Logic and Computation II

Part 6. Automata on infinite objects

Kazuyuki Tanaka

BIMSA

April 17, 2025

Logic and Computation II

- **Part 4. Modal logic**
- **Part 5. Modal μ-calculus**
- **Part 6. Automata on infinite objects**
- **Part 7. Recursion-theoretic hierarchies**

Part 6. Schedule (tentative)

- Apr.15, (1) Second-order arithmetic and analytical hierarchy
- Apr.17, (2) Büchi automata
- Apr.22, (3) The decidability of S1S
- Apr.24, (4) Tree automata
- May 6, (5) The decidability of S2S
- May 8, (6) Finite model theory
- May 13, (7) Parity games

Recap

- Second-order arithmetic Z_2 is a monadic second-order theory, or a two-sorted first-order theory dealing with natural numbers and sets of natural numbers under the condition of full induction and full comprehension.
- The language $\mathcal{L}^2_{\mathrm{OR}}$ of second-order arithmetic is the language of first-order arithmetic $\mathcal{L}_{\mathrm{OR}}$ plus the membership relation symbol $\in$.
- The **analytical hierarchy** of $\mathcal{L}^2_{\mathrm{OR}}$-formulas: For each $j \geq 0$, if $\varphi \in \Sigma^1_j$, then $\forall X_1 \cdots \forall X_k \varphi \in \Pi^1_{j+1}$; if $\varphi \in \Pi^1_j$ then $\exists X_1 \cdots \exists X_k \varphi \in \Sigma^1_{j+1}$. Here, Σ^1_0 and Π^1_0 are arithmetical formulas.
- **Analytical hierarchy $\mathrm{fnc}\text{-}\Sigma^1_n$, $\mathrm{fnc}\text{-}\Pi^1_n$, by function quantifiers**: For each $i \geq 0$, if φ is $\mathrm{fnc}\text{-}\Pi^1_i$, then $\exists f \varphi$ is $\mathrm{fnc}\text{-}\Sigma^1_{i+1}$. If φ is $\mathrm{fnc}\text{-}\Sigma^1_i$, then $\forall f \varphi$ is $\mathrm{fnc}\text{-}\Pi^1_{i+1}$.
For any Σ^1_i (or Π^1_i) formula , there exists an equivalent $\mathrm{fnc}\text{-}\Sigma^1_i$ (or $\mathrm{fnc}\text{-}\Pi^1_i$) formula and vice versa.
- **Normal form theorem for analytical formulas**: For each $i \geq 1$, for any Σ^1_i (or Π^1_i) formula , there exists an equivalent $\mathrm{fnc}\text{-}\Sigma^1_i$ (or $\mathrm{fnc}\text{-}\Pi^1_i$) formula whose arithmetical part is Σ^0_1 or Π^0_1.

Examples

- $\underbrace{\forall x \ f(x) \leq f(x+1)}_{f:\omega\to\omega \text{ is nondecreasing}} \in \Pi^0_1$
- $\underbrace{\exists b \ \forall x \ f(x) \leq b}_{f:\omega\to\omega \text{ is bounded}} \in \Sigma^0_2$
- $\exists X \Big((\forall n \exists m > n \ m \in X) \land f \text{ is bounded on } X \Big) \in \Sigma^1_1$
- $\underbrace{\neg (\exists f \forall x (f(x+1) \prec f(x)))}_{\prec \text{ is well-founded}} \in \text{fnc-}\Pi^1_1$

Example

Rewrite a Π^1_1 formula $\forall f \exists x \forall y \exists z R(x,y,z,f)$ in the normal form.

$$\begin{aligned}
\forall f \exists x \forall y \exists z R(x,y,z,f) &\Leftrightarrow \forall f \forall g \exists x \exists z R(x, g(x), z, f) \\
&\Leftrightarrow \forall f \forall g \exists x R(\pi_0(x), g(\pi_0(x)), \pi_1(x), f) \\
&\Leftrightarrow \forall f \exists x R(\pi_0(x), \pi_0 \circ f(\pi_0(x)), \pi_1(x), \pi_1 \circ f)
\end{aligned}$$

Definition 6.4 (The system of Recursive Comprehension Axioms)

RCA_0 consists of the following axioms.
(1) Basic Axioms of Arithmetic: Same as $\mathsf{Q}_<$.
(2) Δ^0_1 comprehension axiom (Δ^0_1-CA): For any $\varphi(x) \in \Sigma^0_1$ and $\psi(x) \in \Pi^0_1$,

$$\forall x(\varphi(x) \leftrightarrow \psi(x)) \to \exists X \forall x(x \in X \leftrightarrow \varphi(x)).$$

(3) For any $\varphi(x) \in \Sigma^0_1$, $\varphi(0) \land \forall x(\varphi(x) \to \varphi(x+1)) \to \forall x \varphi(x)$.

- RCA_0 is a conservative extension of first-order arithmetic $\mathsf{I}\Sigma_1$.

Definition 6.4 (The system of Arithmetical Comprehension Axioms)

ACA_0 is obtained from RCA_0 by replacing Δ^0_1 comprehension with Σ^0_1 comprehension [i].

- ACA_0 is a conservative extension of first-order arithmetic PA.

[i] Arithmetical comprehension can be achieved by repeatedly applying Σ^0_1 comprehension to parameters.

What is the normal form for the set quantifier hierarchies Σ^1_i, Π^1_i? The following lemma shows that the inner arithmetical part is not Σ^0_2 or Π^0_2.

Lemma 6.8 (Compactness)

For any Π^0_i formula $\varphi(X)$, the formula $\exists X\, \varphi(X)$ is Π^0_i $(i = 1)$.
For any Σ^0_i formula $\varphi(X)$, the formula $\exists X\, \varphi(X)$ is Σ^0_i $(i = 1, 2)$.

Proof.

- We identify a set X with the infinite binary sequence ξ representing its characteristic function. Then a Π^0_1 formula $\varphi(X)$ can be expressed as $\forall x\, R(\xi \restriction x)$ (R is primitive recursive).
  - Let T be a tree $\{t : \forall s \subseteq t\ R(s)\}$. T is also primitive recursive. We can see that $\varphi(X)$ is equivalent to $\xi \in [T]$, where $[T]$ is the set of all infinite paths of tree T.
  - Thus, $\exists X\, \varphi(X)$ is equivalent to $[T] \neq \varnothing$, which is equivalent to the Π^0_1 formula expressing that "T is infinite $(\forall n \exists t \in \{0,1\}^n t \in T)$".

§6.3. S1S and Büchi automata

- S1S is a restricted subsystem of manadic second order arithmetic only with a successor function.
- The decidability of S1S dates back to Büchi in the paper below.
  He translated S1S formulas to non-deterministic automata on infinite strings, now known as Büchi automata, which are also useful for formal verification.
- Today, we will introduce such automata, as well as other acceptance conditions for infinite strings, such as Muller and Rabin conditions.

ON A DECISION METHOD IN RESTRICTED SECOND ORDER ARITHMETIC

J. RICHARD BÜCHI

University of Michigan, Ann Arbor, Michigan, U.S.A.

Let SC be the interpreted formalism which makes use of individual variables t, x, y, z, . . . ranging over natural numbers, monadic predicate variables q( ), r( ), s( ), i( ), . . . ranging over arbitrary sets of natural numbers, the individual symbol 0 standing for zero, the function symbol ′ denoting the successor function, propositional connectives, and quantifiers for both types of variables. Thus SC is a fraction of the restricted second order theory of natural numbers, or of the first order theory of real numbers. In fact, if predicates on natural numbers are interpreted as binary expansions of real numbers, it is easy to see that SC is equivalent to the first order theory of [Re, +, Pw, Nn], whereby Re, Pw, Nn are, respectively, the sets of non-negative reals, integral powers of 2, and natural numbers.

J.R. Büchi

Büchi automata

Let Ω be a finite set (alphabet) and Ω^ω be the set of ω- words $a_0a_1a_2\cdots$ on Ω. If $|\Omega| > 1$ then Ω^ω is uncountable and has the same cardinality as the real numbers.

Definition 6.9

A **nondeterministic Büchi automaton** (NBA) is a 5-tuple $M = (Q, \Omega, \delta, Q_0, F)$,

(1) Q is a non-empty finite set, whose elements are called **states**.

(2) Ω is a non-empty finite set, whose elements are called **symbols**.

(3) $\delta : Q \times \Omega \to \mathcal{P}(Q)$ is a **transition relation**. $\mathcal{P}(Q)$: the power set of Q.

(4) $Q_0 \subset Q$ is a set of **initial states**.

(5) $F \subset Q$ is a set of **final states**.

$(p, a, q) \in \delta$ represents that M can make a transition from state p to state q for input a. M is **deterministic** (DBA) if δ is a single-valued function (i.e., $\delta : Q \times \Omega \to Q$) and Q_0 is a singleton set.

An infinite run

- A **run** of M on an input ω- word

$$\alpha = a_0 a_1 a_2 \cdots \in \Omega^\omega$$

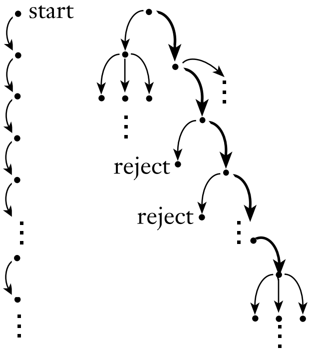
is an infinite sequence of states

$$q_0 q_1 q_2 \cdots \in Q^\omega$$

satisfying:

  - $q_0 \in Q_0$,
  - $(q_i, a_i, q_{i+1}) \in \delta \ (i \geq 0)$.

- If M is deterministic then there is a unique run for any input word.
- If M is non-deterministic, there may be no runs or many runs for an input ω-word, even uncountable many runs.



Accepted of run, word and language

- For an infinite run σ, the set of states that appear infinitely in σ is denoted by $\mathrm{Inf}(\sigma)$. In other words, if $\sigma = q_0 q_1 q_2 \cdots$,

$$\mathrm{Inf}(\sigma) = \bigcap_{n \geq 0} \{q_i \mid i \geq n\}.$$

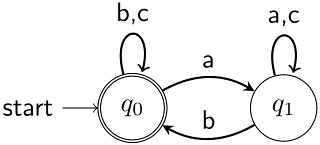
- An infinite run σ is said to be **accepted by** NBA M if $\mathrm{Inf}(\sigma) \cap F \neq \varnothing$, that is, if a state of F occurs infinitely many times in σ.
- An input word α is **accepted by** NBA M if there is an accepted run on α.
- Thus, the **ω-language** $L(M) \subset \Omega^\omega$ **accepted by** by M is defined as

$$L(M) = \{\alpha \in \Omega^\omega \mid \text{there is a run } \sigma \text{ of } M \text{ on } \alpha \text{ such that } \mathrm{Inf}(\sigma) \cap F \neq \varnothing\}.$$

Example

(1) There exists an DBA $M = (Q, \Omega, \delta, q_0, F)$ accepting the following language.

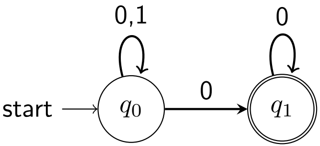
$$\{\alpha \in \{a,b,c\}^\omega \mid \forall n(\alpha(n) = a \to \exists m > n\ \alpha(m) = b)\}.$$



where $Q = \{q_0, q_1\}$, $\Omega = \{a,b,c\}$, $F = \{q_0\}$.

Example

(2) The following NBA accepts the set $(0+1)^*0^\omega$, where "1" appears finitely times.



where $Q = \{q_0, q_1\}$, $\Omega = \{0, 1\}$, $F = \{q_1\}$.

- Note that non-determinism of the Büchi automaton is necessary to guess when the last "1" appears so that the automaton can move to loop in q_1 with input always 0.
- In fact, this language cannot accepted by any DBA.

Definition 6.10

A language accepted by an NBA is called an **ω-regular language**.

Theorem 6.11

The following are equivalent.

- L is an ω-regular language.
- $L = \bigcup_{i<n} U_i V_i^\omega$ for some regular languages $U_i (\subset \Omega^*), V_i \subset (\Omega^+)$.

Proof ($\Rightarrow$) Let $M = (Q, \Omega, \delta, Q_0, F)$ be an NBA that accepts L. By $W_{qq'}$, we denote the language accepted by the finite automaton $M = (Q, \Omega, \delta, \{q\}, \{q'\})$ with the empty word removed, i.e.,

$$W_{qq'} = \{w \in \Omega^+ : q' \in \bar{\delta}(q, w)\}.$$

Each $W_{qq'}$ is clearly regular. And L can be expressed as follows.

$$L = \bigcup_{q_f \in F} W_{q_0 q_f} (W_{q_f q_f})^\omega.$$

($\Leftarrow$) U^ω is ω-regular if U is regular (Consider a finite automaton that accepts U^* as NBA). If U is regular and V is ω-regular, then UV is also ω-regular. If L_i is ω-regular, then $\bigcup_{i \in n} L_i$ is also ω-regular.

Theorem 6.12

The emptiness problem for ω-regular languages is decidable.

Proof.
The empty decision problem is to decide $L(M) \neq \varnothing$. By the theorem in the last page, it is equivalent to decide $\exists i((U_i \neq \varnothing) \wedge (V_i \neq \varnothing))$, which reduces to the emptiness problem of regular languages. The emptiness of regular language is decidable, e.g., from the regular expressions. Thus the emptiness problem for ω-regular languages is decidable.

□

Remark

- The non-emptiness of NBA M is equivalent to reach from some initial state q_0 to some final state q_f and return to q_f infinite many times.
- Therefore, this is a variant of the STconnect problem, which is decidable in polynomial time.

- To show the decidability of S1S, we consider automata over ω-words, that is, NBA. If we show the equivalence of their expressiveness, we can derive the decidability of S1S from the decidability of emptiness of NBA.
- Before proving their equivalence, we need to show the class of ω-regular languages is closed under Boolean operations. It is easy to see the class of ω-regular languages is closed under $\cup$ and $\cap$. The difficulty lies in the closure under complement.
- If an ω-regular language were accepted by a DBA, so is its complement. But, as in the example above, not all ω-regular languages are accepted by some DBA.
- Therefore, we need to consider Muller and Rabin automata, which are stronger than Büchi ones, but whose deterministic machines can imitate non-deterministic ones.

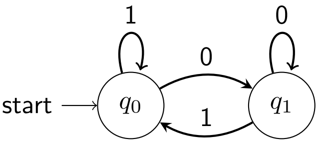
Muller condition

- The acceptance condition of a **Muller automaton** is given by $\mathcal{F} \subseteq \mathcal{P}(Q)$, and a run is accepted iff $\mathrm{Inf}(\sigma) \in \mathcal{F}$.
- Büchi condition $(\mathrm{Inf}(\sigma) \cap F \neq \varnothing)$ can be expressed in terms of the Muller condition

$$\mathcal{F} = \{A \subseteq Q \mid A \cap F \neq \varnothing\}.$$

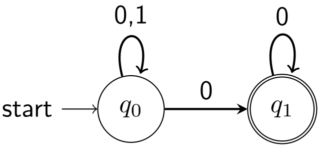
- Non-deterministic / deterministic Muller automata are abbreviated as NMA / DMA.

An DMA accepting $L = (0+1)^*0^\omega$



where $\mathcal{F} = \{\{q_1\}\}$.

An equivalent NBA

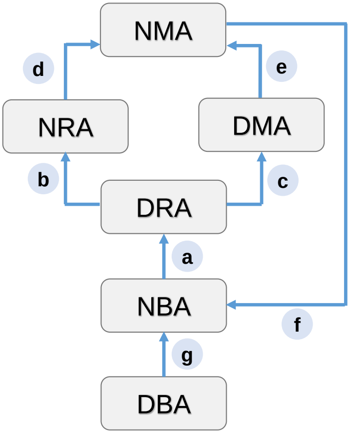


where $F = \{q_1\}$.

Rabin condition

- The acceptance condition of a **Rabin automaton** is given by
$$\mathcal{F} = \{(G_i, R_i) \mid (1 \leq i \leq k)\},$$
where $G_i, R_i \subset Q$.
- A run σ is **accepted**, if there exists i such that $\mathrm{Inf}(\sigma) \cap G_i \neq \varnothing$ and $\mathrm{Inf}(\sigma) \cap R_i = \varnothing$.
- Non-deterministic / deterministic Rabin automata are abbreviated as NRA / DRA.
- When a G_i/R_i state is visited, we say that the i-th green/red signal is on. A green signal is expected to turn on infinitely many times but a red signal only finitely many.
- A Büchi automaton can be simulated by a Rabin automaton with
$$k = 1,\ G_1 = F,\ R_1 = \varnothing.$$
- A Rabin automaton turns into a Muller automaton if
$$\mathcal{F} = \{A \subseteq Q \mid \bigvee_i (A \cap G_i \neq \varnothing \wedge A \cap R_i = \varnothing)\}$$

- (b), (e) and (g) are obvious. (c) and (d) have been explained above. Now, we are going to show (f).
  - Let M be an NMA with an accepting set $\mathcal{F}$. Goal: construct an NBA N to simulate M.
  - For input x, N mimics M by nondeterministically guessing a run σ of M on x.
  - At some point, N nondeterministically predicts that all states of M not in $\mathrm{Inf}(\sigma)$ have appeared and also guesses that $\mathrm{Inf}(\sigma)$ is a certain set $A \in \mathcal{F}$.
  - Then check if A is indeed $\mathrm{Inf}(\sigma)$ as follows:
    - Any state of σ (from that point) is in A, and
    - Let s be the state of N representing that every state of A appeared at least once. Then N accepts the input if s appears inf. many times.
- (a):NBA$\to$ DRA is the most difficult to prove.



In the figure, "XXA $\to$ YYA" means "for any XXA M_1, there exists a YYA M_2 such that $L(M_1) = L(M_2)$".

NBA$\to$ DRA

- It was first prove by McNaughton in 1966, but his construction was doubly exponential. Safra propose a more efficient exponential construction in 1988.

NBA

Given $B = (Q, \Omega, \delta, Q_0, F)$ with $\mid Q \mid = n$

DRA

We want to construct a deterministic Rabin automaton

$$R = (S, \Omega, \delta', S_0, \{(G_1, R_1), (G_2, R_2) \cdots (G_{2n}, R_{2n})\})$$

that accepts the same language.

Goal (Safra's Theorem)

Any NBA with n states can be simulated with a DRA consisting of $2^{O(n \log n)}$ states and n pairs of acceptance conditions. Therefore, it can also be simulated with a DMA with the same number of states.

Thank you for your attention!