

# Serial vs Hierarchical Pattern Recognition

s.butterfill@warwick.ac.uk

The distinction between serial and hierarchical patterns is analogous to that between finite state and phrase structure grammars (Chomsky 1956), but can be understood informally for present purposes. Serial patterns involve relations among consecutive chunks of behaviour and so can be identified by relatively simple learning mechanisms that track statistical relations between pairs of consecutive chunks (Conway and Christiansen 2001). Some serial patterns arise from the fact that pursuit of a goal characteristically involves a sequence of adjacent chunks; for example, eating a burger involves grasping it and then repeatedly bringing it to the mouth. Hierarchical patterns involve relations among non-consecutive chunks. Such patterns sometimes occur because pursuit of a goal involves subgoals that can be performed in different orders, as when making a burger involves preparing the bun and fillings individually before bringing them together. Patterns in the behavioural chunks that constitute burger-making activities cannot be learnt only by tracking relations between pairs of consecutive chunks.

The distinction between serial and hierarchical patterns occurs in thinking about speech perception, where the organisation of phonemes into words counts as a sequential pattern (ignoring, for simplicity, additional structure required to accommodate variability in the allophones that realise phonemes), whereas organisation of words into clauses counts as an hierarchical pattern.

## 1 Sequential patterns

Consider the following sequence of mundane actions:

... push door, walk through frame, walk to desk, grasp paper, orient to pen, reach for pen, write on paper, fold paper, place pen, reach for headphones, push button, ...

Here the flow of movements is described in terms of phoneme-like chunks that infants and adults can readily identify (Saylor et al. 2007). Infants (from around 9 months) and adults can also group these chunks into word-like units such as the letter-writing sequence in the above example: these are chunks that seem to belong together and to be separate from previous and ensuing chunks (Saylor et al. 2007). The boundaries of these units are generally ‘intention-relevant’ (Saylor et al. 2007), corresponding to the completion of a goal or subgoal.

In some cases, subjects have prior knowledge of intention or purpose which guides the identification of word-like behaviour units (Zacks 2004). However, ascribing intention is not generally necessary for isolating these units. There are at least two types of non-intentional cue to their boundaries. First, commencement and completion of a goal or subgoal typically coincide with dramatic changes in the physical features of the movements (Zacks, Tversky, and Iyer 2001). Baldwin and Baird express this idea graphically with the notion of a ‘ballistic trajectory that provides a temporal contour or ‘envelope’ demarcating one intentional act from the next’ (Baldwin et al. 2001). Research using schematic animations has shown that adults use a variety of movement features to group behavioural chunks into word-like units (Zacks 2004; Hard, Tversky, and Lang 2006). The second non-intentional cue is statistical. Chunks of behaviour that are all steps to a single goal or subgoal are more likely to occur in succession than chunks not so related; thus transitional probabilities in the sequence of chunks could in principle be used to identify intention-relevant units, much as phonemes can be grouped into words by means of tracking transitional probabilities (Saffran, Newport, and Aslin 1996; Baldwin et al. 2008; Gómez and Gerken 2000). In fact, Baldwin and colleagues demonstrated that adults can learn to group small chunks of behaviour into larger word-like units on the basis of statistical features alone (Baldwin et al. 2008). And since the statistical learning mechanism required for discerning such units is automatic (Fiser and Aslin 2001), domain-general (Kirkham, Slemmer, and Johnson 2002) and present in other species including monkeys (Hauser, Newport, and Aslin 2001) and rats (Toro and Trobalón 2005), it seems plausible to suppose that sensitivity to intention-relevant units in sequences of behavioural chunks can be found in individuals that lack the ability to ascribe intentions or goals.

In short, domain-general learning mechanisms enable chunks of behaviour to be grouped into word-like units on the basis of ballistic envelopes and transitional probabilities. Just as words are semantically relevant units yet can be discerned without knowledge of semantics, so these behavioural units have boundaries corresponding to the fulfilment of goals and intentions but can be identified without knowledge of goals or intentions.

Note that discerning pattern in sequences of behaviour is not limited to intention-

relevant units: behaviour reading may group behavioural chunks into units that have no clear relevance to a goal, as studies requiring subjects to segment behaviours using reversed animations and other manipulations show (Hard, Tversky, and Lang 2006; Baldwin et al. 2008; Kurby and Zacks 2008). Similarly, use of statistical cues to group phonemes into words does not yield only meaningful words. Discerning intention-irrelevant units may be undesirable overgeneration or it may have predictive value. The definition of behaviour-reading allows for the exploitation of any useful patterns, not just those induced by goals and intentions.

## 2 Hierarchical patterns

In discussing sequential patterns we were concerned with grouping together strings of consecutive behavioural chunks. But there are also patterns in behaviour connecting non-consecutive chunks. For example, making a burger involves several steps whose order is only loosely constrained, where some of these steps can be omitted or replaced (veggie burger, hamburger) and where steps can be interspersed with irrelevant actions (answering the phone). Grouping together all and only the behavioural units involved in making a burger therefore involves discerning hierarchical structure. Clearly this is possible, but is it possible independently of ascribing intention?

Several researchers who appear to be addressing this question are actually concerned only with identifying word-like units (Zacks, Tversky, and Iyer 2001; Hard, Lozano, and Tversky 2006; Hard, Tversky, and Lang 2006; Kurby and Zacks 2008). They label this ‘hierarchical structure’ because flowing motion is first being divided into phoneme-like behavioural chunks whose organisation then defines boundaries of larger word-like units. We can thus regard the chunks as parts of larger units, which justifies the label ‘hierarchical’. However, discerning word-like units only requires sequential learning—that is, it only requires tracking associations between consecutive chunks. Sensitivity to these ‘hierarchical structures’ does not imply that one can also detect patterns involving optional components or variable ordering, such as those involved in making a burger or preparing a nettle. Following Conway and Christiansen (2001), I therefore reserve the label ‘hierarchical’ for patterns involving relations among potentially non-consecutive chunks.

To my knowledge the question of learning hierarchical patterns in behaviour has not been directly addressed. We do know from research in other domains that hierarchical patterns can be learnt by adults (Newport and Aslin 2004) and infants from 12-months or earlier (Saffran et al. 2008). Learning hierarchical patterns differs from learning sequential patterns in that it requires attention and cannot

occur when subjects are distracted by a secondary task (Cohen, Ivry, and Keele 1990) and there appear to be constraints on the types of pattern that can be learnt (Newport and Aslin 2004).

There is uncertainty on whether non-humans are capable of learning hierarchical patterns of any kind. Evidence for an absence of hierarchical learning has been offered for cotton-top tamarins (Fitch and Hauser 2004) and rats (Toro and Trobalón 2005), and there is no convincing evidence for such learning in nonhumans generally (Corballis 2007). Inability to learn hierarchical patterns would limit the power of behaviour reading based on domain-general learning mechanisms to activities that are often uniformly realised (little or no variation in the steps taken or the order in which they are performed) and performed without interruption.

In short, acting on hierarchically structured intentions and goals tends to create statistical and prosodic cues in the pattern of behaviours. Discerning these cues makes it possible to recover the structure of the behaviour without knowledge of the goals or intentions responsible for them. Such patterns can be discerned by humans, perhaps including infants, but probably not by other animals. Learning these patterns is not automatic and requires attention to patterns of behaviour.

### **3 Other cues to hierarchical structure**

Human infants use prosodic cues such as changes in pitch and pauses used by infants to identify clause boundaries (Soderstrom, Nelson, and Jusczyk 2005; Seidl and Cristià 2008). It is possible that somewhat similar cues in behaviour—changes in motion features and pauses—make it possible to identify where hierarchically structured behaviours begin and end. The availability of such cues to human infants may be boosted by ‘motionese’. Brand, Baldwin, and Ashburn (2002) found that ‘mothers spontaneously modified their infant-directed actions in a number of ways ... a higher level of interactiveness, greater repetitiveness and movements that were larger in scale but reduced in complexity.’ These modifications might in principle assist infants’ efforts to detect hierarchical patterns in behaviour, amplifying and inserting motion features and pauses to reduce infants’ dependence on statistics.

Richard Byrne has suggested that identifying the parts of hierarchically structured behaviours may be facilitated by comparison between different occasions (Byrne 2003). In his study of Rwandan mountain gorillas’ preparation of nettles for eating, a complex task involving several sub- and sub-sub-goals, he found that repetitions, omissions, and substitutions indicate boundaries of significant units,

as do points where one-off (and therefore probably extraneous) behaviours are interwoven.

The existence of such cues suggests that, in special cases or where assisted by experts who use motionese, it may be possible in principle to identify plan-induced hierarchical patterns in behaviours without knowledge of intention and independently of any domain-general capacity for recognising hierarchical patterns.

## References

- Baldwin, Dare, Annika Andersson, Jenny R. Saffran, and Meredith Meyer. 2008. "Segmenting Dynamic Human Action via Statistical Structure." *Cognition* 106 (3): 1382–1407.
- Baldwin, Dare, Jodie A. Baird, Megan M. Saylor, and M. Angela Clark. 2001. "Infants Parse Dynamic Action." *Child Development* 72 (3): 708–17.
- Brand, Rebecca J., Dare Baldwin, and Leslie A. Ashburn. 2002. "Evidence for 'Motionese': Modifications in Mothers' Infant-Directed Action." *Developmental Science* 5 (1): 72–83.
- Byrne, Richard W. 2003. "Imitation as Behaviour Parsing." *Philosophical Transactions: Biological Sciences* 358 (1431): 529–36.
- Chomsky, N. 1956. "Three Models for the Description of Language." *Information Theory, IRE Transactions on* 2 (3): 113–24.
- Cohen, Asher, Richard I. Ivry, and Steven W. Keele. 1990. "Attention and Structure in Sequence Learning." *Journal of Experimental Psychology: Learning* 16 (1): 17–30.
- Conway, Christopher M., and Morten H. Christiansen. 2001. "Sequential Learning in Non-Human Primates." *Trends in Cognitive Sciences* 5 (12): 539–46.
- Corballis, Michael C. 2007. "Recursion, Language, and Starlings." *Cognitive Science: A Multidisciplinary Journal* 31 (4): 697–97.
- Fiser, J. zsef, and Richard N. Aslin. 2001. "Unsupervised Statistical Learning of Higher-Order Spatial Structures from Visual Scenes." *Psychological Science* 12 (6): 499–504.
- Fitch, W. Tecumseh, and Marc D. Hauser. 2004. "Computational Constraints on Syntactic Processing in a Nonhuman Primate." *Science* 303 (5656): 377–80.
- Gómez, Rebecca, and LouAnn Gerken. 2000. "Infant Artificial Language Learning and Language Acquisition." *Trends in Cognitive Sciences* 4 (5): 178–86.

- Hard, Bridgette Martin, Sandra C. Lozano, and Barbara Tversky. 2006. "Hierarchical Encoding of Behavior: Translating Perception into Action." *Journal of Experimental Psychology: General*. Vol 135(4) 2006: 588–608.
- Hard, Bridgette Martin, Barbara Tversky, and David S. Lang. 2006. "Making Sense of Abstract Events: Building Event Schemas." *Memory & Cognition* 34: 1221–35.
- Hauser, Marc D., Elissa L. Newport, and Richard N. Aslin. 2001. "Segmentation of the Speech Stream in a Non-Human Primate: Statistical Learning in Cotton-Top Tamarins." *Cognition* 78 (3): B53–64–B53–64.
- Kirkham, Natasha Z., Jonathan A. Slemmer, and Scott P. Johnson. 2002. "Visual Statistical Learning in Infancy: Evidence for a Domain General Learning Mechanism." *Cognition* 83 (2): B35–42–B35–42.
- Kurby, Christopher A., and Jeffrey M. Zacks. 2008. "Segmentation in the Perception and Memory of Events." *Trends in Cognitive Sciences* 12 (2): 72–79.
- Newport, Elissa L., and Richard N. Aslin. 2004. "Learning at a Distance I. Statistical Learning of Non-Adjacent Dependencies." *Cognitive Psychology* 48 (2): 127–62.
- Saffran, Jenny R., Marc D. Hauser, Rebecca Seibel, Joshua Kapfhamer, Fritz Tsao, and Fiery Cushman. 2008. "Grammatical Pattern Learning by Human Infants and Cotton-Top Tamarin Monkeys." *Cognition* 107 (2): 479–500.
- Saffran, Jenny R., Elissa L. Newport, and R. N. Aslin. 1996. "Statistical Learning by 8-Month-Old Infants." *Science* 274 (5294): 1926–8.
- Saylor, Megan M., Dare Baldwin, Jodie A. Baird, and Jennifer LaBounty. 2007. "Infants' on-Line Segmentation of Dynamic Human Action." *Journal of Cognition and Development* 8 (1): 113–13.
- Seidl, Amanda, and Alejandrina Cristià. 2008. "Developmental Changes in the Weighting of Prosodic Cues." *Developmental Science* 11 (4): 596–606.
- Soderstrom, Melanie, Deborah G. Kemler Nelson, and Peter W. Jusczyk. 2005. "Six-Month-Olds Recognize Clauses Embedded in Different Passages of Fluent Speech." *Infant Behavior and Development* 28 (1): 87–94.
- Toro, Juan M., and Josep B. Trobalón. 2005. "Statistical Computations over a Speech Stream in a Rodent." *Perception & Psychophysics* 67: 867–75.
- Zacks, Jeffrey M. 2004. "Using Movement and Intentions to Understand Simple Events." *Cognitive Science* 28 (6): 979–1008.
- Zacks, Jeffrey M., Barbara Tversky, and Gowri Iyer. 2001. "Perceiving, Remembering, and Communicating Structure in Events." *Journal of Experimental Psychology: General*. Vol. 130(1) 130 (1): 29–58.