G. DE LEVE, *Generalized Markovian Decision Processes*—Parts I and II. Mathematisch Centrum, Amsterdam, 1964. 116 pp. and 117 pp.

Review by CYRUS DERMAN

Columbia University

The notion of a Markovian Decision Process is a natural outgrowth of ideas set in motion by the developments in sequential analysis, inventory theory, dynamic programming, and Markov processes. Typically, a Markovian Decision Process is a mathematical model of a dynamic system observed continuously or periodically and controlled at the time of observation by the making of one of a possible number of decisions. The state of the system and the decision interact to determine the evolution of the system. More specifically, a state and decision invoke certain transition probabilities that govern the system's fluctuations. Different decisions invoke different transition probabilities. The transition probabilities are functions only of the present state and the decision made; i.e., independent of the history of the system. It is assumed that costs associated with being in a state and making a decision are known and are constant throughout the evolution of the process. An over-all cost criterion such as average cost per unit time or expected discounted cost is usually of interest. The problem to be solved is that of determining an optimal policy for governing the system where a policy is a complete prescription over time for making decisions. Some recent computational breakthroughs for obtaining optimal policies when the number of possible states is finite, implying a certain practicality to the model, have stimulated research in this area.

These volumes treat the case where the state space is an N-dimensional Cartesian space, the observations on the systems are made continuously, and the over-all cost criterion is the average cost per unit time. It is shown (unsuccessfully, in our opinion) that an optimal policy, when it exists, is in the form of a mapping from the state space to the decision set. Under policies of this form representations for the average cost per unit time are obtained. Optimal policies and computational algorithms for obtaining optimal policies are discussed. Fundamental questions such as whether a dynamic system governed by a policy of the above form is a strong Markov process and whether the number of decisions made in a finite interval of time is finite are also considered.

To develop the subject in such generality as attempted here is a formidable task. Even to define the processes involved as legitimate mathematical entities requires careful treatment. Because of the introduction of continuous time parameter the intricacies of strong Markov processes arise, and since the state space is nondenumerable, a familiarity with the Doeblin treatment of Markov processes is essential for obtaining the right conditions.

The author is to be commended for his courage to undertake such a project.

However, it is our impression that the combination of the above difficulties with a cumbersome notation, inadequate organization, and, possibly, over-complicated formulation make for two volumes which are extremely difficult to read. We regret that more time and editorial work did not go into the preparation of this work (and a third volume to follow) in order that these results might be more accessible to those working in this area.