NORMAL FORMS FOR CONNECTEDNESS IN CATEGORIES

CLAUDIO GUTIÉRREZ

Abstract. The paper gives a simple result on the existence of normal forms for the following equivalence relation between objects of a category: $A \sim B$ if and only if there are maps $A \longrightarrow B$ and $B \longrightarrow A$, under the hypothesis that the category has epi-mono factorizations and each object has finitely many sub-objects and quotient-objects.

Applications to algebra, logic, automata theory, databases are presented.

General abstract principles are useful in identifying patterns and in research. This paper presents one such principle, a simple rewriting result which holds in general categories. It is essentially the proof of facts like the existence of bases for finitely generated systems (vector spaces, systems of axioms, algebras), that a deterministic finite automata can be reduced to a minimal one, that some database queries can be refined and minimized, *etc.* It turns out that the scope of the principle is much wider than these examples.

We present the background material and framework in Section 1. The statement and the proof of the principle is in Section 2. Then, Section 3 presents some well known cases where it is applied. Finally, Section 4 shows new results proved with the help of this principle.

We will suppose a light knowledge of category theory as in the first chapters of [10]. For rewriting theory we recommend [9] and the book [3].

Throughout the paper we will use the letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ for categories, $A, B, C \ldots$ for objects of some category, $\twoheadrightarrow$ will denote epimorphisms, $\hookrightarrow$ monomorphisms, and $\cong$ isomorphisms in the corresponding category.

1. Preliminaries

1.1. **The equivalence relation $\sim$.** Let $\mathcal{C}$ be a category. We are interested in the following relation which occurs in some categories.

Definition 1. Let A, B be objects of $\mathcal{C}$. Then $A \sim B$ if and only if there are arrows $A \longrightarrow B$ and $B \longrightarrow A$.

Using the properties of a category, it follows that $\sim$ is an equivalence relation. For example, in a partial order viewed as a category, the relation $\sim$ means equality of the elements of the partial order; in the category of sets

1991 *Mathematics Subject Classification.* Primary 68Q42, Secondary 18A32, 03B35.

Key words and phrases. Rewriting, Normalization, Confluence, Categories.

with arrows injective functions, $A \sim B$ means the sets A and B have the same cardinality (Cantor-Bernstein Theorem); in the category of formulas as objects and implication as arrow, $\sim$ represents logical equivalence. More interesting examples will be presented in Sections 3 and 4.

The purpose of this note is to show that under certain finiteness conditions the equivalence classes of $\sim$ have normal (canonical) forms, and moreover, they can be obtained by a canonical rewrite system.

1.2. **Finite Categories.** We will restrict our attention to categories whose objects contain only finitely-many nested sub-objects, as formalized in the definition below. This is a slightly more general condition for an object than having a finite number of sub-objects (see remark below).

By a *proper* monomorphism (resp. epimorphism) we will mean one which is not an isomorphism.

Definition 2. Let $\mathcal{C}$ be a category. An object A of $\mathcal{C}$ is *sub-finite* if there is an integer $n_A \geq 0$ such that every chain of proper monomorphisms of the form

$$A_k \xrightarrow{m_k} A_{k-1} \longrightarrow \cdots \longrightarrow A_1 \xrightarrow{m_1} A, \tag{1}$$

has length $k \leq n_A$ (*i.e.*, if there is a chain of monos like (1) with $k > n_A$, then some m_j must be an isomorphism). The *sub-rank* of A, $\mathrm{sr}(A)$, is the minimal n_A.

The corresponding dual statements are the definitions of *quotient-finite* and quotient-rank, $\mathrm{qr}(A)$.

An object A is *finite* if it is sub-finite and quotient-finite. A category $\mathcal{C}$ is *finite* if all its objects are finite.

Remark. The statement "A is sub-finite" implies "A has finitely many sub-objects", but is not equivalent to it. The implication follows from the fact that in a chain of proper monos like (1), all $f_1, \ldots, f_k$, where $f_j = m_1 \cdots m_j : A_j \longrightarrow A$, must be different sub-objects of A. On the other hand, *e.g.*, a vector space of finite dimension $n \geq 2$ over an infinite field is sub-finite in the sense above, but has infinitely many sub-objects.

Almost all categories whose objects are intuitively "finite", are *finite* in the sense above: finite sets, finite groups, finite rings, finite algebras in general, finite geometries, finite graphs (directed, undirected, labeled), matroids, finite-dimensional vector spaces. But there are some categories that, although intuitively finite, are not, such as the free category with only one object, and one arrow f besides the identity, or the natural numbers with arrows $n \longrightarrow m$ if $n < m$ (every object is sub-finite, but not quotient-finite).

There are some simple, but useful, consequences of an object being finite.

Lemma 1.
1. *If B is sub-finite and $m : A \longrightarrow B$ is a proper monomorphism, then A is finite and* $\mathrm{sr}(A) < \mathrm{sr}(B)$.
2. *If A is quotient-finite and $e : A \longrightarrow B$ is a proper epimorphism, then B is quotient-finite and* $\mathrm{qr}(A) > \mathrm{qr}(B)$

Proof. Both statements follow directly from the definition of sr and qr and simple counting. □

Lemma 2. 1. *If A is sub-finite and $m : A \longrightarrow A$ is a monomorphism, then m is an isomorphism.*
2. *If A is quotient-finite and $e : A \longrightarrow A$ is an epimorphism, then e is an isomorphism.*

Proof. (1) Consider the chain $\cdots \xrightarrow{m} A \xrightarrow{m} A \xrightarrow{m} A$. Because A is finite, m must be isomorphism.

(2) is the dual of (1). □

Lemma 3. 1. *If A is sub-finite and there are monomorphisms $m_1 : A \longrightarrow B$ and $m_2 : B \longrightarrow A$, then $A \cong B$.*
2. *If A is quotient-finite and there are epimorphisms $e_1 : A \longrightarrow B$ and $e_2 : B \longrightarrow A$, then $A \cong B$.*

Proof. (1) $g = m_1 m_2 : B \longrightarrow B$ is mono, hence, by Lemma 2, isomorphism. So $m_1(m_2 g^{-1}) = 1_B$. Using the general fact that a monomorphic retraction is an isomorphism, it follows that $m_1 : A \longrightarrow B$ is an isomorphism.

(2) is the dual of (1). □

1.3. **The rewriting relation $\Longrightarrow$.** Roughly speaking, the relation $A \Longrightarrow B$ will reduce A to a smaller structure B which contains all the essential information of A. For example, in vector spaces, a set of generators A not linearly independent has a proper subset B of it which represents the same vector space; a finite automata A which is not minimal has superfluous vertices and edges which can be deleted to get a smaller automata B equivalent to A. In all these cases, we want to "reduce" the object until one is found that has no superfluous elements, one that is *irreducible*.

Recall that a *sub-object* of an object A is an isomorphic-equivalence class of monomorphisms $S \hookrightarrow A$, where $f : S \hookrightarrow A$ and $g : S' \hookrightarrow A$ are in the same class if and only if there is an isomorphism $h : S \longrightarrow S'$ such that $f = hg$. A quotient-object is the dual of a sub-object.

Definition 3. Let $\mathcal{C}$ be any category and A, B objects of $\mathcal{C}$. Define $A \Longrightarrow B$ if and only if B is both, a quotient- and a sub-object of A, that is, there is an epimorphism e and a monomorphism m such that

$$A \xrightarrow{e} B \xrightarrow{m} A.$$

To avoid trivial cases, we will ask also $A \ncong B$.

Note that $\Longrightarrow$ is defined modulo isomorphism (because sub- and quotient-objects are defined modulo isomorphism). By $\stackrel{*}{\Longrightarrow}$ we will denote the reflexive-transitive clousure of $\Longrightarrow$, *i.e.*, $A \stackrel{*}{\Longrightarrow} B$ if and only if either $A \cong B$ or there is a finite chain $A \Longrightarrow \cdots \Longrightarrow B$. Also by $A \stackrel{*}{\Longleftrightarrow} B$ we will denote the symmetric-reflexive-transitive closure of $\Longrightarrow$.

2. The Normalization Lemma

Lemma 4 (Normalization). *Assume $\mathcal{C}$ is a finite category with epi-mono factorization. Then the following statements hold:*

1. *The relation $\Longrightarrow$ in $\mathcal{C}$ is* sound *and* complete *for $\sim$,* i.e., *$A \overset{*}{\Longleftrightarrow} B$ if and only if $A \sim B$.*
2. *The relation $\Longrightarrow$ in $\mathcal{C}$ is* confluent[1], i.e., *if $A \overset{*}{\Longleftrightarrow} B$, then there exists C with $A \overset{*}{\Longrightarrow} C \overset{*}{\Longleftarrow} B$.*
3. *The relation $\Longrightarrow$ in $\mathcal{C}$ is* terminating, i.e., *there is no infinite chain $A \Longrightarrow A_1 \Longrightarrow A_2 \Longrightarrow \dots$*
4. *The relation $\sim$ in $\mathcal{C}$ has normal forms,* i.e., *for each $\sim$-equivalence class of objects there is a unique canonical representative (up to isomorphism).*

Proof. (1) First, $A \overset{*}{\Longleftrightarrow} B$ implies $A \sim B$. This is an easy proof by induction on the length n of the sequence $A \overset{*}{\Longleftrightarrow} B$. Recall that for $n = 0$, we have $A \cong B$, and for the inductive step, note that $A \Longrightarrow B$ implies $A \sim B$ by definition.

Second, $A \sim B$ implies $A \overset{*}{\Longleftrightarrow} B$, follows from (2), which we are going to prove next using the implication we proved above.

(2) is proved by an induction on $n = \mathrm{qr}(A) + \mathrm{qr}(B)$. Using what we proved and the definition of $\overset{*}{\Longrightarrow}$, it is enough to prove the following statement:

If $A \sim B$, then there exists an object C with

$$A \twoheadrightarrow C \hookrightarrow A$$
$$B \twoheadrightarrow C \hookrightarrow B.$$

From $A \sim B$ we have the diagram

$$A \xrightarrow{f} B \xrightarrow{g} A, \tag{2}$$

and because f and g have epi-mono factorizations, we also have

$$A \overset{e_f}{\twoheadrightarrow} A_1 \overset{m_f}{\hookrightarrow} B \quad \text{and} \quad B \overset{e_g}{\twoheadrightarrow} B_1 \overset{m_g}{\hookrightarrow} A$$

for some objects A_1 and B_1 and e_f, e_g epimorphisms and m_f, m_g monomorphisms.

Now suppose $n = 0$. Then $\mathrm{qr}(A) = \mathrm{qr}(B) = 0$. It follows that any arrow $A \longrightarrow B$ is mono (factorize it as epi-mono: then its epi-component must be an iso). The same for arrows $B \longrightarrow A$. We can apply Lemma 3 to Eq. (2) to get $A \cong B$. Choose $C = A$ or $C = B$.

Suppose $n = \mathrm{qr}(A) + \mathrm{qr}(B) > 0$. There are four possible cases:

(a) e_f and e_g are isomorphisms. Then f and g are mono, hence from Eq. (2) and Lemma 3 if follows $A \cong B$. Choose $C = A$ or $C = B$.

[1] The literature of Rewriting sometimes calls this statement 'Church-Rosser', reserving 'confluence' for the apparently weaker statement: $A \overset{*}{\Longleftarrow} D \overset{*}{\Longrightarrow} B$ implies there is C such that $A \overset{*}{\Longrightarrow} C \overset{*}{\Longleftarrow} B$. It turns out that both statements are equivalent.

(b) e_f is isomorphism but e_g is not. Then f is mono and $\mathrm{qr}(B_1) < \mathrm{qr}(B)$. Note that now we can apply the Induction Hypothesis to $A \sim B_1$ in order to get an object C with $A \twoheadrightarrow C \hookrightarrow A$ and $B_1 \twoheadrightarrow C \hookrightarrow B_1$. So we have,

$$A \twoheadrightarrow C \hookrightarrow A$$

$$B \twoheadrightarrow B_1 \twoheadrightarrow C \hookrightarrow A \overset{f}{\hookrightarrow} B.$$

(c) e_g is isomorphism but e_g is not. This is similar to case (b).

(d) Neither e_f nor e_g are isomorphisms. Then $\mathrm{qr}(A_1) < \mathrm{qr}(A)$ and $\mathrm{qr}(B_1) < \mathrm{qr}(B)$. We can apply the Induction Hypothesis to $A \sim B_1$ and to $B \sim A_1$ in order to get an object C with

$$A \twoheadrightarrow C \hookrightarrow A$$
$$B_1 \twoheadrightarrow C \hookrightarrow B_1$$

and an object D with

$$B \twoheadrightarrow D \hookrightarrow B$$
$$A_1 \twoheadrightarrow D \hookrightarrow A_1.$$

So, composing arrows from the above diagrams, we have $C \longrightarrow A \longrightarrow A_1 \longrightarrow D$ and $D \longrightarrow B \longrightarrow B_1 \longrightarrow C$, that is $C \sim D$. Also using $A_1 \twoheadrightarrow D$ and $B_1 \twoheadrightarrow C$ we get

$$\mathrm{qr}(D) + \mathrm{qr}(C) \leq \mathrm{qr}(A_1) + \mathrm{qr}(B_1) < \mathrm{qr}(A) + \mathrm{qr}(B) = n.$$

So by Induction Hypothesis on $C \sim D$ we get and object O with

$$C \twoheadrightarrow O \hookrightarrow C$$
$$D \twoheadrightarrow O \hookrightarrow D.$$

Composing arrows we finally conclude that there is an object O with

$$A \twoheadrightarrow C \twoheadrightarrow O \hookrightarrow C \hookrightarrow A$$
$$B \twoheadrightarrow D \twoheadrightarrow O \hookrightarrow D \hookrightarrow B,$$

which is what we wanted to prove.

(3) follows immediately from the definition of $\Longrightarrow$ and the fact that the objects are finite.

(4) First, let us show that each object A has a unique representative. Apply successively $\Longrightarrow$ to A until it is no more reducible. The process finishes by Part 3, and the result is unique: if $A_1 \overset{*}{\Longleftarrow} A \overset{*}{\Longrightarrow} A_2$ with A_1 and A_2 irreducible, then by Part 2 there is C with $A_1 \overset{*}{\Longrightarrow} C \overset{*}{\Longleftarrow} A_2$; but A_1 and A_2 are irreducible, hence $A_1 \cong C \cong A_2$. Finally note that the unique resulting object, call it $\mathrm{nf}(A)$, is $\sim$-equivalent to A by Part 1.

In order to prove that any two objects A, B in a $\sim$-equivalence class have the same representative, just observe that $\mathrm{nf}(A) \sim A \sim B \sim \mathrm{nf}(B)$, so $\mathrm{nf}(A)$ and $\mathrm{nf}(B)$ are both normal forms for A, hence isomorphic. $\square$

3. Applications I: Known results

We show that several well known normalization results are essentially applications of the Normalization Lemma for certain categories.

3.1. **Finite generators, independence and bases.** It is interesting to discuss what are the essential hypothesis under which standard results about finite sets of generators hold.

Let U be a set, and $\langle\ \rangle : \mathcal{P}(U) \longrightarrow \mathcal{P}(U)$ be an operator which satisfies the following conditions for $A, B \subseteq U$:

1. $A \subseteq \langle A \rangle$,
2. $\langle\langle A \rangle\rangle = \langle A \rangle$,
3. If $A \subseteq B$ then $\langle A \rangle \subseteq \langle B \rangle$.

Notice that all "generator" operators satisfy (1)-(3). We could read $\langle A \rangle$ as "subspace generated by A" in a vector space, "free algebra of terms generated by A" in algebra of a fixed signature, "sentences derivable from A" in a deductive system, *etc.* The primitive notion is $\langle\ \rangle$. A is *independent* if $\langle A \setminus \{x\} \rangle \neq \langle A \rangle$ for every $x \in A$. *A generates M* if $M \subseteq \langle A \rangle$. We want to find normal forms for the equivalence relation "A_1 is equivalent to A_2" if and only if $\langle A_1 \rangle = \langle A_2 \rangle$.

Let $S \subseteq U$ be a *finite* set, and let $\mathcal{C}$ be the category whose objects are subsets of S and whose arrows $S_1 \xrightarrow{f} S_2$ are functions $\bar{f} : S_1 \longrightarrow \mathcal{P}(S_2)$ such that $x \in \langle \bar{f}(x) \rangle$ (*i.e.*, sends x to a set that generates it). The composition of two arrows $S_1 \xrightarrow{f} S_2 \xrightarrow{g} S_3$ is given by $(g \circ f)(s) = \bigcup_{x \in \bar{f}(s)} \bar{g}(x)$. It is not difficult to check that this data forms a category (here the properties (1)-(3) above are needed). Some elementary facts about $\mathcal{C}$:

Lemma 5. *Let S_1, S_2 be objects of $\mathcal{C}$.*

1. *$S_1 \sim S_2$ if and only if $\langle S_1 \rangle = \langle S_2 \rangle$,*
2. *$S_1 \xrightarrow{f} S_2$ is* mono *if and only if for all $x \in S_1$ $\bar{f}(x) \not\subseteq \bar{f}(S \setminus \{x\})$.*
3. *$S_1 \xrightarrow{f} S_2$ is* epi *if and only if* $\mathrm{Im}(f) = S_2$.
4. *If $S_1 \cong S_2$ then $|S_1| = |S_2|$.*

Clearly $\mathcal{C}$ is finite and has images: if $S_1 \xrightarrow{f} S_2$ then $\mathrm{Im}(f) = \bigcup_{s \in S_1} \bar{f}(s)$. So the Normalization Lemma applies. A normal form in $\mathcal{C}$ is a $\Longrightarrow$-irreducible object. Observe that irreducibility implies independence, but the converse is not necessarily true as the following example shows. Consider the set $S = \{p, q, p \wedge q\}$ and $\langle\ \rangle$ to be logical deducibility. Then both $\{p, q\}$ and $\{p \wedge q\}$ are independent, *but* $\{p, q\}$ is not $\Longrightarrow$-irreducible: $\{p \wedge q\}$ is a sub- and quotient-object of $\{p, q\}$.

If we consider generating sets, the above shows that the concept of normal form is stronger than that of *base* (a set which is independent and generating) in the finite case. In fact, normal forms can exist in cases where "bases do not exist" (meaning usually that there are sets of independent generators of different sizes). Two such examples are Subsection 3.1.2 below and free modules over arbitrary rings with identity.

Let us see how the above machinery applies in two examples.

3.1.1. *Existence of bases for finite-dimensional vector spaces.* Let V be a finite-dimensional vector space, and $S \subseteq V$ a finite set. Define $\langle S \rangle$ as the set of vectors generated by S. Then clearly $\langle\ \rangle$ satisfies conditions (1)-(3) above. So S has a normal form, hence an independent subset $B \subset S$, and $\langle B \rangle = \langle S \rangle$ by Lemma 5(1). Also, Lemma 5(4) shows that that any two such bases (normal forms, hence isomorphic) have the same cardinality.

3.1.2. *Independent set of axioms.* Tarski, in his article "Fundamental Concepts of the Methodology of Deductive Sciences" (1930, see [11]), devotes one section, "Independent Sets of Sentences; Basis of a Set of Sentences" to the issue we have been discussing. In this case, we have a finite set S of sentences (in a fixed deductive system), and $\langle A \rangle$ is the set of sentences which are logically deducible from the set A. From the above discussion it follows that a normal form for S is not only an independent set $A \subseteq S$ which generates S, but also has to be of minimal size. (Tarski in his article uses "base" as synonym of "independent and generator".)

3.2. **Minimization of finite deterministic automata.** We will sketch here the discussion in [6], 3.4 and 3.5. An *automaton* is a quintuple $A = (S, \Sigma, M, s, F)$, where S is a finite set of states, Σ is a finite alphabet, $M : S \times \Sigma \longrightarrow S$ is a map, $s \in S$ is the initial state and $F \subseteq S$ is the set of final states. Consider automata over a fixed alphabet Σ as objects. Define an arrow $A \longrightarrow B$ between two automata $A = (S^A, \Sigma, M^A, s^A, F^A)$ and $B = (S^B, \Sigma, M^B, s^B, F^B)$ as a map $\varphi : S^A \longrightarrow S^B$ such that:

1. For every $\sigma \in \Sigma$, $\varphi(M^A(-, \sigma)) = M^B(\varphi(-), \sigma)$,
2. $\varphi(s^A) = s^B$,
3. $s^A \in F^A$ if and only if $\varphi(s^A) \in F^B$.

It can be proved that this is a finite category with epi-mono factorization. Also it holds $A \sim B$ if and only if $L(A) = L(B)$, that is, the automata recognize the same language. Hence, by the Normalization Lemma, there are normal forms, which are precisely *minimal automata*.

3.3. **Minimization of conjunctive queries in relational data bases.** For general background on databases see [1]. We follow the notation in [4], were conjunctive queries were introduced and minimization proved. Fix a relational language L. A *conjunctive query* is an expression of the form

$$(x_1, \ldots, x_k).\exists x_{k+1} x_{k+2} \ldots x_m . A_1 \wedge A_2 \wedge \cdots \wedge A_r \tag{3}$$

where each A_i is an atomic formula, *i.e.*, has the form $R^p_j(y_1, \ldots, y_p)$, where each y_i is either a variable x_q, $q \leq m$, or a constant a_q, and R^p_j a relational symbol.

Are there normal forms for this class of expressions? In [4], it is answered affirmatively: "For every conjunctive query there is a minimal equivalent query, unique up to isomorphism, that can be obtained from the original query by folding" (folding is essentially our rewriting rule $\Longrightarrow$).

The proof given is essentially the Normalization Lemma above. Consider the following category: Objects are conjunctive queries (on a fixed language L). For Q and Q' conjunctive queries as in (3), an arrow $Q \longrightarrow Q'$ is a function $h : FV(Q) \cup C(Q) \longrightarrow FV(Q') \cup C(Q')$ where FV denotes free variables and C constants, such that

1. $h(c) = c$ if c is a constant.
2. $(h(x_1), \dots, h(x_k)) = (x'_1, \dots, x'_k)$.
3. $R(h(x_1), \dots, h(x_p)) \in \{A'_1, \dots, A'_{r'}\}$ for each $R \in \{A_1, \dots, A_r\}$.

It can be proved that $Q_1 \sim Q_2$ if and only if the database queries Q_1 and Q_2 are equivalent. The category is finite and has epi-mono factorization. Hence the Normalization Lemma applies.

3.4. **Minimization of tableaux queries.** Soon after [4], in [2] the so called *Tableaux Queries* were introduced. Fix a language L of constants and variables. A *tableau* is a matrix (whose elements are in L) in which the columns correspond to the attributes of the universe in a fixed order. The first row of the matrix is called the *summary* of the tableaux. The remaining rows are called *rows*. Also the same variable must not appear in two different columns, and a distinguished variable symbol (*i.e.*, one which appears in the summary) must not appear in a column unless it also appears in the summary of that column. Informally, think of a tableau as a conjunctive query like (3), where the tuple $(x_1, \dots, x_k)$ is the summary and the A_i's are the rows.

The category has as objects tableaux. An arrow $T_1 \longrightarrow T_2$ between tableaux is a *containing mapping* [2], a map from the set of symbols in T_1 to the set of symbols in T_2, $h : S(T_1) \longrightarrow S(T_2)$, such that:

1. preserves distinguished variables and constants,
2. maps rows to rows.

It can be proved that $T_1 \sim T_2$ if and only if T_1 and T_2 represent the same tableau query. Again, it is easy to see that this category is finite and has epi-mono factorization. Hence there are normal forms, *i.e.*, minimal tableaux.

4. Applications II: New results

The theory of allegories, ALL, is a general calculus of relations introduced in [5]. Representable allegories, RALL, are those allegories that can be represented by sets of binary relations. In [5] it was proved that the equational theory of RALL is decidable. With the help of the Normalization Lemma we proved that there are normal forms for the terms in both theories (ALL and RALL) and showed as a corollary that the equational theory of ALL is also decidable. We will sketch the main ideas below.

4.1. **Normal forms for the equational theory of representable allegories.** The main tool to get the above results is the fact that the terms in the theory of allegories have a a nice graph-theoretical representation.

Let X be a set of labels. Define D_X as the set of all connected, directed graphs, with edges labeled by elements of X, with two distinguished vertices, the *start* s, the *finish* f, allowing multiple edges between vertices, and edges from a vertex to itself. Define by 1 the graph $\bullet_{s,f}$, which has one vertex (its start and finish) and no edges. Denote by 2_X the set of graphs in D_X with two distinct vertices and one edge. Graphically, a graph in 2_X looks like ${}_s\bullet\xrightarrow{x}\bullet_f$ or ${}_s\bullet\xleftarrow{x}\bullet_f$ for some $x \in X$.

For $g, g_1, g_2 \in D_X$, we define the following operations. The *parallel composition*, $g_1 \| g_2$, is defined as the graph obtained by (1) identifying the starts of the graphs g_1, g_2 (this is the new start), and (2) identifying the finish of the graphs g_1, g_2 (the new finish). The *serial composition*, $g_1 | g_2$, is the graph obtained by identifying the finish of g_1 with the start of g_2, and defining the new start to be s_{g_1} and the new finish f_{g_2}. The *converse* of g, denoted by g^{-1}, is obtained from g by just interchanging its start and finish. It is important to note that there is no label change.

Now define G_X as the class of graphs in D_X generated by 1 and 2_X by the above operations. The category $\mathbf{G}_X$ is defined as follows: objects, the elements of G_X; arrows, graph-homomorphisms preserving start, finish, direction and labels of edges.

The terms (over the set X) in the theory of allegories are built from $X \cup \{1\}$ and the operations $\cap$, ;, $(\ )^o$. To each term t, it is possible to associate naturally a graph $g_t \in G_X$ by the correspondence $1 \mapsto \bullet$, $x \mapsto ({}_s\bullet\xrightarrow{x}\bullet_f)$, $x^o \mapsto ({}_s\bullet\xleftarrow{x}\bullet_f)$, $\cap \mapsto \|$, $(\ )^o \mapsto (\ )^{-1}$ and $; \mapsto |$. Then we have:

Theorem 1 (Freyd-Scedrov). *The equation $r = t$ holds in the equational theory of representable allegories if and only if $g_r \longrightarrow g_t$ and $g_t \longrightarrow g_r$ in $\mathbf{G}_X$ (i.e., $g_r \sim g_r$ in $\mathbf{G}_X$).*

The category $\mathbf{G}_X$ is finite, but unfortunately has no epi-mono factorization. But we can complete $\mathbf{G}_X$ with images preserving the relation $\sim$. If we define the new set of objects by $\bar{G}_X = \{\varphi(g) : \varphi \text{ is an arrow of } \mathbf{G}_X\}$, where $\varphi(g)$ is the graph-theoretical image of g, and the arrows in $\bar{\mathbf{G}}_X$ as the arrows of $\mathbf{G}_X$ plus the obvious new ones, then the new completed category $\bar{\mathbf{G}}_X$ remains finite, has epi-mono factorization, and still holds for it Theorem 1 above. Hence the Normalization Lemma applies, getting normal forms for the theory in the form of graphs.

4.2. **Normal forms and decidability of the equational theory of allegories.** The same argument above can be done for the equational theory of (pure) allegories. In [7] it was introduced a category which captures equality of terms in this theory. The idea is similar to the representable case, now the morphisms are a little bit more involved (for details see [8]). Using the Normalization Lemma we were able to prove:

Theorem 2. 1. *The equational theory of allegories has normal forms.*
2. *The equational theory of allegories is decidable.*

The decision procedure is simple: consider two terms r, t of the theory. Translate them to graphs g_r, g_t respectively. Reduce g_r and g_t to their respective normal forms. Check if these resulting graphs are isomorphic.

References

[1] S. Abiteboul, R. Hull, V. Vianu, **Foundations of Databases**, Addison-Wesley Publ. Co., 1995.

[2] A.V. Aho, Y. Sagiv, J.D. Ullman, *Equivalences among relational expressions*, SIAM Journal of Computing, Vol.8, No.2, May 1979.

[3] F. Baader and T. Nipkow, **Term Rewriting and All That**, Cambridge University Press, 1998.

[4] A.K. Chandra and P.M. Merlin, *Optimal implementation of conjunctive queries in relational data bases*, Proceedings of the 9th. Annual ACM Symposium on the Theory of Computing, 1977.

[5] P. Freyd and A. Scedrov, **Categories and Allegories**, North Holland Math. Library, Vol. 39, 1990.

[6] A. Ginzburg, **Algebraic theory of Automata**, ACM Monograph Series, Academic Press, 1968.

[7] C. Gutiérrez, *The decidability of the Equational Theory of Allegories*, in RELMICS IV, 4th International Seminar on Relational Methods in Computer Science, Warsaw, September 1998.

[8] C. Gutiérrez, *The Arithmetic and Geometry of Allegories: Complexity and Normal Forms for a Fragment of the Theory of Relations*, PhD Dissertation, Wesleyan University, Middletown, U.S.A., 1999.

[9] N. Dershowitz and J.-P. Jouannaud. *Rewrite Systems*, in Handbook of Theoretical Computer Science, Vol. B, ed. J. van Leeuwen, MIT Press, 1990.

[10] S. MacLane, **Categories for the working mathematician**, Springer-Verlag, 1971.

[11] A. Tarski, **Logic, Semantics, Metamathematics**, Papers from1923 to 1938, Second Edition, Ed. J. Corcoran, Hackett Publ. Co., 1983.

Department of Mathematics, Wesleyan University, Middletown, CT 06459, U.S.A.

E-mail address: `cgutierrez@wesleyan.edu`