Residuality and Learning for Register Automata

Presenter → Joshua Moerman
RWTH Aachen University

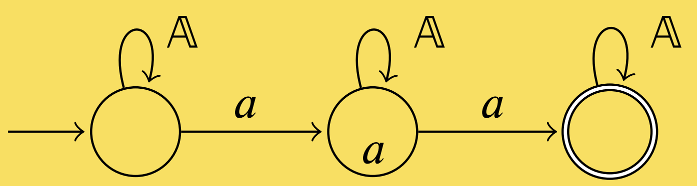
Join at Fr. 18 Sept. 10:52!

Matteo Sammartino ← Collaborator
Royal Holloway University London

REGISTER AUTOMATA & DATA LANGUAGES

Data languages are formal languages over **infinite alphabets**. Can model XML, resource allocations, data flow, etc. These languages can be accepted by **register automata**, i.e. automata with finite memory to store symbols.

Example: $\mathscr{L} = \{ uavaw \mid u,v,w \in \Sigma^*, a \in \Sigma \}$, i.e. some atom occurs twice. Can only be accepted nondeterministically!



For any $\mathscr{L}$, we define the **derivative** $w^{-1}\mathscr{L}$ as $\{ u \mid wu \in \mathscr{L} \}$. The set of all derivatives is $\mathrm{Der}(\mathscr{L})$.

RESULTS FOR LEARNING

L* automata learning for deterministic register automata is shown by [Sakamoto, 1997]. More recently, this topic has become popular and resulted in many phd theses.

In [Moerman et al, 2017] we show how to use nominal sets to generalise both L* and NL* to register automata. However, the class of automata for which nominal NL* worked, was left as **open problem**.

We investigate residual register automata. NL* can be modified to always terminate for RRA.

MAIN THEOREM:

Given a data language $\mathscr{L}$, TFAE:

1. $\mathscr{L}$ is accepted by a residual register automaton
2. There is an orbit-finite $J \subseteq \mathrm{Der}(\mathscr{L})$ which generates $\mathrm{Der}(\mathscr{L})$
3. The set $\mathcal{JI}(\mathrm{Der}(\mathscr{L}))$ is orbit-finite and generates $\mathrm{Der}(\mathscr{L})$

Canonical construction à la Myhill-Nerode

CONTRIBUTIONS

Canonical nondeterministic register automata / Learnability results / nominal lattice theory / Decidable universality

Joshua Moerman and Matteo Sammartino. Residual nominal automata. In CONCUR 2020.

THEORY & TECHNIQUES

Nominal sets / sets with atoms: very convenient theory for register automata

Lattice theory: join-irreducible elements

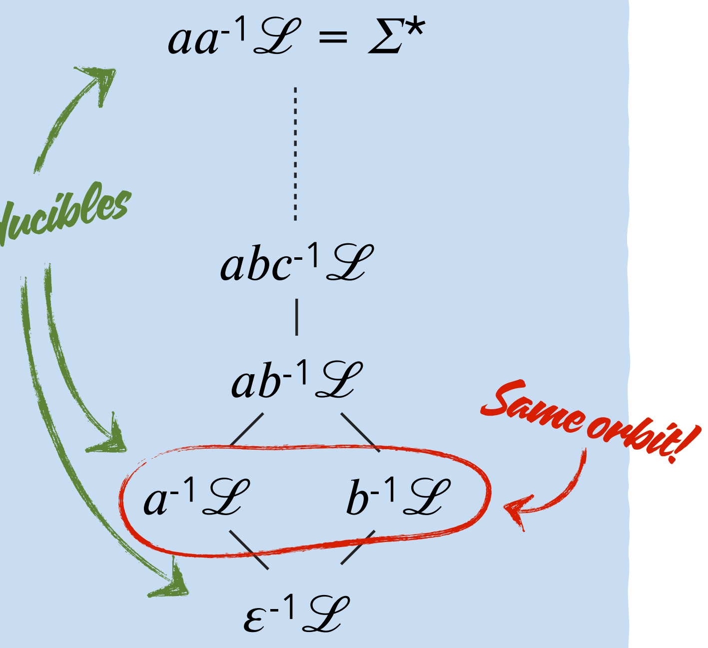
Residuality: every state accepts a derivative language:

$$\forall q \ \exists w \ \ \mathscr{L}(q) = w^{-1}\mathscr{L}(\mathcal{A})$$

Each state is *at some point* directly observable

EXAMPLE

Join irreducibles



Here $\mathrm{Der}(\mathscr{L})$ has **infinitely** many orbits. But we **only need 3** to generate all!