RJ10425 (A0712-004) December 7, 2007
Computer Science

IBM Research Report

The Structure of Inverses in Schema Mappings

Ronald Fagin, Alan Nash
IBM Research Division
Almaden Research Center
650 Harry Road
San Jose, CA 95120-6099

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties). Copies may be requested from IBM T. J. Watson Research Center , P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at http://domino.watson.ibm.com/library/CyberDig.nsf/home .

The Structure of Inverses in Schema Mappings

Ronald Fagin and Alan Nash*

ABSTRACT

A schema mapping is a specification that describes how data structured under one schema (the source schema) is to be transformed into data structured under a different schema (the target schema). The notion of an inverse of a schema mapping is subtle, because a schema mapping may associate many target instances with each source instance, and many source instances with each target instance. In PODS 2006, Fagin defined a notion of the inverse of a schema mapping. This notion is tailored to the types of schema mappings that commonly arise in practice (such as those specified by "source-to-target tuple-generating dependencies"). We resolve the key open problem of the complexity of deciding whether there is an inverse. We also explore a number of interesting questions, including: What is the structure of an inverse? When is the inverse unique? How many non-equivalent inverses can there be? When does an inverse have an inverse? How big must an inverse be? Surprisingly, these questions are all interrelated. Finally, we give greatly simplified proofs of some known results about inverses. What emerges is a much deeper understanding about this fundamental operation.

1. INTRODUCTION

Schema mappings are high-level specifications that describe the relationship between two database schemas. A schema mapping is defined to be a triple $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$, where $\mathbf{S}$ (the *source schema*) and $\mathbf{T}$ (the *target schema*) are sequences of distinct relation symbols with no relation symbols in common and Σ is a set of database dependencies that specify the association between source instances and target instances. The most important case, in both theory and practice, arises when Σ is a finite set of source-to-target tuple-generating dependencies (s-t tgds) We refer to a schema mapping specified by s-t tgds as an *s-t tgd mapping*. These mappings have also been used in data integration scenarios under the name of GLAV (global-and-local-as-view) assertions [Lenzerini 2002]. Our main focus in this paper is on inverses for s-t tgd mappings.

Since schema mappings form the essential building blocks of such crucial data inter-operability tasks as data exchange and data integration (see the surveys [Kolaitis 2005, Lenzerini 2002]), several different operators on schema mappings have been singled out as deserving study in their own right [Bernstein 2003]. The composition operator and the inverse operator have emerged as two of the most fundamental operators on schema mappings. The composition operator has been investigated in depth [Fagin, Kolaitis, Popa and Tan 2005, Madhavan and Halevy 2003, Melnik 2004, Nash, Bernstein and Melnik 2007]; however, progress on the study of the inverse operator was not made until recently. Even finding the exact semantics of this operator is a delicate task, since in spite of the traditional use of the name "mapping", a schema mapping is not simply a function that maps an instance of the source schema to an instance of the target schema. Instead, for each source instance, the schema mapping may associate many target instances. Furthermore, for each target instance, there may be many corresponding source instances.

How should the inverse be defined in our context? Let us associate with the schema mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ the set S_{12} of ordered pairs (I, J) such that I is a source instance, J is a target instance, and (I, J) satisfies Σ_{12} (written $(I, J) \models \Sigma_{12}$). Perhaps the most natural definition of the inverse of the schema mapping $\mathcal{M}_{12}$ would be a schema mapping $\mathcal{M}_{21}$ that is associated with the set $S_{21} = \{(J, I) : (I, J) \in S_{12}\}$. This reflects the standard algebraic definition of an inverse, and is the definition that [Melnik 2004] and [Melnik, Bernstein, Halevy and Rahm 2005] give for the inverse. In those papers, this definition was intended for a generic model management context, where mappings can be defined in a variety of ways, including as view definitions, relational algebra expressions, etc. However, as discussed in [Fagin 2006], this definition does not work well in our context, since the set S_{21} above is not associated with schema mappings defined by s-t tgds or natural source-to-target modifications of s-t tgds.

[Fagin 2006] showed that the identity mapping cannot be obtained by composing an s-t tgd mapping with any other schema mapping. The closest we can come with such a composition is the *copy mapping*, which is specified by s-t tgds that "copy" the source instance to the target instance. The inverse defined in [Fagin 2006], which we study in this paper, is defined essentially as follows: the composition of a mapping and its inverse is the copy mapping. This is the natural adaptation to the setting of s-t tgds of the principle that the composition of a mapping and its inverse should be the identity.

Fagin showed how to construct an inverse of an s-t tgd mapping that is itself an s-t tgd mapping when such an inverse exists. He also developed a number of tools for the study of inverses of s-t tgd mappings. He showed that deciding invertibility of an s-t tgd mapping is coNP-hard, and left open the question as to whether it is even decidable. We give a matching coNP upper bound, which shows that deciding invertibility is coNP-complete.

[Fagin, Kolaitis, Popa and Tan 2007] introduced and studied the notion of a quasi-inverse of a schema mapping. This notion is a principled relaxation of the notion of an inverse of a schema mapping; intuitively, it is obtained from the notion of an inverse by not differentiating between instances that are equivalent for data-exchange purposes. During their development, they obtained a number of results not just for quasi-inverses, but also for inverses.

*Contact information: IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120. Email: fagin@almaden.ibm.com, anash@us.ibm.com.

In particular, they showed that a certain simple combinatorial condition (the *subset property*) is a necessary and sufficient condition for an s-t tgd mapping to be invertible. They also gave an algorithm for constructing a *canonical candidate inverse* for an s-t tgd mapping. It is specified by using what they called *s-t tgds with constants and inequalities*. These are like s-t tgds, but there may also be constant formulas and inequalities in the premise. They showed that if an s-t tgd mapping is invertible, then its canonical candidate inverse is indeed an inverse.

We define *normal inverses*, that are specified by special cases of s-t tgds with constants and inequalities. The canonical candidate inverse is a normal inverse. Hence, if an s-t tgd mapping has an inverse, then it has a normal inverse. Normal inverses are especially nice, in that if I is a source instance, $\mathcal{M}$ is an s-t tgd mapping specified by Σ, and $\mathcal{M}'$ is a normal inverse of $\mathcal{M}$ that is specified by Σ', then the result of chasing I with Σ and then chasing the result by Σ' gives back exactly I (this is not true of arbitrary inverses). We focus our study mainly on normal inverses.

In addition to our result mentioned earlier where we resolve the complexity of the deciding if an s-t tgd mapping is invertible, we obtain a number of other new results about inverses, that we now discuss.

Unique inverses. As we show, no schema mapping has a unique inverse. What about a unique normal inverse? This is possible, and we give a characterization of those s-t tgd mappings with a unique normal inverse.

In the full case (where the s-t tgds have no existential quantifiers) there is an especially interesting story (which we show does not hold in the nonfull case). Let us say that a full s-t tgd mapping $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$ is *onto* if every target instance is the result of chasing some source instance with Σ. We show that if a full s-t tgd mapping is invertible and onto, then it has a unique normal inverse. What about the converse? We show that the converse fails. What if we enrich the language of possible inverses? Following [Fagin, Kolaitis, Popa and Tan 2007], we define *disjunctive tgds with inequalities* by allowing inequalities in the premise and disjunctions in the conclusion (such mappings were shown to be necessary to express quasi-inverses of full s-t tgd mappings in [Fagin, Kolaitis, Popa and Tan 2007]). We show that a full s-t tgd mapping $\mathcal{M}$ has a unique inverse specified by disjunctive tgds with inequalities if and only if $\mathcal{M}$ is invertible and onto. Furthermore, we show that $\mathcal{M}$ satifies these conditions if and only if $\mathcal{M}$ is a slight generalization of the copy mapping.

Inverse of an inverse. Surprisingly, it turns out to be rare that a normal inverse of an s-t tgd mapping is itself invertible. We show that if $\mathcal{M}$ is a full s-t tgd mapping with an invertible normal inverse, then, once again, $\mathcal{M}$ is a slight generalization of the copy mapping. By combining this result with our results about unique inverses, we obtain the unexpected result that a full s-t tgd mapping $\mathcal{M}$ has an invertible normal inverse if and only if $\mathcal{M}$ has a unique inverse specified by disjunctive tgds with inequalities. We also show that this latter theorem does not hold if we remove the restriction that $\mathcal{M}$ be full.

The size of an inverse. How big does a normal inverse need to be? We show that there is a family of full, invertible s-t tgd mappings $\mathcal{M}$ such that the size of the smallest normal inverse of $\mathcal{M}$ is exponential in the size of $\mathcal{M}$. Therefore, we broaden the class of normal mappings by allowing not just inequalities but also Boolean combinations of equalities in the premises, and we call these mappings *Boolean normal*. Allowing Boolean normal mappings does not increase the expressive power of normal mappings, but allows a more compact representation. Indeed, we show that every invertible full s-t tgd mapping has a Boolean normal inverse of polynomial size (and in fact, we give a polynomial-time algorithm for generating this Boolean normal inverse).

Is there a relationship between the number of normal inverses and the size of the minimal Boolean normal inverse? We cannot bound the number of normal inverses in terms of the size of the minimal Boolean normal inverse, since there are examples with an infinite number of inequivalent normal inverses. However, we show that if there are only a small number of inequivalent normal inverses, then the minimal number of constraints in a Boolean normal inverse is small. Specifically, we show that if $\mathcal{M}$ is a full s-t tgd mapping, with k source relation symbols and with exactly $m \geq 1$ inequivalent normal inverses, then $\mathcal{M}$ has a Boolean normal inverse with at most $k + \log_2(m)$ constraints.

Simpler proofs of known results. We give greatly simplified proofs of two results whose previous proofs were quite complex. [Fagin 2006] introduced the *unique-solutions property*, which says that no two distinct source instances have the same set of solutions. (A *solution* for a source instance I with respect to a schema mapping $\mathcal{M}$ is a target instance J such that (I, J) satisfies the constraints of $\mathcal{M}$.) He showed that the unique-solutions property is a necessary condition for a schema mapping to have an inverse. He gave a complicated proof that for LAV mappings (those specified by s-t tgds with a singleton premise), the unique-solutions property is not only a necessary condition but also a sufficient condition for invertibility. We give a simple proof of this result. Second, we give a simple proof of the result in [Fagin, Kolaitis, Popa and Tan 2007] that for invertible s-t tgd mappings $\mathcal{M}$, the canonical candidate inverse of $\mathcal{M}$ is indeed an inverse of $\mathcal{M}$.

2. PRELIMINARIES

Schemas and Schema Mappings. A *schema* $\mathbf{S}$ is a finite sequence $(R_1, \ldots, R_k)$ of relation symbols, each of a fixed arity. An *instance* I over $\mathbf{S}$ (which we may call an $\mathbf{S}$-*instance*) is a sequence $(R_1^I, \ldots, R_k^I)$, where each R_i^I is a finite relation of the same arity as R_i. We shall often use R_i to denote both the relation symbol and the relation R_i^I that interprets it.

A *schema mapping* is a triple $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$ consisting of a source schema $\mathbf{S}$, a target schema $\mathbf{T}$, and a set Σ of constraints. We say that $\mathcal{M}$ is *specified by* Σ. If Σ is a finite set of s-t tgds, then we may refer to $\mathcal{M}$ as an *s-t tgd mapping*. When $\mathbf{S}$ and $\mathbf{T}$ are clear from context, we will sometimes say Σ when we should say $(\mathbf{S}, \mathbf{T}, \Sigma)$, and talk about a set of constraints, when we should talk about a schema mapping.

Instances and Formulas. We consider instances over a two-sorted universe of *values*, which can be *constants* or *(labelled) nulls*. We assume that there is a countably infinite set C of constants and a countably infinite set N of nulls, where C and N are disjoint. We write $\mathrm{dom}(I)$ for the (active) domain of an instance I. We assume that every instance I is finite, and has values in $C \cup N$ (that is, $\mathrm{dom}(I) \subset C \cup N$). We say that I is an *instance over* $\mathbf{S}$, or an $\mathbf{S}$-instance, if the relation symbols of I and $\mathbf{S}$ are the same, with the same arities. In the context of a schema mapping $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$, we may refer to an $\mathbf{S}$-instance as a *source instance*, and a $\mathbf{T}$-instance as a *target instance*. We say that a source instance I is *ground* if $\mathrm{dom}(I) \subset C$.

If P is an m-ary relation symbol in $\mathbf{S}$, and $x_1, \ldots, x_m$ are variables, not necessarily distinct. then $P(x_1, \ldots, x_m)$ is a *relational atom*, or simply *atom* (over $\mathbf{S}$). We may refer to it as a P-*atom*. In the context of a schema mapping $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$, we may refer to a P-atom where P is in $\mathbf{S}$ as a *source atom*, and a P-atom where P is in $\mathbf{T}$ as a *target atom*. If P is an m-ary relation symbol in $\mathbf{S}$, and $c_1, \ldots, c_m$ are values (constants or nulls), not necessarily distinct. then $P(c_1, \ldots, c_m)$ is a *fact* (over $\mathbf{S}$). We may refer to it as a P-*fact*. We sometimes identify an instance with its set of facts.

We will refer to formulas that use the const predicate; the intended interpretation of const is that $\mathrm{const}(x)$ should hold precisely if x is assigned to a constant.

If δ is a conjunction of relational atoms (but no const formulas), then we define I_δ to be an instance obtained from δ as follows. For each variable v, assign a fixed constant c_v, and let the facts of I_δ consist of the facts $P(c_{v_1}, \ldots, c_{v_k})$ where $P(v_1, \ldots, v_k)$ is an atom in δ. For example, if δ is $P(x, y) \wedge Q(y)$, then I_δ is the instance $\{P(c_x, c_y), Q(c_y)\}$. If δ is a conjunction of relational atoms and const formulas, then we define I_δ as follows. For each variable v such that $\mathrm{const}(v)$ is in δ, assign a fixed constant c_v, and for each remaining variable v assign a fixed null n_v. Define I_δ be the facts that result by taking each relational atom in δ and doing the replacement we just described. For example, if δ is $P(x, y) \wedge Q(y) \wedge \mathrm{const}(x)$, then I_δ is the instance $\{P(c_x, n_y), Q(n_y)\}$. It is sometimes convenient to allow δ to contain also inequalities of the form $x \neq y$. In that case, we simply ignore the inequalities in defining I_δ.

A *renaming* of variables is a one-to-one function that maps variables to variables. A *weak renaming* of variables is a function (not necessarily one-to-one) that maps variables to variables.

Define a *prime atom* to be one that contains precisely the variables $x_1, x_2, \ldots, x_k$ for some k, and where the initial appearance of x_i precedes the initial appearance of x_j if $i < j$. For example, $P(x_1, x_2, x_1, x_3, x_2)$ is a prime atom, but $Q(x_2, x_1)$ and $R(x_2, x_3)$ are not. Note that for every relational atom, there is a unique renaming of variables to obtain a prime atom.

Constraints. All sets of constraints we consider are finite, unless otherwise specified. We consider constraints of several forms. A *source-to-target tuple-generating dependency (s-t tgd)* is a constraint of the form $\forall\bar{x}\forall\bar{y}(\alpha(\bar{x}, \bar{y}) \rightarrow \exists\bar{z}\beta(\bar{x}, \bar{z}))$, where α is a conjunction of source atoms and β is a conjunction of target atoms (we assume that the source schema $\mathbf{S}$ and the target instance $\mathbf{T}$ are given). Furthermore, there is a safety condition that every variable in $\bar{x}$ appears in both α and β. We will generally omit writing the $\forall\bar{x}\forall\bar{y}$ part. If $\bar{z}$ is empty, we say that φ is *full*.

Homomorphisms. Let J, J' be two instances. A function h that maps values to values is a *homomorphism* from J to J' if for every constant c, we have that $h(c) = c$, and for every relation symbol R and each tuple $(a_1, \ldots, a_n) \in R^J$, we have that $(h(a_1), \ldots, h(a_n)) \in R^{J'}$. We then write $J \rightarrow J'$. The instances J and J' are said to be *homomorphically equivalent* if there are homomorphisms from J to J' and from J' to J. We then write $J \leftrightarrow J'$.

Solutions and Universal Solutions. Let $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$ be a schema mapping. We say that J is a *solution* for I (under $\mathcal{M}$) if $(I, J) \models \Sigma$. We write $\mathrm{Sol}(\mathcal{M}, I)$ to denote the solutions for I under $\mathcal{M}$. We say that a solution U for the ground instance I is a *universal solution* [Fagin, Kolaitis, Miller and Popa 2005] if $U \rightarrow J$ for every solution J for I.

Composition and Inverse. We recall the concept of the *composition* of two schema mappings, introduced in [Fagin, Kolaitis, Popa and Tan 2005, Melnik 2004], and the concept of an *inverse* of a schema mapping, introduced in [Fagin 2006].

Let $\mathcal{M}_{12} = (\mathbf{S}_1, \mathbf{S}_2, \Sigma_{12})$ and $\mathcal{M}_{23} = (\mathbf{S}_2, \mathbf{S}_3, \Sigma_{23})$ be schema mappings. The *composition* $\mathcal{M}_{12} \circ \mathcal{M}_{23}$ is a schema mapping $(\mathbf{S}_1, \mathbf{S}_3, \Sigma_{13})$ such that for every $\mathbf{S}_1$-instance I and every $\mathbf{S}_3$-instance J, we have that $(I, J) \models \Sigma_{13}$ if and only if there is an $\mathbf{S}_2$-instance K such that $(I, K) \models \Sigma_{12}$ and $(K, J) \models \Sigma_{23}$. When the schemas are understood from the context, we will often write $\Sigma_{12} \circ \Sigma_{23}$ for the composition $\mathcal{M}_{12} \circ \mathcal{M}_{23}$.

Let $\widehat{\mathbf{S}}$ be a replica of the source schema $\mathbf{S}$, that is, for every relation symbol R of $\mathbf{S}$, the schema $\widehat{\mathbf{S}}$ contains a relation symbol $\widehat{R}$ that is not in $\mathbf{S}$ and has the same arity as R. We also assume that $\widehat{R}$ and $\widehat{S}$ are distinct when R and S are distinct. If A is a relational atom $R(x_1, \ldots, x_k)$, then $\widehat{A}$ is the relational atom $\widehat{R}(x_1, \ldots, x_k)$. Similarly, if F is a fact $R(c_1, \ldots, c_k)$, then $\widehat{A}$ is the fact $\widehat{R}(c_1, \ldots, c_k)$. If I is an instance of $\mathbf{S_1}$, define $\widehat{I}$ to be the corresponding instance of $\widehat{\mathbf{S_1}}$. Thus, $\widehat{I}$ consists precisely of the facts $\widehat{F}$ such that F is a fact of I. If I is a ground instance, then we may also refer to $\widehat{I}$ as a ground instance.

The *copy mapping* is the schema mapping $\mathrm{Id} = (\mathbf{S}, \widehat{\mathbf{S}}, \Sigma_{\mathrm{Id}})$, where Σ_{Id} consists of the s-t tgds $R(\bar{x}) \rightarrow \widehat{R}(\bar{x})$ as R ranges over the relation symbols in $\mathbf{S}$. Thus, $(I_1, I_2) \models \Sigma_{\mathrm{Id}}$ if and only if $\widehat{I_1} \subseteq I_2$.

Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be a schema mapping. We say that a schema mapping $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ is an *inverse* of $\mathcal{M}_{12}$ if for all ground instances I and J, we have that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $\widehat{I} \subseteq J$.

Chasing. If $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is an s-t tgd mapping, then *chasing* I with Σ produces a target instance U such that U is a universal solution for I under $\mathcal{M}$ [Fagin, Kolaitis, Miller and Popa 2005]. We may write $U = \mathrm{chase}_{12}(I)$. and say that U is *the* result of the chase. For definiteness, we use the version of the chase as defined in [Fagin, Kolaitis, Popa and Tan 2005], although it does not really matter, since whatever version of the chase we use, the results are all homomorphically equivalent. Similarly, we may write $\mathrm{chase}_{21}(I)$ for the result of chasing I with Σ_{21}. We shall also extend this notation to cases where Σ_{12} or Σ_{21} are not simply sets of s-t tgds, but where we also allow const formulas and inequalities in the premises.

3. DECIDING INVERTIBILITY

In [Fagin 2006] it is shown that deciding invertibility is coNP-hard, and it was left open as to whether it is even decidable. In this section, we prove a matching coNP upper bound, which shows that deciding invertibility is coNP-complete.

An s-t tgd mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ has the *subset property* if $I \subseteq I'$ whenever $\mathrm{Sol}(\mathcal{M}_{12}, I') \subseteq \mathrm{Sol}(\mathcal{M}_{12}, I)$. It was shown in [Fagin, Kolaitis, Popa and Tan 2007] that the subset property (which they called the $(=,=)$-subset property) is a necessary and sufficient condition for invertibility of an s-t tgd mapping. [Fagin, Kolaitis, Miller and Popa 2005] showed that if $\mathcal{M}_{12}$ is an s-t tgd mapping, then the solutions of a source instance I are exactly the homomorphic images of $\mathrm{chase}_{12}(I)$. It follows easily that there is a "homomorphic version" of the subset property, namely, that $I \subseteq I'$ whenever $\mathrm{chase}_{12}(I) \rightarrow \mathrm{chase}_{12}(I')$. This homomorphic version of the subset property is very convenient for our purposes.

We shall make use of the following proposition, whose proof (like almost all proofs in this paper) appears in the Appendix. Note that the second condition in this proposition is a special case of the homomorphic version of the subset property.

PROPOSITION 3.1. *For an s-t tgd mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$, the following are equivalent:*

1. *$\mathcal{M}_{12}$ is invertible.*
2. *For every relational atom A and instance I,*

$$\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I) \textit{ implies } I_A \subseteq I.$$

3. *For every relational atom A and instance I with at most $n_1 n_2$ facts,*

$$\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I) \textit{ implies } I_A \subseteq I$$

where n_1 is the number of facts in $\text{chase}_{12}(I_A)$ and n_2 is the maximal number of relational atoms in a premise of Σ_{12}.

Proposition 3.1 gives us a very simple proof of the desired coNP upper bound on the problem of deciding invertibility of s-t tgd mappings.

THEOREM 3.2. *The problem of deciding if an s-t tgd mapping is invertible is coNP-complete*

PROOF. The proof of coNP-hardness is in [Fagin 2006]. We now show the coNP upper bound. We make use of the equivalence of (1) and (3) in Proposition 3.1. To check that $M_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is not invertible, guess a relational atom A, an instance I such that $I_A \not\subseteq I$ where I has at most $n_1 n_2$ facts, and a homomorphism $h : \text{chase}_{12}(I_A) \to \text{chase}_{12}(I)$ where n_1 is the number of facts in $\text{chase}_{12}(I_A)$ and n_2 is the maximal number of facts in a premise of Σ_{12}. $\square$

4. STRUCTURE OF INVERSES

In this section, we study a class of mappings (that we call *normal*), which are an especially attractive choice for inverses of s-t tgd mappings. If an s-t tgd mapping has an inverse, then it has a normal inverse, because the canonical candidate inverse (defined later) is normal. Since we are interested in inverses $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ of s-t tgd mappings $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$, the normal mappings of interest to us have source schema $\mathbf{S_2}$ and target schema $\widehat{\mathbf{S_1}}$.

DEFINITION 4.1. A constraint is *normal* if it is of the form $\alpha \wedge \chi_A \wedge \eta \to A$, where α is a conjunction of source atoms, A is a target atom, χ_A is the conjunction of the formulas $\text{const}(x)$ for every variable x of A, and η is a conjunction (possibly empty) of inequalities of the form $x \neq y$ for distinct variables x, y of A. Further, there is the safety condition that every variable in A must appear in α. As usual, we suppress the leading universal quantifiers. A schema mapping is said to be normal if all of its constraints are normal.

Notice that we *require* the const predicate on all variables in A, but just *allow* inequalities on variables in A. Note also that every normal constraint is *full* (has no existential quantifiers).

Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ be schema mappings. Let us say that Σ_{21} is *too strong (for $\mathcal{M}_{12}$)* if there are ground instances I and J such that $\widehat{I} \subseteq J$ but $(I, J) \not\models \Sigma_{12} \circ \Sigma_{21}$. So Σ_{21} is not too strong precisely if whenever there are ground instances I and J such that $\widehat{I} \subseteq J$, then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$. If Σ_{12} is a set of s-t tgds, and Σ_{21} is arbitrary, then it follows from a result in [Fagin 2006] that Σ_{21} is not too strong precisely if $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$ for every ground instance I. Let us say that Σ_{21} is *too weak (for $\mathcal{M}_{12}$)* if there are ground instances I and J such that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ but $\widehat{I} \not\subseteq J$. So Σ_{21} is not too weak precisely if whenever there are ground instances I and J such that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, then $\widehat{I} \subseteq J$. It follows immediately from the definition of inverse that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if Σ_{21} is not too strong and not too weak.

If Σ_{21} is not too strong, then for all ground instance I and J where $\widehat{I} \subseteq J$, there is an instance K "in the middle" such that $(I, K) \models \Sigma_{12}$ and $(K, J) \models \Sigma_{21}$. We may say that K *witnesses* that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$. The next proposition says that if $\mathcal{M}_{21}$ is a normal inverse of $\mathcal{M}_{12}$, then an arbitrary universal solution can play the role of this witness. This is a quite useful as a tool in proving properties of normal inverses.

PROPOSITION 4.2. *Assume that $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ is a normal inverse of the s-t tgd mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$. Let I be a ground instance, and let U be an arbitrary universal solution for I with respect to $\mathcal{M}_{12}$. Then $(U, \widehat{I}) \models \Sigma_{21}$, and U witnesses $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ when $\widehat{I} \subseteq J$.*

We now give an example that shows that if $\mathcal{M}_{21}$ is not normal, then there may be no universal solution for I that witnesses $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$.

EXAMPLE 4.3. Let $\mathbf{S_1}$ consist of a unary relation symbol S, and let $\mathbf{S_2}$ consist of a binary relation symbol T. Let Σ_{12} consist of the single s-t tgd $S(x) \to \exists y (T(x, y) \wedge T(y, x))$, and let Σ_{21} consist of the single s-t tgd $T(x, y) \wedge T(y, x) \to \widehat{S}(x)$. Note that this latter s-t tgd is not a normal constraint, since it does not have the formula $\text{const}(x)$ in its premise. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$.

Let I be an arbitrary nonempty ground instance, and let U be an arbitrary universal solution for I with respect to Σ_{12}. Assume that $S(c)$ is a fact of I. Then U must contain facts $T(c, n)$ and $T(n, c)$ for some null n. If $(U, J) \models \Sigma_{21}$, then necessarily J contains the fact $\widehat{S}(n)$, which is not a fact of $\widehat{I}$. So U does not witness $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$. However, if we take K to be an instance whose facts are $\{T(c, c) : S^I(c)\}$, then it is easy to see that K witnesses $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$. It is then straightforward to see that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. Thus, $\mathcal{M}_{21}$ is a (nonnormal) inverse of $\mathcal{M}_{12}$ where no universal solution witnesses $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$.

Now let us "normalize" $\mathcal{M}_{21}$ to obtain the normal inverse $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$, where Σ'_{21} consists of the single constraint $T(x, y) \wedge T(y, x) \wedge \text{const}(x) \to \widehat{S}(x)$. Then, as Proposition 4.2 tells us, every universal solution for I witnesses $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$.

We now discuss another nice property of normal inverses. It was proven in [Fagin 2006] that if $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ are both full s-t tgd mappings, then $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if $\text{chase}_{21}(\text{chase}_{12}(I)) = \widehat{I}$ for every ground instance I. The next theorem says that this strong property (that $\text{chase}_{21}(\text{chase}_{12}(I)) = \widehat{I}$ for every ground instance I) holds for normal inverses $\mathcal{M}_{21}$, even when $\mathcal{M}_{12}$ is not full.

THEOREM 4.4. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an s-t tgd mapping and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ a normal mapping. Then $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if $\text{chase}_{21}(\text{chase}_{12}(I)) = \widehat{I}$ for every ground instance I.*

It is straightforward to verify that if $\mathcal{M}_{12}$ and $\mathcal{M}_{21}$ are as in Example 4.3, then $\text{chase}_{21}(\text{chase}_{12}(I)) \not\subseteq \widehat{I}$. It is more challenging to find an example where $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ but $\widehat{I} \not\subseteq \text{chase}_{21}(\text{chase}_{12}(I))$. An example (from [Fagin, Kolaitis, Popa and Tan 2007]) is in the Appendix.

We now introduce some useful new tools for the study of normal inverses.

4.1 Essential Conjunctions and Essential Atoms

Let Σ_{12} be a finite set of s-t tgds. Assume that A is a relational atom, and δ is a conjunction $\alpha \wedge \chi \wedge \eta$, where α is a conjunction of relational atoms, χ is a conjunction (possibly empty) of const formulas $\text{const}(x)$ for variables x in A, and η is a conjunction (possibly empty) of inequalities of the form $x \neq y$ for distinct variables x, y in A. Let us say that δ is *relevant for A (with respect to Σ_{12})* if $I_\delta \to \text{chase}_{12}(I_A)$. Note that the inequalities play no role, but are allowed for notational convenience. Let us say that δ is *demanding for A (with respect to Σ_{12})* if for every ground instance I such that $I_\delta \rightarrow \mathrm{chase}_{12}(I)$, necessarily $I_A \subseteq I$. We say that δ is *essential for A (with respect to Σ_{12})* if δ is both relevant and demanding for A with respect to Σ_{12}. It is a consequence of this definition that if δ is essential for A, then either (1) δ contains no const formulas and has exactly the same variables as A, or (2) δ contains precisely the formulas $\mathrm{const}(x)$ for every variable x of A. (The fact that in both cases, every variable in A appears in δ is Proposition 4.6 below.) Both cases are possible, because there are two meanings of I_δ: one when δ has no const formulas (which behaves as if δ has the formulas $\mathrm{const}(x)$ for every variable x of A), and one when it does have const formulas.

When Σ_{12} is full, then we are interested in the case where δ is simply a relational atom. In that case, if δ is demanding, then we call δ a *demanding atom*, and similarly we define a *relevant atom* and an *essential atom*. The reason we are interested in demanding atoms (and essential atoms) in the full case is because of the following proposition.

PROPOSITION 4.5. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be a full s-t tgd mapping. Assume that A is a source atom. Then every demanding conjunction for A contains a demanding atom for A, and every essential conjunction for A contains an essential atom for A.*

Let us say that the s-t tgd mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ has the *constant-propagation property* if for every ground instance I, every member of the active domain of I is a member of the active domain of $\mathrm{chase}_{12}(I)$ (that is, $\mathrm{dom}(I) \subseteq \mathrm{dom}(\mathrm{chase}_{12}(I))$). It is shown in [Fagin 2006] that if $\mathcal{M}_{12}$ is invertible, then it has the constant-propagation property. Similarly, we have the following proposition.

PROPOSITION 4.6. *Assume that A is a source atom, and δ is an essential conjunction for A with respect to the set Σ_{12} of s-t tgds. Then every variable in A appears in δ.*

It is easy to see that Proposition 4.6 has the following immediate corollary.

COROLLARY 4.7. *Assume that A is a source atom, and B is an essential atom for A with respect to the set Σ_{12} of full s-t tgds. Then the variables in B are exactly the variables in A.*

Recall that a *weak renaming* is a function that maps variables to variables (the word "weak" refers to the fact that the function is not necessarily one-to-one). If φ is a formula, and f is a weak renaming, let φ^f be the result of replacing every variable x in φ by $f(x)$. We may refer to φ^f as a *weak renaming of* φ. If φ is a normal constraint, then we say that f is *consistent with the inequalities of* φ if $f(x)$ and $f(y)$ are distinct for each inequality $x \neq y$ in the premise of φ. The next theorem characterizes normal inverses of s-t tgd mappings in terms of the notions of *demanding*, *relevant*, and *essential*.

THEOREM 4.8. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an s-t tgd mapping and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S}}_\mathbf{1}, \Sigma_{21})$ be a normal mapping. Then $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if*

1. *Every constraint φ in Σ_{21} is of the form $\delta \rightarrow \widehat{A}$, where δ^f is demanding for A^f for every weak renaming f consistent with the inequalities of φ.*
2. *For each source atom A, there is a relevant conjunction δ for A such that $\delta \rightarrow \widehat{A}$ is a weak renaming of a constraint in Σ_{21}. (By Part (1), this relevant conjunction is essential.)*

The proof of Theorem 4.8 proceeds by showing that part (1) of Theorem 4.8 holds precisely if Σ_{21} is not too strong, and part (2) of Theorem 4.8 holds precisely if Σ_{21} is not too weak.

DEFINITION 4.9. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be a schema mapping, where Σ_{12} is a finite set of s-t tgds. Let e be a partial function whose domain is all prime source atoms that have an essential conjunction. If the prime source atom A has an essential conjunction, then $e(A)$ is an essential conjunction for A that is a conjunction of relational atoms and the formulas $\mathrm{const}(x)$ for each variable x of A. If A has no essential conjunction, then $e(A)$ is undefined. Let Σ_{21}^e consist of all formulas $e(A) \wedge \eta_A \rightarrow \widehat{A}$, where A is a prime source atom and where $e(A)$ is defined, and η_A consists of all inequalities of the form $x \neq y$ where x and y are distinct variables of A.

The next theorem shows how we can construct an inverse out of essential conjunctions.

THEOREM 4.10. *Let $\mathcal{M}_{12}$ be an s-t tgd mapping. The following are equivalent.*

1. *$\mathcal{M}_{12}$ is invertible.*
2. *For every source atom A, there is an essential conjunction for A.*
3. *$\mathcal{M}_{21}^e$ is an inverse of $\mathcal{M}_{12}$, for every partial function e as in Definition 4.9.*
4. *$\mathcal{M}_{21}^e$ is an inverse of $\mathcal{M}_{12}$, for some partial function e as in Definition 4.9.*

In the full case, we can replace "essential conjunction" by "essential atom" in part (2) of Theorem 4.10.

4.2 The Canonical Candidate Inverse

DEFINITION 4.11. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an s-t tgd mapping. For each source atom A, let I_A be, as before, the instance containing the fact obtained by replacing each variable v in A by a distinct constant c_v. Let V_A be the result of chasing I_A with Σ_{12}. Let ν_A be the conjunction of relational atoms obtained by replacing every constant c_v of V_A by the variable v, and replacing every null n of V_A by a new variable v_n (that does not appear in A). Let χ_A be the conjunction of the formulas $\mathrm{const}(x)$ for each variable x in A, and let ω_A be the conjunction of ν_A and χ_A. Let η_A be the conjunction of all inequalities of the form $x \neq y$ where x and y are distinct variables in A, The *canonical candidate inverse* [Fagin, Kolaitis, Popa and Tan 2007] of an invertible s-t tgd mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is the normal mapping $\mathcal{M}_{21}^c = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21}^c)$ where Σ_{21}^c contains, for every prime source atom A, the constraint $\nu_A \wedge \chi_A \wedge \eta_A \rightarrow \widehat{A}$. Note that because of the constant-propagation property for invertible s-t tgd mappings, every variable in A appears in ν_A, so these constraints are well-defined.

It is shown in [Fagin, Kolaitis, Popa and Tan 2007] that if $\mathcal{M}_{12}$ is an invertible s-t tgd mapping, then the canonical candidate inverse of $\mathcal{M}_{12}$ is indeed an inverse of $\mathcal{M}_{12}$. The proof in [Fagin, Kolaitis, Popa and Tan 2007] is quite complicated. We will now give a proof, based on the following proposition, that is much simpler (given our machinery).

PROPOSITION 4.12. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an invertible s-t tgd mapping. Let A be a source atom. Then ω_A, as defined in Definition 4.11, is an essential conjunction for A.*

Thus, the role of the essential conjunction for A that is required in part (2) of Theorem 4.10 can be played by ω_A.

THEOREM 4.13. *[Fagin, Kolaitis, Popa and Tan 2007] Assume that $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is an invertible s-t tgd mapping. Then the canonical candidate inverse of $\mathcal{M}_{12}$ is indeed an inverse of $\mathcal{M}_{12}$.*

PROOF. Let e be the function that assigns to each prime source atom A the formula ω_A. By Proposition 4.12, we know that $e(A)$ is an essential conjunction for A. So by Theorem 4.10, we know that $\mathcal{M}^e_{21}$ is an inverse of $\mathcal{M}$. But $\mathcal{M}^e_{21}$ is the canonical candidate inverse of $\mathcal{M}_{12}$, and so the canonical candidate inverse of $\mathcal{M}_{12}$ is an inverse of $\mathcal{M}_{12}$, as desired. $\square$

5. UNIQUE INVERSES

Say that two schema mappings $(\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $(\mathbf{S_1}, \mathbf{S_2}, \Sigma'_{12})$ are *equivalent* if Σ_{12} and Σ'_{12} are logically equivalent. In the following theorem (and later), when we speak of "uniqueness", we mean uniqueness up to equivalence.

THEOREM 5.1. *No schema mapping has a unique inverse.*

Therefore, if we wish to study uniqueness of inverses, we must restrict our attention to particular classes (such as normal inverses). We have seen that normal inverses are an important class (in particular, every invertible s-t tgd mapping has a normal inverse) We now give an example of an s-t tgd mapping with a unique normal inverse, and another with multiple normal inverses.

EXAMPLE 5.2. Let $\mathcal{M}_{12}$ = ($\mathbf{S_1}$, $\mathbf{S_2}$, Σ_{12}), where $\mathbf{S_1}$ consists of the unary relation symbol R, where $\mathbf{S_2}$ consists of the unary relation symbol S, and where Σ_{12} consists of the tgd $R(x) \rightarrow S(x)$. Let Σ_{21} consist of the normal constraint $S(x) \wedge \mathrm{const}(x) \rightarrow \widehat{R}(x)$, and let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$. It is easy to see that $\mathcal{M}_{21}$ is a normal inverse of $\mathcal{M}_{12}$. As we shall discuss shortly, $\mathcal{M}_{21}$ is the unique normal inverse of $\mathcal{M}_{12}$.

EXAMPLE 5.3. Let $\mathbf{S_1}$ consist of the unary relation symbol R, and let $\mathbf{S_2}$ consist of the binary relation symbol S. Let Σ_{12} consist of the tgd $R(x) \rightarrow S(x, x)$. Let Σ_{21} consist of the normal constraint $S(x, x) \wedge \mathrm{const}(x) \rightarrow \widehat{R}(x)$, and let Σ'_{21} consist of the normal constraint $S(x, y) \wedge \mathrm{const}(x) \rightarrow \widehat{R}(x)$. Let $\mathcal{M}_{12}$ = ($\mathbf{S_1}$, $\mathbf{S_2}$, Σ_{12}), let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$, and let $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$. It is straightforward to verify that $\mathcal{M}_{21}$ and $\mathcal{M}'_{21}$ are inequivalent normal inverses of $\mathcal{M}_{12}$.

Because of these two examples (but where the focus was on unique inverses specified by tgds), [Fagin 2006] says, "It might be interesting to examine the question of when there is a unique inverse mapping specified in a given language."

The next theorem gives a necessary and sufficient condition, based on our notions of "essential" and "demanding", for an invertible s-t tgd mapping to have a unique normal inverse.

THEOREM 5.4. *An invertible s-t tgd mapping has a unique normal inverse if and only if for every source atom A, if δ is an essential conjunction for A, and δ' is a demanding conjunction for A, both with formulas $\mathrm{const}(x)$ for exactly the variables x that appear in A, then $I_\delta \rightarrow I_{\delta'}$.*

Assume that δ and δ' are both essential for A. By Proposition 4.6, it follows that δ and δ' each have all the variables in A. It is then not hard to show from Theorem 5.4 that if an invertible s-t tgd mapping has a unique normal inverse, and if δ and δ' are both essential for A, then I_δ and $I_{\delta'}$ are homomorphically equivalent.

Let us say that a full s-t tgd mapping is *onto* if every target instance is the result of chasing some source instance. That is, the full s-t tgd mapping $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is onto if for every target instance J there is a source instance I such that $\mathrm{chase}_{12}(I) = J$. Note that the mapping $\mathcal{M}_{12}$ of of Example 5.2 is onto, whereas the mapping $\mathcal{M}_{12}$ of Example 5.3 is not onto.

THEOREM 5.5. *A full s-t tgd mapping that is invertible and onto has a unique normal inverse.*

For example, the mapping $\mathcal{M}_{12}$ of Example 5.2, which is invertible and onto, has a unique normal inverse by Theorem 5.5.

Does the converse hold? That is, is every full s-t tgd mapping with a unique normal inverse necessarily onto? The next example shows that this is false.

EXAMPLE 5.6. Let $\mathbf{S_1}$ consist of four unary relation symbols P_i, for $1 \leq i \leq 4$, and let $\mathbf{S_2}$ consist of the four unary relation symbols Q_i, for $1 \leq i \leq 4$ and the unary relation symbol R. Let Σ_{12} consist of the full s-t tgds $P_i(x) \rightarrow Q_i(x)$, for $1 \leq i \leq 4$, along with the full s-t tgds $P_1(x) \wedge P_2(x) \rightarrow R(x)$ and $P_3(x) \wedge P_4(x) \rightarrow R(x)$. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$. The mapping $\mathcal{M}_{12}$ is not onto, since the target instance whose set of facts is $\{Q_1(0), Q_2(0)\}$ is not a universal solution for any source instance I (such an instance I must contain the facts $P_1(0), P_2(0)$, and so every solution for I must also contain the fact $R(0)$). Let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$, where $\Sigma_{21} = \left\{ Q_i(x) \wedge \mathrm{const}(x) \rightarrow \widehat{P}_i(x) : 1 \leq i \leq 4 \right\}$. Although $\mathcal{M}_{12}$ is not onto, it is shown in the Appendix that $\mathcal{M}_{12}$ has a unique normal inverse, namely $\mathcal{M}_{21}$.

Although being invertible and onto is not a necessary and sufficient condition for a full s-t tgd mapping to have a unique inverse, is there a language with a richer set of constructs where this is true? We now give such a language.

DEFINITION 5.7. A *disjunctive tgd with inequalities* is a constraint of the form $\alpha \wedge \eta \rightarrow \exists \bar{y} \beta$, where α is a conjunction of source atoms, β is a disjunction of conjunctions of target atoms, and η is a conjunction (possibly empty) of inequalities of the form $x \neq y$ for distinct free variables x, y of β. Note that this is the same restriction on inequalities that we have for normal mappings: the inequalities must involve only free variables in the conclusion. Further, there is the safety condition that every free variable in β must appear in α. Again, we suppress the leading universal quantifiers.

Disjunctive tgds with inequalities were defined in [Fagin, Kolaitis, Popa and Tan 2007], where they were shown to be rich enough to specify quasi-inverses of quasi-invertible full s-t tgd mappings. It was also shown there that inequalities in the premise and both disjunctions and existential quantifiers in the conclusion are needed in general to specify quasi-inverses of quasi-invertible full s-t tgd mappings. Note that const formulas are not allowed. Every invertible full s-t tgd mapping has an inverse specified in this language, even without the disjunctions, namely the canonical candidate inverse with the const formulas dropped. (The reason it is all right to drop the const formulas is because of a simple result in [Fagin, Kolaitis, Popa and Tan 2007] that const formulas play no role in the inverse of full s-t tgd mappings; this is Proposition A.7 in the Appendix.)

Recall that the *copy mapping*, that is used to define the inverse, is the schema mapping $\mathrm{Id} = (\mathbf{S}, \widehat{\mathbf{S}}, \Sigma_{\mathrm{Id}})$, where Σ_{Id} consists of the s-t tgds $R(\bar{x}) \rightarrow \widehat{R}(\bar{x})$ as R ranges over the relation symbols in $\mathbf{S}$. We now define a *p-copy mapping* (where the *p* stands for "partial" or "permutation") that is a generalization of the copy mapping.

DEFINITION 5.8. The schema mapping $(\mathbf{S}, \mathbf{T}, \Sigma)$ is a *p-copy mapping* if:

1. Every member of Σ is of the form
$$P(x_1, \ldots, x_k) \rightarrow Q(x_{f(1)}, \ldots, x_{f(k)}),$$
where P is a source relation symbol, Q is a target relation symbol, $x_1, \ldots, x_k$ are distinct variables, and f is a permutation of $\{1, \ldots k\}$.

2. Every source relation symbol appears in exactly one premise of Σ.

3. Every target relation symbol appears in exactly one conclusion of Σ.

For example, assume that $\mathbf{S_1}$ consists of the binary relation symbol P_1 and the ternary relation symbol P_2, and $\mathbf{S_2}$ consists of the binary relation symbol Q_1 and the ternary relation symbol Q_2. Assume that Σ_{12} consists of the s-t tgds $P_1(x, y) \rightarrow Q_1(y, x)$ and $P_2(x, y, z) \rightarrow Q_2(y, x, z)$. Then $(\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is a p-copy mapping.

The next theorem says that disjunctive tgds with inequalities form a rich enough language that a full s-t tgd mapping has a unique inverse in this language if and only if it is invertible and onto.

THEOREM 5.9. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be a full s-t tgd mapping. The following are equivalent.*

1. *$\mathcal{M}_{12}$ has a unique inverse specified by disjunctive tgds with inequalities.*
2. *$\mathcal{M}_{12}$ is invertible and onto.*
3. *$\mathcal{M}_{12}$ is equivalent to a p-copy mapping.*

Note that we cannot replace (2) in the statement of the theorem by simply "$\mathcal{M}_{12}$ is onto", because of the schema mapping with source relation symbols P and R and the single target relation symbol Q, that is specified by the tgds $P(x) \rightarrow Q(x)$, $R(x) \rightarrow Q(x)$. This schema mapping is clearly onto but not invertible.

Let us reconsider $\mathcal{M}_{12}$ from Example 5.6. It has a unique normal inverse, but since $\mathcal{M}_{12}$ is not equivalent to a p-copy mapping, it follows from Theorem 5.9 that $\mathcal{M}_{12}$ does not have a unique inverse specified by disjunctive tgds with inequalities. In addition to $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ from Example 5.6, another inverse is specified by Σ_{21} along with the disjunctive tgd $R(x) \rightarrow (\widehat{P_1}(x) \vee \widehat{P_3}(x))$.

Define a *near p-copy mapping* to be a full s-t tgd mapping $\mathcal{M} = (\mathbf{S}, \mathbf{T}, \Sigma)$ where (i) for each member σ of Σ, the premise and conclusion of σ are each singletons, with the same variables in the premise as in the conclusion, and with the variables in the conclusion all distinct, and where (ii) every member of $\mathbf{T}$ appears in the conclusion of exactly one member of Σ, and every member of $\mathbf{S}$ appears in the premise of at most one member of Σ. Thus, a near p-copy mapping may differ from being a p-copy for two reasons. First, the variables in premise are not necessarily distinct. Second, some member of $\mathbf{S}$ may fail to appear in Σ. By the *normalized version* of an s-t tgd mapping, we mean the mapping that results by adding to the premise of every tgd the formulas $\mathrm{const}(x)$ for every variable x that appears in the conclusion. Returning again to Example 5.6, we see that the unique normal inverse $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ is the normalized version of a near p-copy mapping (it is only "near", because the relation symbol R does not appear in Σ_{21}). This is not a coincidence. As a consequence of a later result (Theorem 8.2) that relates the number of normal inverses to the number of constraints in an inverse, we obtain the following result.

THEOREM 5.10. *If a full s-t tgd mapping has a unique normal inverse $\mathcal{M}_{21}$, then $\mathcal{M}_{21}$ is equivalent to the normalized version of a near p-copy mapping.*

We close this section with a explanation of why const formulas are not allowed in the language for inverses used in Theorem 5.9. Would the theorem still be true if we were to enrich the language for inverses still further to be disjunctive tgds with inequalities and constants? It turns out that allowing both const formulas and existential quantifiers makes uniqueness hopeless. For example, consider the schema mapping $\mathcal{M}_{12}$ of Example 5.2. Let σ_1 be the constraint $S(x) \wedge \mathrm{const}(x) \rightarrow \widehat{R}(x)$, and let σ_2 be the constraint $S(x) \rightarrow \exists y \widehat{R}(y)$. In addition to the inverse $\mathcal{M}_{21}$ given in Example 5.2, which is specified by σ_1, another inverse is specified by $\{\sigma_1, \sigma_2\}$. The constraint σ_2 is not logically implied by the constraint σ_1, because of the const formula in the premise of σ_1 but not σ_2. More generally, if there were an s-t tgd mapping $\mathcal{M}'_{12}$ with a unique inverse specified by disjunctive tgds with inequalities and constants, then from the implication (1) $\Rightarrow$ (3) of Theorem 5.9, it would follows that $\mathcal{M}'_{12}$ is equivalent to a p-copy mapping. But then the obvious generalization of the construction we just gave for a second inverse of $\mathcal{M}_{12}$ of Example 5.2 would show that $\mathcal{M}'_{12}$ has inequivalent inverses specified by disjunctive tgds with inequalities and constants.

6. INVERSE OF THE INVERSE

In this section, we consider the question as to when a normal inverse of a schema mapping is itself invertible. Surprisingly, it turns out to be rare that a normal inverse of an s-t tgd mapping is invertible. We begin with the full case.

THEOREM 6.1. *Let $\mathcal{M}_{12}$ be a full s-t tgd mapping. Then $\mathcal{M}_{12}$ has an invertible normal inverse if and only if $\mathcal{M}_{12}$ is equivalent to a p-copy mapping.*

The following theorem, which gives an unexpected connection between unique inverses and invertible inverses, follows immediately from Theorems 5.9 and 6.1.

THEOREM 6.2. *Let $\mathcal{M}_{12}$ be a full s-t tgd mapping. The following are equivalent.*

1. *$\mathcal{M}_{12}$ has a unique inverse specified by disjunctive tgds with inequalities.*
2. *$\mathcal{M}_{12}$ has an invertible normal inverse.*
3. *$\mathcal{M}_{12}$ is equivalent to a p-copy mapping.*

We now show by example that Theorem 6.2 fails when we drop the assumption that $\mathcal{M}_{12}$ be full.

EXAMPLE 6.3. Let $\mathbf{S_1}$ consist of the unary relation symbols P_1 and P_2, and let $\mathbf{S_2}$ consist of the unary relation symbols Q_1 and Q_2. Let Σ_{12} consist of the s-t tgds $P_1(x) \rightarrow Q_1(x)$ and $P_2(x) \rightarrow \exists y(Q_2(x) \wedge Q_1(y))$. Let Σ'_{12} consist of the normal constraints $P_1(x) \wedge \mathrm{const}(x) \rightarrow Q_1(x)$ and $P_2(x) \wedge \mathrm{const}(x) \rightarrow Q_2(x)$. Let Σ_{21} consist of the normal constraints $Q_1(x) \wedge \mathrm{const}(x) \rightarrow$

$\widehat{P_1}(x)$ and $Q_2(x) \wedge \mathrm{const}(x) \rightarrow \widehat{P_2}(x)$. Let Σ'_{21} consist of the normal constraints $Q_1(x) \wedge \mathrm{const}(x) \rightarrow \widehat{P_1}(x)$ and $Q_2(x) \wedge Q_1(y) \wedge \mathrm{const}(x) \rightarrow \widehat{P_2}(x)$. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$, let $\mathcal{M}'_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma'_{12})$, let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$, and let $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$. It is straightforward to verify that $\mathcal{M}_{21}$ and $\mathcal{M}'_{21}$ are inequivalent normal inverses of $\mathcal{M}_{12}$, and $\mathcal{M}'_{12}$ is a normal inverse of $\mathcal{M}_{21}$. So condition (2) of Theorem 6.2 holds, since $\mathcal{M}_{21}$ is a normal inverse of $\mathcal{M}_{12}$, and $\mathcal{M}'_{12}$ is an inverse of $\mathcal{M}_{21}$. However, condition (1) of Theorem 6.2 fails, since $\mathcal{M}_{12}$ has two inequivalent normal inverses, namely $\mathcal{M}_{21}$ and $\mathcal{M}'_{21}$. Furthermore, it is not hard to see that condition (3) of Theorem 6.2 fails also.

7. THE SIZE OF AN INVERSE

In this section, we consider the question of whether there is a polynomial-size inverse in some language. We show that for s-t tgd mappings, the size of the smallest normal inverse may be exponential. We also show, however, that if we expand the language to allow Boolean combinations of equalities rather than simply conjunctions of inequalities in the premise, then in the full case, there is always a polynomial-size inverse (that can be computed in polynomial time).

THEOREM 7.1. *There is a family of full s-t tgd mappings, each of which is invertible, but where the size of the smallest normal inverse is exponential in the size of the schema mapping.*

DEFINITION 7.2. A constraint is *Boolean normal* if it is of the form $\alpha \wedge \chi_A \wedge \theta \rightarrow A$, where α is a conjunction of source atoms, A is a target atom, χ_A is the conjunction of the formulas $\mathrm{const}(x)$ for every variable x of A, and θ is a Boolean combination (possibly empty) of equalities $x = y$ for variables x, y of A. Further, there is the safety condition that every variable in A must appear in α. Again, we suppress the leading universal quantifiers. A schema mapping is said to be Boolean normal if all of its constraints are Boolean normal.

Thus, we obtain the definition of "Boolean normal" from the definition of "normal" by allowing Boolean combinations of equalities in the premise, rather than simply conjunctions of inequalities. It is easy to see that every Boolean normal schema mapping is equivalent to a normal schema mapping. That is, allowing Boolean combinations of equalities in the premise, rather than simply conjunctions of inequalities. does not increase the expressive power. However, allowing Boolean combinations of equalities in the premise does potentially allow a more compact representation. In fact, the next theorem, in combination with the previous theorem, shows that this does indeed happen.

THEOREM 7.3. *There is a polynomial-time algorithm such that if the input is a schema mapping $\mathcal{M}_{12}$ specified by a finite set of full s-t tgds, then the output is a polynomial-size Boolean normal schema mapping that is an inverse of $\mathcal{M}_{12}$ if $\mathcal{M}_{12}$ has an inverse.*

It is open as to whether such a polynomial-time algorithm exists in the nonfull case. It is even open in the nonfull case as to whether or not there always exists a Boolean normal inverse of polynomial size if an inverse exists.

8. RELATING THE LENGTH OF AN INVERSE TO THE NUMBER OF INVERSES

If $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is a schema mapping, where Σ_{12} is a set of constraints, define the *length* of $\mathcal{M}_{12}$ to be the number of constraints in Σ_{12}. In this section, we show that for each full s-t tgd mapping $\mathcal{M}_{12}$, there is a relationship between the minimal length of a Boolean inverse for $\mathcal{M}_{12}$ and the number of normal inverses of $\mathcal{M}_{12}$. We first show that we cannot bound the number of inverses in terms of the minimal length of a Boolean normal inverse, since there is a full s-t tgd mapping with infinitely many distinct normal inverses. We then show that we can bound the minimal length of a Boolean normal inverse in terms of the number of inverses (and the number of relation symbols).

We begin by giving an example of a full s-t tgd mapping with infinitely many distinct normal inverses.

EXAMPLE 8.1. Let $\mathbf{S_1}$ consist of the unary relation symbol P, and let $\mathbf{S_2}$ consist of the binary relation symbol Q. Let Σ_{12} consist of the s-t tgd $P(x) \rightarrow Q(x, x)$. Let Σ^k_{21} consist of the normal constraint

$$Q(x, y_1) \wedge Q(y_1, y_2) \wedge \ldots \wedge Q(y_{k-1}, y_k) \wedge Q(y_k, x) \wedge \mathrm{const}(x) \rightarrow P(x).$$

Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$, and let $\mathcal{M}^k_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma^k_{21})$. It is straightforward to verify that for every ground instance I and for each $k \geq 1$ we have $\mathrm{chase}^k_{21}(\mathrm{chase}_{12}(I)) = \widehat{I}$ (where $\mathrm{chase}^k_{21}(J)$ is the result of chasing J with Σ^k_{21}). It therefore follows from Theorem 4.4 that $\mathcal{M}^k_{21}$ is an inverse of $\mathcal{M}_{12}$ for every k. It is also straightforward to verify that Σ^k_{21} and $\Sigma^{k'}_{21}$ are not logically equivalent if $k \neq k'$. So $\mathcal{M}_{12}$ has infinitely many inequivalent normal inverses.

The next theorem says that we can bound the minimal length of a Boolean normal inverse in terms of the number of inverses (and the number of relation symbols).

THEOREM 8.2. *Let $\mathcal{M}_{12}$ be a full s-t tgd mapping, with k source relation symbols. Assume that $\mathcal{M}_{12}$ has exactly $m \geq 1$ inequivalent normal inverses. Then $\mathcal{M}_{12}$ has a Boolean normal inverse of length at most $k + \log_2(m)$.*

Note in particular that if the s-t tgd mapping $\mathcal{M}_{12}$ has a unique normal inverse (so that $m = 1$ in Theorem 8.2) then $\mathcal{M}_{12}$ has a Boolean normal inverse of length at most k, where k is the number of source relation symbols. This is the key to proving Theorem 5.10.

It is an open problem as to whether a version of Theorem 8.2 holds in the nonfull case.

9. INVERTIBILITY IN THE LAV CASE

Recall that a schema mapping has the *unique-solutions property* if no two distinct source instances have the same set of solutions. [Fagin 2006] showed that the unique-solutions property is a necessary condition for a schema mapping to have an inverse. [Fagin 2006] also showed that for LAV mappings (those specified by s-t tgds with a singleton premise), the unique-solutions property is not only a necessary condition but also a sufficient condition for invertibility. The proof of this latter result was quite complicated. In this section, we give a very simple proof.

Just as we defined a homomorphic version of the subset property in Section 3, there is a homomorphic version of the unique-solutions property, namely, that $I = I'$ whenever $\mathrm{chase}_{12}(I) \leftrightarrow \mathrm{chase}_{12}(I')$. Note that it follows immediately from the two homomorphic versions that the subset property implies the unique-solutions property.

We now give our greatly simplified proof that the unique solutions property characterizes invertibility in the LAV case.

THEOREM 9.1. *[Fagin 2006] A LAV s-t tgd mapping is invertible if and only if it has the unique-solutions property.*

PROOF. We just noted that the subset property implies the unique-solutions property. Since satisfying the subset property is equivalent to invertibility, the "only if" direction follows (even when the s-t tgd mapping is not LAV).

Assume now that $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is a LAV mapping that satisfies the unique-solutions property. We now show that $\mathcal{M}_{12}$ satisfies the subset property, and so is invertible. Assume that I and I' are such that $\mathrm{chase}_{12}(I) \rightarrow \mathrm{chase}_{12}(I')$. Then

$$\mathrm{chase}_{12}(I \cup I') = \mathrm{chase}_{12}(I) \cup \mathrm{chase}_{12}(I') \leftrightarrow \mathrm{chase}_{12}(I')$$

where the equation follows from the fact that $\mathcal{M}_{12}$ is LAV. Then by the homomorphic version of the unique-solutions property, $I \cup I' = I'$ and therefore $I \subseteq I'$. This shows that $\mathcal{M}_{12}$ satisfies the homomorphic version of the subset property, as desired. $\square$

10. CONCLUDING REMARKS AND OPEN PROBLEMS

In addition to resolving the key problem left open in [Fagin 2006] as to the complexity of deciding if an s-t tgd mapping has an inverse, and also providing greatly simplified proofs of some known results, we have explored a number of interesting issues, about the structure of inverses, unique inverses, number of inverses, inverses of inverses, and sizes of inverses. We have shown that in the full case, these issues are, surprisingly, quite interrelated. We have also shown that in the nonfull case, these tight interconnections do not hold. As we noted in Sections 7 and 8, there remain open problems about the size or length of inverses in the nonfull case. Perhaps the most interesting open problem is whether every invertible s-t tgd mapping (not necessarily full) has a polynomial-size Boolean inverse, and if so, whether there is a polynomial-time algorithm for producing it.

11. REFERENCES

BERNSTEIN, P. A. 2003. Applying Model Management to Classical Meta-Data Problems. In *Conference on Innovative Data Systems Research (CIDR)*. 209–220.

FAGIN, R. 2006. Inverting Schema Mappings. In *ACM Symposium on Principles of Database Systems (PODS)*. 50–59. To appear, ACM Trans. on Database Systems.

FAGIN, R., KOLAITIS, P. G., MILLER, R. J., AND POPA, L. 2005. Data Exchange: Semantics and Query Answering. *Theoretical Computer Science (TCS) 336,* 1, 89–124. A preliminary version of this paper appeared in the 2003 ICDT conference.

FAGIN, R., KOLAITIS, P. G., POPA, L., AND TAN, W.-C. 2005. Composing Schema Mappings: Second-order Dependencies to the Rescue. *ACM Transactions on Database Systems (TODS) 30,* 4, 994–1055. A preliminary version of this paper appeared in the 2004 PODS conference.

FAGIN, R., KOLAITIS, P. G., POPA, L., AND TAN, W.-C. 2007. Quasi-Inverses of Schema Mappings. In *ACM Symposium on Principles of Database Systems (PODS)*.

KOLAITIS, P. G. 2005. Schema Mappings, Data Exchange, and Metadata Management. In *ACM Symposium on Principles of Database Systems (PODS)*. 61–75.

LENZERINI, M. 2002. Data Integration: A Theoretical Perspective. In *ACM Symposium on Principles of Database Systems (PODS)*. 233–246.

MADHAVAN, J. AND HALEVY, A. Y. 2003. Composing Mappings Among Data Sources. In *International Conference on Very Large Data Bases (VLDB)*. 572–583.

MELNIK, S. 2004. *Generic Model Management: Concepts and Algorithms*. Lecture Notes in Computer Science, vol. 2967. Springer.

MELNIK, S., BERNSTEIN, P. A., HALEVY, A., AND RAHM, E. 2005. Applying Model Management to Executable Mappings. In *SIGMOD*. 167–178.

NASH, A., BERNSTEIN, P. A., AND MELNIK, S. 2007. Composition of Mappings Given by Embedded Dependencies. *ACM Transactions on Database Systems (TODS) 32,* 1. A preliminary version of this paper appeared in the 2005 PODS conference.

APPENDIX

A. PROOFS

A.1 Proofs for Section 3

Proof of Proposition 3.1. We shall show that (1) and (2) are equivalent, and then show that (2) and (3) are equivalent.

Assume (1) holds. So $\mathcal{M}_{12}$ satisfies the subset property. Pick an atom A and an instance I such that $\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I)$. By the homomorphic version of the subset property, it follows that $I_A \subseteq I$. Therefore, (2) holds.

Now assume that (1) fails. Therefore, the homomorphic version of the subset property fails. Hence, there are I and I' such that $\mathrm{chase}_{12}(I) \rightarrow \mathrm{chase}_{12}(I')$ and $I \not\subseteq I'$. Since $I \not\subseteq I'$, we can assume (by renaming constants if needed) that there is an atom A such that $I_A \subseteq I$ but $I_A \not\subseteq I'$. Since $I_A \subseteq I$, we have $\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I)$. Since also $\mathrm{chase}_{12}(I) \rightarrow \mathrm{chase}_{12}(I')$, we have that $\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I')$, witnessing that (2) fails (where I' plays the role of I).

Now notice that if $\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I)$ then necessarily also $\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I')$ for some $I' \subseteq I$ with at most $n_1 n_2$ facts, since $\mathrm{chase}_{12}(I_A)$ maps into at most n_1 facts in $\mathrm{chase}_{12}(I)$ and each of those facts can be introduced into $\mathrm{chase}_{12}(I)$ by firing a single tgd in Σ_{12} on at most n_2 facts in I. It follows that (2) and (3) are equivalent. $\Box$

A.2 Proofs for Section 4

Proof of Proposition 4.2. Since $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, we know that $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$ and therefore there exists some K such that $(I, K) \models \Sigma_{12}$ and $(K, \widehat{I}) \models \Sigma_{21}$. Let U be an arbitrary universal solution for I with respect to $\mathcal{M}_{12}$. Then there is a homomorphism $h : U \rightarrow K$ that is the identity on I. Pick a constraint $\varphi \in \Sigma_{21}$; by our normality assumption, it must be of the form

$$\alpha(\bar{x}, \bar{y}) \wedge \chi_A(\bar{x}) \wedge \eta(\bar{x}) \rightarrow \widehat{A}(\bar{x}).$$

Assume that U satisfies the premise of φ on $\bar{a}, \bar{b}$. Then $K \models \alpha(h(\bar{a}), h(\bar{b}))$. Since $U \models \chi_A(\bar{a})$, we have $h(\bar{a}) = \bar{a}$ and therefore $K \models \alpha(\bar{a}, h(\bar{b}))$. Since $(K, \widehat{I}) \models \Sigma_{21}$, we must have $\widehat{I} \models \widehat{A}(\bar{a})$. This shows that $(U, \widehat{I}) \models \varphi$. Since φ is an arbitrary member of Σ_{21}, it follows that $(U, \widehat{I}) \models \Sigma_{21}$, as desired. Since $\widehat{I} \subseteq J$, it follows easily that $(U, J) \models \Sigma_{21}$. Since $(I, U) \models \Sigma_{12}$ and $(U, J) \models \Sigma_{21}$, we have that U witnesses $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, as desired. $\Box$

LEMMA A.1. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an s-t tgd mapping and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S}}_\mathbf{1}, \Sigma_{21})$ be a normal mapping. Then Σ_{21} is not too strong if and only if $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) \subseteq \widehat{I}$ for every ground instance I.*

PROOF. Assume first that $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) \subseteq \widehat{I}$ for every ground instance I. We must show that whenever there are ground instances I and J such that $\widehat{I} \subseteq J$, then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$. Let I and J be ground instances such that $\widehat{I} \subseteq J$. Let $U = \mathrm{chase}_{12}(I)$, and let $U' = \mathrm{chase}_{21}(U)$. So $(I, U) \models \Sigma_{12}$ and $(U, U') \models \Sigma_{21}$. Also, by assumption, $U' \subseteq \widehat{I}$. Since also $\widehat{I} \subseteq J$, it follows that $U' \subseteq J$. Since $(U, U') \models \Sigma_{21}$ and $U' \subseteq J$, we see that $(U, J) \models \Sigma_{21}$. Since $(I, U) \models \Sigma_{12}$ and $(U, J) \models \Sigma_{21}$, it follows that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, as desired.

Assume now that Σ_{21} is not too strong. So $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$. Let $U = \mathrm{chase}_{12}(I)$, and let $U' = \mathrm{chase}_{21}(U)$. We need only show that $U' \subseteq \widehat{I}$. The argument in the proof of Proposition 4.2 shows that $(U, \widehat{I}) \models \Sigma_{21}$. Since $\mathcal{M}_{21}$ is a normal mapping, it is easy to see that the result of chasing an arbitrary instance with Σ_{21} is a ground instance. In particular, U' is a ground instance. Since U' is the result of chasing U with Σ_{21}, and U' is ground, a standard property of the chase tells us that for every instance J such that $(U, J) \models \Sigma_{21}$, necessarily $U' \subseteq J$. If we take J to be $\widehat{I}$, then we see that $U' \subseteq \widehat{I}$, as desired. $\Box$

LEMMA A.2. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an s-t tgd mapping and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S}}_\mathbf{1}, \Sigma_{21})$ be a normal mapping. Then Σ_{21} is not too weak if and only if $\widehat{I_A} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I_A))$ for every source atom A.*

PROOF. Assume first that Σ_{21} is not too weak. So whenever there are ground instances I and J such that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, then $\widehat{I} \subseteq J$. Let I be I_A, and let J be $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I_A))$ Then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$. So $\widehat{I} \subseteq J$, that is, $\widehat{I_A} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I_A))$, as desired.

Assume now that $\widehat{I_A} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I_A))$ for every source atom A. By renaming constants if needed, it follows that if F is a ground fact, then $\widehat{F} \in \mathrm{chase}_{21}(\mathrm{chase}_{12}(\{F\}))$. Let I and J be ground instances such that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$; we must show that $\widehat{I} \subseteq J$. Let F be an arbitrary fact in I. Since $\widehat{F}$ is a ground fact in $\mathrm{chase}_{21}(\mathrm{chase}_{12}(\{F\}))$ and since $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, it follows by a standard property of the chase (that we also invoked in the proof of Lemma A.1) that that $\widehat{F} \in J$. Since $\widehat{F}$ is an arbitrary fact in $\widehat{I}$, we see that $\widehat{I} \subseteq J$, as desired. $\Box$

Proof of Theorem 4.4. Assume first that $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) = \widehat{I}$ for every ground instance I. By Lemmas A.1 and A.2, we know that Σ_{21} is not too strong and not too weak. So $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$.

Conversely, assume that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. Therefore, Σ_{21} is not too weak. Let I be a ground instance. Then $\widehat{I} = \cup_{F \in I} \widehat{F}$, which by Lemma A.2 is contained in $\cup_{F \in I} \mathrm{chase}_{21}(\mathrm{chase}_{12}(F))$. It is clear that

$$\cup_{F \in I} \mathrm{chase}_{21}(\mathrm{chase}_{12}(F)) \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(\cup_{F \in I} F)),$$

that is, $\cup_{F \in I} \mathrm{chase}_{21}(\mathrm{chase}_{12}(F)) \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I)$. Combining these inclusions, we see that $\widehat{I} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$. By Lemma A.1 we have that $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) \subseteq \widehat{I}$. Therefore, we have $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) = \widehat{I}$, as desired. $\Box$

Example where $\widehat{I} \not\subseteq \mathbf{chase_{21}(chase_{12}(}I\mathbf{))}$.

EXAMPLE A.3. We now give an example from [Fagin, Kolaitis, Popa and Tan 2007] where $\mathcal{M}_{21}$ is a inverse of $\mathcal{M}_{12}$ but where there is a ground instance I such that $\widehat{I} \not\subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$. Let $\mathbf{S_1}$ consist of the unary relation symbol P, and let $\mathbf{S_2}$ consist of the binary relation symbol Q. Let Σ_{12} consist of $P(x) \rightarrow \exists y Q(x, y)$, and let Σ_{21} consist of the constraints $Q(x, y) \rightarrow P(y)$ and $Q(x, y) \wedge \mathrm{const}(y) \rightarrow P(x)$. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$.

We now show that $\mathcal{M}_{21}$ is a inverse of $\mathcal{M}_{12}$. To do this, we need to show that if I and J are ground instances, then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $\widehat{I} \subseteq J$.

First, let I be a ground instance that consists of n facts $P(x_1)$, ..., $P(x_n)$, and let K be $\{Q(x_i, x_i) : 1 \leq i \leq n\}$. It is easy to see that $(I, K) \models \Sigma_{12}$ and $(K, \widehat{I}) \models \Sigma_{21}$. Hence, $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$, which implies that if $\widehat{I} \subseteq J$, then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$.

Next, assume that I and J are ground instances such that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$; we shall show that $\widehat{I} \subseteq J$. Since $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, there is K such that $(I, K) \models \Sigma_{12}$ and $(K, J) \models \Sigma_{21}$. Suppose I consists of n facts $P(x_1), \ldots, P(x_n)$. Since $(I, K) \models \Sigma_{12}$, we know that K contains $\{Q(x_i, y_i) \mid 1 \leq i \leq n\}$, for some choices of $y_1, \ldots, y_n$. There are two cases:

- *Case 1.* Some y_i is not a constant. Then J contains $P(y_i)$, and so is not ground. Hence, this case is not possible.
- *Case 2.* Every y_i is a constant. Then J contains $P(x_i)$, $1 \leq i \leq n$, and so $I \subseteq J$, as desired.

This concludes the proof that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. However, let $I = \{P(0)\}$; it is easy to see that $\widehat{I} \not\subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$.

Proof of Proposition 4.5. Assume by way of contradiction that δ is a demanding conjunction for A, but that no atom B of δ is demanding for A. So for every atom B of δ, there is a ground instance J_B such that $I_B \rightarrow \mathrm{chase}_{12}(J_B)$ and $I_A \not\subseteq J_B$. Note that since every member of I_B is a constant, and $I_B \rightarrow \mathrm{chase}_{12}(J_B)$, necessarily $I_B \subseteq \mathrm{chase}_{12}(J_B)$. Let I be the union of the instances J_B. So $I_B \subseteq \mathrm{chase}_{12}(I)$. Since this is true for every atom B of δ, it follows that $I_\delta \subseteq \mathrm{chase}_{12}(I)$, and so $I_\delta \rightarrow \mathrm{chase}_{12}(I)$. Since for every B we have that $I_A \not\subseteq J_B$, and I_A is a singleton set, it follows that $I_A \not\subseteq I$. So we have that $I_\delta \rightarrow \mathrm{chase}_{12}(I)$ and $I_A \not\subseteq I$. This contradicts the assumption that δ is demanding for A. Therefore, δ contains a demanding atom for A, as desired.

We have shown that each demanding conjunction for A contains a demanding atom for A. We now show that each essential conjunction for A contains an essential atom for A. Let δ be an essential conjunction for A. So δ is a demanding conjunction for A, and hence, by what we just showed, δ contains a demanding atom B for A. Since δ is relevant for A, it follows easily that B is relevant for A. Since B is both relevant and demanding for A, we see that B is essential for A, as desired. □.

Proof of Proposition 4.6. Assume that A is $P(v_1, \ldots, v_k)$, where $v_1, \ldots, v_k$ are variables, not necessarily distinct. Assume that the variable v_i does not appear in δ; we shall derive a contradiction.

Let d be a new constant, and let I be obtained from I_A by replacing every occurrence of c_{v_i} in I_A by d. Since δ is relevant for A, we know that there is a homomorphism $h : I_\delta \rightarrow \mathrm{chase}_{12}(I_A)$. So for the same homomorphism h, we have $h : I_\delta \rightarrow \mathrm{chase}_{12}(I)$, since c_{v_i} does not appear in I_δ. Since $I_\delta \rightarrow \mathrm{chase}_{12}(I)$, even though $I_A \not\subseteq I$, it follows that δ is not demanding for A, which contradicts the assumption that δ is essential for A. □

We now relate the notions of "not too strong" and "not too weak" to the notions of "demanding" and "relevant".

THEOREM A.4. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an s-t tgd mapping and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S}}_\mathbf{1}, \Sigma_{21})$ be a normal mapping. Then*

1. *Σ_{21} is not too strong if and only if every constraint in Σ_{21} is of the form $\delta \rightarrow \widehat{A}$, where δ^f is demanding for A^f for every weak renaming f consistent with the inequalities of φ.*
2. *Σ_{21} is not too weak if and only if for each source atom A, there is a relevant conjunction δ for A such that $\delta \rightarrow \widehat{A}$ is a weak renaming of a constraint in Σ_{21}.*

PROOF. (1) Assume first that there is a constraint φ in Σ_{21} of the form $\delta \rightarrow \widehat{A}$ and a weak renaming f consistent with the inequalities of φ such that δ^f is not demanding for A^f. Let δ' be δ^f, and let A' be A^f. Then $\delta' \rightarrow \widehat{A'}$ is a normal constraint that is a logical consequence of Σ_{21}. Since δ' is not demanding for A', there is an instance I such that $I_{\delta'} \rightarrow \mathrm{chase}_{12}(I)$, yet $I_{A'} \not\subseteq I$. Since $I_{\delta'} \rightarrow \mathrm{chase}_{12}(I)$, it follows that $\widehat{I_{A'}}$ is the result of chasing $\mathrm{chase}_{12}(I)$ with $\delta' \rightarrow \widehat{A'}$. So $\widehat{I_{A'}} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$ and therefore $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) \not\subseteq \widehat{I}$. By Lemma A.1, Σ_{21} is too strong.

Conversely, assume that Σ_{21} is too strong. Then, by Lemma A.1, there is a ground instance I such that $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) \not\subseteq \widehat{I}$. It follows that there must be a constraint of the form $\delta \rightarrow \widehat{A}$ in Σ_{21} such that the result of chasing $\mathrm{chase}_{12}(I)$ with $\delta \rightarrow \widehat{A}$ produces a fact not in I. By renaming constants in I if needed, this tells us that there is a weak renaming f such that $I_{(\delta^f)} \rightarrow \mathrm{chase}_{12}(I)$ and $I_{(A^f)} \not\subseteq I$. Hence, δ^f is not demanding for A^f.

(2) Assume first that Σ_{21} is not too weak. Pick a source atom A. By Lemma A.2, we know that $\widehat{I_A} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I_A))$. So there must be a normal constraint $\varphi \in \Sigma_{21}$ that fires on $\mathrm{chase}_{12}(I_A)$ to introduce $\widehat{I_A}$. Hence, there must be a weak renaming $\delta \rightarrow \widehat{A}$ of a constraint in Σ_{21} such that $I_\delta \rightarrow \mathrm{chase}_{12}(I_A)$. So δ is relevant for A.

Conversely, assume that for each source atom A, there is a relevant conjunction δ for A such that $\delta \rightarrow \widehat{A}$ is a weak renaming of a constraint $\varphi \in \Sigma_{21}$. Pick a source atom A. Then $I_\delta \rightarrow \mathrm{chase}_{12}(I_A)$ because δ is relevant for A and therefore we have $\widehat{I_A} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I_A))$ because φ fires on $\mathrm{chase}_{12}(I_A)$ to introduce $\widehat{A}$. By Lemma A.2, Σ_{21} is not too weak. □

Proof of Theorem 4.8. This follows immediately from Theorem A.4.

Proof of Theorem 4.10. The implication (2) ⇒ (3) follows from Theorem 4.8. The implications (3) ⇒ (4), and (4) ⇒ (1), are immediate. The implication (1) ⇒ (2) follows by Proposition 4.12. □

Proof of Proposition 4.12. It is clear that ω_A is relevant for A. We now show that ω_A is demanding for A, which completes the proof. Since $\mathcal{M}$ is invertible, we know that $\mathcal{M}$ satisfies the subset property. Assume that $I_{\omega_A} \rightarrow \mathrm{chase}_{12}(I)$ for some ground instance I; we must show that $I_A \subseteq I$. Now $I_{\omega_A} = \mathrm{chase}_{12}(I_A)$. So $\mathrm{chase}_{12}(I_A) \rightarrow \mathrm{chase}_{12}(I)$. By the implication (1) ⇒ (3) of Proposition 3.1, it follows that $I_A \subseteq I$, as desired. □

A.3 Proofs for Section 5

Recall that if $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ are schema mappings, then $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if for every pair I, J of *ground* instances, we have that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $\widehat{I} \subseteq J$. Therefore, for pairs (J_1, J_2) where J_2 is not a ground instance, the pair (J_1, J_2) satisfying or not satisfying Σ_{21} plays no role whatever in determining whether or not $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. Based on this intuition, let us say that Σ_{21} and Σ'_{21} are *weakly equivalent* if whenever J_1 is arbitrary and J_2 is a ground instance (contains no nulls, but only constants), then $(J_1, J_2) \models \Sigma_{21}$ if and only if $(J_1, J_2) \models \Sigma'_{21}$.[1] We may also say that $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ and $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$ are then weakly equivalent. Note that if $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ and $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$ are both normal mappings, then they are weakly equivalent if and only if they are equivalent. This is because if $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ is normal, and $(J_1, J_2) \models \Sigma_{21}$, then J_2 is a ground instance.

We capture the intuition about the irrelevance of pairs (J_1, J_2) where J_2 is not a ground instance in the following simple proposition.

[1] This notion arises also in the full version of [Fagin, Kolaitis, Popa and Tan 2007].

PROPOSITION A.5. *Let $\mathcal{M}_{12}$ be a schema mapping, and let $\mathcal{M}_{21}$ and M'_{21} be weakly equivalent schema mappings. Then $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$.*

PROOF. By symmetry, we need only show that if $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, then $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$. Let I, J be ground instances. Since $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, we know that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $\widehat{I} \subseteq J$. To show that $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$, we need only show that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $(I, J) \models \Sigma_{12} \circ \Sigma'_{21}$.

Now $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if there is J' such that $(I, J') \models \Sigma_{12}$ and $(J', J) \models \Sigma_{21}$. Since J is ground, we have that $(J', J) \models \Sigma_{21}$ if and only if $(J', J) \models \Sigma'_{21}$. Hence, $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if there is J' such that $(I, J') \models \Sigma_{12}$ and $(J', J) \models \Sigma'_{21}$, which happens if and only if $(I, J) \models \Sigma_{12} \circ \Sigma'_{21}$. Therefore, $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $(I, J) \models \Sigma_{12} \circ \Sigma'_{21}$. This was to be shown. $\square$

We now make use of Proposition A.5 to show that no schema mapping has a unique inverse.

Proof of Theorem 5.1. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an invertible schema mapping. Assume that $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ is an inverse of $\mathcal{M}_{12}$. Let J^* be some target instance, and let K^* be some source instance that contains a null value. Let $P = \{(J_1, J_2) : (J_1, J_2) \models \Sigma_{21}\}$. Let $P' = P \cup \left\{(J^*, \widehat{K^*})\right\}$ if $(J^*, \widehat{K^*})$ is not in P, and P' be the set difference $P \backslash \left\{(J^*, \widehat{K^*})\right\}$ if $(J^*, \widehat{K^*})$ is in P, Define Σ'_{21} by having $(J_1, J_2) \models \Sigma'_{21}$ if and only if $(J_1, J_2) \in P'$. Let $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$. By construction, we see that Σ_{21} and Σ'_{21} are not logically equivalent but are weakly equivalent. It follows from Proposition A.5 that $\mathcal{M}'_{21}$ is another inverse of $\mathcal{M}_{12}$. $\square$.

Because of Proposition A.5, it is natural, as far as inverse is concerned, to not distinguish between schema mappings that are weakly equivalent. We now show that, in contrast to Theorem 5.1, there is a schema mapping that has a unique inverse up to weak equivalence.

THEOREM A.6. *There is an invertible schema mapping $\mathcal{M}_{12}$ such that all inverses of $\mathcal{M}_{12}$ are weakly equivalent.*

PROOF. Let $\mathbf{S_1}$ and $\mathbf{S_2}$ be disjoint (and nonempty) schemas, and let f be a one-to-one mapping from all ground instances of $\mathbf{S_1}$ onto all instances (with or without nulls) of $\mathbf{S_2}$. There is such a mapping, since there is a countably infinite number of ground instances of $\mathbf{S_1}$, and there is a countably infinite number of instances of $\mathbf{S_2}$. Define Σ_{12} by letting $(J_1, J_2) \models \Sigma_{12}$ if and only if J_1 is a ground instance, J_2 is a target instance, and $J_2 = f(J_1)$. Define Σ_{21} by letting $(K_1, K_2) \models \Sigma_{21}$ if and only if $\widehat{I} \subseteq K_2$, where $I = f^{-1}(K_1)$. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$. We begin by showing that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. We must show that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $\widehat{I} \subseteq J$. Assume first that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$. Then there is J' such that $(I, J') \models \Sigma_{12}$ and $(J', J) \models \Sigma_{21}$. Since $(I, J') \models \Sigma_{12}$, we know that $J' = f(I)$. Since $(J', J) \models \Sigma_{21}$, it follows that $\widehat{I} \subseteq J$, as desired. Conversely, assume that $\widehat{I} \subseteq J$. Let $J' = f(I)$. Then $(I, J') \models \Sigma_{12}$ and $(J', J) \models \Sigma_{21}$, and so $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, as desired.

Now let $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$ be an arbitrary inverse of M_{12}. We must show that $\mathcal{M}'_{21}$ is weakly equivalent to $\mathcal{M}_{21}$. Assume first that $(K_1, K_2) \models \Sigma_{21}$, where K_2 is a ground instance. We must show that $(K_1, K_2) \models \Sigma'_{21}$. Since $(K_1, K_2) \models \Sigma_{21}$, we know that $\widehat{I} \subseteq K_2$, where $I = f^{-1}(K_1)$. Since $\widehat{I} \subseteq K_2$, and since $\mathcal{M}'_{21}$ is an inverse of M_{12}, there is J such that $(I, J) \models \Sigma_{12}$ and $(J, K_2) \models \Sigma'_{21}$. Since $(I, J) \models \Sigma_{12}$, we know that $J = f(I) = K_1$. Therefore, since $(J, K_2) \models \Sigma'_{21}$, it follows that $(K_1, K_2) \models \Sigma'_{21}$, as desired.

Conversely, assume that $(K_1, K_2) \models \Sigma'_{21}$, where K_2 is a ground instance. We must show that $(K_1, K_2) \models \Sigma_{21}$. Let $I = f^{-1}(K_1)$. So $(I, K_1) \models \Sigma_{12}$. Since also $(K_1, K_2) \models \Sigma'_{21}$, it follows that $(I, K_2) \models \Sigma_{12} \circ \Sigma'_{21}$. Therefore, since $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$, we know that $\widehat{I} \subseteq K_2$. Hence, $(K_1, K_2) \models \Sigma_{21}$, as desired. $\square$

The following proposition, which was proven in [Fagin, Kolaitis, Popa and Tan 2007], says that for full s-t tgd mappings, const formulas play no role in specifying an inverse.

PROPOSITION A.7. *[Fagin, Kolaitis, Popa and Tan 2007] Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be a full s-t tgd mapping. Let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$, where Σ_{21} is a set of s-t tgds with constants and inequalities. Let $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$, where Σ'_{21} is obtained from Σ_{21} by removing every const formula. Let I and J be ground instances. Then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ if and only if $(I, J) \models \Sigma_{12} \circ \Sigma'_{21}$.*

Proof of Theorem 5.4. Assume first that $\mathcal{M}_{12}$ is an invertible s-t tgd mapping with a unique normal inverse. Let A be a source atom. Assume that δ is essential for A, and δ' is demanding for A, and both have formulas $\mathrm{const}(x)$ for exactly the variables x that appear in A, Assume that we do not have $I_\delta \rightarrow I_{\delta'}$; we shall derive a contradiction. Assume without loss of generality that δ' has no inequalities as conjuncts (if necessary, remove them). Let e be as in Definition 4.9 with $e(A) = \delta$. It follows from Theorem 4.10 that $\mathcal{M}^e_{21}$ is an inverse of $\mathcal{M}_{12}$. Let σ' be $\delta' \wedge \eta_A \rightarrow \widehat{A}$, where η_A is a conjunction of the inequalities $x \neq y$ for distinct variables x, y of A. Let $\Sigma_{21} = \Sigma^e_{21} \cup \{\sigma'\}$. Let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$. It follows from Theorem 4.8 that $\mathcal{M}_{21}$ is also an inverse of $\mathcal{M}_{12}$.

We now show that $\mathcal{M}^e_{21}$ and $\mathcal{M}_{21}$ are not equivalent, which gives our desired contradiction. Let $I = I_{\delta'}$. Let J be the result of chasing I with Σ^e_{21}. Clearly $(I, J) \models \Sigma^e_{21}$. We now show that $(I, J) \not\models \sigma'$, and so $(I, J) \not\models \Sigma_{21}$. Note that because of the structure of Σ^e_{21}, it follows that J is a ground instance.

Let $\widehat{A}(\bar{c})$ be the result of chasing I with σ'. We need only show that $\widehat{A}(\bar{c})$ does not appear in J. Let σ be an arbitrary member of Σ^e_{21}. By construction of Σ^e_{21}, we know that σ is of the form $e(A) \wedge \eta_{A'} \rightarrow \widehat{A'}$, where A' is a prime source atom, and where $\eta_{A'}$ is a conjunction of the inequalities $x \neq y$ for distinct variables x, y of A'. We must show that the result of chasing I with σ does not produce $\widehat{A}(\bar{c})$. There are two cases.

Case 1: A' involves a different relation symbol than A. So certainly the result of chasing I with σ does not produce $\widehat{A}(\bar{c})$.

Case 2: A' involves the same relation symbol as A. There are two subcases.

Subcase 2a: A' equals A. Since there is no homomorphism from I_δ to $I_{\delta'}$, that is, from I_δ to I, it follows that σ does not fire on I.

Subcase 2b: A' is different from A. Then the equality pattern of the variables in A' is different from the equality pattern of the variables in A. Hence, the result of chasing I with σ again does not produce $\widehat{A}(\bar{c})$.

We now prove the converse. Assume that for every source atom A, if δ is an essential conjunction for A, and δ' is a demanding conjunction for A, both with formulas $\mathrm{const}(x)$ for exactly the variables x that appear in A, then $I_\delta \rightarrow I_{\delta'}$. Let e be as in Definition 4.9. So $\mathcal{M}^e_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma^e_{21})$ is an inverse of $\mathcal{M}_{12}$, by Theorem 4.10. Let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ be an arbitrary normal inverse of $\mathcal{M}_{12}$. We need only show that Σ_{21}^e and Σ_{21} are logically equivalent.

We first show that Σ_{21} logically implies Σ_{21}^e. Let σ be an arbitrary member of Σ_{21}^e. Then σ is of the form $e(A) \wedge \eta_A \rightarrow \widehat{A}$. Let δ be $e(A)$. By part (2) of Theorem 4.8, we know that there is an essential conjunction δ' for A such that $\delta' \rightarrow \widehat{A}$ is a weak renaming of a constraint in Σ_{21}. Since δ and δ' are both essential for A, it follows by assumption that I_δ and $I_{\delta'}$ are homomorphically equivalent. It is not hard to see that this implies that Σ_{21} logically implies σ. Since σ is an arbitrary member of Σ_{21}^e, it follows that Σ_{21} logically implies Σ_{21}^e, as desired.

We now show that Σ_{21}^e logically implies Σ_{21}. Let σ be an arbitrary member of Σ_{21}. By part (1) of Theorem 4.8, we know that σ is of the form $\delta' \rightarrow \widehat{A}$, where A is a source atom and where $(\delta')^f$ is demanding for A^f for every weak renaming f consistent with the inequalities of σ. For each weak renaming f consistent with the inequalities of σ, let τ_f be obtained from σ^f by adding to the premise of σ^f (if it is not already there) each inequality $x \neq y$ for every pair x, y of distinct variables in the conclusion of σ^f. It is fairly straightforward to see that σ is logically equivalent to the set of all such formulas τ_f. So to prove that Σ_{21}^e logically implies Σ_{21}, we need only show that Σ_{21}^e logically implies each such constraint τ_f.

Now τ_f is a normal constraint of the form $\delta'' \wedge \eta_{A'} \rightarrow \widehat{A'}$, where δ'' is demanding for A' (since as we said, $(\delta')^f$ is demanding for A^f), and $\eta_{A'}$ is the conjunction of all inequalities $x \neq y$ for distinct variables x, y of A'. By further renaming variables if needed, we can assume that A' is a prime atom. Now there is an essential conjunction δ for A' such that $\delta \wedge \eta_{A'} \rightarrow \widehat{A'}$ is a normal constraint in Σ_{21}^e. Let us denote this constraint by γ. Since δ is essential for A, and δ'' is demanding for A, it follows by assumption that $I_\delta \rightarrow I_{\delta''}$. It follows easily that γ logically implies τ_f. So Σ_{21}^e logically implies τ_f, as desired. $\square$

In the next proposition, we give a sufficient condition for a unique normal inverse.

PROPOSITION A.8. *Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ be an invertible s-t tgd mapping. Assume that for every source atom A,*

1. *$\mathrm{chase}_{12}(I_A)$ is a singleton, and*
2. *every demanding conjunction δ' for A with formulas $\mathrm{const}(x)$ precisely for the variables x of A has $\mathrm{chase}_{12}(I_A) \rightarrow I_{\delta'}$.*

Then $\mathcal{M}_{12}$ has a unique normal inverse.

PROOF. We shall make use of Theorem 5.4. Let A be arbitrary source atom. Assume that δ is an essential conjunction for A, and δ' is a demanding conjunction for A, both with formulas $\mathrm{const}(x)$ for exactly the variables x that appear in A; we must show that $I_\delta \rightarrow I_{\delta'}$. Since δ is relevant for A, we have $I_\delta \rightarrow \mathrm{chase}_{12}(I_A)$. By assumption, we have $\mathrm{chase}_{12}(I_A) \rightarrow I_{\delta'}$. So $I_\delta \rightarrow I_{\delta'}$, as desired. $\square$

Proof of Theorem 5.5. By Theorem 5.9 (whose proof does not depend on Theorem 5.5), the fact that $\mathcal{M}_{12}$ is invertible and onto implies that $\mathcal{M}_{12}$ is equivalent to a p-copy mapping. We now use Proposition A.8 to show that a p-copy mapping has a unique normal inverse.

If the p-copy mapping has the tgd

$$P(x_1, \ldots, x_k) \rightarrow Q(x_{f(1)}, \ldots, x_{f(k)}),$$

and if $y_1, \ldots, y_k$ are variables, not necessarily distinct, then let us refer to the atoms $P(y_1, \ldots, y_k)$ and $Q(y_{f(1)}, \ldots, y_{f(k)})$ as *buddies*.

Let A be a source atom. Clearly the first condition of Proposition A.8 holds. Now let δ' be a demanding conjunction for A with formulas $\mathrm{const}(x)$ precisely for the variables x of A. Let γ be the conjunction of the buddies of the relational atoms in δ'. It is easy to see that $I_{\delta'} = \mathrm{chase}_{12}(I_\gamma)$. Since δ' is demanding for A, it follows that $I_A \subseteq I_\gamma$. So $\mathrm{chase}_{12}(I_A) \subseteq \mathrm{chase}_{12}(I_\gamma)$. Clearly $\mathrm{chase}_{12}(I_\gamma) = I_{\delta'}$. Hence, $\mathrm{chase}_{12}(I_A) \subseteq I_{\delta'}$. so by Proposition A.8, it follows that $\mathcal{M}_{12}$ has a unique normal inverse. $\square$

Proof of claim in Example 5.6. We shall use Proposition A.8 to show that $\mathcal{M}_{12}$ has a unique normal inverse. Let A be a source atom. Clearly the first condition of Proposition A.8 holds. Now let δ' be a demanding conjunction for A. By symmetry of the roles of the source atoms, we can assume without loss of generality that A is the source atom $P_1(x)$. Let δ' be a demanding conjunction for A with const formula $\mathrm{const}(x)$ (and no other const formula). We now show that δ' must contain A. Assume not; we shall derive a contradiction.

Let c be the constant such that $I_A = \{P_1(c)\}$, and let d be a constant different from c. Let I consist of the facts $P_i(d)$ for $1 \leq i \leq 4$, along with the facts $P_i(c)$ for $2 \leq i \leq 4$. So $\mathrm{chase}_{12}(I)$ consists of the facts $Q_i(d)$ for $1 \leq i \leq 4$, along with the facts $Q_i(c)$ for $2 \leq i \leq 4$, along with the facts $R(c)$ and $R(d)$. Now $I_{\delta'}$ contains only one constant, namely the constant c, and possibly also null values. Let h be a function where $h(c) = c$ and $h(n) = d$ for every null n. Since δ' does not contain $Q_1(x)$, it follows that $I_{\delta'}$ does not contain $Q_1(c)$. So $I_{\delta'}$ contains some subset of $\{Q_2(c), Q_3(c), Q_4(c)\}$, possibly along with some facts $Q_i(n)$ for some nulls n and for $1 \leq i \leq 4$, possibly along with $R(c)$, and possibly some facts $R(n)$ for some nulls n. Hence, h is a homomorphism that maps $I_{\delta'}$ to $\mathrm{chase}_{12}(I)$. Since $I_{\delta'} \rightarrow \mathrm{chase}_{12}(I)$ but $I_A \not\subseteq I$, this contradicts the assumption that δ' is demanding for A. This contradiction shows that δ' must contain A. Hence, $\mathrm{chase}_{12}(I_A) \subseteq I_{\delta'}$. so by Proposition A.8, it follows that $\mathcal{M}_{12}$ has a unique normal inverse. $\square$

We now give a variant of Proposition 4.2 that holds for inverses specified by disjunctive tgds with inequalities.

PROPOSITION A.9. *Assume that $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is a full s-t tgd mapping, $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ is an inverse of $\mathcal{M}_{12}$, and Σ_{21} is a set of disjunctive tgds with inequalities. Let I be a ground instance, and let $U = \mathrm{chase}_{12}(I)$. Then $(U, \widehat{I}) \models \Sigma_{21}$, and U witnesses $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$ when $\widehat{I} \subseteq J$.*

PROOF. Since $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, we know that $(I, \widehat{I}) \models \Sigma_{12} \circ \Sigma_{21}$ and therefore there exists some K such that $(I, K) \models \Sigma_{12}$ and $(K, \widehat{I}) \models \Sigma_{21}$. Since U is a universal solution for I with respect to $\mathcal{M}_{12}$. there is a homomorphism $h : U \rightarrow K$ that is the identity on I. Pick a constraint $\varphi \in \Sigma_{21}$; by assumption, it must be of the form

$$\alpha(\bar{x}, \bar{y}) \wedge \wedge \eta(\bar{x}) \rightarrow \psi(\bar{x}),$$

where $\eta(\bar{x})$ is a conjunction of inequalities (possibly empty) among the variables in $\bar{x}$, and where $\psi(\bar{x})$ is an existentially quantified disjunction of conjunctions with free variables $\bar{x}$. Assume that U satisfies the premise of φ on $\bar{a}, \bar{b}$. Then $K \models \alpha(h(\bar{a}), h(\bar{b}))$. Now every member of U is a constant, since $U = \mathrm{chase}_{12}(I)$ and Σ_{12} is full. Therefore $h(\bar{a}) = \bar{a}$ and $h(\bar{b}) = \bar{b}$. so $K \models \alpha(\bar{a}, \bar{b})$. Since $(K, \widehat{I}) \models \Sigma_{21}$, we must have $\widehat{I} \models \psi(\bar{a})$. This shows that $(U, \widehat{I}) \models \varphi$. Since φ is an arbitrary member of Σ_{21}, it follows that $(U, \widehat{I}) \models \Sigma_{21}$, as desired. Since $\widehat{I} \subseteq J$, it follows easily that $(U, J) \models \Sigma_{21}$. Since $(I, U) \models \Sigma_{12}$ and $(U, J) \models \Sigma_{21}$, we have that U witnesses $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, as desired. $\square$

Proof of Theorem 5.9.

We begin by showing that (2) $\Rightarrow$ (1). Assume that (2) holds. Since $\mathcal{M}_{12}$ is invertible, Theorem 4.13 tells us that the canonical candidate inverse is indeed an inverse of $\mathcal{M}_{12}$, so $\mathcal{M}_{12}$ has a normal inverse. By Proposition A.7, the const formulas are irrelevant, and so $\mathcal{M}_{12}$ has a inverse specified by tgds with inequalities, and hence by disjunctive tgds with inequalities. Now assume that $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ and $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$ are both inverses of $\mathcal{M}_{12}$, where Σ_{21} and Σ'_{21} are disjunctive tgds with inequalities. We must show that Σ_{21} and Σ'_{21} are logically equivalent. We now show that Σ_{21} logically implies Σ'_{21}. By symmetry, we have that Σ'_{21} logically implies Σ_{21}. Assume that $(J, K) \models \Sigma_{21}$. By replacing each null in (J, K) by a new constant, we obtain (J', K') where every entry of every tuple is a constant, such that (J', K') is isomorphic to (J, K) (but where the isomorphism may map constants into either constants or nulls, and may map nulls into either constants or nulls). Since Σ_{21} has no const formulas, it follows easily that $(J', K') \models \Sigma_{21}$. Since $\mathcal{M}_{12}$ is onto, there is a ground instance I such that $J' = \mathrm{chase}_{12}(I)$. So $(I, J') \models \Sigma_{12}$. Since also $(J', K') \models \Sigma_{21}$, we have that $(I, K') \models \Sigma_{12} \circ \Sigma_{21}$. Therefore, since $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, we have that $\widehat{I} \subseteq K'$. Hence, since $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$, we have that $(I, K') \models \Sigma_{12} \circ \Sigma'_{21}$. So by Proposition A.9, we have that $(J', K') \models \Sigma'_{21}$. Since Σ'_{21} has no const formulas, we have as before $(J, K) \models \Sigma'_{21}$. This was to be shown.

We now show that (1) $\Rightarrow$ (3). Assume that (1) holds. We must show that $\mathcal{M}_{12}$ is equivalent to a p-copy mapping. By Theorem 4.10 and Proposition 4.5, we know that every source atom has an essential target atom. Let e be a function that maps every source atom A onto a target atom $e(A)$ that is essential for A. By Theorem 4.10, we know that $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma^e_{21})$ is an inverse of A. Let $\mathcal{M}^E_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma^E_{21})$ be the result of removing all $\mathrm{const}(x)$ conjuncts in $(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma^e_{21})$. Note that every member of Σ^E_{21} is of the form $B \wedge \eta_A \rightarrow \widehat{A}$, where η_A consists of all inequalities of the form $x \neq y$ where x and y are distinct variables of A. Note that by Corollary 4.7, the variables in A and B are the same, so η_A has inequalities among all distinct variables of B also. By Proposition A.7 we know that $\mathcal{M}^E_{21}$ is an inverse of $\mathcal{M}_{12}$. Since (3) holds, $\mathcal{M}^E_{21}$ is the unique inverse that is specified by disjunctive tgds with inequalities. We now prove the following claim:

Claim 1: $\mathrm{chase}_{12}(I_A)$ is a singleton for each source atom A.

Assume that A is a source atom where $\mathrm{chase}_{12}(I_A)$ is not a singleton; we shall derive a contradiction. Since $\mathcal{M}_{12}$ is invertible, we know that $\mathrm{chase}_{12}(I_A)$ is nonempty (otherwise, we would have $\mathrm{chase}_{12}(I_A) = \mathrm{chase}_{12}(\emptyset)$, and this gives a violation of the unique solutions property). Denote $e(A)$ by B. So Σ^E_{21} contains the formula $B \wedge \eta_A \rightarrow \widehat{A}$. Let ν_A be as in Definition 4.11. Form Σ_{21} from Σ^E_{21} by replacing $B \wedge \eta_A \rightarrow \widehat{A}$ by $\nu_A \wedge \eta_A \rightarrow \widehat{A}$, and let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$. It follows from Theorem 4.8 that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. We now show that Σ_{21} is not logically equivalent to Σ^E_{21}.

Assume that B is the atom $Q(x_1, \ldots, x_t)$, and that F is a fact $Q(a_1, \ldots, a_t)$. Let us say that F is an *exact match* for B if $a_i = a_j$ if and only if x_i and x_j are the same variable, for all i, j. Similarly, we define what it means for one atom to be an exact match for another atom. Let J consist of a single fact F that is an exact match for B. We now show that $(J, \emptyset) \not\models \Sigma^E_{21}$, but $(J, \emptyset) \models \Sigma_{21}$. So Σ_{21} is not logically equivalent to Σ^E_{21}, as desired. The fact that $(J, \emptyset) \not\models \Sigma^E_{21}$ follows from the fact that Σ^E_{21} contains the formula $B \wedge \eta_A \rightarrow \widehat{A}$, and J contains a fact that is an exact match for B. It remains to show that $(J, \emptyset) \models \Sigma_{21}$. Let σ be a member of Σ_{21} except for $\nu_A \wedge \eta_A \rightarrow \widehat{A}$. Since J consists of a single fact that is an exact match for B. it follows that σ does not fire on J, because otherwise F would be an exact match for the atom B' in the premise of σ, and so B' and B would be an exact match, which is not possible since they are essential for atoms that are not an exact match for each other. So $(J, \emptyset) \models \sigma$. We now show that $(J, \emptyset) \models \nu_A \wedge \eta_A \rightarrow \widehat{A}$. To show this, we must show that $\nu_A \wedge \eta_A \rightarrow \widehat{A}$ does not fire on J. If it were to fire on J, then there would be a mapping h on the variables in ν_A that maps each atom in ν_A onto F and (because of η_A) is one-to-one on the variables in ν_A. Let B' be a member of ν_A other than B. Since B and B' map onto the same fact F, it follows that B' is, like B, a Q-stom. Assume that B' is $Q(x_{i_1}, \ldots, x_{i_t})$, where each x_{i_r} is in $\{x_1, \ldots, x_t\}$. Assume that F is the fact $Q(a_1, \ldots, a_t)$. Now $h(x_{i_r}) = a_r = h(x_r)$ for each r, since both B and B' map onto F. Since h is one-to-one on variables, it follows that x_{i_r} and x_r are the same variable for each r. So B' and B are the same atom, a contradiction. This contradiction shows that $\nu_A \wedge \eta_A \rightarrow \widehat{A}$ does not fire on J, as desired. This concludes the proof that Σ_{21} is not logically equivalent to Σ^E_{21}.

Since $\mathcal{M}_{21}$ and $\mathcal{M}^E_{21}$ are both inverses of $\mathcal{M}_{12}$, specified by tgds with inequalities, even though Σ_{21} and Σ^E_{21} are not logically equivalent, this contradicts our assumption that (3) holds. This proves Claim 1.

Define Σ'_{12} to consist of all s-t tgds of the form $A \rightarrow \nu_A$, where A is a prime atom with all variables distinct. Note by Claim 1 that ν_A is a singleton atom. It is clear that Σ_{12} logically implies Σ'_{12}. Later, we shall show that Σ_{12} is logically equivalent to Σ'_{12}. First, we prove another claim.

Claim 2: Assume that A and B are atoms, and that B is essential for A with respect to Σ_{12}. Then Σ'_{12} logically implies the s-t tgd $A \rightarrow B$.

Assume that Claim 2 were false; we shall derive a contradiction. Assume that A is a P-atom. Define ν'_A like ν_A, except that the chase is with Σ'_{12} instead of σ_{12}. Let B' be ν'_A. Note that B' is a singleton atom, because it arises only by firing the s-t tgd in Σ'_{12} whose premise is the P-atom with all variables distinct. Since Σ'_{12} does not logically imply the s-t tgd $A \rightarrow B$, we know that B is different from B'. Since B' is derived as the result of a chase with Σ'_{12}, and since Σ_{12} logically implies Σ'_{12}, it follows that B' is in ν_A. So ν_A contains at least the two distinct atoms B and B'. This contradicts Claim 1, which is our desired contradiction.

Claim 3: Σ_{12} is logically equivalent to Σ'_{12}.

We already noted that Σ_{12} logically implies Σ'_{12}, so Claim 3 is proven if we show that Σ'_{12} logically implies Σ_{12}. Assume not; we shall derive a contradiction. We can assume without loss of generality that every member of Σ_{12} has a singleton conclusion. Let $\alpha \rightarrow B$ be a member of Σ_{12} that is not a logical consequence of Σ'_{12}. If α were to contain an atom A such that B is essential for A with respect to Σ_{12}, then from Claim 2 it would follow that Σ'_{12} logically implies the s-t tgd $A \rightarrow B$, and so Σ'_{12} would logically imply $\alpha \rightarrow B$, a contradiction. Hence, α does not contain an atom A such that B is essential for A with respect to Σ_{12}.

Assume that B is the atom $Q(x_1, \ldots, x_t)$ where $x_1, \ldots, x_t$ are variables. Let τ be an arbitrary member of Σ_{12} of the form $\delta \rightarrow Q(z_1, \ldots, z_t)$, where $z_1, \ldots, z_t$ are variables, not necessarily distinct, and where x_i and x_j are the same variable whenever z_i and z_j are the same variable. Let h_τ be a function with domain the variables in τ such that $h_\tau(z_i) = x_i$ for each i (this is well-defined, since x_i and x_j are the same variable whenever z_i and z_j are the same variable), and where h_τ maps each variable in the premise of τ that is not in the conclusion of τ onto a new variable. Let τ' be the image of τ under h_τ. Thus, τ' is a weak renaming of τ. It is straightforward to verify that if I is a source instance and the chase of I with τ produces a fact F that is an exact match for B, then the chase of I with τ' produces F.

By construction, the conclusion of τ' is B. Assume that the premise of τ' is δ'. Let ψ_τ be the formula $\exists \mathbf{y}\widehat{\delta'}$, where $\mathbf{y}$ consists of the variables in τ' that are not in B. Let Z consist of all such formulas ψ_τ. In particular, Z contains $\exists \mathbf{y}\widehat{\alpha}$, where $\mathbf{y}$ consists of all variables in α that are not in B. Now Z is finite, since its size is at most the number of members of Σ_{12}. Let γ be the formula $B \wedge \eta \rightarrow z$, where η is the conjunction of all inequalities of the form $x_i \neq x_j$ where x_i and x_j are distinct variables in B, and where z is the disjunction of the members of Z. Let $\Sigma_{21} = \Sigma_{21}^E \cup \{\gamma\}$, and let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$.

We now show that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, and that Σ_{21} is not logically equivalent to Σ_{21}^E. To show that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, we must show that for all ground instances I and J:

$$(I, J) \models \Sigma_{12} \circ \Sigma_{21} \text{ if and only if } \widehat{I} \subseteq J. \tag{1}$$

Since $\mathcal{M}_{21}^E$ is an inverse of $\mathcal{M}_{12}$, we know that for all ground instances I and J:

$$(I, J) \models \Sigma_{12} \circ \Sigma_{21}^E \text{ if and only if } \widehat{I} \subseteq J. \tag{2}$$

Now Σ_{21} logically implies Σ_{21}^E, since Σ_{21} is a superset of Σ_{21}^E. It follows easily that $\Sigma_{12} \circ \Sigma_{21}$ logically implies $\Sigma_{12} \circ \Sigma_{21}^E$. So if $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$, then $(I, J) \models \Sigma_{12} \circ \Sigma_{21}^E$, which, by (2), implies that $\widehat{I} \subseteq J$. Assume now that $\widehat{I} \subseteq J$; we must show that $(I, J) \models \Sigma_{12} \circ \Sigma_{21}$. Let $J^* = \text{chase}_{12}(I)$. Since $(I, J^*) \models \Sigma_{12}$, we need only show that $(J^*, J) \models \Sigma_{21}$. Now $(J^*, J) \models \Sigma_{21}^E$, by Proposition A.9. Therefore, since $\Sigma_{21} = \Sigma_{21}^E \cup \{\gamma\}$, we need only show that $(J^*, J) \models \gamma$. Assume that γ fires on J^*. Then there is a one-to-one mapping h (one-to-one because of η) from the variables of B to constants, that maps B onto a fact F of J^*. So F is an exact match for B. Since F is in J^*, there is a member τ of Σ_{12} that generates F in the chase of I with Σ_{12}. Let τ' and δ' be as before. Since F is an exact match for B, it follows from an earlier comment that the chase of I with τ' generates F. It follows fairly easily that $\exists \mathbf{y}\delta'$ is satisfied in I under h, so $\exists \mathbf{y}\widehat{\delta'}$ is satisfied in $\widehat{I}$ under h. Since $\widehat{I} \subseteq J$, it follows that $\exists \mathbf{y}\widehat{\delta'}$ is satisfied in J under h. But $\exists \mathbf{y}\widehat{\delta'}$ is a disjunct in the conclusion of γ. Therefore, $(J^*, J) \models \gamma$, as desired.

We now show that Σ_{21} is not logically equivalent to Σ_{21}^E. It is clear that $(I_B, \emptyset) \not\models \gamma$, and so $(I_B, \emptyset) \not\models \Sigma_{21}$. We now show that $(I_B, \emptyset) \models \Sigma_{12}^E$. Since, as we showed, there is no atom A such that B is essential for A with respect to Σ_{12}, no member of Σ_{21}^E has B in its premise. So for each member $B' \wedge \eta' \rightarrow A'$ of Σ_{21}^E, there is no mapping h that maps B' onto B and satisfies η'. It follows that $(I_B, \emptyset) \models \Sigma_{12}^E$, as desired.

We have shown that $\mathcal{M}_{12}$ has two distinct, inequivalent inverses given by disjunctive tgds with inequalities, namely $\mathcal{M}_{21}^E$ and $\mathcal{M}_{21}$. This contradiction shows that Claim 3 holds.

We now state and prove our final claim.

Claim 4: For every target relation symbol Q, there is exactly one member of Σ'_{12} whose conclusion is a Q-atom. Every variable in this conclusion is distinct.

To prove this claim, we begin by showing that there must be at least one member of Σ_{12} whose conclusion is a Q-atom and where every variable in this conclusion is distinct. Assume not; we shall derive a contradiction. Let γ be the formula $Q(x_1, \ldots, x_t) \wedge \eta \rightarrow \beta$, where the variables $x_1, \ldots, x_t$ are distinct, where η is a conjunction of the inequalities $x_i \neq x_j$ whenever $i \neq j$, and where β is an arbitrary disjunction of source atoms whose variables altogether are exactly $x_1, \ldots, x_t$.[2] Let $\Sigma_{21} = \Sigma_{21}^E \cup \{\gamma\}$, and let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$.

We now show that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, and that Σ_{21} is not logically equivalent to Σ_{21}^E. To show that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$, we must show that (1) holds whenever I and J are ground instances. Let $J^* = \text{chase}_{12}(I)$. As in the proof of Claim 2, since $\Sigma_{21} = \Sigma_{21}^E \cup \{\gamma\}$, we need only show that $(J^*, J) \models \gamma$. Since by assumption there is no member of Σ_{12} whose conclusion is a Q-atom and where every variable in this conclusion is distinct, it follows easily that J^* does not contain a Q-fact $Q(c_1, \ldots, c_t)$, where $c_1, \ldots, c_t$ are distinct constants. So γ does not fire on J^*, and so $(J^*, J) \models \gamma$, as desired.

We now show that Σ_{21} is not logically equivalent to Σ_{21}^E. Let J consist of the singleton fact $Q(c_1, \ldots, c_t)$, where $c_1, \ldots, c_t$ are distinct constants. Then $(J, \emptyset) \not\models \gamma$, and so $(J, \emptyset) \not\models \Sigma_{21}$. However, $(J, \emptyset) \models \Sigma_{21}^E$, since every member of Σ_{21}^E has a premise of the form $B \wedge \eta'$, where B is (up to renaming of variables) a conclusion of Σ_{12}, and by assumption, there is no member of Σ_{12} whose conclusion is a Q-atom with all distinct variables.

We have shown that $\mathcal{M}_{12}$ has two distinct, inequivalent inverses given by disjunctive tgds with inequalities, namely $\mathcal{M}_{21}^E$ and $\mathcal{M}_{21}$. This contradiction shows that there must be at least one member σ of Σ'_{12} whose conclusion is a Q-atom and where every variable in this conclusion is distinct.

We now show that there can be no other member σ' of Σ'_{12} whose conclusion is a Q-atom. Assume not; we shall derive a contradiction. Assume that the premise of σ is a P-atom and the premise of σ' is a P'-atom. Since σ and σ' are different, we know that P and P' are different, by construction of Σ'_{12}. Since every variable in the conclusion of σ is distinct, there is a mapping h that maps the variables in σ to the variables in σ' that maps the conclusion of σ onto the conclusion of σ'. Let A be the P-atom that is the result of applying h to the premise of σ. So the chase of I_A with Σ'_{12} is $I_{B'}$, where B' is the conclusion of σ'. This contradicts the fact that conclusion of σ' is essential for the premise of σ'. This contradiction shows that the only member of Σ_{12} whose conclusion is a Q-atom is σ, where every variable in the conclusion is distinct. This completes the proof of Claim 4.

Now the variables in the source and target of each member of Σ'_{12} are the same by Corollary 4.7, since the target is essential for the source with respect to Σ_{12}. By construction, for every source relation symbol P, there is exactly one member of Σ'_{12} whose premise is a P-atom, and every variable is distinct in this P-atom. By Claim 4, for every target relation symbol Q, there is exactly one member of Σ'_{12} whose conclusion is a Q-atom, and every variable is distinct in this Q-atom. It follows that $\mathcal{M}_{12}$ is a p-copy mapping. Also, by Claim 3, Σ_{12} is logically equivalent to Σ'_{12}. So (3) holds, as desired. This completes the proof that (1) $\Rightarrow$ (3).

We conclude the proof by showing that (3) $\Rightarrow$ (2). From (3), we know that there is a schema mapping $\mathcal{M}'_{12}$ that is equivalent to $\mathcal{M}_{12}$ and that is a p-copy mapping. Clearly, $\mathcal{M}'_{12}$ is invertible and onto. It follows easily that $\mathcal{M}_{12}$ is invertible and onto, as desired. $\square$

LEMMA A.10. *Assume that δ is relevant for A and demanding for A'. Then A and A' are the same atom.*

PROOF. Since δ is relevant for A, we know that $I_\delta \rightarrow \text{chase}_{12}(I_A)$. Therefore, since δ is demanding for A', we have $I_{A'} \subseteq I_A$. Since I_A and $I_{A'}$ are both singletons, we have $I_{A'} = I_A$, so A and A' are the same atom, as desired. $\square$

Proof of Theorem 5.10. Assume that $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$ is a full s-t tgd mapping with a unique normal inverse $\mathcal{M}_{21} =$

[2] A disjunction is required if no source atom has arity at least t.

$(\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$, and $\mathcal{M}_{21}$ is not equivalent to the normalized version of a near p-copy mapping. Assume without loss of generality that Σ_{21} is of minimal length (and also of minimal size in terms of the number of characters) among the various logically equivalent sets of normal constraints logically equivalent to Σ_{21}. Since $\mathcal{M}_{21}$ has q unique normal inverse, it follows from Theorem 8.2 (where $m = 1$) that $\mathcal{M}_{21}$ has length at most k, where k is the number of source relation symbols.

Let P be an arbitrary source relation symbol, and let A_P be a P-atom with all variables distinct. If $\widehat{P}$ is in the conclusion of no member of Σ_{21}, then $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I_{A_P}))$ contains no $\widehat{P}$-fact, so Σ_{21} is too weak. Therefore, $\widehat{P}$ is in the conclusion of some member of Σ_{21}. Since also Σ_{21} has at most k constraints, it follows that $\widehat{P}$ is in the conclusion of exactly one member of Σ_{21}. Let σ_P be the member of Σ_{21} whose conclusion is a $\widehat{P}$-atom. Every variable in the conclusion of σ_P is distinct, or else $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I_{A_P}))$ does not contain $\widehat{I_{A_P}}$, so Σ_{21} is too weak. Let B_P be a P-atom with all variables the same. Now σ_P has no inequalities, or else $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I_{B_P}))$ does not contain $\widehat{I_{B_P}}$, so Σ_{21} is too weak.

The proof of Theorem 8.2 shows that A_P is good, that is, that $\mathrm{chase}_{12}(I_{A_P})$ is a singleton. This singleton is the only relevant atom (with respect to Σ_{12}) for A_P, so it follows fairly easily from part (2) of Theorem 4.8 (and the assumption that Σ_{21} is of minimal size) that the premise of σ_P contains only a single relational atom. This relevant atom is also essential for P, as noted in Theorem 4.8. So the premise of σ_P has a single relational atom, that is essential for the conclusion of σ_P. Note that this is also true about each weak renaming of σ_P (that is, the atom B' in the premise of the weak renaming is essential for the conclusion A' of the weak renaming. This is because A' must have an essential atom, and the only candidate is B'.

Let Q be an arbitrary relation symbol in $\mathbf{S_2}$, and let B_Q be a Q-atom with all variables the same. We now show that at most one member of Σ_{21} can have Q appear in its premise. Assume that σ_P and $\sigma_{P'}$ both have Q appear in its premise; we must show that P and P' are the same. Then B_Q is essential for both B_P and $B_{P'}$ (this follows from our earlier comment about weak renamings of σ_P). Hence, by Lemma A.10, it follows that B_P and $B_{P'}$ are the same atom, so P and P' are the same, as desired. By Corollary 4.7, the variables in the premise and conclusion of σ_P are the same.

It follows from what we have shown that $\mathcal{M}_{21}$ is equivalent to the normalized version of a near p-copy mapping. This was to be shown. $\square$.

A.4 Proofs for Section 6

LEMMA A.11. *Let $\mathcal{M}_{12}$ be a full s-t tgd mapping, and $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ a normal inverse for $\mathcal{M}_{12}$. Let A be a source atom and B a target atom where $\widehat{I_A} \subseteq \mathrm{chase}_{21}(I_B)$. Then B is demanding for A with respect to Σ_{12}.*

PROOF. Assume that $I_B \subseteq \mathrm{chase}_{12}(I)$; we must show that $I_A \subseteq I$. Let $U = \mathrm{chase}_{12}(I)$. We know from Proposition 4.2 that $(U, \widehat{I}) \models \Sigma_{21}$. Since Σ_{21} is full, this implies further that $\mathrm{chase}_{21}(U) \subseteq \widehat{I}$. Since $I_B \subseteq \mathrm{chase}_{12}(I)$ and $\widehat{I_A} \subseteq \mathrm{chase}_{21}(I_B)$, it follows that $\widehat{I_A} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) = \mathrm{chase}_{21}(U)$. Since also $\mathrm{chase}_{21}(U) \subseteq \widehat{I}$, we have that $\widehat{I_A} \subseteq \widehat{I}$, and so $I_A \subseteq I$, as desired. $\square$

Proof of Theorem 6.1. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$. Let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$ be an invertible normal inverse of $\mathcal{M}_{12}$. Whenever we speak of relevant, demanding, or essential in this proof, we mean with respect to Σ_{12}. We shall reserve A and A' for source atoms (with relation symbols in $\mathbf{S_1}$), and B and B' for target atoms (with relation symbols in $\mathbf{S_2}$).

Claim 1: If B is a relevant atom for a source atom A, then $\mathrm{chase}_{21}(I_B) = \widehat{I_A}$.

We now prove Claim 1. Assume that B is a relevant atom for A. Now $\mathrm{chase}_{21}(I_B)$ is nonempty, since otherwise $\mathrm{chase}_{21}(I_B) = \mathrm{chase}_{21}(\emptyset)$, and so $\mathcal{M}_{21}$ would violate the unique-solutions property and so be not invertible. Since $\mathcal{M}_{21}$ is full, we know that that $\mathrm{chase}_{21}(I_B)$ has no nulls, and so every fact in $\mathrm{chase}_{21}(I_B)$ is of the form $\widehat{I_{A'}}$ for some atom A'. The claim is proven if we show that whenever $\widehat{I_{A'}} \subseteq \mathrm{chase}_{21}(I_B)$, then A' is the same atom as A. So assume that $\widehat{I_{A'}} \subseteq \mathrm{chase}_{21}(I_B)$. By Lemma A.11, where the role of A is played by A', we know that B is demanding for A'. Since B is also relevant for A, it follows from Lemma A.10 that A' is the same atom as A, as desired.

Claim 2: Each source atom A has exactly one relevant atom B, and B is essential for A.

We now prove Claim 2. Since $\mathcal{M}_{12}$ is invertible, it follows from Theorem 4.10 and Proposition 4.5, that A has some essential atom B. So B is relevant for A. Assume that A has another relevant atom B'; we shall derive a contradiction. By Claim 1, we have that $\mathrm{chase}_{21}(I_B)$ and $\mathrm{chase}_{21}(I_{B'})$ both equal $\widehat{I_A}$, and so are equal. This violates the unique-solutions property for $\mathcal{M}_{21}$, and so $\mathcal{M}_{21}$ is not invertible, which gives our desired contradiction.

Let us denote the unique relevant atom for A by B_A. For the next claim, recall that if φ is a formula, and f is a weak renaming, then φ^f is the result of replacing every variable x in φ by $f(x)$.

Claim 3: Let f be a weak renaming. Then $(B_A)^f = B_{A^f}$.

We now prove Claim 3. Assume that $A^f = A'$. Since B_A is relevant for A, it is clear that the result $(B_A)^f$ of weakly renaming B_A using f is relevant for A'. That is, $(B_A)^f$ is relevant for A'. By definition, the unique relevant atom for A' is $B_{A'}$. Therefore, $(B_A)^f = B_{A'} = B_{A^f}$, as desired.

Claim 4: Let B be a target atom. Then B is relevant for some source atom.

We now prove Claim 4. We prove it first when every variable in B is distinct. Since $\mathcal{M}_{21}$ is invertible, we know that $\mathrm{chase}_{21}(I_B) \neq \emptyset$, by the unique solutions property. So there is some member $\delta \to \widehat{A}$ of Σ_{21} that fires on I_B. So $\mathrm{chase}_{21}(I_\delta)$ includes $\widehat{I_A}$. By Claim 1, we have that $\mathrm{chase}_{21}(I_{B_A}) = \widehat{I_A}$. So $\mathrm{chase}_{21}(I_{B_A}) \subseteq \mathrm{chase}_{21}(I_\delta)$. Since $\mathcal{M}_{21}$ is invertible, it satisfies (the homomorphic version of) the subset property (although the subset property and its homomorphic version in Proposition 3.1 are shown to be equivalent to invertibility for s-t tgd mappings, this holds also for normal mappings, by the same proof). So $I_{B_A} \subseteq I_\delta$. Therefore, δ has B_A as a conjunct. Since the constraint $\delta \to \widehat{A}$ of Σ_{21} fires on I_B, there is a homomorphism from B_A to B. Since every variable in B is distinct, it follows that B_A and B are the same up to a renaming of variables. Therefore, since B_A is relevant for A, we know that B is relevant for some atom obtained by renaming the variables of A. This completes the proof of Claim 3 when all of the variables in B are distinct.

Let B' be a target atom where the variables need not be distinct. Let B be an atom where all of the variables are distinct and where B' is obtained from B by a weak renaming f, that is, $B' = (B_A)^f$. Since all of the variable in B are distinct, it follows from what we have shown that B is relevant for some source atom A, and so B is simply B_A. Hence, by Claim 3, we know that B' is B_{A^f}. So B' is relevant for A^f.

Claim 5: Let A be a source atom with all variables distinct. Then every variable in B_A is distinct.

We now prove Claim 5. Since B_A is essential for A, it follows from Proposition 4.6 that B_A has exactly the same variables as A. We now show that every variable in B_A is distinct. Assume not; we shall derive a contradiction. Let B' be an atom with the same relation symbol as B_A but with every variable distinct. So B' has strictly more variables than A. Since also every variable in A is distinct, and every variable in B' is distinct, it follows that the arity of B' is strictly bigger than the arity of A. By Claim 4, B' is relevant for some source atom A'. Since by Claim 2 we know that A' has a unique relevant atom, and this atom is essential for A', it follows that B' is essential for A'. So by Proposition 4.6, we know that B' and A' have the same variables. Therefore, since every variable in B' is distinct, the arity of A' is at least the arity of B', which as we noted is strictly bigger than the arity of A. So the arity of A' is strictly bigger than the arity of A. Since B_A is obtained from B' by a weak renaming, and B' is relevant for A', it follows that B_A is relevant for an atom A'' obtained from A' by a weak renaming. Since B_A is demanding for A, it follows from Lemma A.10 that A'' and A are the same atom. But this is impossible, since A'' has the same arity as A', and the arity of A' is strictly bigger than the arity of A. This is our desired contradiction. This completes the proof of Claim 5.

Let Σ'_{12} consist of all of the constraints $A \to B_A$, where A is a prime atom with all variables distinct. Let $\mathcal{M}'_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma'_{12})$. We now show that $\mathcal{M}'_{12}$ is a p-copy mapping and is equivalent to $\mathcal{M}_{12}$.

Let $A \to B_A$ be a member of Σ'_{12}. By construction, every variable in A is distinct. As noted earlier, B_A has exactly the same variables as A, and by Claim 5, every variable in B_A is distinct. By construction, every source relation symbol appears in exactly one premise of Σ'_{12}. To complete the proof that $\mathcal{M}'_{12}$ is a p-copy mapping, all that is left to show is that every target relation symbol appears in exactly one conclusion of Σ_{21}.

Let Q be an arbitrary target relation symbol, and let B' be a Q-atom with every variable distinct. By Claim 4, we have that B' is relevant for some source atom A', and so B' equals $B_{A'}$. Assume that A' is a P-atom. Let A be the prime P-atom with all variables distinct. Let f be a weak renaming where A' is A^f. By Claim 3, we know that $B_{A'}$, that is, B', is $(B_A)^f$. Hence, since B' is a Q-atom, so is B_A. So Q appears in some conclusion of Σ'_{21}.

We now show that Q cannot be in more than one conclusion in Σ'_{12}. Say Q were in the conclusion of the member of Σ'_{12} whose premise has relation symbol P and also in the conclusion of the member of Σ'_{12} whose premise has relation symbol P'. Let F be the fact $P(0,\ldots,0)$, where every variable is set to 0. Similarly, let F' be the fact $P'(0,\ldots,0)$, where every variable is set to 0. Then the result of chasing F with Σ_{12} is $Q(0,\ldots,0)$, and identically the result of chasing F' with Σ_{12} is $Q(0,\ldots,0)$. This is a contradiction of the unique-solutions property. This concludes the proof that $\mathcal{M}'_{12}$ is a p-copy mapping.

We close by showing that Σ_{12} and Σ'_{12} are logically equivalent. Clearly Σ_{12} logically implies Σ'_{12}. We now show that Σ'_{12} logically implies Σ_{12}. We first show that each of the constraints $A' \to B_{A'}$ is a logical consequence of Σ'_{12}. Let A be an atom with the same relation symbol as A' and with all variables distinct. So there is a renaming f where A' is A^f. By Claim 3, we know that $(B_A)^f = B_{A'}$. So $A' \to B_{A'}$ is $(A \to B_A)^f$. Therefore $A' \to B_{A'}$ is a logical consequence of $A \to B_A$, and so of Σ'_{12}, as desired.

Assume now that $\varphi \to B$ is another member of Σ_{12}. By Claim 4, we know that B is B_A for some source atom A. Since $\mathrm{chase}_{12}(I_A)$ is I_{B_A}, and $\mathrm{chase}_{12}(I_\varphi)$ includes I_{B_A}, it follows that $\mathrm{chase}_{12}(I_A) \subseteq \mathrm{chase}_{12}(I_\varphi)$. So by the subset property, $I_A \subseteq I_\varphi$. Therefore, A is in φ. So $\varphi \to B$ is a logical consequence of $A \to B_A$, and so is a logical consequence of Σ'_{12}. $\square$

A.5 Proofs for Section 7

Proof of Theorem 7.1. The family is parameterized by the positive integer k. Let $\mathbf{S_1} = \{P_0,\ldots,P_k\}$, and let $\mathbf{S_2} = \{P'_0,\ldots,P'_k, Q_0,\ldots,Q_{k-1}\}$. Assume that all of the relation symbols in $\mathbf{S_1}$ and $\mathbf{S_2}$ are $4k$-ary.

Let $x_1,\ldots,x_{4k}$ be distinct variables. Let S_1 consist of the s-t tgds $P_i(x_1,x_2,\ldots,x_{4k}) \to P'_i(x_1,x_2,\ldots,x_{4k})$, for $0 \le i \le k$. Define $\bar{x}^i$, for $0 \le i \le k-1$, by letting $x^i_{4i+2} = x_{4i+1}$, $x^i_{4i+4} = x_{4i+3}$, and $x^i_j = x_j$ if $j \notin \{4i+2, 4i+4\}$. For example, $\bar{x}^0$ is $(x_1,x_1,x_3,x_3,x_5,x_6,\ldots,x_{4k-1},x_{4k})$, and $\bar{x}^1$ is $(x_1,x_2,x_3,x_4,x_5,x_5,x_7,x_7,x_9,x_{10},\ldots,x_{4k-1},x_{4k})$. Let S_2 consist of the s-t tgds $P_{i+1}(\bar{x}^i) \to P'_0(\bar{x}^i)$, for $0 \le i \le k-1$. Let S_3 consist of the s-t tgds $P_0(\bar{x}^i) \to Q_i(\bar{x}^i)$, for $0 \le i \le k-1$. Let $\Sigma_{12} = S_1 \cup S_2 \cup S_3$, and let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$.

We begin by showing that $\mathcal{M}_{12}$ is invertible. Let T_1 consist of the s-t tgds $P'_j(x_1,x_2,\ldots,x_{4k}) \to \widehat{P_j}(x_1,x_2,\ldots,x_{4k})$, for $1 \le j \le k$ (note that we do not include the case $j = 0$). Let T_2 consist of the formula

$$\begin{array}{lll} P'_0(x_1,x_2,\ldots,x_{4k}) & \wedge & ((x_1 \ne x_2) \vee (x_3 \ne x_4)) \\ & \wedge & ((x_5 \ne x_6) \vee (x_7 \ne x_8)) \\ & \wedge & \ldots \\ & \wedge & ((x_{4k-3} \ne x_{4k-2}) \vee (x_{4k-1} \ne x_{4k})) \\ & \to & \widehat{P_0}(x_1,x_2,\ldots,x_{4k}). \end{array}$$

Let T_3 consist of the s-t tgds $Q_i(\bar{x}^i) \to \widehat{P_0}(\bar{x}^i)$, for $0 \le i \le k-1$. Let $\Sigma_{21} = T_1 \cup T_2 \cup T_3$, and let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$.

Note that chase_{21} is well-defined, even in the presence of the formula in T_2.

We now show that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$. It is sufficient to show that $\widehat{I} = \mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$ for each ground instance I (this is because the analogue of Theorem 4.4 holds, by the same proof). We first show that $\widehat{I} \subseteq \mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$. If $\widehat{P}_i(\bar{a})$ is a fact of $\widehat{I}$ (and so $P_i(\bar{a})$ is a fact of I), and if $1 \le i \le k$, then we see from the tgds in S_1 and T_1 that $\widehat{P}_i(\bar{a})$ is in $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$. So assume that $\widehat{P_0}(\bar{a})$ is a fact of $\widehat{I}$ (and so $P_0(\bar{a})$ is a fact of I), There are two cases.

Case 1: There is i with $0 \le i \le k-1$ such that $a_{4i+2} = a_{4i+1}$ and $a_{4i+4} = a_{4i+3}$. Then the s-t tgd $P_0(\bar{x}^i) \to Q_i(\bar{x}^i)$ in S_3 and the s-t tgd $Q_i(\bar{x}^i) \to \widehat{P_0}(\bar{x}^i)$ in T_3 guarantee that $\widehat{P_0}(\bar{a})$ is a fact of $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$.

Case 2: There is no i with $0 \le i \le k-1$ such that $a_{4i+2} = a_{4i+1}$ and $a_{4i+4} = a_{4i+3}$. Then the s-t tgd $P_0(x_1,x_2,\ldots,x_{4k}) \to P'_0(x_1,x_2,\ldots,x_{4k})$ in S_1 and the formula in T_2 guarantee that $\widehat{P_0}(\bar{a})$ is a fact of $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$.

We now show the reverse inclusion, that $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I)) \subseteq \widehat{I}$. Because $\mathcal{M}_{12}$ is LAV, and because of the LAV-like form of $\mathcal{M}_{21}$, we see that each fact of $\mathrm{chase}_{21}(\mathrm{chase}_{12}(I))$ is obtained by chasing a single fact $P_i(a_1,a_2,\ldots,a_{4k})$ of I with a single member σ_1 of Σ_{12}, and then chasing the single tuple that results from this chase by a single member σ_2 of Σ_{21}. We must show that the result of this second chase is either empty or is the fact $\widehat{P}_i(a_1,a_2,\ldots,a_{4k})$.

We now consider cases.

Case 1: σ_1 is the s-t tgd

$$P_0(x_1,x_2,\ldots,x_{4k}) \to P'_0(x_1,x_2,\ldots,x_{4k})$$

of S_1. Assume that σ_1 was applied to the fact $P_0(a_1,a_2,\ldots,a_{4k})$ of I to obtain the fact $P'_0(a_1,a_2,\ldots,a_{4k})$. The only member of

Σ_{21} whose premise contains P'_0 is the formula in T_2, and so we may assume that σ_2 is this formula. Since the conclusion of σ_2 is $\widehat{P_0}(x_1, x_2, \ldots, x_{4k})$, it follows that chasing $P'_0(a_1, a_2, \ldots, a_{4k})$ with σ_2 gives either the empty set or the fact $\widehat{P_0}(a_1, a_2, \ldots, a_{4k})$, as desired.

Case 2: σ_1 is the s-t tgd

$$P_i(x_1, x_2, \ldots, x_{4k}) \rightarrow P'_i(x_1, x_2, \ldots, x_{4k})$$

of S_1, for some i with $1 \leq i \leq k$. Assume σ_1 was applied to the fact $P_i(a_1, a_2, \ldots, a_{4k})$ of I to obtain the fact $P'_i(a_1, a_2, \ldots, a_{4k})$. The only member of Σ_{21} whose premise contains P'_i is the tgd $P'_i(x_1, x_2, \ldots, x_{4k}) \rightarrow \widehat{P}_i(x_1, x_2, \ldots, x_{4k})$, and so we may assume that σ_2 is this tgd. Clearly, chasing $P'_i(a_1, a_2, \ldots, a_{4k})$ with σ_2 gives either the empty set or the fact $\widehat{P}_i(a_1, a_2, \ldots, a_{4k})$, as desired.

Case 3: σ_1 is the s-t tgd

$$P_{i+1}(\bar{x}^i) \rightarrow P'_0(\bar{x}^i)$$

of S_2, for some i with $0 \leq i \leq k-1$. Assume that σ_1 was applied to the fact $P_{i+1}(a_1, a_2, \ldots, a_{4k})$ of I to obtain the fact $P'_0(a_1, a_2, \ldots, a_{4k})$. So necessarily $a_{4i+2} = a_{4i+1}$ and $a_{4i+4} = a_{4i+3}$. The only member of Σ_{21} whose premise contains P'_0 is the formula in T_2, and so we may assume that σ_2 is this formula. But this formula is not fired by $P'_0(a_1, a_2, \ldots, a_{4k})$, since $a_{4i+2} = a_{4i+1}$ and $a_{4i+4} = a_{4i+3}$. So this case is not possible.

Case 4: σ_1 is the s-t tgd

$$P_0(\bar{x}^i) \rightarrow Q_i(\bar{x}^i)$$

of S_3, for some i with $0 \leq i \leq k-1$. Assume that σ_1 was applied to the fact $P_0(a_1, a_2, \ldots, a_{4k})$ of I to obtain $Q_i(a_1, a_2, \ldots, a_{4k})$. The only member of Σ_{21} whose premise contains Q_i is the s-t tgd $Q_i(\bar{x}^i) \rightarrow \widehat{P_0}(\bar{x}^i)$ of T_3, so we may assume that σ_2 is this s-t tgd. It follows easily that the result of chasing $Q_i(a_1, a_2, \ldots, a_{4k})$ with σ_2 is $\widehat{P_0}(a_1, a_2, \ldots, a_{4k})$, as desired.

This concludes the proof that $\text{chase}_{21}(\text{chase}_{12}(I)) \subseteq \widehat{I}$, which was the final step in the proof that $\mathcal{M}_{21}$ is an inverse of M_{12}.

We now show that the size of the smallest normal inverse of $\mathcal{M}_{12}$ is exponential in the size of $\mathcal{M}_{12}$. Assume that $\mathcal{M}'_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma'_{21})$ is a normal inverse of $\mathcal{M}_{12}$. It follows from Theorem 4.4 that for every ground instance I:

$$\widehat{I} = \text{chase}'_{21}(\text{chase}_{12}(I)). \qquad (3)$$

Let us refer to $4i+1$ and $4i+3$ as *buddies*, for $0 \leq i \leq k-1$. Let $\bar{a} = (a_1, \ldots, a_{4k})$ be a $4k$-tuple of constants. Let us call $\bar{a}$ *special* if:

1. for each pair i_1, i_2 of buddies, exactly one of the equalities $a_{i_1} = a_{i_1+1}$ or $a_{i_2} = a_{i_2+1}$ holds; and
2. these are the only equalities among members of $\bar{a}$ (that is, if $a_i = a_j$ for distinct values i, j, then there is an odd t such that $\{i, j\} = \{t, t+1\}$).

Let the *equality profile* of $\bar{a}$ be the $2k$-tuple $(\delta_1, \delta_3, \delta_5, \ldots, \delta_{4k-1})$ where $\delta_i = 0$ if $a_i = a_{i+1}$, and $\delta_i = 1$ if $a_i \neq a_{i+1}$. Let us say that an equality profile is *special* if it is the equality profile of a special tuple $\bar{a}$.

For simplicity in what follows, when we say that an inequality $x_i \neq x_j$ *appears in* a formula, we mean that either the inequality $x_i \neq x_j$ or the inequality $x_j \neq x_i$ actually appears. Let σ be a member of Σ'_{21} whose conclusion is of the form $\widehat{P_0}(\bar{x})$, where $\bar{x} = (x_{m_1}, x_{m_2}, \ldots, x_{m_{4k}})$. For each odd number j with $1 \leq j \leq 4k-1$, let us say that j is of *type 0 with respect to* σ if either (a) x_{m_j} and $x_{m_{j+1}}$ are the same variable, or else (b) they are different variables and the inequality $x_{m_j} \neq x_{m_{j+1}}$ does not appear in the premise of σ. If j is not of type 0 with respect to σ, then let us say that it is of *type 1 with respect to* σ. Thus, j is of type 1 precisely if x_{m_j} and $x_{m_{j+1}}$ are distinct variables and the inequality $x_{m_j} \neq x_{m_{j+1}}$ appears in the premise of σ.

Let $\boldsymbol{\delta} = (\delta_1, \delta_3, \delta_5, \ldots, \delta_{4k-1})$ be a special equality profile, and let $\bar{a}$ be a special $4k$-tuple of constants with equality profile $\boldsymbol{\delta}$. Let I be a ground instance whose only fact is $P_0(\bar{a})$. Now $\text{chase}_{12}(I)$ consists of the single fact $P'_0(\bar{a})$ (the s-t tgds in S_2 cannot be applied in the chase since $\bar{a}$ is special). It follows from (3) that there must be a member $\sigma_{\boldsymbol{\delta}}$ of Σ'_{21} such that the chase of $P'_0(\bar{a})$ with $\sigma_{\boldsymbol{\delta}}$ produces $P_0(\bar{a})$. It is clear that $\sigma_{\boldsymbol{\delta}}$ must have the following properties:

1. The conclusion of $\sigma_{\boldsymbol{\delta}}$ is of the form $\widehat{P_0}(\bar{x})$, where $\bar{x} = (x_{m_1}, x_{m_2}, \ldots, x_{m_{4k}})$;
2. variables x_{m_r} and x_{m_s} can be the same variable only if $a_{m_r} = a_{m_s}$;
3. i is of type 0 with respect to $\sigma_{\boldsymbol{\delta}}$ for each i where $\delta_i = 0$; and
4. the only relation symbol that appears in the premise of $\sigma_{\boldsymbol{\delta}}$ is P'_0.

We now show that for each odd i with $1 \leq i \leq 4k-1$, we have that i is of type δ_i with respect to $\sigma_{\boldsymbol{\delta}}$. We already have that if $\delta_i = 0$, then i is of type 0 with respect to $\sigma_{\boldsymbol{\delta}}$ (this follows from the third condition above for $\sigma_{\boldsymbol{\delta}}$). So we need only show that if $\delta_i = 1$, then i is of type 1 with respect to $\sigma_{\boldsymbol{\delta}}$. Let $i_1 = i$, and let i_2 be the buddy of i. Since $\bar{a}$ is special, and since $\delta_{i_1} = 1$, it follows that $\delta_{i_2} = 0$. Therefore, as noted before, i_2 is of type 0 with respect to $\sigma_{\boldsymbol{\delta}}$. Assume that i_1 is also of type 0 with respect to $\sigma_{\boldsymbol{\delta}}$; we shall derive a contradiction.

Let h be a one-to-one mapping from the variables in $\sigma_{\boldsymbol{\delta}}$ to constants, where in particular $h(x_{m_i}) = a_{m_i}$ for each i with $1 \leq i \leq 4k$. This function is well-defined by the second condition about $\sigma_{\boldsymbol{\delta}}$. Let J_1 be the target instance that consists of all of the facts $P'_0(h(y_1), \ldots, h(y_{4k}))$, where the atom $P'_0(y_1, \ldots, y_{4k})$ appears in the premise of $\sigma_{\boldsymbol{\delta}}$. Obtain J_2 from J_1 by replacing each occurrence of a_{i_1+1} by a_{i_1}. Define $\bar{a}' = (a'_1, \ldots, a'_{4k})$ by letting $a'_{i_1+1} = a_{i_1}$, and letting $a'_j = a_j$ if $j \neq i_1 + 1$. Note that $a'_{i_2+1} = a'_{i_2}$, since $a_{i_2+1} = a_{i_2}$ (because $\delta_{i_2} = 0$), and so $a'_{i_2+1} = a_{i_2+1} = a_{i_2} = a'_{i_2}$.

Define h' by letting $h'(y) = h(y)$ if y is not x_{i_1+1}, and letting $h'(x_{i_1+1}) = h(x_{i_1})$, that is, $h'(x_{i_1+1}) = a_{i_1}$. So J_2 consists of all of the facts $P'_0(h'(y_1), \ldots, h'(y_{4k}))$, where the atom $P'_0(y_1, \ldots, y_{4k})$ appears in the premise of $\sigma_{\boldsymbol{\delta}}$. We now show that h' respects each of the inequalities of $\sigma_{\boldsymbol{\delta}}$, that is, that if $y \neq y'$ is an inequality that appears in the premise of $\sigma_{\boldsymbol{\delta}}$, then $h'(y) \neq h'(y')$. There are three cases.

Case 1: $\{y, y'\}$ does not contain x_{i_1+1}. Then $h'(y) = h(y)$ and $h'(y') = h(y)$. Now $h(y) \neq h(y')$, since y and y' are distinct variables. Therefore, $h'(y) \neq h'(y')$, as desired.

Case 2: $\{y, y'\}$ contains x_{i_1+1} but not x_{i_1}. Assume without loss of generality that y is x_{i_1+1}. Then $h'(y) = h(x_{i_1})$ and $h'(y') = h(y')$. Since y' is not x_{i_1}, we know that $h(y') \neq h(x_{i_1})$. So $h'(y') = h(y') \neq h(x_{i_1}) = h'(y)$. Therefore, $h'(y) \neq h'(y')$, as desired.

Case 3: $\{y, y'\} = \{x_{i_1}, x_{i_1+1}\}$. This case is not possible, since i_1 is of type 0 with respect to $\sigma_{\boldsymbol{\delta}}$, and so the inequality $x_{m_{i_1}} \neq x_{m_{i_1+1}}$ does not appear in the premise of $\sigma_{\boldsymbol{\delta}}$.

Since J_2 consists of all of the facts $P'_0(h'(y_1), \ldots, h'(y_{4k}))$, where the atom $P'_0(y_1, \ldots, y_{4k})$ appears in the premise of $\sigma_{\boldsymbol{\delta}}$, and since h' respects each of the inequalities of $\sigma_{\boldsymbol{\delta}}$, it follows that the chase of J_2 with $\sigma_{\boldsymbol{\delta}}$ contains $\widehat{P_0}(\bar{a}')$. Form the source instance I_2

from the target instance J_2 by replacing each fact $P_0'(\bar{b})$ by $P_0(\bar{b})$. Let τ_1 be the s-t tgd $P_0(x_1, x_2, \ldots, x_{4k}) \to P_0'(x_1, x_2, \ldots, x_{4k})$ is J_2. Clearly, the chase of I_2 with τ_1 is J_2.

Since i_1 and i_2 are buddies, there is s with $0 \leq s \leq k-1$ such that $\{i_1, i_2\} = \{4s+1, 4s+3\}$. Let I_2' be the set difference $I_2 \setminus P_0(\bar{a}')$, and let J_2' be the set difference $J_2 \setminus P_0'(\bar{a}')$. Then the chase of I_2' with τ_1 is J_2'. Let I_2'' consist of the fact $P_{s+1}(\bar{a}')$. Let τ_2 be the s-t tgd $P_{s+1}(\bar{x}^i) \to P_0'(\bar{x}^s)$. Since $a'_{i_1} = a'_{i_1+1}$ (by construction) and $a'_{i_2} = a'_{i_2+1}$ (as noted earlier), the chase of I_2'' with τ_2 contains the fact $P_0'(\bar{a}')$. Let $I_3 = I_2' \cup I_2''$. So the chase of I_3 with $\{\tau_1, \tau_2\}$ contains $J_2' \cup \{P_0'(\bar{a}')\}$, which contains J_2. Since τ_1 and τ_2 are members of Σ_{12}, it follows that the chase of I_3 with Σ_{12} contains J_2. Since the chase of I_3 with Σ_{12} contains J_2, and the chase of J_2 with $\sigma_{\boldsymbol{\delta}}$ contains $\widehat{P_0}(\bar{a}')$, it follows that $\mathrm{chase}'_{21}(\mathrm{chase}_{12}(I_3))$ contains $\widehat{P_0}(\bar{a}')$, which is not in $\widehat{I_3}$. Therefore, $\widehat{I_3} \neq \mathrm{chase}'_{21}(\mathrm{chase}_{12}(I_3))$, which contradicts (3) when I is I_3.

We just showed that if $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_5, \ldots, \delta_{4k-1})$ is a special equality profile, then there is a member $\sigma_{\boldsymbol{\delta}}$ of Σ'_{21} such that for each odd i with $1 \leq i \leq 4k-1$, we have that i is of type δ_i. Since $\sigma_{\boldsymbol{\delta}}$ and $\sigma_{\boldsymbol{\delta'}}$ are different when $\boldsymbol{\delta} \neq \boldsymbol{\delta'}$, it follows that Σ'_{21} has at least as many members as there are special equality profiles. Clearly, there are 2^k distinct special equality profiles. So Σ'_{21} has at least 2^k members. This is sufficient to prove the theorem. $\square$

Proof of Theorem 7.3. Let $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$, where Σ_{12} is a finite set of full s-t tgds. For each member $\varphi(\bar{x}) \to (A_1 \wedge \ldots \wedge A_r)$ of Σ_{12}, where each A_i is an atom, let Σ'_{12} contain the s-t tgds $\varphi(\bar{x}) \to A_1, \ldots, \varphi(\bar{x}) \to A_r$. Thus, Σ'_{12} is a finite set of full s-t tgds, each with a singleton conclusion, that is logically equivalent to Σ_{12}.

We now give a procedure to augment Σ'_{12} to a set Σ''_{12}. For each member σ of Σ'_{12}, whose premise consists only of P-atoms for some single relational symbol P, define an equivalence relation $\mathcal{E}_\sigma$ on the variables that appear in σ as follows. Assume that P is t-ary. For each i with $1 \leq i \leq t$, let Y_i be the set of all variables that appear in the ith position of some atom in the premise of σ. Let $\mathcal{E}_\sigma$ be the most refined equivalence relation (largest number of equivalence classes) such that each Y_i is a subset of an equivalence class of $\mathcal{E}_\sigma$. It is easy to see that each equivalence class of $\mathcal{E}_\sigma$ is a union of Y_i's. For each equivalence class, select a unique representative, and let $[x]$ denote the representative of the equivalence class containing x. Form $\sigma^\dagger$ from σ by replacing each variable x by $[x]$. Since $\sigma^\dagger$ is a "special case" of σ (that is, $\sigma^\dagger$ is obtained from σ by identifying some variables), it follows that $\sigma^\dagger$ is a logical consequence of σ. If σ is a member of Σ'_{12} whose premise contains a P-atom and a Q-atom for two different relation symbols P and Q, let $\sigma^\dagger$ be σ. Let $U = \{\sigma^\dagger : \sigma \in \Sigma'_{12}\}$, and let $\Sigma''_{12} = \Sigma'_{12} \cup U$. Since Σ''_{12} consists of Σ'_{12} along with some logical consequences of Σ'_{12}, it follows that Σ''_{12} is logically equivalent to Σ'_{12}. Since also Σ'_{12} is logically equivalent to Σ_{12}, it follows that Σ''_{12} is logically equivalent to Σ_{12}. By renaming variables if needed, we can assume that no two distinct members of Σ''_{12} have a variable in common. Furthermore, we can assume that for each member σ of Σ''_{12}, there is another member $\sigma^\diamond$ of Σ''_{12} that is obtained from σ by renaming the variables in a one-to-one manner and with a disjoint set of variables from σ (we add $\sigma^\diamond$ to Σ''_{12} if needed). It is easy to see that there is a polynomial-time procedure for generating Σ''_{12} from Σ_{12}.

We now give a polynomial-time procedure for generating a set Σ_{21} that specifies an inverse (if there is an inverse). Let us say that a member σ of Σ''_{12} is *special* if the premise contains a single atom, and if every variable in the premise appears in the conclusion (and hence the same variables appear in the premise and the conclusion). Let σ be a special member of Σ''_{12}. Assume that σ is $P(x_{z_1}, \ldots, x_{z_t}) \to Q(x_{i_1}, \ldots, x_{i_k})$. So the conclusion of σ is a Q-atom. Let τ be an arbitrary member of Σ''_{12}, other than σ, such that the conclusion of τ is a Q-atom. Assume that the conclusion of τ is $Q(x_{j_1}, \ldots, x_{j_k})$. Recall that σ and τ have no variables in common. Let $\mathcal{E}^\tau$ be the most refined equivalence relation (largest number of equivalence classes) on the variables in σ and τ such that x_{i_ℓ} and x_{j_ℓ} are in the same equivalence class, for $1 \leq \ell \leq k$. Let θ^1_τ be a conjunction of equalities among the variables in σ, where the equality $x_{i_r} = x_{i_s}$ is an atom in θ^1_τ precisely if x_{i_r} and x_{i_s} are in the same equivalence class of $\mathcal{E}^\tau$. For each equivalence class E of $\mathcal{E}^\tau$, select a unique representative. If this equivalence class E contains a variable in σ, then choose the representative of E to be a variable in σ. (The only times that the equivalence class E does not contain a variable in σ is when E consists of a variable in the premise of τ but not in the conclusion of τ.) Let $[x]^\tau$ denote the representative of the equivalence class of $\mathcal{E}^\tau$ containing x. Let us refer to the variables $[x_{i_1}]^\tau, \ldots, [x_{i_k}]^\tau$ as *distinguished*. Let us say that a P-atom $P(x_{w_1}, \ldots, x_{w_t})$ in the premise of τ is *distinguished* if $[x_{w_\ell}]^\tau$ is distinguished for $1 \leq \ell \leq t$. If A is the distinguished P-atom $P(x_{w_1}, \ldots, x_{w_t})$, define γ_A to be the conjunction of the equalities $[x_{w_\ell}]^\tau = [x_{z_\ell}]^\tau$ for $1 \leq \ell \leq t$. Define θ^2_τ to be the disjunction of the formulas γ_A for each distinguished P-atom of τ. If this disjunction is empty (because τ has no distinguished P-atom), then θ^2_τ is the empty disjunction, which is logically equivalent to *False*. Let θ_τ be the formula $\theta^1_\tau \to \theta^2_\tau$. Note that if τ has no distinguished P-atoms, then θ_τ is logically equivalent to $\neg\theta^2_\tau$. Let θ be the conjunction of the formulas θ_τ, over all members τ of Σ_{12} other than σ, where the conclusion of τ is a Q-atom. We now define σ^* to be $Q(x_{i_1}, \ldots, x_{i_k}) \wedge \theta \to \widehat{P}(x_{z_1}, \ldots, x_{z_t})$. Note that (the hatted version of) the premise of σ is the conclusion of σ^*, and the conclusion of σ is a part of the premise of σ^*. Let Σ_{21} consist of all of the formulas σ^*, where σ is a special member of Σ_{12}. Let $\mathcal{M}_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21})$.

Assume that θ is a Boolean combination of equalities, and f is a weak renaming of variables. Let us say that θ *holds under* f if the Boolean expression that results by replacing each equality $x = y$ by *True* when $f(x)$ and $f(y)$ are the same variable, and replacing each equality $x = y$ by *False* when $f(x)$ and $f(y)$ are different variables, evaluates to *True*. Similarly, if g is a function that maps variables to constants, then say that θ *holds under* g if the Boolean expression that results by replacing each equality $x = y$ by *True* when $g(x)$ and $g(y)$ are the same constant, and replacing each equality $x = y$ by *False* when $g(x)$ and $g(y)$ are different constants, Let us say that f and g *agree on equalities* if for each x, we have that $f(x) = f(y)$ if and only if $g(x) = g(y)$. Clearly, if f and g agree on equalities, then θ holds under f if and only if θ holds under g. As before, if φ is a formula, let φ^f be the result of replacing every variable x in φ by $f(x)$. If A is an atom, let A^g be the fact that arises by replacing every variable x in φ by $g(x)$,

Claim: For every constraint σ^* in Σ_{21}, which must be of the form $\beta \wedge \theta \to \widehat{\alpha}$, where α is a source atom, β is a target atom with the same variables as α, and θ is a Boolean combination of equalities among the variables, and for every weak renaming f, we have that θ holds under f if and only if β^f is an essential atom for α^f.

Note that σ^* is derived from σ in Σ''_{12}, where σ is $\alpha \to \beta$. Assume that α is a P-atom and β is a Q-atom. We now prove the Claim. Assume first that β^f is essential for α^f; we wish to show that θ holds under f. To show this, we must show that if τ is a member of Σ''_{12} other than σ, and the conclusion of τ is a Q-atom, then θ_τ holds under f. Thus, assume that θ^1_τ holds under f; we must show that θ^2_τ holds under f. Now the conclusion of

σ^f is β^f. Since θ_τ^1 holds under f, it follows that σ^f and τ^f have the same conclusion. So the conclusion of τ^f is β^f. Let g be a function that maps variables into constants and that agrees with f on equalities. Let I be an instance whose facts are the facts A^g for each atom A in the premise of τ. So the chase of I with τ is β^g. Since β^f is essential for α^f, it follows that α^g is a fact in I. So α^g is A^g for some atom A in the premise of τ. It follows that γ_A, as defined earlier, holds under g, and so θ_τ^2 holds under g. Since f and g agree on equalities, this implies that θ_τ^2 holds under f, as desired.

Assume now that θ holds under f; we must show that β^f is an essential atom for α^f. Let g be a function that maps variables into constants and that agrees with f on equalities. So θ holds under g. Let I be an instance where $\mathrm{chase}_{12}(I)$ contains β^g; we need only show that α^g is a fact in I. It is easy to see that the result of chasing with Σ_{12} and Σ''_{12} are the same. So the result of chasing I with Σ''_{12} contains β^g. Hence, there is a constraint τ in Σ''_{12} that fires on I and produces β^g. If τ is σ, then α^g is in I, as desired. If τ is not σ, it is straightforward to verify that θ_τ^1 holds under g. Since also θ holds under g, this implies that θ_τ^2 holds under g. So there is some distinguished atom A in the premise of τ such that γ_A holds under g. Hence, A^g and α^g are the same fact. Since τ fires on I to produce β^g, there is a homomorphism h that maps the premise of τ into I and that maps the conclusion of τ onto β^g. Hence, h must agree with g on the variables in α, and hence on the variables in A, since A is distinguished. So A^g is in I. But we showed that A^g and α^g are the same fact. So α^g is in I, as desired. This concludes the proof of the Claim.

Assume that $\mathcal{M}_{12}$ has an inverse. We now use the Claim to prove that $\mathcal{M}_{21}$ is an inverse of $\mathcal{M}_{12}$.

Assume that σ^* is a member of Σ_{21}, and σ^* is $\beta \wedge \theta \rightarrow \widehat{\alpha}$. Let k be the number of variables that appear in σ^*. Define the set T_{σ^*} as follows. For each weak renaming f of the variables in σ^* such that the range of f is in $\{x_1, \ldots, x_k\}$ and such that θ holds for f, let T_{σ^*} contain the constraints $\beta^f \wedge \eta_f \rightarrow \widehat{\alpha^f}$, where η_f is the conjunction of the inequalities $f(x) \neq f(y)$ where x and y are variables of σ^* and where $f(x)$ and $f(y)$ are different variables. (The assumption that range of f is in $\{x_1, \ldots, x_k\}$ is only to assure that T_{σ^*} be finite.) It is straightforward to see that σ^* is logically equivalent to T_{σ^*}. Let Σ'_{21} be the union of the sets T_{σ^*} over all σ^* in Σ_{21}, and let $\mathcal{M}'_{21} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma'_{12})$. We need only show that $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$. Obtain Σ''_{21} from Σ'_{21} by adding to the premise of every member τ of Σ'_{21} the conjuncts $\mathrm{const}(x)$ where x is a variable that appears in τ. Let $\mathcal{M}''_{21} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma''_{21})$ By Proposition A.7, we know that $\mathcal{M}'_{21}$ is an inverse of $\mathcal{M}_{12}$ if and only if $\mathcal{M}''_{21}$ is an inverse of $\mathcal{M}_{12}$. So we need only show that $\mathcal{M}''_{21}$ is an inverse of $\mathcal{M}_{12}$. Note that by construction, $\mathcal{M}''_{21}$ is normal.

We now use Theorem 4.8 to show that $\mathcal{M}''_{21}$ is an inverse of $\mathcal{M}_{12}$. Since each T_{σ^*} was obtained by considering weak renamings f such that θ holds for f, it follows easily from the Claim that for every member φ of Σ''_{21}, the premise of φ is essential for the conclusion of φ. Hence, the first condition of Theorem 4.8 holds (when Σ''_{21} plays the role of Σ_{21}). We now show that the second condition also holds. Let A be a source atom. Since $\mathcal{M}_{12}$ is invertible, we know by Proposition 4.12 that ω_A is essential for A, and so contains an atom B that is essential for A (with respect to Σ_{12}).

It follows from the construction of Σ''_{12} that there is a member σ of Σ''_{12} with a singleton premise (and a singleton conclusion) such that the chase of I_A is the same with σ as it is with Σ_{12}. Write σ as $\alpha \rightarrow \beta$. So there is a weak renaming f such that α^f is A and β^f is B. We now show that every variable in α appears in β, and so σ is special. Assume that some variable x appears in α but not in β; we shall derive a contradiction. Let f' be a weak renaming that is like f except that $f'(x)$ is a new variable. So $\alpha^{f'}$ is different from A, although $\beta^{f'}$ is the same as β^f, that is, B. So $\mathrm{chase}_{12}(I_{\alpha^{f'}})$ contains I_B, even though $I_A \not\subseteq I_{\alpha^{f'}}$. This contradicts the fact that B is essential for A. Hence, σ is special, as desired.

So there is θ such that σ^* is $\beta \wedge \theta \rightarrow \widehat{\alpha}$, and σ^* is in Σ_{21}. Since β^f (namely, B) is essential for α^f (anmely, A), it follows from the Claim that θ holds under f. So there is δ such that the only atom in δ is B, and $\delta \rightarrow \widehat{A}$ is a weak renaming of a constraint in Σ''_{21}. Hence, the second condition of Theorem 4.8 holds (when Σ''_{21} plays the role of Σ_{21}), as desired. This completes the proof that $\mathcal{M}''_{21}$ is an inverse of $\mathcal{M}_{12}$. By making use of Proposition A.7, we can add const formulas to the premises of member of $\mathcal{M}''_{21}$ to obtain a Boolean normal inverse of $\mathcal{M}_{12}$. $\square$

A.6 Proofs for Section 8

Proof of Theorem 8.2. Assume that $\mathcal{M}_{12} = (\mathbf{S_1}, \mathbf{S_2}, \Sigma_{12})$. Let us say that the source atom B is *good* if $\mathrm{chase}_{12}(I_B)$ has exactly one member. Let us say that B is *bad* if B is not good. Let b be the number of bad prime source atoms. We now show that $2^b \leq m$ (where m is the number of inequivalent normal inverses of $\mathcal{M}$) and that $\mathcal{M}$ has a Boolean normal inverse of length $k + b$. Since $2^b \leq m$, we have that $b \leq \log_2(m)$, and so $k + b \leq k + \log_2(m)$. The theorem follows.

For each source relation symbol P, let A_P be the prime P-atom $P(x_1, \ldots x_r)$ where $x_1, \ldots, x_r$ are distinct. If B is a P-atom $P(y_1, \ldots, y_r)$, let φ_B be the formula that is the conjunction of the equalities $x_i = x_j$ for each i, j where y_i and y_j are the same variable along with the inequalities of the form $x_i \neq x_j$ for each i, j where y_i and y_j are different variables. Intuitively, φ_B completely describes the equality pattern of the variables in B. Let θ_P be the disjunction of the formulas φ_B where B is a good P-atom. Let σ_P be the formula $\omega_{A_P} \wedge \theta_P \rightarrow \widehat{A_P}$, where ω_{A_P} is defined as in Definition 4.11.

Let $B_1, \ldots, B_b$ be precisely the bad prime source atoms (they may involve various relation symbols). If B_i is $P(y_1, \ldots, y_r)$, define η_i to be the conjunction of the inequalities of the form $y_i \neq y_j$ where y_i and y_j are distinct variables. By Proposition 4.12, we know that ω_{B_i} is essential for B_i with respect to Σ_{12}, for each i. Since B_i is bad, it follows that ω_{B_i} is a conjunction of more than one atom. By Proposition 4.5, we know that some atom C_i in ω_{B_i} is essential for B_i with respect to Σ_{12}, for each i. Let ψ_i^0 be the constraint $C_i \wedge \eta_i \rightarrow \widehat{B_i}$, and let ψ_i^1 be the constraint $\omega_{B_i} \wedge \eta_i \rightarrow \widehat{B_i}$.

Let $\mathbf{v} = (v_1, \ldots, v_b)$ be an arbitrary $\{0,1\}$-vector of length b. Define $\Sigma_{21}^{\mathbf{v}}$ to consist of the k formulas σ_P (one for each source relation symbol P) along with the b constraints $\psi_i^{v_i}$ for $1 \leq i \leq b$. Let $\mathcal{M}_{21}^{\mathbf{v}} = (\mathbf{S_2}, \widehat{\mathbf{S_1}}, \Sigma_{21}^{\mathbf{v}})$. We now show that each $\mathcal{M}_{21}^{\mathbf{v}}$ is an inverse of $\mathcal{M}_{12}$, and that $\mathcal{M}_{21}^{\mathbf{v}}$ and $\mathcal{M}_{21}^{\mathbf{v}'}$ are not equivalent if $\mathbf{v} \neq \mathbf{v}'$. Since the number of vectors $\mathbf{v}$ is 2^v, this shows that $2^b \leq m$. Further, since each $\mathcal{M}_{21}^{\mathbf{v}}$ is a Boolean normal inverse of length $k + b$, this shows that $\mathcal{M}$ has a Boolean normal inverse of length $k + b$ (in fact, it has at least 2^b Boolean normal inverse of length $k + b$). This is sufficient to complete the proof.

Fix $\mathbf{v} = (v_1, \ldots, v_b)$. We begin by showing that $\mathcal{M}_{21}^{\mathbf{v}}$ is an inverse of $\mathcal{M}_{12}$. We now define a function e that maps each prime source atom B to an essential conjunction $e(B)$ with respect to Σ_{12}. For the bad prime source atom B_i, we let $e(B_i) = C_i$ if $v_i = 0$, and $e(B_i) = \omega_{B_i}$ if $v_i = 1$. By construction, $e(B_i)$ is essential for B_i if $v_i = 0$, and by Proposition 4.12, we know that $e(B_i)$ is essential for B_i if $v_i = 1$. For each good prime source atom A, we let $e(B) = \omega_B$. Again by Proposition 4.12, we know that $e(B)$ is then essential for B. So by Theorem 4.10, $\mathcal{M}_{21}^e$ is an inverse of $\mathcal{M}$. We now show that $\mathcal{M}_{21}^e$ is equivalent to $\mathcal{M}_{21}^{\mathbf{v}}$, which completes the proof that $\mathcal{M}_{21}^{\mathbf{v}}$ is an inverse of $\mathcal{M}_{12}$.

For each prime source atom B where B is bad, $\mathcal{M}_{21}^{e}$ and $\mathcal{M}_{21}^{\mathbf{v}}$ contain the same constraint with conclusion $\widehat{B}$. Let us now consider the good prime source atoms B. The formula σ_P is logically equivalent to the set consisting of all of the formulas $\omega_{A_P} \wedge \varphi_B \rightarrow \widehat{A_P}$, where B is a good prime P-atom. Assume that B is a good prime source atom. Let σ_1 be the formula $\omega_{A_P} \wedge \varphi_B \rightarrow \widehat{A_P}$, and let σ_2 be the formula $\omega_B \wedge \eta_B \rightarrow \widehat{B}$, where as before η_B is the conjunction of all inequalities of the form $x \neq y$ where x and y are distinct variables in B, By construction, σ_2 is the unique member of $\mathcal{M}_{21}^{e}$ with conclusion $\widehat{B}$. So to complete the proof that $\mathcal{M}_{21}^{e}$ is equivalent to $\mathcal{M}_{21}^{\mathbf{v}}$, we need only show that the formula σ_1 is logically equivalent to the formula σ_2.

Assume that B is the good atom $P(y_1, \ldots, y_r)$, where $y_1, \ldots, y_r$ are variables, not necessarily distinct. Let ψ_B be the formula obtained from σ_1 by replacing the variable x_i by y_i, for $1 \leq i \leq r$. We now show that ψ_B is logically equivalent to both σ_1 and σ_2, which implies that σ_1 and σ_2 are logically equivalent, as desired. In forming ψ_B, two variables x_i and x_j in σ_1 are replaced by the same variable precisely if y_i and y_j are the same variable, which holds precisely if the equality $x_i = x_j$ appears in φ_B. It follows easily that ψ_B is logically equivalent to σ_1. We now show that ψ_B is logically equivalent to σ_2.

It is easy to see that the conclusions of ψ_B and σ_2 are the same, and that the result of replacing the variable x_i by y_i, for $1 \leq i \leq r$, in φ_B is equivalent to η_B. Let τ_B be the result of replacing the variable x_i by y_i in ω_{A_P}, for $1 \leq i \leq r$. So we need only show that τ_B is equivalent to ω_B. Now the conjunct(s) of τ_B must be in ω_B, by properties of the chase with s-t tgds. Since ω_B is a singleton (because B is good), it follows easily that τ_B is the same as ω_B. This concludes the proof that $\mathcal{M}_{21}^{e}$ is equivalent to $\mathcal{M}_{21}^{\mathbf{v}}$,

We conclude the proof by showing that $\mathcal{M}_{21}^{\mathbf{v}}$ and $\mathcal{M}_{21}^{\mathbf{v}'}$ are not equivalent if $\mathbf{v} \neq \mathbf{v}'$. Say $\mathbf{v} \neq \mathbf{v}'$, and that $\mathbf{v} = (v_1, \ldots, v_b)$ and $\mathbf{v}' = (v_1', \ldots, v_b')$. So there is i with $1 \leq i \leq b$ such that $v_i \neq v_i'$ Assume without loss of generality that $v_i = 0$ and $v_i' = 1$. We now show that $(I_{C_i}, \emptyset)$ satisfies $\Sigma_{21}^{\mathbf{v}'}$ but not $\Sigma_{21}^{\mathbf{v}}$ This of course shows that $\mathcal{M}_{21}^{\mathbf{v}}$ and $\mathcal{M}_{21}^{\mathbf{v}'}$ are not equivalent. Clearly ψ_i^0 fires on I_{C_i}, and so $(I_{C_i}, \emptyset)$ does not satisfy ψ_i^0. Hence, $(I_{C_i}, \emptyset)$ does not satisfy $\Sigma_{21}^{\mathbf{v}}$, because $\Sigma_{21}^{\mathbf{v}}$ contains ψ_i^0. We now show that no member of $\Sigma_{21}^{\mathbf{v}'}$ fires on I_{C_i}. Since B_i is bad, we know that ω_{B_i} has some other atom A in addition to C_i as a conjunct. Since C_i is essential for B_i, it follows from Proposition 4.6 that B_i and C_i have the same variables. Since Σ_{12} is full, every variable in A is in B_i, and hence in C_i. Assume that C_i is $Q(y_1, \ldots, y_m)$. Then I_{C_i} consists of the fact $Q(c_{y_1}, \ldots, c_{y_m})$. If ψ_i^1 were to fire on I_{C_i}, then there would be a homomorphism h from the premise of ψ_i^1 to I_{C_i}. Since C_i is part of the premise of ψ_i^1, we must have $h(y_i) = c_{y_i}$ for $1 \leq i \leq m$. Since h must map A onto $Q(c_{y_1}, \ldots, c_{y_m})$, and since every variable in A is among $y_1, \ldots, y_m$, it is easy to see that A must be C_i, which is a contradiction. So ψ_i^1 does not fire on I_{C_i}. We now show that no other member of $\Sigma_{21}^{\mathbf{v}'}$ fires on I_{C_i}. If some member $\psi_j^{v'_j}$ were to fire on I_{C_i} where $j \neq i$, then because of the inequalities in $\psi_j^{v'_j}$, it would follow that some member of $\mathrm{chase}_{12}(I_{B_j})$ is of the form $Q(c_1, \ldots, c_m)$, where $c_k = c_\ell$ if and only if $y_k = y_\ell$. So there would be a homomorphism $I_{C_i} \rightarrow \mathrm{chase}_{12}(I_{B_j})$. Since C_i is demanding for B_i, it follows that $I_{B_i} \subseteq I_{B_j}$. But this is impossible, since $i \neq j$. So no member $\psi_j^{v'_j}$ fires on I_{C_i} where $j \neq i$. A similar argument shows that no σ_P fires on I_{C_i}. So no member of $\Sigma_{21}^{\mathbf{v}'}$ fires on I_{C_i}, and hence $(I_{C_i}, \emptyset)$ satisfies $\Sigma_{21}^{\mathbf{v}'}$, as desired. $\square$