Towards high confidence medical device software

Zhihao Jiang, Houssam Abbas, Kuk Jin Jang and Rahul Mangharam

Department of Electrical and Systems Engineering, University of Pennsylvania, USA

1 The Promise and Challenges of Medical Devices

Medical devices play an essential role in the care of patients around the world, and can have a life-saving effect. To cite one example, an estimated 3 million peo- ple worldwide have implanted pacemakers (a heart rhythm adjustment device), with ˜600,000 added annually. In the US, 8˜00,000 people have an implanted defib- rillator (another heart rhythm management device), with 10,000 added monthly. Clinical trials have presented evidence that patients implanted with defibrillators have a mortality rate reduced by up to 31%.

Financially, the medical device market is worth $289 billion. Of that, $110 bil- lion is in the US alone, with this number pro jected to reach $133 billion in 2016. Examples include everything from adhesive bandages to drug infusion pumps, surgical robots, deep brain stimulation systems and devices still undergoing basic research like the artificial pancreas. These are safety-critical technologies com- bining hardware and software, each of which must be rigorously verified to be efficacious and safe.

According to the US Food and Drug Administration, in 1996, 10% of all medical device recalls were caused by software-related issues. This percentage rose to an average of 15% of recalls from 2008 to 2012. Implanted cardiac pace- makers and defibrillators have approximately 80,000-100,000 lines of software code [3] which essentially makes all sensing, control and actuation decisions autonomously within the human body, over the 5-7 year device lifetime. The primary challenge of high-confidence medical device software is to guarantee the device will never drive the patient into an unsafe condition even though we do not have complete understanding of the physiological plant.

1.1 Human-in-the-loop Medical Devices

We may classify devices as open-loop or closed-loop. A closed-loop device like

a pacemaker is in a feedback loop with the organ(s) it effects (see Fig.1): it monitors certain physiological variables like heart rate, and delivers therapy,

in the form of low-energy electrical pulses, to maintain a healthy heart rate. Another example is the artificial pancreas, which monitors blood glucose levels and delivers therapy, in the form of insulin, to maintain safe glucose levels. An

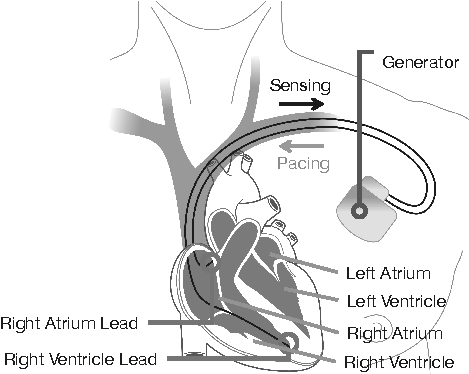
open loop device on the other hand, either, (a) like a drug infusion pump, does not

measure any physiological variables: the therapy it delivers is pre-programmed and non-reactive; or (b) only measures physiological signals but does not deliver

therapy, as in the case of a blood pressure monitor.

Closed-loop devices require very little physician intervention after the dis- charge visit, and hence permit a better lifestyle. Because they are constantly monitoring the physiological variables, they permit a more timely delivery of therapy. The complex run-time diagnoses needed for closed loop performance, and the intricate therapy delivered, has driven most diagnosis and therapy func- tions into software. This software is life-critical and verification methods should provide a high confidence in its correctness.

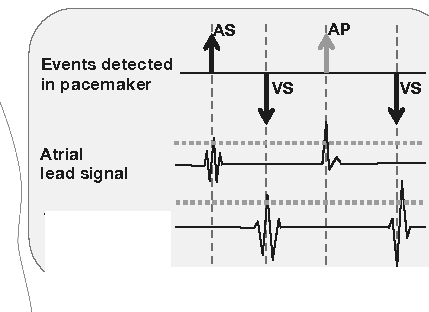
Pacemaker Pulse



**Ventricular**

lead signal

Fig. 1. Pacemaker operating in a closed-loop with the heart. The leads sense cardiac electrophysiological activity from inside the heart tissue (AS/ VS = Atrial/Ventricular Sense event) and actuate the heart (AP/ VP = Atrial/ Ventricular Pacing event



to maintain a desired heart rate.

Verifying the safety and efficacy of closed-loop software, by definition, re­ quires that the device be connected to the organ(s) it is affecting. For example, in the case of a pacemaker, that would be the heart of a living patient. But with the advent of computer models of physiological functions, such as those encom­ passed by the Physiome project or presented later in this article, the *model-based design ( MBD) of closed-loop medical devices* presents efficient complementary approaches that are actively researched in various disciplines of engineering and computer science. In MBD, the device (or a model thereof) is connected to a *model* of the "physiological plant" it interacts with. By high confidence verifi­ cation, researchers mean that under all possible behaviors of the physiological models, the device will act correctly.

There are two major differences between modeling physiology and model­ ing man-made systems: first, physiology is much more complex and less well­ understood than man-made systems like cars and airplanes, and spans several scales from the molecular to the entire human body. Secondly, the variability be­ tween humans is orders of magnitude larger than that between two cars coming off the assembly line. Using the pacemaker as an example of closed-loop device, and the heart as the organ to be modeled, we present several of the challenges and early results in model-based verification.

2 Life-critical Closed-loop Software

The heart is a specialized muscle that pumps oxygenated blood to the rest of the body. It is composed of four chambers: two upper chambers called the left and right atrium, and two lower chambers called the left and right ventricle, are synchronized respectively. In a healthy resting adult, the heart rate is 60 to 100 beats per minute (each ventricular contraction is a beat). The contractions of the heart are controlled by the waves of spontaneous electric depolarization that

traverse it regularly. A spontaneous electric current originates in the Sino-Atrial

(SA) node in the right atrium and propagates throughout the atria, causing

them to contract. It then propagates down to the ventricles along well-defined conduction pathways, causing the ventricles to contract in turn. The SA node is thus termed the natural pacemaker of the heart.

Under certain diseased conditions, the heart rate drops below what is needed to maintain adequate blood flow to the body. This clinical condition is called

bradycardia. When such a heart rate drop is due to abnormalities in the elec-

trical conduction system, an implanted pacemaker might be recommended as treatment. A pacemaker is implanted near the left collar of the patient as shown in Fig.1, and has two leads: one connects to the right atrium, the other to the

right ventricle.

The leads act as both sensors and effectors: if the pacemaker senses a missing electric spike on either lead, indicative of a missed contraction, it will send an electric signal to the corresponding chamber to provoke contraction, thus acting as an artificial pacemaker. The algorithms for detecting missed beats are complex and implemented in software which runs on the pacemaker itself. Part of the difficulty of performing that detection comes from the great variability in heart rates between patients and indeed within a single patient across time. Moreover, because the pacemaker is limited to sensing electrical activity through its two leads, different phenomena can manifest themselves identically to the pacemaker, thus making detection even harder.

In Endless Loop Tachycardia (ELT), this ambiguity causes the pacemaker

to actually induce dangerously elevated heart rates (tachycardia), which would

not have arisen had the heart been operating on its own. This is an example

of a adverse closed-loop condition : a dangerous situation that arises as a result

of the interaction between heart and device. No amount of open-loop device

testing and verification can reveal this condition - hence the need for closed-loop

validation of medical devices, and for physiological heart models that enable early and affordable closed-loop validation.

3 Choosing the right model for the job

Heart models of different kinds have been developed for a range of applications and Fig. 2 shows four heart modeling approaches that emphasize the electric, mechanical, cellular and fluid flow aspects of cardiac function. Several of these modeling approaches employ over 4 million finite elements or 100,000 ordinary differential equations to describe the dynamics and take several hours to simulate a single cardiac cycle.

Cellular models describe the generation and spread of electrical action po-

tentials (i.e., voltages) at the moleculo-cellular level [5]. At the cellular level, the flow of charged ions into and out of the cardiac cell is responsible for the change

in voltage across the cell membrane. Cellular models of electrical activity are used to study how activity across ion channels affect the relation between elec- trical and mechanical behaviors of heart tissue, as well as to study drug therapies that affect the ion channels properties.

Anatomical models are developed using imaging technologies like MRI, and

seek to re-create detailed anatomical structures like fiber orientations and the distribution and extent of scar tissue. These structures affect the heart’s oper- ation by determining muscle contraction and modifying the speed and paths of electrical conduction throughout the heart. Thus anatomical models provide a foundation for whole heart modeling efforts that we cover next. They are also used to simulate the effects of certain medical devices like stents and artificial valves.

Whole heart models use a continuum approximation of the cellular models

of electrical propagation with the structure obtained from anatomical models. Researchers have developed Partial Differential Equations models of electrical

activity in the whole heart to analyze the mechanisms of various arrhythmias.

Researchers at Johns Hopkins [6] further used these models to predict the onset

of arrhythmias and propose potential therapies. Such electro-mechanical models

help evaluate the mechanical effects of different arrhythmias on blood flow.

Electro-Physiological (EP) heart models help study the timing properties

of the generation and propagation of electrical signals through the electrical conduction system, and can accurately diagnose most arrhythmia. Unlike the previous heart models which are built from the cellular level on up, EP models

are developed by conducting a clinical EP testing procedure, and are amenable

to model checking [1], a powerful verification technique pioneered in the semi-

conductor industry. EP testing is a common method to diagnose arrhythmias: a physician inserts catheters with electrodes into the patient’s heart through the veins and measures the local electrical activity around the electrodes. The

physician uses the patterns of electrical activity and its timing characteristics to

diagnose the heart’s condition. In particular, electrical timing parameters of ac- tion potentials in a tissue region like conduction delay, rest period and refractory period are measured, and any abnormal conduction paths are detected.

In [2], researchers from the University of Pennsylvania developed a heart

model based on clinical EP testing. Since the pacemaker only looks at the timing

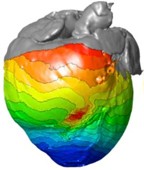
of events as input, this EP model only seeks to model the correct timing of

**Cellular Model**



**Anatomical Model**

**Electrophysiology Model**



**Electro-mechanical Model**

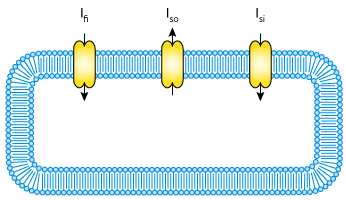
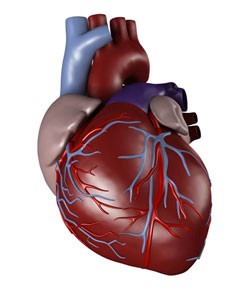
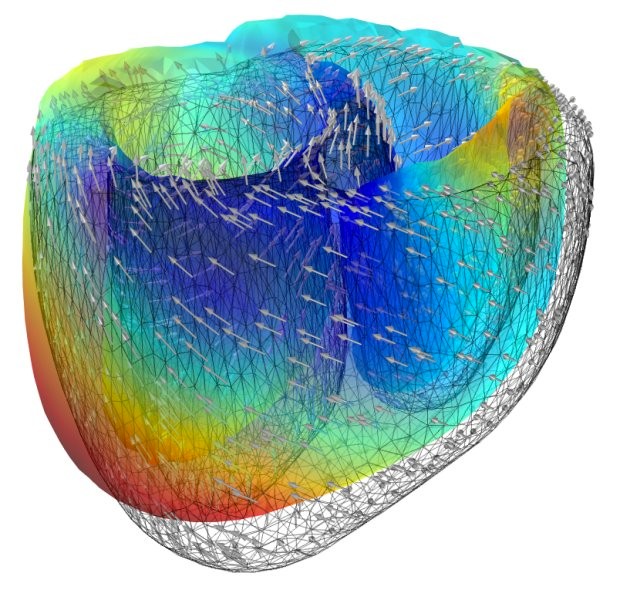


Fig. 2. Modelling different phenomena in the heart

electrical activity in select tissue of the heart. Specialized tissue like the SA node generates electric events and is modeled by a finite timed automaton known as

node automaton. The node automaton models the timing of signal generation,

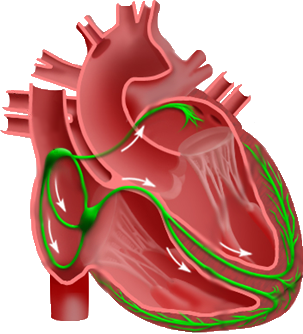
blocking and transmission in heart structures like the AV node. The rest of the tissue is abstracted as variable conduction delays as path automata between node automata. Different heart conditions can be modeled by varying the number and connectivity of note/path automata and their parameters (Fig. 3.(b)). Moreover, pacing applied to the heart can be represented as an external activation signal to the node automata. EP models have been validated by physicians, and have been used for model checking pacemaker software.

**AS AS**

**SA node**

**AV node**

**VS VS**



**His-Purkinje System**

**(a) (b) (c)**

Fig. 3. Timed automata (b) capture the timing relations between electrical events in the heart, and can produce sequences of sense events (c) to the pacemaker.

4 Closed-loop model checking of device software

Short of clinical trials (which we cover later in this article), current validation practice for closed-loop medical devices focuses on open-loop testing and reviews of the design process. In testing, a set of input sequences is fed to the device, and the device’s output is checked for correctness, typically by comparing it to a pre-defined expected output. Such open-loop testing does not test the effect of the device on the organ: e.g., we can’t test how the heart rate changes following a pacing by the pacemaker. Thus, we need a heart model that can interact with the device. If we use, say, a high-fidelity PDE-based model such as the electro- mechanical models described above to react to the device and generate input sequences, there are an infinite number of heart rhythms that such a model can generate, and testing only uses a finite subset of those. Thus, testing (whether open-loop or closed-loop with a heart model) is necessarily an incomplete tech- nique, and safety violations of the device may be missed during testing.

The timed automata-based EP models are amenable to model checking, a

technique that mathematically explores all possible executions of the heart

model and device software combination against specified requirements (e.g. the pacemaker will not pace the heart beyond an upper rate limit). Model checking is widely used in the semiconductor industry to verify chip designs at various levels of abstraction, in particular at the Register Transfer Level (RTL). Violations of the requirements are returned by the model checking tool as an execution trace, which can be analyzed and used to improve the system. To capture the variabil- ity in the heart’s behavior (more generally, in the physiological phenomena of

interest), the heart model is non-deterministic: for example, rather than speci-

fying that the conduction delay in the AV node is always 0.14ms, we allow it to be any value in the correct physiological range [0.12, 0.2] ms. The model checker will symbolically explore all executions corresponding to all values in this range

(rather than select a few) in search for requirements violations. Subtle errors in the design of safety-critical systems that often elude conventional simulation and testing techniques can be (and have been) found in this way. Because it has been proven cost-effective and integrates well with conventional design methods, model checking has being adopted as a standard procedure for the quality as- surance of automotive and avionics systems, but has yet to enter the world of medical devices.

Endless Loop Tachycardia (ELT) is one example of a safety hazard that arises in the interaction between pacemaker and heart, shown in Fig. 4. The ELT starts with an early ventricular contraction (PVC), which is a common scenario even in a healthy person. The electrical signal travels from the ventricle to the atrium (red arrows in Fig. 4), triggering an atrial sense (AS), i.e. the pacemaker senses an event in the right atrium. As a result, the pacemaker paces the ventricle (VP) after a pre-programed delay (AVI), which triggers ventricle to

atrium conduction again and this VP → AS → VP loop persists. The ventricular

rate during ELT is determined by the conduction delay from the ventricle to the atrium and the programed delay in the pacemaker, which is very fast. The

healthy condition in Fig. Fig. 4 demonstrates the same input-output sequence

**ELT**

**A channel**

**V channel**

PVC

AS AS AS AS

**Events**

VS

**Healthy**

**A channel**

**V channel**

AVI

AVI AVI

VP VP VP

9

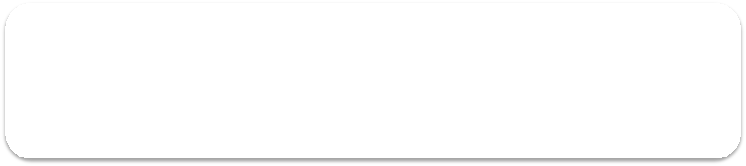
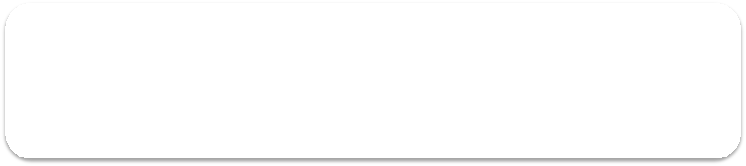


Fig. 4. Endless-loop Tachycardia (ELT) and a healthy heart condition mapped to the

same input-output execution of the pacemaker (middle sequence). The heart model

should have the details to resolve this ambiguity.

as ELT. In it, the pacemaker paces the ventricle after each atrial event (AS) generated by the SA node (as opposed to an AS conducted from the ventricle as in ELT). This is correct pacing at an appropriate rate and maintains adequate blood flow. The heart model should be able to generate these two conditions to assess the pacemaker’s ELT elimination capabilities and that they do not affect other healthy heart functions.

5 Multiscale Modeling

From the example above, we can see that in order to perform model checking on the closed-loop system, the heart model should not only cover all possible inputs to the pacemaker specified in the requirement, but also have enough details to resolve ambiguities of executions that may introduce false-positives and false-negatives. Fig. 5 shows a collection of heart conditions modeled by the EP heart models. These models are definitely not exhaustive. Model checking the pacemaker model with all these heart models will not guarantee absolute safety. By using physiological abstraction rules (R1-R7 in Fig. 5), the heart conditions can be generalized and expand the possible inputs to the pacemaker. The result

is a heart model Hall with only two node automata correspond to the inputs

to the pacemaker. By allowing both node automata to be able to send inputs

to the pacemaker [0, ∞]msec after the last input, the heart model Hall covers

all possible inputs to the pacemaker. However, Hall cannot distinguish the ELT

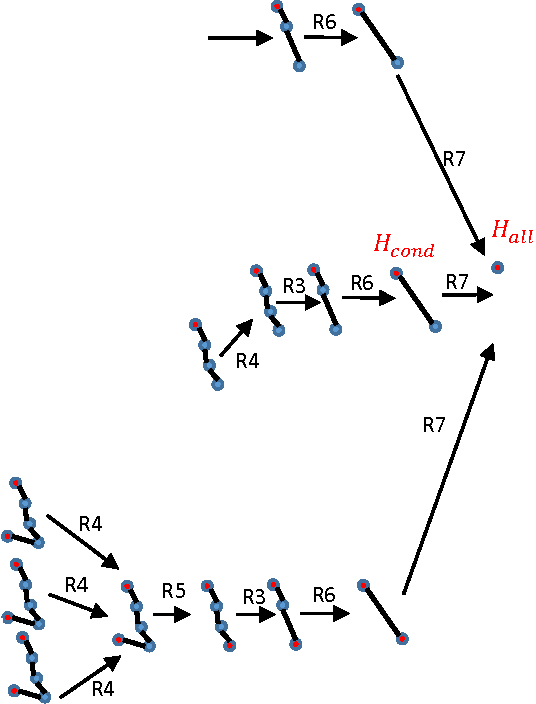
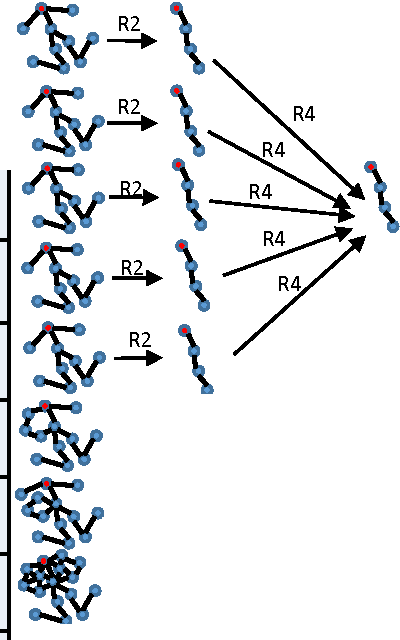
condition from the healthy condition due to the lack of representation of ventricle

to atrium conduction. The heart model Hcond models the conduction between

the atrium and the ventricle with a path automata, thus is the appropriate heart model to evaluate ELT. Similarly, more complex properties will lead to appropriately detailed models along the Model Abstraction Tree.

**Heart**

**Conditions**



Normal Sinus Rhythm

Bradycardia

**1---------1**

AV Block

Bundle Branch Block

Sinus

Tachycardia

Atrial

Flutter

AVNRT Atrial

Fibrillation

Premature

Ventricle

Contraction

$

R3

!'!o.$-"\

$"\:\ •

$

$--R2.\_..

-V-en-tr-icle---i\_ - Tachycardia

**V 4V**

1-

ibrillation · **...!j,V**

L-Ven-tri-cle . \_

**Fig.** 5. Multi-scale modeling of the heart. The heart model at a higher level (further to the right) contains all possible inputs to the pacemaker from the heart models at previous levels

Model Checking

Simulation-based Testing

Platform Testing

•  Non-deterministic UPAAL Model

•  Abstract Heart and Device models

•  Deterministic Simulink Model

•  Increased complexity and interactions

• Hardware implementation of verified models

• Heart-on-a-Chip Platform

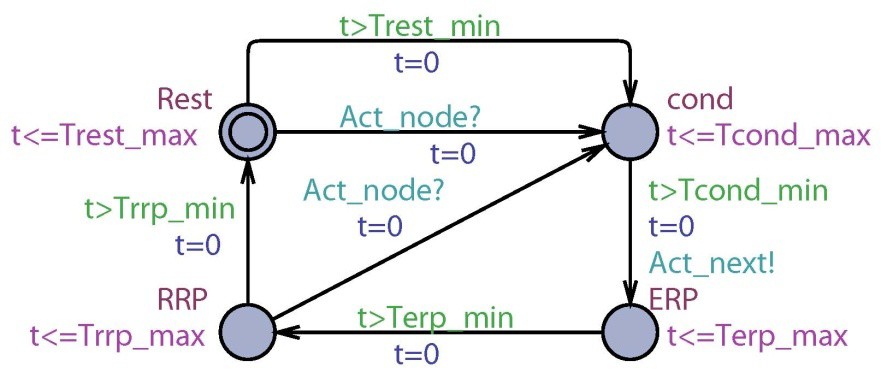
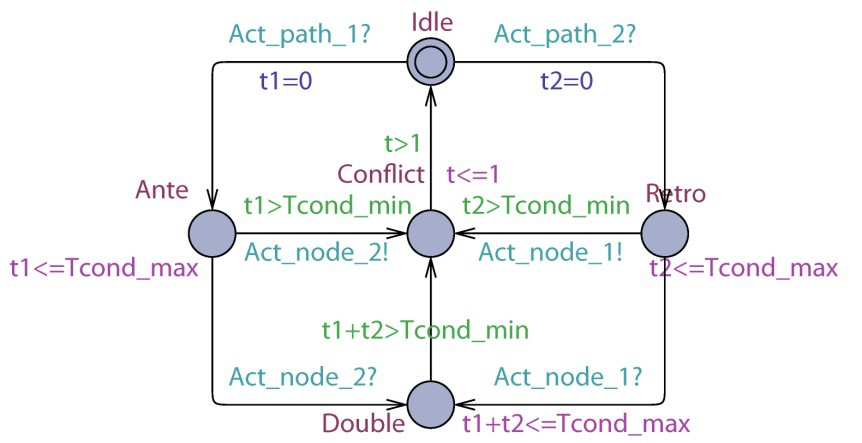
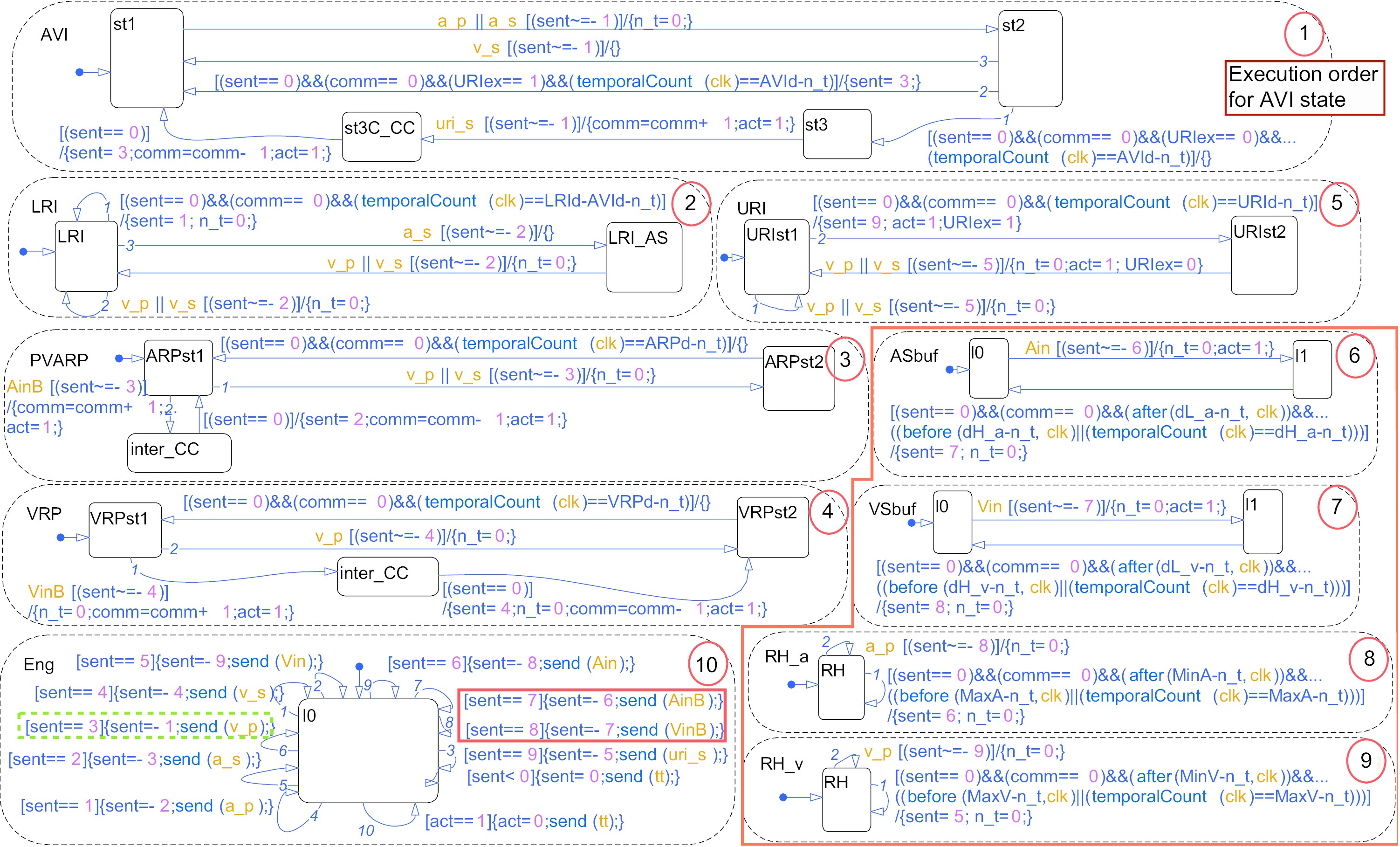
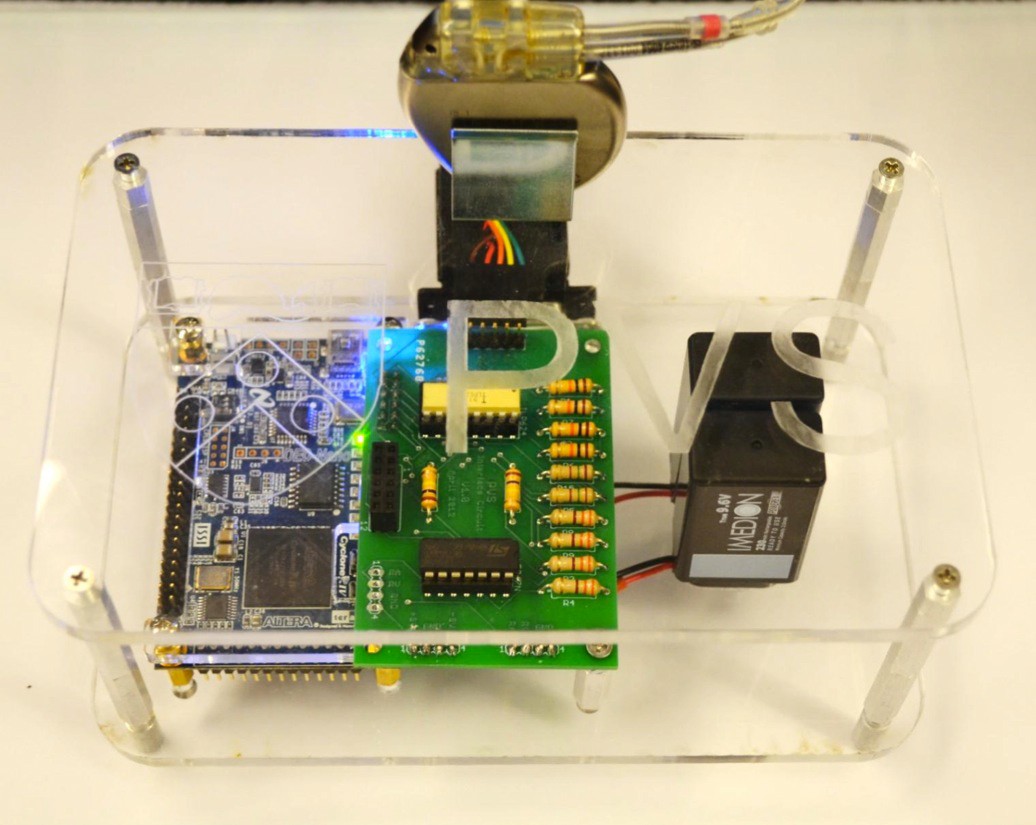


Fig. 6. Model translation framework. The pacemaker design is verified using model

checking and automatically translated into code implementation. Heart models are

available at all levels

6 From Verified Model to Verified Code

During model checking, the abstract model of the pacemaker is verified against safety requirements. The abstract pacemaker model is then automatically syn- thesized into simulation models and into a code implementation (Fig. 6) us- ing the UPP2SF tool we developed. This automatic synthesis provides rigorous traceability throughout the development process and ensures that the verified model is translated into verified code. Similarly the heart model is translated from timed automata to Simulink and also synthesized into a Heart-on-a-Chip for platform level testing.

7 Model-Based Clinical Trials

The final step before the introduction of a medical device to market is usually the

clinical trial. The randomized clinical trial (RCT) is considered to be the “gold

standard” for guaranteeing that a medical intervention is safe and efficacious

[4]. In the case of new medical devices classified as “significant risk”, an RCT is even mandated by the FDA before allowing the device on the market. There are many variations of RCTs, but they all share a basic outline, an example of which we now illustrate.

Suppose that a manufacturer of medical devices is designing a new pacemaker

that’s supposed to assist in treating certain abnormal cardiac rhythms, or ar-

rhythmias. Both the hardware and software are tested by the company to ensure

it satisfies certain specifications, perhaps using model-based methods outlined above. The device may then be implanted and tested on animals. But up to this point, the effect of the device on humans has not been observed. Observations of interest are not merely restricted to whether the device operates as intended or not. We want to know (and regulators require evidence of ) whether it can be implanted safely; whether it has unexpected side effects, or whose magnitude is too large; and, very importantly, whether it treats the targeted arrhythmias better than current medical care, thus justifying its release on the market.

The RCT seeks to answer these questions by comparing two groups of pa-

tients: one which is implanted with the investigational device, aka the treatment

group, and one which is on standard medical care, aka the control group. The

assignment of a patient to the treatment or control group is done at random,

which guarantees the validity of the statistical tests used to analyze the results,

and helps ensure that the two groups are comparable in terms of the physiolog-

ical factors that might affect the outcome of the intervention. Both treatment and control groups are then monitored for a pre-determined amount of time,

at the end of which the rate of treated arrhythmias is evaluated in each group (this is the so-called clinical outcome of the trial). Finally, statistical methods like the t-test are applied to determine whether the difference in rates between

the groups, if any, is significant, i.e. is unlikely to be due to chance alone.

The recognized superiority of RCTs comes at a cost: an RCT is in general a ma jor effort involving patients, medical investigators, ethics boards, biostatis-

ticians, regulators, companies, clinical centers, and large sums of money easily running in the millions of dollars. In addition, technical errors can arise at almost every step of the trial planning, jeopardizing the validity of the results. Finally, even if the trial is well-planned, poor execution, unexpected events or even just pure chance can lead to the wrong conclusions.

The application of computer models to the medical domain presented above have largely centered on the design, verification and deployment of a given device instance, and have mostly eschewed matters related to the clinical trial. There is however now an opportunity to use these computer models to assist in the planning and conduct of RCTs (see, e.g., the Avicenna pro ject). We call this the Model-Based Clinical Trial (MBCT). Broadly speaking, we define an MBCT to

be a trial in which the sub jects are computer models of the biological phenomena

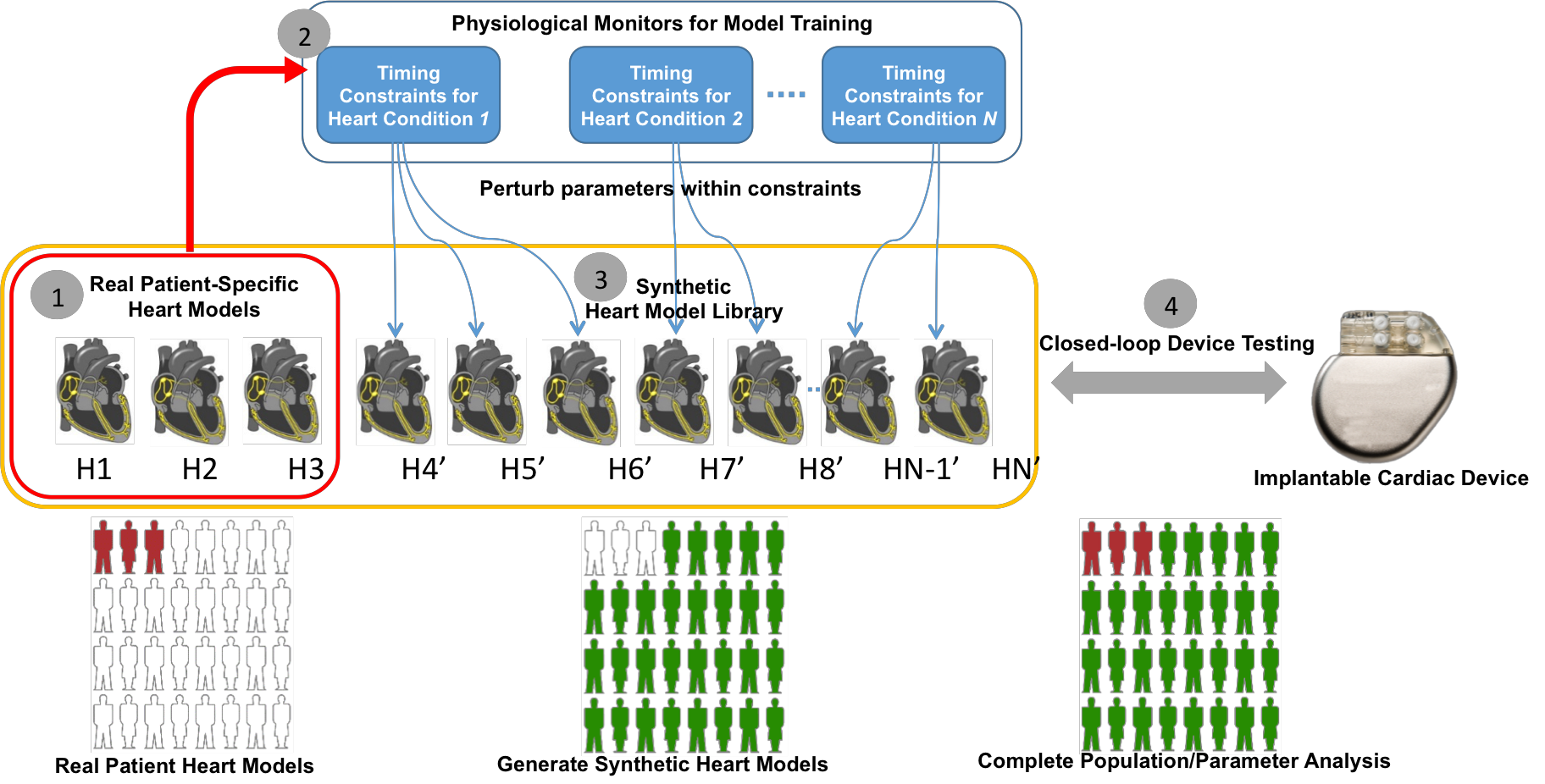


Fig. 7. Model-based Clinical Trials

being studied, including the effect of the device, rather than humans. An MBCT

is not a replacement for a clinical trial: rather, it will allow us to run very large-

scale targeted simulated trials to better inform our conduct of an actual RCT. For example, we can study how variations in a patient’s physiological parameters, like speed of propagation of the electric current in the heart, affects the safety and efficacy of the device. This is doable in MBCT and provides valuable insight into which patients should be enrolled in the trial (and in whom the device is most efficacious). Another application would be to get tighter estimates of statistical quantities like effect size needed before the conduct of the trial. Unlike drug trials, the results of medical device trials depend on the skill of the physician operating the device or implanting it. An MBCT that models the effects of physician errors, like added noise on dislodged pacemaker leads, could inform the trial investigators as to how much training is necessary for the physicians involved in the trial.

For an MBCT, it is not sufficient to only validate that the model structure

can produce physiologically correct behavior. The model cohort as a whole must

also present the right variability to effectively be treated as a group of patients. (The model cohort is the group of models enrolled in the trial). Our current approach is to learn the constraints on parameters for the parameterized model

using a sample of real patients’ data, as indicated by markers 1 and 2 in Fig.

7. Ideally this sample of patients is a cohort from a previous trial. These learned

constraints are then used to generate more instances of the model (marker 3 ), and this constitutes our model cohort. The device is then connected to these models (marker 4 ) and the outcomes of interest are evaluated, such as incidence of adverse events.

Early efforts tying physiological modeling to clinical trials also include the

UVA/PADOVA diabetes model which is used to generate simulated patients.

New diabetes control algorithms are evaluated on simulated patients instead of animals to assess their efficacy. These applications of MBCT usher a new era of exciting research challenges for the closed-loop verification, validation and testing medical devices at scale. MBCT have the potential to reduce the scope, cost and probability of failure of clinical trials of medical devices with complex hardware and software. MBCT when viewed as a rapid certification toolchain to speed up medical device approvals is gaining increasing traction both within the regulatory environment and the medical device industry.

References

[1] E. Clarke, O. Grumberg, and D. Peled. Model Checking. MIT Press, 1999.

[2] Z. Jiang, M. Pajic, and R. Mangharam. Cyber-Physical Modeling of Im-

plantable Cardiac Medical Devices. Proceedings of the IEEE, 100(1):122

–137, Jan. 2012.

[3] P. L. Jones. Senior Systems/Software Engineer, Office of Science and Engi-

neering Laboratories, US FDA. Personal communication, 2010.

[4] L. M. Friedman and C. D. Furberg and D. L. DeMets. Fundamentals of

Clinical Trials. Springer, 2010.

[5] A. Pullan, L. Cheng, and M. Buist. Mathematically Modelling the Electrical

Activity of the Heart: From Cell to Body Surface and Back Again. World

Scientific, 2005.

[6] N. A. Trayanova and P. M. Boyle. Advances in modeling ventricular ar-

rhythmias: from mechanisms to the clinic. Wiley Interdisciplinary Reviews:

Systems Biology and Medicine, 6(2):209–224, 2014.