1

Introduction

"All growth depends upon activity."
—Calvin Coolidge

"From time to time, it is worth wandering around the fuzzy border regions of what you do, if only to remind yourself that no human activity is an island."
—Julian Baggini

"People love chopping wood. In this activity one immediately sees results."
—Albert Einstein

"Nothing is more terrible than activity without insight."
—Thomas Carlyle

"We are what we repeatedly do."
—Aristotle

Learning and understanding observed activities is at the center of many fields of study. An individual's activities affect that individual, those around him, society, and the environment. In the past, theories about behavior and activities were formed based on limited observation. Over the past decade, the maturing of sensor technologies has made it possible to automate activity learning. Sensors have become miniature, low power, low cost, and high capacity. At the same time, we have witnessed tremendous progress in the areas of wireless networks, data processing, and machine learning. These advances have shifted researchers' focus from low level data collection and transmission to high-level information collection, inference, and recognition.

Activity Learning: Discovering, Recognizing, and Predicting Human Behavior from Sensor Data, First Edition. Diane J. Cook and Narayanan C. Krishnan.

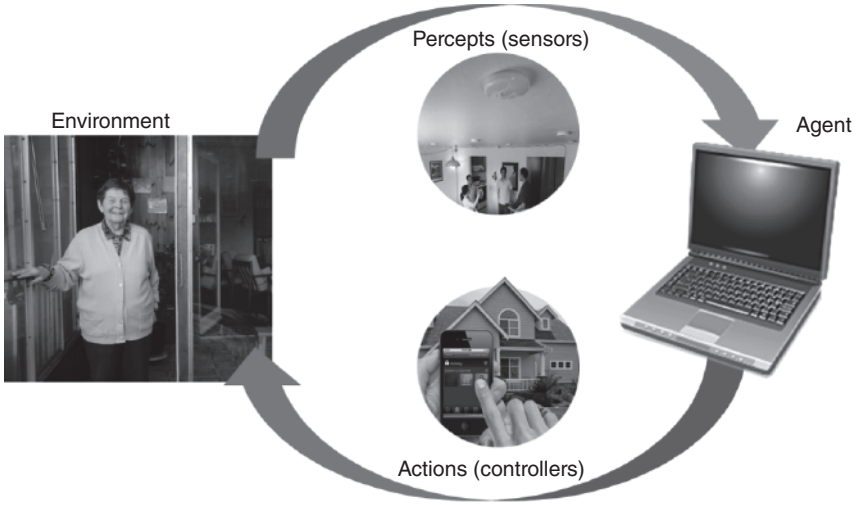

FIGURE 1.1 *The role of activity learning in the design of an intelligent agent.*

Activity learning plays a key role in the design of intelligent agents. Russell and Norvig[1] define an agent as an entity that perceives its environment through sensors and acts upon the environment through actuators. The techniques we describe in this book add intelligence and activity-awareness to this concept of an agent, shown in Figure 1.1. The state of individuals and their environment is perceived by sensors and stored as raw data. The activity learning agent processes this data in order to analyze it, make inferences from it, and describe it in terms of recognizable activities. Activity patterns can then be used to make a more informed decision about actions that can be taken to change the environment in a way that achieves the agent's goal.

Because activity learning technologies have become more robust, automated modeling and tracking of activities have become integral components of numerous solutions to real-world problems. For example, surveillance and security systems try to recognize and predict activities in order to address terrorist threats. Ambient-assisted living exploits activity discovery and monitoring to support independent living for individuals with cognitive and physical impairments. Activity-aware services have transformed ideas such as intelligent meeting rooms, automated homes, and personalized digital assistants from science fiction to everyday reality. Activity learning has also recently made its debut as a core component in products including game consoles and fitness smart phone apps.

As a result of the technology push and application pull, the number and diversity of projects that are designing or utilizing activity learning is exploding. Activity discovery and recognition is found in settings where data can be captured by cameras, body sensors, phones, or buildings. Entire conferences and journals are dedicated to reporting the latest advances in the field. Over the last 3 years, there have been two dozen workshops specifically devoted to activity recognition. The interest and enthusiasm for this topic is still increasing.

The goal of this text is to provide an in-depth look at computational approaches to activity learning from sensor data. Activity learning here refers to several aspects of activity processing. These include features that can be used to represent and model activities from sensor data. Also included are basic techniques for learning activity concepts and recognizing them in offline or streaming mode from sensor data. A companion to the idea of learning activity concepts from labeled data is the idea of automatically discovering activity patterns from unlabeled data, which we describe in detail. The third major component of activity learning that we include is the prediction or forecasting of activities that will occur at a time point in the future.

This book contains a pragmatic introduction to the primary algorithms and approaches to activity learning. Because the emphasis is on computational approaches, we will provide an algorithmic perspective. We detail the methods using pseudocode and provide implementations of the techniques in the source code that accompanies this book. We also illustrate the methods using examples of sensor data collected in real-world settings.

We note that the book is intended to be an algorithmic reference. It is written, where possible, to be independent of any specific sensor device or class of devices. Some of the examples used to illustrate the techniques are based on data collected from a particular type of sensor such as environment sensors, smart phone sensors, or cameras. However, the data features that are used to represent and recognize the data can be extracted and abstracted from a wide variety of low-level data collection sources. Here is a brief overview of the remaining chapters in the book.

Chapter 2 introduces the terminology that will be used throughout the text. The definitions of actions, interactions, and activities are provided and related to each other. A discussion of activity classifications and activity ontology generation is provided.

Chapter 3 provides an overview of sensors that are commonly used for observing and learning activities. In addition to describing sensor types and capabilities, the chapter includes descriptions of high-level features that can be extracted from low-level sensor data.

Chapter 4 introduces machine learning techniques that will be used throughout the remainder of the text. In addition to supervised classification techniques, the chapter also discusses approaches to dimensionality reduction that is valuable when learning concepts from high-dimensional sensor features.

Chapter 5 presents concepts and approaches for activity recognition. Additionally, methods for segmenting data into distinct activities are described together with methods to process streaming data. The chapter also includes performance metrics that can be used to evaluate and compare alternative activity recognition algorithms.

Chapter 6 describes algorithmic techniques for discovering activity patterns from sensor data. These techniques include zero-shot learning, sequence mining, clustering, and topic models. Methods for evaluating the results of activity discovery are summarized.

Chapter 7 covers activity prediction. Activity prediction includes identifying the next activity in a sequence as well as forecasting the next point in time when a particular activity will occur. As with activity recognition and discovery, metrics for evaluating the performance of activity prediction are included.

Chapter 8 discusses activity learning "in the wild." This refers to computational techniques that can be used to address problems that are encountered when activity learning is used in complex, real-world situations. These problems include obtaining sufficient high-quality labeled activity sensor data, transferring data or learned activity models to new settings with new users, mapping sensor sequences to multiple activity labels, and learning activities in the presence of multiple individuals.

Chapter 9 provides an overview of activity learning applications that are currently being investigated. Ongoing work is described that is related to activity recognition for health monitoring and intervention, marketing and activity-aware automation, surveillance and emergency management, describing routine behaviors in terms of activities, and analyzing human dynamics for entire populations.

Chapter 10 provides a conclusion of the book. This chapter poses some grand challenge problems for activity learning as well as uses for activity learning in other fields.

Additional materials including chapter updates, data sets, implementation of algorithms described in the book, and links to other tools and ongoing activity learning research projects are available online. These resources can be found at http://eecs.wsu.edu/~cook/albook.