Cognitive Psychology 76 (2015) 30–77

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Inferring action structure and causal relationships in continuous sequences of human action

Daphna Buchsbaum [a,*], Thomas L. Griffiths [b], Dillon Plunkett [b], Alison Gopnik [b], Dare Baldwin [c]

[a] University of Toronto, Department of Psychology, 100 St. George Street, 4th Floor, Sidney Smith Hall, Toronto, ON M5S 3G3, Canada
[b] University of California, Berkeley, Department of Psychology, 3210 Tolman Hall # 1650, Berkeley, CA 94720-1650, USA
[c] Department of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227, USA

A R T I C L E I N F O

Article history:
Accepted 22 October 2014
Available online 18 December 2014

Keywords:
Statistical learning
Causal inference
Action segmentation
Rational analysis
Bayesian inference
Event segmentation

A B S T R A C T

In the real world, causal variables do not come pre-identified or occur in isolation, but instead are embedded within a continuous temporal stream of events. A challenge faced by both human learners and machine learning algorithms is identifying subsequences that correspond to the appropriate variables for causal inference. A specific instance of this problem is action segmentation: dividing a sequence of observed behavior into meaningful actions, and determining which of those actions lead to effects in the world. Here we present a Bayesian analysis of how statistical and causal cues to segmentation should optimally be combined, as well as four experiments investigating human action segmentation and causal inference. We find that both people and our model are sensitive to statistical regularities and causal structure in continuous action, and are able to combine these sources of information in order to correctly infer both causal relationships and segmentation boundaries.

1. Introduction

Human social reasoning depends on understanding the relationship between actions, goals and outcomes. In order to understand the reasons behind others' behavior, we must be able to distinguish

* Corresponding author. Fax: +1 (617) 209 1068.
E-mail addresses: Buchsbaum@psych.utoronto.ca (D. Buchsbaum), tom_griffiths@berkeley.edu (T.L. Griffiths), dillonplunkett@berkeley.edu (D. Plunkett), gopnik@berkeley.edu (A. Gopnik), baldwin@uoregon.edu (D. Baldwin).

http://dx.doi.org/10.1016/j.cogpsych.2014.10.001
0010-0285/

the unique actions we see others performing, and recognize the effects of these actions. Imagine watching someone coming home and opening their front door. To understand this simple scene, an observer needs to identify meaningful behaviors from within the continuous stream of motion they see, such as "exiting the car", "coming up the stairs" and "opening the door", which are themselves composed of smaller motion elements such as "standing up", "closing the car door", "taking a step", "reaching for the doorknob", and so on.

Determining which subsequences of motion go together hierarchically, and what outcomes they produce, is also an important instance of the more general problem of causal variable discovery (a similar problem – determining how spatially distributed observations should be encoded as variables – is discussed by Goodman, Mansinghka, & Tenenbaum (2007)). Consider the case of learning which actions are necessary to open a door by observing multiple performances, embedded in everyday scenes such as the one above. A learner might notice that people almost always grasp and then turn a doorknob before the door opens, but sometimes they pull a handle instead. They frequently insert a key into a lock and then turn it before trying the doorknob, but not always. Often, other actions precede the door opening as well – putting down groceries, fumbling around in a purse, ringing a doorbell, sliding a bolt – which of these are causally necessary and which are incidental? While this ambiguity can make causal learning more challenging, the presence of statistical variation can actually aid inference. Motions that do not consistently precede outcomes are less likely to be causally necessary. Motions that reliably appear together and, in fact, predict each other, are more likely to be coherent units, corresponding to intentional, goal-directed action.

There is now a large body of evidence suggesting that both infants and adults can use statistical patterns in spoken language to help solve the related problem of segmenting words from continuous speech (for a partial review, see Gómez & Gerken, 2000). Recently, Baldwin, Andersson, Saffran, and Meyer (2008) demonstrated that a similar sensitivity to statistical regularities in continuous action sequences may play an important role in action processing. However, a key difference between action segmentation and word segmentation is that intentional actions usually have effects in the world. In fact, many of the causal relationships we experience result from our own and others' actions, suggesting that understanding action may bootstrap learning about causation, and vice versa. Here we present a combination of experimental and computational approaches investigating how the ability to segment action and to infer its causal structure functions and develops.

We first introduce a Bayesian analysis of action segmentation and causal inference, which provides a rational analysis of how statistical and causal cues to segmentation should optimally be combined. Next, we present four experiments investigating how both people and our model use statistical and causal cues to action structure. Our first experiment demonstrates that people are able to segment statistically determined actions using only the co-occurrence patterns between motions. This experiment is also the first to demonstrate that the continuous boundary judgment measures used in event segmentation research align with the sequence discrimination measures traditionally used in the statistical segmentation literature. Our second experiment demonstrates that people experience these actions as coherent, meaningful, and most importantly, causal sequences. Our third experiment shows that people are able to extract the correct causal variables from within a longer action sequence, and that they find causal sequences to be more coherent and meaningful than other sequences with equivalent statistical structure. Our fourth experiment demonstrates that, when statistical and causal cues conflict, both sets of cues influence segmentation and causal inference, suggesting that action structure and causal structure are learned jointly and simultaneously, and demonstrates that these results are not accounted for by simpler heuristic models. We conclude by discussing our results in the context of broader work, as well as its implications for more generalized human statistical learning abilities.

2. Background

Many if not most of the causal outcomes we witness are the result of intentional human action. We must be able to distinguish the unique actions we see other people performing and recognize their effects in order to understand the reasons behind others' behavior, and in order to potentially bring