University of Florence

PhD in Smart Computing
XXXII cycle

Progress Report

first year

Tommaso Papini
tommaso.papini@unifi.it
DT21263

Research topics: Model-based quantitative analysis for on-line diagnosis, prediction, scheduling and compliance evaluation in partially observable systems
Advisor: Prof. Enrico Vicario
Supervisory committee: Dr. Mieke Massink, Prof. Mirco Tribastone

Research and results

In this section, the main research conducted and the most relevant results will be shown. The main topic of the PhD research is that of model-based quantitative analysis, especially in the scenario of partially observable systems.

Before the start of the first year of the PhD, a period of five months as a Research Fellow at University of Florence has been conducted, during which, under the supervision of Prof. Enrico Vicario, research activity has been started, following the same topics. In particular, during this period we produced the conference paper [17], which focused on the performance evaluation of a mutual exclusion protocol (the Fischer's protocol) exploiting a technique for steady-state evaluation of Markov Regenerative Processes (MRP) [16]. Part of this work was also the implementation of the steady-state technique for MRPs described in [16] exploiting the APIs of the Oris tool [7].

The PhD period started with the investigation and implementation of a technique for the transient analysis of MRPs under *enabling restriction*, which characterises all those MRPs that have, at any given time, at most one GEN (GENerally distributed transition) . In particular, the technique implemented has been studied from [12] and implemented through the Oris tool APIs.

During the first year of the PhD the LINFA (Logistica INtelligente del FArmaco) project has also been followed. The LINFA project is a regional project funded by the Tuscany region that aims to develop a software for decision support for hospital staff members for drugs restocking. Drugs restocking can in fact be a hard and expensive process and by exploiting data processing and forecasting technique it can be made easier and cheaper. For this reason, techniques for model-based forecasting and decision support has been investigated. In particular, techniques that exploit Markov Decision Processes (MDP) [2] has been investigated and later implemented through the PRISM model checking tool [15] and a Java framework that generates an actualised PRISM model each time a drugs restock has to be issued, following the idea of *models@runtime* [6].

A compositional technique for transient analysis of MRPs has then been investigated. The idea was to combine both the technique for transient analysis of MRPs under enabling restriction shown in [12] and the technique for transient analysis of MRPs under *bounded regeneration*, which characterises all the MRPs that has no cycles between any two regenerations, shown in [14], exploiting non-deterministic analysis. In particular, this compositional technique would first perform non-deterministic analysis on an MRP for each of its regenerations and classify them depending on which of the two conditions (enabling restriction and bounded regeneration) are satisfied. Depending on the result of this classification, the correct transient technique would then be applied to compute local and global kernels for that specific regeneration epoch, exploiting the fact that kernel rows, corresponding to different regenerative epochs, can be evaluated independently and thus with different techniques. When the whole local and global kernels have been computed, transient solution can be evaluated through the evaluation of Markov renewal equations. In order to evaluate those regenerations where none of the two conditions are satisfied, approximate evaluation has also been studied: the approximate technique proposed is based on the technique for transient analysis based on stochastic state classes [14] and implements a guided transient analysis in order to explore the most "relevant" epochs with a finer grain, while the least "relevant" ones with a coarser grain, focusing the available budget of precision on the most "relevant" ones. This compositional technique has been implemented exploiting the Oris tool APIs. Results of this work, along with experimentation with the implemented technique, have been published in [3].

Another investigation, regarding the analysis of assembly lines, has been pursued. Assembly lines analysis has become more and more important in these last years, in order to exploit techniques of data processing to maximise throughput and efficiency of the assembly lines, for example by dynamically adapting the production during runtime, according to the agenda of

Industrie 4.0 [13]. The scenario investigated is that of an assembly line of sequential workstation, with transfer blocking and no buffering capacity, where each workstation can implement a complex workflow, with sequential, alternative of cyclic phases. At any given time, the assembly line under analysis could be inspected by an external observer, such as a human observer or a polling system, producing an observation where the status of each workstation and the specific phase of the producing ones could be observed. Such an observer, however, would only be able to partially observe the assembly line: firstly because some phases of the same workstation could be similar enough to produce the same observation (ambiguity on the logical state) but also because the external observer has no information regarding the time elapsed in the currently producing phases (ambiguity on the remaining times). In this context, we derived performance measure for the analysis of the assembly line: in particular we derived the *Time To Done (TTD)*, representing the time until a certain workstation finishes working on a product, the *Time To Idle (TTI)*, representing the time until a certain workstation is available to accept a product from the previous one, and the *Time To Start Next (TTSN)*, representing the time until a certain workstation start the production of a new product. Upper and lower bounds evaluation for the CDF (Cumulative Distribution Function) for the three measures was possible in a compositional fashion. In particular, a lower and an upper bounds were computed in a naive way, while a second lower bound was derived in a more educated way thanks to the positive correlation of remaining times in the producing phases and to the Key Renewal Theorem [22]. This work has been implemented through the Oris tool API, thanks to which it was possible to validate the approach through a series of experiments that showed how the proposed approach results more feasible and more scalable than simulation. The detailed work, along with experimental results, have been published in [5].

Lastly, the field of Activity Recognition (AR) [9, 20, 25] for Ambient Assisted Living (AAL) [10] and the field of dataset creation for AAL AR has been investigated. This investigation was part of an European secondment programme called REMIND and has been conducted while at the University of Jaén (UJA), Spain, and it focused on exploring and comparing the different techniques for AAL AR developed by the STLab (Software Technologies Lab) [24] in Florence and by the Sinbad^2 research group [23] in Jaén. In particular, techniques for AAL AR based on stochastic models [4, 8] and on fuzzy logic [18], as well as several works on the creation of datasets for AAL AR [19, 21], have been studied and compared and joint proposals have been suggested in order to take advantage of the different techniques studied.

Courses attended

The following section reports a list of exams passed, seminars, tutorials, or summer schools attended. The Smart Computing PhD programme requires at least 9 credits by the end of the first year and at least 18 credits at the end of the second year.

Exams

- **GPU Programming Basics** (Marco Bertini, UniFi): 3 credits
- **Fuzzy Logic & Fuzzy Systems** (Beatrice Lazzerini, UniPi): 3 credits

Seminars

- **ProPPA: Probabilistic Programming Process Algebra** (Anastatis Georgoulas, IMT Lucca): 1 credit
- **Modelling, analysis and design of cyber-physical systems** (Ezio Bartocci, UniFi): 0.5 credits

Summer schools

- **Summer School on Optimization, Big Data and Applications (OBA)** (Veroli, Italy): 5 credits

Current total credits

The number of current total credits achieved by the end of the first year of PhD is 12.5.

Publications

The followings are all the published papers:

- **Title:** Performance Evaluation of Fischer's Protocol through Steady-State Analysis of Markov Regenerative Processes [17]
  Authors: Stefano Martina, Marco Paolieri, Tommaso Papini, Enrico Vicario
  Conference: Modeling, Analysis and Simulation of Computer and Telecommunication Systems, MASCOTS 2016

- **Title:** Exploiting Non-deterministic Analysis in the Integration of Transient Solution Techniques for Markov Regenerative Processes [3]
  Authors: Marco Biagi, Laura Carnevali, Marco Paolieri, Tommaso Papini, Enrico Vicario
  Conference: International Conference on Quantitative Evaluation of Systems, QEST 2017

- **Title:** An Inspection-Based Compositional Approach to the Quantitative Evaluation of Assembly Lines [5]
  Authors: Marco Biagi, Laura Carnevali, Tommaso Papini, Kumiko Tadano, Enrico Vicario
  Conference: European Workshop on Performance Engineering, EPEW 2017

Conferences and workshops

The followings are all the conferences and workshops attended:

- **International Conference on Quantitative Evaluation of Systems** (QEST 2017), Berlin (Germany), September 5-7 2017
- **European Workshop on Performance Engineering** (EPEW 2017), Berlin (Germany), September 7-8 2017
- **International Workshop on Practical Applications of Stochastic Modelling** (PASM 2017), Berlin (Germany), September 9 2017

Research visits to external institutions

The followings are all the research visits done to external institutions:

- **Destination:** Ageing Lab Foundation and University of Jaén, Jaén, Spain
  Period: May 20th - June 19th 2017
  Research and results: This period was part of an European secondment programme called REMIND and has been conducted while at the Ageing Lab Foundation [1] and at the University of Jaén (UJA), both in Jaén, Spain. The main goal was to explore and compare the different techniques for AAL AR and for the creation of datasets for AAL AR developed by the STLab [24] in Florence and by the Sinbad^2 research group [23] in Jaén. In particular, techniques for AAL AR based on stochastic models [4, 8] and on fuzzy logic [18], as well as several works on the creation of datasets for AAL AR [19, 21], have been studied and compared and joint proposals have been suggested in order to take advantage of the different techniques studied. A final report has been produced on this work, containing a review of the studied methods as well as joint proposals taking advantage of the strong points of the methods shown.

Research plan for the next year

One of the main topics that will be further investigated during the next year is the analysis of assembly line. At the moment, the work produced [5] sets a good base for investigation of analysis techniques for assembly lines, but lacks many aspects in order to be usable with real assembly lines. The first restriction that should be surpassed is the lack of buffers between workstations. In real assembly lines there is usually a buffer of fixed size, so one of the first investigations should be towards the introduction of buffering capacity, considering the same capacity for every workstation or even different sizes. More performance measures should also be introduced, such as the time until the production from the whole assembly line of a specific product in the line, or the time until the next N products will be completed by the assembly line. These kind of measures would be complex than the ones already proposed in [5], but would also allow a more accurate analysis of the assembly line. Evaluation of these additional measures should also be derived in a compositional fashion, as done already in [5], in order to keep the technique computationally feasible. Lastly, a better upper bound for the measures proposed should be derived, similarly to the lower bound derived in [5].

Another topic that will be investigated is that of smart drugs restocking for the LINFA project. The work conducted at the moment is still far from being usable in a real ward scenario. This mainly due the problem of state space explosion, since the problem to be modelled is quite complex and many aspects would need to be modelled, rendering it infeasible to analyse. One of the first directions is thus that of optimising the prediction model itself, in order to model more aspects while maintaining the number of generated states feasible. The PRISM model checker used up until now doesn't provide any solution for state space optimisation, so alternative modelisation tools might be investigated, such as the Storm tool [11]. Then, the restrictions currently present in the model should one by one be surpassed. An example would be to introduce different healthcare protocols for different kinds of patients, including different phases with different drugs consumption rates, thus personalising more the sojourn of patients in the ward. This could be done by employing techniques of process mining [26], allowing to generate models automatically from historical data, such as a dataset of patients from a real ward.

Lastly, the topic of AR, especially in the scenario of AAL, will be further investigated. The idea is to expand on the work of [4, 8] and on the investigation conducted in Jaén in order to enhance the AR model. Among the possible enhancements there is the possibility of adding support for continuous sensors, such as a thermometer or an accelerometer, which is now lacking in the current model. Another important research activity for this topic is that of finding good datasets for AAL AR in order to apply techniques of process mining more accurately and refining the existing model, possibly following the recommendations shown in [19].

References

[1] Ageing Lab Foundation. Ageing Lab Foundation. `http://www.ageinglab.com/`.

[2] Bellman, R. A Markovian decision process. *Journal of Mathematics and Mechanics* (1957), 679–684.

[3] Biagi, M., Carnevali, L., Paolieri, M., Papini, T., and Vicario, E. Exploiting Non-deterministic Analysis in the Integration of Transient Solution Techniques for Markov Regenerative Processes. In *International Conference on Quantitative Evaluation of Systems* (2017), Springer, pp. 20–35.

[4] Biagi, M., Carnevali, L., Paolieri, M., Patara, F., and Vicario, E. A Stochastic Model-Based Approach to Online Event Prediction and Response Scheduling. In *European Workshop on Performance Engineering* (2016), Springer, pp. 32–47.

[5] Biagi, M., Carnevali, L., Papini, T., Tadano, K., and Vicario, E. An Inspection-Based Compositional Approach to the Quantitative Evaluation of Assembly Lines. In *European Workshop on Performance Engineering* (2017), Springer, pp. 152–166.

[6] Blair, G., Bencomo, N., and France, R. B. Models@ run. time. *Computer 42*, 10 (2009).

[7] Bucci, G., Carnevali, L., Ridi, L., and Vicario, E. Oris: a tool for modeling, verification and evaluation of real-time systems. *International Journal on Software Tools for Technology Transfer (STTT) 12*, 5 (2010), 391–403.

[8] Carnevali, L., Nugent, C., Patara, F., and Vicario, E. A continuous-time model-based approach to activity recognition for ambient assisted living. In *International Conference on Quantitative Evaluation of Systems* (2015), Springer, pp. 38–53.

[9] Chen, L., Nugent, C. D., Biswas, J., and Hoey, J. *Activity recognition in pervasive intelligent environments*, vol. 4. Springer Science & Business Media, 2011.

[10] Cook, D. J., Augusto, J. C., and Jakkula, V. R. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing 5*, 4 (2009), 277–298.

[11] Dehnert, C., Junges, S., Katoen, J.-P., and Volk, M. A storm is coming: A modern probabilistic model checker. *arXiv preprint arXiv:1702.04311* (2017).

[12] German, R., Logothetis, D., and Trivedi, K. S. Transient analysis of Markov regenerative stochastic Petri nets: A comparison of approaches. In *Petri Nets and Performance Models, 1995., Proceedings of the Sixth International Workshop on* (1995), IEEE, pp. 103–112.

[13] Hermann, M., Pentek, T., and Otto, B. Design principles for industrie 4.0 scenarios. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on* (2016), IEEE, pp. 3928–3937.

[14] Horváth, A., Paolieri, M., Ridi, L., and Vicario, E. Transient analysis of non-Markovian models using stochastic state classes. *Performance Evaluation 69*, 7 (2012), 315–335.

[15] Kwiatkowska, M., Norman, G., and Parker, D. PRISM 4.0: Verification of Probabilistic Real-time Systems. In *Proc. 23rd International Conference on Computer Aided Verification (CAV'11)* (2011), G. Gopalakrishnan and S. Qadeer, Eds., vol. 6806 of *LNCS*, Springer, pp. 585–591.

[16] Logothetis, D., Trivedi, K. S., and Puliafito, A. Markov regenerative models. In *Computer Performance and Dependability Symposium, 1995. Proceedings., International* (1995), IEEE, pp. 134–142.

[17] Martina, S., Paolieri, M., Papini, T., and Vicario, E. Performance Evaluation of Fischer's Protocol through Steady-State Analysis of Markov Regenerative Processes. In *Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS), 2016 IEEE 24th International Symposium on* (2016), IEEE, pp. 355–360.

[18] Medina, J., Espinilla, M., Moya, F., and Nugent, C. Activity recognition by means of rule-based inference engine based on fuzzy linguistic approach. In *13th Scandinavian Conference on Artificial Intelligence Halmstad* (November 2015 2015).

[19] Patara, F., Nugent, C. D., and Vicario, E. Recommendations for the Creation of Datasets in Support of Data Driven Activity Recognition Models. In *International Conference on Smart Homes and Health Telematics* (2015), Springer, pp. 79–91.

[20] Patterson, D. J., Liao, L., Fox, D., and Kautz, H. Inferring high-level behavior from low-level sensors. In *UbiComp* (2003), Springer, pp. 73–89.

[21] Quesada, F. J., Moya, F., Medina, J., Martínez, L., Nugent, C., and Espinilla, M. Generation of a partitioned dataset with single, interleave and multioccupancy daily living activities. In *International Conference on Ubiquitous Computing and Ambient Intelligence* (2015), Springer, pp. 60–71.

[22] Serfozo, R. *Basics of applied stochastic processes*. Springer Science & Business Media, 2009.

[23] Sinbad^2 team. Sinbad^2. http://sinbad2.ujaen.es/.

[24] STLab team. STLab. https://stlab.dinfo.unifi.it/.

[25] Turaga, P., Chellappa, R., Subrahmanian, V. S., and Udrea, O. Machine recognition of human activities: A survey. *IEEE Transactions on Circuits and Systems for Video Technology 18*, 11 (2008), 1473–1488.

[26] Van der Aalst, W., Weijters, T., and Maruster, L. Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering 16*, 9 (2004), 1128–1142.