NOTES ON SEPARABILITY

1. General observations

Let L and K be languages. We write $L|K$ if there is a regular R with $L \subseteq R$ and $R \cap K = \emptyset$. For a rational transduction $T \subseteq \Sigma^* \times \Gamma^*$, we define

$$TL = \{v \in \Gamma^* \mid \exists (u,v) \in T,\ u \in L\}, \quad T^{-1} = \{(v,u) \in \Gamma^* \times \Sigma^* \mid (u,v) \in T\}.$$

Lemma 1. *Let T be a rational transduction. Then $L|TK$ if and only if $T^{-1}L|K$.*

Proof. Suppose $L \subseteq R$ and $R \cap TK = \emptyset$ for some regular R. Then clearly $T^{-1}L \subseteq T^{-1}R$ and $T^{-1}R \cap K = \emptyset$. Therefore, the regular set $T^{-1}R$ witnesses $T^{-1}L|K$. Conversely, if $T^{-1}L|K$, then $K|T^{-1}L$ and hence, by the first direction, $(T^{-1})^{-1}K|L$. Since $(T^{-1})^{-1} = T$, this reads $TK|L$ and thus $L|TK$. $\square$

Proposition 2. *Let $\mathcal{C}$ be a full trio generated by the language G (i.e. it consists of languages TG for rational transductions T). Then regular separability for $\mathcal{C}$ can be reduced to the following problem:*

Given: *A language L from $\mathcal{C}$.*
Question: *Does $L|G$?*

Proof. Since $\mathcal{C}$ is generated by G, the input for regular separability for $\mathcal{C}$ comprises two rational transductions T_1 and T_2 and we are asked whether $T_1G|T_2G$. According to lemma 1, the latter is equivalent to $T_2^{-1}T_1G|G$. Since $T_2^{-1}T_1$ is a rational transduction, $T_2^{-1}T_1G$ is a member of $\mathcal{C}$ and this is an instance as required. $\square$

2. VASS languages and CCF/BPP languages

Let $D \subseteq \{a, \bar{a}\}^*$ be the one-sided Dyck language over one pair of parentheses. If $f_i \colon \{a, \bar{a}\}^* \to \{a_i, \bar{a}_i\}^*$ is the morphism with $f_i(a) = a_i$ and $f_i(\bar{a}) = \bar{a}_i$, then we define $D_n = f_1(D) \sqcup\!\sqcup \cdots \sqcup\!\sqcup f_n(D)$, where $\sqcup\!\sqcup$ is the shuffle operator. Observe that D_n is a CCF language (i.e. BPP language). Now proposition 2 tells us that regular separability of VASS languages can be reduced to the following problem:

Given: VASS language L and $n \in \mathbb{N}$.
Question: Does $L|D_n$?

Recall that every VASS language can be written as $h(g^{-1}(D_n) \cap R)$ for some regular R and alphabetic homomorphisms g, h. Now lemma 1 tells us that $h(g^{-1}(D_n) \cap R)|D_n$ iff $g^{-1}(D_n) \cap R|h^{-1}(D_n)$. Since $g^{-1}(D_n)$ and $h^{-1}(D_n)$ are also CCF/BPP languages, we can reduce regular separability of VASS languages to the following problem:

Given: CCF/BPP languages K, L, regular R.
Question: Does $K \cap R|L$?

3. Separation from palindromes

Generalizing the notation above, we write $X|Y$ for subsets $X, Y \subseteq M$ of a monoid M if X and Y are separable by a recognizable subset of M. Moreover, we denote $\Delta = \{(m,m) \mid m \in M\}$.

Lemma 3. *If $X = \bigcup_{i=1}^{n} X_i$ and $Y = \bigcup_{j=1}^{m} Y_j$, then $X|Y$ if and only if $X_i|Y_j$ for every $i \in [1,n]$ and $j \in [1,m]$.*

Proposition 4. *Let M be a monoid and $X, Y \subseteq M$. Then $X|Y$ if and only if $(X \times Y)|\Delta$.*

Proof. One direction is immediate: If $X|Y$ with a recognizable separator $S \subseteq M$, then $X \times Y$ is separated from Δ with the recognizable set $S \times (M \setminus S)$.

Suppose $(X \times Y)|\Delta$ witnessed by $S \subseteq M \times M$ with $X \times Y \subseteq S$ and $S \cap \Delta = \emptyset$. We can write $S = \bigcup_{i=1}^{n} R_i \times T_i$ for recognizable subsets $R_i, T_i \subseteq M$ for $i \in [1,n]$. Note that then $(R_i \times T_i) \cap \Delta = \emptyset$ and thus $T_i \subseteq M \setminus R_i$. Moreover, we have $X \subseteq \bigcup_{i=1}^{n} R_i$ and $Y \subseteq \bigcup_{i=1}^{n} T_i \subseteq \bigcup_{i=1}^{n} (M \setminus R_i)$.

For any $I \subseteq [1,n]$, let $R_I = \bigcap_{i \in I} R_i \cap \bigcap_{i \in [1,n] \setminus I} (M \setminus R_i)$ and $X_I = X \cap R_I$, and $Y_I = Y \cap R_I$. We claim that for any $I, J \subseteq [1,n]$, we have $X_I|Y_J$. Since $X = \bigcup_{\emptyset \neq I \subseteq [1,n]} X_I$ and $Y = \bigcup_{I \subsetneq [1,n]} Y_I$, the proposition then follows from lemma 3.

Suppose $I = J$. We shall prove that then either $X_I = \emptyset$ or $Y_J = \emptyset$, which clearly implies $X_I|Y_J$. Toward a contradiction, assume that there are $x \in X_I$ and $y \in Y_J$. Since $X \times Y \subseteq S$, there is an $i \in [1,n]$ with $x \in R_i$ and $y \in T_i \subseteq (M \setminus R_i)$. The former implies $i \in I$, and the latter $i \notin J$, contradicting $I = J$.

Suppose $I \neq J$. If there is an $i \in I \setminus J$, then $X_I \subseteq R_I \subseteq R_i$ and $Y_J \subseteq R_J \subseteq M \setminus R_i$, meaning that R_i witnesses $X_I|Y_J$. On the other hand, if $i \in J \setminus I$, then $X_I \subseteq M \setminus R_i$ and $Y_I \subseteq R_i$, so that $M \setminus R_i$ witnesses $X_I|Y_J$. □

4. Approximants of $\mathbb{Z}$-VASS

Fix $n \in \mathbb{N}$ and let $e_1, \ldots, e_n \in \mathbb{Z}^n$ be the unit vectors. Let $\Sigma = \{a_i, \bar{a}_i \mid i \in \{1, \ldots, n\}\}$ and $\varphi\colon \Sigma^* \to \mathbb{Z}^n$ the morphism with $\varphi(a_i) = e_i$ and $\varphi(\bar{a}_i) = -e_i$. Let $Z \subseteq \Sigma^*$ be the set of all $w \in \Sigma^*$ with $\varphi(w) = 0$.

We want to understand the regular languages R with $R \cap Z = \emptyset$. We begin with two types of such regular languages. For $k \in \mathbb{N}$, let $\mu_k\colon \mathbb{Z}^n \to (\mathbb{Z}/k\mathbb{Z})^n$ the projection modulo k. The language

$$M_k = \{w \in \Sigma^* \mid \mu_k(\varphi(w)) \neq 0\}$$

is clearly regular and satisfies $M_k \cap Z = \emptyset$. For each $k \in \mathbb{N}$, let

$$I_{y,k} = \{w \in \Sigma^* \mid \langle \varphi(w), y \rangle > 0,\ \langle \varphi(f), y \rangle \geq -k \text{ for every factor } f \text{ of } w\}.$$

Here, $\langle x, y \rangle$ denotes the usual scalar product of vectors x and y. The number k allows words in $I_{y,k}$ to have factors with a scalar product that is negative, but not too far below zero. Observe that, again, $I_{y,k}$ is disjoint from Z and regular. The main result of this section is that the sets of the form M_k and $I_{y,k}$ yield a complete description of regular sets that are disjoint from Z:

Theorem 5. *Let $R \subseteq \Sigma^*$ be a regular language with $R \cap Z = \emptyset$. Then $R = \bigcup_{i=1}^{r} R_i$ where each R_i is a regular subset of either some M_k or some $I_{y,k}$.*

For the proof of theorem 5, we need two ingredients. The first is the well-known Farkas Lemma [2, Corollary 7.1d]. It tells us that non-solvability of an equation system is certified by a hyperplane separating all the left-hand sides from the right-hand side.

Lemma 6 (Rational Farkas Lemma)**.** *Let $A \in \mathbb{Q}^{n\times m}$ and $b \in \mathbb{Q}^n$. Then exactly one of the following holds:*

(1) There is a solution $x \in \mathbb{Q}^m$ to the equation $Ax = b$ with $x \geq 0$.
(2) There is a vector $y \in \mathbb{Q}^n$ with $y^tA \geq 0$ and $y^tb < 0$.

The following is the known fact that abelian groups are subgroup separable. It essentially says that non-membership in a finitely generated subgroup is certified by a morphism into a finite group.

Lemma 7. *If G is a finitely generated abelian group, H is a finitely generated subgroup, and $g \notin H$, then there is a finite abelian group F and a morphism $\psi\colon G \to F$ with $\psi(H) = 0$ and $\psi(g) \neq 0$.*

Proof. Dividing by H allows us to assume that $H = 0$, $g \neq 0$. Since G is finitely generated abelian, we have $G \cong \mathbb{Z}^n \oplus \bigoplus_{i=1}^m \mathbb{Z}/a_i\mathbb{Z}$. If $n = 0$, we are done. Otherwise, choose a number k that is larger than each of the torsion-free components of g. Then the projection to $(\mathbb{Z}/k\mathbb{Z})^n \oplus \bigoplus_{i=1}^m(\mathbb{Z}/a_i\mathbb{Z})$ maps g to an element $\neq 0$. □

Let $I_y = \{w \in \Sigma^* \mid \langle\varphi(w), y\rangle > 0\}$. Although this language is not regular, each of its regular subsets is contained in some $I_{y,k}$.

Lemma 8. *If a regular language is included in I_y, then it is included in $I_{y,k}$ for some $k \geq 0$.*

Proof. Consider a loop in an automaton for $R \subseteq I_y$. The φ-image z of the label of this loop must satisfy $y^tz \geq 0$, since otherwise pumping would yield a word outside of I_y. This implies that we can choose k so that R is contained in $I_{y,k}$. □

A *linear path scheme* is a tuple $S = (u_0, v_1, u_1, \ldots, v_m, u_m)$ of words from Σ^*. We call S *linearly independent* if the vectors $\varphi(v_1), \ldots, \varphi(v_m)$ are linearly independent. Its *language* is $L(S) = u_0v_1^*u_1\cdots v_m^*u_m$.

Lemma 9. *Consider a linearly independent linear path scheme*

$$S = (u_0, v_1, u_1, \ldots, v_m, u_m)$$

so that $L(S) \cap Z = \emptyset$. Then at least one of the following holds:

(1) $L(S) \subseteq M_k$ for some $k \geq 0$.
(2) $L(S) \subseteq I_y$ for some $y \in \mathbb{Z}^n$.

Proof.
- Let $A \in \mathbb{Z}^{n\times m}$ be the matrix whose columns are $\varphi(v_1), \ldots, \varphi(v_m)$ and let $-b \in \mathbb{Z}^n$ be the vector $\varphi(u_0\cdots u_m)$. Since $L(S) \cap Z = \emptyset$, we know that $Ax = b$ has no solution in $\mathbb{N}^m$.
- Suppose $Ax = b$ has no solution in $\mathbb{Z}^m$. According to lemma 7, there is a finite abelian group F and a morphism $\mu\colon \mathbb{Z}^n \to F$ such that μ maps all columns of A to 0 and $\mu(b) \neq 0$. In particular, μ maps all φ-images of words in $L(S)$ to $F \setminus \{0\}$. Since every morphism from $\mathbb{Z}^n$ into a finite group factorizes over some μ_k, this yields a set M_k containing $L(S)$.

- Suppose $Ax = b$ has a solution in $\mathbb{Z}^m$. Then there is no solution $x \in \mathbb{Q}^m$ with $x \geq 0$: Otherwise, since the independence of the columns of A make the solution unique, this solution would belong to $\mathbb{Q}_+^m \cap \mathbb{Z}^m = \mathbb{N}^m$, which is impossible.

  We can therefore apply lemma 6 and obtain a $y \in \mathbb{Q}^n$ with $y^t A \geq 0$ and $y^t b < 0$. We can clearly choose it with $y \in \mathbb{Z}^n$. Then clearly $L(S) \subseteq I_y$.

□

Proof sketch for theorem 5.
- Construct regular sets $R_1, \ldots, R_\ell$ and linearly independent linear path schemes $S_1, \ldots, S_\ell$ such that
  (1) $R = R_1 \cup \cdots \cup R_\ell$ and
  (2) for each $i = 1, \ldots, \ell$, we have $\Psi(R_i) = \Psi(L(S_i))$.

  This can be done either (i) using skeleton decompositions or (ii) using a Parikh annotation for R [3] and, in either case, using the idea of [1, Prop. 9.1] to make the period vectors linearly independent while guaranteeing Parikh-equivalence.
- Since $\Psi(L(S_i)) \subseteq \Psi(R)$, we have $L(S_i) \cap Z = \emptyset$ and thus lemma 9 tells us that we either have $L(S_i) \subseteq M_k$ for some $k \geq 0$ or $L(S_i) \subseteq I_y$ for some $y \in \mathbb{Z}^n$.

  In the first case, we also have $R_i \subseteq M_k$ since $\Psi(R_i) \subseteq \Psi(L(S_i))$. In the second case, we have $R_i \subseteq I_y$ (again, because $\Psi(R_i) \subseteq \Psi(L(S_i))$) and thus $R_i \subseteq I_{y,k}$ for some $k \geq 0$ according to lemma 8.

□

Theorem 5 tells us that regular separability of languages in $\mathcal{C}$ from $\mathbb{Z}$-VASS languages reduces to the following problem:

Given: A language L from $\mathcal{C}$ and $n \in \mathbb{N}$.
Question: Does there exist k and a finite set $F \subseteq \mathbb{Z}^n$ such that $L \subseteq M_k \cup \bigcup_{y \in F} I_{y,k}$?

5. Soft-bodied $\mathbb{Z}$-VASS separability

For an alphabet X, let $X^\oplus$ denote the set of mappings $X \to \mathbb{N}$. Moreover, let $\Psi_X(\cdot)\colon X^* \to X^\oplus$ be the Parikh map. For an alphabet $Y \subseteq X$, the map $\pi_Y\colon X^* \to Y^*$ is the projection onto Y^*.

We want to solve the *$\mathbb{Z}$-VASS separation problem*:

Input: A regular language $R \subseteq X^*$ and semilinear sets $U_1, U_2 \subseteq X^\oplus$.
Question: Is there a regular language S with $R \cap \Psi_\Sigma^{-1}(U_1) \subseteq S$ and $S \cap R \cap \Psi_\Sigma^{-1}(U_1) = \emptyset$?

We show the following:

Proposition 10. *The $\mathbb{Z}$-VASS separation problem can be reduced to separability of semilinear sets by recognizable sets.*

The first step is to define Parikh annotations.

Definition 11. *Let $L \subseteq X^*$ be a language and $\mathcal{C}$ be a language class. A* Parikh annotation *(PA) for L in $\mathcal{C}$ is a tuple $(K, C, P, (P_c)_{c \in C}, \varphi)$, where*

- *C, P are alphabets such that X, C, P are pairwise disjoint,*
- *$K \subseteq C(X \cup P)^*$ is in $\mathcal{C}$,*
- *φ is a morphism $\varphi\colon (C \cup P)^\oplus \to X^\oplus$,*

- *P_c is a subset $P_c \subseteq P$ for each $c \in C$,*

such that

1. *$\pi_X(K) = L$ (the* projection property*),*
2. *$\varphi(\pi_{C\cup P}(w)) = \Psi(\pi_X(w))$ for each $w \in K$ (the* counting property*), and*
3. *$\Psi(\pi_{C\cup P}(K)) = \bigcup_{c\in C} c + P_c^{\oplus}$ (the* commutative projection property*).*

Intuitively, a Parikh annotation describes for each w in L one or more Parikh decompositions of $\Psi(w)$. The symbols in C represent constant vectors and symbols in P represent period vectors. Here, the symbols in $P_c \subseteq P$ correspond to those that can be added to the constant vector corresponding to $c \in C$. Furthermore, for each $x \in C \cup P$, $\varphi(x)$ is the vector represented by x.

For this application of Parikh annotations, we need a further property. A Parikh annotation $(K, C, P, (P_c)_{c\in C}, \varphi)$ for L is said to be *pseudo-bounded* if on each of the sets P_c, one can establish a linear order $(P_c, <)$ such that

$$cp_1^* \cdots p_n^* \subseteq \pi_{C\cup P}(K),$$

where $P_c = \{p_1, \ldots, p_n\}$ and $p_1 < \cdots < p_n$.

In other words, in a pseudo-bounded PA, when projecting to the annotation alphabet $\{c\} \cup P_c$, every description of a Parikh image appears as some word from the bounded language $cp_1^* \cdots p_n^*$. It is not hard to see that regular languages admit pseudo-bounded Parikh annotations. See [3, Lemma 9.3.5, p. 151] for a (short) proof.

Lemma 12. *Given a regular language R, one can construct a regular pseudo-bounded Parikh annotation for R.*

Reduction. We now prove proposition 10. Let $(K, C, P, (P_c)_{c\in C}, \varphi)$ be a pseudo-bounded regular Parikh annotation for R. For each $i \in \{1, 2\}$ and $c \in C$, define

$$T_{i,c} = \{\mu \in P_c^{\oplus} \mid \varphi(c + \mu) \in U_i\}.$$

Then $T_{i,c}$ is clearly Presburger-definable and hence semilinear. We establish proposition 10 by showing the following.

Lemma 13. *$R \cap \Psi_\Sigma^{-1}(U_1)$ and $R \cap \Psi_\Sigma^{-1}(U_2)$ are separable by a regular language if and only if for each $c \in C$, the sets $T_{1,c}$ and $T_{2,c}$ are separable by a recognizable set.*

Proof. We begin with the "if" direction. Let $S_c \subseteq P_c^{\oplus}$ be a recognizable separator of $T_{1,c}$ and $T_{2,c}$, meaning $T_{1,c} \subseteq S_c$ and $S_c \cap T_{2,c} = \emptyset$. Define

$$S = \{\pi_X(w) \mid \exists c \in C\colon cw \in K,\ \Psi(\pi_{P_c}(w)) \in S_c\},$$

which is clearly regular because each S_c is recognizable. We claim that S is a separator for $R \cap \Psi_X^{-1}(U_1)$ and $R \cap \Psi_X^{-1}(U_2)$.

Suppose $u \in R \cap \Psi_X^{-1}(U_1)$. By the projection property, there is a $c \in C$ and $cw \in K$ such that $\pi_X(w) = u$. Since $\Psi(u) \in U_1$, the counting property entails

$$\varphi(c + \Psi(\pi_{P_c}(w))) = \Psi(u) \in U_1,$$

which implies $\Psi(\pi_{P_c}(w)) \in T_{1,c} \subseteq S_c$ and thus $u \in S$.

Now assume $u \in S$. Then we can write $u = \pi_X(w)$ such that for some $c \in C$, we have $cw \in K$ and $\Psi(\pi_{P_c}(w)) \in S_c$. The latter implies $\Psi(\pi_{P_c}(w)) \notin T_{2,c}$. By definition of $T_{2,c}$, this means $\varphi(c + \Psi(\pi_{P_c}(w))) \notin U_2$ and therefore, by the counting property,

$$\Psi(u) = \varphi(c + \Psi(\pi_{P_c}(w))) \notin U_2.$$

Thus, $S \cap \Psi_X^{-1}(U_2) = \emptyset$ and S separates $R \cap \Psi_X^{-1}(U_1)$ and $R \cap \Psi_X^{-1}(U_2)$.

For the "only if" direction, suppose S is a regular language with $R \cap \Psi_X^{-1}(U_1) \subseteq S$ and $R \cap \Psi_X^{-1}(U_2) \cap S = \emptyset$. First, we modify S slightly and obtain for each $c \in C$ the regular language

$$S'_c = \{cw \in K \mid \pi_X(w) \in S,\ \pi_{P_c}(w) \in p_1^* \cdots p_n^*\},$$

where $P_c = \{p_1, \ldots, p_n\}$ such that $p_1 < \cdots < p_n$. Observe that the language $\pi_{P_c}(S'_c)$ is included in $p_1^* \cdots p_n^*$, which implies that its image $S_c = \Psi(\pi_{P_c}(S'_c)) \subseteq P_c^{\oplus}$ is recognizable. We claim that S_c separates $T_{1,c}$ and $T_{2,c}$.

Let $\mu \in T_{1,c}$. Then $\varphi(c + \mu) \in U_1$. Moreover, by pseudo-boundedness of the Parikh annotation, we have a word $cw \in K$ such that $\pi_{P_c}(w) \in p_1^* \cdots p_n^*$ and $\Psi(\pi_{P_c}(w)) = \mu$. Since $\pi_X(cw) \in R$ (projection property) and $\Psi(\pi_X(cw)) = \varphi(c + \mu) \in U_1$, we have $\pi_X(cw) \in S$ and hence $cw \in S'_c$. Thus $\mu = \Psi(\pi_{P_c}(w)) \in S_c$.

Now let $\mu \in S_c$, which means there is a $cw \in K$ with $\pi_X(cw) \in S$ and $\Psi(\pi_{P_c}(w)) = \mu$. Together with $\pi_X(cw) \in R$, the fact $\pi_X(cw) \in S$ tells us that $\Psi(\pi_X(cw)) \notin U_2$ and thus

$$\varphi(c + \mu) = \Psi(\pi_X(cw)) \notin U_2,$$

hence $\mu \notin T_{2,c}$ by definition. □

6. Counter-examples

For most language classes encoutered for infinite-state systems, it turns out that emptiness of intersection is decidable if and only if regular separability is decidable. This raises the question of whether the two problems are in fact reducible to each other. Here, we present counter examples to this.

6.1. **Separation decidable, but not intersection.** We present two full trios $\mathcal{H}$ and $\mathcal{C}$ such that given L_1 from $\mathcal{H}$ and L_2 from $\mathcal{C}$, it decidable whether $L_1 | L_2$, but undecidable whether $L_1 \cap L_2 = \emptyset$.

For $w \in \{0,1\}^*$, let $\nu(w)$ be the number represented by w in binary. Furthermore, for a subset $S \subseteq \mathbb{N}$, let $2^S = \{2^n \mid n \in S\}$. For $L \subseteq \{0,1\}^*$, let $P_L = \mathbb{N} \backslash 2^{\mathbb{N}} \cup \{2^{\nu(w)} \mid w \in L\}$.

Let $\mathcal{H}$ be the class of higher-order pushdown languages. A subset $P \subseteq \mathbb{N}$ is a *para-pushdown predicate* if $P = P_L$ for some L from $\mathcal{H}$. We will use the following fact:

Lemma 14. *If $L \subseteq \{0,1\}^*$ is an order-i pushdown language, then $\{10^{\nu(w)} \mid w \in L\}$ is an order-$(i+1)$ pushdown language.*

To define $\mathcal{C}$, we use an automata model. A *para-pushdown automaton* has a set of states Q, a finite set of edges $E \subseteq Q \times \times \Sigma^* \times \{0,1\} \times Q$, an initial state q_0, and a finite set F of *acceptance pairs*: The elements of F are pairs (q, P), where $q \in Q$ and P is a para-pushdown predicate. A *configuration* is a pair $(q, n) \in Q \times \mathbb{N}$. Taking an edge (q, w, m, q') means going from q to q', reading w from the input, and adding m to the counter. A run is *accepting* if it ends in a configuration (q, n) such that there is an acceptance pair (q, P) with $n \in P$. Let $\mathcal{C}$ be the class of languages accepted by para-pushdown automata.

Intersection. Let us argue that intersection for $\mathcal{H}$ and $\mathcal{C}$ is undecidable. We reduce from intersection of context-free languages. Given two context-free languages L_1 and L_2, consider the languages

$$K_1 = \{10^{2^{\nu(w)}} \mid w \in L_1\}, \qquad K_2 = \{10^n \mid n \in P_{\bar{L}_2}\},$$

where $\bar{L}_2 = \{10^{\nu(w)} \mid w \in L_2\}$. Note that K_1 is an order-3 pushdown language by lemma 14 applied twice. Moreover $\bar{L}_2$ is an order-2 pushdown language by lemma 14, so that K_2 is a para-pushdown language. Finally, we have $K_1 \cap K_2 = \emptyset$ if and only if $L_1 \cap L_2 = \emptyset$.

Regular separability. Let us argue that regular separability for $\mathcal{H}$ and $\mathcal{C}$ is decidable. For a subset $S \subseteq \mathbb{N}$, let $a^S = \{a^m \mid m \in S\}$.

Clearly, a language L is in $\mathcal{C}$ if and only if it is of the form $L = T_1L_1 \cup \cdots \cup T_nL_n$, where $L_i = a^{P_i}$ for some para-pushdown predicate P_i. Hence, given K from $\mathcal{H}$ and L from $\mathcal{C}$, it suffices to decide whether $K|T_iL_i$ for each $i \in [1,n]$. According to lemma 1, that is the case if and only if $T_i^{-1}K|L_i$. Since $T_i^{-1}K$ is in $\mathcal{H}$ as well and satisfies $T_i^{-1}K \subseteq a^*$, it suffices to decide the following problem:

Given: A language $K \subseteq a^*$ in $\mathcal{H}$ and a para-pushdown predicate $P \subseteq \mathbb{N}$.
Question: Does $K|a^P$ hold?

That is decidable because of the following lemma: Recall that finiteness is decidable for higher-order pushdown languages and para-pushdown predicates are decidable.

Lemma 15. *Let $K \subseteq a^*$ and P be a para-pushdown predicate. Then $K|a^P$ if and only if K is finite and disjoint from a^P.*

This, in turn, follows from the following observation:

Lemma 16. *Let $S_0 \subseteq 2^{\mathbb{N}}$ and $\mathbb{N}\backslash 2^{\mathbb{N}} \subseteq S_1$. Then a^{S_0} and a^{S_1} are regularly separable if and only if S_0 is finite and disjoint from S_1.*

Proof. The "if" is obvious. For the "only if", we argue that no infinite subset of a^* is regularly separable from a^{S_1}. That is because a^{S_1} intersects every infinite regular set: Since $\mathbb{N} \setminus 2^{\mathbb{N}} \subseteq S_1$, a regular set $R \subseteq a^*$ that is disjoint from a^{S_1} has to satisfy $R \subseteq a^{2^{\mathbb{N}}}$, which implies that R is finite. ☐

6.2. **Intersection decidable, but not separation.** Here, we present two language classes $\mathcal{C}_0$ and $\mathcal{C}_1$ such that given L_0 from $\mathcal{C}_0$ and L_1 from $\mathcal{C}_1$, it is decidable whether $L_0 \cap L_1 = \emptyset$, but undecidable whether $L_0|L_1$.

Let $\mathcal{C}$ be a language class. A subset $S \subseteq \mathbb{N}$ is a *pseudo-$\mathcal{C}$ predicate* if $S = \nu(L) = \{\nu(w) \mid w \in L\}$ for some language L in $\mathcal{C}$. A *pseudo-$\mathcal{C}$ counter machine* has a set of states Q, a finite set of edges $E \subseteq Q \times \times \Sigma^* \times \{0,1\} \times Q$, an initial state q_0, and a finite set F of *acceptance pairs*: The elements of F are pairs (q, P), where $q \in Q$ and P is a pseudo-$\mathcal{C}$ predicate. A *configuration* is a pair $(q, n) \in Q \times \mathbb{N}$. Taking an edge (q, w, m, q') means going from q to q', reading w from the input, and adding m to the counter. A run is *accepting* if it ends in a configuration (q, n) such that there is an acceptance pair (q, P) with $n \in P$.

Let $\mathcal{L}$ be a full trio where emptiness is decidable, but infinity is undecidable. For example, one can take the languages accepted by lossy channel systems. Moreover, let Reg denote the class of regular languages. Let

- $\mathcal{C}_0$ be the class of languages accepted by pseudo-$\mathcal{L}$ counter machines and

- $\mathcal{C}_1$ be the class of languages accepted by pseudo-Reg counter machines.

Decidability of intersection. Suppose we are given L_0 from $\mathcal{C}_0$ and L_1 from $\mathcal{C}_1$. We can clearly write each L_i as a finite union of sets of the form Ta^S, where S is a pseudo-$\mathcal{L}$ precicate for $i = 0$ and S is a pseudo-Reg predicate if $i = 1$. Hence, to decide whether $L_0 \cap L_1 = \emptyset$, it suffices to decide whether $T_0a^{S_0} \cap T_1a^{S_1} = \emptyset$, where S_0 is a pseudo-$\mathcal{L}$ predicate and S_1 is a pseudo-Reg predicate. Note that $T_0a^{S_0} \cap T_1a^{S_1} = \emptyset$ if and only if $a^{S_0} \cap T_1^{-1}T_0a^{S_1}$. The rational transduction $T_1^{-1}T_0$ is a rational subset of $a^* \times a^*$ and if we identify words in a^* with natural numbers, it is semilinear. That means $T_1^{-1}T_0a^{S_1}$ corresponds to a set of numbers that is definable in first-order logic in the structure $(\mathbb{N}, +, \leq, S_1)$. Since S_1 is a pseudo-Reg set, this structure is automatic with respect to the usual binary encoding. Therefore, the set of binary representations of words in $T_1^{-1}T_0a^{S_1}$ is effectively regular. Hence, we have $T_1^{-1}T_0a^{S_1} = a^{\nu(R)}$ for some regular set $R \subseteq \{0,1\}^*$. Now deciding whether $a^{S_0} \cap a^{\nu(R)} = \emptyset$ amounts to checking whether a given language from $\mathcal{L}$ and a given regular language intersect. That is decidable because emptiness is decidable for $\mathcal{L}$.

Remark 1. As a simpler example, one could also take $\mathcal{C}_0 = \mathcal{C}_1$ to be the class of languages of pseudo-$\mathcal{L}$ counter machines, where $\mathcal{L}$ is the class of languages of lossy channel systems. Undecidability of separability of course continues to hold. For intersection, one needs to observe that $T_1^{-1}T_0a^{S_1}$ is not only definable in first-order logic of $(\mathbb{N}, +, \leq, S_1)$, but in the Σ_1 fragment of $(\mathbb{N}, +, 0, 1, S_1)$, where S_1 occurs only positively. Then, constructions from automatic structures show that the binary representations of $T_1^{-1}T_0a^{S_1}$ are accepted by a lossy channel system. Then, one can use decidability of intersection for lossy channel systems.

Undecidability of regular separability. We reduce the infinity problem of $\mathcal{L}$, so suppose we are given L from $\mathcal{L}$. Since $\mathcal{L}$ is closed under rational transductions, we may assume that $L \subseteq 10^*$ and hence $\nu(L) \subseteq 2^{\mathbb{N}}$. The language $K_0 = a^{\nu(L)}$ is clearly in $\mathcal{C}_0$. Moreover, the language $K_1 = a^{\mathbb{N}\setminus 2^{\mathbb{N}}} = a^{\nu(1\{0,1\}^*1\{0,1\}^*)}$ is clearly in $\mathcal{C}_1$, because $1\{0,1\}^*1\{0,1\}^*$ is regular. According to lemma 16, it follows that K_0 and K_1 are regularly separable if and only if K_0 is finite and disjoint from K_1. Since K_0 and K_1 are disjoint, we have $K_0|K_1$ if and only if K_0 is finite, which is the case if and only if L is finite.

References

[1] Samuel Eilenberg and Marcel-Paul Schützenberger. "Rational sets in commutative monoids". In: *Journal of Algebra* 13.2 (1969), pp. 173–191. DOI: `10.1016/0021-8693(69)90070-2`.
[2] Alexander Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, 1986.
[3] Georg Zetzsche. "Monoids as Storage Mechanisms". PhD thesis. Technische Universität Kaiserslautern, 2016.