RESEARCH (YEAR 2)

System Design is a widely used methodology concerned with aiding the prototyping process in order to reduce costs and time. It has several names depending on the field of application such as System Engineering, Model Based Design etc, but the underlying idea is always the same. Another useful methodology which is closely linked with the System Design is the *System Identification*. They are different faces of the same medal. The common idea is to build a mathematical object/computational model, called (in silico) model, which is able to simulate the behavior of a real system in order to describe it (System Identification) or to build it (System Design) [3].

My research is focused on System Design problems aimed at estimating the parameters of a target model so that it satisfies some requirements/specifications. That methodology is commonly used in Synthetic and System Biology where the systems (and therefor the models) are mainly stochastic and highly noisy, or to describe Cyber-Physical-Systems [10, 11] or more generally complex systems. The use of quantitative formal methods in System Design framework has increased rapidly in the last few years, starting from Biology and continuing in the industrial field such as aerospace and automotive sectors, where the quality and compliance of the end product are of paramount importance. In the automotive field, in particular, the models are described by using block diagram tools (i.e Simulink, Gt Suite) which are mainly deterministic and highly complex to verify. In this research I have considered both the stochastic and the deterministic cases. In the first case Chemical Reaction Networks (CRNs, [4]) have been considered and the requirements have been expressed by using Signal Temporal Logic (STL, [5]). The goal was to maximize the probability of satisfaction of target formulas. The reason for considering such Temporal Logic Formalism to describe the requirements is twofold. Firstly, this formalism is able to capture the emergent properties of a system, which is a natural approach from the modeling point of view. Secondly, it is suitable for expressing the requirements in a rigorous way, so that computers can directly interpret them. Several model checkers capable of verifying different temporal logic formalism exist.

In [1] a multi-objective approach to the design of CRNs using STL has been considered. It is quite common in the System Design Process to have conflicting requirements, where conflicting means that increasing the satisfiability of a given requirement entails the decrease of another. The multi-objective approach manages that situation by introducing a very natural relation of dominance. One model is preferable to (dominates) another one if its performance with respect to each requirement is better; otherwise the two models are in a non-dominated relation. The multi-objective approach has the goal of finding all the non dominated points that compose the so called Pareto Front (see. [6]). After having proved (by means of simulations) that this approach is reasonable for the CRNs, a combined approach of quantitative and qualitative semantics to find the Pareto front has been considered. Using that combination we have defined a new relation of Pareto dominance and the results show that this definition can speed up the performance of a genetic algorithm to reach the Pareto front in terms of less evaluations. An interesting outcome of this research consists in showing the limitation of the robustness semantics that we call *the length scale problem*. The criticism is referred to its sensitiveness with respect to the length scale of the atomic predicates. In fact, the robustness semantics combines the values of predicates contained in a formula (i.e a conjunctive clause) in a way that may not be fully meaningful if their robustness length scale is different. This entails that the falsification algorithm considers only the predicates with larger robustness length scale and ignores others producing sub-optimal results, (see [1]).

The other part of my research has been focused on the falsification of block diagram models (such as Simulink) which are largely used in control applications (automotive, aerospace, etc). They are difficult to verify because of their complexity (consider that a basic industrial model could contain several switch blocks, look-up tables and encapsulated hybrid systems). For this reason standard model checking techniques are most of the time unfeasible. Those kinds of model take as inputs time varying quantities (such as temporal series i.e throttle and brake angle dynamics) and produce as outputs temporal series (such as engine speed, gear dynamics etc). The falsification problem consists in finding an input, if any, which causes a bug in the target model, expressed as falsification of a given STL formula. That input is called counterexample. It is possible to determine counterexamples by means of optimization where the simple idea consists in minimizing the robustness of the STL specification by changing the inputs temporal series and stopping the process as soon as a counterexample has been found (it will have negative robustness). We have tackled this problem by handling the model as a black box, meaning that is not possible to look inside the model but it is only possible to observe input and output temporal series. On the one hand, this choice is a limitation because no abstraction or model transformation is possible and this increases the difficulty of falsifying the system. On the other hand, the general applicability of the proposed process increases enormously because of its independence of

the target model. The main difficulties here consist in efficiently parameterizing the input temporal series and considering optimization algorithms which are able to find a counterexample with the lowest number of evaluations. In [2] we have modified the fixed control point parameterization [8] and used a Gaussian Process Upper Confidence Bound (GP-UCB, [9]) optimizer in order to overcome the two aforementioned difficulties. The results in terms of the minimum number of evaluations needed to find a counterexample are good if compared with the state of the art S-Taliro, [8]. The other advantages of using Gaussian Processes (or other statistical techniques) is their ability to provide statistical guarantees about the non falsifiability of black box models which admit inputs in a countable set. In fact, the falsifiability of black box models is a semi-decidible problem given that no abstraction techniques can be applied. The GPs can link the number of failed attempts at falsifying a model with the probability that this model is not falsifiable. This statistical property is closely related to the assumption of regularity we are making about the black box model, or similarly with the relation between the variation of the robustness semantics with respect to the variation in the input signals. For example, in the automotive field where Simulink is widely used, even if the models are highly complex we know that for sufficiently close input signals the system will generate outputs with similar robustness. The input functions which violate that assumption will be of measure zero (after having defined an appropriate measure).

REFERENCES

[1] Bortolussi, Luca, et al. Logic-Based Multi-objective Design of Chemical Reaction Networks, *International Workshop on Hybrid Systems Biology*, Springer International Publishing, 2016.

[2] Silvetti, Simone and Mariapia Marchi Validation of Automotive Control Applications using Formal Methods and metamodeling techniques, *Proceeding of the International CAE Conference*,2016.

[3] Bortolussi, Luca, and Guido Sanguinetti. Learning and Designing Stochastic Processes from Logical Constraints *QEST*. Vol. 8054. 2013.

[4] Gillespie, Daniel T. Exact stochastic simulation of coupled chemical reactions. *journal of physical chemistry* 81.25 (1977): 2340-2361.

[5] Maler, Oded, and Dejan Nickovic. Monitoring temporal properties of continuous signals. *Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems*, Springer Berlin Heidelberg, 2004. 152-166.

[6] Deb, Kalyanmoy. Multi-objective optimization using evolutionary algorithms. Vol. 16. John Wiley & Sons, 2001.

[7] Bardh Hoxha, Houssam Abbas, and Georgios Fainekos. Benchmarks for temporal logic requirements for automotive systems. *Proc. of Applied Verification for Continuous and Hybrid Systems*, 2014.

[8] Fainekos, Georgios E., et al. Verification of automotive control applications using s-taliro. *American Control Conference (ACC)* , IEEE, 2012.

[9] Srinivas, Niranjan, et al. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory 58.5*, 2012: 3250-3265.

[10] Lee, Edward A. Cyber physical systems: Design challenges. *11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, 2008.

[11] Jensen, Jeff C., Danica H. Chang, and Edward A. Lee. A model-based design methodology for cyber-physical systems. *7th International Wireless Communications and Mobile Computing Conference*, IEEE, 2011.