Verifying Continuous Time Markov Chains

Adnan Aziz
ECE
UT Austin

Kumud Sanwal
Bell Labs
AT&T

Vigyan Singhal
CBL
Cadence

Robert Brayton
EECS
UC Berkeley

Abstract. We present a logical formalism for expressing properties of continuous time Markov chains. The semantics for such properties arise as a natural extension of previous work on discrete time Markov chains to continuous time. The major result is that the verification problem is decidable; this is shown using results in algebraic and transcendental number theory.

Introduction

Recent work on formal verification has addressed systems with stochastic dynamics. Certain models for discrete time Markov chains have been investigated in [6, 3]. However, a large class of stochastic systems operate in continuous time. In a generalized decision and control framework, continuous time Markov chains form a useful extension [9]. In this paper we propose a logic for specifying properties of such systems, and describe a decision procedure for the model checking problem. Our result differs from past work in this area [2] in that quantitative bounds on the probability of events can be expressed in the logic.

1 Continuous Markov Chains

We will consider models of the form $M = (S, \Lambda, A, \theta)$, where $S = \{s_1, s_2, \ldots, s_n\}$ is a finite set of *states*, Λ is the *transition rate matrix*, A is a finite set of *outputs*, and $\theta : S \rightarrow A$ is the *output function*. A *path* through M is a map from $\mathbb{R}^+$ to S (here $\mathbb{R}^+$ denotes the set of non-negative reals); $S^{\mathbb{R}^+}$ is the set of all paths.

The transition rate matrix Λ is an $|\, S \,| \times |\, S \,|$ matrix. The off diagonal entries are non-negative rationals; the diagonal element a_{jj} is constrained to be $-(\sum_{i \neq j} a_{ji})$.

At state s_j, the probability of making a transition to state s_k (where $k \neq j$) in time dt is given by $a_{jk}\, dt$. This is the basis for formulating a stochastic differential equation for the evolution of the probability distribution whose solution is

$$D(t) = e^{\Lambda t} \cdot D_0$$

Here D_0 is a column vector of dimension $|S|$, with the constraint that $\sum_i [D_0]_i = 1$.

Technically, with any state s in M we associate a natural *probability space* $\mathcal{P}_M^s = (U^s, \mathcal{C}^s, \mu^s)$, where the set of all paths starting at s is the *universe* U^s, and the Borel sigma field on U^s gives the associated space of *events* $\mathcal{C}^s$, i.e.

the class of subsets of U^s to which probabilities can be assigned. The transition rate matrix Δ yields the probability measure $\mu^s : \mathcal{C}^s \to [0,1]$; by the measure extension theorem [10], μ^s is well defined. Given a set β of functions from $\mathbb{R}^+$to A, we will abuse notation and refer to the probability of β when we mean the probability of the set of all state sequences starting at s which map under θ to elements in β.

We will not dwell on the technicalities of measure theory; all the sets of paths defined later in this paper will be readily seen to be events, i.e. elements of $\mathcal{C}^s$.

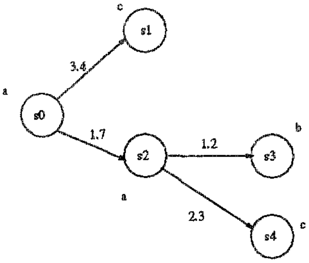

Fig. 1. A continuous time Markov chain: $S = \{s_0, s_1, s_2, s_3\}$; only edges with positive weights are shown.

2 CSL syntax and semantics

Let $M = (S, A, A, \theta)$ be a continuous Markov chain. In this section, we develop formal syntax and semantics for CSL (Continuous Stochastic Logic). This logic is inspired by the logic CTL [4], and its extensions to discrete time stochastic systems (pCTL [6]), and continuous time non-stochastic systems (tCTL [1]).

There are two types of formulae in CSL: state formulae (which are true or false in a specific state), and path formulae (which are true or false along a specific path). A state formula is given by the following syntax:

1. a for $a \in A$
2. If f_1 and f_2 are state formula, then so are $\neg f_1, f_1 \vee f_2$
3. If g is a path formula, then $Pr_{>c}(g)$ is a state formula. (c is a rational between 0 and 1 expressed as the ratio of two binary coded integers).

Path formulas are formulas of the form

- $f_1 U_{[a_1,b_1]} f_2 U_{[a_2,b_2]} \cdots f_n$, where $f_1, f_2, \ldots f_n$ are state formulas, and $a_1, b_1, \ldots, a_{n-1}, b_{n-1}$ are non-negative rationals expressed as the ratio of two binary coded integers.

CSL is the set of state formulae that are generated by the above rules.

Let f be a state formula, and g be a path formula. We now define the satisfaction relation ($\models_M$) using induction on the length of the formula. For a state formula f we use $[\![f]\!]_M$ to denote the set of states satisfying f.

1. f is of the form **a**: $s \models_M f$ iff $\theta(s) = a$.
2. f is of the form $(\neg f_1)$: $s \models_M f$ iff $s \not\models_M f$.
3. f is of the form $(f_1 \vee f_2)$: $s \models_M f$ iff $s \models_M f_1$ or $s \models_M f_2$.
4. f is of the form $Pr_{>c}(g)$: $s \models_M f$ iff $\mu^s(\{\pi \in S^{\mathbb{R}^+} \mid \pi \models_M g\}) > c$.
5. g is a path formula of the form $f_1 U_{[a_1,b_1]} f_2 U_{[a_2,b_2]} \cdots f_n$: $\pi \models_M g$ iff there exist $t_1, \ldots, t_{n-1}$ such that $(\forall i)\ [(a_i \leq t_i \leq b_i)\ \wedge\ (\forall t' \in [t_{i-1}, t))\,(\pi(t) \in [\![f_i]\!]_M)]$ (where t_{-1} is defined to be 0 for convenience).

Example 1. The formula $\phi = Pr_{>0.03}(aU_{[0.0,4.0]}b)$ is a state formula for the example in figure 1. It formally expresses that with probability greater that 0.03, the system will remain in a state where the output is a before making a transition before 4.0 seconds have elapsed to a state where the output is b.

The probability of the set of paths starting at s_0 on which the output is a before becoming b before time t_1 is given by the following integral:

$$\int_0^{t_1} \int_0^{t_1 - \tau_1} e^{-(r_1+r_2)\cdot\tau_1} \cdot e^{-(r_3+r_4)\cdot\tau_2} \cdot (r_2/(r_1+r_2)) \cdot (r_3/(r_3+r_4)) \cdot d\tau_1 \cdot d\tau_2$$

Setting t_1 equal to 4.0, and taking the rates $r_1 = 1.0, r_2 = 2.0, r_3 = 3.0$, and $r_4 = 3.0$, this simplifies to $(1/45) - (e^{-12}/18) + (e^{-20}/30)$. Observe that $e^{-12}/18 > e^{-20}/30$; hence the probability is bounded above by $(1/45)$ which is less than 0.03, and so ϕ is false at s_0.

3 CSL model checking

The CSL model checking problem is as follows: given a continuous Markov chain M, a state s in the chain, and a CSL formula f, does $s \models_M f$? In this section we establish that the model checking problem for CSL is decidable.

Theorem 1. CSL model checking is decidable.

Proof. The non-trivial step in model checking is to model check formula of the form $Pr_{>c}(g)$. In order to do this we need to be able to effectively reason about the measure of the set $\{\pi \in S^{\mathbb{R}^+} \mid \pi(0) = s_0 \wedge \pi \models_M g\}$ under μ^{s_0}.

First, we review some elementary algebra. An *algebraic complex number* is any complex number which is the root of a polynomial with rational coefficients. Properties of the algebraic numbers are derived in [8]; of particular interest to us is the fact that they constitute a field, and that the real and imaginary parts of an algebraic number are also algebraic.

We will denote the set of complex numbers which are of the form $\sum_j \eta_j e^{\delta_j}$ where the η_j and δ_j are algebraic by $E_{\mathcal{A}}$. This set is a ring, and is referred to as the *transcendental extension* of A by e [8].

Tarksi [11] proved that the theory of the field of complex numbers (i.e. the theory of the structure $<\mathbb{C}, +, \times, 0, 1>$) was decidable; an effective (in the recursion theoretic sense) procedure for converting formulas to a logically equivalent quantifier free form was given. Consequences of this result include the existence of effective procedures for determining the number of distinct roots of a polynomial, and testing the equality of algebraic numbers defined by formulas.

We now demonstrate how to measure the set of paths which start at a designated state and satisfy a specified path formula. Consider a path formula of the form $\psi_0 U_{[a_1,b_1]} \psi_1 U_{[a_2,b_2]} \psi_2 \cdots$

First, consider the case where the time intervals $[a_1, b_1]$, $[a_2, b_2], \ldots$ are non overlapping.

We define the following matrices.

- a transition matrix $Q_{i,i}$ obtained from Λ, that treats $[\![\psi_i]\!]_M^c$ as an absorbing set of states. This is obtained by using
$$\begin{aligned} q(j,k) &= \lambda_{j,k} \ \text{ if } j \in [\![\psi_i]\!]_M \\ &= 0 \quad \text{ if } j \in [\![\psi_i]\!]_M^c \end{aligned}$$
this enables us to model the transitions where the Markov chain remains in $[\![\psi_i]\!]_M$.
- a transition matrix $Q_{i-1,i}$ obtained from Λ, that treats $[\![\psi_{i-1}]\!]_M^c \cap [\![\psi_i]\!]_M^c$ as an absorbing set of states. For this we use
$$\begin{aligned} q(j,k) &= \lambda_{j,k} \ \text{ if } j \in [\![\psi_i]\!]_M \cup [\![\psi_{i-1}]\!]_M \\ &= 0 \quad \text{ if } j \in [\![\psi_i]\!]_M^c \cap [\![\psi_{i-1}]\!]_M^c \end{aligned}$$
this allows us to model the transitions from $[\![\psi_{i-1}]\!]_M$ to $[\![\psi_i]\!]_M$.
- An indicator matrix I_i for $[\![\psi_i]\!]_M$, such that
$$\begin{aligned} I_i(j,k) &= 1 \ \text{ if } j = k \in [\![\psi_i]\!]_M \\ &= 0 \ \text{ otherwise} \end{aligned}$$

Hence, the probability of a formula of the form

$$f_1 = \psi_0 U_{[a_1,b_1]} \psi_1 U_{[a_2,b_2]} \psi_2 \cdots U_{[a_n,b_n]} \psi_n \tag{1}$$

is given by

$$\begin{aligned} \mu^s(f_1) = \pi_s \cdot P_{0,0}(a_1) \cdot I_0 \cdot P_{0,1}(b_1 - a_1) \cdot I_1 \cdot P_{1,1}(a_2 - b_1) \cdot \\ I_1 \cdot P_{1,2}(b_2 - a_2) \cdot I_2 \cdots P_{n-1,n}(b_n - a_n) \cdot I_n \cdot \underline{1} \end{aligned} \tag{2}$$

where $P_{l,m}(t), t \geq 0$ is the one step transition matrix for time t corresponding to the rate matrix $Q_{l,m}$, π_s is the starting probability distribution, which in our case has unity for state s and zeros otherwise, and $\underline{1}$ is the column vector whose elements are all 1. For a finite state Markov chain with a transition rate matrix Q, this matrix is given by

$$P(t) = e^{Qt}$$

Note that Q is composed of rational entries, and the arguments of $P_{i-1,i}$ are rationals (since a_i, b_i are rational). This observation leads to the following lemma:

Lemma 2. Each element of the $P_{l,m}(t)$ matrices may be expressed as $\sum_j \eta_j e^{\delta_j}$ where η_j and δ_j are algebraic complex numbers.

Proof. Any square matrix B can always be expressed in Jordan canonical form [7], i.e. in the form $C \cdot J \cdot C^{-1}$. Here J is an upper block diagonal matrix as shown below:

$$\begin{bmatrix} J_1 & 0 & \cdots & & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ 0 & \cdots & J_3 & \cdots & 0 \\ & & & \ddots & \\ 0 & & \cdots & & J_n \end{bmatrix}$$

The diagonal entries of each J_i are the eigenvalues of B, and the remaining entries of J_i are unity, as shown below:

$$\begin{bmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ & & & \ddots & \\ 0 & & \cdots & 0 & \lambda_i \end{bmatrix}$$

The size of J_i is equal to the multiplicity of λ_i. Since the eigenvalues are the solutions of the characteristic equation of B and the entries of B are rationals, the eigenvalues are, by definition, algebraic complex numbers. Similarly, the entries of C, C^{-1} are also algebraic complex numbers.

The matrix e^{Bt} is equal to $C \cdot e^{Jt} \cdot C^{-1}$ and e^{Jt} is of the following form:

$$\begin{bmatrix} e^{J_1 t} & 0 & \cdots & & 0 \\ 0 & e^{J_2 t} & 0 & \cdots & 0 \\ 0 & \cdots & e^{J_3 t} & \cdots & 0 \\ & & & \ddots & \\ 0 & & \cdots & & e^{J_n t} \end{bmatrix}$$

The sub-matrix $e^{J_i t}$ is of the form

$$\begin{bmatrix} e^{\lambda_i t} & te^{\lambda_i t} & (t^2 e^{\lambda_i t})/2! & \cdots & (t^{m_i} e^{\lambda_i t})/(m_i)! \\ 0 & e^{\lambda_i t} & te^{\lambda_i t} & \cdots & (t^{m_i - 1} e^{\lambda_i t})/(m_i - 1)! \\ & & & \ddots & \\ & & & & e^{\lambda_i t} \end{bmatrix}$$

By inspection, the elements of $e^{J_i t}$ are members of $E_{\mathcal{A}}$. Since $E_{\mathcal{A}}$ is a ring, it is closed under products and sums. Hence the lemma follows. It also follows that $\mu^s(f_1)$ is a member of $E_{\mathcal{A}}$ i.e. equal to an expression of the form $\sum_k \eta_k e^{\delta_k}$ where the η_k, δ_k are algebraic. ■

Consider again the expression for $\mu^s(f_1) = \sum_k \eta_k e^{\delta_k}$. The δ_k's are algebraic; since are effective procedures for checking the equality of algebraic numbers, $\mu^s(f_1)$ can be effectively simplified to an expression of the form $\sum_{k'} \eta_{k'} e^{\delta_{k'}}$ where the $\eta_{k'}$'s are non zero, and the $\delta_{k'}$'s are distinct.

In order to decide if $\mu^s(f_1) > c$, we exploit a celebrated theorem of transcendental number theory [8].

Theorem 3 (Lindemann-Weirstrass:). *Let $c_1, \ldots c_n$ be pairwise distinct algebraic numbers belonging to $\mathbb{C}$. Then there exists no equation $a_1 e^{c_1} + \cdots + a_n e^{c_n} = 0$ in which $a_1, \ldots, a_n$ are algebraic numbers and are not all zero.*
Historical note:This result implies the transcendence of π (take $n = 2$, $c_1 = 2, c_2 = i\pi$); it was the first proof of the non-algebraic nature of π. For a highly readable account of the development of this theorem, refer to [5].

Suppose the expression $\sum_{k'} \eta_{k'} e^{\delta_{k'}}$ is degenerate, i.e. it consists of a single term of the form η_0. Then the expression denotes an algebraic number, and it can be effectively checked if it is greater than c.

If it is not degenerate, invoking the Lindemann-Weirstrass theorem and noting that c is rational, we see that $\mu^s(f_1)$ can not be equal to c and so $\mid \mu^s(f_1) - c \mid > 0$.

Decidability of model checking follows from the following lemma.

Lemma 4. Given a transcendental real r of the form $\sum_j \eta_j e^{\delta_j}$ where the η_j and δ_j are algebraic complex numbers, and the δ_j's are pairwise distinct, there is an effective procedure to test if $r > c$ for rational c.

Proof. Suppose a sequence of algebraic numbers $S_1, S_2, \ldots$ such that $|r - S_k| < 2^{-k}$ can be effectively constructed. Let $|r - c| = a > 0$. By the triangle inequality, $|r - c| \leq |r - \mathtt{Re}(S_k)| + |\mathtt{Re}(S_k) - c|$. Hence $|r - \mathtt{Re}(S_k)| + |\mathtt{Re}(S_k) - c|$ is bounded away from 0 by a. Since r is real, $|r - \mathtt{Re}(S_k)| \leq |r - (S_k)| < 2^{-k}$, and $|r - \mathtt{Re}(S_k) + |\mathtt{Re}(S_k) - c|$ is bounded away from 0 by a, for sufficiently large k, it must be that $|\mathtt{Re}(S_k) - c| > 2^{-k}$. The sign of of $\mathtt{Re}(S_k) - c$ is the sign of $r - c$.

In order to construct the sequence $S_1, S_2, \ldots$, we use the fact that e^z can be approximated with an error of less than ϵ (when $\epsilon < 1$) by taking the first $\lceil (3 \cdot \mid z \mid^2 / \epsilon) \rceil + 1$ terms of the Maclaurin expansion for e^z. This can be extended to obtain an upper bound on the number of terms to sum for an expression of the form $\sum_j \eta_j e^{\delta_j}$ (which being the finite sum of algebraic numbers is algebraic) in order to achieve an error of less than ϵ. ∎

Now consider the case where the successive intervals where the transitions are desired ($[a_i, b_i]$, $i = 1, 2, \ldots$ are allowed to overlap. Since a formula is finite, we can have a finite number of overlapping intervals. A key observation is that the finite number of overlaps allows us to partition the time in a finite number of non-overlapping intervals and write the probability of the specification (set of acceptable paths) as a sum of the probabilities of disjoint events. This enables us to write $\mu^s(f_1)$ as the sum of exponentials of algebraic complex numbers, weighted by algebraic coefficients. To illustrate this, consider the formula

$$f_2 = \psi_0 U_{[a_1,b_1]} \psi_1 U_{[a_2,b_2]} \psi_2$$

where $0 < a_1 < a_2 < b_1 < b_2$. In this case, we may realize f_2 as one of four disjoint cases and hence we can write

$$\mu^s(f_2) = \mu^s(\psi_0 U_{[a_1,a_2]} \psi_1 U_{[a_2,b_1]} \psi_2) + \mu^s(\psi_0 U_{[a_1,a_2]} \psi_1 U_{[b_1,b_2]} \psi_2) \\ + \mu^s(\psi_0 U_{[a_2,b_1]} \psi_1 U_{[b_1,b_2]} \psi_2) + \mu^s(\psi_0 U_{[a_2,b_1]} \psi_1 U_{[a_2,b_1]} \psi_2) \qquad (3)$$

The first three terms are equivalent to the case with non-overlapping intervals. The last term involves having both the $[\![\psi_0]\!]_M \to [\![\psi_1]\!]_M$ and $[\![\psi_1]\!]_M \to [\![\psi_2]\!]_M$ transitions in the same interval $[a_2, b_1]$ in the correct order. This may be evaluated by integrating the probabilities over the time of the first transition.

$$\mu^s(\psi_0 U_{[a_2,b_1]} \psi_1 U_{[a_2,b_1]} \psi_2) = \pi_s P_{0,0}(a_2) I_0 \int_{a_2}^{b_1} P_{0,0}(t - a_2) I_0 Q_{0,1} I_1 P_{1,2}(b_1 - t) I_2 dt$$

It is clear that since the integrand involved algebraic terms and and exponentials in algebraic complex numbers and t, the definite integral with rational limits can be written in the form of a sum of exponentials of algebraic numbers with algebraic coefficients. Hence, this term is in $E_{\mathcal{A}}$. The other three terms in equation 3 correspond to forms equivalent to the non-overlapping intervals case, and hence already satisfy the decidability criteria. ■

4 Conclusions and Future Work

We have defined a logic for specifying properties of finite state continuous time Markov chains. The model checking problem for this logic was shown to be decidable through a combination of results in algebraic and transcendental number theory. In practise we believe that a model checker can be built using conventional numerical methods for computing probabilities of events in continuous Markov chains.

In the future, we intend to study synthesis of specifications in the logic. We are planning to use some of the techniques used in this paper to derive decidability results for analyzing dynamical systems which evolve using exponential laws.

References

1. R. Alur, C. Courcoubetis, and D. Dill. Model Checking for Real-Time Systems. In *Proc. IEEE Symposium on Logic in Computer Science*, pages 414–425, 1990.
2. R. Alur, C. Courcoubetis, and D. Dill. Model Checking for Probabilistic Real Time Systems. In *Proc. of the Colloquium on Automata, Languages, and Programming*, pages 115–126, 1991.

3. C. Courcoubetis and M. Yannakakis. Verifying Temporal Properties of Finite State Probabilistic Programs. In *Proc. IEEE Symposium on the Foundations of Computer Science*, pages 338–345, 1988.
4. E. A. Emerson. Temporal and Modal Logic. In J. van Leeuwen, editor, *Formal Models and Semantics*, volume B of *Handbook of Theoretical Computer Science*, pages 996–1072. Elsevier Science, 1990.
5. J. H. Ewing. *Numbers*. Springer-Verlag, 1991.
6. H. Hansson and B. Jonsson. A Logic for Reasoning about Time and Reliability. *Formal Aspects of Computing*, 6:512–535, 1994.
7. T. Kaliath. *Linear Systems*. Prentice-Hall, 1980.
8. I. Niven. *Irrational Numbers*. John-Wiley, 1956.
9. S. Ross. *Stochastic Processes*. Wiley, 1983.
10. H. L. Royden. *Real Analysis*. Macmillan Publishing, 1989.
11. A. Tarski. *A Decision Procedure for Elementary Algebra and Geometry*. University of California Press, 1951.