Journal of Mathematical Psychology 59 (2014) 50–64

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

Markovian interpretations of dual retrieval processes

C.F.A. Gomes, C.J. Brainerd *, K. Nakamura, V.F. Reyna

Department of Human Development, Cornell University, Ithaca, NY 14853, United States

HIGHLIGHTS

- We develop mathematical machinery for Markov models of dual retrieval processes.
- Dual processes are measured for learning, forgetting and reminiscence.
- The Markov models are applied to a corpus of 230 recall experiments.
- Application of the model to test theoretical hypotheses is illustrated.

ARTICLE INFO

Article history:
Available online 12 August 2013

Keywords:
Markov chains
Recollection
Reconstruction
Familiarity
Fuzzy-trace theory

ABSTRACT

A half-century ago, at the dawn of the all-or-none learning era, Estes showed that finite Markov chains supply a tractable, comprehensive framework for discrete-change data of the sort that he envisioned for shifts in conditioning states in stimulus sampling theory. Shortly thereafter, such data rapidly accumulated in many spheres of human learning and animal conditioning, and Estes' work stimulated vigorous development of Markov models to handle them. A key outcome was that the data of the workhorse paradigms of episodic memory, recognition and recall, proved to be one- and two-stage Markovian, respectively, to close approximations. Subsequently, Markov modeling of recognition and recall all but disappeared from the literature, but it is now reemerging in the wake of dual-process conceptions of episodic memory. In recall, in particular, Markov models are being used to measure two retrieval operations (direct access and reconstruction) and a slave familiarity operation. In the present paper, we develop this family of models and present the requisite machinery for fit evaluation and significance testing. Results are reviewed from selected experiments in which the recall models were used to understand dual memory processes.

In 1959, after spending a summer studying with J. Kemeny, a mathematician whose specialty was Markov chains (e.g. Kemeny & Snell, 1959), Estes published a comprehensive treatment of that topic for learning and memory researchers. The proximal reason for that essay was to expand the mathematical foundations of stimulus sampling theory to deal with what, at the time, was known as the small N problem. In stimulus sampling theory, a set of N stimulus elements is randomly sampled on each trial of a learning experiment, and following reinforcement or feedback or a study trial, the elements change their conditioning states in an all-or-none fashion. As long as N was assumed to be large, as it traditionally had been, the law of large numbers applied, and the ordinary mathematics of probability theory could be used to derive quantitative predictions from stimulus sampling theory. However, there were certain types of experiments, such as discrimination learning, for which it seemed plausible that N must be small—so that ordinary probability theory could no longer be used. In such circumstances, predictions could be derived by implementing stimulus sampling theory in finite Markov chains, and Estes developed the necessary mathematical machinery to accomplish that. During the ensuing years, those developments were widely used by other authors (e.g. Bower, 1961; Greeno & Steiner, 1964; Restle, 1962; Waugh & Smith, 1962), but not to derive predictions from stimulus sampling theory. Instead, finite Markov chains provided a natural framework for handling the all-or-none learning demonstrations that began to appear in the literature at about the same time as Estes' essay (e.g. Estes, 1960; Rock, 1958; Trabasso, 1963).

All-or-none learning refers to discrete changes in empirical response probabilities, rather than in theoretical entities, such as stimulus elements. Specifically, at the level of individual items, the probability of a correct response only takes on a small set of values, such as 0, some constant $0 < p < 1$, and 1. In the span of a few years, experiments in which subjects received multiple opportunities to study a focal list and to retrieve it produced data of that sort. Some tasks, such as associative recognition (Bower, 1961) and concept identification (Restle, 1962; Trabasso, 1963), were found to be one-stage Markovian; that is, variability in the probability

* Corresponding author.
E-mail addresses: cf365@cornell.edu (C.F.A. Gomes), cb299@cornell.edu (C.J. Brainerd), kn252@cornell.edu (K. Nakamura), vr53@cornell.edu (V.F. Reyna).

0022-2496/$ – see front matter
http://dx.doi.org/10.1016/j.jmp.2013.07.003

of a correct response consisted of an initial unlearned state U, in which the probability was some constant $0 < p < 1$, and a terminal learned state L, in which the probability was 1. Other tasks, such as associative recall (Greeno, 1968) and free recall (Waugh & Smith, 1962), were found to be two-stage Markovian; that is, variability in the probability of a correct response consisted of an initial unlearned state U, in which the probability was 0, a partially learned state P, in which the probability was some constant $0 < p < 1$, and a terminal learned state L, in which the probability was 1. Using the developments in Estes (1959), it was a simple matter to derive the sampling distributions of a plethora of statistics for these models (e.g., trial number of the last error, trial number of first success, and auto-correlation of errors) and to use them to conduct rigorous quantitative evaluations of one- and two-stage Markov models for the indicated paradigms. Consequently, Markov chains dominated mathematical modeling work on basic memory paradigms for some years (for reviews, see Brainerd, Howe, & Desrochers, 1982; Greeno, 1974). Ultimately, they faded from view as more complex modeling efforts were mounted, with the appearance of global matching models of memory being a prominent case in point (for a review, see Clark & Gronlund, 1996).

Recently, however, there has been a resurgence of Markovian analyses of memory, for reasons not unlike those that motivated Estes (1960) essay—namely, the need for a mathematical framework that will help us gain leverage on important theoretical problems. In this instance, the theoretical motivation is supplied by dual-process conceptions of episodic memory. According to such conceptions (e.g. Jacoby, 1991; Mandler, 1980), two forms of retrieval control memory for the events of our lives. The first is a recollective form, in which retrieval of episodic memories is accompanied by realistic phenomenology—vivid restatement of events' prior occurrences and details of the contexts in which events were experienced. The second is a nonrecollective form, which confers high confidence that the events were experienced but is not accompanied by realistic phenomenology. The distinction between recollective and nonrecollective forms of retrieval has been exploited, particularly over the last three decades, to explain memory deficits in patients with brain lesions (Schacter & Tulving, 1994) and neurodegenerative diseases (Bugaiska, Morson, Moulin, & Souchay, 2011), brain activity during encoding and retrieval (Diana, Yonelinas, & Ranganath, 2007; Ranganath, 2010), memory development (Gomes & Brainerd, 2013; Reyna & Mills, 2007), and the effects of various experimental variables on memory, such as list length (Cary & Reder, 2003), word frequency (Gardiner & Java, 1990; Guttentag & Carroll, 1997), time of testing (Knowlton & Squire, 1995; Tulving, 1985), and subjects' level of attention during study (Gardiner & Parkin, 1990; Yonelinas, 2001), among others (for a review, see Yonelinas, 2002)).

Dual-process distinctions have been studied primarily with old/new item recognition. To do that, old/new recognition has been enriched with metacognitive tasks that supposedly disentangle its recollective and nonrecollective components. Remember/know judgments (Tulving, 1985) are far and away the most commonly used procedure, but confidence judgments, which allow recollective and nonrecollective components of the receiver operating characteristic (ROC) to be estimated (Yonelinas, 1994), and inclusion versus exclusion instructions, which allow the recollective and nonrecollective parameters of the process-dissociation model (Jacoby, 1991) to be estimated, have also been used. There are serious validity challenges to these recognition methodologies, however, and that has led to renewed interest in Markov models *of recall* as a response to the challenges—all of which forms the substance of the present paper.

The paper consists of four main sections. The first sketches traditional recognition-based approaches to measuring dual memory processes, summarizes recent validity criticisms of those approaches, and discusses the need for alternative measurement procedures that respond to the criticisms. The second section presents a recall-based response that we have implemented in recent experimentation. This approach uses fuzzy-trace theory's (FTT) distinction between verbatim and gist traces of experience (e.g. Brainerd & Reyna, 2004; Reyna & Brainerd, 2011) to interpret the states of two-stage Markov chains in such a way that measures of dual memory processes can be extracted from raw recall data, without the need of additional metacognitive tasks. In the third section, the mathematical machinery for this family of models (matrix representations and fit tests) is developed. In the fourth section, two worked illustrations of how these techniques can be applied experimentally to answer theoretical questions about episodic memory are presented. The first focuses on how factor analysis of parameter spaces over large numbers of data sets can be used to test the theoretical assumptions that lie behind this family of models. The second shows how the model can be extended to study dual memory processes in an important domain in which virtually no data on them are currently available, namely child-to-adult development.

1. Dual-process conceptions of recognition and recall

That episodic memory is controlled by two distinct forms of retrieval, recollective and nonrecollective, is, of course, one of the touchstones of contemporary memory research. Strong (1913) is usually credited with being the first to provide an experimental demonstration of this distinction, but current interest can be traced to a theoretical paper by Mandler (1980). Strong reported a surprisingly modern recognition experiment, in which his subjects studied a list of target words and then responded to an old/new recognition test. After that, subjects were asked to review the items that they had accepted as old, both hits and false alarms, to introspect on the phenomenology that each provoked, and to write a brief description of that phenomenology. Despite great variability in the specifics of what subjects wrote, Strong noticed that the responses were of two basic sorts. Sometimes, old responses were justified by describing realistic details that had been associated with their presentation (e.g., that a subject's stomach had growled or that a subject had remembered a dental appointment), whereas at other times, old responses were justified by assertions of confidence that were not backed up by realistic details (e.g., "just know it was on the list", "must have seen it"). Strong also noticed that the first type of report, now called recollection, was overwhelmingly stimulated by hits, whereas false alarms primarily stimulated the second type of report, now called familiarity. In his later paper, Mandler reviewed experimental demonstrations of the recollection-familiarity distinction and provided compelling anecdotal illustrations, such as the well-known butcher on the bus example. He also developed a simple multinomial model of how recollection and familiarity affect hit rates, in which their effects were jointly independent and additive.

The modern literature on recollection and familiarity preserves two key features of Strong's (1913) original demonstration: Old/new recognition is the focal memory paradigm, and the two forms of retrieval are measured by asking subjects to introspect on the phenomenological qualities of remembering (called metacognitive judgment nowadays). Tulving's (1985) remember/know procedure is the most popular form of metacognitive judgment. For instance, when we recently reviewed the published literature on this procedure, we found that it contained over 1000 data sets. Two other procedures that have been extensively used are Jacoby's (1991) process-dissociation methodology and Yonelinas (1994) dual-process ROC methodology. Together, these procedures have spawned vast behavioral and neuroscience literatures