

# Languages, automata and computation II

## Tutorial 1 – Semilinear sets

Winter semester 2024/2025

We consider  $\mathbb{N}^d$  as a finitely generated free commutative additive monoid. We can define several operations on subsets  $A, B$  of  $\mathbb{N}^d$ :

- Union  $A \cup B$ , intersection  $A \cap B$ , and complement  $\mathbb{N}^d \setminus A$ .
- Addition  $A + B := \{a + b \mid a \in A, b \in B\}$ .
- Iteration  $\mathbb{N}A := A^* := \bigcup_{n \in \mathbb{N}} nA$ , where  $nA = A + A + \dots + A$  ( $n$  times).
- Projection  $\pi_i(A) := \{(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_d) \mid (a_1, \dots, a_d) \in A\}$ .

By convention,  $0A = \{0\}$  if  $A$  is nonempty, otherwise  $0\emptyset = \emptyset$ .

### Semilinear sets

A subset  $L$  of  $\mathbb{N}^d$  is *linear* if it is of the form  $b + P^*$  for some *basis vector*  $b \in \mathbb{N}^d$  and a finite set of *periods*  $P \subseteq \mathbb{N}^d$ . A subset  $S$  of  $\mathbb{N}^d$  is *semilinear* if it is a finite union of linear sets,  $S = L_1 \cup \dots \cup L_n$ .

**Exercise 1** (Closure properties). Show that the semilinear sets are closed under union, addition, iteration, and projection.

**Exercise 2.** Find a subset of  $\mathbb{N}$  which is not semilinear.

**Exercise 3.** Show that the set  $\{(n, 2^n) \in \mathbb{N}^2 \mid n \in \mathbb{N}\}$  is not semilinear.

**Exercise 4.** Let  $X \subseteq \mathbb{N}$  be an arbitrary (finite or infinite) set of natural numbers. Show that its iteration  $X^*$  is a semilinear set.

**Exercise 5.** Find a set  $X \subseteq \mathbb{N}^2$  such that  $X^*$  is not semilinear.

### Rational sets

In analogy with the rational subsets of the free monoid  $\Sigma^*$ , we say that a subset  $R$  of the free commutative monoid  $\mathbb{N}^d$  is *rational* if it can be generated from finite sets using union “ $\cup$ ”, addition “ $+$ ” and iteration “ $_*$ ”.

**Exercise 6** (Rational = semilinear). Show that a subset of  $\mathbb{N}^d$  is rational if, and only if, it is semilinear.

In view of the exercise, we can see the basis-periods representation of a semilinear set as a normal form for rational sets.

## Büchi sets

We explore an automata-based approach to represent subsets of  $\mathbb{N}^d$ . For simplicity we deal with  $d = 1$ , however the whole approach generalizes to arbitrary  $d$ . Consider the alphabet  $\Sigma = \{0, 1\}$ . We see a word  $w = a_0 \cdots a_n \in \Sigma^*$  as a binary encoding (least digit first) of the natural number

$$[w]_2 = a_0 + a_1 \cdot 2 + a_2 \cdot 2^2 + \cdots + a_n \cdot 2^n \in \mathbb{N}. \quad (1)$$

Clearly every natural number as a (nonunique) binary encoding. Similarly, for every dimension  $d \geq 1$  let  $\Sigma_d := \Sigma^d$ ; then, a word  $w \in \Sigma_d^*$  encodes a  $d$ -tuple of natural numbers  $[w]_2 \in \mathbb{N}^d$ . For instance, the word  $w = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  encodes the pair  $(6, 5)$ . In this way a language  $L \subseteq \Sigma_d^*$  encodes a subset  $[L]_2$  of  $\mathbb{N}^d$ , and reciprocally every subset  $X$  of  $\mathbb{N}^d$  can be encoded as  $X = [L]_2$  for some language  $L \subseteq \Sigma_d^*$ . A *Büchi subset of  $\mathbb{N}^d$*  is one which is encoded by some regular language.

**Exercise 7.** Show that the set of numbers of the form  $2^n$  with  $n \in \mathbb{N}$  is a Büchi set.

**Exercise 8.** Show that the set of triples of natural numbers  $(a, b, c)$  s.t.  $a + b = c$  is a Büchi subset of  $\mathbb{N}^3$ . Generalise the argument to triples of  $d$ -tuple of natural numbers.

**Exercise 9.** Show that Büchi sets are closed under union, addition, intersection, complement, iteration, and projection.

As a consequence, semilinear sets are a strict subset of Büchi sets.