Safety Verification of Random Ordinary Differential Equations
随机常微分方程安全性验证

Bai Xue, Martin Fränzle, Naijun Zhan, Sergiy Bogomolov and Bican Xia

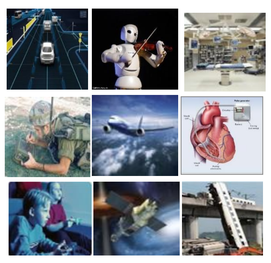
IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

联系方式（薛白、13552483249、xuebai@ios.ac.cn）

1. 研究背景

1.1 信息物理融合系统（CPS）

（计算进程与物理过程结合体，集传感、通信、计算与控制于一身的智能系统）



- **其应用越来越广泛、对人们生活和国家安全的影响越来越大**
- **安全攸关系统：其失效会带来人员伤亡、重大经济损失、对环境的严重破坏等灾难性后果**
- **“如何设计让人们放心使用的信息物理融合系统” 是计算机科学、控制及数学的一个巨大挑战**
- **基于严格数学基础的形式验证被公认为安全攸关系统开发和保证的有效方法，也是工业设计软件的核心功能之一**

1.2 可达集分析

CPS系统是离散控制和连续变化行为相互叠加的混成系统，混成自动机数学模型被广泛用来描述这种混成系统，此模型中通常用（常 / 时滞 / 随机等）微分方程来描述系统的连续行为。近二十年，基于混成自动机对CPS系统形式验证，成为了计算机科学的主要关注热点之一。目前，形式验证方法主要基于模型检验和定理证明两种技术。模型检验是通过穷举待验证系统模型的状态空间来检测系统行为是否满足给定性质的一种自动验证方法，依赖系统状态的可达性分析。定理证明主要解决如何利用逻辑和数学推理的手段来验证系统的关键性质，其核心问题是不变式生成，最终仍可归结为可达性分析问题。从而，基于混成自动机形式验证CPS系统的关键问题是可达性分析，而其中系统连续行为可达集计算是主要挑战。计算准确可达集一般不可能，尤其是非线性连续行为，从而往往借助于通过计算可达集的上近似（Over / Outer / External Approximation）或下近似（Under / Inner / Internal Approximation）来对系统进行形式验证。

- **上近似是可达集的超集，一般用于安全性验证**。上近似与危险区域不相交，则可达集与危险区域不相交，系统安全。如图 1 所示。然而，上近似与危险区域相交，则不能确定系统是否安全。

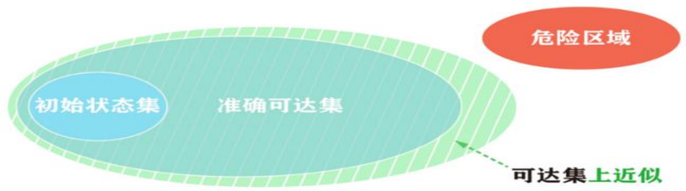


图1： 可达集上近似之系统安全验证图示

- **下近似是可达集的子集，一般用于**
  - **不安全检验**：涉及向前可达集（Forward Reachable Sets）下近似的计算，可用于发现系统bug等。如下近似与危险区域相交，则可达集与危险区域相交，系统不安全。如图 2 所示。
  - **确定使系统安全运行的安全初始状态集合**：涉及向后可达集（Backward Reachable Sets）下近似的计算，用于路径规划等。从安全初始状态集合下近似出发一定能使系统安全运行；而从上近似出发，则不一定。如图 3 所示。

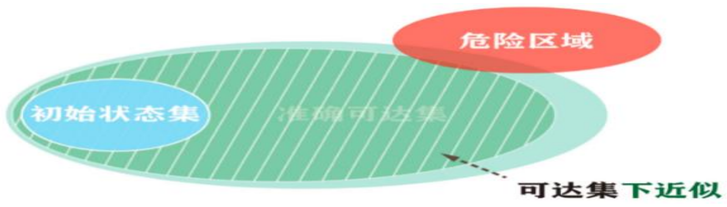


图2： 可达集下近似之系统不安全检验图示

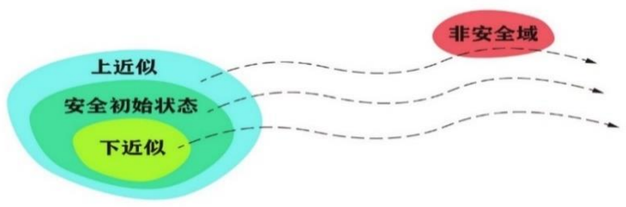


图3： 可达集下近似之系统安全初始状态图示

（黑色虚线表示系统轨道）

1.3 随机常微分方程

在物理世界中，随机现象广泛存在。事实上，实际系统总是存在随机因素的干扰，如气流、电磁等外部环境的干扰，以及系统内部结构参数的扰动等。当这些扰动相对较小、可忽略时，确定性数学模型如（扰动）常 / 时滞微分方程等常用来刻画实际系统连续行为。然而，当这些随机因素不可忽略时，将对系统动力学行为产生质的影响，此类系统为随机系统。随机系统形式验证不同于确定性系统形式验证：确定性系统形式验证涉及到计算使系统一定安全或者一定不安全的可达集上、下近似，而随机系统形式验证具有概率属性，即计算使得系统以一定概率安全或者不安全的可达集上、下近似。

随机微分方程（Stochastic Differential Equations， SDEs）和随机常微分方程（Random Ordinary Differential Equations，RODEs）是刻画随机系统连续动力学行为的两个常用数学模型。Itô 型SDEs

$$d\,x(t) = a\big(x(t)\big)dt + \sigma\big(x(t)\big)d\,w(t) \tag{1}$$

其中$w(t)$是m维矢量高斯随机过程，$a(x)$是n维矢量多项式函数，$\sigma(x)$是$n \times m$维矩阵函数。方程（6）的积分形式为：

$$x(t) = x(t_0) + \int_{t_0}^{t} a(x(s))\,ds + \lim_{|\Delta|\to 0}\sum_{i=1}^{l}\sigma(x(s_{i-1}))(w(s_i) - w(s_{i-1}))$$

其中$t_0 = s_0 < \cdots < s_l = t$ 和$\Delta = \max_{1\le i\le l}(s_i - s_{i-1})$。RODEs的形式一般为：

$$\frac{d\,x(t)}{d\,t} = f\big(x(t), W(t,w)\big) \tag{2}$$

其中$W(t,w)$是定义在概率空间$(\Omega, \mathcal{F}, P)$上的m维随机过程。$W(t,w)$可以是高斯随机过程，也可以是分形布朗运动、（复合）泊松过程等一般Lévy过程，更可以是有界的随机过程。因此，RODE（2）比SDEs（1）有更强的随机行为表达能力，可更准确的刻画一些实际随机系统的连续动力学行为，从而被广泛应用于生物、医学、物理学、航空航天等领域。

RODEs与SDEs有着本质的不同，RODEs是具有随机过程的常微分方程，然而SDEs不是常微分方程。相较于SDEs的形式验证理论及方法，RODEs的形式验证尚处于初始阶段。

2. 问题

目前，RODE（2）的形式验证研究主要集中在$W(\cdot, w) \equiv$常量时。当$W(\cdot, w)$随时间变化时，RODE （2）的形式验证理论及方法仍旧空白。我们在此工作中填充此空白。

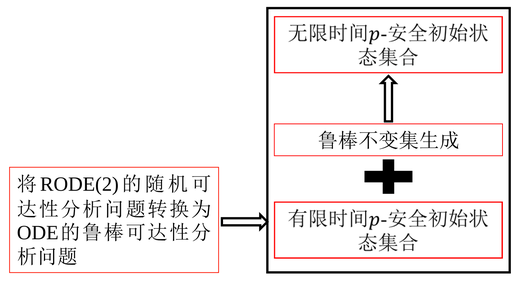
定义：给定安全概率阈值$p \in (0,1)$, 一个安全区域X和一个目标区域TR, 一个初始状态x_0是p-安全的，

1. 有限时间$[0,T]$: 如果从初始状态x_0出发，系统（2）在时间$[0,T]$内在安全区域X中运行并且在时间T时刻到达目标区域TR的概率大于等于p, 即：
$$P\big(\{w \in \Omega \mid \forall s \in [0,T].\, \phi_{x_0}^{W_w}(s) \in X \wedge \phi_{x_0}^{W_w}(T) \in \mathrm{TR}\}\big) \ge p,$$
其中$\phi_{x_0}^{W_w}(\cdot): [0,T] \to \mathbb{R}^n$是系统（2）在随机路径$W(\cdot, w)$下的轨道。
2. 无限时间$[0,\infty)$: 如果从初始状态x_0出发，系统（2）一直在安全区域X中运行的概率大于等于p, 即：
$$P\big(\{w \in \Omega \mid \forall s \in [0,\infty).\, \phi_{x_0}^{W_w}(s) \in X\}\big) \ge p,$$
其中$\phi_{x_0}^{W_w}(\cdot): [0,\infty) \to \mathbb{R}^n$是系统（2）在随机路径$W(\cdot, w)$下的轨道。

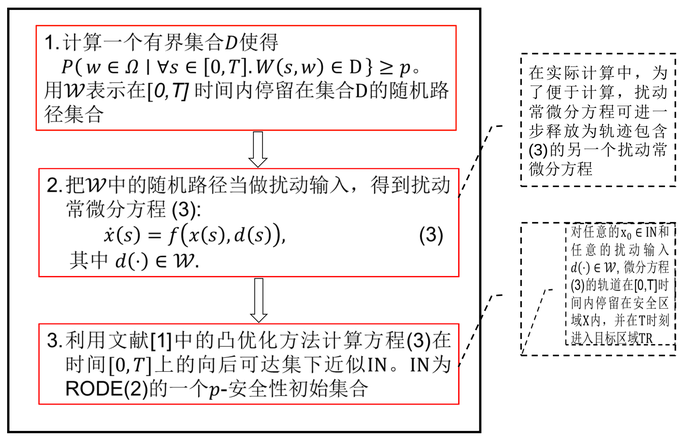
p-安全性问题：计算有限时间和无限时间上的p-安全初始状态集合。

3. 解决方案

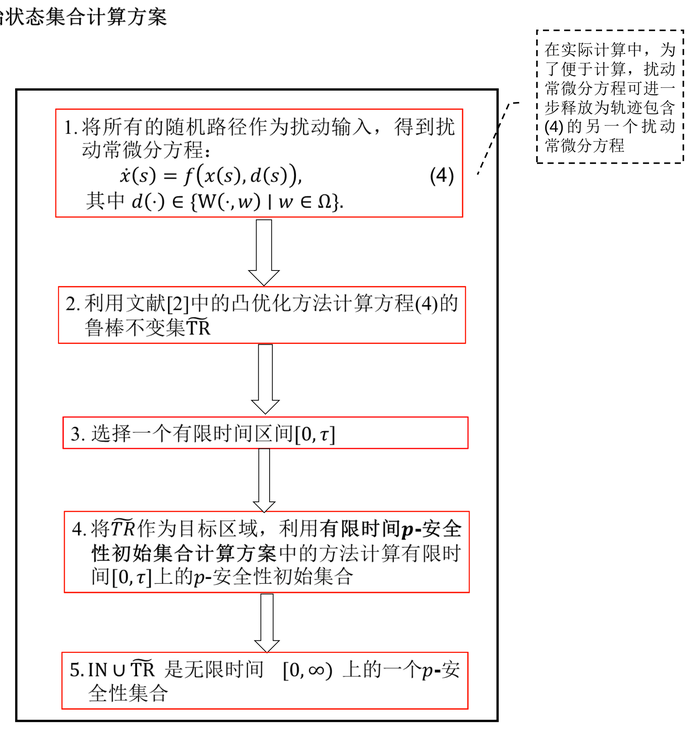
总的方案：将p-安全初始状态集合计算问题转化为扰动常微分方程的鲁棒可达集下近似问题。如下图所示。



1）有限时间**p**-安全初始状态集合计算方案



2）无限时间**p**-安全初始状态集合计算方案



4. 相关文献

[1] Bai Xue, Martin Fränzle and Naijun Zhan. Inner-Approximating Reachable Sets for Polynomial Systems with Time-Varying Uncertainties. IEEE Transactions on Automatic Control, 65(4): 1468-1483, 2020.
[2] Bai Xue, Qiuye Wang, Naijun Zhan and Martin Fränzle. Robust invariant sets generation for state-constrained perturbed polynomial systems. HSCC 2019, pp. 128--137, 2019