Why is it hard to obtain a dichotomy for consistent query answering?

Gaëlle Fontaine
Department of Computer Sciences
University of Chile
Email: gaelle@dcc.uchile.cl

Abstract—A database may for various reasons become inconsistent with respect to a given set of integrity constraints. To overcome the problem, a formal approach to querying such inconsistent databases has been proposed and since then, a lot of efforts have been spent to classify the complexity of consistent query answering under various classes of constraints. It is known that for the most common constraints and queries, the problem is in \( \text{coNP} \) and might be \( \text{coNP}-\text{hard} \), yet several relevant tractable classes have been identified. Additionally, the results that emerged suggested that given a set of key constraints and a conjunctive query, the problem of consistent query answering is either in \( \text{PTime} \) or is \( \text{coNP}-\text{complete} \). However, despite all the work, as of today this dichotomy remains a conjecture.

The main contribution of this paper is to explain why it appears so difficult to obtain a dichotomy result in the setting of consistent query answering. Namely, we prove that such a dichotomy w.r.t. common classes of constraints and queries, is harder to achieve than a dichotomy for the constraint satisfaction problem, which is a famous open problem since the 1990s.

I. INTRODUCTION

A. Querying inconsistent databases

One way to control databases is to impose integrity constraints upon them, that is, semantic properties that the database must obey. However, in many situations control can be lost (e.g. in the context of data integration or exchange). This gives rise to inconsistent databases, which no longer satisfy the constraints.

To overcome the problem, one option is to restore consistency using data cleaning. The approach consists in arbitrarily transforming the database into a well-behaved one. Another approach, introduced by Arenas et al [4], is to directly query the original database, as inconsistent as it is. The consistent answer of a query \( q \) on an inconsistent database \( D \) is then defined as the intersection of the answers of \( q \) on all the consistent databases that differs from \( D \) in a “minimal way”.

The approach is elegant and undoubtedly more principled than data cleaning. Unsurprisingly, the abstraction of the method is counterbalanced by a high computational complexity. Since the seminal work of Arenas et al [4], the computational complexity of consistent query answering has been studied for various classes of constraints. Initially, the focus was on functional constraints, inclusion dependencies and denial constraints (see the overviews [6], [11]). More recently, other classes of constraints such as LAV constraints, GAV constraints, tuple-generating-dependencies (tgds) and equality-generating-dependencies (egds) [22], [3], [23] have also been considered. Those constraints play a central role in data integration [19] and data exchange [13], [2].

As an attempt to classify the complexity of consistent query answering, Afrati and Kolaitis [1] raised the question of the existence of a dichotomy result for the problem of consistent query answering under a set of key constraints. They conjectured that given a conjunctive query \( q \) and a set of key constraints \( \Sigma \), the problem of consistent query answering of \( q \) under \( \Sigma \) should either be in \( \text{PTime} \) or \( \text{coNP}-\text{complete} \). Recall that if \( \text{PTime} \neq \text{NP} \), there are infinitely many intermediate problems in \( \text{coNP} \) that are neither \( \text{coNP}-\text{complete} \) nor belong to \( \text{PTime} \) [18]. A dichotomy conjecture states that the considered class of problems does not contain any intermediate problem.

The question has been actively explored recently; yet only few results, and in very restricted settings, have been obtained. The first of these results is a necessary and sufficient condition for first-order rewriting of acyclic conjunctive queries without self-joins [24] (note that first-order rewritability implies tractability for consistent query answering). Given that condition, Kolaitis and Pema [16] proved a dichotomy theorem for queries containing only two atoms and no self-joins.

Part of this work was done while doing a postdoctoral stay at the University of California Santa Cruz, supported by NSF Grant IIS-0905276. The author is currently funded by Fondecyt grant 3130491 of Conicyt.
Even with such strong restrictions, the proof turned out to be involved.

We show that there is actually a very good reason for the difficulties encountered. We prove that a dichotomy result for consistent query answering would imply a solution for a famous long-standing open problem, namely the dichotomy conjecture for the constraint satisfaction problem.

B. Constraint satisfaction problem

The constraint satisfaction problem (CSP) is a fundamental topic in computer science, the main reason being that CSP provides a common framework for a wide range of problems arising in theoretical computer science and artificial intelligence. An instance of CSP is determined by a set of variables, a set of values and a set of constraints. The goal is to assign a value to each variable in such a way that the constraints are satisfied.

In general, CSP is in NP and there are families of instances (e.g. boolean satisfiability) that are known to be NP-complete. An impressive amount of effort has been devoted to isolate tractable cases and develop heuristics. The classes of instances that received the most attention are the non-uniform constraint satisfaction problems. Each of those classes is characterized by a fixed set of allowed constraint relations; examples include boolean satisfiability, graph coloring and systems of equations.

The first major result [20] concerning non-uniform CSP establishes that every boolean non-uniform CSP is either polynomial or is NP-complete, where an instance of CSP is said to be boolean if its set of values contains exactly two elements. Feder and Vardi [14] postulated that the result holds for arbitrary non-uniform CSP, that is, each non-uniform CSP is either solvable in polynomial time or is NP-complete. This conjecture is known as the dichotomy conjecture for CSP and is the most important open problem in the field.

Initially, and despite the considerable attention received by the problem, progress was slow. However, after the adoption of an algebraic approach, some significant results have been obtained. The most recent developments include a dichotomy theorem for non-uniform CSP over sets of values with three elements [8] and a dichotomy theorem for non-uniform conservative CSP [7], [5], that is, non-uniform CSP over a constraint language containing all unary relations. The proofs of those results are highly complex.

C. Linking two conjectures about separation

Our goal is to explain why it appears so difficult to obtain a dichotomy result in the setting of consistent query answering. We do so by proving that if such a dichotomy result holds, then so does the dichotomy conjecture for CSP. We were not able to prove such a result in the setting described by Afrati and Kolaitis (i.e. key constraints and conjunctive queries). The solution is to turn our attention to GAV constraints and unions of conjunctive queries (UCQ), which are common well-studied classes of constraints and queries.

The main result establishes that if the dichotomy conjecture holds for consistent query answering of UCQs w.r.t a set of GAV constraints, then so does the dichotomy conjecture for CSP. Given the fact that the dichotomy conjecture for CSP is still open and that a proof would be the most fundamental breakthrough in the study of CSP, our result means that there is very little hope in pursuing a dichotomy result for consistent query answering in its most general form.

Concerning key constraints, even though we do not have a result similar to our main theorem, we prove that a dichotomy result for consistent query answering of UCQs with constants w.r.t. key constraints, would yield to an alternative proof of the dichotomy theorem for conservative CSP. Considering the time and the effort spent to obtain a dichotomy for conservative CSP, this shows that a dichotomy for consistent query answering in the setting described above is a highly difficult task.

Our third result establishes that a dichotomy result for consistent query answering of UCQs w.r.t. egds, would yield to an alternative proof of the dichotomy theorem for conservative CSP. Compared to our second result, this shows that if we are willing to consider egds constraints instead of key constraints, then we do not need constants in the queries.

The three results presented provide a formal explanation of the difficulty of proving a dichotomy for consistent query answering; they also emphasize the close connection between consistent query answering and CSP. It does not mean though that no further investigation of a dichotomy for consistent query answering in restricted settings should be pursued and that no meaningful understanding will be gained.

D. Related work

Links between the dichotomy conjecture for (non-uniform) CSP and a possible dichotomy result for problems arising in database theory have been previously explored. Feder and Vardi [14] proved that the logic MMSNP and non-uniform CSP are polynomially equivalent. Hence, the dichotomy conjecture holds for CSP if it holds for MMSNP. Finally, let us mention the results of Calvanese et al [10] establishing a connection between
the tractable instances of CSP and the instances of query rewriting that admit a perfect rewriting in polynomial time. Those results do not prove though that a dichotomy theorem in one setting implies a dichotomy result in the other setting.

**Organisation of the paper** In Section II, we introduce the basics of consistent query answering and CSP. In Section III, we present our main result, namely that a dichotomy result for consistent query answering of UCQs w.r.t. GAV constraints, implies a dichotomy theorem for CSP. Finally, in Section IV, we mention two other results establishing a connection between conservative CSP and consistent query answering of UCQs w.r.t. key constraints and egds. Concluding remarks can be found in Section V.

**II. Preliminaries**

**A. Consistent query answering**

A schema $\sigma$ is a finite sequence $(R_1, \ldots, R_k)$, where each $R_i$ is a relation symbol of arity $n_i > 0$. A database $D$ over the schema $\sigma$ assigns to each relation symbol $R_i$ a finite $n_i$-ary relation $R_i^D$. The active domain is the set of all elements that occur in any of the relations $R_i^D$. Databases can be seen as first-order structures by taking the domain to be the active domain.

If $(a_1, \ldots, a_n)$ belongs to $R_i^D$, we say that $R_i(a_1, \ldots, a_n)$ is a fact of $D$. Each database can be identified with the set of its facts.

A set of constraints $\Sigma$ is a set of first-order formulas over $\sigma$. A database is consistent w.r.t. $\Sigma$ if it satisfies the formulas in $\Sigma$. Otherwise, the database is inconsistent. In this paper, we focus on the following constraints.

**Definition II.1.** [19] A tuple-generating dependency (tgds) or equality-generating dependency (egd) is a first-order formula of the form

$$\forall x \exists y (\phi(x) \rightarrow \psi(x, y)),$$

where $\phi$ and $\psi$ are conjunctions of atomic formulas and $x$ and $y$ are tuples of variables. Such a tgd is a local-as-view dependency (LAV) if $\phi$ consists of a single atomic formula.

A global-as-view dependency (GAV) is a tgd of the form

$$\forall x (\phi(x) \rightarrow \psi(x')),$$

where $x$ and $x'$ are tuples of variables and the variables in $x'$ occur in $x$.

An equality-generating dependency (egd) is a first-order formula of the form

$$\forall x (\phi(x) \rightarrow y = z),$$

where $\phi$ is a conjunction of atomic formulas, $x$ is a tuple of variables and $y$ and $z$ are variables occurring in $x$.

A key constraint is a first-order formula of the form

$$\forall x, y, z (R(x, y) \land R(x, z) \rightarrow y = z),$$

where $x$ and $y$ and $z$ are tuples of variables.

For the sake of readability, we will drop the universal quantifiers when writing constraints.

Tuple-generating-dependencies (tgds) and egds play a fundamental role in data exchange [13], [2] and data integration [19]; they are used to express the relationship between a local source database and a global target database. Typically the relation symbols occurring on the left side of the implication of a tgd belong to the schema of the source database, while the symbols occurring on the right side belong to the schema of the target database. Hence, a tgd specifies how conditions verified by the source imply conditions on the target.

Among the class of tgds, two important subclasses have been extensively studied: the LAV (local-as-view) dependencies and the GAV dependencies. In the case of GAV, since only one relation symbol occurs on the right side of the implication, each relation of the target database is defined in terms of the relations in the source database. In the case of LAV, relations of the source are described in terms of the relations of the target.

Our main result is concerned with the problem of querying databases that do not satisfy a given set of GAV constraints. The approach of querying inconsistent databases introduced by Arenas et al [4] has been developed around the notion of repair. Intuitively, a database is a repair of an inconsistent database if it satisfies the constraints and differs from the original database in a “minimal way”. Several notions of minimality have been introduced, giving rise to different definitions of repairs. Here, we opt for a standard notion of minimality, based on the set inclusion order. If $D$ and $E$ are databases, we denote by $D \oplus E$ the symmetric difference of $D$ and $E$, i.e. the set $D \setminus E \cup E \setminus D$.

**Definition II.2** (Repair). Let $\Sigma$ be a set of constraints. A database $E$ is a repair of a database $D$ w.r.t. $\Sigma$ if $E \models \Sigma$ and for all databases $E'$ such that $E' \models \Sigma$ and $E' \oplus D \subseteq E \oplus D$, we have $E = E'$.

The queries that we consider in this work are unions of conjunctive queries (UCQ). Recall that a conjunctive query (CQ) is a formula of the form

$$q(x) \equiv y \phi(x, y),$$

1Note that in order to simplify notations, we assumed that the variables in $x$ occur in the first positions of $R$. In general, this does not need to be the case.
where $\phi$ is a conjunction of atomic formulas. If a variable $x$ occurs in $x$ and not in $y$, $x$ is a free variable. A conjunctive query with constants is a CQ for which we allow the use of constants in the atomic formulas. We stick to the usual convention that the interpretation of a constant on a database is the constant itself.

Conjunctive queries are the most fundamental class of queries in database theory and form the core of all practical query languages. UCQs are disjunctions of conjunctive queries; they are easily seen to be equivalent to the existential and positive fragment of first-order logic. A UCQ is boolean to the existential and positive fragment of first-order logic. A UCQ is a boolean query, we write $\text{UCQ}(q,D,\Sigma) = \top$ if it does not contain any free variable. If $D$ is a database and $q$ an UCQ, we denote by $q(D)$ the set of tuples that belong to the evaluation of $q$ over $D$. The answers of a query on an inconsistent database $D$ are obtained by evaluating the query over all the repairs of $D$ and taking the intersection.

**Definition II.3** (Consistent query answering). Let $\Sigma$ be a set of constraints, $D$ a database and $q$ a query. The consistent answers of $q$ on $D$ w.r.t. $\Sigma$, denoted by $\text{CQA}(q,D,\Sigma)$, is defined as the set
\[
\bigcap \{q(E) : E \text{ is a repair of } D \text{ w.r.t. } \Sigma \}.
\]
If $q$ is a boolean query, we write $\text{CQA}(q,D,\Sigma) = \top$ if $q$ is true in all the repairs of $D$ w.r.t. $\Sigma$. Otherwise, $\text{CQA}(q,D,\Sigma) = \bot$.

The consistent query answering problem of $q$ w.r.t. $\Sigma$, denoted by $\text{CQA}(q,\Sigma)$, is the following problem: given a database $D$ and a tuple, determine whether the tuple is a consistent answer of $q$ on $D$ w.r.t. $\Sigma$. We write $\overline{\text{CQA}}(q,\Sigma)$ for the following problem: given a database $D$ and a tuple, determine whether the tuple is not a consistent answer of $q$ on $D$ w.r.t. $\Sigma$.

As mentioned in the introduction, the complexity of consistent query answering under various classes of constraints has been deeply investigated in the last two decades. Since here we only consider constraints that are GAV, egds or keys, we simply recall that in each of those cases, the consistent query answering problem is known to be in $\text{coNP}$ [12], [21].

Afrati and Kolaitis [1] pushed further the investigation of the complexity of consistent query answering by conjecturing the existence of a dichotomy result. Although the original conjecture was stated for key constraints and conjunctive queries, we give here a more general formulation.

**Definition II.4** (Dichotomy conjecture). Let $C$ be a class of constraints and let $Q$ be a class of queries such that for all subsets $\Sigma$ of $C$ and for all queries $q \in Q$, $\text{CQA}(q,\Sigma)$ is in $\text{coNP}$. The dichotomy conjecture w.r.t. $C$ and $Q$ states that for all subsets $\Sigma$ of $C$ and for all queries $q \in Q$, $\text{CQA}(q,\Sigma)$ is either in $\text{PTIME}$ or is $\text{NP}$-complete.

**Conjecture II.5.** [1] The dichotomy conjecture w.r.t. key constraints and conjunctive queries holds.

As mentioned earlier, the most recent contribution to the above conjecture is a dichotomy result for the case of CQs with two atoms and no self-joins [16].

**B. Constraint satisfaction problem**

An instance of the constraint satisfaction is defined by a set of values, a set of variables and a set of constraints, and asks whether there is a way to assign a value to each variable such that the constraints are satisfied. For our purpose, we adopt an equivalent formulation of the constraint satisfaction problem in terms of homomorphisms [17].

Recall that a map $h : \mathbb{A} \to \mathbb{B}$ between two structures is an homomorphism if for all relation symbols $R$ and for all $(a_1, \ldots, a_n) \in R^\mathbb{A}$, $(h(a_1), \ldots, h(a_n))$ belongs to $R^\mathbb{B}$.

Given a map $h : \mathbb{A} \to \mathbb{B}$ and a tuple $a = (a_1, \ldots, a_n)$ of elements in $\mathbb{A}$, we denote by $h(a)$ the tuple $(h(a_1), \ldots, h(a_n))$. Moreover, we denote by $A$ the domain of the structure $\mathbb{A}$ and by $B$ the domain of $\mathbb{B}$.

**Definition II.6.** Let $\mathbb{B}$ be a structure. The (non-uniform) constraint satisfaction problem $\text{CSP}(\mathbb{B})$ is the following problem: given a structure $\mathbb{A}$, determine whether there is a homomorphism $h : \mathbb{A} \to \mathbb{B}$.

**The dichotomy conjecture for CSP** states that for every structure $\mathbb{B}$, $\text{CSP}(\mathbb{B})$ is either in $\text{PTIME}$ or is $\text{NP}$-complete. It was shown [15] that if $\mathbb{B}'$ is the core of $\mathbb{B}$, then $\text{CSP}(\mathbb{B})$ and $\text{CSP}(\mathbb{B}')$ are polynomially equivalent. Moreover, if $\mathbb{B}'$ is a core, then Bulatov et al. [9] established that $\text{CSP}(\mathbb{B}')$ is tractable (resp. $\text{NP}$-complete) iff $p\text{Hom}(\mathbb{B}')$ is tractable (resp. $\text{NP}$-complete), where $p\text{Hom}(\mathbb{B}')$ is the pointed homomorphism problem as defined below. Hence, in order to prove the dichotomy conjecture, we may restrict ourselves to the study the pointed homomorphism problem.

**Definition II.7** (Pointed homomorphism problem). Let $\mathbb{B}$ be a structure. We define the pointed homomorphism problem $p\text{Hom}(\mathbb{B})$ as the following problem: given a structure $\mathbb{A}$ and a partial homomorphism $f : \mathbb{A} \to \mathbb{B}$, determine whether there is a homomorphism $g : \mathbb{A} \to \mathbb{B}$ extending $f$.

The dichotomy conjecture for the pointed homomorphism problems states that for every structure $\mathbb{B}$, $p\text{Hom}(\mathbb{B})$ is either in $\text{PTIME}$ or is $\text{NP}$-complete.
Finally, we recall the dichotomy result proved by Bulatov [7] for CSP over a schema containing all unary relations. In terms of homomorphisms, the result is formulated as follows.

**Definition II.9.** Let $\mathbb{B}$ be a structure. The conservative homomorphism satisfaction problem $\text{cHom}(\mathbb{B})$ is the following problem: given a structure $\mathbb{A}$ and a set of constraints $\Sigma$, determine whether there is a homomorphism $h: \mathbb{A} \to \mathbb{B}$ such that for all $a \in \mathbb{A}$, $h(a)$ belongs to $L_a$.

**Theorem II.10.** [7], [5] The dichotomy conjecture for the conservative homomorphism satisfaction problems holds.

### III. Main result

Our main result establishes a connection between the dichotomy conjecture for CSP and the dichotomy conjecture for consistent query answering of UCQs w.r.t. GAV constraints.

**Theorem III.1.** If the dichotomy conjecture for consistent query answering of UCQs w.r.t. GAV constraints holds, then so does the dichotomy conjecture for the constraint satisfaction problems.

By Proposition II.8, in order to prove Theorem III.1, it is sufficient to show that if the dichotomy conjecture for consistent query answering of UCQs w.r.t. GAV constraints holds, then so does the dichotomy conjecture for the pointed homomorphism problems. This is a direct consequence of Proposition III.2 proved below.

**Proposition III.2.** For each structure $\mathbb{B}$, we can compute a boolean UCQ $q$ and a set $\Sigma$ of GAV constraints such that $\text{pHom}(\mathbb{B})$ and $\text{CQA}(q, \Sigma)$ are polynomially equivalent, i.e., there is a polynomial reduction from $\text{pHom}(\mathbb{B})$ to $\text{CQA}(q, \Sigma)$ and vice-versa.

**Proof.** Let $\mathbb{B}$ be a structure. We need to define a UCQ $q$ and a set of constraints $\Sigma$ such that $\text{pHom}(\mathbb{B})$ and $\text{CQA}(q, \Sigma)$ are polynomially equivalent. Before doing so, we give some intuition about the roles played by each constraint and by the query. For the sake of the explanation, we only focus on one reduction, from $\text{pHom}(\mathbb{B})$ to $\text{CQA}(q, \Sigma)$.

Suppose that we want to check for the existence of a homomorphism from a given structure $\mathbb{A}$ to the structure $\mathbb{B}$. We associate with $\mathbb{A}$ a database $D$ that contains all the relations $R^\mathbb{A}$ and a unary relation $S^\mathbb{A}$ consisting of the domain of $\mathbb{A}$. Then, we will define the constraints $\Sigma$ in such a way that each repair $E$ of $D$ encodes a partial map $f^E : \mathbb{A} \to \mathbb{B}$. Moreover, if $q$ is false in $E$, this will ensure that $f^E$ is a homomorphism and its domain is the domain of $\mathbb{A}$.

The way we encode a partial map in a repair $E$ is by introducing a unary relation $N_c$ for each $c \in B$. The fact $N_c(a) \in E$ means that $E$ holds for a repair $E$ if the map $f^E$ sends the element $a$ to an element that is not $c$. For all $b \in B$, we abbreviate the formula

$$\bigwedge \{N_c(x) : c \in B, c \neq b\}$$

by $\phi_b(x)$. Hence, $\phi_b(a)$ holds in a repair $E$ if $f^E$ maps $a$ to $b$.

Let $R$ be a relation symbol of arity $n$ and let $b = (b_1, \ldots, b_n)$ be a tuple in $B^n$. If $R(b) \notin \mathbb{B}$, we let $\psi_{R(b)}$ be the following constraint

$$\phi_{b_1}(x_1) \land \cdots \land \phi_{b_n}(x_n) \rightarrow C_R(x_1, \ldots, x_n)$$

In the databases in which $q$ (that we will define later) is false, we will think of $C_R$ as being a subset of the complement of the relation $R$. Hence, the meaning of the constraint $\psi_{R(b)}$ is as follows. If for all $1 \leq i \leq n$, the map $f^E$ maps the element $a_i$ to the element $b_i$ and $R(b_1, \ldots, b_n) \notin \mathbb{B}$, then the tuple $(a_1, \ldots, a_n)$ must belong to the complement of $R$. That is, the map $f^E$ is a homomorphism.

For all $b \in B$, we define $\chi_b$ as the constraint

$$\phi_b(x) \land S(x) \rightarrow O(x).$$

In the databases in which $q$ is false, the interpretation of $O$ is the empty set. Recall that if $\phi_b(a)$ holds in a repair $E$ of a database $D$, it means that the map $f^E$ maps $a$ to $b$. The formulas $\chi_b$s basically say that the set $S^E$ and the domain of the map associated with $E$ have an empty intersection.

Moreover, using the minimality condition of the repairs, we will show that this implies that those two sets not only have an empty intersection, but actually form a partition of $S^D$.

Next, we let $\Sigma$ be the following set of constraints

$$\{\chi_b : b \in B\} \cup \{\psi_{R(b)} : R(b) \notin B\}.$$

We define $q$ as the query

$$\exists x O(x) \lor \exists x S(x) \lor \bigvee \{\exists x (R(x) \land C_R(x)) : R \in \sigma\}.$$

Given a database $D$, the query $q$ is false in a repair $E$ iff $O$ and $S$ are empty in $E$ and for all relation symbols $R$, the intersection $R \cap C_R$ is empty in $E$. The fact that the intersection $R \cap C_R$ is empty in $E$ means that $C_R$ is a subset of the complement of $R$. 

...
The intuition behind the fact that $S$ is empty is a bit more complicated. Recall that in a repair $E$ of a database $D$ in which $q$ is false, the constraints $C_Q$ ensure that the set $S$ and the domain of the map $f$ form a partition of $S$. In that case, the fact that $S$ is empty means that all the elements of $S$ have an image, or more informally, that the domain of the map associated with $E$ is “big enough” for our purpose.

In order to prove that $pHom(E)$ and $CQA(q, \Sigma)$ are polynomially equivalent, we have to show that

(a) there is a polynomial reduction form $pHom(E)$ to $CQA(q, \Sigma)$,
(b) there is a polynomial reduction from $CQA(q, \Sigma)$ to $pHom(E)$.

Before proving that (a) and (b) holds, we proceed with the following claims.

**Claim 1.** Let $E$ be a repair of a database $D$ w.r.t. $\Sigma$. Then

$$O^D \subseteq O^E, \quad N^E_b \subseteq N^D_b, \quad S^E \subseteq S^D, \quad R^E = R^D, \quad C^E_R \subseteq C^D_R,$$

for all $b \in B$ and all relation symbols $R$. In particular, if $E \models \phi_0(a)$ for some $a$ and $b$, then $D \models \phi_0(a)$.

**Proof.** Intuitively, the claim follows from the facts that $R$ does not occur in $\Sigma$, and $N$ only occur on the left sides of logical implications and $C_R$ and $O$ only occur on the right sides of logical implications.

Formally, let $E$ be a repair of $D$ w.r.t. $\Sigma$. We define $E_0$ as the following database

$$O^{E_0} = O^D \cup O^E, \quad S^{E_0} = S^D \cap S^E, \quad N^E_b = N^D_b \cap N_b^E, \quad R^{E_0} = R^D, \quad C^{E_0}_R = C^E_R \cup C^D_R,$$

for all $b \in B$ and all relation symbols $R$. We can check that if $\Sigma$ is true in $E$, then $\Sigma$ remains true in $E_0$. Moreover, by definition of $E_0$, $D \oplus E_0 \subseteq D \oplus E$. Since $E$ is a repair of $D$ w.r.t. $\Sigma$, $E = E_0$. The claim follows.

**Claim 2.** Let $E$ be a repair of a database $D$ w.r.t. $\Sigma$. Suppose that $q$ is false in $E$. Let $f$ be a map such that for all $a \in dom(f)$ and for all $b \in B$, $D \models \phi_0(a)$ if $b = f(a)$.

We define $A^D$ as the structure with domain $S^D$ and for all relation symbols $R$ of arity $n$,

$$R^{A^D} = R^D \cap (S^D)^n.$$

Then there is a homomorphism $g : A^D \rightarrow \mathbb{B}$ such that for all $a \in dom(f)$, $g(a) = f(a)$.

**Proof.** We denote by $A^D$ the domain of $A^D$. We start by proving that

for all $a \in A^D$, there is $b \in B$ such that $\phi_0(a)$ and if $a \in dom(f)$, then $b_a = f(a)$.

Let $a$ be an element of $A^D$. Suppose for contradiction that there is no $b$ such that $\phi_0(a)$ holds in $E$. We define $E_0$ as the instance obtained by adding the tuple $S(a)$ to the database $E$.

We prove that $E_0$ is true in $E_0$. Since $\Sigma$ is true in $E$ and $E_0$ is obtained by adding $S(a)$ to $E$, $\Sigma$ can only be false in $E_0$ if

$$\phi_0(a) \land S(a) \rightarrow O(a)$$

is false in $E_0$. Since $\Sigma$ is true in $E$ and $E_0$ is obtained by adding $S(a)$ to $E$, $\Sigma$ can only be false in $E_0$ if

$$D \oplus E_0 \subseteq D \oplus E$$

Since $\Sigma$ is true in $E_0$, this contradicts the fact that $E$ is a repair and this finishes the proof of (2).

Next we prove (3). Suppose that for some $a \in dom(f)$, we have $E \models \phi_0(a)$. By Claim 1, this implies that $D \models \phi_0(a)$. By (1), this can only happen if $b_a = f(a)$. This finishes the proof of (3).

It follows from (2) and (3) that we may pick a map $g : A^D \rightarrow \mathbb{B}$ such that for all $a \in dom(f)$, $g(a) = f(a)$ and for all $a \in A^D$, $E \models \phi_0(a)$. We prove that $g$ is a homomorphism.

Suppose for contradiction that $g$ is not a homomorphism. That is, there are a relation symbol $R$ of arity $n$ and a tuple $a = (a_1, \ldots, a_n)$ such that $R(a)$ belongs to $A^D$ but $R(g(a))$ does not belong to $\mathbb{B}$. By definition of $g$,

$$\phi_{g(a_1)}(a_1) \land \cdots \land \phi_{g(a_n)}(a_n)$$

holds in $E$. Since $\Sigma$ is true in $E$ and $R(g(a))$ does not belong to $\mathbb{B}$, $\psi_R(g(a))$ given by

$$\phi_{g(a_1)}(x_1) \land \cdots \land \phi_{g(a_n)}(x_n) \rightarrow C_R(x_1, \ldots, x_n)$$

is true in $E$. Together with (4), we obtain that $C_R(a) \rightarrow E$.

Since $R(a)$ belongs to $A^D$, the tuple $R(a)$ belongs to $D$. By Claim 1, this implies that $R(a)$ belongs to $E$. Putting everything together, we have

$$C_R(a) \in E \land R(a) \in E.$$
This contradicts the fact that \( q \) is false in \( E \). \( \square \)

We start by proving (a). That is, there is a polynomial reduction form \( \text{pHom}(\mathbb{B}) \) to \( \text{CQA}(q, \Sigma) \). Let \( \mathbb{A} \) be a structure. Let \( f \) be a partial homomorphism from \( \mathbb{A} \) to \( \mathbb{B} \). We let \( D_0 \) be the following database

\[
\begin{align*}
S_{D_0} &= A, \\
O_{D_0} &= \emptyset, \\
N^D_{D_0} &= \{ a \in A : a \in \text{dom}(f), f(a) \neq b \}, \\
C^D_{R_{D_0}} &= A^n \setminus R^k, \\
R^{D_0} &= R^k,
\end{align*}
\]

where \( b \in B \) and \( R \) is a relation symbol of arity \( n \). In order to prove (a), it is sufficient to show that

\[ \text{CQA}(q, \Sigma, D_0) \models \bot \iff (\mathbb{A}, f) \in \text{pHom}(\mathbb{B}). \quad (5) \]

For the direction from left to right of (5), suppose that the consistent answer of \( q \) is false. Let \( E_0 \) be a repair of \( D_0 \) w.r.t. \( \Sigma \) in which \( q \) is false. By Claim 2, there is an homomorphism \( g_0 : \mathbb{A}^{E_0} \to \mathbb{A} \) such that for all \( a \in \text{dom}(f) \), \( g_0(a) = f(a) \).

Hence, in order to prove that \((\mathbb{A}, f) \) belongs to \( \text{pHom}(\mathbb{B}) \), it is sufficient to prove that \( \mathbb{A}^{D_0} \) is equal to \( \mathbb{A} \). This follows from the definitions of \( D_0 \) and \( \mathbb{A}^{D_0} \).

Now we show the direction from right to left of (5). Suppose that there is a homomorphism \( g_1 : \mathbb{A} \to \mathbb{B} \) such that for all \( a \in \text{dom}(f) \), \( g_1(a) = f(a) \). We define \( F_0 \) as the following database

\[
\begin{align*}
S^{F_0} &= \emptyset, \\
O^{F_0} &= \emptyset, \\
C^{F_0}_{R} &= A^n \setminus R^k, \\
R^{F_0} &= R^k, \\
N^{F_0}_b &= \{ a \in A : g_1(a) \neq b \},
\end{align*}
\]

where \( b \in B \) and \( R \) is a relation symbol of arity \( n \). It is a simple exercise to prove that \( \Sigma \) is true in \( F_0 \). Intuitively, each constraint \( \psi_b \) is true because \( S^{F_0} \) is empty. Each constraint \( \psi_{R(b)} \) (where \( R(b) \notin \mathbb{B} \)) is true because \( g_1 \) is a homomorphism and \( C^{F_0}_R \) contains the complement of \( R^k \).

Since \( \Sigma \) is true in \( F_0 \), there exists a repair \( F_1 \) of \( D_0 \) w.r.t. \( \Sigma \) such that

\[ D_0 \oplus F_1 \subseteq D_0 \oplus F_0. \quad (6) \]

We show that \( q \) is false in \( F_1 \). This will imply that \( \text{CQA}(q, \Sigma, D_0) \models \bot \).

By definition, the query \( q \) is false in \( F_1 \) iff \( O^{F_1} = \emptyset \), \( S^{F_1} = \emptyset \) and for all relation symbols \( R \), \( R^{F_1} \cap C^{F_1}_R \) is empty. Since \( O^{F_0} = O^{D_0} \), it follows from (6) that \( O^{F_1} = O^{D_0} \). That is, \( O^{F_1} = \emptyset \).

Next we prove that for all relation symbols \( R \),

\[ R^{F_1} \cap C^{F_1}_R = \emptyset. \quad (7) \]

Let \( R \) be a relation symbol of arity \( n \). Since \( R^{F_0} = R^{D_0} \) and \( C^D_R = C^{D_0}_R \), we have (6) that \( R^{F_1} = R^{D_0} \) and \( C^{F_1}_R = C^{D_0}_R \). This means that \( R^{F_1} = R^k \) and \( C^{F_1}_R = A^n \setminus R^k \). It follows that (7) holds.

In order to prove that \( q \) is false in \( F_1 \), it remains to show that \( S^{F_1} = \emptyset \). Suppose for contradiction that for some \( a \in A \), \( S(a) \) belongs to \( F_1 \). Since for all \( a \in \text{dom}(f) \), \( g_1(a) = f(a) \), it follows from the definition of \( F_0 \) and \( D_0 \) that for all \( b \in B \), \( N^{F_0}_b \subseteq N^{D_0}_b \). Together with (6), this implies that for all \( b \in B \),

\[ N^{F_0}_b \subseteq N^{F_1}_b. \quad (8) \]

It also follows from the definition of \( F_0 \) that for all \( a \in A \),

\[ \phi_{g_1(a)}(a) = \bigwedge \{ N^0(b) : b \in B, b \neq g_1(a) \} \]

holds in \( F_0 \). Together with (8), we obtain that \( \phi_{g_1(a)}(a) \) holds in \( F_1 \). Recall that we assume that \( S(a) \) belongs to \( F_1 \). Hence,

\[ \phi_{g_1(a)}(a) \land S(a) \]

holds in \( F_1 \). Since \( \chi_{g_1(a)} \) given by

\[ \phi_{g_1(a)}(x) \land S(x) \to O(x) \]

true in \( F_1 \), this implies that \( O(a) \) belongs to \( F_1 \). We proved earlier that \( O^{F_1} = \emptyset \), which is a contradiction. This finishes the proof that \( S^{F_1} = \emptyset \).

Next we prove (b). That is, there is a polynomial reduction from \( \overline{\text{CQA}}(q, \Sigma) \) to \( \text{pHom}(\mathbb{B}) \). Let \( D \) be a database. We define \( X_0 \) as the set

\[ \{ a : \text{for some } b \in B, D \models \neg N^g_b \land \phi_b(a) \}. \]

Note that for all \( a \in X_0 \), there is a unique \( b \in B \) such that \( D \models \neg N^g_b \land \phi_b(a) \). Indeed, suppose for contradiction that for some \( c \neq b \), we have \( D \models \neg N^g_c \land \phi_c(a) \) and \( D \models \neg N^g_b \land \phi_b(a) \). By definition of \( \phi_c \) and since \( b \neq c \), \( D \models \phi_c(a) \) implies that \( a \) belongs to \( N^D_b \), which is a contradiction.

For all \( a \in X_0 \), we let \( f^D(a) \) be the unique element \( b \in B \) such that

\[ D \models \neg N^g_b \land \phi_b(a). \]

Next, we define \( X \) as the set

\[ \{ a : \text{for some } b \in B, D \models \phi_b(a) \}. \]

Note that \( X_0 \subseteq X \) and if \( a \) belongs to \( X \setminus X_0 \), \( D \models \neg N^g_b(a) \) for all \( b \in B \). We will make sure that the domain of structure associated with \( D \) is a subset of
X. Intuitively, $X_0$ contains the elements $a$ that can only be mapped to $f^D(a)$, while the elements in $X \setminus X_0$ can have an arbitrary image.

We define $A^D$ as in Claim 2. That is, the domain $A^D$ of $A^D$ is the set $S^D$ and for all relation symbols $R$ of arity $n$,

$$R^A^D = R^D \cap (S^D)^n.$$ 

In order to prove (b), we exhibit a set of three conditions (the satisfaction of which can be checked in polynomial time) such that if $D$ satisfies one of those conditions, then it is clear that the consistent answer of $q$ is true; and if $D$ does not satisfy any of those conditions, then

$$(A^D, f^D) \in \text{pHom}(\mathbb{B}) \iff CQA(q, D; \Sigma) = \bot.$$ 

This will show that there is a polynomial reduction from $CQA(q, \Sigma)$ to $\text{pHom}(\mathbb{B})$.

The three conditions are given by:

(i) for some relation symbol $R$, $R^D \cap C_R^D \neq \emptyset$,

(ii) $S^D \setminus X \neq \emptyset$,

(iii) $O^D \neq \emptyset$.

We prove that

(A) if (i), (ii) or (iii) holds, then $CQA(q, D; \Sigma) = \top$,

(B) if neither (i) nor (ii) nor (iii) holds, then

$$(A^D, f^D) \in \text{pHom}(\mathbb{B}) \iff CQA(q, D; \Sigma) = \bot.$$ 

We start by showing (A). We pick a repair $G_0$ of $D$ w.r.t. $\Sigma$. Suppose that (i) holds. That is, for some relation symbol $R$ and some tuple $a$, $R(a)$ and $C_R(a)$ belong to $D$. By Claim 1, this implies that $R(a)$ and $C_R(a)$ belong to $G_0$. Hence, $q$ is true in $G_0$.

Next suppose that (ii) holds. Suppose that there exists $a$ such that $S(a)$ belongs to $D$ and $a$ does not belong to $X$. Let $G_1$ be the database obtained by adding the tuple $S(a)$ to the database $G_0$. We prove that $\Sigma$ is true in $G_1$.

Since $\Sigma$ is true in $G_0$ and $G_1$ is obtained from $G_0$ by adding $S(a)$, the only way for $\Sigma$ to be false in $G_1$ is if the constraint

$$S(a) \land \phi_b(a) \rightarrow O(a)$$

is false in $\Sigma$, for some $b \in B$. Suppose that $S(a) \land \phi_b(a)$ holds in $G_1$ for some $b \in B$. Since $\phi_b(a)$ holds in $G_1$, it follows from Claim 1 that $\phi_b(a)$ holds in $D$. Hence, $a$ belongs to $X$, which is a contradiction. This finishes the proof that $\Sigma$ is true in $G_1$.

By definition of $G_1$, $D \oplus G_1 \subseteq I \oplus G_0$. Since $G_0$ is a repair of $D$ w.r.t. $\Sigma$, this can only happen if $G_0 = G_1$. Hence, $S(a)$ belongs to $G_0$. By definition of $q$, this implies that $q$ is true in $G_0$.

Now assume that (iii) holds. That is, $O^D \neq \emptyset$. By Claim 1, this means that $O^{G_0} \neq \emptyset$. Hence, $q$ is true in $G_0$.

Next we prove (B). Suppose that neither (i) nor (ii) nor (iii) holds. For the direction from right to left of (9), suppose that the consistent answer of $q$ is false. It follows from Claim 2 that there is a homomorphism $h_0 : A^D \rightarrow \mathbb{B}$ such that for all $a \in \text{dom}(f^D)$, $h_0(a) = f^D(a)$. Hence, $(A^D, f^D)$ belongs to $\text{pHom}(\mathbb{B})$.

For the direction from left to right of (9), suppose that there is a homomorphism $h_1 : A^D \rightarrow \mathbb{B}$ such that for all $a \in \text{dom}(f^D)$, $h_1(a) = f^D(a)$. We let $H_0$ be the following database

$$R^{H_0} = R^D,$$

$$C_{R}^{H_0} = C_R^D \cup \{a \in (A^D)^n : h_1(a) \not\in R^\mathbb{B}\},$$

$$S_{H_0} = \emptyset,$$

$$O_{H_0} = \emptyset,$$

$$N_{b}^{H_0} = \{a \in A^D : h_1(a) \neq b\},$$

where $b \in B$ and $R$ is a relation symbol of arity $n$. It is easy to show that $\Sigma$ is true in $H_0$. Basically, this follows from the facts that $S^{H_0}$ is empty and that $h_1$ is a homomorphism. Since $\Sigma$ is true in $H_0$, there is a repair $H_1$ of $D$ w.r.t. $\Sigma$ such that $D \oplus H_1 \subseteq D \oplus H_0$.

We show that $q$ is false in $H_1$. We start by proving that $\exists xO(x)$ is false in $H_1$. Since (iii) does not hold, $O^D$ is empty. Together with $O^{H_0} = \emptyset$ and $D \oplus H_1 \subseteq D \oplus H_0$, we obtain that $O^{H_1} = \emptyset$.

Next we prove that $\exists xS(x)$ is false in $H_1$. Suppose that there is a fact $S(a)$ in $H_1$. We will derive a contradiction by showing that

$$\phi_{h_1(a)}(a) \land S(a) \rightarrow O(a)$$

is false in $H_1$, which is impossible as $H_1$ is a repair w.r.t. $\Sigma$.

We proved previously that $O^{H_1} = \emptyset$. We also assume that $S(a)$ belongs to $H_1$. Hence, (27) is false iff $\phi_{h_1(a)}(a)$ holds in $H_1$. Since $D \oplus H_1 \subseteq D \oplus H_0$, in order to show that $\phi_{h_1(a)}(a)$ holds in $H_1$, it is sufficient to prove that

$$H_0 \models \phi_{h_1(a)}(a) \land D \models \phi_{h_1(a)}(a).$$

The fact that $\phi_{h_1(a)}(a)$ holds in $H_0$ follows from the definition of $H_0$. Next, we prove that $\phi_{h_1(a)}(a)$ is true in $D$.

Since $S^{H_0} = \emptyset$, $S(a)$ belongs to $H_1$ and $D \oplus H_1 \subseteq D \oplus H_0$, $S(a)$ must belong to $D$. Since (ii) does not hold, $a$ belongs to $X$.

• If $a$ belongs to $X \setminus X_0$, then for all $b \in B$, $N_{b}^{H_0}(a)$ holds in $D$. In particular, $\phi_{h_1(a)}(a)$ is true in $D$. 

8
• If \( a \) belongs to \( X_0 \), then \( \phi_{f^D(a)}(a) \) holds in \( D \). Since \( f^D(a') = h_1(a') \) for all \( a' \in \text{dom}(f^D) \), this implies that \( \phi_{h_1(a)}(a) \) is true in \( D \).

This finishes the proof of (11) and the proof that \( \exists x S(x) \) is false in \( H_1 \).

Since \( O^{H_1} = \emptyset \) and \( S^{H_1} = \emptyset \), in order to show that \( q \) is false in \( H_1 \), it remains to prove that for all relation symbols \( R \),

\[ \exists x (R(x) \land C_R(x)) \]

is false in \( H_1 \). Suppose for contradiction that there exists a tuple \( a \) such that \( R(a) \) and \( C_R(a) \) belong to \( H_1 \). We prove that this implies

\[ R(a) \in D \text{ and } C_R(a) \in H_0. \tag{12} \]

By Claim 1, since \( R(a) \) belongs to \( H_1 \), \( R(a) \) belongs to \( D \). Since \( D \oplus H_1 \subseteq D \oplus H_0 \) and \( C_R^D \subseteq C_R^H \), we have \( C_R^{H_1} \subseteq C_R^{H_0} \). In particular, if \( C_R(a) \) belongs to \( H_1 \), then \( C_R(a) \) belongs to \( H_0 \). Hence, (12) holds.

Since (i) does not hold, (12) can only happen if

\[ R(a) \in D \text{ and } C_R(a) \in H_0 \text{ and } C_R(a) \notin D. \]

By definition of \( H_0 \), this means that \( h_1(a) \) does not belong to \( R^B \). Since \( h_1 \) is a homomorphism, it follows that \( R(a) \) does not belong to \( \mathbb{A}^D \). This contradicts the fact that \( R(a) \) belongs to \( D \).

IV. OTHER RELATED RESULTS

As mentioned in the introduction, we were not able to adapt the proof of Proposition III.2 to the setting of key constraints. However, if we restrict our attention to conservative CSP, we can prove a similar result.

**Theorem IV.1.** There is a key constraint \( \phi \) such that for each structure \( \mathbb{B} \), we can compute a boolean UCQ \( q \) using constants such that \( c\text{Hom}(\mathbb{B}) \) and \( CQA(q, \phi) \) are polynomially equivalent.

As a consequence, a dichotomy result for consistent query answering w.r.t. keys and UCQs with constants would provide an alternative proof for the dichotomy theorem for conservative CSP.

**Proof.** Let \( \mathbb{B} \) be a structure. Before defining \( q \) and \( \phi \), we give some intuition and for that purpose, we only focus on the reduction from \( c\text{Hom}(\mathbb{B}) \) to \( CQA(q, \phi) \).

Fix a structure \( \mathbb{A} \) and a family \( \mathcal{L} = \{ L_a \subseteq B : a \in A \} \). Suppose that we want to check whether \( (\mathbb{A}, \mathcal{L}) \in c\text{Hom}(\mathbb{B}) \).

We associate with \( (\mathbb{A}, \mathcal{L}) \) a database \( D \). The database \( D \) contains all the relations \( R^A \) and for each \( (a, b) \) such that \( b \in L_a \), \( D \) contains the fact \( F(a, b) \). In other words, the presence of \( F(a, b) \) in \( D \) means that we are allowed to map \( a \) to \( b \). The key \( \phi \) is defined in such a way that each repair \( E \) of the database encodes a map \( f^E : \mathbb{A} \to \mathbb{B} \) such that \( f^E(a) = b \) iff \( F(a, b) \in E \). So the key must express that for each \( a \), there is at most one element \( b \) such that \( F(a, b) \in E \). We let \( \phi \) be the following key

\[ F(x, u) \land F(x, v) \to u = v. \]

Next, the query \( q \) is defined such that \( q \) is false in a repair \( E \) iff then \( f^E \) is a homomorphism. For all \( (\mathbb{R}(b) \) with \( b = (b_1, \ldots, b_n) \), we let \( q_R(b) \) be the following conjunctive query

\[ \exists x_1, \ldots, x_n (R(x_1, \ldots, x_n) \land F(x_1, b_1) \land \cdots \land F(x_n, b_n)). \]

We define \( q \) by

\[ \bigwedge \{ q_R(b) : R(b) \notin \mathbb{B} \}. \]

We can show that \( q \) is false in a repair \( E \) iff \( f^E \) is a homomorphism. The proof that \( c\text{Hom}(\mathbb{B}) \) and \( CQA(q, \phi) \) are polynomially equivalent is in appendix.

If we accept to trade keys for egds, we can prove a similar result without using constants in the queries.

**Theorem IV.2.** For each structure \( \mathbb{B} \), we compute a boolean UCQ \( q \) and a set of egds \( \Sigma \) such that \( c\text{Hom}(\mathbb{B}) \) and \( CQA(q, \Sigma) \) are polynomially equivalent.

**Proof.** Let \( \mathbb{B} \) be a structure. We give some intuition about the constraints \( \Sigma \) and the query \( q \) that we introduce, and we focus first on the reduction from \( c\text{Hom}(\mathbb{B}) \) to \( CQA(q, \Sigma) \). Fix a structure \( \mathbb{A} \) and a family \( \mathcal{L} = \{ L_a \subseteq B : a \in A \} \). We want to check whether \( (\mathbb{A}, \mathcal{L}) \) belongs to \( c\text{Hom}(\mathbb{B}) \).

We define a database \( D^B \) containing the facts \( R(a) \) for all \( a \in R^A \) and the facts \( F_b(a) \), where \( a \in A \) and \( b \in L_a \). Moreover, \( D^B \) contains a special fact \( Q(\bot_0) \) where \( \bot_0 \notin A \cup B \). The idea is that in each repair \( E \), either \( Q^E \) is empty or \( E \) encodes a map \( f^E : \mathbb{A} \to \mathbb{B} \) such that \( f^E(a) = b \) iff \( F_b(a) \in E \).

If \( Q^E \neq \emptyset \), the way we ensure that \( E \) encodes a map is by introducing for all \( b, c \in B \) such that \( b \neq c \), the egd \( \phi_{b,c} \) given by

\[ F_b(x) \land F_c(x) \land Q(y) \to x = y. \]

Since \( Q^E \) is not empty and for all \( b \), \( Q^E \cap F_b = \emptyset \), the constraints \( \phi_{b,c} \) express that for each \( a \), there is at most one \( b \) such that \( F_b(a) \) belongs to \( E \). If \( Q^E \) consists of exactly one element and for all \( b \), \( Q^E \cap F_b = \emptyset \), we say that \( Q^E \) is well-behaved.

In general, if we are given an arbitrary database \( D \) (and not a database of the form \( D^B \)), there is no guarantee that in each repair \( E \) of \( D \), either \( Q^E \) is empty.
or $Q^E$ is well-behaved. We enforce this by introducing the following constraint and query. We let $\phi$ be the egd given by:

$$Q(x) \land Q(y) \rightarrow x = y.$$  

The egd $\phi$ ensures that $Q$ has at most one element in each repair. Next, we define $q_1$ as the query

$$\bigvee \{ (Q(x) \land F_b(x)) : b \in B \}.  \tag{1}$$

If a repair $E$ satisfies $\phi$ and falsifies $q_1$, then either $Q^E = \emptyset$ or $Q^E$ is well-behaved.

Next we introduce a query $q_2$ such that $q_2$ is false in a repair $E$ encoding a map $f^E$ (as defined above) iff $f^E$ is an homomorphism. For all $R(b)$ with $b = (b_1, \ldots, b_n)$, we let $q_{R(b)}$ be the following conjunctive query

$$\exists x_1, \ldots, x_n (R(x_1, \ldots, x_n) \land F_{b_1}(x_1) \land \cdots \land F_{b_n}(x_n))$$

and we let $q_2$ by the query given by

$$\bigvee \{ q_{R(b)} : R(b) \not\in \mathbb{B} \}.  \tag{2}$$

Let $E$ be a repair for which there is a map $f^E : A \rightarrow \mathbb{B}$ such that $f^E(a) = b$ iff $F_{b}(a) \in E$. We can prove that $q_2$ is true in $E$ iff $f^E$ is not a homomorphism.

Finally, we define $\Sigma$ as the set of constraints

$$\{ \phi_{b,c} : b, c \in B, b \neq c \} \cup \{ \phi \}$$

and we let $q$ be the query $q_1 \lor q_2$. To summarize our informal intuition: in the repairs $E$ of a database $D$ in which $Q^E \neq \emptyset$ and $q_1$ is false, $Q^E$ is well-behaved, $E$ encodes a map $f^E : A \rightarrow \mathbb{B}$ and $f^E$ is a homomorphism iff $q_2$ is false.

In the repairs $E$ of $D$ in which $Q^E$ is empty, we can show that $q$ is false in $E$ iff $q_2$ is false in $E$ iff each map $f : A \rightarrow \mathbb{B}$ satisfying $f_{\alpha(a)}(a) \in D$ (for all $a \in A$), is a homomorphism. A proof that $\text{chom}(\mathbb{B})$ and $\text{CQA}(q, \Sigma)$ are polynomially equivalent is in appendix.

\section{V. Conclusion}

We proved that if the dichotomy conjecture holds for consistent query answering with respect to GAV constraints and unions of conjunctive queries, then so does the dichotomy conjecture for CSP. One question left open is whether a similar result could be achieved for other classes of constraints and queries. The case of key constraints and conjunctive queries would be of particular interest, as this is the setting of the original dichotomy conjecture stated by Afrati and Kolaitis [1].

Another open question is whether we can prove the opposite implication of our main result. That is, is it true that if there is a dichotomy result for CSP, then there is a dichotomy result for consistent query answering w.r.t. given classes of constraints and queries?

\section*{Acknowledgment}

The author thanks Phokion Kolaitis for suggesting this line of research and Phokion Kolaitis and Balder ten Cate for many useful discussions during the early stages of this work. Many thanks also to Pablo Barceló and Amélie Gheerbrant for comments on earlier versions of the paper.

\section*{References}


VI. APPENDIX

A. Proof of Proposition IV.1

Proposition VI.1. There is a key constraint $\phi$ such that for each structure $\mathcal{B}$, we can compute in polynomial time a boolean UCQ $q$ using constants such that $c\text{Hom}(\mathcal{B})$ and $\overline{\text{CQA}}(q, \phi)$ are polynomially equivalent.

Note that if $\Sigma$ is a set of keys, a repair of a database $D$ w.r.t. $\Sigma$ is always a subset of $D$.  

Proof. Let $\mathcal{B}$ be a structure. We let $\phi$ be the following key

$$F(x, u) \land F(x, v) \rightarrow u = v.$$  

For all $R(b)$ with $b = (b_1, \ldots, b_n)$, we let $q_{R(b)}$ be the following conjunctive query

$$\exists x_1, \ldots, x_n(R(x_1, \ldots, x_n) \land F(x_1, b_1) \land \cdots \land F(x_n, b_n)).$$

We define $q$ by

$$\bigvee \{q_{R(b)} : R(b) \notin \mathcal{B}\}.$$  

In order to prove that $c\text{Hom}(\mathcal{B})$ and $\overline{\text{CQA}}(q, \phi)$ are polynomially equivalent, we have to show

(a) there is a polynomial reduction from $c\text{Hom}(\mathcal{B})$ to $\overline{\text{CQA}}(q, \phi)$,

(b) there is a polynomial reduction from $\overline{\text{CQA}}(q, \phi)$ to $c\text{Hom}(\mathcal{B})$.

The proof that (a) and (b) hold is based on the following claim.

CLAIM 3. Let $D$ be a database. We define $\mathcal{A}^D$ as the set

$$\{a : \text{ for some } b \in B, F(a, b) \in D\}.$$  

We define $\mathcal{A}^D$ as the structure with domain $\mathcal{A}^D$ and

$$R^D = R^D \cap (\mathcal{A}^D)^n,$$

for all $R$ of arity $n$. For all $a \in \mathcal{A}^D$, we let $L_a^D$ be the set $\{b \in B : F^D(a, b)\}$ and we let $L^D$ be the set $\{L_D^s : s \in \mathcal{A}^D\}$. Then,

$$\text{CQA}(q, D, \phi) = \bot \iff (\mathcal{A}^D, L^D) \in c\text{Hom}(\mathcal{B}).$$  

(13)

Proof. Suppose first that $\mathcal{A}^D$ is empty. Then it is clear that $(\mathcal{A}^D, L^D)$ belongs to $c\text{Hom}(\mathcal{B})$. Moreover, it can easily be seen that in case $\mathcal{A}^D$ is empty, $\text{CQA}(q, D, \phi)$ is false.

So assume that $\mathcal{A}^D$ is not empty. For the implication from left to right, suppose that $\text{CQA}(q, D, \phi) = \bot$. Hence, there is a repair $E$ of $D$ such that $E \not\models q$. First, we show that

for all $a \in \mathcal{A}^D$, there is a unique $b$ s.t. $F(a, b) \in E$.  

(14)

Since $\phi$ is true in $E$, for all $a \in \mathcal{A}^D$, there is at most one element $b$ such that $F(a, b)$ belongs to $E$.

Next suppose for contradiction that for some $a \in \mathcal{A}^D$, there is no $b$ such that $F(a, b)$ belongs to $E$. By definition of $\mathcal{A}^D$, there exists $b_0 \in B$ such that $F(a, b_0)$ belongs to $D$. We let $E_0$ be the database obtained from the database $E$ by adding the tuple $F(a, b_0)$. The key constraint $\phi$ remains true in $E_0$ and moreover, $E \subseteq E_0 \subseteq D$. This contradicts the fact that $E$ is a repair of $D$. This finishes the proof of (14).

It follows that there is a unique map $f : \mathcal{A}^D \rightarrow \mathcal{B}$ such that

$$f(a) = b : \text{ for all } a \in \mathcal{A}, F(a, f(a)) \in E.$$  

(15)

In order to show that $(\mathcal{A}^D, L^D)$ belongs to $c\text{Hom}(\mathcal{B})$, it is sufficient to prove that for all $a \in \mathcal{A}^D$, $f(a)$ belongs to $L_a^D$ and $f$ is a homomorphism.

We start by proving that for all $a \in \mathcal{A}^D$, $f(a)$ belongs to $L_a^D$. Let $a$ be an element of $\mathcal{A}^D$. Since $F(a, f(a))$ belongs to $E$, $F(a, f(a))$ also belongs to $D$. Therefore, $f(a)$ belongs to $L_a^D$.

Next we prove that $f$ is a homomorphism. Suppose for contradiction that $f$ is not a homomorphism. That is, there is a tuple $a = (a_1, \ldots, a_n)$ and a relation symbol $R$ such that $R(a)$ belongs to $\mathcal{A}^D$ and $R(f(a))$ does not belong to $\mathcal{B}$. By definition of $\mathcal{A}^D$, if $R(a)$ belongs to $\mathcal{A}^D$, then $R(a)$ belongs to $D$. Since $R$ does occur in the constraint $\phi$, this implies that $R(a)$ belongs to $E$.

Together with (22), we obtain

$$E \models R(a) \land F(a_1, f(a_1)) \land \cdots \land F(a_n, f(a_n)).$$

That is, $g_{R(f(a))}$ is true in $E$. Since $R(f(a))$ does not belong to $\mathcal{B}$, this implies that $q$ is true in $E$, which is a contradiction. This finishes the proof that $f$ is a homomorphism.

We show now the implication from right to left of (13). Assume that there is a homomorphism $g : \mathcal{A}^D \rightarrow \mathcal{B}$ such that for all $a \in \mathcal{A}^D$, $g(a) \in L_a^D$. We define $X^D$ as the set

$$\{r \notin \mathcal{A}^D : \text{ for some } s, F(r, s) \in D\}.$$  

We pick an arbitrary map $h$ with domain $X^D$ such that for all $a \in X^D$, $F^D(a, h(a))$ holds. We define the database $G$ by

$$F^G = \{((a, g(a)) : a \in \mathcal{A}^D\} \cup \{(a, h(a)) : a \in X^D\},$$

$$R^G = R^D,$$

for all relation symbols $R$. The database $G$ is a repair of $D$ with respect to $\phi$. Hence, in order to prove the implication from right to left of (13), it is sufficient to show that $q$ is false in $G$. 

11
Suppose for contradiction that $q$ is true in $G$. By definition of $q$, this means that there is a tuple $R(b) \notin B$ with $b = (b_1, \ldots, b_n)$ such that $g(R(b))$ is true in $G$. That is, there exists $a_1, \ldots, a_n$ such that

$$G \vDash R(a_1, \ldots, a_n) \land F(a_1, b_1) \land \ldots \land F(a_n, b_n).$$

We prove that this implies that

$$R(a_1, \ldots, a_n) \in A^D \text{ and } R(g(a_1), \ldots, g(a_n)) \notin B,$$

which contradicts the fact that $g$ is a homomorphism.

For all $1 \leq i \leq n$, since $b_i$ belongs to $B$ and $F(a_i, b_i)$ belongs to $G$, $a_i$ belongs to $A^D$. Since $(a_1, \ldots, a_n)$ belongs to $(A^D)^n$ and $R(a_1, \ldots, a_n)$ belongs to $G$, $R(a_1, \ldots, a_n)$ belongs to $A^D$.

In order to prove (16), it remains to show that $R(g(a_1), \ldots, g(a_n))$ does not belong to $B$. Recall that we proved that for all $1 \leq i \leq n$, $a_i$ belongs to $A^D$. By definition of $F^G$, if $a_i$ belongs to $A^D$ and $F(a_i, b_i)$ belongs to $G$, then $b_i = g(a_i)$. Recall also that $R(b)$ does not belong to $B$. Together with $b_i = g(a_i)$, this implies that $R(g(a_1), \ldots, g(a_n))$ does not belong to $B$.

We start by showing (b). That is, there is a polynomial reduction from $CQA(q, \Sigma)$ to $cHom(B)$. Let $D$ be a database. It follows from the previous claim that

$$CQA(q, D, \phi) = \bot \iff (A^D, L^D) \in cHom(B).$$

This implies that (b) holds.

Next we prove (a). That is, there is a polynomial reduction from $cHom(B)$ to $CQA(q, \Sigma)$. Let $\mathcal{A}$ be a structure and for all $a \in A$, let $L_a$ be a subset of $B$. We let $\mathcal{L}$ be the set $\{L_a : a \in A\}$. Without loss of generality, we may assume that for all $a \in A$, $L_a \neq \emptyset$. We define a database $D_0$ by

$$F^{D_0} = \{(a, b) \in A \times B : b \in L_a\},$$

$$R^{D_0} = R^\mathcal{A},$$

for all relation symbols $R$. In order to prove (a), it is sufficient to show that

$$CQA(q, D_0, \phi) = \bot \iff (A, \mathcal{L}) \in cHom(B). \quad (17)$$

It follows from the claim that

$$CQA(q, D_0, \phi) = \bot \iff (A^D, L^D_0) \in cHom(B).$$

It also follows from the definition of $D_0$ that $A^D_0 = A$ and for all $a \in A$, $L^D_a = L_a$. Together with the previous equivalence, we obtain (17).

$\square$

B. Proof of Proposition IV2

**Proposition VI.2.** For each structure $\mathcal{B}$, we can compute a boolean UCQ $q$ and a set of egds $\Sigma$ such that $cHom(\mathcal{B})$ and $CQA(q, \Sigma)$ are polynomially equivalent.

Note that if $\Sigma$ is a set of egds, a repair of a database $D$ w.r.t. $\Sigma$ is always a subset of $D$.

**Proof.** Let $\mathcal{B}$ be a structure. If $b, c \in B$, we define $\phi_{b,c}$ as the following egd:

$$F_b(x) \lor F_c(x) \land Q(y) \rightarrow x = y$$

and we let $\phi$ be the egd given by:

$$Q(x) \land Q(y) \rightarrow x = y.$$

We define $\Sigma$ as the set of following egds:

$$\{\phi_{b,c} : b, c \in B, b \neq c\} \cup \{\phi\}.$$

For all $b \in B$, we define $\psi_b$ as the query

$$\exists x(Q(x) \land F_b(x)).$$

For all $R(b)$ with $b = (b_1, \ldots, b_n)$, we let $q_{R(b)}$ be the following conjunctive query

$$\exists x_1, \ldots, x_n (R(x_1, \ldots, x_n)) \land F_{b_1}(x_1) \land \ldots \land F_{b_n}(x_n).$$

Next we define $q_1$ by

$$\bigvee \{\psi_b : b \in B\}$$

and we define $q_2$ by

$$\bigvee \{q_{R(b)} : R(b) \notin B\}.$$

We let $q$ be the query $q_1 \lor q_2$.

In order to prove that $cHom(\mathcal{B})$ and $CQA(q, \Sigma)$ are polynomially equivalent, we have to show that

(a) there is a polynomial reduction from $cHom(strB)$ to $CQA(q, \Sigma)$,

(b) there is a polynomial reduction from $CQA(q, \Sigma)$ to $cHom(\mathcal{B})$.

We now proceed with the proof that (a) and (b) hold.

We start with the following Claim.

**Claim 4.** Let $D$ be a database. If $\text{dom}(B) = \{b_1, \ldots, b_k\}$, we define $A^D$ as the set

$$F_{b_1}^D \cup \cdots \cup F_{b_k}^D.$$

We define $A^D$ as the structure with domain $A^D$ and

$$R^{\mathcal{A}} = R^\mathcal{A} \cap (A^D)^n,$$

for all relation symbols $R$ of arity $n$. Assume that $Q^D \neq \emptyset$ and $\Sigma$ is true in $D$. Then, if $q_1$ is false in $D$, there is
Proof. It follows from the definition of $\mathbb{A}^D$ that there is a map $f : \mathbb{A}^D \to \mathbb{B}$ such that for all $a \in A^D$, $F_b(a)$ belongs to $D$, where $b = f^D(a)$. Moreover, $f^D$ is a homomorphism \iff $D \neq q_2$.

We abbreviate $f^D$ by $f$. The formula $q_2$ is true in $D$ iff there is a tuple $R(b) \notin \mathbb{B}$ with $b = (b_1, \ldots, b_n)$ such that $q_{R(b)}$ is true in $D$. The query $q_{R(b)}$ is true in $D$ iff there exists a tuple $a = (a_1, \ldots, a_n)$ such that $D \models R(a_1, \ldots, a_n) \land F_{b_1}(a_1) \land \cdots \land F_{b_n}(a_n)$. Observe that since $F_{b_i}(a_i) \in D$, $a_i$ belongs to $A^D$ for all $1 \leq i \leq n$. Hence, by definition of $\mathbb{A}^D$, $R(a_1, \ldots, a_n) \in D$ iff $R(a_1, \ldots, a_n) \in \mathbb{A}^D$. Moreover, it follows from the unicity of $f$ that for all $1 \leq i \leq n$, $F_{b_i}(a_i) \in D$ iff $b_i = f(a_i)$.

Putting everything together, we obtain that $q_2$ is true in $D$ iff there are a tuple $R(b) \notin \mathbb{B}$ and a tuple $(a_1, \ldots, a_n)$ such that $R(a_1, \ldots, a_n) \in \mathbb{A}^D, f(a_1) = b_1, \ldots, f(a_n) = b_n$, where $b = (b_1, \ldots, b_n)$. This happens iff $f$ is not a homomorphism.

Claim 5. Let $E$ be a repair of a database $D$ w.r.t. $\Sigma$ such that $Q^E = \emptyset$. Then, for all $b \in B$ and for all relation symbols $R$, $R^E = R^D$ and $F_b^E = F_b^D$.

Proof. Let $G$ be the following database

\[
\begin{align*}
Q^G &= \emptyset, \\
R^G &= R^D, \\
F_b^G &= F_b^D,
\end{align*}
\]

where $b \in B$ and $R$ is a relation symbol. Since $Q^E = \emptyset$, we have $E \subseteq G \subseteq D$. Moreover, since $Q^G = \emptyset$, it is easy to check that $\Sigma$ is true in $G$. As $E$ is a repair of $D$ w.r.t. $\Sigma$, this can only happen if $E = G$. The claim follows.

We start by proving (a). That is, there is a polynomial reduction from $cHom(\mathbb{B})$ to $CQA(q, \Sigma)$. Let $\mathbb{A}$ be a structure and for all $a \in A$, let $L_a$ be a subset of $A$. Without loss of generality, we may assume that for all $a \in A$, $L_a \neq \emptyset$. We let $L$ be the set \{ $L_a : a \in A$ \}. We define a database $D_0$ by

\[
\begin{align*}
Q_{D_0} &= \{ \bot_0 \}, \\
F_{b_0} &= \{ a \in A : b \in L_a \}, \\
R_{D_0} &= R^L,
\end{align*}
\]

where $b \in B$ and $R$ is a relation symbol. In order to show (a), it is sufficient to prove that

\[
CQA(q, D_0, \Sigma) = \bot \iff (\mathbb{A}, L) \in cHom(\mathbb{B}).
\]

Suppose that $CQA(q, D_0, \Sigma) = \bot$. Hence, there is a repair $E_0$ of $D_0$ such that $E_0 \not\subseteq q$. We make the following case distinction.

- Suppose that $Q^E_0 = \emptyset$. Let $f : \mathbb{A} \to \mathbb{B}$ be an arbitrary map such that for all $a \in A$, $f(a) \in L_a$. We prove that $f$ is a homomorphism. Suppose for contradiction that $f$ is not a homomorphism. Hence, there are tuples $b = (b_1, \ldots, b_n)$ and $a = (a_1, \ldots, a_n)$ such that $a \in R^b, b \notin R^B$ and $f(a_i) = b_i$, for all $1 \leq i \leq n$. As $f(a_i) = b_i$, $b_i$ belongs to $L_a$. By definition of $F_{b_0}$, this implies that $F_{b_0}(a_i) \in D_0$. Together with the facts $a \in R^b$ and $R_{D_0} = R^L$, we obtain

\[
D_0 \models R(a_1, \ldots, a_n) \land F_{b_1}(a_1) \land \cdots \land F_{b_n}(a_n).
\]

Since $Q^E_0 = \emptyset$, by Claim 5, this implies that $E_0 \models R(a_1, \ldots, a_n) \land F_{b_1}(a_1) \land \cdots \land F_{b_n}(a_n)$. That is, $q_{R(b)}$ is true in $E_0$. Since $b \notin R^B$, this implies that $q$ is true in $E_0$, which is a contradiction.

- Next suppose that $Q^E_0 \neq \emptyset$. It follows from Claim 4 that there is a homomorphism $f^{E_0} : \mathbb{A}^{E_0} \to \mathbb{B}$ such that for all $a \in A^{E_0}$, $F_b(a)$ belongs to $E_0$, where $b = f^{E_0}(a)$. Hence, in order to prove that $(\mathbb{A}, L) \in cHom(\mathbb{B})$, it is sufficient to show that $\mathbb{A}^{E_0} = \mathbb{A}$ and for all $a \in A$, $f^{E_0}(a) \in L_a$.

We prove that for all $a \in a$, $f^{E_0}(a)$ belongs to $L_a$. Let $a$ be an element of $A$ and let $b$ be the image of $f^{E_0}(a)$. Since $F_b(a)$ belongs to $E_0$ and $E_0$ is a subset of $D_0$, $F_b(a)$ belongs to $D_0$. By definition of $F_{D_0}$, $b$ belongs to $L_a$. 

13
Next we show that $A_{E_0}^A = A$. By definition of $A_{E_0}^A$, this is equivalent to show that $A_{E_0}^A$ is equal to $A$. Since $E_0$ is a subset of $D_0$, it is immediate that $A_{E_0}^A$ is a subset of $A_{D_0}^A$. Moreover, since for all $b$, $F_{b_{D_0}}$ is a subset of $A$, it is clear that $A_{D_0}^A$ is a subset of $A$. Hence, $A_{E_0}^A$ is a subset of $A$.

Now suppose for contradiction that $A_{E_0}^A$ is a proper subset of $A$. That is, for some $a \in A$, there is no $b$ such that $F_{b_{D_0}}(a)$ belongs to $E_0$. Since $L_a \not= \emptyset$, there exists $b_0 \in A$ such that $F_{b_0}(a)$ belongs to $D_0$. We let $E_1$ be the database obtained from the database $E_0$ by adding the tuple $F_{b_0}(a)$. The constraint $\Sigma$ remain true in $E_1$ and moreover, $E_0 \subseteq E_1 \subseteq D_0$. This contradicts the fact that $E_0$ is a repair of $D_0$.

This finishes the proof that $A_{E_0}^A = A$.

We show now the implication from right to left of (19). Assume that there is a homomorphism $g: A \rightarrow B$ such that for all $a \in A$, $g(a) \in L_a$. We have to find a repair $G_0$ of $D_0$ with respect to $\Sigma$ in which $q$ is false. We define the database $G_0$ by

\[
Q^{G_0} = \{ \bot \}, \quad F_{b_{G_0}} = \{ a \in A : g(a) = b \}, \quad R^{G_0} = R^{D_0},
\]

where $b \in B$ and $R$ is a relation symbol. The instance $G_0$ is a repair of $D_0$ with respect to $\Sigma$. We show that $q$ is false in $G_0$.

Since for all $b$, $Q^{G_0} \cap F_{b_{G_0}}$ is empty, $q_1$ is false in $G_0$. Next we prove that $q_2$ is false in $G_0$. Since $q_1$ is false in $G_0$ and $Q^{G_0} \not= \emptyset$, it follows from Claim 4, that in order to prove that $q_2$ is false in $G_0$, it is enough to show that $f^{G_0}$ is a homomorphism. As $g$ is a homomorphism, it is sufficient to prove that $f^{G_0} = g$. Recall that $f^{G_0}$ is the unique map such that for all $a \in A^{G_0}$, $F_{g(a)}(a)$ belongs to $G_0$, where $b = f^{G_0}(a)$. By definition of $G_0$,

\[ F_{g(a)}(a) \in G_0, \text{ for all } a \in A. \]

Hence, $f^{G_0} = g$ and this finishes the proof that $q$ is false in $G_0$.

Next we prove (b). That is, there is a polynomial reduction from $\overline{CQA}(q, \Sigma)$ to $cHom(\mathbb{B})$. Let $D_1$ be a database. We let $A_{D_1}$ and $\{ L_a^{D_1} : a \in A_{D_1} \}$ be as defined in Claim 4. That is,

\[
A_{D_1} = \bigcup \{ F_b : b \in B \}, \quad R^{A_{D_1}} = R^{D_1} \cap (A_{D_1})^n, \quad L_a = \{ b \in B : F_b(a) \in D_1 \},
\]

where $R$ is a relation symbol of arity $n$ and $a \in A$. In order to make notation easier, we abbreviate $A_{D_1}$ by $A^1$, $A_{D_1}^{L_a}$ by $A^1_a$, and $L_a^{D_1}$ by $L_{a}^1$. We let $L_1$ be the set $\{ L_a^1 : a \in A^1 \}$.

In the proof we make use of the notion of $Q$-compatibility that we define as follows. We say that an element $x$ is $Q$-compatible if $x$ belongs to $Q^{D_1}$ and for all $a \in A^1 \setminus \{ x \}$, there is a unique $b$ such that $F_b(a)$ belongs to $D_1$. The intuition behind the notion of $Q$-compatibility is as follows: a database $D$ admits a $Q$-compatible element iff in each repair $E$, $Q^E$ is not empty. We prove this property later.

It is clear that $CQA(q, D_1, \Sigma) = \bot$ iff
\begin{itemize}
  \item either there is a repair $H_1$ such that $Q^{H_1} = \emptyset$ and $H_1 \not= q$
  \item or there is a repair $H_2$ such that $Q^{H_2} \not= \emptyset$ and $H_2 \not= q$
\end{itemize}

We will show that
\begin{itemize}
  \item[(A)] there is a repair $H_1$ such that $H_1 \not= q$ and $Q^{H_1} = \emptyset$ iff $D_1 \not= q_2$ and there is no $x$ $Q$-compatible,
  \item[(B)] there is a repair $H_2$ such that $H_2 \not= q$ and $Q^{H_2} \not= \emptyset$ iff $(A^1, L_1) \in cHom(\mathbb{B})$ and
\end{itemize}

\[ D_1 \models \exists x (Q(x) \land \neg F_{b_{0}}(x) \land \cdots \land \neg F_{b_{k}}(x)). \] (20)

Recall that $\{ b_1, \ldots, b_k \}$ is the domain of $\mathbb{B}$.

Provided that (A) and (B) hold, we obtain that $CQA(q, D_1, \Sigma) = \bot$ iff
\begin{itemize}
  \item either $D_1 \not= q_2$ and there is no $x$ $Q$-compatible,
  \item or $(A^1, L_1) \in cHom(\mathbb{B})$ and (20) holds.
\end{itemize}

Since checking for the existence of a $Q$-compatible element, the satisfaction of (20) and the fact that $D \not= q_2$ can be performed in polynomial time, this equivalence shows that there is a polynomial reduction from $\overline{CQA}(q, \Sigma)$ to $cHom(\mathbb{B})$.

Hence, in order to prove (b), it is sufficient to show that (A) and (B) hold. We start by proving that (A) holds. We do so by showing that
\begin{itemize}
  \item[(i)] there is a repair $H_1$ such that $Q^{H_1} = \emptyset$ iff there is no $Q$-compatible element,
  \item[(ii)] if $H_1$ is a repair such that $Q^{H_1} = \emptyset$, $H_1 \not= q$ iff $D_1 \not= q_2$.
\end{itemize}

First we prove (ii). Let $H_1$ be a repair such that $Q^{H_1} = \emptyset$. For the implication from right to left, suppose that $q_2$ is false in $D_1$. Since $H_1 \subseteq D_1$, it is clear that $q_2$ is false in $H_1$. Moreover, as $Q^{H_1} = \emptyset$, $q_1$ is also false in $H_1$. Hence, $H_1 \not= q$.

Next we prove the implication from left to right of (ii), suppose that $q$ is false in $H_1$. Since $Q^{H_1} = \emptyset$, it follows from Claim 5 that
\[
R^{H_1} = R^{D_1} \text{ and } F_{b_{D_1}}^{H_1} = F_{b_{D_1}}^{D_1}, \] (21)
for all relation symbols \( R \) and for all \( b \in B \). Since \( q \) is false in \( H_1 \), \( q_2 \) is false in \( H_1 \). Observe that the only symbols occurring in \( q_2 \) are the relation symbols \( R \) and \( F_b \). Together with (21) and the fact that \( q_2 \) is false in \( H_1 \), this means that \( q_2 \) is false in \( D_1 \).

Now we prove (i). That is, there is a repair \( H_1 \) such that \( Q^{H_1} = \emptyset \) iff there is no \( Q \)-compatible element. For the direction from left to right, let \( H_1 \) be a repair such that \( Q^{H_1} = \emptyset \) and suppose for contradiction that there is an element \( \bot_1 \) that is \( Q \)-compatible. We define \( I_1 \) as the database obtained by adding the fact \( Q(\bot_1) \) to \( H_1 \).

Using the fact that \( \bot_1 \) is \( Q \)-compatible, we check that the constraints \( \Sigma \) remain true in \( I_1 \). Since \( \bot_1 \) is the only element in \( Q^{I_1} \), the egd \( \phi \) is true in \( I_1 \). Next let \( b, c \in B \) be such that \( b \neq c \). We have to prove that \( \phi_{b,c} \) is true in \( I_1 \). Suppose that

\[
I_1 \models F_b(a) \land F_c(a) \land Q(a').
\]  

We have to prove that \( a = a' \). Suppose for contradiction that \( a \neq a' \). By definition of \( I_1 \), \( I_1 \models Q(a') \) implies \( a' = \bot_1 \). Hence, \( a \neq \bot_1 \). Together with the fact that \( \bot_1 \) is \( Q \)-compatible, this implies that there is a unique element \( b_0 \in B \) such that \( F_{b_0}(a) \in D_1 \). Since \( I_1 \) is a subset of \( D_1 \), (22) implies that \( F_{b_0}(a) \) and \( F_{b_0}(a) \) belong to \( D_1 \), which contradicts the unicity of \( b_0 \). This finishes the proof that the constraints of \( \Sigma \) are true in \( I_1 \).

Moreover, since \( Q^{H_1} = \emptyset \), we have \( H_1 \subseteq I_1 \subseteq D_1 \). This contradicts the fact that \( H_1 \) is a repair of \( D_1 \) w.r.t. \( \Sigma \).

Next we prove the implication from right to left of (ii). Suppose that there is no element \( Q \)-compatible. We define \( H_1 \) as the following database:

\[
Q^{H_1} = \emptyset, \\
R^{H_1} = R^{D_1}, \\
F_b^{H_1} = F_b^{D_1},
\]

where \( b \in B \) and \( R \) is a relation symbol. We show that \( H_1 \) is a repair such that \( Q^{H_1} = \emptyset \). Suppose for contradiction that \( H_1 \) is not a repair. Since \( H_1 \models \Sigma \), there is a repair \( I_2 \) such that \( H_1 \subseteq I_2 \subseteq D_1 \). By definition of \( H_1 \), this can only happen if there is a fact of the form \( Q(\bot_2) \) in \( I_2 \).

We prove that \( \bot_2 \) is \( Q \)-compatible, which is a contradiction. Recall that \( \bot_2 \) is \( Q \)-compatible iff \( \bot_2 \) belongs to \( Q^{D_1} \) and for all \( a \in A^1 \setminus \{\bot_2\} \), there is a unique \( b \) such that \( F_b(a) \in D_1 \).

By definition, \( Q(\bot_2) \) belongs to \( I_2 \) and since \( I_2 \subseteq D_1 \), \( Q(\bot_2) \) belong to \( D_1 \). Next we prove that for all \( a \in A^1 \setminus \{\bot_2\} \), there is a unique \( b_0 \) such that \( F_{b_0}(a) \).

Take \( a \in A^1 \setminus \{\bot_2\} \). For all \( c, d \in B \) such that \( c \neq d \),

\[
F_c(a) \land F_d(a) \land Q(\bot_2) \rightarrow \bot_2 = a
\]

holds in \( I_2 \). Since \( a \neq \bot_2 \), this implies that there is a unique \( b_0 \) such that \( F_{b_0}(a) \in I_2 \). Together with \( F_{b_0}(a) \) and \( F_{b_0}(a) \) for all \( c \in B \), this means that there is a unique \( b_0 \) such that \( F_{b_0}(a) \in D_1 \). This finishes the proof that (A) holds.

We show now that (B) holds. That is, there is a repair \( H_2 \) such that \( H_2 \neq q \) and \( Q^{H_2} \neq \emptyset \) iff \( (\mathcal{A}^1, \mathcal{L}^1) \in \text{chom}(\mathcal{B}) \) and

\[
D_1 \models \exists x(Q(x) \land \neg F_{b_0}(x) \land \cdots \land \neg F_{b_0}(x)).
\]  

First we prove the direction from left to right of (B). Suppose that there is a repair \( H_2 \) such that \( H_2 \neq q \) and \( Q^{H_2} \neq \emptyset \). Let \( \bot_3 \) be an element in \( Q^{H_2} \). We start by showing that \( (\mathcal{A}^1, \mathcal{L}^1) \in \text{chom}(\mathcal{B}) \). Since \( Q^{H_2} \neq \emptyset \) and \( q \) is false in \( H_2 \), it follows from Claim 4 that there is a homomorphism \( f^{H_2} : (\mathcal{A}^1, \mathcal{L}^1) \rightarrow \mathcal{B} \) such that \( F_{b_0}(a) \) belongs to \( H_2 \), if \( a \in A^{H_2} \) and \( b = f^{H_2}(a) \). In order to prove that \( (\mathcal{A}^1, \mathcal{L}^1) \in \text{chom}(\mathcal{B}) \), it is enough to show that

\[
\mathcal{A}^{H_2} = \mathcal{A}^{D_1},
\]

and for all \( a \in A^{H_2} \), \( f^{H_2}(a) \in L_a^3 \).

By definitions of \( \mathcal{A}^{H_2} \) and \( \mathcal{A}^{D_1} \), (24) holds iff \( A^{H_2} = A^{D_1} \). Since \( \Sigma \) is a set of egds and \( H_2 \) is repair of \( D_1 \), \( H_2 \) is a subset of \( D_1 \). Hence, \( A^{H_2} \subseteq A^{D_1} \).

Suppose for contradiction that \( A^{H_2} \) is a proper subset of \( A^{D_1} \). That is, there is a fact \( F_{b_0}(a) \) in \( D_1 \) and there is no \( c \in B \) such that \( F_c(a) \in H_2 \). We let \( H_3 \) be the database obtained from \( H_2 \) by adding the tuple \( F_{b_0}(a) \). Since there is no \( c \in B \) such that \( F_c(a) \in H_2, \Sigma \) remains true in \( H_3 \). Moreover,

\[
H_2 \subseteq H_3 \subseteq D_1.
\]

This contradicts the fact that \( H_2 \) is a repair and proves (24).

Next, we show (25). Let \( a \) be an element of \( A^{H_2} \). By definition of \( f^{H_2} \), if \( b = f^{H_2}(a) \), \( F_b(a) \) belongs to \( H_2 \). Since \( H_2 \) is a subset of \( D_1 \), this implies that \( F_b(a) \) belongs to \( D_1 \). By definition of \( L_a^3 \), this implies that \( b \) belongs to \( L_a^3 \). This finishes the proof that \( (\mathcal{A}^1, \mathcal{L}^1) \in \text{chom}(\mathcal{B}) \).

Next, we show (23) by proving that

\[
D_1 \models Q(\bot_3) \land \neg F_{b_0}(\bot_3) \land \cdots \land \neg F_{b_0}(\bot_3).
\]  

Since \( \bot_3 \) belongs to \( Q^{H_2} \), \( \bot_3 \) also belongs to \( Q^{D_1} \). Suppose for contradiction that for some \( b \) in \( B \), \( F_b(\bot_3) \) belongs to \( D_1 \). Hence, \( \bot_3 \) belongs to \( A^{D_1} \). By (24), \( \bot_3 \) belongs to \( A^{H_2} \). That is, for some \( c \in B \), \( F_c(\bot_3) \)
belongs to $H_2$. Since $\bot_3$ belongs to $Q^{H_2}$, this implies that
\[ \exists x (Q(x) \land F_c(x)) \]
is true in $H_2$. This is not possible, as $q$ is false in $H_2$. This finishes the proof of (26) and the proof of the implication from left to right of (B).

Now we prove the implication from right to left of (B). Suppose that $(\mathcal{A}^1, \mathcal{L}^1)$ belongs to $cHom(\mathcal{B})$ and that there is an element $\bot_4$ such that $D_1 \models Q(\bot_4) \land \neg F_{b_1}(\bot_4) \land \cdots \land \neg F_{b_k}(\bot_4)$. (27)

We pick a homomorphism $g_1 : \mathcal{A}^1 \to \mathcal{B}$ such that for all $a \in A^1$, $g_1(a)$ belongs to $L^a_1$. We define $J_1$ as the following subset of $D_1$:

\[
\begin{align*}
Q^{J_1} &= \{ \bot_4 \}, \\
R^{J_1} &= R^{D_1}, \\
F_{b_i}^{J_1} &= \{ a \in A^1 : g_1(a) = b \},
\end{align*}
\]

where $R$ is a relation symbol and $b \in B$. The database $J_1$ is a repair of $D_1$ w.r.t. $\Sigma$. Next we prove that $J_1 \not\models q$.

First, we prove that $q_1$ is false in $J_1$. Suppose for contradiction that $q_1$ is true in $J_1$. Then there are $b \in B$ and $a \in A^1$ such that

\[ J_1 \models Q(a) \land F_b(a). \]

Since $J_1$ is a subset of $D_1$, $D_1 \models Q(a) \land F_b(a)$, which contradicts (27).

Next we prove that $q_2$ is false in $J_1$. Since $Q^{J_1} \neq \emptyset$ and $q_1$ is false in $J_1$, it follows from Claim 4 that there is a unique map $f_{J_1} : \mathcal{A}^{J_1} \to \mathcal{B}$ such that for all $a \in A^{J_1}$ with $b = f_{J_1}(a)$, $F_b(a)$ belongs to $J_1$. By definition of $J_1$, this implies that $f_{J_1} = g_1$. Moreover, we obtain from Claim 4 that

\[ f_{J_1} \text{ is a homomorphism } \iff J_1 \not\models q_2. \]

Since $g_1$ is a homomorphism and $f_{J_1} = g_1$, $q_2$ is false in $H_1$. \qed