Relational Colour Refinement for Non-Relational Signatures

Bachelor's Thesis

Theodor Teslia

August 28, 2025

RWTH Aachen University

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

1 Introduction

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

2 Relational Colour Refinement

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

3 Relational Colour Refinement for structures with functions

3.1 Naive Encoding of functions

A simple way to apply relational colour refinement to non-relational structures is, to encode the functions in the signature as a relation. Formally we transform a signature σ that includes function symbols to a new signature σ': For every relation symbol $R \in \sigma$, we introduce a relation symbol $R \in \sigma'$ with the same arity and for every function symbol $f \in \sigma$ with arity k, we introduce a relational symbol $R_f \in \sigma'$ of arity $k+1$.

Semantically, a structure $\mathfrak{A}$ of signature σ can then be encoded as a structure $\mathfrak{A}'$ of signature σ' and with the same universe as $\mathfrak{A}$. For every relational symbol $R \in \sigma$ we set $R^{\mathfrak{A}'} := R^{\mathfrak{A}}$ and for every function symbol $f \in \sigma$ of arity k there exists a relation symbol $R_f \in \sigma'$ and we set $R_f^{\mathfrak{A}} := \{(\mathbf{x}, y) : f^{\mathfrak{A}}(\mathbf{x}) = y\}$ where $\mathbf{x}$ is a tuple of arity k.

This procedure encodes a non-relational structure as a relational one, on which Relational Colour Refinement can now be performed. As such we say, that the Naive Relational Colour Refinement (nRCR) distinguishes two structures $\mathfrak{A}$ and $\mathfrak{B}$ if, and only if, RCR distinguishes their naive encodings $\mathfrak{A}'$ and $\mathfrak{B}'$. However, this results in a very weak logical characterisation, that does not allow nesting of terms, namely the nesting-free-fragment of $\mathsf{GF}(\mathsf{C})$.

Definition 1 ($\mathsf{nfGF}(\mathsf{C})$)**.** Consider the definition of $\mathsf{GF}(\mathsf{C})$ given in **??**. We obtain the nesting-free fragment, by allowing $f(\mathbf{x}) = y$ as a further atomic formula. Concretely, the only allowed atomic formulae are of the form $R(x_1, \ldots, x_\ell)$, $x = y$ and $f(x_1, \ldots, x_\ell) = y$, where f has arity ℓ, $\mathrm{free}(f(x_1, \ldots, x_\ell) = y) = \{x_1, \ldots, x_\ell\}$ and $\mathrm{gd}(f(\mathbf{x}) = y) = 0$.

The remaining definitions stay the same.

Theorem 2. *The two following statements are equivalent:*

1. *nRCR distinguishes $\mathfrak{A}$ and $\mathfrak{B}$.*
2. *There exists a sentence $\varphi \in \mathsf{nfGF}(\mathsf{C})$ such that $\mathfrak{A} \models \varphi$ and $\mathfrak{B} \not\models \varphi$.*

Proof. 1. $\Rightarrow$ 2.: By definition, $\mathfrak{A}$ and $\mathfrak{B}$ are distinguished by nRCR if, and only if, $\mathfrak{A}'$ and $\mathfrak{B}'$ are distinguished by RCR. Using the result of [1], we obtain a sentence $\varphi' \in \mathsf{GF}(\mathsf{C})$ that distinguishes the encoded structures. Via a structural induction on the formula, we can now translate φ' into a formula $\varphi \in \mathsf{nfGF}(\mathsf{C})$ This can be achieved by expanding formulae $R_f(x_1, \ldots, x_\ell, y)$ to $f(x_1, \ldots, x_\ell) = y$ for function symbols $f \in \sigma$ and letting everything else stay the same.

2. $\Rightarrow$ 1.: When considering $\mathsf{nfGF}(\mathsf{C})$, one can find that the transformation done at the end of the first direction can be applied in reverse. This then leads to a distinguishing sentence in $\mathsf{GF}(\mathsf{C})$ and with [1] to a distinguishing colouring of the encoded structures, which by definition is a distinguishing colouring for the structures themselves. $\square$

While the above theorem results in a nice characterisation of the naive encoding, the nesting of terms is often very desired when using functions. However, it can be shown that nesting is too powerful for such a naive encoding.

Consider the two structures $\mathfrak{A}$ and $\mathfrak{B}$ of signature $\sigma = \{f/1\}$ which can be seen in Figure 1. Formally they are defined as

$$\begin{aligned}
\mathfrak{A} = (A = \{&a_1, a_2, a_3, a_4, a_5, a_6\}, \\
f^{\mathfrak{A}} = \{& \\
& a_1 \mapsto a_3,\ a_3 \mapsto a_2,\ a_2 \mapsto a_1, \\
& a_4 \mapsto a_5,\ a_5 \mapsto a_6,\ a_6 \mapsto a_4 \\
\})&
\end{aligned}
\quad \text{and} \quad
\begin{aligned}
\mathfrak{B} = (B = \{&b_1, b_2, b_3, b_4, b_5, b_6\}, \\
f^{\mathfrak{B}} = \{& \\
& b_1 \mapsto b_3,\ b_3 \mapsto b_5,\ b_5 \mapsto b_6, \\
& b_6 \mapsto b_4,\ b_4 \mapsto b_2,\ b_2 \mapsto b_1 \\
\})&
\end{aligned}$$

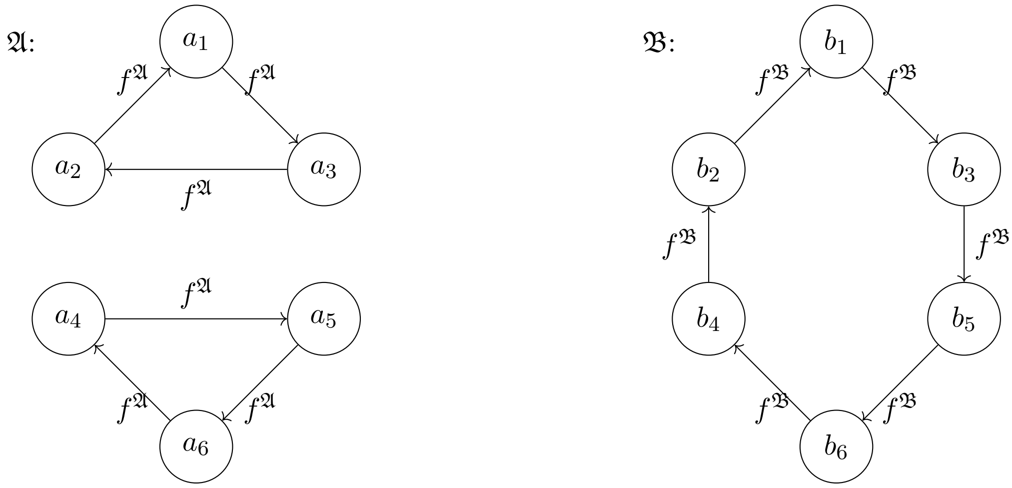

Figure 1: Two σ-structures $\mathfrak{A}$ and $\mathfrak{B}$

Consider the formula $\varphi = \exists x.(f(f(f(x))) = x)$ which utilizes term nesting to find a cycle with length three. It is obvious that $\mathfrak{A} \models \varphi$ and $\mathfrak{B} \not\models \varphi$. However, when encoding the two structures with the naive method described above, one finds that nRCR cannot distinguish them. Therefore, term nesting is too powerful for the naive encoding.

A method that allows for the nesting of terms will be described in the following section.

3.2 Using the transitive expansion

Let

$$\begin{aligned}\mathcal{I}(n,m) = \{(k,l,p) \in [n]^3 \quad : \quad & k+p < k+l \leq n \wedge \\ & k + r \cdot l + p = m \text{ for some } r \in \mathbb{N}\}.\end{aligned}$$

The set will represents the possible ways, to decompose a path into a cycle and the path to and from it. This means, that the triple (k, ℓ, p) will represent a path, that has a beginning part of length k, then a cycle of length ℓ and a last part that consists of the first p elements of the cycle. One can see that in a structure $\mathfrak{A}$ with a unary function f and n elements, any path along of f with length $m > n$ can be decomposed into a triple in the set $\mathcal{I}(n, m)$.

Lemma 3. *Let $\psi(x_1, x_2) \coloneqq f^m(x_1) = x_2$. Then there exists a formula $\vartheta(x_1, x_2) \in \mathsf{GF(C)}$ such that for any $\mathfrak{A}$ with $\|\mathfrak{A}\| = n$ it holds*

$$\mathfrak{A}, a_1, a_2 \models \psi(x_1, x_2) \textit{ if, and only if, } \mathfrak{A}, a_1, a_2 \models \vartheta(x_1, x_2)$$

and for any $f^{m'}(x)$ that appears in ϑ, $m' \leq n$.

Proof. If $m \leq n$, we let $\vartheta \coloneqq \psi$ and the claim follows.

Otherwise, we define

$$\vartheta(x_1, x_2) \coloneqq \bigvee_{(k,\ell,p)\in\mathcal{I}(n,m)} \zeta_{(k,\ell,p)}(x_1, x_2)$$

where

$$\begin{aligned}\zeta_{(k,\ell,p)}(x_1, x_2) \coloneqq & f^{k+p}(x_1) = x_2 \wedge f^k(x_1) = f^{k+\ell}(x_1) \\ & \wedge \mathrm{E}_f^{k,\ell}(x_1) \\ & \wedge \bigwedge_{\ell' < \ell} f^k(x_1) \neq f^{k+\ell'}(x_1)\end{aligned}$$

and

$$E_f^{k,\ell}(t(x_1)) = \begin{cases} \top & \text{if } k = 0 \\ f^{k-1}(t(x_1)) \neq f^{k-1+\ell}(t(x_1)) & \text{otherwise.} \end{cases}$$

Due to the definition of $\mathcal{I}(n,m)$ it is obvious that only $f^{m'}$ with $m' \leq n$ appears.

We now proceed to the proof of the equivalence. For the purpose of readability, we will use $f_{\mathfrak{A}}$ instead of $f^{\mathfrak{A}}$.

We will show that if $\mathfrak{A}, a_1, a_2 \models \vartheta(x_1, x_2)$, then $\mathfrak{A}, a_1, a_2 \models \psi(x_1, x_2)$. Let $\mathfrak{A}, a_1, a_2 \models \vartheta(x_1, x_2)$. By definition of ϑ, there are $(k, \ell, p) \in \mathcal{I}(n,m)$ with $\mathfrak{A}, a_1, a_2 \models \zeta_{(k,\ell,p)}(x_1, x_2)$. In particular $f_{\mathfrak{A}}^k(a_1) = f_{\mathfrak{A}}^{k+\ell}(a_1)$. It follows that

$$f_{\mathfrak{A}}^k(a_1) = f_{\mathfrak{A}}^{k+\ell}(a_1) = f_{\mathfrak{A}}^{k+2\ell}(a_1) = f_{\mathfrak{A}}^{k+3\ell}(a_1) = \cdots = f_{\mathfrak{A}}^{k+r\cdot\ell}(a_1)$$

for all $r \in \mathbb{N}$. By using the definition of $\mathcal{I}(n,m)$, we get

$$a_2 = f_{\mathfrak{A}}^{k+p}(a_1) = f_{\mathfrak{A}}^{k+r\cdot\ell+p}(a_1) = f_{\mathfrak{A}}^m(a_1).$$

From this we can deduce $\mathfrak{A}, a_1, a_2 \models \psi(x_1, x_2)$, where $\psi(x_1, x_2)$ has the form $f^m(x_1) = x_2$.

Now we prove that if $\mathfrak{A}, a_1, a_2 \models \psi(x_1, x_2)$, then $\mathfrak{A}, a_1, a_2 \models \vartheta(x_1, x_2)$. Let $\mathfrak{A}, a_1, a_2 \models \psi(x_1, x_2)$. By assumption $m > n$ and by the pigeonhole principle there have to be distinct i, j such that $f_{\mathfrak{A}}^i(a_1) = f_{\mathfrak{A}}^j(a_1)$. Choose such i, j such that they are lexicographically minimal.

Now choose $k \coloneqq i$, $\ell \coloneqq j - i$ and $p \coloneqq (m-i) \mod (j-i) = (m-i) \mod \ell$. Obviously $(k, \ell, p) \in \mathcal{I}(n,m)$ and what remains to be shown is that $\mathfrak{A}, a_1, a_2 \models \zeta_{(k,\ell,p)}(x_1, x_2)$. For that, we consider the parts of the conjunction and show for each one that it is satisfied.

$f^{k+p}(x_1) = x_2$: We use the fact that $a = b \mod c \Leftrightarrow b = r \cdot c + a$ for some $r \in \mathbb{N}$. Then

$$f_{\mathfrak{A}}^{k+p}(a_1) = f_{\mathfrak{A}}^{i+(m-i)-r\cdot\ell}(a_1) = f_{\mathfrak{A}}^{i+r\cdot\ell+m-i-r\cdot\ell}(a_1) = f_{\mathfrak{A}}^m(a_1) = a_2.$$

Therefore $\mathfrak{A}, a_1, a_2 \models f^{k+p}(x_1) = x_2$.

$f^k(x_1) = f^{k+\ell}(x_1)$: Consider that

$$f_{\mathfrak{A}}^k(a_1) = f_{\mathfrak{A}}^i(a_1) = f_{\mathfrak{A}}^j(a_1) = f_{\mathfrak{A}}^{j+i-i}(a_1) = f_{\mathfrak{A}}^{i+j-i}(a_1) = f_{\mathfrak{A}}^{k+\ell}(a_1).$$

This leads to $\mathfrak{A}, a_1, a_2 \models f^k(x_1) = f^{k+\ell}(x_1)$.

$\mathrm{E}_f^{k,\ell}(x_1)$: This has to be satisfied, otherwise $f_{\mathfrak{A}}^{k-1}(a_1) = f_{\mathfrak{A}}^{k-1+\ell}(a)$, but then $(k-1, \ell)$ would be lexicographically smaller than (i, j).

The same reasoning applies to $\bigwedge_{\ell' < \ell} f^k(x_1) \neq f^{k+\ell'}(x_1)$. If it weren't satisfied, there would be a (i, j') with $j' < j$ and $f_{\mathfrak{A}}^i(a_1) = f_{\mathfrak{A}}^{i+j'}(a_1)$ which would be lexicographically smaller than (i, j).

Thus we have shown that every subformula of the conjunction and therefore the formula is being fulfilled. □

The above proof allows for the translation of formulae $f^m(x) = y$ to a formula $\vartheta(x, y)$ that is equivalent for structures with n elements. A natural extension would be, to allow alternation of functions, for example formulae like $g^m(f^{m'}(x)) = y$. This is also possible and will be proved in the following proof.

Lemma 4. *Let $\psi(x_1, x_2) \coloneqq t(x_1) = x_2$ be an atomic formula. Then there exists a formula $\vartheta_t(x_1, x_2) \in$ $\mathsf{GF}(\mathsf{C})$, such that for any structure (of a fitting signature) $\mathfrak{A}$ with $\|\mathfrak{A}\| = n$ it holds*

$$\mathfrak{A}, a_1, a_2 \models \psi(x_1, x_2) \textit{ if, and only if, } \mathfrak{A}, a_1, a_2 \models \vartheta_t(x_1, x_2).$$

Furthermore, $\vartheta_t(x_1, x_2)$ is of the form $\vartheta_t(x_1, x_2) = \bigvee \Phi(x_1, x_2)$ where all $\varphi(x_1, x_2) \in \Phi(x_1, x_2)$ are of the form

$$t'(x_1) = x_2 \wedge \bigwedge \Psi(x_1)$$

for some term $t'(x_1)$, and for every $f^m(s(x))$ that appears in ϑ_t, $m \leq n$.

Proof. We prove this via an induction on the term $t(x_1)$.

Base case: If $t(x_1) \coloneqq f^m(x_1)$ for a unary function symbol f and $m \in \mathbb{N}$, we use the formula constructed in the proof of Theorem 3. It can easily be verified that it is in the correct form.

Inductive step: Assume that $t(x_1) \coloneqq g^m(s(x_1))$ for a unary function symbol g, $m \in \mathbb{N}$ and term s. By induction hypothesis, we have a formula $\vartheta_s(y_1, y_2) = \bigvee \Phi_s(y_1, y_2)$ in the above defined form with $\mathfrak{A}, a_1, a_2 \models s(y_1) = y_2$ if, and only if, $\mathfrak{A}, a_1, a_2 \models \vartheta_s(y_1, y_2)$.

If $m \leq n$, we set

$$\vartheta_t(x_1, x_2) = \bigvee \Phi'(x_1, x_2),$$

where $\Phi'(x_1, x_2) \coloneqq \{g^m(t'(y_1/x_1)) = x_2 \wedge \bigwedge \Psi(y_1/x_1) : t'(y_1) = y_2 \wedge \bigwedge \Psi(y_1) \in \Phi_s(y_1, y_2)\}$.

Now assume $m > n$.

Then we set

$$\vartheta_t(x_1, x_2) = \bigvee_{(k,\ell,p) \in \mathcal{I}(n,m)} \bigvee \Phi'_{(k,\ell,p)}(x_1, x_2),$$

where

$$\begin{aligned}\Phi'_{(k,\ell,p)} \coloneqq \{g^{k+p}(t'(y_1/x_1)) = x_2 \wedge g^k(t'(y_1/x_1)) = g^{k+l}(t'(y_1/x_1))& \\ \wedge \mathrm{E}_g^{k,l}(t'(y_1/x_1)) \wedge \bigwedge_{\ell' < \ell} g^k(t'(y_1/x_1)) \neq g^{k+\ell'}(t'(y_1/x_1))& \\ \wedge \Psi(y_1/x_1) : t'(y_1) = y_2 \wedge \bigwedge \Psi(y_1) \in \Phi_s(y_1, y_2)\}&\end{aligned}$$

By using the above definitions, we get $\mathfrak{A}, a_1, a_2 \models s(y_1) = y_2$ if, and only if, $\mathfrak{A}, a_1, a_2 \models \varphi_s(y_1, y_2)$ for some $\varphi_s \in \Phi_s$ where $\varphi_s(y_1, y_2)$ is of the form $t'(y_1) = y_2 \wedge \bigwedge \Psi(y_1)$. Therefore

$$\mathfrak{A}, a_1, a_2 \models s(y_1) = y_2 \text{ if, and only if, } \mathfrak{A}, a_1, a_2 \models t'(y_1) = y_2 \wedge \bigwedge \Psi(y_1). \tag{1}$$

We now proof that

$$\mathfrak{A}, a_1, a_2 \models t(x_1) = x_2 \text{ if, and only if, } \mathfrak{A}, a_1, a_2 \models \vartheta_t(x_1, x_2).$$

Assume $m \leq n$. Let $\mathfrak{A}, a_1, a_2 \models \vartheta_t$. Then there is some $\varphi(x_1, x_2) \coloneqq g^m(t'(y_1/x_1)) = x_2 \wedge \bigwedge \Psi(y_1/x_1)$ such that $\mathfrak{A}, a_1, a_2 \models \varphi(x_1, x_2)$. We then get

$$\begin{aligned} & \mathfrak{A}, a_1, a_2 \models g^m(t'(y_1/x_1)) = x_2 \wedge \bigwedge \Psi(y_1/x_1) \\ \Leftrightarrow\ & \mathfrak{A}, a_1, a_2, a_3 \models g^m(x_3) = x_2 \wedge \bigwedge \Psi(y_1/x_1) \wedge t'(y_1/x_1) = x_3 \text{ for all } a_3 \in A \\ \overset{\text{Equation (1)}}{\Leftrightarrow}\ & \mathfrak{A}, a_1, a_2, a_3 \models g^m(x_3) = x_2 \wedge s(x_1) = x_3 \text{ for all } a_3 \in A \\ \Leftrightarrow\ & \mathfrak{A}, a_1, a_2 \models g^m(s(x_1)) = x_2. \end{aligned}$$

Now let $m > n$. Then there is a

$$\begin{aligned}\varphi(x_1, x_2) \coloneqq g^{k+p}(t'(y_1/x_1)) = x_2 \wedge g^k(t'(y_1/x_1)) = g^{k+l}(t'(y_1/x_1))& \\ \wedge \mathrm{E}_g^{k,l}(t'(y_1/x_1)) \wedge \bigwedge_{\ell' < \ell} g^k(t'(y_1/x_1)) \neq g^{k+\ell'}(t'(y_1/x_1))& \\ \wedge \Psi(y_1/x_1)&\end{aligned}$$

for some $(k,\ell,p) \in \mathcal{I}(n,m)$ with $\mathfrak{A}, a_1, a_2 \models \varphi(x_1,x_2)$. And now

$$\begin{aligned}
&\mathfrak{A}, a_1, a_2 \models \psi(x_1,x_2)\\
\Leftrightarrow\ & A, a_1, a_2, a_3 \models g^{k+p}(x_3) = x_2 \wedge g^k(x_3)) = g^{k+l}(x_3)\\
&\qquad \wedge \mathrm{E}_g^{k,l}(x_3) \wedge \bigwedge_{\ell'<\ell} g^k(x_3) \neq g^{k+\ell'}(x_3)\\
&\qquad \wedge \Psi(y_1/x_1) \wedge t'(y_1/x_1) = x_3 \text{ for all } a_3 \in A\\
\overset{\text{Theorem 3}}{\Leftrightarrow}\ & \mathfrak{A}, a_1, a_2, a_3 \models g^m(x_3) = x_2 \wedge t'(y_1/x_1) = x_3 \wedge \Psi(y_1/x_1) \text{ for all } a_3 \in A\\
\overset{\text{Equation (1)}}{\Leftrightarrow}\ & \mathfrak{A}, a_1, a_2, a_3 \models g^m(x_3) = x_2 \wedge s(x_1) = x_3 \text{ for any } a_3 \in A\\
\Leftrightarrow\ & \mathfrak{A}, a_1, a_2 \models g^m(s(x_1)) = x_2.
\end{aligned}$$

The other direction follows in both cases, as only equivalent steps have been used and it is obvious that the disjunction of a set is being fulfilled, if a formula of the set is satisfied.

Therefore we have finished the proof. □

To obtain our characterising result for structures with (unary) functions, we have to define how the functions should be encoded.

Definition 5 (Transitive Expansion)**.** Let $\sigma := \sigma_{\mathrm{Rel}} \dot\cup \sigma_{\mathrm{Func}}$ be a signature with relation symbols σ_{Rel} and unary function symbols σ_{Func} and let $\mathfrak{A}$ be a structure of signature σ with $\|\mathfrak{A}\| = n$. For readability, we define the family of sets $\mathrm{Alters}_n^0(\sigma) := \emptyset$ and

$$\mathrm{Alters}_n^k(\sigma) := \mathrm{Alters}_n^{k-1}(\sigma) \cup \{f_1^{m_1} f_2^{m_2} \dots f_k^{m_k} : f_1 f_2 \dots f_k \in (\sigma_{\mathrm{Func}})^k, 0 \leq m_i \leq n \text{ for } 1 \leq i \leq k\}$$

For an arbitrary $k \in \mathbb{N}$, we define the transitive expansion with alternation depth k as a structure $\widetilde{\mathfrak{A}}$ of signature $\widetilde{\sigma}$, where

$$\widetilde{\sigma} := \sigma_{\mathrm{Rel}} \dot\cup \{F_\alpha : \alpha \in \mathrm{Alters}_n^k(\sigma)\}$$

and the F_α are binary relations. Semantically, we have

$$F_\alpha^{\widetilde{\mathfrak{A}}} := \{(a,b) : \alpha^{\mathfrak{A}}(a) = b\}.$$

We now can define the algorithm for relational colour refinement for (unary) functions.

Definition 6 (RCR for structures with unary functions)**.** Let σ be a signature with relation and unary function symbols and let $\mathfrak{A}$ and $\mathfrak{B}$ be structures of signature σ.

We say that $\mathfrak{A}$ and $\mathfrak{B}$ are being distinguished by RCR with alternation depth k (RCR_k), if $\|\mathfrak{A}\| \neq \|\mathfrak{B}\|$ or the transitive expansions with alternation depth k, $\widetilde{\mathfrak{A}}$ and $\widetilde{\mathfrak{B}}$, are being distinguished by RCR.

To show that this definition may be sensible, we want to execute RCR_1 on the structures $\mathfrak{A}$ and $\mathfrak{B}$ from Figure 1. First we compute $\widetilde{\sigma}$ as $\{F_{f^i} : 0 \leq i \leq 6\}$ and by performing the translation we obtain:

$$\begin{aligned}
\widetilde{\mathfrak{A}} = (A, F_{f^0}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_1),(a_2,a_2),(a_3,a_3),(a_4,a_4),(a_5,a_5),(a_6,a_6)\},\\
F_{f^1}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_3),(a_2,a_1),(a_3,a_2),(a_4,a_5),(a_5,a_6),(a_6,a_4)\},\\
F_{f^2}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_2),(a_2,a_3),(a_3,a_1),(a_4,a_6),(a_5,a_4),(a_6,a_5)\},\\
F_{f^3}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_1),(a_2,a_2),(a_3,a_3),(a_4,a_4),(a_5,a_5),(a_6,a_6)\},\\
F_{f^4}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_3),(a_2,a_1),(a_3,a_2),(a_4,a_5),(a_5,a_6),(a_6,a_4)\},\\
F_{f^5}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_2),(a_2,a_3),(a_3,a_1),(a_4,a_6),(a_5,a_4),(a_6,a_5)\},\\
F_{f^6}^{\widetilde{\mathfrak{A}}} &= \{(a_1,a_1),(a_2,a_2),(a_3,a_3),(a_4,a_4),(a_5,a_5),(a_6,a_6)\})
\end{aligned}$$

and

$$
\begin{aligned}
\widetilde{\mathfrak{B}} = (B, F_{f^0}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_1),(a_2,a_2),(a_3,a_3),(a_4,a_4),(a_5,a_5),(a_6,a_6)\},\\
F_{f^1}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_3),(a_2,a_1),(a_3,a_5),(a_4,a_2),(a_5,a_6),(a_6,a_4)\},\\
F_{f^2}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_5),(a_2,a_3),(a_3,a_6),(a_4,a_1),(a_5,a_4),(a_6,a_2)\},\\
F_{f^3}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_6),(a_2,a_5),(a_3,a_4),(a_4,a_3),(a_5,a_2),(a_6,a_1)\},\\
F_{f^4}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_4),(a_2,a_6),(a_3,a_2),(a_4,a_5),(a_5,a_1),(a_6,a_3)\},\\
F_{f^5}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_2),(a_2,a_4),(a_3,a_1),(a_4,a_6),(a_5,a_3),(a_6,a_5)\},\\
F_{f^6}^{\widetilde{\mathfrak{B}}} &= \{(a_1,a_1),(a_2,a_2),(a_3,a_3),(a_4,a_4),(a_5,a_5),(a_6,a_6)\})
\end{aligned}
$$

By using [1], we know that RCR distinguishes $\widetilde{\mathfrak{A}}$ and $\widetilde{\mathfrak{B}}$ if, and only if, there is a formula $\widetilde{\varphi} \in \mathsf{GF}(\mathsf{C})$ of signature $\widetilde{\sigma}$ that distinguishes them. Notice that $F_{f^0}^{\widetilde{\mathfrak{A}}} = F_{f^3}^{\widetilde{\mathfrak{A}}} = F_{f^6}^{\widetilde{\mathfrak{A}}}$, $F_{f^1}^{\widetilde{\mathfrak{A}}} = F_{f^4}^{\widetilde{\mathfrak{A}}}$ and $F_{f^2}^{\widetilde{\mathfrak{A}}} = F_{f^5}^{\widetilde{\mathfrak{A}}}$, while only $F_{f^0}^{\widetilde{\mathfrak{B}}} = F_{f^6}^{\widetilde{\mathfrak{B}}}$. Therefore the sentence

$$\exists^{\geq 6}(x,y).\, \Big(F_{f^1}(x,y) \wedge F_{f^4}(x,y)\Big) \in \mathsf{GF}(\mathsf{C})$$

is satisfied by $\widetilde{\mathfrak{A}}$, but not $\widetilde{\mathfrak{B}}$.

We see, that this procedure distinguishes structures, that were not distinguished by nRCR. To formalise this, we want to characterise this algorithm logically, as well as combinatorially.

3.2.1 Logical characterisation of RCR_k

Definition 7 (Alternation bounded $\mathsf{GF}(\mathsf{C})$)**.** The fragment of $\mathsf{GF}(\mathsf{C})$ with an alternation bound of k ($\mathsf{GF}(\mathsf{C})_k$) is $\mathsf{GF}(\mathsf{C})$ with the constraint that for all formulae $\varphi \in \mathsf{GF}(\mathsf{C})_k$ of signature σ and every term t that appears in φ, there is an $n \in \mathbb{N}$ and an $\alpha \in \mathrm{Alters}_n^k(\sigma)$ such that $\alpha = t$. Atomic formulae are defined as usual, that is, the formulae $R(t_1(x_1), t_2(x_2), \ldots, t_n(x_n))$ and $t_1(x_1) = t_2(x_2)$ for terms $t_1, t_2, \ldots, t_n$ and variables $x_1, x_2, \ldots, x_n$ are atomic formulae.

Theorem 8. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two structures of the same signature σ with relation and unary function symbols and let $k \in \mathbb{N}$ The two following statements are equivalent:*

1. RCR_k *distinguishes $\mathfrak{A}$ and $\mathfrak{B}$.*
2. *There exists a sentence $\varphi \in \mathsf{GF}(\mathsf{C})_k$ such that $\mathfrak{A} \models \varphi$ and $\mathfrak{B} \not\models \varphi$.*

Proof. We prove that *1.* implies *2.*. Let $\mathfrak{A}$ and $\mathfrak{B}$ be distinguished by RCR_k. If they are of different sizes, assume without loss of generality that

$$\|\mathfrak{A}\| = n > n' = \|\mathfrak{B}\|.$$

Then define $\varphi := \exists^{\geq n} x. \top \in \mathsf{GF}(\mathsf{C})_k$, which obviously distinguishes the structures.

Now assume $\|\mathfrak{A}\| = \|\mathfrak{B}\| = n$. By definition, RCR distinguishes $\widetilde{\mathfrak{A}}$ and $\widetilde{\mathfrak{B}}$. When using the proof from [1], we obtain a formula $\widetilde{\varphi} \in \mathsf{GF}(\mathsf{C})$ of signature $\widetilde{\sigma}$ that distinguishes the expansions. This formula $\widetilde{\varphi}$ can then be translated to a formula $\varphi \in \mathsf{GF}(\mathsf{C})_k$ of signature σ.

For every atomic subformula $R(\mathbf{x})$, where $R \in \sigma_{\mathrm{Rel}}$, let the formula stay the same. For every atomic subformla $F_\alpha(x,y)$, where $\alpha \in \mathrm{Alters}_n^k(\sigma)$, replace it by the formula $\alpha(x) = y$. Obviously, if a structure's expansion satisfied $\widetilde{\varphi}$, it also satisfies φ and vice versa.

Therefore, we get a formula $\varphi \in \mathsf{GF}(\mathsf{C})_k$ that distinguishes $\mathfrak{A}$ and $\mathfrak{B}$.
Now we prove that *2.* implies *1.*. Let $\varphi \in \mathsf{GF}(\mathsf{C})_k$ such that $\mathfrak{A} \models \varphi$ and $\mathfrak{B} \not\models \varphi$.

□

4 Relational Colour Refinement for symmetric structures

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

5 Conclusion

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

References

[1] Benjamin Scheidt and Nicole Schweikardt. Color Refinement for Relational Structures, January 2025. arXiv:2407.16022, doi:10.48550/arXiv.2407.16022.