Appeared in: STACS 2002, Lecture Notes in Computer Science, vol. 2285, 489-500, 2002.

A further step towards a theory of regular MSC languages

Dietrich Kuske

Department of Mathematics and Computer Science
University of Leicester
LEICESTER
LE1 7RH, England
d.kuske@mcs.le.ac.uk

Abstract. This paper resumes the study of regular sets of Message Sequence Charts initiated by Henriksen, Mukund, Narayan Kumar & Thiagarajan [10]. Differently from their results, we consider infinite MSCs. It is shown that for bounded sets of infinite MSCs, the notions of recognizability, axiomatizability in monadic second order logic, and acceptance by a deterministic Message Passing Automaton with Muller acceptance condition coincide. We furthermore characterize the expressive power of first order logic and of its extension by modulo-counting quantifiers over bounded infinite MSCs.
Complete proofs can be found in the Technical Report [15].

1 Introduction

Message sequence charts (MSCs) form a popular visual formalism used in the software development. In its simplest incarnation, an MSC depicts the desired exchange of messages and corresponds to a single partial-order execution of the system. Several methods to specify sets of MSCs have been considered, among them MSC-graphs or High-level MSCs (HMSCs) that generate sets of MSCs by concatenating "building blocks", (Büchi-)automata that accept the linear extensions of MSCs, and logics. In general, these formalisms have different expressive power.

In [1], Alur & Yannakakis show that the collection of MSCs generated by a "bounded" ("locally synchronized" in the terminology of [10]) MSC-graph can be represented as a string language recognizable by a finite deterministic automaton. Based on this observation, Henriksen et al. [10] study sets of MSCs whose linear extensions form a regular string language. I will call these sets of MSCs "recognizable". The notion of recognizability has proven to be robust and fruitful in different settings like strings, trees, Mazurkiewicz traces and other classes of partial orders (both finite and infinite). The robustness is reflected by the fact that, in all these settings, recognizable sets can be presented by finite-state devices, by congruences of finite index, or by sentences of monadic second order logic. The main results in [10] show similar equivalences for sets of B-bounded finite MSCs (an MSC is B-bounded if in any execution, any buffer will contain at most B messages at any given time). In particular, they prove the equivalence of the following three concepts for sets K of B-bounded finite MSCs:

1. The set of linear extensions of K can be accepted by a finite deterministic automaton.
2. There is a sentence φ of the monadic second order logic such that K is the set of B-bounded finite MSCs that satisfy φ.
3. Some finite nondeterministic message passing automaton accepts K.

This result was sharpened in [20] where it is shown that deterministic message passing automata suffice.

The main focus of this paper is the extension of these results to sets of infinite MSCs. These infinite MSCs occur naturally as executions of systems that are not meant to stop, e.g., distributed operating systems or telecommunication networks. In the first part, we will extend the equivalence between the first and the second statement. We will also consider two fragments of monadic second order logic, namely first-order logic FO and its extension by modulo-counting quantifiers $\mathrm{FO}{+}\mathrm{MOD}(n)$ [22]. We describe the expressive power of these logics in the spirit of Büchi's theorem: for a set of B-bounded possibly infinite MSCs K, the following statements are equivalent (Theorems 3.3 and 4.6)

1. The set of linear extensions of K is recognizable (n-solvable, aperiodic, resp.).
2. The set K is axiomatizable by a sentence of monadic second order logic (of the logic $\mathrm{FO}{+}\mathrm{MOD}(n)$, first-order logic, resp.) relative to all possibly infinite MSCs.

The proof of the implication 1→2 relies on a first-order interpretation of a B-bounded MSC in any of its linearisations as well as on the fact that the set of all linearisations of B-bounded MSCs is aperiodic. This allows us to use results from [3, 17] and [22] that characterize the expressive power of the logics in question for infinite words. The proof 2→1 for *finite* MSCs from [10] uses a first-order interpretation of the lexicographically least linear extension of t in the finite MSC t. This proof method does not extend to the current setting since in general no linear extension of order type ω can be defined in an infinite MSC. To overcome this problem, we use ideas from [23] by choping an infinite MSC into its finite and its infinite part. It turns out that the infinite part is the disjoint union of infinite posets to which the "classical" method from [10] is applicable.

The second part of the paper gives a characterization of recognizable sets of infinite MSCs in terms of message passing automata. To this aim, we extend the model from [10] by a Muller-acceptance condition. It is shown that for a set of B-bounded possibly infinite MSCs K, the following statements are equivalent (Theorem 5.7).

1. The set of linear extensions of K is recognizable.
3. Some finite deterministic message passing automaton with Muller-acceptance condition accepts K.

The proof of the implication 3→1 is an obvious variant of similar proofs for finite automata for words (cf. [24]), asynchronous automata for traces [25], or asynchronous cellular automata for pomsets without autoconcurrency [8]. Mukund et al. proved the implication 1→3 for finite MSCs. In order to do so, they had to reprove several results from the theory of Mazurkiewicz traces in the more complex realm of MSCs. Differently, my proof for infinite MSCs refers to deep results in the theory of Mazurkiewicz traces directly, in particular to the theory of asynchronous mappings [4, 6]. These results are applicable since any recognizable set of MSCs can be represented as a set of traces up to an easy relabeling. This constitutes a newly discovered relation between Mazurkiewicz traces and MSCs that differs fundamentally from those used e.g. in [21] for the investigation of race condition and confluence properties and in [10] for some undecidability results. This new observation has in my opinion several nice aspects: (1) it simplifies the proof, (2) it also results in smaller message passing automata for finite MSCs, and (3) it highlights the similarity of MSCs and Mazurkiewicz traces and the unifying role that Mazurkiewicz traces can play in the theory of distributed systems. This last point is also stressed by the fact that similar proof techniques have been used, e.g., in [2, 7, 14, 8, 16, 19].

2 Notation

Let $\mathcal{P}$ be a finite set of processes (or agents) which communicate with each other through messages via reliable FIFO-channels. Let Σ be the set of communication actions $p!q$ and $p?q$ for $p, q \in \mathcal{P}$ distinct. The action $p!q$ is to be read as "p sends to q" and $p?q$ is to be read as "p receives from q". Hence $p\theta q$ is performed by the process p, denoted $\mathrm{proc}(p\theta q) = p$. Following [10], we shall not be concerned with the internal actions of the agents which is no essential restriction since the results of this paper can be extended to deal with internal actions. We will also not consider the actual messages that are sent and received.

A Σ-labeled poset is a structure $t = (V, \leq, \lambda)$ where $(V, \leq)$ is a partially ordered set, $\lambda : V \to \Sigma$ is a mapping, $\lambda^{-1}(\sigma) \subseteq V$ is linearly ordered for any $\sigma \in \Sigma$, and any $v \in V$ dominates a finite set. A subset $X \subseteq V$ is an *order ideal* if it is downwards closed, i.e., if $x \in X$, $v \in V$ and $v \leq x$ imply $v \in X$. For $A \subseteq \Sigma$, let $\pi_A(t)$ denote the restriction of t to $\lambda[A]$, i.e., to the A-labeled nodes. For $v \in V$, we write $\mathrm{proc}(v)$ as a shorthand for $(\mathrm{proc} \circ \lambda)(v)$. Furthermore, we define $\downarrow v = \{u \in V \mid u \leq v\}$. In order to define message sequence charts, for a Σ-labeled poset t, we define the relations $\sqsubseteq_{\mathcal{P}}$ and $\sqsubseteq$ as follows:

- $v \sqsubseteq_{\mathcal{P}} w$ iff $\mathrm{proc}(v) = \mathrm{proc}(w)$ and $v \leq w$.
- $v \sqsubseteq w$ iff $\lambda(v) = p!q$, $\lambda(w) = q?p$, and $|\downarrow v \cap \lambda^{-1}(p!q)| = |\downarrow w \cap \lambda^{-1}(q?p)|$ for some $p, q \in \mathcal{P}$ distinct.

Definition 2.1. *A* message sequence chart or MSC *for short is a Σ-labeled poset $t = (V, \leq, \lambda)$ satisfying*

- $\leq = (\sqsubseteq_{\mathcal{P}} \cup \sqsubseteq)^\star$,
- $\mathrm{proc}^{-1}(p) \subseteq V$ *is linearly ordered for any $p \in \mathcal{P}$, and*
- $|\lambda^{-1}(p!q)| = |\lambda^{-1}(q?p)|$ *for any $p, q \in \mathcal{P}$ distinct.*

*The MSC t is B-*bounded *for some $B \in \mathbb{N}$ if, for any $v \in V$, we have $|\downarrow v \cap \lambda^{-1}(p!q)| - |\downarrow v \cap \lambda^{-1}(q?p)| \leq B$. By MSC^∞, we denote the set of all message sequence charts while MSC denotes the set of finite MSCs. Furthermore, MSC_B and MSC_B^∞ denote the sets of B-bounded (finite) MSCs. Finally, $\downarrow\mathrm{MSC}$ denotes the set of order ideals in finite MSCs, and $\downarrow\mathrm{MSC}_B^\infty$ etc are defined similarly.*

Let $w = w_0 w_1 w_2 ...$ be a finite or infinite word over Σ. We write $w\lceil_i = w_0 w_1 ... w_i$ for the prefix of w of length $i+1$ where $0 \leq i < |w|$. The word w is *proper* if, for any $0 \leq i < |w|$ and any $p, q \in \mathcal{P}$ distinct, we have $|w\lceil_i|_{p!q} \geq |w\lceil_i|_{q?p}$ (where $|w|_a$ denotes the number of occurrences of the letter a in the word w). From a proper word w, we can define the Σ-labeled poset $\mathrm{MSC}(w)$ as follows:

- $V = \mathrm{dom}(w) = \{i \in \mathbb{N} \mid 0 \leq i < |w|\}$,
- $\lambda(i) = \lambda_w(i) = w_i$, and
- $\leq = R^\star$ where $(i,j) \in R$ iff $i < j$ and either $\mathrm{proc}(w_i) = \mathrm{proc}(w_j)$ or $w_i = p!q$, $w_j = q?p$ and $|w\lceil_i|_{p!q} = |w\lceil_j|_{q?p}$.

It is easily seen that $\mathrm{MSC}(w)$ is an element of $\downarrow\mathrm{MSC}^\infty$, i.e., an ideal in an MSC. The proper word w is B-bounded if, for any $0 \leq i < |w|$, we have $|w\lceil_i|_{p!q} - |w\lceil_i|_{q?p} \leq B$. Note that an MSC t is B-bounded iff any proper word w with $\mathrm{MSC}(w) = t$ is B-bounded. On the other hand, there are proper B-bounded words w for which $\mathrm{MSC}(w)$ is not B-bounded (e.g., $w = (p!q\, q?p)^\omega$). We call a proper word w *complete* provided $|w|_{p!q} = |w|_{q?p}$ for $p, q \in \mathcal{P}$ or, equivalently, if $\mathrm{MSC}(w)$ is an MSC.

For an ideal in a MSC t, let $\mathrm{Lin}_\omega(t)$ denote the set of all proper words w with $\mathrm{MSC}(w) = t$. For $K \subseteq \downarrow\mathrm{MSC}^\infty$, let $\mathrm{Lin}_\omega[K]$ denote the union of all sets $\mathrm{Lin}_\omega(t)$ for $t \in K$. Using this set of words, we define below recognizable sets of MSCs.

Recall that a set L of finite words over Σ is recognizable (by a finite deterministic automaton) iff there exists a finite monoid S and a homomorphism $\eta : \Sigma^\star \to S$ such that $L = \eta^{-1}\eta(L)$. The set L is aperiodic if S can be assumed to be aperiodic (i.e., groupfree). Finally, we call L n-solvable (for some $n \in \mathbb{N}$) if we can assume that any group in S is solvable and has order dividing some power of n. Similarly, one can define recognizable, n-solvable, and aperiodic sets of words in Σ^∞: $L \subseteq \Sigma^\infty$ is recognizable iff there exists a finite monoid S and a homomorphism $\eta : \Sigma^\star \to S$ such that, for any $u_i, v_i \in \Sigma^\star$ with $\eta(u_i) = \eta(v_i)$ and $u_0 u_1 u_2 \cdots \in L$, we get $v_0 v_1 v_2 \cdots \in L$. The set L is said to be n-solvable or aperiodic if the monoid S can be assumed to satisfy the corresponding conditions.

Definition 2.2. *A set $K \subseteq \downarrow\mathrm{MSC}^\infty$ is* recognizable, n-solvable, *or* aperiodic *if $\mathrm{Lin}_\omega[K]$ is recognizable, n-solvable, or aperiodic, respectively.*

In [10], it was shown that any recognizable set of finite MSCs is B-bounded for some $B \in \mathbb{N}$. The proof for ideals in possibly infinite MSCs goes through verbatimly (note that any n-solvable or aperiodic set is recognizable):

Proposition 2.3 (cf. [12, Prop. 3.2]). *Let $K \subseteq \downarrow\mathrm{MSC}^\infty$ be recognizable. Then K is bounded, i.e., there exists a positive integer B such that $L \subseteq \downarrow\mathrm{MSC}^\infty_B$.*

The sets of recognizable, n-solvable, and aperiodic word languages have been characterized in terms of fragments of monadic second order logic [3, 17, 22]. In the first part of this paper, these results will be extended to subsets of $\downarrow\mathrm{MSC}^\infty$.

Formulas of the monadic second order language MSO over Σ involve first order variables $x, y, z...$ for nodes and set variables $X, Y, Z...$ for sets of nodes. They are built up from the atomic formulas $\lambda(x) = \sigma$ for $\sigma \in \Sigma$, $x \leq y$, and $x \in X$ by means of the boolean connectives $\neg, \wedge$ and the quantifier $\exists$ (both for first order and for set variables).

A first order formula is a formula without set variables. Formulas without free variables are called sentences. The satisfaction relation $t \models \varphi$ between Σ-labeled posets t and formulas φ is defined canonically. Let $\mathfrak{X}$ be a set of Σ-labeled posets. A set $K \subseteq \mathfrak{X}$ is *axiomatizable relative to* $\mathfrak{X}$ iff there is a sentence φ such that $K = \{t \in \mathfrak{X} \mid t \models \varphi\}$.

Formulas of the logic $\mathrm{FO}{+}\mathrm{MOD}(n)$ are built up from the atomic formulas $\lambda(x) = a$ and $x \leq y$ be the connectives $\wedge$ and $\neg$ and the first-order quantifiers $\exists$ and $\exists^m$ for $0 \leq m < n$. A Σ-labeled poset $t = (V, \leq, \lambda)$ satisfies $\exists^m \varphi(x)$ if the number of nodes $v \in V$ such that $t \models \varphi(v)$ is finite and congruent $m \mod n$. Thus, syntactically, $\mathrm{FO}{+}\mathrm{MOD}(n)$ is not a fragment of MSO, but since MSCs have width at most $|\mathcal{P}|$, any $\mathrm{FO}{+}\mathrm{MOD}(n)$-definable set of MSCs can alternatively be defined in MSO.

3 From logics to monoids

Proposition 3.1. *The sets of B-bounded and proper (complete, resp.) words in $\Sigma^\star$ and Σ^ω are aperiodic.*

Proof. The set of B-bounded, proper and finite words L can be accepted by a finite deterministic automaton with $(B+1)^{|\mathcal{P}|^2} + 1$ states. If $uv^{B+1}w \in L$, then $|v|_{p!q} = |v|_{q?p}$ for $p, q \in \mathcal{P}$ distinct. Hence $uv^B w$ and $uv^{B+2}w$ belong to L, i.e., L is aperiodic. From this result, the remaining assertions follow easily. □

Hence the set of B-bounded, proper, and finite words is first-order axiomatizable relative to $\Sigma^\star$ [18]. Using a similar idea, one can define the relation R from the definition of $\mathrm{MSC}(w)$ in the B-bounded and proper word $w \in \Sigma^\infty$. This then allows to interpret $\mathrm{MSC}(w)$ in the B-bounded and proper word w by a first-order formula, i.e., we obtain the following result:

Proposition 3.2. *Let $B \in \mathbb{N}$. There exists a first-order formula $\varphi(x, y)$ such that for any B-bounded and proper word $w \in \Sigma^\infty$, we have*

$$(\mathrm{dom}(w), \varphi^w, \lambda_w) \cong \mathrm{MSC}(w)$$

where $\varphi^w = \{(i, j) \mid 0 \leq i, j < |w|, w \models \varphi(i, j)\}$.

Let ψ be some sentence of FO, $\mathrm{FO}{+}\mathrm{MOD}(n)$, or MSO. Replace in this sentence any subformula of the form $x \leq y$ by $\varphi(x, y)$ from the lemma above. Since FO is a fragment of all the other logics, the resulting sentence belongs to FO, $\mathrm{FO}{+}\mathrm{MOD}(n)$, or MSO, respectively. By Prop. 3.2 and 3.1, the set of proper words w with $\mathrm{MSC}(w) \models \psi$ is axiomatizable in the respective logic. Hence we can apply the algebraic characterizations of these sets from [3, 17, 22] and obtain

Theorem 3.3. *Let $K \subseteq {\downarrow}\mathrm{MSC}^\infty$ be bounded and axiomatizable by a sentence of* MSO / $\mathrm{FO}{+}\mathrm{MOD}(n)$ / FO. *Then K is recognizable / n-solvable / aperiodic, respectively.*

4 From monoids to logics

Henriksen *et al.* [10] showed that any recognizable set K of *finite* MSCs is axiomatizable in monadic second order logic relative to MSC_B for some B. Their proof strategy follows the idea from [9] to interpret the lexicographically least linear extension of an MSC t in t. This interpretation is possible for infinite MSCs, too, as the following lemma shows.

Let $\preceq$ be a fixed linear order on Σ. Let furthermore X be a subset of an MSC t. Then $\preceq$ defines the lexicographic order on the linear extensions of $t\restriction_X$. Since these linear extensions are well orders, the lexicographic order is a well order, too. Hence there is a lexicographically least linear extension of $t\restriction_X$ that we denote by $\mathrm{lexNF}(t\restriction_X)$.

Lemma 4.1. *There exists a first-order formula φ with two free variables such that for any $t = (V, \leq, \lambda) \in \downarrow\mathrm{MSC}^\infty$ and $X \subseteq V$, we have $(X, \varphi^{t\restriction_X}, \lambda\restriction_X) \cong \mathrm{lexNF}(t\restriction_X)$.*

The proof is an immediate adaptation of the corresponding proof for Mazurkiewicz traces, cf. [5].

The proof from [10] for finite MSCs then refers to Büchi's theorem for finite words. We cannot continue that way since the lexicographically least linear extension of, ***e.g.***, $\mathrm{MSC}((p!q\, q?p)^\omega)$ with $p!q \preceq q?p$, is no ω-word (it equals $(p!q)^\omega(q?p)^\omega$). Even worse, the MSC $\mathrm{MSC}((p!q\, q?p\, p'!q'\, q'?p')^\omega$ has no definable linear extension of order type ω at all. Hence we cannot generalize the proof from [10] to infinite MSCs. To overcome this problem, we will chop an MSC into pieces, consider these pieces independently, and combine the results obtained for them.

Now let $t = (V, \leq, \lambda) \in \downarrow\mathrm{MSC}^\infty$. Then $\mathrm{alph}(t) = \lambda[V]$ and $\mathrm{alphInf}(t) = \{\sigma \in \Sigma \mid \lambda^{-1}(\sigma)$ is infinite$\}$. Let Y be the largest filter in t with $\lambda[Y] \subseteq \mathrm{alphInf}(t)$ and let $X = V \setminus Y$. Then the *finitary part of* t is defined by $\mathrm{Fin}(t) = t\restriction_X$ and the *infinitary part of* t by $\mathrm{Inf}(t) = t\restriction_Y$. Then $\mathrm{Fin}(t)$ is an ideal in a finite MSC while in general $\mathrm{Inf}(t)$ is only a Σ-labeled poset.

Let $E \subseteq \Sigma^2$ contain all pairs $\sigma, \tau \in \Sigma$ with $\mathrm{proc}(\sigma) = \mathrm{proc}(\tau)$ or $\{\sigma, \tau\} = \{p!q, q?p\}$ for some $p, q \in \mathcal{P}$. Then (Σ, E) is an undirected graph.

Lemma 4.2. *Let $t = (V, \leq, \lambda) \in \downarrow\mathrm{MSC}_B^\infty$, and $A = \mathrm{alphInf}(t)$. Let $(A_i)_{1\leq i\leq n}$ be the connected components of the graph (A, E). Then we have for $1 \leq i \leq n$:*

1. *the Σ-labeled poset $t_i = \pi_{A_i}(\mathrm{Inf}(t))$ is directed (i.e., it does not contain two disjoint filters),*
2. *any linear extension of t_i is of order type ω, and*
3. $\mathrm{Inf}(t) = \dot{\bigcup}_{1\leq i\leq n} t_i$.

Note that the MSC $\mathrm{MSC}(((p!q)(q?p))^\omega)$ is directed, but it has a linear extension of type $\omega + \omega$, namely $(p!q)^\omega(q?p)^\omega$. Thus, to prove the second statement of the lemma above, one indeed needs the B-boundedness of t.

By $u \sqcup\!\sqcup v$ we denote the set of all shuffles of the words u and v and $K \sqcup\!\sqcup L = \bigcup\{u \sqcup\!\sqcup v \mid u \in L, v \in L\}$. Now suppose $t \in \downarrow\mathrm{MSC}_B^\infty$ and let $\mathrm{alphInf}(t) = A$. Then, by Lemma 4.2(3), $\mathrm{Inf}(t)$ is the disjoint union of the Σ-labeled posets $t_i = \pi_{A_i}(\mathrm{Inf}(t))$ for A_i a connected component of (A, E). We define $\mathrm{NFInf}(t) = \sqcup\!\sqcup_{1\leq i\leq n} \mathrm{lexNF}(t_i)$. Then, by Lemma 4.2(2), $\mathrm{NFInf}(t) \subseteq \Sigma^\omega$. Finally, by Lemma 4.2(2), we obtain that $\mathrm{NFInf}(t) \subseteq \mathrm{Lin}_\omega(\mathrm{Inf}(t))$.

Definition 4.3. *Let* $t \in {\downarrow}\mathrm{MSC}_B^\infty$*,* $A = \mathrm{alphInf}(t)$ *and let* $\$ \notin \Sigma$*. Then* $\mathrm{NF}(t) = \mathrm{lexNF}(\mathrm{Fin}(t)) \cdot \{\$\} \cdot \mathrm{NFInf}(t)$ *is the* set of normal forms of t*. For* $K \subseteq {\downarrow}\mathrm{MSC}_B^\infty$*, we set* $\mathrm{NF}[K] = \bigcup_{t \in K} \mathrm{NF}(t)$.

Since $\mathrm{NFInf}(t)$ is a subset of Σ^ω, we get $\mathrm{NF}(t) \subseteq (\Sigma \cup \$)^\infty$. Furthermore, the restriction of any word in $\mathrm{NF}(t)$ to Σ is a linear extension of t, i.e., $\mathrm{NF}(t) \subseteq \mathrm{Lin}_\omega(t) \sqcup\!\sqcup \{\$\}$. Using Prop. 3.1 and 3.2, one obtains that $\mathrm{NF}[{\downarrow}\mathrm{MSC}_B^\infty]$ is first-order axiomatizable relative to $(\Sigma \cup \{\$\})^\infty$.

Let $u \in \Sigma^\infty$ and $k \in \mathbb{N}$. Then the set of all first-order sentences φ of quantifier depth at most k is the *k-first-order theory of* u. A set of first-order sentences of quantifier depth at most k is a *complete k-first-order theory* if it is the k-first-order theory of some word $u \in \Sigma^\infty$. Since, up to logical equivalence, there are only finitely many first-order sentences of quantifier depth at most k, there are only finitely many complete k-first-order theories. Furthermore, each complete k-first-order theory T is characterized by one first-order sentence γ_T of quantifier depth k, i.e., for any word $u \in \Sigma^\infty$, we have $u \models \gamma$ for all $\gamma \in T$ iff $u \models \gamma_T$ (cf. [13, Thm. 3.3.2]). For notational convenience, we will identify the characterizing sentence γ_T and the complete k-first-order theory T.

Let $K \subseteq {\downarrow}\mathrm{MSC}_B^\infty$ be aperiodic. Then $\mathrm{Lin}_\omega[K]$ is first-order axiomatizable relative to Σ^∞. Let $k \geq 2$ be the least integer such that $\mathrm{Lin}_\omega[K] \sqcup\!\sqcup \{\$\}$ and $\mathrm{NF}[{\downarrow}\mathrm{MSC}_B^\infty]$ are first-order axiomatizable relative to $(\Sigma \cup \{\$\})^\infty$ by a sentence of quantifier depth at most k. Let T be a complete k-first-order theory and $A \subseteq \Sigma$. Then set $K_{T,A} := \{t \in K \mid \mathrm{lexNF}(\mathrm{Fin}(t)) \models T, \mathrm{alphInf}(t) = A\}$ and $X_{T,A} := \mathrm{NF}[K_{T,A}]$.

Lemma 4.4. *In the first order language of* $(\Sigma \cup \{\$\})$*-labeled linear orders with one constant* c*, there exists a sentence* φ *of quantifier depth* k *such that* $(v, c) \models \varphi$ *iff* $v \in X_{T,A}$ *and* $\lambda(c) = \$$ *for* $v \in (\Sigma \cup \{\$\})^\infty$*.*

Apart from an extension of Mezei's theorem to languages of infinite words we now have all the ingredients for the proof of Thm. 4.6. To formulate the mentioned extension, we need the notion of an aperiodic extension: a finite monoid S' is an *aperiodic extension* of a monoid S if there is a surjective homomorphism $\eta : S' \to S$ such that $\eta^{-1}(f)$ is an aperiodic semigroup for any idempotent element $f \in S$. Note that aperiodic extensions of finite / n-solvable / aperiodic monoids are finite / n-solvable / aperiodic.

Theorem 4.5. *Let* $\Sigma = \Sigma_1 \dot\cup \Sigma_2$ *be an alphabet. Let* $L \subseteq \Sigma^\infty$ *be recognized by a homomorphism into* $(S, \cdot)$ *and suppose that* $u_1 \sqcup\!\sqcup u_2 \cap L \neq \emptyset$ *implies* $u_1 \sqcup\!\sqcup u_2 \subseteq L$ *for any* $u_i \in \Sigma_i^\infty$ *(*$i = 1, 2$*). Then* L *is a finite union of sets* $L_1 \sqcup\!\sqcup L_2$ *where* $L_i \subseteq \Sigma_i^\infty$ *is recognized by an aperiodic extension of* $(S, \cdot)$*.*

Putting all these results (and their obvious extensions to the more expressive logics) together, one obtains the following converse of Thm. 3.3:

Theorem 4.6. *Let* $K \subseteq {\downarrow}\mathrm{MSC}^\infty$ *be recognizable /* n*-solvable / aperiodic. Then* K *is bounded and axiomatizable by a sentence of* MSO */* $\mathrm{FO}{+}\mathrm{MOD}(n)$ */* FO*, respectively.*

5 Deterministic message passing automata

In this section, we will extend the the automata-theoretic characterizations of recognizable word languages. Message passing automata, the automata model that we consider, reflect the concurrent behavior of an MSC. It was introduced by Henriksen et al. [10] and is similar to asynchronous cellular automata from the theory of Mazurkiewicz traces. We will extend results from [10, 20] to infinite MSCs. Since the proofs rely on the theory of Mazurkiewicz traces, we first investigate the relation between these traces and MSCs.

5.1 The key observation

A *dependence alphabet* is a pair (Γ, D) where Γ is a finite set and D is a reflexive and symmetric dependence relation. A *trace over* (Γ, D) is a Γ-labeled partial order $(V, \leq, \lambda')$ such that

- $(\lambda'(x), \lambda'(y)) \notin D$ whenever $x, y \in V$ are incomparable, and
- $(\lambda'(x), \lambda'(y)) \in D$ whenever y is an upper neighbor of x (denoted $x \multimap y$).

The set of all traces over (Γ, D) is denoted by $\mathbb{R}(\Gamma, D)$, the set $\mathbb{M}(\Gamma, D)$ comprises the finite traces.

The key observation that is announced by the title of this section is that any recognizable set of MSCs is the “relabeling” of a monadically axiomatizable set of traces over a suitable dependence alphabet. Recall that any recognizable set of MSCs is bounded. The bound B influences the chosen dependence alphabet as defined in the following paragraph.

For a positive integer $B \in \mathbb{N}$, let $\Gamma = \Sigma \times \{0, 1, \dots, B-1\}$. On this alphabet, we define a dependence relation D as follows: $(p_1\theta_1 q_1, n_1)$ and $(p_2\theta_2 q_2, n_2)$ are dependent iff

1. $p_1 = p_2$, or
2. $\{(p_1\theta_1 q_1, n_1), (p_2\theta_2 q_2, n_2)\} = \{(p!q, n), (q?p, n)\}$ for some $p, q \in \mathcal{P}$ and $0 \leq n < B$.

For $t = (V, \leq, \lambda) \in {\downarrow}\mathrm{MSC}_B^\infty$, we define a new Γ-labeling λ' by

$$\lambda'(v) = (\lambda(v), |{\downarrow}v \cap \lambda^{-1}\lambda(v)| \mod B),$$

i.e. the first component of the label is the old label and the second counts modulo B the number of occurrences of the same action in the past of v. We then define $\mathrm{tr}(t) = (V, \leq, \lambda')$. First, one shows that $\{\mathrm{tr}(t) \mid t \in {\downarrow}\mathrm{MSC}_B^\infty\}$ is a first-order axiomatizable set of traces in $\mathbb{R}(\Gamma, D)$:

Lemma 5.1. *The set* $\mathrm{tr}[{\downarrow}\mathrm{MSC}_B^\infty]$ *is the set of all traces* $s \in \mathbb{R}(\Gamma, D)$ *satisfying*

I. $\pi_{\{(p!q,n),(q?p,n)\}}(s)$ *is a prefix of* $((p!q, n)(q?p, n))^\omega$ *for* $p, q \in \mathcal{P}$ *and* $0 \leq n < B$.
II. $\pi_{\{(\sigma,n) \mid 1 \leq n < B\}}(s)$ *is a prefix of* $((\sigma, 1)(\sigma, 2) \dots (\sigma, B-1)(\sigma, 0))^\omega$ *for* $\sigma \in \Sigma$.
III. If $v, w \in V$ *with* $v \multimap w$, *then* $\mathrm{proc}(\lambda'(v)) = \mathrm{proc}(\lambda'(w))$ *or* $\lambda'(v) = (p!q, n)$ *and* $\lambda'(w) = (q?p, n)$ *for some* $p, q \in \mathcal{P}$ *and* $0 \leq n < B$.

Since all these properties are first-order expressible and since one can interpret the MSC t in the trace $\mathrm{tr}(t)$, we obtain

Proposition 5.2. *Let $K \subseteq {\downarrow}\mathrm{MSC}_B^\infty$ by monadically axiomatizable relative to ${\downarrow}\mathrm{MSC}_B^\infty$. Then $\mathrm{tr}[K] \subseteq \mathbb{R}(\Gamma, D)$ is monadically axiomatizable relative to $\mathbb{R}(\Gamma, D)$.*

This proposition can be used for an alternative proof of parts of Theorem 4.6; it works perfectly well for recognizable languages, can in some cases be used for n-solvable languages (if B divides some power of n), and is of no use whatsoever for aperiodic languages (since the relabeling cannot be defined in first-order logic).

So far, we transformed any monadically axiomatizable set of bounded MSCs into a monadically axiomatizable set of traces. In order to make use of this transformation, we need the following definitions and results from the theory of Mazurkiewicz traces.

Let $s = (V, \leq, \lambda')$ be a trace over (Γ, D) and let $A \subseteq \Gamma$. Then $\partial_A(t)$ is the least ideal of t such that the complementary filter does not contain any A-labeled vertex. Let $\gamma \in \Gamma$. Then $D(\gamma) = \{\delta \in \Gamma \mid (\gamma, \delta) \in D\}$. Furthermore, $t\gamma$ is the unique trace $(V \dot\cup \{\star\}, \leq', \rho)$ with $t\gamma{\restriction}_V = t$, $\rho(\star) = \gamma$, and $\star \in \max(t\gamma)$. A mapping $\mu : \mathbb{M}(\Gamma, D) \to A$ is *asynchronous* if, for any $\Delta_1, \Delta_2 \subseteq \Gamma$, any $\gamma \in \Gamma$, and any $t \in \mathbb{M}(\Gamma, D)$,

1. $\mu(\partial_{\Delta_1 \cup \Delta_2}(t))$ is completely determined by $\mu(\partial_{\Delta_1}(t))$, $\mu(\partial_{\Delta_2}(t))$, and the sets Δ_1 and Δ_2, and
2. $\mu(\partial_\gamma(t\gamma)$ is completely determined by $\mu(\partial_{D(\gamma)}(t))$ and the letter γ.

Theorem 5.3 ([25, 4]). *Let (Γ, D) be a dependence alphabet and $L \subseteq \mathbb{M}(\Gamma, D)$. Then L is monadically axiomatizable if, and only if, there exists an asynchronous mapping μ into some finite set such that $L = \mu^{-1}\mu(L)$.*

This result was used to construct a deterministic asynchronous cellular automaton that accepts a given recognizable language of finite traces. Diekert & Muscholl [6] use the same concept of an asynchronous mapping to construct a deterministic asynchronous cellular automaton with Muller acceptance condition that accepts a given recognizable set of infinite traces. In order to state their result, we need some more notations:

Let (Γ, D) be a dependence alphabet, $t = (V, \leq, \lambda') \in \mathbb{R}(\Gamma, D)$ a trace, and $\gamma \in \Gamma$. Let $\mu_\gamma^\infty(t) \subseteq A$ be the set of all $a \in A$ for which there are infinitely many nodes $v \in V$ with $\lambda'(v) = \gamma$ and $\mu(t{\restriction}_{\downarrow v}) = a$.

Theorem 5.4 ([6, 9]). *Let (Γ, D) be a dependence alphabet and $L \subseteq \mathbb{R}(\Gamma, D)$ be monadically axiomatizable. Then there exists a finite set A, a set $\mathcal{T} \subseteq \prod_{\gamma \in \Gamma} 2^A$ of Γ-tuples of subsets of A, and an asynchronous mapping $\mu : \mathbb{M}(\Gamma, D) \to A$ such that for $t \in \mathbb{R}(\Gamma, D)$, we have: $t \in L \iff (\mu_\gamma^\infty(t))_{\gamma \in \Gamma} \in \mathcal{T}$.*

5.2 The construction of deterministic message passing automata

A *message passing automaton with Muller-acceptance condition* is a structure $\mathcal{A} = ((\mathcal{A}_p)_{p \in \mathcal{P}}, \Delta, s_{in}, \mathcal{S})$ where

1. Δ is a finite set of messages,
2. each component $\mathcal{A}_p$ is of the form $(S_p, \rightarrow_p)$ where
   - S_p is a finite set of local states,
   - $\rightarrow_p \subseteq S_p \times \Sigma_p \times \Delta \times S_p$ where $\Sigma_p = \{\sigma \in \Sigma \mid \mathrm{proc}(\sigma) = p\}$ is a local transition relation,
3. $s^{in} \in \prod_{p \in \mathcal{P}} S_p$ is the global initial state, and
4. $\mathcal{S} \subseteq \prod_{p \in \mathcal{P}} 2^{S_p}$ is a Muller acceptance condition.

Let $(s, a, m, s') \in \rightarrow_p$ be a local transition of process p. Suppose a is a send event, i.e., $a = p!q$ for some process q. Then the transition (s, a, m, s') denotes that the process p can perform the action $a = p!q$ in state s; it changes its local state to s' and sends a message m into the FIFO-channel from process p to process q. By enlarging the set of messages and local states (if necessary), we can assume that $m = s'$ for any send action (in particular, $\Delta = \bigcup_{p \in \mathcal{P}} S_p$). Now suppose that $a = p?q$ is a receive action. Then the transition (s, a, m, s') denotes that the process p can change its local state from s to s' when reading the message m from the channel that connects p and q.

A message passing automaton is *deterministic* if

- $(s, p!q, m_1, s_1), (s, p!q, m_2, s_2) \in \rightarrow_p$ imply $s_1 = s_2$ and $m_1 = m_2$
- $(s, q?p, m_1, s_1), (s, q?p, m_1, s_2) \in \rightarrow_p$ imply $s_1 = s_2$.

The message passing automaton is *complete*, if

- there exists a transition $(s, p!q, m, s')$ for any $s \in S_p$ and $q \in \mathcal{P} \setminus \{p\}$
- there exists a transition $(s, q?p, m, s')$ for any $s \in S_p$, $m \in \Delta$, and $q \in \mathcal{P} \setminus \{p\}$

Let $t = (V, \leq, \lambda)$ be an ideal in an MSC and let $\mathcal{A}$ be a message passing automaton. Let furthermore $r : V \rightarrow \bigcup_{p \in \mathcal{P}} S_p$ be a mapping and $v \in V$. We define a second mapping $r^- : V \rightarrow \bigcup_{p \in \mathcal{P}} S_p$: if there is $u < v$ with $\mathrm{proc}(u) = \mathrm{proc}(v)$, let u be maximal with this property and let $r^-(v)$ denote $r(u)$. If v is the minimal event performed by the process $\mathrm{proc}(v)$, let $r^-(v) = s^{in}_{\mathrm{proc}(v)}$. Then $r^-(v)$ is the local state of process $\mathrm{proc}(v)$ *before* the execution of v; this process is in state $r(v)$ *after* performing v.

A *run* of $\mathcal{A}$ on t is a mapping $r : V \rightarrow \bigcup_{p \in \mathcal{P}} S_p$ satisfying for any $v \in V$:

1. If $\lambda(v) = p!q$, then there is a transition $(r^-(v), p!q, m, r(v))$ in $\rightarrow_p$ for some message m (which turns out to be $r(v)$ by our assumption).
2. Now let $\lambda(v) = p?q$. Since t is an ideal in an MSC, there is a unique matching node $u \in V$ with $u \sqsubseteq v$. We require that $(r^-(v), p?q, r(u), r(v)) \in \rightarrow_p$.

Let $r : V \rightarrow \bigcup_{p \in \mathcal{P}} S_p$ be a run of a Muller message passing automaton on $t = (V, \leq, \lambda) \in {\downarrow}\mathrm{MSC}^\infty$. For $p \in \mathcal{P}$, let $X_p \subseteq S_p$ be the set of all $s \in S_p$ such that, for any $v \in V$ with $\mathrm{proc}(v) = p$, there exists $w \in V$ with $v \leq w$, $\mathrm{proc}(w) = p$, and $r(w) = s$ (and $\{s^{in}_p\}$ if no such v exists). The run r is *successful* provided $(X_p)_{p \in \mathcal{P}} \in \mathcal{S}$. A set $K \subseteq {\downarrow}\mathrm{MSC}^\infty$ is *accepted by $\mathcal{A}$ relative to* $\mathfrak{X} \subseteq {\downarrow}\mathrm{MSC}^\infty$ if, for any $t \in \mathfrak{X}$, $t \in L$ iff t is accepted by $\mathcal{A}$.

Above, we associated to any monadically axiomatizable subset of ${\downarrow}\mathrm{MSC}^\infty_B$ a monadically axiomatizable set of traces. By Theorem 5.4, we therefore get an asynchronous mapping. Next, we construct a message passing automaton from an asynchronous mapping.

Proposition 5.5. *Let $\mu : \mathbb{M}(\Gamma, D) \to A$ be some asynchronous mapping into a finite set A. Then there exists a complete deterministic message passing automaton with Muller acceptance condition $\mathcal{A}$ with local state space S and a function $f : S \to A$ such that for the run r of $\mathcal{A}$ on $t = (V, \leq, \lambda) \in \downarrow\mathrm{MSC}_B^\infty$, we have $f(r(v)) = \mu(\mathrm{tr}(\downarrow v))$.*

Now one can show that relative to $\downarrow\mathrm{MSC}_B^\infty$, message passing automata and monadic second order have the same expressive power:

Proposition 5.6. *A set $K \subseteq \downarrow\mathrm{MSC}_B$ can be accepted by the message passing automaton $\mathcal{A}$ relative to $\downarrow\mathrm{MSC}_B$ iff it is monadically axiomatizable relative to $\downarrow\mathrm{MSC}_B$.*

The construction of a formula from an MPA follows the wellknown pattern of [24, 25, 8]. The other implication follows easily from Prop. 5.5 and 5.2 together with Theorem 5.4.

In order to extend the result to $\downarrow\mathrm{MSC}^\infty$, one first observes that $\downarrow\mathrm{MSC}_B$ can be accepted by a deterministic message passing automaton relative to $\downarrow\mathrm{MSC}_{B+1}$ since it is monadically axiomatizable relative to $\downarrow\mathrm{MSC}$. Note that an ideal in an MSC $t \in \downarrow\mathrm{MSC}^\infty$ is NOT B-bounded iff it contains a principal ideal that is not B-bounded. This allows to show that $\downarrow\mathrm{MSC}_B^\infty$ can be accepted by a complete deterministic message passing automaton with Muller acceptance condition. This is used in the proof of our main result:

Theorem 5.7. *Let $K \subseteq \downarrow\mathrm{MSC}^\infty$. Then K is recognizable iff it is bounded and there exists a Muller message passing automaton that accepts K.*

Proof. The set K is monadically axiomatizable and B-bounded for some B. Hence there exists a deterministic message passing automaton $\mathcal{A}_1$ that accepts K relative to $\downarrow\mathrm{MSC}_B^\infty$. So $K = L(\mathcal{A}_1) \cap \downarrow\mathrm{MSC}_B^\infty$ is the intersection of two sets that can be accepted by deterministic message passing automata. □

One can prove the corresponding statement from [20] for finite MSCs accordingly. Our construction requires $m2^{O((n^2B)^2\log(n^2B))}$ local states (where n is the number of processes $|\mathcal{P}|$, B is the bound on the MSCs in the language K, and m is the size of the syntactic monoid of $\mathrm{Lin}[K]$). Recall that the construction from [20] needed $2^{2^{O(n^2B)}m\log m}$ local states.

References

1. R. Alur and M. Yannakakis. Model checking of message sequence charts. In *CONCUR'99*, Lecture Notes in Comp. Science vol. 1664, pages 114–129. Springer, 1999.
2. A. Arnold. An extension of the notions of traces and of asynchronous automata. *Informatique Théorique et Applications*, 25:355–393, 1991.
3. J.R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.
4. R. Cori, Y. Métivier, and W. Zielonka. Asynchronous mappings and asynchronous cellular automata. *Information and Computation*, 106:159–202, 1993.
5. V. Diekert and Y. Métivier. Partial commutation and traces. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages Volume 3*, pages 457–533. Springer, 1997.

6. V. Diekert and A. Muscholl. Deterministic asynchronous automata for infinite traces. *Acta Informatica*, 31:379–397, 1994.
7. M. Droste and P. Gastin. Asynchronous cellular automata for pomsets without autoconcurrency. In *CONCUR'96*, Lecture Notes in Comp. Science vol. 1119, pages 627–638. Springer, 1996.
8. M. Droste, P. Gastin, and D. Kuske. Asynchronous cellular automata for pomsets. *Theoretical Comp. Science*, 247:1–38, 2000. (Fundamental study).
9. W. Ebinger and A. Muscholl. Logical definability on infinite traces. *Theoretical Comp. Science*, 154:67–84, 1996.
10. J.G. Henriksen, M. Mukund, K. Narayan Kumar, and P.S. Thiagarajan. Towards a theory of regular MSC languages. Technical report, BRICS RS-99-52, 1999. The results of this technical report appeared in the extended abstracts [12, 11].
11. J.G. Henriksen, M. Mukund, K. Narayan Kumar, and P.S. Thiagarajan. On message sequence graphs and finitely generated regular MSC languages. In *ICALP'00*, pages 675–686. Springer, 2000.
12. J.G. Henriksen, M. Mukund, K. Narayan Kumar, and P.S. Thiagarajan. Regular collections of message sequence charts. In *MFCS 2000*, Lecture Notes in Computer Science vol. 1893. Springer, 2000.
13. W. Hodges. *Model Theory*. Cambridge University Press, 1993.
14. D. Kuske. Asynchronous cellular automata and asynchronous automata for pomsets. In *CONCUR'98*, Lecture Notes in Comp. Science vol. 1466, pages 517–532. Springer, 1998.
15. D. Kuske. Another step towards a theory for regular MSC languages. Technical Report 2001-36, Department of Mathematics and Computer Science, University of Leicester, 2001.
16. D. Kuske and R. Morin. Pomsets for local trace languages: Recognizability, logic and Petri nets. In *CONCUR 2000*, Lecture Notes in Comp. Science vol. 1877, pages 426–411. Springer, 2000.
17. R.E. Ladner. Application of model theoretic games to discrete linear orders and finite automata. *Information and Control*, 33:281–303, 1977.
18. R. McNaughton and S. Papert. *Counter-Free Automata*. MIT Press, Cambridge, USA, 1971.
19. R. Morin. On regular message sequence chart languages and relationships to Mazurkiewicz trace theory. In *FOSSACS01*, Lecture Notes in Comp. Science vol. 2030, pages 332–346. Springer, 2001.
20. M. Mukund, K. Narayan Kumar, and M. Sohoni. Synthesizing distributed finite-state systems from MSCs. In C. Palamidessi, editor, *CONCUR 2000*, Lecture Notes in Computer Science vol. 1877, pages 521–535. Springer, 2000.
21. A. Muscholl and D. Peled. Message sequence graphs and decision problems on Mazurkiewicz traces. In *MFCS'99*, Lecture Notes in Computer Science vol. 1672, pages 81–91. Springer, 1999.
22. H. Straubing, D. Thérien, and W. Thomas. Regular languages defined with generalized quantifiers. *Information and Computation*, 118:289–301, 1995.
23. P.S. Thiagarajan and I. Walukiewicz. An expressively complete linear time temporal logic for mazurkiewicz traces. In *LICS'97*, pages 183–194. IEEE Computer Society Press, 1997.
24. W. Thomas. Automata on infinite objects. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, pages 133–191. Elsevier Science Publ. B.V., 1990.
25. W. Thomas. On logical definability of trace languages. In V. Diekert, editor, *Proceedings of a workshop of the ESPRIT BRA No 3166: Algebraic and Syntactic Methods in Computer Science (ASMICS) 1989*, Report TUM-I9002, Technical University of Munich, pages 172–182, 1990.