ω-regular languages and monadic decidability

Alexey Slizkov

Languages

Fix a finite alphabet Σ.

A *language* is any $L \subseteq \Sigma^*$.

e.g., $ab^* = \{a, ab, abb, abbb, \ldots\}$.

Automata

A *deterministic finite automaton* (DFA) is given by its

- a finite set Q of states, with a selected *initial state* $q_0 \in Q$
- a transition function $\delta\colon Q \times \Sigma \to Q$
- a set of *accepting states* $F \subseteq Q$.

Observation: Any word $w \in \Sigma^*$ can be uniquely *read* by the DFA, leading to a state $A(w) \in Q$. If $A(w) \in F$, we say that A *accepts* w, otherwise it *rejects* it.

Automata

We say that a A *recognizes* $L \subseteq \Sigma^*$, if

$L = \{w \in \Sigma^* \mid A \text{ accepts } w\}$.

Definition. A language $L \subseteq \Sigma^*$ is *regular*, iff it is the language of some DFA: $\exists A\ L = \{w \in \Sigma^* \mid A \text{ accepts } w\}$

Example 1: the language

$a(b|c)^*a = \{aa, aba, aca, abba, abca, acba, acca, \dots\}$ is regular.

Example 2: the language $\{a^n b^n \mid n \in \mathbb{N}\} = \{\varepsilon, ab, aabb, \dots\}$ is not regular (can be proved).

Nondeterministic automata

A *nondeterministic finite automaton* (NFA) is given by its

- a finite set Q of states, with a selected *intial state* $q_0 \in Q$
- a nondeterministic transition function $\delta \colon Q \times \Sigma \to \mathcal{P}Q$
- a set of *accepting states* $F \subseteq Q$.

We say that A accepts w, if *there exists* a path $q_0 \overset{w}{\rightsquigarrow} q_t$, $q_t \in F$.

DFAs and NFAs are equivalent!

A DFA-regular language L is clearly also NFA-regular. It turns out that the converse also holds:

The new set of states is $Q' := \mathcal{P}Q$, still finite. The transitions are the induced ones: for $q', q'' \in Q'$,

$$(q' \overset{a}{\mapsto} q'') :\iff q'' = A[\underbrace{q' \times \{a\}}_{\in \mathcal{P}Q \times \Sigma}].$$

Then the constructed automaton A' is deterministic, and $L(A) = L(A')$.

Regular languages

So, we could as well define regular languages via NFAs, we will get the same class. AFAs (alternating automata) also give the same class. Expressions with concatenation, disjunction, Kleene star, (disjunction) also give the same class. There are many theorems which all indicate that the class of regular languages is very robust and fundamental.

There are also excellent closure properties.

Closure properties of Reg

If A and B are regular languages, then:

- $A \cap B$ is also regular: the *product automaton* gives the corresponding DFA: $Q' := Q_A \times Q_B$,

$$\delta'((a,b),c) = (a',b') \overset{\triangle}{\iff} \left[\begin{array}{l} \delta_A(a,c) = a' \wedge \delta_B(b,c) = b', \\ \delta_A(a,c) = a' \wedge b = b', \\ a = a' \wedge \delta_B(b,c) = b' \end{array} \right. .$$

Closure properties of Reg

If A and B are regular languages, then:

- $A \cap B$ is also regular
- $A \cup B$ is also regular: we can merge the two automata into a new *nondeterministic* one

Closure properties of Reg

If A and B are regular languages, then:

- $A \cap B$ is also regular
- $A \cup B$ is also regular
- $\Sigma^* \backslash A$ is also regular: take $F' := Q \backslash F$. It is important to use *deterministic* automata here in the proof

Closure properties of Reg

If A and B are regular languages, then:

- $A \cap B$ is also regular
- $A \cup B$ is also regular
- $\Sigma^* \backslash A$ is also regular
- $A^* := \{\varepsilon\} \cup A \cup AA \cup \underbrace{AAA}_{\{w_1 w_2 w_3 | w_1 \in A \wedge w_2 \in A \wedge w_3 \in A\}} \cup \dots$ is also regular

Closure properties of Reg

If A and B are regular languages, then:

- $A \cap B$ is also regular
- $A \cup B$ is also regular
- $\Sigma^* \backslash A$ is also regular
- $A^* := \{\varepsilon\} \cup A \cup AA \cup \underbrace{AAA}_{\{w_1 w_2 w_3 | w_1 \in A \wedge w_2 \in A \wedge w_3 \in A\}} \cup \ldots$ is also regular
- people are happy

First-order logic

Recall that for a logical structure with some domain A and signature σ, the *first order logic* allows us to quantify over variables, e.g. (for A being a graph and σ containing only the edge relation):

$$\underbrace{\forall v\ \exists u\ (\forall w\ E(u,v) \wedge E(v,w) \vee u = w) \vee (\forall w\ E(w,w))}_{\text{this is an FO sentence}}.$$

This is weak enough to even have some reasonable chances to be decidable. For example, the FO theory of a free group is *decidable* (F_2, F_3, $F_{\aleph_0}$, they are elementarily equivalent, but this decidability is extremely hard to prove). But recall, say, PA...

Second-order logic

The second order allows us to quantify over *subsets* of A, and even *relations* of arity > 1. E.g., for two infinite graphs on $\mathbb{N}$ and $\mathbb{N}$, the property of them being isomorphic is SO-expressible ($\mathbb{D} = \mathbb{N} \sqcup \mathbb{N}$:

$$\exists f \subseteq \mathbb{D}^2 \left[\underbrace{(\forall u\ \exists! v\ (u, v) \in f)}_{f \text{ is a function graph}} \wedge \underbrace{(\forall v\ \exists! u\ (u, v) \in f)}_{f \text{ is a bijection}} \right.$$
$$\left. \wedge \underbrace{\forall u_1\ \forall u_2\ ((u_1, u_2) \in E_1 \iff (f(u_1), f(u_2)) \in E_2)}_{\text{gives a graph isomorphism}} \right].$$

SO logic is *extremely powerful*. Decidability is impossible completely: even *empty* signature over any infinite domain gives a very-very undecidable theory (Turing degree of SOA).

Monadic second-order logic

The *monadic* SO restricts us to quantify over 1-ary relations, i.e. subsets, only. Say, three-colorability of an infinite graph is MSO-expressible:

$$\exists A \subseteq V \, \exists B \subseteq V \, \exists C \subseteq V \left[\underbrace{(\forall v \, v \in A \wedge v \notin B \wedge v \notin C \vee \dots)}_{A \sqcup B \sqcup C = V} \right.$$

$$\left. \wedge \, \forall u \, \forall v \, (u, v) \in E \implies \neg(u \in A \wedge v \in A \vee \dots) \right]$$

(but this property is in fact firstorderizable, since 3-colorability of every finite subgraph is equivalent to 3-colorability of the whole graph, exercise)

So MSO is the largest decidability hope we can potentially have.

But of course, for most signatures, this is still insanely powerful, e.g., $\mathrm{MSO}(\mathbb{N}, +)$ can already interpret SOA.

S1S and wS1S

The *weak monadic* SO restricts us to quantify over finite subsets only.

S1S is the MSO theory of $(\mathbb{N}, \leq)$.

wS1S is the WMSO theory of $(\mathbb{N}, \leq)$.

Theorem. wS1S is decidable.

Decidability of wS1S

Proof sketch (I find the idea very beautiful)! (notice that equivalent to $(\mathbb{N}, +1)$, MSO is *powerful*. also, let's handle FO quantifiers as WMSO, with new predicate of being "element")

For every subformula of our given sentence, we will inductively construct the regular language it expresses, as a language over alphabet Σ^r:

$$\forall S\ \exists m\ (\forall T\ \ m \in S \lor \underbrace{T \subseteq S}_{(1|3)^* \subseteq 4^*} \lor \underbrace{\exists n\ \ \underbrace{m = n+1}_{0^*120^* \subseteq 4^*}}_{01^* \subseteq 2^*})$$

These are regular languages, so we will store them as their automata.

Decibability of wS1S

We have different cases depending on what type of subexpression we have:

- if we see conjunction, we intersect the languages
- if we see disjunction, we take the union of the languages
- $\exists$ projects the language into a smaller alphabet, remains regular
- $\neg$ corresponds to complements
- $\forall$ is just a combination of $\exists$ and $\neg$
- we know the languages for atomic predicates "$m = S(n)$" and "m is a signletone set"

This completes the proof (sketch)!

ω-regular languages

Definition. An ω-*language* is now a set of *infinite* strings: $L \subseteq \Sigma^\omega$

:)

L is called an ω-regular language, if it is recognized by a **nondeterministic** *Büchi automaton*.

A deterministic Büchi automaton is a finite automaton which eats ω-strings, with the following acceptance condition: **the set of terminal states is visited infinitely often**.

It turns out that deterministic and nondeterministic Büchi automata are not equivalent, and *it is important to use the more powerful, nondeterministic ones*. In this case, we will have the closure properties!

Theorem. S1S is decidable. The proof is overall similar to WS1S, but one of the clojure properties is highly nontrivial: the negation of an ω-regular language is an ω-regular language. The proof of this fact is beautiful and uses infinite Ramsey theorem, among others!

This is one of the strongest decidable theories known, and essentially the only monadic one (But **S2S** is even more powerful, and still decidable. It is the theory of infinite binary tree instead of an infinite "bamboo". The decidability was very hard to establish, was first proved by Rabin, but now a more natural approach is known. This uses tree automata, and even some Borel determinacy plays a role!)

Thank you for attention!!!

I hope you enjoyed the overview :)