Max-Plus Algebra and Discrete Event Systems

J. Komenda * S. Lahaye ** J.-L. Boimond **
T. van den Boom ***
* Institute of Mathematics, Brno Branch, Czech Academy of Sciences, Prague, Czech Republic
** LARIS, Angers University, Angers, France
*** Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands

1. EMERGENCE OF MAX-PLUS APPROACH. A SYSTEM THEORY TAILORED FOR SYNCHRONIZATION.

The emergence of a system theory for certain discrete event systems (DES), in which max-plus algebra and similar algebraic tools play a central role, dates from the early 80s.

Its inspiration stems certainly from the following observation: synchronization, which is a very nonsmooth and nonlinear phenomenon with regard to "usual" system theory, can be modeled by linear equations in particular algebraic structures such as max-plus algebra and other idempotent semiring structures Cuninghame-Green [1979], Cohen et al. [1983].

Two important features must be mentioned for this approach, so-called max-plus linear system theory:

- most of the contributions have used as a guideline the "classical" linear system theory;
- it is turned towards DES performance related issues as opposed to logical aspects considered in other approaches such as automata and formal language theory by including timing aspects in DES description.

A consideration has significantly contributed to the promotion and the scope definition of the approach: a class of ordinary Petri nets, namely the timed event graphs (TEGs) has been identified to capture the class of stationary max-plus linear systems Cohen et al. [1985] and subsequent publications by Max Plus team [2]. TEGs are timed Petri nets in which each place has a single input transition and a single output transition. A single output transition means that no conflict is considered for the tokens consumption in the place, in other words, the attention is restricted to DES in which all potential conflicts have been solved by some predefined policy. Symmetrically, a single input transition implies that there is no competition in supplying tokens in the place. In the end, mostly synchronization phenomena (corresponding to the configuration in which a transition has several input places) can be considered, and this is the price to pay for linearity.

This new area of linear system theory has benefited from existing mathematical tools related to idempotent algebras such as lattice theory Birkhoff [1940], residuation theory Blyth and Janowitz [1972], graph theory Gondran and Minoux [1979], optimization Zimmermann [1981] and idempotent analysis Kolokoltsov and Maslov [1997] but it is worth mentioning that the progress has probably been impeded by the fact that some fundamental mathematical issues in this area are still open.

The overview of the contributions reveals that main concepts from linear system theory have been step by step specified into max-plus linear system theory. Without aiming to be exhaustive:

- several possible representations have been studied, namely state-space equations, transfer function in event domain Cohen et al. [1983, 1985], time domain Caspi and Halbwachs [1986], and two-dimensional domain using series in two formal variables Cohen et al. [1986] (with more details in Cohen et al. [1989]);
- performance analysis and stability are mostly based on the interpretation of the eigenvalue of the state-matrix in terms of cycle-time, with its associated eigenspace and related cyclicity property Baccelli et al. [1992], Gaubert [1997];
- a wide range of control laws have been adapted such as:
  - open-loop structures overcoming system output tracking [Baccelli et al. 1992, chap. 5.6], Cofers and Garg [1996], Menguy et al. [2000] or model reference tracking Libeaut and Loiseau [1996],
  - closed-loop structures taking into account disturbances and model-system mismatches Cottereau et al. [1999], Lieders and Santos-Mendes [2002], possibly including a state-observer Hardouin et al. [2010],
  - model predictive control scheme [De Schutter and van den Boom 2001, van den Boom and De Schutter 2002] with emphasis on stability in [Necoara et al. 2007].
For a large survey on max-plus linear systems theory, we refer to books Baccelli et al. [1992], Gunawardena [1998], Heidergott et al. [2006], Butkovic [2010], to manuscript Gaubert [1992] and surveys Cohen et al. [1989], Gaubert [1997], Cohen et al. [1999], Akian et al. [2003].

2. SOME EXTENSIONS FOCUSED ON SYNCHRONIZATION IN DES.

A natural generalization of deterministic max-plus-linear systems are stochastic max-plus-linear systems, which have been studied for more than two decades. Ergodic theory of stochastic timed event graphs have already been developed in Baccelli et al. [1992], where most of the theory is covered. In particular, asymptotic properties of stochastic max-plus-linear systems are studied therein in terms of the so-called Lyapunov exponents that correspond to the asymptotic mean value of the norm of the state variables. In the case the underlying event graph is strongly connected the Lyapunov exponent is the unique value to which the mean value converges. For general event graphs there is a maximal Lyapunov exponent.

Uncertainty can also be considered through intervals defining the possible values for parameters of the system. In Lhommeau et al. [2004] TEGs, in which the number of initial tokens and the time delays are only known to belong to intervals, are represented over a semiring of intervals and robust controllers are designed.

Another way of extending techniques for linear systems is to consider that parameters of the models may vary, that is study non-stationary linear systems. This possibility has been examined within the max-plus linear setting Brat and Garg [1998], Lahaye et al. [1999, 2004] with contributions mainly focused on representation, control and performance analysis.

Continuous timed event graphs are particular graphs in which fluids hold rather than discrete tokens and the fluid flow through transitions can be limited to a maximum value. Moreover, an initial volume of fluid can be defined in places and times can be associated with places to model fluid transportation times MaxPlus [1991]. Such graphs are relevant for example to approximate the behavior of high throughput manufacturing systems in which the number of processed parts is very large. In parallel, a similar approach called network calculus has been developed by considering computer network traffic as a flow (based on the use of ‘leaky buckets’) to approximate the high number of conveyed packets Cruz [1991]. Extension on fluid timed event graphs with multipliers in a new algebra, analogous to the min-plus algebra, has been proposed in Cohen et al. [1995], Cohen et al. [1998].

Switching Max-Plus-Linear (SMPL) systems are discrete-event systems that can switch between different modes of operation [van den Boom and De Schutter 2006]. The switching allows to change the structure of the system, to break synchronization, or to change the order of events. In each mode the system is described by a max-plus-linear state equation and a max-plus-linear output equation. Note that regular max-plus-linear systems are a subclass of SMPL system, namely with only one mode. In van den Boom and De Schutter [2011], authors describe the commutation between different (max,+) linear modes and shows that an SMPL system can be written as a piecewise affine system which allows for using similar techniques in such seemingly different classes of systems. SMPL models can be used to describe the dynamics of various semi-cyclic discrete event systems.

Another extension of the class of systems that can be modeled in max-plus algebras consists in considering hybrid Petri nets, and more particularly, hybrid TEGs that consist of a discrete part (a TEG) and a continuous part (a continuous TEG). It has been shown in Komenda et al. [2001] that a linear model can be obtained based on counter function if only one type of the interface between continuous and discrete part is present. However, for application to just in time control this constraint can be relaxed as it has been shown in Hamaci et al. [2006].

P-time Petri nets form an important extension of Petri nets, where the timing of places/transitions is nondeterministic. They have been studied in max-plus algebras in Declerck and Alaoui [2004], Declerck and Alaoui [2005]. They find their applications e.g. in modeling of electroplating lines or chemical processes, where both upper and lower bound constraints processing time are required, see e.g. Spacek et al. [1999].

3. CONSIDERING DES WITH SHARED RESOURCES.

An important aspect of DES is that different processes share and compete for resources. Resource allocation policies have to be considered to solve the corresponding conflicts. Timed discrete event systems (TDES) are mostly modeled by timed extension of Petri nets or by timed extension of finite automata. There have been several different approaches to the investigation of resource sharing phenomenon, in conjunction with other phenomena of TDES (synchronization and concurrency).

Resource allocation policies have been studied based on classical linear systems (with underlying TEG models) in the idempotent semiring min-plus with counter functions Cohen et al. [1997]. More recently, conflicts among timed event graphs have been studied in Addad et al. [2012], where conflicting timed event graphs (CTEGs) have been proposed with some fairly restrictive assumptions. Resource allocation policies studied in Addad et al. [2012] are mainly FIFO and cyclic (periodic) policy. The authors provide an analytic evaluation of an upper bound on the cycle time of CTEG as a function of the cycle time of each TEG and of the times of the conflict places. However, this direct modeling approach has not been yet applied to control problems.

There is a completely different approach that is based on the so-called max-plus automata, which have been introduced by S. Gaubert in Gaubert [1995] to generalize both max-plus-linear systems and logical automata. Resource sharing is then modeled by the underlying finite automaton instead of a Petri net. Unlike the first approach, where different resource allocation policies are handled one by one, this approach allows for simultaneous modeling of different resource allocation policies within a single model as long as these policies can be represented by a regular
language. This means that as an automaton-based model it can handle several such policies at the same time within a single model without having to rebuild the model each time the policy is changed.

The approach based on max-plus automata has enabled a detailed study of performance evaluation Gaubert [1995] of DES with shared resources. S. Gaubert has presented results about the worst case, optimal case and mean case performance of max-plus automata. More sharp results are naturally obtained for deterministic max-plus automata, but most of the paper is focused on general nondeterministic max-plus automata. It is to be noted that not all max-plus automata can be determined and their determinization, i.e. existence of a deterministic max-plus automaton having the same behavior (recognizing the same formal power series), is still an open problem and is not even known if determinization of a given nondeterministic max-plus automaton is decidable. In Gaubert [1995] a characterization of deterministic max-plus automata is provided based on the concept of projectively finite semigroups, which can be used as a semi-algorithm for determinization (with no guarantee of success).

More recently, supervisory control theory of max-plus automata has been investigated with the ultimate goal to extend this theory from logical to timed automata. Although max-plus automata can be viewed as general timed automata, they have strong expressive power in terms of timed Petri nets as shown in Gaubert and Mairesse [1999]: Supervisory control theory of max-plus automata with complete observations has been proposed in Komenda et al. [2009], where the basic elements of supervisory control, such as supervisor, closed-loop system and controllability are extended from logical to max-plus automata. However, it follows from results presented therein that rational (finite state) controllers can only be obtained for systems (plants) which have behaviors at the same time max-plus and min-plus rational. This is a serious fundamental obstacle, the class of series that are at the same time max-plus and min-plus rational coincides with the class of unambiguous series (series recognized by an unambiguous automaton) as has been shown in Lombardy and Sakarovitch [2006]. Although unambiguous series is less restrictive property than deterministic series (series recognized by a deterministic max-plus automaton), a typical approach for imposing unambiguity is to determinize a max-plus automaton. More complete picture about rationality issues extended to more general setting is presented in Lahaye et al. [2015]. More specifically, minimally permissive and just-after-time supervisors are studied in order to guarantee a minimal required behavior and to delay the system as late as possible so that sequences of event occur later than prescribed dates, which is important for applications in transportation networks (e.g. improving train connections in railway systems), but also in manufacturing systems and communication networks. It has been shown that finite state controllers for both control problems exist if the system-series and the specification (reference-series) are both unambiguous. This assumption is met for several classes of practically relevant max-plus automata, e.g. those modeling a type of manufacturing systems such as safe Flow-shops and Job-shops. Another class of timed system called timed weighted systems has been studied in Su et al. [2012]. Timed weighted systems are simply modular automata (collection of local automata) endowed with a so called mutual exclusion function as well as a time-weighted function. Timed weighted systems can be understood as a synchronous product of max-plus automata (which is not made explicit) and durations of events are described by time-weighted function.

In our opinion max-plus automata form a gateway to the general timed automata, because systems modeled by max-plus automata exhibit most of decidability and determinization issues that are present for general timed automata, while they are conceptually simpler, which allows for better grasping the core of these fundamental problems.

On the other hand, max-plus automata can so far efficiently represent only safe timed Petri net models, although some extensions to general bounded non safe timed Petri nets are being explored.

The Model Predictive Control (MPC) design method can be applied to (switching) max-plus linear systems [van den Boom and De Schutter 2006]. MPC for conventional (non-DES) systems is very popular in the process industry [Maciejowski 2002] and a key advantage of MPC is that it can accommodate constraints on the inputs and outputs. For every cycle the future control actions are optimized by minimizing a cost function over a prediction window subject to constraints. If the cost function and the constraints are piecewise affine functions in the input, output, and state variables, the resulting optimization problem will be a mixed-integer linear programming (MILP) problem, for which fast and reliable algorithms exist. An alternative approach is to use optimistic optimization [Xu et al. 2014]. In De Schutter and van den Boom [2001] MPC for regular max-plus-linear systems was studied using the just-in-time cost function with constraints that were monotonically nondecreasing in the output. In that case the problem turns out to be a linear programming problem. Results for MPC of stochastic (switching) max-plus linear systems are given in van den Boom and De Schutter [2012] and Farahani et al. [2016].

In scheduling problems the system matrix of the SMPL model becomes affine in max-plus binary scheduling variables [van den Boom et al. 2013]. This scheduling technique has been applied in Kersbergen et al. [2016], where a railway traffic management algorithm has been derived that can determine new conflict-free schedules and routes for a railway traffic network when delays occur.

4. MAX-PLUS-GEOMETRIC THEORY.

Classical linear algebra as well as control of linear systems are known to be closely related to geometry. It was then natural that much effort has been put on generalizing these well established connections to the max-plus algebra and max-plus-linear systems. First attempts to formulate basic concepts of max-plus-geometric theory go back to 1990’s, cf. Cohen et al. [1996], where projection onto images of operators and parallel to the kernels of operators have been studied. Another important concept, convexity, a powerful tool in optimization has been investigated and extended to the max-plus framework, because max-
plus-convex sets play an important role in optimization problems. Minkowski theorem in classical linear algebra shows that a non-empty compact convex subset of a finite dimensional space is the convex hull of its set of extreme points. In Gaubert and Katz [2007] a max-plus analogue is proven, which extends this result to max-plus convex sets. Such sets arise in natural way in several domains, ranging from max-plus-linear systems, abstractions of timed automata to solutions of a Hamilton-Jacobi equation associated with a deterministic optimal control problem.

Concerning control of max-plus-linear systems, an important concept of \((A, B)\) invariant spaces has been investigated in Katz [2007], which generalizes some invariance properties from the Wonham’s geometric approach Wonham [1974] to linear systems. Let us point out at this point that the Wonham’s geometric approach is well known for being at the very origin at the whole domain of discrete event systems, namely supervisory control theory developed in early 1980’s in parallel with the theory of max-plus-linear systems.

Recently, an interesting connection has been discovered between geometrical approach to classical max-plus-linear systems Gaubert and Katz [2007] and reachability analysis of timed automata, a very general model of TDES including several continuous variables called clocks that measure time that have elapsed since their last reset and define timing conditions for enabling logical transitions in timed automata.

Interestingly, results from max-plus geometry find their application in reachability analysis of timed automata. Timed automata with infinite clock spaces are abstracted into finite automata called region or zone automata, where the clock space is abstracted by a finite number of geometric zones. The zones are represented by efficient data structures called difference bound matrices (DBM) that represents the bounds on differences between state variables. The reachability of different zones can be studied using max-plus-cones from geometric theory of max-plus-linear systems.

It has been shown in Lu et al. [2012] that max-plus polyhedra are very useful in reachability analysis of timed automata. Every max-plus cone (also called max-plus polyhedron) can actually be expressed as a union of finitely many DBM’s as shown in Adzkiya et al. [2013]. These geometric objects have proven to be extremely useful for both forward and backward reachability analysis. Forward reachability analysis aims at computing the set of possible states that can be reached under the model dynamics, over a set of consecutive events from a set of initial conditions and possibly by choosing control actions [Adzkiya et al. 2015]. Backward reachability analysis is to determine the set of states that enter a given set of final states, possibly by choosing control actions. This is of practical importance in safety control problems consisting in determine the set of initial conditions leading to unsafe states. However, for backward reachability analysis the system matrix has to be max-plus invertible, i.e. in each row and in each column there should be a single finite element (not equal to \(-\infty\)), which is restrictive. The main advantage of using max-plus polyhedra is saving of computational complexity, because time complexity of these approaches is polynomial as for standard DBM based algorithms.

5. APPLICATIONS.

It can appear somewhat surprising that methods based on very particular structure of max-plus algebra can find a large number of applications. But it turns out that max-plus system theory has been applied to a wide variety of domains, such as:

- capacity assessment, evaluation and control of delays in transportation systems Braker [1991], Heidergott et al. [2006], Houssin et al. [2007], Kersbergen et al. [2016] and car traffic Farhi et al. [2011],
- sizing, optimization and production management in manufacturing systems Cohen et al. [1985], Imaev and Judd [2008], Martinez and Castagna [2003],
- performance guarantees in communication networks through so-called network calculus Cruz [1991], J.-Y. Le Boudec [2001],
- high throughput screening in biology and chemistry Brunsch et al. [2012],
- modeling, analysis and control of legged locomotion Lopes et al. [2014, 2016],
- speech recognition Mohri et al. [2002] or image processing Culik and Kari [1997] through weighted automata such as max-plus automata,
- optimization of crop rotation in agriculture Bacaër [2003],
- scheduling of energy flows for parallel batch processes Mutsaers et al. [2012],
- paper handling in printers Alirezaei et al. [2012]

... This diversity is to be emphasized all the more since these applications have sometimes suggested new theoretical questions.

We cannot finish this overview without mentioning the important connections with other fields of research: dynamic programming and optimal control with solutions to Hamilton-Jacobi-Bellmann differential equations Maslov and Kolokoltsov [1994], Quadrat and Max-Plus [1994], statistical mechanics Quadrat and Max-Plus [1997], operations research Zimmermann [2003],...

REFERENCES


