Festschrift in Honor of Joseph Goguen, LNCS 4060, pp 157-172

Complete Categorical Deduction for Satisfaction as Injectivity

Grigore Roşu

Department of Computer Science
University of Illinois at Urbana-Champaign, USA

Abstract. Birkhoff (quasi-)variety categorical axiomatizability results have fascinated many scientists by their elegance, simplicity and generality. The key factor leading to their generality is that equations, conditional or not, can be regarded as special morphisms or arrows in a special category, where their satisfaction becomes injectivity, a simple and abstract categorical concept. A natural and challenging next step is to investigate complete deduction within the same general and elegant framework. We present a categorical deduction system for equations as arrows and show that, under appropriate finiteness requirements, it is complete for satisfaction as injectivity. A straightforward instantiation of our results yields complete deduction for several equational logics, in which conditional equations can be derived as well at no additional cost, as opposed to the typical method using the theorems of constants and of deduction. At our knowledge, this is a new result in equational logics.

1 Introduction

Equational logic is an important paradigm in computer science. It admits complete deduction and is efficiently mechanizable by rewriting: CafeOBJ [15], Maude [12] and Elan [9] are equational specification and verification systems in the OBJ [21] family that can perform millions and tens of millions of rewrites per second on standard PC platforms. It is expressive: Bergstra and Tucker [5, 6] showed that any computable data type can be characterized by means of a finite equational specification, and Goguen and Malcolm [17], Wand [41], Broy, Wirsing and Pepper [11], and many others showed that equational logic is essentially strong enough to easily describe virtually all traditional programming language features. It has simple semantic models: its models are algebras, straightforward and intuitive structures. We suggest Goguen and Malcolm [19] and Padawitz and Wirsing [31] as good references for many-sorted equational logic, its completeness, as well as applications to computer science.

There are many variants and generalizations of equational logics, ranging from unsorted [7] to many-sorted [19, 31], to partial [32], to order-sorted [20, 40], to membership [27, 10], to local [13], to hidden [18, 34] equational logics, and so on. A major challenge is to develop a uniform common framework for all these variants, that allows one to formulate and prove at least some of their important properties, such as Birkhoff axiomatizability, complete deduction and

Craig interpolation. Whether this is possible or not is open, but what is certain is the existence of elegant categorical equational variants by Banaschewski and Herrlich [4], Andréka, Németi and Sain [2, 3, 30], Adámek and Rosický [1] and many others, in which equations are viewed as epimorphisms and their satisfaction as injectivity, and that these allow very general treatments of variety and quasi-variety results. We also adopt this categorical view in the present paper.

To emphasize the simplicity and generality of this approach, we mention that everything happens within only one category, denoted by $\mathcal{C}$ in this paper, which has a factorization system $\langle \mathcal{E}, \mathcal{M} \rangle$. The objects of $\mathcal{C}$ are viewed as models and the morphisms in $\mathcal{E}$, which for simplicity will be called *equations*, are viewed as sentences[1]. In order to define our sound (w.r.t. injectivity) four rule inference system for arrows in $\mathcal{E}$, $\mathcal{C}$ is required to additionally have pushouts and enough $\mathcal{E}$-projectives. To show it complete, $\mathcal{C}$ also needs to have directed colimits and to be $\mathcal{E}$-co-well-powered, and some appropriate notions of finiteness for arrows in $\mathcal{E}$ need to be introduced. A related variant by Diaconescu [14], called category-based equational logic, considers equations as pairs of arrows, one for each term, and then gives a set of deduction rules that resembles that of equational logics.

The present paper is part of our efforts to develop a unifying, categorical framework for axiomatizability, deduction and interpolation for equational and coequational logics. In [37] it is shown that the difference w.r.t. injectivity between epimorphisms of free/projective sources and epimorphisms of any sources is exactly as the difference w.r.t. usual satisfaction between unconditional and conditional equations, that is, the first define varieties while the second define quasi-varieties. In [33, 36], equational axiomatizability for hidden equational logic and coalgebra is investigated, and in [38] a categorical generalization of equational interpolation is given. The closest to the present paper is [35], where we also present a complete four rule inference system for equations as epics, but limited to unconditional axioms. In the present paper, due to crucial developments of finiteness concepts and results, especially Proposition 3, we non-trivially extend the results in [35] by eliminating the admittedly frustrating limitation to unconditional axioms, putting thus an end to our quest for complete deduction when satisfaction is injectivity. We show that a four rule inference system for epics is complete provided that all the axioms have finite conditions and the equation to be derived is finite. An interesting characteristic of our deduction system is that it is also complete for conditional equations, and that those can be derived the same way as the unconditional ones. We are not aware of any similar result for any equational paradigm in the literature until [35], where a version of it, restricted to unconditional axioms, was presented.

Section 2 recalls some categorical concepts and introduces our notational conventions. Section 3 revises factorization systems. Section 4 shows how equations, both unconditional and conditional, are equivalent to surjective morphisms and their satisfaction to injectivity; clarifying examples are presented. Section

[1] If one thinks that equations should be *regular epimorphisms* then one can read so instead of "epimorphism." Our results hold for any epimorphisms, so a restriction to regular epimorphisms would be technically artificial and less general.

5 introduces our four rule inference system for arrows and shows how it works on various examples. Finiteness concepts and results are explored in Section 6, which are necessary in Section 7 to show the completeness result. The last section concludes the paper and presents challenges for further research.

2 Preliminaries

The reader is assumed familiar with basic concepts of category theory [26, 23] and equational logics [7, 8, 31, 19]. In this section we introduce our notations and conventions, and recall some less frequent notions. Given a category $\mathcal{C}$, let $|\mathcal{C}|$ denote its class of objects; we use diagrammatic order for composition of morphisms, i.e., if $f\colon A \to B$ and $g\colon B \to C$ then $f;g\colon A \to C$. If the source or the target of a morphism is not important in a certain context, then we replace it by a bullet to avoid inventing new letters; for example, $f\colon A \to \bullet$. In situations where there are more bullet objects, they may be different. If $f\colon A \to B$ and $g\colon A \to C$ have a pushout then we let $f^g\colon C \to \bullet$ and $g^f\colon B \to \bullet$ denote the opposite arrows, up to isomorphism, of f and g in that pushout.

Given a class of morphisms $\mathcal{E}$ in a category $\mathcal{C}$, $P \in |\mathcal{C}|$ is called **$\mathcal{E}$-projective** iff for any $e\colon \bullet \to X$ in $\mathcal{E}$ and any $h\colon P \to X$, there is a g s.t. $g;e = h$. $\mathcal{C}$ **has enough $\mathcal{E}$-projectives** iff for each object $X \in |\mathcal{C}|$ there is some $\mathcal{E}$-projective object P_X and a morphism $e_X\colon P_X \to X$ in $\mathcal{E}$. It is known that any set is $\mathcal{E}$-projective where $\mathcal{E}$ consists of all the surjective functions, that free algebras are $\mathcal{E}$-projective where $\mathcal{E}$ is the class of surjective morphisms, and that the category of algebras has enough $\mathcal{E}$-projectives (for an algebra X, one can take P_X to be the free algebra over the elements in X seen as variables). Dually, I is **$\mathcal{E}$-injective** iff for any $e\colon X \to \bullet$ and any $h\colon X \to I$, there is a g s.t. $e;g = h$. $\mathcal{C}$ is called **$\mathcal{E}$-co-well-powered** iff for any $X \in |\mathcal{C}|$ and any *class* $\mathcal{D}$ of morphisms in $\mathcal{E}$ of source X, there is a *set* $\mathcal{D}' \subseteq \mathcal{D}$ such that each morphism in $\mathcal{D}$ is isomorphic to some morphism in $\mathcal{D}'$; we often call $\mathcal{D}'$ a **representative set** of $\mathcal{D}$.

If X is an object in a category $\mathcal{E}$, then $X \downarrow \mathcal{E}$ is the comma category containing morphisms $e, e', ...\colon X \to \bullet$ in $\mathcal{E}$ as objects and morphisms $h \in \mathcal{E}$ such that $e;h = e'$ as morphisms. Notice that if $\mathcal{E}$ contains only epimorphisms then there is at most one morphism between any two objects in $X \downarrow \mathcal{E}$. The intuition in our framework for the the objects $e, e', ...\colon X \to \bullet$ in the comma category $X \downarrow \mathcal{E}$ will be that of equations over the same source (variables, condition).

3 Factorization Systems

The idea to form subobjects by factoring each morphism f as $e;m$, where e is an epic and m is a mono, seems to go back to Grothendieck [22] in 1957, and was intensively used by Isbell [24], Lambek [25], Mitchell [28], and many others. Lambek was probably the first to explicitly state a diagonal-fill-in property in 1966 [25], called also “orthogonality” by Freyd and Kelly in [16]. One of the first formal definition of a factorization system that we are aware of was given by Herrlich and Strecker [23] in 1973, under the name *factorizable category*, and a comprehensive study of factorization systems, containing different equivalent definitions, was done by Németi [29] in 1982.

Definition 1. *A* **factorization system** *of a category $\mathcal{C}$ is a pair $\langle \mathcal{E}, \mathcal{M} \rangle$, s.t.:*

- *$\mathcal{E}$ and $\mathcal{M}$ are subcategories of epics and monics, respectively, in $\mathcal{C}$,*
- *all isomorphisms in $\mathcal{C}$ are both in $\mathcal{E}$ and $\mathcal{M}$, and*
- *each morphism f in $\mathcal{C}$ can be factored as $e;m$ with $e \in \mathcal{E}$ and $m \in \mathcal{M}$ "uniquely up to isomorphism", that is, if $f = e';m'$ is another factorization of f then there is a unique isomorphism α such that $e;\alpha = e'$ and $\alpha;m' = m$.*

The following are important properties of factorization systems:

Proposition 1. *Let $\langle \mathcal{E}, \mathcal{M} \rangle$ be a factorization system for $\mathcal{C}$, and let $e \in \mathcal{E}$ and $f \in \mathcal{C}$ be morphisms having the same source. Then*

1. **Diagonal-fill-in.** *If $f;m = e;g$ then there is a "unique up to isomorphism" $h \in \mathcal{C}$ such that $e;h = f$ and $h;m = g$, and*
2. **Pushout.** *If the pushout of e and f exists then $e^f \in \mathcal{E}$.*

For the rest of the paper, suppose that $\langle \mathcal{E}, \mathcal{M} \rangle$ is a factorization system for a category $\mathcal{C}$. The proof of the following proposition, which intuitively shows conditions under which "equations can be put together," can be found in [35]:

Proposition 2. *If $X \in |\mathcal{C}|$ and $\mathcal{C}$ has colimits then $X \downarrow \mathcal{E}$ has colimits.*

When $\mathcal{C}$ is $\mathcal{E}$-co-well-powered, colimits in $X \downarrow \mathcal{E}$ also exist for large diagrams $\mathcal{D}$ (whose nodes form a class): one takes the colimit of a representative set of $\mathcal{D}$.

Definition 2. *We let $(\{\gamma_i\}_{i \in I}, e_{\mathcal{D}} \colon X \to X_{\mathcal{D}})$ denote the colimit of $\mathcal{D} \subseteq X \downarrow \mathcal{E}$, and use $e_1 \cup e_2$ instead of $e_{\mathcal{D}}$ if $\mathcal{D}$ consists of only $e_1 \colon X \to \bullet$ and $e_2 \colon X \to \bullet$.*

4 Equations as Epimorphisms

As advocated by Banaschewski and Herrlich [4], by Andréka, Németi and Sain [2, 30], and by many others including the author [37, 35], equations can be regarded as epimorphisms and their satisfaction as injectivity. Readers with different background bases can find/have different explanations or intuitions for these relationships. We next informally give our version which seems closest in spirit to the subsequent results, together with some examples inspired from group theory.

An unconditional equation e over variables $x, y, ...$ is nothing but a binary relation R_e (containing only one pair) on the term algebra $T(x, y, ...)$. This relation generates a congruence C_e, which further generates a surjective morphism of free source $s_e \colon T(x, y, z, ...) \to T(x, y, ...)/_{C_e}$. An algebra satisfies e iff it is $\{s_e\}$-injective. Conversely, the kernel K_s of a surjective morphism of free source $s \colon T(x, y, ...) \to \bullet$ is nothing but a set of equations quantified by $x, y, ...$, and an algebra is $\{s\}$-injective iff satisfies K_s. It is often more convenient to work with sets of equations rather than with individual equations, as perhaps best illustrated by Craig interpolation results that do not hold for individual equations but do hold for sets of equations [39, 38]. In this paper, by equation we also mean a set of individual equations over the same variables, so there is a one-to-one correspondence between equations and epimorphisms of free sources.

Example 1. Let Σ be the unsorted signature consisting of a constant 1, a unary operation $\overline{(\_)}$ and a binary operation $\_\_$, and let us consider the equations $(\forall x)\ x1 = x$, $(\forall x)\ x\overline{x} = 1$, and $(\forall x, y, z)\ x(yz) = (xy)z$. In our notation, these equations correspond to the following three epimorphisms:

$$\begin{array}{ll} \mathsf{axiom}_1 : T_\Sigma(x) \to \bullet & \text{generated by } (x1, x), \\ \mathsf{axiom}_2 : T_\Sigma(x) \to \bullet & \text{generated by } (x\overline{x}, 1), \\ \mathsf{axiom}_3 : T_\Sigma(x, y, z) \to \bullet & \text{generated by } (x(yz), (xy)z), \end{array}$$

where $T_\Sigma(x)$ and $T_\Sigma(x, y, z)$ are the Σ-term algebras over the variable x and over the variables x, y, z, respectively, and an epimorphism $e\colon T_\Sigma(x, y, ...) \to \bullet$ is *generated* by a binary relation R of terms iff e is the natural surjection $T_\Sigma(x, y, ...) \to T_\Sigma(x, y, ...)/_R$ that maps each term to its congruence class. Notice that we could have also merged the first two epics into the epic $\mathsf{axiom}_1 \cup \mathsf{axiom}_2\colon T_\Sigma(x) \to T_\Sigma(x)/_{\{(x1,x),(x\overline{x},1)\}}$. It is known that the algebras satisfying the three equations above are exactly the groups, i.e., the left unit and left inverse equations can be proved from the above. We will focus on these proofs in the next section.

What is less known is that conditional equations can also be viewed as epics and their satisfaction as injectivity. This is explained in detail in [35]. Intuitively, one first factors the term algebra by the condition and then takes the epic generated by the equivalence classes of the conclusion.

Example 2. The conditional equation $(\forall x)\ x = 1$ **if** $xx = 1$ on groups (see Example 1), is in our notation equivalent to the epic

$$\mathsf{axiom}_4 : T_\Sigma(x)/_{(xx,1)} \to \bullet \text{ generated by } (x, 1),$$

where, for simplicity, we have identified equivalence classes with some representatives: $(x, 1)$ should normally be $(\hat{x}, \hat{1})$. A group satisfies this new axiom iff it has no proper square roots of unity iff it is $\{\mathsf{axiom}_4\}$-injective.

In theoretical efforts, it is often technically more easily to abstract freeness by projectivity. We have shown in [37] that there is essentially no difference between projective and free sources of epimorphisms with respect to axiomatizability, and that free objects are usually projective in almost any category. The results in this paper also hold for both situations, but we only discuss projective sources.

For the rest of the paper we assume that $\mathcal{C}$, besides its factorization system $\langle \mathcal{E}, \mathcal{M} \rangle$, also has enough $\mathcal{E}$-projectives. Moreover, for each object $X \in |\mathcal{C}|$ we fix an arbitrary $\mathcal{E}$-projective object P_X and an arbitrary morphism $e_X\colon P_X \to X$ in $\mathcal{E}$. If $\mathcal{C}$ is the category of algebras over some signature and X is the quotient of a free algebra by some congruence, then P_X is usually taken to be the free algebra and e_X to map each term to its congruence class.

Definition 3. *We call the morphisms in* $\mathcal{E}$ **equations**. *If* $e\colon X \to \bullet$ *is an equation then* $e_X : P_X \to X$ *is called its* **condition**. *If* $X = P_X$ *then* e *is called* **unconditional**. *An object* A *in* $\mathcal{C}$ **satisfies** *the equation* $e\colon X \to \bullet$*, written* $A \models e$*, if and only if* A *is* $\{e\}$*-injective.* $\models$ *trivially extends to sets of equations.*

5 Sound Deduction

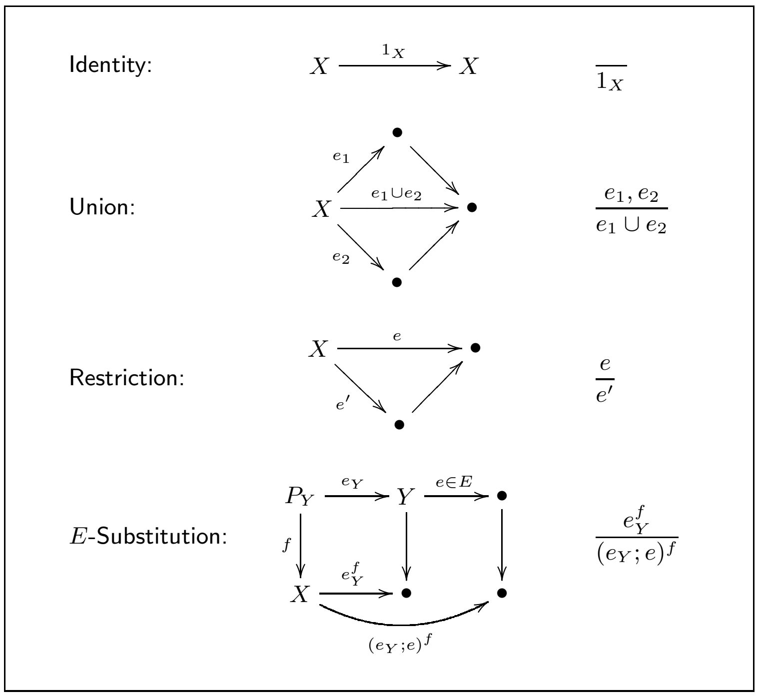


Fig. 1. Categorical inference rules.

In this section we give four inference rules for equations as arrows as defined in the previous section, show that they are sound and give some examples. The first three rules also appeared in [35]. The fourth rule appeared in an over-simplified form in [35] because conditional axioms were not allowed there.

In this section we assume that $\mathcal{C}$, besides a factorization system $\langle \mathcal{E}, \mathcal{M} \rangle$ and enough $\mathcal{E}$-projectives, also has pushouts.

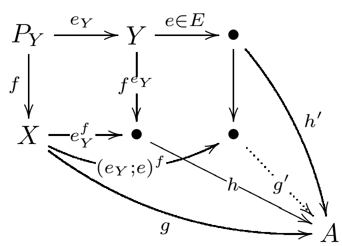
Definition 4. *Given a set of equations* E*, let* $\vdash$ *denote the derivation relation generated by the rules in Fig. 1, where* E-Substitution *is a class of rules, one for each* $f\colon P_Y \to X$*. If the source of* e *is* X *and* $E \vdash e$ *then* e *is called an*

X-derivation of E. *Let $\mathcal{D}_X(E)$ denote the full subcategory of $X \downarrow \mathcal{E}$ of X-derivations of E.*

Note that $E \vdash e$ for each $e \in E$ since one can take $f = e_Y$ in E-Substitution, and also that $\mathcal{D}_X(E)$ can be a class in general because E can be a class. Since equations in E were allowed to have only $\mathcal{E}$-projective sources in [35], E-Substitution was a simple pushout there, for which reason it was called E-Pushout.

Theorem 1. **Soundness.** $E \vdash e$ *implies* $E \models e$.

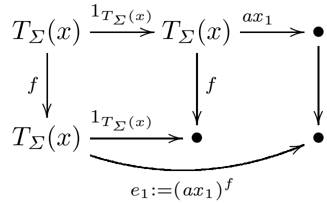
Proof. The soundness of the first three rules is easy; we only show the soundness of E-Substitution. Let us assume that $E \models e_Y^f$, let A be any object such that $A \models E$, and let $g\colon X \to A$ be any morphism.



Since $A \models e_Y^f$, there is a morphism h like in the diagram above, such that $e_Y^f; h = g$. Further, since $A \models e$ there is a morphism h' such that $e; h' = f^{e_Y}; h$. Hence $f; g = (e_Y; e); h'$, so by the pushout property there is a morphism g' such that $(e_Y; e)^f; g' = g$. Therefore, $A \models (e_Y; e)^f$, i.e., $E \models (e_Y; e)^f$.

Example 3. We show that the three arrows $E = \{\mathsf{axiom}_1, \mathsf{axiom}_2, \mathsf{axiom}_3\}$ defined in Example 1 define indeed the groups, that is, that the remaining arrows $g_1\colon T_\Sigma(x) \to \bullet$ generated by $(1x, x)$, and $g_2\colon T_\Sigma(x) \to \bullet$ generated by $(\overline{x}x, 1)$, stating the left-unit and left-inverse axioms of groups, can be derived from E. The table in Figure 2 shows a possible proof, where the first column shows or gives names to newly inferred arrows, the second shows a set of generators of the kernel of the new arrow (a dash "-" means that the set of generators is obvious, so we do not write it to save space), and the third column shows the inference rule used to derive the new arrow (identity is omitted).

To derive e_1, for example, one applies the substitution rule for $e = \mathsf{axiom}_1$ where $f\colon T_\Sigma(x) \to T_\Sigma(x)$ takes x to $\overline{x}x$, using tacitly the identity rule on $1_{T_\Sigma(x)}$:



	Generated by $(\_,\_)$	Inference rule
e_1	$(\overline{x}x)1,\ \overline{x}x$	Substitution : axiom_1
e_2	$(\overline{x}x)\overline{\overline{x}x},\ 1$	Substitution : axiom_2
e_3	$(\overline{x}x)((\overline{x}x)\overline{\overline{x}x}),\ ((\overline{x}x)(\overline{x}x))\overline{\overline{x}x}$	Substitution : axiom_3
e_4	–	Union : $e_1 \cup e_2$
e_5	–	Union : $e_3 \cup e_4$
e_6	$((\overline{x}x)(\overline{x}x))\overline{\overline{x}x},\ \overline{x}x$	Restriction : e_5
e_7	$\overline{x}1,\ \overline{x}$	Substitution : axiom_1
e_8	$\overline{x}(1x),\ (\overline{x}1)x$	Substitution : axiom_3
e_9	–	Union : $e_7 \cup e_8$
e_{10}	$\overline{x}(1x),\ \overline{x}x$	Restriction : e_9
e_{11}	$x\overline{x},\ 1$	Substitution : axiom_1
e_{12}	–	Union : $e_{10} \cup e_{11}$
e_{13}	$\overline{x}((x\overline{x})x),\ \overline{x}x$	Restriction : e_{12}
e_{14}	$x(\overline{x}x),\ (x\overline{x})x$	Substitution : axiom_3
e_{15}	–	Union : $e_{13} \cup e_{14}$
e_{16}	$\overline{x}(x(\overline{x}x)),\ \overline{x}x$	Restriction : e_{15}
e_{17}	$\overline{x}(x(\overline{x}x)),\ (\overline{x}x)(\overline{x}x)$	Substitution : axiom_3
e_{18}	–	Union : $e_{16} \cup e_{17}$
e_{19}	$(\overline{x}x)(\overline{x}x),\ \overline{x}x$	Restriction : e_{18}
e_{20}	–	Union : $e_2 \cup e_{19}$
e_{21}	$((\overline{x}x)(\overline{x}x)\overline{\overline{x}x},\ 1$	Restriction : e_{20}
e_{22}	–	Union : $e_6 \cup e_{21}$
g_2	$\overline{x}x,\ 1$	Restriction : e_{22}
e_{23}	–	Union : $e_{14} \cup g_2$
e_{24}	$x1,\ (x\overline{x})x$	Restriction : e_{23}
e_{25}	$x1,\ x$	Substitution : axiom_1
e_{26}	–	Union : $e_{24} \cup e_{25}$
e_{27}	–	Union : $e_{11} \cup e_{26}$
g_1	$1x,\ x$	Restriction : e_{27}

Fig. 2. Deriving the remaining group properties.

We showed in 29 inference steps that the three axioms define groups. The careful reader may have noticed that we have used unnecessarily many Restriction steps. Indeed, if one does all the substitutions first, followed by all the unions, and then by reductions, then one can prove the above in only 19 steps.

As mentioned before, a benefit of our deduction system is that one can also directly infer conditional equations.

Example 4. In the same context as in Example 3, one can infer the conditional equation $(\forall x)\ x = \overline{x}$ **if** $xx = 1$, which in our notation is the arrow $g_3\colon\ T_\Sigma(x)/_{(xx,1)} \to \bullet$ generated by $(x, \overline{x})$ as in the table in Figure 3. Note that e_{30} was possible since its source was $T_\Sigma(x)/_{(xx,1)}$.

Example 5. In the context of groups without square roots of unity in Example 2, where $E = \{\mathsf{axiom}_1, \mathsf{axiom}_2, \mathsf{axiom}_3, \mathsf{axiom}_4\}$, we can derive the conditional equation $(\forall x, y)\ x = y$ **if** $x\overline{y} = y\overline{x}$, which in our notation is the morphism

$T_\Sigma(x)/_{(xx,1)} \to \bullet$	Generated by $(\_,\_)$	Inference rule
$e_1,..,e_{22},g_2,..,g_1$	same as before	same as before
e_{28}	$\overline{x}(xx),(\overline{x}x)x$	Substitution : axiom_3
e_{29}	–	Union : $g_2 \cup e_{28}$
e_{30}	$1x,\overline{x}1$	Restriction : e_{29}
e_{31}	–	Union : $e_7 \cup e_{30}$
e_{32}	–	Union : $g_1 \cup e_{31}$
g_3	$x,\overline{x}$	Restriction : e_{32}

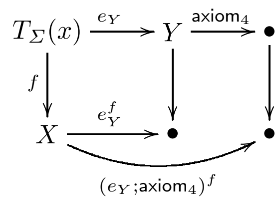
Fig. 3. Inferring a conditional property.

$g_4\colon T_\Sigma(x,y)/_{(x\overline{y},y\overline{x})} \to \bullet$ generated by (x,y). To apply substitution on axiom_4, with $f\colon T_\Sigma(x) \to T_\Sigma(x,y)/_{(x\overline{y},y\overline{x})}$ taking x to $x\overline{y}$, where $Y = T_\Sigma(x)/_{(xx,1)}$ and $X = T_\Sigma(x,y)/_{(x\overline{y},y\overline{x})}$ and where e_Y^f is generated by $((x\overline{y})(x\overline{y}),1)$ and $(e_Y;\mathsf{axiom}_4)^f$ by $(x\overline{y},1)$, we first derive e_Y^f like in Figure 4. The diagram below shows the rel-

$T_\Sigma(x)/_{(x\overline{y},y\overline{x})} \to \bullet$	Generated by $(\_,\_)$	Inference rule
$e_1,...,e_{22},g_2,...,g_1$	same as before	same as before
e_{28}	$(x\overline{y})(x\overline{y}),((x\overline{y})x)\overline{y}$	Substitution : axiom_3
e_{29}	$(x\overline{y})(x\overline{y}),((y\overline{x})x)\overline{y}$	Restriction : e_{28}
e_{30}	$y(\overline{x}x),(y\overline{x})x$	Substitution : axiom_3
e_{31}	–	Union : $e_{29} \cup e_{30}$
e_{32}	–	Union : $g_2 \cup e_{31}$
e_{33}	$(x\overline{y})(x\overline{y}),(y1)\overline{y}$	Restriction : e_{32}
e_{34}	$y1,y$	Substitution : axiom_1
e_{35}	$y\overline{y},1$	Substitution : axiom_2
e_Y^f	$(x\overline{y})(x\overline{y}),1$	Restriction : e_{35}
$(e_Y;\mathsf{axiom}_4)^f$	$x\overline{y},1$	Substitution : axiom_4
e_{36}	$x(\overline{y}y),(x\overline{y})y$	Substitution : axiom_3
e_{37}	$\overline{y}y,1$	similar to g_2
e_{38}	$1y,y$	similar to g_1
e_{39}	–	Union : $(e_Y;\mathsf{axiom}_4)^f \cup e_{36}$
e_{40}	–	Union : $e_{37} \cup e_{39}$
e_{41}	–	Union : $e_{38} \cup e_{40}$
g_4	x,y	Restriction : e_{41}

Fig. 4. xyz

evant morphisms involved in this proof:



6 Finiteness

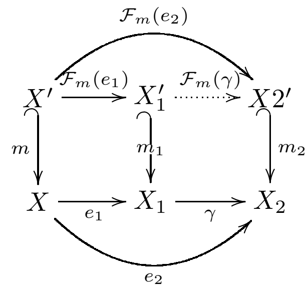
Since derivation of arrows involves a finite number of steps, one cannot expect any deduction system to be complete without some form of finiteness requirements. In this section we first recall the usual categorical concept dedicated to finiteness, then instantiate it to our framework, and then add one more requirement to factorization systems that makes them deal with finiteness smoothly.

A nonempty partially ordered set $(\mathcal{I}, \leq)$ is *directed* provided that each pair of elements has an upper bound. A *directed colimit* in a category $\mathcal{K}$ is a colimit of a diagram $D\colon (\mathcal{I}, \leq) \to \mathcal{K}$, where $(\mathcal{I}, \leq)$ is a directed poset (regarded as a category). An object K of a category $\mathcal{K}$ is *finitely presentable* provided that its hom-functor $Hom(K, \_)\colon \mathcal{K} \to \mathbf{Set}$ preserves directed colimits. It is easy to see that K is finitely presentable iff for each directed colimit $(\{\gamma_i\colon D(i) \to C\}_{i \in |\mathcal{I}|}, C)$ and each morphism $f\colon K \to C$, there is an $i \in |\mathcal{I}|$ and a unique morphism $f_i\colon K \to D(i)$ such that $f_i;\gamma_i = f$.

There are many examples of finitely presentable objects, such as finite sets and posets, finite graphs and automata, finite and discrete topological spaces, algebras presented by finitely many generators and finitely many equations, etc. We refer the interested reader to [1] for many more examples, as well as interesting properties of finitely presentable objects. What is relevant to our paper is that a surjective morphism $e\colon X \to \bullet$ of algebras is finitely presentable in the comma category of surjective morphisms of source X iff its kernel, regarded as a subalgebra of $X \times X$, is finitely generated; in our setting, where equations are surjective morphisms, that means that e stands for a finite set of equations.

Definition 5. *Equation $e\colon X \to \bullet$ is* **finite** *iff it is finitely presentable in $X \downarrow \mathcal{E}$.*

If $\mathcal{D} \subseteq X \downarrow \mathcal{E}$ is a finite diagram of finite equations, then with the notation in Definition 2, by Proposition 1.3 in [1] it follows that $e_\mathcal{D}$ is also finite. In particular, $e_1 \cup e_2$ is finite whenever e_1 and e_2 are finite, so finiteness is preserved by union. We next give conditions under which finiteness is also preserved by pushout. Given a morphism $m\colon X' \to X$ in $\mathcal{M}$, one can build "up to an isomorphism" a functor $\mathcal{F}_m\colon X \downarrow \mathcal{E} \to X' \downarrow \mathcal{E}$ as follows: for each $e\colon X \to \bullet$, let $\mathcal{F}_m(e)\colon X' \to \bullet$ be the epic by which $m;e$ factorizes, and for each $e_1\colon X \to X_1$, $e_2\colon X \to X_2$ and $\gamma\colon X_1 \to X_2$ with $e_1;\gamma = e_2$, let $\mathcal{F}_m(\gamma)$ be the unique "up to isomorphism" morphism given by the diagonal-fill-in property applied to the diagram $\mathcal{F}_m(e_2); m_2 = \mathcal{F}_m(e_1); (m_1; \gamma)$, where $m; e_1$ factors through $\mathcal{F}_m(e_1); m_1$ and $m; e_2$ factors through $\mathcal{F}_m(e_2); m_2$, like in the diagram below:



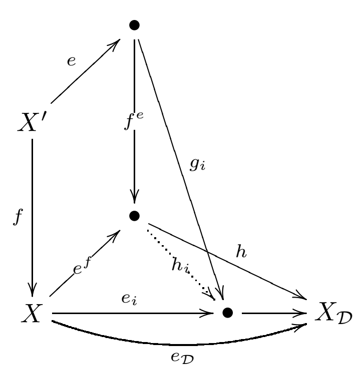
$\mathcal{F}_m(e)$ should be thought of as the restriction of e to X'. Interestingly, $\mathcal{F}_m$ does not preserve colimits in general. For example, if $\mathcal{C}$ is the category of sets then one can take $X = \{a_1, a_2, a_3\}$, $X' = \{a_1, a_3\}$, and $e_1\colon X \to \bullet$ and $e_2\colon X \to \bullet$ such that $e_1(a_1) = e_1(a_2)$ and $e_2(a_2) = e_2(a_3)$, respectively, and note that $(e_1 \cup e_2)(a_1) = (e_1 \cup e_2)(a_3)$, while $(\mathcal{F}_m(e_1) \cup \mathcal{F}_m(e_2))(a_1) \neq (\mathcal{F}_m(e_1) \cup \mathcal{F}_m(e_2))(a_3)$, where m is the inclusion $X' \subset X$. However, $\mathcal{F}_m$ does preserve directed colimits both for sets and algebras. The proof is relatively easy but takes much space, so we let it as an exercise to the interested reader (Hint: work with kernels instead of epics). With the notation above,

Definition 6. *The factorization system $\langle \mathcal{E}, \mathcal{M} \rangle$ is* **reasonable** *provided that $\mathcal{F}_m$ preserves directed colimits for each $m \in \mathcal{M}$.*

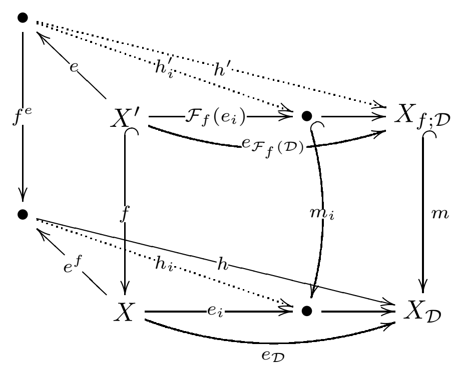
The following important property can be shown:

Proposition 3. *In the context of Proposition 1, if $\langle \mathcal{E}, \mathcal{M} \rangle$ is reasonable and e is finite, then e^f is finite.*

Proof. Due to factorization, it suffices to show the result separately for $f \in \mathcal{E}$ and for $f \in \mathcal{M}$. Let $\mathcal{D} \subseteq X \downarrow \mathcal{E}$ be a directed diagram and let h be a morphism such that $e^f; h = e_{\mathcal{D}}$ (see Definition 2).



If $f \in \mathcal{E}$ then note that $f;\mathcal{D}$ is also a directed diagram and that $e_{f;\mathcal{D}} = f; e_{\mathcal{D}}$. Since e is finite and since $e;(f^e;h) = e_{f;\mathcal{D}}$, there is some $e_i \colon X \to \bullet$ in $\mathcal{D}$ such that $f;e_i$ factors through e, i.e., there is some morphism g_i such that $f;e_i = e;g_i$. By the pushout property of e and f, it follows that there is some morphism h_i such that $e^f;h_i = e_i$ and $f^e;h_i = g_i$. Hence, e_i factors through e^f, so e^f is finite. If $f \in \mathcal{M}$ then, since $\langle \mathcal{E}, \mathcal{M} \rangle$ is reasonable, there is some morphism $m \in \mathcal{M}$ with $f;e_{\mathcal{D}} = e_{\mathcal{F}_f(\mathcal{D})};m$.



Then $e;(f^e;h) = e_{\mathcal{F}_f(\mathcal{D})};m$, so by the diagonal-fill-in property there is a morphism h' such that $e;h' = e_{\mathcal{F}_f(\mathcal{D})}$ and $h';m = f^e;h$. Since e is finite and $\mathcal{F}_f(\mathcal{D})$ is directed, there is some $e_i \in \mathcal{D}$ such that $\mathcal{F}_f(e_i)$ factors through e, so there is an h'_i with $\mathcal{F}_f(e_i) = e;h'_i$. Therefore, $e;(h'_i;m_i) = f;e_i$, where $m_i \in \mathcal{M}$ is such that $f;e_i$ factors as $\mathcal{F}_f(e_i);m_i$, so by the pushout property there is some morphism h_i such that $e^f;h_i = e_i$ and $f^e;h_i = h'_i;m_i$. Hence e_i factors through e^f, so e^f is also finite.

7 Completeness

In this section we fix the following

> Framework: A category $\mathcal{C}$ that
> - admits a reasonable factorization system $\langle \mathcal{E}, \mathcal{M} \rangle$,
> - has enough $\mathcal{E}$-projectives,
> - is $\mathcal{E}$-co-well-powered,
> - has colimits[2],

and show that, under appropriate finiteness conditions, the four rules presented in the previous section are complete wrt satisfaction as injectivity.

The usual notion of closure under inference rules is extended to classes of epics; in particular, $\mathcal{D} \subseteq X \downarrow \mathcal{E}$ is closed under E-substitution iff for any $e \colon Y \to \bullet$ in E and any $f \colon P_Y \to X$, if e_Y^f is in $\mathcal{D}$ then so is $(e_Y;e)^f$. Notice that $\mathcal{D}_X(E)$ is closed under all the four inference rules, so it is non-empty (because of closure

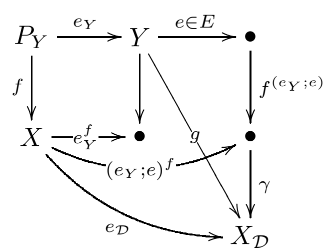
[2] Actually only directed colimits and certain pushouts are needed.

under Identity) and directed (due to closure under Union). If $\mathcal{D}_X(E)$ is not a set then, since $\mathcal{C}$ is $\mathcal{E}$-co-well-powered, it can be replaced by some representative set that it includes, so we can let $e_{\mathcal{D}_X(E)}: X \to X_{\mathcal{D}_X(E)}$ denote its colimit object, as usual (see Definition 2). Then, with the notation in Definition 2,

Theorem 2. *If E contains only equations of finite conditions, then*

1. *$X_\mathcal{D} \models E$ for any non-empty directed diagram $\mathcal{D} \subseteq X \downarrow \mathcal{E}$ closed under* Restriction *and E-*Substitution*;*
2. *For any equation e of source X, $E \models e$ iff $X_{\mathcal{D}_X(E)} \models e$;*
3. **Completeness.** *$E \models e$ implies $E \vdash e$ whenever e is finite.*

Proof. 1. Let $e: Y \to \bullet$ be any equation in E and let $g: Y \to X_\mathcal{D}$ be a morphism. Since P_Y is $\mathcal{E}$-projective and since $e_\mathcal{D} \in \mathcal{E}$, there is a morphism $f: P_Y \to X$ such that $e_Y; g = f; e_D$.



Since e_Y^f is an arrow in the pushout of f and e_Y, $e_\mathcal{D}$ factors through e_Y^f, and since e_Y^f is finite (Proposition 3) and $\mathcal{D}$ is directed and non-empty, there is an e' in $\mathcal{D}$ which factors through e_Y^f. It follows then that $e_Y^f \in \mathcal{D}$ because $\mathcal{D}$ is closed under Restriction, and further that $(e_Y; e)^f \in \mathcal{D}$ because $\mathcal{D}$ is closed under E-Substitution. Thus there is a morphism γ such that $(e_Y; e)^f; \gamma = e_\mathcal{D}$. Notice that $e_Y; (e; f^{(e_Y;e)}; \gamma) = f; (e_Y; e)^f; \gamma = f; e_\mathcal{D} = e_Y; g$, so $e; (f^{(e_Y;e)}; \gamma) = g$ because e_Y is an epimorphism. Therefore, $X_\mathcal{D} \models e$.
2. If $E \models e$ then by *1.*, noticing that $\mathcal{D}_X(E)$ is closed under Restriction and E-Substitution and is directed (because it is closed under Union) and non-empty (because it is closed under Identity), it follows that $X_{\mathcal{D}_X(E)} \models e$. Conversely, if $X_{\mathcal{D}(E)} \models e$ then there is an e' such that $e; e' = e_{\mathcal{D}_X(E)}$. Let $A \models E$ and let $h: X \to A$. Since $A \models \mathcal{D}_X(E)$, for each $e_j \in \mathcal{D}_X(E)$ there is a β_j such that $e_j; \beta_j = h$. Then A together with the morphisms β form a cocone in $\mathcal{C}$ for $\mathcal{D}_X(E)$, so there is a unique $g: X_{\mathcal{D}(E)} \to A$ such that $\gamma_j; g = \beta_j$ for all $e_j \in \mathcal{D}_X(E)$. It follows then that $e; (e'; g) = e_{\mathcal{D}_X(E)}; g = e_j; \gamma_j; g = e_j; \beta_j = h$, that is, $A \models e$.
3. $X_{\mathcal{D}_X(E)} \models e$ by *2.*, so there is an e' such that $e; e' = e_{\mathcal{D}_X(E)}$. Since e is finite and since $\mathcal{D}_X(E)$ is non-empty, there is an e_j in $\mathcal{D}_X(E)$ which factors through e. Since $E \vdash e_j$, by Restriction it follows that $E \vdash e$.

Therefore, under reasonable and necessary finiteness conditions, the four rule inference system can be used to derive any arrow e which is injectively satisfied by all objects satisfying E. On the one hand, this can be regarded as a purely categorical characterizing result, independently from logics. On the other hand, instantiated to equational logics it gives an inference system which can derive any conditional equational semantical consequence *directly*. For example, the Identity rules corresponds to reflexivity; the Union rule corresponds to closures under transitivity and congruence closures over conclusions of conditional equations, assuming that they have (provably) the same hypotheses (note that closure under symmetry is implicit, because kernels or morphisms are symmetric binary relations); the Restriction allows one to retain only a part of the conclusion of a conditional equation, in case one proved more than needed; finally, the E-Substitution rule corresponds as expected to substitution, but note that it can also derive conditional equations (when X is not projective).

8 Conclusion and Future Work

We presented a four rule categorical deduction system for a categorical abstraction of equational logics, in which equations are regarded as epimorphisms and their satisfaction as injectivity. We showed that under reasonable finiteness conditions, the four rule deduction system is complete. The research presented in this paper is part of a project aiming at developing a categorical framework in which axiomatizability, complete deduction and interpolation can be treated uniformly. Birkhoff variety and quasi-variety results for equations regarded as epics and for satisfaction regarded as injectivity are known and considered folklore among category theorists. The results in this paper show that there is also complete deduction within this framework. We are not aware of other similar categorical completeness results in the literature, except previous work by the author [35] where only unconditional axioms where supported and some interesting results by Diaconescu [14] within his category-based equational logic, where equations were regarded as parallel pairs of arrows and his five inference rules were the typical ones for equational deduction.

There is much challenging research to be done. Can the Craig-like interpolation results in [38] be instantiated to the categorical equational logic framework presented in this paper? Can the present results be dualized, hereby obtaining complete deduction for some variant of modal or coalgebraic logics? Would it be possible to implement the four rules and thus develop an arrow-based, perhaps graphical, equational reasoning engine?

Dedication. The author dedicates this paper to his former PhD adviser, Joseph Goguen, to whom he warmly thanks for all his teachings and unforgettable time spent at the University of California at San Diego. The author is also grateful to Joseph Goguen for his enthusiasm in categorical approaches to equational logics, and in particular for his encouragements in writing this material up.

References

1. J. Adámek and J. Rosický. *Locally Presentable and Accessible Categories*. Cambridge University Press, 1994. London Math. Society. Lecture Note Series 189.

2. H. Andréka and I. Németi. A general axiomatizability theorem formulated in terms of cone-injective subcategories. In B. Csakany, E. Fried, and E. Schmidt, editors, *Universal Algebra*, pages 13–35. North-Holland, 1981. Colloquia Mathematics Societas János Bolyai, 29.
3. H. Andréka and I. Németi. Generalization of the concept of variety and quasivariety to partial algebras through category theory. In *Dissertationes Mathematicae*, volume 204. Polish Scientific Publishers, 1983.
4. B. Banaschewski and H. Herrlich. Subcategories defined by implications. *Houston Journal Mathematics*, 2:149–171, 1976.
5. J. Bergstra and J. Tucker. Characterization of computable data types by means of a finite equational specification method. In J. de Bakker and J. van Leeuwen, editors, *Automata, Languages and Programming, Seventh Colloquium*, pages 76–90. Springer, 1980. Lecture Notes in Computer Science, Volume 81.
6. J. Bergstra and J. V. Tucker. Equational specifications, complete term rewriting systems, and computable and semicomputable algebras. *Journal of the Association for Computing Machinery*, 42(6):1194–1230, 1995.
7. G. Birkhoff. On the structure of abstract algebras. *Proceedings of the Cambridge Philosophical Society*, 31:433–454, 1935.
8. G. Birkhoff and J. Lipson. Heterogenous algebras. *Journal of Combinatorial Theory*, 8:115–133, 1970.
9. P. Borovanský, H. Cîrstea, H. Dubois, C. Kirchner, H. Kirchner, P.-E. Moreau, C. Ringeissen, and M. Vittek. Elan. User manual; http://www.loria.fr/equipes/protheo/SOFTWARES/ELAN.
10. A. Bouhoula, J.-P. Jouannaud, and J. Meseguer. Specification and proof in membership equational logic. *Theoretical Computer Science*, 236:35–132, 2000.
11. M. Broy, M. Wirsing, and P. Pepper. On the algebraic definition of programming languages. *ACM Trans. on Prog. Lang. and Systems*, 9(1):54–99, Jan. 1987.
12. M. Clavel, S. Eker, P. Lincoln, and J. Meseguer. Principles of Maude. In J. Meseguer, editor, *Proceedings, First International Workshop on Rewriting Logic and its Applications*. Elsevier Science, 1996. Volume 4, *Electronic Notes in Theoretical Computer Science*.
13. V. Căzănescu. Local equational logic. In Z. Esik, editor, *Proceedings, 9th International Conference on Fundamentals of Computation Theory FCT'93*, pages 162–170. Springer-Verlag, 1993. Lecture Notes in Computer Science, Volume 710.
14. R. Diaconescu. *Category-based Semantics for Equational and Constraint Logic Programming*. PhD thesis, University of Oxford, 1994.
15. R. Diaconescu and K. Futatsugi. *CafeOBJ Report: The Language, Proof Techniques, and Methodologies for Object-Oriented Algebraic Specification*. World Scientific, 1998. AMAST Series in Computing, volume 6.
16. P. Freyd and G. Kelly. Categories of continuous functors. *Journal of Pure and Applied Algebra*, 2:169–191, 1972.
17. J. Goguen and G. Malcolm. *Alg. Semantics of Imperative Programs*. MIT, 1996.
18. J. Goguen and G. Malcolm. A hidden agenda. *Theoretical Computer Science*, 245(1):55–101, August 2000.
19. J. Goguen and J. Meseguer. Completeness of many-sorted equational logic. *Houston Journal of Mathematics*, 11(3):307–334, 1985. Preliminary versions have appeared in: *SIGPLAN Notices*, July 1981, Volume 16, Number 7, pages 24–37; SRI Report CSL-135, May 1982; and Report CSLI-84-15, Stanford, Sep. 1984.
20. J. Goguen and J. Meseguer. Order-sorted algebra I: Equational deduction for multiple inheritance, overloading, exceptions and partial operations. *Theoretical Computer Science*, 105(2):217–273, 1992.

21. J. Goguen, T. Winkler, J. Meseguer, K. Futatsugi, and J.-P. Jouannaud. Introducing OBJ. In J. Goguen and G. Malcolm, editors, *Software Engineering with OBJ: algebraic specification in action*, pages 3–167. Kluwer, 2000.
22. A. Grothendieck. Sur quelques points d'algébre homologique. *Tôhoku Mathematical Journal*, 2:119–221, 1957.
23. H. Herrlich and G. Strecker. *Category Theory*. Allyn and Bacon, 1973.
24. J. R. Isbell. Subobjects, adequacy, completeness and categories of algebras. *Rozprawy Matematyczne*, 36:1–33, 1964.
25. J. Lambek. *Completions of Categories*. Springer-Verlag, 1966. Lecture Notes in Mathematics, Volume 24.
26. S. M. Lane. *Categories for the Working Mathematician*. Springer, 1971.
27. J. Meseguer. Membership algebra as a logical framework for equational specification. In *Proceedings, WADT'97*, volume 1376 of *LNCS*, pages 18–61. Springer, 1998.
28. B. Mitchell. *Theory of categories*. Academic Press, New York, 1965.
29. I. Németi. On notions of factorization systems and their applications to cone-injective subcategories. *Periodica Mathematica Hungarica*, 13(3):229–335, 1982.
30. I. Németi and I. Sain. Cone-implicational subcategories and some Birkhoff-type theorems. In B. Csakany, E. Fried, and E. Schmidt, editors, *Universal Algebra*, pages 535–578. North-Holland, 1981. Colloquia Mathematics Societas János Bolyai, 29.
31. P. Padawitz and M. Wirsing. Completeness of many-sorted equational logic revisited. *Bulletin of the European Association for Theoretical Computer Science*, 24:88–94, Oct. 1984.
32. H. Reichel. *Initial Computability, Algebraic Specifications, and Partial Algebras*. Oxford University Press, 1987.
33. G. Roşu. A Birkhoff-like axiomatizability result for hidden algebra and coalgebra. In B. Jacobs, L. Moss, H. Reichel, and J. Rutten, editors, *Proceedings of the First Workshop on Coalgebraic Methods in Computer Science (CMCS'98), Lisbon, Portugal, March 1998*, volume 11 of *Electronic Notes in Theoretical Computer Science*, pages 179–196. Elsevier Science, 1998.
34. G. Roşu. *Hidden Logic*. PhD thesis, University of California at San Diego, 2000. `http://ase.arc.nasa.gov/grosu/phd-thesis.ps`.
35. G. Roşu. Complete categorical equational deduction. In L. Fribourg, editor, *Proceedings of Computer Science Logic (CSL'01)*, volume 2142 of *Lecture Notes in Computer Science*, pages 528–538. Springer, 2001.
36. G. Roşu. Equational axiomatizability for coalgebra. *Theoretical Computer Science*, 260(1-2):229–247, 2001.
37. G. Roşu. Axiomatizability in inclusive equational logics. *Mathematical Structures in Computer Science*, to appear. `http://ase.arc.nasa.gov/grosu/iel.ps`.
38. G. Roşu and J. Goguen. On equational Craig interpolation. *Journal of Universal Computer Science*, 6(1):194–200, 2000.
39. P. H. Rodenburg. A simple algebraic proof of the equational interpolation theorem. *Algebra Universalis*, 28:48–51, 1991.
40. G. Smolka, W. Nutt, J. Goguen, and J. Meseguer. Order-sorted equational computation. In M. Nivat and H. Aït-Kaci, editors, *Resolution of Equations in Algebraic Structures, Volume 2: Rewriting Techniques*, pages 299–367. Academic, 1989.
41. M. Wand. First-order identities as a defining language. *Acta Informatica*, 14:337–357, 1980.