a logic for synchronous relations

varun ramanathan

ANR DeLTA Meeting, LaBRI 2018

joint work with Pascal Weil and Diego Figueira

overview

1. introduction
   - languages and relations
   - synchronous relations
   - logical formalism
2. the Σ_1 fragment
   - characterization
   - decidability
3. conclusion
   - future work

languages and relations

regular languages of finite words are well studied. We use different tools: automata, logic, algebra. But what about relations?

Reverse: $\{(a, a), (b, b) \ldots (baaba, abaab) \ldots\}$

Prefix: $\{(a, a), (b, b) \ldots (abba, abbababb) \ldots\}$

Subword: $\{(a, a) \ldots (aba, bbaabbbab) \ldots\}$

What are some automata models for studying relations?

Rational relations: multi-tape automata with regular synchronizing language

Synchronous relations: multi-tape automata with synchronous movement of heads (i.e. synchronizing language $(12)^*(1^* + 2^*)$).

synchronous relations

Studied under the name *regular relations* (Libkin, et al, 2003) and *automatic relations* (Blumensath, Grädel, 2000).

Applications: Extended Conjunctive Regular Path Queries (ECPRQs) for graph databases.

Natural question: logical characterization?

Starting point: FO characterization given by Eilenberg et al (1969).

synchronizing a tuple of words

Representing a tuple of words as a single word:

Let $\perp \notin A$, and $A_\perp = A \cup \{\perp\}$. *Synchronize* words as follows:

$$(aabab, aba) \longrightarrow \begin{pmatrix} a \\ a \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} \begin{pmatrix} a \\ \perp \end{pmatrix} \begin{pmatrix} b \\ \perp \end{pmatrix}$$

$$(bbb, ababab) \longrightarrow \begin{pmatrix} b \\ a \end{pmatrix} \begin{pmatrix} b \\ b \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} \begin{pmatrix} \perp \\ b \end{pmatrix} \begin{pmatrix} \perp \\ a \end{pmatrix} \begin{pmatrix} \perp \\ b \end{pmatrix}$$

$$(aab, abaabb, baab) \longrightarrow \begin{pmatrix} a \\ a \\ b \end{pmatrix}\begin{pmatrix} a \\ b \\ a \end{pmatrix}\begin{pmatrix} b \\ a \\ a \end{pmatrix}\begin{pmatrix} \perp \\ a \\ b \end{pmatrix}\begin{pmatrix} \perp \\ b \\ \perp \end{pmatrix}\begin{pmatrix} \perp \\ b \\ \perp \end{pmatrix}$$

the *synchronizing word* of $(w_1, w_2 \dots w_k)$ is denoted by $w_1 \otimes w_2 \dots w_k$ or simply $\bar{w}$.

$\mathbf{R}$ is a synchronous relation if

$$L_{\mathbf{R}} = \{(w_1 \otimes w_2 \cdots \otimes w_k : (w_1, w_2 \dots w_k) \in \mathbf{R}\} \text{ is regular}$$

synchronous relations

$Sync_k(A^*)$: set of synchronous k-ary relations (of words in A^*)

$$\mathbf{Sync}(A^*) = \bigcup_{k \in \mathbb{N}} Sync_k(A^*)$$

Let $A_\Delta = \{a_1 \otimes a_1 \colon a_1 \in A\}$, $A_\times = A \times A$.

- the prefix relation is given by $A_\Delta^*(\{\perp\} \times A)^*$
- the equal length relation is given by $A_\times^*$.

logical formalism for synchronous relations

First order logic with predicate set $\sigma_A = (\preceq, eq, \{\ell_a\}_{a \in A})$.

- $(w_1, w_2) \vDash x \preceq y$ iff w_1 is a prefix of w_2
- $(w_1, w_2) \vDash eq(x, y)$ iff $|w_1| = |w_2|$.
- $w \vDash \ell_a(x)$ iff $w \in A^* a$

Using usual FO semantics, we can describe relations with $\mathbf{FO}[\sigma]$ formulae.

$$(w_1, w_2, w_3) : \forall u((u \preceq w_1 \wedge u \preceq w_2) \iff u \preceq w_3)$$
$$\|\varphi\| = \{(w_1, w_2, w_3) \colon w_3 = w_1 \sqcap w_2\}$$

Theorem (Eilenberg, Shepherdson, Elgot '69)

Let A be an alphabet with at least 2 letters. Then

$$\mathbf{Sync}(A^*) = \|\mathbf{FO}[\sigma_A]\|$$

from FO formulae to synchronous relations

for all $\varphi \in \mathsf{FO}[\sigma_A]$, $\|\varphi\| \in \mathbf{Sync}(A^*)$. Use induction on the structure of φ.

base case: Recall that the prefix and equal length relations are synchronous; $|\ell_a(x)| = A^*a$ is regular hence synchronous.

induction step: Synchronized languages are closed under Boolean set operations. Furthermore, if

$$\varphi(y_1, y_2 \ldots y_{m-1}) : \exists y_m\, \psi(y_1, y_2 \ldots y_m)$$

then

$$\mathbf{R}_\varphi = \pi_{\{1\ldots(m-1)\}}(\mathbf{R}_\psi)$$

Synchronous relations are closed under projections on any subset of the components.

Therefore, $\|\mathsf{FO}[\sigma_A]\| \subseteq \mathsf{Sync}(A^*)$.

from synchronous relations to **FO**–formulae

For the converse, let

- $\mathbf{R}$ synchronous relation
- $\mathscr{A}$ automaton for $L_{\mathbf{R}}$ with states $Q = \{q_1, q_2 \ldots q_n\}$ and transition relation δ.

Let ρ be a run of $\mathscr{A}$ on some synchronized word $\bar{w}$.
idea: to encode ρ as a synchronized word. For every $q_i \in Q$, u_i is a word in $\{0,1\}^*$ such that:

- $u_i[j] = 1$ if ρ contains q_i at position j, 0 otherwise.
- $|u_i| = |\rho|$

Then $\bar{u}_\rho = u_1 \otimes u_2 \ldots u_n$.

For example if ρ is $q_1 \to q_2 \to q_1 \to q_3 \to q_4$ then

$$\bar{u}_\rho = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

We want a formula $\varphi(\vec{x})$ of the form

$$\exists \vec{y}(\psi_{\mathrm{enc}}(\vec{y}) \wedge \psi_{\mathrm{run}}(\vec{x}, \vec{y}))$$

where

- $\psi_{\mathrm{enc}}(\vec{y})$: for encoding
- $\psi_{\mathrm{run}}(\vec{y}, \vec{x})$: for stating the encoding is a valid run. (Here, due to the "next state" formula using the successor relation, ψ_{run} is a Π_2 formula.)

Therefore $\varphi \in \Sigma_3[\sigma_A]$.

Corollary

$\mathbf{Sync}(A^*) = \|\mathbf{FO}[\sigma_A]\| = \|\Sigma_3[\sigma_A]\|$.

the quantifier alternation hierarchy

FO collapses to Σ_3. But what about Σ_1?

- Characterization: What sort of relations are Σ_1 definable?
- Membership: Given a **Sync** relation is it Σ_1–definable?
- What about Σ_2, $\mathscr{B}\Sigma_1$, Δ_2?

Let's jog our memories and go back to the classical FO[<] world.

FO on words: Let B be an alphabet. Recall the first order logic $\mathbf{FO}[<, \mathbf{B}]$ finite total order($<$) and letter predicates ($\mathbf{B} = \{\mathbf{b}\}_{b \in B}$).

$\Sigma_1[<, \mathbf{B}]$ sentences define finite unions of languages of the form

$$B^* b_1 B^* b_2 \ldots B^* b_n B^*$$

where $b_1, b_2 \ldots b_n \in B$.

$\Sigma_1[<, \mathbf{B}, \mathit{last}]$ sentences define finite unions of languages of the form

$$B^* b_1 B^* b_2 \ldots B^* b_n$$

$w = b_1 b_2 \ldots b_n$ is called a *subword* of w', ($w \sqsubseteq w'$) iff $w' \in B^* b_1 B^* b_2 \ldots B^* b_n B^*$

back to relations: $\Sigma_1[\sigma_A]$–definable relations are connected to classical $\Sigma_1[<]$–definable languages.

Σ_1 definable relations

Recall the alphabets

$$A_\Delta = \{a \otimes a : a \in A\} \qquad A_1 = A \times \{\perp\}$$
$$A_\times = A \times A \qquad A_2 = \{\perp\} \times A$$

Theorem

If $\varphi(x_1, x_2) \in \Sigma_1[\sigma]$, then synchronization language of $\|\varphi\|$ is a finite union of languages of the form

$$L_\Delta L_\times L_\perp$$

where

- $L_\Delta \in \|\Sigma_1[<, \mathbf{A}_\Delta]\|$
- $L_\times \in \|\Sigma_1[<, \mathbf{A}_\times, last]\|$
- $L_\perp \in \|\Sigma_1[<, \mathbf{A}_1, last]\| \cup \|\Sigma_1[<, \mathbf{A}_2, last]\|$

Key idea: The notion of subwords can be generalized from words to pair of words.

synchronized subwords

$$\bar{w} = \binom{b}{b}\binom{a}{a}\binom{a}{a}\binom{a}{b}\binom{b}{b}\binom{b}{a}\binom{\perp}{a}\binom{\perp}{b}$$

$$\downarrow$$

$$\underbrace{\binom{b}{b}\binom{a}{a}\binom{a}{a}}_{A_\Delta^* \text{ part}} \cdot \underbrace{\binom{a}{b}\binom{b}{b}\binom{b}{a}}_{A_\times^* \text{ part}} \cdot \underbrace{\binom{\perp}{a}\binom{\perp}{b}}_{A_\perp^* \text{ part}}$$

$$\downarrow$$

$$\underbrace{\binom{a}{a}\binom{b}{b}\binom{b}{b}\binom{a}{a}\binom{b}{b}\binom{a}{a}\binom{b}{b}}_{A_\Delta^* \text{ superword}} \cdot \underbrace{\binom{b}{b}\binom{a}{b}\binom{a}{b}\binom{b}{b}\binom{b}{a}}_{A_\times^* \text{ superword}} \cdot \underbrace{\binom{\perp}{b}\binom{\perp}{a}\binom{\perp}{a}\binom{\perp}{a}\binom{\perp}{b}}_{A_\perp^* \text{ superword}}$$

$$\downarrow$$

$$\bar{w}' = \binom{a}{a}\binom{b}{b}\binom{b}{b}\binom{a}{a}\binom{b}{b}\binom{a}{a}\binom{b}{b}\binom{b}{a}\binom{a}{b}\binom{a}{b}\binom{b}{b}\binom{b}{a}\binom{\perp}{b}\binom{\perp}{a}\binom{\perp}{a}\binom{\perp}{a}\binom{\perp}{b}$$

then $\bar{w}$ is a synchronized subword of $\bar{w}'$, denoted by $\bar{w} \lhd \bar{w}'$.

some corollaries

$\Sigma_1[<]$ sentences describe words which contain some specific subwords.

$\Sigma_1[\sigma_A]$ formulae describe pairs whose synchronized word contains specific synchronized subwords.

$\mathcal{B}\Sigma_1[<]$ sentences describe words having the same set of bounded subwords.

$\mathcal{B}\Sigma_1[\sigma_A]$ formulae describe pairs whose synchronized words have the same set of bounded synchronized subwords.

Corollaries:

- binary $\|\Sigma_1[\sigma_A]\|$ membership decidable
- binary $\mathcal{B}\Sigma_1[<]$ membership decidable

conclusion and future work

To conclude

- $\mathbf{FO}[\sigma_A]$ completely characterizes synchronous relations
- $\mathbf{FO}[\sigma_A]$ collapses to its Σ_3 fragment
- Σ_1 and $\mathscr{B}\Sigma_1$ membership decidable in the binary case

Future work:

- Generalizing the Σ_1 characterization to greater arities.
- Characterizing $\|\Sigma_2\|$ relations as projections of $\|\Pi_1\|$ relations.
- Undecidable to check if a given rational relation is synchronous. But is it decidable to check if it is in Σ_1? This is unknown.

$$\begin{pmatrix} T \\ y \end{pmatrix} \begin{pmatrix} h \\ o \end{pmatrix} \begin{pmatrix} a \\ u \end{pmatrix} \begin{pmatrix} n \\ \perp \end{pmatrix} \begin{pmatrix} k \\ \perp \end{pmatrix}$$