

# A note on the identifiability of parallel and serial processes

J. T. TOWNSEND\*

Purdue University, Lafayette, Indiana 47907

Due to the significant research effort devoted to discovering whether certain psychological processes are serial or parallel, it seems important to establish the degree to which such processes are identifiable and to investigate possible ways in which such knowledge can improve our experiments. General definitions of parallel and serial systems are given, followed by a qualitative summary of identifiability results obtained with special classes of exponential systems. Some of these results are applied to a current experimental paradigm, and possible techniques are suggested to provide stronger serial-parallel tests and acquire more temporal processing information. Finally, the possibility of S's possessing the ability to manipulate his distribution of processing energy is reacknowledged.

Experiments designed to test whether possible hypothetical mechanisms act in a serial or parallel fashion or to which serial or parallel models have been applied are now numerous in the literature and cover several kinds of psychological tasks; for example, a few are: symbol detection or recognition tasks (Estes & Taylor, 1964), immediate memory comparisons (Sternberg, 1966), vigilance-scanning operations (Neisser, 1963), and same-different discrimination experiments (Egeth, 1966). It is, thus, of importance to understand to what extent parallel and serial models are testable in principle and how we can arrange our experiments in order to optimize our chances of identifying whether processing is parallel or serial (if either). This note summarizes an effort to implement this understanding by a formalization within which diverse experiments can be formulated and by consideration of a special class of stochastic systems (Townsend, 1969, 1970, in press).<sup>1</sup> An application of some of the results to a particular experimental paradigm is then discussed. Hopefully, these investigations may also be propaedeutic to further work on identifiability concerning this and other spatiotemporal properties of information-processing systems.

Problems of relating observations to structure and concomitant uniqueness questions are broadly referred to as problems of identifiability. There appear to be no very general theorems on identifiability, but results are known for some special cases. Most notable within experimental psychology is, perhaps, the work of Greeno and Steiner (1964) on identification of states within finite Markov processes. In the present paper, we will mean by identifiability

the prospects for determining the parallel or serial nature of underlying or unobservable processes by examination of input-output relationships and all possible statistics relating to such behavior in organisms. We will employ, when convenient, the transitive verb "mimic" to refer to a parallel (serial) process resembling a serial (parallel) process more usually or "naturally" associated with certain input-output relationships. These notions will be made more precise below.

By parallel processing, we mean that the entities (hereafter called elements) to be processed are worked on all at once, processing beginning on all simultaneously but possibly finishing at different times. We allow processing rates to depend on the past but not the future; thus, the rate of processing Element b may, when there are five elements present, speed up or slow down after two elements are processed, depending on whether Element a or Element c finishes first and even their order of finishing. But given that, say, Element a is finished first and c second, b's rate at this, the third, stage may not depend on whether Element d is completed fourth or fifth. By serial processing (or a serial system), we mean that the elements are processed one at a time, each being completed before another is begun. For the sake of definition, we assume the same order is taken on each trial for a single system but allow different processing rates for the different elements. In order to consider a broader class of systems, we must take account of the possibility of probability mixtures of single systems. In the case of serial systems, this means allowing for different processing orders occurring with some probability distribution on the set of possible permutations and for possible different processing rates associated with the various orders. Similarly, for parallel systems, this allows for the

operation of any of a finite number of parallel systems, each operating with some fixed probability, with complete freedom of processing rates subject to the above constraints for parallel operations. All systems other than the above will be called hybrid. Alternative definitions of "parallel" and "serial" are possible, of course. Those presented here seem to correspond to current usage in experimental literature and to include reasonable constraints on related system structure.

Space limitations preclude exact description of the experiments cited above. However, consideration of the following possible specifications of "element" together with reference to the studies will suffice to illustrate the formulation.

(1) In the Estes and Taylor (1964) paradigm, the elements are simply the letters displayed to S. Processing must at some stage include a comparison of each letter with the signal letter. With this as with other studies one may make an element correspond to a grosser or finer grained entity, depending on theoretical considerations.

(2) An element in an immediate memory comparison (Sternberg, 1964) would be a symbol in the memory set, processing being comparison with a probe stimulus.

(3) Neisser's (1963) type of scanning experiment would include "element" as one of the letters in memory and/or one of the letters in the list to be scanned. The parallel-serial question may be posed with respect to both the letters in memory and those on the test page.

(4) The same-different discrimination tasks (Egeth, 1966) can use "element" to mean a single dimension, processing referring to a comparison of the values on a single dimension in the two presented stimuli.

The system for which the following results were obtained specified that each element be processed with exponential waiting time, the rates or intensity parameters being subject to the foregoing constraints. This means that in a serial system the element completion times are a series (sum) of independent, exponentially distributed durations, and in a parallel system elements are processed independently and exponentially during a single stage, i.e., from the point in time when k are finished to when k+1 are finished. Statistics such as correlation of, say, the k<sup>th</sup> with the (k+1)<sup>th</sup> intercompletion times or of the overall processing times for Element a and Element b will depend on the set of values of processing rates and probabilities of different mixtures of

\*The reading and helpful comments by James G. Greeno are appreciated.

systems. For the sake of economy of discussion, we will limit ourselves to homogeneous systems, that is, to systems operating such that processing rates do not depend on time. It will, however, be apparent that many of the present remarks will be pertinent to extensions to nonhomogeneous exponential systems.

We can now give a more precise definition of identifiability as used to obtain the present results. When two classes of serial and parallel systems are specified, if given a member system of one class, there exist no members of the other which predict the same joint probability distribution of element completion times, that system is said to be identifiable (that is, it cannot be mimicked within the specified class of the other type); otherwise it is nonidentifiable. "Identifiability" is typically defined on a single class of systems (Koopmans, 1950; Greeno & Steiner, 1964), the question being whether or not members of the class can be distinguished from one another. Hence, the present use is an extension to questions of uniqueness between classes.

It is a corollary of these remarks that the following results hold for any observable input-output relationships and statistics; we assume only that one is unable to observe directly the processing and thereby ascertain its spatio-temporal nature. Thus, when identification is possible below, it may depend on some particular observations (for example, minimum completion time or joint intercompletion time distributions), but when nonidentifiability holds, it holds a fortiori for all conceivable experiments.

The results obtained with these systems consisted of three main findings. The first was that within the broadest classes of the above-defined exponential systems, and assuming the same number of elements for each system, there exist mixed parallel systems which unmixed serial systems are incapable of mimicking and mixed serial systems which unmixed parallel systems are incapable of mimicking. For example, if a, b of a two-element system are processed at different rates and neither is processed first with probability one, then there is no way a two-element unmixed parallel system can mimic the associated distribution, in particular the waiting time to first completion. This is because the first completion time for the parallel system is a single exponential, but that for the serial is a probability mixture of two exponentials. Furthermore, it will in general take a mixed system of  $n$  parallel systems (each operating on  $n$  elements) to mimic completely a mixed-order serial system operating on  $n$  elements.

The second finding was that if the serial system were constrained to a single processing rate independent of order of processing and specific elements, then there existed an unmixed parallel system that was identical, irrespective of the serial distribution so long as no permutation possessed zero probability of occurrence. This means that the only possibility of escaping nonidentifiability of such systems lies in encouraging the use of some processing orders over others, the best experimental strategy being for  $S$  always to process elements in a single order. Intuitively, the parallel system is unable to predict that Element  $a$ , say, always finishes first; this result clearly extends to all parallel-system distributions that possess appropriately overlapping nonzero probability density or mass functions. However, it is true that although mathematical equivalence cannot exist when this happens, the exponential parallel system can arbitrarily approximate closely the serial system and yet fulfill our definition of parallel.

The third finding was that for unmixed independent parallel processing (that is, when each element may have its own, possibly unique, processing rate that does not change as a function of when other elements complete processing), then there exist mixed serial processes which completely mimic these. For instance, if  $a, b$  are processed at different but constant exponential rates in a parallel system, then there is a mixed serial system which is equivalent. This mimicking system has the feature of processing a first with a probability reflecting the relative rate of processing  $a$  in the parallel system. In addition, the rate of processing the second element in the serial system is smaller than that of the first.

Now consider a task that requires of  $S$  that he memorize a small (say, one to six) set of letters on each trial and then respond as to whether or not a letter presented a short time later (the probe stimulus) was or was not contained in the memorized set. Conditions are such that performance is almost perfect and the main dependent variable is reaction time, plotted as a function of number of letters in the memorized set. This paradigm has been employed by Sternberg (1966) and others to investigate the question of serial vs parallel nature of the comparison of probe stimulus to letters in the memory set. We will confine our discussion to the serial-parallel issue, although other characteristics, such as self-terminating vs exhaustive processing, are also of interest to the investigator.

The typical finding has been an essentially linear function of mean latency vs number of letters in the memory set. Although writers are often chary of making strong claims based on this kind of result, it has nevertheless generally been interpreted as convincing evidence for an underlying serial process. However, the second finding of the analyses limned in above (Townsend, 1969, in press) shows that there is no mathematically compelling reason for such interpretation. There is a parallel system predicting this and all other results associated with a serial exponential system with constant processing rate. [Atkinson, Holmgren, & Juola (1969) noted a special case of this result.] Furthermore, in this case the conclusion of seriality is not more compelling intuitively than that of parallelism. Such a parallel system is characterized by the property of constant "energy," in the sense that the overall processing rate or mean intercompletion time is constant. To the  $S$  this could mean a spread of attention (not necessarily uniform) over the available elements; when one is completed, he refocuses the excess energy or attention on the remaining elements, in this case memorized letters. Thus, there is a fairly intuitive rationale for the mimicking parallel system.

Moreover, the only possibility for distinguishing between any exponential serial system and any unmixed parallel exponential system is to (1) use elements that will have different processing rates and more than one processing order, or (2) use materials and instructions that will encourage some processing orders over others, especially a single order. Even these strategies depend for success on being able to obtain additional data.<sup>2</sup> For instance, if one could ensure (1)—perhaps by using elements, perhaps words, of unequal difficulty—and request  $S$  to respond yes or no with respect to whether the first word completed was the probe or not, then the accompanying latency would reflect the minimum completion time for a given trial. This distribution is different from that for the same statistic of an unmixed parallel process; for example, the coefficient of variation will be one for the parallel system but will deviate from one for the serial, the magnitude of the difference depending on how different the rates are for the various elements when they are processed first.<sup>3</sup> Another strategy would be that given by (2) and then to gather information pertaining to individual elements' latencies. Good fits of a serial model with a single order of processing would mean that the system is at least capable of acting in a serial-like

manner. The success of these approaches, of course, depends on the feasibility of "tapping" the system in such a manner as to obtain the required information.

It seems likely that within the tolerances currently accompanying psychological experimentation and modeling that some of the results now seen as supporting parallel or serial processes may be just as adequately modeled by the other type of system.<sup>4</sup> Where, then, do these and other such results leave us with respect to testing between serial and parallel processes? The answer to this question will depend on the amount and kind of assumptions about the processes, the extent to which these assumptions constrain the class of possible processes, and the extent to which properties associated with the assumptions are observable. Also, it is important whether we ask about complete identifiability, as defined in this paper referring to equivalence of distributions, or whether we are also concerned with possible approximating systems. An example with few assumptions and few observable properties of the associated distribution is a parallel process with no other constraints, where the only observable datum is the time required to process all of the (unknown) number of elements. Then there will exist a probability mixture of gamma distributed serial processes, possibly with unequal numbers of steps or stages, that will approximate to any desired degree of accuracy the distribution for the parallel system. This conclusion follows from a theorem stating that any positive random variable can be approximated by the general gamma distribution (Kingman, 1966). Obviously, if we had good reason to suspect a particular distribution and a specific number of elements for a given parallel system, then this might rule out the approximating serial processes.

As another example, assume one has good reason to believe that the system is either serial, exponential or parallel, exponential each with constant rate parameter (across elements and processing position). Then the systems could be distinguished simply because they completely specify different stochastic processes which differ in a radical manner. In other cases, especially, perhaps, when the model tests are qualitative, support may go to the intuitively reasonable system, even though there may exist systems of the

opposite type which are equivalent. Such may be the case, for example, for Neisser's results (1963) where he found scanning time to be constant across number of target letters. His interpretation, of course, was that this supported the notion of a parallel process.<sup>5</sup> Although we can find serial processes that predict this result, such processes have the property of a rate of processing that increases with number of elements, a system that seems unlikely for many situations. It should be noted that, in general, only nonstochastic systems predict a constant maximum completion time.

Similarly, with respect to mathematical equivalence, one can construct parallel systems whose component processing distributions are gamma and which are mimicked by hybrid time-sharing systems or serial systems having the property that the number of stages for each element varies probabilistically from trial to trial, and the number of stages is negatively correlated among elements for a given trial. Thus, the parallel system seems more parsimonious and in a given experiment may be the more compelling explanation.

Finally, it should not be forgotten that for some psychological systems, S may be able to manipulate the manner in which he distributes processing energy over the elements to be processed, if not within a single trial, perhaps from trial to trial. Within the class of exponential models mentioned previously, this means the possibility of altering the processing rates anywhere from uniformity across elements to letting all be zero but one at any one "moment," thus proceeding from a parallel system to a serial system. We might also predict a lessening of accuracy on each element for the first extreme, in line with one of Sir William Hamilton's (1859) laws of intelligence: "This law is, that the greater the number of objects to which our consciousness is simultaneously extended, the smaller is the intensity with which it is able to consider each."

#### REFERENCES

- ATKINSON, R. C., HOLMGREN, J. E., & JUOLA, J. F. Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, 1969, 6, 321-326.
- EGETH, H. Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, 1966, 1, 245-252.
- ESTES, W., & TAYLOR, A. A detection method and probabilistic model for assessing information processing from brief visual displays. *Proceedings of the*

- National Academy of Science, 1964, 52, 446-454.
- GREENO, J. G., & STEINER, T. E. Markovian processes with identifiable states: General considerations and application to all-or-none learning. *Psychometrika*, 1964, 29, 309-333.
- HAMILTON, W. *Lectures on metaphysics and logic*. Vol. 1. Edinburgh: Blackwood, 1859.
- KINGMAN, J. F. C. *On the algebra of queues*. London: Methuen, 1966.
- KOOPMANS, T. C. (Ed.) *Statistical inference in dynamic economic models*. New York: Wiley, 1950. (Cowles Commission Monograph No. 10)
- NEISSER, U. Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology*, 1963, 76, 376-385.
- STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- TOWNSEND, J. T. Stochastic representations of parallel and serial processes. Paper presented at Mathematical Psychologists Conference, Indiana University, 1969.
- TOWNSEND, J. T. A result concerning the testability of parallel and serial processes. Paper presented at Mathematical Psychologists Conference, Indiana University, 1970.
- TOWNSEND, J. T. Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical & Statistical Psychology*, in press.

(Accepted for publication February 5, 1971.)

#### NOTES

1. Copies of these papers are available on request and the reader is referred to them for mathematical development, discussion of alternative definitions of "parallel" and "serial," and comments relating to generalization of present results.
2. It should be noted here that mean values of reaction time for different elements or positions in a display, for example, are not sufficient to distinguish even a parallel independent system from a single-path serial system with a constant processing rate.
3. This specific prediction concerning coefficient of variation is, of course, particular to exponential systems. In any case, the minimum completion time distribution for the unmixed parallel system will be different from that for a mixed serial system with different processing rates for the various elements.
4. New developments indicate that parallel-serial identifiability difficulties may not be as severe for certain probability distributions of completion time other than the exponential, at least with respect to complete equivalence. However, this will be of little aid to us until we are better able (and willing) to employ more subtle aspects of our data than simply means. Also, we need to extend the types of results presented here to broad and simple aspects of data, perhaps means and variances, that may be applicable to larger families of completion time distributions (for example, the general gamma distribution).
5. The validity of this conclusion may naturally depend on the extent to which accuracy remains constant across number of target letters.