Chapter 7

Language Equations

7.1 Introduction

The study of language equations, that is, the investigation of solutions to systems of equations in which there are unknowns and fixed language constants, has been the subject of much research in many varied areas. In this chapter, we seek to unify previous results in the theory of language equations initially investigated by Kari [106]. The language operations we consider are shuffle and deletion along trajectories.

We first investigate language equations of the form $L_1 = X \sqcup\!\sqcup_T L_2$, or $L_1 = X \rightsquigarrow_T L_2$, where T is a fixed set of trajectories and L_1, L_2 are fixed languages. Equations of this form have previously been studied by Kari [106]. However, by the closure properties for shuffle and deletion on trajectories, we are able to claim positive decidability results when L_1, L_2 and T are regular.

We also investigate decomposition results for a certain class of trajectories. The problem of decompositions using shuffle on trajectories was initially suggested by Mateescu *et al.* [147] as a possible means of representing complex languages as a combination of simpler languages. For instance, given a language L, if $L = L_1 \sqcup\!\sqcup_T L_2$, and if the combined complexity of L_1, L_2 and T are less than the complexity of L (for some appropriate measure of complexity) then the triple $[L_1, L_2, T]$ can serve as a more compact representation of the language L. Shuffle decompositions for arbitrary shuffle $T = (0+1)^*$ was studied by Câmpeanu *et al.* [21]. When $T = 0^*1^*$, the decomposition of languages into *prime* parts, that is, into languages which cannot be further decomposed, was studied by Salomaa and Yu [176].

We conclude by investigating systems of equations using shuffle on trajectories. We obtain preliminary results showing that the invertibility of shuffle on trajectories allows some analysis of systems of equations rather than simply individual language equations.

Before continuing, we note that the equations we consider in this chapter are known as *implicit language equations* by, e.g., Leiss [131, Sect. 2.6.2]. Implicit language equations are of the form $R = \varphi$, where R is a fixed language and φ is a formula involving constant languages and unknowns connected by language operations. In contrast, *explicit language equations* are of the form $X = \varphi$, where X is an unknown. We consider some explicit language equations in Section 8.11.

7.2 Solving One-Variable Equations

We begin by examining equations with one unknown. We find positive decidability results in these cases, provided that the languages involved are regular.

7.2.1 Solving $X \sqcup\!\sqcup_T L = R$ and $X \leadsto_T L = R$

The following is a result of Kari [106, Thm. 4.6]:

Theorem 7.2.1 *Let L, R be languages over Σ and $\diamond, \star$ be two binary word operations, which are left-inverses to each other. If the equation $X \diamond L = R$ has a solution $X \subseteq \Sigma^*$, then the language*

$$R' = \overline{\overline{R} \star L}$$

is also a solution of the equation. Moreover, R' is a superset of all other solutions of the equation.

By Theorem 7.2.1, Theorem 5.8.1 and Lemma 5.3.1, we note the following corollary:

Corollary 7.2.2 *Let $T \subseteq \{0,1\}^*$. Let T, L, R be regular languages. Then it is decidable whether the equation $X \sqcup\!\sqcup_T L = R$ has a solution X.*

The idea is the same as discussed by Kari [106, Thm. 2.3]: we compute R' given in Theorem 7.2.1, and check whether R' is a solution to the desired equation. Since all languages involved are regular and the constructions are effective, we can test for equality of regular languages. Also, we note the following corollary, which is established in the same manner as Corollary 7.2.2:

Corollary 7.2.3 *Let $T \subseteq \{i,d\}^*$. Let T, L, R be regular languages. Then it is decidable whether the equation $X \rightsquigarrow_T L = R$ has a solution X.*

7.2.2 Solving $L \sqcup\!\sqcup_T X = R$ and $L \rightsquigarrow_T X = R$

The following is also a result of Kari [106, Thm. 4.2]:

Theorem 7.2.4 *Let L, R be languages over Σ and $\diamond, \star$ be two binary word operations, which are right-inverses to each other. If the equation $L \diamond X = R$ has a solution $X \subseteq \Sigma^*$, then the language*

$$R' = \overline{L \star \overline{R}}$$

is also a solution of the equation. Moreover, R' is a superset of all other solutions of the equation.

Thus, the following result is easily shown, by appealing to Theorem 5.8.2:

Corollary 7.2.5 *Let $T \subseteq \{0,1\}^*$. Let T, L, R be regular languages. Then it is decidable whether the equation $L \sqcup\!\sqcup_T X = R$ has a solution X.*

We now consider the decidability of solutions to the equation $L \rightsquigarrow_T X = R$ where T is a fixed set of trajectories, L, R are regular languages and X is unknown. We have the following result, which is an immediate corollary of Theorem 5.8.3:

Corollary 7.2.6 *Let $T \subseteq \{i,d\}^*$. Let T, L, R be regular languages. Then it is decidable whether the equation $L \rightsquigarrow_T X = R$ has a solution X.*

7.2.3 Solving $\{x\} \sqcup\!\sqcup_T L = R$

In this section, we briefly address the problem of finding solutions to equations of the form

$$\{x\} \sqcup\!\sqcup_T L = R$$

where T is a fixed regular set of trajectories, L, R are regular languages, and x is an unknown word. This is a generalization of the results of Kari [106].

Theorem 7.2.7 *Let Σ be an alphabet. Let $T \subseteq \{0,1\}^*$ be a fixed regular set of trajectories. Then for all regular languages $R, L \subseteq \Sigma^*$, it is decidable whether there exists a word $x \in \Sigma^*$ such that $\{x\} \sqcup\!\sqcup_T L = R$.*

Proof. Let $r = \min\{|y| \ : \ y \in R\}$. Given a DFA for R, it is clear that we can compute r by breadth-first search. Then note that $|z| = |x| + |y|$ for all $z \in x \sqcup\!\sqcup_T y$ (regardless of T). Thus, it is clear that if x exists satisfying $\{x\} \sqcup\!\sqcup_T L = R$, then $|x| \leq r$. Our algorithm then simply considers all words x of length at most r, and checks whether $\{x\} \sqcup\!\sqcup_T L = R$ holds. ■

7.2.4 Solving $\{x\} \rightsquigarrow_T L = R$

In this section, we are concerned with decidability of the existence of solutions to the equation

$$\{x\} \rightsquigarrow_T L = R$$

where x is a word in Σ^*, and L, R, T are regular languages. Equations of this form have previously been considered by Kari [106]. Our constructions generalize those of Kari directly.

We begin with the following technical lemma:

Lemma 7.2.8 *Let Σ be an alphabet. Then for all sets of trajectories $T \subseteq \{i, d\}^*$, and for all $R, L \subseteq \Sigma^*$, the following equality holds:*

$$\overline{(\overline{R} \sqcup\!\sqcup_{\tau^{-1}(T)} L)} = \{x \in \Sigma^* \ : \ \{x\} \rightsquigarrow_T L \subseteq R\}.$$

Proof. Let x be a word such that $\{x\} \rightsquigarrow_T L \subseteq R$, and assume, contrary to what we want to prove, that $x \in \overline{R} \sqcup\!\sqcup_{\tau^{-1}(T)} L$. Then there exist $y \in \overline{R}$, $z \in L$ and $t \in \tau^{-1}(T)$ such that $x \in y \sqcup\!\sqcup_t z$. By Theorem 5.8.1,

$$y \in x \rightsquigarrow_{\tau(t)} z.$$

As $\tau(t) \in T$, we conclude that $y \in (\{x\} \rightsquigarrow_T L) \cap \overline{R}$. Thus $\{x\} \rightsquigarrow_T L \subseteq R$ does not hold, contrary to our choice of x. Thus $x \in \overline{(\overline{R} \sqcup\!\sqcup_{\tau^{-1}(T)} L)}$.

For the reverse inclusion, let $x \in \overline{(\overline{R} \sqcup\!\sqcup_{\tau^{-1}(T)} L)}$. Further, assume that $(\{x\} \rightsquigarrow_T L) \cap \overline{R} \neq \emptyset$. In particular, there exist words $z \in L$ and $t \in T$ such that

$$x \rightsquigarrow_t z \cap \overline{R} \neq \emptyset.$$

Let y be some word in this intersection. As $y \in x \rightsquigarrow_t z$, by Theorem 5.8.1, we have that $x \in y \sqcup\!\sqcup_{\tau^{-1}(t)} z$. Thus, $x \in \overline{R} \sqcup\!\sqcup_{\tau^{-1}(T)} L$, contrary to our choice of x. This proves the result. ■

Thus, we can state the main result of this section:

Theorem 7.2.9 *Let Σ be an alphabet. Let $T \subseteq \{i, d\}^*$ be an arbitrary regular set of trajectories. Then the problem "Does there exist a word x such that $\{x\} \rightsquigarrow_T L = R$" is decidable for regular languages L, R.*

Proof. Let L, R be regular languages. We note that if R is infinite, then the answer to our problem is no; there can only be finitely many deletions along the set of trajectories T from a finite word x. Thus, assume that R is finite. Then we can construct the following regular language:

$$P = \overline{(\overline{R} \sqcup\!\sqcup_{\tau^{-1}(T)} L)} - \bigcup_{S \subsetneq R} \overline{(\overline{S} \sqcup\!\sqcup_{\tau^{-1}(T)} L)}.$$

Note that $\subsetneq$ denotes proper inclusion. We claim that $P = \{x : \{x\} \rightsquigarrow_T L = R\}$.

Assume $x \in P$. Then by Lemma 7.2.8, we have that

$$x \in \{x : \{x\} \rightsquigarrow_T L \subseteq R\}, \quad (7.1)$$

$$x \notin \{x : \{x\} \rightsquigarrow_T L \subseteq S \subsetneq R\}. \quad (7.2)$$

Thus, we must have that $\{x\} \rightsquigarrow_T L = R$, since $\{x\} \rightsquigarrow_T L$ is a subset of R, but is not contained in any proper subset of R.

Similarly, if $\{x\} \rightsquigarrow_T L = R$, then by Lemma 7.2.8, we have that $x \in \overline{(\overline{R} \sqcup_{\tau^{-1}(T)} L)}$. But as $\{x\} \rightsquigarrow_T L$ is not contained in any S with $S \subsetneq R$, we have that $x \notin \bigcup_{S \subsetneq R} \overline{(\overline{S} \sqcup_{\tau^{-1}(T)} L)}$. Thus, $x \in P$.

Thus, if R is finite, to decide if a word x exists satisfying $\{x\} \rightsquigarrow_T L = R$, we construct P and test if $P \neq \emptyset$. Since P will be regular, this can be done effectively (as we have noted, if R is infinite, we answer no). ■

7.3 Decidability of Shuffle Decompositions

Say that a language L has a *non-trivial shuffle decomposition* with respect to a set of trajectories $T \subseteq \{0, 1\}^*$ if there exist $X_1, X_2 \neq \{\epsilon\}$ such that $L = X_1 \sqcup_T X_2$.

In this section, we are concerned with giving a class of sets of trajectories $T \subseteq \{0, 1\}^*$ such that it is decidable, given a regular language R, whether R has a non-trivial shuffle decomposition with respect to T. For $T = (0 + 1)^*$, this is an open problem [21, 75]. While we do not settle this open problem, we establish a unified generalization of the results of Kari and Kari and Thierrin [105, 106, 114, 117], which leads to a large class of examples of sets of trajectories where the shuffle decomposition problem is decidable.

Our focus will be on letter-bounded regular sets of trajectories, which we studied in Section 5.5. We will require the following result of Ginsburg and Spanier [54] on bounded regular languages:

Theorem 7.3.1 *Let $L \subseteq w_1^* w_2^* \cdots w_n^*$ be a regular language. Then there exist $N \geq 1$, and $b_{j,k}, c_{j,k} \in \mathbb{N}$ for all $1 \leq j \leq N$ and $1 \leq j \leq n$ such that*

$$L = \bigcup_{j=1}^{N} w_1^{b_{j,1}} (w_1^{c_{j,1}})^* \cdots w_n^{b_{j,n}} (w_n^{c_{j,n}})^*. \tag{7.3}$$

From results due to Ginsburg and Spanier (see Ginsburg [50, Thm. 5.5.2]) and Szilard *et al.* [190, Thm. 2], we have the following result:

Corollary 7.3.2 *Let $L \subseteq \Sigma^*$ be a bounded regular language. Then we can effectively compute $w_1, \ldots, w_n \in \Sigma^*$, $N \geq 1$ and $b_{j,k}, c_{j,k} \in \mathbb{N}$ for all $1 \leq j \leq N$ and $1 \leq k \leq n$ such that (7.3) holds.*

We will also require the following observation:

Lemma 7.3.3 *Let $T \subseteq \{i, d\}^*$ be a letter-bounded regular set of trajectories. Then T is a finite union of i-regular sets of trajectories.*

Proof. Let $m \geq 0$ and $T \subseteq (i^*d^*)^m i^*$. Then by Theorem 7.3.1, there exist $N \geq 1$, $b_{j,k}, c_{j,k} \in \mathbb{N}$ with $1 \leq j \leq N$ and $1 \leq k \leq 2m+1$ such that

$$T = \bigcup_{j=1}^{N} \left(\prod_{k=1}^{m} i^{b_{j,2k-1}} (i^{c_{j,2k-1}})^* d^{b_{j,2k}} (d^{c_{j,2k}})^* \right) i^{b_{j,2m+1}} (i^{c_{j,2m+1}})^*.$$

Let $T_j = \left(\prod_{k=1}^{m} \#_k (d^{b_{j,2k}} (d^{c_{j,2k}})^*) \right) \#_{m+1}$ for all $1 \leq j \leq N$. Let φ_j be defined by $\varphi_j(d) = \{d\}$ and $\varphi_j(\#_k) = i^{b_{j,2k-1}} (i^{c_{j,2k-1}})^*$ for all $1 \leq j \leq m+1$. Then note that $T = \bigcup_{j=1}^{N} \varphi_j(T_j)$. The result thus holds, as $\varphi_j(T_j)$ is i-regular for all $1 \leq j \leq N$. ■

We first require a small detour to demonstrate that given a letter-bounded regular set of trajectories, we can compute m such that $T \subseteq (i^*d^*)^m i^*$ (such an m necessarily exists, as is easily observed).

Lemma 7.3.4 *Let $T \subseteq \{i, d\}^*$ be a letter-bounded regular set of trajectories. Suppose that $T \subseteq w_1^* \cdots w_n^*$ where $w_j \in \{i, d\}^*$, with natural numbers $N, b_{j,k}, c_{j,k}$ with $1 \leq j \leq N$ and $1 \leq k \leq n$ such that*

$$T = \bigcup_{j=1}^{N} w_1^{b_{j,1}} (w_1^{c_{j,1}})^* \cdots w_n^{b_{j,n}} (w_n^{c_{j,n}})^*. \tag{7.4}$$

If $w_\ell \notin i^* + d^*$ for some $1 \leq \ell \leq n$, then $c_{j,\ell} = 0$ for all $1 \leq j \leq N$.

Proof. Suppose that $w_\ell \notin i^* + d^*$ and that there exists j with $1 \leq j \leq N$ and $c_{j,\ell} \neq 0$. Then there exist $u, v \in \{i, d\}^*$ such that $u(w_\ell^{c_{j,\ell}})^* v \subseteq T$. Therefore, for any natural number m, we can choose a word x in T such that more than m blocks of occurrences of i (resp., d) are separated by blocks of occurrences of d (resp., i). Thus, we cannot have that $T \subseteq (i^*d^*)^m i^*$ for any m and, from this, we can easily see that T is not letter-bounded. ■

The following observation is also useful:

Fact 7.3.5 *Let $w = w_1 \cdots w_n \in \Sigma^*$ be a word with $w_i \in \Sigma$. Then for all $i \geq 0$, the following inclusion holds:*

$$w^{\leq i} \subseteq (\prod_{i=1}^{n} w_i^*)^i.$$

In particular, any finite subset of w^* is letter-bounded.

Theorem 7.3.6 *Let $T \subseteq \{i,d\}^*$ be a letter-bounded regular set of trajectories. Then we can effectively calculate $m \geq 1$ such that $T \subseteq (i^*d^*)^m i^*$.*

Proof. As $T \subseteq \{i,d\}^*$ is bounded and regular, by Corollary 7.3.2, we can effectively determine $w_1, \ldots, w_n \in \{i,d\}^*$ such that $T \subseteq w_1^* w_2^* \cdots w_n^*$, $N \geq 1$, and $b_{j,k}, c_{j,k}$ for all $1 \leq j \leq N$ and $1 \leq k \leq n$ such that

$$T = \bigcup_{j=1}^{N} w_1^{b_{j,1}} (w_1^{c_{j,1}})^* \cdots w_n^{b_{j,n}} (w_n^{c_{j,n}})^*. \tag{7.5}$$

If $w_j \in i^* + d^*$ for all $1 \leq j \leq n$, then we can easily find an m to satisfy our conditions. Suppose $w_j \notin i^* + d^*$ for some $1 \leq j \leq n$. Let $S = \{j \ : \ 1 \leq j \leq n, w_j \in i^* + d^*\}$. Then by Lemma 7.3.4, if $k \notin S$ then $c_{j,k} = 0$ for all $1 \leq j \leq N$.

Thus, we can effectively determine, for all $k \notin S$, $\alpha_k = \max\{b_{j,k} \ : \ 1 \leq j \leq N\}$. Now we note that, using Fact 7.3.5,

$$m \leq \left(\sum_{j \notin S} \alpha_j \cdot |w_j| \right) + |S| + 2$$

(the last term reflects the possibility of needing to change an expression of the form $T \subseteq (d^*i^*)^k d^*$ to $T \subseteq (i^*d^*)^{k+1} i^*$). Thus, we can test, for all m in this range, the resulting (decidable) inclusion $T \subseteq (i^*d^*)^m i^*$. ■

We now return to our investigations of shuffle decompositions. We have the following corollary of Theorem 5.5.1.

Corollary 7.3.7 *Let $T \subseteq \{i,d\}^*$ be a letter-bounded regular set of trajectories. Then for all regular languages R, there are only finitely many regular languages L' such that $L' = R \rightsquigarrow_T L$ for some* *language L. Furthermore, given effective constructions for T and R, we can effectively construct a finite set $\mathcal{S}$ of regular languages such that if $L' = R \rightsquigarrow_T L$ for some language L, then $L' \in \mathcal{S}$.*

Proof. Let R be a regular language accepted by a DFA $M = (Q, \Sigma, \delta, q_0, F)$. Let $T \subseteq (i^*d^*)^m i^*$ for some $m \geq 0$ be a regular set of trajectories. By Theorem 7.3.6, such an m is computable. Then by Lemma 7.3.3 and Corollary 7.3.2, there exist $n \geq 0$ and $T_i \in \mathfrak{I}$ for $1 \leq i \leq n$ such that $T = \cup_{i=1}^n T_i$. By (5.5), we know that if $Q_R(T_i, L) = Q_R(T_i, L')$, then $R \rightsquigarrow_{T_i} L = R \rightsquigarrow_{T_i} L'$ for all $1 \leq i \leq n$.

Note that, for all $L \subseteq \Sigma^*$ and all $1 \leq i \leq n$, $Q_R(T_i, L) \subseteq Q^{2m}$. As Q^{2m} is a finite set, there are only finitely many languages of the form $R \rightsquigarrow_{T_i} L$. This set can be obtained by considering all possible choices of sets $Q' \subseteq Q^{2m}$, and constructing the regular language from (5.5) with $Q' = Q_R(T_i, L)$ (duplicates may also then be removed, as we can compare the resulting regular languages).

Let $\mathcal{S}_i$ be the finite set of regular languages of the form $R \rightsquigarrow_{T_i} L$. As

$$R \rightsquigarrow_T L = \bigcup_{i=1}^{n} R \rightsquigarrow_{T_i} L,$$

we have that if L' is of the form $L' = R \rightsquigarrow_T L$, then $L' = \cup_{i=1}^n L_i$ where $L_i \in \mathcal{S}_i$ for all $1 \leq i \leq n$. There are again only finitely many languages in $\{\cup_{i=1}^n L_i \ : \ L_i \in \mathcal{S}_i\}$. This establishes the result. ■

Theorem 7.3.8 *Let $T \subseteq \{0, 1\}^*$ be a letter-bounded regular set of trajectories. Let R be a regular language over an alphabet Σ. Then there exists a natural number $n \geq 1$ such that there are n distinct regular languages Y_i with $1 \leq i \leq n$ such that for any $L \subseteq \Sigma^*$ the following are equivalent:*

(a) there exists a solution $Y \subseteq \Sigma^*$ to the equation $L \sqcup\!\sqcup_T Y = R$;

(b) there exists an index i with $1 \leq i \leq n$ such that $L \sqcup\!\sqcup_T Y_i = R$.

The languages Y_i can be effectively constructed, given effective constructions for T and R. Further, if Y is a solution to $L \sqcup\!\sqcup_T Y = R$, then there is some $1 \leq i \leq n$ such that $Y \subseteq Y_i$.

Proof. Let T, R be given. Let $\mathcal{S}_1(T, R)$ be the finite set of languages of the form $\overline{R} \rightsquigarrow_{\pi(T)} L$ for some $L \subseteq \Sigma^*$. This set is finite and effectively constructible by Corollary 7.3.7. Let $\mathcal{S}(T, R) =$ co-$\mathcal{S}_1(T, R)$.

Let L be arbitrary. Thus, if $L \sqcup\!\sqcup_T Y = R$, then $Y \subseteq X$ for some $X \in \mathcal{S}(T, R)$ by Theorems 5.8.2 and 7.2.4, and $L \sqcup\!\sqcup_T X = R$. Further, each language in $\mathcal{S}(T, R)$ is regular, by Corollary 7.3.7. Thus, (a) implies (b). The implication (b) implies (a) is trivial. ■

The symmetric result also holds:

Theorem 7.3.9 *Let $T \subseteq \{0, 1\}^*$ be a letter-bounded regular set of trajectories. Let R be a regular language over an alphabet Σ. Then there exists a natural number $n \geq 1$ such that there are n distinct regular languages Z_i with $1 \leq i \leq n$ such that for any $L \subseteq \Sigma^*$ the following are equivalent:*

(a) there exists a solution $Z \subseteq \Sigma^*$ to the equation $Z \sqcup\!\sqcup_T L = R$;

(b) there exists an index i with $1 \leq i \leq n$ such that $Z_i \sqcup\!\sqcup_T L = R$.

The languages Z_i can be effectively constructed, given effective constructions for T and R. Further, if Z is a solution to $Z \sqcup\!\sqcup_T L = R$, then there is some $1 \leq i \leq n$ such that $Z \subseteq Z_i$.

We can now give the main result of this section, which states that the shuffle decomposition problem is decidable for letter-bounded regular sets of trajectories:

Theorem 7.3.10 *Let $T \subseteq \{0, 1\}^*$ be a letter-bounded regular set of trajectories. Then given a regular language R, it is decidable whether there exist X_1, X_2 such that $X_1 \sqcup\!\sqcup_T X_2 = R$.*

Proof. Let $\mathcal{S}(T, R)$ be the set of languages described by Theorem 7.3.8 and, analogously, let $\mathcal{T}(T, R)$ be the set of languages described by Theorem 7.3.9.

We now note the result follows since if $X_1 \sqcup\!\sqcup_T X_2 = R$ has a solution $[X_1, X_2]$, it also has a solution in $\mathcal{S}(T, R) \times \mathcal{T}(T, R)$, since $\sqcup\!\sqcup_T$ is monotone. Thus, we simply test all the finite (non-trivial) pairs in $\mathcal{S}(T, R) \times \mathcal{T}(T, R)$ for the desired equality. ■

This result was known for catenation, $T = 0^*1^*$ (see, e.g., Kari and Thierrin [117]). However, it also holds for, e.g., the following operations: insertion ($0^*1^*0^*$), k-insertion ($0^*1^*0^{\le k}$ for fixed $k \ge 0$), and bi-catenation ($1^*0^* + 0^*1^*$).

We also note that if the equation $X_1 \sqcup\!\sqcup_T X_2 = R$ has a solution, where R is a regular language and T is a letter-bounded regular set of trajectories, then the equation also has solution $Y_1 \sqcup\!\sqcup_T Y_2 = R$ where Y_1, Y_2 are regular languages. This result is well-known for $T = 0^*1^*$ (see, e.g., Choffrut and Karhumäki [25]). For $T = (0+1)^*$, this problem is open [21, Sect. 7].

7.3.1 1-thin sets of trajectories

Recall that a language L is *1-thin* if $|L \cap \Sigma^n| \le 1$ for all $n \ge 0$. We now prove that if $T \subseteq \{0,1\}^*$ is a fixed 1-thin set of trajectories, given a regular language R, it is decidable whether R has a shuffle decomposition with respect to T.

Define the right-useful solutions to $L \sqcup\!\sqcup_T X = R$ as

$$use_T^{(r)}(X; L) = \{x \in X \;:\; L \sqcup\!\sqcup_T x \neq \emptyset\}. \tag{7.6}$$

The left-useful solutions, denoted $use_T^{(\ell)}(X; L)$, are defined similarly for the equation $X \sqcup\!\sqcup_T L = R$.

Theorem 7.3.11 *Let $T \subseteq \{0,1\}^*$ be a 1-thin regular set of trajectories. Given a regular language R, it is decidable whether R has a shuffle decomposition with respect to T.*

Proof. Let $L_1 = R \rightsquigarrow_{\tau(T)} \Sigma^*$ and $L_2 = R \rightsquigarrow_{\pi(T)} \Sigma^*$. Then we claim that

$$\exists X_1, X_2 \text{ such that } R = X_1 \sqcup\!\sqcup_T X_2 \iff L_1 \sqcup\!\sqcup_T L_2 = R. \tag{7.7}$$

The right-to-left implication is trivial. To prove the reverse implication, we first show that if $X_1 \sqcup\!\sqcup_T X_2 = R$, then $use_T^{(\ell)}(X_1; X_2) \subseteq L_1$ and $use_T^{(r)}(X_2; X_1) \subseteq L_2$.

We show only that $use_T^{(\ell)}(X_1; X_2) \subseteq L_1$. The other inclusion is proven similarly. Let $x \in use_T^{(\ell)}(X_1; X_2)$. Then there is some $y \in X_2$ such that $x \sqcup\!\sqcup_T y \neq \emptyset$. As $X_1 \sqcup\!\sqcup_T X_2 = R$, we must have that $z \in R$ for all $z \in x \sqcup\!\sqcup_T y$. Thus, by Theorem 5.8.1, $x \in z \rightsquigarrow_{\tau(T)} y \subseteq L_1$. The inclusion is proven. Thus,

$$R = X_1 \sqcup\!\sqcup_T X_2 = use_T^{(\ell)}(X_1; X_2) \sqcup\!\sqcup_T use_T^{(r)}(X_2; X_1) \subseteq L_1 \sqcup\!\sqcup_T L_2.$$

To conclude the proof, we need only establish the inclusion $L_1 \sqcup\!\sqcup_T L_2 \subseteq R$.

Let $x \in L_1$. Thus, there exist $\alpha \in R$, $\beta \in \Sigma^*$ and $t \in T$ such that $x \in \alpha \rightsquigarrow_t \beta$. Thus, $\{\alpha\} = x \sqcup\!\sqcup_t \beta$. Now, as $\alpha \in R = X_1 \sqcup\!\sqcup_T X_2$, there is some $x_1 \in X_1$, $x_2 \in X_2$ and $t' \in T$ such that $\{\alpha\} = x_1 \sqcup\!\sqcup_{t'} x_2$.

Consider now that $|t| = |\alpha| = |t'|$. As T is 1-thin, this implies that $t = t'$. Thus,

$$x \sqcup\!\sqcup_t \beta = x_1 \sqcup\!\sqcup_t x_2,$$

from which it is clear that $x = x_1$ and $x_2 = \beta$. Thus, $x \in X_1$. A similar argument establishes that $L_2 \subseteq X_2$. Therefore $L_1 \sqcup\!\sqcup_T L_2 \subseteq X_1 \sqcup\!\sqcup_T X_2 = R$. Thus, we have established that $R = L_1 \sqcup\!\sqcup_T L_2$ and (7.7) holds. The useful solutions are nontrivial iff $L_1, L_2 \neq \{\epsilon\}$. ■

We note that Theorem 7.3.10 and Theorem 7.3.11 do not apply to all sets of trajectories. Thus, to our knowledge, the question of the decidability of the existence of solutions to $R = X_1 \sqcup\!\sqcup_T X_2$ for a given regular language R is still open in the following cases (for details on literal and initial literal shuffle, see Berard [16]):

(a) arbitrary shuffle: $T = (0+1)^*$;

(b) literal shuffle: $T = (0^* + 1^*)(01)^*(0^* + 1^*)$;

(c) initial literal shuffle: $T = (01)^*(0^* + 1^*)$.

7.4 Solving Quadratic Equations

Let $T \subseteq \{0, 1\}^*$ be a letter-bounded regular set of trajectories. We can also consider solutions X to the equation $X \sqcup\!\sqcup_T X = R$, for regular languages R. This is a generalization of a result due to Kari and Thierrin [114].

Theorem 7.4.1 *Fix a letter-bounded regular set of trajectories $T \subseteq \{0,1\}^*$. Then it is decidable whether there exists a solution X to the equation $X ⧢_T X = R$ for a given regular language R.*

Proof. Let $\mathcal{S}(T, R)$ be the set of languages described by Theorem 7.3.8, and, analogously, let $\mathcal{T}(T, R)$ be the set of languages described by Theorem 7.3.9.

Assume the equation $X ⧢_T X = R$ has a solution. Then we claim that it also has a regular solution. Let X be a language such that $X ⧢_T X = R$. Then, in particular, X is a solution to the equation $X ⧢_T Y = R$, where X is fixed and Y is a variable. Thus, by Theorem 7.3.8, there is some regular language $Y_i \in \mathcal{S}(T, R)$ such that $X ⧢_T Y_i = R$. Further, $X \subseteq Y_i$. Analogously, considering the equation $X ⧢_T Y_i = R$, $X \subseteq Z_j$ for some regular language $Z_j \in \mathcal{T}(T, R)$. Thus, $X \subseteq Y_i \cap Z_j$, and $Z_j ⧢_T Y_i = R$.

Let $X_0 = Y_i \cap Z_j$. Then note that $R = X ⧢_T X \subseteq X_0 ⧢_T X_0 \subseteq Z_j ⧢_T Y_i = R$. The inclusions follow by the monotonicity of $⧢_T$. Thus, $X_0 ⧢_T X_0 = R$. By construction, X_0 is regular.

Thus, to decide whether there exists X such that $X ⧢_T X = R$, we construct the set

$$\mathcal{U}(T, R) = \{Y_i \cap Z_j \ : \ Y_i \in \mathcal{S}(T, R), Z_j \in \mathcal{T}(T, R)\},$$

and test each language for equality. If a solution exists, we answer yes. Otherwise, we answer no.
■

7.5 Existence of Trajectories

In this section, we consider the following problem: given languages L_1, L_2 and R, does there exist a set of trajectories T such that $L_1 ⧢_T L_2 = R$? We prove this to be decidable when L_1, L_2, R are regular languages.

Theorem 7.5.1 *Let $L_1, L_2, R \subseteq \Sigma^*$ be regular languages. Then it is decidable whether there exists a set $T \subseteq \{0,1\}^*$ of trajectories such that $L_1 ⧢_T L_2 = R$.*

Proof. Let

$$T_0 = \{t \in \{0,1\}^* \ : \ \forall x \in L_1, y \in L_2, x \sqcup\!\sqcup_t y \subseteq R\}. \quad (7.8)$$

Note that the following are equivalent definitions of T_0:

$$T_0 = \{t \in \{0,1\}^* \ : \ \forall x \in L_1, y \in L_2, (x \sqcup\!\sqcup_t y \neq \emptyset \Rightarrow x \sqcup\!\sqcup_t y \subseteq R)\}; \quad (7.9)$$

$$T_0 = \{t \in \{0,1\}^* \ : \ \forall x \in L_1 \cap \Sigma^{|t|_0}, y \in L_2 \cap \Sigma^{|t|_1}, (x \sqcup\!\sqcup_t y \subseteq R)\}. \quad (7.10)$$

Then we claim that

$$\exists T \subseteq \{0,1\}^* \text{ such that } (L_1 \sqcup\!\sqcup_T L_2 = R) \iff L_1 \sqcup\!\sqcup_{T_0} L_2 = R.$$

The right-to-left implication is trivial. Assume that there is some $T \subseteq \{0,1\}^*$ such that $L_1 \sqcup\!\sqcup_T L_2 = R$. Let $t \in T$. Then for all $x \in L_1$ and $y \in L_2$, $x \sqcup\!\sqcup_t y \subseteq L_1 \sqcup\!\sqcup_T L_2 = R$. Thus, $t \in T_0$ by definition, and $T_0 \supseteq T$.

Thus, note that $R = L_1 \sqcup\!\sqcup_T L_2 \subseteq L_1 \sqcup\!\sqcup_{T_0} L_2$. It remains to establish that $L_1 \sqcup\!\sqcup_{T_0} L_2 \subseteq R$. But this is clear from the definition of T_0. Thus $L_1 \sqcup\!\sqcup_{T_0} L_2 = R$ and the claim is established.

We now establish that T_0 is regular and effectively constructible; to do this, we establish instead that $\overline{T_0} = \{0,1\}^* - T_0$ is regular.

Let $M_j = (Q_j, \Sigma, \delta_j, q_j, F_j)$ be a complete DFA accepting L_j for $j = 1, 2$. Let $M_r = (Q_r, \Sigma, \delta_r, q_r, F_r)$ be a complete DFA accepting R. Define an NFA $M = (Q, \{0,1\}, \delta, q_0, F)$ where $Q = Q_1 \times Q_2 \times Q_r$, $q_0 = [q_1, q_2, q_r]$, $F = F_1 \times F_2 \times (Q_r - F_r)$, and δ is defined as follows:

$$\delta([q_j, q_k, q_\ell], 0) = \{[\delta_1(q_j, a), q_k, \delta_r(q_\ell, a)] \ : \ a \in \Sigma\} \quad \forall [q_j, q_k, q_\ell] \in Q_1 \times Q_2 \times Q_r,$$

$$\delta([q_j, q_k, q_\ell], 1) = \{[q_j, \delta_2(q_k, a), \delta_r(q_\ell, a)] \ : \ a \in \Sigma\} \quad \forall [q_j, q_k, q_\ell] \in Q_1 \times Q_2 \times Q_r.$$

Then we note that δ has the following property: for all $t \in \{0,1\}^*$,

$$\delta([q_1, q_2, q_r], t) = \{[\delta(q_1, x), \delta(q_2, y), \delta(q_r, x \sqcup\!\sqcup_t y)] \ : \ x, y \in \Sigma^*, |x| = |t|_0, |y| = |t|_1\}.$$

By (7.10), if $t \in \overline{T_0}$ there is some $x, y \in \Sigma^*$ such that $x \in L_1$, $y \in L_2$, $|x| = |t|_0$, $|y| = |t|_1$ but $x \sqcup\!\sqcup_t y \cap \overline{R} \neq \emptyset$. This is exactly what is reflected by the choice of F. Thus, $L(M) = \overline{T_0}$.

Thus, as T_0 is effectively regular, to determine whether there exists T such that $L_1 \sqcup\!\sqcup_T L_2 = R$, we construct T_0 and test $L_1 \sqcup\!\sqcup_{T_0} L_2 = R$. ■

Note that the proof of Theorem 7.5.1 is similar in theme to the proofs of, e.g., Kari [106, Thm. 4.2, Thm. 4.6]: they each construct a maximal solution to an equation, and that solution is regular. The maximal solution is then tested as a possible solution to the equation to determine if any solutions exist. However, unlike the results of Kari, Theorem 7.5.1 does not use the concept of an inverse operation.

We can also repeat Theorem 7.5.1 for the case of deletion along trajectories. The results are identical, with the proof following by the substitution of $T_0 = \{t \in \{i,d\}^* : \forall x \in L_1, y \in L_2, x \rightsquigarrow_t y \subseteq R\}$. The proof that T_0 is regular differs slightly from that above; we leave the construction to the reader. Thus, we have the following result:

Theorem 7.5.2 *Let $L_1, L_2, R \subseteq \Sigma^*$ be regular languages. Then it is decidable whether there exists a set $T \subseteq \{i,d\}^*$ of trajectories such that $L_1 \rightsquigarrow_T L_2 = R$.*

7.6 Undecidability of One-Variable Equations

We now turn to establishing undecidability results for equations with one unknown. We focus on the case where one of the remaining languages is an LCFL, and the other is a regular language.

Let $\Pi_0, \Pi_1 : \{0,1\}^* \to \{0,1\}^*$ be the projections given by $\Pi_0(0) = 0$, $\Pi_0(1) = \epsilon$ and $\Pi_1(1) = 1$, $\Pi_1(0) = \epsilon$. We say that $T \subseteq \{0,1\}^*$ is *left-enabling* (resp., *right-enabling*) if $\Pi_0(T) = 0^*$ (resp., $\Pi_1(T) = 1^*$).

We first show that if a set of trajectories is regular and left- or right-enabling, then it is undecidable whether the corresponding equation has a solution:

Theorem 7.6.1 *Fix $T \subseteq \{0,1\}^*$ to be a regular set of left-enabling (resp., right-enabling) trajectories. For a given LCFL L and regular language R, it is undecidable whether or not $L \sqcup\!\sqcup_T X = R$ (resp., $X \sqcup\!\sqcup_T L = R$) has a solution X.*

Proof. Let T be left-enabling. Let Σ be an alphabet of size at least two and let $\#, \$ \notin \Sigma$. Let $R = (\Sigma^+ + \#^+) \sqcup\!\sqcup_T \$^*$. By the closure properties of $\sqcup\!\sqcup_T$, and the fact that T is regular, R is a regular language. Let $L \subseteq \Sigma^+$ be an arbitrary LCFL and $L_\# = L + \#^+$. Note that $L_\#$ is an LCFL. We claim that

$$L_\# \sqcup\!\sqcup_T X = R \text{ has a solution } \iff L = \Sigma^+. \tag{7.11}$$

This will establish the result, since it is undecidable whether an arbitrary LCFL $L \subseteq \Sigma^+$ satisfies $L = \Sigma^+$.

First, if $L = \Sigma^+$, then note that $X = \$^*$ is a solution for $L_\# \sqcup\!\sqcup_T X = R$. Second, assume that X is a solution for $L_\# \sqcup\!\sqcup_T X = R$. It is clear that for all X,

$$L_\# \sqcup\!\sqcup_T X = R \iff L_\# \sqcup\!\sqcup_T use_T^{(r)}(X; L_\#) = R, \tag{7.12}$$

where $use_T^{(r)}(X, L)$ is defined by (7.6). Thus, we will focus on useful solutions to the equation $L_\# \sqcup\!\sqcup_T X = R$.

Now, we note that, assuming that $use_T^{(r)}(X, L_\#)$ is a solution to $L_\# \sqcup\!\sqcup_T use_T^{(r)}(X; L_\#) = R$, $use_T^{(r)}(X, L_\#)$ cannot contain words with letters from Σ, because words in R do not contain words with both # and letters from Σ.

In particular, let $x \in use_T^{(r)}(X, L_\#) \subseteq X$. Then there exists $y \in L_\#$ (in particular, $y \neq \epsilon$) such that $y \sqcup\!\sqcup_T x \neq \emptyset$. Consider the word $\#^{|y|}$. As y and $\#^{|y|}$ have the same length, we must have that $\#^{|y|} \sqcup\!\sqcup_T x \neq \emptyset$.

Consider any $z \in \#^{|y|} \sqcup\!\sqcup_T x \subseteq L_\# \sqcup\!\sqcup_T X$. As $|y| \neq 0$, $|z|_\# > 0$. As $L_\# \sqcup\!\sqcup_T X = R$, we must have that $z \in (\Sigma^+ + \#^+) \sqcup\!\sqcup_T \$^*$. Thus, $z \in (\# + \$)^+$, and consequently, $x \in (\# + \$)^*$. Thus, $use_T^{(r)}(X, L_\#) \subseteq (\# + \$)^*$.

Let $\Pi_\Sigma : (\Sigma \cup \{\#, \$\})^* \to \Sigma^*$ be the projection onto Σ. Now as T is left-enabling, note that

$\Pi_\Sigma(R) = \Sigma^+$, by definition of $R = (\Sigma^+ + \#^+) \sqcup_T \$^*$. Thus,

$$\begin{aligned}
\Sigma^+ &= \Pi_\Sigma(R) = \Pi_\Sigma(L_\# \sqcup_T X) \\
&= \Pi_\Sigma(L_\# \sqcup_T use_T^{(r)}(X, L_\#)) \subseteq \Pi_\Sigma(L_\# \sqcup_T (\# + \$)^*) \\
&= \Pi_\Sigma((L + \#^+) \sqcup_T (\# + \$)^*) = \Pi_\Sigma((L \sqcup_T (\# + \$)^*) + (\#^+ \sqcup_T (\# + \$)^*)) \\
&= \Pi_\Sigma(L \sqcup_T (\# + \$)^*) \\
&= L \subseteq \Sigma^+.
\end{aligned}$$

The last equality is valid since T is left-enabling, and therefore, for all $x \in L$, there is some $j \geq 0$ such that $x \sqcup_T (\$ + \#)^j \neq \emptyset$. We conclude that $L = \Sigma^+$, and thus, by (7.11), the result follows. The proof in the case that T is right-enabling is similar. ■

Say that a set $T \subseteq \{0, 1\}^*$ of trajectories if *left-preserving* (resp., *right-preserving*) if $T \supseteq 0^*$ (resp., $T \supseteq 1^*$). Note that if T is complete, then it is both left- and right-preserving.

We can give an incomparable result which removes the condition that T must be regular, but must strengthen the conditions on words in T. Namely, T must be left-preserving rather than left-enabling:

Theorem 7.6.2 *Fix $T \subseteq \{0, 1\}^*$ to be a set of left-preserving (resp., right-preserving) trajectories. Given an LCFL L and a regular language R, it is undecidable whether there exists a language X such that $L \sqcup_T X = R$ (resp., $X \sqcup_T L = R$).*

Proof. Let T be left-preserving (the proof when T is right-preserving is similar). It is clear that for all X,

$$L \sqcup_T X = R \iff L \sqcup_T use_T^{(r)}(X; L) = R.$$

Thus, we will focus on useful solutions to our equation.

Let Σ be our alphabet and $\# \notin \Sigma$. Let $L \subseteq \Sigma^+$ be an arbitrary LCFL. Let $L_\# = L + \#^+$. Note that $\epsilon \notin L_\#$ and that $L_\#$ is an LCFL. We claim that $L_\# \sqcup_T use_T^{(r)}(X; L_\#) = \Sigma^+ + \#^+$ if and only if $L = \Sigma^+$ and $use_T^{(r)}(X; L_\#) = \{\epsilon\}$.

First, assume that $L = \Sigma^+$ and $use_T^{(r)}(X; L_\#) = \{\epsilon\}$. Then $L_\# = \Sigma^+ + \#^+$ and

$$\begin{aligned} L_\# \sqcup_T X &= L_\# \sqcup_T use_T^{(r)}(X; L_\#) \\ &= (\Sigma^+ + \#^+) \sqcup_T \{\epsilon\} \\ &= (\Sigma^+ + \#^+), \end{aligned}$$

since $T \supseteq 0^*$.

Now, assume that $L_\# \sqcup_T use_T^{(r)}(X; L_\#) = \Sigma^+ + \#^+$. Let $x \in use_T^{(r)}(X; L_\#)$. Then there exists $y \in L_\#$ ($y \neq \epsilon$) such that $y \sqcup_T x \neq \emptyset$. Consider $\#^{|y|}$. As $|y| = |\#^{|y|}|$, we must have that $\#^{|y|} \sqcup_T x \neq \emptyset$.

For all $z \in \#^{|y|} \sqcup_T x \subseteq L_\# \sqcup_T use_T^{(r)}(X; L_\#)$, as $|y| \neq 0$, $|z|_\# > 0$. Further,

$$z \in L_\# \sqcup_T use_T^{(r)}(X; L_\#) = \Sigma^+ + \#^+.$$

Thus, we must have that $z \in \#^+$ and that $x \in \#^*$. Thus, $use_T^{(r)}(X; L_\#) \subseteq \#^*$.

We now show that $\epsilon \in use_T^{(r)}(X; L_\#)$. As $L_\# \sqcup_T use_T^{(r)}(X; L_\#) = \Sigma^+ + \#^+$, for all $y \in \Sigma^+$, there exist $\alpha \in L_\#$ and $\beta \in use_T^{(r)}(X; L_\#) \subseteq \#^*$ such that $y \in \alpha \sqcup_T \beta$. If $\beta \neq \epsilon$, then $|y|_\# > 0$. Thus $\alpha = y$, and $\beta = \epsilon \in use_T^{(r)}(X; L_\#)$. This also demonstrates that $\Sigma^+ \subseteq L_\#$, which implies that $L = \Sigma^+$.

It remains to show that $use_T^{(r)}(X; L_\#) = \{\epsilon\}$. Let $\#^i \in use_T^{(r)}(X; L_\#)$ for some $i > 0$. Then, there is some $y \in L_\# = \Sigma^+ + \#^+$ such that $y \sqcup_T \#^i \neq \emptyset$.

If $y \in \Sigma^+$, then for all $z \in y \sqcup_T \#^i$, $|z|_\Sigma, |z|_\# > 0$, which contradicts that $z \in \Sigma^+ + \#^+$, since $L_\# \sqcup_T use_T^{(r)}(X; L_\#) = \Sigma^+ + \#^+$. Thus, $y \in \#^+$. But then let $y' \in \Sigma^+$ be chosen so that $|y| = |y'|$. We have that $y' \in L_\#$ as well. We are thus reduced to the first case with y' and $\#^i$, and our assumption that $\#^i \in use_T^{(r)}(X; L_\#)$ is therefore false.

We have established that a (useful) solution to the equation

$$(L + \#^+) \sqcup_T X = (\Sigma^+ + \#^+)$$

exists if and only if $L = \Sigma^+$. Therefore, the existence of such solutions must be undecidable. ■

Note that as $use_T^{(r)}(X; L_\#) = \{\epsilon\}$, Theorem 7.6.2 remains undecidable even if the required (useful) language is required to be a singleton.

We also note that if R and L are interchanged in the equations of the statements of Theorem 7.6.2 or Theorem 7.6.1, the corresponding problems are still undecidable. The proofs are trivial, and are left to the reader.

7.7 Undecidability of Shuffle Decompositions

It has been shown [21] that it is undecidable whether a context-free language has a nontrivial shuffle decomposition with respect to the set of trajectories $\{0, 1\}^*$. Here we extend this result for arbitrary complete regular sets trajectories.

If T is a complete set of trajectories, then any language L has decompositions $L \sqcup\!\sqcup_T \{\epsilon\}$ and $\{\epsilon\} \sqcup\!\sqcup_T L$. Below we exclude these trivial decompositions; all other decompositions of L are said to be nontrivial.

Theorem 7.7.1 *Let T be any fixed complete regular set of trajectories. For a given context-free language L it is undecidable whether or not there exist languages $X_1, X_2 \neq \{\epsilon\}$ such that $L = X_1 \sqcup\!\sqcup_T X_2$.*

Proof. Let $P = (u_1, \ldots, u_k; v_1, \ldots, v_k)$, $k \geq 1$, $u_i, v_i \in \Sigma^*$, $i = 1, \ldots, k$, be an arbitrary PCP instance. We construct a context-free language $L(P)$ such that $L(P)$ has a nontrivial decomposition along the set of trajectories T if and only if the instance P does not have a solution.

Choose $\Omega = \Sigma \cup \{a, b, \#, \flat_1, \flat_2, \natural_1, \natural_2, \$_1, \$_2\}$, where $\{a, b, \#, \flat_1, \flat_2, \natural_1, \natural_2, \$_1, \$_2\} \cap \Sigma = \emptyset$. Let

$$L_0 = \left(\flat_1^+(\Sigma \cup \{a, b, \#\})^* \natural_1^+ \cup \flat_2^+(\Sigma \cup \{a, b, \#\})^* \natural_2^+\right) \sqcup\!\sqcup_T (\$_1^+ \cup \$_2^+). \qquad (7.13)$$

Define

$$\begin{aligned} L_1' &= \{ab^{i_1} \cdots ab^{i_m} \# u_{i_m} \cdots u_{i_1} \# \mathrm{rev}(v_{j_1}) \cdots \mathrm{rev}(v_{j_n}) \# b^{j_n} a \cdots b^{j_1} a \\ &\quad : \ i_1, \ldots, i_m, j_1, \ldots, j_n \in \{1, \ldots, k\}, \ m, n \geq 1\} \end{aligned}$$

and let

$$L_1 = L_0 - [\flat_1^+ L_1' \natural_1^+ \sqcup\!\sqcup_T \$_2^+].$$

Using the fact that T is regular, it is easy to see that a nondeterministic pushdown automaton M can verify that a given word is not in $\flat_1^+ L_1' \natural_1^+ \sqcup\!\sqcup_T \$_2^+$. On input w, using the finite state control M keeps track of the unique trajectory t (if it exists) such that $w \leadsto_{\tau(t)} \$_2^* \in \flat_1^+(\Sigma \cup \{a, b, \#\})^* \natural_1^+$ and $w \leadsto_{\pi(t)} \flat_1^+(\Sigma \cup \{a, b, \#\})^* \natural_1^+ \in \$_2^*$. If $t \notin T$, M accepts. Also if t does not exist, M accepts. Using the stack M can verify that $w \leadsto_{\tau(t)} \$_2^* \notin \flat_1^+ L_1' \natural_1^+$ by guessing where the word violates the definition of L_1'. Note that this verification can be interleaved with the computation checking whether t is in T. Since L_0 is regular, it follows that L_1 is context-free.

Define

$$\begin{aligned} L_2' &= \{ab^{i_1} \cdots ab^{i_m} \# w \# \text{rev}(w) \# b^{i_m} a \cdots b^{i_1} a \\ &\quad : \ w \in \Sigma^*, i_1, \ldots, i_m \in \{1, \ldots, k\}, \ m \geq 1\} \end{aligned}$$

and let

$$L_2 = L_0 - [\flat_1^+ L_2' \natural_1^+ \sqcup\!\sqcup_T \$_2^+].$$

As above it is seen that L_2 is context-free. It follows that also the language

$$L(P) = L_1 \cup L_2 = L_0 - [\flat_1^+ (L_1' \cap L_2') \natural_1^+ \sqcup\!\sqcup_T \$_2^+] \tag{7.14}$$

is context-free.

First consider the case where the PCP instance P does not have a solution. Now $L_1' \cap L_2' = \emptyset$ and (7.13) gives a nontrivial decomposition for $L(P) = L_0$ along the set of trajectories T.

Secondly, consider the case where the PCP instance P has a solution. This means that there exists a word

$$w_0 \in L_1' \cap L_2'. \tag{7.15}$$

For the sake of contradiction we assume that we can write

$$L(P) = X_1 \sqcup\!\sqcup_T X_2, \tag{7.16}$$

where $X_1, X_2 \neq \{\epsilon\}$.

We establish a number of properties that the languages X_1 and X_2 must necessarily satisfy. We first claim that it is not possible that

$$\text{alph}(X_1) \cap \{\flat_i, \natural_i\} \neq \emptyset \text{ and } \text{alph}(X_2) \cap \{\flat_j, \natural_j\} \neq \emptyset \tag{7.17}$$

where $\{i, j\} = \{1, 2\}$. If the above relations would hold, then the completeness of T would imply that $X_1 \sqcup\!\sqcup_T X_2$ has some word containing a letter from $\{\flat_1, \natural_1\}$ and a letter from $\{\flat_2, \natural_2\}$. This is impossible since $X_1 \sqcup\!\sqcup_T X_2 \subseteq L_0$.

Let $\Phi = \{\flat_1, \flat_2, \natural_1, \natural_2\}$. Since $L(P)$ has both words that contain letters $\flat_1$, $\natural_1$ and words that contain letters $\flat_2$, $\natural_2$, by (7.17) the only possibility is that all the letters of Φ "come from" one of the components X_1 and X_2. We assume in the following that

$$\text{alph}(X_2) \cap \Phi = \emptyset. \tag{7.18}$$

This can be done without loss of generality since the other case is completely symmetric (we can just interchange the letters 0 and 1 in T.)

Next we show that

$$\text{alph}(X_2) \cap (\Sigma \cup \{a, b, \#\}) = \emptyset. \tag{7.19}$$

Let $\Pi_\Phi : \Omega^* \to \Phi^*$ be the projection onto Φ. Since $\Pi_\Phi(L(P)) = \flat_1^+ \natural_1^+ \cup \flat_2^+ \natural_2^+$ and X_2 does not contain any letters of Φ, it follows that $\Pi_\Phi(X_1) = \flat_1^+ \natural_1^+ \cup \flat_2^+ \natural_2^+$. Thus if (7.19) does not hold, the completeness of T implies that $X_1 \sqcup\!\sqcup_T X_2$ contains words where a letter from $\Sigma \cup \{a, b, \#\}$ occurs before a letter from $\{\flat_1, \flat_2\}$ or after a letter from $\{\natural_1, \natural_2\}$. Hence (7.19) holds.

Since $X_2 \neq \{\epsilon\}$, the equations (7.18) and (7.19) imply that

$$\text{alph}(X_2) \cap \{\$_1, \$_2\} \neq \emptyset.$$

Since $L(P)$ has words with letters $\$_1$, other words with letters $\$_2$, and no words containing both letters $\$_1$, $\$_2$, using again the completeness of T it follows that

$$\text{alph}(X_1) \cap \{\$_1, \$_2\} = \emptyset. \tag{7.20}$$

Now consider the word $w_0 \in L_1' \cap L_2'$ given by (7.15). We have $\flat_i w_0 \natural_i ⧢_T \$_i \subseteq L_i$, $i = 1, 2$, and let $u_i \in \flat_i w_0 \natural_i ⧢_T \$_i$, $i = 1, 2$, be arbitrary. We can write

$$u_i = x_{i,1} ⧢_{t_i} x_{i,2}, \text{ such that } x_{i,j} \in X_j, t_i \in T, i = 1, 2, j = 1, 2.$$

By (7.18), (7.19) and (7.20) we have

$$X_1 \subseteq (\Phi \cup \Sigma \cup \{a, b, \#\})^* \text{ and } X_2 \subseteq \{\$_1, \$_2\}^*$$

and hence

$$x_{i,1} = \flat_i w_0 \natural_i, \quad x_{i,2} = \$_i, \quad i = 1, 2.$$

Now $x_{1,1} ⧢_{t_1} x_{2,2} \subseteq X_1 ⧢_T X_2$. Let $\eta = \flat_1 w_0 \natural_1 ⧢_{t_1} \$_2 \in x_{1,1} ⧢_T x_{2,2}$, Then $\eta \notin L(P)$ by the choice of w_0 and (7.14). This contradicts (7.16). ■

In the proof of Theorem 7.7.1, whenever the CFL has a nontrivial decomposition along the set of trajectories T, it has a decomposition where the component languages are, in fact, regular. This gives the following result:

Corollary 7.7.2 *Let T be any fixed complete regular set of trajectories. For a given context-free language L it is undecidable whether or not*

(a) there exist regular languages $X_1, X_2 \neq \{\epsilon\}$ such that $L = X_1 ⧢_T X_2$.

(b) there exist context-free languages $X_1, X_2 \neq \{\epsilon\}$ such that $L = X_1 ⧢_T X_2$.

7.8 Undecidability of Existence of Trajectories

We now turn to undecidability results for problems involving the existence of a set of trajectories satisfying a certain equation.

Lemma 7.8.1 *Given an LCFL L, it is undecidable whether there exists $T \subseteq \{0, 1\}^*$ such that $L ⧢_T \{\epsilon\} = \Sigma^*$ (resp., whether $\{\epsilon\} ⧢_T L = \Sigma^*$).*

Proof. We claim that

$$\exists\, T \subseteq \{0,1\}^* \text{ such that } L \sqcup\!\sqcup_T \{\epsilon\} = \Sigma^* \iff L = \Sigma^*.$$

If $L = \Sigma^*$, then $T = 0^*$ satisfies the equation. Assume that there exists T such that $L \sqcup\!\sqcup_T \{\epsilon\} = \Sigma^*$. Then for all $x \in \Sigma^*$, there exist $y \in L$ and $t \in T$ such that $x \in y \sqcup\!\sqcup_t \epsilon$. But this only happens if $x = y$ and $t = 0^{|x|}$. Thus, $x \in L$. Therefore, $L = \Sigma^*$. This establishes the first part of the lemma. The second follows on noting that $T \subseteq \{0,1\}^*$ satisfies $L \sqcup\!\sqcup_T \{\epsilon\} = \Sigma^*$ iff $sym_s(T)$ satisfies $\{\epsilon\} \sqcup\!\sqcup_{sym_s(T)} L = \Sigma^*$. ■

We will require some additional constructs before proving the remaining case. For a set $I \subseteq \mathbb{N}$, let $\Sigma^I = \{x \in \Sigma^* \ : \ |x| \in I\}$.

We show the following undecidability result:

Lemma 7.8.2 *Given an LCFL L, it is undecidable whether there exists $I \subseteq \mathbb{N}$ such that $L = \Sigma^I$.*

Proof. We appeal to Corollary 2.5.4. Let $P(L)$ be true if $L = \Sigma^I$ for some $I \subseteq \mathbb{N}$. Note that P is non-trivial, as, e.g., $P(\{a^n b^n \ : \ n \geq 0\})$ does not hold. Further, $P(\Sigma^*)$ is true, since $\Sigma^* = \Sigma^{\mathbb{N}}$ in our notation. Note that P is preserved under quotient, since if $L = \Sigma^I$ and $a \in \Sigma$ is arbitrary, then $L/a = \Sigma^{I'}$ where $I' = \{x \ : \ x+1 \in I\}$. Thus, we can apply Corollary 2.5.4 and it is undecidable whether $P(L)$ holds for an arbitrary LCFL L. ■

We note that a similar undecidability result was proven by Kari and Sosík [112, Lemma 8.1] for proving undecidability results of the following form: given $R_1, R_2 \in$ REG and $L \in$ CF, does $L \sqcup\!\sqcup_T R_1 = R_2$ hold? Necessary and sufficient conditions on regular sets of trajectories T such that the above problem is decidable are given by Kari and Sosík. The related undecidability proof given by Kari and Sosík uses a reduction from PCP.

Lemma 7.8.3 *Given an LCFL L, it is undecidable whether there exists $T \subseteq \{0,1\}^*$ such that $\Sigma^* \sqcup\!\sqcup_T \{\epsilon\} = L$.*

Proof. Let L be an LCFL. Then we claim that

$$\exists T \subseteq \{0,1\}^* \text{ such that } L = \Sigma^* \sqcup\!\sqcup_T \{\epsilon\} \iff \exists I \subseteq \mathbb{N} \text{ such that } L = \Sigma^I.$$

($\Leftarrow$): If $I \subseteq \mathbb{N}$ is such that $L = \Sigma^I$, then let $T = \{0\}^I$. Then we can easily see that $L = \Sigma^* \sqcup\!\sqcup_T \{\epsilon\}$.
($\Rightarrow$): If $T \subseteq \{0,1\}^*$ exists such that $L = \Sigma^* \sqcup\!\sqcup_T \{\epsilon\}$, then by definition of shuffle on trajectories, we can assume without loss of generality that $T \subseteq 0^*$. Let $I = \{i : 0^i \in T\}$. Then we can see that $L = \Sigma^I$. Therefore, since it is undecidable whether $L = \Sigma^I$, we have established the result. ■

To summarize, we have established the following result:

Corollary 7.8.4 *Given an LCFL L and regular languages R_1, R_2, it is undecidable whether there exists $T \subseteq \{0,1\}^*$ such that (a) $R_1 \sqcup\!\sqcup_T R_2 = L$, (b) $R_1 \sqcup\!\sqcup_T L = R_2$ or (c) $L \sqcup\!\sqcup_T R_1 = R_2$.*

We now turn to deletion along trajectories.

Lemma 7.8.5 *Given an LCFL L, it is undecidable whether there exists $T \subseteq \{i,d\}^*$ such that $L \rightsquigarrow_T \{\epsilon\} = \Sigma^*$.*

Proof. Let Σ be an alphabet of size at least two, and let $L \subseteq \Sigma^*$ be an LCFL. Then we can verify that

$$\exists T \subseteq \{i,d\}^* \text{ such that } L \rightsquigarrow_T \{\epsilon\} = \Sigma^* \iff L = \Sigma^*, T \supseteq i^*.$$

The right-to-left implication is easily verified. For the reverse implication, let $T \subseteq \{i,d\}^*$ be such that $L \rightsquigarrow_T \{\epsilon\} = \Sigma^*$. Let $x \in \Sigma^*$ be arbitrary. Then there exist $y \in L$ and $t \in T$ such that $x \in y \rightsquigarrow_t \epsilon$. By definition, $y = x$ and $t = i^{|x|}$. From this we can see that $L = \Sigma^*$ and $T \supseteq i^*$. ■

Lemma 7.8.6 *Given an LCFL L, it is undecidable whether there exists $T \subseteq \{i,d\}^*$ such that $\Sigma^* \rightsquigarrow_T \{\epsilon\} = L$.*

Proof. It is easy to verify that

$$\exists T \subseteq \{i,d\}^* \text{ such that } L = \Sigma^* \rightsquigarrow_T \{\epsilon\} \iff \exists I \subseteq \mathbb{N} \text{ such that } L = \Sigma^I.$$

($\Leftarrow$): If $I \subseteq \mathbb{N}$ is such that $L = \Sigma^I$, then let $T = \{i\}^I$. Then we can easily see that $L = \Sigma^* \rightsquigarrow_T \{\epsilon\}$.

($\Rightarrow$): If $T \subseteq \{i, d\}^*$ exists such that $L = \Sigma^* \rightsquigarrow_T \{\epsilon\}$, then by definition of deletion along trajectories, we can assume without loss of generality that $T \subseteq i^*$. Let $I = \{j \ : \ i^j \in T\}$. Then we can see that $L = \Sigma^I$. ■

Lemma 7.8.7 *Given a LCFL $L \subseteq \{a, b\}^*$, it is undecidable whether there exists $T \subseteq \{i, d\}^*$ such that $(\overline{a}a + \overline{b}b)^* \rightsquigarrow_T L = (\overline{a} + \overline{b})^*$, where $\{\overline{a}, \overline{b}\}$ is a marked copy of $\{a, b\}$.*

Proof. Let $R_1 = (\overline{a}a + \overline{b}b)^*$ and $R_2 = (\overline{a} + \overline{b})^*$. We show that there exists $T \subseteq \{i, d\}^*$ such that $R_1 \rightsquigarrow_T L = R_2$ holds if and only if $L = \{a, b\}^*$.

Assume that there exists $T \subseteq \{i, d\}^*$ such that $R_1 \rightsquigarrow_T L = R_2$. Let $x \in R_2$ be arbitrary. Then there exist $y \in R_1$, $z \in L$ and $t \in T$ such that $x \in y \rightsquigarrow_t z$. Let $y = \prod_{i=1}^m \overline{y_i} y_i$ where $y_i \in \{a, b\}$. As $R_2 \subseteq \{\overline{a}, \overline{b}\}^*$, and $L \subseteq \{a, b\}^*$, we must have that $t = (id)^m$ and $z = \overline{x}$. Thus, $L = \{a, b\}^*$, since x was chosen arbitrarily in R_2.

The converse equality $R_1 \rightsquigarrow_T L = R_2$ with $L = \{a, b\}^*$ and $T = (id)^*$ is easily verified. Thus, as it is undecidable whether $L = \{a, b\}^*$, the result is established. ■

Thus, we have demonstrated the following result:

Corollary 7.8.8 *Given an LCFL L and regular languages R_1, R_2, it is undecidable whether there exists $T \subseteq \{i, d\}^*$ such that (a) $R_1 \rightsquigarrow_T R_2 = L$, (b) $L \rightsquigarrow_T R_1 = R_2$, or (c) $R_1 \rightsquigarrow_T L = R_2$.*

7.9 Systems of Language Equations

The study of language equations is part of the larger study of systems of language equations and their solutions. In this section, we move from our previous work on the study single language equations to the study of systems of language equations.

Leiss makes the following strong criticism in his monograph on language equations:

> Somewhat related results, for a single equation in a single variable, were reported in [Kari] [106]; however, this paper restricts the class $\text{EX}_A(\text{CONST}; \text{OP}, X_1, \ldots, X_n)$ in

> our terminology [the set of all systems of language equations over the alphabet A in the variables $X_1, \ldots, X_n$, with operations from OP and taking constant languages, and having a solution in, the class of languages CONST] to one where only a single operator occurs, which, moreover, is assumed invertible with respect to words (not languages). Effectively, this excludes the standard language equations; the equations considered in [106] are essentially word equations. Even more damaging for the generality of these results, only a single equation can be treated at a time, since in order to be able to talk about (nontrivial) systems of equations, it is necessary to have at least two variables present in at least one equation. Therefore, this paper does not contribute significantly toward our goal of establishing a theory of language equations. [131, p. 127]

Leiss does not address the results of Kari and Thierrin [117], which extends the criticized work of Kari to deal with decomposition of languages via catenation. This is perhaps because these results only deal with catenation, and not a general set of operations. However the results of Section 7.3 deal with a large class of operations defined by shuffle on trajectories and therefore introduce a situation where "at least two variables [are] present in at least one equation". Thus, the results of Section 7.3 suggest that the framework introduced by Kari [106], and extended by Kari and Thierrin [117], does represent a valid contribution to the theory of language equations.

In this section, we seek to extend this study of language equations even further and directly address the criticisms of Leiss by considering systems of equations involving shuffle on trajectories. We again focus on decidability of the existence of solutions to a system of equations. We feel again that this shows that the equations considered have merit in the theory of language equations.

We consider systems of equations of the following form. Let $n \geq 1$. Let Σ be an alphabet and $R_1, \ldots, R_n$ be regular languages over Σ. Let $X_1, \ldots, X_m$ be variables. Further, let $Y_{i,1}, Y_{i,2} \in \{X_1, \ldots, X_m\} \cup \text{REG}$ for all $1 \leq i \leq n$ (i.e., $Y_{i,j}$ is either a variable or a regular language over Σ). Let $T_i \subseteq \{0,1\}^*$ for all $1 \leq i \leq n$ be regular sets of trajectories, subject to the condition that if $Y_{i,1}, Y_{i,2}$ are both variables, then T_i is letter-bounded. We define our system of equations as follows: for all $1 \leq i \leq n$, let E_i be the equation

$$E_i: \; R_i = Y_{i,1} \sqcup\!\sqcup_{T_i} Y_{i,2}. \tag{7.21}$$

Our problem then is the following: Given the system of equations E_i for $1 \leq i \leq n$, does there exist a solution $[X_1, \ldots, X_m]$?

Theorem 7.9.1 *Let E_i with $1 \leq i \leq n$ be a system of equations as given by (7.21) and the description above. It is decidable whether there exists a solution $[X_1, \ldots, X_m]$ to the system.*

Proof. Let $1 \leq i \leq n$. Let $\mathcal{S}(T_i, R_i)$ and $\mathcal{T}(T_i, R_i)$ be the sets of languages described by Theorems 7.3.8 and 7.3.9, respectively. For all $1 \leq i \leq n$ and $1 \leq j \leq m$, define sets $\mathcal{V}_j^{(i)}$ of languages as follows:

(a) If $Y_{i,1} = X_j$ and $Y_{i,2} \subseteq \Sigma^*$, then $\mathcal{V}_j^{(i)} = \{\overline{\overline{R_i} \rightsquigarrow_{\tau(T_i)} Y_{i,2}}\}$.

(b) If $Y_{i,1} \subseteq \Sigma^*$ and $Y_{i,2} = X_j$, then $\mathcal{V}_j^{(i)} = \{\overline{\overline{R_i} \rightsquigarrow_{\pi(T_i)} Y_{i,2}}\}$.

(c) If $Y_{i,1} = X_j$ and $Y_{i,2} \in \{X_1, \ldots, X_m\} - \{X_j\}$, then $\mathcal{V}_j^{(i)} = \mathcal{S}(T_i, R_i)$.

(d) If $Y_{i,1} \in \{X_1, \ldots, X_m\} - \{X_j\}$ and $Y_{i,2} = X_j$, then then $\mathcal{V}_j^{(i)} = \mathcal{T}(T_i, R_i)$.

(e) If $Y_{i,1} = Y_{i,2} = X_j$, then $\mathcal{V}_j^{(i)} = \{L_1 \cap L_2 \; : L_1 \in \mathcal{S}(T_i, R_i), L_2 \in \mathcal{T}(T_i, R_i)\}$.

(f) If $Y_{i,1}, Y_{i,2} \in (\{X_1, \ldots, X_m\} - \{X_j\}) \cup \text{REG}$, then $\mathcal{V}_j^{(i)} = \{\Sigma^*\}$.

For all $1 \leq j \leq m$, let

$$\mathcal{V}_j = \{\bigcap_{i=1}^{n} Z^{(i)} \; : \; Z^{(i)} \in \mathcal{V}_j^{(i)}\}.$$

Claim 7.9.2 *The system E_i ($1 \leq i \leq n$) has a solution $[X_1, \ldots, X_m]$ iff it has a solution in $\prod_{j=1}^{m} \mathcal{V}_j$ ($= \mathcal{V}_1 \times \mathcal{V}_2 \times \cdots \times \mathcal{V}_m$).*

Proof. ($\Leftarrow$): Trivial.

($\Rightarrow$): Assume there exists a solution $[X_1, \ldots, X_m]$. Let $1 \leq j \leq m$ be arbitrary. We show that as X_j is a solution to each of the equations E_i in which X_j appears, there is also a language $Z_j \in \mathcal{V}_j$ which is a solution, and $X_j \subseteq Z_j$. Let $1 \leq i \leq n$ be chosen so that $X_j \in \{Y_{i,1}, Y_{i,2}\}$. There are five cases:

(a) $X_j = Y_{i,1}$ and $Y_{i,2} \subseteq \Sigma^*$. Then E_i is given by $R_i = X_j \sqcup_{T_i} Y_{i,2}$. By Theorems 7.2.1 and 5.8.1, we have that $X_j \subseteq \overline{\overline{R_i} \rightsquigarrow_{\tau(T_i)} Y_{i,2}}$. Thus, let $Z_j^{(i)} = \overline{\overline{R_i} \rightsquigarrow_{\tau(T_i)} Y_{i,2}}$. We also have that $R_i = Z_j^{(i)} \sqcup_{T_i} Y_{i,2}$ and that $Z_j^{(i)} \in \mathcal{V}_j^{(i)}$.

(b) $X_j = Y_{i,2}$ and $Y_{i,1} \subseteq \Sigma^*$: similar to case (a).

(c) $X_j = Y_{i,1}$ and $Y_{i,2} \in \{X_1, \ldots, X_m\} - \{X_j\}$. Then we have that $X_j \subseteq Z_j^{(i)}$ for some $Z_j^{(i)} \in \mathcal{V}_j^{(i)}$ by Theorem 7.3.8. Further, $R_i = Z_j^{(i)} \sqcup\!\sqcup_{T_i} Y_{i,2}$.

(d) $X_j = Y_{i,2}$ and $Y_{i,1} \in \{X_1, \ldots, X_m\} - \{X_j\}$. This case is similar to case (c).

(e) $X_j = Y_{i,1} = Y_{i,2}$. Then E_i is given by $R_i = X_j \sqcup\!\sqcup_{T_i} X_j$. By the proof of Theorem 7.4.1, we have that $X_j \subseteq Z_j^{(i)}$ for some $Z_j^{(i)} \in \mathcal{V}_j^{(i)}$. Further, $Z_j^{(i)} \sqcup\!\sqcup_{T_i} Z_j^{(i)} = R_i$.

Thus, we have that for all $1 \le i \le n$ and $1 \le j \le m$, if X_j appears in E_i, then $X_j \subseteq Z_j^{(i)}$ for some $Z_j^{(i)} \in \mathcal{V}_j^{(i)}$. Further, replacing X_j by $Z_j^{(i)}$ in E_i also yields a solution. If X_j does not appear in E_i, then $Z_j^{(i)} = \Sigma^*$, and $X_j \subseteq Z_j^{(i)}$. Let

$$Z_j = \bigcap_{i=1}^{n} Z_j^{(i)}.$$

Note that $X_j \subseteq Z_j$ and that $Z_j \in \mathcal{V}_j$.

We now show that replacing X_j with Z_j still results in a solution to the system of equations E_i with $1 \le i \le n$, i.e., that $[X_1, \ldots, X_{j-1}, Z_j, X_{j+1}, \ldots, X_m]$ is a solution to the system. Consider an arbitrary i with $1 \le i \le n$ where X_j appears in E_i. There are again five cases:

(a) $X_j = Y_{i,1}$ and $Y_{i,2} \subseteq \Sigma^*$. Then $R_i = X_j \sqcup\!\sqcup_{T_i} Y_{i,2}$. By Theorems 7.2.1 and 5.8.1, we have that

$$\begin{aligned} R_i &= X_j \sqcup\!\sqcup_{T_i} Y_{i,2} \subseteq Z_j \sqcup\!\sqcup_{T_i} Y_{i,2} \\ &\subseteq Z_j^{(i)} \sqcup\!\sqcup_{T_i} Y_{i,2} = R_i. \end{aligned}$$

The inclusions are due to the monotonicity of $\sqcup\!\sqcup_{T_i}$. Thus, Z_j satisfies E_i.

(b) $X_j = Y_{i,2}$ and $Y_{i,1} \subseteq \Sigma^*$: similar to case (a).

(c) $X_j = Y_{i,1}$ and $Y_{i,2} \in \{X_1, \ldots, X_m\} - \{X_j\}$. Let $1 \le \ell \le m$ be chosen so that $X_\ell = Y_{j,2}$. Then note that

$$\begin{aligned} R_i &= X_j \sqcup\!\sqcup_{T_i} X_\ell \subseteq Z_j \sqcup\!\sqcup_{T_i} X_\ell \\ &\subseteq Z_j^{(i)} \sqcup\!\sqcup_{T_i} X_\ell = R_i. \end{aligned}$$

The inclusions are again by the monotonicity of $\sqcup\!\sqcup_{T_i}$ and the final equality is due to Theorem 7.3.8. Thus, Z_j is a solution to equation E_i.

(d) $X_j = Y_{i,2}$ and $Y_{i,1} \in \{X_1, \ldots, X_m\} - \{X_j\}$. This case is similar to case (c).

(e) $X_j = Y_{i,1} = Y_{i,2}$. Then E_i is given by $R_i = X_j \sqcup\!\sqcup_{T_i} X_j$. Again, we have that

$$\begin{aligned} R_i &= X_j \sqcup\!\sqcup_{T_j} X_j \subseteq Z_j \sqcup\!\sqcup_{T_i} Z_j \\ &\subseteq Z_j^{(i)} \sqcup\!\sqcup_{T_i} Z_j^{(i)} = R_i. \end{aligned}$$

Thus, Z_j is a solution to E_i.

Thus, we have established that if a solution $[X_1, \ldots, X_m]$ exists, each X_j may be replaced by some $Z_j \in \mathcal{V}_j$. This establishes the claim. ■

We now return to our main proof. We know that each set $\mathcal{V}_j^{(i)}$ is finite and contains only regular languages, each of which may be effectively constructed. Thus, there are finitely many effectively regular languages in $\mathcal{V}_j$ for all $1 \le j \le m$, and the set $\prod_{j=1}^{m} \mathcal{V}_j$ consists of finitely many m-tuples of effectively regular languages. We can test each of these m-tuples for equality. This gives an effective procedure for determining whether solutions to this systems of equations exist. ■

We note that the systems we consider cannot be reduced to a single language equation in the manner of Baader and Narendran [13] (see also Baader and Küsters [11]) since our equations do not involve an explicit union operation.

We also note that for systems of equations as given by (7.21), if the system has a solution $[X_1, \ldots, X_m]$, it also has a solution $[Y_1, \ldots, Y_m]$ which consists of regular languages. We refer to the reader to Choffrut and Karhumäki [25] and Polak [167] for a discussion of systems of language equations involving catenation ($T = 0^*1^*$), Kleene closure and union.

7.10 Conclusions

In this chapter, we have considered language equations involving shuffle and deletion on trajectories. Positive decidability results have been obtained when the fixed languages involved in our equations are regular. When context-free languages are involved, undecidability results have been obtained. We have also considered systems of equations involving shuffle on trajectories.

In particular, we have made progress in the problem posed by shuffle decompositions. The question of whether a regular language R has a non-trivial shuffle decomposition $R = X_1 \sqcup\!\sqcup_T X_2$ when $T = (0+1)^*$ remains open. However, for a substantial and practically significant class of sets of trajectories–namely, the regular letter-bounded sets of trajectories–we have positively answered the shuffle decomposition problem.

We have also investigated decidability problems for equations where the unknown is the set of trajectories. While we have solved the decidability problems for regular languages and LCFLs, the constructions used are distinct from those in the remainder of this chapter, as they do not explicitly involve the use of an inverse operation.