universität
innsbruck

An Isabelle/HOL Formalization of Semi-Thue and Conditional Semi-Thue Systems

Dohan Kim

University of Innsbruck, Austria

16th International Conference on Interactive Theorem Proving (ITP 2025), Reykjavik, Iceland, Sept. 29, 2025

Contents

1. Overview of Semi-Thue Systems

2. Inference System for Completion of Semi-Thue Systems

3. Conditional Semi-Thue Systems

4. Types and Properties of Conditional Semi-Thue Systems

5. An Isabelle/HOL Formalization

6. Conclusion

Overview of Semi-Thue Systems

Figure: Axel Thue (1863-1922)

Semi-Thue (String Rewriting) Systems

It is generally accepted that Thue first introduced semi-Thue systems, also known as string rewriting systems in 1910's to solve the word problem of semigroups/monoids.

Overview of Semi-Thue Systems

Semi-Thue Systems

- A *semi-Thue system* (STS) $\mathcal{R}$ over a finite alphabet Σ is a subset of $\Sigma^* \times \Sigma^*$.
- Notation: $u \to_{\mathcal{R}} v$ denotes that u rewrites to v via a rule from $\mathcal{R}$. Also, $\to_{\mathcal{R}}^*$ denotes the reflexive, transitive closure of $\to_{\mathcal{R}}$.
- Example: $\Sigma = \{a, b\}$ and $\mathcal{R} = \{(ab, ba), (aa, a)\}$.
  - Then, $abab \to_{\mathcal{R}} baab \to_{\mathcal{R}} bab \to_{\mathcal{R}} bba$.
- *Thue congruence* induced by $\mathcal{R}$ is the relation $\overset{*}{\leftrightarrow}_{\mathcal{R}}$.
- In the above example, $abab \overset{*}{\leftrightarrow}_{\mathcal{R}} bba$.
- The congruence class $[w]_{\mathcal{R}}$ of a word $w \in \Sigma^*$ is defined as $[w]_{\mathcal{R}} := \{v \in \Sigma^* \mid w \overset{*}{\leftrightarrow}_{\mathcal{R}} v\}$.
- In the above example, $[abab]_{\mathcal{R}} := \{abab, baab, bab, bba, \ldots\}$.

Monoids

Monoids and factor monoids

- Monoids are fundamental algebraic structures widely used in mathematics and computer science.
- A monoid is a set equipped with an associative binary operation and a (two-sided) identity element.
- Σ^* is a monoid, called the *free monoid*.
  - Elements: all finite (possibly empty) strings from Σ
  - Operation: string concatenation
  - Identity: empty string ε
- Given a semi-Thue system $\mathcal{R}$ over Σ, $M_{\mathcal{R}} := \Sigma^* / \overset{*}{\leftrightarrow}_{\mathcal{R}}$ is a monoid, called the *factor monoid* of Σ^* modulo $\overset{*}{\leftrightarrow}_{\mathcal{R}}$.
  - Elements: Congruence classes $[w]_{\mathcal{R}}$ in $\{[w]_{\mathcal{R}} \mid w \in \Sigma^*\}$.
  - Operation: A binary operation $\cdot$ such that $[u]_{\mathcal{R}} \cdot [v]_{\mathcal{R}} = [uv]_{\mathcal{R}}$, where $u, v \in \Sigma^*$.
  - Identity: $[\varepsilon]_{\mathcal{R}}$

Word problem of finitely presented monoids and groups

Word problem of finitely presented monoids

- Given a semi-Thue system $\mathcal{R}$ over Σ, a monoid M is *finitely presented* by $(\Sigma; \mathcal{R})$ if $M \cong M_{\mathcal{R}}$ and both Σ and $\mathcal{R}$ are finite.
- The *word problem* of the monoid $M = (\Sigma; \mathcal{R})$ is the following decision problem: Given two words $u, v \in \Sigma^*$, decide if $u = v$ in M.
- The word problem of a finitely presented monoid $M = (\Sigma; \mathcal{R})$ is undecidable in general, but it is decidable if $\to_{\mathcal{R}}$ is confluent and terminating (i.e., complete).

Groups

- A group is a monoid in which every element is invertible. More specifically,
- A monoid $M = (\Sigma; \mathcal{R})$ is a group if for all $u \in \Sigma$, there is some $v \in \Sigma^*$ such that $uv \overset{*}{\leftrightarrow}_{\mathcal{R}} \varepsilon$ and $vu \overset{*}{\leftrightarrow}_{\mathcal{R}} \varepsilon$.

Completion for Finitely Presented Monoids

Purposes

- If a monoid is presented by a finite terminating semi-Thue system $\mathcal{R}$ but it is not confluent, then one may attempt to construct a finite complete semi-Thue system $\mathcal{R}'$ equivalent to $\mathcal{R}$ using a completion procedure, where $\mathcal{R}$ is confluent if whenever $s \overset{*}{\to}_{\mathcal{R}} u$ and $s \overset{*}{\to}_{\mathcal{R}} v$, there is w such that $u \overset{*}{\to}_{\mathcal{R}} w$ and $v \overset{*}{\to}_{\mathcal{R}} w$.
- It is known that an STS S on Σ^* can be associated with a term rewriting system (TRS) $\mathcal{R}_S$ in such a way that $\mathcal{R}_S := \{\ell(x) \to r(x) \,|\, \ell \to r \in S\}$, where each letter from an alphabet Σ is interpreted as a unary function symbol.
- Knuth-Bendix algorithms for STSs are already known in the literature. There are also known inference systems (e.g. abstract completion) for completion of TRSs.
- Proposed approach: Uses an inference system for completion of STSs (instead of using algorithms). Uses simple linear time string matching and length-lexicographic order (shortlex order) instead of using unification and more complex term ordering.

Inference System for Completion of Semi-Thue Systems

- In the following, $\mathcal{E}$ is a set of equations on Σ^* and $\mathcal{R}$ is a set of string rewriting rules on Σ^* such that for each $(\ell, r) \in \mathcal{R}$, $\ell >_{sl} r$, where $>_{sl}$ is a shortlex order.
- Observe that for each $(\ell, r) \in \mathcal{R}$, ℓ cannot be the empty string because it is not the case that $\varepsilon >_{sl} s$ for any $s \in \Sigma^*$.

Deduce

$$\frac{(\mathcal{E}, \mathcal{R})}{(\mathcal{E} \cup \{s u_3 \approx u_1 t\}, \mathcal{R})}$$

if $(u_1 u_2, s) \in \mathcal{R}$, $(u_2 u_3, t) \in \mathcal{R}$, and $u_2 \neq \varepsilon$.

Simplify

$$\frac{(\mathcal{E} \cup \{u_1 u_2 u_3 \approx s\}, \mathcal{R})}{(\mathcal{E} \cup \{u_1 t u_3 \approx s\}, \mathcal{R})}$$

if $(u_2, t) \in \mathcal{R}$.

Orient

$$\frac{(\mathcal{E} \cup \{s \approx t\}, \mathcal{R})}{(\mathcal{E}, \mathcal{R} \cup \{(s,t)\})} \quad \text{if } s >_{sl} t.$$

Collapse

$$\frac{(\mathcal{E}, \mathcal{R} \cup \{(u_1 u_2 u_3, s)\})}{(\mathcal{E} \cup \{u_1 t u_3 \approx s\}, \mathcal{R})} \quad \text{if } (u_2, t) \in \mathcal{R}.$$

Compose

$$\frac{(\mathcal{E}, \mathcal{R} \cup \{(s, u_1 u_2 u_3)\})}{(\mathcal{E}, \mathcal{R} \cup \{(s, u_1 t u_3)\})} \quad \text{if } (u_2, t) \in \mathcal{R}.$$

Delete

$$\frac{(\mathcal{E} \cup \{s \approx s\}, \mathcal{R})}{(\mathcal{E}, \mathcal{R})}$$

Inference System for Completion of Semi-Thue Systems

Fair non-failing run

- We write $(\mathcal{E},\mathcal{R}) \vdash_{SR} (\mathcal{E}',\mathcal{R}')$ if $(\mathcal{E}',\mathcal{R}')$ can be obtained from $(\mathcal{E},\mathcal{R})$ by applying one of the inference rules in the inference system.
- A (finite) *run* for an initial set of equations $\mathcal{E}$ is a finite sequence $(\mathcal{E}_0,\mathcal{R}_0) \vdash_{SR} (\mathcal{E}_1,\mathcal{R}_1) \vdash_{SR} \cdots \vdash_{SR} (\mathcal{E}_n,\mathcal{R}_n)$, where $\mathcal{E}_0 = \mathcal{E}$ and $\mathcal{R}_0 = \varnothing$.
- A run $(\mathcal{E}_0,\mathcal{R}_0) \vdash_{SR} (\mathcal{E}_1,\mathcal{R}_1) \vdash_{SR} \cdots \vdash_{SR} (\mathcal{E}_n,\mathcal{R}_n)$ *fails* if $\mathcal{E}_n \neq \varnothing$.
- A run $(\mathcal{E}_0,\mathcal{R}_0) \vdash_{SR} (\mathcal{E}_1,\mathcal{R}_1) \vdash_{SR} \cdots \vdash_{SR} (\mathcal{E}_n,\mathcal{R}_n)$ is *fair* if $CP(\mathcal{R}_n) \subseteq \bigcup_{i=0}^{n} \leftrightarrow_{\mathcal{E}_i}$.

Correctness

For every fair non-failing run $(\mathcal{E}_0,\mathcal{R}_0) \vdash_{SR} (\mathcal{E}_1,\mathcal{R}_1) \vdash_{SR} \cdots \vdash_{SR} (\mathcal{E}_n,\mathcal{R}_n)$, $\rightarrow_{\mathcal{R}_n}$ is a complete STS for an initial finite set of equations $\mathcal{E} = \mathcal{E}_0$ on Σ^*.

Conditional Semi-Thue Systems (CSTSs) [Deiß, 1992]

Purposes

- Conditional semi-Thue systems (CSTSs) are extensions of STSs, where each of their rules has the form $(\ell, r) \Leftarrow s_1 \approx t_1, \ldots, s_n \approx t_n$ for $\ell, r, s_1, t_1, \ldots, s_n, t_n \in \Sigma^*$. Here, each rule of CSTS $(\ell, r) \Leftarrow \phi$ can also be denoted as $\ell \to r \Leftarrow \phi$.
- A finitely presented monoid with decidable word problem may not admit a finite complete (unconditional) presentation, but it may admit a finite complete conditional presentation.
- For example, the monoid $M = (\Sigma; \mathcal{R})$, where $\Sigma = \{a, b\}$ and $\mathcal{R} = \{aba \to ba\}$, is finitely presented and have decidable word problem, but does not admit an equivalent monoid presentation with $(\Sigma; \mathcal{R}')$, where $\mathcal{R}'$ is a finite complete STS.
- More specifically, a completion procedure for $\mathcal{R}$ may only yield an infinite complete semi-Thue system $\{ab^n a \to b^n a \mid n \geq 1\}$. However, $\mathcal{R}$ admits an equivalent finite complete (reductive) right-join CSTS $\mathcal{R}'' = \{aba \to ba, abb \to bb \Leftarrow ab \approx b\}$. (Here, $(\Sigma; \mathcal{R}'')$ is a finite complete (reductive) conditional presentation of M.)

Types of Conditional Semi-Thue Systems

Types of CSTSs

- The types of CSTSs depend on the string rewriting relations induced by CSTSs.
- More specifically, the types of CSTSs depend on how conditions are evaluated in the conditional parts of their conditional string rewriting rules. The induced string rewriting relations from CSTSs are structured into *levels*.

Left-Right-Join CSTSs

- The string rewriting relation $\to_{\mathcal{R},lr,j}$ for a *left-right-join* CSTS $\mathcal{R}$ on Σ^* is defined as follows: $t_1 \to_{\mathcal{R},lr,j} t_2$ iff $t_1 \to_{\mathcal{R}_n} t_2$ for some $n \geq 0$. Here, the unconditional STS $\mathcal{R}_n$ are inductively defined as follows:

$$\begin{cases} \mathcal{R}_0 := \varnothing \\ \mathcal{R}_{n+1} := \{ (u\ell v, urv) \mid (\ell, r) \Leftarrow \phi \in \mathcal{R}, \text{and } usv \downarrow_{\mathcal{R}_n} utv \text{ for } \forall s \approx t \in \phi, \text{ and } u, v \in \Sigma^* \} \end{cases}$$

Right-Join CSTSs

- The string rewriting relation $\to_{\mathcal{R},r,j}$ for a *right-join* CSTS $\mathcal{R}$ on Σ^* is defined as follows: $t_1 \to_{\mathcal{R},r,j} t_2$ iff $t_1 \to_{\mathcal{R}_n} t_2$ for some $n \geq 0$. Here, the unconditional STS $\mathcal{R}_n$ are inductively defined as follows:

$$\begin{cases} \mathcal{R}_0 := \varnothing \\ \mathcal{R}_{n+1} := \{(\ell v, rv) \mid (\ell, r) \Leftarrow \phi \in \mathcal{R}, \text{ and } sv \downarrow_{\mathcal{R}_n} tv \text{ for all } s \approx t \in \phi \text{ and } v \in \Sigma^*\} \end{cases}$$

Pure-Join CSTSs

- The string rewriting relation $\to_{\mathcal{R},p,j}$ for a *pure-join* CSTS $\mathcal{R}$ on Σ^* is defined as follows: $t_1 \to_{\mathcal{R},p,j} t_2$ iff $t_1 \to_{\mathcal{R}_n} t_2$ for some $n \geq 0$. Here, the unconditional STS $\mathcal{R}_n$ are inductively defined as follows:

$$\begin{cases} \mathcal{R}_0 := \varnothing \\ \mathcal{R}_{n+1} := \{(\ell, r) \mid (\ell, r) \Leftarrow \phi \in \mathcal{R}, \text{ and } s \downarrow_{\mathcal{R}_n} t \text{ for all } s \approx t \in \phi\} \end{cases}$$

Types of Conditional Semi-Thue Systems

Example

- Consider $\mathcal{R} = \{a\ell \to b\ell, \ell am \to \ell bm, c \to d \Leftarrow a \approx b\}$, where $\Sigma = \{a, b, c, d, \ell, m\}$ with $a > b > c > d > \ell > m$. If $\mathcal{R}$ is a pure-join CSTS, then neither $c\ell \to_{\mathcal{R},p,j} d\ell$ nor $\ell cm \to_{\mathcal{R},p,j} \ell dm$ holds. If $\mathcal{R}$ is a right-join CSTS, then $c\ell \to_{\mathcal{R},r,j} d\ell$ holds, but $\ell cm \to_{\mathcal{R},r,j} \ell dm$ does not hold. If $\mathcal{R}$ is a left-right-join CSTS, then both $c\ell \to_{\mathcal{R},rl,j} d\ell$ and $\ell cm \to_{\mathcal{R},rl,j} \ell dm$ hold.

Reductive CSTSs

- We call a CSTS $\mathcal{R}$ *reductive* if for all $(\ell, r) \Leftarrow \phi \in \mathcal{R}$, $\ell >_{sl} r$, $\ell >_{sl} s_i$, and $\ell >_{sl} t_i$ for all $(s_i, t_i) \in \phi$.
- If $\mathcal{R}$ is a finite reductive right-join CSTS, then $\to_{\mathcal{R},r,j}$ is terminating and decidable.
- If $\mathcal{R}$ is a finite reductive pure-join CSTS, then $\to_{\mathcal{R},p,j}$ is terminating and decidable.

Discussion

Limitation of Left-Right-Join CSTSs

- Non-overlap may not be joinable for left-right-join CSTSs.
- Example: $\mathcal{R} = \{b \to u \Leftarrow i \approx j, c \to v \Leftarrow l \approx m, aic \to ajc, bld \to bmd\}$ over $\Sigma = \{a, b, c, d, i, j, l, m, u, v\}$. Using the shortlex order $>_{sl}$ induced by the precedence $a > b > c > d > i > j > l > m > u > v$, we see that $\mathcal{R}$ is reductive.
- Consider $\mathcal{R}$ as a left-right-join CSTS and a non-overlap $aucd\ {}_{\mathcal{R},rl,j}\!\leftarrow abcd \to_{\mathcal{R},rl,j} abvd$.
- Here, the step $abcd \to_{\mathcal{R},rl,j} aucd$ uses the rules $b \to u \Leftarrow i \approx j$ and $aic \to ajc$, and the step $abcd \to_{\mathcal{R},rl,j} abvd$ uses the rules $c \to v \Leftarrow l \approx m$ and $bld \to bmd$. However, it is not the case that $aucd \downarrow_{\mathcal{R},rl,j} abvd$.
- Therefore, joinability of critical pairs does not suffice to show local confluence for left-right-join CSTSs even if they are reductive.

Confluence of CSTSs

Critical Pairs for Right-Join and Pure-Join CSTSs

- For each pair of not necessarily distinct conditional string rewriting rules from a CSTS $\mathcal{R}$, say $(u_0, v_0) \Leftarrow \forall_{i=1}^{m} u_i \approx v_i$ and $(u_0', v_0') \Leftarrow \forall_{i=1}^{n} u_i' \approx v_i'$,
- A critical pair arises when there is an overlap, i.e., either $u_0 y = x u_0'$ and $|x| < |u_0|$ or $u_0 = x u_0' y$ for some $x, y \in \Sigma^*$. Assume that there is an overlap with such x and y:
- Now, each element in the set of critical pairs w.r.t. $\rightarrow_{\mathcal{R},r,j}$ has the following form:
  - $(v_0 y, x v_0') \Leftarrow \forall_{i=1}^{m} u_i y \approx v_i y \wedge \forall_{i=1}^{n} u_i' \approx v_i'$.
  - $(v_0, x v_0' y) \Leftarrow \forall_{i=1}^{m} u_i \approx v_i \wedge \forall_{i=1}^{n} u_i' y \approx v_i' y$
- Meanwhile, each element in the set of critical pairs w.r.t. $\rightarrow_{\mathcal{R},p,j}$ has the following form:
  - $(v_0 y, x v_0') \Leftarrow \forall_{i=1}^{m} u_i \approx v_i \wedge \forall_{i=1}^{n} u_i' \approx v_i'$.
  - $(v_0, x v_0' y) \Leftarrow \forall_{i=1}^{m} u_i \approx v_i \wedge \forall_{i=1}^{n} u_i' \approx v_i'$

Confluence of CSTSs

Joinability of Critical Pairs

- For a right-join CSTS $\mathcal{R}$, a critical pair $(s_0, t_0) \Leftarrow \forall_{i=1}^{n} s_i \approx t_i$ is *joinable* w.r.t. $\rightarrow_{\mathcal{R},r,j}$ if for any $y \in \Sigma^*$, $s_i y \downarrow_{\mathcal{R},r,j} t_i y$ for all $1 \leq i \leq n$ implies $s_0 y \downarrow_{\mathcal{R},r,j} t_0 y$.
- For a pure-join CSTS $\mathcal{R}$, a critical pair $(s_0, t_0) \Leftarrow \forall_{i=1}^{n} s_i \approx t_i$ is *joinable* w.r.t. $\rightarrow_{\mathcal{R},p,j}$ if $s_i \downarrow_{\mathcal{R},p,j} t_i$ for all $1 \leq i \leq n$ implies $s_0 \downarrow_{\mathcal{R},p,j} t_0$.

Confluence Criteria for Right-Join and Pure-Join CSTSs

- Let $\mathcal{R}$ be a finite reductive right-join CSTS. Then, $\rightarrow_{\mathcal{R},r,j}$ is confluent if and only if all critical pairs of $\mathcal{R}$ are joinable w.r.t. $\rightarrow_{\mathcal{R},r,j}$ [Deiß,1992].
- Let $\mathcal{R}$ be a finite reductive pure-join CSTS. Then, $\rightarrow_{\mathcal{R},p,j}$ is confluent if and only if all critical pairs of $\mathcal{R}$ are joinable w.r.t. $\rightarrow_{\mathcal{R},p,j}$.

Inference system for conditional equational theories

Inference rules for conditional equational theories

- Reflexivity: $\dfrac{}{s \approx s}$
- Symmetry: $\dfrac{s \approx t}{t \approx s}$
- Transitivity: $\dfrac{s \approx t \qquad t \approx u}{s \approx u}$
- Congruence: $\dfrac{s \approx t}{usv \approx utv}$
- Replacement (LR): $\dfrac{\forall (s \approx t) \in \phi : usv \approx utv}{u\ell v \approx urv}$ for all $(\ell, r) \Leftarrow \phi \in \mathcal{R}$.
- Replacement (R): $\dfrac{\forall (s \approx t) \in \phi : su \approx tu}{\ell u \approx ru}$ for all $(\ell, r) \Leftarrow \phi \in \mathcal{R}$.
- Replacement (P): $\dfrac{\forall (s \approx t) \in \phi : s \approx t}{\ell \approx r}$ for all $(\ell, r) \Leftarrow \phi \in \mathcal{R}$.

Above, $\mathcal{R}$ is a CSTS on Σ^*. The Replacement rule (LR/R/P) can be selected.

Inference system for conditional equational theories

Conditional Equational Theories

- We write $\mathcal{R} \vdash_{lr} s \approx t$ if $s \approx t$ is derivable from the inference system consisting of Reflexivity, Symmetry, Transitivity, Congruence, and the Replacement (LR) rule.
- We write $\mathcal{R} \vdash_{r} s \approx t$ if $s \approx t$ is derivable from the inference system consisting of Reflexivity, Symmetry, Transitivity, Congruence, and the Replacement (R) rule.
- We write $\mathcal{R} \vdash_{p} s \approx t$ if $s \approx t$ is derivable from the inference system consisting of Reflexivity, Symmetry, Transitivity, Congruence, and the Replacement (P) rule.

Related Results

- $t_1 \overset{*}{\leftrightarrow}_{\mathcal{R},lr,s} t_2$ iff $\mathcal{R} \vdash_{lr} t_1 \approx t_2$. Also, if $\to_{\mathcal{R},lr,j}$ is confluent, then $\to_{\mathcal{R},lr,s} = \to_{\mathcal{R},lr,j}$.
- $t_1 \overset{*}{\leftrightarrow}_{\mathcal{R},r,s} t_2$ iff $\mathcal{R} \vdash_{r} t_1 \approx t_2$. Also, if $\to_{\mathcal{R},r,j}$ is confluent, then $\to_{\mathcal{R},r,s} = \to_{\mathcal{R},r,j}$.
- $t_1 \overset{*}{\leftrightarrow}_{\mathcal{R},p,s} t_2$ iff $\mathcal{R} \vdash_{p} t_1 \approx t_2$. Also, if $\to_{\mathcal{R},p,j}$ is confluent, then $\to_{\mathcal{R},p,s} = \to_{\mathcal{R},p,j}$.

Word Problem of Monoids defined by CSTSs

Monoids Defined by Finite Reductive CSTSs

- Given a finite reductive right-join CSTS $\mathcal{R}$, $M_{\mathcal{R},r,j} := \Sigma^* / \overset{*}{\leftrightarrow}_{\mathcal{R},r,j}$ is also a monoid.
  - Elements: Congruence classes $[w]_{\mathcal{R},r,j}$ in $\{[w]_{\mathcal{R},r,j} \mid w \in \Sigma^*\}$.
  - Operation: A binary operation $\cdot$ such that $[u]_{\mathcal{R},r,j} \cdot [v]_{\mathcal{R},r,j} = [uv]_{\mathcal{R},r,j}$, where $u, v \in \Sigma^*$ with the identity element $[\varepsilon]_{\mathcal{R},r,j}$.
- Given a finite reductive pure-join CSTS $\mathcal{R}$, the monoid $M_{\mathcal{R},p,j} := \Sigma^* / \overset{*}{\leftrightarrow}_{\mathcal{R},p,j}$ can be defined similarly.

Decision Procedure for the Word Problem of Monoids $M_{\mathcal{R},r,j}$ and $M_{\mathcal{R},p,j}$

- Let $\mathcal{R}$ be a finite reductive right-join (resp. pure-join) CSTS on Σ^*. If $\rightarrow_{\mathcal{R},r,j}$ (resp. $\rightarrow_{\mathcal{R},p,j}$) is confluent, then we can decide whether s and t on Σ^* are the same element in the monoid $M_{\mathcal{R},r,j} := \Sigma^* / \overset{*}{\leftrightarrow}_{\mathcal{R},r,j}$ (resp. $M_{\mathcal{R},p,j} := \Sigma^* / \overset{*}{\leftrightarrow}_{\mathcal{R},p,j}$).

Formalization of the Proposed Completion Procedure of STSs

Using Inductively Defined Predicates in Isabelle/HOL

inductive *sts_compl_step::* "$sts \times sts \Rightarrow sts \times sts \Rightarrow bool$" (**infix** "$\vdash_{SR}$" 55) **where**

deduce: "$(E, R) \vdash_{SR} (E \cup \{(s @ u3, u1 @ t)\}, R)$"

if "$(u1 @ u2, s) \in R$" **and** "$(u2 @ u3, t) \in R$" **and** "$u2 \neq []$"

| *simplifyl:* "$(E \cup \{(u1 @ u2 @ u3, s)\}, R) \vdash_{SR} (E \cup \{(u1 @ t @ u3, s)\}, R)$"

if "$(u2, t) \in R$"

| *simplifyr:* "$(E \cup \{(s, u1 @ u2 @ u3)\}, R) \vdash_{SR} (E \cup \{(s, u1 @ t @ u3)\}, R)$"

if "$(u2, t) \in R$"

| *orientl:* "$(E \cup \{(s, t)\}, R) \vdash_{SR} (E, \{R \cup \{(s, t)\})$" **if** "$s >_{sl} t$"

| *orientr:* "$(E \cup \{(s, t)\}, R) \vdash_{SR} (E, \{R \cup \{(t, s)\})$" **if** "$t >_{sl} s$"

| *collapse:* "$(E, R \cup \{(u1 @ u2 @ u3, s)\}) \vdash_{SR} (E \cup \{(u1 @ t @ u3, s)\}, R)$" **if** "$(u2, t) \in R$"

| *compose:* "$(E, R \cup \{(s, u1 @ u2 @ u3)\}) \vdash_{SR} (E, R \cup \{(s, u1 @ t @ u3)\})$" **if** "$(u2, t) \in R$"

| *delete:* "$(E \cup \{(s, s)\}, R) \vdash_{SR} (E, R)$"

Above, E consists of ordered pairs (instead of unordered pairs) for technical convenience.

Formalization of CSTSs

Formalization of $\rightarrow_{\mathcal{R},r,j}$

definition "$\mathit{csr\_r\_join\_step}\ \mathcal{R} = (\bigcup n.\ \mathit{csr\_r\_join\_step\_n}\ \mathcal{R}\ n)$"

fun $\mathtt{csr\_r\_join\_step\_n}$:: "$\mathit{csts} \Rightarrow \mathit{nat} \Rightarrow \mathit{string\ rel}$" **where**
"$\mathit{csr\_r\_join\_step\_n}\ \mathcal{R}\ 0 = \{\}$" |
"$\mathit{csr\_r\_join\_step\_n}\ \mathcal{R}\ (\mathit{Suc}\ n) =$
$\{(C\langle\!\langle \ell\ @\ w \rangle\!\rangle, C\langle\!\langle r\ @\ w \rangle\!\rangle) \mid C\ \ell\ r\ cs\ w.\ ((\ell, r),\ cs) \in \mathcal{R}\ \wedge$
$(\forall (s_i,\ t_i) \in \mathit{set}\ cs.\ (s_i\ @\ w,\ t_i\ @\ w) \in (\mathit{csr\_r\_join\_step\_n}\ \mathcal{R}\ n)^{\downarrow})\}$”

- In the csr_r_join_step_n function, C is a (string) context and $C\langle\!\langle \ell @ w \rangle\!\rangle$ (resp. $C\langle\!\langle r @ w \rangle\!\rangle$) denotes the application of the context C to the string ℓw (resp. rw). A context for strings is formalized based on the existing formalization of contexts for terms in IsaFoR.

Locales for CSTSs

Locale for Right-Join CSTSs

locale *conditional_r_join_semi_Thue* = $reductive\_r\_join$ + $conditional\_semi\_Thue\ R\ S$
  for $R :: csts$ **and** $S ::$ "$char\ set$" +
  fixes $Thue\_R\_Congruence ::\ sts$
  assumes "$Thue\_R\_Congruence = (csr\_r\_join\_step\ R)^{\leftrightarrow^*}$"
    and "$Thue\_R\_Congruence \subseteq S^* \times S^*$"
begin
$\ldots$

The assumptions in the locale *conditional_r_join_semi_Thue* also represent the assumption $\overset{*}{\leftrightarrow}_{\mathcal{R},r,j} \subseteq \Sigma^* \times \Sigma^*$, which also implies the assumption $\rightarrow_{\mathcal{R},r,j} \subseteq \Sigma^* \times \Sigma^*$.

- Here, S denotes an alphabet Σ.
- The assumption that S is finite and nonempty is declared in the locale *conditional_semi_Thue*.

Locales for CSTSs

Using the Existing Locale *Monoid*

The (existing) locale *monoid* is instantiated using the different parameters for the monoids $M_{\mathcal{R},r,j}$ and $M_{\mathcal{R},p,j}$, respectively:

$monoid\ "S^\star/Thue\_R\_Congruence"\ "([\cdot]_r)"\ "equiv\_r.Class\ \varepsilon"$

$monoid\ "S^\star/Thue\_P\_Congruence"\ "([\cdot]_p)"\ "equiv\_p.Class\ \varepsilon"$

- Above, $[\cdot]_r$ (resp. $[\cdot]_p$) represents an associative binary operator for the monoid $M_{\mathcal{R},r,j}$ (resp. $M_{\mathcal{R},p,j}$).
- Finally, $equiv\_r.Class\ \varepsilon$ (resp. $equiv\_p.Class\ \varepsilon$) represents the congruence class $[\varepsilon]_{\mathcal{R},r,j}$ (resp. $[\varepsilon]_{\mathcal{R},p,j}$) corresponding to the identity element in $M_{\mathcal{R},r,j}$ (resp. $M_{\mathcal{R},p,j}$).

Conclusion

Summary

- Presented and formalized an inference system for a completion procedure of semi-Thue systems, which is adapted from the existing Knuth-Bendix completion procedure of semi-Thue systems and an abstract completion procedure of rewriting systems.
  - Used the simple and efficient (linear-time) string-matching algorithms and ordering (length-lexicographic ordering) for inference rules instead of using more complex unification/matching for terms and their ordering.
- Presented a formalization of conditional semi-Thue systems and provided a new formalized proof of the confluence criterion for right-join and pure-join CSTSs.
- Provided a new formalized decision procedure for the word problem of monoids presented by finite complete reductive right-join and pure-join CSTSs.

Thank you! Questions?

Dohan Kim
University of Innsbruck, Austria

16th International Conference on Interactive Theorem Proving (ITP 2025), Reykjavik, Iceland, Sept. 29, 2025