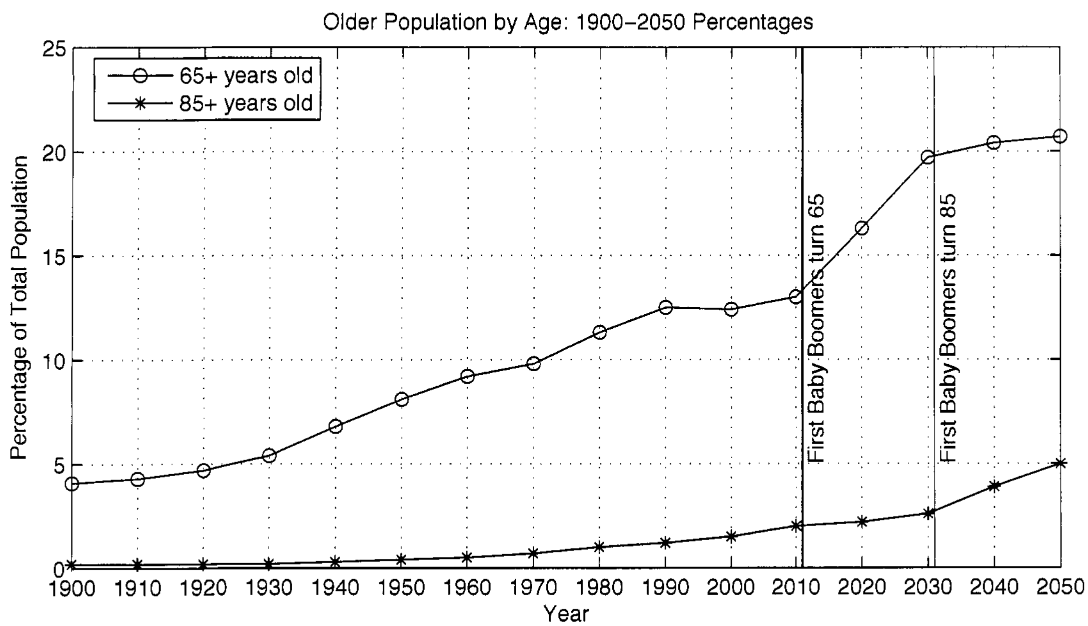


Figure 2.1: Actual and projected North American population

Actual and projected North American population by age showing the aging of the baby boomers. Source: U.S. Census Bureau 2000 census [4]

Recent studies have suggested that the increase in life expectancy may be tapering off, or even decreasing, due to the effects of increased mortality associated with obesity. It is unclear what effect this will have on demographic projections of population, but it is unlikely to completely reverse the effects of a century of improving health [115].

The overall effect of the baby boom and increased life expectancy is an increased number of older and oldest-old Americans in the next 50 years.

2.1.2 Shifting Proportions

While the absolute number of elderly are increasing, society is also going to experience a change in the relative proportion of different age ranges. Figure 2.2 shows how the age of the U.S. population has changed and will be changing in the future.

In the 19th century, the population was distributed roughly linearly. As the age range of a cohort increased, the percentage of the total population that they comprised decreased.

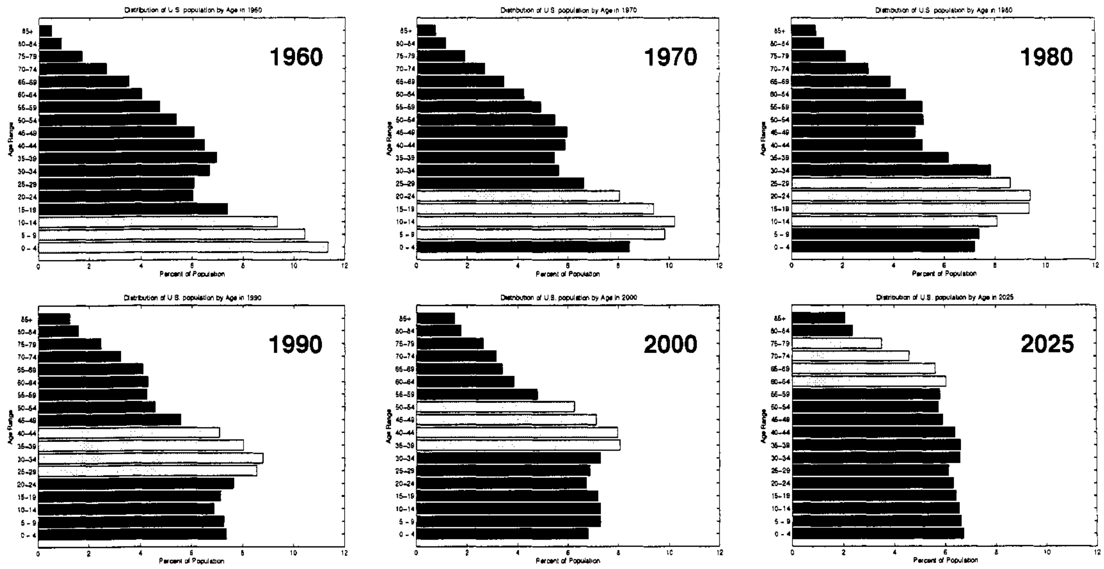


Figure 2.2: Distribution of North American population

Distribution of North American population by age bracket (baby boom generation highlighted) from 1960 to 2000 and with projections through 2025. Source: U.S. Census Bureau [69]

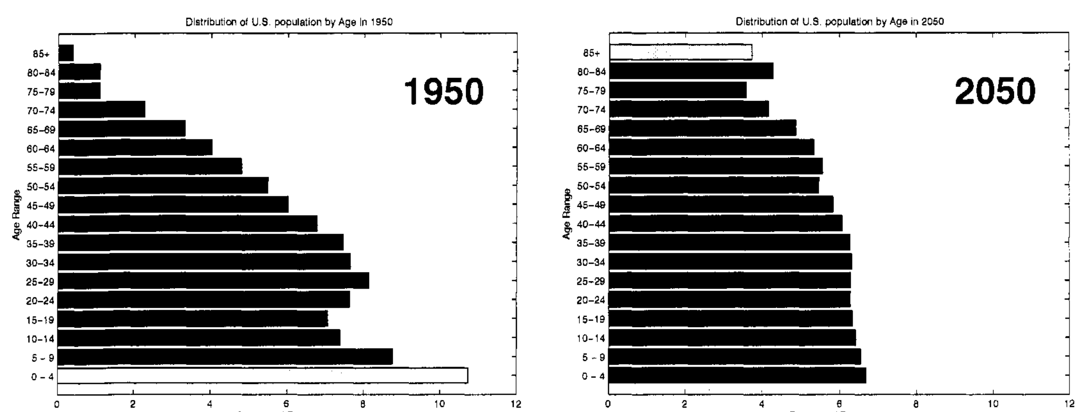


Figure 2.3: Comparison of distribution of North American population

Comparison of distribution of North American population by age bracket with the baby boom generation highlighted. 2050 data is projected. Source: U.S. Census Bureau [69]

The youngest, between ages 0 and 4, comprised approximately 10% of the population pre-1950. The oldest-old comprised approximately 0.5% of the population in 1950 (not shown in figure). In the graph depicting 1960 in figure 2.2, this general linear shape can be observed with two notable exceptions. The first is a pinching of the population coinciding with the Great Depression. The second is a bulging of the population coinciding with the end of WWII – the baby boomers.

In the subsequent years, 1970, 1980, 1990, and 2000, the startling effect of increased life expectancy can be observed. Unlike historical trends, the baby boomers are at the front of a longevity wave. Figure 2.3 shows the sharp contrast between the triangle shape that roughly characterized the 1950 population with a cylinder shape that characterizes the projected 2050 population.

The impact of these trends remains to be seen, but it suggests that social structures which rely on a greater number of young working-age individuals supporting their parent's generation need to be revisited, revised and restructured. In the realm of health care, the increase in the absolute and relative numbers of older Americans is going to challenge our ability to accommodate the larger numbers of people who are going to have predictable needs associated with aging. There is going to be a simultaneous increase in the need for services such as long-term care facilities and a decrease in the number of people available to staff them [109, 132].

2.2 *Alzheimer's Disease*

Although there are many health changes that age causes, one of the most concerning to individuals and families is the potential of suffering from or having a loved-one who suffers from dementia. This fear is not unfounded. In this section we will look in detail at the definition of dementia and AD. We will also examine the metrics by which the disease and interventions are evaluated. These metrics are important as they will naturally lead to an analysis of where technological interventions are warranted. In subsequent sections we will look more closely at the objective and subjective costs of the disease itself.

Dementia is a brain disorder that seriously affects a person's ability to carry out daily activities. The most common form of dementia among older people is AD, which causes a gradual deterioration of the brain [8]. Although it is the most common form of dementia among older people, it is not part of the normal aging process. AD is pathologically well-characterized, but unfortunately it is incurable.

> "The disease usually begins after age 60, and risk goes up with age. While younger people also may get AD, it is much less common. About 5 percent of men and women ages 65 to 74 have AD, and nearly half of those age 85 and older may have the disease." [8]

Being single, living alone, avoidance of social activities, and lack of physical activity all increase the risk for AD [123]. Because age is an independent risk factor, people in the oldest old age group are at the highest risk for AD.

The prevalence of AD in older people coupled with their rising numbers suggests a convergence of trends whose preemption society needs to take seriously.

2.2.1 Background and Pathology

AD is named after Dr. Alois Alzheimer, a German doctor who, in 1906, was the first to identify abnormal clumps (amyloid plaques) and tangled bundles of fibers (neurofibrillary tangles) in the brain of a woman who died of this unusual dementia. Since then the pathology and origins of these structures have been extensively studied and characterized. There is no single cause for Alzheimer's disease although it is known to be initiated by a molecule called beta-amyloid. The source of this molecule stems, at least partially, from one or more of a variety of combinations of genetic mutations in AD sufferers [8, 58].

> "AD is a progressive neurodegenerative and dementing disorder that can be detected clinically only in its end phase. AD is the most widespread type of dementia and affects about 10% of individuals older than 65 years and about 40% of individuals older than 80 years of age. The earliest signs of AD is a subtle decline in memory functions in a state of clear consciousness. Mental capabilities gradually worsen and personality changes appear, followed by deterioration of

> language functions, impairment of visuospatial tasks, and, in the disease's final stages, dysfunction of the motor system in the form of a hypokinetic-hypertonic syndrome." [130]

The disease highly varies in its rate of decline, but nominally takes its course over a 4 to 20 year period [124]. The progression is characterized by a *retrogenesis* of cognitive function which inversely follows the normal development of a child [122]. The retrogenesis spans the range of functioning from slight loss of memory to inability to hold one's head up. A definitive diagnosis of AD based on clinical observation is impossible and requires confirmation by postmortem examination [130].

Although there is no known cure for AD, drugs that prevent the degeneration of brain tissue have clinical value in slowing the progress of the disease. In particular, drugs that inhibit cholinesterase–the enzyme which breaks down the neurotransmitter acetylcholine have been approved for use in the U.S. and several other countries. A variety of other possible clinical treatments are also currently undergoing testing all aimed at slowing the progression of the disease [21].

2.2.2 Clinical Quantification of Alzheimer's Disease

In order to know how to effectively help people with AD, it is necessary to understand more about the progression of Alzheimer's disease, how it is evaluated and where intervention is warranted: clearly, an information technology appliance will not be of any use to an individual with severe motor system dysfunction. The literature on the clinical care of AD provides many insights into the progression of the disease and illuminates possible points at which information technology has the potential for assistance.

AD causes a gradual deterioration of the brain. Accordingly AD is measured on several progressive scales [126]. Current research identifies an additional precursor condition called *Mild Cognitive Impairment* (MCI) which is clinically detectable and is considered by some to be simply the early stages of AD [23]. MCI sufferers however present far less severe symptoms, do not exhibit the behaviors traditionally associated with AD and are able to carry on with most activities of daily living. The most frequent presentation of this clinical

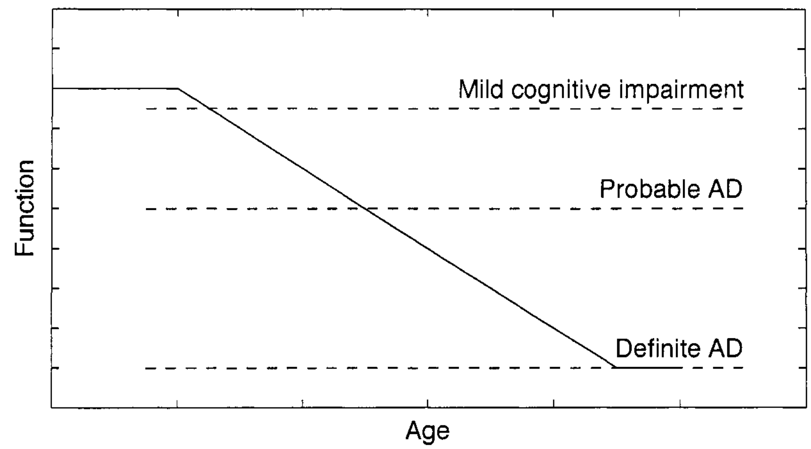


Figure 2.4: A theoretical continuum from normal aging

A theoretical continuum from normal aging through mild cognitive impairment to Alzheimer's disease [117].

condition is forgetfulness [117].

Figure 2.4 shows a qualitative characterization of the relationship between MCI and AD. Notably, this is *not* normal aging, but a pattern of progressive dementia. An individual starts with some baseline of cognitive function at which point the pathology of AD begins and cognitive functions starts to decline. Shortly after this point it is possible to clinically evaluate the individual for MCI, although the symptoms are often so slight they may not be brought to a clinician's attention at all. Sometime later a clinician can identify a patient as having probable AD and then finally, only an autopsy can conclusively identify the presence of AD.

Quantitatively AD can be characterized with several measures. The most frequently mentioned measures in the literature include the Folstein Mini-Mental Status Exam (MMSE), the General Dementia Scale (GDS), the Functional Assessment Stages (FAST), the Blessed Dementia Scale (BDS), and the Boston Naming Test (BNT), all of which are clinical, non-invasive evaluations of cognition. They consist of a combination of interview questions for the patient and their caregivers and small batteries of puzzle-type questions. Several of these cognitive measures have been shown to be comparable in longitudinal studies [121],

Table 2.1: A comparison of quantitative measures of Alzheimer's disease

A comparison of quantitative measures of Alzheimer's disease and their clinical expression [121, 117].

GDS	FAST	FMMSE	BDS	Characteristics	Clinical Diagnosis	Duration
1	1			No decrement	Normal Adult	
2	2	29	35	Subjective deficit in word finding	Normal aged adult or MCI	
3	3	25	29	Deficits noted	MCI or Onset AD	7 yr.
4	4	19	23	Requires assistance in complex tasks	Mild AD	2 yr.
5	5	14	16	Requires assistance in choosing attire	Moderate AD	18 mo.
6	6a			Requires assistance dressing	Moderately Severe AD	5 mo.
	6b			Requires assistance bathing		5 mo.
	6c	5	6	Requires assistance with toileting		5 mo.
	6d			Urinary Incontinence		4 mo.
	6e			Fecal Incontinence		10 mo.
7	7a	0	0	Speech limited to a dozen words	Severe AD	12 mo.
	7b			Speech limited to one word		18 mo.
	7c			Ambulatory ability lost		12 mo.
	7d			Ability to sit up lost		12 mo.
	7e			Ability to smile lost		18 mo.
	7f			Ability to hold up head lost		unknown

and vary primarily in their ability to accurately assess the extremes of AD, both at onset and extreme dementia. However, recognition of MCI has led to recent improvements in the sensitivity of clinical evaluations to the onset of dementia. Table 2.2.2 demonstrates the conversion between some of the scales and the functional assessment of the patient at each level.

2.2.3 Quantifying the Impact of Interventions

In 1983, Dr. M. Powell Lawton, challenged clinicians to attempt to quantify the mental health of Alzheimer's patients [87]. His remarks were based on the idea that it is impossible to effectively measure the impact of environmental changes on the elderly without an objective measure of comparison. Dr. Lawton proposed that quality of life can be represented

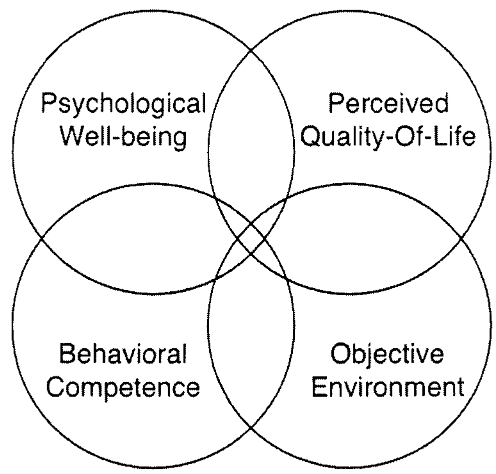


Figure 2.5: Quantifying the "good life"

Diagram from [87]

on four distinct axes: *behavioral competence*, *psychological well-being*, *perceived quality of life* and *objective environment.*

Behavioral competence is a measure of how well an individual is able to function in the world. It includes low-level functioning such as basic health on a cellular level as well as complicated social functioning. Psychological well-being is an internal evaluation of one's inner experience. It includes self-evaluations of anxiety, happiness, and the congruence between goal setting and goal achievement. Perceived quality of life is also a self-evaluation of satisfaction level in relation to friends, family, neighborhood etc. Finally, objective environment is a measure of environmental factors such as local crime rate, and neighborhood traffic which impact quality of life, but are not subjective.

Exploring these areas as guidelines will help to target interventions in an effective way.

2.2.4 Activities of Daily Living

One of the ways to evaluate behavioral competence focuses on the ability of individuals to carry out the *Activities of Daily Living* (ADLs), or the *Instrumental Activities of Daily*

Living (IADLs)[1]. ADLs include activities such as "using the toilet, eating, ambulating, dressing, bathing, and grooming,...[IADLs include] cooking, shopping, using transportation, housekeeping, doing laundry, and financial management behavior [88]." These activities represent a mid-range of complexity of behavioral competence and are more complicated than basic sensory, motor and cognitive functions. From a technological perspective they form an interesting class of tasks because as the environment and capabilities of a person become more complex, it is possible to achieve these goals in different ways. An impairment at one level, such as arthritis, does not necessarily prevent an individual from achieving the goal, as there may be many routes to the same outcome, such as opening mail, which bypass an individual's limitations. These activities are also crucial because the "satisfactory performance of such behaviors is a crucial determinant of the older person's ability to maintain community residence [88]." Or put another way, successful ADL performance keeps people from entering long-term care [13].

Lawton [88] measured the impact of impairment on IADLs, in terms of the amount of time that impaired individuals spent on them versus the amount of time that non-impaired individuals spent on them. Table 2.2 highlights the areas most impacted by impairment in terms of absolute and percentage reduction in time spent in an activity.

By assisting the Alzheimer sufferer with the activities of daily life, AC attempts to raise the level of behavioral competence of an individual. Some potentially challenging daily activities such as using public transportation, cooking and shopping may be achievable with well designed technical solutions. As a result any deployed technologies should naturally be empirically evaluated by measuring patient improvement in relevant areas of behavioral competence.

2.3 The Cost of Aging and Alzheimer's Disease

To this point we have looked at the growing numbers of elderly, the characteristics of AD and how it and interventions are quantified. Now we will briefly turn toward the question of

[1]The distinction in the literature seems to be whether the activity in question is, first, essential for independent living (both ADLs and IADLs must fit this qualification), secondly, whether they are performed daily (ADLs) or less frequently (IADLs).

Table 2.2: Mean minutes reported in selected activities in a 24-hour day from [88].

Only activities showing the largest/smallest time reduction and the largest/smallest % drop are shown.

	Independent Residents	**Impaired Residents**	**Time Reduction**	**% drop**
Housework/home maintenance	68	38	30	44.12
Shopping	22	13	9	40.91
Cooking	69	45	24	34.78
Religious activity (non-services)	10	7	3	30.00
Helping others	10	7	3	30.00
Recreation and hobbies	44	32	12	27.27
...				
Eating	77	77	0	0
Television	205	210	-5	-2.44
Non-Family Social Interaction	54	59	-5	-9.26
Radio	28	33	-5	-17.86
Personal Health Care	53	71	-18	-33.96
Rest	128	200	-72	-56.25

whether or not interventions are warranted.

Because there is no cure, people with AD typically receive some combination of home-care and institutionalized care as the disease progresses. We will look at two types of costs. The first is economic as financial resources are tapped to pay for care-giving services. The second is social as family caregivers make personal sacrifices to care for ailing loved ones and as society gradually loses the AD sufferer.

2.3.1 Economic Costs of Aging

AD is going to be one of the major economic sources of the cost of an aging population. This can seen by examining the results of several studies that have broken done such costs (*e.g.* [1, 80, 136]), and attributing the sources to symptoms of AD.

One recent study in the New England Journal of Medicine studied the source of health-care costs in the United States and was motivated by the following:

"The implications for the delivery and financing of health care will be profound,

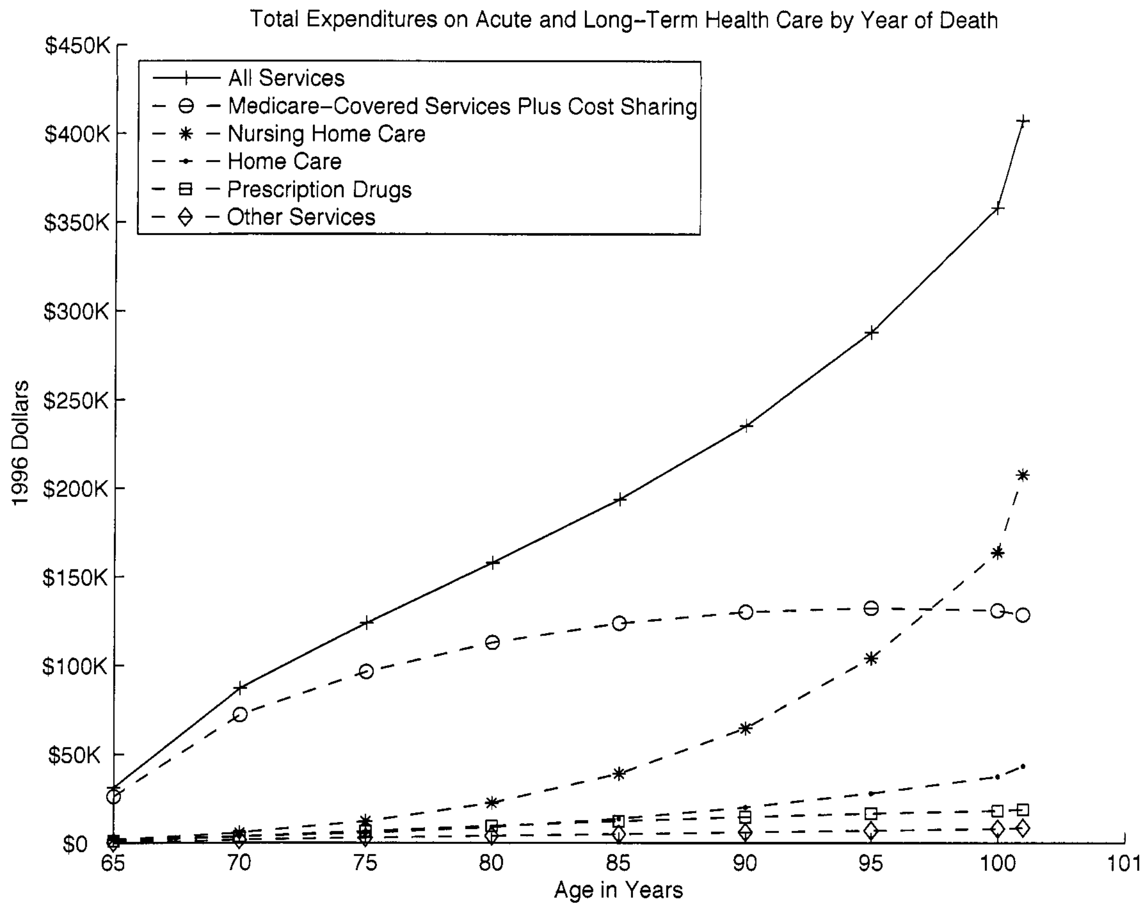


Figure 2.6: Total Expenditures on Acute and Long-Term Health Care by Year of Death

Total Expenditures on Acute and Long-Term Health Care by Year of Death , Source: Spillman, New England Journal of Medicine [136]

> because elderly persons use health care services at a greater rate than younger persons. ... Increases in the number of persons 85 years of age or older, who are most likely to require nursing home and other long-term care, will exert ... pressure on the Medicaid program, which pays for about half the total costs of nursing home care [in the U.S.]." [136]

The study factored the costs in a number of ways and concluded with several pertinent facts (see Figure 2.6). First they identified that increases in longevity increase spending on long-term care. Secondly, the cost of health care with increased longevity increases primarily due to the costs of long-term care, rather than with acute care. Finally they demonstrated that health-care spending on women is consistently higher than that for men even after adjusting for their increased life expectancy.

This last point is particularly relevant because, although the numbers of older Americans will go from approximately 1 in 8 people, to 1 in 5 people in 2030, this increase is not going to be equal across genders. Amongst the oldest-old, there are approximately twice as many women as men for all projected years [5].

The conclusion to be drawn from this study is that the cost of health care for the elderly in the future is going to strain the methods with which we are currently financing it. With current population trends and existing patterns of health care expenditures, one of the biggest impacts that can be made is to extend a person's independence and reduce their need for long-term care, especially for women.

Dementia and Functional Failure: Causes for Care-Facility Admission

We have seen that long-term care consists of the bulk of elderly people's medical bills. Dementia and an inability to carry out ADLs are strongly implicated in the need for long-term care.

There are several risk factors for institutionalization including physical disability, medical morbidity, and dementia [49, 50, 134]. This risk of institutionalization among disabled elderly persons increases further for those who become a "burden" to caregivers, even when a formal caregiver is utilized [46, 73, 86].

The results of an important seven-year longitudinal study into the causes of nursing home admission made the following important observations about the 6,676 participants in the study:

> "[There is] strong evidence that both functional impairment and dementia are important and consistent independent predictors of nursing home admission. It is noteworthy that our analyses demonstrate that the risk of institutionalization increases even with minimal decreases in function as measured by ADLs and

IADLs. The implication of this finding is that interventions with even a modest impact on preserving physical function may significantly influence the decision to enter a nursing home...

[When] analyses controlled for functional impairment and demographic characteristics, most chronic medical conditions did not significantly contribute to nursing home rates. This finding suggests that many medical diagnoses *per se* influence institutionalization almost entirely through their impact on functional status. ... We can therefore hope that as innovations are developed that effectively reduce the functional burden of chronic conditions (e.g., assistive devices), the rate of nursing home admissions may well decrease...

In contrast to the chronic medical conditions evaluated, the risk of nursing home admission significantly increased for individuals with any degree of dementia, even when adjustments for functional status were included....

Overall then, interventions to reduce institutionalization should not only focus on minimizing the progression of dementia, but also on early identification of problems and caregiver support during the earliest periods of cognitive decline." [13]

The Bottom Line

Table 2.3 shows midrange dollar estimates associated with a diagnosis of AD. The numbers were calculated in 1991 by Ernst et. al., and the dollar values were updated to equivalent 2005 dollar values.

The data demonstrate that a diagnosis of AD will cost approximately $50,000 a year. Including the opportunity cost of unpaid caregivers and premature mortality the total bill per patient is $248,723.

If the conservative assumption is made that the Alzheimer's population is steady over time (from 1991 levels) the total discounted costs of all future generations of Alzheimer's patients are in the trillions of dollars [46].

Table 2.3: Estimated Net Costs of Alzheimer's Disease per Person

Adjusted to 2005 dollars (using CPI) (Midrange Estimates) source [46]

Annual direct costs	
Diagnosis (first year only)	$2 074
Nursing Home	$10 825
Long-term mental hospital	$561
Paid home care	$4 490
Regular physician care	$333
Acute care hospitalization	$1 719
Other patient direct costs	$0
Caregiver medical care	$214
Total direct cost (first year only)	$20 216
Total direct cost (second and later years)	$18 142
Annual indirect costs of unpaid home care	$29 887
Total cost first year, excluding morbidity and mortality	$50 103
Total cost second and later years, excluding morbidity and mortality	$48 029
Total discounted direct cost	$68 040
Total discounted (direct cost + unpaid caregiver cost)	$176 685
Total discounted (direct cost + unpaid caregiver cost + disability and premature mortality cost)	$248 723

Other economic data demonstrate that the cost of caring for a patient with dementia increases with functional loss [152].

Clearly any aids that will enable people with dementia to retain functional status and out of long-term care will greatly benefit society as a whole.

2.3.2 Social Costs of Alzheimer's Disease

Economic costs for AD are overwhelming, but they aren't the only reason why addressing the disease is worthwhile. There are also high associated social costs. The gradual loss of the cognitive faculties of the patient with AD are accompanied by anxiety, agitation and violence, sleep/wake disturbances, suspiciousness/delusions/hallucinations, depression, and inappropriate sexuality [148]. If technological interventions can ameliorate these symptoms or delay their onset everyone will benefit [125].

Caregivers pay a high cost when taking care of dependents

In addition to the patient with AD, the stress on caregivers and family members is immense [95]. Caregivers make 46% more physician visits and take 71% more prescribed drugs than non-caregivers and are more likely to be hospitalized [46]. One-third live in poverty or near-poverty [7]. Two thirds of working Alzheimer caregivers reported that they missed work because of caregiving responsibilities. 14% gave up work or retired. 13% cut back hours or took a new job. 8% turned down a promotion. 7% lost job benefits [6].

> "Adverse effects on the mental health of the caregiver, especially in the form of depression, have been reported to be much higher than either age- or gender-based population norms or demographically matched non-caregiving control groups. AD caregivers reporting symptoms of depression have ranged from 18% to 55%, with female caregivers more likely to experience depression than males." [19]

So as much as technological interventions promise to assist the person suffering from AD, it may be just as useful to target interventions toward the caregivers themselves. In either case the potential for auxiliary benefits to the caregiving network of the patient with AD is also promising.

2.4 Targeting Technology

2.4.1 Objective Scale Target

The natural question then for AC is: At which level of dementia should technological intervention be directed? [105] From a technological standpoint, we will see that AC builds models of behavior that correspond to normal functioning of an individual through observation. This reduces the burden of data entry or programming from the user and, assuming the learning works, makes the technology more practical for an elderly user. This also implies that individuals should begin using AC at the very earliest stages of dementia, perhaps even before they are clinically diagnosed with something like AD or MCI. A desiderata, therefore, of AC is that it should operate in modes that provide value to a general elderly audience, so that they will be motivated to use AC before it becomes essential for independent functioning.

At the other extreme, after GDS stage 6, a great deal of the pathology of AD in particular takes a physical form and requires intricate physical interaction with the patient. Such physical interaction is outside the scope of a purely information-based AC device.

Related to the desire to motivate the elderly to use AC solutions before they need them, is a desire to create technological solutions that avoid any social stigma. This might relate to the form factor of the user-interface or to the way in which the solution is targeted. For example, using a specialized user interface would make the rest of the world aware that a user has a cognitive limitation, but using something like a PDA or cell-phone as a user interface masks the reason for its use. The user may just be making a phone call, or they may be receiving cognitive assistance.

The AC project should, therefore, focus its efforts on the subset of the dementia scale that corresponds to GDS stages 1-4. Although this is only "Mild" AD at worst, GDS stages 2-4 represent approximately nine years, or half of the nominal amount of time that a patient is afflicted with AD. Finally, since GDS stage 4 "represents the first stage in which a majority of subjects manifest the degeneration over time which is characteristic of AD[121]," unless we can effectively work with patients at this level, we have fallen short of our goal of decisively helping Alzheimer patients in particular.

2.4.2 User Interface and the Alzheimer's Patient

In any task with which AC aspires to assist, the act of interfacing with a cognitively impaired person promises to be a challenging design task. Such an interface must cope gracefully with the many complex demands of the targeted user group. Some of the basic complexities are a result of the fact that Alzheimer's patients are frequently elderly and even without AD, reduced vision, reduced hearing and reduced manual dexterity are typical. In addition AC users will have multiple cognitive deficits [67], so designing any computer system for AD must treat these limitations as central to the design process. Hopper [67] in table 2.4 summarizes the results of a study which sought to characterize Alzheimer patient functioning. Hopper also concluded several additional points which are relevant to the user interface (UI) of AC:

Table 2.4: AD Patient Performance

Performance of AD Patients (Percent Correct) on Items from the Arizona Battery for Communicative Disorders of Dementia and the Functional Linguistic Communication Inventory. Columns left blank indicate that patients at this stage were not tested either because of ceiling or floor effects on test items. Adapted from [67]

Test Item	**GDS 3**	**GDS 4**	**GDS 5**	**GDS 6**	**GDS 7**
Answer multiple choice questions		75	64	10	8
Answer two-choice questions		91	86	20	33
Answer yes/no questions		66	66	53	
Follow commands	99	93	82	44	
Follow two-step command		100	93	40	
Follow one-step command		100	100	60	25
Answer comparative questions	92	97	83	69	
Reading comprehension – words		97	81	50	
Reading comprehension – sentences	100	92	67	34	
Correct misinformation		100	84	67	39

- "Reducing demands on free recall may facilitate the expression of conceptual knowledge."

- "In the early stages of AD [GDS stage 3], comprehension and expression of grammar and syntax are relatively spared, although occasional errors may occur across modalities. Other aspects of communicative function, including auditory and reading comprehension remain areas of strength"

- "[GDS stage 3 Alzheimer's patients] were highly accurate on ... picture to word matching task, achieving a mean score of 11.4 out of a possible 13 correct."

- "Not only were IADLs such as managing finances and writing letters affected more [than the simpler ADLs], they also were among the earliest reported signs of AD. Even so, many individuals with mild AD ... were able to use the telephone, do laundry, and take medications."

- The presence of impaired speech and inconsistent grammar [67] suggests that using voice recognition to understand an Alzheimer patient will be difficult.

AC should focus its attention on the activities in which impaired individuals have seen the greatest loss of functioning in order to return their level of behavioral competence to that of independent adults. Alternatively it would be valuable if AC could assist in areas that support many of the highly impacted IADLs, such as transportation and financial management. Phelan et.al. provided evidence that ADL improvement was achieved through tools to help self-manage ADL accomplishment [118] suggesting that this is an achievable goal.

Synthesizing the guidance on UI design from table 2.4 we see that early stage Alzheimer's patients show a high degree of ability to make accurate choices and follow instructions. AC technologies should reduce the complexity of commands and choices, and open-ended questions should be avoided, but graphic icons presented to the user, simple requests for information from the user, and text displays all seem to be feasible and reasonable methods for human-computer interaction, not withstanding the typical non-cognitive deficits of the elderly. The AC UI can expect that its users will be able to carry out step by step instructions, although more complicated multi-task instructions may be beyond the capabilities of the intended users. Finally, since the ability to correct misinformation remains high in early stage AD, the prospect of learning from a patient seems promising.

2.5 In Summary, the Challenge

The growing numbers of elderly, in absolute and relative numbers, threaten to strain the social structures that have served to maintain quality of life for generations of elderly. One of the many challenges associated with the aging baby boom is accommodating the health care needs of the 50% of baby boomers which will be over age 85 in 2031.

The mental cognition and quality of life of people with MCI and AD, in particular, can be measured and evaluated. Research suggests that according to these measures the most effective interventions are aimed at improving functional status of individuals as measured through IADL and ADL performance.

As barriers to independence are lowered, quality of life improves. With independence comes decreased reliance on caregivers and long-term care facilities. This subsequently can have the effect of lowering the economic impact of aging which, research has demonstrated, is driven primarily by the cost of long-term health care.

Therefore, an ideal Assisted Cognition solution will take into consideration the caregiver as well as the patient, will promote independence, social interaction and physical well-being.

Chapter 3

ASSISTED COGNITION AS ASSISTIVE TECHNOLOGY

Assistive technology can take many forms. Traditionally it has been in the form of physical assistance as seen in wheelchairs, canes, special elevators, better ergonomics, etc. However, *Assisted Cognition* refers to a different type of assistive technology. This is assistive technology that attempts to merge sensor networks and artificial intelligence to produce systems that compensate for cognitive deficiencies in their users. The cognitive deficiencies might span a range from errors made by high functioning people to those made by people suffering from advanced dementia.

Different points on the spectrum have different challenges. A high functioning individual may make subtle errors which are hard to detect but easy to communicate to the user when discovered. Cognitively disabled individuals may make clear errors, but pose a more difficult user interface challenge in communicating the correct course of action. Additionally, as an individual suffers from greater dementia, physical assistance frequently becomes a significant component of the user's needs, such that a reasoning system will necessarily require physical augmentation in order to help.

3.1 Goals

Given these factors, Assisted Cognition potentially has the ability to be useful for a spectrum of assistance [97]:

- **Logging**: Keeping track of what happened over a given period of time. This might be as simple as instrumenting particular physical events, such as knowing when the telephone is used as a proxy for social contact, or instrumenting the front door to immediately detect wandering behavior. It might also entail complex probabilistic evaluations of combinations of occasionally faulty sensors to develop an overall picture of accomplished activities. This type of information would be useful for remote caregivers, or for self-analysis by the patient themselves. Depending on the quality and focus of the logs it may also be helpful for clinicians who are treating the patient medically.

- **Rating**: Not just knowing what has happened, but evaluating how well an activity was performed. This is more involved than simply determining if the user, for example, did the laundry. It involves a judgment of whether the correct steps were carried out, whether the right tools were employed, and if they were employed with the right sequence and timing. This may also include a fine-grained grading which details a spectrum of successful activity accomplishment. This is potentially a challenging task because as an activity is done less and less correctly it may be harder to determine what the intended activity was. At an extreme, for example, if a person walks into a kitchen and takes no more action, it would be impossible to rate the intended activity without external information.

- **Trending**: Evaluating how performance is changing over time. If a person is getting better at performing an activity then perhaps they are learning new skills. If the person's performance is degrading, then that might be an indication of cognitive decline or the onset of MCI [77]. Trending could be a complicated analysis that takes into account frequency of activity accomplishment and acceptable variations in "identical" performance of an activity.

- **Evaluation**: Reasoning about activity performance. This presumes the ability to rate an activity, but additionally includes an analysis of how the activity was substandard and what could be done to improve performance or efficiency.

- **Guidance**: Providing feedback about the results of evaluation. This might require sophisticated reasoning of which aspects of performance are most important to improve. This also implies an understanding of the best way to communicate activity deficiencies and a way to partition feedback to the user into comprehensible and mentally digestible chunks.

- **Actuation**: Performing an activity for the user. This is a complicated decision to make because on the one hand a system that removes all burden of action from the user threatens to hasten cognitive decline. On the other hand there are obvious times when a system could take positive steps to improve the immediate safety and security of the user: locking the doors at night after the user has gone to sleep, or turning off the stove if the user has left the house without doing so. The question of whether the action should be a reminder to the user or the action itself may complicate this process.

3.2 Related Work

There are a number of research groups that have made progress on developing assistive technology with an element of cognitive reasoning:

3.2.1 Location, Navigation and Way-Finding

Location sensing technologies are naturally a precursor to navigation tracking and support. Hightower and Borriello present a survey of these technologies, place them in a conceptual framework, and discuss their performance in [62]. Most location context research traces its conceptual origins to work at Xerox PARC that integrated location information with applications, typically to support information workers [3, 128, 145].

One different but popular class of location-aware applications are tour guide projects, one vision of which is put forth under the name of "Cyberguides" [2]. Several systems of this class have been attempted including Campus-Aware [27] and the GUIDE project [31].

The Place Lab initiative [127] is a project designed to make outdoor Wi-Fi localization ubiquitous through mass collaboration. Such a system would augment or replace other specialized location sensing hardware so that any location service such as those explored by the ActiveCampus project [55, 56] can be made broadly available.

Outdoor localization on highly resource constrained devices based on radio signals has been proposed and explored in the RightSPOT project [84].

Location sensing technologies can be combined in a process known as sensor fusion in order to provide more reliable or accurate position information [12, 22, 39, 65]. A single sensor can be augmented with a user model, possibly learned, to improve it's accuracy [11, 34, 59, 76, 85]. Data from multiple users can help speed the learning of user models [10]. Then sensor fusion, and user modeling can be combined with sophisticated inference technologies and external knowledge to achieve even better results [64, 93, 131].

Some of these inference technologies include, the abstract hidden Markov model which uses hierarchical representations. These have been demonstrated to efficiently infer a person's goal in an indoor environment from camera information [25]. These models have been, extended to include memory nodes, which enables the transfer of context information over multiple time steps [24]. Complementary to this work, Liao et.al. [91], presented a discriminative model to automatically classify significant places and activities based on the framework of relational probabilistic models [52, 141].

Nursebot

Nursebot [29] is a project at Carnegie Mellon University that provides a robotic platform for delivering navigation assistance to the elderly. This robot is envisioned to operate under the auspices of a community living home and helps users make it to appointments on time and provides directions to get there as well. Nursebot has the advantage of being able to provide an element of physical assistance as well since it is embodied in a physical artifact.

Nursebot requires caregiver support to update its understanding of the world, for example, the remote location of people and the state of their schedules. Since it is a robotic platform it only makes use of sensors embedded on the robot, although, in principle extending the robot's knowledge to a network of distributed sensors would be possible. As with many robotic platforms it also suffers from a short working time due to short battery life.

Nursebot's navigation is based on robotic mapping technology and laser range finders. Potential destinations are identified on a known map. When an elderly individual indicates a potential destination, Nursebot plans a route to the location and executes the plan. The plan is updated according to real-time laser range finder inputs which help to regularly update the position of the robot and the robot's knowledge of the people in the environment.

IMP

Closely related to Nursebot is IMP [104], a walker that has been augmented with a laser range finder and navigational reasoning. It operates using a map that has been built using SLAM technology [102]. When a user wants to go somewhere in the mapped facility, they can indicate their destination on an attached computer and a path-planning algorithm will guide the user to the destination using an arrow. On board sensors monitor progress and assist the user in getting to their destination.

One of the design decisions that this system made was to navigate a person directly through the use of a displayed arrow. This puts a high burden on the navigational system and sensor suite to avoid leading its users into dangerous environments that can not be sensed by the walker. This is probably not a large concern for controlled environments such as nursing homes, but would be a problem for transportation assistance outdoors.

3.2.2 ADL Tracking and Support

There are many research groups that are applying mobile and ubiquitous computing to the goal of aging-in-place.

Helal and Mann at the University of Florida's Mobile and Pervasive Computing Laboratory have several projects that are directed at cognitive assistance for the elderly, including meal preparation assistance [18] and preliminary research into using a cell-phone as a cognitive assistance [28, 60].

Jimison and Pavel, at the Oregon Health Sciences University, have done work on using computer-based tasks to identify changes in the cognitive status of elders [74]. Such a system would support early recognition of MCI or other health problems that might impact cognition.

Wearable Activity Recognition

There has been much recent work in wearable computing that has been along the same lines as BARISTA, namely justifying the value of various sensing modalities and inference techniques for interpreting various aspects of context. This work ranges from unimodal evaluations of activity recognition and social context from audio [32, 137, 138] , video [48], and accelerometers [75], to multi-modal sensing which included these and other sensors [81] and some which has attempted to optimize sensor selection for arbitrary activity recognition [33]. A theme of much of this recent work is that "heavyweight" sensors such as machine vision can be replaced by large numbers of tiny, robust, easily worn sensors [66, 139].

Autominder

Autominder [98] is a planning assistance system that is designed to help a user meet their scheduling goals in the presence of conflicting or overlapping goals. Autominder was originally part of the Nursebot project and supported the process of helping Nursebot's users identify where they needed and wanted to be. One of the key aspects of Autominder's plan reminder system is a balance that is struck between trying to alert a user to the potential for missing an appointment versus not being overly aggressive in its reminding behavior.

Similar to Nursebot and IMP one of the primary limitations of this system is that it doesn't have knowledge of the physical environment and requires a user or caretaker to input information about the plan state in the physical world. (*e.g.* such as when a medication has been taken, or the user has completed using the bathroom)

House_n

Stephen Intille at the MIT House_n project has focused on three aspects of cognitive assistance.

The first is the development of an instrumented condominium for studying activity recognition in a naturalistic environment, called the PlaceLab [71]. This space is designed as a living laboratory from which experiments in activity recognition can be conducted in a naturalistic manner. From this line of research a variety of activity recognition research has been published which explore how simple sensors can be used to recognize activities as they are being performed [14][140]. This work has focused primarily on making activity logging a practical success and has focused on accelerometers and simple switches as the environmental sensors.

The second research direction are efforts to provide "just-in-time" guidance for people to make decisions that will have a positive impact on their health [70]. It presumes that a context sensitive computer system can identify some activity of a user and give them feedback quickly enough that they can adjust their activity to a more healthy one. The first work along these lines has focused on very simple state recognition system that determines whether people are taking stairs versus an adjacent escalator and projects messages designed to encourage taking the stairs.

The final research direction is the development of tools that support researchers in determining what activity an individual is actually engaged in. Such work is essential to evaluating activity recognition systems. This style of sampling what activity a user is engaged in is called context-sensitive experience sampling because it uses limited sensors to determine when targeted activities might be underway and then uses those cues to ask the user what activity he or she is engaged in.

COACH

Another activity recognition system that addresses logging, rating, evaluation, and guidance is the COACH system (Cognitive Orthosis for Assisting aCtivities in the Home). This system has been developed my Alex Mihailidis at the University of Toronto to assist people with advanced dementia with hand washing.

COACH is an adaptive device which learns using Markov decision processes for how best to guide a user through the process of washing hands. The sensor input for this task is an overhead camera located over the sink which is processed, primarily for the location of relevant objects, and then becomes the source of information for the decision process.

This work is noteworthy, in particular for its use of verbal prompts that increase in specificity as the user becomes less and less likely to achieve the goal of hand washing. Of all the systems mentioned here it targets individuals with the most severe forms of dementia.

Although the system promises to be extensible to other activities, at this time it only recognizes and prompts the hand-washing activity [99].

STAR

Researchers at Carnegie Mellon University have been developing logging applications that, like Intille's work, focus on using simple sensors to identify the location and future motion of people in an apartment. Using an apartment outfitted with 49 simple sensors and a trained motion model they were able to determine the location of three people in the apartment correctly between 85% and 98% of the time. [151]

Intel Research Seattle

Intel Research Seattle has been working with a multi-sensor board that provides many more sensors than researchers have typically evaluated up to this point. This worn sensor has a video camera, audio sensor, light sensor, barometer, accelerometer, and digital compass. The focus of this work has been on determining which of the many hundreds of features that these sensors provide give the most information regarding the activities that the wearer is undertaking.[33]

They are also close collaborators with the RFID based ADL recognition systems described in this thesis.

Digital Family Portraits

A different class of work has focused on taking streams of sensor information, identifying the activities that are currently underway and attempts to digest them for use in communicating the status of an elder. This work assumes the capability of logging and is studying how to provide appropriate guidance. In this case the guidance isn't to the user of the system, but rather it is focused on the care givers of the observed individual. The focus of this type of work so far has been on user-interface and design.

Researchers at Georgia Tech have developed the concept of the digital family portrait that is designed to compensate for the casual day-to-day contact that remote family members are unable to experience with an elderly family member who in some way relies on the remote relative for care.

This device takes the form factor of a picture frame with a picture of the individual that is being monitored in the center. Around the outside are icons that aggregate aspects of activity for a given day and present an ambient sense of how well the day is going for the senior [108].

Computer Supported Cooperative Care

Intel Research Seattle [36] has taken this concept and extended it to include more depictions of sensing information with a focus on provided user-centered design techniques. Rather than just displaying aggregate information about a person's day, this system allows the remote user to drill down on any given day to find specific information about the source of the aggregation, or any messages the senior might want to provide to their caregiver. All of this is done under the auspices of privacy policies that are controlled by the senior.

3.2.3 Commercial Cognitive Support

PEAT

PEAT [90, 114] is a commercial product that has many of the same goals as Autominder. It is built on a PDA platform and it's goal is to help individuals who experience difficulty formulating and following a plan. Like Autominder, it is more than a calendaring system as it actually schedules, alerts and follows up with a user's execution, although the rescheduling is somewhat less sophisticated than Autominder.

PEAT also requires users to enter data about their schedule and continuously update their schedule with performance information. It also has no sense of a user's context other than what the user has provided through the PDA interface.

Bath Institute for Medical Engineering

This group has developed a number of prototype commercial applications of cognitive devices and collaborates with Dementia Voice, a dementia services center for the southwest of England, and Housing 21, a UK housing association [18].

Some of their projects include:

- **Cooker Monitor**: An instrumented stove that monitors for dangerous situations such as gas leaks, smoke, or burning pans. This stove reacts to dangerous situations by shutting down the gas flow and sending a text message alert.

- **Locator**: A misplaced object finder.

- **NightLight**: A light that turns on automatically when an individual gets up. This is triggered by weight in the bed rather than motion.

- **Tap Monitor**: An instrumented faucet that controls temperature and prevents flooding.

- **Time Orientation**: A programmable display designed for displaying textual and image-based reminders.

Wandering Alert and Monitoring Systems

A number of companies have attempted, with varying success, to create wander alert systems. Some have been stand alone systems such as Digital Angel [42] and others have been integrated into smart assisted-living environments such as with Elite-Care assisted living homes [45]. The company Independent Living has an installable system that promises to monitor both ADLs and wandering and alert a caregiver when programmed parameters are exceeded [94].

3.3 Summary

In summary, there are a number of different ways in which assistive technology can assist users and the user's care networks. They span the range of assistance from logging activities to actually performing activities for the individual. Many research groups are making advances on assistive technology that, like AC , include elements of cognitive support for their users. As the user base for cognitive assistance grows, more of these research efforts will likely find their way out of the lab and into commercial consumer-grade technology.

Chapter 4

ASSISTING OUTDOOR NAVIGATION

A central theme in ubiquitous computing is building rich predictive models of human behavior from low-level sensor data. One strand of such work concerns tracking and predicting a person's movements in outdoor settings using GPS [11, 22, 39, 62]. But movement and location are only a small part of a person's state. Ideally we would recognize and predict the high-level cognitive intentions and complex behaviors that cause particular physical movements through space. Such higher-order models would both enable the creation of new computing services that autonomously respond to a person's unspoken needs, and support much more accurate predictions about future behavior at all levels of abstraction and beyond just physical location.

4.1 The Case for Outdoor Navigation

We have seen in previous chapters that successful independent living requires an individual to be able to complete IADLs with minimal assistance and prompting (e.g., [13]). Any IADL that takes place outside the home has, as a prerequisite, the ability to successfully navigate to some location. Shopping, going to a doctor's office, attending social events such as church or clubs all presume that an individual can get somewhere. Mobility is key to leading an independent life.

For the elderly or people with cognitive disabilities, driving themselves is frequently not an option. This might be because they have additional challenges such as poor vision, or other uncompensated physical disabilities. They might not be able to drive because they are unable to afford a vehicle, or the scope of their cognitive abilities might preclude safe driving. For many of these individuals, mobility in the community means using public transportation. It is key to their social life, their employment, and their ability to receive goods and services.

Public transportation can also have a substantial cognitive load associated with it, particularly for routes that are taken on an irregular schedule or require navigating a multi-modal public transportation infrastructure. This type of trip may be too daunting, too dangerous or too complicated for someone with a cognitive deficiency.

There is often no choice in these situations, but for them to give up their potential future independence and be under direct supervision of their care givers or family members; a healthy individual is needed to detect situations where a mistake made by a cognitively disabled person may cause distress or harm.

Thus, the inability to safely use public transportation harms their quality of life as well as that of their formal and informal support network [36, 87, 88]. However, if impaired individuals had effective compensatory cognitive aids to help them use public transportation, their independence and safety would improve, they would have new opportunities for socialization and employment, and stress on their families and care givers would be reduced.

Fortunately, our public transportation infrastructure is quickly becoming one of the best sources of sensor data information. Many municipalities provide real time information about the location of buses [83, 110]. Online trip planners, bus schedules, ferry schedules, maps and fare information are increasingly available. These sources are rapidly evolving from a simple static reference for people to a dynamic machine-readable source of information. Once the information is machine-readable, new opportunities to integrate the data into complicated automatic reasoning systems become possible. The work that is described in this and the next chapter fits this category.

Because of the cross-cutting value which transportation provides for IADL completion, and because advanced sensor systems and real-time information is increasingly available for these infrastructures, outdoor public transportation navigation is an excellent target for an Assisted Cognition solution.

4.1.1 More Than a Route Planner

This idea is substantially different than the direction services that are currently available on the web [54, 96]. Those services provide driving directions from one location to another location. The GPS enhanced analog which is available in car navigation systems, monitors a users progress on such a route and provides information on upcoming changes in direction. ("Turn right on 1st Ave").

In contrast to these services, the solution we are proposing is user-centric, not vehicle-centric. It does not travel with a car. It is designed to reside on and stay with a person. This implies that the device has to function across multiple modalities of transportation including walking, driving and using buses and other public transportation services.

Additionally, because our target audience is cognitively disabled, this solution should not require a user to program a starting or ending destination. The starting destination is always available because the system is aware of the user's current location through a position sensor and rather than programming a destination address, the user selects one of a small number of possible final destinations from an iconic list. These destinations are learned and never require a user to use the portable device to enter address information.

4.2 Design Challenges

Although our proposed solution would have value to a particular audience, there are some design decisions that could doom such a system to failure that are worth discussing.

4.2.1 Accuracy of Positioning Systems

Although there are a number of different positioning technologies, which are available for both indoor and outdoor use [61], the only truly ubiquitous positioning technology that is currently deployed is GPS. GPS provides accuracy on the order of 5m 95% of the time [61]. While this is excellent for positioning an oil tanker that is in the middle of the Pacific Ocean, it is not sufficient resolution to determine which side of the street an individual is currently standing on. Furthermore, GPS receivers experience substantial warm-up latency when transiting from an indoor to an outdoor environment while the electronics lock on to available satellite signals. Although more accurate GPS-based systems exist, they are not sized practically for a device that is regularly carried by a user.

The combination of coarse granularity (at the scale of a user) and latency makes some approaches to navigation untenable.

4.2.2 Following Arrows Considered Harmful

Our early attempts at solving this problem demonstrated that a navigation aid that attempts to explicitly route an individual with the use of an arrow is fraught with problems. This is the technique that some related work in the literature suggests for routing a user [104], but we believe that this will ultimately fail outside the most controlled environments.

Using an arrow to guide an individual is a dangerous proposition. It will not take much time in the real world before such an arrow will point someone into the middle of a busy intersection or off the edge of a steep incline. Using the arrow as an interface removes too much reasoning from the user. It would be easy for one to follow the arrow, have one's mind wander, forget about the surroundings and meander into harm.

Fundamentally an arrow-based solution assumes two things that are unlikely to be true anytime in the near future: highly accurate positioning systems and comprehensive knowledge bases.

4.2.3 Comprehensive Knowledge

Even with a perfectly accurate positioning system, a prototypical arrow-based guidance system is still not well-founded. Producing directional guidance will have to be based on some knowledge of the environment. There will never be sufficient information to generate such directional guidance.

Simply having a map is not good enough because the device can not accept responsibility for making bad decisions based on inaccurate or incomplete map data. A perfect and complete map coupled with a perfect positioning system still won't address challenges like timing an individual to cross the street when the light is green or, further, avoiding a car that is inadvertently ignoring the corresponding red light. Similar problems arise in situations such as trying to reason about a locked door. An arrow-based navigation system would have to have information about the user's access to a key.

These are all challenges related to trying to provide situational reasoning assistance which may be fundamentally insurmountable. The solution is to present relevant information to the user in a way that acknowledges the uncertainty in the inference and requires active participation from the user. Rather than pointing the user across a busy intersection, presenting a command to "Cross 1st Ave when safe" keeps the onus of safety on the user. Rather than worrying about whether a user has a key, presenting a route to the user and reminding them that they need a key to get through a relevant door would solve the problem. Rather than reasoning about the presence of bus fare, telling the user the cost of executing the public transportation plan that the system has calculated for them eliminates the need for hard-to-obtain knowledge.

4.3 The Activity Compass Vision

In the next chapter we will focus on the technical details of learning how a person uses different kinds of transportation in the community. We use GPS data to infer and predict a user's transportation *mode*, such as walking, driving, or taking a bus. The learned model can predict mode transitions, such as boarding a bus at one location and disembarking at another. We will show that the use of such a higher-level transportation model can also increase the accuracy of location prediction, which is important in order to handle GPS signal loss or preparing for future physical delivery of services.

A key to inferring high-level behavior is fusing a user's historic sensor data with general commonsense knowledge of real-world constraints. Real-world constraints include, for example, that buses only take passengers on or off at bus stops, that cars are left in parking lots, and that cars and buses can only travel on streets, *etc.*. We will incrementally present a unified probabilistic framework that accounts for both sensor error (in the case of GPS, loss of signal, triangulation error, or multi-path propagation error) and commonsense rules.

Although this work has broad applications to ubiquitous computing systems, we will call our motivating application the Activity Compass. This is a device that helps guide a cognitively impaired person safely through the community [112]. The system notes when the user departs from a familiar routine (for example, gets on the wrong bus) and provides proactive alerts or calls for assistance. The Activity Compass and it's successor, "Opportunity Knocks" (OK) [1] are examples of Assisted Cognition solutions which use probabilistic models of human behavior to help users [78].

This system holds promise for a wide range of individuals. The ultimate target is people with Alzheimer's disease, but mentally retarded individuals and individuals with traumatic brain injury are also potential beneficiaries of this system. The latter two groups serve as a valuable transitional model, since they may be unable to use public transportation due to short-term confusion or memory lapses, but generally show stable levels of cognitive ability over time. They frequently are employed, and are either using specialized transportation services or using public transportation with marginal efficacy. Ultimately we hope to be able to help even high functioning people who inevitably make mistakes.

The follow-on name to the Activity Compass, "Opportunity Knocks," is derived from the desire to provide our users with a source of computer generated opportunities that remind them of previous transportation routes and correct simple errors before they become dangerous errors. When the system has determined that an especially important opportunity has made itself available, it plays a sound like a door knocking to get the user's attention. Less critical opportunities are simply displayed if the user expresses interest. We desire to support existing cognitive capacities, not replace them, by helping users to remain engaged in their transportation decisions. This ameliorates the risk of hastening cognitive decline and removes some liability in the case of device failure.

Although in this thesis we focus on a system that assists cognitively impaired people, the techniques we present can be applied to any user-centric location-based service that would benefit from probabilistically predicted location information (e.g., just-in-time traffic information for specific routes, home climate and appliance control, or reminders for errands-of-convenience).

[1] In other publications the Activity Compass refers to work that does not take destination into account when reasoning. OK reasons with destinations in mind. In this thesis, the distinction is less dramatic.

4.4 Scenario: Eileen goes to Ted's House

In order to ground our system, we present a fictitious running example that will help illustrate the most important features of our system:

Eileen has a physical therapist at a nearby university campus, whom she visits on a bi-weekly basis. After one such visit, Eileen finds herself exiting the building uncertain of which way to proceed. After a few minutes of hesitation, she reaches for her phone and invokes OK. OK offers images of three destinations that she typically travels to after the therapist visit: her home, a grocery store, and the house of her friend Ted. Eileen selects her home and the system suggests her typical route: it provides instructions to find the nearest bus stop and tells her to wait for bus number 372.

Bus number 68 arrives first. Since this is the bus that Eileen normally takes to the *grocery store,* she accidentally boards it instead. Its route initially coincides with that of number 372; while OK can identify that she is on *a* bus, it is unable to detect the *identity* of the bus. It remains silent as it observes that Eileen is moving toward home in the expected manner. When the bus suddenly turns west after some time, following the bus route to the grocery store, her phone makes a knocking sound and alerts her that she should get off at the next stop. At that point, it directs her back a few hundred feet to a bus stop where she can board the next 372 bus. This time she gets on a correct bus and arrives home safely.

4.5 System Architecture

In order to support Eileen in the way we describe in the previous scenario, several technical pieces have to be composed. First, we describe the overall architecture of the system before discussing the individual components in detail.

Figure 4.1 diagrams our overall system architecture. The data flow of our system starts at a sensor beacon that is carried by a user. The sensor samples the environmental context of the user and forwards this information over a secure Bluetooth connection to the cell phone. The cell phone initially acts as a network access point and again forwards the context information to a remote server over the high-speed GPRS data network. The remote server, which is running the OK software, uses the sensor information in conjunction with

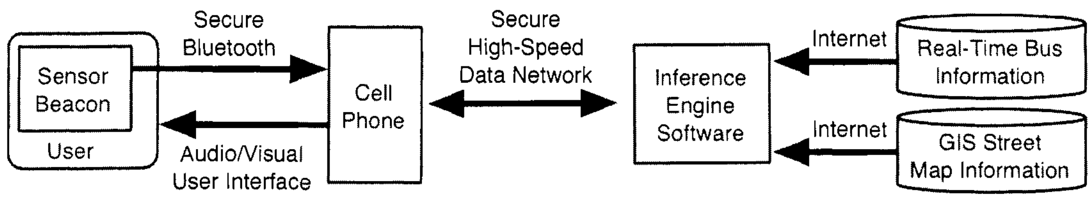


Figure 4.1: Architectural Diagram of Opportunity Knocks.

Geographic Information Systems' (GIS) databases to localize the user. When the software has sufficient confidence in the position of the user, it is then able to suggest opportunities about which the user may want to know. These opportunities are sent back to the cell phone for display through the user interface. If an urgent opportunity, such as a plan for recovering from boarding the wrong bus, is recognized, the phone proactively alerts by making a door-knock sound; otherwise the phone remains passive with information available for reference by the user. If the user selects an opportunity, such as a route to a frequent destination, the cell phone requests supporting information from the server, which may require referencing real-time information about bus schedules.

4.5.1 Cell Phone

We chose a cell phone as our client hardware because of its role as a defacto standard for a portable computing device. It has inherent value that is related to its primary function as a phone and for many people it is as common to carry as a wallet or a purse. As a result, it is likely to be a familiar, non-stigmatizing method of delivering assistive services. In the cell phone market there is a spectrum of products available, which spans from a traditional phone on one end, to a Personal Digital Assistant (PDA) on the other end. We opted for devices that were more like traditional phones rather than "smart-phones" because of their ubiquity, simple interface, and limited maintenance requirements. Cell phones also offer the promise of a cross-platform development environment which would enable an application written for the J2ME (Java 2 Micro-Edition) platform to work on any compliant phone.

Our system currently uses a Nokia 6600 cell phone. The Nokia 6600 phone is a GSM phone that has a wide-range of features required by our system. First it supports the J2ME Mobile Information Device Profile (MIDP) 2.0 that provides support for secure networking, serial port connection support, and the Application Management System — a push registry that enables authorized applications to be launched remotely. Some model specific features of the phone that we utilize include a high-resolution (176 x 208 pixels), high color (16-bit) screen, a digital camera, Bluetooth support, and high-speed data network capabilities (GPRS). Under continuous operation our system lasts approximately 4 hours.

4.5.2 Sensor Beacon

The sensor beacon, which our users are required to carry, is a physically separate unit from the phone. We intend for users to place the sensor beacon in a purse, on their belt, in a backpack or on a wheelchair while transiting. In the future, it appears imminent that at least a simple GPS sensor will be incorporated into the phone itself [47] eliminating the need for a separate sensor beacon.

Currently, however, OK utilizes two different beacon implementations, both of which broadcast exclusively GPS information. One is a commercial package available from Socket Communications Inc. [135], shown at the top of Figure 4.2. This device measures 50x90x16mm, and contains a rechargeable six-hour battery.

The second is a custom-made device shown at the bottom of Figure 4.2 that utilizes a Bluetooth serial profile broadcaster and connects to an ATmega 128 processor. The ATmega processor functions as a communication gateway, controlling the multiplexing of several sensors, packaging of the data and sending it to the cell phone via Bluetooth. Our custom system will enable prototyping new sensors (e.g., digital compass, accelerometer, Wi-Fi localizer) in response to new research and our user studies.

4.5.3 Alternative Design Strategies

This system architecture is based on distributed design. This enables specialized devices to be connected in a way that leverages each of their strengths: The GPS beacon provides location, the phone provides communication, and the desktop provides computing horsepower.

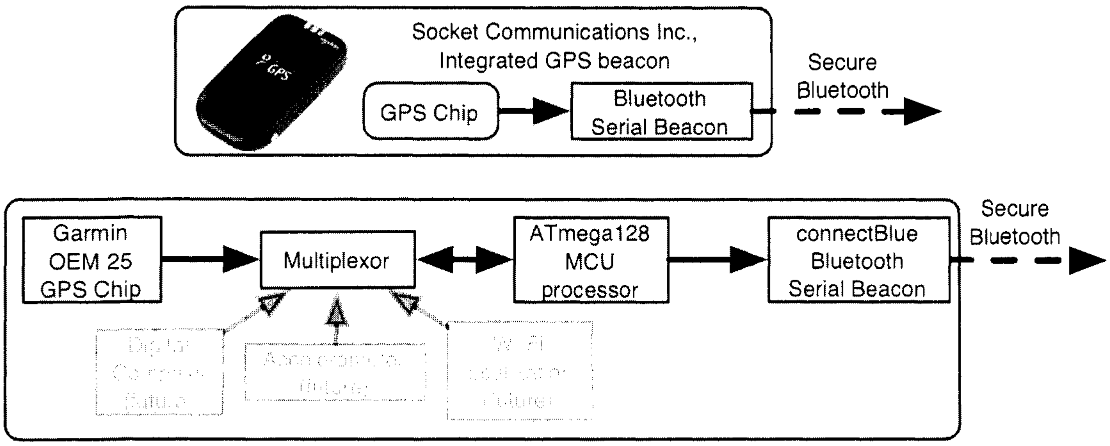


Figure 4.2: Opportunity Knocks Sensor Beacon Designs

Top: picture and schematic of off-the-shelf sensor beacon. Bottom: schematic of custom-built sensor beacon.

The drawback of such an architecture, however, is that it is prone to many single points of failure. If the GPS unit loses synchronization with the satellites, position estimation begins to degrade. The Bluetooth implementations for most platforms are not designed to be fault tolerant - once a connection is lost the entire application has to be restarted just to reinitialize the Bluetooth connection. If the cell-phone loses connectivity a similar problem occurs. Many of these problems would be avoided if all of the components were built into one device. Unfortunately such a device is not available today.

4.5.4 User Interface Concept

Based on exploratory interviews with members of our target community, we have focused on a simple user experience. When the user desires transportation assistance, she refers to her phone and observes up to four images of predicted destinations (in Section 5.5 we describe how this selection is made). If she would like to go to one, she selects it. If the system has observed the user going to this destination in different ways, for example by foot and by bus, it will prompt her for the method she would currently like to take. The previously observed route is then provided in text form.

The system will not present destinations to which the user hasn't previously traveled, but it will allow the user to select a familiar destination even if it has never observed the user getting there from the current location. In this case OK presents a route that is based on a real-time bus route planning service provided by the local transit authority (e.g., [82, 146]).

Notably in the course of this interaction the user didn't have to provide any information about where she was, and only a very small amount of information about where she wanted to go, yet the system was still able to route her effectively.

There are two occasions in which the phone might become proactive and make a knocking sound. The first is when the system has high confidence that a novel or erroneous event has occurred (independently tunable). The second is when the system identifies that you are at a new significant location. We discuss the details of the second alert in the next section.

Position to Place

Our system interfaces with the user by suggesting destinations that it has high confidence she is heading toward, and then routing her to that destination. It would be insufficient to present destinations as GPS latitude/longitude positions, and infeasible to require the user to enter a description for every interesting position on a cell-phone keypad.

Ideally, we would like to produce place descriptions automatically. This, however, is recognized as a difficult open problem [38, 63]. When attempting to create a meaningful label for a place, it is clear that the purpose of the labeling and the perspective of the labeler quickly dominate the proposed ontology. Should a description of place focus on demographics, land use, administrative use, functional use, or personal memories of the place? What happens when multiple ontologies define a region in different ways, or don't even separate the region in the same way? And which way is the best way for the current user?

To solve this problem we investigated a novel use of the camera phone. Since our system is monitoring the user's location, it is able to recognize when she has spent a sufficient amount of time in a location to call it significant. When this condition is met, the camera phone alerts the user to take a picture that captures her location. In the future, whenever the system wants to refer to that location, rather than trying to call it something in particular, it simply uses the photo to identify the spot. The advantage of such a system is that the user can decide what is meaningful about their location and can take a picture which reflects that.

4.6 Summary

In this chapter we have described an argument for a valuable Assisted Cognition solution. It addressed the problem of public transportation usage for the cognitively disabled as an enabling tool for the completion of many IADLs. We have presented our design decisions and the rationale behind them and presented a distributed system architecture as an implementation. Our design uses commercial off-the-shelf technology that, at least as a prototype, can meet the needs of our proposed transportation scenarios.

Chapter 5

MODELING OUTDOOR ACTIVITIES

5.1 Modeling User Transportation

Given the design goals and perspectives introduced in the previous chapter, we now proceed to develop a system to model user transportation, first for inference and eventually for prediction and assistance.

Our approach is built on recent successes in particle filters, a variant of Bayes filters for estimating the state of a dynamic system [44]. In particular we will show how the idea of graph-constrained particle filtering introduced in [93] can be used to integrate information from street maps. Our extensions to this technique include rich user transportation state models and the inclusion of multiple kinds of commonsense background knowledge. We introduce a three-part model in which a low-level filter continuously corrects systematic sensor error, a particle filter uses a switching state-space model for different transportation modes (and further for different velocity bands within a transportation mode), and a street map guides the particles through the high-level transition model of the graph structure. We will additionally show how to apply Expectation-Maximization (EM) to learn typical transportation patterns of a user in a completely unsupervised manner. The transition probabilities learned from real data significantly increase the model's predictive quality and robustness to loss of GPS signal.

5.1.1 Tracking on a Graph

Our approach tracks a person's location and mode of transportation using street maps such as the ones being used for route planning and GPS-based car tracking. We make the assumption that whenever a person is transiting, they are traveling on or next to a mapped road or path. More specifically, our model of the world is a graph $G = (V, E)$ that has a set,

V, of vertices and a set, E, of directed edges. Edges correspond to straight sections of roads and foot paths, and vertices are placed in the graph to represent either an intersection, or to accurately model a curved road as a set of short straight edges. To estimate the location and transportation mode of a person we apply Bayes filters, a probabilistic approach for estimating the state of a dynamic system from noisy sensor data. We will now briefly describe Bayes filters in the general case, show how to project the different quantities of the Bayes filter onto the structure represented in a graph, and then discuss our extensions to the state space model.

Bayesian Filtering on a Graph

Bayes filters address the problem of estimating the state, x_t, of a dynamic system from sensor measurements. Uncertainty is handled by representing all quantities involved in the estimation process using random variables. The key idea of Bayes filters is to recursively estimate the posterior probability density over the state space conditioned on the data collected so far. The data consists of a sequence of observations, $z_{1:t}$, and the posterior over the state x_t at time t is computed from the previous state, x_{t-1}, using the following update rule (see [16, 44] for details):

$$p(x_t \,|\, z_{1:t}) \propto p(z_t \,|\, x_t) \int p(x_t \,|\, x_{t-1}) \; p(x_{t-1} \,|\, z_{1:t-1}) dx_{t-1} \tag{5.1}$$

The term $p(x_t \mid x_{t-1})$ is a probabilistic *motion model* of the object dynamics (the user's transportation motion in this case), and $p(z_t \mid x_t)$ describes the likelihood of making observation z_t given the location x_t, called the *sensor model.*

In the context of location estimation, the state, x_t, typically describes the position and velocity of the object in 2D-space. When applying Bayesian filtering to a graph, the state of an object becomes a triple $x_t = \langle e, d, v \rangle$, where $e \in E$ denotes on which edge the object resides, d indicates the distance of the object from the start vertex of edge e, and v indicates the velocity along the edge [93]. The motion model $p(x_t \,|\, x_{t-1})$ considers that the objects are constrained to motion on the graph and may either travel along an edge, or, at the endpoint of the edge, switch to a neighboring edge. To compute the probability of motion from one edge to another, the graph is annotated with transition probabilities, $p(e_j \mid e_i)$, that describe the probability that the object transits to edge, e_j, given that the previous edge was e_i and an edge transition took place. Without other knowledge, this probability is a uniform distribution over all neighboring edges of e_i.

Our work builds on graph-based Bayesian tracking by hierarchically extending the state model. We add a higher level of abstraction that contains the transportation information and a lower level sensor error variable. The resulting state, x_t, consists of the variables shown in Fig. 5.1. The presence of a bus stop near the person is given by the binary variable b_t, and the presence of a parking lot is modeled by p_t. The mode of transportation, denoted m_t, can take on one of three different values

$$m_t \in \{BUS, FOOT, CAR\}.$$

v_t denotes the motion velocity, and the location of the person at time t is represented by $l_t = \langle e, d \rangle$. o_t denotes the expected sensor error, which in our current model compensates for systematic GPS offsets. Finally, at the lowest level of the model, raw GPS sensor measurements are represented by gps_t.

Tracking such a combined state space can be computationally demanding. Fortunately, Bayes filters can make use of the independences between the different parts of the tracking problem. Such independences are typically displayed in a graphical model like Fig. 5.1. A dynamic Bayes net [40, 106], such as this one, consists of a set of variables for each time point t, where an arc from one variable to another indicates a causal influence. Although all of the links are equivalent in their causality, Fig. 5.1 represents causality through time with dashed arrows. In an abstract sense the network can be as large as the maximum value of t (perhaps infinite), but if one desires to do online inferencing under the assumptions that the dependencies between variables do not change over time, and that the state space conforms to the first-order Markov independence assumption, it is only necessary to represent and reason about two time slices at a time. In the figure the slices are numbered $t-1$ and t. The variables labeled gps are directly observable, and represent the position and velocity readings from the GPS sensor (where a possible value for the reading includes "loss of

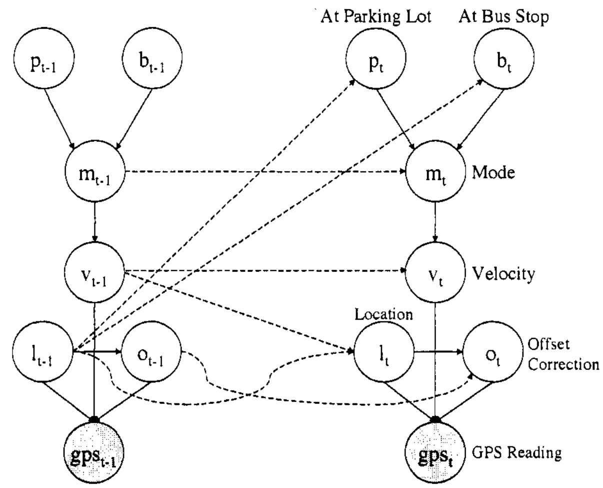


Figure 5.1: Dynamic Bayes net model of the transportation domain

Two-slice dynamic Bayes net model of the transportation domain, showing dependencies between the observed and hidden variables. Observed variables are shaded. Intra-temporal links are solid, inter-temporal links are dashed.

signal"). All of the other variables — sensor error, velocity, user location, mode, and the presence of a parking lot or bus stop location — are hidden variables whose values must be inferred from the raw GPS readings.

The dependencies between the nodes in Fig. 5.1 can be quite complex. The GPS reading at each time point is influenced by the local sensor error and the user's actual velocity and location. The location at time t only depends on the person's previous location and the motion velocity. Note that GPS is explicitly not considered to provide the true user location; urban interference, map reference point errors, GPS error and sensor failure all cause the true location to be a hidden variable. The sensor offset correction node o_t is used to reason about errors in the GPS readings that are systematic over time and location. This node maintains a probability distribution over corrections to the GPS signal, which are caused by multi-path propagation error and/or dynamic satellite geometry. The node updates its belief state by comparing GPS readings to the street map to gradually adjust to local variations in signal offset.

A more complex relationship governs how the mode of transportation influences the instantaneous velocity (see figure 5.2). The influence of mode on velocity is complicated by the fact that the range of possible instantaneous velocities for each mode overlap. For example, movement at 7 km/hr may be a brisk walk or a slowly moving car or bus. To simplify the relationship between mode and velocity we model the continuous velocities using the Gaussian mixture shown in Fig. 5.2. A separate unsupervised Expectation-Maximization (EM) process determined the parameters of these probability densities using real velocity data. Our model assumes that velocities are drawn randomly from these Gaussians, where the probability of drawing from a particular Gaussian depends on the mode. For example, the walking mode draws a speed from the left-most cluster with high probability. In the bus mode, the person has a uniform chance of being in each of the three slowest velocity clusters. In our current approach, the probabilities for the Gaussians in the different transportation modes were set manually based on external knowledge. Learning the weights of the mixture components depending on the transportation mode (and eventually location) is left for future research.

In our model, the motion mode at time t only depends on the previous mode and the presence of a parking lot or bus stop. For example, the person can only get on a bus if the node b_t indicates the presence of a bus stop. The values of the bus stop and parking lot nodes depend on the location of the person, as indicated by the arrows in the model shown in Fig. 5.1. Learning mode and location transition probabilities is an important aspect of our approach and will be discussed in Sect. 5.2.

Particle Filter Based Implementation

Particle filters provide a sample-based implementation of general Bayes filters [44]. They represent posterior distributions over the state space with temporal sets, S_t, of n weighted samples:

$$S_t = \{\langle x_t^{(i)}, w_t^{(i)} \rangle \mid i = 1, \ldots, n\}$$

Here each $x_t^{(i)}$ is a sample (or state), and the $w_t^{(i)}$ are non-negative numeric factors called *importance weights*, which sum to one. Like Kalman filters, particle filters apply the re-

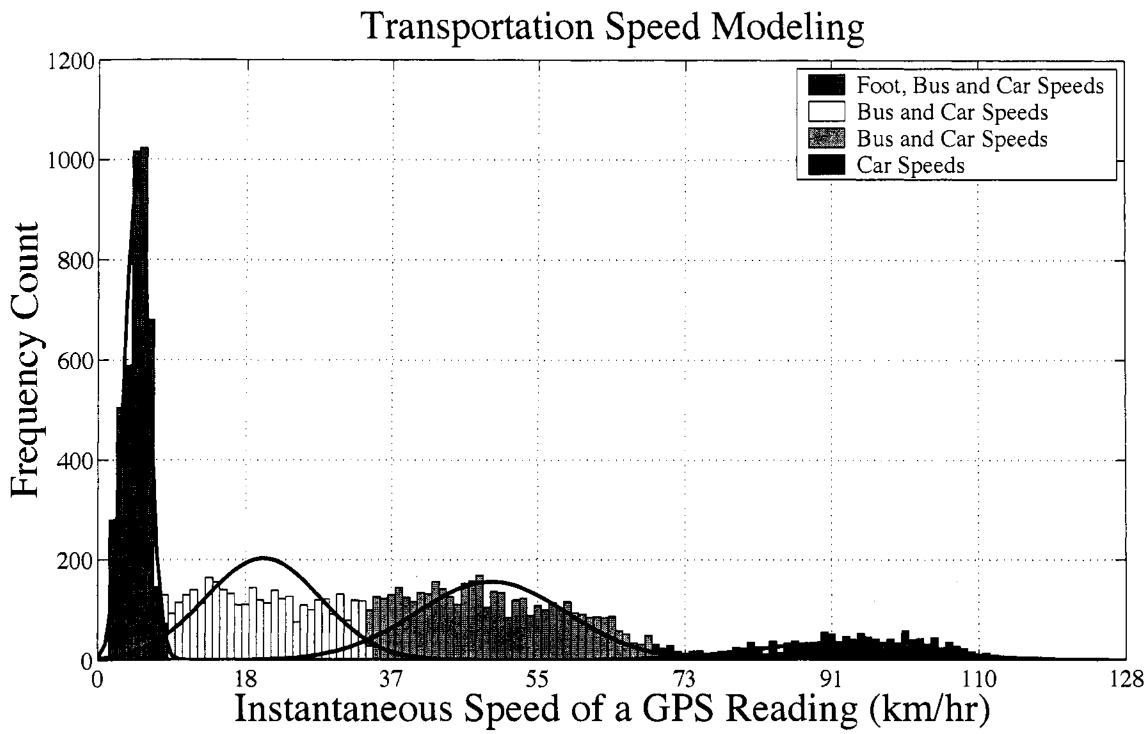


Figure 5.2: Gaussian mixture model for transportation velocities.

Gaussian mixture model for the dependency of transportation mode on velocities. The Gaussians were learned using EM based on previously collected velocity data. The frequencies of the raw velocity values are indicated by the bins. Different transportation modes are modeled by sampling with different probability from the four Gaussians.

cursive Bayes filter update to estimate posteriors over the state space, but unlike Kalman filters, particle filters are not restricted to unimodal posterior distributions. [1] The basic particle filter updates the posterior according to the following sampling procedure, often referred to as sequential importance sampling with re-sampling (SISR, see also [44]):

- **Sampling:** Draw n samples $x_{t-1}^{(i)}$ from the previous set and generate n new samples

[1]We consider multi-hypothesis tracking to be a viable alternative to our particle filter based implementation. Multi-hypothesis tracking overcomes the restrictive assumption of the plain Kalman filter by estimating a state using multiple Kalman filters [15]. An implementation of such an approach is demonstrated in [92].

$x_t^{(j)}$ using the distribution

$$p(x_t \mid x_{t-1}).$$

The new samples now represent the density given by the product

$$p(x_t \mid x_{t-1})p(x_{t-1} \mid z_{1:t-1})$$

This density is the so-called *proposal distribution* used in the next step.

- **Importance sampling:** Assign each sample $x_t^{(j)}$ an importance weight according to the likelihood of the observation, z_t, given the sample,

$$w_t^{(j)} = p(z_t \mid x_t^{(j)}).$$

- **Re-sampling:** Multiply / discard samples by drawing samples with replacement according to the distribution defined through the importance weights $w_t^{(j)}$.

It can be shown that this procedure in fact approximates the Bayes filter update (5.1), using a sample-based representation [41, 44].

The application of particle filters to the problem of location and mode estimation using the network shown in Fig. 5.1 is rather straightforward. Each particle $x_t^{(i)}$ represents an instantiation of the random variables describing the transportation mode m_t, the location l_t, and the velocity v_t. The parking lot and bus stop variables p_t and b_t are extracted from each sample location l_t. Finally, o_t is determined globally for all particles by estimating the offset between GPS readings and the street map. The update steps of the particle filter can be implemented as follows. The temporal sampling step corresponds to advancing each particle according to the motion model: First the transportation mode is chosen according to the previous transportation mode and the presence of bus stops or parking lots. This gives us m_t. Then we randomly pick a velocity from the velocity model for the specific mode m_t. The velocity is used to advance the position of the person on the graph. If the sampled velocity implies a transition to another edge, the next edge e_t is drawn with probability $p(e_t \mid e_{t-1}, m_t)$ (see [93] for more information on edge transitions). After these sampling steps, the resulting states represent the predicted location, velocity, and transportation mode. The importance sampling step is implemented by weighting each sample according to the likelihood of observing the current signal from the GPS sensor given the new location of the sample. The re-sampling step of the particle filter algorithm does not have to be changed.

5.2 Parameter Learning

One of the advantages of modeling the world with a graph is the ability to record behavioral data about edge transitions. The discrete nature of such transitions facilitates unsupervised learning of hierarchical model parameters. We have an intuitive prior expectation of how state transitions occur between and within edges: edge transitions occur uniformly among the edge's neighbors, and mode transitions vary according to the presence of a bus stop or parking lot.

Learning in this context means adjusting the model parameters to better fit the training data, typically to better model an individual user or the environment. Learning parameters for specific individuals captures idiosyncratic motion patterns — the movements the user commonly makes, as opposed to the logically possible set of movements. Since our model also includes transportation mode, learning also means changing our prior expectations about which edges mode transitions occur on. Bus stops and parking locations are conceptual locations where mode transitions may occur. Our model enables learning of the commonly used subset of these locations, to highlight where a user frequently parks her car, for example. The learned model supports better tracking and prediction than the prior model, and is the foundation upon which high-level understanding of the user's behavior is built.

We now describe how to learn the parameters of our graph model using data collected by a person moving through the community. Our motivating application of the Activity Compass forces us to learn the transportation modes in an *unsupervised manner.* When deployed, Activity Compass users must not be required, for example, to keep a diary for several weeks of their transportation modes in order to create a supervised training set. Hence, the most obvious difficulty is that we have to learn the motion model based solely on a map and a stream of non-continuous and noisy GPS sensor data.

A general approach for solving such learning problems is the well-known Expectation-Maximization (EM) algorithm [20, 120]. In our application, EM is based on the observation that learning the model parameters would be easy *if* we knew the person's true location and transportation mode at any point in time. Unfortunately, location and transportation mode are hidden variables, *i.e.* they cannot be observed directly but have to be inferred from the raw GPS measurements. EM solves this problem by iterating between an Expectation step (E-step) and a Maximization step (M-step). In a nutshell, each E-step estimates expectations (distributions) over the hidden variables using the GPS observations along with the current estimate of the model parameters. Then in the M-step the model parameters are updated using the expectations of the hidden variables obtained in the E-step. The updated model is then used in the next E-step to obtain more accurate estimates of the hidden variables. EM theory tells us that in each iteration the estimation of the parameters will be improved and it will eventually converge to a local optimum. In the following we give a more detailed description of how to apply EM theory in our domain.

5.2.1 *E-step:*

Let Θ denote the parameters of the graph-based model we want to estimate and Θ^i denote the estimation thereof at the i-th iteration of the EM algorithm. The model parameters contain all conditional probabilities needed to describe the dynamic system shown in Fig. 5.1. The E-step estimates

$$p(x_{1:t} \mid z_{1:t}, \Theta^{(i-1)}), \tag{5.2}$$

i.e. the posterior distribution over the trajectories of the person given the observations and parameters updated in the previous iteration. Here $x_{1:t}$ and $z_{1:t}$ are the states and observations, respectively. Since it is not possible to find a closed-form solution for the posterior over $x_{1:t}$, we have to resort to an approximate approach [89]. Observe that when we do particle filtering using the motion model with parameter $\Theta^{(i-1)}$, the particle distribution at each time t along with the history of particles is an approximation for $p(x_{1:t} \mid z_{1:t}, \Theta^{(i-1)})$.

Hence, the desired expectation can be computed using the graph-based particle filter described in Sect. 5.1.1. Before we give implementation details for the E-step, let us take a closer look at the M-step.

5.2.2 M-step:

The goal of the M-step is to maximize the expectation of $\log p(z_{1:t}, x_{1:t} \mid \Theta)$ over the distribution of $x_{1:t}$ obtained in the E-step by updating the parameter estimations. Because the distribution of $x_{1:t}$ is represented by the history of particles, the estimation of the parameters at the i-th EM iteration is computed by summing over all trajectories:

$$\begin{aligned} \Theta^{(i)} &= \operatorname*{argmax}_{\Theta} \sum_{j=1}^{n} \log p(z_{1:t}, x_{1:t}^{(j)} \mid \Theta) \\ &= \operatorname*{argmax}_{\Theta} \sum_{j=1}^{n} (\log p(z_{1:t} \,|\, x_{1:t}^{(j)}, \Theta) + \log p(x_{1:t}^{(j)} \,|\, \Theta)) \end{aligned} \tag{5.3}$$

Here, n is the number of particles, $x_{1:t}^{(j)}$ is the state history of the j-th particle, and assuming the following independence condition:

$$p(z_{1:t} \mid x_{1:t}^{(j)}, \Theta) = p(z_{1:t} \mid x_{1:t}^{(j)}),$$

i.e., observations are independent of model transition parameters if the state trajectory is known and the sensor model is fixed. We arrive at the following:

$$\Theta^{(i)} = \operatorname*{argmax}_{\Theta} \sum_{j=1}^{n} (\log p(z_{1:t} \,|\, x_{1:t}^{(j)}) + \log p(x_{1:t}^{(j)} \,|\, \Theta)) \tag{5.4}$$

$$= \operatorname*{argmax}_{\Theta} \sum_{j=1}^{n} \log p(x_{1:t}^{(j)} \mid \Theta) \tag{5.5}$$

For simplicity, we assume that all the particles have equal weight (at least after they are resampled). It is straightforward to extend our derivation to the case of different weights.

Our approach is in fact a direct extension of the Monte Carlo EM algorithm [147]. The only difference is that we allow particles to evolve with time. It has been shown that when the number of particles n is large enough, Monte Carlo EM estimation converges to the theoretical EM estimation [89].

5.2.3 Implementation Details

Even though EM can be used to learn all parameters Θ of the model described in Sect. 5.1.1, we are mostly interested in learning those parts of the model that describe the typical motion patterns of a user. All other parameters are fixed beforehand and not adjusted to a specific user. An advantage of this approach is that it requires much less training data than learning all parameters at once.

The motion patterns of a specific user are described by the location transitions on the graph and the mode transitions at the different locations. For the learning process, we have to initialize these probabilities to some reasonable values:

$p(e_t \mid e_{t-1}, m_{t-1})$ is the transition probability on the graph conditioned on the mode of transportation just prior to transitioning to the new edge. This conditional probability is initialized to a uniform distribution across all outgoing edges, with the exception of bus routes, which have a strong bias forcing buses to follow the route (bus routes can be obtained from GIS sources such as [37]). With this exception, our model has no preference for a specific path of the person.

$p(m_t \mid m_{t-1}, e_{t-1})$ is the mode transition probability. This probability depends on the previous mode m_{t-1} and the location of the person, described by the edge e_{t-1}. For example, each person has typical locations where she gets on and off the bus. Mode transitions are initialized with commonsense knowledge (*e.g.*, one may not switch from a bus to a car without first being on foot), and with knowledge of bus stops. Parking lots are uniformly distributed across our map with no biases toward actual parking lots.

A straightforward implementation of the E-step given in (5.2) is to generate the expectation over state trajectories by storing the history of each particle (see [44] for a discussion). To do so, at each re-sampling phase, the history of old samples needs to be copied to the new samples [2]. Then at the last time step, we have a set of samples with their histories. At the M-step, we update the model parameters simply by counting over the particle histories. For example, to get $p(e_j \mid e_i, BUS)$, we count the number of times when a particle in BUS mode transits from edge e_i to e_j and then normalize the counts over all edges following e_i and BUS. This approach, although easy to implement, suffers from two drawbacks. First, it is not efficient. When the data log is fairly long, saving the histories for all the particles needs a large amount of space and history replication becomes slow. Second, and more importantly, since the number of samples is finite, the repetition of the re-sampling will gradually diminish the number of different histories and eventually decrease the accuracy of the particle based approximation [44].

We can overcome these problems by observing that we are only interested in learning the discrete transitions between edges and modes, *e.g.*, the probability of transiting from edge e_i to edge e_j in BUS mode. The discreteness of these transitions allows us to apply the well-known Baum-Welch algorithm [120], an EM algorithm for hidden Markov models (HMM). The Monte Carlo version of the Baum-Welch algorithm [142] performs at each iteration both a forward and a backward (in time) particle filtering step. At each forward and backward filtering step, the algorithm counts the number of particles transiting between the different edges and nodes. To obtain probabilities for the different transitions, the counts of the forward and backward pass are normalized and then multiplied at the corresponding time slices.

To show how it works, we define:

$\alpha_t(e_t, m_t)$ is the number of particles on edge e_t and in mode m_t at time t in the *forward* pass of particle filtering.

$\beta_t(e_t, m_t)$ is the number of particles on edge e_t and in mode m_t at time t in the *backward* pass of particle filtering.

$\xi_{t-1}(e_t, e_{t-1}, m_{t-1})$ is the probability of transiting from edge e_{t-1} to e_t at time $t-1$ and in

[2] Unnecessary copy operations can be avoided by using tree data structures to manage pointers describing the history of particles.

mode m_{t-1}.

$\psi_{t-1}(m_t, m_{t-1}, e_{t-1})$ is the probability transiting from mode m_{t-1} to m_t on edge e_{t-1} at time $t-1$.

A short derivation gives us [120, 142],

$$\xi_{t-1}(e_t, e_{t-1}, m_{t-1}) \propto \alpha_{t-1}(e_{t-1}, m_{t-1}) p(e_t \mid e_{t-1}, m_{t-1}) \beta_t(e_t, m_{t-1}) \tag{5.6}$$

and

$$\psi_{t-1}(m_t, m_{t-1}, e_{t-1}) \propto \alpha_{t-1}(e_{t-1}, m_{t-1}) p(m_t \mid m_{t-1}, e_{t-1}) \beta_t(e_{t-1}, m_t) \tag{5.7}$$

After we have ξ_{t-1} and ψ_{t-1} for all the t from 2 to T, we update the parameters as: [3]

$$\begin{aligned} p(e_t \mid e_{t-1}, m_{t-1}) &= \frac{\text{expected number of transitions from } e_{t-1} \text{ to } e_t \text{ in mode } m_{t-1}}{\text{expected number of transitions from } e_{t-1} \text{ in mode } m_{t-1}} \\ &= \frac{\sum_{t=2}^{T} \xi_{t-1}(e_t, e_{t-1}, m_{t-1})}{\sum_{t=2}^{T} \sum_{e_t \in Neighbors\ of\ e_{t-1}} \xi_{t-1}(e_t, e_{t-1}, m_{t-1})} \end{aligned} \tag{5.8}$$

and similarly

$$\begin{aligned} p(m_t \mid m_{t-1}, e_{t-1}) &= \frac{\text{expected number of transitions from } m_{t-1} \text{ to } m_t \text{ on edge } e_{t-1}}{\text{expected number of transitions from } m_{t-1} \text{ on edge } e_{t-1}} \\ &= \frac{\sum_{t=2}^{T} \psi_{t-1}(m_t, m_{t-1}, e_{t-1})}{\sum_{t=2}^{T} \sum_{m_t \in \{BUS, FOOT, CAR\}} \psi_{t-1}(m_t, m_{t-1}, e_{t-1})} \end{aligned} \tag{5.9}$$

The complete implementation is depicted in Table 5.1. As the number of particles increases, the approximation converges to the theoretical EM estimation. Fortunately, our approach is very efficient in this regard, since our model parameters are associated with the number of edges and modes in the graph, not the number of particles.

In addition to the user specific parameters, our model requires the specification of other parameters, such as motion velocity and the GPS sensor model. The motion velocity is modeled as a mixture of Gaussians from which velocities are drawn at random. The probabilities of the mixture components depend on the current motion mode and can be learned

[3]Usually we also need a prior number for each transition. These are hand set in this model.

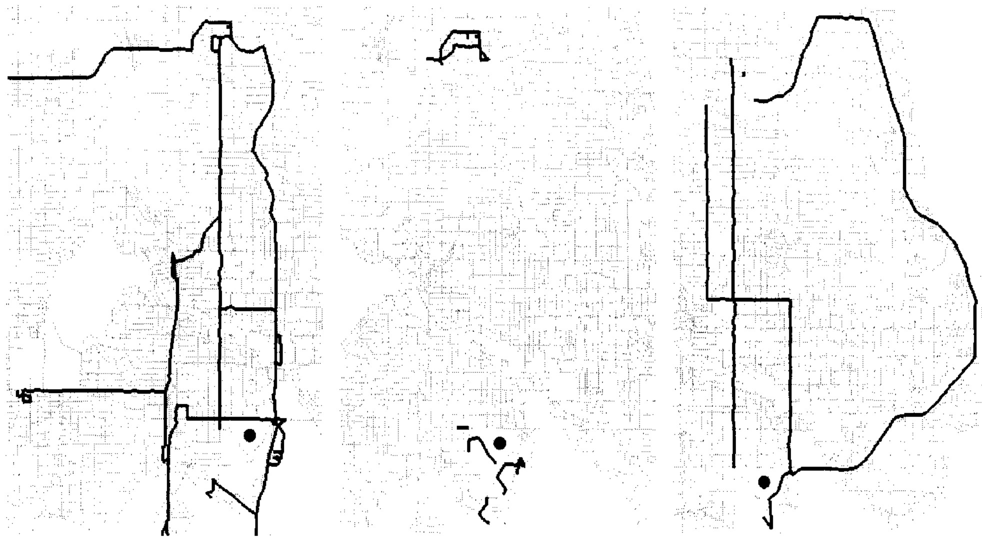


Figure 5.3: Training Data used for Model Learning Experiment

Car (left), Foot (middle), and Bus (right) training data used for experiments. The black dot is a common map reference point on the University of Washington campus.

beforehand using data labeled with the correct mode of motion. We use a standard model to compute the likelihood $p(z_t \mid x_t)$ of a GPS sensor measurement z_t given the location x_t of the person [62].

5.3 Experiment #1

To validate our formalisms, we conducted an experiment using real position data. Our test data set consists of logs of GPS data collected by the author. The data contains position and velocity information collected at approximately 5 second intervals during periods of time in which the author was moving about outdoors. This data was hand labeled with one of three modes of transportation: foot, bus, or car. This labeling was useful for validating the results of our unsupervised learning, but was not used by the EM learning process described in sec 5.2.

From this data set, we chose 29 episodes representing a total of 12 hours of logs. This subset consists of all of portions of the data set that were bounded by GPS signal loss, *i.e.*

Table 5.1: EM-based parameter learning algorithm

Model Initialization: Initialize the model parameters $p(e_t|e_{t-1}, m_{t-1})$ and $p(m_t|m_{t-1}, e_{t-1})$.

E-step:

1. Generate n uniformly distributed samples and set time $t = 1$.

2. Perform forward particle filtering:

   (a) Sampling: generate n new samples from the existing samples using the current parameter estimation $p(e_t|e_{t-1}, m_{t-1})$ and $p(m_t|m_{t-1}, e_{t-1})$.

   (b) Importance sampling: reweight each sample based on observation z_t.

   (c) Re-sampling: copy / discard samples according to their importance weights.

   (d) Count and save $\alpha_t(e_t, m_t)$

   (e) Set $t = t + 1$ and repeat (2a)-(2d) until $t = T$.

3. Generate n uniformly distributed samples and set $t = T$.

4. Perform backward particle filtering:

   (a) Compute backward parameters $p(e_{t-1}|e_t, m_t)$, $p(m_{t-1}|m_t, e_t)$ from $p(e_t|e_{t-1}, m_{t-1})$ and $p(m_t|m_{t-1}, e_{t-1})$

   (b) Sampling: generate n new samples from the existing samples using the backward parameter estimation.

   (c) Importance sampling: reweight each sample based on observation z_t.

   (d) Re-sampling: multiply / discard samples according to their importance weights.

   (e) Count and save $\beta_t(e_t, m_t)$

   (f) Set $t = t - 1$ and repeat (4b)-(4e) until $t = 1$.

M-step

1. Compute $\xi_{t-1}(e_t, e_{t-1}, m_{t-1})$ and $\psi_{t-1}(m_t, m_{t-1}, e_{t-1})$ using (5.6) and (5.7) and then normalize.

2. Update $p(e_t|e_{t-1}, m_{t-1})$ and $p(m_t|m_{t-1}, e_{t-1})$ using (5.8) and (5.9).

Loop Repeat E-step and M-step using updated parameters until model converges.

Table 5.2: Results of Estimating Transportation Mode

Mode estimation quality of different algorithms.

Model	Cross-Validation Prediction Accuracy
Decision Tree with Speed and Variance	55%
Prior Graph Model, w/o bus stops and bus routes	60%
Prior Graph Model, w/ bus stops and bus routes	78%
Learned Graph Model	84%

had no intermediate loss of signal of more than 30 seconds, and that contained a change in the mode of transportation at some point in the episode. These episodes were divided chronologically into three groups which formed the sets for three-fold cross-validation for our learning. Fig. 5.3 shows one of the cross-validation groups used for training. The street map was provided by the US Census Bureau [26] and the locations of the bus stops come from the King County GIS office [37].

5.3.1 Mode Estimation and Prediction

One of the primary goals of our approach is learning a motion model that predicts transportation routes, conditioned on the mode of transportation. We conducted an experiment to validate our model's ability to correctly learn the mode of transportation at any given instant. For comparison, we also trained a decision tree model using supervised learning on the data [101]. We provided the decision tree with two features: the current velocity and the standard deviation of the velocity in the previous sixty seconds. Using the data annotated with the hand-labeled modes of transportation, the task of the decision tree was to output the transportation mode based on the velocity information. We used three-fold cross-validation groups to evaluate all of the learning algorithms. The results are summarized in the first row of Table 5.2. The first result indicates that 55% of the time the decision tree approach was able to accurately estimate the current mode of transportation on the test data. Next, we used our Bayes filter approach without learning the model parameters, *i.e.* with uniform transition probabilities. Furthermore, this model did not consider the locations of bus stops or bus routes (we never provided parking locations to the algorithm). In contrast to the decision tree, the Bayes filter algorithm integrates information over time, thereby increasing the accuracy to 60%. The benefit of additionally considering bus stops and bus routes becomes obvious in the next row, which shows a mode accuracy of 78%. Finally, using EM to learn the model parameters increases the accuracy to 84% on test data not used for training. Note that this value is very high given the fact that often a change of transportation mode cannot be detected instantaneously.

Table 5.3: Prediction Accuracy of Mode Transition Changes.

Model	Precision	Recall
Decision Tree with Speed and Variance	2%	83%
Prior Graph Model, w/o bus stops and bus routes	6%	63%
Prior Graph Model, w/ bus stops and bus routes	10%	80%
Learned Graph Model	40%	80%

A similar comparison can be done looking at the techniques' ability to predict not just instantaneous modes of transportation, but also *transitions* between transportation modes. Table 5.3 shows each technique's accuracy in predicting the qualitative change in transportation mode within 60 seconds of the actual transition — for example, correctly predicting that the person got off a bus. Precision is the percentage of time when the algorithm predicts a transition that an actual transition occurred. Recall is the percentage of real transitions that were correctly predicted. Again, the table clearly indicates the superior performance of our learned model. Learning the user's motion patterns significantly increases the precision of mode transitions, *i.e.* the model is much more accurate at predicting transitions that will actually occur.

An example of the modes of transportation predicted after training on one cross-validation set is shown in Fig. 5.4.

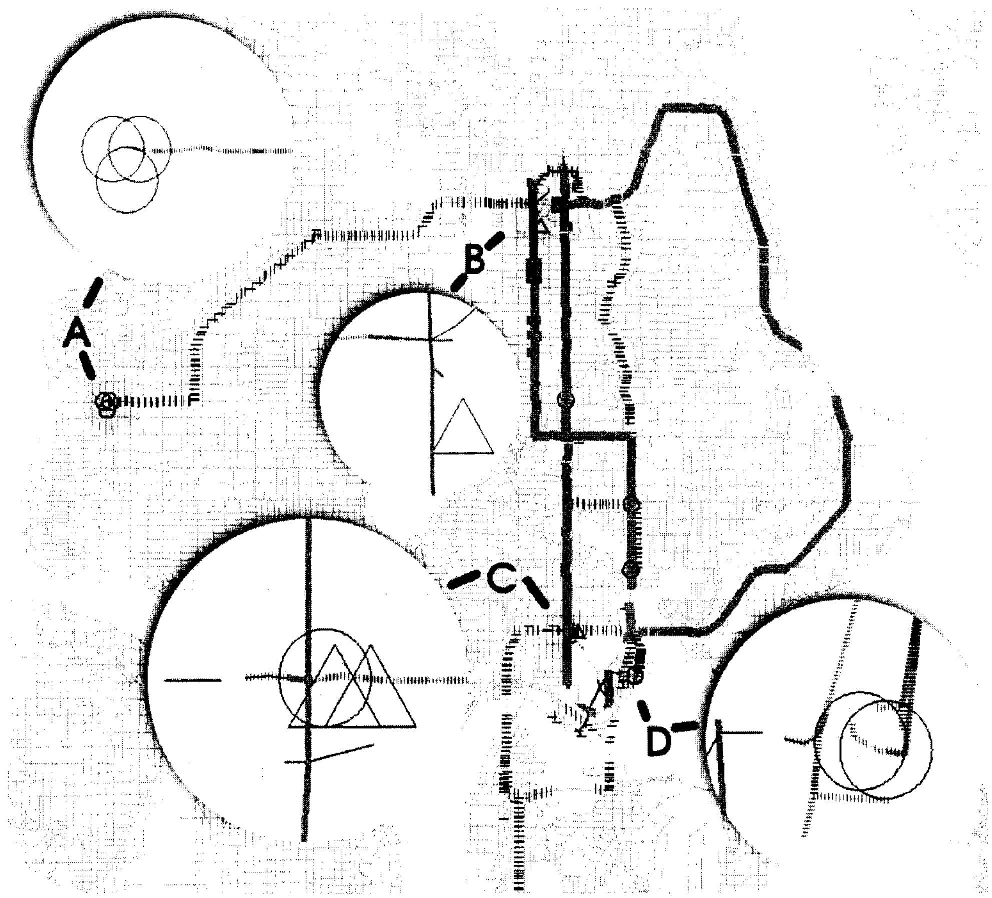


Figure 5.4: Results of learning

This map shows the learned transportation behavior based on one cross-validation set containing nineteen episodes. Shown are only those edges and mode transitions that the learned model predicts with high probabilities. Thick gray lines indicate learned bus routes, thin black lines indicate learned walking routes, and cross-hatches indicate learned driving routes. Circles indicate parking spots, and the triangles show the subset of bus stops that were learned to have a high probability of transitioning to or from foot. There are four call-outs to show detail. **(A)** shows a frequently traveled road ending in three distinct parking spaces. This route and the parking spots indicate the correctly learned car trips between the author's home and church. **(B)** shows a frequently traveled foot route that enters from the northeast, ending at one of the frequently used bus stops of the author. The main road running east-west is an arterial road providing access to the highway for the author. **(C)** shows an intersection at the northwest of the University of Washington campus. There are two learned bus stops. The author frequently takes the bus north and south from this location. This is also a frequent car drop off point for the author, hence the parking spot indication. Walking routes extend west to a shopping area and east to the campus. **(D)** shows a major university parking lot. Foot traffic walks west toward campus.

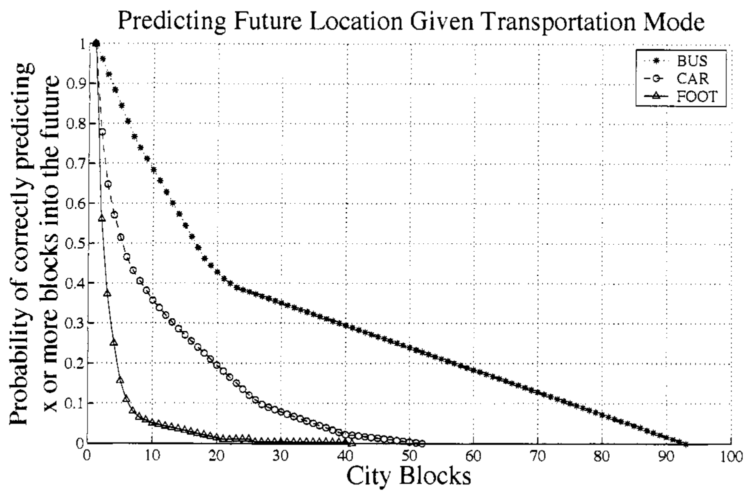


Figure 5.5: Location Prediction Capabilities of the Learned Model.

5.3.2 Location Prediction

The location prediction capabilities of our approach are illustrated in Fig. 5.5 and 5.6. In Fig. 5.5, the learned model was used to predict the location of the person into the future. This was done by providing the ground truth location and transportation mode to the algorithm and then predicting the most likely path based on the transition probabilities learned from the training data. The figure shows the percentage of trajectories that were predicted correctly, given different prediction horizons. Prediction length was measured in city blocks. For example, in 50% of the cases, the location of the person was predicted correctly for 17 blocks when the person was on the bus. In 30% of the cases, the prediction was correct for 37 blocks, and 75 blocks were predicted correctly in 10% of the cases. Note that the linear drop in bus route prediction probability is due to the fact that the data contained several correctly predicted episodes of a 92 block long bus trip. Obviously, long term distance prediction is much less accurate when a person walks. This is due to the higher variability of walking patterns and the fact that people typically do not walk for

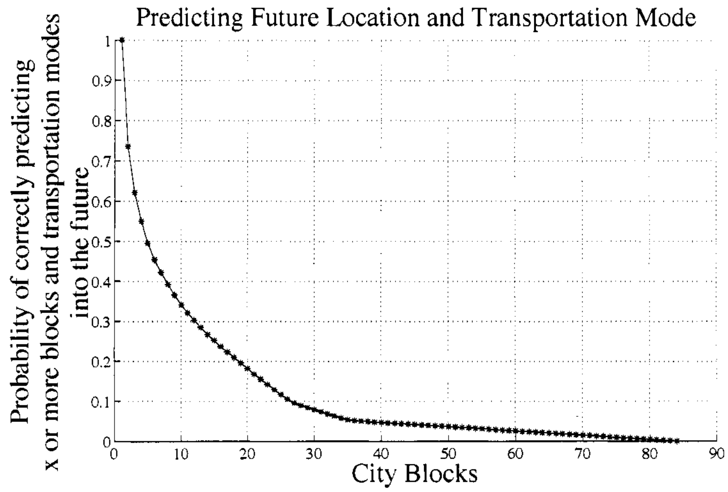


Figure 5.6: Location and Mode Prediction Capabilities of the Learned Model.

many city blocks, thereby making a long term prediction impossible.

In Fig. 5.6, the learned model was used to predict both the location and the transportation mode of the person into the future. This was done by providing the ground truth location to the algorithm and then predicting the most likely path and sequence of transportation mode switches based on the transition probabilities learned from the training data. The graph shows that in 50% of the cases, the model is able to correctly predict the motion and transportation mode of the person for five city blocks. This result is extremely promising given that the model was trained and tested on subsets of 29 episodes.

5.4 *Experiment #1 Summary*

The work presented so far in this chapter helps lay the foundation for reasoning about high-level descriptions of human behavior using sensor data. We showed how complex behaviors such as boarding a bus at a particular bus stop, traveling, and disembarking can be recognized using GPS data and general commonsense knowledge, without requiring additional sensors to be installed throughout the environment. We demonstrated that good predictive user-specific models can be learned in an unsupervised fashion.

The key idea of our approach is to apply a graph-based Bayes filter to track a person's location and transportation mode on a street map annotated with bus route information. The location and transportation mode of the person is estimated using a particle filter. We showed how the EM algorithm along with frequency counts from the particle filter can be used to learn a motion model of the user. A main advantage of this unsupervised learning algorithm is the fact that it can be applied to raw GPS sensor data.

5.5 The Complete Inference Engine

In previous chapters we have described the desired behavior of our system — a behavior that depends on the system being able to learn and reason about its user's transportation routines. In particular, we required the following:

- the system should learn about its user's transportation routines in an unsupervised and unobtrusive manner;

- the system should be able to predict likely destinations the user may want to go to at any given moment in time;

- the system should be able to recognize anomalous behavior; in particular, if told where the user is going (by the user who requests directions or by a care taker or job coach who specifies the destination), the system should be able to detect, as early as possible, when the user strays from one of the usual paths that lead to that destination.

Because of the inherent uncertainties about human behavior as well as the possible errors from the maps and GPS measurements, we have adopted a probabilistic approach that can handle potential errors and uncertainties in a statistically sound way. Neither the model we have described so far in this chapter nor a model from the recent literature proposed in [11, 113] are sufficient to meet the needs of our system. They both have fundamental inadequacies with respect to the set of requirements laid out above. Once we examine them

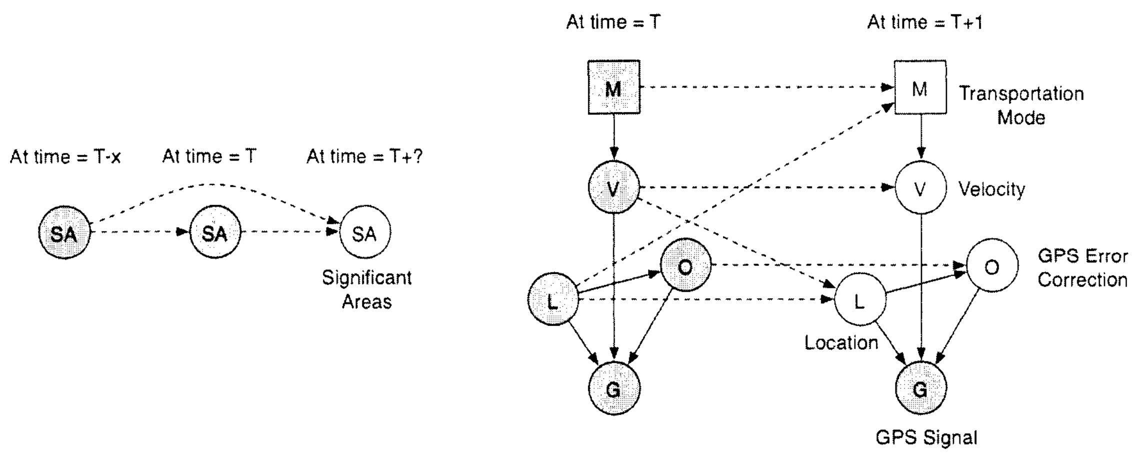


Figure 5.7: Two Proposed Techniques for Modeling Transportation Behavior

On the left, a second order Markov Model (2MM) from [11] showing dependencies between observed and hidden variables. Shaded nodes are observed. All links are inter-temporal. On the right, a redrawn two-slice Dynamic Bayesian Network (2TBN) from figure 5.1 showing dependencies between observed and hidden variables. Shaded nodes are observed. Intra-temporal causal links are solid, inter-temporal links are dashed.

we will continue by discussing a more comprehensive model that subsumes the other two and provides the required functionality. In Section 5.6 we will evaluate the new model with respect to our needs.

5.5.1 Previous Models

Ashbrook, et al. have proposed using a second order Markov model (2MM – see Figure 5.7-left) as a predictive tool for reasoning about likely destinations toward which a user may next be traveling [11]. The system logs continuous GPS signals, extracts places where the user seems to have stopped for a significant amount of time and then clusters them into significant locations. The optimal radius for a significant location is chosen after manual inspection of results for different radii. These results become the basis for training a second order Markov model. The authors have demonstrated that given the last two significant locations visited by the user, the system was able to generate a small and accurate set of the next most likely destinations.

In contrast to our desired behavior, this model is not able to refine estimates of the current goal using GPS information observed when moving from one significant location to another. Since significant locations might be long distances away this causes an unacceptable lag in noticing unusual behavior and significant amounts of GPS information are disregarded. This model also has no timing mechanism, so there is no way to judge when destinations will be reached or to react when too much time has passed. Finally, since the model only considers two previous locations, complex plans involving multiple significant locations cannot easily be reasoned about.

The model proposed earlier in this chapter is a two slice Dynamic Bayesian Network (2TBN – see Figure 5.7-right) for inferring a user's transportation mode from continuous GPS signals (see also [113]). A Dynamic Bayesian Network is an extension of Bayesian Networks that allows for time–changing variables (details in [106]). Given a representation of the street maps, we saw that the system was able to accurately infer a user's most likely position, compensating for GPS sensor errors. The system was also able to infer locations of parking lots accessed by the user as well as bus routes and bus stop locations, all of which improved its accuracy. Finally, it estimated a user's street to street transition probabilities in an unsupervised manner and was able to use the information to further improve its accuracy.

The 2TBN could easily be adapted to detect when the user strays from a frequently traveled path. But the biggest shortcoming of this model stems from the fact that the system does not explicitly reason about the ultimate goals of a trip. Therefore, the system cannot predict the likely destinations toward which the user may be heading. Moreover, neither model, even if told a user's destination, can reason about the likely paths the user might take and, subsequently, cannot detect when the user strays from a correct path.

5.5.2 A New Hierarchical Model

To account for the inability of these models to meet our desiderata we use a new *hierarchical* Dynamic Bayesian Network model representing transportation routines introduced by Liao et.al. in [92]. The new model subsumes the capabilities of the previous models and bridges

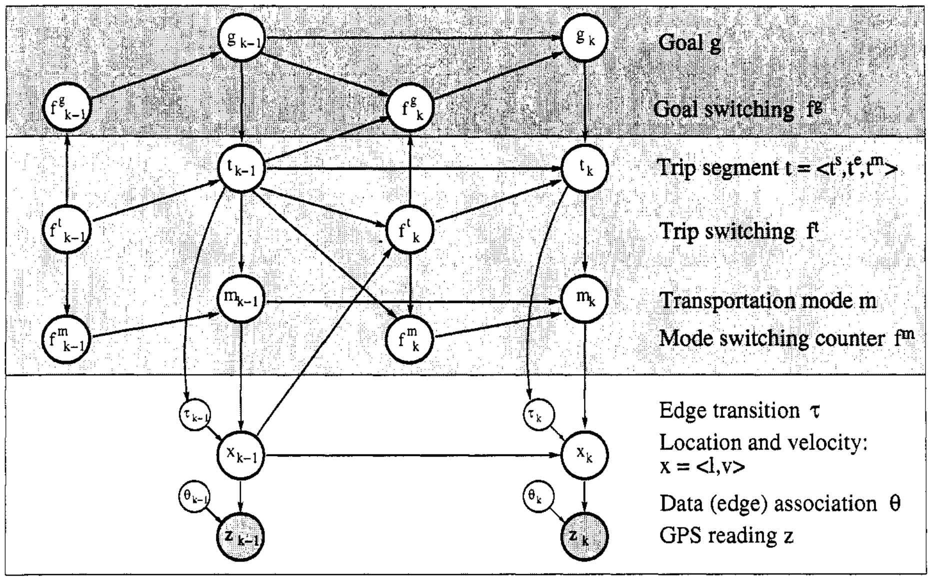


Figure 5.8: Hierarchical activity model

Hierarchical activity model representing a person's outdoor activities. The top level estimates the current goal, the middle layer represents segments of a trip and mode of transportations and the lowest layer estimates the person's location on the street map. Figure courtesy of Liao [92]

the gap between the raw sensor measurements and the abstract goal intentions of a user. A brief discussion of this model follows; refer to [92] for full technical details of the model structure, inference and training.

Figure 5.8 shows the graphical structure of the new model. At the very highest level of this model, goals g_k (subscript k indicates the discrete time step) are explicitly represented as significant locations. Transitions between goals have specific probability distributions independently of the routes by which they are reached. Each goal destination influences the choice of which *trip segments* the user takes. Trip segments are sequences of motion in which the transportation mode is constant. Each trip segment t includes its start location t_k^s, end location t_k^e, and the mode of transportation, t_k^m, the person uses during the segment. Each trip segment biases the expectation over the mode of transportation and the changes in location. The mode of transportation m, in turn, determines the location and velocity distribution of the user. At the bottom level, we denote by $x_k =< l_k, v_k >$ the location and motion velocity of the person. Edge transition τ_k indicates the next street when passing an intersection and data association θ_k "snaps" a GPS measurement onto some streets around it. The switching nodes f_k^g, f_k^t and f_k^m indicate when changes in a variable's value can happen.

An efficient algorithm based on Rao-Blackwellised particle filters [24, 25, 43] has been developed to perform online inference for this model. At the lowest level, location tracking on the street map is done using graph-based Kalman filtering which is more efficient than the grid-based Bayesian filter and traditional particle filtering [44] used for the 2TBN model. At the highest level, the joint distribution of goals and trip segments is updated analytically using exact inference techniques. As a result, this model makes it possible to reason about high level goals (or significant locations) explicitly. The contribution of this model is that it considers not only previous significant locations visited but also the current location and the path taken so far to reason about likely destinations.

The parameters in the model are estimated in an unsupervised manner. This is a three step process. In a first pass through the data, the possible goals for a user are discovered by observing when the user stays at a location for a long time. Then in a second pass, the usual parking spots and bus stops are inferred using an Expectation-Maximization algorithm [120] similar to the learning of the 2TBN earlier in this chapter. Finally, the transition matrices at all levels are re-estimated simultaneously using a second Expectation-Maximization procedure with the full model. The learning process does not require any labeled data and therefore, requires no intervention from the user.

To detect abnormal events, the approach uses two models with different transition parameters. The first tracker assumes the user is behaving according to his personal historical trends and uses the learned hierarchical model for tracking. The second tracker assumes a background model of activities and uses an untrained prior model that accounts for general physical constraints but is not adjusted to the user's past routines. The trackers are run in parallel, and the probability of each model given the observations is calculated. When the user is following his ordinary routine the learned hierarchical model should have a higher probability, but when the user does something unexpected the second model should become more likely.

To compute the probability of each model, we use the concept of *Bayes factors*, which are a standard tool for comparing the quality of dynamic models based on measurements [149]. The Bayes factor H_k is computed recursively as

$$\begin{aligned} H_k &\equiv \frac{P(z_{1:k} \mid M_{prior})}{P(z_{1:k} \mid M_{learned})} \\ &= \frac{P(z_k \mid z_{1:k-1}, M_{prior})}{P(z_k \mid z_{1:k-1}, M_{learned})} \cdot \frac{P(z_{1:k-1} \mid M_{prior})}{P(z_{1:k-1} \mid M_{learned})} \\ &= \frac{P(z_k \mid z_{1:k-1}, M_{prior})}{P(z_k \mid z_{1:k-1}, M_{learned})} \cdot H_{k-1}, \end{aligned} \tag{5.10}$$

where $P(z_k \mid z_{1:k-1}, M_{prior})$ and $P(z_k \mid z_{1:k-1}, M_{learned})$ are the likelihoods of the observation z_k given the untrained and the hierarchical model (the likelihoods are extracted as a side-product of tracking). From a Bayes factor, we can compute the probability of abnormal behavior:

$$\begin{aligned} P_k(\text{Abnormal}) \equiv P_k(\text{correct model} &= M_{prior}) \\ = \frac{P(z_{1:k} \mid M_{prior})}{P(z_{1:k} \mid M_{prior}) + P(z_{1:k} \mid M_{learned})} &= \frac{H_k}{H_k + 1} \end{aligned} \tag{5.11}$$

The last step follows directly from (5.10).

5.5.3 Errors versus Novel Behavior

The above approach can detect unexpected events, but cannot distinguish errors from deliberate novel behavior. An important contribution of OK, however, is the ability to differentiate these cases using knowledge of the user's destination. This is possible because there are times when the system knows where the user is going, *e.g.*, if the user asks for directions to a destination, if a care-giver or job coach indicates the "correct" destination, or if the system has access to a location enabled date-book. In those situations we can *clamp* the value of the goal node in our model and reinterpret the low level observations. When the observations diverge significantly from the clamped high level predictions, the system is able to signal a possible error. Unlike in the 2TBN model, this model is capable of spotting anomalous behavior even if the user is following a well-trodden path, provided that path does not lead to the specified destination. This is what enabled us, in Section 4.4, to alert Eileen that she should get off bus number 68 and switch to 372, even though she takes both routes frequently.

Similarly to Equation (5.11), the probability of erroneous behavior given the user's input g (*i.e.*, the true goal and/or true trip segment) is

$$P_k(\text{Error} \mid g) \quad = \quad \frac{\hat{H}_k}{\hat{H}_k + 1} \tag{5.12}$$

where the Bayes factor $\hat{H}_k$ is now defined as

$$\hat{H}_k \quad \equiv \quad \frac{P(z_{1:k} \mid M_{prior})}{P(z_{1:k} \mid M_{learned}, g)}. \tag{5.13}$$

Here, $P(z_{1:k} \mid M_{learned}, g)$ is the likelihood given the clamped model.

In practice, when we track users for a long time, the probability of an error can grow very small and it can take too long for an observed error to cause this probability to cross the recognition threshold. To combat this lag, one could specify a floor that limits the error probability (e.g., 0.01 in our experiments) or compute the Bayes factors using the n most recent measurements.

5.6 *Experiment #2*

5.6.1 Experimental Methodology

In order to test our destination-enhanced model, we had a user carry a WAAS-enabled GPS logger with him continuously for 24 hours a day for 30 days. We then performed a three way cross validation procedure on the data without any manual labeling. The learned model correctly identifies six common goals, frequently used bus stops and parking lots, as shown in Figure 5.9 (left). Furthermore, our system is able to estimate the transition probabilities between goals, trip segments and streets. Using those transition matrices, we calculate the most likely trajectories on the street map between the goals, as shown in Figure 5.9 (right).

We tested our system using the learned model on a scenario similar to that in Section 4.4.

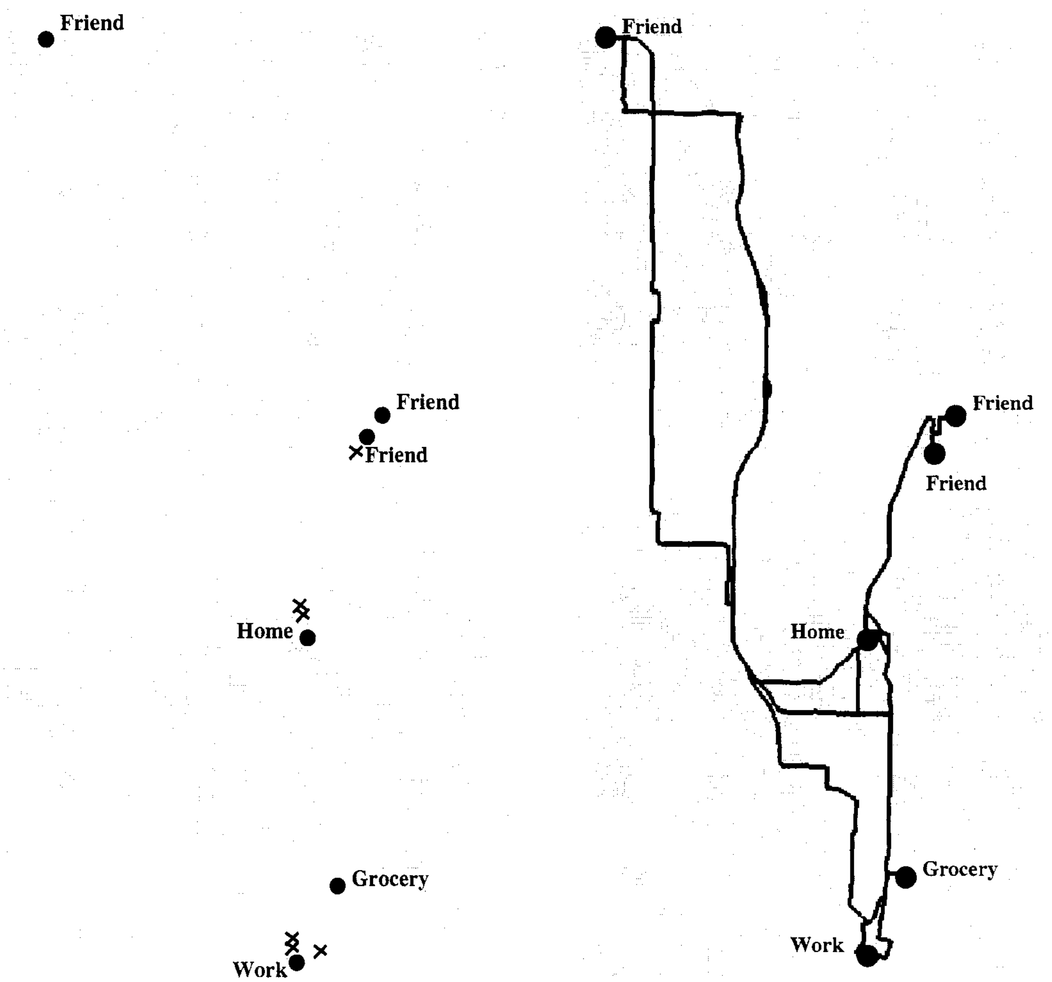


Figure 5.9: Learning Destinations and Routes

The left picture shows the street map along with the goals (dots), and usual bus stops and parking lots (cross marks). The right picture shows the most likely trajectories between the goals.

The results are shown in Figures 5.10 through 5.15. These figures present a sequential panel of experimental results. The left side of each panel displays a representation of the reasoning process that the inference engine is undertaking. The right portion of each panel displays what the users saw at each stage of the experiment, and the bottom portion holds a text description of the frame.

5.6.2 Model Clamping for Error Detection

In Figures 5.10–5.15, we have shown that OK is able to detect errors even when the user was on a frequently taken route. The system achieves this by letting the user explicitly select a destination, which we call *model clamping*. Figure 5.16 shows the impact of model clamping on inference results.

On the top we use the same data as in Section 5.6.1. In this example for the first 700 seconds both models have approximately equal belief that the user is not making an error, but when the bus took a turn that the user had never taken to get home, the probability of errors in the clamped model instantly and dramatically jumped. In contrast, the unclamped model cannot determine an error occurred because the user had taken that route to get to other destinations.

On the right is a foot experiment in which the user left his office and proceeded to walk in a direction away from the parking lot. When the destination is not specified, the tracker has a fairly steady level of confidence in the user's path (there are lots of previously observed paths from his office), but when the destination is specified, the system initially sees behavior consistent with walking toward the parking lot, and then as the user begins to turn away at time 125, the tracker begins to doubt the success of the user's intentions.

5.7 Summary

In this chapter we have presented a system called "Opportunity Knocks" (OK) that utilizes a rich model of user motion and behavior based on GPS sensor information to provide transportation assistance to people with mild cognitive disabilities. The primary function of the system is to route an individual from their current location to a chosen destination,

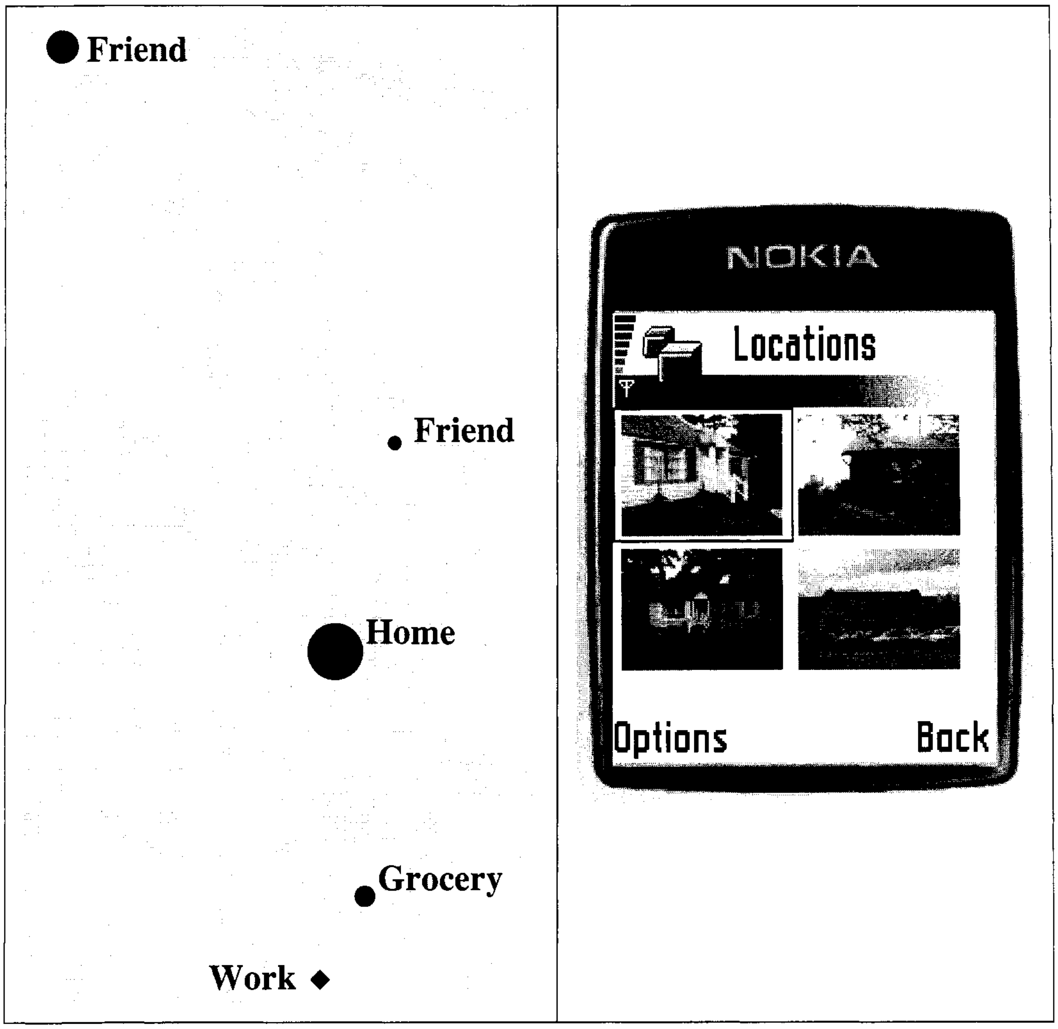


Figure 5.10: Experimental Results Part 1

This frame shows the user at "Work" while the system attempts to identify his most likely destinations. Circle areas correspond to relative likelihood. The phone displayed images of the most likely destinations left to right, top to bottom: home, friend 1, friend 2, grocery store. At this point, the user pretended to be confused and referred to OK for assistance.

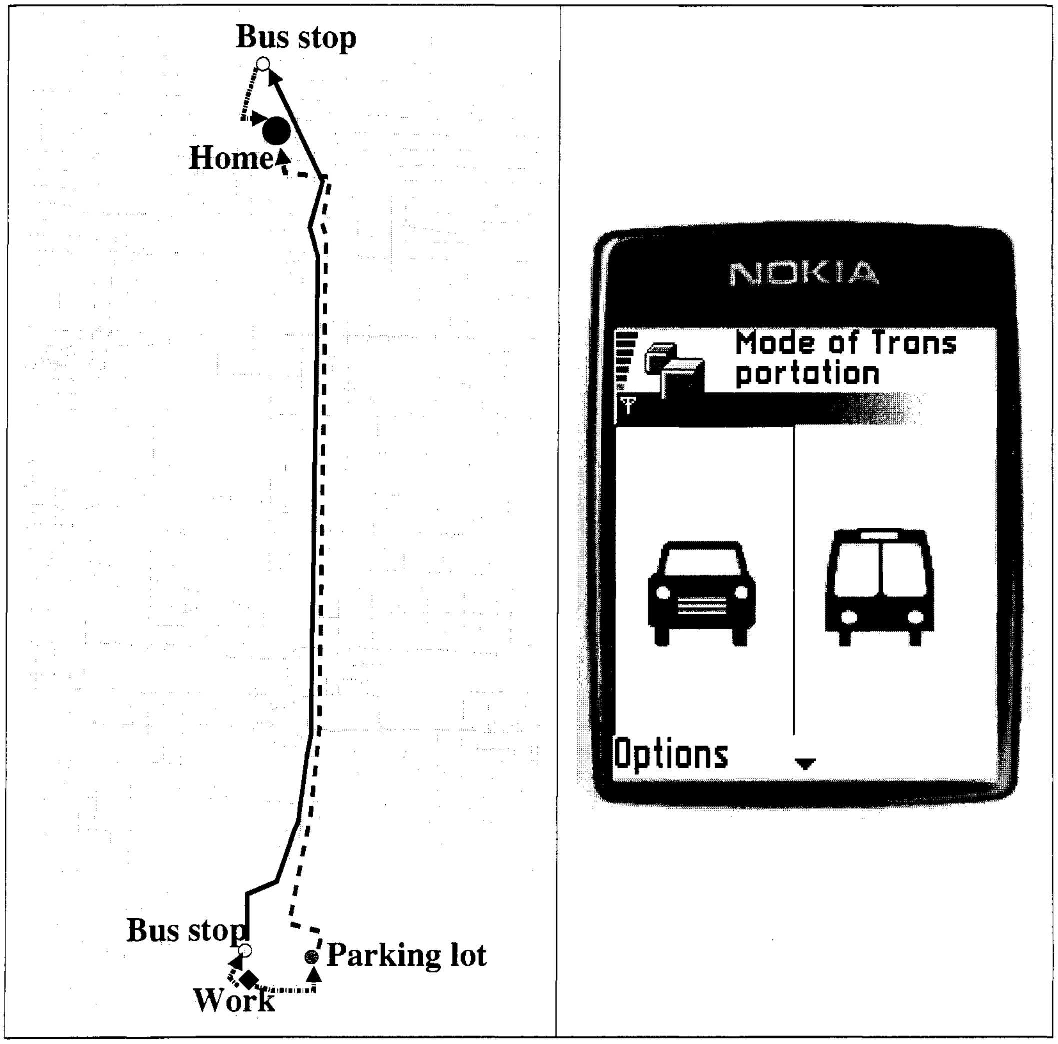


Figure 5.11: Experimental Results Part 2

After the user indicated that he'd like to go home, the system identified two routes that he usually takes, a bus route shown in solid lines and a car pool route shown in dashed lines. The system asked the user which way he would like to proceed.

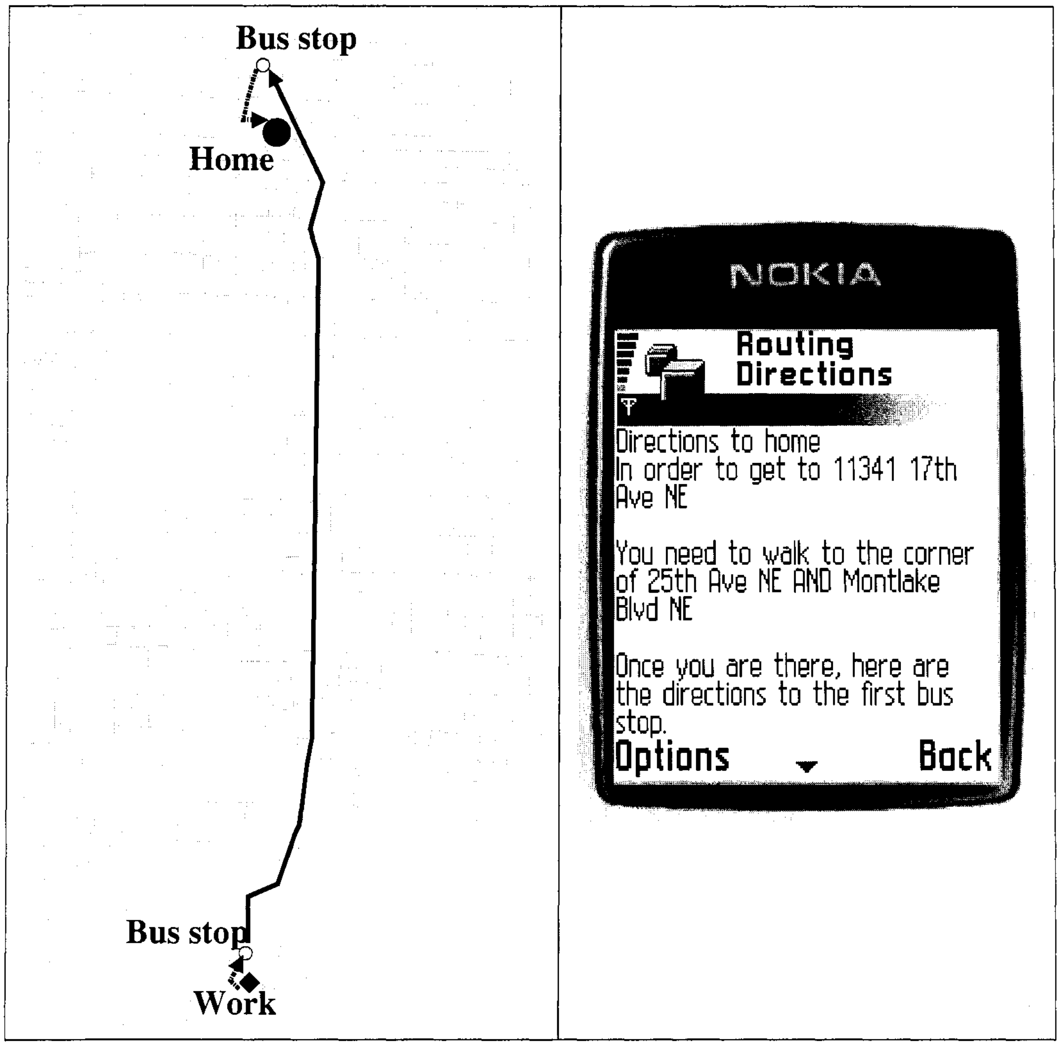


Figure 5.12: Experimental Results Part 3

The user selected the bus route and OK presented a text description of the learned route. The user proceeded to the bus stop, and boarded a bus. The bus that the user boarded however was going to his friend's house, a familiar, but incorrect route considering the expressed intention to go home.

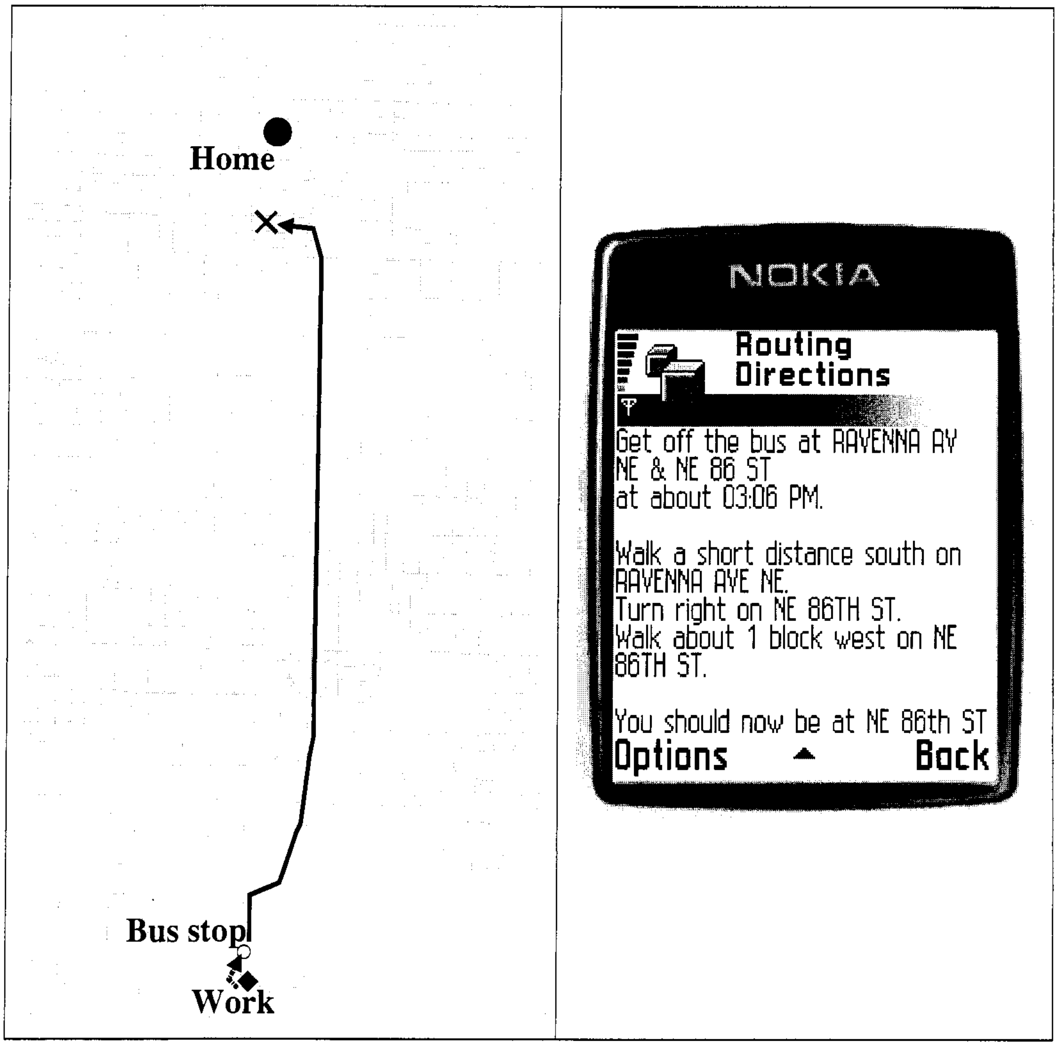
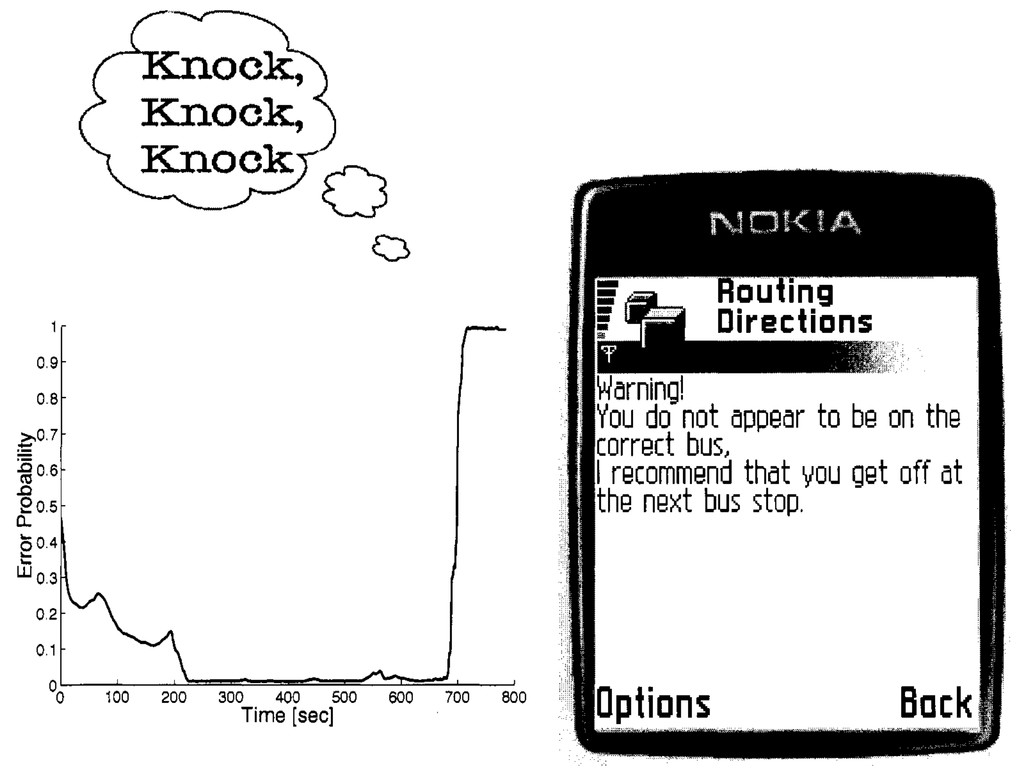


Figure 5.13: Experimental Results Part 4

The user rode the incorrect bus and the system monitored his progress. The system was unable to identify that the user was on the wrong bus because the routes coincided for the first portion of the bus ride. Before getting to the correct bus stop for going home, the system observed that the user had departed from the expected trip segment and turned west.



This graph shows how the system reasons about errors. When the error probability rises above 50% confidence, the system believes an abnormal condition exists. OK waits until a clamped curve breaks 80% confidence before alerting.

Figure 5.14: Experimental Results Part 5

When the bus diverted from the correct route, the system identified the behavior condition as an error. This was possible *even though* the user was on a frequently taken route. Because the user has explicitly selected a goal, OK identified an actual error (not just a novel behavior) had occurred. In response it proactively made its door knocking alert noise and showed a corresponding message

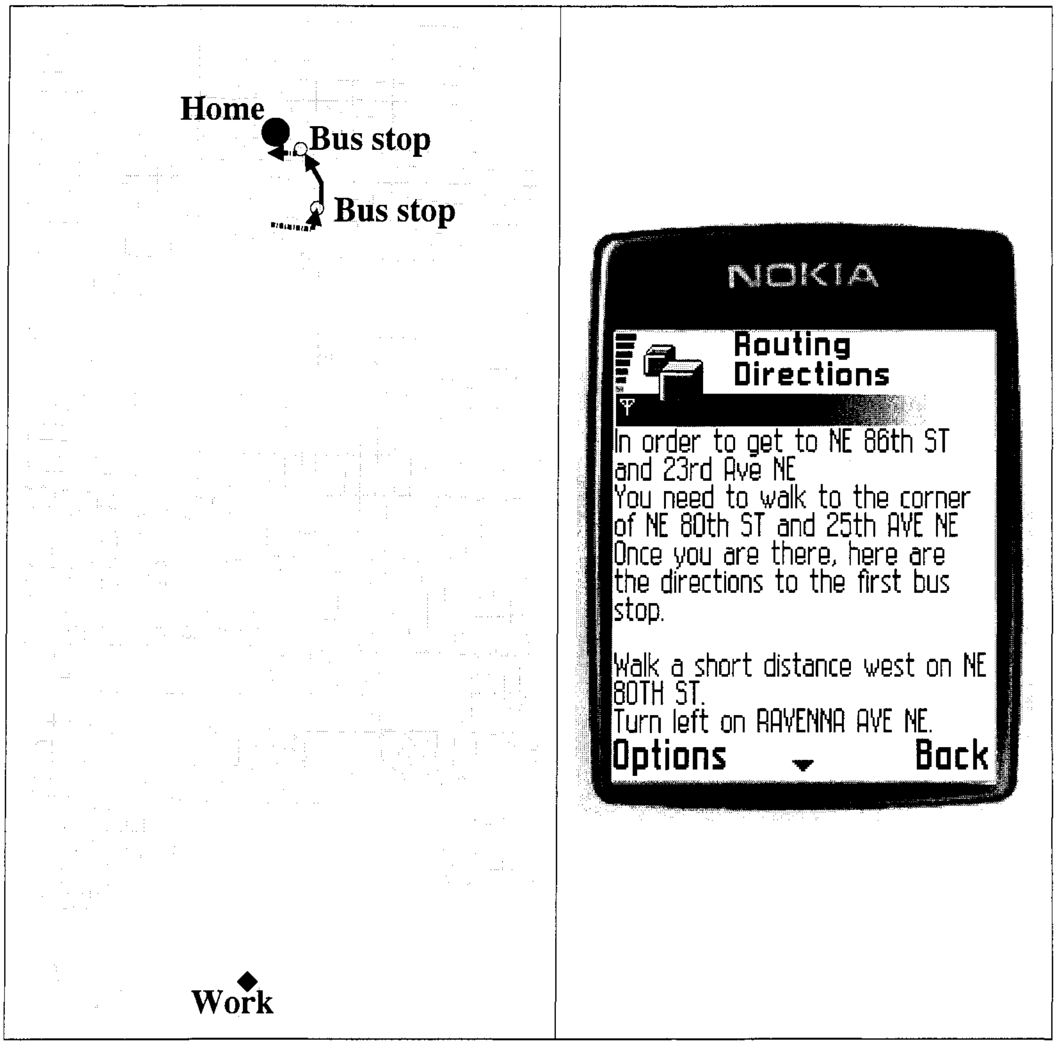


Figure 5.15: Experimental Results Part 6

Once off the incorrect bus, the user reselects home as the destination. This time the system has no history of the user ever going home from the current location. As a result OK queries a real-time bus planning system for a route home. The user is directed to walk back to the arterial road and catch a different bus that is going the correct way.

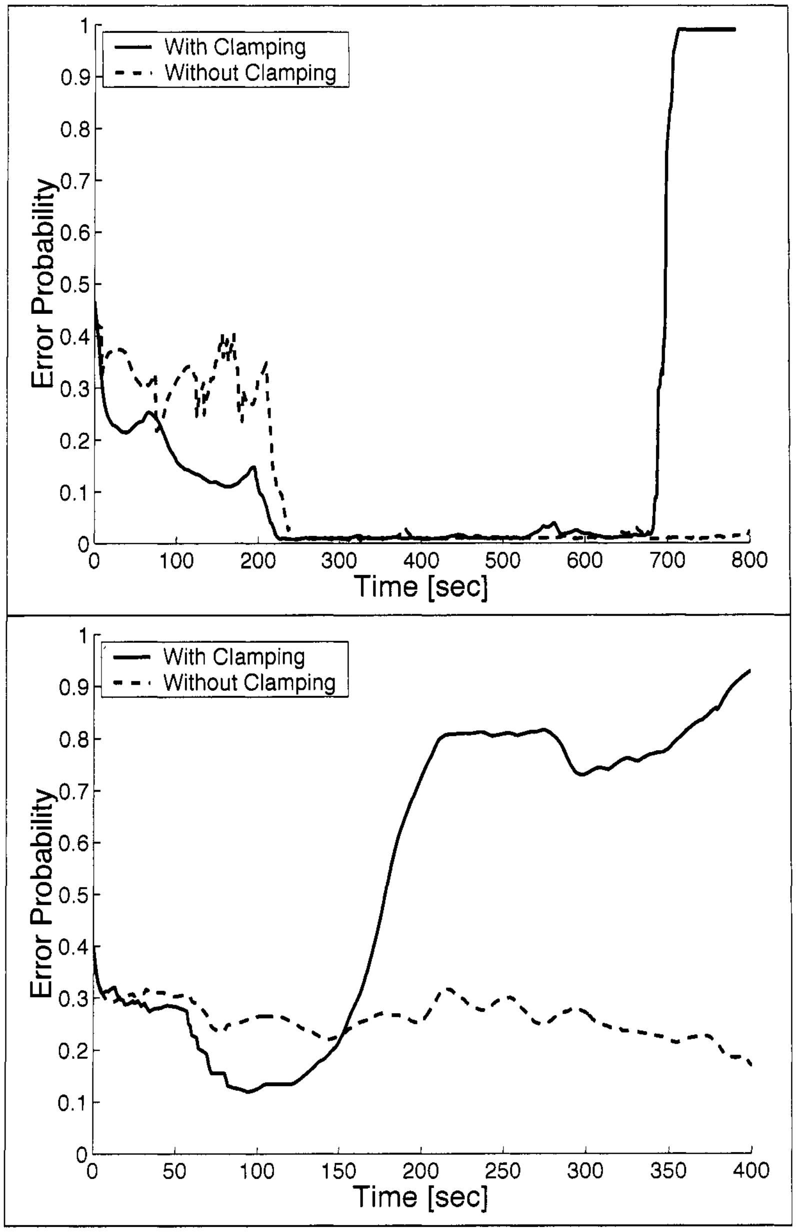


Figure 5.16: Impact of Model Clamping on Error Detection.

Probability of errors is shown in comparison to time in the presence and absence of a known destination. The dotted line shows the probability of errors when the user's destination is not known. The solid line shows the probability of errors when the destination is known (and clamped). The bus experiment is on the top (error made at time 700) and the foot experiment is on the bottom (with a gradual error beginning at time 125).

but unlike existing route planning systems, it is user-centric, not vehicle-centric and requires very little user input. Instead it relies on observed user history as a basis for predicting likely destinations and identifying novel and erroneous behavior.

Our system utilizes a Bluetooth GPS beacon that talks to a cell-phone, which in turn exchanges information with a remote inference engine. The software on the remote engine runs a new hierarchical dynamic Bayesian network that is able to explicitly reason about how high-level destinations will affect many levels of transportation decisions by the user, down to the street level.

We are able to use the camera function of the phone as a method of labeling places to eliminate the need for a user to manually translate positions to places before the system can communicate about them with the user.

Finally, we have experimentally shown that this system, in conjunction with real-time transit information, has promise for effectively providing transportation assistance in the face of mild confusion, memory lapses, and inattention.

Chapter 6

BARISTA: MODELING INDOOR ACTIVITIES

Up to this point we have argued that enabling outdoor activity recognition is a valuable and practical approach to take to support the independence of people who make cognitive errors. Indoor activity recognition also has many potential benefits. Many ADLs are themselves indoor activities and therefore enabling a computer to recognize them is the first step at creating technological aids for their performance. For example a computer that understands that you are done cooking can promote safety by alerting you that the stove is still on, but it could also adjust your dining room lighting, alert family members that dinner is ready or update a nutritional log with data about the meal. In this chapter we are going to look at the feasibility of performing fine-grained *indoor* activity recognition.

ADL monitoring is an ongoing, important activity in health care. For example, in the United States, any nursing home that receives Medicare funds has to record and report ADLs. Trained caregivers spend a great deal of time measuring and tracking ADL accomplishment for persons under their care. However, manual monitoring is time-consuming, error-prone, and invasive. Automated aids that can address these issues and reduce the record-keeping burden on caregivers are of great interest.

Previous work has approached activity recognition from several different directions. Early work on plan recognition [79, 150] had the insight that much behavior follows stereotypical patterns, but lacked well-founded and efficient algorithms for learning and inference, or any way to ground their theories in directly sensed experience. More recent work on plan-based behavior recognition [35] uses more robust probabilistic inference algorithms, but still does not directly connect to sensor data.

Most, if not all, systems that recognize home activities have been limited in the variety of activities they recognize, their robustness to noise, and their ease of use. In particular previous work on activity recognition has (i) used sensors that can provide only a very coarse idea of what is going on — for example, by detecting movement in a room, one might infer that an activity associated with that room is happening [53, 57]; (ii) required deployment of an extensive custom sensing apparatus to monitor each task [17, 99, 100, 133, 143]; (iii) and/or relied upon solutions to deep technical problems such as machine vision [72, 103].

We present an approach to activity recognition that addresses these problems in three ways:

- We provide a simple and flexible framework for modeling activities in terms of the manipulation of physical objects.

- Recent radio frequency identification (RFID) tag technologies have allowed wearable computers, in particular, to have unambiguous knowledge about the objects in their environment. We utilize this approach and treat RFIDs as the primary sensor.

- We employ efficient probabilistic reasoning algorithms that can quickly and robustly infer long sequences of activities from unsegmented data.

In this chapter we will examine the advantages and challenges of this framework for performing indoor activity recognition. We will present the results of two experiments. In the first experiment, we successfully applied our system to a known difficult problem from health care: recognizing multiple ADLs in a home environment. We used data from non-researchers performing activities in a real-home outfitted with RFID tags. In the second experiment we focused on the analysis and identification of what features of our model were valuable for activity recognition. We evaluated a sequence of increasingly powerful probabilistic graphical models for activity recognition using data collected from 10 mornings in the same household. Through this second experiment, we demonstrated the advantages of adding additional complexity and concluded with a model that can reason tractably about *aggregated object instances* and uses a technique called *abstraction smoothing* that generalizes gracefully from instances to object classes. Using these models we demonstrate situations in which both aggregation and abstraction are crucial for disambiguating activities. We then validate the use of these techniques on data collected from a morning household routine.

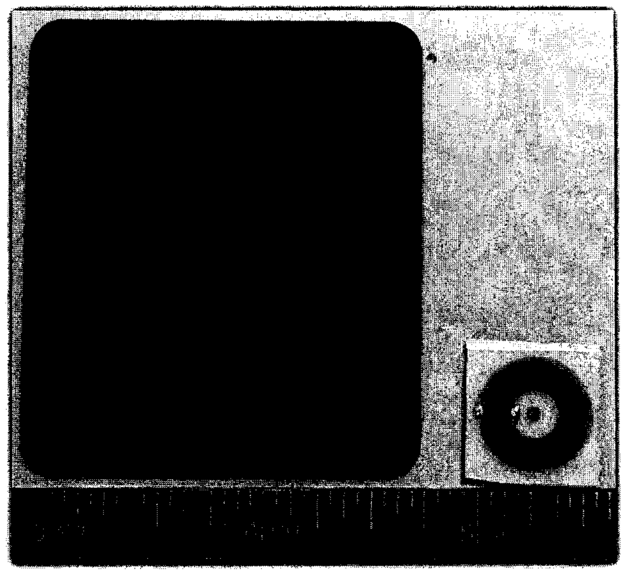


Figure 6.1: RFID tags

An example of the RFID tag types used in this experiment next to a ruler for scale.

6.1 *Vision and Background*

As we will argue in more detail below, RFID technology provides an appealing alternative to existing activity recognition technologies. RFID tags are tiny, inexpensive, semi-passive transceivers that can be attached to practically any object, and then sensed to determine when that object is moved or manipulated.

Our approach is therefore object-based activity recognition. It relies on the "invisible human hypothesis" of [119] which states that activities are well characterized by the objects that are manipulated during their performance, and can be recognized from streams of sensor data about object touches. In fact, every ADL we have been able to imagine involves the physical manipulation of one or more objects in the environment. For example, cooking involves utensils, appliances, and packages of food; dressing involves items of clothing; and so on.

6.1.1 Sensing using RFID

Central to this approach is the assumption of a world in which many consumer goods are tagged with RFIDs. Currently, except for a few trials, RFIDs are only utilized for supply-chain management at the palette level. However, recent advances in miniaturization and manufacturing have dramatically improved the functionality and reduced the cost of RFID transceivers. They can now be purchased off-the-shelf and cost roughly \$0.40 each. As such they are expected replace product barcodes in the near future [144]. Thus, much of the necessary infrastructure will exist "for free" in coming years.

The tags have the form factor of postage stamps (including adhesive backing), have no batteries and can withstand day-to-day use for years. Fig. 6.1 shows a few examples of the tags used in our experiments. Unlike barcodes they do not require line-of-sight in order to be read and they identify specific globally unique instances of objects, rather than merely classes of objects. For example, a bar code might identify a can of orange juice of a specific brand and size, but an RFID tag will identify which of the millions of those cans is being used.

RFID deployment involves tagging tens to hundreds of objects in the environment, and entering each tag identifier into a database. This can be done incrementally; the more tags, the broader and deeper the potential coverage of ADLs. As mentioned above, market forces are pushing toward the near-universal use of RFID tags on essentially all products, from clothing to foodstuffs.[1] Such pre-existing tags could then be used for applications such as ours by using global databases to map tag IDs to types of objects.

ADLs that would be difficult or impossible to detect using either coarse location sensors or state of the art machine vision can often be recognized on the basis of contact with a tagged object. For instance, consider trying to determine if a person is reading. Location alone is clearly inadequate, while reliably recognizing the act of "reading" from a video stream under a wide range of orientations, positions, and lighting conditions is far beyond the capabilities of machine vision for the foreseeable future. On the other hand, if all the

[1] The EPCglobal organization is coordinating the development of universal international standards for RFID tags; information about this effort and the state of the art of RFID technology can be obtained from http://www.epcglobalinc.org.

books, magazines, and newspapers in the home were tagged, determining when a person was reading could be done quite reliably.

We are working with a short-range RFID reader built into the palm of a glove that can determine the objects that the user is touching with a 0% false positive rate [144]. The prototype glove, built by Intel Research Seattle [68], is shown in figure 6.2. The antenna is connected to an Intermec RFID reader, which is packaged with a Crossbow Mica Mote radio, a USB-based power supply, and a rechargeable battery. All components except the antenna itself are housed in the small box on the back of the glove. The reader samples the environment once every half second; any RFID seen is broadcast to an HP iPaq 5400, utilized as a wearable computer. The iPaq either stores the data on-board for future analysis or forwards it via WiFi to the inference engine running on a workstation. The reader lasts for two hours at this duty cycle.

While it would be problematic for many applications to use such a glove, this is a temporary problem. When a portable RFID glove as a UI device was proposed three years ago [129], it measured roughly 160 cm^3, and was not wireless. Our latest version is a little over 30 cm^3, including wireless. The ultimate end system would require technology which is less obtrusive than a glove, for example a bracelet or ring which offered the same technologies. A different form-factor would enable activities which involve touching liquids and harsher environments and is feasible given the rate of progress in miniaturizing this technology.

6.1.2 Research Contributions

The research contributions of our first experiment include the following:

1. We develop a framework for modeling everyday activities that handles activities with partially-ordered substeps, continuous times, and probabilistic associations between substeps and the manipulation of physical objects. While the inference in the first experiment uses hand-crafted models (and indeed, we argue that basic models are easy to create in our formalism by simple introspection or traditional knowledge-engineering) we will briefly allude to our related work on learning models from natural

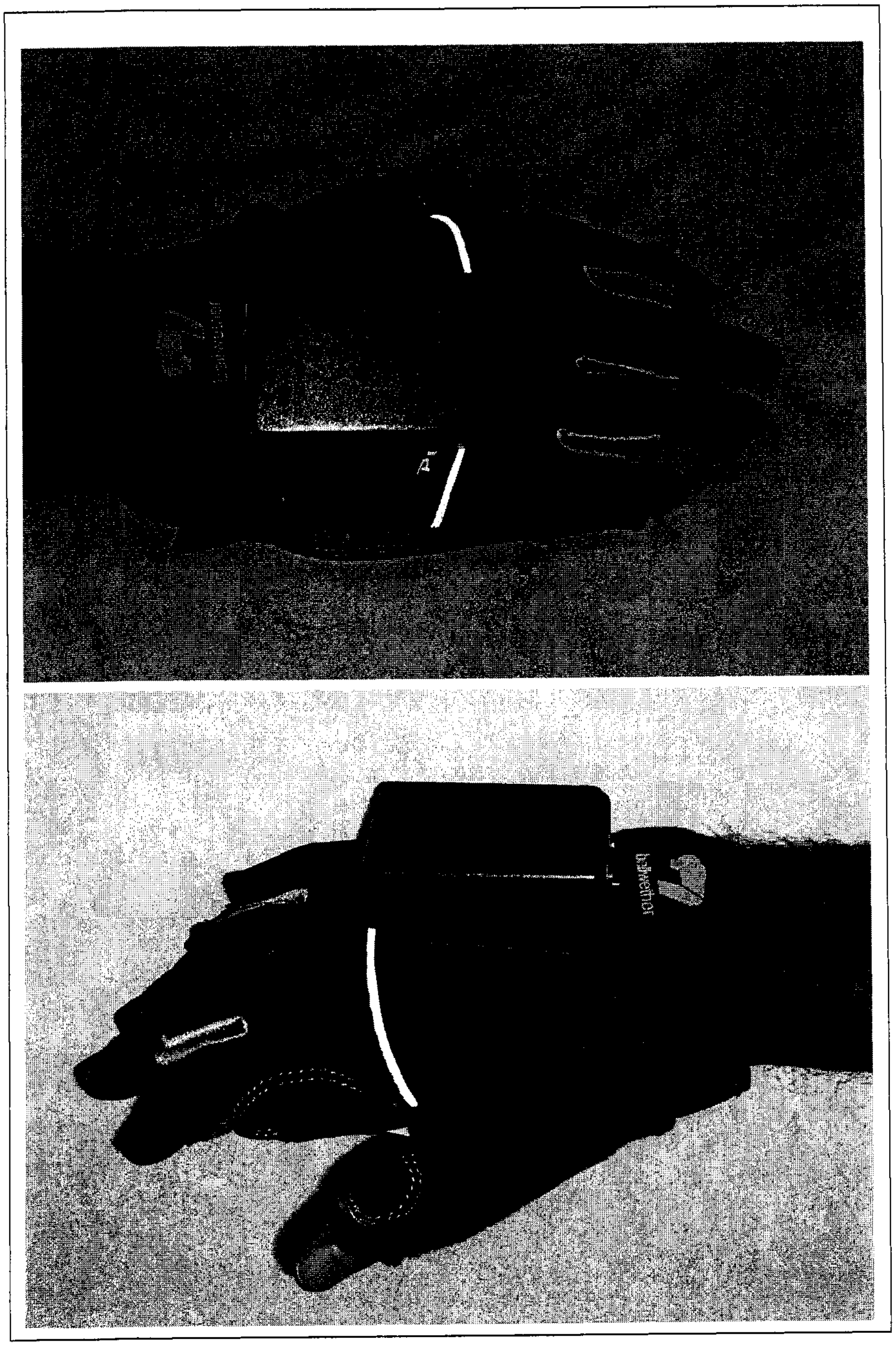


Figure 6.2: RFID glove

Two views of the right-handed RFID glove. The left-hand glove was symmetrically identical.

language descriptions, which has allowed us to created fine-grained libraries containing over 20,000 different activities. By avoiding intricate causal structure for activities while associating as many relevant objects as possible, our models are both robust and accurate.

2. We describe techniques for efficient implementing inference using particle filtering [107]. In order to efficiently handle the object probabilities we introduce a new re-sampling method based on estimating the expected probability of an activity based on possible future objects that are seen before the activity completes.

3. We deployed and evaluated our system in a real home over a six week period, including a final evaluation involving 14 non-researcher users. Our early results are promising. We show that nine ADLs, which no known prior work has addressed, were accurately inferred by our system. For four more ADLs, we were able for the first time to move beyond qualitative presentation to quantitative analysis.

In the second experiment we explore object-interaction based behavior recognition in a more challenging and realistic setting. We address the problem of activity recognition in the presence of complexities such as interleaved and interrupted activities. In particular, we put forward the following contributions related to accurately tracking and distinguishing activities based on the objects that are touched during their performance:

- We demonstrate accurate activity recognition in the presence of interleaved and interrupted activities.

- We accomplish this using automatically learned models.

- We demonstrate the ability to disambiguate activities when the models share common objects.

- We show that object-based activity recognition can be improved by distinguishing object instances from object classes, and learning dependencies on aggregate features.

- We introduce abstraction smoothing, a form of relational learning, that can provide robustness and prevent drastic loss in recognition accuracy in the face of expected variation in activity performance.

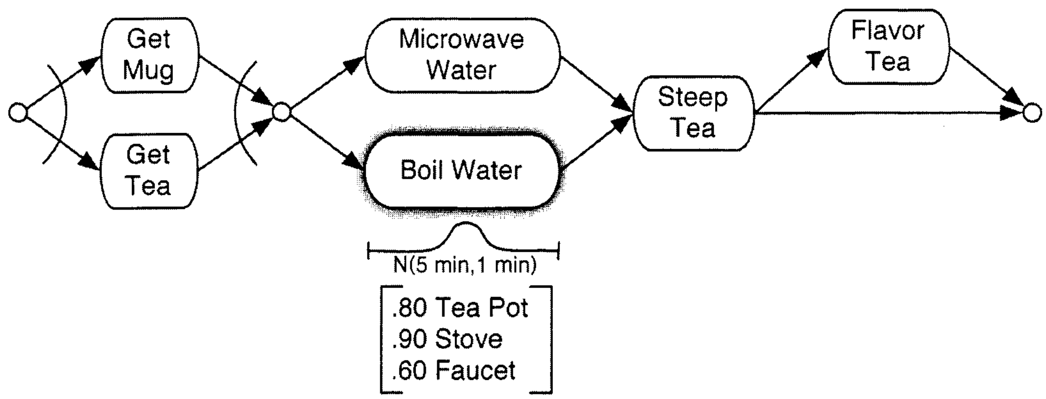


Figure 6.3: Making tea represented as an activity graph.

6.2 *BARISTA Experiment #1*

6.2.1 Modeling Activities

The goal of representing activities of daily living in terms of the gross manipulation of physical objects requires us to face the problem of developing a formal model that satisfies a number of constraints: first, the model should easily express significant properties of and distinctions between activities, while remaining robust to unimportant variations in activity performance; second, the parameters of the model should be be easily estimated; and third, the model should be implementable in a manner that supports efficient and scalable inference.

Descriptions of activities from a wide variety of sources, including healthcare literature, instructional manuals, and recipes typically break an activity down in a set of steps, where each step involves manipulating one or more objects over some period of time. Although textual descriptions usually present the steps in a total order, the underlying logical dependencies between steps often form only a partial order, and include alternative and optional steps. The kinds of objects used in a step is usually flexible, and it is not difficult to form a coarse estimate on the probability of object use on the basis of the description. For example, while making a cup of tea (used as running example in this section), we might estimate that the probability of using a spoon to stir the tea is 75%, allowing for cases where one uses a different utensil or none at all. Such considerations led us to develop the general model that is described in this paper.

Fig. 6.3 gives an example of modeling the activity "making tea" in four stages: getting out the supplies, heating the water, steeping the tea, and flavoring (*i.e.* adding sugar or lemon to the tea). The first stage consists of two steps which must both occur but in any order: this is indicated by a conjunctive arc across arrows the first pair of outgoing arrows and the following pair of incoming arrows. This is an example of a *partial ordering* constraint. The disjunctive choice of which of two ways to heat the water — using the microwave or using the stove — is indicated by a set of plain arcs. The fact that the flavoring step is optional is represented by a disjunctive arc that bypasses the step.

Each step also has a Gaussian duration. Duration information can provide important constraints for distinguishing activities that use similar objects. For example, washing your hands at the kitchen sink takes about a minute, while washing dishes at the kitchen sink takes about ten minutes. Finally, each step includes a set of objects that are expected to be used. (In Fig. 6.3 the duration and object information is only shown for the Boil Water step.) The value associated with each object is termed an *object use probability*, and is the estimate of the probability that the object is manipulated *at least once* before the step completes. Also included in the model but not shown in the illustration are prior probabilities on each activity as a whole and on choice transitions within a model (such as the probability of including the optional "flavor tea" step); by default these are uniform across choices.

Although the diagram is similar to a Markov model, it is differentiated by its use of partial orders and object use probabilities. This diagram also assumes small discrete time steps to approximate continuous time. While partial orderings could be handled by a disjunction of linear total orderings, the way object use probabilities are specified requires special attention. A Markov or semi-Markov model specifies an *instantaneous* probability that an observation is made at *each* time step. By contrast, object use probabilities give the chance that *at least one* observation of a given type is made over *all* the time steps that the agent is engaged in a particular step of an activity.

We believe that object use probabilities are more appropriate than instantaneous probabilities for modeling the kinds of activities we are concerned with for two main reasons: First, for many activities seeing an object more than once does not greatly change your intuitive belief about what is occurring. Second, as we argued above, reasonable values for completion probabilities are often easy to estimate and hold for a wide range of individuals. Using instantaneous probabilities, however, would require us to estimate the probability that an object tag is seen at each time point during the step. This value depends upon many factors, such as the number of times the object is touched, the duration of each touch, and esoteric details of the tag reader. Creating a useful and general apriori estimate of such a value is difficult. Conceivably, a system could learn such parameters for a particular individual and sensor apparatus, at the cost of gathering and labeling extensive training data for each user and the resulting model which is fragile to variations in routine.

The models of ADLs used in this experiment were created by hand. In related work we are investigating data mining techniques for automatically extracting activity models from natural language texts available on the Web [116]. Web sites such as "ehow.com" contain step-by-step instructions for tens of thousands of activities — ranging from complex mechanical operations (*e.g.*, how to repair a dishwasher) to simple ADLs (*e.g.*, how to brush your teeth). Lightweight syntactic parsing can be used to extract the steps and associated objects for an activity, and statistics on the co-occurrence of the names of activities and objects can be used as (rough) estimates of object use probabilities.

6.2.2 Inference

We now turn to the problem of computing the probability distribution for the current activity given the history of observations, also called *filtering*, and the related task of computing the most likely *sequence* of activities that explains the data.

Inference in activity graphs can be conceived of as a two-stage process. First, we will

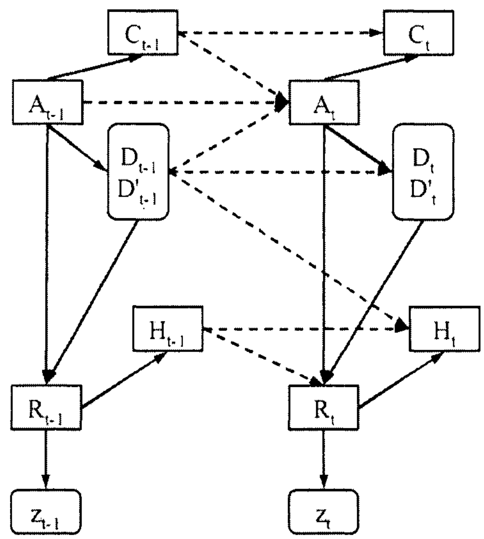


Figure 6.4: The dynamic Bayesian network encoding of an activity graph.

describe how an activity graph can be translated to a DBN (the same formalism as was used in "Opportunity Knocks"). With the DBN we specify the set of variables that describe the system's "state of belief" (that is, a joint probability distribution) at any point in time, and how those variables influence each other instantaneously and between adjacent time points. Second, we will describe how an approximate inference algorithm for DBNs can be efficiently implemented to solve both the filtering and most-likely-sequence tasks.

The reason for (conceptually) introducing the DBN is that it is a well-established formalism that provides a precise, abstract way to specify the probabilistic dependencies that should be respected by any method for implementing inference. The way in which these models capture timing constraints of activities are also well-characterized (e.g. [111]). Many exact and approximate inference algorithms (such as particle filtering) have been defined for DBNs, so we do not need to start from scratch in devising a way to reason about activity graphs. In practice, however, we do not need to explicitly translate the activity graph into a DBN; we can, in one step, turn the activity graph into the data structures used by the particle filter.

The DBN consists of a set of variables indexed by time point. Each variable can influence variables at the same and next time points. In concept the DBN extends infinitely into the future, but only two adjacent slices need to be explicitly represented at any one time.[2]

Fig. 6.4 illustrates the DBN structure for the translation of an activity graph. The variables are A, the current activity step; D', the total duration of the current activity step (from the past into the future); D, the time remaining in the current activity step; C, a set of "credits" used to handle partially-ordered activities (as described below); R, the objects currently being touched; H, the history of all objects that have been touched during the current activity step; and z, the current sensor readings (a set of RFID tags). The arcs specify the influences from one variable to another. A DBN is called a "generative" model, because it can be used to generate a probability distribution over observations (z). Inference means finding the most likely distribution over the hidden variables (such as A) given the actual observations.

Not shown in the figure are the conditional probability tables (CPTs) for each node, that give the probability distribution over values for the variables given the values of its parents (incoming nodes). Note that the graphical structure of the DBN shows us the general structure of influences between variables, but nothing about the particular activity model. The activity model is represented by the particular values stored in the conditional probability tables.

In order to complete our description of the DBN we need to describe the general form of CPTs so that the DBN respects the semantics of the activity graph. When an activity step A begins both D' and D are initialized to amount of time the step will require, by randomly sampling from the Gaussian associated with the particular value of A. D then counts down by one unit at each time slice. When D reaches 0, the activity step has completed and the value of A must switch at the next time slice, at which point D' and D are reinitialized. The arcs in the DBN between A_{t-1}, D_{t-1}, A_t, D'_t, and D_t indicate the logical dependencies in this process. When an activity includes partially-ordered steps, we allow a transition from a step within the partial order to any other step within the partial order whose immediate

[2]In this version of a DBN time is assumed to be discrete as some fine level of granularity. There are also versions of DBNs where time is fully continuous, and the "next" slice refers to the time of the next observation; since our sensors are clocked to take readings at fixed intervals, the difference is not significant.

predecessors have all been completed. The completed steps within a partial order are stored in the variable C. Thus when A_t switches, its new value may depend on C_{t-1} as well as on A_{t-1} and D_{t-1}.

The arc from R to z models the accuracy of the RFID tag reader: that is, the probability of missing a reading when an object is touched. The history H_t of objects touched during the current step is determined by its previous value H_{t-1} and the objects being touched at the current instance R_t. In the special case when a new activity begins — that is, when $D_{t-1} = 0$ — H_{t-1} is initialized to the empty set.

The final part of the network we must specify is R_t, the instantaneous probability of touching an object. As we argued above, we do not want to require the user to specify such instantaneous probabilities, but rather object use probabilities that are only defined for completed activity steps. But the abstract DBN model must allow us to compute $P(R_t|z_1, ..., z_t)$ since a DBN represents a complete joint distribution over all variables. To do this, we have to determine the conditional probability $P(R_t|H_{t-1}, A_t, D_t, D'_t)$, which is the probability of touching a specific object given the current activity, the relative time spent in this activity, and the objects touched since the activity started (cf. the DBN in Figure 6.4). Our system determines this value by assuming that the probability of touching an object for the first time during an activity depends on the object use probability of the activity and the time left in the activity. All values needed to compute this probability are contained in the variables H_{t-1}, A_t, D_t, and D'_t. If H_{t-1} already contains the object in R_t, then we assume that the probability of touching an object again is independent of the actual activity. As a result, the object use probability is counted only once during an activity; when the object is being touched for the first time. In our experiments we found that this approach successfully avoids the problem of generating overly peaked posterior distributions when an object is being touched for extended periods of time.

Now we turn to the particular inference algorithm we employed for the first BARISTA experiment: particle filters, an approach based on evolving a set of samples (particles) of the joint probability distribution over time. At any moment the probability of a particular activity step can be estimated by counting the number of particles that sample that step. The most likely sequence given all the observations can be estimated as the history of the most likely particle at the final time slice.[3]

In order to make particle filtering efficient we employ two techniques, one old and one new. The first is to reduce the number of particles needed by representing some of the variables as distributions, rather than single quantities. In particular, we can keep a distribution of the time remaining in an activity step by storing just the time at which the particle last switched to a new value for A. Then from the sample of A, the difference between the last-switch time and the current clock time, and the Gaussian associated with the activity step, one can calculate the probability distribution for D_t. At each time slice a particle will *split* into two particles, where one particle stays in the current activity step and the other progresses to a next activity step; the *weight* of each particle is determined by the distribution over D_t. This simple technique immediately reduces the number of particles needed for accurate inference by more than an order of magnitude (from 20,000 to 1,000).

The second technique to improve efficiency is based on the observation that it is not necessary to re-weight and re-sample every particle at every time slice. Instead, we simply update the observation H based on any new sensor readings, but do not re-weight or re-sample until one of the following occurs:

- When a particle "splits" and switches to a new value for A, the new particle is re-weighted according to the object use probabilities for the previous activity and history.
- When the number of particles exceeds a threshold (due to particles splitting), re-weight and re-sample all particles according to the *expected probability* that the particle will have at the completion of the current step.

We call this technique *sporadic re-sampling.* With it we can process the data from a hour of sensor data from the experiments described below in a few seconds.

6.2.3 Experiment #1

Of the more than 20 ADLs that caregivers typically monitor, we chose 14 for evaluation. We eliminated the rest due to the structure of our experiment, where subjects performed

[3] By "history" we mean the sequence of all values assumed by each variable in the particle over its complete life — not the variable H, the observation history.

Table 6.1: Tested ADLs and their descriptions.

No.	ADL	Task Sheet Description
1	Personal Appearance	Please touch up your personal appearance.
2	Oral Hygiene	Please take care of your oral hygiene as if you are about to go to bed.
3	Toileting	Please use the toilet.
4	Washing	Please wash up.
5	Housework	Please clean something.
6	Safe use of Appliances	Please use an appliance.
7	Use of heating	Please adjust the thermostat.
8	Care of clothes and linen	Please use the washer and/or dryer.
9	Preparing simple snack	Please prepare something to eat.
10	Preparing simple drink	Please make yourself a drink.
11	Use of telephone	Please use the Telephone to get your horoscope.
12	Leisure Activity	Please watch TV, read, play cards, or listen to music for a few minutes.
13	Caring for an infant	Please care for the baby.
14	Taking Medication	Please take some pills.(These were candy pills, of course.)

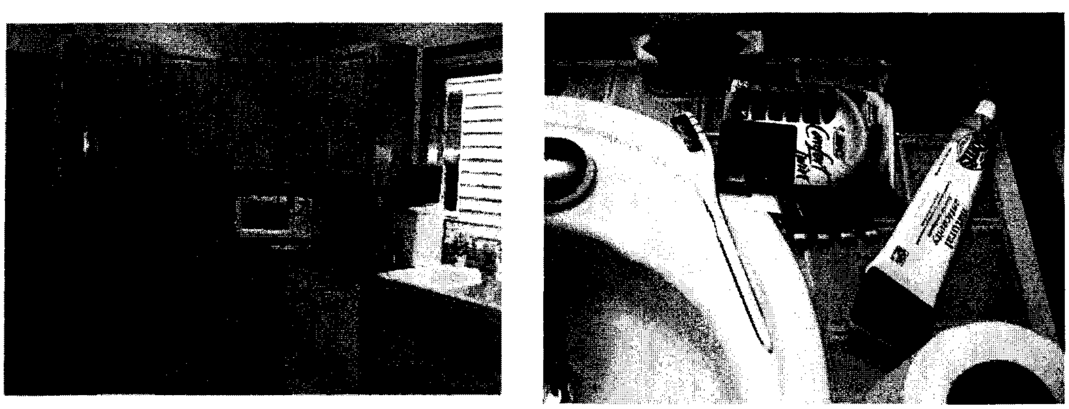


Figure 6.5: Left: kitchen of experiment home. Right: examples of tagged items in the bathroom: after-shave lotion, toothpaste, razors.

the tasks in another person's home, so ADLs such as "bathing" could not feasibly be tested. Others were omitted because they required travel outside the home (*e.g.*, "shopping"). We stress that it was the focused nature of the experiment, not any limitation of the toolkit, which led to this exclusion. The 14 ADLs we tested (Table 6.1) are, to our knowledge, 11 more than any other system has attempted.

We felt it vital to use a real home that was subject to the wear and clutter of daily living. The home belonged to a colleague, where he lived with his wife and their 2-year-old child. We instrumented it with 108 tags in a few hours. We did this before the activity models were written, so as to not be biased by its later designation of "key" objects. By tagging as many objects as possible, we also hoped to avoid having subjects feel they were "steered" to a narrow set of objects. Fig. 6.5 shows the kitchen area with some tags circled, and a close-up of tagged objects in the bathroom.

At various times over the next 6 weeks we tested 14 subjects (3 male and 11 female, with ages ranging from 25 to 63). We did not include the homeowner or his family in the test group. Although the test subjects were not in their own homes, this arrangement gave us a practical compromise between testing on a broad range of different people and the desire to have the same baseline environment for this initial study.

After a tour of the house and a demonstration of the RFID glove (Fig. 6.2), subjects were informed of the nature of the experiment. Subjects were given a package of 14 task sheets, one for each ADL. Table 6.1 shows the tasks as described on each sheet. The task sheets also had pictures of where some objects could be found in the house to avoid "treasure hunts". We kept activity descriptions as broad and as close to the medical description as possible.

Subjects randomly selected 12 of the 14 tasks. They then went into the house, and performed those 12 tasks, in any order, without observation. This took from 20 to 60 minutes, depending upon the subject. Subjects could engage in other household activities as they wished. As they touched tags in the course of performing their tasks, they would hear a "beep", as the system indicated that it had recorded a tag touch. Since this round of experiments was not meant to test the efficacy of the glove itself, subjects were asked to touch something several times if they saw that they were touching a tagged object and did not hear a beep. This was sometimes necessary due to the prototypical nature of the glove; however, only about 1 touch of out 60 needed to be repeated. While performing tasks, subjects wrote on a sheet of paper which task they were doing. After leaving the house, subjects gave one experimenter the sheet. This was kept separate from the other experimenters until after the results were processed.

As an additional test of the robustness of the system, the tags were left in the house while subjects were not present. With the exception of a few tags that were awkwardly placed on edges all tags stayed usable throughout the entire experiment.

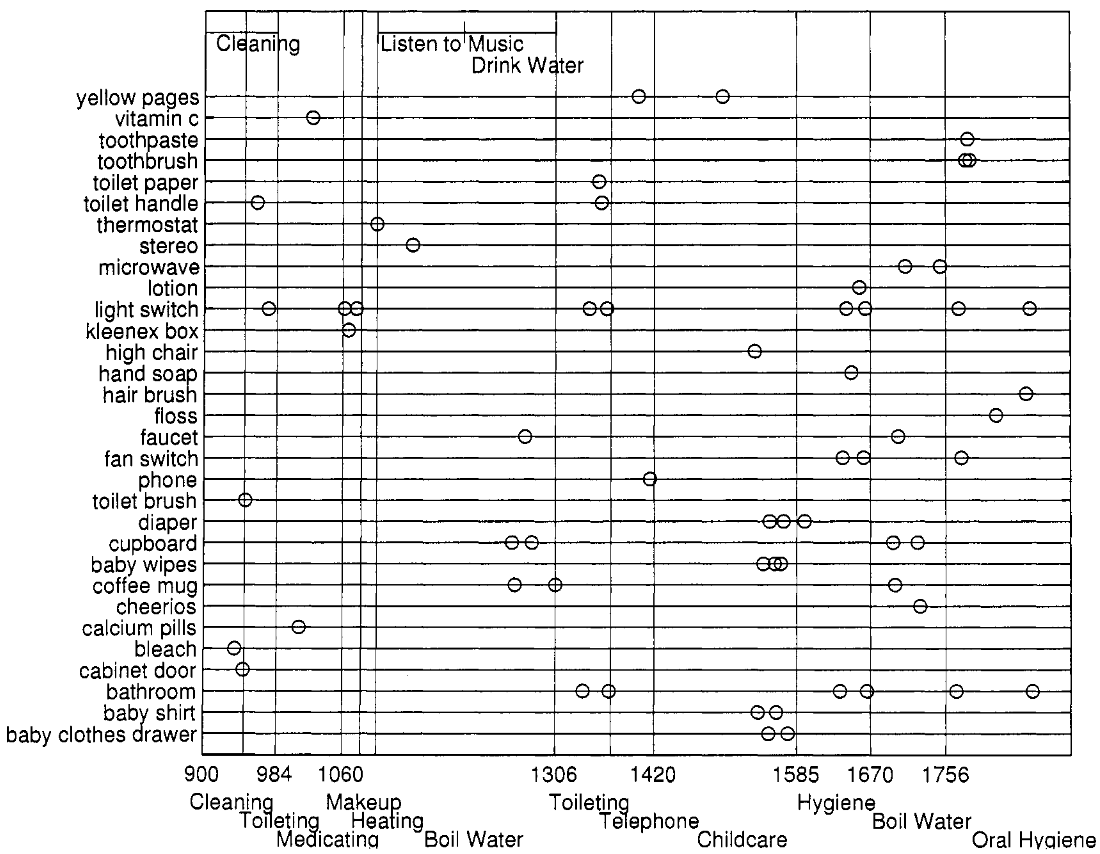


Figure 6.6: An recorded trace of a user in a home. The horizontal axis is the time in seconds (beginning at timestamp 900). The vertical axis lists the tags that were recorded for this particular session, and a circle indicates the time of a tag reading. The horizontal lines and captions at the bottom of the graph indicate the segmentation and most likely activity sequence inferred by our system, and errors (missed or mistaken activities) are noted at the top of the graph.

Table 6.2: Summary of in-home experiment. Key: TP = true positive, FP = false positive, FN = false negative.

ADL No.	TP	FP	FN	Precision	Recall
1	11	1	1	92	92
2	7	3	2	70	78
3	8	3	3	73	73
4	3	0	6	100	33
5	8	0	4	100	75
6	21	4	6	84	78
7	8	0	3	100	73
8	7	0	2	100	78
9	6	2	4	75	60
10	9	5	5	64	64
11	11	0	3	100	79
12	7	0	5	100	58
13	13	0	1	100	93
14	9	0	2	100	82
TOTAL	**128**	**18**	**47**	**88**	**73**

We gave the complete tag sequence for each subject (without manually segmenting between activities), along with models for the 14 activities to the inference engine. The engine returned a log of the most likely sequence of activities that would explain the readings. Fig. 6.6 illustrates the recorded trace and inferred activity sequence for one of the trials.

We compared the logs generated by the system with the activity sequence reported by the subjects immediately after the experiment, which was treated as ground truth. A true activity, correctly, recognized is counted as a true positive (TP); an inferred activity that did not actually occur is a false positive (FP); and an unrecognized true activity is a false negative (FN). We then computed the standard metrics of precision and recall. Precision is TP/(TP+FP), and recall is TP/(TP+FN). Precision and recall are termed "positive predictive value" and "sensitivity", respectively, in the medical community.

Table 6.2 summarizes the results. As one can see, the system did well on average, with a precision of 88% and recall of 73%. Given the ambiguous and overlapping activity definitions (consider the ADLs personal appearance (#1), oral hygiene (#2), and washing up (#4)), we believe this is strong validation for our basic approach.

To place these numbers in further context, note that while ADL inferencing has often been investigated, 9 of the ADLs inferred here are inferred for the first time. For 4 more (meal preparation, toileting, heating control, and medication taking), this is the first time that any quantitative results are reported. For the one ADL where previous quantitative results were presented, hand washing, our system's precision/recall of 100/33% are below the 95%/84% reported by Mihailidis [99]. However, that work targets the single activity of hand washing, and requires cameras to be installed in the bathroom. The performance on hand-washing is actually our worst case.

We now discuss the reasons for this result, and a few others of interest. The radio waves used by most RFID tags are absorbed by water and metal; metal can also attenuate the tag antenna. These factors cause the detection rate to plummet for tags which are too close to those substances, which particularly affected the activities of washing hands (#4), making a snack (#9), and making a drink (#10). Using newer RF tags which are optimized for placement near liquids and metal could mitigate this.

In some cases, the model involves so few observations as to not be easily distinguishable from noise. For example, the only observable for playing music (activity #12) was the stereo. The single observation that results is not enough to convince the activity inference engine that a new task has begun. Adding more relevant tags could solve this (*e.g.*, tagging the CDs).

Activities with common prefixes posed a more subtle problem. For instance, activities for personal appearance (activity #1), oral hygiene (#2), toileting (#3) and washing (#4) all begin with entering the bathroom and possibly turning on the light. Our models replicate the nodes for these sub-activities for each of the four activities. When bathroom and light are detected, each of these activities are equally likely (nominally 25%). When disambiguating objects are then seen (*e.g.*, toothbrush and toothpaste), the inference engine concludes that the oral hygiene activity is the correct one. If the user then uses the toilet without leaving the bathroom first, the inference engine fails to detect the second activity. To solve this problem, we are working on an extension to the activity model that allow different activities to share sub-steps.

Although it is possible for our approach to learn parameters without supervision (in particular, duration of sub-activities) using techniques such as expectation maximization (EM) [120], the initial experiments we report here did not involve learning. For example, the duration our model specified for boiling water was too short. As a result the system would sometimes prematurely conclude that boiling was complete, and erroneously jump ahead to the most likely next activity — yielding a false positive.

6.2.4 Summary of Experiment #1

The first experiment in this chapter demonstrated first steps toward building practical and well-founded systems for monitoring a wide variety of activities of daily living in the home. It builds upon the AI community's success in applying dynamic Bayesian networks and particle filtering inference algorithms to behavior recognition problems, together with recent advances in sensing technology. This experiment provides support for the hypothesis that we can engineer robust and general probabilistic models of high-level behavior by focusing on the manipulation of physical objects.

As we have noted before, a major part of our overall project that is not dealt with in this first experiment is learning activity models from sensor data (as opposed to knowledge engineering or mining models from natural language texts). In the next section we will discuss techniques to learn parameters for simpler models automatically.

6.3 BARISTA Experiment #2

In this second experiment we take a step back from the previous focus on accuracy of ADL inference to try and understand what components of a probabilistic model are most valuable for this task. The previous experiment had the capability to monitor many aspects of an activity, but surprisingly didn't necessarily benefit from all the flexibility. In particular the duration modeling in the hand-models was uniform across all steps in all activities and still recognition accuracy was very high.

Table 6.3: Activities performed during data collection

1	Using the bathroom
2	Making Oatmeal
3	Making Soft-Boiled Eggs
4	Preparing Orange Juice
5	Making coffee
6	Making tea
7	Making or answering a phone call
8	Taking out the trash
9	Setting the table
10	Eating Breakfast
11	Clearing the table

6.3.1 Methodology

Problem Domain

For the purpose of this second experiment, we focused on a morning routine in a small home. In order to investigate the properties that we were interested in, activities were chosen such that they shared objects and by their ability to be naturally interleaved during execution. Table 6.3 lists the 11 activities which were included. To create our data set, one of the authors performed each activity 12 times in two contexts: Each activity was performed by itself twice, and then on 10 mornings all of the activities were performed together in a variety of patterns.

In order to capture the identity of the objects being manipulated, the kitchen was outfitted with 60 RFID tags placed on every object touched by the user during a practice trial. The list of tagged objects is shown in table 6.4.

In this experiment the user simultaneously wore two RFID gloves (unlike in the first experiment in which one glove was used). The time and id of every object touched was sent wirelessly by the glove to a database for analysis. The mean length of the 10 interleaved runs was 27.1 minutes ($\sigma = 1.7$) and object touches could be captured at approximately 10 per second. The mean length of each uninterrupted portion of the interleaved tasks was 74 seconds. Most tasks were interleaved with, or interrupted, by others during the 10 full data

Table 6.4: Tagged Objects

The 60 objects tagged for this experiment. Parentheses denote multiple instances.

bowl, coffee container, coffee grinder, coffee tamper, cupboard(6), dishwasher, door(2), drawer(2), egg carton, espresso cup(2), espresso handle, espresso steam knob, espresso switches, faucet(2), freezer, milk, hand soap, juice, juice pitcher, kettle, measuring cup-half, measuring cup-one, measuring scoop, milk steaming pitcher, mug, oatmeal, refrigerator, salt, saucepan, cooking spoon, stove control(2), sugar, table cup(4), table plate(4), table spoon(4), tea bag, tea box, telephone, toilet flush handle, toilet lid, vanilla syrup

collection sessions.

This data represents a significant increase in inference difficulty over experiment #1 by introducing three confounding factors:

- We collected data in the presence of multiple people. Although only one was instrumented, four people were present in the home during all data collection adding uncontrolled variability to the data collection.

- The activities were not performed sequentially or in isolation from each other. Whenever there was a pause in an activity, progress was attempted in other activities (such as when waiting for water to boil) and some activities interrupted others at uncontrolled times (such as answering the phone).

- The activities that the user performed had objects in common. This made interleaved activity recognition much more difficult than simply associating a characteristic object with an activity (such as a vacuum cleaner indicating a vacuuming activity)

6.3.2 Models Which Improve Accuracy

In order to justify the inference model that we ultimately developed we proceeded systematically by first focusing on *accuracy* and then on *robustness.* We developed the simplest possible probabilistic model, evaluated its performance and then augmented it with features that were sufficient to disambiguate errors. In this section we will present the techniques we used to improve *accuracy* by describing the three baseline models that we used and a fourth model that incorporated reasoning with aggregate features. The models increase in complexity by adding representational power. In subsequent sections we will present abstraction techniques that we used to improve *robustness.*

Baseline Model A: Independent Hidden Markov Models

As a baseline we modeled the activities individually as 11 independent one-state Hidden Markov Models (HMMs) (see figure 6.7). Used in a generative context, each state emits an *object-X-touched* event or a *no-object-touched* event at each tick of the clock. Each state's emission probability was trained on the 12 examples of a user performing the corresponding activity. After training, the probability of emitting a *no-object-touched* event was equalized across all HMMs so that the timing characteristic of the model was completely captured by the self-transition probability.

To infer the activity being performed at each second, each HMM was presented with a 74 second window of data ending at the query second (This was the mean amount of atomic time spent performing any portion of an activity in the training data.) This produced a log-likelihood value for each model at each tick of the clock. The activity model with the highest log likelihood was used as the system's estimate of the current activity.

This model was trained and tested on data in which object types were equalized so that there was no distinction made between spoon #1 and spoon #2, for example, but both appeared identically as a "spoon".

Baseline Model B: Connected HMMs

As a second baseline we connected the states from the 11 independent HMMs of baseline A in order to be able to learn about and subsequently smooth the transitions between activities. We retrained this HMM using the 10 examples of the user performing the 11 interleaved activities. The *no-object-touched* emission probability was again equalized across all states (see figure 6.8).

This HMM was evaluated over the entire data window and the Viterbi algorithm [120]

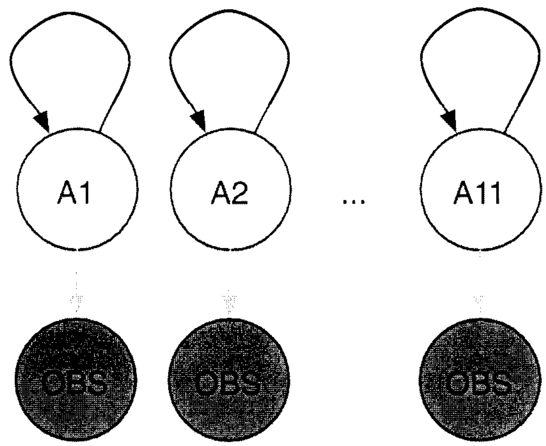


Figure 6.7: Model A

A state diagram for baseline model A consisting of 11 independent one-state HMMs. The log likelihood of each HMM was calculated on overlapping successive short windows of data.

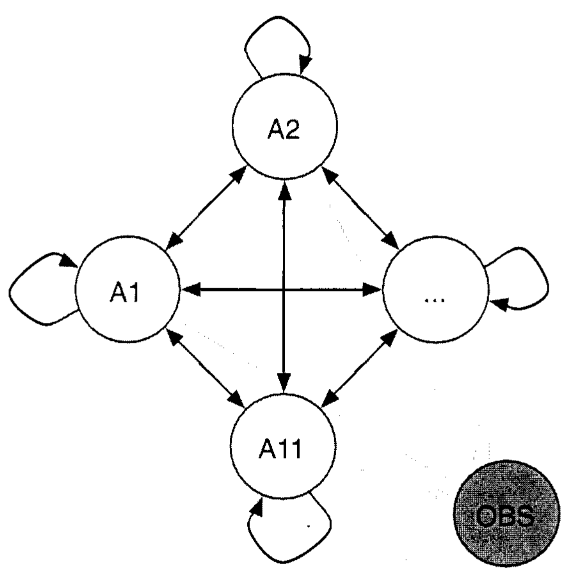


Figure 6.8: Model B

A state diagram for baseline model B consisting of an 11 state HMM. At any moment in time the user is in exactly one of the 11 states. The most likely state sequence was recovered using the Viterbi algorithm over the entire data sequence.

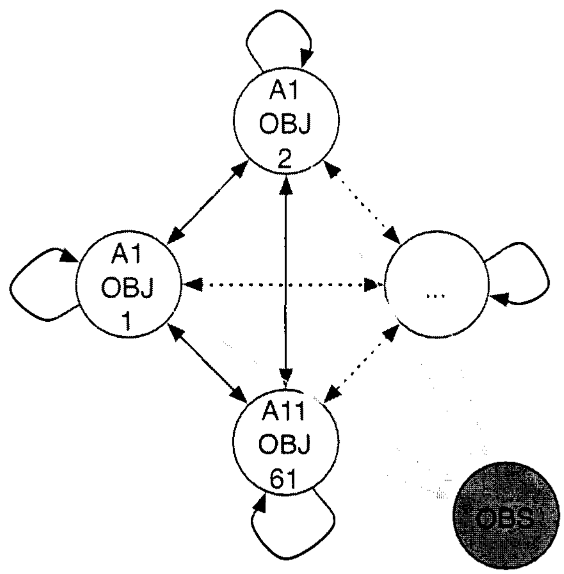


Figure 6.9: Model C

A state diagram for model baseline C consisting of an 671 (61 tags x 11 activities) state HMM. At any moment in time the user is in exactly one of the 671 states. The most likely state sequence was recovered using the Viterbi algorithm over the entire data sequence.

was used to recover the activity at every time point given by the maximum likelihood path through the state-space. Again, this model was trained and tested on data in which object types were equalized to eliminate distinctions between instantiations of objects.

Baseline Model C: Object-Centered HMMs

As a third baseline we split the states in baseline B into a separate state for each activity and each object that may be observed. This allowed this model to capture some information about how objects were used at different points in the execution of an activity (i.e. early versus later) at the expense of introducing more trainable parameters. We retrained this HMM using the 10 examples of the user performing the 11 interleaved activities. The *no-object-touched* emission probability was equalized across all states. The conditional probability table associated with state observations was degenerate for this model since each state could only emit an observation of one particular object or *no-object-touched* (see figure 6.9).

This HMM was also evaluated over the entire data window using the Viterbi algorithm and was trained and tested on data with object types equalized.

Aggregate Model D: Dynamic Bayes Net with Aggregates

For our fourth model we chose to examine the effect of reasoning about aggregate information. The specific feature that we wanted to model was how many objects of a given type were touched during the course of the current activity. This aggregate can only be computed if globally unique object instances can be identified. This choice was motivated by the desire to differentiate setting the table from eating breakfast. Both of these activities generate a sensor stream in which many of the same objects are touched (a spoon for example), but the difference is captured in whether one spoon is touched four times or four spoons are touched once.

Figure 6.10 shows a DBN which accomplishes this. Unlike figures 6.7, 6.8, and 6.9, which are state diagrams, this figure is a dependency diagram in which the state of the system is factored into independent random variables indicated by nodes in the graph. It has been rolled out in time showing nodes from three different time steps and dependency diagrams for models A-C are also shown for comparison. The top row of model D has the same dependency structure as models A-C.

Model D adds an additional deterministic state variable, the boolean exit node "E", which captures dynamic information about when an activity has changed. It is true if the state has changed from one time step to another.

The gray box denotes an aggregation template which is instantiated for each class of objects with multiple object instantiations. Each "Obj" node in the template is a deterministic boolean node indicating whether a given instantiation of an object has been touched since the last time the activity changed. The "Obj" nodes are aggregated by a summation node, "+". When the DBN changes activities ("Exits"), the system explicitly reasons about the number of different instantiations of the objects in the template that were touched. This is captured by the dependence of the aggregate distribution node, "AD", on

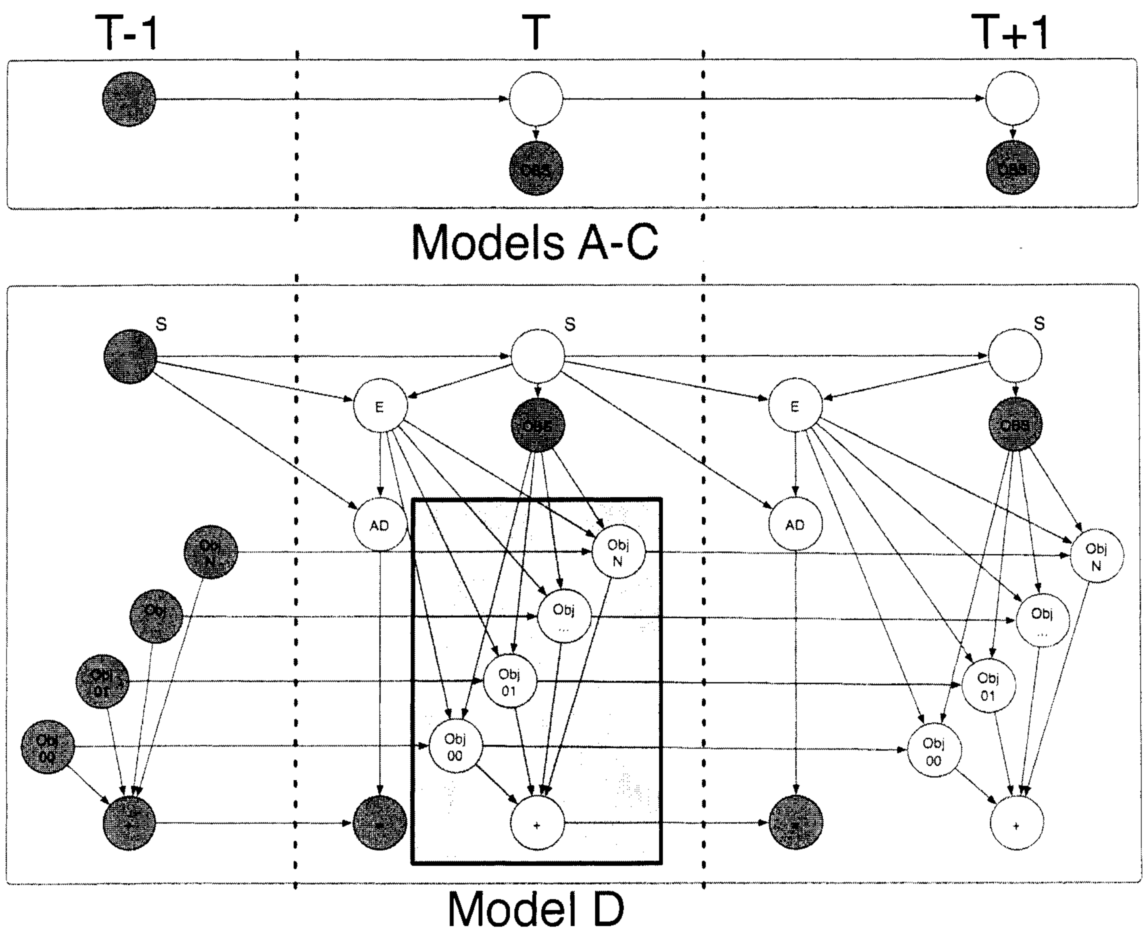


Figure 6.10: Model D

A dependency diagram showing the DBN used for inference in comparison to the baseline models, rolled out for three time steps. Time steps are separated by a vertical dotted line. Observed variables are shaded, hidden variables are not. All variables are discrete multinomials.

Table 6.5: Summary of model representational power .

	Exponential Timing Distribution	Inter-Activity Transitions	Intra-Activity Transitions	Aggregate Info.
Model A Ind. HMMs				
Model B Conn. HMMs	√	√		
Model C Object HMMs	√	√	√	
Model D Agg. DBN	√	√		√

the state node,"S". The "AD" node is constrained to be equal to the summation node from the previous time step through the observed equality node, "=". The equality node is always observed to be true and coupled with its conditional probability table force "AD" to equal "+".

The overall effect is that the probability of the model inferring whether or not an activity has ended is mediated by an expectation over the number of objects in each class which have been touched.

This model, including the aggregate distribution for activities, was also automatically learned from the 10 labeled training traces, but in contrast to models A, B and C, this model differentiates between instantiations of objects so that, for example, spoon #1 and spoon #2 created different observations.

Model Summary

The various features of these models are summarized in table 6.5. In this table "Exponential Timing" distributions refer to the fact that the model expects the length of an uninterrupted portion of an activity to occur with a duration which is distributed according to an exponential distribution. The parameters of the distribution are learned from the data. This timing distribution is a result of the structure of the HMMs and DBNs used.

"Inter-Activity" transitions refers to the ability of the model to represent the tendency of certain activities to follow or interrupt other activities more or less often. "Intra-Activity Transitions" refers to the ability of the model to represent biases about when in the course of an activity certain objects are used. (e.g., one uses a kettle early in the process of making tea). Finally "Aggregate Information" refers to the ability of the model to represent aggregations over individual objects.

6.3.3 Accuracy Experiments

Our accuracy experiments were conducted with leave-one-out cross validation across the 10 interleaved runs. We calculated two accuracy metrics. The first was what percentage of the time the model correctly inferred the true activity. This metric is biased against slight inaccuracies in the start and end times and will vary based on the time granularity with which the experiments were conducted. We also evaluated our models using a string edit distance measure. In this case we treated the output of the inference as a string over an 11 character alphabet, one character per activity, with all repeating characters merged. We calculated the minimum string edit distance between the inference and the ground truth. A string edit distance of 1, means that the inferred activity sequence either added a segment of an activity that didn't occur (insertion), it missed a segment that did occur (deletion), or it inserted an activity that didn't occur into the middle of an activity (reverse splicing). A perfect inference will have a string edit distance of 0. The string edit distance is biased against rapid changes in the activity estimate and is tolerant of inaccuracies in the start and end time of activities. Table 6.6 summarizes the results of experiments.

Baseline Model A

Figure 6.11 shows a short portion of the inference in which model A performed badly and demonstrates some of the key reasons why inference with this model only obtained 68% accuracy. In this figure ground truth is indicated with a thin line and the inference is indicated by a dot at each time slice. The model clearly needs to be smoothed to prevent rapid switching between activities. This was the primary motivation for using model B.

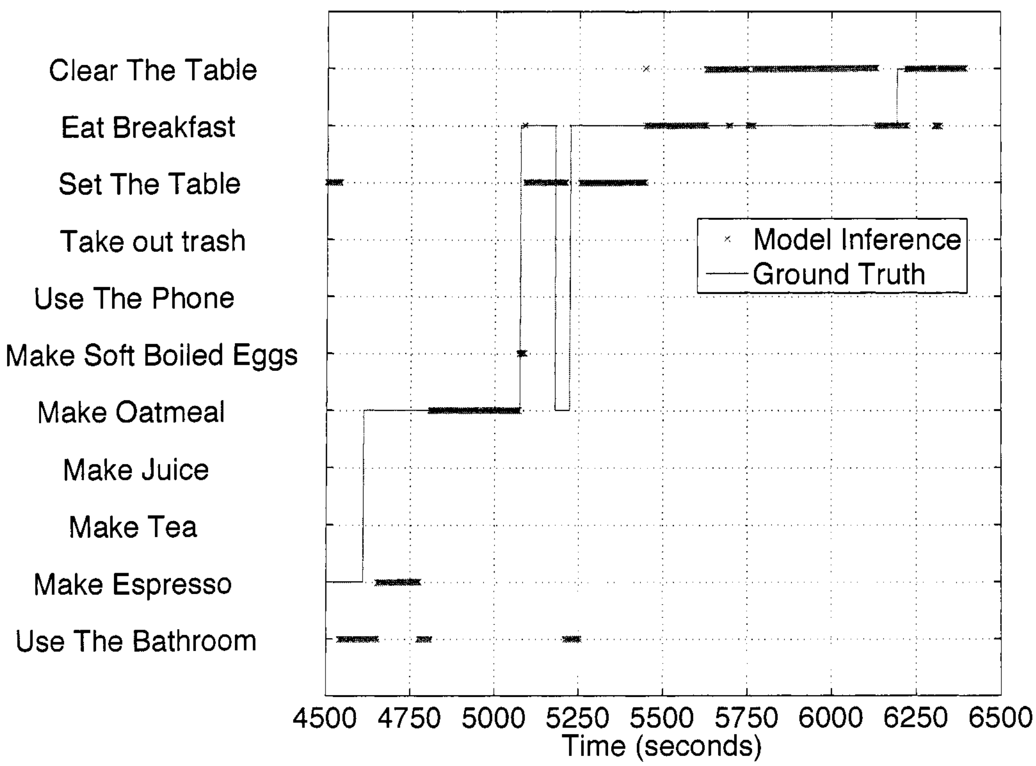


Figure 6.11: Model A Results (Independent HMMs)

Small segment of inference with various models. Ground truth is indicated by the thin line. Inference is indicated by the dots.

Table 6.6: Summary of accuracy results

	Time-Slice Accuracy μ (σ)	Edit Distance μ (σ)
Model A Ind. HMMs	68% (5.9)	12 (2.9)
Model B Connected HMMs	88% (4.2)	9 (6.2)
Model C Object HMMs	87% (9.3)	14 (10.4)
Model D Agg. DBN	88% (3.1)	7 (2.2)

Baseline Model B

Figure 6.12 shows a short portion of the inference performed by model B. In this graph the benefits of learning a distribution over the inter-activity transitions is displayed in the relatively smooth trace. The tendency of model A to jump between activities was eliminated. There are two places in the figure however in which the model inferred the wrong activity. In both cases the activity "Eat Breakfast" is confused with the activity "Clear the Table". These two activities weren't differentiable by the objects that were used in their performance: both activities used plates, spoons and cups. Two possibilities emerged for representations in which these two activities could be disambiguated. The first was to allow the model to learn a tendency for objects to be used earlier or later in the activity – this inspired model C. Second was to allow the model to differentiate activities based on the number of objects which were touched – this capability inspired model D.

Object Model C

Figure 6.13 shows a short portion of the inference performed by model C. For comparison, the same segment of time is shown as in the other figures. However, this model did considerably worse, according to the string edit distance metric, in areas of inference not shown in the figure. What the graph demonstrates however is that this model is able to correct for the errors of model B. It clearly makes an inter-activity error by inferring a transition from

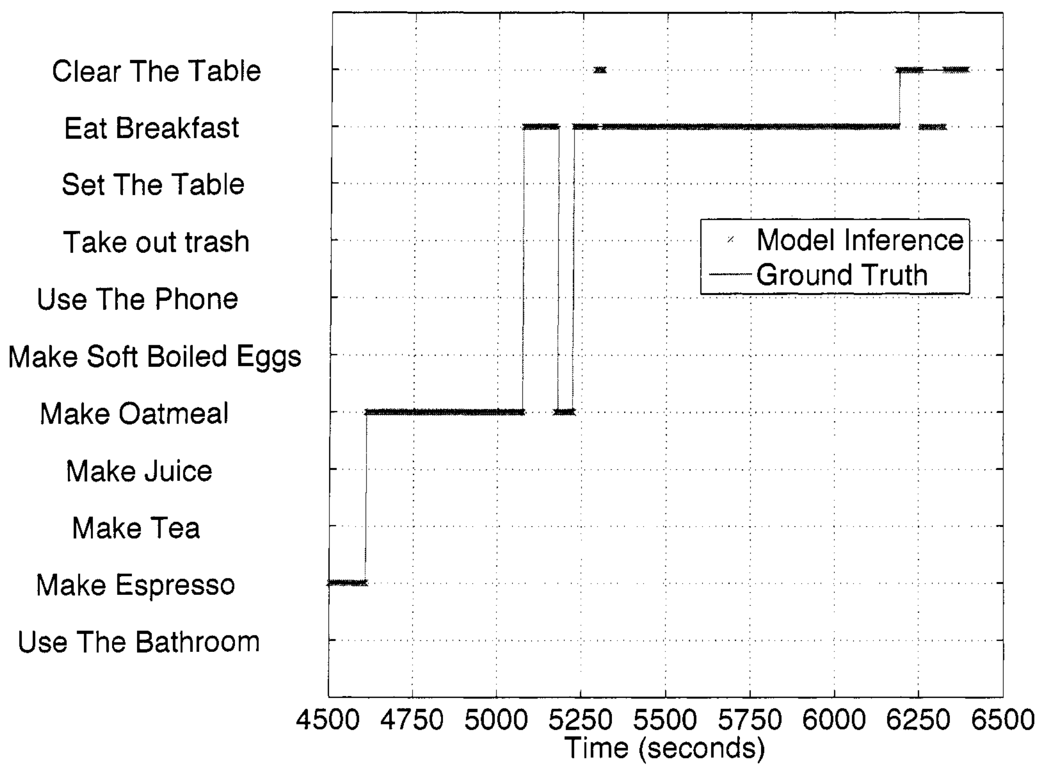


Figure 6.12: Model B Results (Connected HMMs)

Small segment of inference with various models. Ground truth is indicated by the thin line. Inference is indicated by the dots.

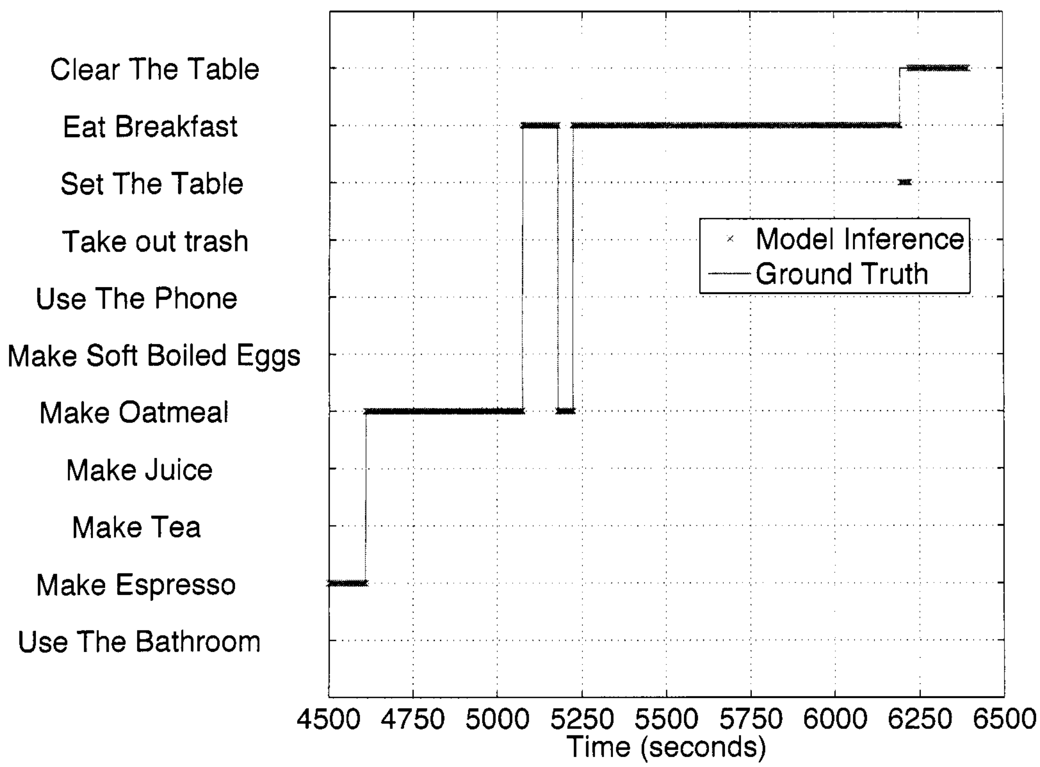


Figure 6.13: Model C Results (Object-Centered HMMs)

Small segment of inference with various models. Ground truth is indicated by the thin line. Inference is indicated by the dots.

eating breakfast to setting the table then to clearing the table. Given a sufficient amount of training data, this model will not make such an error. However there are a huge number of parameters required to specify this model. This model required 671^2 transition parameters to be learned and our data set simply did not have enough data to capture statistically significant information for every transition. Unlike model B, which captures information about transitions for activity A_i to A_j, model C represents information about transitions from $A_{i_{Obj_j}}$ to $A_{k_{Objm}}$.

Aggregate Model D

Figure 6.14 shows a short portion of the inference performed by model D. In this graph learning a distribution over the aggregate number of object instantiations eliminates the ambiguity in Figures 6.11-6.13 without requiring an inordinate number of parameters to be introduced into the model.

6.3.4 Models Which Improve Robustness

One of the concerns with the previous models is how well they will respond if someone used a functionally similar object but one which was nonetheless unexplained by the model. For example, in our model we trained on cooking oatmeal using a cooking spoon. Our inference should not fail if the user performed the same task using a table spoon. Likewise if the user makes tea in a cup rather than a mug, that should be a less likely, but still plausible alternative. To solve this problem we introduce the concept of *abstraction smoothing.*

Abstraction Smoothing Over Objects

In order to perform smoothing over objects we created a relational model inspired by [9]. Unlike the full power of that work, however, we used a single hierarchical object relation rather than a lattice. The hierarchy that we used was mined with supervision from an internet shopping site [51] (see figure 6.15). The name of each object was entered into the shopping search engine and the hierarchy that was returned for that object was inserted into the global object tree. In the case of objects with multiple hierarchies, one was manually

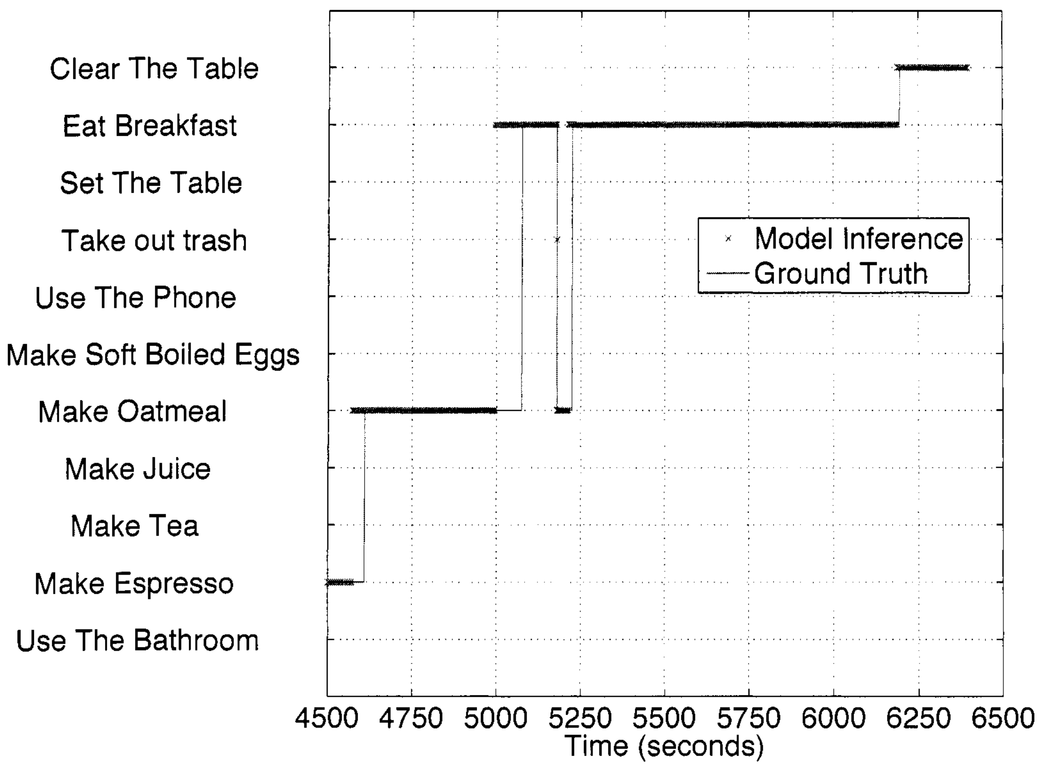


Figure 6.14: Model D Results (Aggregate DBN)

Small segment of inference with various models. Ground truth is indicated by the thin line. Inference is indicated by the dots.

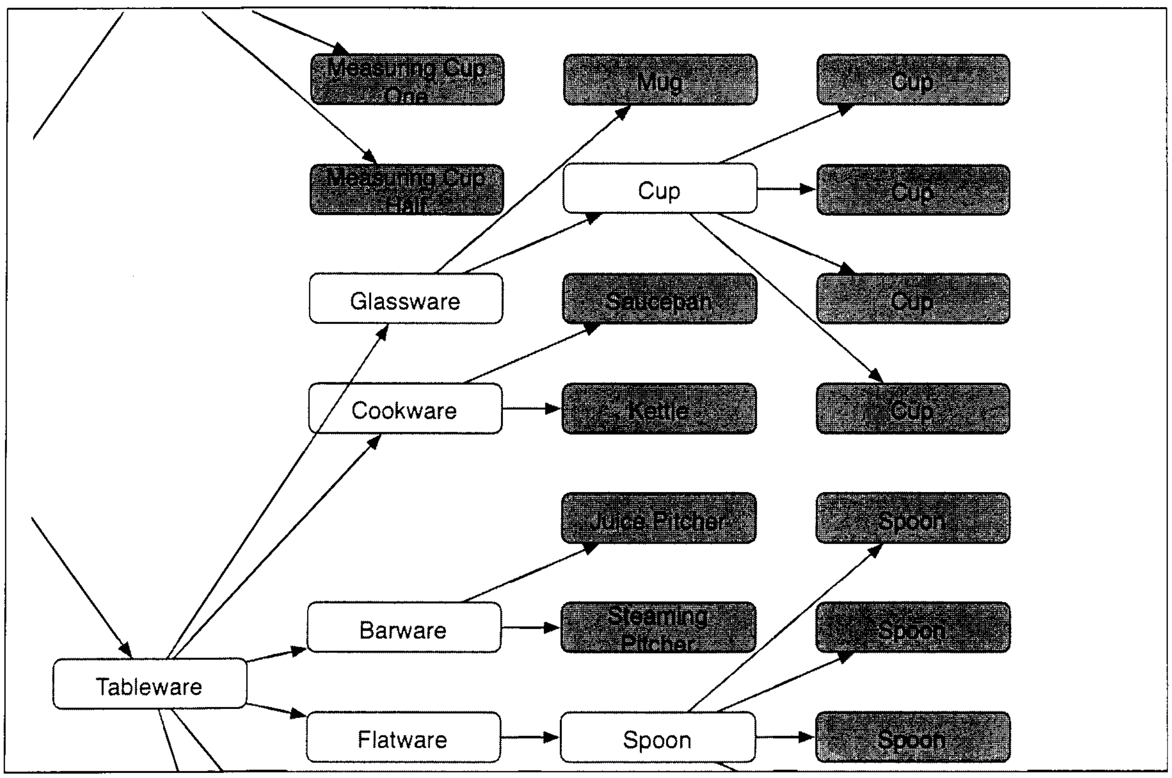


Figure 6.15: Object Abstraction Hierarchy

A portion of the object abstraction hierarchy mined from an Internet shopping site. Objects in our training data are shaded. Abstractions are not shaded.

selected.

The semantics that we applied to the resulting tree were that objects that were close to each other in the graph were functionally similar. To specify the notion of "close", we weighted all edges on the graph equally and created an all-pairs functional equivalence metric according the following formula:

$$P(O_i \rightarrow O_j) = \frac{\exp(-\frac{Dist(O_i,O_j)}{2})}{\sum_j \exp(-\frac{Dist(O_i,O_j)}{2})} \tag{6.1}$$

Where $Dist(O_i, O_j)$ is the shortest-path distance between O_i and O_j on the graph. This says that when object O_i is expected in the model, it will be substituted by object O_j with probability $P(O_i \rightarrow O_j)$. The likelihood of substituting one object for another falls off exponentially with distance in the hierarchy.

Figure 6.16 shows how the abstraction function is inserted graphically into model D. This changes the semantics of the model somewhat so that now the node O_i represents the object that "should" be touched in order to complete this activity and what is observed is O_j. The conditional probability table associated with the dependency is captured by equation 6.1.

Table 6.7: Accuracy metrics for when an object in the test trace is replaced by a functionally similar object.

	Model A Ind. HMMs	Model B Single HMM	Model D Aggregates w/ Abstract.
Mean Accuracy	52.5%	77.4%	81.2%
Net Change	-15.1%	-10.9%	-6.4%
Mean Edit Dist.	24.7	35.6	8.8
Net Change	+12.7	+26.6	+1.1

Robustness Experiments

To validate how well this technique worked when objects were substituted, we reran our experiments with abstraction smoothing added to model D. This resulted in an insignificant

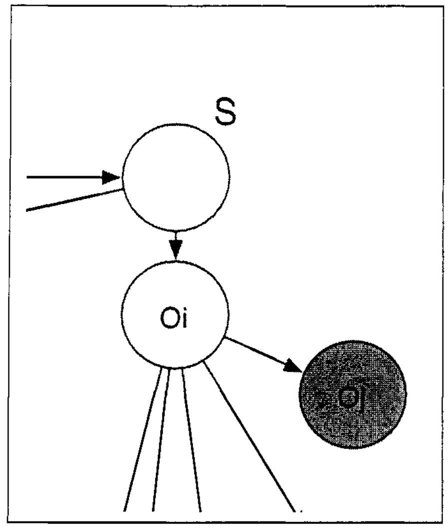


Figure 6.16: Model Changes to Support Abstraction Smoothing

The O_i and O_j nodes replace the previous "OBS" node in the DBN with $P(O_i \rightarrow O_j)$ embodied in the CPT.

decrease in accuracy of 0.1% and -.1 in edit distance.

Then we reran our experiments with the same data streams in which all instances of an object were replaced by another instance of an object. Table 6.7 shows our results for one scenario.

Whereas abstraction smoothing doesn't greatly harm normal activity recognition it greatly increases robustness to object substitution as Table 6.7 demonstrates. The metrics from this table were generated by replacing a mug in all of the testing sequences with a cup. In two of the baseline models accuracy is dramatically lowered, but the abstraction model suffers a relatively modest decrease in accuracy, especially according to the string edit metric.

6.4 Summary

In this chapter we described two experiments. The first examined how feasible activity recognition in a home environment was. We demonstrated strong results indicating that an RFID based activity system is plausible, deployable and informative.

The second experiment evaluated a progression of increasingly sophisticated models we applied to an activity recognition problem generated by an RFID-glove. We made a methodological point of proceeding in a bottom-up manner, starting with the simplest possible approach, and adding complexity only when it was necessary to improve accuracy and illuminate relevant features related to activity recognition. Such an approach makes clear how and why each feature of the system contributes to successful recognition. A second benefit is that the final system supports highly efficient inference and learning.

We saw that the popular approach of training separate HMMs (Model A) for each activity, and then performing classification by selecting the HMM that accords the current observation the highest likelihood (for example, see [33]), performs poorly once activities interact in time or when object use is shared among activities. This led us to consider a combined HMM model, where the observations still correspond to using an object of a given type. Although this greatly improves accuracy by learning transitions between activities, we saw that it could not distinguish between activities that used the same objects. Needing to distinguish among activities that use the same objects occurs often in modeling household activities — not only in kitchen activities, as in our experiments, but in many others ADLs: for example, doing laundry versus getting dressed, reading a book versus organizing a bookshelf, *etc.*

Modeling observations as touches of particular objects (as an RFID-glove does) does not, by itself, solve the activity recognition problem, because we would then lose any ability to generalize in those cases where object identity is irrelevant. The heart of the matter is that many activities are naturally non-Markovian, *i.e.*, they are best modeled in terms of the *history* of objects that were used during their execution. This led to our DBN model (D) that maintained such a history, by using memory nodes, and which allowed us to infer the aggregate feature of the number of objects used.

Differentiating between a model which allows for objects to be used at different temporal points in an activity versus a model which reasons about aggregations of objects (Model C vs Model D) allowed us to tease apart different ways in which two activities which used the same objects could be disambiguated. Our results demonstrated that both model C and model D had power to disambiguate these activities. Using aggregations of object touches however performed the best without requiring as much training data.

Finally, we saw that the use of smoothing over an abstraction hierarchy of object types greatly enhances the robustness of the system with no significant change in accuracy, and no significant additional computational burden.

In summary, these results argue that key features for modeling activities of everyday life on the basis of object interaction (and, very likely, on the basis of other kinds of direct sensor data) include a way to capture transition probabilities between activities; a way to compute aggregate features over the history of instantiated events in the activity; and a way to handle abstract classes of events. With these features alone, it is possible to perform quite fine-grained activity recognition. The activity recognition is successful even when the activities are interleaved and interrupted, when the models are automatically learned, when the activities have many objects in common and when the user deviates from the expected way of performing an activity.

Chapter 7

FUTURE DIRECTIONS

As sensors in the environment proliferate and as networking becomes available in smaller and smaller devices there seems to be a certain future for activity recognition in many forms. In this thesis we have looked at two different types of activity recognition, outdoors and indoors, using two different types of sensors GPS and RFID, respectively. In this chapter I would like to look at the future of these technologies.

7.1 Opportunity Knocks' Assumptions

OK is a very specialized type of activity recognition. Recognizable transportation plans were narrowly defined. The target user population and target application were also narrowly defined. Future work should consider relaxing each of these assumptions.

7.1.1 More Complicated Transportation Plans

OK assumed a particular form of transportation plan. It was a movement from a starting place to an ending destination, connected by segments of three different kinds of transportation. The only sub-goals consisted of transportation change locations (bus stops, parking lots etc.). From the perspective of an activity recognition system, the activity that was being recognized was quite rigidly defined. Fortunately it was successfully able to capture quite a bit of real transportation behavior.

A challenge for future work is to broaden the definition of a transportation plan to include more of the variation that occurs in the real world. An interesting extension would be to include the idea of sub-goals. These sub-goals could be either facilitating the current activity or not.

Facilitating sub-goals are places that people stop on the way to another location but that are premeditated and in some sense necessary. Places like this include, ATMs on the way to a movie, mailboxes on the way to work, and fast-food drive-throughs on the way home. To a certain degree OK can already learn this type of behavior. For example, if the sub-goal is a regular part of a person's routine then OK would just learn it in the course of its normal operation. The existing system simply learns that the route you take from home to the theater is either the one that goes by the ATM or the one that goes directly.

This is somehow not a satisfying solution, though, because it denies the computer any opportunity to imbue high-level context information into the stop at the sub-goal. If the computer knew that the ATM was a "short stop" or understood the manner in which it facilitated the current activity it could potentially aggregate statistics across all trips from home to the theater even when they included a sub-goal. It could also infer about new times when you might stop by the ATM even though you had never stopped at an ATM on that route before. Such reasoning would seem to be important for extending applications beyond way-finding recovery.

The challenge of reasoning with such intermediate locations is that it introduces complexity. Now instead of inferring about possible end goals, (assuming the start goal is known from GPS) a system would have to entertain the possibility of many different combinations of sub-goals and end-goals. Multiple sub-goals causes an exponential growth in the combinations. This complexity could be reduced in an interesting way with an understanding of why a person makes a stop at a sub-goal and then only including it as an option when that condition exists.

The other class of sub-goals don't immediately facilitate the current activity. This would be a situation where you stop at an ATM on the way to work, not because you need the money for work but because your spouse needs cash for some unrelated event at a future time. Accommodating these "random" stops of convenience is difficult because one would like to remove them from the learning process in order to improve statistical aggregation, but still allow them to become "routine" stops of convenience once there is enough evidence to support them as such.

Differentiating between the two types of stops might require the introduction of a hidden random variable that captures the notion that the sub-goal is facilitating the current activity.

Then aggregation can occur in different ways that reflect the state of that variable.

Typing locations

The previous discussion also suggests the idea of extending the notion of goals from simply physical locations to places that are typed and then reasoning with a relational model over types rather than necessarily instantiating the destination immediately. There is much active work along these lines [92].

But typing locations presents another difficult problem, that of identifying what the types of locations should be. Different locations are typed in different ways for different people at different times. One person's "office" might be another person's "bank" or still another person's "ATM". How each person uses each location depends a lot on how they type the location.

One of the advantages of typing locations is that it holds the promise of transferring model parameters among people by creating a layer of indirection from the actual physical locations of places and instead reasons about them with abstract labels. The assumption that one has to make in transferring model parameters among people is that different people will travel among places in the same way if they agree on their semantic labeling. A correlated and very interesting problem would be to see if the parameters transfer between people using abstract labels even if the same physical locations are labeled in contradictory ways. An extreme and simple example would be two apartment managers who might live at the location where the other works. "Home" and "Work" are opposite locations for these two people, but the patterns of travel between home and work would presumably transfer.

There are many different potential sources of type label information. One possibility is to type places based on how people use them. This would be an implicit typing. With only time and location as input, implicit typing would mean categorizing places in terms of the timing and frequency of visitation. Another source of information is GIS databases which impose their own ontology on the world to support their application. A third source would be a mass collaboration database that allows people to type places with free-form descriptions. Naturally hybrid approaches also seem plausible.

Other Types of Transportation Plans

Not all transportation consists of moving from one destination to another. Exercising, for example, frequently involves traveling a circuit outdoors and arriving at your starting point. Or alternatively spending a fair amount of time in roughly one outdoor location, such as playing basketball. Running errands looks like a successive sequence of sub-goals with no end goal.

Different types of careers also involve different sensor signatures. Someone who works as a delivery person, or as a bus driver, or long-haul truck driver completely reverses the assumed transportation model. For these people going to work means constant motion, not arriving at a destination. Plumbers, Cable T.V. repairmen, contractors go to "homes" when they go to "work."

Capturing this type of transportation requires generalizing the transportation model to capture alternative patterns.

7.1.2 Relaxing the Targeted User Population

OK assumed a user that made good decisions the majority of the time such that statistical analysis of their behavior could reveal meaningful patterns. These same users were then assumed to make occasional errors which could be corrected in light of past behavior. This user model is compelling because it corresponds with the behavior of people with MCI or early stage Alzheimer's Disease. It also rings somewhat true of all people. Everyone can relate to being distracted and missing a bus stop, highway exit, or getting lost. OK would need to change, however to support more and less cognitively able individuals.

If the user made more errors rather than just on an occasional basis, the OK learning mechanism would begin to break down. The learning mechanism would aggregate the common portions of the errors, learn them and start to encourage the user to repeat the same error patterns.

Someone with traumatic brain injury who is unable to navigate but can use a map or other descriptive tool, could utilize OK by turning off learning and relying on definitive routes previously loaded into OK. This would allow the system to make very certain decisions about when a user was making an error and immediately tell them or alert a care-giver. This is like putting someone inside a moving invisible fence that can sound an alarm when it is breached.

On the other end of the spectrum are people who are higher functioning. These are people for whom deviations from routing behavior are intentional and should not indicate an error. Working with such a user population is much easier because inference can always assume the user is correct. Such users will likely demand significant new value for the small burden that a location system like OK entails. This would likely take the form of new location-based applications.

7.1.3 Enhancing the application

There are other applications that an OK-like system could support.

- Location Awareness: Communicating current and likely future locations to remote parties
- Car Finding: Helping high functioning individuals find their car in large parking lots
- Sub-goal recommendations: Suggesting previously unknown ATMs that a user might want to stop at, because an ATM stop is predicted.
- Physical item reminders: Reminders for bringing things to a predicted destination
- Location based to-do lists: Reminders for accomplishing things when you are in a particular location

7.1.4 Other Opportunity Knocks Ideas

There are several other directions OK research could explore.

- User Interface Design: An analysis of user-interface decisions needs to be conducted for different subsets of the target population.

- Human Subjects Evaluation: Is a system like OK going to provide real value to people?

- Information Integration: Many other sources of information are available that could be effectively leveraged for improved accuracy or enhanced functionality. Calendars, weather reports, and traffic reports, for example, all contain information that influences where and how people travel. Weekends indicate that different transportation plans are in effect for many people.

- Different modes of transportation: OK could potentially differentiate ferry routes, airline routes, subway routes, trolley routes, monorail routes, etc.

- New sensors: Incorporating knowledge about specifically what bus or car one is riding in would enable much more accurate estimates of future trajectories. This could be accomplished through specifically instrumenting vehicles. Incorporating new location technologies could enable OK to continue working indoors. Adding a digital compass to the device would help to orient a lost user.

- Not all transportation occurs on roads. There are places where a graph representation of the world fails to accurate capture the motion that occurs there. Extremely large parking lots, shopping malls, parks, etc., all have structures that prevent easy graph-based inference. Extending the current model to allow for open space with free-form motion seems appropriate.

- Power management: When a person is sitting still there is little need for the full range of processing capabilities to be on-line. Current battery life for our system hovers around four hours. Aggressive management of power consumption could extend that further.

7.2 Multiple People

7.2.1 Outdoors

Reasoning with multiple people simultaneously in an outdoor environment has a lot of interesting potential. Clearly environments in which multiple people are cooperating outdoors already have a lot of effort devoted to coordination, both technically advanced and some not so advanced. Tasks in which people intentionally coordinate there behavior include military exercises, transfer of physical goods, team delivery of goods and services to multiple delivery points. Communication with a central coordination point is traditionally an effective and efficient way of managing resources.

Recognizing carpooling would be an interesting direction. This could be extended with services that automatically informed all parties concerned about transportation status. It would require the ability to recognize that two people are moving together even though they are only roughly collocated.

It also begs the question of trying to understand whether coordinated movement actually expressed some sort of coordinated plan. On the one hand a carpool seems to require some relationship between the individuals that could be supported with communications (automatically), but a bus could also look like a big carpool and it is neither interesting nor appropriate to report the status of every bus rider to every other bus rider.

From a sensor perspective it would be interesting to try and discover that two agents are working in tandem. In any densely populated environment this seems nearly impossible without some sort of additional information. If an object were tagged then one could coordinate the motion of the object with a person and identify that the object has been transferred to another person. For some objects this might be interesting, like a key, paperwork, or deliverable packages, but for other objects this is not so interesting. It doesn't seem interesting to track individual bills as they transit among people because they are somewhat interchangeable.

The entire realm of social contact mining would require identifying when multiple people were in close proximity for the purposes of socializing, exchanging information and/or conducting business. Again, this is very difficult to identify in general public spaces such as a coffee shop. It seems hard to identify from a single interaction whether or not two people are just eating lunch near each other or whether they are engaged in a social endeavor.

For some venues however it might be possible to identify such a collaboration. This would require additional information about the location such as the fact that a group of people have assembled in a meeting or conference room. Or that the space is a lecture hall. The reverse is also true. Knowing that a location is a coffee shop decreases the likelihood that two people are actually socializing.

In both situations time plays an important role. In a coffee shop, a chance encounter might not mean anything, but multiple repeated encounters of the same type might carry more information. Again, deciding whether or not this is important is more difficult in a coffee shop where the regulars return on schedule day after day, but really don't know each other. In contrast, however, repeated visits to a classroom does indicate something more important.

Another aspect of time that is important is whether or not the social contact begins and ends at the same time. Social events seem likely to end with all parties dispersing. Start times probably have more variability in their collocation initiation.

Key aspects of multiple person activity recognition depend a great deal on the type of activity that is being detected. A class of activities requires additional knowledge about the objects involved. Another class (social) requires additional knowledge about the locations involved and both seem like they could benefit from some observation of repeated encounters as well as duration and co-occurrence of starting and ending times.

7.3 BARISTA

7.3.1 Multiple People Indoors

Inferring how and if multiple people are coordinating their indoor activities in a system like BARISTA seems very challenging based on the assumptions of our system. The primary assumption was that the objects that people touch are the primary source of knowledge about what activities are in progress.

Given just two people and two sensor streams possible explanations for the observations are that individual independent activities are in progress, that a single coordinated activity is in progress, or some combination of multiple activities with some combination of coordinated effort. As more people are added to the system, identifying who are coordinating their activities with whom grows much worse. Decomposing these sensors streams into individual activities requires strong expectations about observations an activity is likely to generate. These expectations are likely to come from some sort of a "recipe" that describes how an activity is accomplished.

7.3.2 Activities are Hard to Define

Once an activity recognition system begins to scale up to hundreds of activities, it becomes challenging to split them into meaningful exclusive categories. Some activities are independent of who does them, for example, changing the state of the world. Opening or closing a specific door, cooking spaghetti or cleaning the floor are fairly straightforward to describe. But depending on how an activity is described it could much more difficult.

Take for example the activity of washing a load of laundry. One might describe this activity in many ways that are appropriate for different activities. Perhaps in an ADL recognition context it is important to recognize "washing one's clothes". This places restrictions on which clothes are being washed. Two people can coordinate to "wash clothes" or two people can coordinate to "wash Mary's clothes". These are both activities that might be accomplished simultaneously which aren't at all exclusive. Exclusivity of activities was an important simplification made in BARISTA. In fact many activities may have taken place in the course of a single load of laundry. Clothes were washed, Mary's clothes were washed, clothes were dried, laundry was done, cleaning (in a sense) was done.

One solution to this problem of identifying activities would be to put all activities in a strict ontology. This would be some sort of a structure that places all activities into exclusive bins such that at any given time one is only engaged in one activity. This is how BARISTA was structured.

As soon as the possibility of accomplishing multiple activities becomes possible, the possibility of doing all activities simultaneously must be considered. This is true both in the sense that multiple activities are occurring simultaneously as well as in the sense that multiple activities are both being accomplished by the same actions. Based on our attempts to model this problem we found that solving this problem with exact inference is beyond the capabilities of today's inference engines.

One possible solution would be to approach this problem from a query-based perspective instead. Instead of trying to detail every possible type of activity that might be of interest. An inference engine might be better served by just indexing an incoming data stream and then attempting to answer specific questions after the fact. Such as "was the laundry done today?", or "Does Mary have clean clothes?". In such cases commonly asked queries could be cached or computed in near real-time so that up-to-date status could be relayed when it was needed.

In some sense the ontology approach is implicitly accomplishing this. The ontology is a way of defining the queries that are able to be asked in the future and then processing the data in advance to collect sufficient statistics to answer a set of possible queries. In effect an ontology is saying that the inference engine is able to identify which of the exclusively defined activities was in progress at any given time. The struggle is in trying to balance the flexibility for the query asker - that is allowing the largest number of possibly future queries to be answered, with tractability.

7.3.3 Other Directions for BARISTA

- Reasoning with specific instances of objects is an interesting direction to take activity inference. We began to explore this in our aggregation model, but there is more interesting work to be done along these lines. Especially when an activity is only an activity when you use the same exact object at particular times in an activity.

- Different types of RFID tags promise to offer new capabilities. If an RFID tag could sense small amounts of information about the world around it, that information could be leveraged by the inference engine. For example, a tea pot that indicated there was hot water inside would be of great help in deciphering the process of making tea in the presence of multiple people and an unexpected sequence of observations.

- Leveraging location would reduce the number of activities that could be considered at any given time. It shouldn't be necessary to consider the possibility that you are mowing the lawn when you touch a lawn mower in the store, or that you are brushing your teeth when you are detailing your car, in the driveway, with a toothbrush.

Chapter 8

CONCLUSIONS

In this thesis we have looked at the future of aging in the industrialized world and seen that the absolute and relative numbers of the future elderly are going to strain existing social structures. The amount of fiscal and social capital that is going to be required to take care of the predictable challenges of aging is daunting. Since one of the leading causes of these costs is functional impairment leading to admission in a long-term care facility, a viable solution is to maintain the independence of these individuals.

Advances in both sensor technology, and probabilistic reasoning offer some hope of how we can promote independence in the face of cognitive decline. By developing information systems that augment cognition in the same way that physical devices compensate for physical disabilities we may be able to maintain higher quality of life for elderly people and their care-givers. At the same time we can reduce costs to these families and their various governments.

We have looked at two possible avenues for assistance. The first is using an outdoor activity recognition system based on GPS to help people who make occasional cognitive errors recover safely. The second is an indoor activity recognition system based on a wearable computing platform and RFID tags that is designed to monitor which activities are occurring in a home.

In the outdoor case, we demonstrated that such a system could successfully be built now, that the reasoning that such a system can do is both accurate and valuable and that user interface innovations make such a system usable without extensive user programming.

In the indoor case, we saw that a single technology can subsume many previous activity recognition techniques in a way that is robust, easily deployable and accurate at a moderately fine level of detail. We identified particular aspects of object interaction that are valuable for activities disambiguation and suggested some ways of dealing with functional object substitution.

In summary we have shown that it is possible to use current and next generation sensors to recognize activities of individuals with sufficient detail that cognitive errors can be recognized, and corrected.

BIBLIOGRAPHY

[1] Scitovsky AA. "The high cost of dying" revisited. *Milbank Quarterly*, 72(4):561–591, 1994. PMID: 7997219.

[2] Gregory D. Abowd, Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. Cyberguide: a Mobile Context-Aware Tour Guide. *Wirel. Netw.*, 3(5):421–433, 1997.

[3] Norman Adams, Rich Gold, Bill N. Schilit, Michael Tso, and Roy Want. An Infrared Network for Mobile Computers. In *Proc. USENIX Symposium on Mobile & Location-Independent Computing*, pages 41–52. USENIX Association, August 1993.

[4] Administration on Aging. http://www.aoa.gov/prof/Statistics/ online_stat_data/AgePop2050.asp.

[5] Administration on Aging. A Profile of Older Americans. Technical report, U.S. Department of Health and Human Services, 2003. http://www.aoa.gov/prof/Statistics/profile/2003/2003profile.pdf.

[6] Alzheimer's Association and the National Alliance for Caregiving. Families Care: Alzheimer's Caregiving in the United States. Technical report, Alzheimer's Association and the National Alliance for Caregiving, 2004. http://www.alz.org/Resources/FactSheets/Caregiverreport.pdf.

[7] Alzheimer's Association of Cleveland website. http://www.alzclv.org/Statistics.htm.

[8] Alzheimer's Disease Education and Referral Center. http://www.alzheimers.org/pubs/adfact.html.

[9] Corin Anderson, Pedro Domingos, and Dan Weld. Relational Markov Models and their Application to Adaptive Web Navigation. In *Proc. of the Eighth Intl. Conf. on Knowledge Discovery and Data Mining*, pages 143–152. ACM Press, 2002.

[10] D. Ashbrook and T. Starner. Using GPS to Learn Significant Locations and Predict Movement Across Multiple Users. *Personal and Ubiquitous Computing*, 7(5), 2003.

[11] Daniel Ashbrook and Thad Starner. Learning Significant Locations and Predicting User Movement with GPS. In *Proc. of the Sixth IEEE Intl. Symposium on Wearable Computers*, Seattle, WA, October 2002.

[12] Paramvir Bahl and Venkata N. Padmanabhan. RADAR: An In-Building RF-Based User Location and Tracking System. In *INFOCOM (2)*, pages 775–784, 2000.

[13] J. Banaszak-Holl, A.M. Fendrick, N.L. Foster, A.R. Herzog, M.U. Kabeto, D.M. Kent, W.L. Straus, and K.M. Langa. Predicting Nursing Home Admission: Estimates from a 7-year Follow-up of a Nationally Representative Sample of Older Americans. *Alzheimer Disease Association Disorders*, 18(2):83–9, April-June 2004. PMID:15249852, NLM:8704771.

[14] Ling Bao and Stephen S. Intille. Activity Recognition from User-Annotated Acceleration Data. In Alois Ferscha and Friedemann Mattern, editors, *Pervasive Computing, Second International Conference, PERVASIVE 2004, Vienna, Austria, April 21-23, 2004, Proceedings*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer, 2004.

[15] Y. Bar-Shalom and X.-R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques.* Yaakov Bar-Shalom, 1995.

[16] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation.* John Wiley, 2001.

[17] T. Barger, M. Alwan, S. Kell, B. Turner, S. Wood, and A. Naidu. Objective Remote Assessment of Activities of Daily Living: Analysis of Meal Preparation Patterns. Technical report, Medical Automation Research Center, University of Virginia Health System, 2002.

[18] Bath Institue for Medical Engineering website. http://www.bath.ac.uk/bime/projects.htm.

[19] R. A. Beeson. *Loneliness and Depression in Caregivers*, pages 449–54. Humana Press, 2004.

[20] J.A. Bilmes. A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical Report ICSI-TR-97-021, University of Berkeley, 1998.

[21] J. P. Blass. *Metabolism and Alzheimer's Disease*, pages 39–47. Humana Press, 2004.

[22] P. Bonnifait, P. Bouron, P. Crubillé, and D. Meizel. Data Fusion of Four ABS Sensors and GPS for an Enhanced Localization of Car-like Vehicles. In *Proc. of the IEEE International Conf. on Robotics & Automation*, 2001.

[23] M. Bruscoli and S. Lovestone. Is MCI Really Just Early Dementia? A Systematic Review of Conversion Studies. *International Psychogeriatrics*, 16(2):129–40, June 2004. PMID:15318760,NLM:9007918.

[24] H. Bui. A General Model for Online Probabilistic Plan Recognition. In *Proc. of the International Joint Conf. on Artifical Intelligence*, 2003.

[25] H. Bui, S. Venkatesh, and G. West. Policy Recognition in the Abstract Hidden Markov Model. In *Journal of Artificial Intelligence Research*, 2002.

[26] US Census Bureau. Census 2000 TIGER/Line Data. http://www.esri.com/data/download/census2000-tigerline/, 2000.

[27] Jenna Burrell, Geri K. Gay, Kiyo Kubo, and Nick Farina. Context-Aware Computing: A Test Case. In *Proc. of the 4th international conference on Ubiquitous Computing*, pages 1–15. Springer-Verlag, 2002.

[28] W. Mann C. Giraldo, S. Helal. mPCA - A Mobile Patient Care-Giving Assistant for Alzheimer Patients. In *Proceedings of UbiCog '02: First International Workshop on Ubiquitous Computing for Cognitive Aids*, Gothenberg, Sweden, 2002.

[29] Carnegie Mellon University: Nursebot project website. http://www.cs.cmu.edu/~nursebot/.

[30] Census 1997. Aging in the United States: Past, Present and Future. Technical report, National Institute on Aging, U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census, 1997. http://www.census.gov/ipc/prod/97agewc.pdf.

[31] Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday, and Christos Efstratiou. Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences. In *CHI*, pages 17–24, 2000.

[32] T. Choudhury and A. Pentland. Sensing and Modeling Human Networks Using the Sociometer. In *Proc. of the Seventh IEEE Intl. Symposium on Wearable Computers*, pages 216–222. IEEE, October 2003.

[33] Tanzeem Choudhury, Jonathan Lester, and Gaetano Borriello. A Hybrid Discriminative/Generative Approach for Modeling Human Activities. In *Proc. of the Nineteenth Intl. Joint Conf. on A.I.*, Edinburgh, UK, 2005. Morgan-Kaufmann Publishers.

[34] M.; Cielniak, G.; Bennewitz and W. Burgard. Where is ...? learning and utilizing motion patterns of persons with mobile robots. In *Proc. of the Seventeenth Intl. Joint Conf. on A.I.*, 2003.

[35] D. Colbry, B. Peintner, and M. Pollack. Execution Monitoring with Quantitative Temporal Bayesian Networks. In *6th Intl. Conf. on AI Planning and Scheduling*, April 2002.

[36] S. Consolvo, P. Roessler, B.E. Shelton, A. LaMarcha, B. Schilit, and S. Bly. Technology for Care Networks of Elders. *IEEE Pervasive Computing Mobile and Ubiquitous Systems: Successful Aging*, 3(2):22–29, Apr-Jun 2004. http://www.intel-research.net/Publications/Seattle/022320041335_230.pdf.

[37] King County. GIS (Graphical Information System). http://www.metrokc.gov/gis/mission.htm, 2003.

[38] Andy Crabtree, Tom Rodden, Terry Hemmings, and Steve Benford. Finding a Place for UbiComp in the Home. In Anind Dey, Albrecht Schmidt, and Joseph F. McCarthy, editors, *Proc. of UBICOMP 2003*, volume LNCS 2864, pages 208–226. Springer-Verlag, 2003.

[39] Y. Cui and S.S. Ge. Autonomous Vehicle Positioning with GPS in Urban Canyon Environments. In *Proc. of the IEEE International Conf. on Robotics & Automation*, 2001.

[40] T. Dean and K. Kanazawa. Probabilistic Temporal Reasoning. In *Proc. of the National Conf. on Artificial Intelligence*, 1988.

[41] P. Del Moral and L. Miclo. Branching and Interacting Particle Systems Approximations of Feynam-Kac Formulae with Applications to Non-linear Filtering. In *Seminaire de Probabilites XXXIV*, number 1729 in Lecture Notes in Mathematics. Springer-Verlag, 2000.

[42] Digital Angel website. http://www.digitalangelcorp.com/.

[43] A. Doucet, J de Freitas, K. Murphy, and S. Russell. Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks. In *Proc. of the Conf. on Uncertainty in Artificial Intelligence*, 2000.

[44] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo in Practice.* Springer-Verlag, New York, 2001.

[45] Elite Care website. http://www.elite-care.com/index.html.

[46] R. L. Ernst and J. W. Hay. The US economic and social costs of Alzheimer's disease revisited. *American Journal of Public Health*, 84(8):1261–4, August 1994. PMID:8059882, NLMID:1254074.

[47] F.C.C. U.S. Federal Communication Commission E-911 Website, 2004. http://www.fcc.gov/911/enhanced.

[48] P. Fitzpatrick and C.C. Kemp. Shoes as a Platform for Vision. In *Proc. of the Seventh IEEE Intl. Symposium on Wearable Computers*, pages 231–234. IEEE, October 2003.

[49] D.J. Foley, A.M. Ostfeld, L.G. Branch, R.B. Wallace, J.J. McGloin, and J.C. Cornoni-Huntley. The Risk of Nursing Home Admission in Three Communities. *Journal of Aging Health*, 4(2):155–73, May 1992. PMID: 10117872.

[50] V.A. Freedman, L.F. Berkman, S.R. Rapp, and A.M. Ostfeld. Family Networks: Predictors of Nursing Home Entry. *American Journal of Public Health*, 84(5):843–5, May 1994. PMID: 8179059.

[51] Froogle Shopping Site. http://froogle.google.com/.

[52] L. Getoor, N. Friedman, D. Koller, and A. Pfeffer. Learning Probabilistic Relational Models. In S. Dzeroski and N. Lavrac, editors, *Relational Data Mining*. Springer-Verlag, 2001.

[53] A. Glascock and D. Kutzik. Behavioral Telemedicine: A New Approach to the Continuous Nonintrusive Monitoring of Activities of Daily Living. *Telemedicine Journal*, 6(1):33–44, 2000.

[54] Google. Google Directions, 2005. http://maps.google.com/maps?q=seattle+to+irvine+ca.

[55] W. G. Griswold, P. Shanahan, S. W. Brown, R. Boyer, M. Ratto, R. B. Shapiro, and T. M. Truong. ActiveCampus - Experiments in Community-Oriented Ubiquitous Computing. Technical Report CS2003-0750, UC San Diego, June 2003.

[56] William G. Griswold, Robert Boyer, Steven W. Brown, and Tan Minh Truong. A Component Architecture for an Extensible, Highly Integrated Context-Aware Computing Infrastructure. In *Proc. of the 25th International Conf. on Software Engineering*, pages 363–372. IEEE Computer Society, 2003.

[57] V. Guralnik and K. Haigh. Learning models of human behaviour with sequential patterns. In *AAAI Workshop on Automation as Caregiver*, pages 24–30, 2000.

[58] J. Hardy. *Genetics of Alzheimer's Disease and Related Disorders*, pages 3–20. Humana Press, 2004.

[59] Ramaswamy Hariharan and Kentaro Toyma. Project Lachesis: Parsing and Modeling Location Histories. In *GIScience*, 2004.

[60] A. Helal, W. Mann, C. Giraldo, Y. Kaddoura, C. Lee, and H. Zabbadani. Smart Phone Based Cognitive Assistant. In *Proceedings of UbiHealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, Seattle, Washington, USA, Oct 2003.

[61] J. Hightower and G. Borriello. A Survey and Taxonomy of Location Systems for Ubiquitous Computing, 2001.

[62] J. Hightower and G. Borriello. Location Systems for Ubiquitous Computing. *Computer*, 34(8), 2001. IEEE Computer Society Press.

[63] Jeffrey Hightower. From Position to Place. In *Proc. of the 2003 Workshop on Location-Aware Computing*, pages 10–12, October 2003. part of the 2003 Ubiquitous Computing Conf.

[64] Jeffrey Hightower and Gaetano Borriello. Accurate, Flexible and Practical Location Estimation for Ubiquitous Computing: A Case Study of Particle Filters. In *Proc. of UBICOMP 2004*. Springer-Verlag, 2004.

[65] Jeffrey Hightower, Barry Brumitt, and Gaetano Borriello. The Location Stack: A Layered Model for Location in Ubiquitous Computing. In *Proc. of the Fourth IEEE Workshop on Mobile Computing Systems and Applications*. IEEE Computer Society, 2002.

[66] Jason L. Hill and David E. Culler. Mica: A Wireless Platform for Deeply Embedded Networks. *IEEE Micro*, 22(6):12–24, 2002.

[67] T. Hopper, K. Bayles, and E. Kim. Retained Neuropsychological Abilities of Individuals with Alzheimer's Disease. *Seminars in Speech and Language*, 22(4):261–273, 2001.

[68] Intel Research Seattle. Intel Research Seattle Guide Project Website, 2003. http://seattleweb.intel-research.net/projects/guide/.

[69] International Database. http://www.census.gov/ipc/www/idbnew.html.

[70] S.S. Intille. Ubiquitous Computing Technology for Just-in-Time Motivation of Behavior Change. In *Proceedings of UbiHealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, Seattle, Washington, USA, 2003.

[71] Stephen S. Intille, Kent Larson, J. Beaudin, E. Munguia Tapia, P. Kaushik, J. Nawyn, and T.J. McLeish. The PlaceLab: a Live-In Laboratory for Pervasive Computing Research. In *Pervasive Computing, Third International Conference, PERVASIVE*

2005, Video Program, Lecture Notes in Computer Science, page video. Springer, May 2005.

[72] Tony Jebara and Alex Pentland. Action Reaction Learning: Automatic Visual Analysis and Synthesis of Interactive Behaviour. In *ICVS '99: Proc. of the First Intl. Conf. on Computer Vision Systems*, pages 273–292, London, UK, 1999. Springer-Verlag.

[73] A.M. Jette, S. Tennstedt, and S. Crawford. How Does Formal and Informal Community Care Affect Nursing Home Use? *Journal of Gerontology: Social Science*, 50(1):S4–S12, Jan 1995. PMID:7757829, Medline:95277547.

[74] Holly Brügge Jimison, Misha Pavel, James McKanna, and Jesse Pavel. Unobtrusive Monitoring of Computer Interactions to Detect Cognitive Status in Elders. *IEEE Transactions on Information Technology in Biomedicine*, 8(3):248–252, 2004.

[75] K. K. Van Laerhoven and H. Gellersen. Spine Versus Porcupine: A Study in Distributed Wearable Activity Recognition. In *Proc. of the Eighth IEEE Intl. Symposium on Wearable Computers*, volume 1, pages 142–149. IEEE, October 2004.

[76] Jong Hee Kang, William Welbourne, Benjamin Stewart, and Gaetano Borriello. Extracting Places from Traces of Locations. In *The Second ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots*, 2004.

[77] S. Katz. Assessing Self-Maintenance: Activities of Daily Living, Mobility and Instrumental Activities of Daily Living. *Journal of the American Geriatrics Society*, 31(12):721–726, 1983.

[78] Henry Kautz, Larry Arnstein, Gaetano Borriello, Oren Etzioni, and Dieter Fox. The Assisted Cognition Project. In *Proceedings of UbiCog '02: First International Workshop on Ubiquitous Computing for Cognitive Aids*, Gothenberg, Sweden, 2002.

[79] Henry A. Kautz and James F. Allen. Generalized Plan Recognition. In *AAAI*, pages 32–37, 1986.

[80] P Kemper and CM Murtaugh. Lifetime Use of Nursing Home Care. *N Engl J Med*, 324(9):595–600, 1991.

[81] N. Kern, S. Antifakos, B. Schiele, and A. Schwaninger. A Model for Human Interruptability: Experimental Evaluation and Automatic Estimation from Wearable Sensors. In *Proc. of the Eighth IEEE Intl. Symposium on Wearable Computers*, volume 1, pages 158–165. IEEE, October 2004.

[82] King County Department of Transportation. Trip planner. http://tripplanner.metrokc. gov/, 2004.

[83] King County Department of Transportation. Real Time Bus Positions. http://tracker-map.metrokc.gov/tracker-map-launch.jsp, 2005.

[84] John Krumm, Gerry Cermak, and Eric Horvitz. RightSPOT: A Novel Sense of Location for Smart Personal Object. In Anind Dey, Albrecht Schmidt, and Joseph F. McCarthy, editors, *Proc. of UBICOMP 2003*, volume LNCS 2864. Springer-Verlag, 2003.

[85] John Krumm, Lyndsay Williams, and Greg Smith. SmartMoveX on a Graph - An Inexpensive Active Badge Tracker. In *Proc. of the 4th international conference on Ubiquitous Computing*, pages 299–307. Springer-Verlag, 2002.

[86] K.M. Langa, M.E. Chernew, M.U. Kabeto, A.R. Herzog, M.B. Ofstedal, R.J. Willis, R.B. Wallace, L.M. Mucha, W.L. Straus, and A.M. Fendrick. National Estimates of the Quantity and Cost of Informal Caregiving for the Elderly with Dementia. *Journal of General Internal Medicine*, 16(11):770–8, November 2001. PMID:11722692, NLM:8605834.

[87] M. Powell Lawton. Environment and Other Determinants of Well-Being in Older People. *The Gerontologist*, 23(4):349–357, August 1983. NLM:0375327;PMID:6352420.

[88] M. Powell Lawton. Aging and Performance of Home Tasks. *Human Factors*, 32(5):527–536, October 1990. NLM:0374660;PMID:2074107.

[89] R. Levine and G. Casella. Implementations of the Monte Carlo EM Algorithm. *Journal of Computational and Graphical Statistics*, 10, 2001.

[90] Richard Levinson. The Planning and Execution Assistant and Trainer (PEAT). In *The Journal of Head Trauma Rehabilitation*, April 1997.

[91] L. Liao, D. Fox, and H. Kautz. Location-Based Activity Recognition using Relational Markov Networks. In *Proc. of the Nineteenth Intl. Joint Conf. on A.I.*, 2005.

[92] Lin Liao, D. Fox, and H. Kautz. Learning and Inferring Transportation Routines. In *Proc. of the 19th National Conf. on Artificial Intelligence (AAAI)*, 2004.

[93] Lin Liao, Dieter Fox, Jeffrey Hightower, Henry Kautz, and Dirk Schulz. Voronoi Tracking: Location Estimation using Sparse and Noisy Sensor Data. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 723–728. IEEE/RSJ, October 2003.

[94] Living Independently website. http://www.livingindependently.com/.

[95] N. L. Mace and P. V. Rabins. *Managing Symptoms in the 36-Hour Day.* The Johns Hopkins University Press, Baltimore, 1981.

[96] Mapquest. Mapquest Directions, 2005. http://www.mapquest.com/directions/.

[97] Matthai Philipose. From personal communication.

[98] C. E. McCarthy and M. E. Pollack. A Plan-Based Personalized Cognitive Orthotic. In *6th International Conference on AI Planning and Scheduling*, April 2002.

[99] A. Mihailidis, J.C. Barbenel, and G. Fernie. The Efficacy of an Intelligent Cognitive Orthosis to Facilitate Handwashing by Persons with Moderate-to-Severe Dementia. *Nueropsychological Rehabilitation*, 00(00):00–01, 0 0000.

[100] A. Mihailidis, B. Carmichael, and J. Boger. The Use of Computer Vision in an Intelligent Environment to Support Aging-in-Place, Safety, and Independence in the Home. *IEEE Transactions on Information Technology in BioMedicine*, 8(3):238–247, 2004.

[101] T. Mitchell. *Machine Learning.* McGraw-Hill, 1997.

[102] Michael Montemerlo, Sebastian Thrun, Daphne Koller, and Ben Wegbreit. FastSLAM: a Factored Solution to the Simultaneous Localization and Mapping Problem. In *Eighteenth national conference on Artificial intelligence*, pages 593–598, Menlo Park, CA, USA, 2002. American Association for Artificial Intelligence.

[103] D. Moore, I. Essa, and M. Hayes. Exploiting Human Actions and Object Context for Recognition Tasks. In *Proceedings of IEEE International Conference on Computer Vision 1999 (ICCV.99)*, March 1999.

[104] A.C. Morris, R.R. Donamukkala, A. Kapuria, A.M. Steinfeld, J. Matthews, J. Dunbar-Jacobs, and S. Thrun. A Robotic Walker that Provides Guidance. In *Proceedings of the 2003 IEEE Conference on Robotics and Automation (ICRA '03)*, May 2003.

[105] J.C. Morris, S. Edland, C. Clark, D. Galasko, E. Koss, R. Mohs, G. van Belle, G. Fillenbaum, and A. Heyman. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). *Neurology*, 43(12):2457–2465, December 1993. NLM:0401060;PMID:8255439.

[106] K. Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning.* PhD thesis, UC Berkeley, Computer Science Division, 2002.

[107] K. Murphy and S. Russell. *Rao-blackwellized particle filtering for dynamic Bayesian networks*, pages 499–516. Springer Verlag, 2001.

[108] Elizabeth D. Mynatt, Jim Rowan, Sarah Craighill, and Annie Jacobs. Digital Family Portraits: Supporting Peace of Mind for Extended Family Members. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 333–340, New York, NY, USA, 2001. ACM Press.

[109] Progress Report on Alzheimer's Disease 2000. online at http://www.alzheimers.org, 2000.

[110] NextBus Inc. Real Time Bus Predictions. http://www.nextbus.com/predictor/stopSelector.jsp, 2005.

[111] Takayuki Osogami and Mor Harchol-Balter. A Closed-Form Solution for Mapping General Distributions to Minimal PH Distributions. In Peter Kemper and William H. Sanders, editors, *Computer Performance Evaluation / TOOLS*, volume 2794 of *Lecture Notes in Computer Science*, pages 200–217. Springer, 2003.

[112] Don Patterson, Oren Etzioni, Dieter Fox, and Henry Kautz. The Activity Compass. In *Proceedings of UbiCog '02: First International Workshop on Ubiquitous Computing for Cognitive Aids*, Gothenberg, Sweden, 2002.

[113] Donald J. Patterson, Lin Liao, Dieter Fox, and Henry Kautz. Inferring High-Level Behavior from Low-Level Sensors. In Anind Dey, Albrecht Schmidt, and Joseph F. McCarthy, editors, *Proceedings of UBICOMP 2003: The Fifth International Conference on Ubiquitous Computing*, volume LNCS 2864, pages 73–89. Springer-Verlag, 2003.

[114] PEAT: The Planning and Execution Assistant and Training System. http://www.brainaid.com.

[115] A. Peeters, J.J. Barendregt, F. Willekens, J.P. Mackenbach, A. Al-Mamun, and L. Bonneux. Obesity in Adulthood and its Consequences for Life Expectancy: a Life-Table Analysis. *Annals Internal Medicine*, 138(1):24–32, January 2003. NLM:0372351,PMID:12513041ed,pii:200301070-00008,medline:22401331.

[116] Mike Perkowitz, Matthai Philipose, Donald J. Patterson, and Kenneth P Fishkin. Mining Models of Human Activities from the Web. In *Proceedings of WWW2004: The Thirteenth International World Wide Web Conference*, 2004.

[117] R. C. Petersen. *Conceptual Overview*, pages 1–14. Oxford University Press, 2004.

[118] E.A. Phelan, B. Williams, B. Penninx, J.P. LoGerfo, and S.G. Leveille. Activities of Daily Living Function and Disability in Older Adults in a Randomized Trial of the Health Enhancement Program. *The Journals of Gerontology: Series A : Biological sciences and medical sciences*, 59A(8):838–843, August 2004.

[119] Matthai Philipose, Kenneth P. Fishkin, Mike Perkowitz, Donald J Patterson, Dirk Hahnel, Dieter Fox, and Henry Kautz. Inferring Activities from Interactions with Objects. *IEEE Pervasive Computing: Mobile and Ubiquitous Systems*, 3(4):50–57, 2004.

[120] L. R. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. In A. Waibel and K.-F. Lee, editors, *Readings in Speech Recognition*, pages 267–296. Kaufmann, San Mateo, CA, 1990.

[121] B. Reisberg, S. Ferris, M. deLeon, A. Kluger, E. Franssen, J. Borenstein, and R. Alba. The Stage Specific Temporal Course of Alzheimer's Disease: Functional and Behavioral Concomitants Based Upon Cross-Sectional and Longitudinal Observation. *Progress in Clinical and Biological Research*, 317:23–41, 1989. NLM:7605701;PMID:2690101.

[122] B. Reisberg, S. Kenowsky, E. Franssen, S. Auer, and L. Souren. Towards a Science of Alzheimer's Disease Management: a Model Based Upon Current Knowledge of Retrogenesis. *International Psychogeriatrics*, 11(1):7–23, March 1999. MEDL:99205594;PMID:10189596.

[123] B. Z. Richter. *Potential and Established Risk Factors for Alzheimer's Disease*, pages 65–71. Humana Press, 2004.

[124] B. Z. Richter and R.W. Richter. *Treatment of Alzheimer's Disease:Basic Considerations*, pages 177–8. Humana Press, 2004.

[125] Jane M. Sanders. Sensing the Subtleties of Everyday Life. *Research Horizons*, Winter 2000.

[126] L. S. Scheider. *Rating Scales and Outcome Variables Used in Clinical Trials*, pages 337–49. Humana Press, 2004.

[127] Bill Schilit, Anthony LaMarca, Gaetano Borriello, William Griswold, David McDonald, Edward Lazowska, Anand Balachandran, Jason Hong, and Vaughn Iverson. Challenge: Ubiquitous Location-Aware Computing and the Place Lab Initiative. In *Proc. of WMASH 2003: The First ACM International Workshop on Wireless Mobile Applications and Services on WLAN*, September 2003.

[128] Bill N. Schilit, Norman Adams, Rich Gold, Michael Tso, and Roy Want. The PARCTAB Mobile Computing System. In *Proc. Fourth Workshop on Workstation Operating Systems (WWOS-IV)*, pages 34–39. IEEE, October 1993.

[129] A. Schmidt, H-W. Gellersen, and C. Merz. Enabling Implicit Human Computer Interaction - A Wearable RFID-Tag Reader. In *Proceedings of ISWC 2000*, pages 193–194, 2000.

[130] C. Schultz, K. Del Tredici, and H. Braak. *Neuropathology of Alzheimer's Disease*, pages 21–32. Humana Press, 2004.

[131] Dirk Schulz, Dieter Fox, and Jeffrey Hightower. People Tracking with Anonymous and ID-Sensors using Rao-Blackwellised Particle Filters. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI)*, 2003. To appear.

[132] David Shenk. *The Forgetting: Alzherimer's: Portrait of an Epidemic.* Doubleday, New York, NY, 2001.

[133] Yifan Shi, Yan Huang, David Minnen, Aaron Bobick, and Irfan Essa. Propagation Networks for Recognition of Partially Ordered Sequential Action. In *Proceedings of IEEE CVPR04*, 2004.

[134] G.E. Smith, P.C. O'Brien, R.J. Ivnik, E. Kokmen, and E.G. Tangalos. Prospective Analysis of Risk Factors for Nursing Home Placement of Dementia Patients. *Neurology*, 57(8):1467–73, Oct 2001. PMID:11673591.

[135] Socket Communications Inc. Socket website, 2004. http://www.socketcom.com.

[136] Brenda C. Spillman and James Lubitz. The Effect of Longevity on Spending for Acute and Long-Term Care. *New England Journal of Medicine*, 342(19):1409–1415, 2000.

[137] M. Stager, P. Lukowicz, N. Perera, and G. Troster. Implementation and Evaluation of a Low-Power Sound-Based User Activity Recognition System. In *Proc. of the Eighth IEEE Intl. Symposium on Wearable Computers*, volume 1, pages 138–141. IEEE, October 2004.

[138] M. Stager, P. Lukowicz, N. Perera, T. von Buren, G. Troster, and T. Starner. Sound-Button: Design of a Low Power Wearable Audio Classification System. In *Proc. of the Seventh IEEE Intl. Symposium on Wearable Computers*, pages 12–17. IEEE, October 2003.

[139] Emmanuel Munguia Tapia, Stephen S. Intille, and Kent Larson. Activity Recognition in the Home Using Simple and Ubiquitous Sensors. In *Pervasive*, pages 158–175, 2004.

[140] Emmanuel Munguia Tapia, Stephen S. Intille, and Kent Larson. Activity Recognition in the Home Using Simple and Ubiquitous Sensors. In Alois Ferscha and Friedemann Mattern, editors, *Pervasive Computing, Second International Conference, PERVASIVE 2004, Vienna, Austria, April 21-23, 2004, Proceedings*, volume 3001 of *Lecture Notes in Computer Science*, pages 158–175. Springer, 2004.

[141] Ben Taskar, Pieter Abbeel, and Daphne Koller. Discriminative probabilistic models for relational data. In *Proc. of the Conf. on Uncertainty in Artificial Intelligence*, 2002.

[142] S. Thrun, J. Langford, and D. Fox. Monte Carlo Hidden Markov Models: Learning Non-Parametric Models of Partially Observable Stochastic Processes. In *Proc. of the International Conference on Machine Learning (ICML)*, 1999.

[143] Q. Tran and E. Mynatt. What Was I Cooking? Towards Deja Vu Displays of Everyday Memory. Technical Report Technical Report GIT-GVU-TR-03-33, Georgia Institute of Technology, 2003.

[144] R. Want. A Key to Automating Everything. *Scientific American*, 290(1):56–65, 2003.

[145] Roy Want, Bill Schilit, Norman Adams, Rich Gold, Karin Petersen, John Ellis, David Goldberg, and Mark Weiser. The PARCTAB Ubiquitous Computing Experiment. Technical Report CSL-95-1, Xerox Palo Alto Research Center, March 1995.

[146] Washington Metro Area Transit Authority. The Ride Guide. http://rideguide.wmata.com, 2004.

[147] G. Wei and M.A. Tanner. A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithms. *Journal of the American Statistical Association*, 85, 1990.

[148] M. F. Weiner. *Management of Emotional and Behavioral Symptoms in AD*, pages 363–69. Humana Press, 2004.

[149] Mike West and Jeff Harrison. *Bayesian Forecasting and Dynamic Models.* Springer-Verlag, second edition, 1999.

[150] Robert Wilensky. *Planning and Understanding: A Computational Approach to Human Reasoning.* Addison-Wesley, Reading, MA, 1983.

[151] Daniel H. Wilson and Chris Atkeson. Simultaneous Tracking and Activity Recognition (STAR) Using Many Anonymous, Binary Sensors. In *Proceedings of PERVASIVE 2005*, pages 62 – 79, 2005.

[152] A.W. Wimo and B. Winblad. *The Economic Burden of Alzheimer's Disease and Dementia*, pages 57–64. Humana Press, 2004.

VITA

Education

University Of Washington — **Seattle, WA**

Ph.D., Computer Science — Fall 2005

"Assisted Cognition: Activity Inference for the Cognitively Disabled"

Advisors: Professor Henry Kautz, Assistant Professor Dieter Fox

M.S., Computer Science — March 2001

Cornell University — **Ithaca, NY**

M.Eng, Electrical Engineering — January 1995

"Automatic ECG Monitoring using Neural Networks"

B.S, Computer Science, with Distinction, — Spring 1994

Awards

Fellowship: UW CSE Educators Fellowship. — 2004-2005

1-year graduate school tuition and stipend, competitive departmental award.

Honor: UW CSE Bob Bandes Teaching Assistant Award. — 2001

Cash award to the best student instructor, competitive departmental award.

Fellowship: National Defense Science and Engineering Graduate Fellowship. 1999-2002

3-year graduate school tuition and stipend, competitive national award.

Scholarship: Association of Naval Engineering Scholarship. 1994-1995

Partial graduate school tuition and benefits, competitive national award.

Scholarship: Naval Reserve Officer Training Corps Scholarship (NROTC). 1990-1994

4-year undergraduate tuition, benefits and stipend, competitive national award.

Award: Eagle Scout Award. 1989

Highest scouting award from the Boy Scouts of America.

Research Experience

University Of Washington **Seattle, WA**

Research Assistant, Artificial Intelligence. 2001-2005

Developed graphical model techniques to recognize high-level activities from low-level RFID and GPS sensor streams.

Advisors: Professors Henry Kautz and Dieter Fox

Research Assistant, Artificial Intelligence. 2001-2005

Developed graphical model techniques to recognize high-level activities from low-level RFID and GPS sensor streams.

Advisors: Professors Henry Kautz and Dieter Fox

Research Assistant, Computational Biology. 1999-2001

Developed machine learning techniques to recognize intron/exon boundaries in pre-mRNA.

Advisor: Professor Larry Ruzzo

Intel Research Seattle **Seattle, WA**

Summer Research Intern, Activity Recognition. 2003 & 2004

Designed and implemented a modeling language and statistical inference engine for probabilistic activity recognition from RFID sensors.

Publications

Refereed Conference Papers

Opportunity Knocks: a System to Provide Cognitive Assistance with Transportation Services.

Donald J. Patterson, L. Liao, K. Gajos, M. Collier, N. Livic, K. Olson, S. Wang, D. Fox, H. Kautz.

Published in the proceedings of the Sixth International Conference on Ubiquitous Computing (UBICOMP 2004), September 2004.Acceptance Rate: 18%

Mining Models of Human Activities from the Web

Mike Perkowitz, Matthai Philipose, Donald J. Patterson, Ken Fishkin.

Published in the proceedings of The Thirteenth International World Wide Web Conference (WWW 2004), May 2004. Acceptance Rate: 14.6%

Inferring High-Level Behavior from Low-Level Sensors

Donald J. Patterson, Lin Liao, Dieter Fox, Henry Kautz.

Published in the proceedings of the Fifth International Conference on Ubiquitous Computing (UBICOMP 2003), October 2003. Acceptance Rate: 14%

pre-mRNA Secondary Structure Prediction Aids Splice Site Prediction

Donald J. Patterson, Ken Yasuhara, Walter L. Ruzzo.

Published in the proceedings of the Pacific Symposium on Biocomputing (PSB 2002), January 2002.

Refereed Symposium Papers

Contextual Computer Support for Human Activity

Donald J. Patterson, Dieter Fox, Henry Kautz, Kenneth P. Fishkin, Mike Perkowitz, Matthai Philipose.

Published in the proceedings of the 2004 AAAI Spring Symposium: Interaction between Humans and Autonomous Systems over Extended Operations, March 2004.

Refereed Magazine Articles

Inferring Activities from Interactions with Objects

Matthai Philipose, Kenneth P. Fishkin, Mike Perkowitz,Donald J. Patterson, Dirk Hahnel, Dieter Fox, Henry Kautz.

Oct-Dec 2004 IEEE Pervasive Computing Magazine

Refereed Workshop Papers

Behavior Recognition in Assisted Cognition

Lin Liao, Donald J. Patterson, Dieter Fox, Henry Kautz.

Published in the proceedings of the 2004 AAAI Workshop on Supervisory Control of Learning and Adaptive Systems, July 2004

Guide: Towards Understanding Daily Life via Auto-Identification and Statistical Analysis

Matthai Philipose, Kenneth P. Fishkin, Dieter Fox, Henry Kautz, Donald J. Patterson, Mike Perkowitz.

Published in the proceedings of the UbiHealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications, October 2003.

Expressive, Tractable and Scalable Techniques for Modeling Activities of Daily Living

Donald J. Patterson, Dieter Fox, Henry Kautz, Matthai Philipose.

Published in the proceedings of the UbiHealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications, October 2003.

Research on Statistical Relational Learning at the University of Washington

Pedro Domingos, Yeuhi Abe, Corin Anderson, AnHai Doan, Dieter Fox, Alon Halevy, Geoff Hulton, Henry Kautz, Tessa Lau, Lin Liao, Jayant Madhavan, Mausam,Donald J. Patterson, Matthew Richardson, Sumit Sanghai, Daniel Weld, Steve Wolfman.

Published in the proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI) Workshop on Learning Statistical Models.

Intelligent Ubiquitous Computing to Support Alzheimer's Patients: Enabling the Cognitively Disabled

Donald J. Patterson, Oren Etzioni, Dieter Fox, Henry Kautz.

Published in the proceedings of UbiCog '02: First International Workshop on Ubiquitous Computing for Cognitive Aids,September 2002.

The Activity Compass

Donald J. Patterson, Oren Etzioni, Dieter Fox, Henry Kautz.

Published in the proceedings of UbiCog '02: First International Workshop on Ubiquitous Computing for Cognitive Aids, September 2002.

Auto-Walksat: A Self-Tuning Implementation of Walksat

Donald J. Patterson, Henry Kautz.

Published in the proceedings of SAT2001: Workshop on Theory and Application of Satisfiability Testing, June 2001.

Technical Reports

Sporadic State Estimation for General Activity Inference

Intel Research Seattle

Donald J. Patterson, Dieter Fox, Henry Kautz,Matthai Philipose.

Modeling Details of the Activity Tracker

Intel Research Seattle

Donald J. Patterson.

The Probabilistic Activity Toolkit: Towards Enabling Activity-Aware Computer Interfaces

Intel Research Seattle

M. Philipose, K. Fishkin, M. Perkowitz, Donald J. Patterson, Dirk Hahnel.

Under Preparation

Learning and Inferring Transportation Routines

Invited for fast-track review for Artificial Intelligence Journal (AIJ)

Lin Liao, Donald J. Patterson, Dieter Fox, Henry Kautz.

Invited Talks

Fine-grained Model-based Activity Recognition from RFID Sensors

NIPS 2004: Workshop on Activity Recognition and Discovery, December 2004.

A Modeling Language for Activity Recognition

Intel Research Cambridge, September 2004.

Future Technology for the Aging

Two presentations made to congressional staff and policy makers at Senate offices in Washington D.C. in conjunction with the CAST Conference, March 2004

Teaching and Advising Experience

University Of Washington

Guest Lecturer: CSE 593: Artificial Intelligence. Fall 2004

Taught graphical model theory to graduate class of 25 computer science Ph.D. students.

Research Lead: Led a team of two graduate students Fall 2004

and four undergraduate students in a research project which culminated in publication in UBICOMP 2004.

Teaching Assistant: CSE 590hk: Technology for Alzheimer's Disease. Assisted in the teaching and administration of an interdisciplinary seminar on assistive technologies. Spring 2002

Teaching Assistant: CSE 593: Artificial Intelligence. Assisted in the teaching and administration of this graduate class of 25 computer science Ph.D. students Fall 2001

Head Teaching Assistant: CSE 143: Computer Programming II. Led 6 TA's and a recitation section of 35 undergraduate students in this class on programming in C++. Bob Bandes teaching award granted as a result of this class. Fall 2000

University Of Maryland Extension Campus

Instructor: Math 100: Transitional Mathematics. Taught two night classes per week, developed curriculum, graded student work and held office hours for this class of 15 American students living in Italy. Winter 1998

Work Experience

Seattle Institute For Biomedical And Clinical Research Seattle, WA

Data Modeler, Nephrology Research Group 2002-2005

Worked with a team of medical doctors to create data abstractions to model and analyze Veteran Administration patient databases. This was in support of epidemiological studies of kidney disease.

U.S. Navy

Operations Officer, La Maddalena, Italy, 1997 - 1999
USS SIMON LAKE (AS-33),
Supervised an 80-person department responsible for computing infrastructure, radio and satellite operation, and navigation for the 1500-person crew of a forward-deployed submarine tender.

Strike Officer, Yokosuka, Japan, 1995 - 1997
USS CURTIS WILBUR (DDG-54),
Supervised a 15-person division responsible for the maintenance and operation of the Tomahawk missile system on a forward-deployed destroyer.

Service

Program Committee:

The 20th National Conf. on Artificial Intelligence 2005

External Reviewer:

The 7th International Conf. on Ubiquitous Computing 2005

External Reviewer:

IEEE Transactions on Information Technology in BioMedicine 2005

External Reviewer:

The 3rd International Conf. on Pervasive Computing 2005

External Reviewer:

The 6th International Conf. on Ubiquitous Computing 2004

Committee Chair:

UW CSE Graduate Student Recruiting Chair 2002-2003

Student Coordinator:

UW CSE Graduate Student Coordinator 2001-2002

Social Coordinator:

UW CSE Graduate Student Social Coordinator 2000-2001

Press Coverage

MIT Technology Review : October 2004

"Portable Pathfinder"

The Futurist: September/October 2004

"AI Helps Keep Seniors Mobile"

IEEE Computer: April 2004

"Inventing Wellness Systems for Aging in Place"

Wired Online: March 19, 2004

"RFID Keeps Track of Seniors"

New Scientist Online: March 17, 2004

"RFID Chips Watch Grandma Brush Teeth"

USA Today: December 4, 2002

"Parents, Athletes Put GPS to Work"

Newsweek: September 23, 2002

"Gray Market for Gadgets"

Focus: September 2002

"Digitaler Betreuer fur Alzheimer-Patienten"

University Week: July 25, 2002

"Prompted to Live"

Minnesota Public Radio: July 24, 2002

"Artificial Intelligence and Alzheimer's"

USA Today: July 23, 2002

"Surveillance casts an eye to the future"

Wired Online: June 24, 2002

"AI to Assist Alzheimer's Patients"

References

Henry Kautz

Professor, University of Washington,

kautz@cs.washington.edu, (206) 543-1896

Dieter Fox

Assistant Professor, University of Washington,

fox@cs.washington.edu, (206) 685-2517

Gaetano Borriello

Professor, University of Washington,

Director of Intel Research Seattle

gaetano@cs.washington.edu, (206) 685-9432

John Krumm

Researcher, Microsoft Research

jckrumm@microsoft.com