Automata and Logics for Concurrent Systems: Realizability and Verification
Benedikt Bollig

To cite this version:

HAL Id: tel-01242861
https://hal.archives-ouvertes.fr/tel-01242861
Submitted on 14 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Automata and Logics for Concurrent Systems: Realizability and Verification

Habilitation Thesis
Mémoire d’Habilitation à Diriger des Recherches

submitted by

Benedikt Bollig

June 2015
Abstract. Automata are a popular tool to make computer systems accessible to formal methods. While classical finite automata are suitable to model sequential boolean programs, models of concurrent systems involve several interacting processes and extend finite-state machines in various respects. This habilitation thesis surveys several such extensions, including pushdown automata with multiple stacks, communicating automata with fixed, parameterized, or dynamic communication topology, and automata running on words over infinite alphabets. We focus on two major questions of classical automata theory, namely realizability (asking whether a specification has an automata counterpart) and model checking (asking whether a given automaton satisfies its specification).

The defense took place on Tuesday, June 2, 2015, 2 p.m., at ENS Cachan, France. The jury members are:

- Prof. Ahmed Bouajjani  
  Reviewer
- Prof. Volker Diekert  
  Examiner
- Prof. Paul Gastin  
  Examiner
- Prof. Leonid Libkin  
  Reviewer
- Prof. Anca Muscholl  
  Reviewer
- Prof. Jean-Marc Talbot  
  Examiner

Author
Benedikt Bollig
Laboratoire Spécification et Vérification
CNRS & ENS Cachan
France
email: bollig@lsv.ens-cachan.fr
Thank you, Merci, Danke!

Many people have helped me over the last years, and I would like to use the opportunity to express my gratitude to all of them. The list of people I mention here cannot be exhaustive, but I am aware that many more crossed and influenced my road in one way or another.

My foremost thanks, and my sincerest appreciation, go to my mentor Paul Gastin. I am extremely grateful for his continuous support, advice, and encouragement. Since I defended my PhD, he has had, without any doubt, the most significant impact on my evolution as a researcher. When writing a definition or proof, the best guideline for me has always been to first think about how Paul would write it. And as if all that were not enough, he gave me the opportunity to work with such brilliant PhD students: I am so proud to count Aiswarya Cyriac, Benjamin Monmege, and Akshay Sundararaman among my family, in an academic sense and in terms of friendship. I would consider myself lucky if they have learned half as much from me as I have learned from them.

I am flattered that Ahmed Bouajjani, Volker Diekert, Paul Gastin, Leonid Libkin, Anca Muscholl, and Jean-Marc Talbot readily accepted to be a jury member or a reviewer of my habilitation thesis. I can only hope that this document fulfils the expectations that come with such a high-calibre committee.

I am very grateful to Benjamin Monmege for taking the trouble to read this thesis and for his many accurate, subtle, and pertinent comments.

During my years at LSV, I have been so fortunate to share my office with adorable and courteous people: Nicolas Markey, Graham Steel (together with a curious little cat whose name I do not know), and Claudine Picaronny have created an enjoyable and stimulating atmosphere every day.

I also spent a considerable amount of time in Christoph Haase’s office. I sincerely thank Christoph for his support and encouragement, the various thought-provoking discussions we had on emerging trends in computer science, and many excellent (among some not so good) “Kalauers”\(^1\).

\(^1\)http://de.wikipedia.org/wiki/Kalauer
I would like to thank all my co-authors for joining me on this exciting journey that is research. Most results presented here would not have been possible without them. A special mention goes to Dietrich Kuske, Madhavan Mukund, K. Narayan Kumar, and Thomas Schwentick. Besides sharing their immense knowledge with me, they made some memorable research stays possible. Normann Decker, Peter Habermehl, Martin Leucker, and Daniel Thoma kindly let me take part in the LeMon project, which has been another invaluable and fruitful experience.

I could not imagine a better vehicle to undertake my research journey than the LSV, which is blessed with kind and talented people. I am particularly grateful to the MExICo team under the guidance of Stefan Haar, to Stefan Göller and Stefan Schwwoon for enlightening and pleasurable conversations on computer science and football science, and to the remarkably efficient and friendly administrative and technical staff for their support.

I will never forget the warm welcome that I received when I joined the LSV ten years ago. I am very much indebted to Myrto Arapinis, Nathalie Bertrand, Patricia Bouyer-Decitre, Stéphanie Delaune, Marie Duflot-Kremer, Catherine Forestier, Steve Kremer, Nicolas Markey, Arnaud Sangnier, and Tali Sznajder, who have had a lasting impression on me, and certainly on the LSV as a whole.

My heartfelt thanks go to my beloved sister, I am so happy she is around. I am deeply grateful to my parents-in-law for their unfailing support and immense generosity. A special thought goes to my friends Darius Dlugosz, Carsten Kern, and Martin Leucker, who are far away and yet so close, and with whom all this started.

_I dedicate this thesis to my mother, who has always been there for me, and to the memory of my father, who will always be right by my side, every step I take._

_Le dernier mot est réservé pour toi, Virginie. Merci, infiniment, pour ton amour inconditionnel et pour Lucas, le plus beau cadeau qu’on ait pu me faire._

Benedikt Bollig
Cachan, May 2015
# Contents

1 Introduction

1.1 Automata Models and Their Behavior ....................... 2
1.2 Specification Formalisms .................................. 5
1.3 Realizability .............................................. 6
1.4 Model Checking .......................................... 7
1.5 Contributions and Outline .................................. 8
1.6 Further Contributions ...................................... 14
1.7 Notation ................................................... 15

2 Static Recursive Shared-Memory Systems .................. 17

2.1 (Multiply) Nested Words and Their Automata .......... 17
2.2 Contexts, Phases, Scopes, and Ordered Stacks ............ 21
2.3 Monadic Second-Order Logic ................................ 23
2.4 The Language Theory of Nested-Word Automata .......... 24
2.5 Nested Traces and Their Automata ......................... 26
2.6 Realizability of NWA Specifications ....................... 29
2.7 Realizability of MSO Specifications ....................... 33
2.8 Temporal Logic for Nested Words ......................... 34
2.9 Perspectives ................................................. 38

3 Parameterized Message-Passing Systems .................. 39

3.1 Topologies ................................................ 40
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Message Sequence Charts</td>
<td>42</td>
</tr>
<tr>
<td>3.3</td>
<td>Underapproximation Classes</td>
<td>43</td>
</tr>
<tr>
<td>3.4</td>
<td>Communicating Automata</td>
<td>45</td>
</tr>
<tr>
<td>3.5</td>
<td>Fixed-Topology Communicating Automata</td>
<td>47</td>
</tr>
<tr>
<td>3.6</td>
<td>Parameterized Communicating Automata (PCAs)</td>
<td>49</td>
</tr>
<tr>
<td>3.7</td>
<td>Logical Characterizations of PCAs</td>
<td>50</td>
</tr>
<tr>
<td>3.8</td>
<td>Model Checking</td>
<td>55</td>
</tr>
<tr>
<td>3.9</td>
<td>Perspectives</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>Dynamic Sequential Systems</td>
<td>59</td>
</tr>
<tr>
<td>4.1</td>
<td>(Fresh-)Register Automata and Session Automata</td>
<td>61</td>
</tr>
<tr>
<td>4.2</td>
<td>Closure Properties of Session Automata</td>
<td>64</td>
</tr>
<tr>
<td>4.3</td>
<td>Class Register Automata</td>
<td>66</td>
</tr>
<tr>
<td>4.4</td>
<td>Automata vs. Logic over Data Words</td>
<td>71</td>
</tr>
<tr>
<td>4.5</td>
<td>Decision Problems</td>
<td>75</td>
</tr>
<tr>
<td>4.6</td>
<td>Perspectives</td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>Dynamic Message-Passing Systems</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Dynamic Message Sequence Charts</td>
<td>80</td>
</tr>
<tr>
<td>5.2</td>
<td>Dynamic Communicating Automata</td>
<td>82</td>
</tr>
<tr>
<td>5.3</td>
<td>Register MSC Graphs</td>
<td>87</td>
</tr>
<tr>
<td>5.4</td>
<td>Executability of Register Message Sequence Graphs</td>
<td>88</td>
</tr>
<tr>
<td>5.5</td>
<td>Guarded Register MSC Graphs</td>
<td>92</td>
</tr>
<tr>
<td>5.6</td>
<td>Perspectives</td>
<td>93</td>
</tr>
<tr>
<td>6</td>
<td>Static Timed Shared-Memory Systems</td>
<td>95</td>
</tr>
<tr>
<td>6.1</td>
<td>Timed Automata with Independently Evolving Clocks</td>
<td>96</td>
</tr>
<tr>
<td>6.2</td>
<td>The Existential Semantics</td>
<td>99</td>
</tr>
<tr>
<td>6.3</td>
<td>The Universal Semantics</td>
<td>102</td>
</tr>
<tr>
<td>6.4</td>
<td>The Reactive Semantics</td>
<td>104</td>
</tr>
<tr>
<td>6.5</td>
<td>Perspectives</td>
<td>107</td>
</tr>
<tr>
<td>7</td>
<td>Conclusion</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>127</td>
</tr>
</tbody>
</table>
Computer systems are ubiquitous. They have pervaded almost every aspect of modern human life and are becoming more and more complex. A countless number of processes run in parallel and access data stored on servers that are physically distributed all over the globe. Indeed, mobile computing, ad-hoc networks, multicore systems are only a few of many omnipresent keywords that stand for a shift towards concurrent systems. To put it bluntly:

The world IS concurrent. It IS parallel.\(^1\)

However, understanding concurrency is a challenging task. The interplay of several processes becomes unpredictable, and even carefully written and seemingly correct programs may eventually reveal serious bugs. Unfortunately, even subtle errors may entail serious damage to hardware, transport systems, business processes, electronic devices, etc.

This is where formal methods come into play. The term formal methods gathers tools and methods for a mathematically founded analysis of computer systems. They may be adopted at the early stages of system design, but also be applied to check an existing system against a requirements specification. This thesis aims at contributing to the mathematical foundations of concurrent systems, their formal modeling, specification, verification, and synthesis.

In the following, we first describe different models of concurrent systems to take into account applications in various domains. We will focus on automata, which are a versatile tool to model different phenomena of concurrency, yet, at the same time, offer a unifying view of all of them. We then describe high-level specification

\(^1\)Joe Armstrong [Arm07].
formalisms (such as logics, regular expressions, and also automata). Though those are also equipped with a precise mathematical semantics, they allow for a more abstract, declarative view of a system. Two natural questions arise in this context, which are usually referred to as realizability and model checking:

**Realizability:** Given a requirements specification in terms of a formula (regular expression, automaton, respectively), is there an automaton that recognizes all, but only those, behaviors that satisfy the specification?

**Model Checking:** Given a system model in terms of an automaton, and a requirements specification in terms of a formula (regular expression, automaton, respectively), do all executions of the automaton satisfy the specification?

Both questions are actually closely related: various model-checking procedures rely on automata-theoretic techniques and the fact that a specification is “realizable” as an automaton. This thesis tries to contribute to both of these aspects of formal verification: realizability and model checking. In the following, we examine some characteristics of automata, specification formalisms, realizability, and model checking in more detail.

### 1.1 Automata Models and Their Behavior

Automata are a popular model of computer systems, making them accessible to formal methods and, in particular, synthesis and automated verification techniques such as model checking. While classical finite-state automata are suitable to model sequential boolean programs, models of concurrent systems, involving several interacting processes, extend finite-state machines in several respects. Roughly, we may classify a system (or a system model) according to the following characteristics:

**Form of Communication.** When there is only one process, we deal with a sequential system. When there are several processes, then interprocess communication may be accomplished, e.g., via shared variables or message passing. In this document, we consider boolean programs so that shared-variable communication usually gives rise to finite-state systems. Thus, classical methods are applicable for their analysis. On the other hand, message passing via a priori unbounded channels leads to undecidability of basic verification questions. However, putting some restrictions on the system (e.g., imposing a channel bound or restricting the system architecture) will allow us to infer positive results for both realizability and model checking.

In Figure 1.1, we illustrate the way of modeling the behavior that is generated by automata for the various types of communication. One single sequential process will produce sequences of events, as depicted on the left-hand side. One such execution can be seen as a word, i.e., the sequence of actions that we observe. More generally, it may be seen as a graph whose nodes are arranged chronologically, following a total order generated by the edge relation. Each node represents an event, and
with each event, we associate the action that it executes. For example, the actions $a, b, c$ used in the figure may stand for writing or reading a variable, outputting the value of a variable, etc. The figure in the middle models an execution as it may be produced by a shared-memory system where processes share access to resources such as variables. This kind of graph is usually referred to as a *Mazurkiewicz trace*. In the example, we deal with two processes, sharing two events that both execute an action $c$. That is, $c$ may stand for reading or writing a variable that is shared by both processes. Finally, in a message-passing environment, processes do not share resources in a synchronous manner, but use an asynchronous communication medium such as channels. The causal ordering implied by the channel policy may then be depicted as an additional edge between a send and the corresponding receive event. This is what happens in the graph on the right, which is called a *message sequence chart (MSC)*. In this example, we deal with actions $a$ (for sending) and $b$ (for receiving), while $c$ may be some internal action.

![Graphs showing different forms of communication](image)

Figure 1.1: Forms of communication

As advocated in [AG14], modeling an execution of a concurrent system in terms of an acyclic (i.e., partially ordered) graph actually offers many advantages over the interleaving approach (which assumes an order on concurrent events). Since graphs visualize communication links between events, corresponding specification languages are usually more expressive and more intuitive.

**System Architecture.** The system architecture, connecting processes and arranging them in a certain way (e.g., as a pipeline or as a tree), may be static and fixed (i.e., known), or static but unknown, or it may change dynamically during a system execution. These three cases are schematically depicted in Figure 1.2. In the particular case where the topology is static but unknown, we deal with a *parameterized setting*. So, one will be interested in questions such as “Is the system correct no matter what the topology or number of processes is?” or “Can one transform a specification into a system that is correct for all system architectures?”.

In the particular case where the topology is static but unknown, we deal with a *parameterized setting*. So, one will be interested in questions such as “Is the system correct no matter what the topology or number of processes is?” or “Can one transform a specification into a system that is correct for all system architectures?”.

More and more applications (let us just mention mobile computing) have dynamic architectures which allow processes to connect to/disconnect from processes at execution time. There has been a wide range of techniques for the verification of parameterized and dynamic systems. There are also close connections with the theory of words over infinite alphabets, where the infinitary part can be used to
model an unbounded supply of process identifiers.

Finite-State, Recursive, and Timed Processes. Processes themselves can be finite-state (i.e., be modeled as finite-state automata), recursive (i.e., be modeled as a pushdown automaton), or timed (i.e., be modeled as a timed automaton). Figure 1.3 illustrates three process behaviors, one for each case. The behavior of a recursive process is depicted as a nested word, where some (well-nested) edges connect the call and return position of a function call. A timed execution, on the other hand, provides additional time stamps telling us when an event has taken place.

In this thesis, we survey automata models for some combinations of the above characteristics. In particular, we cover automata for concurrent recursive shared-memory systems, parameterized and dynamic message-passing systems, and concurrent timed systems. In turn, these automata generalize the well-studied notions of Zielonka automata [Zie87] (towards recursion), communicating automata [BZ83] (towards dynamic and parameterized architectures), and timed automata [AD94] (towards concurrency).
1.2 Specification Formalisms

A specification describes the desired (or forbidden) behavior of a system. While automata serve as models of programs and, therefore, provide a low-level view of a system, requirements specifications are usually written in a high-level language, abstracting from implementation details and being closer to human way of thinking. Popular examples include regular expressions, monadic second-order (MSO) logic, as well as temporal logics. The expressive equivalence of finite automata and MSO logic, known as the Büchi-Elgot-Trakhtenbrot Theorem, constitutes one of the most beautiful cornerstones in the landscape of theoretical computer science [Büc60, Elg61, Tra62]. This can also be claimed for the coincidence of languages expressible in first-order (FO) logic and those definable in linear-time temporal logic (LTL), which is referred to as Kamp’s Theorem [Kam68]. Beautiful on their own, those logical connections have led to practically usable verification tools [Hol03, HJJ+95].

The Büchi-Elgot-Trakhtenbrot Theorem and Kamp’s Theorem have counterparts in concurrent settings [Tho90, TW97, DG02, DG06]. However, some classical results concerning the connection of specification formalisms with automata are not transferable. We will see that a generic MSO logic over graphs is, in general, too expressive to capture automata for static message-passing or recursive systems (while the second-order quantifier-alternation hierarchy collapses over words). Moreover, the transfer of temporal logics to richer structures and graphs than words is far from obvious. Indeed, there have been several proposals of temporal logics over Mazurkiewicz traces (see [GK07] for an overview), and a number of temporal logics have recently been considered for concurrent recursive processes [Ati10, ABKS12, LTN12]. Any such logic comes with tailored algorithms for its application in verification. However, it is fair to state that no canonical logic has been identified so far (unlike in the word case where LTL can certainly be considered the yardstick logic for linear-time model checking). Previously defined temporal logics for concurrent recursive systems do not even enjoy a theoretical justification as it is provided for the word case in terms of Kamp’s Theorem.

For concurrent recursive programs, we will, therefore, consider all temporal logics, as defined in the book by Gabbay et al. [GHR94], in a unifying framework. Note that this approach has already led to a versatile framework for temporal logics over Mazurkiewicz traces [GK07, GK10]. We also provide automata-based specification formalisms (some of which can be considered or defined as regular expressions). The difference with system models is that automata specifications have a global state space offering, in a sense, a high-level view. As an example, automata with multiple pushdown stacks but one state space will serve as specifications for concurrent-system models where several pushdown automata, with separate state spaces, communicate via shared variables. In a dynamic setting, we introduce automata specifications with registers that can take values from infinite domains (for example, to model process identifiers). This will reveal new connections with the theory of formal languages over infinite alphabets.
1.3 Realizability

The realizability problem arises when we are given a specification that we would like to transform into an implementation in terms of an automata model. So, the first question to ask is if there is such an implementation at all. In the previous section, we already mentioned the Büchi-Elgot-Trakhtenbrot Theorem saying that, roughly speaking, all regular specifications are realizable. In Figure 1.4, three equivalent specifications are given (from top to bottom: LTL formula, regular expression, MSO formula) as well as an implementation in terms of a finite automaton.

\[ G(a \rightarrow Fb) \]
\[ ((b + c)^*(\varepsilon + a(a + c)^*b))^* \]
\[ \forall x.(a(x) \rightarrow \exists y.(x \leq y \land b(y))) \]

Figure 1.4: Realizability of sequential finite-state systems

While, in terms of regular languages, any specification has an implementation, the situation is more difficult when we move to the concurrent setting. Let us look at a well studied setting of concurrent shared-memory systems with finite-state processes. Suppose that we have two processes, 1 and 2, where 1 executes action \( a \), and 2 executes action \( b \). Consider Figure 1.5. Is it justified to say that the concurrent automaton on the right-hand side, with no communication between processes 1 and 2, is an implementation of the regular expression \( \alpha = ab(ab)^* \)? There are (at least) two arguments against it. Expression \( \alpha \) implies that any execution performs as many \( a \)'s as \( b \)'s. But this is not realizable under the architecture that we consider (recall that 1 and 2 do not communicate). Moreover, \( b \)'s have to be preceded by \( a \)'s, which is not realizable for the same reason. Note that these problems are inherent to the specifications and will not vanish with a more clever implementation. So, what is the “right” formalism for the specification of those systems? Or, put differently, what are the valid specifications, when we assume that \( ab(ab)^* \) is not realizable?

Consider Figure 1.6, and suppose that the system provides a third action, call it \( c \), that is executed simultaneously by both processes. The concurrent automaton on the right-hand side might then be seen as an implementation of the regular expression on the left. First, the action \( c \) allows both processes to synchronize after each \( a \)-step (\( b \)-step, respectively), so that, in particular, there are as many \( a \)'s as \( b \)'s in any execution. Second, the order of executing \( a \) and \( b \) is not fixed anymore by the specification. Indeed, a classical result of Mazurkiewicz trace theory due
to Zielonka states that regular specifications that are closed under permutation of such independent actions can be effectively translated into a concurrent automaton [Zie87]. Note that the specification from Figure 1.5 is actually not closed in that sense. Similar results were obtained in the static message-passing setting [HMK+05, Kus03, GKM06]. In this thesis, we obtain a generalization of Zielonka's theorem for concurrent recursive programs, using Zielonka's result as a black box.

\[(ab + ba)c^* \]

Figure 1.6: Realizability of concurrent finite-state systems

Recall that the realizability question amounts to asking whether a specification can be translated into an automaton. Thus, we would like to distinguish algorithmically between the specifications from Figure 1.5 (not realizable) and Figure 1.6 (realizable). The question is decidable for finite-state systems [Mus94, PWW98], but, unfortunately, it is undecidable in the recursive case, which follows from the undecidability of the nonemptiness problem [Ram00]. We will, therefore, identify sufficient decidable criteria that still guarantee realizability. The idea is to restrict the specifications. For example, we may impose a bound on the number of context switches, a notion introduced in [QR05], where each context allows only one process to call or return from a procedure. This amounts to underapproximating the complete system behavior. Note that the final implementation can exhibit executions that do not fit into this restriction anymore. We will come back to that point in the subsequent section about model checking.

Other possibilities are specification formalisms that do not distinguish between linearizations but are interpreted directly on a graph. Though this does not necessarily guarantee realizability, it already rules out design errors that can a priori be avoided. In this thesis, we consider MSO logic and temporal logic interpreted over behavior graphs (rather than their linearizations). Over some general classes of graphs (subsuming the behavior of message-passing systems and shared-memory systems), we provide results in the spirit of the Büchi-Elgot-Trakhtenbrot Theorem, showing an effective equivalence between automata and (the existential fragment of) MSO logic.

1.4 Model Checking

While realizability asks for an implementation of a given specification, model checking is the process of verifying the specification against a given system model. It has its root in the seminal works by Clarke & Emerson [CE81] and Queille & Sifakis [QS82]. Model checking is by now very well understood as far as finite-state sequential systems are concerned, and theoretical results have turned into tools
successfully employed in practice (e.g., [Hol03]). On the other hand, processes communicating via message passing through unbounded channels and shared-variable recursive processes have an undecidable control-state reachability problem [BZ83, Ram00]. However, similarly to the case of realizability, under- or overapproximating the behavior of a system will still allow us to check certain system properties. In other words, while exploring the set of possible behaviors, one restricts to a subset that is susceptible to containing the relevant critical executions.

For example, in a message-passing system, we may consider only those behaviors that do not exceed a predefined channel bound, and in a concurrent recursive setting, one may impose a bound on contexts (cf. previous section), each allowing only one process to evolve. The context restriction is all the more important (and practically justified) as many system errors can be detected within a few context-switches, as it was argued in [QR05]. Subsequently, other restrictions, relaxing the context-bound and being mutually incomparable, were defined and shown to have a decidable model-checking problem. They impose bounds on the number of phases [LMP07], scopes [LN11], or restrict to certain ordered executions [BCCC96]. In [MP11], it was observed that all of them induce behaviors of bounded tree-width, giving a general explanation of decidability. Indeed, tree-automata techniques play an important role in model checking concurrent recursive programs.

Note that underapproximation techniques can be useful for verifying both positive (or, liveness) and negative (or, safety) specifications. A positive specification contains all the behaviors that a system should exhibit, while a negative specification represents faulty executions. If, now, the underapproximation admits all positive behaviors, then this is actually the case for the original system. Similarly, when the underapproximation intersects the negative specification, then this also applies to the system. Dually, an overapproximation can be used for the following reasoning. If it does not subsume all the positive behaviors, then the system does not exhibit all of them either. Moreover, if the overapproximation does not intersect with the negative behaviors, we can safely assume that the actual system will never exhibit any bad behavior. Overapproximation techniques are usually quite different from underapproximation methods. They have been applied successfully to message-passing systems (for example, assuming lossy channels) [AJ93], and concurrent recursive programs [BET03].

In this work, the focus is on underapproximation techniques for model checking parameterized message-passing systems and context-bounded recursive systems. Our results rely on very different approaches. In the first case, we exploit a realizability result for translating formulas into a system model (which is then compared to the actual system model via product construction). In the latter case, we exploit Hanf’s theorem as well as aforementioned tree-automata techniques.

1.5 Contributions and Outline

The large majority of the results presented in this document have been obtained in collaboration with several co-authors, whom I gratefully acknowledge. All results
have been published in peer-reviewed conference proceedings and journals. Theorems that go back to publications that have been co-authored by myself after my PhD are highlighted in gray. Most proofs are omitted. However, we usually give the proof ideas and references to papers containing full proofs.

Next, we give a summary of the contributions along with the outline of this thesis. Recall that we deal with (a subset of)

\[
\{ \text{realizability} \} \cup \{ \text{model checking} \} \cup \{ \text{static} \} \cup \{ \text{dynamic} \} \cup \{ \text{parameterized} \} \cup \{ \text{finite-state} \} \cup \{ \text{recursive} \} \cup \{ \text{timed} \} \cup \{ \text{sequential} \} \cup \{ \text{shared-memory} \} \cup \{ \text{message-passing} \}
\]

systems. An attribution to chapters can be found in Table 1.1. Chapters are self-contained and can be read independently of each other, though we often point out analogies with other chapters.

Table 1.1: Systems covered by this document

<table>
<thead>
<tr>
<th>Chapter</th>
<th>realizability/verification</th>
<th>static &amp; fixed</th>
<th>recursive</th>
<th>shared-memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 2</td>
<td>realizability/verification</td>
<td>static &amp; fixed</td>
<td>recursive</td>
<td>shared-memory</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>realizability/verification</td>
<td>parameterized</td>
<td>finite-state</td>
<td>message-passing</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>realizability/verification</td>
<td>dynamic</td>
<td>finite-state</td>
<td>sequential</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>realizability/verification</td>
<td>dynamic</td>
<td>finite-state</td>
<td>message-passing</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>verification</td>
<td>static &amp; fixed</td>
<td>timed</td>
<td>shared-memory</td>
</tr>
</tbody>
</table>

Chapter 2. Chapter 2 studies static recursive shared-memory systems. Both realizability and model checking are addressed. As specification formalisms, we consider (E)MSO logic and MSO-definable temporal logics, as well as (multiply) nested-word automata (NWAs), a kind of multi-stack automata with one single state space. As a model of an implementation, we introduce nested-trace automata (NTAs), which have distributed state spaces and generalize Zielonka automata to a recursive setting. Our main contributions in this chapter read as follows:

- We show that Zielonka’s theorem carries over to concurrent recursive systems: any global specification (given by an NWA) that is closed under permutation of independent events can be translated into an equivalent NTA. As closure under independence rewriting is undecidable, we provide a variant of that theorem. It states that certain bounded representations, which can be considered as incomplete specifications, have a distributed counterpart in terms of an NTA. Here, “bounded” may refer to the number of contexts or phases. We show that the property of being a bounded representation is, in each case, decidable.

- We give logical characterizations of NTAs where behaviors are bounded. This extends work by Thomas [Tho90], who had established expressive equivalence
of MSO logic and (non-recursive) Zielonka automata. We complement our result by showing that MSO logic is strictly more expressive than NTAs over the domain of all nested words. Actually, this is obtained as a corollary from a characterization of NTAs in terms of EMSO logic and strictness of the monadic second-order quantifier-alternation hierarchy over multiply nested words.

- We determine the precise complexity of the emptiness problem for ordered NWAs. More precisely, the problem is shown to be 2-EXPTIME-complete.

- We show that model checking NTAs is decidable in elementary time for every MSO-definable temporal logic over nested words. Actually, the problem is solvable in \((n+2)\)-EXPTIME where \(n\) is the maximal level of the MSO modalities in the monadic second-order quantifier-alternation hierarchy. We also establish an \(n\)-EXPSPACE lower bound for some MSO-definable temporal logic.

The chapter refers to the following references:


Chapter 3. In this chapter, we provide a simple yet natural model of parameterized message-passing systems, which we call parameterized communicating automata (PCAs). PCAs are a conservative extension of the classical communicating automata [BZ83]. They also have a static communication topology and accept message sequence charts (MSCs), which look like the graph depicted on the right in Figure 1.1. However, the precise communication topology is henceforth unknown so that it becomes a parameter of the system. Most work in the literature has focused on the verification of such systems, i.e., on the question “Does the given system satisfy some property independently of the topology or the number of processes?”.
Here, we will initiate a study of the language theory of parameterized systems and, in particular, the realizability question. Combined with emptiness-checking algorithms, this eventually allows us to show decidability of model checking parameterized systems against MSO logic over MSCs, which is more expressive than the (temporal) logics that have been previously considered.

Our main contributions from this chapter can be summarized as follows:

- We provide several results in the spirit of the Büchi-Elgot-Trakhtenbrot Theorem, which roughly read as follows: Given a set of communication topologies \( \mathcal{T} \) (e.g., the class of pipelines, ranked trees, grids, or rings) and a high-level specification \( \varphi \) from some logic, there is a PCA that accepts precisely the models of \( \varphi \) when it is run on a topology from \( \mathcal{T} \).

- Inspired by the restrictions introduced for recursive systems, we study context-bounded verification of PCAs to overcome the inherent undecidability coming with that model. Using the previous logical characterization and an emptiness-checking algorithm, we obtain decidability of model checking of context-bounded PCAs.

The chapter refers to the following articles:


**Chapter 4.** This chapter considers sequential systems rather than concurrent ones. However, it lays some foundations that will be exploited later in a dynamic concurrent setting. In particular, it prepares some mechanism to deal with an unbounded number of processes and an infinite supply of process identifiers. Accordingly, finite automata are extended to run over infinite alphabets, whose elements are often referred to as data values. To do so, an automaton is equipped with registers as storing capabilities for data values. Many such models have been defined in the literature, each having its advantages and drawbacks. Here, we introduce two further models. One generalizes previously introduced models, one is a restriction, which we call session automata. More precisely, this chapter deals with the following issues:
• We consider the model-checking problem for session automata. Unlike previous approaches, which can only handle weak fragments of data MSO logic (featuring an equality predicate for data values), we allow the full MSO logic as a specification language. On the other hand, we restrict the automata model to gain decidability of model checking. Moreover, we show that session automata, unlike many other classes of register automata, form a robust language class: they are closed under boolean operations (complementation wrt. the class of certain bounded words) and have a decidable inclusion problem.

• Moreover, we introduce class register automata, a “realistic” model of register automata. Though it has an undecidable emptiness problem, it allows one to reflect reading capabilities of automata (such as receive a message or return from a procedure) that will later be transferred to a concurrent setting. We show that the expressiveness of class register automata is between EMSO logic and MSO logic, solving the realizability problem for EMSO logic, and demonstrating that, somehow, class register automata recognize only “regular” languages over infinite alphabets.

The chapter covers the references indicated below. Note that, in the present manuscript, we present a restricted framework compared to the original publications. The latter deal with stacks, which we omit here for simplicity, and multidimensional data words, while we restrict to one data value per event.


Chapter 5. In this chapter, we study dynamic message-passing systems. Like Chapter 3, this extends classical communicating automata, but in an orthogonal way. Moreover, we consider different questions since the results from Chapter 3 do not carry over to the dynamic setting. We focus on realizability, but we also touch the model-checking problem. Our contributions are as follows:

• We introduce a dynamic version of communicating automata. Like the static variant, dynamic communicating automata (DCAs) feature concurrent processes that communicate via message exchange. However, new processes may be generated at runtime and the number of processes that participate in an
execution is a priori unbounded. Processes are identified by unique process identifiers (pids), which they can store in registers and pass to other processes via messages. In a natural way, DCAs recognize dynamic MSCs, which are MSCs extended by dynamic process creation.

- We study the realizability questions for high-level MSCs with registers, which are inspired by the model of session automata presented in Chapter 4. Now, dynamic process creation raises new problems and questions. In order for a language to be realizable, one has to guarantee that processes that communicate with one another know each other. In particular, there must be a way of communicating pids in such a way that a sender holds the pid of a receiver in one of its registers at the time of communication. To our knowledge, this question has not been considered before. Unfortunately, realizability (actually, we consider a variant called implementability) of high-level MSCs in terms of DCAs is undecidable. Therefore, we study a decidable sufficient criterion, which we call executability. We then define a subclass of high-level MSCs for which executability and implementability are equivalent. The proof of that result yields an effective construction of a DCA out of a high-level MSC, provided that the latter is executable.

The chapter covers the following references:


Chapter 6. This chapter studies timed systems with a static communication topology. The model we introduce is kind of unusual, since it does not assume that clocks of the different processes evolve at the same speed. Indeed, it is often more appropriate to relax the assumption of perfectly synchronous clocks. So, we extend the classical notion of a timed automaton with an equivalence relation on the set of clocks. Roughly speaking, we assume that there is no way to predict the relative clock speed of clocks that are not equivalent (or, do not belong to the same process). Though relative clock speeds are not predictable, or, not robust, we are looking for a way to verify such timed automata with independently evolving clocks. The idea is to consider an under- and an overapproximation of the actual system behavior, which includes behaviors that are executable under every and, respectively, some evolution of local times. Note that we had already considered underapproximations in Chapters 2 and 3. However, the reason there to look at behavioral restrictions was undecidability of the general problem, rather than unpredictability. The chapter refers to the following article:
Conclusion 7. In the last chapter, we conclude and give general directions and challenges for future work.

1.6 Further Contributions

Some of the author’s contributions are perfectly in the scope of this document, but have not been considered in order to keep the presentation simple and homogenous. These contributions read as follows:

- We establish a logical characterization of Muller communicating automata, producing infinite executions. The result makes use of an extension of Hanf’s Theorem to the infinitary quantifier [BK08].

- We also extend the logical characterization to a timed message-passing setting. Here, every process is an event-clock automaton (a subclass of timed automata). In addition, timing constraints can be imposed on the delay between sending and receiving a message. While timed automata accept timed words, event-clock message-passing automata accept timed MSCs, where each event comes with an additional time stamp. As a specification language, MSO logic is naturally extended by timing constraints [ABG13].

- We consider PDL logic for both realizability and model checking of message-passing systems. Concerning realizability, we show that every formula can be translated into a communicating automaton. Moreover, model checking communicating automata and high-level MSCs against PDL specifications are shown to be PSPACE-complete [BKM10].

- As an alternative approach to realizability, we sketch a learning framework for communicating automata. Based on Angluin’s L*-algorithm, this approach allows one to infer message-passing systems simply by classifying MSCs as positive and negative [BK08].

The references refer to the following papers:


1.7 Notation

The set of natural numbers is denoted by \( \mathbb{N} = \{0, 1, 2, 3, \ldots \} \) and the set of positive natural numbers by \( \mathbb{N}_{>0} = \{1, 2, 3, \ldots \} \). For \( n \in \mathbb{N} \), we let \([n] \overset{\text{def}}{=} \{1, \ldots, n\}\). An alphabet is any nonempty set. In this thesis, we will deal with both finite and infinite alphabets. Given an alphabet \( \Sigma \), we denote by \( \Sigma^* \) the set of finite words (or, strings) over \( \Sigma \). The empty word is denoted \( \varepsilon \). Moreover, \( \Sigma^+ \overset{\text{def}}{=} \Sigma^* \setminus \{\varepsilon\} \).

For two sets \( A \) and \( B \), we let \([A \rightarrow B]\) be the set of partial mappings from \( A \) to \( B \). Let \( f : A \rightarrow B \) be a total or partial function, \( a \in A \), and \( b \in B \). By \( f[a \mapsto b] \), we denote the function that maps \( a \) to \( b \) and coincides with \( f \) on all other arguments. We sometimes write \( f_a \) to denote \( f(a) \). We let \( \text{dom}(f) \) stand for the domain of \( f \), i.e., the set of elements \( a \in A \) such that \( f(a) \) is defined. We may identify \( f \) with the set \( \{(a, f(a)) \mid a \in \text{dom}(f)\} \).
In this chapter, we consider recursive systems with a static and fixed (i.e., known) communication topology. A system is recursive if a process may suspend its activity and invoke (or, call) a subtask, before proceeding (or, returning). A stack is used to store the current configuration of the process, and to recover it, once the subtask has been completed. The programs we consider are boolean, i.e., possible variables range over a finite domain. In particular, the set of states and the alphabet of stack symbols are finite. Boolean programs, in turn, may arise as abstractions from programs with infinite-domain variables.

As discussed in the introduction, recursive systems may be defined in several equivalent ways. The classical definition is in terms of a pushdown automaton with explicit stacks. A run is then a sequence of configurations, which keep track of the current stack contents. Alternatively, automata may run on words enriched with nesting relations, one for each process. A nesting relation connects a call position with the corresponding return position. In that case, a transition of the automaton that performs a return will depend on the state reached after executing the corresponding call position. It is then sufficient to define a run as a simple sequence of states (without explicitly mentioning the stack contents). This is the approach adopted in this thesis.

2.1 (Multiply) Nested Words and Their Automata

Nested words [AM09a] extend a word (or, total order) by one or more nesting relations. They reflect executions of concurrent recursive programs, with the understanding that the nesting relation connects a function call with its corresponding return position. We assume that the event labeling tells us whether we deal with
a call or a return so that the nesting relation is actually uniquely given.

A distributed (visibly pushdown) alphabet is a tuple $\tilde{\Sigma} = (\Sigma, P, \text{type}, \delta)$ where $\Sigma$ is a finite alphabet, $P$ is a nonempty finite set of processes, the mapping $\text{type} : \Sigma \to \{\text{call}, \text{ret}, \text{int}\}$ indicates the type of an action (call, return, internal), and $\delta : \Sigma \to 2^P$. The intuition is that $\delta(a)$ is the set of processes that are involved in the execution of $a \in \Sigma$. We require that $\delta(a) \neq \emptyset$ for all $a \in \Sigma$, and $|\delta(a)| = 1$ for all $a \in \Sigma_{\text{call}} \cup \Sigma_{\text{ret}}$. The latter condition implies that synchronization is achieved via internal actions only. Note that, however, the concrete distribution of internal actions only matters in Section 2.5, when we define nested Mazurkiewicz traces.

**Definition 2.1 (nested word).** A (multiply) nested word over the distributed alphabet $\tilde{\Sigma}$ is a triple $(E, \prec, \lambda)$ where

- $E$ is a nonempty finite set of events,
- $\lambda : E \to \Sigma$ is the event-labeling function, and
- $\prec = \prec_{+1} \cup \prec_{cr} \subseteq E \times E$ is an acyclic edge relation (denoting direct successors or matching a call with a return)

such that the properties 1.–4. below are satisfied. Hereby, for $\tau \in \{\text{call}, \text{ret}, \text{int}\}$, we let $E_\tau \overset{\text{def}}{=} \{e \in E \mid \tau = \text{type}(\lambda(e))\}$. Moreover, given $p \in P$, set $E_p \overset{\text{def}}{=} \{e \in E \mid p \in $\delta(\lambda(e))$\}$.

1. $\prec \overset{\text{def}}{=} \prec^*$ is a total order on $E$ (with strict part $<$), and $\prec_{+1}$ is the direct successor relation of $\prec$,
2. $\prec_{cr}$ induces a bijection between $E_{\text{call}}$ and $E_{\text{ret}}$,
3. for all $(e, f) \in \prec_{cr}$, there is $p \in P$ such that $e \in E_p$ and $f \in E_p$, and
4. for all $(e_1, f_1), (e_2, f_2) \in \prec_{cr}$ and $p \in P$ such that $e_1 \in E_p$, $e_2 \in E_p$, and $e_1 < e_2 < f_1$, we have $f_2 < f_1$. \hfill \Box

Condition 3. ensures that a call-return edge is restricted to one process, and condition 4. guarantees that calls and returns of one process follow a stack policy (in other words, $\prec_{cr} \cap (E_p \times E_p)$ is well-nested for every $p \in P$). Note that the sets $\prec_{+1}$ and $\prec_{cr}$ are not necessarily disjoint.

The set of nested words over $\tilde{\Sigma}$ is denoted by $\text{NW}(\tilde{\Sigma})$. When $\tilde{\Sigma}$ is clear from the context, we may just write $\text{NW}$. We do not distinguish isomorphic nested words.

The term visibly stresses the fact that an action uniquely determines the type and stack involved in its execution. Visibility actually reveals the nesting structure if only the sequence of labels from $\Sigma$ is given. Let us formalize this and let $W = (E, \prec, \lambda)$ be a nested word. Assume, without loss of generality, that $E = \{e_1, \ldots, e_n\}$ with $e_1 \prec_{+1} e_2 \prec_{+1} \ldots \prec_{+1} e_n$. The string of $W$ is defined as $\text{string}(W) \overset{\text{def}}{=} \lambda(e_1) \ldots \lambda(e_n) \in \Sigma^+$. Now, given any word $w \in \Sigma^+$, there is at most one (up to isomorphism) nested word $W$ over $\tilde{\Sigma}$ such that $\text{string}(W) = w$.  

18
If it exists, then we denote it by \textit{nested}(w) and we call \(w\) \textit{well-nested}, since it is guaranteed that calls and returns are well bracketed, for any process/stack. Note that, for any nested word \(W\), we have \textit{nested}(\text{string}(W)) = W.

\textbf{Example 2.2.} Consider the distributed alphabet \(\hat{\Sigma} = (\Sigma, P, \text{type}, \delta)\) given by \(\Sigma = \{a_1, b_1, a_2, b_2, c\}\), \(P = \{1, 2\}\), \(\text{type}(a_1) = \text{type}(a_2) = \text{call}\), \(\text{type}(b_1) = \text{type}(b_2) = \text{ret}\), and \(\text{type}(c) = \text{int}\), as well as \(\delta(a_1) = \delta(b_1) = \{1\}\), \(\delta(a_2) = \delta(b_2) = \{2\}\), and \(\delta(c) = \{1, 2\}\). Figure 2.1 shows the nested word \(W_{2.1}\) over \(\hat{\Sigma}\). Straight edges represent the relation \(<_{+1}\), whereas curved edges represent \(<_{cr}\). We have \(\text{string}(W_{2.1}) = a_1ca_2a_1ca_2cb_1cb_1b_2b_2\).

![Figure 2.1: The multiply nested word \(W_{2.1}\)](image)

Now, we consider automata that run on multiply nested words and are suitable as a model of recursive programs. As explained before, these automata run directly over nested words so that the stacks are only implicit.

\textbf{Definition 2.3 (NWA).} A \textit{nested-word automaton (NWA)} over \(\hat{\Sigma}\) is a tuple \(A = (S, \Gamma, \Delta, \iota, F)\) where

\begin{itemize}
  \item \(S\) is the nonempty finite set of states,
  \item \(\Gamma\) is the nonempty finite set of stack symbols,
  \item \(\iota \in S\) is the initial state,
  \item \(F \subseteq S\) is the set of final states, and
  \item \(\Delta = \Delta_{\text{call}} \uplus \Delta_{\text{ret}} \uplus \Delta_{\text{int}}\) is the transition relation, partitioned into
    \begin{itemize}
      \item \(\Delta_{\text{call}} \subseteq S \times \Sigma_{\text{call}} \times \Gamma \times S\),
      \item \(\Delta_{\text{ret}} \subseteq S \times \Sigma_{\text{ret}} \times \Gamma \times S\), and
      \item \(\Delta_{\text{int}} \subseteq S \times \Sigma_{\text{int}} \times S\).
    \end{itemize}
\end{itemize}

Reading an event \(e\) of a nested word, transition \((s, a, s') \in \Delta_{\text{int}}\) lets the automaton move on from the current state \(s\) to the target state \(s'\) if \(e\) is labeled with letter \(a\). The same applies to a transition \((s, a, A, s') \in \Delta_{\text{call}} \cup \Delta_{\text{ret}}\). However, in addition, this transition allows \(A\) to associate with a call event stack symbol \(A\). The symbol can then be retrieved at the corresponding return position. In a sense, this is equivalent to reading a stack symbol previously pushed. More precisely, we require
that, whenever $e \triangleleft_{cr} f$, the stack symbol chosen at $e$ is the same as the stack symbol employed by the transition that is taken at position $f$.

Let $W = (E, \triangleleft, \lambda)$ be a nested word. Let us, again, assume that $E = \{e_1, \ldots, e_n\}$ with $e_1 \triangleleft e_2 \triangleleft \cdots \triangleleft e_n$. A run of $A$ on $W$ is a pair $(\rho, \sigma)$ of mappings $\rho : E \rightarrow S$ and $\sigma : E_{\text{call}} \cup E_{\text{ret}} \rightarrow \Gamma$ such that,

- for all $(e, f) \in \triangleleft_{cr}$, we have $\sigma(e) = \sigma(f)$, and
- for every $i \in \{1, \ldots, n\}$ (letting $\rho(e_0) = i$),
  \[
  \begin{cases}
  (\rho(e_{i-1}), \lambda(e_i), \rho(e_i)) \in \Delta_{\text{int}} & \text{if } e_i \in E_{\text{int}} \\
  (\rho(e_{i-1}), \lambda(e_i), \sigma(e_i), \rho(e_i)) \in \Delta_{\text{call}} \cup \Delta_{\text{ret}} & \text{if } e_i \in E_{\text{call}} \cup E_{\text{ret}}. 
  \end{cases}
  \]

The run $(\rho, \sigma)$ is accepting if $\rho(e_n) \in F$. The set of multiply nested words for which there is an accepting run is denoted by $L(A)$.

We say that $A$ is deterministic if

- for all $(s, a) \in S \times \Sigma_{\text{call}}$, there is at most one pair $(A, s') \in \Gamma \times S$ such that $(s, a, A, s') \in \Delta_{\text{call}}$.
- for all $(s, a, A) \in S \times \Sigma_{\text{ret}} \times \Gamma$, there is at most one $s' \in S$ such that $(s, a, A, s') \in \Delta_{\text{ret}}$, and
- for all $(s, a) \in S \times \Sigma_{\text{int}}$, there is at most one $s' \in S$ such that $(s, a, s') \in \Delta_{\text{int}}$.

**Example 2.4.** An NWA $A_{2.2} = (\{s_0, \ldots, s_5\}, \{\$\}, \Delta, s_0, \{s_5\})$ over the distributed alphabet $\tilde{\Sigma}$ from Example 2.2 is given in Figure 2.2. Transitions are self-explanatory. For example, $\Delta_{\text{int}}$ contains $(s_3, c, s_4)$, and $\Delta_{\text{ret}}$ contains $(s_4, b_2, \$, s_5)$, which is indicated by the edge from $s_4$ to $s_5$ with label $b_2\$. Note that $A_{2.2}$ accepts those nested words that (i) contain a subsequence $a_1a_2$ or $a_2a_1$, (ii) have a call and a subsequent return phase, separated by some action $c$, and (iii) schedule returns of process 1 before those of process 2. The nested word over $\tilde{\Sigma}$ from Figure 2.1 is accepted by $A_{2.2}$. Note that $A_{2.2}$ is not deterministic.

![Figure 2.2: Nested-word automaton $A_{2.2}$](image)

Unfortunately, in the presence of at least two processes/stacks, most basic verification problems for NWAs are undecidable, such as nonemptiness:

**Theorem 2.5 ([Ram00]).** The following problem is undecidable:

**INSTANCE:** $\tilde{\Sigma}$; NWA $A$ over $\tilde{\Sigma}$

**QUESTION:** $L(A) \neq \emptyset$?
2.2 Contexts, Phases, Scopes, and Ordered Stacks

In the seminal paper [QR05], Qadeer and Rehof exploited the fact that errors of recursive programs typically occur already within a few contexts where a context refers to an execution involving only one process. Imposing a bound on the number of contexts indeed renders many verification problems decidable. In other words, instead of looking at all possible executions, we consider an underapproximation of the actual system behavior. This view is particularly suitable when checking positive specifications: if a system $A$ should exhibit all behaviors given by a set $\text{Good}$, then it is sufficient to show that $\text{Good} \subseteq L(U)$ is true for a restricted version $U$ of $A$ such that the problem $\text{Good} \subseteq L(U)$ is decidable and that we may adjust incrementally so as to converge to $L(A)$. On the other hand, given a negative specification (or, safety property) $\text{Bad}$, we can also draw conclusions from $L(U) \setminus \text{Bad} = \emptyset$, while showing complete absence of bad behaviors would require an overapproximation of the system behavior.

Next, we recall the notion of a context as well as other restrictions that have been defined in the literature. Let $W = (E, \prec, \lambda) \in \text{NW}(\Sigma)$ be a nested word. An interval of $W$ is a set $I \subseteq E$ such that $I = \emptyset$ or $I = [e, f] = \{ g \in E \mid e \leq g \leq f \}$ for some $e, f \in E$. In a context, only one designated process is allowed to call or return, while a phase only restricts return operations:

- A context of $W$ is an interval $I$ such that $I \cap (E_{\text{call}} \cup E_{\text{ret}}) \subseteq E_p$ for some $p \in P$.
- A phase of $W$ is an interval $I$ such that $I \cap E_{\text{ret}} \subseteq E_p$ for some $p \in P$.

Note that every context is a phase, while the converse does not hold in general.

**Definition 2.6 (k-context word [QR05] and k-phase word [LMP07]).**

Let $W = (E, \prec, \lambda) \in \text{NW}(\Sigma)$ be a nested word and $k \geq 1$. We say that $W$ is a $k$-context word ($k$-phase word) if there are contexts (phases, respectively) $I_1, \ldots, I_k$ of $W$ such that $E = I_1 \cup \ldots \cup I_k$.

While every $k$-context word is a $k$-phase word, but not the other way around, two orthogonal restrictions have been defined in terms of bounded scopes and ordered nested words, which we consider next. A scope-bounded word restricts the number of contexts between a push and the corresponding pop operation.

**Definition 2.7 (k-scope word [LN11]).** Let $W = (E, \prec, \lambda) \in \text{NW}(\Sigma)$ be a nested word and $k \geq 1$. We call $W$ a $k$-scope word if, for all $(e, f) \in \prec_{\text{cr}}$, there exist contexts $I_1, \ldots, I_k$ of $W$ such that we have $(e, f) = I_1 \cup \ldots \cup I_k$.

Finally, in an ordered word, we refer to a total ordering $\preceq$ on $P$ (with irreflexive part $\prec$). Then, a pop operation can only be performed by the process that is minimal among all processes with a nonempty stack.
**Definition 2.8 (ordered word [BCCC96]).** Let \( W = (E, \preceq, \lambda) \in \text{NW}(\overline{\Sigma}) \). We call \( W \) an ordered word (wrt. \( \preceq \)) if, for all \( p, p' \in P, e, f \in E_p, \) and \( f' \in E_{\text{ret}} \cap E_{p'} \) such that \( e \preceq_{\text{cr}} f \) and \( e < f' < f \), we have \( p' \preceq p \). \( \Diamond \)

**Example 2.9 (continues Example 2.2).** Figure 2.3 illustrates the concepts introduced above by means of the nested word \( W_{2.1} \) from Figure 2.1. The shadowed areas are dedicated to process 2. Thus, \( W_{2.1} \) is a 6-context word, a 5-scope word, and a 2-phase word. All these bounds are optimal. Moreover, \( W_{2.1} \) is an ordered word under the assumption \( 1 < 2 \) (but not for \( 2 < 1 \)). \( \Diamond \)

![Figure 2.3: Illustration of context, phase, and scope](image)

Let \( R \overset{\text{def}}{=} \{ k\text{-cnt}, k\text{-ph}, k\text{-scp} \mid k \geq 1 \} \cup \{ \text{ord} \} \) be the set of possible “restrictions”. For \( \theta \in R \), the set of \( \theta \)-words is denoted by \( \text{NW}_\theta(\overline{\Sigma}) \). In particular, by \( \text{NW}_{\text{ord}}(\overline{\Sigma}) \), we denote the set of ordered words, silently assuming an order on \( P \). Moreover, for an NWA \( A \), we let \( L_\theta(A) \overset{\text{def}}{=} L(A) \cap \text{NW}_\theta(\overline{\Sigma}) \). When convenient, we omit the reference to \( \overline{\Sigma} \) and write, for example, \( \text{NW}_{k\text{-ph}} \).

**Example 2.10.** Consider the NWA \( A = A_{2.2} \) from Figure 2.2. We have that \( L(A) = L_{2\text{-ph}}(A) = L_{\text{ord}}(A) \). On the other hand, \( L_{k\text{-cnt}}(A) \) and \( L_{k\text{-scp}}(A) \) are strictly included in \( L(A) \), for all \( k \geq 1 \). \( \Diamond \)

Let us define the respective nonemptiness problems:

**Problem 2.11. Ord-Nonemptiness:**
- **Instance:** \( \overline{\Sigma} \); NWA \( A \) (over \( \overline{\Sigma} \))
- **Question:** \( L_{\text{ord}}(A) \neq \emptyset ? \)

**Problem 2.12. Context-Nonemptiness:**
- **Instance:** \( \overline{\Sigma} \); NWA \( A \); \( k \geq 1 \)
- **Question:** \( L_{k\text{-cnt}}(A) \neq \emptyset ? \)

**Problem 2.13. Phase-Nonemptiness:**
- **Instance:** \( \overline{\Sigma} \); NWA \( A \); \( k \geq 1 \)
- **Question:** \( L_{k\text{-ph}}(A) \neq \emptyset ? \)

**Problem 2.14. Scope-Nonemptiness:**
- **Instance:** \( \overline{\Sigma} \); NWA \( A \); \( k \geq 1 \)
- **Question:** \( L_{k\text{-scp}}(A) \neq \emptyset ? \)

Indeed, all these problems are decidable with varying complexities. We assume that the parameter \( k \) is given in unary. For a comparison between the effects of a unary and a binary encoding, see [BD13].

**Theorem 2.15 ([QR05, LTKR08]).** Context-Nonemptiness is NP-complete.
Theorem 2.16 ([LMP07, LMP08b]). Phase-Nonemptiness is 2-Exptime-complete.

Theorem 2.17 ([LN11]). Scope-Nonemptiness is Pspace-complete.

Theorem 2.18 ([ABH08]). Ord-Nonemptiness is 2-Exptime-complete.

In [MP11], a uniform argument for decidability of the above problems was given by Madhusudan and Parlato: for $\theta \in \{k\text{-cnt}, k\text{-ph} \mid k \geq 1\} \cup \{\text{ord}\}$, the class of $\theta$-words has bounded tree width. By Courcelle’s theorem, this implies that nonemptiness of NWAs and even model-checking of NWAs against MSO properties is decidable. It was then shown that scope-bounded words have bounded tree width, too [LP12].

An alternative unifying approach is given by Cyriac, Gastin, and Narayan Kumar [CGNK12] in terms of the notion of split-width. Since split-width is tailored to nested words, it is simpler than tree width and allows for a more intuitive understanding of a class of bounded multi-threaded programs. Note that split-width also led to generalizations of the above-mentioned classes and other existing work on recursive message-passing systems [AGNK14b, Cyr14].

2.3 Monadic Second-Order Logic

We give a short account of monadic second-order (MSO) logic over nested words. The logic is built over countably infinite supplies \(\{x, y, x_1, x_2, \ldots\}\) of first-order and \(\{X, Y, X_1, X_2, \ldots\}\) of second-order variables. First-order variables are interpreted as events, second-order variables as sets of events. As usual, the predicates available in MSO logic depend on the signature of a structure, which is given in terms of the distributed alphabet \(\tilde{\Sigma} = (\Sigma, P, \text{type}, \delta)\).

Definition 2.19. The formulas from \(\text{nwMSO}(\tilde{\Sigma})\) are built according to the following grammar:

\[
\varphi ::= a(x) \mid x <_{\alpha+1} y \mid x <_{\alpha} y \mid x = y \mid x \in X \mid \\
\neg \varphi \mid \varphi_1 \lor \varphi_2 \mid \exists x. \varphi \mid \exists X. \varphi
\]

where \(a \in \Sigma\), \(x, y\) are first-order variables, and \(X\) is a second-order variable.

Throughout this document, we use the usual abbreviations such as \(\varphi_1 \land \varphi_2\) for \(\neg(\neg \varphi_1 \lor \neg \varphi_2)\), \(\varphi \rightarrow \psi\) for \(\neg \varphi \lor \psi\) and \(\forall x. \varphi\) for \(\neg \exists x. \neg \varphi\).

A formula \(\varphi \in \text{nwMSO}(\tilde{\Sigma})\) is interpreted over a nested word \(W = (E, <, \lambda) \in \text{NW}(\tilde{\Sigma})\) wrt. a mapping \(I\). The purpose of the latter is to interpret free variables. It maps any first-order variable \(x\) to an event \(I(x) \in E\) and any second-order
variable $X$ to a set of events $\mathcal{I}(X) \subseteq E$. We write $W \models_\mathcal{I} \varphi$ if formula $\varphi$ is evaluated to true when the free variables of $\varphi$ are interpreted according to $\mathcal{I}$. In particular, we have $W \models_\mathcal{I} a(x)$ if $\lambda(\mathcal{I}(x)) = a$, and $W \models_\mathcal{I} x \prec_{\text{cr}} y$ if $\mathcal{I}(x) \prec_{\text{cr}} \mathcal{I}(y)$. The remaining operators are as expected. When $\varphi$ is a sentence, i.e., a formula without free variables, we may omit the index $\mathcal{I}$ and simply write $W \models \varphi$. For an introduction to MSO logic on (classical) words and its semantics, we refer the reader to [Tho97].

The fragment $\text{nwEMSO}(\tilde{\Sigma})$ contains the existential MSO (EMSO) formulas, which are of the form $\exists X_1 \ldots \exists X_n.\varphi$ where $\varphi$ does not make use of second-order quantification.

2.4 The Language Theory of Nested-Word Automata

Let us examine closure properties of NWAs as well as their expressive power compared to MSO logic. One may call a language (or automata) class robust when it is effectively closed under all Boolean operations and has a decidable emptiness problem. A robust language class is amenable to verification tasks such as model checking, since the inclusion problem is decidable.

It is easy to see that NWAs are closed under union and intersection. This has to be seen in contrast to the fact that the class of context-free languages (CFLs) is not closed under intersection. The intuitive reason for this is that, in CFLs, actions are no longer visible. Indeed, augmenting words by the nesting structure makes sure that, while simulating two NWAs, the automaton for intersection performs push and pop operations at the same positions of the input word.

However, NWAs do not form a robust automata class, as their emptiness problem is undecidable (Theorem 2.5). This even applies when we restrict to two stacks. However, in the presence of at most two stacks, we can establish a logical characterization:

**Theorem 2.20 ([Bol08]).** Suppose $|P| = 2$ and let $L \subseteq \text{NW}(\tilde{\Sigma})$. The following statements are effectively equivalent:

1. There is an NWA $A$ over $\tilde{\Sigma}$ such that $L(A) = L$.
2. There is a sentence $\varphi \in \text{nwEMSO}(\tilde{\Sigma})$ such that $L(\varphi) = L$.

In [Bol08], it is also shown that the monadic quantifier-alternation hierarchy over nested words is infinite, by an encoding of grids into nested words. We obtain, as a corollary, that NWAs are not complementable:
We have seen that restricting the domain of nested words appropriately, renders the emptiness problem for NWAs decidable. It has been shown that above restrictions actually give rise to robust language classes. The next theorem states that, under any of the restrictions introduced above, NWAs are complementable:

**Theorem 2.22 (complementability [LMP07, LNP14a, LNP14b]).** Let $\theta \in \mathfrak{R}$. Then, for all NWAs $A$ over $\tilde{\Sigma}$, the following hold:

1. There exists an NWA $A'$ over $\tilde{\Sigma}$ such that $L(A') = \text{NW}_{\theta}(\tilde{\Sigma}) \setminus L(A)$.
2. There exists an NWA $A'$ over $\tilde{\Sigma}$ such that $L(A') = \text{NW}(\tilde{\Sigma}) \setminus \text{NW}(A)$.

For every $\theta \in \mathfrak{R}$, the proof more or less follows the following schema: A nested word $W \in \text{NW}_\theta(\Sigma)$ is encoded by means of an injective mapping as a ranked tree $\text{tree}(W)$. One constructs a tree automaton $B$ recognizing $\text{tree}(\text{NW}_\theta(A))$ and a tree automaton $B_A$ recognizing $\text{tree}(L_\theta(A))$. Since tree automata are complementable, there is a tree automaton $B_A'$ for the complement of $L(B_A)$. From $B_A' \times B$, one then extracts an NWA $A'$ such that $L(A') = \text{tree}^{-1}(L(B_A' \times B)) = \text{NW}_\theta(\Sigma) \setminus L(A)$.

Note that, in Theorem 2.22, statement 1. is implied by 2. and vice versa by the following lemma:

**Lemma 2.23.** For all $\theta \in \mathfrak{R}$, there are deterministic NWAs $A_\theta$ and $A'_\theta$ over $\tilde{\Sigma}$ such that $L(A_\theta) = \text{NW}_\theta(\tilde{\Sigma})$ and $L(A'_\theta) = \text{NW}(\tilde{\Sigma}) \setminus \text{NW}_\theta(\tilde{\Sigma})$.

**Proof (sketch).** The lemma is easily seen for contexts and phases. In the case of ordered words, one has to keep track of the stacks that currently have a stack symbol pushed. For $k$-scope words, one verifies that all call-return edges span over at most $k$ contexts (some call-return edge spans over more that $k$ contexts, respectively). □

While phase-bound and ordered NWAs are in general not determinizable [LMP07], the other restrictions allow for the following theorem:

**Theorem 2.24 ([LMP10a, LNP14a]).** Let $\theta \in \{ k\text{-cnt} \mid k \geq 1 \}$. Then, for all NWAs $A$ over $\tilde{\Sigma}$, there exists a deterministic NWA $A'$ over $\tilde{\Sigma}$ such that $L(A') = L_\theta(A)$.

Moreover, all of the above restrictions give rise to a characterization in terms of MSO logic $\text{nwMSO}(\tilde{\Sigma})$. The proof essentially uses Theorem 2.22, which reduces negation in MSO logic to complementation in NWAs.
Table 2.1: Closure, decidability, and expressiveness properties for NWA

<table>
<thead>
<tr>
<th></th>
<th>⋃</th>
<th>⋂</th>
<th>Compl.</th>
<th>Det.</th>
<th>Emptiness</th>
<th>Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>RegL</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>NLOG</td>
<td>MSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Büc60, Elg61, Tra62]</td>
</tr>
<tr>
<td>CFL</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>P</td>
<td>ψMatchFO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[LST99]</td>
</tr>
<tr>
<td>1-NWA (</td>
<td>P</td>
<td>= 1)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[AM09a]</td>
</tr>
<tr>
<td>2-NWA (</td>
<td>P</td>
<td>= 2)</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[AM09a]</td>
</tr>
<tr>
<td>k-context NWA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>NP-c</td>
<td>MSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[LMP07, LMP08]</td>
</tr>
<tr>
<td>k-phase NWA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>2-EXPTIME-c</td>
<td>MSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[LMP07, LMP08]</td>
</tr>
<tr>
<td>k-scope NWA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>PSPACE-c</td>
<td>MSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[LNP14a]</td>
</tr>
<tr>
<td>ordered NWA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>2-EXPTIME-c</td>
<td>MSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[LNP14a]</td>
</tr>
</tbody>
</table>

Theorem 2.25 ([LMP07, LNP14a, LNP14b]). Let \( \theta \in \mathcal{R} \). Then, for all \( L \subseteq \text{NW}_\theta(\Sigma) \), the following statements are equivalent:

- There is an NWA \( A \) over \( \hat{\Sigma} \) such that \( L(A) = L \).
- There is a sentence \( \varphi \in \text{nwMSO}(\hat{\Sigma}) \) such that \( L(\varphi) = L \).

From Theorems 2.15–2.18 and Theorem 2.25, one can infer that model checking bounded NWAs against MSO properties is decidable: Given \( \theta \in \mathcal{R} \), an NWA \( A \) over \( \hat{\Sigma} \), and a sentence \( \varphi \in \text{nwMSO}(\hat{\Sigma}) \), do we have \( L_\theta(A) \subseteq L(\varphi) \)?

Table 2.1 summarizes decidability, closure, and expressiveness properties for NWA. In view of these results, it is justified to call NWA a robust automata class when considered over the classes of ordered words and k-context/phase/scope words. This makes them appealing in several verification settings as well as for language-theoretic considerations. Their robustness is underpinned by the results from the next section, which lift classical theorems from Mazurkiewicz trace theory to the recursive setting.

2.5 Nested Traces and Their Automata

Multiply nested words account for the interweaving of calls and returns performed by several processes. However, they do not reflect the inherent concurrency of events that are involved in distinct processes. A natural model of executions of concurrent recursive programs are nested traces, which combine nested words and Mazurkiewicz traces. Thus, we fix a distributed alphabet \( \hat{\Sigma} = (\Sigma, P, \text{type}, \delta) \) (cf.
Section 2.5). The definition of a nested trace is very similar to that of a nested word (cf. Definition 2.1 on page 18), but instead of just one direct-successor relation $\triangleleft_{+1}$, there is one such relation $\triangleleft_p$ for every process $p \in P$.

**Definition 2.26 (nested trace [BGH09]).** A nested trace over the distributed alphabet $\bar{\Sigma}$ is a triple $(E, \triangleleft, \lambda)$ where

- $E$ is a nonempty finite set of events,
- $\lambda : E \rightarrow \Sigma$ is the event-labeling function, and
- $\triangleleft = \triangleleft_{cr} \cup \bigcup_{p \in P} \triangleleft_p \subseteq E \times E$ is an acyclic edge relation

such that properties 1.–4. below are satisfied. Here, the sets $E_p$, $E_{\text{call}}$, and $E_{\text{ret}}$ are given as in Definition 2.1.

1. For all $p \in P$, $\triangleleft_p \overset{\text{def}}{=} \triangleleft_p \overset{\text{def}}{=}$ is a total order on $E_p$ (with strict part $\triangleleft_p$), and $\triangleleft_p$ is the direct successor relation of $\triangleleft_p$.
2. $\triangleleft_{cr}$ induces a bijection between $E_{\text{call}}$ and $E_{\text{ret}}$.
3. for all $(e, f) \in \triangleleft_{cr}$, there is $p \in P$ such that $e \in E_p$ and $f \in E_p$, and
4. for all $(e_1, f_1), (e_2, f_2) \in \triangleleft_{cr}$ and $p \in P$ such that $e_1 \in E_p$, $e_2 \in E_p$, and $e_1 <_p e_2 <_p f_1$, we have $f_2 <_p f_1$. $\diamond$

The set of nested traces over $\bar{\Sigma}$ is denoted by $NTr(\bar{\Sigma})$. Again, we do not distinguish isomorphic nested traces.

**Example 2.27.** Consider the distributed alphabet $\bar{\Sigma}$ from Example 2.2 on page 19, with $P = \{1, 2\}$. A nested trace over $\bar{\Sigma}$ is depicted in Figure 2.4. $\diamond$

![Figure 2.4: The nested trace $T_{2.4}$](image)

Next, we define a distributed automata model running over nested traces, originally introduced in [BGH09]. Here, *distributed* means that every process has its own local state space and transition relation. More precisely, there is a transition relation for every action $a \in \Sigma$. A transition involving $a$ can access the local state of any process from $\delta(a)$, and modify it. Since $a \in \Sigma_{\text{call}} \cup \Sigma_{\text{ret}}$ implies $|\delta(a)| = 1$, executing call and return actions updates the local state of only one single component.
Definition 2.28 (NTA [BGH09]). A nested-trace automaton (NTA) over $\tilde{\Sigma}$ is a tuple $C = ((S_p)_{p \in P}, \Gamma, \Delta, t, F)$. Here, the $S_p$ are disjoint finite sets of local states ($S_p$ containing the local states of process $p$), and $\Gamma$ is the nonempty finite set of stack symbols. Given a set $P' \subseteq P$, let $S_{P'}$ denote the cartesian product $\prod_{p \in P'} S_p$. The tuple $e \in S_P$ is a global initial state, and $F \subseteq S_P$ is the set of global final states. Finally, $\Delta = \Delta_{\text{call}} \cup \Delta_{\text{ret}} \cup \Delta_{\text{int}}$ is the transition relation, partitioned into

- $\Delta_{\text{call}} \subseteq \bigcup_{p \in P'} (S_p \times \Sigma_{\text{call}} \times \Gamma \times S_p)$,
- $\Delta_{\text{ret}} \subseteq \bigcup_{p \in P'} (S_p \times \Sigma_{\text{ret}} \times \Gamma \times S_p)$, and
- $\Delta_{\text{int}} \subseteq \bigcup_{a \in \Sigma_{\text{int}}} (S_{\delta(a)} \times \{a\} \times S_{\delta(a)})$.

Let $S = \bigcup_{p \in P} S_p$. For $s \in S$ and $p \in P$, we let $s_p$ be the $p$-th component of $s$ (if it exists). A run of the NTA $C$ on a nested trace $T = (E, \prec, \lambda)$ will include a mapping $\rho : E \rightarrow S$ such that $\rho(e) \in S_{\delta(\lambda(e))}$ for all $e \in E$. Intuitively, for process $p \in \delta(\lambda(e))$, the component $\rho(e)_p$ is the state that $p$ reaches after executing $e$. Before we specify when $\rho$ is actually a run, let us define another mapping $\rho^- : E \rightarrow S$ that also satisfies $\rho^-(e) \in S_{\delta(\lambda(e))}$ for all $e \in E$. Intuitively, $\rho^-(e)$ collects the current source states of those processes that are involved in executing $e$. We let $\rho^-(e) = (s_p)_{p \in \delta(\lambda(e))}$ where

$$s_p = \begin{cases} t_p & \text{if } e \text{ is } \prec_p \text{-minimal} \\ \rho(f)_p & \text{if } f \prec_p e. \end{cases}$$

Recall that, hereby, $\rho(f)_p$ denotes the $p$-th component of $\rho(f) \in S_{\delta(\lambda(f))}$. This component indeed exists since $p \in \delta(\lambda(f))$. Now, we call the pair $(\rho, \sigma)$, with $\sigma : E_{\text{call}} \cup E_{\text{ret}} \rightarrow \Gamma$, a run of $C$ on $T$ if,

- for all $(e, f) \in \prec_{\text{cr}}$, we have $\sigma(e) = \sigma(f)$, and
- for all $e \in E$, it holds

$$\begin{align*}
(\rho^-(e)_p, \lambda(e), \sigma(e), \rho(e)_p) & \in \Delta_{\text{call}} & \text{if } e \in E_{\text{call}} \cap E_p \\
(\rho^-(e)_p, \lambda(e), \sigma(e), \rho(e)_p) & \in \Delta_{\text{ret}} & \text{if } e \in E_{\text{ret}} \cap E_p \\
(\rho^-(e), \lambda(e), \rho(e)) & \in \Delta_{\text{int}} & \text{if } e \in E_{\text{int}}.
\end{align*}$$

To determine if run $(\rho, \sigma)$ is accepting, we look at the global final state reached at the end of a run. It collects, for all $p \in P$, the local state associated with the $\prec_p$-maximal event, or $t_p$ if $E_p = \emptyset$. Formally, let $t = (t_p)_{p \in P} \in S_P$ be given by

$$t_p = \begin{cases} \rho(e)_p & \text{if } e \text{ is } \prec_p \text{-maximal} \\ t_p & \text{if } E_p = \emptyset. \end{cases}$$

We call $(\rho, \sigma)$ accepting if $t \in F$. Finally, we denote by $L(C)$ the set of nested traces over $\tilde{\Sigma}$ that come with an accepting run of $C$. 

28
Example 2.29. Again, we assume the distributed alphabet $\tilde{\Sigma}$ from Example 2.2, with $P = \{1, 2\}$. Consider the NTA $C_{2.5} = ((S_p)_{p \in P}, \{\$\}, \Delta, \iota, F)$ over $\tilde{\Sigma}$ depicted in Figure 2.5. Its components are given by $S_1 = \{s_0, s_1, s_2\}$, $S_2 = \{t_0, t_1, t_2\}$, $\iota = (s_0, t_0)$, $F = \{(s_2, t_2)\}$. The new feature (compared to NWAs) are the synchronizing transitions from $\Delta_{\text{int}}$, which contains $((s_i, s_i), c, (s_i, s_i))$ for all $i \in \{0, 1, 2\}$ as well as $((s_1, t_1), c, (s_2, t_2))$. Note that $C_{2.5}$ is eventually forced to execute “concurrent” occurrences of $a_1$ and $a_2$: When one process moves on to $s_1$ or $t_1$, then executing $c$ is no longer possible unless the other process catches up. The nested trace $T_{2.4}$ on page 27 is accepted by $C_{2.5}$. Note that, there are indeed concurrent occurrences of $a_1$ and $a_2$. 

A special case of NTAs is given when $\Sigma = \Sigma_{\text{int}}$. Then, an NTA is precisely an asynchronous automaton (also Zielonka automaton) [Zie87], and a nested trace is actually a Mazurkiewicz trace. However, according to our terminology, we rather call it a trace automaton. Since all actions from $\Sigma$ have the same type, the language of a trace automaton may be seen as a set of Mazurkiewicz traces [DR95].

2.6 Realizability of NWA Specifications

Note that NTAs are a truly concurrent model as processes may move independently unless they perform synchronizing actions from $\Sigma_{\text{int}}$. On the other hand, NWAs possess one single state space, and the global control may enforce an order even on a priori independent actions. Yet, there are tight connections between NTAs and NWAs. To reveal that relation, we will first study more closely the relation between nested words and nested traces.

First, we can naturally associate with a nested trace $T = (E, \sqsubseteq, \lambda)$ over $\tilde{\Sigma}$ its linearizations, each of which fixes an ordering of incomparable (wrt. $\leq$) events. Intuitively, a linearization is one possible scheduling policy of a nested trace. Formally, a linearization of $T$ is any nested word of the form $W = (E, \ll', \lambda) \in \text{NW}(\tilde{\Sigma})$ such that $\leq \subseteq \ll'$. Note that this implies $\llcall = \ll'_{\text{call}}$. For example, the nested word from Figure 2.1 on page 19 is a linearization of the nested trace from Figure 2.4 on page 27. By $\text{lin}(T)$, we denote the set of linearizations of $T$. Given
$W \in \text{NW}(\tilde{\Sigma})$, there is a unique nested trace $T \in \text{NT}(\tilde{\Sigma})$ such that $W \in \text{lin}(T)$. We denote this trace by $\text{trace}(W)$. This is extended to languages $L \subseteq \text{NW}(\tilde{\Sigma})$, and we set $\text{trace}(L) = \{ \text{trace}(W) \mid W \in L \}$. The relation between the mappings $\text{string}$ and $\text{nested}$ (from Section 2.1) as well as $\text{lin}$ and $\text{trace}$ is illustrated in Figure 2.6.

Since $\text{nested}(\text{string}(W)) = W$ for any nested word $W$, we may consider the set of linearizations of $T$ as a string language over $\Sigma$. This will facilitate some definitions when we consider languages up to a congruence relation taking into account that some actions are independent (those that do not share processes), while others are not (those that have at least one process in common).

Given the distributed alphabet $\tilde{\Sigma}$, we define an independence relation $I_{\tilde{\Sigma}} \overset{\text{def}}{=} \{(a, b) \in \Sigma \times \Sigma \mid \delta(a) \cap \delta(b) = \emptyset\}$. Note that $I_{\tilde{\Sigma}}$ is irreflexive and symmetric. Its complement, the dependence relation $D_{\tilde{\Sigma}} = (\Sigma \times \Sigma) \setminus I_{\tilde{\Sigma}}$, is, therefore, reflexive and symmetric.

With this, let $\sim_{\tilde{\Sigma}} \subseteq \Sigma^* \times \Sigma^*$ be the least congruence relation that satisfies $ab \sim_{\tilde{\Sigma}} ba$ for all $(a, b) \in I_{\tilde{\Sigma}}$. For example, if $\tilde{\Sigma}$ is the distributed alphabet from Example 2.2 on page 19, then $\{ a_1a_2cb_1b_2, a_2a_1cb_1b_2, a_1a_2cb_1b_2, a_2a_1cb_1b_2 \}$ is an equivalence class of $\sim_{\tilde{\Sigma}}$. The equivalence relation $\sim_{\tilde{\Sigma}}$ is lifted in the natural way to nested words: we let $W \sim_{\tilde{\Sigma}} W'$ if $\text{string}(W) \sim_{\tilde{\Sigma}} \text{string}(W')$. We say that $L \subseteq \text{NW}(\tilde{\Sigma})$ is $\sim_{\tilde{\Sigma}}$-closed if we have $L = [L]_{\tilde{\Sigma}} \overset{\text{def}}{=} \{ W \in \text{NW}(\tilde{\Sigma}) \mid W \sim_{\tilde{\Sigma}} W' \text{ for some } W' \in L \}$.

We will consider an NWA $\mathcal{A}$ to be a specification of a system, and we are looking for a realization or implementation of $\mathcal{A}$, which is provided by an NTA $\mathcal{C}$ such that $L(\mathcal{C}) = \text{trace}(L(\mathcal{A}))$. Actually, specifications often have a “global” view of the system, and the difficult task is to distribute the state space onto the processes, which henceforth communicate in a restricted manner that conforms to the predefined system architecture $\tilde{\Sigma}$. Note that, unlike $\text{lin}(L(\mathcal{C}))$, the language $L(\mathcal{A})$ is not necessarily $\sim_{\tilde{\Sigma}}$-closed. However, $\mathcal{A}$ may yet be considered as an incomplete specification so that we can still ask for an NTA $\mathcal{C}$ such that $L(\mathcal{C}) = \text{trace}(L(\mathcal{A}))$.

Note that it is easy to come up with an NWA recognizing the linearizations of the nested traces for a given NTA. Here, the state space of the NWA is the cartesian product of the local state spaces.

**Lemma 2.30.** Let $\mathcal{C}$ be an NTA over $\tilde{\Sigma}$. There is an NWA $\mathcal{A}$ over $\tilde{\Sigma}$ such that $L(\mathcal{A}) = \text{lin}(L(\mathcal{C}))$.

We are, however, interested in the other direction of transforming a given NWA into an NTA. As a preparation, we now recall two well-known theorems from Mazurkiewicz trace theory. The first one, Zielonka’s celebrated theorem, applies to distributed alphabets such that $\Sigma = \Sigma_{\text{int}}$. It will later be lifted to general distributed alphabets.
Theorem 2.31 ([Zie87]). Suppose $\Sigma = \Sigma_{\text{int}}$. Let $\mathcal{A}$ be an NWA over $\hat{\Sigma}$ such that $L(\mathcal{A})$ is $\sim_{\hat{\Sigma}}$-closed. Then, there is an NTA $\mathcal{C}$ over $\hat{\Sigma}$ such that $L(\mathcal{C}) = \text{trace}(L(\mathcal{A}))$.

Note that the theorem actually yields a deterministic trace automaton (we omit the definition). The doubly exponential complexity (in the number of processes) of Zielonka’s construction has been reduced to singly exponential [GM06, GGMW10].

Moreover, under the assumption $\Sigma = \Sigma_{\text{int}}$, closure under $\sim_{\hat{\Sigma}}$ is a decidable criterion:

Theorem 2.32 ([Mus94, PWW98]). The following is Pspace-complete:

Instance: $\hat{\Sigma}$ such that $\Sigma = \Sigma_{\text{int}}$; NWA $\mathcal{A}$ over $\hat{\Sigma}$

Question: Is $L(\mathcal{A})$ $\sim_{\hat{\Sigma}}$-closed?

Note that, in these theorems, the NWA and the NTA do not employ any stack so that we actually deal with a finite and a trace automaton, respectively. However, one can lift Zielonka’s theorem to arbitrary distributed alphabets:

Theorem 2.33 ([BGH09]). Let $\mathcal{A}$ be an NWA over $\hat{\Sigma}$ such that the language $L(\mathcal{A})$ is $\sim_{\hat{\Sigma}}$-closed. There is an NTA $\mathcal{C}$ over $\hat{\Sigma}$ such that $L(\mathcal{C}) = \text{trace}(L(\mathcal{A}))$.

Proof (sketch). We interpret the given NWA $\mathcal{A} = (S, \Delta, \iota, F)$ as an NWA $\mathcal{B}$ over a new distributed alphabet $\hat{\Omega} = (\Omega, P, \delta', \text{type}')$. Here, $\Omega = \Sigma \times \Gamma$ and, for all $(a, A) \in \Sigma \times \Gamma$, $\text{type}'((a, A)) = \text{int}$ and $\delta'((a, A)) = \delta(a)$. A transition $(s, a, A, s') \in \Delta_{\text{call}} \cup \Delta_{\text{ret}}$ in $\mathcal{A}$ is considered as a transition $(s, (a, A), s')$ in the new NWA $\mathcal{B}$, and a transition $(s, a, s') \in \Delta_{\text{int}}$ is translated to $(s, (a, A), s')$ with $A$ arbitrary. One can now apply Theorem 2.31 to obtain, from $\mathcal{B}$, an NTA $\mathcal{C}$ over $\hat{\Omega}$. Finally, we reinterpret $\mathcal{C}$ as an NTA $\mathcal{C}'$ over the original alphabet $\hat{\Sigma}$. In particular, a call or return transition $(s, (a, A), s')$ becomes $(s, a, A, s')$. \hfill \blacksquare

Remark 2.34. The size of $\mathcal{C}$ is at most doubly exponential in $|\mathcal{A}|$ and triply exponential in $|\Sigma|$, when we use the construction from [GM06].

Actually, Theorem 2.33 is a corollary of a more general statement, which uses the notion of a lexicographic normal form and is also used to prove a different result on NWA realizability (Theorem 2.38 below).

Theorem 2.33 demonstrates that NWAs, though they have a global view of the system in terms of one single state space, are suitable specifications for NTAs provided they recognize a $\sim_{\hat{\Sigma}}$-closed language. Unfortunately, it is in general undecidable if the language of a given NWA is $\sim_{\hat{\Sigma}}$-closed. This can be easily shown by a reduction from the undecidable emptiness problem (cf. Theorem 2.5 on page 20).

Theorem 2.35. The following problem is undecidable:

Instance: $\hat{\Sigma}$; NWA $\mathcal{A}$ over $\hat{\Sigma}$

Question: Is $L(\mathcal{A})$ $\sim_{\hat{\Sigma}}$-closed?
Therefore, we will consider restrictions to \( \theta \)-words, for suitable \( \theta \in \mathfrak{R} \). This will allow us to define decidable sufficient criteria for the transformation of an NWA into an NTA. We will state a Zielonka-like theorem that is tailored to this restriction. In the theorem, we require that an NWA \( \text{represents} \) the \( \theta \)-words of a system, while the final implementation can produce executions that do not fit into the \( \theta \)-restriction.

A \( \theta \)-representation is a set of nested words that does not distinguish between locally equivalent \( \theta \)-words. Here, two nested words \( W, W' \in \text{NW}(\Sigma) \) are said to be locally equivalent, written \( W \sim_{\Sigma}^\text{loc} W' \), if there are \( u, v \in \Sigma^* \) and \( (a, b) \in I_{\Sigma} \) such that \( \text{string}(W) = uabv \) and \( \text{string}(W') = ubav \).

**Definition 2.36.** Let \( \theta \in \mathfrak{R} \) and \( L \subseteq \text{NW}(\Sigma) \). We call \( L \) a \( \theta \)-representation if \( L \subseteq \text{NW}_\theta(\Sigma) \) and, for all \( W, W' \in \text{NW}_\theta(\Sigma) \) such that \( W \sim_{\Sigma}^\text{loc} W' \), we have \( W \in L \) iff \( W' \in L \).

**Example 2.37.** For the NWA \( A_{2.2} \) from page 20, we have that \( L(A_{2.2}) \) is both a 2-phase representation and an ordered representation. It is, however, neither a \( k \)-context nor a \( k \)-scope representation, for any \( k \geq 1 \).

Next, we present our Zielonka theorem suited to \( \theta \)-representations. Hereby, we require that \( \theta \) be a member of the set \( \mathfrak{R}^- \overset{\text{def}}{=} \{ k\text{-cnt}, k\text{-ph} \mid k \geq 1 \} \). We do not know if the following result holds for bounded-scope or ordered representations.

**Theorem 2.38 ([BGH09]).** Let \( \theta \in \mathfrak{R}^- \) and let \( A \) be an NWA over \( \tilde{\Sigma} \) such that \( L(A) \) is a \( \theta \)-representation. Then, there is an NTA \( C \) over \( \tilde{\Sigma} \) such that \( L(C) = \text{trace}(L(A)) \).

The proof of Theorem 2.38 relies on the definition of a lexicographic normal form that allows one to apply Ochmański’s Theorem [Och95]. The crux is that reordering independent events in order to obtain the lexicographic normal form shall not affect membership in the given \( \theta \)-representation. This, however, requires an extension of the underlying distributed alphabet. Note that the theorem was stated in [BGH09] only for the phase-restriction, but proving it for contexts is along the same lines.

Indeed, being a \( \theta \)-representation is a decidable criterion (for all \( \theta \in \mathfrak{R} \), including “bounded scope” and “ordered”). It is also decidable whether the \( \theta \)-restriction of a given NWA is \( \sim_{\tilde{\Sigma}} \)-closed (which implies that it is a \( \theta \)-representation).

**Theorem 2.39.** The following problems are decidable (for every \( \theta \in \mathfrak{R} \)) in elementary time:

**Instance:** \( \tilde{\Sigma} \); NWA \( A \) over \( \tilde{\Sigma} \)

**Question 1:** Is \( L_\theta(A) \sim_{\tilde{\Sigma}} \)-closed?

**Question 2:** Is \( L_\theta(A) \) a \( \theta \)-representation?
Proof (sketch). We will give the rough idea of the proof, which is inspired by [Mus94, PWW98] where similar problems are addressed in a stack-free setting. Consider first Question 1. Using Theorem 2.22, we transform the given NWA $A$ into an NWA $A'$ for the complement, i.e., such that $L(A') = \text{NW}(\bar{\Sigma}) \setminus L_\theta(A)$. Afterwards, we build an NWA $B$ recognizing all nested words $W \in \text{NW}(\bar{\Sigma})$ where string($W$) is of the form $uabv$ such that $(a, b) \in I_\bar{\Sigma}$ and $uabv \in \text{string}(W')$ for some $W' \in L(A')$. Then, $L_\theta(A)$ is not $\sim_{\bar{\Sigma}}$-closed iff $L_\theta(B) \neq \emptyset$. The latter is decidable due to Theorems 2.15–2.18. To solve Question 2, we build $A'$ such that $L(A') = \text{NW}_\theta(\bar{\Sigma}) \setminus L(A)$ (again, using Theorem 2.22).

2.7 Realizability of MSO Specifications

On a higher level than NWAs, concurrent recursive programs can be specified in logic, be it MSO logic or temporal logic (for the latter, see the subsequent section). Similarly to the word case, the logic $\text{ntMSO}(\bar{\Sigma})$ is defined just as $\text{nwMSO}$ but, instead of $\ll_{+1}$, it features a collection of predicates $\ll_p$, one for every process $p \in P$. Given a sentence $\varphi \in \text{ntMSO}(\bar{\Sigma})$, we denote by $L(\varphi)$ the set of nested traces $T$ over $\bar{\Sigma}$ such that $T \models \varphi$.

Again, we will recall a well known result for concurrent programs without stacks, and then lift it to the recursive setting.

Theorem 2.40 ([Tho90]). Suppose $\Sigma = \Sigma_{\text{int}}$ and let $L \subseteq \text{NTr}(\bar{\Sigma})$. The following statements are effectively equivalent:

1. There is an NTA $C$ such that $L(C) = L$.

2. There is a sentence $\varphi \in \text{ntMSO}(\bar{\Sigma})$ such that $L(\varphi) = L$.

Note that this theorem cannot be lifted to nested structures without imposing any restriction. This is due to the fact that NWAs (and, therefore, NTAs) are not complementable (Theorem 2.21), while MSO logic is closed under negation. However, it will turn out that restricting to $k$-context/$k$-phase traces also helps in this case.

Of course, given a restriction $\theta \in \mathcal{R}$, we first have to clarify what we mean by a $\theta$-trace. There are at least two reasonable possibilities. For example, we may call a nested trace $T$ a $\theta$-trace if all linearizations of $T$ are $\theta$-words. Alternatively, we may require that some linearization of $T$ is a $\theta$-word. We will choose the latter, existential, definition, as it captures more nested traces. Note that there are similar options and definitions in the setting of MSCs (cf. Chapter 3).

Thus, we call $T \in \text{NTr}(\bar{\Sigma})$ a $\theta$-trace if $\text{lin}(T) \cap \text{NW}_\theta(\bar{\Sigma}) \neq \emptyset$. We adopt other notations for nested words and let $\text{NTr}_\theta(\bar{\Sigma})$ be the set of $\theta$-traces. Moreover, for an NTA $C$, we let $L_\theta(C) \overset{\text{def}}{=} L(C) \cap \text{NTr}_\theta(\bar{\Sigma})$. For a sentence $\varphi \in \text{ntMSO}(\bar{\Sigma})$, the set $L_\theta(\varphi)$ is defined accordingly.
Example 2.41. The nested trace $T_{2,4}$ is a 2-phase trace (it admits the 2-phase linearization $W_{2,1}$), a 4-context trace, a 3-scope trace, and an ordered trace (since its linearization $W_{2,1}$ is ordered). ⊓⊔

As a preparation of a logical characterization of NTAs, we show that they are complementable for some restrictions of nested traces. This is an analogue of Theorem 2.22 for NWAs. As we rely on Theorem 2.38, we have to restrict to the set $R^- = \{k\text{-cnt}, k\text{-ph} \mid k \geq 1\}$, i.e., to bounded contexts or bounded phases.

Lemma 2.42. Let $\theta \in R^-$ and let $C$ be an NTA over $\tilde{\Sigma}$. Then, there is an NTA $C'$ such that $L(C') = \text{NTr}_\theta(\tilde{\Sigma}) \setminus L(C)$.

Proof. From the given NTA $C$ be an NTA, using Lemma 2.30 and Theorem 2.22, we get an NWA $A$ over $\tilde{\Sigma}$ such that $L(A) = \text{NW}_\theta(\tilde{\Sigma}) \setminus \text{lin}(L(C))$. Observe that $L(A)$ is a $\theta$-representation. By Theorem 2.38, there is an NTA $C'$ over $\tilde{\Sigma}$ such that $L(C') = \text{trace}(L(A))$. One easily verifies that we actually have $L(C') = \text{NTr}_\theta(\tilde{\Sigma}) \setminus L(C)$. ■

In particular, Lemma 2.42 implies that there is an NTA recognizing the set $\text{NTr}_\theta(\tilde{\Sigma})$. The following theorem constitutes a generalization of Theorem 2.40 adapted to $\theta$-traces, where $\theta \in R^-$. Theorem 2.43 ([BGH09]). The following implications are effective:

1. For every NTA $C$ over $\tilde{\Sigma}$, there is a sentence $\varphi \in \text{ntMSO}(\tilde{\Sigma})$ such that $L(\varphi) = L(C)$.
2. Let $\theta \in R^-$. For every sentence $\varphi \in \text{ntMSO}(\tilde{\Sigma})$, there is an NTA $C$ over $\tilde{\Sigma}$ such that $L(C) = L_{\theta}(\varphi)$.

Proof (sketch). The proof of 1. follows the standard construction for the translation of automata into logic. One guesses an assignment of states and stack symbols to events in terms of existentially quantified second-order variables. Then, a first-order kernel checks if the assignments actually correspond to an accepting run.

For the proof of 2., we proceed by structural induction. Hereby, the only critical case is negation, which will be taken care of by Lemma 2.42. For simplicity, we suppose that there are no free variables. To get an automaton for $\neg \varphi$, suppose that we already have an NTA $C$ such that $L(C) = L_{\theta}(\varphi)$. By Lemma 2.42, there is an NTA $C'$ such that $L(C') = \text{NTr}_\theta(\tilde{\Sigma}) \setminus L(C)$. We have $L(C') = L_{\theta}(\neg \varphi)$ so that we are done. ■

2.8 Temporal Logic for Nested Words

In this section, we consider the model-checking problem for phase-bounded NWAs. Note that we can already infer, from the previous results, that it is decidable
whether all (phase-bounded) nested words accepted by a given NWA satisfy a given MSO formula. However, since the complexity is inherently nonelementary, it is worth looking at temporal logics for nested words.

There has been a whole bunch of papers considering the model-checking problem for temporal logics over (multiply) nested words [Ati10, BCGZ14, CGNK12, LTN12, ABKS12, BS14]. These works differ in the choice of the behavioral restriction described before (context-, phase-, scope-bounded, ordered), but also in the concrete temporal logic adopted for the model-checking task. While [Ati10, ABKS12] consider properties over strings such as classical LTL rather than nested words, [LTN12] introduces a temporal logic that allows one to identify related call and return positions of a given process and to distinguish between linear successors (referring to the word structure) and abstract successors (involving the nesting edges). However, as a matter of fact, there is so far no agreement on a canonical temporal logic for nested words, not even for those with one single nesting relation [AAB+08, AEM04].

We, therefore, consider the class of all temporal logics as defined in the book by Gabbay, Hodkinson, and Reynolds [GHR94], which subsumes virtually all existing formalisms. The unifying feature of these temporal logics is that their modalities are defined in MSO logic. In [BCGZ14], we showed that satisfiability and model checking for any MSO-definable temporal logic are decidable in EXPTime when restricting to phase-bounded executions. We improved this in [BKM13] showing that the problems are still elementary if $k$ is part of the input. This is an important issue, as an elementary procedure allows for a gradual adjustment of $k$ at the cost of only an elementary blow-up. In the following, we will review the main results of [BKM13].

We fix a distributed alphabet $\Sigma = (\Sigma, P, \text{type}, \delta)$. For a natural number $m \in \mathbb{N}$, we call $\varphi \in \text{nwMSO}(\Sigma)$ an $m$-ary modality if its free variables consist of one first-order variable $x$ and $m$ set variables $X_1, \ldots, X_m$. The idea is that variable $X_i$ is interpreted as the set of positions where the $i$-th argument of the modality holds.

**Definition 2.44.** A temporal logic (over $\Sigma$) is given by a triple $\mathcal{L} = (M, \text{arity}, [\cdot])$ including

- a finite set $M$ of modality names,
- a mapping $\text{arity} : M \to \mathbb{N}$, and
- a mapping $[\cdot] : M \to \text{nwMSO}(\Sigma)$ such that, for $M \in M$ with $\text{arity}(M) = m$, $[M]$ is an $m$-ary modality.

The syntax of $\mathcal{L}$, i.e., the set of formulas $\varphi \in \text{Form}(\mathcal{L})$, is given by

$$\varphi ::= a \mid \neg \varphi \mid \varphi \lor \varphi \mid M(b, \ldots, \varphi)_{\text{arity}(M)}$$

where $a$ ranges over $\Sigma$ and $M$ ranges over $M$. 35
The semantics of a formula $\varphi \in \text{Form}(\Sigma)$ wrt. a nested word $W = (E, <, \lambda)$ over $\Sigma$ is defined inductively as a set $[\varphi]_W \subseteq E$, containing the events of $W$ that satisfy $\varphi$. Formally, $[-]_W$ is given as follows:

- $[a]_W \overset{\text{def}}{=} \{ e \in E \mid \lambda(e) = a \}$
- $[\neg \varphi]_W \overset{\text{def}}{=} E \setminus [\varphi]_W$
- $[\varphi_1 \lor \varphi_2]_W \overset{\text{def}}{=} [\varphi_1]_W \cup [\varphi_2]_W$
- $[M(\varphi_1, \ldots, \varphi_m)]_W \overset{\text{def}}{=} \{ e \in E \mid W \models_{I_\varphi} [M] \}$
  where $I_\varphi$ is such that $x$ is mapped to $e$, and $X_I$ to $[\varphi_i]_W$ for all $i \in \{1, \ldots, m\}$

**Example 2.45.** As mentioned above, a wide range of temporal logics have been defined for nested words and concurrent systems. Their until operator usually depends on what is considered a path between two word positions and, more specifically, on a notion of successor. In the classical setting of words without nesting relations, one naturally considers the direct successor following the linear order. The situation is less clear in the presence of one or more nesting relations. In [AAB+08], Alur et al. identify three different kinds of successors in singly nested words, namely the linear, call, and abstract successor. Each of them comes with a separate until operator.

Towards nested words with multiple nesting relations, Atig et al. consider only the linear successor [Ati10, ABKS12], while La Torre and Napoli also define modalities that correspond to linear, call, and abstract successors [LTN12]. As an example, we demonstrate how to deal with the abstract until in our framework. An abstract $p$-path in a nested word is a path that does not choose a linear direct successor from a call position or to a return position (wrt. $<_\text{cr} \cap (E_p \times E_p)$). The temporal-logic formula $\varphi U_p^a \psi$ for the abstract until says that, starting from $x$, there is an abstract $p$-path (represented by a second-order variable $Y$) to some event $x'$. Hereby, we require that $\varphi$ is satisfied along the path ($Y \subseteq X_1$) until $x'$ satisfies $\psi$ ($x' \in X_2$).

Formally, the binary modality $U_p^a$ is defined by

$$[U_p^a](x, X_1, X_2) =$$

$$\exists Y. \exists x'. (Y \subseteq X_1 \land x' \in X_2 \land$$

$$\forall z. (z \in Y \lor z = x') \rightarrow (z = x \lor \exists y. (y \in Y \land \varphi_p(y, z))) .$$

Hereby,

$$\varphi_p(y, z) \overset{\text{def}}{=} (p(y) \land y <_\text{cr} z) \lor (y <_{+1} z \land \neg \text{call}_p(y) \land \neg \text{ret}_p(z))$$

where $p(x) \overset{\text{def}}{=} \bigvee_{a \in \Sigma} a(x)$ and, for $\tau \in \{\text{call}, \text{ret}\}$, $\tau_p(x) \overset{\text{def}}{=} p(x) \land \bigvee_{a \in \Sigma} a(x)$. $\text{type}(a)=t$.

Indeed, all the modalities considered in [AAB+08, Ati10, ABKS12, LTN12] are MSO-definable. However, they appear to be just a few of many other possibilities. For example, one may define an abstract path including two or more nesting relations, or include past-time counterparts of until modalities, which are not present.
in [LTN12]. Such extensions can be realized in our framework by giving their definition in MSO. An elementary upper bound of the satisfiability and model-checking problem follows immediately from the result stated below, without changing anything in the decidability proof.

\[
\text{Let } L \text{ be a temporal logic over } \tilde{\Sigma}. \text{ For a temporal-logic formula } \varphi \in \text{Form}(\mathcal{L}), \text{ we let } L(\varphi) \text{ be the set of nested words } W \text{ such that } e \in [\varphi]_W \text{ where } e \text{ is the (unique) minimal event of } W. \text{ Now, the corresponding model-checking problem is defined as follows:}
\]

**Problem 2.46.** NWA-Model-Checking($\tilde{\Sigma}, \mathcal{L}$):

**Instance:** NWA $\mathcal{A}$ over $\tilde{\Sigma}; \varphi \in \text{Form}(\mathcal{L}); k \geq 1$ (encoded in unary)

**Question:** Do we have $L_{k\text{-ph}}(A) \subseteq L(\varphi)$?

For $n \geq 0$, let $M\Sigma_n(\tilde{\Sigma})$ be the set of nwMSO($\tilde{\Sigma}$)-formulas of the form

\[
\exists X_1 \forall X_2 \ldots \exists / \forall X_n. \varphi
\]

where $\varphi$ is a first-order formula, i.e., it does not contain any second-order quantifier, and the $X_i$ are blocks of second-order variables. A temporal logic $\mathcal{L} = (M, \text{arity}, [-])$ is called $M\Sigma_n(\tilde{\Sigma})$-definable if, for all $M \in M$, we have $[M] \in M\Sigma_n(\tilde{\Sigma})$.

**Theorem 2.47 ([BKM13]).** Let $n \geq 0$ and let $\mathcal{L}$ be some $M\Sigma_n(\tilde{\Sigma})$-definable temporal logic. Then, NWA-Model-Checking($\tilde{\Sigma}, \mathcal{L}$) is in $(n + 2)$-EXPTIME.

**Theorem 2.48 ([BKM13]).** There is a distributed alphabet $\tilde{\Sigma}$ such that, for all $n \geq 1$, there is an $M\Sigma_n(\tilde{\Sigma})$-definable temporal logic $\mathcal{L}$ for which the problem NWA-Model-Checking($\tilde{\Sigma}, \mathcal{L}$) is $n$-EXPSPACE-hard.

Two key ideas are pursued in the proof of the upper bound. First, we translate, in polynomial time, a temporal logic formula into an MSO formula in a certain normal form. The construction is based on Hanf’s locality theorem and independent of the number of phases. Second, we show that an MSO formula in normal form can be transformed into a tree automaton in $(n + 1)$-fold exponential space. The tree automaton works on tree encodings of multiply nested words and can then be checked for emptiness. One of its key ingredients is a tree automaton recovering the direct successor relation of the encoded multiply nested words. We show that such an automaton can be computed in polynomial space, avoiding the generic doubly exponential construction given in [LMP07].
2.9 Perspectives

Most results presented in this chapter rely on the specific restriction of the domain of nested words to bounded contexts/ phases, scopes, or to ordered words. It will be worthwhile to study a more generic setting by bounding the split-width. The model-checking question has been well-studied here as well [CGNK12, AGNK14b, Men14], but not much is known about realizability.

Note that some realizable specifications will inevitably yield implementations in terms of NWAs that are non-deterministic and suffer from deadlocks. One should, therefore, study classes of NWAs that are arguably more “realistic” meaning, in particular, that they are deterministic and deadlock-free [SEM03, ADGS13]. A natural question is then to ask for a specification formalism that guarantees such realistic implementations.

It also remains to study realizability in the realm of recursive processes communicating through FIFO channels [LMP08a, HLMS12]. The model-checking question is by now well understood, in particular thanks to the split-width technique [AGNK14b, Cyr14]. To the best of our knowledge, realizability questions for such communicating recursive processes have not been considered. The quest for controllers, whose study has been initiated in [AGNK14a], seems to be closely related.
While, in Chapter 2, the communication topology was static and fixed once for all in terms of a distributed alphabet $\tilde{\Sigma}$, we now consider systems that can be run on any topology from a (possibly infinite) class of topologies. Thus, the topology becomes a parameter of the system. We will study parameterized systems in the realm of message passing.

There has been a large body of literature on parameterized concurrent systems (e.g., [EN03, EK04, LMP10b, ABQ11, DSZ11, AHH13, EGM13, Esp14, AJKR14]), with a focus on verification: Does the given system satisfy a specification independently of the topology or the number of processes? A variety of different models have been introduced, covering a wide range of communication paradigms such as broadcasting, rendez-vous, token-passing, etc. However, it is justified to say that, so far, there is no such thing as a canonical or “robust” model of parameterized concurrent systems. In particular, expressiveness issues have not been considered. One reason for that may be that the semantics of parameterized systems is often described in terms of strings that represent interleavings of independent events. Since such interleavings blur the natural direct-successor relation of a single process, it is hard to find logical or algebraic formalisms over strings that match the expressive power of automata.

Similarly to nested traces from the previous chapter, the semantics that we present here does not rely on interleavings but keeps a maximum of information on an execution, such as a direct-successor relation for each process. This will finally allow us to come up with classes of parameterized message-passing systems that can be considered robust, since they satisfy at least one of the following properties:

(a) They have a decidable emptiness problem.
They are closed under boolean operations.

They enjoy a natural logical characterization.

In particular, we provide several Büchi-Elgot-Trakhtenbrot Theorems, which roughly read as follows:

Given a formula \( \varphi \) and a class \( \mathcal{T} \) of topologies, there is a system \( A \) such that, whenever \( A \) is run on a topology from \( \mathcal{T} \), it accepts exactly the behaviors that are a model of \( \varphi \).

Depending on the logic and the class of topologies, such a result will require further restrictions on the behavior of a system, similar in spirit to the restrictions that we considered in Chapter 2 in the realm of multiply nested words.

### 3.1 Topologies

As we are in a parameterized setting, we assume that there are a finite but unbounded number of processes. We consider that processes are equipped with interfaces, through which they can talk to other processes. When interfaces of different processes are plugged together, we obtain a topology.

Though the number of processes may be unbounded, we will assume that there is a fixed nonempty finite set \( \mathcal{N} = \{a, b, c, \ldots\} \) of interface names (or, simply, interfaces). When we consider pipelines or rings, then we may set \( \mathcal{N} = \{a, b\} \) where \( a \) refers to the right neighbor and \( b \) to the left neighbor of a process. A pipeline with four processes is depicted in Figure 3.1 and a ring with five processes in Figure 3.2. They allow a process to execute an action \( a! \), which sends a message to its right neighbor, or an action \( b! \), which sends a message to the left neighbor. The complementary receive actions are \( a? \) and \( b? \), which receive from the right and left neighbor, respectively. For grids, we will need two more names, which refer to adjacent processes above and below (Figure 3.4). Ranked trees require an interface for each of the (boundedly many) children of a process, as well as pointers to the father process (see Figure 3.3 for a binary tree).

Actually, we assume that two processes \( p \) and \( q \) that are adjacent in a topology communicate through channels. More precisely, there are two FIFO channels between \( p \) and \( q \), one for messages sent from \( p \) to \( q \), and one for messages from \( q \) to \( p \). Thus, an edge \( \circlearrowleft a \circlearrowright b \) in the topology shall be understood as

![Diagram](attachment:image.png)

As every process can only talk directly to \( |\mathcal{N}| \) neighbors, topologies are structures of bounded degree. This excludes star topologies or unranked trees, but takes into account such natural cases as pipelines, ranked trees, rings, and grids.
Definition 3.1 (topology). A topology over $\mathcal{N}$ is a pair $\mathcal{T} = (P, \rightarrow)$ where $P$ is the nonempty finite set of processes and $\rightarrow \subseteq P \times \mathcal{N} \times \mathcal{N} \times P$ is the edge relation. We write $p \overset{a,b}{\rightarrow} q$ for $(p, a, b, q) \in \rightarrow$, which signifies that the $a$-interface of $p$ points to $q$, and the $b$-interface of $q$ points to $p$. We require that, whenever $p \overset{a,b}{\rightarrow} q$, the following hold:

(a) $p \neq q$ (there are no self loops),
(b) $q \overset{b,a}{\rightarrow} p$ (adjacent processes are mutually connected), and
(c) for all $a', b' \in \mathcal{N}$ and $q' \in P$ such that $p \overset{a',b'}{\rightarrow} q'$, we have $a = a'$ iff $q = q'$ (an interface points to at most one process, and two distinct interfaces point to distinct processes).

The automata and logics that we are going to define will not be able to distinguish isomorphic topologies. An exception is Section 3.5, where topologies are fixed.

Example 3.2. In this chapter, the focus will be on the specific topology classes that are illustrated in Figures 3.1–3.4. We define pipelines and rings as topologies over $\{a, b\}$, and binary trees and grids as topologies over $\{a, b, c, d\}$. However, whenever convenient, we assume that pipelines and rings are topologies over $\{a, b, c, d\}$ as well, though they do not make use of $c$ and $d$. Note that several of the results in this chapter actually extend to more general classes.

- A pipeline is a topology over $\{a, b\}$ of the form $\left(\{1, \ldots, n\}, \rightarrow\right)$ where $n \geq 2$ and $\rightarrow = \{(i, a, b, i+1) \mid i \in [n-1]\} \cup \{(i+1, b, a, i) \mid i \in [n-1]\}$. Note that a pipeline is uniquely determined by its number of processes.
• A *ring* is like a pipeline where the endpoints are glued together. I.e., it is a topology of the form \( \{1, \ldots, n\}, \cdots \) where \( n \geq 3 \) and the edge relation is given by \( \cdots = \{(i, a, b, (i \bmod n) + 1) \mid i \in [n]\} \cup \{(i \bmod n) + 1, b, a, i) \mid i \in [n]\} \). Like a pipeline, since we do not distinguish isomorphic topologies, a ring is uniquely determined by its number of processes. A *ring forest* is simply a disjoint union of rings.

• A *(binary-)tree topology* is a topology \((P, \cdots)\) over \(\{a, b, c, d\}\) where \(P\) is a prefix-closed subset of \(\{0, 1\}^*\) such that \(|P| \geq 2\), and \(\cdots = \{(u, a, b, u0) \mid u, u0 \in P\} \cup \{(u0, b, a, u) \mid u, u0 \in P\} \cup \{(u, c, d, u1) \mid u, u1 \in P\} \cup \{(u1, d, c, u) \mid u, u1 \in P\} \).

• A *grid* is a topology of the form \([m] \times [n], \cdots\) where \(\max\{m, n\} \geq 2\) and
\[
\cdots = \\{(i, j), a, b, (i, j + 1) \mid i \in [m] \text{ and } j \in [n - 1]\}
\cup \{(i, j + 1), b, a, (i, j) \mid i \in [m] \text{ and } j \in [n - 1]\}
\cup \{(i, j), c, d, (i + 1, j) \mid i \in [m - 1] \text{ and } j \in [n]\}
\cup \{(i + 1, j), d, c, (i, j) \mid i \in [m - 1] \text{ and } j \in [n]\}.
\]

3.2 Message Sequence Charts

As already mentioned, we do not consider an interleaving semantics but rather keep a maximum of information about the type of an event and the causal dependencies between events. One possible behavior of a system is depicted as a *message sequence chart (MSC)*. An MSC consists of a topology (over the given set of interfaces) and a set of events, which represent the communication actions executed by a system. Events are located on the processes and connected by process and message edges, which reflect causal dependencies.

**Definition 3.3 (MSC).** A message sequence chart (MSC) over \(\mathcal{N}\) is a tuple \(M = (P, \cdots, E, \prec, \pi)\) where

• \((P, \cdots)\) is a topology over \(\mathcal{N}\),

• \(E\) is the nonempty finite set of events,

• \(\prec \subseteq E \times E\) is the acyclic edge relation, which is partitioned into \(\prec_{\text{proc}}\) and \(\prec_{\text{msg}}\), and

• \(\pi : E \to P\) determines the location of an event in the topology; for \(p \in P\), we let \(E_p = \{e \in E \mid \pi(e) = p\}\).

We require that the following hold:

• \(\prec_{\text{proc}}\) is a union \(\bigcup_{p \in P} \prec_p\) where each \(\prec_p \subseteq E_p \times E_p\) is the direct-successor relation of some total order on \(E_p\).
• there is a partition $E = E! \uplus E?$ such that $\prec_{\text{msg}}$ induces a bijection between $E!$ and $E?$, 

• for all $(e, f) \in \prec_{\text{msg}}$, there are $a, b \in \mathcal{N}$ such that $\pi(e) \xrightarrow{a \ b} \pi(f)$ (communication is restricted to adjacent processes), and 

• for all $(e, f), (e', f') \in \prec_{\text{msg}}$ such that $\pi(e) = \pi(e')$ and $\pi(f) = \pi(f')$, we have $e \prec_{\text{proc}} e'$ iff $f \prec_{\text{proc}} f'$ (FIFO).

Let $\Sigma \overset{\text{def}}{=} \{a! \mid a \in \mathcal{N}\} \cup \{a? \mid a \in \mathcal{N}\}$. We define a mapping $\ell_M : E \to \Sigma$ that associates with each event of the MSC $M$ the type of action that it executes: for $(e, f) \in \prec_{\text{msg}}$ and $a, b \in \mathcal{N}$ such that $\pi(e) \xrightarrow{a \ b} \pi(f)$, we set $\ell_M(e) = a!$ and $\ell_M(f) = b?$. The set of MSCs (over the fixed set $\mathcal{N}$) is denoted by $\text{MSC}$. Like for topologies, we usually do not distinguish isomorphic MSCs.

Example 3.4. Two example MSCs are depicted in Figure 3.5. As we will see, they represent executions of a token-ring protocol. In particular, their underlying topologies are rings (of size 5). The process labelings $s_1, s_2, s'_2, s_3$ can be ignored for the moment. The events are the endpoints of message arrows, which represent $\prec_{\text{msg}}$. Process edges are implicitly given; they connect successive events located on the same (top-down) process line. Finally, the mapping $\ell_M$ is illustrated on a few events.

3.3 Underapproximation Classes

Similarly to multiply nested words (Chapter 2), we will run quickly into undecidability and non-complementability without imposing any restriction on the behavior of a parameterized system. There are several natural restrictions for MSCs. The simplest one is to assume a class of topologies $\mathcal{F}$ among the topologies over $\mathcal{N}$ (e.g., the class of pipelines, trees, rings, grids, or rings) and to restrict ourselves to the set $\text{MSC}_{\mathcal{F}}$, defined as the set of MSCs $(P, \longrightarrow, E, \prec, \pi) \in \text{MSC}$ such that $(P, \longrightarrow) \in \mathcal{F}$. 

Figure 3.5: Two MSCs with a ring topology
### 3.3.1 Channel-Bounded MSCs

Some of our (positive) results will deal with systems that have (existentially) $k$-bounded channels, for some $k \geq 1$ (see, e.g., [LM04, GKM06, GKM07]). Intuitively, an MSC is $k$-bounded if it can be scheduled in such a way that, along the execution, there are never more than $k$ messages in each channel. Formally, we define boundedness via linearizations. A linearization of an MSC $M = (P, \rightarrow, E, \preceq, \pi)$ is any total order $\preceq \subseteq E \times E$ satisfying $\preceq^* \subseteq \preceq$. Then, $\preceq$ is called $k$-bounded if, for all $f \in E$ and all $p, q \in P$ and $a, b \in N$ such that $p \xrightarrow{b} q$, we have $|\{e \in E \mid e \preceq f, \pi(e) = p, \text{ and } \ell_M(e) = a!\}| - |\{e \in E \mid e \preceq f, \pi(e) = q, \text{ and } \ell_M(e) = b?\}| \leq k$.

In other words, in any prefix of $\preceq$, there are no more than $k$ pending messages, in every “channel” $(p, q)$.

**Definition 3.5 ($k$-bounded MSC).** For a natural number $k \geq 1$, an MSC is said to be $k$-bounded if it has some $k$-bounded linearization.

A stronger restriction is that of rendez-vous communication, where a send event and its receive event are executed simultaneously. We then say that an MSC is 0-bounded.

**Definition 3.6 (0-bounded MSC).** An MSC $(P, \rightarrow, E, \preceq, \pi) \in MSC$ is called 0-bounded if the graph $(E, \preceq \cup \preceq^{-1}_{\text{msg}})$, does not admit a cycle that uses at least one $\preceq_{\text{proc}}$-edge.

Intuitively, message edges in a 0-bounded MSC can always be drawn horizontally. For $k \in \mathbb{N}$, the set of $k$-bounded MSCs is denoted $MSC_{\leq k}$. Note that, for all $k \in \mathbb{N}$, we have $MSC_{\leq k} \subseteq MSC_{\leq k+1}$. The MSCs from Figure 3.5 are both 0-bounded.

### 3.3.2 Context-Bounded MSCs

We will see that even restricting to pipelines and 0-bounded MSCs will not be enough to obtain a “robust” class of parameterized automata that is closed under complementation. Inspired by the notion of a context bound in the realm of multiply nested words (cf. Chapter 2), we now introduce context-bounded MSCs.

Actually, the efficiency of distributed algorithms and protocols is usually measured in terms of two parameters: the number $n$ of processes, and the number $k$ of contexts. It is, therefore, natural to bound any of these parameters in order to turn parameterized communicating automata (as defined below) into a robust model. Note that bounding the number of processes is known as “cut-off” (see, e.g., [EN03, AKR+14]). A context, on the other hand, restricts communication of a process to patterns such as “send a message to each neighbor and receive a message from each neighbor”. The trade-off between $n$ and $k$ is often in favor of a smaller $k$, so that it is all the more justified to impose a bound on $k$.

Here, we actually consider more relaxed definitions where, in every context, a process may perform an unbounded number of actions. In an interface-context, for example, a process can send and receive an arbitrary number of messages to/from a
fixed neighbor. A second context-type definition allows for arbitrarily many sends to all neighbors, or receptions from a fixed neighbor.

Recall that $\Sigma = \{a! \mid a \in \mathcal{N}\} \cup \{a? \mid a \in \mathcal{N}\}$. A word $w \in \Sigma^*$ is called an

- $(s\oplus r)$-context if $w \in \{a! \mid a \in \mathcal{N}\}^*$ or $w \in \{a? \mid a \in \mathcal{N}\}^*$,
- $(s1+r1)$-context if $w \in \{a!,b?\}^*$ for some $a,b \in \mathcal{N}$,
- $(s\oplus r1)$-context if $w \in \{a! \mid a \in \mathcal{N}\}^*$ or $w \in \{b?\}^*$ for some $b \in \mathcal{N}$,
- $\text{intf}$-context if $w \in \{a!,a?\}^*$ for some $a \in \mathcal{N}$.

The context type $s1\oplus r$ ($w \in \{a!\}^*$ for some $a \in \mathcal{N}$ or $w \in \{b? \mid b \in \mathcal{N}\}^*$) is dual to $s\oplus r1$, and we only consider the latter case. All results for $s\oplus r1$ presented below are valid for $s1\oplus r$.

Let $k \geq 1$ be a natural number and $ct \in \{s\oplus r,s1+r1,s\oplus r1,\text{intf}\}$ be a context type. We say that $w \in \Sigma^*$ is $(k,ct)$-bounded if there are $w_1,\ldots,w_k \in \Sigma^*$ such that $w = w_1 \ldots w_k$ and $w_i$ is a ct-context, for all $i \in [k]$. To lift this definition to MSCs $M = (P,\rightarrow, E, \triangleleft, \pi)$, we define the projection $M|_p \in \Sigma^*$ of $M$ to a process $p \in P$. Let $e_1 \triangleleft_{\text{proc}} e_2 \triangleleft_{\text{proc}} \ldots \triangleleft_{\text{proc}} e_n$ be the unique process-order preserving enumeration of all events of $E_p$. We let $M|_p = \ell_M(e_1)\ell_M(e_2)\ldots\ell_M(e_n)$. In particular, $E_p = \emptyset$ implies $M|_p = \varepsilon$.

**Definition 3.7 ((k, ct)-bounded MSC [BGS14]).** Let $k \geq 1$ be a natural number and $ct \in \{s\oplus r,s1+r1,s\oplus r1,\text{intf}\}$. We say that an MSC $M = (P,\rightarrow, E, \triangleleft, \pi) \in \text{MSC}$ is $(k,ct)$-bounded if $M|_p$ is $(k,ct)$-bounded, for all $p \in P$. \hfill \Box

Let $\text{MSC}_{(k,ct)}$ denote the set of all $(k,ct)$-bounded MSCs.

**Example 3.8.** As every process in the MSCs from Figure 3.5 executes only two events, both MSCs are $(2,ct)$ bounded, for all context types $ct$. Now, consider the set $L$ of MSCs depicted in Figure 3.10 on page 55, where we assume that the dotted areas can be arbitrarily large. Contexts are depicted as rectangles. Every MSC in $L$ is $(3,s\oplus r1)$-bounded: processes 1, 2, and 3 use three contexts, while 4 and 5 use only one context. However, for all $k \geq 1$, there is an MSC in $L$ that is not $(k,\text{intf})$-bounded: process 4 switches unboundedly often between sending through $a$ and sending through $b$. \hfill \Box

### 3.4 Communicating Automata

Next, we introduce our automata model of a parameterized message-passing system, which can be run on any topology over the fixed set $\mathcal{N}$. The idea is that each process can execute actions of the form $(a!,m)$, which emits a message $m$ through interface $a$, or $(a?,m)$, which receives $m$ from interface $a$.

Given a topology $T = (P,\rightarrow)$ over $\mathcal{N}$, an automaton will actually run identical subautomata on processes of the same type. We define $\text{type} : P \rightarrow 2^\mathcal{N}$ by $\text{type}(p) \overset{\text{def}}{=} \{a \in \mathcal{N} \mid \text{there are } b \in \mathcal{N} \text{ and } q \in P \text{ such that } p \xrightarrow{a} b \xrightarrow{} q\}$. Thus, $\text{type}(p)$ contains those interfaces of $p$ that are connected to some other process in $T$ (we assume that $T$ is clear from the context).
Definition 3.9 (CA). A communicating automaton (CA) over the set of interface names $\mathcal{N}$ is a tuple $A = (S, \iota, \text{Msg}, \Delta, \mathcal{F})$ where

- $S$ is the finite set of states,
- $\iota : 2^\mathcal{N} \to S$ assigns to every possible type in a topology an initial state,
- $\text{Msg}$ is a nonempty finite set of messages,
- $\Delta \subseteq S \times (\Sigma \times \text{Msg}) \times (S \setminus \iota(2^\mathcal{N}))$ is the transition relation, and
- $\mathcal{F}$ is the acceptance condition (which will be specified below). \qed

Here, $\iota(2^\mathcal{N}) \overset{\text{def}}{=} \{ \iota(A) \mid A \subseteq \mathcal{N} \}$. Thus, processes do not return to initial states. This is a purely technical issue which allows us to consider several automata models in a unifying framework.

Let $M = (P, \leftarrow, E, \lessdot, \pi)$ be an MSC. A run of $A$ on $M$ will be a mapping $\rho : E \to S$ satisfying some requirements. As in the previous chapter, $\rho(e)$ can be seen as the local state of $\pi(e)$ after executing $e$. To determine when $\rho$ is a run, we define another mapping, $\rho^- : E \to S$, denoting the source states of a transition: whenever $f \lessdot_{\text{proc}} e$, we let $\rho^-(e) = \rho(f)$; moreover, if $e$ is $\lessdot_{\text{proc}}$-minimal, we let $\rho^-(e) = \iota(\text{type}(\pi(e)))$. With this, we say that $\rho$ is a run of $A$ on $M$ if, for all $(e, f) \in \lessdot_{\text{msg}}$, there are $a, b \in \mathcal{N}$ and a message $m \in \text{Msg}$ such that $\pi(e) \overset{a,b}{\rightarrow} \pi(f)$, $(\rho^-(e), (a^! m), \rho(e)) \in \Delta$, and $(\rho^-(f), (b^? m), \rho(f)) \in \Delta$.

Whether a run is accepting or not depends on the variant of CA and, in particular, its acceptance condition. Those will be defined below. In any case, acceptance will depend on the mapping $\lambda_{M, \rho} : P \to S$ that collects, for every process $p$ the final state $\lambda_{M, \rho}(p)$ in which $p$ terminates. Formally, $\lambda_{M, \rho}(p) = \iota(\text{type}(p))$ if $E_p = \emptyset$. Otherwise, $\lambda_{M, \rho}(p) = \rho(e)$ where $e$ is the $\lessdot_{\text{proc}}$-maximal event from $E_p$.

Example 3.10. The CA $A_{\text{token}}$ over $\{a, b\}$ from Figure 3.6 describes a simplified version of the IEEE 802.5 token-ring protocol. In the protocol, a single binary token, which can carry a value from $m \in \{0,1\}$, circulates in a ring. Initially, the token has value 1. A process that has the token may emit a message and pass it along with the token to its a-neighbor. We will abstract the concrete message away and only consider the token value. Since all processes in a ring have type $\{a, b\}$, we have a single initial state $s_0$ for all of them, i.e., $\iota(\{a, b\}) = s_0$. Whenever a process receives the token from its b-neighbor, it will forward it to its a-neighbor, while

- leaving the token value unchanged (the process then ends in $s_2$ or $s_3$), or
- changing its value from 1 to 0, to signal that the message has been received (the process then ends in $s'_2$).

We did not define the acceptance condition formally yet. However, note that it will allow us to require that (i) there is exactly one process that terminates in
state $s_1$ and (ii) no process terminates in a state from $\{t_1, t_2, t_3\}$ (i.e., no process stops halfways). With this condition, MSC $M_1$ from Figure 3.5 will be accepted by $A_{\text{token}}$, while $M_2$ is not accepted, since it requires two processes to end in $s_1$, which violates the acceptance condition.

\[
\begin{array}{c}
\text{Figure 3.6: The PCA } A_{\text{token}}
\end{array}
\]

### 3.5 Fixed-Topology Communicating Automata

Recall that we aim at Büchi-Elgot-Trakhtenbrot Theorems for (parameterized) CAs. To put the parameterized setting into perspective, we will first recall corresponding results in the case of fixed topologies.

**Definition 3.11 (fixed-topology CA).** Let $T = (P, \rightarrow)$ be a topology over $\mathcal{N}$. A fixed-topology CA over $T$ is a CA $A = (S, \circ, \text{Msg}, \Delta, \mathcal{F})$ over $\mathcal{N}$ where the acceptance condition $\mathcal{F}$ is a set of mappings from $P$ into $S$.

Let $A = (S, \circ, \text{Msg}, \Delta, \mathcal{F})$ be a fixed-topology CA over $T = (P, \rightarrow)$. Moreover, suppose $M = (P, \rightarrow, E, \prec, \pi) \in \text{MSC}_{\{T\}}$ is an MSC (with topology $T$) and $\rho : E \rightarrow S$ a run of $A$ on $M$. Then, $\rho$ is accepting if $\lambda_{M, \rho} \in \mathcal{F}$. By $L(A)$, we denote the set of MSCs $M = (P, \rightarrow, E, \prec, \pi) \in \text{MSC}_{\{T\}}$ such that there is an accepting run of $A$ on $M$.

Once we fix a topology $T = (P, \rightarrow)$ (over $\mathcal{N}$), a corresponding MSO logic comes naturally. In particular, it features a predicate $p(x)$ to express that some event is executed by process $p \in P$.

**Definition 3.12.** For a topology $T = (P, \rightarrow)$, the syntax of the logic $\text{MSO}_T$ is given by the grammar

\[
\varphi ::= p(x) \mid a!(x) \mid a?(x) \mid x \preceq_{\text{proc}} y \mid x \preceq_{\text{msg}} y \mid x = y \mid x \in X \mid \neg \varphi \mid \varphi \lor \varphi \mid \exists x. \varphi \mid \exists X. \varphi
\]

where $p \in P$ and $a \in \mathcal{N}$.

As usual, we let $\text{EMSO}_T$ denote the set of formulas of the form $\exists X_1 \ldots \exists X_n. \varphi$ such that $\varphi$ does not contain any second-order quantification. For an MSC $M =$
\((P, \longrightarrow, E, \triangleleft, \pi) \in \text{MSC}_{\{T\}}\), a formula \(\varphi \in \text{MSO}_T\), and an interpretation \(I\) of variables, the satisfaction relation \(M \models I \varphi\) (\(M \models \varphi\) if \(\varphi\) is a sentence) is defined as expected. For example, \(M \models I p(x)\) if \(\pi(I(x)) = p\), \(M \models I a!(x)\) if \(\ell_M(I(x)) = a!\), and \(M \models I x \triangleleft_{\text{msg}} y\) if \(I(x) \triangleleft_{\text{msg}} I(y)\). Note that, since the topology is fixed, \(a!(x)\) and \(a?!(x)\) can actually be expressed using the predicates of the form \(p(x)\) and \(x \triangleleft_{\text{msg}} y\). We let \(L(\varphi)\) denote the set of MSCs \(M \in \text{MSC}_{\{T\}}\) such that \(M \models \varphi\).

We are now ready to present some previously known logical characterizations of fixed-topology CAs. Indeed, fixed-topology CAs over \(T\) and the logic EMSO\(_T\) are expressively equivalent:

**Theorem 3.13 ([BL06]).** Let \(T\) be a topology and \(L \subseteq \text{MSC}_{\{T\}}\). The following statements are equivalent:

1. There is a fixed-topology CA \(A\) over \(T\) such that \(L(A) = L\).
2. There is a sentence \(\varphi \in \text{EMSO}_T\) such that \(L(\varphi) = L\).

**Proof (sketch).** The direction 1. \(\Rightarrow\) 2. is standard.

The proof of the transformation 2. \(\Rightarrow\) 1. is based on Hanf’s normal form, which states that any first-order formula (over structures of bounded degree) is logically equivalent to a boolean combination of the form “there are at least \(n\) occurrences of neighborhood \(N\) of radius \(r\)” (for a uniform \(r \in \mathbb{N}\), which depends on the formula) [Han65, Lib04]. Here, the neighborhood of an event comprises all events that have distance at most \(r\) in the Gaifman graph \((E, \triangleleft \cup \triangleleft^{-1})\) of the underlying MSC. As MSCs are structures of bounded degree (every event has at most three neighbors), there are, up to isomorphism, only finitely many such neighborhoods. Now, to prove 2. \(\Rightarrow\) 1., one first constructs a (fixed-topology) CA that “recognizes” neighborhoods of radius \(r\). Another fixed-topology CA then counts these neighborhoods up to some threshold and determines, based on Hanf’s normal form, if the original formula is satisfied.

It was shown in [BK12] that the construction of Hanf’s normal form can be done in triply exponential time. This implies that Theorem 3.13 is actually effective and that the translation of a formula into an automaton takes only elementary time.

**Theorem 3.14 (cf. [BK12]).** The transformation of a fixed-topology CA from a formula as stated in Theorem 3.13 is effective and can be carried out in elementary time.

It was also shown in [BL06] that fixed-topology CAs over the pipeline \(T\) with two processes are not closed under complementation so that MSO\(_T\) is strictly more expressive. However, Genest, Kuske and Muscholl showed that the assumption of existentially bounded channels yields a robust model of message-passing systems:
Theorem 3.15 ([GKM06]). Let $\mathcal{T}$ be a topology, $k \in \mathbb{N}$, and $L \subseteq \text{MSC}_{\{\mathcal{T}\}} \cap \text{MSC}_{\exists k}$. The following statements are equivalent:

- There is a fixed-topology CA $A$ over $\mathcal{T}$ such that $L(A) = L$.
- There is a sentence $\varphi \in \text{MSO}_{\mathcal{T}}$ such that $L(\varphi) = L$.

In particular, it was shown in [GKM06] that there is a fixed-topology CA $A$ over $\mathcal{T}$ such that $L(A) = \text{MSC}_{\{\mathcal{T}\}} \cap \text{MSC}_{\exists k}$.

Note that, when restricting MSCs further to universally-$k$-bounded MSCs (all linearizations are $k$-bounded), one gets a similar logical characterization and a class of CAs that is determinizable [HMK+05]. On the other hand, determinizability fails in the case of existentially bounded channels [GKM07].

### 3.6 Parameterized Communicating Automata (PCAs)

It is our aim to lift the above logical characterizations of fixed-topology CAs to a parameterized setting. Recall that, when we fixed a topology, the acceptance condition could directly speak about all of its processes (Definition 3.11). However, when we have to cope with a class of topologies, this is no longer possible. It is then natural to specify acceptance in terms of a formula from a tailored monadic second-order logic, which scans the final configuration reached by a system: the underlying topology together with the local final states in which the processes terminate. If $S$ is the finite set of such local states, the formula thus defines a set of $S$-labeled topologies, i.e., structures $(P, \longrightarrow, \lambda)$ where $(P, \longrightarrow)$ is a topology and $\lambda : P \to S$. The corresponding logic $\text{tMSO}(S)$ is given by the grammar

$$
\mathcal{F} ::= \, u \xrightarrow{a | b} v \mid (\lambda(u) = s) \mid u = v \mid u \in U \mid \neg \mathcal{F} \mid \mathcal{F} \lor \mathcal{F} \mid \exists u. \mathcal{F} \mid \exists U. \mathcal{F}
$$

where $a, b \in \mathcal{N}$, $s \in S$, $u$ and $v$ are first-order variables (interpreted as processes), and $U$ is a second-order variable (ranging over sets of processes). As usual, we assume an infinite supply of variables. Satisfaction $(P, \longrightarrow, \lambda) \models \mathcal{F}$ for an $S$-labeled topology $\mathcal{T}$ and a sentence $\mathcal{F} \in \text{tMSO}(S)$ is defined as expected.

**Definition 3.16 (PCA).** A parameterized communicating automaton (PCA, for short) over $\mathcal{N}$ is a CA $A = (S, \iota, \text{Msg}, \Delta, \mathcal{F})$ over $\mathcal{N}$ where the acceptance condition $\mathcal{F}$ is a sentence from $\text{tMSO}(S)$.

Let $A = (S, \iota, \text{Msg}, \Delta, \mathcal{F})$ be a PCA over $\mathcal{N}$, let $M = (P, \longrightarrow, E, \preceq, \pi)$ be any MSC over $\mathcal{N}$, and suppose $\rho : E \to S$ is a run of $A$ on $M$. Recall that the mapping $\lambda_{M, \rho} : P \to S$ yields, for each process $p$, the state in which $p$ terminates. With this, $\rho$ is accepting if $(P, \longrightarrow, \lambda_{M, \rho}) \models \mathcal{F}$. By $L(A)$, we denote the set of MSCs $M$ such that there is an accepting run of $A$ on $M$.

A (syntactic) subclass of PCAs is obtained when the acceptance condition is only allowed to count terminal states of non-idle processes, up to some threshold:
Definition 3.17 (weak PCA). A PCA $A = (S, \iota, \text{Msg}, \Delta, F)$ is called weak if the acceptance condition $F$ is a boolean combination of sentences of the form $\exists^2 n. \lambda(u) = s$ where $n \in \mathbb{N}$ and $s \in S \setminus \iota(2^N)$.

Here, the formula $\exists^2 n. \lambda(u) = s$ is an abbreviation of

$$\exists u_1, \ldots, u_n. \bigwedge_{1 \leq i < j \leq n} u_i \neq u_j \land \bigwedge_{1 \leq i \leq n} \lambda(u_i) = s$$

which says that there are at least $n$ processes that terminate in state $s$. As $s \notin \iota(2^N)$, these processes have to be non-idle, meaning that they execute at least one event. Thus, unlike a PCA, a weak PCA is not allowed to talk about the topology but only about the events of an MSC. In particular, it cannot express that “the topology has five processes” but only “five processes are non-idle”. This is a priori weaker, but also makes sense, since it reflects the intuition that a PCA accepts behaviors rather than topologies.

The set of PCAs over $\mathcal{N}$ is denoted by $\text{PCA}_\mathcal{N}$, the class of weak PCAs over $\mathcal{N}$ by $\text{wPCA}_\mathcal{N}$.

Example 3.18. We resume Example 3.10 on page 46 describing the CA $A_{\text{token}}$ over $\{a, b\}$. To get a PCA, it remains to specify an appropriate acceptance condition $F$. Recall that we required that (i) exactly one process terminates in $s_1$ and (ii) no process terminates in a state from $\{t_1, t_2, t_3\}$. Formally, this is achieved using

$$F = \exists^1 u. \lambda(u) = s_1 \land \neg \exists^2 u. \lambda(u) = s_1 \land \bigwedge_{1 \leq i \leq 3} \neg \exists^1 u. \lambda(u) = t_i.$$ 

As $F$ is a boolean combination of statements of the form $\exists^2 n. \lambda(u) = s$ with $s \neq s_0$, the PCA $A_{\text{token}}$ is actually a weak PCA. Consider again the MSCs $M_1$ and $M_2$ from Figure 3.5 on page 43. We have $M_1 \in L(A_{\text{token}})$ and $M_2 \notin L(A_{\text{token}})$.

3.7 Logical Characterizations of PCAs

Since an MSC contains two types of varying objects, namely processes and events, it is natural to introduce a two-valued logic to reason about them. Accordingly, we have variables $u, v, w, \ldots$ and $U, V, W, \ldots$ to range over processes and sets of processes, respectively, as well as variables $x, y, z, \ldots$ and $X, Y, Z, \ldots$, ranging over events and sets of events, respectively.

Definition 3.19. The syntax of the logic $\text{MSO}_\mathcal{N}$ (over the fixed set of interface names $\mathcal{N}$) is given as follows:

$$\varphi ::= a!(x) \mid a?(x) \mid a \in \text{type}(\pi(x)) \mid x \prec_{\text{proc}} y \mid x \prec_{\text{msg}} y \mid x = y \mid x \in X \mid \neg \varphi \mid \varphi \lor \psi \mid \exists x. \varphi \mid \forall x. \varphi \mid \psi$$

$$\psi ::= x@u \mid u \overset{a}{\rightarrow} v \mid u = v \mid u \in U \mid \exists u. \varphi \mid \forall U. \varphi$$

where $a, b \in \mathcal{N}$.
All predicates are more or less self-explanatory apart from $x \sim y$, saying that $x$ and $y$ are located on the same process, and $x@u$, saying that $x$ is located on $u$. Note that $\sim$ can be expressed in terms of $\prec^\ast_{\text{proc}}$, since $x \sim y$ is equivalent to $(x \prec^\ast_{\text{proc}} y) \lor (y \prec^\ast_{\text{proc}} x)$. However, some of our results hold for logic fragments that include $\sim$ but discard $\prec^\ast_{\text{proc}}$. For the same reason, we include the predicate $a \in \text{type}(\pi(x))$, which is actually expressible using the formulas of type $\psi$.

There are indeed several natural fragments of MSO$_N$. When we do not allow formulas of the form $\psi$, then we obtain weak MSO, denoted wMSO$_N$. The fragment wEMSO$_N$ of wMSO$_N$ consists of the ($\psi$-free) formulas of the form $\exists x_1 \ldots \exists x_n. \varphi$ where $\varphi$ does not contain any second-order quantification. We will also refer to the fragment wEMSO$_N(\prec_{\text{proc}}, \sim, \prec_{\text{msg}})$ of wEMSO$_N$ where the predicate $\prec^\ast_{\text{proc}}$ is not allowed.

Let us formally define when $M \models_\mathcal{I} \varphi$ for an MSC $M = (P, \to, E, \llcorner, \lrcorner, \pi)$, a formula $\varphi \in \text{MSO}_N$, and an interpretation function $\mathcal{I}$, which maps

- a first-order event variable $x$ to an event $\mathcal{I}(x) \in E$,
- a first-order process variable $u$ to a process $\mathcal{I}(u) \in P$,
- a second-order event variable $X$ to a set $\mathcal{I}(X) \subseteq E$, and
- a second-order process variable $U$ to a set $\mathcal{I}(U) \subseteq P$.

We proceed by induction, but consider only some cases:

- $M \models_\mathcal{I} a!(x)$ if $\ell_M(\mathcal{I}(x)) = a!$,
- $M \models_\mathcal{I} a?(x)$ if $\ell_M(\mathcal{I}(x)) = a?$,
- $M \models_\mathcal{I} a \in \text{type}(\pi(x))$ if $a \in \text{type}(\pi(\mathcal{I}(x)))$,
- $M \models_\mathcal{I} x \sim y$ if $\pi(\mathcal{I}(x)) = \pi(\mathcal{I}(y))$,
- $M \models_\mathcal{I} x@u$ if $\pi(\mathcal{I}(x)) = \mathcal{I}(u)$,
- $M \models_\mathcal{I} u \overset{a \rightarrow b}{\rightarrow} v$ if $\mathcal{I}(u) \overset{a \rightarrow b}{\rightarrow} \mathcal{I}(v)$.

The other formulas are interpreted as expected. When $\varphi$ is a sentence, i.e., it does not have free variables, then satisfaction does not depend on the interpretation function so that we write $M \models \varphi$ instead of $M \models_\mathcal{I} \varphi$. In that case, the set of MSCs $M \in \text{MSC}$ such that $M \models \varphi$ is denoted by $L(\varphi)$.

**Example 3.20.** We continue the token-ring protocol from Examples 3.10 and 3.18. Recall that $\mathcal{N} = \{a, b\}$. We would like to express that there is a process that emits a message and gets an acknowledgment that results from a sequence of forwards through interface $a$. We first let $\text{fwd}(x, y) \equiv x \overset{a \rightarrow b}{\rightarrow} y \land \exists z. (x \prec_{\text{proc}} z \prec_{\text{msg}} y)$ where $x \overset{a \rightarrow b}{\rightarrow} y$ is a shorthand for $\exists u. \exists v. (x@u \land y@v \land u \overset{a \rightarrow b}{\rightarrow} v)$. It is well known that the transitive closure of the relation induced by $\text{fwd}(x, y)$ is definable in MSO$_N$, too. Let $\text{fwd}^+(x, y)$ be the corresponding formula. It expresses that there is a
sequence of events leading from $x$ to $y$ that alternately takes process and message edges, hereby following the causal order. With this, the desired formula is

$$
\varphi = \exists x, y, z.(x \preceq_{\text{proc}} y \land x \prec_{\text{msg}} z \land x \overset{a}{\underset{b}{\rightarrow}} z \land \text{fwd}^+(z, y)) \in \text{MSO}_{\mathcal{N}}.
$$

Consider again Figure 3.5 on page 43. We have $M_1 \models \varphi$ and $M_2 \not\models \varphi$, as well as $L(A_{\text{token}}) \subseteq L(\varphi)$. 

We now turn to logical characterizations of PCAs. We first consider weak PCAs, which turn out to be closely related to wEMSO$_\mathcal{N}$. However, note that restrictions on the logic and the topologies are necessary [Bol14]: over $\mathcal{N} = \{a, b, c, d\}$, there is a weak first-order formula without $\preceq^*_\text{proc}$ and $\sim$ that is not realizable by a weak PCA on the class of ring forests. Thus, we are left with a small margin for positive results. However, the following theorem is an analogon to Theorem 3.13, which was stated there for fixed topologies:

**Theorem 3.21 ([Bol14]).** Suppose $\mathcal{N} = \{a, b, c, d\}$. Let $\mathcal{T}$ be any of the following topology classes: pipelines, trees, grids, or rings. Moreover, let $L \subseteq \text{MSC}_\mathcal{T}$. The following statements are equivalent:

- There is $A \in \text{wPCA}_{\mathcal{N}}$ such that $L(A) \cap \text{MSC}_\mathcal{T} = L$.
- There is $\varphi \in \text{wEMSO}_{\mathcal{N}}(\preceq_{\text{proc}}, \sim, \prec_{\text{msg}})$ such that $L(\varphi) \cap \text{MSC}_\mathcal{T} = L$.

We leave open if Theorem 3.21 holds for the full logic wEMSO$_\mathcal{N}$. The answer is positive iff Theorem 3.13 still holds when we include the predicate $\preceq^*_\text{proc}$, but this is an open problem. However, we can safely include $\preceq^*_\text{proc}$ when we restrict to existentially bounded MSCs:

**Theorem 3.22 ([Bol14]).** Suppose $\mathcal{N} = \{a, b, c, d\}$. Let $\mathcal{T}$ be any of the following topology classes: pipelines, trees, grids, or rings. Moreover, let $k \in \mathbb{N}$ and $L \subseteq \mathcal{M} \overset{\text{def}}{=} \text{MSC}_\mathcal{T} \cap \text{MSC}_{3k}$. The following statements are equivalent:

- There is $A \in \text{wPCA}_{\mathcal{N}}$ over $\mathcal{N}$ such that $L(A) \cap \mathcal{M} = L$.
- There is a sentence $\varphi \in \text{wEMSO}_{\mathcal{N}}$ such that $L(\varphi) \cap \mathcal{M} = L$.

**Proof (sketch).** Like in the proof of Theorem 3.13, we make use of a normal form of first-order logic. However, due to the predicate $\preceq^*_\text{proc}$, the structures we consider do not have bounded degree anymore. Fortunately, there is a normal form due to Schwentick and Barthelmann [SB99] that does not rely on that assumption (like Gaifman’s normal form which, however, seems to be more difficult to apply in
Figure 3.7: Neighborhoods of e with radius r = 3

our setting). Their normal form essentially (modulo existential quantification) says that all events satisfy a formula $\varphi$ that only talks about a neighborhood of a certain radius $r$. Unlike in Hanf’s normal form, the size of a neighborhood is unbounded. This is illustrated in Figure 3.7 for several signatures. However, every neighborhood spans over a bounded sub-topology (unless we include a predicate $\prec^*$ in the logic that refers to the full partial order). Topologies, in turn, are structures of bounded degree so that there are finitely many such sub-topologies. For sub-topology $T$, using Theorem 3.15 on page 49 by Genest, Kuske, and Muscholl, we can construct a fixed-topology CA $A_T$ that evaluates $\varphi$ on $T$. The weak PCA we are looking for now glues these finitely many fixed-topology CAs together. More precisely, every process guesses a sub-topology $T$ and starts $A_T$. To make sure that the guess was correct, it communicates it to any communication partner. As neighbors also have to start their own fixed-topology CA, processes have to simulate several of them simultaneously. However, a weak PCA is not able to detect topology neighborhoods by itself. It needs some help from the underlying class of topologies. Informally speaking, we require that paths that form a cycle in some topology form a cycle everywhere, in any other topology from the underlying class. This is satisfied by the classes of pipelines, grid, and trees. For rings, some additional work is needed to get the result. Actually the theorem holds for any unambiguous class of topologies (see [Bol14] for details). Note that this proof works similarly for Theorem 3.21.

Let us now turn to (full) PCAs. Again, there is no chance to get an MSO characterization even when we restrict to 0-bounded and certain context-bounded MSCs. Actually, PCAs are not closed under complementation:

**Theorem 3.23 ([BGK14]).** Suppose $\mathcal{N} = \{a, b\}$. Let $ct \in \{s \oplus r, s_1+r_1\}$, let $\mathcal{X}$ be the set of pipelines, and set $\mathcal{M} \overset{\text{def}}{=} \text{MSC}_{\mathcal{X}} \cap \text{MSC}_{\mathcal{Z}} \cap \text{MSC}_{(1,ct)}$. There is $A \in \text{PCA}_{\mathcal{N}}$ such that, for all $A' \in \text{PCA}_{\mathcal{N}}$, we have $L(A') \cap \mathcal{M} \neq M \setminus L(A)$. 

53
Proof (sketch). Even under the context-type restrictions $s \oplus r$ and $s_1 + r_1$, MSCs over pipelines can encode grid-like structures (cf. Figures 3.8 and 3.9). For grids, there have been similar results, showing that graph acceptors are not complementable [Tho96, MST02]. By the grid encoding, the non-complementability result can be transferred to our setting.

The next theorem does not rely on a particular class of topologies. However, recall that we fixed some finite set $\mathcal{N}$ of interface names.

**Theorem 3.24 ([BGK14]).** Let $k \geq 1$ and $ct \in \{s \oplus r_1, \text{inf}\}$. Moreover, let $L \subseteq \mathcal{M} \overset{\text{def}}{=} \text{MSC}_{s_0} \cap \text{MSC}_{(k,ct)}$. The following statements are equivalent:

- There is $A \in \text{PCA}_{\mathcal{N}}$ such that $L(A) \cap \mathcal{M} = L$.
- There is a sentence $\varphi \in \text{MSO}_{\mathcal{N}}$ such that $L(\varphi) \cap \mathcal{M} = L$.

Proof (sketch). Again, the construction of a formula from an automaton is rather standard, though we have to cope with a two-sorted logic.

The other direction is by induction on the structure of the formula. In particular, negation translates to complementability of (context-bounded) PCAs, which is the key technical issue in the proof. Complementation relies on a disambiguation construction. From the given PCA $A$, one constructs another PCA that is unambiguous and equivalent to $A$ on $\text{MSC}_{(k,ct)}$. Unambiguous here means that every MSC comes with exactly one run (accepting or not). The main idea for this construction is to divide the events of an MSC from $\text{MSC}_{(k,ct)}$ into zones, as illustrated in Figure 3.10. We only illustrate the more complicated case $ct = s \oplus r_1$ (the case $\text{inf}$ is similar). An interval is a set of successive (i.e., not interrupted) events that belong to one and the same process. Then, a zone consists of an interval of successive send events of one process, together with the corresponding receive events, as long as the latter form intervals themselves on the corresponding processes.

One can show that, for each MSC in $\text{MSC}_{(k,s \oplus r_1)}$, there is a zone partitioning such that every process traverses at most $K = k \cdot (|\mathcal{N}|^2 + 2|\mathcal{N}| + 1)$ different zones. The crucial point is now that, on $\text{MSC}_{(k,s \oplus r_1)}$, zones can be computed deterministically by a PCA. During that procedure, a sending process will maintain a summary of all possible “global transitions” induced by a zone, by simulating all possible transitions.
of the receiving processes. The acceptance condition will then check in terms of a tMSO formula whether the summaries can be glued together towards an accepting run. Once we have an unambiguous PCA, complementation is accomplished just by negating the acceptance condition.

Interestingly, determinization procedures have been used to obtain complementability and MSO characterizations for context-bounded and scope-bounded multi-pushdown automata [LMP10a, LNP14a] (cf. Chapter 3). A pattern that we share with these approaches is that of computing summaries in a deterministic way. Overall, however, we have to use quite different techniques, which is due to the fact that, in our model, processes evolve asynchronously.

3.8 Model Checking

From the previous section, we conclude that context-bounded PCAs with rendezvous communication form a robust automata model: they are closed under all boolean operations and enjoy a logical characterization in terms of an unrestricted MSO logic. But there is more: the emptiness problem is decidable for some topology classes, which is stated in the following theorem.

**Theorem 3.25 ([BGS14]).** Suppose $\mathcal{N} = \{a, b, c, d\}$. Let $ct \in \{s@r1, intf\}$, and let $\mathcal{T}$ be any of the following topology classes: pipelines, trees, or rings. The following problem is decidable:

**Instance:** $\mathcal{A} \in \text{PCA}_{\mathcal{N}}$; $k \geq 1$

**Question:** $L(\mathcal{A}) \cap \text{MSC}_{\mathcal{T}} \cap \text{MSC}_{\exists 0} \cap \text{MSC}_{(k, ct)} \neq \emptyset$?

**Proof (sketch).** The proof relies on the zone partitioning that we used for Theorem 3.24. Consider the case of pipelines. Roughly speaking, we construct a finite automaton (running over pipelines) that guesses a run of the given PCA in terms
of zones and state summaries. As we can assume that every process traverses a bounded number of zones, we actually deal with a finite automaton over a finite alphabet. The finite automaton can now check if its guess actually corresponds to an accepting run. Thus, it accepts all those pipelines that allow for a context-bounded run of the given PCA. Emptiness of that exponentially big (in \( k \)) finite automaton can be checked in \textit{Pspace}. Accordingly, for tree topologies, we construct a tree automaton. In the case of rings, we essentially use the construction for pipelines. However, some additional work is needed to rule out cyclic dependencies that would contradict the run definition.

Over pipelines and rings, the decision problem from Theorem 3.25 is \textit{Pspace}-complete when the acceptance condition is presented as a finite automaton (instead of an MSO formula). Over trees, it is \textit{EXPTIME}-complete when the acceptance condition is given in terms of a tree automaton. In all cases, we assume that the context bound \( k \) is encoded in unary.

From Theorems 3.24 and 3.25, we deduce that context-bounded model checking of PCAs against MSO properties is decidable:

\begin{theorem}[[BGS14, BGK14]]. Let \( ct \in \{s \oplus r1, intf\} \) and let \( \Xi \) be any of the following topology classes over \( \mathcal{N} = \{a, b, c, d\} \): pipelines, trees, or rings. The following problem is decidable:

\begin{itemize}
  \item \textsc{Instance:} \( A \in \text{PCA}_\mathcal{N} \); \( \varphi \in \text{MSO}_\mathcal{N} \); \( k \geq 1 \)
  \item \textsc{Question:} \( L(A) \cap \text{MSC}_\Xi \cap \text{MSC}_{\exists 0} \cap \text{MSC}_{(k,ct)} \subseteq L(\varphi) \)?
\end{itemize}
\end{theorem}

Let us conclude by comparing the above model-checking result with related work. Maybe the most interesting feature in view of previous results on model checking parameterized systems is that our MSO logic is \textit{unrestricted}, while other approaches rely on dropping temporal operators such as the next modality from LTL [EN03, AJKR14, JB14]. However, as demonstrated in Example 3.20, next modalities make a lot of sense once the behavior of a parameterized system is modeled as a partial order or MSC rather than a string.

In general, there are several orthogonal ways to get decidability of model checking in a parameterized setting. For example, nonemptiness of PCAs becomes decidable over rings when tokens are unary (i.e., \(|MSG| = 1\)) [EN03, AJKR14]. Note that our token-ring protocol from Example 3.10 on page 46 assumes a binary token \( MSG = \{0, 1\} \), for which nonemptiness is already undecidable in general. However, as we have seen, we obtain decidability when restricting to a bounded number of contexts. Point-to-point communication yields an undecidable model, too, unless processes (i) choose a communication partner nondeterministically [AKR+14], (ii) use broadcasting [DSZ10, DSZ11], (iii) restrict the topologies in a suitable manner (e.g., bounded depth) [Mey08, EGM13, AKR+14], or (iv) bound the number of contexts (our approach).
3.9 Perspectives

The results presented in this chapter can only be called a first step towards a “regular” language theory of parameterized concurrent systems.

Note that Theorems 3.25 and 3.26 apply to concrete “regular” classes of topologies, namely pipelines, trees, and rings. It will be interesting to see if this can be extended to classes of bounded tree/clique width. This is the approach from [AKR+14], though for topologies of unbounded degree.

One should actually try to extend our results to automata models that take account of communication topologies of unbounded degree such as star topologies or, more generally, unranked trees. Such automata may be equipped with registers so that processes can remember, at any time, some of their neighbors [DST13, BCH+13] (see also Chapters 4 and 5). Moreover, many distributed protocols are able to exchange and compare process identifiers. Actually, the logic MSO_N is so powerful that it is capable of tracing back the origin of a pid in such a system. This will possibly allow us to verify also leader-election protocols and alike. Similar ideas will be applied in Chapter 4 to automata with registers, which can store values from an infinite alphabet (such as pids) and compare them for equality.

Finally, there seem to be close connections with the area of distributed algorithms, which should be explored further. Indeed, there are algorithms that evaluate a given process architecture wrt. logical specifications [GW10] or construct a map of an anonymous graph [CDK10], which is similar to what a (weak) PCA does in the proof of Theorem 3.22.
Like the previous chapter, this chapter also deals with systems whose behavior involves an unbounded number of processes. However, the systems modeled here may be considered dynamic, since there will be a feature that allows them to create fresh process identifiers. We will first consider sequential systems, meaning that there is a global control to which all participating processes have access. Consequently, our system model will have one single state space. A study of a model of dynamic concurrent systems can be found in Chapter 5.

An action that is performed by a dynamic sequential system is taken from an infinite alphabet. The alphabet is the cartesian product of a finite part $A$, denoting the type or label of an action, and an infinite part $D$, denoting the pid of the process executing the action. In a more general setting, the elements of the infinite set are often referred to as data values. As our systems are sequential, a behavior may be described as a string over $A \times D$, which is commonly called a data word. Similarly to words over a distributed (visibly) alphabet (cf. Chapter 2), which induce nesting relations, a data word comes with a natural graph structure in terms of edges between two successive word positions as well as between two position that carry the same data value.

Recall that we are interested in modeling systems as automata. The study of automata over infinite alphabets has its origins in the seminal work by Kaminski and Francez [KF94]. Their finite-memory automata (more commonly called register automata) equip finite-state machines with registers in which data values can be stored and be reused later. Register automata preserve some of the nice properties of finite automata: they have a decidable emptiness problem and are closed under union and intersection. On the other hand, they are neither determinizable nor closed under complementation. There are actually several variants of register
automata, which all have the same expressive power but differ in the complexity of decision problems (see [DL09]).

Subsequently, many more automata models have been considered, aiming at a good balance between expressivity and decidability [NSV04, DL09, KZ10, BL10]. Motivated by the study of logics over data words, Bojanczyk et al. introduced data automata [BDM+11]. While register automata are a one-way model, data automata read a data word twice: a finite-state transducer first scans its projection onto $A$, outputting a word of equal length over some finite alphabet $A'$. A second device, a finite automaton over $A'$, then checks every projection to positions with the same data value.

Data automata are expressively equivalent to the one-way model of class memory automata [BS10]. A class memory automaton does not use registers either. Instead, when reading a word position with, say, some data value $d$, it can access the state taken after executing the prior position with $d$. This is very much in the spirit of nested-word automata where, at some return position, one can access the state (or, stack symbol) associated with the corresponding call (cf. Chapter 2). We come back to that issue in Section 4.3. Class memory automata [BS10] feature a notion of freshness. A transition of a class memory automaton explicitly allows a data value to be declared as fresh, i.e., to occur for the first time in the history of a run. Freshness is an important notion in programming languages when names are supposed to be unique. In our context, it is crucial, since we require that processes are identifiable by unique pids. Freshness for register automata was introduced in [Tze11]. Like ordinary register automata, fresh-register automata preserve some of the good properties of finite automata. However, they inherit negative results of register automata and have, for example, an undecidable universality/equivalence problem.

In this chapter, we are aiming at a robust automata model for data words, i.e., a model that is closed under complementation (in a restricted sense), has a decidable equivalence problem, and enjoys a logical characterization. Moreover, it should be suitable as a model of dynamic sequential systems. In particular, this means that we are looking for a one-way model so that (variants of) data automata, alternating automata, or pebble automata are out of the question. In Section 4.1, we propose session automata. Like register automata, session automata are a syntactical restriction of fresh-register automata, but in an orthogonal way. While register automata drop the feature of global freshness (referring to the history) and keep a local variant (referring to the registers), session automata discard local freshness, while keeping the global one. In Section 4.3, we will then look at a more general model that unifies several of the above-mentioned models but still comes with a solution to the realizability problem.

\footnote{In [BHL14], we were also aiming at an application in the context of automata learning, where decidability of equivalence is crucial. However, this issue is beyond the scope of the thesis.}
4.1 (Fresh-)Register Automata and Session Automata

Let, in the following, $A$ be a finite alphabet and $D$ be an infinite alphabet. We will assume that $A$ and $D$ are disjoint.

For a set $\mathcal{R}$, we let $\mathcal{R}^\oplus \overset{\text{def}}{=} \{ \mathcal{r}^\oplus \mid \mathcal{r} \in \mathcal{R} \}$, $\mathcal{R}^\odot \overset{\text{def}}{=} \{ \mathcal{r}^\odot \mid \mathcal{r} \in \mathcal{R} \}$, and $\mathcal{R}^\uparrow \overset{\text{def}}{=} \{ \mathcal{r}^\uparrow \mid \mathcal{r} \in \mathcal{R} \}$. Below, we introduce automata with a set of registers $\mathcal{R}$. Transitions will be labeled with an element from $\mathcal{R}^\oplus \cup \mathcal{R}^\odot \cup \mathcal{R}^\uparrow$, which determines a register and the operation that is performed on it. More precisely, $\mathcal{r}^\oplus$ writes a globally fresh value into $\mathcal{r}$, $\mathcal{r}^\odot$ writes a locally fresh value into $\mathcal{r}$, and $\mathcal{r}^\uparrow$ uses the value that is currently stored in $\mathcal{r}$. For $\pi \in \mathcal{R}^\oplus \cup \mathcal{R}^\odot \cup \mathcal{R}^\uparrow$, we let $\text{reg}(\pi) = r$ if $\pi \in \{ \mathcal{r}^\oplus, \mathcal{r}^\odot, \mathcal{r}^\uparrow \}$.

Similarly,

$$\text{op}(\pi) = \begin{cases} \oplus & \text{if } \pi \text{ is of the form } \mathcal{r}^\oplus \\ \odot & \text{if } \pi \text{ is of the form } \mathcal{r}^\odot \\ \uparrow & \text{if } \pi \text{ is of the form } \mathcal{r}^\uparrow. \end{cases}$$

Definition 4.1 (fresh-register automaton, cf. [Tze11]). A fresh-register automaton (FRA) over $A$ and $D$ is a tuple $\mathcal{A} = (S, \mathcal{R}, \iota, F, \Delta)$ where

- $S$ is the nonempty finite set of states,
- $\mathcal{R}$ is the nonempty finite set of registers,
- $\iota \in S$ is the initial state,
- $F \subseteq S$ is the set of final states, and
- $\Delta$ is a finite set of transitions.

A transition is a tuple of the form $(s, (a, \pi), s')$ where $s, s' \in S$ are the source and target state, respectively, $a \in A$, and $\pi \in \mathcal{R}^\oplus \cup \mathcal{R}^\odot \cup \mathcal{R}^\uparrow$. We call $(a, \pi)$ the transition label.

For a transition $(s, (a, \pi), s') \in \Delta$, we also write $s \xrightarrow{(a,\pi)} s'$. When taking this transition, the automaton moves from state $s$ to state $s'$ and reads a symbol of the form $(a, d) \in A \times D$. If $\pi = \mathcal{r}^\uparrow \in \mathcal{R}^\uparrow$, then $d_i$ is the data value that is currently stored in $\mathcal{r}$. If $\pi = \mathcal{r}^\odot \in \mathcal{R}^\odot$, then $d$ is some globally fresh data value that has not been read in the whole history of the run; $d$ is then written into register $\mathcal{r}$. Finally, if $\pi = \mathcal{r}^\oplus \in \mathcal{R}^\oplus$, then $d$ is some locally fresh data value that is currently not stored in the registers; it will henceforth be stored in register $\mathcal{r}$.

Let us formally define the semantics of $\mathcal{A}$. A configuration is a triple $\gamma = (s, \tau, U)$ where $s \in S$ is the current state, $\tau : \mathcal{R} \rightarrow D$ is a partial mapping denoting the current register assignment, and $U \subseteq D$ is the set of data values that have been used so far. We say that $\gamma$ is final if $s \in F$. As usual, we define a transition relation over configurations and let $(s, \tau, U) \xrightarrow{(a,d)} (s', \tau', U')$, where $(a, d) \in A \times D$, if there is a transition $s \xrightarrow{(a,\pi)} s'$ such that the following hold:
\[\begin{align*}
&\begin{cases}
  d = \tau(\text{reg}(\pi)) & \text{if } \text{op}(\pi) = \uparrow \\
  d \notin \tau(\mathcal{R}) & \text{if } \text{op}(\pi) = \odot \\
  d \notin U & \text{if } \text{op}(\pi) = \otimes,
\end{cases}
\end{align*}\]

1. \(\text{dom}(\pi') = \text{dom}(\pi) \cup \{\text{reg}(\pi)\}\) and \(U' = U \cup \{d\}\).

2. \(\text{dom}(\pi') = \text{dom}(\pi) \cup \{\text{reg}(\pi)\}\) and \(U' = U \cup \{d\}\).

3. \(\pi'(\text{reg}(\pi)) = d\) and \(\pi'(r) = \pi(r)\) for all \(r \in \text{dom}(\pi) \setminus \{\text{reg}(\pi)\}\).

A run of \(A\) on a data word \((a_1, d_1) \ldots (a_n, d_n) \in (A \times D)^*\) is a sequence
\[\gamma_0 \xrightarrow{(a_1, d_1)} \gamma_1 \xrightarrow{(a_2, d_2)} \ldots \xrightarrow{(a_n, d_n)} \gamma_n\]
for suitable configurations \(\gamma_0, \ldots, \gamma_n\) with \(\gamma_0 = (i, \emptyset, \emptyset)\). The run is accepting if \(\gamma_n\) is a final configuration. The language \(L(A) \subseteq (A \times D)^*\) of \(A\) is then defined as the set of data words for which there is an accepting run. Note that FRAs cannot distinguish between data words that are equivalent up to permutation of data values: for \(w, w' \in (\Sigma \times D)^*\), we write \(w \equiv w'\) if \(w = (a_1, d_1) \ldots (a_n, d_n)\) and \(w' = (a_1, d'_1) \ldots (a_n, d'_n)\) such that, for all \(i, j \in [n]\), we have \(d_i = d_j\) if \(d'_i = d'_j\). For instance, \((a, 4)(b, 2)(b, 4) \equiv (a, 2)(b, 5)(b, 2)\). We call \(L \subseteq (\Sigma \times D)^*\) a data language if, for all \(w, w' \in (\Sigma \times D)^*\) such that \(w \approx w'\), we have \(w \in L\) iff \(w' \in L\). In particular, \(L(A)\) is a data language for every FRA \(A\).

We obtain natural subclasses of fresh-register automata when we restrict the transition labels \((a, \pi) \in A \times (\mathcal{R}^\odot \cup \mathcal{R}^\uparrow)\) in the transitions.

**Definition 4.2 (register automaton [KF94])**. A register automaton (RA) is an FRA \((S, \mathcal{R}, i, F, \Delta)\) where every transition label is from \(A \times (\mathcal{R}^\odot \cup \mathcal{R}^\uparrow)\).

**Definition 4.3 (session automaton [BHLM14])**. A session automaton (SA) is an FRA \((S, \mathcal{R}, i, F, \Delta)\) where every transition label is from \(A \times (\mathcal{R}^\odot \cup \mathcal{R}^\uparrow)\).

The following example will demonstrate that RAs and SAs are incomparable wrt. expressive power, and that FRAs are strictly more expressive than both restrictions.

**Example 4.4.** Consider the set of labels \(A = \{\text{req, ack}\}\) and the set of data values \(D = \mathbb{N}\), representing an infinite supply of pids. We model a simple (sequential) system where processes can approach a server and make a request, indicated by \text{req}, and where the server can acknowledge these requests, indicated by \text{ack}. More precisely, \((\text{req}, d) \in A \times D\) means that the process with pid \(d\) performs a request, which is acknowledged when the system executes \((\text{ack}, d)\). Consider the following data languages:

- \(L_1 = \text{“there are at most two open requests at a time”}\)
- \(L_2 = \text{“a process waits for acknowledgment before its next request”}\)
- \(L_3 = \text{“every acknowledgment is preceded by a request”}\)
- \(L_4 = \text{“requests are acknowledged in the order they are received”}\)

62
• $L_5$ = “every process makes at most one request”
• $L_6$ = “all requests take place before all acknowledgments”

Figure 4.1 depicts an RA that recognizes the data language $L_1 \cap \ldots \cap L_4$. That is, it models a server that can store two requests at a time and will acknowledge them in the order they are received. For example, it accepts the data word (req, 8)(req, 4)(ack, 8)(req, 3)(ack, 4)(req, 8)(ack, 3)(ack, 8). Note that a transition label containing $r_i^\circ \lor r_i^\uparrow$ actually refers to two transitions, one using $r_i^\circ$ and one using $r_i^\uparrow$.

When, in addition, we want to guarantee that every process makes at most one request, we need the freshness operator. Figure 4.1 depicts the SA $A_{4.1}$ recognizing $L_1 \cap \ldots \cap L_5$. We obtain $A_{4.2}$ from $A_{4.1}$ by replacing any occurrence of $r_i^\circ \lor r_i^\uparrow$ with $r_i^\circ$. While (req, 8)(req, 4)(ack, 8)(req, 3)(ack, 4)(req, 8)(ack, 3)(ack, 8) is no longer contained in $L(A_{4.2})$, (req, 8)(req, 4)(ack, 8)(req, 3)(ack, 4)(req, 8)(ack, 3)(ack, 3) is still accepted. Note that $L(A_{4.2})$ is indeed not recognizable by a (non-fresh) RA.

To separate FRAs from RAs and SAs, simply consider data language $L_5$. Clearly, it is not recognized by any RA nor by any SA, while an FRA recognizing it is given in Figure 4.3 (where $r$ is the only register).
However, it is easily seen that there is no FRA that recognizes $L_5 \cap \ldots \cap L_6$. To address this weakness of FRAs, we will later introduce an automata model that actually captures that language (Section 4.3).

From Example 4.4, we deduce the following:

**Corollary 4.5.** FRAs are strictly more expressive than both RAs and SAs, while SAs and RAs are incomparable wrt. expressive power.

### 4.2 Closure Properties of Session Automata

Closure properties of SAs and decidability will be established by means of a normal form of a data word. The crucial observation is that, in an SA, data equality in a data word only depends on the transition labels that generate it. In this section, we suppose that the set of registers of a session automaton is of the form $\mathcal{R} = \{1, \ldots, k\}$ so that we can assume a natural total ordering on it. In the following, let $\Omega \overset{\text{def}}{=} (\mathbb{N}_{>0})^\circ \cup (\mathbb{N}_{>0})^\uparrow$.

Suppose an SA reads a sequence $u = (a_1, \pi_1) \ldots (a_n, \pi_n) \in (A \times \Omega)^*$ of transition labels. We call $u$ a symbolic word. It “produces” a data word iff a register is initialized before it is used. Formally, we say that $u$ is well-formed if, for all positions $j \in [n]$ with $\text{op}(\pi_j) = \uparrow$, there is $i \in [n]$ such that $\pi_i = \text{reg}(\pi_j)^\circ$. Let $\text{WF} \subseteq (A \times \Omega)^*$ be the set of all well-formed words. With $u \in (A \times \Omega)^*$, we can associate an equivalence relation $\sim_u \subseteq [n] \times [n]$, letting $i \sim_u j$ iff

- $\text{reg}(\pi_i) = \text{reg}(\pi_j)$, and
- $i \leq j$ and there is no position $k \in \{i + 1, \ldots, j\}$ such that $\pi_k = \text{reg}(\pi_i)^\circ$, or $j \leq i$ and there is no position $k \in \{j + 1, \ldots, i\}$ such that $\pi_k = \text{reg}(\pi_i)^\circ$.

If $u$ is well-formed, then the data values of any data word $w = (a_1, d_1) \ldots (a_n, d_n)$ that is “accepted via” $u$ conform with the equivalence relation $\sim_u$. That is, we have $d_i = d_j$ iff $i \sim_u j$. We call $w$ a concretization of $u$. Let $\text{data}(u)$ denote the set of all concretizations of $u$. This is extended to sets $L \subseteq (A \times \Omega)^*$ of well-formed words, and we set $\text{data}(L) \overset{\text{def}}{=} \bigcup_{u \in L \cap \text{WF}} \text{data}(u)$. Note that, here, we first filter the well-formed words before applying the operator. Now, let $A = (S, \mathcal{R}, \iota, F, \Delta)$ be an SA. In the obvious way, we may consider $A$ as a finite automaton over $A \times (\mathcal{R}^\circ \cup \mathcal{R}^\uparrow)$. We then obtain a regular language $L_{\text{symb}}(A) \subseteq (A \times \Omega)^*$. It is not difficult to verify that $L(A) = \text{data}(L_{\text{symb}}(A))$.  

![Diagram](image-url)
Though we have a symbolic representation of data languages recognized by SAs, it is in general difficult to compare their languages, since quite different symbolic words may give rise to the same concretizations. For example, we have $\text{data}((a, 1^\circ)(a, 1^\circ)(a, 1^\downarrow)) = \text{data}((a, 1^\circ)(a, 2^\circ)(a, 2^\downarrow))$. However, we can associate, with every data word, a (symbolic) normal form, producing the same set of concretizations. Intuitively, the normal form uses the first (according to the natural total order) register whose current data value is not used anymore. In the above example, $(a, 1^\circ)(a, 1^\circ)(a, 1^\downarrow)$ would be in symbolic normal form: the data value stored at the first position in register 1 is not reused so that, at the second position, register 1 has to be overwritten. For that reason, $(a, 1^\circ)(a, 2^\circ)(a, 2^\downarrow)$ is not in symbolic normal form. In contrast, $(a, 1^\circ)(a, 2^\circ)(a, 2^\downarrow)(a, 1^\uparrow)$ is in normal form, since register 1 is read at the end of the word.

Let $w = (a_1, d_1) \ldots (a_n, d_n) \in (A \times D)^*$ be a data word and $d \in \{d_1, \ldots, d_n\}$. By first$(d)$, we denote the position where $d$ occurs for the first time, i.e., such that $d_j = d$ and there is no $k < j$ such that $d_k = d$. Accordingly, we define last$(d)$ to be the last position where $d$ occurs. We define the symbolic normal form $\text{snf}(w) = (a_1, \pi_1) \ldots (a_n, \pi_n) \in (A \times \Omega)^*$ of $w$ inductively, along with sets $\text{Free}(i) \subseteq \mathbb{N}$ indicating the registers that are reusable after executing position $i \in [n]$. Setting $\text{Free}(0) = \mathbb{N}_{>0}$, we define

$$\pi_i = \begin{cases} 
\min(\text{Free}(i - 1)) & \text{if } i = \text{first}(d_i) \\
\text{reg}(\pi_{\text{first}(d_i)}) & \text{otherwise},
\end{cases}$$

and

$$\text{Free}(i) = \begin{cases} 
\text{Free}(i - 1) \setminus \min(\text{Free}(i - 1)) & \text{if } i = \text{first}(d_i) \neq \text{last}(d_i) \\
\text{Free}(i - 1) \cup \{\text{reg}(\pi_i)\} & \text{if } i = \text{last}(d_i) \\
\text{Free}(i - 1) & \text{otherwise}.
\end{cases}$$

We canonically extend $\text{snf}$ to data languages $L$, setting $\text{snf}(L) = \{\text{snf}(w) \mid w \in L\}$.

**Example 4.6.** Let $w = (a, 8)(b, 4)(a, 8)(c, 3)(a, 4)(b, 3)(a, 9)$. Then, we have that $\text{snf}(w) = (a, 1^\circ)(b, 2^\circ)(a, 1^\uparrow)(c, 1^\circ)(a, 2^\circ)(b, 1^\uparrow)(a, 1^\circ)$. \hfill $\Diamond$

![Figure 4.4: Relation between data words and symbolic normal forms](image)

The relation between the mappings $\text{data}$ and $\text{snf}$ is illustrated in Figure 4.4. One easily verifies that $L = \text{data}(\text{snf}(L))$, for all data languages $L$. Therefore, equality of data languages reduces to equality of their symbolic normal forms:

**Lemma 4.7.** Let $L$ and $L'$ be data languages. Then, $L = L'$ iff $\text{snf}(L) = \text{snf}(L')$. 

65
Of course, symbolic normal forms may use any number of registers so that the set of symbolic normal forms is a language over an infinite alphabet as well. However, given an SA $A$, the symbolic normal forms that represent the language $L(A)$ do with a bounded (i.e., finite) number of registers. Indeed, an important notion in the context of SAs is the bound of a data word. Intuitively, the bound of $w = (a_1, d_1) \ldots (a_n, d_n) \in (A \times D)^*$ is the minimal number of registers that an SA needs to execute $w$. Or, in other words, the bound is the maximal number of overlapping scopes, where a scope is an interval delimiting the occurrences of one particular data value. Formally, a scope of $w$ is any set $I \in \{1, \ldots, n\}$ such that there is $d \in \{d_1, \ldots, d_n\}$ with $I = \{\text{first}(d), \ldots, \text{last}(d)\}$. Given a natural number $k \geq 1$, we say that $w$ is $k$-bounded if every position $i \in \{1, \ldots, n\}$ is contained in at most $k$ scopes. One can verify that a data word is $k$-bounded if $\text{snf}(w)$ is a word over the alphabet $A \times \Omega_k$ where $\Omega_k \overset{\text{def}}{=} \{1, \ldots, k\} \cup \{1, \ldots, k\}^\dagger$.

Let $\text{DW}_k$ denote the set of $B$-bounded data words. A data language $L$ is $k$-bounded if $L \subseteq \text{DW}_k$. Note that the set of all data words is not $k$-bounded, for any $k$.

**Theorem 4.8 ([BHLM14]).** Let $A = (S, \mathcal{R}, \iota, F, \Delta)$ be an SA with $\mathcal{R} = \{1, \ldots, k\}$. Then, $L(A)$ is $k$-bounded. Moreover, $\text{snf}(L(A))$ is a regular language over the finite alphabet $A \times \Omega_k$. A corresponding automaton can be effectively computed. Its number of states is at most exponential in $k$ and linear in $|S|$.

Using Theorem 4.8, we obtain the following closure properties of SA:

**Theorem 4.9 ([BHLM14]).** We have the following closure properties:

- SAs are closed under union and intersection.
- SAs are complementable for $k$-bounded data words: Given an SA $A$ with $k$ registers, there is an SA $B$ with $k$ registers such that $L(B) = \text{DW}_k \setminus L(A)$.

Note that, while FRA are not complementable wrt. the set of all data words, they are complementable for $k$-bounded data words. The reason is that, given an FRA $A$, one can construct an SA $B$ such that $L(B) = L(A) \cap \text{DW}_k$.

### 4.3 Class Register Automata

Example 4.4 on page 62 revealed a certain weakness of FRAs concerning their expressiveness and, therefore, modeling power. In this section, we will present an automata model that captures the language $L_3 \cap \ldots \cap L_6$ in question. Despite its expressive power (which entails undecidability of its emptiness problem), it
has a natural mode of operation, taking into account the local control flow of a process. Moreover, it captures previously defined automata (as well as others from the literature) as special cases. Our model is named class register automata, since it combines register automata and class memory automata [BS10]. Though basic decision problems are undecidable for that model, we will later see that its expressive power lies between EMSO and MSO logic so that it is still, in a sense “regular”.

Similarly to nested words and message sequence charts (cf. Chapters 2 and 3), our new automata model will access a graph structure that we add on top of a data word. Actually, we associate with a data word \( w = (a_1, d_1) \ldots (a_n, d_n) \in (A \times D)^* \) two relations, \( \prec_{+1} \subseteq [n] \times [n] \) and \( \prec_{\oplus 1} \subseteq [n] \times [n] \). We let \( \prec_{+1} = \{(i, i+1) \mid i \in [n-1]\} \). Moreover, for all \( i, j \in [n] \), we let \( i \prec_{\oplus 1} j \) iff \( i < j \), \( d_i = d_j \), and there is no \( k \in \{i + 1, \ldots, j - 1\} \) such that \( d_i = d_k \). Thus, \( i \prec_{\oplus 1} j \) denotes the fact that \( i \) and \( j \) are successive positions carrying the same data value.

**Example 4.10.** In Figure 4.5, we illustrate the two relations associated with a data word \( w \) over \( \Sigma = \{a, b\} \) and \( D = \mathbb{N} \). Straight edges represent the relation \( \prec_{+1} \), while curved edges represent \( \prec_{\oplus 1} \).

![Figure 4.5: The relations \( \prec_{+1} \) and \( \prec_{\oplus 1} \)](image)

A run of a class register automaton on a data word \( w = (a_1, d_1) \ldots (a_n, d_n) \) proceeds as follows. Like an FRA, it comes with a finite set of registers, say \( \mathcal{R} \). Like in an FRA, data values may be written into these registers and be reused later. The main difference is that, when reading a position \( j \) in the data word, the automaton can access the contents of some register \( r \) both

- at position \( i \) with \( i \prec_{+1} j \) (denoted \( r^r \)), and
- at position \( i \) with \( i \prec_{\oplus 1} j \) (denoted \( r^\ominus \)).

This looks all the more reasonable, since \( \prec_{\oplus 1} \) reflects the control flow of one single process. Moreover, it is in the spirit of nested-word automata (cf. Chapter 2). So, let \( \mathcal{R}^- \overset{\text{def}}{=} \{r^r \mid r \in \mathcal{R}\} \) and \( \mathcal{R}^\ominus \overset{\text{def}}{=} \{r^\ominus \mid r \in \mathcal{R}\} \). Actually, a transition will be constrained by a guard, which is a boolean combination of atoms \( \theta_1 = \theta_2 \) where \( \theta_1, \theta_2 \in \{c\} \cup \mathcal{R}^- \cup \mathcal{R}^\ominus \). Here, \( c \) stands for “current data value”. The set of guards (over \( \mathcal{R} \)) is denoted by \( \text{Guard}(\mathcal{R}) \). When executing a transition, registers may obtain new data values: either the data value currently read, or one of the data values previously stored in registers, or a “guessed” data value: any
data value that is in the \( \ell \)-neighborhood (as defined below) of the given position, where \( \ell \in \mathbb{N} \) is a radius. This is specified by an update, which is a partial function 
\[ \nu : \mathcal{R} \to \mathcal{R}^{-} \cup \mathcal{R}^{\ominus} \cup \{0\} \]. The set of updates is denoted by \( \text{Upd}(\mathcal{R}) \).

**Definition 4.11 (CRA [Bol11])**. A class register automaton (CRA) over \( A \) and \( D \) is a tuple \( \mathcal{A} = (S, \mathcal{R}, \Delta, \iota, F, G) \) where

- \( S \) is the nonempty finite set of states,
- \( \mathcal{R} \) is the nonempty finite set of registers,
- \( \iota \in S \) is the initial state,
- \( F \subseteq S \) is the set of local final states,
- \( G \subseteq S \) is the set of global final states, and
- \( \Delta \subseteq 2^{S} \times S \times \text{Guard}(\mathcal{R}) \times A \times \text{Upd}(\mathcal{R}) \times S \) is the finite set of transitions. \( \Box \)

We call \( \mathcal{A} \) non-guessing if every update is a partial function \( \nu : \mathcal{R} \to \mathcal{R}^{-} \cup \mathcal{R}^{\ominus} \cup \{0\} \). This means that each “guess” is restricted to the current data value.

Again, we define the semantics of \( \mathcal{A} \) via the notion of a configuration. Here, a configuration is a pair \( \gamma = (s, \tau) \) where \( s \in S \) is the current state and \( \tau : \mathcal{R} \to D \) is a partial mapping denoting the current register assignment. We do not need to keep track of the set of used data values, since a run will be defined as a mapping of word positions to configurations. As before, configuration \( \gamma \) is final if \( s \in F \). It is initial if \( s = \iota \) and \( \tau \) is undefined everywhere.

Let \( (a_1, d_1) \ldots (a_n, d_n) \in (A \times D)^* \) be a data word. For \( i \in [n] \) and \( \ell \in \mathbb{N} \), let \( \mathcal{D}_{\ell}(i) \) denote the set \( \{d_j \mid j \in [n] \} \) has distance at most \( \ell \) from \( i \) in the (undirected) graph \( ([n], \triangleleft_{\mathcal{R}} \cup \triangleleft_{\mathcal{R})} \cup \triangleleft_{\mathcal{R}}^{-} \cup \triangleleft_{\mathcal{R}}^{\ominus}) \) of those data values that are in the \( \ell \)-neighborhood of \( i \) in \( w \). A run of \( \mathcal{A} \) is a sequence of configurations \( \gamma_0 \gamma_1 \ldots \gamma_n \) satisfying some properties. First, \( \gamma_0 \) has to be initial. Suppose \( \gamma_i = (s_i, \tau_i) \). Then, we require that, for all \( i \in [n] \), there are a guard \( \Phi \in \text{Guard}(\mathcal{R}) \), an update \( \nu \in \text{Upd}(\mathcal{R}) \), and \( T \subseteq S \) such that

- \( (T, s_{i-1}, \Phi, a_i, \nu, s_i) \in \Delta \),
- if there is no \( j \in [n] \) such that \( j \triangleleft_{\mathcal{R}} i \), then we have both \( T = \emptyset \) and \( \nu(r) \notin \mathcal{R}^{\ominus} \) for all \( r \in \mathcal{R} \),
- \( j \triangleleft_{\mathcal{R}} i \) implies \( s_j \in T \), for all \( j \in [n] \) (note that there is at most one such \( j \)),
- guard \( \Phi \) is evaluated to true on the basis of its atoms: \( \theta_1 = \theta_2 \) is true iff \( \text{val}_i(\theta_1) = \text{val}_i(\theta_2) \in D \) where
  - \( \text{val}_i(c) = d_i \),
  - \( \text{val}_i(r^-) = \tau_{i-1}(r) \) (i.e., if \( \tau_{i-1}(r) \) is undefined, then so is \( \text{val}_i(r^-) \)), and
For example, \( \text{val}_1(r) = \begin{cases} \tau_j(r) & \text{if } j \not<_{\oplus 1} i \\ \text{undef} & \text{if there is no } j \text{ such that } j \not<_{\oplus 1} i \end{cases} \) 

(note that, in the latter two cases, the valuation may be undefined and, therefore, not be in \( D \) so that \( \theta_1 = \theta_2 \) is not satisfied),

- for all \( r \in \mathcal{R} \), we have

\[
\tau_i(r) = \begin{cases} 
\tau_{i-1}(\hat{r}) & \text{if } \nu(r) = \hat{r}^- \\
\tau_i(r) = \tau_j(\check{r}) & \text{if } j \not<_{\oplus 1} i \text{ and } \nu(r) = \hat{r}^\oplus \\
\tau_i(r) \in \mathcal{D}_\ell(i) & \text{if } \nu(r) = \ell \in \mathbb{N} \\
\tau_i(r) \text{ undefined} & \text{if } \nu(r) \text{ undefined.}
\end{cases}
\]

The run is accepting if \( s_n \in G \) and, for all \( i \in [n] \) such that \( i \) is \( <_{\oplus 1} \)-maximal, \( s_i \in F \). As usual, the set of data words that admit an accepting run of \( \mathcal{A} \) is denoted by \( L(\mathcal{A}) \). Clearly, \( L(\mathcal{A}) \) is a data language.

Example 4.12. Recall from Example 4.4 on page 62 that there is no FRA recognizing the language defined as the intersection of the following data languages over \( A = \{\text{req, ack}\} \) and \( D = \mathbb{N} \):

- \( L_3 = \) “every acknowledgment is preceded by a request”
- \( L_4 = \) “requests are acknowledged in the order they are received”
- \( L_5 = \) “every process makes at most one request”
- \( L_6 = \) “all requests take place before all acknowledgments”

Formally, the language in question is

\[ L = \{ w \in (A \times D)^* \mid w \approx (\text{req}, 1) \ldots (\text{req}, n)(\text{ack}, 1) \ldots (\text{ack}, n) \text{ for some } n \in \mathbb{N} \} \]

For example, \((\text{req}, 8)(\text{req}, 5)(\text{req}, 3)(\text{req}, 4)(\text{ack}, 8)(\text{ack}, 5)(\text{ack}, 3)(\text{ack}, 4) \in L \).

The (non-guessing) CRA \( \mathcal{A}_{1,6} \) from Figure 4.6 recognizes \( L \), i.e., \( L(\mathcal{A}_{1,6}) = L \). We have \( F = \{ s_2 \} \) and \( G = \{ s_0, s_2 \} \), i.e., \( s_2 \) is the only local final state (indicated by the outgoing arrow), and \( s_0 \) and \( s_2 \) are the global final states. The updates are written in the form of an assignment. In a run, state \( s_1 \) is taken after executing requests, while \( s_2 \) is assigned to acknowledgment positions. Moreover, the CRA has two registers, \( r_1 \) and \( r_2 \). Suppose it reads \((\text{req}, d_1) \ldots (\text{req}, d_n)(\text{ack}, d_{n+1}) \ldots (\text{ack}, d_{2n}) \).

In the request phase, the current data values are respectively stored in register \( r_1 \). Moreover, \( r_2 \) takes the data values at the \( <_{\oplus 1} \)-predecessor position. When entering the second phase (going from \( s_1 \) to \( s_2 \)), the guard \( r_2^\oplus \not= r_2^\ominus \) will ensure that position \( n + 1 \) is the unique \( <_{\oplus 1} \)-predecessor position where \( r_2 \) is undefined. Thus, \( d_1 = d_{n+1} \). Then, the loop in \( s_2 \) makes sure that every position \( n + i \) with \( i \geq 2 \) matches the request position where the contents of \( r_2 \) equals that of \( r_1 \) at position \( n + i - 1 \). Since \( s_2 \) is the only local final state, every request has to be followed by a matching acknowledgment. Altogether, we have \( d_1 \ldots d_n = d_{n+1} \ldots d_{2n} \). \( \diamond \)
Note that some interesting automata models over data words can be identified as a special case of CRAs.

- A class memory automaton (CMA) \([BS10]\) is a CRA \(A = (S, \mathcal{R}, \Delta, \iota, F, G)\) such that, for all transitions \((T, s, \Phi, a, \nu, s') \in \Delta\), we have that \(\Phi = \text{true}\) and \(\nu\) is undefined everywhere. In other words, registers are not used at all, while states of \(\preceq_{\preceq_{\neq_1}}\) can still be accessed.

- A quasi-FRA is a non-guessing CRA \(A = (S, \mathcal{R}, \Delta, \iota, F, G)\) such that \(F = S\) and, for all transitions \((T, s, \Phi, a, \nu, s') \in \Delta,\)
  
  1. \(T = S\) or \(T = \emptyset\),
  2. if \(T = S\), then \((\emptyset, s, \Phi, a, \nu, s') \in \Delta\),
  3. for all atoms \(\theta_1 = \theta_2\) occurring in \(\Phi\), we have \(\theta_1, \theta_2 \in \{\varepsilon\} \cup \mathcal{R}^-\), and
  4. for all registers \(r \in \mathcal{R}, \nu(r)\) is undefined or contained in \(\mathcal{R}^- \cup \{0\}\).

Condition 1. implies that the concrete state at a \(\preceq_{\preceq_{\neq_1}}\)-predecessor cannot be accessed anymore. However, one may specify that a data value is globally fresh \((T = \emptyset)\). By Condition 2., one cannot say that a data value is “non-fresh”. Conditions 3. and 4. prevent the automaton from accessing registers at \(\preceq_{\preceq_{\neq_1}}\)-predecessors.

- A quasi-RA is a quasi-FRA \(A = (S, \mathcal{R}, \Delta, \iota, F, G)\) such that, for all transitions \((T, s, \Phi, a, \nu, s') \in \Delta,\) both \((\emptyset, s, \Phi, a, \nu, s')\) and \((S, s, \Phi, a, \nu, s')\) are contained in \(\Delta\). In other words, it cannot be enforced anymore that a data value is globally fresh.

Actually, many variants of register automata have been defined in the literature. In the configuration of an RA, a register assignment is always an injective mapping, which is relaxed in quasi-RAs. As a consequence, the latter use guards to compare data values. Quasi-RAs are similar to the model presented in \([CHJ^{+}11]\) in the context of automata learning. That model also uses guards to compare register contents. Transitions in the model from \([Seg06]\), on the other hand, are guarded by a subset \(R\) of registers. The meaning is that \(R\) is exactly the set of registers that contain the data value that is currently read. Though syntactically different, those models are expressively equivalent, and the same holds for quasi-FRAs and
FRAs. Note that, however, the complexity of decision problems do not carry over. For example, nonemptiness for quasi-RAs is PSPACE-complete, while it is NP-complete for RAs (cf. Section 4.5). The expressive power of the various automata models over data words (and logics as presented in the next section) is illustrated in Figure 4.7. In particular, it was shown in [BS10] that CMAs are strictly more expressive than RAs. The fact that CMAs are strictly weaker than CRAs is witnessed by the language from Example 4.12. In fact, a simple pumping argument shows that this language cannot be recognized by a CMA.

In the next section, we will establish that the expressive power of CRAs is between EMSO and MSO logic (with a predicate for comparing data values for equality), solving the realizability problem for EMSO logic. Note that, in the article [Bol11], we introduced CRAs in a more general framework, allowing some flexibility in the choice of the signature (rather than only $\langle \cdot \rangle_1$ and $\langle \cdot \rangle_2$). This allowed us to treat dynamic communicating automata, which will be introduced in Chapter 5, as a special case.

4.4 Automata vs. Logic over Data Words

We will now consider MSO logic over data words. It is quite natural to extend classical MSO logic over words with a binary predicate $x \sim y$ to express that the data values at positions $x$ and $y$ coincide. Alternatively, one may include a predicate $x \triangleleft_1 y$ relating two successive positions with the same data values. In MSO logic, one will be expressible in terms of the other.

Definition 4.13. The set $d\text{MSO}(A, D)$ of data MSO formulas (over $A$ and $D$) is built from the following grammar:

$$
\varphi ::= a(x) \mid x \triangleleft_{+1} y \mid x \triangleleft_1 y \mid x = y \mid x \in X \mid \\
\neg \varphi \mid \varphi_1 \lor \varphi_2 \mid \exists x. \varphi \mid \exists X. \varphi
$$

where $a \in A$, $x, y$ are first-order variables, and $X$ is a second-order variable.

The semantics of $d\text{MSO}(A, D)$ is as expected. Let us only pick some cases: for a data word $w = (a_1, d_1) \ldots (a_n, d_n) \in (A \times D)^*$ and positions $i, j \in [n]$, we write

- $w, i \models a(x)$ if $a_i = a$,
- $w, i, j \models x \triangleleft_{+1} y$ if $i + 1 = j$, and
- $w, i, j \models x \triangleleft_1 y$ if $d_i = d_j$ and, for all $k \in \{i+1, \ldots, j-1\}$, we have $d_i \neq d_k$.

For a sentence $\varphi \in d\text{MSO}(A, D)$, we let $L(\varphi)$ denote the set of data words $w$ such that $w \models \varphi$. Like for automata, $L(\varphi)$ is a data language.

We will often write $d\text{MSO}$ instead of $d\text{MSO}(A, D)$, since we fixed $A$ and $D$.

The predicate $x \sim y$ is defined as an abbreviation for $(x \triangleleft_{+1}^* y) \lor (y \triangleleft_{+1}^* x)$, where, in turn, $\triangleleft_{+1}$ is the dMSO-definable reflexive transitive closure of $\triangleleft_1$. 

71
The fragments \( \text{dFO} \) and \( \text{dEMSO} \) are defined as expected. Moreover, \( \text{dEMSO}_2 \) will denote the fragment of \( \text{dEMSO} \) that makes use of only two first-order variables. Again, we may explicitly mention all the binary predicates (apart from \( x = y \) and \( x \in X \)) that we allow. For example, the logic \( \text{dEMSO}(\triangleleft_{+1}, \triangleleft_{\ominus}, \sim) \) includes the binary predicate \( \sim \), which is per se not expressible in \( \text{dEMSO} = \text{dEMSO}(\triangleleft_{+1}, \triangleleft_{\ominus}) \).

**Example 4.14.** We give a \( \text{dFO} \) formula \( \varphi \) over \( A = \{\text{req}, \text{ack}\} \) and \( D = \mathbb{N} \) such that \( L(\varphi) \) equals \( \{w \in (A \times D)^* \mid w \approx (\text{req}, 1) \ldots (\text{req}, n)(\text{ack}, 1) \ldots (\text{ack}, n) \text{ for some } n \in \mathbb{N}\} \) from Example 4.12. Recall that there is a CRA recognizing \( L(\varphi) \), but no CMA and no FRA can accept \( L(\varphi) \). We define \( \varphi \) as the conjunction \( \varphi_1 \land \ldots \land \varphi_4 \) where

\[
\varphi_1 = \forall x. \exists y. (\text{req}(x) \rightarrow \text{ack}(y) \land x \triangleleft_{\ominus} y)
\]

says that every request is acknowledged (before the same process sends another request),

\[
\varphi_2 = \forall x. \exists y. (\text{ack}(x) \rightarrow \text{req}(y) \land y \triangleleft_{\ominus} x)
\]

says that every acknowledgment has a corresponding request,

\[
\varphi_3 = \neg \exists x. \exists y. (\text{ack}(x) \land \text{req}(y) \land x \triangleleft_{+1} y)
\]

says that all requests precede all acknowledgments, and

\[
\varphi_4 = \forall x. \forall y. \left( \text{req}(x) \land \text{req}(y) \land x \triangleleft_{+1} y \implies \exists x', \exists y'. (\text{ack}(x') \land \text{ack}(y') \land x' \triangleleft_{\ominus} x' \land y' \triangleleft_{\ominus} y')(\text{ack}(x) \land \text{ack}(y) \land x \triangleleft_{+1} y)\right)
\]

guarantees that two (successive) requests are acknowledged in the order they were received.

There have been some logical characterizations of automata over data words. A prominent one considers \textit{data automata} \([\text{BDM}^+11]\), which were shown to be expressively equivalent to CMAs in \([\text{BS}10]\).

**Theorem 4.15 (\([\text{BDM}^+11]\)).** CMAs are expressively equivalent to the data logic \( \text{dEMSO}_2(\triangleleft_{+1}, \triangleleft_{\ominus}, \sim, <) \).

A logical characterization of RAs has been obtained in \([\text{CLP}11]\) via a semantical restriction:

**Definition 4.16 (\([\text{CLP}11]\)).** The set \( \text{rgMSO} \) of rigid guarded MSO formulas contains any dMSO sentence of the form \( \exists X_1 \ldots X_m. \varphi \) where \( \varphi \) is built from the grammar

\[
\varphi ::= \alpha(x) \mid \alpha(x, y, X_1, \ldots, X_m) \land x \sim y \mid x = y \mid x \in X \mid 
\neg \varphi \mid \varphi_1 \lor \varphi_2 \mid \exists x. \varphi \mid \exists X. \varphi
\]

with \( \alpha \) being generated by the same grammar (without quantification over the variables \( X_1, \ldots, X_m \)) such that \( \min(x, y) \) is uniquely determined from both \( \max(x, y) \) and \( X_1, \ldots, X_n \) (in the expected manner).
Theorem 4.17 ([CLP11]). RAs are expressively equivalent to the logic rgMSO.

Further automata models and logics have been considered in the literature. For example, a logical characterization has been established for weak data automata, which are a semantical restriction of data automata and expressively equivalent to the logic dEMSO$_2(<_{1+}, \sim)$ [KST12]. It seems that a logical characterization for FRAs has not been considered so far.

Now, we turn to SAs and CRAs. We first identify a fragment of dMSO$(A, D)$ that is expressively equivalent to SAs.

**Definition 4.18.** A session MSO (sMSO) formula is a dMSO sentence of the form

$$\exists X_1 \ldots \exists X_m. (\alpha \land \forall x. \forall y.(x \sim y \leftrightarrow \beta))$$

where $\alpha$ and $\beta$ are from dMSO($<_{1+}$), i.e., they are classical MSO formulas containing neither $\lhd_{\oplus 1}$ nor $\sim$.

**Example 4.19.** The sentence

$$\varphi_1 = \forall x. \forall y.(x \sim y \leftrightarrow x = y)$$

is an sMSO formula. Its semantics $L(\varphi_1)$ is the set of data words in which every data value occurs at most once. Moreover,

$$\varphi_2 = \forall x. \forall y.(x \sim y \leftrightarrow \text{true})$$

is an sMSO formula, and $L(\varphi_2)$ is the set of data words where all data values coincide. As a last example, let

$$\varphi_3 = \exists X. \forall x. \forall y.(x \sim y \leftrightarrow (\exists z.z \in X \land (x < z \leq y \lor y < z \leq x))).$$

Then, $L(\varphi_3)$ is the set of 1-bounded data words. Intuitively, the variable $X$ represents the set of word positions where a fresh data value is introduced.

**Theorem 4.20 ([BHLM14]).** Let $L \subseteq (A \times D)^*$. The following are equivalent:

- There is an SA $A$ such that $L(A) = L$.
- There is a sentence $\varphi \in \text{sMSO}$ such that $L(\varphi) = L$.

**Proof (sketch).** The construction of a formula from an SA follows the classical scheme: by means of existential second-order quantifiers, we guess an assignment of transitions to word positions. In the first-order part $\alpha$, it is then verified if this assignment corresponds to an accepting run. Moreover, formula $\beta$ checks if data
equality in the data word at hand corresponds to the symbolic word produced by the automaton.

For the other direction, the crucial observation is that every sMSO formula defines a $k$-bounded data language where we can choose $k$ to be the number of states of a deterministic finite (word) automaton that recognizes the models of $\beta$ (where free variables are handled by an extension of the alphabet $A$).

Thanks to the closure of SA under complementation for bounded data words, we can establish another logical characterization in terms of full dMSO logic. However, we have to restrict in advance to the class of $k$-bounded data words, for some $k$. Recall that this is in the spirit of the logical characterizations in Chapters 2 and 3 where, similarly, one has to constrain the domain of nested words and MSCs to obtain logical characterizations in terms of MSO logic.

**Theorem 4.21.** Let $k \geq 1$ and $L \subseteq \text{DW}_k$. The following statements are (effectively) equivalent:

- There is an SA $A$ such that $L(A) = L$.
- There is a sentence $\varphi \in \text{dMSO}$ such that $L(\varphi) = L$.

This implies that, considered over $k$-bounded data words, the formalisms dMSO logic, CRA, FRA, and SA all have the same expressive power, while RAs are strictly weaker. Next, we turn to the relation between logic and CRAs. Recall that, basic decision problems being undecidable, we were interested in the realizability problem. It is solved positively for dEMSO specifications:

**Theorem 4.22 ([Bol11]).** For every sentence $\varphi \in \text{dEMSO}$, there is a CRA $A$ such that $L(A) = L(\varphi)$.

**Proof (sketch).** We adopt the technique from [BL06] (Theorem 3.13 on page 48). The idea is to construct an automaton that computes, in a data word, the sphere around a given event. The theorem then follows by Hanf’s normal form. Using registers, the technique from [BL06] can indeed be adapted to deal with data words. Note that the guessing mode is necessary to be able to anticipate neighborhoods before verifying them.

Moreover, one can show that, despite their expressive power, CRAs recognize only data languages that are definable in dMSO:
Theorem 4.23 ([Bol11]). For every CRA $A$, there is a sentence $\varphi \in \text{dMSO}$ such that $L(\varphi) = L(A)$.

Figure 4.7 summarizes and compares the expressive power of the various formalisms considered in this chapter. Here, $\rightarrow$ means “strictly included” and $\rightarrow$ means “included” (strict inclusion being an open problem).

\begin{center}
\begin{tikzpicture}
  \node (dMSO) {dMSO} ;
  \node (CRA) [below of=dMSO] {CRA} ;
  \node (dEMSO) [below of=CRA] {dEMSO($\triangleq_{+1, \triangleq_{\oplus 1}}$)} ;
  \node (CMA) [below of=dEMSO] {CMA = dEMSO$_2(\triangleq_{+1, \triangleq_{\oplus 1}, \sim, <})$} ;
  \node (FRA) [left of=CMA] {FRA} ;
  \node (SA) [below of=FRA] {SA = sMSO} ;
  \node (RA) [below of=FRA] {RA = r$_{\text{gMSO}}$} ;

  \path (dMSO) -- (CRA) ;
  \path (CRA) -- (dEMSO) ;
  \path (dEMSO) -- (CMA) ;
  \path (CMA) -- (FRA) ;
  \path (FRA) -- (SA) ;
  \path (FRA) -- (RA) ;
\end{tikzpicture}
\end{center}

Figure 4.7: A hierarchy of automata and logics over data words

4.5 Decision Problems

Let us turn to decision problems, namely nonemptiness, inclusion, and model checking. For a class $\mathcal{C} \in \{\text{FRA,RA,SA,CRA,CMA}\}$ of automata over data words and a logic fragment $\mathcal{L}$, we consider the following problems$^2$:

\begin{align*}
\textbf{Problem 4.24. Nonemptiness($\mathcal{C}$):} & \quad \text{Input:} \text{Automaton } A \text{ from } \mathcal{C} \\
\text{Question:} & \quad L(A) \neq \emptyset \ ?
\end{align*}

\begin{align*}
\textbf{Problem 4.25. Inclusion($\mathcal{C}$):} & \quad \text{Input:} \text{Automata } A \text{ and } B \text{ from } \mathcal{C} \\
\text{Question:} & \quad L(A) \subseteq L(B) \ ?
\end{align*}

\begin{align*}
\textbf{Problem 4.26. ModelChecking($\mathcal{C}, \mathcal{L}$):} & \quad \text{Input:} \text{Automaton } A \text{ from } \mathcal{C} ; \varphi \in \mathcal{L} \\
\text{Question:} & \quad L(A) \subseteq L(\varphi) \ ?
\end{align*}

We start with results that had been known for FRAs and RAs:

\begin{align*}
\textbf{Theorem 4.27 ([KF94, Tze11]).} & \quad \text{The problem Nonemptiness(FRA) and the problem Nonemptiness(RA) are both NP-complete.}
\end{align*}

$^2$Recall that $A$ and $D$ are fixed. For the forthcoming results, it actually does not make a difference if $A$ is part of the input or any fixed finite alphabet.
The NP upper bound directly carries over to SAs. The nonemptiness problem for SAs is actually NP-complete, too. The lower bound is shown by a reduction from 3-CNF-SAT (cf. [BCH+13]).

**Theorem 4.28.** \textsc{Nonemptiness(SA)} is NP-complete.

Note that the problem becomes considerably harder when we move up in the hierarchy of automata for data words.

**Theorem 4.29 ([BDM+11]).** \textsc{Nonemptiness(CMA)} is decidable, and as hard as Petri-net reachability.

As a corollary from Theorem 4.29 and the effectiveness of Theorem 4.15, one obtains that satisfiability of $d$EMSO$_2(\triangle_1, \triangle_1, \sim, <)$ is decidable. It is easy to see that dropping the two-variable restriction renders the satisfiability problem undecidable. By Theorem 4.22, this means that nonemptiness for CRAs is undecidable. Thus, the gain in expressive power that CRAs provide comes at the expense of an undecidable nonemptiness problem:

**Theorem 4.30.** \textsc{Nonemptiness(CRA)} is undecidable.

Next, we consider the inclusion problem. Unlike the nonemptiness problem, it is already undecidable for the simple model of RAs:

**Theorem 4.31 ([KF94, Tze11]).** \textsc{Inclusion(RA)} is undecidable.

Therefore, the inclusion problem is undecidable for any other stronger automata model such as FRAs and CMAs. However, we can establish decidability for SAs, which is a corollary from Theorem 4.8:

**Theorem 4.32 ([BHL14]).** \textsc{Inclusion(SA)} is decidable.

The situation in model checking is similar to that of inclusion when we choose the full dMSO logic as a specification. It is undecidable for RAs, which follows from undecidability of the satisfiability problem for dMSO.

**Theorem 4.33.** \textsc{ModelChecking(RA,dMSO)} is undecidable.

To obtain decidability, one therefore has to restrict the logic. From Theorems 4.29 and 4.15, it follows that the model-checking problem of CMAs against formulas from the logic $d$FO$_2(\triangle_1, \triangle_1, \sim, <)$ is decidable.
Theorem 4.34 ([BDM⁺11]). ModelChecking(CMA,dFO₂(\texttt{\textless}₁⁺₁, \texttt{\textless}₁₁, \texttt{\textless};,<)) is decidable.

For the full dMSO logic, model checking is decidable for SAs, which follows from Theorems 4.21 and 4.32:

Theorem 4.35 ([BCGNK12]). ModelChecking(SA,dMSO) is decidable.

Note that the model-checking problem was actually solved in [BCGNK12] for a more powerful model than SAs, including stacks.

4.6 Perspectives

Though we saw already a certain number of different automata models running on data words, our study is far from being exhaustive. There are actually many more interesting models such as alternating automata [DL09] or walking automata [MMP13]. The latter, however, are rather “recognizers” than “generators” of behaviors so that they cannot be seen as system models. In this chapter, we introduced the model of CRAs that captures, in a unifying framework, many of those models from the literature that can actually be seen as system models. This, however, comes at the price of undecidability of basic verification questions. Still, it would be interesting to identify subclasses of CRAs that extend other existing models but allow one to model realistic protocols such as leader-election algorithms. To this end, one may resort to symbolic techniques or to well-quasi orderings as a technique to establish decidability for infinite-state systems. The latter have actually been successfully applied in the realm of data words [Fig10].
In this chapter, we consider dynamic message-passing systems. In a sense, it combines Chapters 3 and 4, which considered static message-passing systems and, respectively, dynamic sequential systems. Processes will communicate, via message passing, in a dynamic environment, which allows them to create new processes during runtime. This framework is inspired by concurrent programming languages such as Erlang. More precisely, each process has a unique process identifier (pid). It also disposes of a finite number of registers, in which it can store pids of other processes. Now, a process can communicate with all processes that it knows, i.e., whose pids are present in its registers. The knowledge topology is subject to changes, though. When sending a message, a process may include, in the message, processes that it knows. In turn, the receiving process may overwrite some of its registers using the pids attached to the message. Moreover, a process can spawn a new process, which obtains a unique, fresh pid (cf. fresh-register automata from Chapter 4). The latter is henceforth stored in one of the registers of the spawning process, which can then propagate the new pid to other processes.

In the following, we first extend the notion of message sequence charts (MSCs, cf. Chapter 3) towards a dynamic version, and then introduce dynamic communicating automata (DCAs) as an extension of the static (fixed-topology) version from Chapter 3. While communicating automata accept MSCs, DCAs will accept dynamic MSCs. On the specification side, we introduce a conservative extension of the session automata from the previous chapter to handle dynamic MSCs. These register message sequence graphs (rMSGs) allow the specifier a global view of the system. In this chapter, we study mainly the realizability question. Given an rMSG, is there a DCA recognizing the same set of behaviors? Since our dynamic framework properly extends the static setting, these questions are undecidable. For this reason, we will consider executability, a non-trivial sufficient criterion for
implementability (essentially, realizability modulo message refinement), which we then show to be decidable. Essentially, an rMSG is executable if, in each scenario, a process \( p \) that sends a message to another process, say, with pid \( q \), “knows” \( q \). Here, knowledge refers to the possibility of passing, during the execution, the process identity of \( q \) to \( p \). In a second step, we discover a restriction of rMSGs, for which executability and implementability coincide.

A first step towards MSCs over an evolving set of processes was made in [LMM02], where MSO model checking is shown decidable for fork-and-join MSC grammars. Our rMSGs are similar to these grammars, but take into account pids as message contents and distinguish messages and process creation. Moreover, (implementable) subclasses can be identified more easily. Nevertheless, several of our results apply to the formalism from [LMM02] once the latter is adjusted to our setting. In [BGP08], an MSC semantics was given for the \( \pi \)-calculus, but the problems studied there are very different from ours and do not distinguish between a specification and an implementation.

5.1 Dynamic Message Sequence Charts

We modify MSCs from Chapter 3 (Definition 3.3 on page 42) to deal with process creation. Let \( P \) be a nonempty (finite or infinite) set of processes. Later, \( P \) may be instantiated with a finite set of registers or the infinite set \( P \equiv \mathbb{N} \) of pids. A process can perform send and receive actions, but it is also allowed to spawn new processes. Moreover, it may attach processes/pids to a message. Thus, we fix a finite ranked alphabet\(^1\) \( A \). For any set \( B \), let \( A(B) \equiv \{ a(b_1, \ldots, b_n) \mid a \in A, n = \text{arity}(a), \text{and } b_1, \ldots, b_n \in B \} \). Then, the set of messages is given as \( A(P) \).

The following definition is similar to Definition 3.3. However, since communication topologies are henceforth dynamic, an MSC does not include a topology but certain spawn events. We will first introduce partial MSCs, which feature an arbitrary number of “existing” processes (i.e., those that are not spawned). Those processes are analogous to free variables in logic formulas.

**Definition 5.1 (partial MSC).** A partial MSC over \( A \) and \( P \) is a tuple \( M = (E, \triangleleft, \pi, \mu) \) where

- \( E \) is a nonempty finite set of events,
- \( \triangleleft = \triangleleft_{\text{new}} \cup \triangleleft_{\text{proc}} \cup \triangleleft_{\text{msg}} \) is the acyclic edge relation,
- \( \pi : E \rightarrow P \) maps each event to a process—for \( p \in P \), we let \( E_p \equiv \{ e \in E \mid \pi(e) = p \} \), and
- \( \mu : \triangleleft_{\text{msg}} \rightarrow A(P) \) maps each message edge to a message.

We require that the following hold:

\(^1\)A ranked alphabet is an alphabet \( A \) that comes with a mapping \( \text{arity} : A \rightarrow \mathbb{N} \).
1. \( \prec_{\text{proc}} \) is a union \( \bigcup_{p \in P} \prec_p \) where each \( \prec_p \subseteq E_p \times E_p \) is the direct-successor relation of some total order on \( E_p \).

2. there is a partition \( E = E_1 \uplus E_2 \uplus E_{\text{spawn}} \uplus E_{\text{start}} \) such that
   - \( \prec_{\text{msg}} \) induces a bijection between \( E_1 \) and \( E_2 \),
   - \( \prec_{\text{new}} \) induces a bijection between \( E_{\text{spawn}} \) and \( E_{\text{start}} \),
   - \( \prec_{\text{msg}} \) and \( \prec_{\text{new}} \) are subsets of \( \bigcup_{p \neq q} (E_p \times E_q) \),
   - there is no event \( e \in E_{\text{start}} \) that has a \( \prec_{\text{proc}} \)-predecessor, and
   - for all \( (e, f), (e', f') \in \prec_{\text{msg}} \) such that \( \pi(e) = \pi(e') \) and \( \pi(f) = \pi(f') \), we have \( e \prec_{\text{proc}} e' \) iff \( f \prec_{\text{proc}} f' \) (FIFO).

The set of partial MSCs over \( A \) and \( P \) is denoted by \( \text{pMSC}(A, P) \).

Note that the partition \( E = E_1 \uplus E_2 \uplus E_{\text{spawn}} \uplus E_{\text{start}} \) is uniquely determined by the fact that \( \prec \) is a partition \( \prec_{\text{new}} \uplus \prec_{\text{proc}} \uplus \prec_{\text{msg}} \). We let \( \text{Pids}(M) \overset{\text{def}}{=} \{ p \in P \mid E_p \neq \emptyset \} \). By \( \text{Free}(M) \overset{\text{def}}{=} \{ p \in \text{Pids}(M) \mid E_{\text{start}} \cap E_p = \emptyset \} \), we denote the set of free processes of \( M \). Intuitively, free processes of a partial MSC \( M \) have not been initiated in \( M \). Moreover, \( \text{Bnd}(M) \overset{\text{def}}{=} \text{Pids}(M) \setminus \text{Free}(M) \) denotes the set of bound processes.

For every \( p \in \text{Pids}(M) \), there are a unique minimal and a unique maximal event in the total order \( (E_p, \prec_p^{\text{new}}) \), which we denote by \( \min_p(M) \) and \( \max_p(M) \), respectively.

By \( \text{MsgPar}(M) \), we denote the set of processes \( p \in P \) that occur as a parameter in some message, i.e., such that there is \( a(p_1, \ldots, p_n) \in \mu(\prec_{\text{msg}}) \) with \( p \in \{ p_1, \ldots, p_n \} \).

Now, we can define (complete rather than partial) dynamic MSCs:

**Definition 5.2 (MSC).** A dynamic MSC (or, simply, MSC) is a partial MSC \( M = (E, \prec, \pi, \mu) \) such that \( \text{Free}(M) \) is a singleton set.

Note that the partial order \( (E, \prec^*) \) associated with \( M \) has a unique minimal event, which we denote by \( \text{init}(M) \). The set of MSCs over \( A \) and \( P \) is denoted by \( \text{dMSC}(A, P) \).

The notion of free processes in a partial MSC is important in the context of automata, where we start with a fixed number of processes. Actually, we suppose that, at the beginning of an execution, we start with one single free process, though we could extend the framework to an arbitrary fixed number of initial processes.

**Example 5.3.** An MSC \( M \) over \( A = \{ a, b, c \} \) and \( P \) (recall that \( P = \mathbb{N} \)) is shown in Figure 5.1. Here, \( \text{arity}(a) = 1 \) and \( \text{arity}(b) = \text{arity}(c) = 0 \). We have \( \text{Free}(M) = \{ 8 \} \) and \( \text{Bnd}(M) = \{ 3, 4, 5 \} \).

As usual, we do not distinguish isomorphic structures. Moreover, none of the formalisms that we are going to introduce next will be able to distinguish between MSCs that differ only in pids. Given a partial MSC \( M \), we denote by \( [M] \) the set of partial MSCs that can be obtained from \( M \) by renaming of pids (in the expected manner—we omit the formal definition). This is extended to sets \( L \), and we let \( [L] \).
denote the union of sets $[M]$ with $M$ ranging over $L$. When $L = [L]$ holds, then we say that $L$ is closed. Note that a closed set of MSCs is in the spirit of a data language (cf. Chapter 4), which is a “closed” set of data words.

### 5.2 Dynamic Communicating Automata

Next, we present our model of an implementation of a dynamic message-passing system. As mentioned before, we essentially deal with communicating automata, where processes execute send and receive actions (cf. Definition 3.9 on page 46). However, processes can henceforth spawn new processes and exchange pids along with messages.

#### 5.2.1 The Model

Let us give the formal definition of dynamic communicating automata.

**Definition 5.4 (DCA [BCH+13]).** A dynamic communicating automaton (DCA) over the ranked alphabet $A$ (and $\mathbb{P}$) is a tuple $A = (S, R, i, F, \Delta)$ where

- $S$ is the nonempty finite set of states
- $i \in S$ is the initial state,
- $F \subseteq S$ is the set of final states,
- $R$ is the nonempty finite set of registers, and
- $\Delta$ is the set of transitions.

A transition is of the form $(s, \alpha, s')$ where $s, s' \in S$, and $\alpha$ is an action, possibly

- $r := \text{spawn}(t, r')$ where $r, r' \in R$ and $t \in S$,
- $r!a(r_1, \ldots, r_n)$ where $r \in R$ and $a(r_1, \ldots, r_n) \in A(R \cup \{\text{self}\})$, or
- $r?a(r_1, \ldots, r_n)$ where $r \in R \cup \{\ast\}$, and $a(r_1, \ldots, r_n) \in A(R \cup \{-\})$ such that $r_i = r_j \in R$ implies $i = j$. 

As usual, we write a transition \((s, \alpha, s') \in \Delta\) as \(s \xrightarrow{\alpha} s'\).

The basic feature of a DCA is the transmission of pids via messages. When a process executes \(r!a(r_1, \ldots, r_n)\), it sends a message to the process whose pid is stored in register \(r\). The message consists of label \(a\) as well as \(n = \text{arity}(a)\) many pids stored in registers \(r_1, \ldots, r_n\) (or the sender’s pid if \(r_i = \text{self}\)). Executing \(r?a(r_1, \ldots, r_n)\), a process receives a message from the process whose pid is stored in \(r\) (selective receive) or, in case \(r = \ast\), from any process (non-selective receive). The message must be of the form \(a(p_1, \ldots, p_n) \in A(P)\). In the resulting configuration, the receiving process updates its local registers \(r_1, \ldots, r_n\) to \(p_1, \ldots, p_n\), respectively, unless \(r_i = -\), in which case \(p_i\) is neglected. Finally, a process executing action \(r := \text{spawn}(t, r')\) spawns a new process, whose fresh pid is henceforth stored in register \(r\). The new process starts in state \(t\). Its registers are a copy of the registers of the spawning process, except for \(r_0\), which is set to the pid of the spawning process.

**Example 5.5.** An example DCA is depicted in Figure 5.2. It realizes a peer-to-peer protocol. Hereby, \(A = \{\text{ack}, \text{com}\}\) with \(\text{arity}(\text{ack}) = 1\) and \(\text{arity}(\text{com}) = 0\). In that protocol, a spawn action rather plays the role of joining a host and making a request. The request is either forwarded to another host, or acknowledged (ack), in which case a connection between the user and the latest host is established so that they can henceforth communicate (com).

The first part of the DCA, comprising the states \(s_0, \ldots, s_3\), is only executed by the very first, i.e., the requesting process. All other processes start in \(t_1\). Going from \(s_0\) to \(s_1\), the initial process, say with pid 1, creates a new process, say with pid 2. Pid 2 is henceforth stored in register \(r\), but this is actually irrelevant, since it will later be overwritten without being used. The new process, on the other hand, starts its execution in \(t_1\). It creates/joins another process, say with pid 3. While process 3 stores the pid 2 in \(r'\), it inherits the contents of \(r\), i.e., its register assignment will be \(\{r \mapsto 1, r' \mapsto 2\}\). In this manner, the pid of the initial process is forwarded to any other process. When, at some point, we have \(n \geq 3\) processes, the register assignment of process \(n\) is \(\{r \mapsto 1, r' \mapsto n - 1\}\) (for \(n = 2\), we have \(\{r \mapsto 1\}\)). In particular, \(r\) holds the pid of the initial process, which had sent the initial request. Now, suppose that process \(n\) is capable of satisfying the request. It will then go from \(t_1\) to \(t_2\) and send an acknowledgment to the process whose pid is stored in \(r\), namely the initial process. Along with \text{ack}, process \(n\) sends its own pid, indicated by \text{self}. Though \(1\) does not know \(n\), it may receive that message, executing the non-selective receive associated with the transition from \(s_1\) to \(s_2\). As a result, register \(r\) of process 1 takes the value \(n\). Now, both processes know each other (their respective pids are stored in register \(r\)) and can exchange messages of type \text{com}, using selective receives.

### 5.2.2 Semantics

Let us describe the semantics of DCAs formally. Like for nested-word automata (Chapter 2) and communicating automata (Chapter 3), there are essentially two
A5.2 ways of defining the semantics: one running directly on MSCs, and one accepting their linearizations. Again, we stick to the former.

Let $\tau : \mathcal{R} \rightarrow \mathbb{P}$ be a partial mapping, assigning identities to registers. For $p \in \mathbb{P}$ and $m = a(r_1, \ldots, r_n) \in A(\mathcal{R} \uplus \{\text{self}\})$ such that $\{r_1, \ldots, r_n\} \subseteq \text{dom}(\tau) \cup \{\text{self}\}$, we let $m[\tau, p]$ denote the message $a(p_1, \ldots, p_n) \in \text{Msg}$ where

$$p_i = \begin{cases} \tau(r_i) & \text{if } r_i \in \mathcal{R} \\ p & \text{if } r_i = \text{self} \end{cases}.$$

Intuitively, $m[\tau, p]$ is the message generated by a process with pid $p$ and register contents $\tau$, when executing an action of the form $r!a(r_1, \ldots, r_n)$.

Now, when a process with register contents $\tau : \mathcal{R} \rightarrow \mathbb{P}$ executes $r?a(r_1, \ldots, r_n)$ (recall that $a(r_1, \ldots, r_n) \in A(\mathcal{R} \uplus \{-\})$ such that $r_i = r_j \in \mathcal{R}$ implies $i = j$) and receives message $a(p_1, \ldots, p_n) \in A(\mathbb{P})$, then its new register contents will be $\tau[r_1, \ldots, r_n \mapsto p_1, \ldots, p_n]$, which is defined to be the partial mapping $\tau' : \mathcal{R} \rightarrow \mathbb{P}$ given by

$$\tau'(r) = \begin{cases} p_i & \text{if } r = r_i \\ \tau(r) & \text{if } r \notin \{r_1, \ldots, r_n\} \end{cases}.$$

Let $\mathcal{A} = (S, \mathcal{R}, \iota, F, \Delta)$ be a DCA and $M = (E, \triangleleft, \pi, \mu) \in \text{dMSC}(\mathcal{A}, \mathbb{P})$. To facilitate the definition of a run of $\mathcal{A}$ on $M$, we assume that $E$ contains some additional
event $e$, without particular labeling, such that $e \prec_{\text{proc}} \text{init}(M)$. A run of $A$ will be a pair $(\sigma, \tau)$, where $\sigma : E \rightarrow S$ and $\tau : E \rightarrow [R \rightarrow P]$. In the following, we will write $\sigma_e$ and $\tau_e$ for $\sigma(e)$ and $\tau(e)$, respectively. Then, $(\sigma, \tau)$ is a run if it respects the following conditions:

1. $\sigma_e = \iota$ and $\tau_e$ is undefined everywhere.

2. For all $e_1, e_2, f \in E$ with $e_1 \prec_{\text{proc}} e_2$ $\prec_{\text{new}} f$, the relation $\Delta$ contains a transition

$$\sigma_{e_1} \xrightarrow{r := \text{spawn}(s,r')} \sigma_{e_2}$$

such that $\sigma_f = s$, $\tau_{e_2} = \tau_{e_1}[r \mapsto \pi(f)]$, and $\tau_f = \tau_{e_1}[r' \mapsto \pi(e_1)]$.

3. For all $e_1, e_2, f \in E$ with $e_1 \prec_{\text{proc}} e_2 \prec_{\text{msg}} f$, we have $\tau_{e_1} = \tau_{e_2}$, and there is a transition

$$\sigma_{e_1} \xrightarrow{r \lambda(a_1, \ldots, a_n)} \sigma_{e_2}$$

such that $\{r, r_1, \ldots, r_n\} \subseteq \text{dom}(\tau_{e_1}) \cup \{\text{self}\}$, $\tau_{e_1}(r) = \pi(f)$, and $\mu(e_2, f) = a(r_1, \ldots, r_n)[\tau_{e_1}, \pi(e_2)]$.

4. For all $e_1, e_2, f \in E$ with $e_1 \prec_{\text{proc}} e_2$ $\prec_{\text{msg}} e_2$, and $\mu(f, e_2) = a(p_1, \ldots, p_n)$, there is a transition

$$\sigma_{e_1} \xrightarrow{r \lambda(a_1, \ldots, a_n)} \sigma_{e_2}$$

such that $r = \ast$ or $\tau_{e_1}(r) = \pi(f)$, and $\tau_{e_2} = \tau_{e_1}[r_1, \ldots, r_n \mapsto p_1, \ldots, p_n]$. The run $(\sigma, \tau)$ is accepting if $\{\sigma(\max_p(M)) \mid p \in Pids(M)\} \subseteq F$. By $L(A)$, we denote the set of MSCs $M$ over $A$ and $P$ such that there is an accepting run of $A$ on $M$. Note that $L(A)$ is closed, i.e., $L(A) = \overline{L(A)}$.

5.2.3 Realizability vs. Implementability

Like for communicating automata, there is a discrepancy between the languages that are recognized by DCAs and languages that one would actually consider to be implementable. More precisely, there are languages $L$ that are not the language of a DCA, but for which there is a DCA implementing them up to some refinement. The refinement allows a DCA to attach more information to a message than the specification provides, for example additional pids. This is formalized as follows. Let $A, B$ be finite ranked alphabets and let $h : B \rightarrow A$. We say that the pair $(B, h)$ is a refinement of $A$ if, for all $b \in B$, $\text{arity}(h(b)) \leq \text{arity}(b)$. We can extend $h$ to a mapping $h : \text{dMSC}(B, P) \rightarrow \text{dMSC}(A, P)$ as follows: for an MSC $M = (E, \prec, \pi, \mu) \in \text{dMSC}(B, P)$, we let $h(M) = (E, \prec, \pi, \mu') \in \text{dMSC}(A, P)$ where $\mu'(e, f) = h(b)(p_1, \ldots, p_{\text{arity}(h(b))})$ whenever $\mu(e, f) = h(b)(p_1, \ldots, p_n)$. The mapping is then further extended to sets of MSCs as expected.
Definition 5.6 (realizability and implementability). A set $L \subseteq dMSC(A, P)$ is called

- **realizable** if $L = L(A)$ for some DCA $A$,

- **implementable** if there are a refinement $(B, h)$ of $A$ and a DCA $A$ over $B$ and $P$ such that $L = h(L(A))$. \hfill \diamondsuit

For both, realizability and implementability, it is necessary that the sender $p$ of a message knows the receiver $q$ at the time of sending, i.e., $q$ should be stored in some register of $p$. Note that this aspect does not arise in simple communicating automata (be it with fixed or parameterized topology).

**Example 5.7.** Consider Figures 5.4 and 5.5. The MSC language $\{M_1\}$ is not implementable, as process 1 does not know 2 when sending message $a$. However, $\{M_2\}$ is implementable (and even realizable), as 2 may know 1: when spawning 2, process 0 can communicate the pid 1 to 2.

![Figure 5.4: Implementability](image1)

The language $\{M_3\}$ is not realizable: as process 0 knows neither 2 nor 3 when it receives the messages, it has to use non-selective receives. But then, the DCA also accepts $M_4$. On the other hand, $\{M_3, M_4\}$ is realizable. However, $\{M_3\}$ and $\{M_4\}$ are both implementable, since we can attach additional information to the message contents $a$ sent from 2 and 3 to 0, such as “1st/2nd message to be received”, respectively. Alternatively, the first message may be refined to contain the pid of the sender of the second message so that, for the latter, 0 can use a selective receive. \hfill \diamondsuit

![Figure 5.5: Realizability vs. Implementability](image2)
5.3 Register MSC Graphs

DCAs serve as a model of an implementation, reflecting low-level primitives—such as spawn, send, and receive—that are provided by concurrent programming languages. As such, DCAs are inherently hard to come up with or to analyze, all the more since basic decision problems are undecidable. In this section, we provide a specification formalism that allows for a high-level view of a system. We define register message sequence graphs (rMSGs), which extend the well-studied message sequence graphs (MSGs) without process creation (see, for example, [GMP03, AEY05, HMK+05, GMSZ06, GKM06]). In a sense, MSGs are to finite automata what rMSGs are to session automata (cf. Chapter 4). Actually, rMSGs combine MSGs with session automata: the transition labels are MSCs whose processes are registers. Free processes are executed by the pids indicated by the corresponding registers, while bound processes obtain a fresh pids, which are henceforth stored in the allotted registers.

Like MSGs, rMSGs compose single behaviors towards bigger ones via concatenation. Let $M = (E, <, \pi, \mu)$ and $M' = (E', <', \pi', \mu')$ be partial MSCs over $A$ and $P$. The concatenation $M \circ M'$ glues identical processes together. Thus, it is defined if $\text{Pids}(M) \cap \text{Bnd}(M') = \emptyset$. In that case, $M \circ M' \overset{\text{def}}{=} (E \oplus E', \vec{\pi} \cup \pi', \mu \cup \mu')$ where

- $\vec{c}_{\text{proc}} = c_{\text{proc}} \cup c'_{\text{proc}} \cup \{(\max_p(M), \min_p(M')) | p \in \text{Pids}(M) \cap \text{Pids}(M')\}$,
- $\vec{c}_{\text{msg}} = c_{\text{msg}} \cup c'_{\text{msg}}$, and
- $\vec{c}_{\text{new}} = c_{\text{new}} \cup c'_{\text{new}}$.

Next, we define the new formalism to describe sets of MSCs. As already mentioned, it is analogous to session automata, but the transitions are labelled with partial MSCs.

**Definition 5.8 (rMSG).** A register message sequence graph (rMSG) over $A$ (and $P$) is a tuple $H = (\mathcal{L}, \mathcal{R}, \mathcal{L}_0, \mathcal{L}_{\text{acc}}, r_0, T)$ where

- $\mathcal{L}$ is the nonempty finite set of locations,
- $\mathcal{L}_0 \subseteq \mathcal{L}$ is the set of initial locations,
- $\mathcal{L}_{\text{acc}} \subseteq \mathcal{L}$ is the set of accepting locations,
- $\mathcal{R}$ is the nonempty finite set of registers with initial register $r_0 \in \mathcal{R}$, and
- $T$ is the finite set of transitions.

A transition is a triple $(\ell, M, \ell') \in \mathcal{L} \times \text{pMSC}(A, \mathcal{R}) \times \mathcal{L}$, usually written $\ell \xrightarrow{M} \ell'$, such that $\text{MsgPar}(M) \cap \text{Bnd}(M) = \emptyset$ (which will guarantee an unambiguous interpretation of message parameters).
Like in (fresh-)register automata, we associate MSCs with an rMSG through the notion of runs, which we will define after some preparation. Henceforth, a register assignment is an injective partial mapping $\tau : \mathcal{R} \rightarrow \mathbb{P}$ from the set of registers to the set of pids.

A configuration of $\mathcal{H}$ is a triple $\gamma = (\ell, \tau, U)$ where $\ell \in \mathcal{L}$ is the current location, $\tau : \mathcal{R} \rightarrow \mathbb{P}$ is a register assignment, and $U \subseteq \mathbb{P}$ is the set of pids that have been used so far. We say that $\gamma$ is initial if $\ell \in \mathcal{L}_0$, $\tau = \{r_0 \mapsto p\}$, and $U = \{p\}$ for some $p \in \mathbb{P}$. It is final if $\ell \in \mathcal{L}_{\text{acc}}$. Via the global transition relation, we can switch from one configuration to another, while reading a partial MSC from the set $p\text{MSC}(A, \mathbb{P})$. More precisely, we let $(\ell, \tau, U) \xrightarrow{M'} (\ell', \tau', U')$ if there are a transition $\ell \xrightarrow{M} \ell'$ and a bijection $\xi : Bnd(M) \rightarrow Bnd(M')$ such that the following hold:

1. $\text{Free}(M) \cup \text{MsgPar}(M) \subseteq \text{dom}(\tau)$,
2. $\tau' = \tau[\xi]$ and $M' = \tau'(M)$, and
3. $U' = U \uplus Bnd(M')$.

Here, $\tau'(M)$ is defined as expected: every register occurrence $r \in \mathcal{R}$ in $M$ is replaced by $\tau'(r) \in \mathbb{P}$. Condition 1. says that free processes can be instantiated. Condition 2. makes sure that registers remain unchanged unless they are overwritten for a new process. Finally, Condition 3. guarantees that bound processes obtain fresh pids.

A run of $\mathcal{H}$ is a sequence $\rho = \gamma_0 \xrightarrow{M_1} \gamma_1 \xrightarrow{M_2} \ldots \xrightarrow{M_n} \gamma_n$ with $n \geq 1$. It generates the MSC $M(\rho) \overset{\text{def}}{=} M_1 \circ \ldots \circ M_n$. Note that $M(\rho)$ is indeed well-defined and has exactly one free pid. We say that $\rho$ is accepting if $\gamma_n$ is final. Finally, $L(\mathcal{H}) \overset{\text{def}}{=} \{M(\rho) \mid \rho \text{ is an accepting run of } \mathcal{H}\}$ is the language of $\mathcal{H}$.

**Example 5.9.** The rMSG $\mathcal{H}_{5.6}$ from Figure 5.6 below gives a high-level view of the peer-to-peer protocol considered in Example 5.5 on page 83. Recall that $A = \{\text{ack}, \text{com}\}$ (acknowledgment, communication) with $\text{arity(ack)} = 1$ and $\text{arity(com)} = 0$. The initial register is $r_0$. A request is forwarded to new processes along with the pid $p$ of the initial process. At some point, a process acknowledges the request, sending its own pid $q$ to the initial process. Processes $p$ and $q$ may then communicate and exchange messages. A generated MSC is depicted in Figure 5.3 on page 84. Note that $L(\mathcal{H}_{5.2})$ is realizable, since we have $L(A_{5.2}) = L(\mathcal{H}_{5.6})$, i.e., $L(\mathcal{H}_{5.6})$ is recognized by a DCA without any message refinement.

---

**5.4 Executability of Register Message Sequence Graphs**

Like for classical MSGs, the language of an rMSG is not necessarily realizable, nor implementable. Even worse, undecidability of realizability and implementability carries over to the dynamic setting. More precisely, we consider the following decision problems:
**Problem 5.10.** rMSG-Realizability:  
**Input:** $A$; rMSG $\mathcal{H}$ over $A$  
**Problem:** Is $L(\mathcal{H})$ realizable?

**Problem 5.11.** rMSG-Implementability:  
**Input:** $A$; rMSG $\mathcal{H}$ over $A$  
**Problem:** Is $L(\mathcal{H})$ implementable?

Unfortunately, both problems are undecidable, which is a consequence of the fact that both problems are undecidable in the case of a fixed number of processes:

**Theorem 5.12 ([AEY05, HMK+05]).** Both problems, rMSG-Realizability and rMSG-Implementability, are undecidable.

In the following, we focus on the implementability question. To overcome undecidability, we introduce a criterion, called executability, that is sufficient for implementability but turns out to be decidable. As a second step, we then identify a fragment of rMSGs for which executability and implementability coincide so that implementability is actually decidable for that class.

Let $M = (E, \prec, \pi, \mu) \in \text{dMSC}(A, P)$ be an MSC. Intuitively, $M$ is executable if a process $p$ may know the pid of a process $q$ at the time of (i) sending a message to $q$, or (ii) sending the pid $q$ to some other process. Here, process $p$ may know $q$ if it is possible to pass $q$ to $p$ along the message flow. Let us be more precise. Given $q \in Pids(M)$ and an event $e \in E$ of $M$, we write $q \leadsto_M e$ if there is a path from the minimal event $\min_q(M)$ of $q$ to $e$ in $M$. This path might involve the reversal of the spawn edge that started $q$. That is, $q \leadsto_M e$ if $(\min_q(M), e) \in (\prec \cup \prec_{\text{new}}^{-1})^*$. In a sense, $q \leadsto_M e$ indicates that the process executing $e$ is aware of process $q$. Next, we formally define executability of MSCs.

**Definition 5.13 (executability).** Let $M = (E, \prec, \pi, \mu) \in \text{dMSC}(A, P)$. A message $(e, f) \in \prec_{\text{msg}}$ with contents $\mu(e, f) = a(p_1, \ldots, p_n)$ is executable if

- $\{p_1, \ldots, p_n\} \subseteq Pids(M)$, and
- $q \leadsto_M e$ for every $q \in \{\pi(f), p_1, \ldots, p_n\}$.

Moreover, $M$ is executable if each of its messages is executable. Finally, an rMSG $\mathcal{H}$ is executable if each MSC from $L(\mathcal{H})$ is executable.
Let $M$ be an MSC and $\mathcal{H}$ be an rMSG. One can verify that

- $M$ is executable iff $\{M\}$ is implementable, and
- $\mathcal{H}$ is executable if it is implementable (while the converse might fail).

**Example 5.14.** For illustration, consider Figures 5.4 and 5.5 on page 86. There, $M_2, M_3, M_4$ are executable, while $M_1$ is not. Moreover, the rMSG $\mathcal{H}_{5.6}$ is executable (recall that it is actually implementable).

Executability leads us to another decision problem:

**Problem 5.15.** rMSG-Executability:

**Input:** $A$, rMSG $\mathcal{H}$ over $A$

**Question:** Is $L(\mathcal{H})$ executable?

Unlike implementability, executability of rMSGs is decidable:

**Theorem 5.16 ([BCH+13]).** rMSG-Executability is in Pspace.

*Proof (sketch).* We will sketch the proof idea. The problem is reduced to a reachability problem in a finite transition system. Essentially, the transition system is a finite unfolding of the given rMSG $\mathcal{H}$. In particular, transitions are labeled with the finitely many partial MSCs that occur in $\mathcal{H}$. The locations of $\mathcal{H}$ are extended to contain information about which processes know each other, or, equivalently, about the $\rightsquigarrow$-paths that exist between processes. We only need to take into account the active processes, i.e., those processes that possibly execute further actions. As these active processes can be identified with their respective registers, all the information we need can be maintained within a finite alphabet.

More precisely, a state of the transition system that we are going to define is a pair $(\ell, CS)$ where $\ell$ is a location of $\mathcal{H}$ and $CS$ is the current communication structure. The latter is a graph whose nodes are some of the (finitely many) registers. There is an edge $r \rightsquigarrow r'$ if, informally speaking, the process associated with $r$ is known by (or, was communicated to) the process associated with $r'$. Now, there is a simulation relation between the “real”, infinitely many configurations of $\mathcal{H}$ and the finitely many states of the transition system. This simulation relation is correct in the following sense. If a path to $(\ell, CS)$ in the transition system simulates a run $\rho$ of $\mathcal{H}$, say, to configuration $(\ell, \tau, U)$, then, for all distinct $r, r' \in \text{dom}(\tau)$, we have that $\tau(r) \rightsquigarrow_{M(\rho)} \max_{\tau(r')} (M(\rho))$ iff $r \rightsquigarrow r'$ is an edge in the communication structure $CS$.

Now, rMSG $\mathcal{H}$ is not executable iff there is an accepting path in the transition system using a transition $(\ell, CS) \xrightarrow{M} (\ell', CS')$ such that the (symbolic) partial MSC $M$ is not executable wrt. $CS$. Figure 5.7 illustrates a part of the transition...
Figure 5.7: Part of the finite symbolic transition system associated with $\mathcal{H}_{5.6}$

The system for the executable rMSG $\mathcal{H}_{5.6}$ from Figure 5.6. Note that the crossed transition is actually not in the transition system and just added for illustration: since $r_0$ does not know $r_1$ in the current communication structure, it could not send a message to $r_1$. From the existence of an accepting path using such a transition, we could actually deduce that the underlying rMSG is not executable.

There is an alternative proof of decidability of rMSG-EXECUTABILITY, though it does not directly give us a complexity upper bound. Indeed, executability is a property that is definable in MSO logic. In turn, model checking rMSGs against MSO properties is decidable, which can be shown along the same lines as for session automata (cf. Theorem 4.35 on page 77).

**Definition 5.17.** The set $\text{MSO}(A, P)$ of MSO formulas over $A$ and $P$ is built from the following grammar:

$$
\varphi ::= x \triangleleft_{\text{new}} y \mid x \triangleleft_{\text{proc}} y \mid x \triangleleft_{\text{msg}}^{a(x_1, \ldots, x_n)} y \mid x = y \mid x \in X \mid \\
\neg \varphi \mid \varphi_1 \lor \varphi_2 \mid \exists x. \varphi \mid \exists X. \varphi
$$

where $a \in A$ with $\text{arity}(a) = n$.

The only predicate that needs explanation is $x \triangleleft_{\text{msg}}^{a(x_1, \ldots, x_n)} y$. It says that events $x$ and $y$ exchange a message of the form $a(p_1, \ldots, p_n)$ such that event $x_i$ is located on $p_i$, for all $i \in [n]$.\(^2\) With this, the satisfaction relation $M \models \varphi$ for an MSC $M$ over $A$ and $P$ and a sentence $\varphi \in \text{MSO}(A, P)$ is defined as expected. We are now ready to state that MSO model checking of rMSGs is decidable.

\(^2\)Alternatively, we could choose a logic in the style of Definition 3.19 on page 50, which quantifies over processes. Since a process can be identified with one of its event, this would not change the expressive power.
Theorem 5.18. The following problem is decidable:

**Instance:** \( A ; \) rMSG \( \mathcal{H} \) over \( A \) and \( \mathcal{P} ; \varphi \in \text{MSO}(A, \mathcal{P}) \)

**Question:** \( M \models \varphi \) for all \( M \in L(\mathcal{H}) \)?

The theorem generalizes a similar model-checking result for MSGs (without process creation) [Mad01, MM01].

Let us come back to the alternative proof of decidability of executability of rMSGs. By Theorem 5.18, it is sufficient to come up with an MSO \((A, \mathcal{P})\)-formula that defines executability. We let

\[
\varphi_{\text{exec}} = \bigwedge_{a \in A} \forall x, y, x_1, \ldots, x_n. (x \prec_{\text{msg}}^{a(x_1, \ldots, x_n)} y \rightarrow (y \leadsto^* x \wedge \bigwedge_{i \in [n]} x_i \leadsto^* x))
\]

where \( (y \leadsto^* x) = \exists y' (y' \prec_{\text{proc}}^* y \wedge (y', x) \in (\prec \cup \prec_{\text{new}})^*). \) Then, \( M \) is executable iff \( M \models \varphi_{\text{exec}} \). This also proves decidability of rMSG-EXECUTABILITY (Problem 5.15 and Theorem 5.16).

### 5.5 Guarded Register MSC Graphs

Now that we have a decidable criterion that is sufficient for implementability, we are interested in a fragment of rMSGs for which the criterion and implementability coincide so that the latter actually becomes decidable. To do so, we adapt definitions that were given for MSGs [GMSZ06] and guarantee their implementability. Guarded rMSGs are based on the notion of a leader process, which determines the next transition to be taken in an rMSG.

**Definition 5.19 (guarded rMSG).** An rMSG \( \mathcal{H} = (\mathcal{L}, \mathcal{R}, \mathcal{L}_0, \mathcal{L}_{\text{acc}}, r_0, T) \) over \( A \) is said to be guarded if

1. for all partial MSCs \( M = (E, \prec, \pi, \mu) \in p\text{MSC}(A, \mathcal{R}) \) that occur in \( \mathcal{H} \), \( (E, \prec^*) \) has a unique minimal element \( e \)—we let \( \text{first}(M) = \pi(e) \), and

2. there is a mapping \( \text{leader} : \mathcal{L} \rightarrow \mathcal{R} \) such that, for all transitions \( \ell \xrightarrow{M} \ell' \),
   - \( \text{leader}(\ell) = \text{first}(M) \),
   - \( \text{leader}(\ell') \in \text{Pids}(M) \), and,
   - for all \( r \in \text{Pids}(M) \), \( \max_e(M) \prec^* \max_{\text{leader}(\ell')}(M) \).

Intuitively, the last condition of 2. says that all processes of \( M \) terminate before \( \text{leader}(\ell') \) terminates.
Example 5.20. The rMSG $H_{5.6}$ from Example 5.9 on page 88 is guarded. The leaders as required in Definition 5.19 are given by $\text{leader}(\ell_i) = r_i$ for $i \in \{0, 1, 2\}$, and $\text{leader}(\ell_i) = r_0$ for $i \in \{3, 4\}$.

Indeed, executability is both sufficient and necessary for implementability when we restrict to the class of guarded rMSGs.

\begin{theorem} ([BCH+13]). A guarded rMSG is implementable iff it is executable. Moreover, if it is implementable, then an equivalent DCA can be constructed in exponential time. \end{theorem}

\textit{Proof (sketch).} The proof is by construction of a DCA from an executable rMSG. We give only a rough description of it. Essentially, we start from the rMSG enriched with communication structures, as described in the proof of Theorem 5.16. This enriched structure is then projected onto the processes (i.e., registers). However, the nondeterminism in the rMSG can lead to potentially inconsistent runs if each process is allowed to choose the next transition. The guardedness property helps us to avoid this problem. We let the leader process nondeterministically choose the next transition. This decision can be communicated to all other active processes, thanks to the connectedness of the partial MSCs. During the simulation of each process, we also attach the current communication structure to the messages so that each process is aware of the processes that it knows and that other processes know. This is essential to infer which registers have to/must not be updated. By guardedness of the underlying rMSG, it will be guaranteed that the communication structure sent/received is always up-to-date. \hfill \blacksquare

5.6 Perspectives

In [BCH+13], we actually presented a richer version of rMSGs, which is equipped with a branching operator (the formalism is called branching high-level MSCs). This operator allows one to divide a current computation into subcomputations, splitting the registers, and to merge the subcomputations once they terminate. Branching high-level MSCs were inspired by branching automata [LW00, LW01], which serve as recognizers of series-parallel pomsets. In terms of sequentializations, adding branching capabilities may be seen as using an additional stack. Some protocols such as the leader-election protocol, which are perfectly implementable in terms of DCAs, can only be specified globally using pushdown stacks. The seemingly close relation with series-parallel pomsets may help us to establish an algebraic approach to dynamic concurrent systems. A related question for future research concerns extensions of Zielonka’s theorem to a dynamic setting (cf. also its generalization to a recursive setting in Chapter 2). We may consider rMSGs as analogon of a finite automaton and study its closure under independent events. One may then ask if this guarantees implementability in terms of a DCA.
In this chapter, we turn towards quantitative systems. More precisely, we consider concurrent systems with a static and fixed communication structure, whose processes obey timing constraints. Timed automata [AD94] are a well-studied formalism to describe sequential systems with timing. Networks of timed automata have then been extensively studied as models of distributed timed systems (see, e.g., [LPY95, BHR06, GL08, CCJ06, BCH12]). In most cases, it was assumed that all processes evolve at the same speed, having knowledge of a global time.

However, it is not always justified to assume that all processes proceed at the same speed. Clock evolution may depend on temperature, workload, clock precision, etc., and drifting clocks are all the more a realistic assumption, since processes are more and more often executed on physically distributed computers. In this section, we study timed automata whose clocks evolve independently. More precisely, the set of clocks is partitioned into equivalence classes, and equivalent clocks are running at the same speed, while others do not necessarily. We will model timed systems with one global state space, focussing on the new aspects of independently evolving clocks. There are actually several natural distributed versions of that model (see [DL07, ABG+14], for example), involving several processes. In particular, one may assume that every clock is assigned to an owner process, and consider that two clocks are equivalent if they belong to the same process. Actually, the semantics of the distributed model considered in [ABG+14] is defined in terms of the timed systems defined below.
In the following, we let $\mathbb{R}_{\geq 0}$ denote the set of non-negative real numbers and set $\mathbb{R}_{> 0} = \mathbb{R}_{\geq 0} \setminus \{0\}$. For an alphabet $\Sigma$, we let $\Sigma_e \overset{\text{def}}{=} \Sigma \cup \{\varepsilon\}$.

Let us formally define our model of timed automata with independently evolving clocks, as well as their semantics.

**Definition 6.1 (icTA).** A timed automaton with independently evolving clocks (icTA) over $P$ is a tuple $\mathcal{A} = (S, \Sigma, C, \Delta, I, F, \sim)$ where $(S, \Sigma, C, \Delta, I, F)$ is a classical timed automaton \[AD94\] and $\sim$ is an equivalence relation to determine which clocks run at the same speed. Let us be more precise:

- $S$ is the nonempty finite set of states,
- $\Sigma$ is the finite alphabet of actions,
- $C$ is the nonempty finite set of clocks,
- $\Delta \subseteq S \times \Sigma \times \text{Constr}(C) \times 2^C \times S$ is the finite set of transitions (where the set Constr($C$) of clock constraints is specified below),
- $I \in S$ is the initial state,
- $F \subseteq S$ is the set of final states, and
- $\sim \subseteq C \times C$ is an equivalence relation.

In the above definition, the set Constr($C$) of clock constraints over $C$ is given by the grammar

$$\varphi ::= \text{true} \mid \text{false} \mid x \preceq c \mid \lnot \varphi \mid \varphi_1 \lor \varphi_2 \mid \varphi_1 \land \varphi_2$$

where $x$ ranges over $C$, $\preceq \in \{<, \leq, >, \geq, =\}$, and $c \in \mathbb{N}$. A (clock) valuation is a mapping $\nu : C \to \mathbb{R}_{\geq 0}$ assigning to each clock its current value. Satisfaction of a clock constraint $\varphi$ wrt. a valuation $\nu \in \text{Val}(C)$ is defined as expected. If $\nu$ satisfies $\varphi$, we write $\nu \models \varphi$. The set of clock valuations over $C$ is denoted by $\text{Val}(C)$.

A transition $(s, a, \varphi, R, s') \in \Delta$ shall be read as follows: being in state $s$, the automaton can move on to state $s'$ and perform $a \in \Sigma$ provided the current clock values satisfy constraint $\varphi$. When the transition is taken, clocks from the set $R \subseteq C$ are reset.

When $x \sim y$, then the clocks $x$ and $y$ evolve at the same speed, whereas $x \not\sim y$ means that they evolve independently. Thus, if $\sim = C \times C$, then we deal with a classical timed automaton. By $[x] \overset{\text{def}}{=} \{y \in C \mid x \sim y\}$, we denote the equivalence

\[\text{Note that our model from [ABG14]}\] includes invariants (in terms of clocks constraints) on states, which were necessary to obtain some negative results in the distributed setting. They are, however, not essential for the presentation in this chapter so that we omit them for simplicity.
class of \(x\). Since \([x]\) can be seen as a set of clocks that belong to the same process, we let \(\mathcal{P} \stackrel{\text{def}}{=} \{[x] \mid x \in \mathcal{C}\}\) denote the set of equivalence classes induced by \(\sim\) (a corresponding index is omitted, as \(\sim\) will be clear from the context). Moreover, we let \(p\) and \(q\) range over \(\mathcal{P}\).

Equivalent clocks evolve according to some local time. In turn, local time is modeled relative to some global time. More precisely, for \(p \in \mathcal{P}\), a local time function is a mapping \(\tau_p : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}\) (from global time to local time), with \(\tau_p(0) = 0\), that is strictly increasing, diverging, and continuous. We set \(\text{Rates}\) to be the set of tuples \((\tau_p)_{p \in \mathcal{P}}\) where each \(\tau_p\) is a local time function. Note that \(\tau \in \text{Rates}\) can also be seen as a function \(\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}\). Each of the two diagrams below illustrates an element from \(\text{Rates}\), where we assume that there are two equivalence classes, \(p\) and \(q\).

For a valuation \(\nu : \mathcal{C} \rightarrow \mathbb{R}_{\geq 0}\) and \(t = (t_p)_{p \in \mathcal{P}} \in \mathbb{R}_{\geq 0}^{\mathcal{P}}\), let \(\nu + t\) be the valuation defined by \((\nu + t)(x) = \nu(x) + t_{[x]}\). Intuitively, when time passes, we add, to every clock \(x\), the time elapse that corresponds to the local time of \([x]\).

Usually, the semantics of a timed automaton is defined in terms of an infinite automaton. A configuration of the infinite automaton keeps track of the current state and the current clock valuation. In our setting, since the semantics will be defined wrt. local times, we have to keep track of the current reference point in the global-time scale. Thus, for \(\tau \in \text{Rates}\), we define an infinite automaton \(A_{\tau} = (\mathcal{S}, \Sigma, \rightarrow, \bar{\iota}, F)\) where

- \(\mathcal{S} = S \times \text{Val}(\mathcal{C}) \times \mathbb{R}_{\geq 0}\) is the set of configurations,
- \(\bar{\iota} = (\iota, \{x \mapsto 0 \mid x \in \mathcal{C}\}, 0)\) is the initial configuration, and
- \(F = F \times \text{Val}(\mathcal{C}) \times \mathbb{R}_{\geq 0}\) is the set of final configurations.

Moreover, \((s, \nu, t), a, (s', \nu', t') \in \mathcal{S} \times \Sigma_c \times \mathcal{S}\) is contained in the transition relation \(\rightarrow\) (and we write \((s, \nu, t) \xrightarrow{a} (s', \nu', t')\)) if \(t \leq t'\), there are \(\varphi \in \text{Constr}(\mathcal{C})\) and \(R \subseteq \mathcal{C}\) such that

- \((s, a, \varphi, R, s') \in \Delta,\)
- \(\nu + \tau(t') - \tau(t) \models \varphi,\) and

97
\[ \nu' = (\nu + \tau(t') - \tau(t))[R] \] (which behaves like \( \nu + \tau(t') - \tau(t) \) on all clocks from \( C \setminus R \), and maps all clocks from \( R \) to 0).

The language \( L(A_\tau) \) of \( A_\tau \) is then defined as expected as a subset of \( \Sigma^* \). The \( \tau \)-semantics of \( A \) is \( L(A, \tau) \overset{\text{def}}{=} L(A_\tau) \).

However, the precise local time rate \( \tau \) may be difficult to predict or simply be unknown. This is why we introduce two more semantics (actually three as we will see later) that do not depend on concrete time rates. One is an overapproximation of the \( \tau \)-semantics, while the other is an underapproximation. The choice of the precise semantics depends on the system property that one may wish to verify.

**Definition 6.2 (semantics).** Let \( A = (S, \Sigma, C, T, s_0, F, \sim) \) be an icTA. We let

\[ L^\exists(A) \overset{\text{def}}{=} \bigcup_{\tau \in \text{Rates}} L(A, \tau) \]

\[ L^\forall(A) \overset{\text{def}}{=} \bigcap_{\tau \in \text{Rates}} L(A, \tau) \]

be the existential and, respectively, universal semantics of \( A \).

**Example 6.3.** Consider the icTA \( A \) from Figure 6.1, with clocks \( x \) and \( y \) such that \( x \not< y \). Suppose first that both clocks evolve at the same speed as global time, i.e., we assume \( \text{id} = (\tau[x], \tau[y]) \in \text{Rates} \) where \( \tau[x] \) and \( \tau[y] \) are both the identity function on \( \mathbb{R}_{\geq 0} \). Then, we actually deal with an ordinary timed automaton, and it is easily seen that \( A \) accepts \( a, ab, \) and \( b \). By definition, all these words are also in the existential semantics. On the other hand, \( L(A, \text{id}) \) does not contain \( c \): since there are no resets in \( A \), accepting \( c \) would require \( y \) to run faster than \( x \) at some point, which contradicts the assumption of synchronous clocks. We deduce that \( c \) is not contained in the universal semantics of \( A \). However, it is in the existential one: we can choose time rates such that \( x < 1 < y \) is eventually satisfied, activating the transition from \( s_0 \) to \( s_5 \). The same rates witness the fact that \( b \) is not contained in the universal semantics. We still have to check containment of \( a \) and \( ab \) in \( L^\forall(A) \). Clearly, \( a \in L^\forall(A) \): since local time rates are strictly increasing, \((0 < x < 1) \land (0 < y < 1) \) is eventually satisfied so that the transition from \( s_0 \) to \( s_1 \) can always be taken. Finally, \( ab \in L^\forall(A) \) also holds, though by a more subtle argument. First, note that guard \( y \leq 1 \leq x \) on the transition from \( s_1 \) to \( s_3 \) and guard \( x < 1 = y \) on the transition from \( s_2 \) to \( s_3 \) are complementary: one of them will eventually be satisfied, but not both of them. As local time rates are given in advance, there is always a way to reach the final state \( s_3 \) while reading \( ab \).

Summarizing, we have

- \( L(A, \text{id}) = \{a, ab, b\} \),
- \( L^\exists(A) = \{a, ab, b, c\} \), and
- \( L^\forall(A) = \{a, ab\} \).

As suggested above, we may now use \( L^\exists(A) \) to check safety properties: given a regular set \( \text{Bad} \subseteq \Sigma^* \) of undesired behaviors, the corresponding model-checking
problem asks whether \( L_3(A) \cap \text{Bad} = \emptyset \). Dually, we use \( L_\nu(A) \) to check liveness properties: given a regular set \( \text{Good} \) of behaviors that the system should exhibit, no matter what the local time functions are, we would like to have \( \text{Good} \subseteq L_\nu(A) \). So, the next question to ask is if the above problems are solvable, say, when \( \text{Bad} \) and \( \text{Good} \) are given as regular languages. While the answer is positive in the case of the existential semantics, we will show that nonemptiness (and universality) of the universal semantics are undecidable.

Note that very similar models with clock drifts have already been studied [Pur00, DDMR04, DL07, SFK08]. However, the semantics considered there rather correspond to the existential semantics and the questions studied were actually quite different.

### 6.2 The Existential Semantics

In this section, we argue that the existential semantics is always regular and can be effectively computed. Therefore, it can be used to check safety properties of a given icTA.

**Theorem 6.4** ([ABG+14]). Let \( A \) be an icTA. Then, \( L_3(A) \) is a regular language that can be effectively represented as a finite automaton.

In the rest of this section, we sketch the proof of Theorem 6.4.

Note that the above result was known for timed automata with perfectly synchronous clocks. Actually, the proof of Theorem 6.4 is done via a conservative extension of the region automaton associated with an ordinary timed automaton. The main idea of the region automaton is to discretize clock valuations, i.e., to classify them into finitely many equivalence classes, called (clock) regions. The regions of a classical timed automaton with two clocks are depicted on the left-hand side of Figure 6.2. The clue is that

- regions are coarse enough to give rise to finitely many equivalence classes,
• regions are fine enough to be faithfully evaluated over the guards that occur in the underlying finite automaton, and
• equivalent regions are crossing the same regions when letting time elapse and resetting clocks from time to time. In particular, the time-elapse successors for a given region are uniquely determined.

Note that the region partitioning actually depends on the largest constant that a clock is compared with in the automaton (in the example, 2 for clock $y$ and 3 for clock $x$). Note that the left-hand side of Figure 6.2 indeed assumes that clocks $x$ and $y$ are synchronous. Speaking in terms of icTAs, they are equivalent, i.e., $x \sim y$. When $x \not\sim y$, the region equivalence is coarser. As illustrated on the right-hand side of Figure 6.2, the diagonal regions are abandoned. The reason is that $x$ and $y$ may evolve at different speeds. In particular, from the initial region, induced by $\{x \mapsto 0, y \mapsto 0\}$, any other region is “reachable”. Actually, we must not include diagonal regions, as the induced region-successor relation would not be a partial order anymore and spoil forthcoming constructions.

Let us be (slightly) more formal, and let $\mathcal{A} = (S, \Sigma, C, \Delta, \iota, F, \sim)$ be an icTA. Let $p \in \mathcal{P}$. Given a clock valuation $\nu \in \text{Val}(C)$, we define its $p$-restriction $\nu_p : p \to \mathbb{R}_{\geq 0}$ by $\nu_p(x) = \nu(x)$ for all $x \in p$. We say that two clock valuations $\nu$ and $\nu'$ over $C$ are equivalent if, for all $p \in \mathcal{P}$, the restrictions $\nu_p$ and $\nu'_p$ are equivalent in the classical sense wrt. the clocks in $p$. Recall that this equivalence depends on the largest constants the clocks are compared with. We write $[\nu]$ for the equivalence class of $\nu$. By $\text{Regions}(\mathcal{A}) \overset{\text{def}}{=} \{[\nu] \mid \nu \in \text{Val}(C)\}$, we denote the set of (clock) regions induced by $\mathcal{A}$. Note that $\text{Regions}(\mathcal{A})$ is finite.

Let $\gamma, \gamma' \in \text{Regions}(\mathcal{A})$ be two clock regions. We say that $\gamma'$ is accessible from $\gamma$, written $\gamma \sqsubseteq \gamma'$, if either $\gamma = \gamma'$ or there are $\nu \in \gamma$, $\nu' \in \gamma'$, and $t \in \mathbb{R}_{\geq 0}$ such that $\nu' = \nu + t$. Note that $\sqsubseteq$ is a partial-order relation on $\text{Regions}(\mathcal{A})$. The direct-successor relation, written $\gamma \sqsupset \gamma'$, is, as usual, defined by $\gamma \sqsubseteq \gamma', \gamma \not\sqsubseteq \gamma'$, and $\gamma'' = \gamma$ or $\gamma'' = \gamma'$ for all clock regions $\gamma''$ with $\gamma \sqsubseteq \gamma'' \sqsubseteq \gamma'$. The relation $\sqsupset$ is represented on the right-hand side of Figure 6.2 by arrows pointing from one to another region.

\footnote{We assume the reader is familiar with the region equivalence for ordinary timed automata and simply refer to [AD94] or to the left-hand side of Figure 6.2.}
We will now associate with an icTA $A = (S, \Sigma, C, \Delta, \iota, F)$ a finite automaton recognizing its existential semantics (the correctness proof is omitted, though). A modified, game-based version of that finite automaton will later be used to define a semantics that underapproximates the universal semantics. We define the finite automaton $A_3 = (S, \bar{\iota}, F)$ over $\overline{\Sigma}$ by

- $S = S \times \text{Regions}(A)$,
- $\bar{\iota} = (\iota, \{x \mapsto 0 \mid x \in C\})$, and
- $F = F \times \text{Regions}(A)$.

Moreover, the transition relation $\delta \subseteq S \times \overline{\Sigma} \times S$ is partitioned into a set $\delta_0$ of discrete transitions and a set $\delta_1$ of timed transitions:

- The set $\delta_0$ contains $(s, \gamma) \xrightarrow{a} (s', \gamma')$ if there are $\nu \models \varphi$ and $(s, a, \varphi, R, s') \in \Delta$ such that $\nu[R] \in \gamma'$.
- The set $\delta_1$ contains $(s, \gamma) \xrightarrow{a} (s', \gamma')$ if $a = \varepsilon$, $s = s'$, and $\gamma \subseteq \gamma'$.

**Example 6.5.** Part of the region automaton $A_3$ belonging to the icTA $A$ from Figure 6.1 is depicted in Figure 6.3. Discrete transitions origin in rounded boxes, time-elapse transitions in rectangular boxes. The extract shows that $a$ and $ab$ are contained in the existential semantics of $A$ and that $A_3$ admits even two accepting runs on $ab$. The latter observation corresponds to the fact that $ab$ is accepted for local time rates that are crossing different regions.

Now, Theorem 6.4 follows by the following lemma:

**Lemma 6.6.** We have $L(A_3) = L_\exists(A)$. 

101
6.3 The Universal Semantics

Next, we show that, unlike the existential semantics, the universal semantics comes with an undecidable nonemptiness problem.

**Theorem 6.7 ([ABG+14]).** The following problem is undecidable:

**Instance:** icTA \( A \)

**Question:** \( L_\varphi(A) \neq \emptyset \)?

In the rest of this section, we prove Theorem 6.7. The proof is by reduction from Post’s correspondence problem (PCP).

**Problem 6.8.** Post’s correspondence problem (PCP):

**Input:** Finite alphabet \( \Sigma \) and morphisms \( f, g : \Sigma^* \to \{0, 1\}^* \)

**Question:** Is there \( w \in \Sigma^+ \) such that \( f(w) = g(w) \)?

Let \( \Sigma = \{a_1, \ldots, a_k\} \), where \( k \geq 1 \), and \( f, g \) be an instance of the PCP. We will determine an icTA \( A \) with set of clocks \( \{x, y\} \) and \( x \not= y \) such that \( L_\varphi(A) = \{w \in \Sigma^+ | f(w) = g(w)\} \), i.e., the universal semantics of \( A \) contains precisely the solutions to the PCP instance. One ingredient of the construction is an encoding of sequences over \( \{0, 1\} \) in terms of local time functions. Let \( p = \{x\} \) and \( q = \{y\} \) be the equivalence classes induced by \( \sim \), and suppose \( \tau = (\tau_p, \tau_q) \in Rates \). Actually, \( \tau \) will encode a word in \( \{0, 1, 2\}^\omega \). We do this using \((1 \times 1)\)-square regions. If the rate function leaves this region by the upper boundary or by the right boundary, then we write 0 or 1, respectively. If it leaves the square by the end-point \((1, 1)\), then we write 2. A new square region is started at the point where the rate function left the old one. Thus, the direction sequences partition the space of time rates. This is illustrated in Figure 6.4, where we obtain \( dir(\tau) = 101 \ldots \)

![Figure 6.4: Binary sequence associated with local times](image)

Let us formalize the encoding. With \( \tau \), we associate a sequence \( t-dir(\tau) = t_1t_2 \ldots \in (\mathbb{R}_{\geq 0})^\omega \) of time instances and a sequence \( dir(\tau) = d_1d_2 \ldots \in \{0, 1, 2\}^\omega \) of directions.
as follows: for \( i \geq 1 \), we let first (assuming \( t_0 = 0 \)) \( t_i = \min \{ t > t_{i-1} \mid \tau_r(t) - \tau_r(t_{i-1}) = 1 \} \) for some \( r \in \{ p, q \} \). With this, we set

\[
d_i = \begin{cases} 
0 & \text{if } \tau_p(t_i) - \tau_p(t_{i-1}) < 1 \text{ and } \tau_q(t_i) - \tau_q(t_{i-1}) = 1 \\
1 & \text{if } \tau_q(t_i) - \tau_q(t_{i-1}) < 1 \text{ and } \tau_p(t_i) - \tau_p(t_{i-1}) = 1 \\
2 & \text{otherwise}
\end{cases}
\]

We will now construct the icTA \( A = (S, \Sigma, C, T, \iota, F, \sim) \), with \( \Sigma = \{ a_1, \ldots, a_k \} \), \( C = \{ x, y \} \), and \( x \neq y \), such that \( L_\psi(A) = \{ w \in \Sigma^+ \mid f(w) = g(w) \} \). We proceed in two steps:

**Step 1:** We construct icTAs

\[
A_f = (S, \Sigma, C, \Delta, \iota, F, \sim) \\
A_g = (S', \Sigma, C, \Delta', \iota', F', \sim)
\]

(with the same \( \Sigma, C, \) and \( \sim \)) such that, for all \( \tau \in Rates \):

\[
L(A_f, \tau) = \{ w \in \Sigma^+ \mid f(w).2 \not\leq dir(\tau) \} \\
L(A_g, \tau) = \{ w \in \Sigma^+ \mid g(w).2 \leq dir(\tau) \}
\]

Here, \( \not\leq \) denotes the prefix relation on words and \( u.v \) denotes the concatenation of \( u \) and \( v \).

**Step 2:** We build an icTA \( A = A_f \cup A_g \), which simply branches nondeterministically into \( A_f \) or \( A_g \).

With this, one can easily show the following:

**Claim 6.9.** We have

\[
L_\psi(A) = \{ w \in \Sigma^+ \mid f(w) = g(w) \}.
\]

*Proof.* For the inclusion from left to right, let \( w \in L_\psi(A) \). Then, \( w \in \Sigma^+ \) and, for all \( \tau \in Rates \), \( w \in L(A_f, \tau) \) or \( w \in L(A_g, \tau) \), i.e., \( f(w).2 \not\leq dir(\tau) \) or \( g(w).2 \leq dir(\tau) \). In particular, there is \( \tau \in Rates \) with \( dir(\tau) \in f(w).2.\{0,1,2\}^\omega \) and such that \( f(w).2 \not\leq dir(\tau) \) or \( g(w).2 \leq dir(\tau) \). As \( f(w), g(w) \in \{0,1\}^* \), we have \( f(w) = g(w) \).

For the inclusion from right to left, let \( w \in \Sigma^+ \) such that \( f(w) = g(w) \). Let \( \tau \in Rates \). Trivially, we have \( f(w).2 \not\leq dir(\tau) \) or \( f(w).2 \leq dir(\tau) \). As \( f(w) = g(w) \), the latter implies \( g(w).2 \leq dir(\tau) \). Therefore, \( w \in L(A, \tau) \). We conclude that \( w \in L_\psi(A) \).

The construction of \( A_f \) and \( A_g \) is not difficult and can be found in [ABG+14]. Let us just mention that, to “detect” this encoding in automata, we use the guards

\[ x < 1 \land y = 1 \text{ (for 0)}, \]
as well as their negations. Then, applying a new square, as described above, corresponds to resetting both clocks simultaneously.

This concludes the proof of Theorem 6.7. A very similar construction yields undecidability of the universality problem:

\[
\begin{align*}
  y < 1 & \land x = 1 \text{ (for 1)}, \\
  x = 1 & \land y = 1 \text{ (for 2)},
\end{align*}
\]

\[\text{Remark 6.11.}\] The essence of the proof technique presented above is universal in the sense that it can be easily carried over to other settings such as word transducers. A word transducer over two finite alphabets \(\Sigma\) and \(\Gamma\) is a finite automaton \(A\) whose (finitely many) transitions have labels from \(\Sigma^* \times \Gamma^*\). In the expected manner, \(A\) determines a (rational) relation \(R(A) \subseteq \Sigma^* \times \Gamma^*\). A very well known decision problem is universality, which asks whether \(R(A) = \Sigma^* \times \Gamma^*\) holds. The problem can easily be shown undecidable by a reduction from the PCP. Slightly adapting our proof technique, one can now show that the existential version of the problem is undecidable, too: is there a word \(u \in \Sigma^*\) such that, for all \(v \in \Gamma^*\), we have \((u, v) \in R(A)\)? The problem seems natural: just like in the timed setting, the alphabet \(\Gamma\) may model a set of signals, emitted from an environment, that trigger behaviors from \(\Sigma\). When, now, the environment is not under control of the system, checking liveness specifications indeed amounts to asking the existential version of the better known problem. However, we are not aware of that undecidability result in the literature, where only universality seems to be considered.

6.4 The Reactive Semantics

The undecidability results around the universal semantics are, of course, unsatisfactory, since they imply that liveness properties cannot be checked for icTAs. In this section, we will present the reactive semantics. The reactive semantics is an underapproximation of the universal semantics. Like the existential semantics, it is always a regular set that can be effectively represented as a finite automaton. Thus, it will allow us to verify icTAs against (certain) liveness properties.

To motivate the reactive semantics, it is important to understand the reason for undecidability of the universal semantics and decidability in the existential case. To do so, it will be useful to consider both, the existential and the universal semantics, as a game. We may think of a two-player game, played by Player 0 and Player 1, where Player 0 controls the transitions of the icTA, while Player 1 controls the
time evolution. In the existential semantics, both players cooperate, and we may actually think of one single player who chooses, given a word $w \in \Sigma^*$, both, time rates and a run through the automaton accepting $w$. In the universal semantics, Player 0 aims at “accepting” the word, while Player 1 wants to “reject” the word. In other words, Player 1 tries to choose time rates in a way that does not allow Player 0 to navigate the icTA into a final state while reading $w$. Note that the game is not turned-based since Player 1 has to reveal its time rates entirely in advance. Thus, we actually deal with an imperfect-information game: Player 1 has no means to react upon the moves played by Player 0, so that it is, in a sense, blind. Now, imperfect information is actually often the reason for undecidability (see, e.g., [PR79, PRA01]). So, we will try to modify the game for the universal semantics to make it a turn-based, perfect-information game. In other words, we will equip Player 1 with more power.

In the following, we will first define a kind of reachability game, and then associate a game with an icTA and an input word. The reactive semantics will be defined as the set of words for which Player 0 has a winning strategy. This set can be represented as the language of an alternating two-way automaton, which shows its regularity.\(^3\)

**Definition 6.12 (game).** A game is a tuple $G = (S_0, S_1, \Sigma, \Delta, i, F)$ where

- $S_0$ and $S_1$ are disjoint sets of states, which belong to Player 0 and Player 1, respectively, and we let $S = S_0 \cup S_1$,
- $\Sigma$ is a finite alphabet,
- $i \in S$ is the initial state,
- $F \subseteq S$ is the set of final states, which are winning for Player 0, and
- $\Delta \subseteq S \times \Sigma_e \times S$ is the transition relation.

We require that, for all states $s \in S$, there are $a \in \Sigma_e$ and $s' \in S$ such that $(s, a, s') \in \Delta$.

A (positional) strategy for Player $i$, with $i \in \{0, 1\}$, is a mapping $\sigma : S_i \rightarrow \Sigma_e \times S$ such that $\sigma(s) = (a, s')$ implies $(s, a, s') \in \Delta$, for all $(s, a, s') \in S_i \times \Sigma_e \times S$. Let $\sigma_0$ and $\sigma_1$ be strategies for Player 0 and Player 1, respectively. The play of $G$ induced by $\sigma_0$ and $\sigma_1$ is the unique infinite sequence $(s_0, a_1, s_1, a_2, s_2, \ldots)$, with $s_i \in S$ and $a_i \in \Sigma_e$, such that $s_0 = i$ and, for all $i \geq 0$, $s_i \in S_0$ implies $\sigma_0(s_i) = (a_{i+1}, s_{i+1})$, and $s_i \in S_1$ implies $\sigma_1(s_i) = (a_{i+1}, s_{i+1})$. We say that $\sigma_0$ is winning for $w \in \Sigma^*$ if, for all strategies $\sigma_1$ for Player 1, the play $(s_0, a_1, s_1, a_2, s_2, \ldots)$ induced by $\sigma_0$ and $\sigma_1$ contains a position $n \geq 0$ such that both $s_n \in F$ and $w = a_1 \cdots a_n$ (i.e., $w$ is the concatenation of $a_1, \ldots, a_n$).

Now, we will associate, with a given icTA $A = (S, \Sigma, C, \Delta, i, F, \sim)$, the game $G(A) = (S_0, S_1, \Sigma, \Delta, i, F)$, which is defined as follows:

\(^3\)In the original article [ABG+14], the semantics was presented directly in terms of an alternating two-way automaton.
\[ S_0 = (S \times \text{Regions}(A) \times \{0\}) \cup \{\uparrow_0\}, \]
\[ S_1 = (S \times \text{Regions}(A) \times \{1\}) \cup \{\uparrow_1\}, \]
\[ \bar{t} = (t, \{x \mapsto 0 \mid x \in C\}, 0), \] and
\[ F = F \times \text{Regions}(A) \times \{0, 1\}. \]

Assume that \(\delta_0\) and \(\delta_1\) are the sets of discrete and, respectively, time-elapse transitions of the finite automaton \(A_\Xi\) (cf. Section 6.2). Then, \((s, \gamma, pl) \xrightarrow{a} (s', \gamma', pl') \in S \times \Sigma_c \times S\) is contained in \(\bar{\Delta}\) if one of the following holds:

- \(pl = 0\) and \(pl' = 1\) and \((s, \gamma) = (s', \gamma')\), or
- \(pl = 0\) and \(pl' = 0\) and \(((s, \gamma), a, (s', \gamma')) \in \delta_0\), or
- \(pl = 1\) and \(pl' = 0\) and \(((s, \gamma), a, (s', \gamma')) \in \delta_1\).

Moreover, \(\bar{\Delta}\) contains transitions

- \((\uparrow_0, \varepsilon, \uparrow_0)\) for all \(s \in S_0\), and
- \((\uparrow_1, \varepsilon, \uparrow_1)\) and \(((s, \gamma, 1), \varepsilon, \uparrow_1)\) for all \(s \in S\) and \(\subseteq\)-maximal \(\gamma\).

**Definition 6.13.** Let \(A = (S, \Sigma, C, \Delta, t, F, \sim)\) be an icTA. The reactive semantics of \(A\) is defined as the set

\[ L_{\text{react}}(A) \overset{\text{def}}{=} \{ w \in \Sigma^* \mid \text{Player 0 has a winning strategy for } w \text{ in } \mathcal{G}(A) \} . \]
Example 6.14. Part of the game $\mathcal{G}(A)$ of the icTA $A$ from Figure 6.1 on page 99 is depicted in Figure 6.5. Here, rounded states belong to Player 0, and rectangular states belong to Player 1. We will argue that $a$ is the only word for which Player 0 has a winning strategy in $\mathcal{G}(A)$. To accept $a$, Player 0 will go from $r_0$ to $r_1$ (which is the only possible choice). Player 1 will have to respond by moving to $r_2$ whereupon Player 0 can conclude by going to the accepting configuration $r_3$, while reading $a$. However, Player 0 does not have a winning strategy for $ab$: whatever $a$-transition Player 0 will choose in $r_2$, Player 1 may go to $r_4$ or, respectively, $r_5$, so that reading $b$ is no longer possible. We deduce that $L_{\text{react}}(A) = \{a\}$. ◊

Theorem 6.15 ([ABG+14]). Let $A$ be an icTA. Then, $L_{\text{react}}(A)$ is a regular language that can be effectively represented as a finite automaton.

Proof (sketch). The reactive semantics $L_{\text{react}}(A)$ is the language of some alternating two-way automaton (over words). The automaton has existential transitions (those taken by Player 0) and universal transitions (those taken by Player 1). The fact that we need two-way transitions comes from the ε-moves. More precisely, the alternating two-way automaton does never go backward, but may stay at the same position when taking a transition.

6.5 Perspectives

There is an obvious difference between the model presented in this chapter and those from Chapters 2, 3, and 5: The semantics of a system is defined in terms of a word language, though we deal with a concurrent system. The main reason is that a pure graph and the associated partial order do not necessarily reflect causal dependencies when timing constraints come into play. There are some approaches to “marrying” the partial-order semantics and timed words (see, e.g., [BJLY98, LNZ05, BCH12]). Unfortunately, they are not directly applicable to our setting, since time stamps are in a sense meaningless in icTAs, as explained at the beginning of this chapter. However, since the partial-order approach is important and convenient for specifications, this aspect should be investigated more closely.

It will also be worthwhile to study the realizability problem, which we did not tackle for this timed case. In the distributed setting with several timed automata, the problem may read as follows: given a regular language $L$ and a process topology, is there a distributed timed automaton $A$ such that $L_{\exists}(A) = L_{\forall}(A) = L$? If the answer is affirmative, then $A$ may be seen as a robust implementation of $L$. Here, the process topology will fix the set of processes and their reading capabilities such as “clocks of $p$ can be read by $q”$ (in terms of guards). This framework is then similar to [Gen05] where untimed message-passing systems have been studied.
In this thesis, we conducted a study of automata models of concurrent programs, covering systems of various kinds. Those may involve recursive procedure calls, feature message-passing or shared-variable communication, and either come with a static and fixed topology or involve an unbounded number of processes, be it in a dynamic or parameterized setting. Our study was driven by applicability in formal design and verification, which stands or falls with the robustness of the formalisms considered. Whether an automata model is robust may be evaluated on the basis of the following criteria: decidability of the emptiness problem, closure under boolean operations, or existence of equivalent logical and algebraic characterizations. Not only are those properties interesting from an algorithmic point of view, but they also provide us with a deeper understanding of an automata class.

Unfortunately, even simple systems such as message-passing programs are inherently not robust: they have an undecidable emptiness problem and are not closed under complementation. However, limiting the behavior in a nontrivial meaningful way allows us to recover those properties. In this thesis, we introduced some automata models (and reviewed known ones in unifying frameworks) that turned out to be robust under some natural behavioral restrictions:

- **nested-trace automata** as a model of concurrent recursive programs,
- **parameterized communicating automata** as a model of message-passing systems with static but unknown process topology, and
- **session automata**, which are models of sequential dynamic systems.

We showed that all these models have a decidable emptiness problem and are expressively equivalent to monadic second-order logic when their behaviors are re-
stricted suitably to a bounded number of contexts or sessions, respectively. Moreover, we provided several decidability results for

- **dynamic communicating automata**, a model of message-passing systems with dynamically evolving communication topology, and

- **distributed timed automata**, which reflect concurrent systems whose processes communicate via independently evolving clocks.

Those results relied on different ideas. In the former case, we introduced a sufficient criterion for implementability to get some decidability, whereas the latter model came with natural under- and overapproximative semantics, though those were quite different from the idea of bounding the number of contexts.

In the future, since more and more systems are designed for an open world, parameterized and dynamic systems are particularly worth being studied further. Though they have indeed received increasing interest in recent years, there is, by now, no canonical approach to modeling and verifying such systems. A long-term goal should, therefore, be a unifying theory that lays algebraic, logical, and automata-theoretic foundations to support and facilitate the study of parameterized and dynamic concurrent systems. As we have seen, such theories indeed exist in non-parameterized settings where the number of processes and the way they are connected are fixed in advance. However, parameterized and dynamic systems lack such foundations and often restrict to very particular models with specialized verification techniques. This is especially true as far as algebraic approaches are concerned that may provide high-level expressions for parameterized systems. Algebraic studies of languages over infinite alphabets can be found in [BPT03, FK03, Boj13]. Algebraic approaches to nested-word languages have been proposed in [Cyr10, BS12]. Though it is not clear how all those techniques fit into the synthesis and verification of parameterized and dynamic systems, it will be worthwhile to explore possible connections.

In addition, several extensions/variations of the notions from this thesis are worth being studied:

- While we considered finite, terminating behaviors, the study of infinite structures would allow us to model reactive systems. There has been a lot of work on such extensions for concurrent systems (e.g., [DM94, Mus94, DGP95, Kus03, BK08]), and we believe that many of the results in this thesis go through in an infinite setting with minor adjustments.

- A substantial change in direction would be to consider branching-time behavior instead of a linear-time semantics. Recall that we defined the semantics of a system in terms of all (accepted) behaviors. As a consequence, a specification can only talk about one fixed behavior, but not about the branching structure of a system. While the branching behavior of a sequential system can be seen as a tree, that of a concurrent system is better reflected by an event structure [WN95]. Though that setting is substantially more complicated, there have been positive model-checking results [Pen97, Mad03]. An
unresolved question is whether those results can be transferred to recursive, dynamic, or parameterized systems. A starting point may be dynamic sequential systems (cf. Chapter 4) where a branching-time behavior is naturally represented by a data tree [BMSS09]. Recursive systems, in turn, would give rise to nested trees [ACM06].

- Models of concurrent systems that communicate with an (uncontrollable) environment are usually dealt with in the context of distributed games and distributed control/synthesis (cf. [SF07, GS13, GGMW13, MW14] for recent advances). Note that there are also some intimate connections with the above-mentioned branching-time setting [MTY05]. Parameterized synthesis has been studied in [JB14], and it would be worthwhile to explore possible extensions to a recursive or dynamic setting.


117


121


<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>alphabet</td>
<td>15</td>
</tr>
<tr>
<td>finite</td>
<td>15</td>
</tr>
<tr>
<td>infinite</td>
<td>15</td>
</tr>
<tr>
<td>ranked</td>
<td>80</td>
</tr>
<tr>
<td>automaton</td>
<td></td>
</tr>
<tr>
<td>asynchronous</td>
<td>29</td>
</tr>
<tr>
<td>branching</td>
<td>93</td>
</tr>
<tr>
<td>class memory (CMA)</td>
<td>70</td>
</tr>
<tr>
<td>class register (CRA)</td>
<td>68</td>
</tr>
<tr>
<td>data</td>
<td>72</td>
</tr>
<tr>
<td>fresh-register (FRA)</td>
<td>61</td>
</tr>
<tr>
<td>nested-trace (NTA)</td>
<td>28</td>
</tr>
<tr>
<td>nested-word (NWA)</td>
<td>19</td>
</tr>
<tr>
<td>register (RA)</td>
<td>62</td>
</tr>
<tr>
<td>session (SA)</td>
<td>62</td>
</tr>
<tr>
<td>timed</td>
<td>96</td>
</tr>
<tr>
<td>Zielonka</td>
<td>29</td>
</tr>
<tr>
<td>clock region</td>
<td>100</td>
</tr>
<tr>
<td>clock valuation</td>
<td>96</td>
</tr>
<tr>
<td>closed</td>
<td>82</td>
</tr>
<tr>
<td>communicating automaton (CA)</td>
<td>46</td>
</tr>
<tr>
<td>dynamic (DCA)</td>
<td>82</td>
</tr>
<tr>
<td>fixed-topology</td>
<td>47</td>
</tr>
<tr>
<td>parameterized (PCA)</td>
<td>49</td>
</tr>
<tr>
<td>weak parameterized</td>
<td>50</td>
</tr>
<tr>
<td>communication structure</td>
<td>90</td>
</tr>
<tr>
<td>concatenation</td>
<td>87</td>
</tr>
<tr>
<td>concretization</td>
<td>64</td>
</tr>
<tr>
<td>context</td>
<td>21</td>
</tr>
<tr>
<td>data language</td>
<td>62</td>
</tr>
<tr>
<td>data word</td>
<td>59</td>
</tr>
<tr>
<td>symbolic</td>
<td>64</td>
</tr>
<tr>
<td>well-formed symbolic</td>
<td>64</td>
</tr>
<tr>
<td>dependence relation</td>
<td>30</td>
</tr>
<tr>
<td>distributed alphabet</td>
<td>18</td>
</tr>
<tr>
<td>domain</td>
<td>15</td>
</tr>
<tr>
<td>executable</td>
<td>89</td>
</tr>
<tr>
<td>game</td>
<td>105</td>
</tr>
<tr>
<td>implementable</td>
<td>86</td>
</tr>
<tr>
<td>independence relation</td>
<td>30</td>
</tr>
<tr>
<td>interface name</td>
<td>40</td>
</tr>
<tr>
<td>Mazurkiewicz trace</td>
<td>3, 29</td>
</tr>
<tr>
<td>message sequence chart (MSC)</td>
<td>3</td>
</tr>
<tr>
<td>message sequence graph (MSG)</td>
<td>87</td>
</tr>
<tr>
<td>message sequence chart (MSC)</td>
<td>42</td>
</tr>
<tr>
<td>$(k, ct)$-bounded</td>
<td>45</td>
</tr>
<tr>
<td>$k$-bounded</td>
<td>44</td>
</tr>
<tr>
<td>branching high-level</td>
<td>93</td>
</tr>
<tr>
<td>dynamic</td>
<td>81</td>
</tr>
<tr>
<td>linearization of</td>
<td>44</td>
</tr>
<tr>
<td>partial</td>
<td>80</td>
</tr>
<tr>
<td>model checking</td>
<td>2</td>
</tr>
<tr>
<td>MSO logic</td>
<td></td>
</tr>
<tr>
<td>over dynamic MSCs</td>
<td>91</td>
</tr>
</tbody>
</table>
over MSCs, 47, 50
over nested traces, 33
over nested words, 23
over topologies, 49

nested trace, 27
  linearization of, 29
nested word, 18
  k-context, 21
  k-phase, 21
  k-scope, 21
  ordered, 22

phase, 21
Post’s correspondence problem, 102
process identifier (pid), 79, 80

realizability, 2
realizable, 86
register message sequence graph (rMSG), 87
  executable, 89
  guarded, 92
  representation, 32
  restriction, 22

semantics
  existential, 98
  reactive, 106
  universal, 98
series-parallel pomset, 93
specification, 5

temporal logic over nested words, 35
topology, 41
  grid, 42
  pipeline, 41
  ring, 42
  ring forest, 42
  tree, 42
transition label, 61

word, 15
  empty, 15