Deliverable D2.2

Algorithms for Games on Evolving Structures

Wolfgang Thomas (RWTH),
Axel Haddad (UMONS),
Christof Löding (RWTH), Jiří Srba (AAU)

Project	Cassting — Collective Adaptive Systems SynThesIs with Non-zero-sum Games		
Project id.	FP7-601148		
Workpackage	WP2	**Nature**	R (report)
Deliverable	D2.2	**Dissemination**	PU (public)
Date submitted	March 2015	**Date due**	March 2015
Version	1.0		

;

SEVENTH FRAMEWORK PROGRAMME

Algorithms for Games on Evolving Structures

Wolfgang Thomas, Axel Haddad, Christof Löding, Jiří Srba

March 2015

1 Introduction

The present report is a follow-up of the deliverable “Games on Evolving Structures, Interactions with Open Environments”, submitted only six months ago. Thus the progress on which we report here can only be moderate given the short time that elapsed. However, as indicated by the present title, the focus shifted and now addresses mainly algorithmic questions rather than pursuing the emphasis on conceptual models as in the previous report.

Also it turned out that the view on “evolving structures” should not be restricted to the change of game arenas only (which was the perspective adopted in devising the work package and in the previous deliverables). We realized in our collaborations and in particular in the past two Cassting workshops that a more appropriate view should be more comprehensive and should address “adaptivity” jointly for the game models *and* for the corresponding strategies. Evolving structures in the game theoretic approach are best studied in a way that changes of game arenas and strategies go hand in hand. In the present report we illustrate in Section 4 the aspect of evolvement of strategies, using concepts of learning. This opens a perspective to the above-mentioned more comprehensive view.

The remainder of this report is structured into the following sections:

- Section 2: Quantitative sabotage games
- Section 3: Stability in dynamic betworks
- Section 4: Adaptivity in strategy construction
- Section 5: Dynamic networks of hybrid systems

followed by a Conclusion.

2 Quantitative Sabotage Games

In this subject we continued work that has a somewhat long tradition in Cassting and is now extended by new features.

The original setting as developed by van Benthem [vB05] is given by a finite directed graph G with designated start vertex s and target vertex t. Two players are involved: The player Walker takes a path through G, starting in s and proceeds edge by edge, aiming at reaching t, whereas after each of his moves the adversary, called Saboteur, takes out an edge from G. Walker wins the game if he has a strategy to reach t whatever choices Walker takes for eliminating edges. The complexity of solving these finite games was determined in pre-Cassting work of the Aachen team ([LR03, Roh05]). A transfer of some of these these results to a probabilistic setting (where Sabotuer behaves as “nature” according to a probability distribution on the edges) was done by the Aachen team in [KRT12].

This setting was modified in joint work of the Brussels and Mons teams to cover quantitative objectives in the deliverable of [BGHM14] of September 2014. Here the framework of classical zero-sum two-player games was studied for reachability objectives combined with the concept of mean-payoff. This work on the problem ”how to reach a target while optimising a payoff” was continued with a submitted article [BGHM15] constructing a pseudo-polynomial time algorithm for solving total-payoff games.

An important aspect in that work is the insight that studying finite horizon quantitative games is not simpler than studying infinite horizon quantitative games – a fact that may justify why we chose to study a long-run average payoff (mean-payoff) for our quantitative sabotage games.

Indeed, extending the results of the Aachen team, the teams of Mons and Brussels launched the study of sabotage games with behavior of infinite lengths. To that purpose, the process of sabotage has been changed to incorporate possibly unbounded behaviors. Our choice was to handle this by introducing quantities. Whereas Walker continues to move along the edges, Saboteur has now a budget of tokens that he puts on vertices of the graphs, and is allowed to move one token at every step. The first player wants to minimize its average number of tokens he sees during the whole play.

If the budget is fixed a priori, the problem of knowing whether the first player has a strategy to obtain an average value below a given threshold seems to be solvable in polynomial time. The static version of this problem, where the second player simply drops his tokens on the graph and leave them as they are forever, is coNP-complete. In the general case, we show that the problem is EXPTIME-complete.

An EXPTIME algorithm can be obtained by studying the configuration graph of this game and solving the induced Mean-Payoff game. The EXPTIME-hardness is obtained by studying a variant of those games where Walker wants to never see a token. We show that such safety games can be encoded into a quantitative sabotage games, and furthermore we reduce an EXPTIME-complete problem, finding the winner in a two-player version of the SAT problem (as Stockmeyer and Chandra did in their classical work of 1979), into this variant.

Reference: [BGHM14, BGHM15]

3 Stability in Dynamic Networks

This work builds on preparatory work on so-called “connectivity games” of the Aachen team over several years (before Cassting) that have been described in detail (and with references not repeated here) in the previous deliverable of September 2014. It involves, besides Cassting, a collaboration with J. Gross (KHT, Sweden), an expert in communications networks, thus enabling us to create bridges to industrial practice.

As described in the previous report of September 2014, we study stability and delay in dynamic networks under adversarial conditions. The network is modelled as an undirected graph in which the nodes are connected by communication links. A list of source and destination pairs determines the desired traffic that has to be routed through the network. We assume that there is a primary user that can block parts of the communication links (subject to some constraints). Such a network is called stable if there is a way of routing the traffic from the sources to the destinations such that the number of packets in the network remains bounded, no matter how the primary user blocks the communication links.

In [TLRG14] we have given a characterization of the stability region of such a network in terms of multi-commodity flows (MCF). This characterization states that the network is stable if there is an MCF with throughput 1 (the commodities corresponding the the individual source and destination pairs for the traffic). The routing strategy that we construct in [TLRG14] guarantees to keep the network stable whenever such an MCF exists, and if there is no such MCF, then the primary user has a way to block the network such that over time the number of packets grows unboundedly. However, problems arise when transferring this routing strategy to an applicable real-world routing algorithm. For the routing algorithm to work, each node is required to have perfect information of the available channels in the entire network in order to make forwarding decisions. This is due to the fact that the characterizing MCFs depend on the active blocking pattern, and the strategy is routing along these flows.

To overcome this problem, we have studied the behaviour of the backpressure routing scheme [TE92] in our setting. Backpressure routing is a routing scheme that only requires local information for the forwarding decisions in each node of the network. Roughly, the packets are forwarded to the neighbours which currently have the smallest queue sizes (the queues are storing the packets that have to be forwarded at a node). Based on the results published in [TLRG14], we can show that backpressure routing is stable under almost the same conditions as the general stability criterion. The existence of an MCF with throughput 1 has to be strengthened to the existence of an MCF with throughput strictly greater than 1. The algorithm from [TLRG14] that checks for a given network whether there exists an MCF with throughput one computes an MCF with maximal throughput, and thus can also be used to check for this new criterion. This result is appealing from a practical point of view as well as from a theoretical point of view. From the practical point of view, we have shown that stability can be ensured (under slightly stronger conditions) by a local routing scheme that does not require global knowledge of the current status of the network. From the theoretical point of view, it is very interesting to see how the existence of some stable routing strategy that is constructed for the purpose, can be used to show that a fixed routing scheme that does not directly depend on the network and the behavior of the primary user can be shown to be stable, too (again under slightly stronger conditions). Whether backpressure routing is stable if there is an MCF with throughput 1 but no MCF with throughput greater than 1 remains an open question.

These results, together with the results from [TLRG14], are currently compiled in a manuscript that will be submitted as a journal paper.

Reference: [TLRG14]

4 Adaptivity in Strategy Construction

As mentioned in the Introduction, the challenge of dealing with evolving structures as game arenas must simultaneously address the question of adapting strategies to meet changes of the game arena. In discussions among the Cassting teams, techniques of learning were raised as a way to implement such adaptations. The Aalborg team came up with work that fits this approach, and we mention it here to indicate the present attempts to integrate strategy adaptation into scenarios of evolving game arenas.

The framework considered by the Aalborg team is given by (priced) timed games; i.e., two-player quantitative games involving an environment assumed to be completely antogonistic. Classical analysis consists in the synthesis of strategies ensuring safety, time-bounded or cost-bounded reachability objectives. Assuming a randomized environment, the (priced) timed game essentially defines an infinite-state Markov (reward) decision process. In this setting the objective is classically to find a strategy that will minimize the expected reachability cost, but with no guarantees on worst-case behaviour. The Aalborg team devised efficient methods for computing reachability strategies that will both ensure worst case time-bounds as well as provide (near-) minimal expected cost. The method employed extends the synthesis algorithms of the synthesis tool Uppaal-Tiga with suitably adapted reinforcement learning that allows to improve a strategy in iterations of learning steps. The resulting algorithm exhibits improvements of several orders of magnitude w.r.t. previously known automated methods.

Reference: [DJL$^+$14]

5 Dynamic Networks of Hybrid Systems

The Aalborg team continued work on enhancing the tool UPPAAL-SMC by features that allow users a dynamically modify a network of timed automata on the fly—by creating new templates. In the previous deliverable the results of 2013 were reported, referring to the paper [DLLP13]

Let us recall the scenario studied in this work: Multiprocessing systems are capable of running multiple processes concurrently. Such systems have established themselves as the defacto standard for operating systems. At the core of an operating system is the ability to execute programs and as such there must be a primitive for instantiating new processes - also programs are allowed to die/terminate. Operating systems may allow the executing programs to split (spawn) into more computational threads in order to let programs take advantage of concurrent execution as well. One of the most used modelling languages, Timed Automata, is based on multiple automata interacting; thus they easily model the concurrent execution of programs. However, this language assumes a fixed size system in the sense that automata cannot be created at will but must be instantiated when the overall system is created. This is in contrast with the fact that developers are able to create threads when needed.

In [DLL+14] we establish a logic-based technique to incorporate spawning of threads into UPPAAL SMC. The modelling language, Dynamic Networks of Stochastic Hybrid Automata, is essentially Timed Automata extended with a spawning primitive and a tear-down primitive. The dynamic creation of threads has the side-effect that it is no longer possible to use ordinary logics to specify behaviours of individual threads - because the threads no longer have unique names. We propose an extension of Metric Temporal Logic with means for quantifying over the dynamically created threads. This makes it possible to zoom in on individual threads and specify requirements to their future behaviour. Furthermore, we present a monitoring procedure for the logic based on rewriting formulas. The presented modelling language and the specification language have been implemented in UPPAAL SMC version 4.1.18.

Reference: [DLL+14]

6 Conclusion

The progress reported in this deliverable essentially covers work of the past six months, following the previous delivarable “Games over Evolving Structures, Interactions with Open Environments” of September 2014.

In the shift of research towards algorithmic solutions as indicated in the title of this report, we obtained results in four areas:

- Sabotage games with infinite behavior,
- Algorithms for stability of networks,
- Adaptivity in strategy constructions,
- Dynamic networks of hybrid systems.

In the first area, work of the Aachen team was taken up by the teams of Mons and Brussels, in the second one, the Aachen team extended previous Cassting results towards practical algorithms, in the third, the Aalborg team opened a new facet of evolving systems by developing a method for adaptivity of strategy construction which has an impact on our research regarding evolving game arenas, and in the fourth, the Aalborg team extended the tool UPPAAL-SMC by features of process creation.

Summarizing, the time of six months since completion of our last report has been used productively by all Cassting teams for interesting new results, with good prospects for the remainder of the Cassting project.

List of Cassting Publications

[BGHM14] Thomas Brihaye, Gilles Geeraerts, Axel Haddad, and Benjamin Monmege. To reach or not to reach? Efficient algorithms for total-payoff games. *CoRR*, abs/1407.5030, 2014. URL: `http://arxiv.org/abs/1407.5030`.

[BGHM15] Thomas Brihaye, Gilles Geeraerts, Axel Haddad, and Benjamin Monmege. To reach or not to reach? Efficient algorithms for total-payoff games, 2015. Submitted to CAV'15.

[DJL$^+$14] Alexandre David, Peter G. Jensen, Kim Guldstrand Larsen, Axel Legay, Didier Lime, Mathias Grund Sørensen, and Jakob H. Taankvist. On time with minimal expected cost! In Franck Cassez and Jean-François Raskin, editors, *Proceedings of the 12th International Symposium on Automated Technology for Verification and Analysis (ATVA'14)*, volume 8837 of *Lecture Notes in Computer Science*, pages 129–145. Springer, November 2014. `doi:10.1007/978-3-319-11936-6_10`.

[DLL$^+$14] Alexandre David, Kim G. Larsen, Axel Legay, Guangyuan Li, and Danny Bøgsted Poulsen. Quantified dynamic metric temporal logic for dynamic networks of stochastic hybrid automata. In *Proceedings of the 14th International Conference on Application of Concurrency to System Design (ACSD'14)*, pages 32–41, June 2014. `doi:10.1109/ACSD.2014.21`.

[TLRG14] Simon Tenbusch, Christof Löding, Frank G. Radmacher, and James Gross. Guaranteeing stability and delay in dynamic networks based on infinite games. In *11th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2014, Philadelphia, PA, USA, October 28-30, 2014*, pages 461–469. IEEE, 2014. `doi:10.1109/MASS.2014.41`.

List of references

[DLLP13] Alexandre David, Kim G. Larsen, Axel Legay, and Danny Bøgsted Poulsen. Statistical model checking of dynamic networks of stochastic hybrid automata. *ECEASST*, 66, 2013. URL: `http://journal.ub.tu-berlin.de/eceasst/article/view/893`.

[KRT12] Dominik Klein, Frank G. Radmacher, and Wolfgang Thomas. Moving in a network under random failures: A complexity analysis. *Science of Computer Programming*, 77(7-8):940–954, 2012. `doi:10.1016/j.scico.2010.05.009`.

[LR03] Christof Löding and Philipp Rohde. Solving the sabotage game is PSPACE-hard. Technical Report AIB-05-2003, RWTH Aachen, 2003. URL: `http://aib.informatik.rwth-aachen.de/2003/`.

[Roh05] Philipp Rohde. *On games and logics over dynamically changing structures*. PhD thesis, RWTH Aachen University, 2005. URL: `http://darwin.bth.rwth-aachen.de/opus/volltexte/2006/1380/index.html`.

[TE92] L. Tassiulas and A. Ephremides. Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks. *IEEE Transactions on Automatic Control*, 37(12):1936–1948, dec 1992.

[vB05] Johan van Benthem. An essay on sabotage and obstruction. In Dieter Hutter and Werner Stephan, editors, *Mechanizing Mathematical Reasoning, Essays in Honor of Jörg H. Siekmann on the Occasion of His 60th Birthday*, volume 2605, pages 268–276. Springer, 2005. `doi:10.1007/978-3-540-32254-2_16`.