Usefulness of information for regular languages

Given a problem, we shall consider a supplementary information (advice) useful, if it enables to find a simpler solution of the problem.

- **Problem** - decide whether an input w belongs to a regular language L
- **Solution** - a deterministic finite automaton A accepting the language L
- **Complexity of the solution** - the number of states of the automaton A ($sc(A)$)
- **Advice** - the input w belongs to L_{adv} accepted by an automaton A_{adv}

The advice enables us to solve a new problem instead of the original problem.

- **New problem** - decide whether the input w belongs to a regular language L_{new} accepted by an automaton A_{new}
- $\mathbf{L = L_{new} \cap L_{adv}}$

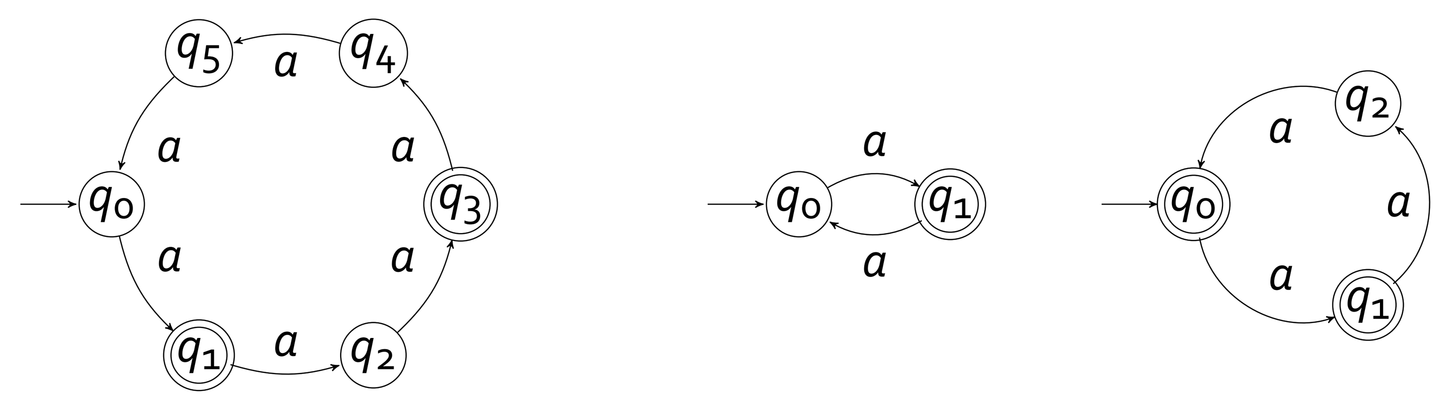
We consider the advice useful if the solution of the new problem is simpler than the best solution of the original problem. Verifying the correctness of the advice should also be simpler than the best solution of the original problem.

- **Best solution** - the minimal automaton A_{min} accepting L
- $\mathbf{sc(A_{new}) < sc(A_{min})}$, $\mathbf{sc(A_{adv}) < sc(A_{min})}$

If these conditions hold, we say that the automata A_{adv} and A_{new} form a decomposition of A_{min} and L is decomposable into L_{adv} and L_{new}.

Example of a decomposition

- Problem - decide whether the length of an input w is 1 or 3 modulo 6 (L).
- Advice - the length of the input w is odd (L_1).
- New language - it suffices to find out whether the length of the input w is 1 or 2 modulo 3 (L_2).



$L = \{a^{6k+1}, a^{6k+3} \mid k \in \mathbb{N}\}$ $L_1 = \{a^{2k+1}\}$ $L_2 = \{a^{3k}, a^{3k+2}\}$

Figure 1. The language L is decomposable into languages L_1 and L_2. It holds $L = L_1 \cap L_2$

Unary languages [1]

Unary languages have been characterized upon decomposability [1] - sufficient and necessary conditions for decomposability of unary languages have been found. These conditions are related to **the number of states in the cycle** of the minimal automaton for a given language, **the number of states in the initial tail** and the position of the **accepting states**.

References

[1] Giovanni Pighizzini, Branislav Rovan, and Šimon Sádovský.
Usefulness of information and decomposability of unary regular languages.
Information and Computation, page 104868, 2022.

[2] Vincent Hlaváč.
Power of supplementary information for unary regular languages.
Master's thesis, Comenius University in Bratislava, 2025.

[3] Andrej Ravinger.
Usefulness of information for bounded regular languages.
Master's thesis, Comenius University in Bratislava, 2025.

Family of languages $\mathcal{L}_A$ [2]

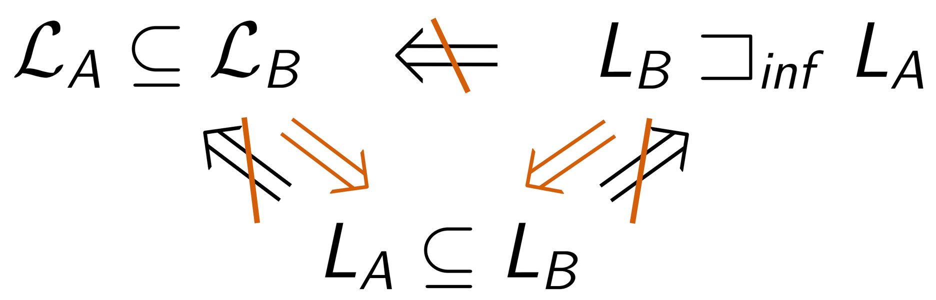
We say that $L(A)$ **provides useful information** for L if there exists a decomposition of the minimal automaton accepting the language L, which contains the automaton A.

We denote by $\mathcal{L}_A$ the family of problems such that $L(A)$ provides useful information for them.

Results concerning cardinality of $\mathcal{L}_A$ follow.

- Every **infinite unary regular language** $L(A)$ except a^* provides useful information for **infinitely many problems**, i.e., $\mathcal{L}_A$ is infinite.
- Neither a^* nor any finite unary regular language $L(B)$ provide useful information for any problem, i.e., $\mathcal{L}_B$ is empty.

The following figure describes relations between the inclusion of families $\mathcal{L}_A$ and $\mathcal{L}_B$, the inclusion of languages $L(A)$ and $L(B)$ and the relation *provide useful information* $\sqsupset_{inf}$.



For all implications from the figure, that do not hold, we found infinite sequences of counterexamples.

The upper implication is related to the transitivity of the relation *provide useful information*. Thus we proved the unintuitive fact that the relation *provide useful information* is **not transitive**.

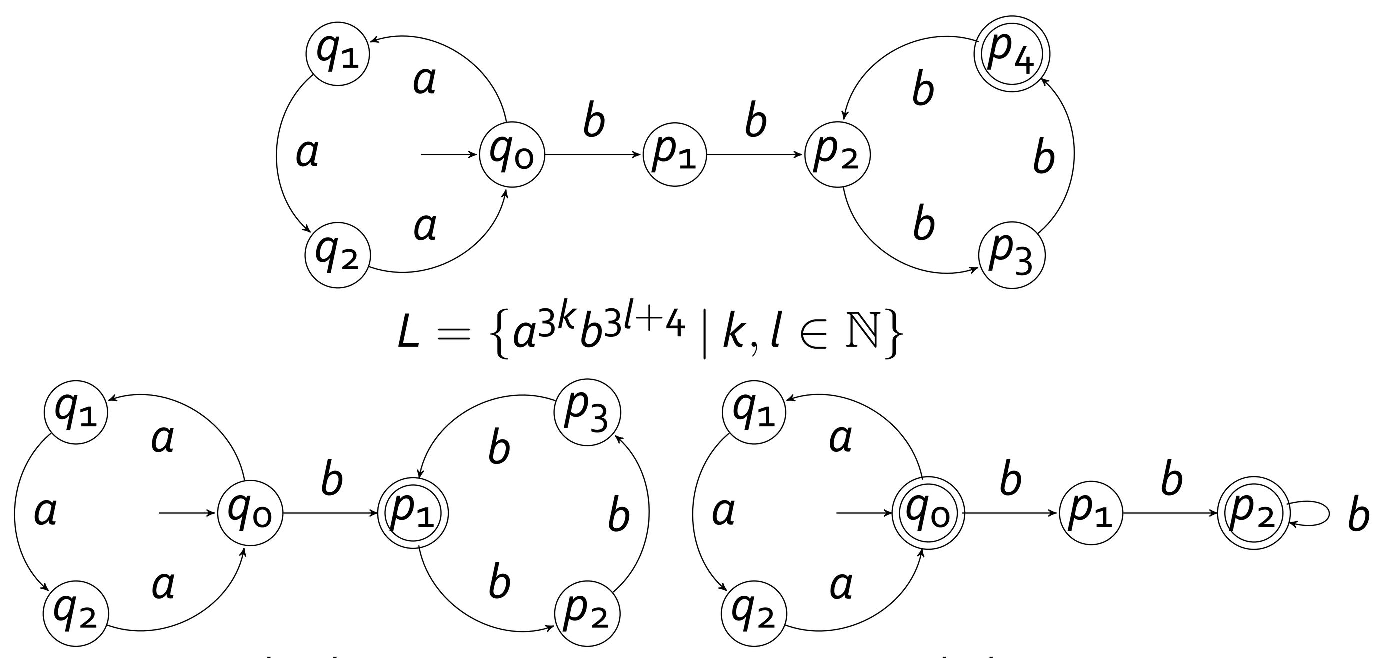
Languages bounded by a^*b^* [3]

Languages *bounded by* a^*b^* are languages that are a subset of a^*b^*.

Simple bicyclic languages are regular languages bounded by a^*b^*, that are a concatenation of two unary langauges, one over $\{a\}$ and one over $\{b\}$, i.e., $L = L^a L^b$, where $L^a \subseteq a^*$ and $L^b \subseteq b^*$.

Simple bicyclic languages have been characterized upon decomposability into languages bounded by a^*b^*. The sufficient and necessary conditions are related to the **decomposability of the two unary languages** in the concatenation as well as to the **structure of the automaton** accepting the language.

An example of such decomposition is in the following figure.



$L = \{a^{3k}b^{3l+4} \mid k, l \in \mathbb{N}\}$

$L_1 = \{a^{3k}b^{3l+1} \mid k, l \in \mathbb{N}\}$ $L_2 = \{a^{3k}b^{l} \mid k, l \in \mathbb{N}, l \neq 1\}$

Figure 2. A decomposition of a simple bicyclic language

For each regular language bounded by a^*b^* There exists a **partition into simple bicyclic languages**, i.e., $L = L_1 \cup \cdots \cup L_n$, where $L_1, \ldots, L_n$ are simple bicyclic.

Several necessary and sufficient conditions for decomposability have been found. These conditions are related to the **decomposability of the simple bicyclic languages** in the partition and to the **structure of the automaton** accepting the language.

Vincent Hlaváč, Andrej Ravinger, Branislav Rovan
Fakulta matematiky, fyziky a informatiky. Univerzita Komenského v Bratislave