Reset reproduction of article published in *Computational Intelligence*, Vol. 7, No. 1 (February 1991), pp. 1–10. Reprinted July 1994.

Markov Analysis of Qualitative Dynamics*

Jon Doyle
Laboratory for Computer Science, Massachusetts Institute of Technology
545 Technology Square, Cambridge, Massachusetts 02139

Elisha P. Sacks
Department of Computer Science, Princeton University, Princeton, New Jersey 08544

Abstract

Commonsense sometimes predicts events to be likely or unlikely rather than merely possible. We extend methods of qualitative reasoning to predict the relative likelihoods of possible qualitative behaviors by viewing the dynamics of a system as a Markov chain over its transition graph. This involves adding qualitative or quantitative estimates of transition probabilities to each of the transitions and applying the standard theory of Markov chains to distinguish persistent states from transient states and to calculate recurrence times, settling times, and probabilities for ending up in each state. Much of the analysis depends solely on qualitative estimates of transition probabilities, which follow directly from theoretical considerations and which lead to qualitative predictions about entire classes of systems. Quantitative estimates for specific systems are derived empirically and lead to qualitative and quantitative conclusions, most of which are insensitive to small perturbations in the estimated transition probabilities. The algorithms are straightforward and efficient.

1 Introduction

Qualitative dynamical reasoning seeks to predict the global behavior of a complex dynamic system by partitioning its state space into a manageable number of regions and characterizing its behavior by the sequences of regions that it can go through. Although considerable progress has been made toward automating such reasoning, some important prediction problems have not been addressed. In particular, this methodology is too weak to describe the limiting behavior of dynamic systems. For example, a damped pendulum eventually must approach equilibrium either directly below or directly above its pivot (Fig. 1). The first possibility is almost certain, whereas the second almost never occurs. Qualitative simulation discovers both equilibria, but neither can determine their relative likelihoods nor rule out the possibility that the pendulum will spin forever. Yet qualitative considerations suffice for both conclusions, independent of the numeric details of the system.

Such limiting behaviors are global characteristics of a system. To understand them, we must look beyond individual transitions to sequences of transitions. We must assign each sequence a probability ranging from impossible to definite. The probability of a particular limiting behavior equals the total probability of the subset of possible histories in which the corresponding sequence of transitions occurs. For example, the probability that the pendulum approaches its unstable equilibrium when released in arbitrary position is zero because the set of sequences that either start or terminate at the unstable equilibrium has measure

*This paper extends and improves on the earlier paper [3]. Authors listed alphabetically.

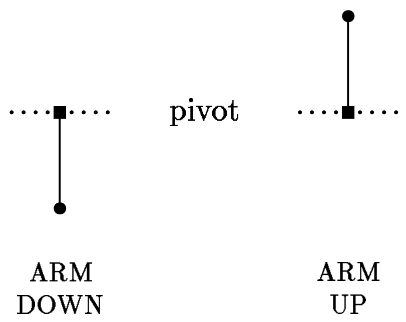

Figure 1: Equilibria of a damped pendulum.

zero. Calculating the probabilities is straightforward in systems whose exact limiting behavior is known for all initial conditions. The challenge is to estimate the probabilities when the limiting behavior is unknown. This can occur in qualitative reasoning where the underlying equations are incompletely specified or in quantitative reasoning about intractable systems.

In this paper, we describe a method for predicting the relative likelihoods of the limiting behaviors of such dynamic systems. The method provides a formal justification for commonsense conclusions about relative likelihoods and an efficient algorithm for deriving them. It also provides numeric likelihood estimates for fully specified systems. The method rests upon the simplifying assumption that the system forms a *Markov chain* over its transition graph, i.e. that the next state of the system is a time-independent probabilistic function of its current state. This assumption agrees with the standard qualitative reasoning model in which the next qualitative state of a system depends only on its current state, not on its past. It extends that model by assigning probabilities to branching states instead of treating all branches uniformly. Markov theory is then employed to derive the probabilities of the transition sequences.

The central step of the analysis, in which we assign transition probabilities and derive probability estimates for the possible asymptotic behaviors, applies to every extant form of qualitative dynamics, including ones generated by qualitative simulation. For concreteness, we illustrate the method by using the classical mathematical theory of dynamic systems to derive a set of qualitative states and transition graph from the phase space of a system, following Sacks [14].

Our method can derive useful results at many levels of detail, ranging from the abstract level of the qualitative reasoning formalisms in the AI literature to fully specified ordinary differential equations. It can process qualitative probability estimates in the $\{0, (0, 1), 1\}$ quantity space, symbolic estimates such as p or $q+r$, and numeric estimates. Markov theory blends the available information into a unifying framework that provides the best possible conclusions about asymptotic behavior given the evidence. Qualitative information leads to qualitative predictions about entire classes of systems, such as all damped pendulums or all instances of a parameterized equation. Quantitative information leads to qualitative and quantitative predictions about individual systems.

Markov theory provides some sorts of essentially qualitative information that qualitative simulation does not, including a partition into persistent and transient states (transient states are always improbable as asymptotic behaviors) and a partition of the persistent states into the probable and the improbable. Many of these facts follow directly from qualitative estimates of transition probabilities, and may be derived through purely qualitative algorithms. Other qualitative conclusions, though derived from numeric estimates of transition probabilities, are insensitive to small perturbations in these estimates. The theory also provides quantitative refinements of these qualitative conclusions, including the mean and variance of settling times. Unlike the qualitative conclusions, the quantitative results are in some cases sensitive to variations in the input probabilities. The algorithms are straightforward, principally consisting of a topological sort of the transition graph and a few matrix operations on the transition probabilities, and require time at most cubic in the number of regions. The numeric analysis goes through for symbolic probability estimates, although at the price of exponential-time worst-case performance.

The next section describes our approach to dynamics, which is based on the classical mathematical theory of dynamical systems, and shows how the sorts of dynamics employed in other AI approaches to qualitative reasoning may be translated into ours. The following section shows how to model dynamic systems as Markov chains. There we state and defend the requisite simplifying assumptions. Section 4 describes the algorithms for analyzing Markov chains. Section 5 demonstrates our methods on several examples, including the damped pendulum and the quadratic map. The final section draws conclusions and discusses some possible extensions and generalizations.

2 Qualitative dynamics in phase space

Our qualitative dynamics builds upon the phase space representation developed by Poincaré. The phase space for a system of first-order differential equations

$$x_i' = f_i(x_1, \ldots, x_n);\ i = 1, \ldots, n \tag{1}$$

is the Cartesian product of the x_i's domains. One can convert higher-order equations to first-order ones by introducing new variables as synonyms for higher derivatives. Points in phase space represent states of the system. Curves on which the equations (1) are satisfied, called *trajectories,* represent solutions. The topological and geometric properties of trajectories characterize the qualitative behavior of solutions. For instance, a point trajectory, called a *fixed point,* indicates an equilibrium solution, whereas a closed curve indicates a periodic solution. A fixed point is *stable* if every nearby trajectory approaches it asymptotically and *unstable* otherwise. More generally, the *basin* of a fixed point is the set of trajectories that approach it asymptotically.

A *phase diagram* for a system depicts its phase space and trajectories graphically. For example, the standard model for a damped pendulum is

$$\theta'' + \frac{\mu}{m}\theta' + \frac{g}{l}\sin\theta = 0,$$

with θ the angle between the arm and the vertical, l the length of the (weightless rigid) arm, m the mass of the bob, g the gravitational constant, and μ the damping coefficient (Fig. 2). The phase diagram of the pendulum appears in Figure 3. The phase space is

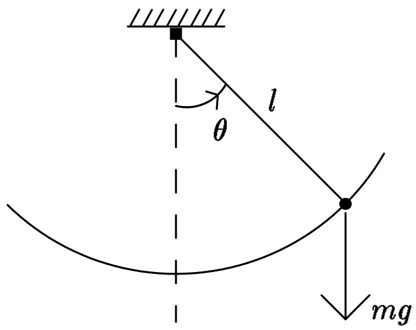

Figure 2: A damped pendulum.

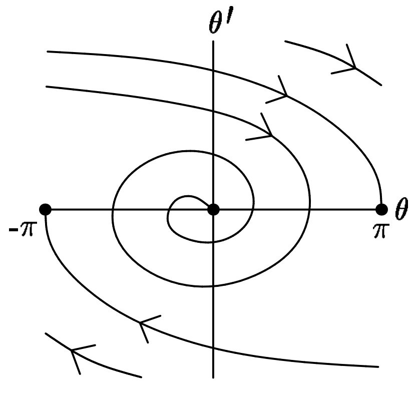

Figure 3: Phase diagram for the damped pendulum.

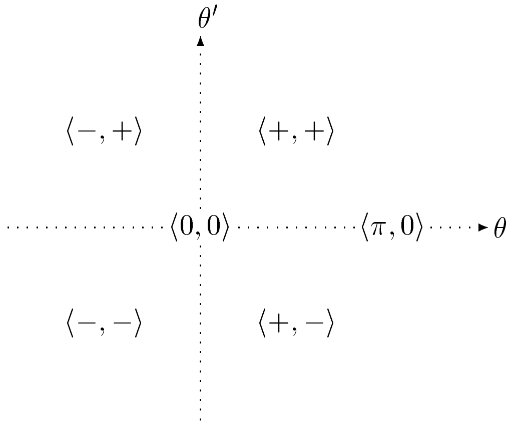

Figure 4: Phase space regions of the qualitative states of the damped pendulum.

cylindrical, since angles that differ by 2π are physically indistinguishable. Two trajectories spiral toward the unstable fixed point at $(\pi, 0)$; the rest spiral toward the stable fixed point at $(0, 0)$.

A complete qualitative description of a system consists of a partition of its phase space into sets of qualitatively equivalent trajectories. The equivalence criterion depends on the problem task. Mathematicians generally focus on topological equivalence, whereas coarser relations are more useful in engineering applications. We follow standard AI practice and equate all trajectories that go through a specific sequence of regions in phase space. Our qualitative dynamics consists of a partition of phase space into regions along with a graph of possible transitions between regions. Sacks [12, 13] presents a system that automatically identifies such regions and the possible transitions between them for second-order systems of ordinary differential equations. Most of the ideas extend directly to larger systems.

Sacks [14] shows how to translate traditional qualitative reasoning into our qualitative dynamics without loss of information or increase in complexity. Qualitative states correspond to rectangular regions in phase space, and qualitative simulation amounts to finding the possible transitions between regions. A transition occurs from region A to region B if the derivative of the system on the boundary between the regions points into B. A transition occurs from a region to a neighboring fixed point unless all the eigenvalues of the fixed point have positive real part, as explained in the next section. For example, automatic analysis of the damped pendulum equation results in six qualitative states corresponding to four rectangles ($\langle \pm, \pm \rangle$) and two fixed points ($\langle 0, 0 \rangle$,$\langle \pi, 0 \rangle$), as shown in Figure 4. (The complete dynamics includes the degenerate rectangles $\langle 0, \pm \rangle$ and $\langle \pm, 0 \rangle$, which we ignore for simplicity. Including them complicates the results without affecting the analysis or the conclusions.) Figure 5 shows the transition graph over these qualitative states.

3 Transforming flows into Markov chains

A Markov chain is a process whose state at time $t + 1$ is a time-independent probabilistic function of its state at time t. We model dynamic systems as Markov chains whose states are regions and fixed points in phase space, or equivalently qualitative states. The transition

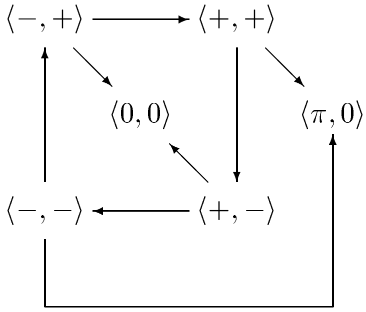

Figure 5: Qualitative state transition graph of the damped pendulum.

probabilities express the likelihood of the system's state moving from one region or fixed point to another in unit time. We employ Markov theory to infer properties of the trajectories from properties of the transition probabilities. Each inference applies to all systems with the necessary properties. Inferences that require only qualitative properties apply to an entire class of systems, whereas inferences that rest on numeric probability estimates apply only to "nearby" systems, as explained in the next section.

Determination of suitable qualitative states and time scales for the Markov analysis is a difficult problem. We discuss some of the issues involved below, but do not present any new methods for automatic model construction, as the probabilistic analysis techniques studied here may be applied in conjunction with any model construction method. Extant approaches to automatic model construction do not always yield the best models, but in many cases the errors introduced by suboptimal models can be treated exactly as ordinary stochastic behavior. The strength of the Markov analysis is that it provides the best predictions possible given the model used.

Transition probabilities have one meaning for transitions between regions and another meaning for transitions involving fixed points. The transition probability from region A to region B expresses the fraction of points in A whose corresponding trajectories are in B after one time unit. For nondegenerate A of finite measure, this is just

$$\frac{\mu(A \cap \phi^{-1}(B))}{\mu(A)} \tag{2}$$

where μ is Lebesgue measure and ϕ is the flow over unit time. Regions which have infinite or zero measure, such as $\langle +, + \rangle$ and $\langle 0, - \rangle$, call for different treatment. One approach to regions of infinite measure, which suffices for any physically realizable system, is to replace each infinite region in equation (2) with a large, but finite, subregion. This amounts to defining the transition probability as a limit of equation (2) over a monotone increasing sequence of finite subsets. One could similarly replace degenerate regions with small nondegenerate regions surrounding them, which amounts to defining the transition probabilities over monotone decreasing supersets.

Our treatment of degenerate regions extends directly to fixed points. Intuitively, the transition probability from region A to fixed point p expresses the fraction of points in A whose trajectories "reach" p in unit time. To justify this approach, we must resolve the conflict between the theory of ordinary differential equations, which implies that smooth systems cannot reach a fixed point from another state in finite time, and everyday experience, which indicates the contrary. One can side with the theory and claim that systems merely *appear* to reach a fixed point because of our perceptual limitations, or one can side with commonsense and claim that ordinary differential equations are an imperfect model of reality. Theorists can interpret the word "reach" as entering some ϵ-neighborhood; others can take it literally. Indeed, since stable fixed points attract all points in some neighborhood, estimates of asymptotic probability are not misguided by identifying transition probabilities to a neighboring region with transition probabilities to the stable fixed point.

The measure-theoretic construction interprets qualitative states as lumped states, so that the transition probabilities represent the imprecision in the qualitative model of the dynamic system. If we were able to choose as regions the actual attractors and basins of the system, there would be no imprecision and the transition rules would be perfectly deterministic. The great difficulty of determining the optimal set of regions for analysis helps motivate the stochastic approach to analyzing behaviors. Additional factors that transition probabilities can model include (1) uncertainty about initial conditions which induces a distribution of possible trajectories, (2) uncertainty about the parameters of the model equations, and (3) uncertainties (or noise sources) explicitly occurring in the system's equations, as in stochastic differential equations.

Qualitative probability estimates follow directly from the transition graph and fixed point types of a system. The transition probability is zero from a region to an unreachable region by definition of the graph. By the stable manifold theorem [5, p. 13], the dimension of the basin of a fixed point equals the number of eigenvalues of its Jacobian matrix that have negative real parts.[1] Unstable fixed points have positive eigenvalues, so their basins form lower-dimensional, hence measure zero, subspaces. This implies that the transition probability into an unstable fixed point from any other region is always zero by equation (2). Stable fixed points have basins of positive measure because they attract all points in some neighborhood. Hence, there is a positive transition probability from a neighboring region to a stable fixed point.

Numeric estimates of transition probabilities are derivable by numeric simulations or physical experiments that sample representative points in each region and count how many go to each region. One can also obtain subjective estimates from domain experts. Both sampling and subjective estimates may contain errors, but as we show below, the qualitative analysis is insensitive to minor errors in these probabilities.

In the pendulum example, the fixed point $(0,0)$ is stable and $(\pi,0)$ is unstable. The basin of $(\pi,0)$ is a one-dimensional curve in the two-dimensional phase space (Fig. 3) whereas the basin of $(0,0)$ has positive measure. The transitions to $(\pi,0)$ have probability zero and those to $(0,0)$ have positive probability. These qualitative estimates suffice to prove that the pendulum comes to rest at $(0,0)$ with probability one. The exact arrival time depends on the specific transition probabilities, which vary from system to system. We show how to derive both the qualitative and the quantitative information in the next section.

The qualitative analysis of the pendulum applies to any system with the same transition graph and types of fixed points. It is independent of whether the pendulum is underdamped or overdamped, that is whether the eigenvalues of $(0,0)$ are real or complex. In the real case,

[1] This theorem applies to hyperbolic fixed points. The number of positive real parts at an arbitrary fixed point suffices for our analysis.

trajectories eventually approach $(0,0)$ directly from within $\langle -,+\rangle$ or $\langle +,-\rangle$, while in the complex case, they spiral inward, forever cycling between the outer regions. This distinction is immaterial according to both the theoretical and the commonsense views described above. Commonsense asserts that real trajectories always come to rest at $(0,0)$ after a finite number of cycles. Theory claims that they enter and remain in a small neighborhood of $(0,0)$; for the conclusions reached by the transition model, it is irrelevant whether trajectories continue to spiral within that neighborhood.

We assume that the transition probabilities remain constant over time and that they depend only on the qualitative state of origin, independent of past history. The first assumption holds for *autonomous* equations that are free of their independent variable. One can reduce any general system to an autonomous one by treating the independent variable, t, as a state variable governed by the equation $t' = 1$. The second assumption holds to the extent that the future trajectory of the system is insensitive to its distant past. The most questionable case is that of conservative systems in which the volume of each region in phase space is preserved for all time by the flow, causing small differences between trajectories to retain their significance forever. Conservative systems pose problems for qualitative reasoning generally, not just for our stochastic analysis, as the regions of interest must be chosen carefully. Fortunately, most realistic systems are dissipative, hence volume shrinking, causing differences between trajectories to shrink exponentially. Figure 6 illustrates the two cases.

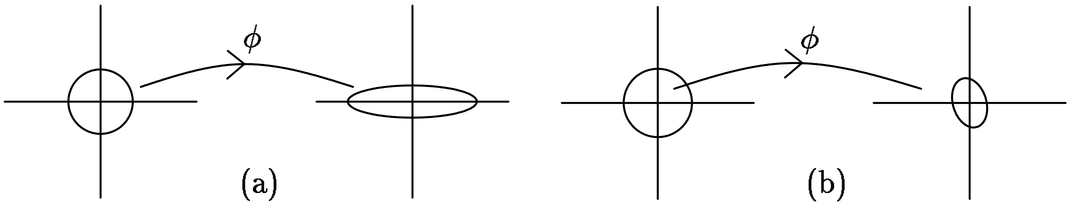

Figure 6: (a) Volume preserving flow. (b) Dissipative flow.

Time-dependent transition probabilities imply that the current partition of phase space is too coarse: differences within a prior region express themselves in the current region because the distances between points in the prior region are too great to damp out in one time step. One approach to dealing with such time-dependence involves iterative improvement of the model, following Hsu [7] or Sacks [13] (though unlike that work, we have not automated this refinement process). If one observes time-dependent behavior in constructing the transition probabilities, one subdivides or otherwise refines the set of regions and starts over. In principle, the process ends when the chain assumption appears correct for all regions, but in practice the choice of when to accept a model as satisfactory involves a tradeoff of model complexity against model accuracy. The aptness of the chain assumption can also be tested against empirical observations or long term numeric simulations.

Like the choice of regions, the choice of the time unit over which transition probabilities are measured influences the accuracy of the model. Too long a time scale, and all transitions may appear possible; too short, and all regions may appear to be fixed points. More to the point, the choice of a useful time unit depends on the choice of regions. Indeed, for some sets of regions, there may be no time unit useful for all regions of phase space. The problem here is analogous to the problem of choosing good grid points for numeric integration or for spline approximation to functions. We offer no new methods for choosing sampling intervals.

4 Analysis of Markov chains

In this section we first summarize the elements of the theory of Markov chains and then describe how to organize the analytic algorithms to yield qualitative and quantitative conclusions. Readers familiar with Markov theory may skip to Section 4.5. Readers unfamiliar with Markov theory may find more details in Feller [4], Kemeny and Snell [8], or Roberts [11]. For simplicity, we will treat only finite Markov chains, and so restrict attention to systems whose qualitative dynamics involves only finitely many regions of interest.

Let $S = \{s_1, \ldots, s_n\}$ be the set of states of the qualitative dynamics, that is, the set of nodes of the dynamic digraph. Each of these will also be a *state* of the Markov chain. We describe the entire chain by specifying, for each nonexclusive choice of states s_i and s_j, the transition probability p_{ij} that if the system is in state s_i at one instant, it will be in state s_j after one time unit has passed. We write P to mean the $n \times n$ *transition matrix*

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

of all transition probabilities. P is also called a *stochastic matrix,* which means that all entries are nonnegative and that each row sums to 1. Each row of P is called a *probability vector*. The *transition digraph* of a stochastic matrix is the graph over the states with a directed arc from s_i to s_j iff $p_{ij} \neq 0$.

The probability that the chain is in state s_j at time t given that it starts in state s_i at time 0 is written $p_{ij}^{(t)}$ and called a *higher-order transition probability.* This probability is the i, j entry of P^t, the t'th power of P. If we start the Markov chain at random, where the probabilities of starting in each state are given by an *initial probability vector* $p^{(0)} = (p_1^{(0)}, \ldots, p_n^{(0)})$, then the probabilities of being in particular states at time t, $p^{(t)} = (p_1^{(t)}, \ldots, p_n^{(t)})$, are given by the equation $p^{(t)} = p^{(0)} P^t$.

A set C of states is *closed* if $p_{ij} = 0$ for all $s_i \in C$ and $s_j \notin C$, that is, if once in C the chain can never leave C. A closed set C is *ergodic* if no proper subset is closed. A state is ergodic if it is in some ergodic set, and is *transient* otherwise. A state that forms an ergodic set by itself is called an *absorbing* state. Chains whose states form a single ergodic set are called ergodic chains, and chains in which each ergodic set is a singleton are called absorbing chains.

The mathematical analysis of the asymptotic behavior of a Markov chain is divided into two parts: the behavior before entering an ergodic set, and the behavior after entering one. One then combines these sub-analyses to get the overall asymptotic behavior. In the first step, one creates an absorbing chain by lumping all states in each ergodic set into a single compound state. The transition probability from a transient state s to a compound state c is the sum of the transition probabilities from s to the members of c. The transition probability from c to other states is 0 by definition. The main result of the analysis is the long-term probability of entering each ergodic set when starting in each of the transient states. In the second step, one analyzes each ergodic set as a separate ergodic chain, unaffected by the other states. The result of the analysis is the long-term probability of being in each of the states of the set. Combining these results yields the long-term probability of being in each of the states of the chain. This is just the product of the probability of entering an ergodic set containing that state (this is zero if the state is transient) times the probability of appearing in that state once in the ergodic set. Stability analysis shows that these asymptotic probabilities are insensitive to variations in the input probabilities.

The separation of a Markov chain into ergodic sets and transient states is accomplished by topologically sorting the strongly connected components in the transition graph. Strongly connected components with no outgoing arcs, the minimal components in the sorted graph, comprise ergodic sets. All other strongly connected components consist of transient states. For example, the transition graph for the damped pendulum (Fig. 2) has three strongly connected components: $\{\langle 0,0\rangle\}$, $\{\langle \pi,0\rangle\}$, and the remaining nodes. The first two are ergodic; the third is transient.

4.1 Analysis of absorbing chains

The first step in analyzing an absorbing chain is to reorder its states so that the absorbing states appear first, thus converting the transition matrix to the canonical form

$$P = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix} \tag{3}$$

with I an identity matrix, R the matrix of transition probabilities from transient states to absorbing states, and Q the matrix of transition probabilities from transient states to transient states. The matrix $I-Q$ has an inverse,[2] denoted by N. The quantities of interest may all be computed from N and R. The probability that the system eventually enters an absorbing state is 1. The expected number of steps τ_i from the ith transient state until absorption equals the sum of the entries in the ith row of N. The variance σ_i of this mean is given by the ith entry in $(2N-I)\tau - \tau^{\langle 2\rangle}$, where $\tau_i^{\langle 2\rangle} = (\tau_i)^2$. The probability of reaching the jth absorbing state from the ith transient state is b_{ij}, where $B = \{b_{ij}\}$ is given by $B = NR$.

4.2 Analysis of ergodic chains

The analysis in this section applies to *regular* ergodic chains, that is, chains that contain at least one nonzero diagonal element in their transition matrix. Absorbing states are the trivial case of regular chains. All the examples in this paper are regular. One can expect qualitative analysis of continuous dynamics to yield regular chains. The analysis must construct at least one region of positive measure. One can always make the chain regular by choosing a time scale short enough that most points in this region cannot escape in one step, thus giving the region a positive probability of transition to itself. This usually happens

[2] This I is an identity matrix of the same dimensions as Q. It may differ in dimension from the identity matrix in equation (3).

in practice. Even when the natural time scale for modeling yields a non-regular chain, the analysis involves only a few additional calculations, not any new ideas or additional computational complexity. See Roberts [11] for details.

The powers P^t of the transition matrix of an ergodic chain approach a stochastic matrix W as t approaches infinity. The rows of W are identical, each consisting of the unique probability vector $w = (w_1, \ldots, w_n)$ satisfying the n equations

$$\begin{cases} \sum_{i=1}^{n} w_i p_{ij} = w_j; \quad j = 1, \ldots, n-1 \\ \sum_{i=1}^{n} w_i = 1. \end{cases}$$

This implies that as t approaches infinity the probability of being in state i at step t approaches w_i independently of the initial state. Correspondingly, the expected period of recurrence of state i is just $1/w_i$.

4.3 Stability of the analysis

By combining the two preceding analyses, we see that when each of the ergodic chains in P is regular, the powers P^t approach a limiting stochastic matrix P^∞ as $t \to \infty$, with p_{ij}^∞ representing the asymptotic probability of the system being in state j when starting in state i. If state j is transient, then $p_{ij}^\infty = 0$. Otherwise j is in some ergodic set J, and p_{ij}^∞ is the asymptotic probability of entering J (which we might write as b_{iJ}) times the asymptotic probability of being in j within J (which we might write as $w_j(J)$).

The classification of states into transient and ergodic states is stable under any variation in probabilities that does not change a positive probability into a zero probability or vice versa. Since the system ends up in ergodic states with probability 1, this means that the basic classification of asymptotic behaviors is very stable. In addition, the asymptotic probabilities P^∞ are stable with respect to smaller perturbations in the transition probabilities P. To see this, notice that the matrix P^∞ is a continuous function of P. Since the set of all $n \times n$ stochastic matrices is a compact, convex subspace of Euclidean space, the function P^∞ is absolutely continuous over this subspace, with a Lifschitz constant of nC, where C is the maximum absolute value of any of the partial derivatives of P^∞ with respect to entries in P. Hence for any $\epsilon > 0$, varying the entries in P by less than ϵ/nC will not cause changes larger than ϵ in the entries in P^∞.

Unlike the probabilities of asymptotic behaviors, the settling times τ and the relative likelihoods b_{iJ}/b_{iK} of absorption by different ergodic sets J and K can be sensitive to perturbations in P, especially when the direct probabilities of absorption are low (that is, when the norm of R is close to 0). The pendulum example of Section 5.1 illustrates these sensitivities in parametric form.

4.4 Initial conditions

Most treatments of qualitative prediction presuppose ignorance about the exact initial state of the system. Such ignorance is naturally viewed as assuming a uniform distribution over initial states, in which each state is equally likely to be the starting state. In this case, we get asymptotic probabilities for ending in each state by averaging the asymptotic probabilities of transitions into that state from all states, so that

$$p_i^\infty = \sum_{j=0}^{n} p_{ji}^\infty.$$

We can model knowledge about the initial states as a probability distribution $\Delta(s_i)$ over qualitative states. This distribution may be viewed as derived from a distribution δ over phase space, with $\Delta_i \equiv \Delta(s_i) = \int_{s_i} \delta$. Knowledge of initial conditions modifies the probabilities of asymptotic behaviors very simply, with $p^\infty = \Delta \cdot P^\infty$, that is

$$p_i^\infty = \sum_{j=0}^{n} \Delta(s_j) p_{ji}^\infty.$$

For example, a common case is when the initial state is known to occur in some subset S' of S, with the modeler ignorant about which state in S' it is. This might be modeled as a distribution uniform over S' and zero over $S - S'$. In this case, the asymptotic probabilities are given by equation (4), in which $k = |S'|$ and states have been relabeled so that all states in S' appear before states in $S - S'$.

$$p_i^\infty = k^{-1} \sum_{j=0}^{k} p_{ji}^\infty \tag{4}$$

4.5 Analytic procedure and algorithms

We divide the computation of predictions into qualitative and quantitative stages. The qualitative stage derives the basic judgments of asymptotic probability or improbability solely from information widely available in qualitative reasoning formalisms. We assign positive and zero transition probabilities from adjacent regions to attractors and other fixed points respectively, using the dimensional analysis explained in Section 3. We assign a zero transition probability to all pairs of states that have no connecting edge in the transition graph, such as $\langle -, - \rangle$ and $\langle +, + \rangle$ in Figure 5. We assign a positive probability to transitions between adjacent regions, such as $\langle -, - \rangle$ and $\langle -, + \rangle$. This decision is justified by the continuous dependence of trajectories on initial conditions and by the definition of the transition graph [14]. Using these probabilities, we partition the associated Markov chain into ergodic sets and transient states, drawing the qualitative conclusions that the former, but not the latter, persist asymptotically. We employ the topological sorting algorithm of Tarjan [1, Sec. 5.5] to find the ergodic sets. When nonuniform initial conditions are specified, we filter the sorted graph to find the ergodic states reachable from initial states with positive probabilities. The entire analysis takes time and space linear in the size of the transition graph.

The quantitative stage derives symbolic or numeric refinements of the qualitative judgements when corresponding estimates of the transition probabilities are available. It obtains the mean and variance of the settling times of the system from each transient state and the asymptotic distribution of ergodic states. All these quantities are given by simple matrix equations and can be computed in a constant number of matrix multiplications, inversions, and triangularizations. For numeric probability estimates, straightforward implementations

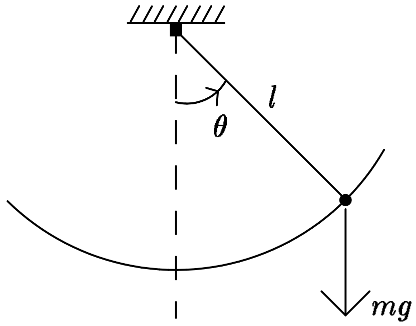

Figure 7: A damped pendulum.

of these operations require $O(n^3)$ time and $O(n^2)$ space for a graph with n nodes [1, Chap. 6]. Hsu [7] implements and demonstrates the calculations on several examples including chaotic ones.

The analysis may be carried out using probability estimates that are polynomials in indeterminate parameters, although at the price of exponential-time worst-case performance, since inverses of matrices with symbolic entries can have entries exponential in the size of the original matrix. The examples presented in the next section illustrate the use of parametric analysis in judging the sensitivity of relative asymptotic probabilities under variation in transition probabilities. A full sensitivity analysis using standard statistical techniques might constitute a third stage of the analytic procedure, but we have not automated it.

5 Examples

Each of the following three examples illustrates a different aspect of the stochastic analysis. The first example makes precise the analysis of the damped gravitational pendulum, and adds calculations of settling times and an analysis of their sensitivity to the earlier determination of asymptotic behavior. The second example, that of a charged pendulum in the presence of two other charges, is representative of a large class of everyday systems in which there are several asymptotic behaviors of nonzero probability. The analysis of the charged pendulum calculates these probabilities, and examines the dependence of their sizes on the magnitude of the charges. The third example, the quadratic map, illustrates the applicability of the stochastic analysis to discrete systems, including those exhibiting chaotic behaviors.

5.1 The gravitational pendulum

With the Markov theory in hand, we can make precise our intuitive analysis of the transition graph of the pendulum equation

$$\theta'' + \frac{\mu}{m}\theta' + \frac{g}{l}\sin\theta = 0,$$

shown in Figure 7. We begin with a qualitative analysis. Dimensional analysis provides the signs of the transition probabilities, as discussed in Section 2. The transition probabilities

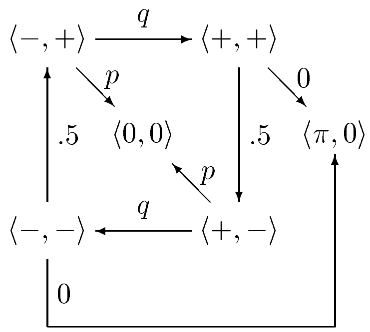

Figure 8: Prototypical transition probabilities for the pendulum with $q = (1-p)/2$. Self-loop probabilities are not shown.

into $\langle \pi, 0 \rangle$ are zero. The transition probabilities into $\langle 0, 0 \rangle$ are positive in the damped case ($\mu > 0$) and zero in the undamped case ($\mu = 0$). In the damped case, the transition graph forms an absorbing chain with absorbing states $\langle 0, 0 \rangle$ and $\langle \pi, 0 \rangle$. The pendulum eventually approaches $\langle 0, 0 \rangle$ with probability one from any transient state; it cannot cycle between the transient states forever. In the undamped case, the transition graph decomposes into three ergodic chains: $\{\langle 0, 0 \rangle\}$, $\{\langle \pi, 0 \rangle\}$, and $\{\langle +, + \rangle, \langle +, - \rangle, \langle -, - \rangle, \langle -, + \rangle\}$. Trajectories in the third chain, which comprises essentially the entire phase space, oscillate around the origin forever.

Given additional information about the transition probabilities, we can estimate the absorption times for the damped case and the asymptotic distribution of states in the undamped case. For example, suppose that the transition probability into $\langle 0, 0 \rangle$ is p and that all other probabilities are equally distributed (Fig. 8). The matrix of absorption times is given by equation (5) for $p > 0$.

$$\tau = p^{-1} \begin{pmatrix} 2-p \\ 2+p \\ 2-p \\ 2+p \end{pmatrix} \tag{5}$$

The equation shows that the settling time of the pendulum increases toward ∞ as friction decreases toward zero. Equation (5) does not apply in the limiting case of $p = 0$, since $\langle 0, 0 \rangle$ ceases to be an absorbing state. Under our equiprobability assumption, trajectories are equally likely to be in any of the four regions in the third ergodic chain at any given time.

5.2 The two-charge pendulum

The gravitational pendulum model presupposes that the force on the bob is independent of the bob's location. The model for the variable attraction between a positively charged pendulum bob and two negative charges is more complicated (Fig. 9). Each negative charge exerts a force

$$f(\alpha, k) = k(d+l)l^{-5}(\sin\alpha)(d^2 + dl + 2l^2 - 2l(d+l)\cos\alpha)^{-3/2}$$

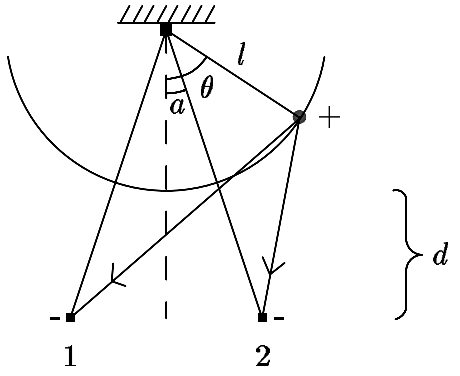

Figure 9: A positively charged pendulum attracted by two negative charges.

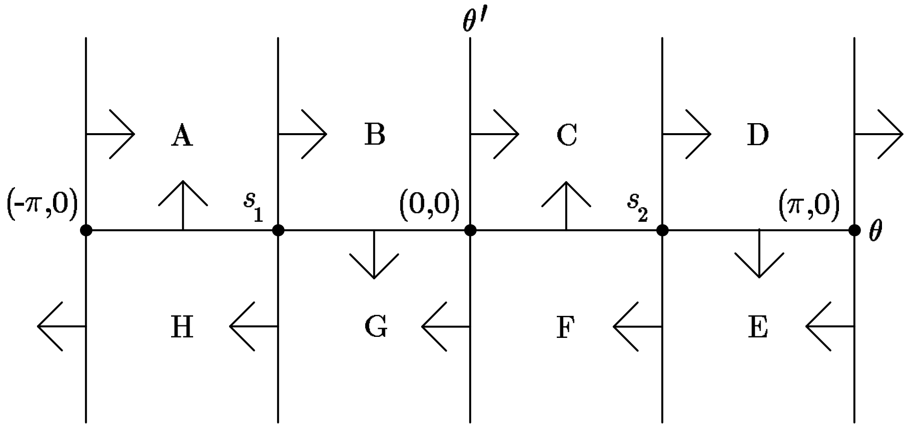

Figure 10: Qualitative dynamics for the two-charge pendulum.

along the line between it and the bob, with α the angle between that line and the pendulum, k the coefficient of electrostatic attraction between the bob and the charge, l the length of the arm, and d the vertical distance from the bob's orbit to the line connecting the two negative charges. The two-charge pendulum obeys the equation

$$\theta'' + \frac{\mu}{m}\theta' + f(\theta + a, k_1) + f(\theta - a, k_2) = 0$$

with a the angle between each pole and the vertical, k_1 and k_2 corresponding to the left and right charges, $\mu > 0$ the damping coefficient, and m the mass of the bob. Figure 10 contains the qualitative dynamics for the case of equal charges ($k_1 = k_2$). Saddles appear at $(\pi, 0)$ and $(0, 0)$ where the charges cancel each other. A sink appears where each charge is strongest. The pendulum can spin (A-B-C-D and E-F-G-H), oscillate around both negative charges (A-B-C-D-E-F-G-H), oscillate around the left charge (A-B-G-H), or oscillate around the right charge (C-D-E-F). It can also switch from spinning to oscillating and from oscillating around both charges to oscillating around either charge. When the charges are unequal, the unstable equilibria move away from $\langle 0, 0\rangle$ and $\langle \pi, 0\rangle$ and the stable equilibria are positioned asymmetrically around the θ' axis, but otherwise the qualitative dynamics appears just as in Figure 10.

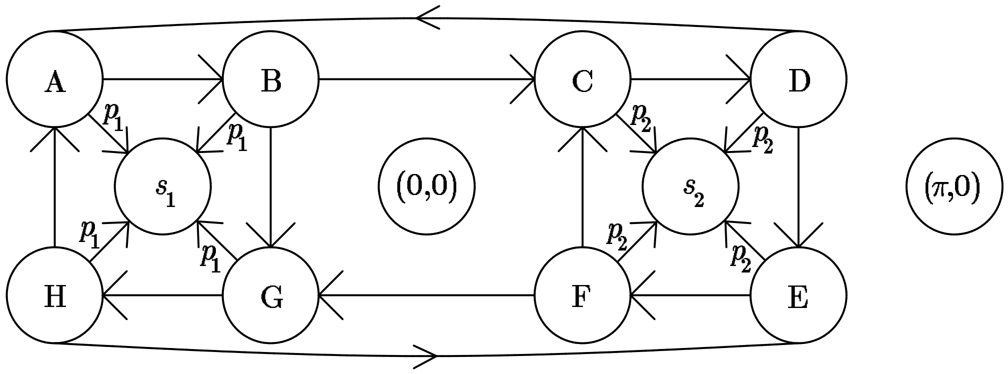

Figure 11: Transition probabilities for the two-charge pendulum. The probabilities from the adjacent regions of s_i are p_i. The unmarked transitions of each node have equal probabilities.

The qualitative analysis of the two-charge pendulum resembles that of the gravitational pendulum. Dimensional analysis determines that the transition probabilities are zero into the saddles and positive into the sinks. The transition graph forms an absorbing chain whose absorbing states are the fixed points. The pendulum eventually approaches one of the sinks with probability one from the remaining, transient states. It is more likely to end up at the sink with the larger charge. If the charges are equal, it is equally likely to end up at either one.

To estimate the relative likelihoods of the two sinks, we assign transition probabilities of p_1 and p_2 to transitions into s_1 and s_2 from adjacent regions and assume that the remaining transitions are equally distributed (Fig. 11). Let p_1^∞ and p_2^∞ represent the asymptotic probabilities of appearing in states s_1 and s_2, averaged over all possible transient starting states. Calculating the ratio $r^\infty = p_1^\infty / p_2^\infty$ yields

$$r^\infty = \left(\frac{p_1}{p_2}\right)\left(\frac{(p_1-1)p_2^2-(3p_1+21)p_2+2p_1-14}{(p_1^2-3p_1+2)p_2-p_1^2-21p_1-14}\right)$$

The dependence of r^∞ on p_1 and p_2 agrees with our intuitions. The ratio increases monotonically from 0 as p_1 increases from 0 to 1 and decreases monotonically from ∞ as p_2 increases from 0 to 1. It equals 0 when $p_1 = 0$, ∞ when $p_2 = 0$, and 1 when p_1 equals p_2.

5.3 The quadratic map

Markov analysis handles discrete dynamic systems as well as continuous ones. The evolution law of a discrete system maps states to their immediate successors. Given initial state x_0 and map f, the system is in state $x_1 = f(x_0)$ at time 1, $x_2 = f(f(x_0))$ at time 2, and $x_i = f^i(x_0)$ at time i. The set of iterates $\{x_i\}$ is called the *trajectory* of x_0. Dynamic systems theory seeks to determine the qualitative properties of a system's trajectories from its evolution law. For example, a trajectory is *fixed* if $f(x) = x$ and *periodic* with period p if $f^p(x) = x$. Devaney [2] provides a good elementary introduction to discrete dynamic systems, including the example below.

Discrete systems are useful in modeling population dynamics. Here x_0 denotes an initial population, x_i denotes the population after i generations, and $f(x)$ expresses the birth rate.

May [10] studies the system $f(x) = ax - bx^2$ with a and b positive parameters that represent natural reproduction and the negative effects of overcrowding. The birth rate increases from 0 to a maximum then decreases to 0 as x increases from 0 to a/b. Even this system, arguably the simplest nonlinear one possible, can exhibit extremely complicated behavior.

Following May, we substitute bx/a for x to obtain the canonical form

$$f(x) = ax(1 - x),$$

which is called the *quadratic map*. Trajectories that leave the interval $[0, 1]$ approach $-\infty$ monotonically and $f(1) = 0$, implying that large enough populations always die out. The problem is to identify and characterize the trajectories that remain in $(0, 1)$ forever, which represent the stable patterns of population fluctuation. The answer depends on a. The point $(a - 1)/a$ is always fixed. For $a < 3$, all trajectories in $(0, 1)$ approach $(a - 1)/a$ asymptotically. This means that populations below a certain size approach a stable equilibrium, whereas larger populations die out. New phenomena arise within $(0, 1)$ for $3 \leq a \leq 4$, including periodic orbits with arbitrarily large periods and chaos, but trajectories still cannot escape because $f \leq 1$. In physical terms, the population can vary erratically, but cannot die out. For $a > 4$, almost all trajectories escape and the remainder form a chaotic Cantor set; almost all initial populations die out, but some persist and vary erratically.

Rigorous derivation of these conclusions requires great mathematical expertise. More complicated equations defy analysis altogether. Markov analysis, although more limited, provides many of the same answers and applies uniformly to all equations. We define a Markov chain with two states $[0, 1]$ and its complement $C = (-\infty, 0) \cup (1, \infty)$. For $a \leq 4$, we obtain two absorbing states because trajectories never cross between regions. For $a > 4$, we obtain an absorbing chain with absorbing state C, implying that trajectories in $[0, 1]$ eventually escape to C with probability 1. In physical terms, the population survives for $a < 4$ and dies out otherwise.

Markov theory also predicts the extinction time for $a > 4$, which equals the absorption time into C. The transition probability from $[0, 1]$ to C equals the measure of the set of points in $[0, 1]$ that map directly into C. This set is the subinterval

$$\left[\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{a}},\quad \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{a}}\,\right],$$

so the transition probability is $p = \sqrt{1 - 4/a}$. The expected absorption time equals $1/p$, as explained in Section 4.1. Figure 12 compares the predicted mean absorption times with the average absorption times derived by numerically simulating trajectories for 20000 pseudorandom initial points. The small relative error indicates that the chain assumption applies well.

In a more refined analysis, we could investigate the behavior of the quadratic map within the interval $[0, 1]$ by partitioning that interval into small subintervals. Markov analysis would estimate the distribution of states within each region. This approach provides a statistical understanding of systems whose individual trajectories defy analysis. For example, it demonstrates that most iterates of the map $4x(1 - x)$ on $[0, 1]$ cluster near the endpoints, a result that Lasota and Mackey [9] confirm by analytic means. In contrast,

a	Markov prediction	Observed time	% error
4.1	6.40	6.49	1.4
4.2	4.58	4.58	0.0
4.5	3.00	2.98	0.6
5	2.24	2.22	0.9
6	1.73	1.75	1.1

Figure 12: Stochastically predicted time for the quadratic map to leave $[0, 1]$ compared with the average observed time for 20000 pseudorandom starting points.

the chaotic behavior of the quadratic map implies that some trajectories cross between every pair of regions, thus reducing its qualitative dynamics to a complete graph, devoid of predictive power.

6 Conclusions and future work

We apply the theory of Markov chains to estimate the relative likelihoods of possible behaviors of a system, thereby filling a serious gap in the predictions of qualitative simulation. Our method provides a formal justification and an efficient algorithm for commonsense and quantitative reasoning about relative likelihoods. It enables us to draw the best possible conclusions from the available information. We can determine the possible long term behaviors of a system directly from its qualitative dynamics in a simple qualitative analysis that runs in linear time. More detailed information, such as the likelihoods of the possible behaviors and the expected settling times for each initial state, requires estimates of the transition probabilities between qualitative states. The estimates can be numeric or symbolic; the analysis is formally identical in both cases, but has $O(n^3)$ time-complexity in the former and exponential time-complexity in the latter. We exhibit the utility of our method in several examples, and analyze the robustness of its conclusions to perturbations in the transition probabilities. The likelihoods of the long term behaviors are never sensitive to perturbations in the transition probabilities, whereas the expected settling times can be sensitive.

Markov analysis of dynamical systems has also been pursued by Hsu [7], whose analysis of "generalized cell-to-cell mappings" is similar to our analysis of dynamics over qualitative states. Hsu develops algorithms which yield the same numerical information as ours for numerically formulated systems. These algorithms are more limited than ours, however, in that they cannot provide purely qualitative or symbolic conclusions from purely qualitative or symbolic information about the dynamic system.

Our current analysis is only a first step towards full exploitation of the stochastic approach to qualitative reasoning. We have not fully explored the potential of Markov theory. Further investigation may yield simple ways of determining other qualitative properties of systems through application of known techniques, or purely qualitative algorithms for obtaining qualitative predictions of relative likelihoods. One might also relax the chain assumption underlying our treatment and instead view the qualitative dynamics as describing a more general Markov process in which transitions depend on past states. There is a rich theory of these processes which may support many of the same conclusions as above in the more general setting.

Incorporating global properties of flows into stochastic analysis is another topic for future research. Our current analysis derives the qualitative dynamics of a system from local properties of its flow: where it vanishes and whether it crosses certain curves. Sacks [13] shows how to increase the accuracy of the qualitative dynamics by ruling out locally consistent behaviors that violate global constraints, thus reducing the size of a transition graph. For example, his program uses an energy argument to prove that a pendulum cannot spin forever. One could make this argument within the qualitative dynamics by partitioning phase space along level curves of the pendulum's total energy. Energy conservation then rules out transitions from a region to higher energy regions. This dynamics eliminates the spurious behaviors of wobbling and unbounded oscillation, which the original dynamics permits. It facilitates stochastic analysis by reducing the size of the transition graph and by providing trapping regions for sinks without the estimation problems discussed in Section 5.2. Energy conservation also implies that the pendulum cannot spin forever without resort to the chain assumption, thus lightening the burden on stochastic analysis. Deriving and exploiting the relations between the stochastic model and the global phase flow is a more ambitious task.

Incorporating global analysis [15] with stochastic analysis is another direction for future research. Global analysis considers not just one flow, but a class of flows. This is useful when we do not know the exact equations describing a system, and wish to make predictions based on what we do know about them. One global concept is that of *structurally stable* system, all of whose perturbations have the same qualitative behavior. Global analysis formalizes such notions by considering measures over classes of flows. When the measures considered are probability densities over systems, the stochastic analysis presented above can be directly extended to the case of indeterminate dynamics. For example, the set of flows for the Lottka-Volterra model of population sizes of competing species divides into four classes of flows, each with qualitatively different dynamics [6]. Each of these has a small number of attractors, and can be analyzed for expected asymptotic behavior just as was done with the charged pendulum in Section 5.2. If we can estimate (by some means) the probability that the actual system falls into each class, we can then calculate the overall probability of each possible asymptotic behavior in each of the different dynamics.

Automating the model refinement procedure for stochastic qualitative analysis is a major challenge. To do so requires answering some fundamental questions about the stability of the stochastic predictions under changes in the underlying qualitative graph, such as dividing phase space into overly-fine regions. Yip [16] presents some methods for sampling and observing the behaviors of dynamical systems. Perhaps they can be extended to the model construction task. Machine learning techniques may also prove relevant.

Acknowledgments

The authors thank William Long, Ramesh Patil, Peter Szolovits, Michael Wellman, and the anonymous referees for helpful comments on drafts of this paper, and Kenneth Yip for valuable suggestions. Jon Doyle is supported by the National Library of Medicine under National Institutes of Health Grant R01 LM04493. Elisha Sacks is supported by the National Science Foundation under grant No. IRI–9008527.

References

[1] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The Design and Analysis of Computer Algorithms.* Addison-Wesley Publishing Company, 1974.

[2] R. L. Devaney. *An Introduction to Chaotic Dynamical Systems.* Benjamin/Cummings, Menlo Park, California, 1986.

[3] J. Doyle and E. P. Sacks. Stochastic analysis of qualitative dynamics. In N. S. Sridharan, editor, *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 1187–1192, San Mateo, CA, 1989. Morgan Kaufmann.

[4] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume I. John Wiley and Sons, New York, 1957.

[5] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag, New York, 1986.

[6] M. W. Hirsch and S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press College Division, Orlando, Florida, 1974.

[7] C. S. Hsu. *Cell-to-Cell Mapping.* Springer-Verlag, New York, 1987.

[8] J. G. Kemeny and J. L. Snell. *Finite Markov Chains.* Springer-Verlag, New York, 1976.

[9] A. Lasota and M. C. Mackey. *Probabilistic Properties of Deterministic Systems.* Cambridge University Press, Cambridge, 1985.

[10] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467, 1976.

[11] F. S. Roberts. *Discrete Mathematical Models.* Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1976.

[12] E. P. Sacks. Piecewise linear abstraction of intractable dynamic systems. *International Journal of Artificial Intelligence in Engineering*, 3(3):151–155, July 1988.

[13] E. P. Sacks. Automatic qualitative analysis of ordinary differential equations using piecewise linear approximations. *Artificial Intelligence*, 41(3):313–364, Jan. 1990.

[14] E. P. Sacks. A dynamic systems perspective on qualitative simulation. *Artificial Intelligence*, 42(2-3):349–362, 1990.

[15] S. Smale. *The Mathematics of Time: Essays on Dynamical Systems, Economic Processes, and Related Topics.* Springer-Verlag, New York, 1980.

[16] K. M.-K. Yip. Generating global behaviors using deep knowledge of local dynamics. In *Proceedings of the National Conference on Artificial Intelligence*, pages 280–285. American Association for Artificial Intelligence, 1988.

Contents