Combining Existential Rules with the Power of CP-Theories

Tommaso Di Noia
Dip. di Elettrotecnica ed Elettronica
Politecnico di Bari, Italy
tommaso.dinoia@poliba.it

Thomas Lukasiewicz
Department of Computer Science
University of Oxford, UK
thomas.lukasiewicz@cs.ox.ac.uk

Maria Vanina Martinez **Gerardo I. Simari**
Dept. of Computer Science and Engineering
Univ. Nacional del Sur and CONICET, Argentina
{mvm,gis}@cs.uns.edu.ar

Oana Tifrea-Marciuska
Department of Computer Science
University of Oxford, UK
oana.tifrea@cs.ox.ac.uk

Abstract

The tastes of a user can be represented in a natural way by using qualitative preferences. In this paper, we explore how ontological knowledge expressed via existential rules can be combined with CP-theories to (i) represent qualitative preferences along with domain knowledge, and (ii) perform preference-based answering of conjunctive queries (CQs). We call these combinations ontological CP-theories (OCP-theories). We define skyline and k-rank answers to CQs based on the user's preferences encoded in an OCP-theory, and provide an algorithm for computing them. We also provide precise complexity (including data tractability) results for deciding consistency, dominance, and CQ skyline membership for OCP-theories.

1 Introduction

From its inception, the Web has been centered around the idea of linking information to make it more accessible and useful for the users. Its fast growth, however, made it hard for end users to satisfy their information needs due to the issues related to information overload. Too much information is currently available on the Web, and new personalized information filtering techniques are needed that are able to retrieve documents/data/resources that best fit users' interests and preferences, i.e., that best match users' profiles.

Moreover, the Web has recently evolved at an increasing pace towards the so-called Web 3.0, where classical linked information lives together with ontological knowledge and social interactions of users. While the former may allow for more precise and rich results in search and query answering tasks, the latter can be used to enrich the user profile, and it paves the way to more sophisticated personalized access to information. This requires new techniques for ranking search results, fully exploiting ontological and user-centered data, i.e., user preferences. Such techniques are also very useful for computer applications (e.g., preferences over different data sources) and agents (e.g., planning with preferences).

The study of preferences has been carried out in many different areas, such as philosophy, economics, and choice theory. They can be modeled in both qualitative and quantitative ways, where quantitative preferences are associated with a number representing their worth, while qualitative preferences are related to each other via pairwise comparisons.

In this paper, we focus on ranking answers for conjunctive queries (CQs) to Datalog+/– ontologies, based on user preferences encoded in CP-theories [Wilson, 2004], which generalize CP-nets [Boutilier *et al.*, 1999]. We chose CP-theories for their natural, concise, and flexible representation of qualitative preferences – along with their subclasses, they are widely used to represent and reason with qualitative preferences. We chose the Datalog+/– ontology language for its expressive power and intuitive nature – its fragments generalize many other ontology languages, such as the *DL-Lite* family [Calvanese *et al.*, 2007] of description logics (DLs) [Calì *et al.*, 2012a]. The integration between the ontology and the CP-theory is tight: on the one hand, CP-theory outcomes are constrained by the ontology, and, on the other hand, they directly inform how answers to CQs are ranked.

The main contributions of this paper are briefly as follows:

- We introduce ontological CP-theories (OCP-theories), which combine Datalog+/– with CP-theories, modeling preferences over ground atoms in Datalog+/– ontologies.
- We define skyline and k-rank answers for CQs to OCP-theories. We also provide an algorithm for computing such answers based on the preferences encoded in an OCP-theory.
- We analyze the computational complexity of deciding consistency, dominance, and CQ skyline membership for OCP-theories, providing (generic and concrete) precise complexity results for different types of combined complexity.
- We also provide several tractability results in the data complexity for the case where query answering in the underlying classical ontology is tractable in the data complexity.

2 Preliminaries

Before introducing OCP-theories, we recall the basics on Datalog+/– and classical CP-theories.

2.1 Datalog+/–

We now recall the basics of Datalog+/– [Calì *et al.*, 2012a], namely, relational databases, dependencies, (Boolean) conjunctive queries ((B)CQs), and ontologies.

General. Consider the following sets: a set Δ of *constants*, a set Δ_N of *labeled nulls*, and a set $\mathcal{V}$ of regular *variables*. A *term* t is a constant, null, or variable. An *atom* has the form $p(t_1, \ldots, t_n)$, where p is an n-ary predicate, and $t_1, \ldots, t_n$ are terms. Conjunctions of atoms are often identified with the sets of their atoms. An *instance* I is a (possibly infinite) set of atoms $p(\mathbf{t})$, where $\mathbf{t}$ is a tuple of constants and nulls. A *database* D is a finite instance that contains only constants. A *homomorphism* is a substitution $h : \Delta \cup \Delta_N \cup \mathcal{V} \to \Delta \cup \Delta_N \cup \mathcal{V}$ that is the identity on Δ. We assume the reader is familiar with *conjunctive queries (CQs)*. The answer to a CQ q over an instance I is denoted $q(I)$. A Boolean CQ (BCQ) q has a positive answer over I, denoted $I \models q$, if $q(I) \neq \emptyset$.

Dependencies. A *tuple-generating dependency (TGD)* or *existential rule* σ is a first-order formula $\forall \mathbf{X}\, \varphi(\mathbf{X}) \to \exists \mathbf{Y}\, p(\mathbf{X}, \mathbf{Y})$, where $\mathbf{X} \cup \mathbf{Y} \subset \mathcal{V}$, $\varphi(\mathbf{X})$ is a conjunction of atoms, and $p(\mathbf{X}, \mathbf{Y})$ is an atom; $\varphi(\mathbf{X})$ is the *body* of σ, denoted $body(\sigma)$, while $p(\mathbf{X}, \mathbf{Y})$ is the *head* of σ, denoted $head(\sigma)$. For clarity, we consider single-atom-head TGDs; however, our results can be extended to TGDs with a conjunction of atoms in the head. An instance I satisfies σ, written $I \models \sigma$, if the following holds: whenever there exists a homomorphism h such that $h(\varphi(\mathbf{X})) \subseteq I$, then there exists $h' \supseteq h_{|\mathbf{X}}$, where $h_{|\mathbf{X}}$ is the restriction of h on $\mathbf{X}$, such that $h'(p(\mathbf{X}, \mathbf{Y})) \in I$. A *negative constraint (NC)* ν is a first-order formula of the form $\forall \mathbf{X}\, \varphi(\mathbf{X}) \to \bot$, where $\mathbf{X} \subset \mathcal{V}$, $\varphi(\mathbf{X})$ is a conjunction of atoms and is called the *body* of ν, denoted $body(\nu)$, and $\bot$ denotes the truth constant *false*. An instance I satisfies ν, written $I \models \nu$, if there is no homomorphism h such that $h(\varphi(\mathbf{X})) \subseteq I$. Given a set Σ of TGDs and NCs, I satisfies Σ, written $I \models \Sigma$, if I satisfies each TGD and NC of Σ. For brevity, we omit the universal quantifiers in front of TGDs and NCs, and use a comma instead of $\wedge$ to denote conjunction. As another component, Datalog+/– allows for special types of *equality-generating dependencies (EGDs)*. Since they can also be modeled via NCs, we omit them here and refer to [Calì *et al.*, 2012a] for their details.

Conjunctive Query Answering. Given a database D and a set Σ of TGDs and NCs, the answers we consider are those that are true in *all* models of D and Σ. Formally, the *models* of D and Σ, denoted $mods(D, \Sigma)$, is the set of instances $\{I \mid I \supseteq D,\, I \models \Sigma\}$. The *answer* to a CQ q w.r.t. D and Σ is the set of tuples $ans(q, D, \Sigma) = \bigcap_{I \in mods(D,\Sigma)} \{\mathbf{t} \mid \mathbf{t} \in q(I)\}$. The answer to a BCQ q is *positive*, denoted $D \cup \Sigma \models q$, if $ans(q, D, \Sigma) \neq \emptyset$. The problem of *CQ answering* is defined as follows: given a database D, a set Σ of TGDs and NCs, a CQ q, and a tuple of constants $\mathbf{t}$, decide whether $\mathbf{t} \in ans(q, D, \Sigma)$. Following Vardi's taxonomy (1982), the *combined complexity* of CQ answering is calculated by considering all the components, i.e., the database, the set of dependencies, and the query, as part of the input. The *data complexity* is calculated by only considering the database as part of the input. The *bounded-arity combined complexity* (or *ba-combined complexity*) is calculated by assuming that the arity of the underlying schema is bounded by an integer constant. Notice that in the context of DLs, whenever we refer to the combined complexity in fact we refer to the *ba*-combined complexity, as the arity of the underlying schema is at most two. The *fixed-program combined complexity* (or *fp-combined complexity*) is calculated by considering the set of TGDs and NCs as fixed.

hotel

	id	city	conn	class
$\mathbf{t}_1$	h_1	rome	*c*	*e*
$\mathbf{t}_2$	h_2	rome	*w*	*l*
$\mathbf{t}_3$	h_3	rome	*c*	*e*

book

	id	user	price
$\mathbf{t}_4$	h_1	*b*	30
$\mathbf{t}_5$	h_2	*b*	40
$\mathbf{t}_6$	h_3	*j*	35

review

	id	user	feedback
$\mathbf{t}_7$	h_1	*b*	*n*
$\mathbf{t}_8$	h_2	*b*	*p*
$\mathbf{t}_9$	h_3	*j*	*p*

reviewer

	user	age
$\mathbf{t}_{10}$	*b*	20
$\mathbf{t}_{11}$	*j*	30

friend

	user	user
$\mathbf{t}_{12}$	*b*	*a*
$\mathbf{t}_{13}$	*j*	*a*

Figure 1: Database D.

Datalog+/– Ontologies. A *Datalog+/– ontology* $O = (D, \Sigma)$, where $\Sigma = \Sigma_T \cup \Sigma_{NC}$, consists of a finite database D over Δ, a finite set Σ_T of TGDs, and a finite set Σ_{NC} of NCs.

Example 1 A simple Datalog+/– ontology $O = (D, \Sigma)$ is as follows: the database D is shown in Fig. 1, modeling the domain of an online hotel booking service, integrated with a social network. The *hotel* relation contains information about the hotel, such as its location, if it has WiFi or cable connection, and its class; the *review* relation is used to register the users' feedback after their stay. Intuitively, D encodes, e.g., that *h1*, *h2*, and *h3* are three hotels, and p (positive) and n (negative) represent users' feedback, *w* (WiFi) and *c* (cable) are available internet connections, and hotels are either *e* (economy) or *l* (luxury) class. In our example, Alice (a) is planning her holidays and looking for a hotel in Rome. She is connected to John (j) and Bob (b) in the social network. Both John and Bob are active users, as they already booked and reviewed some hotels. The set Σ encodes relationships among predicates, e.g., that a reviewer always has friends (if s/he has no friends, then he is not allowed to review).

$$
\begin{aligned}
\Sigma = \{&\mathit{review}(I,U,F) \to \exists C,O,S\ \mathit{hotel}(I,C,O,S),\\
&\mathit{review}(I,U,F) \to \exists A\ \mathit{reviewer}(U,A),\\
&\mathit{review}(I,U,F) \to \exists P\ \mathit{book}(I,U,P),\\
&\mathit{reviewer}(U,A) \to \exists F\ \mathit{friend}(U,F),\\
&\mathit{friend}(A,B) \to \mathit{friend}(B,A),\\
&\mathit{friend}(A,A) \to \bot\}.
\end{aligned}
$$

2.2 CP-Theories

We now recall CP-theories from [Wilson, 2004]. We assume a finite set of *variables* $X \in \mathcal{X}$, each having a finite set of *values* x, denoted $Dom(X)$, also called the *domain* of X. A *value* u for $U \subseteq \mathcal{X}$ associates with every $X \in U$ a value of X. The set of all values u for U is denoted $Dom(U)$, also called the *domain* of U. We denote with $\top$ the only value of $U = \emptyset$. A *conditional preference* φ on $\mathcal{X}$ has the form $u\colon x \succ x'[W]$, where $U \subseteq \mathcal{X}$, $u \in Dom(U)$, $X \in \mathcal{X} - U$, $x, x' \in Dom(X)$, and $W \subseteq \mathcal{X} - (U \cup \{X\})$; intuitively, it means that given u and $t \in Dom(T_\varphi)$, where $T_\varphi = \mathcal{X} - (U \cup \{X\} \cup W)$, we prefer x to x', irrespective of the value of W. A *conditional preference theory* (or *CP-theory*) on $\mathcal{X}$ is a finite set of conditional preferences on $\mathcal{X}$.

An *outcome* is any value $o \in Dom(\mathcal{X})$. Interpretations of CP-theories are strict total orders on $Dom(\mathcal{X})$, i.e., irreflexive and transitive binary relations π on $Dom(\mathcal{X})$ that are total (i.e., for any two distinct $o, o' \in Dom(\mathcal{X})$, either $(o, o') \in \pi$ or $(o', o) \in \pi$). For conditional preferences $\varphi = u\colon x \succ x'[W]$, let $\varphi^\star = \{(tuxw, tux'w') \mid t \in Dom(T_\varphi), w, w' \in Dom(W)\}$. An interpretation π *satisfies* (or *is a model of*) φ, denoted $\pi \models \varphi$, if $\pi \supseteq \varphi^\star$. We say π *satisfies* (or *is a model of*) a CP-theory Γ if it satisfies all $\varphi \in \Gamma$. A CP-theory is *consistent* if it has a model. For two outcomes o and o', we say o is *dominated* by o' in Γ, denoted $o' \succ o$, if (o', o) belongs to all models π of Γ. An outcome o is *undominated* in Γ if there is no outcome o' such that $o' \succ o$. Let $\Gamma^\star$ denote the transitive closure of $\cup_{\varphi \in \Gamma} \varphi^\star$. Then, for all outcomes o and o', it holds that $o \succ o'$ iff $(o, o') \in \Gamma^\star$.

CP-nets [Boutilier *et al.*, 2004] (resp., TCP-nets [Brafman *et al.*, 2006]) can be represented via conditional preferences of the form $u : x \succ x'[W]$ with $W = \emptyset$ (resp., $|W| \leqslant 1$).

Example 2 Alice would like to be online during the holidays to share pictures and comments, so she always prefers a WiFi connection to a cable one (statement (1) below), even though she can live also with the latter. When choosing a hotel, Alice also looks at what her friends already did, and she checks the feedback that they left on the social network. Her preference always goes to hotels with a positive feedback irrespective of who left the feedback and the hotel's connection type (2). Alice knows that Bob is usually very objective in his opinions so, for positive feedbacks, she prefers Bob's reviews to John's (3). She also knows that Bob cannot live without being online with his phone, so Bob does not consider a WiFi connection optional. This is why when there is no WiFi connection, Bob's feedback is biased and tends to be negative. Hence, for hotels without WiFi, Alice prefers John's review (4). For the same reason, for hotels with WiFi, she prefers Bob's feedback to John's (5).

Based on the above description of Alice's preferences, we introduce three variables C (connection), R (reviewer), and F (feedback) with the domains $Dom(C) = \{w, c\}$, $Dom(R) = \{b, j\}$, and $Dom(F) = \{p, n\}$, respectively, and we model Alice's preferences by the CP-theory

$$\begin{aligned}\Gamma \;=\; \{&(1)\,\top\colon w \succ c[\emptyset],\ (2)\,\top\colon p \succ n[\{C, R\}],\\ &(3)\,p\colon b \succ j[\emptyset],\ (4)\,cn\colon j \succ b[\emptyset],\ (5)\,w\colon b \succ j[\emptyset]\},\end{aligned}$$

which has the following two models:

$$\begin{aligned}&pwb \succ pcb \succ pwj \succ pcj \succ nwb \succ nwj \succ ncj \succ ncb,\\ &pwb \succ pwj \succ pcb \succ pcj \succ nwb \succ nwj \succ ncj \succ ncb.\end{aligned}$$

For example, the outcome $o = pwb$ represents wireless connection and positive feedback from reviewer Bob.

3 OCP-Theories

We now introduce ontological CP-theories (or OCP-theories), which extend CP-theories by ontologies. They informally define preferences between conjunctions of atoms relative to an ontology. W.l.o.g., the set Δ_N of nulls is the set of all ground terms constructed from the set Δ of constants and a set $\mathcal{F}$ of functions used to skolemize all existential variables in TGDs.

Definition 1 (OCP-Theory) Let O be a Datalog+/– ontology O over Δ. Let $\mathcal{X}$ be a finite set of variables, each $X \in \mathcal{X}$ being associated with a predicate p from O, denoted $pred(X)$, and as domain $Dom(X)$ with a finite set of at least two different ground atoms $p(c_1, \ldots, c_k)$, with $c_1, \ldots, c_k \in \Delta \cup \Delta_N$. Let $Dom^+(X)$ be the set of all (non-ground) atoms $p(t_1, \ldots, t_k)$ with terms $t_1, \ldots, t_k$ over Δ, $\mathcal{V}$, and $\mathcal{F}$. Then, an *ontological conditional preference* over $\mathcal{X}$ has the form

$$\upsilon\colon \xi \succ \xi'\,[W], \tag{1}$$

where (i) $\upsilon \in Dom^+(U)$ for some $U \subseteq \mathcal{X}$, (ii) $\xi, \xi' \in Dom^+(X)$ for some $X \in \mathcal{X} - U$, and (iii) $W \subseteq \mathcal{X} - (U \cup \{X\})$. We say $\upsilon\theta\colon \xi\theta \succ \xi'\theta\,[W]$, where θ is a substitution, is a *ground instance* of (1), if $\upsilon\theta \in Dom(U)$ and $\xi\theta, \xi'\theta \in Dom(X)$. Let Γ be a finite set of preferences of the form (1). Then, (O, Γ) is an *ontological CP-theory* (or *OCP-theory*).

The following example illustrates OCP-theories.

Example 3 (Hotel Booking cont'd) Consider OCP-theory (O, Γ) given by ontology O of Example 1 and the following set Γ of ontological conditional preferences:

$$\begin{aligned}\Gamma = \{&\top\colon \mathit{hotel}(I, C, w, S) \succ \mathit{hotel}(I, C, c, S)[\emptyset],\\ &\top\colon \mathit{review}(I, U, p) \succ \mathit{review}(I, U, n)[\{C_O, R_O\}],\\ &\mathit{review}(I, U, p)\colon \mathit{reviewer}(b, A) \succ \mathit{reviewer}(j, A')[\emptyset],\\ &\mathit{hotel}(I, C, c, S)\ \mathit{review}(I, U, n)\colon\\ &\qquad\qquad \mathit{reviewer}(j, A) \succ \mathit{reviewer}(b, A')[\emptyset],\\ &\mathit{hotel}(I, C, w, S)\colon \mathit{reviewer}(b, A) \succ \mathit{reviewer}(j, A')[\emptyset]\}\,.\end{aligned}$$

It is defined over the variables C_O, R_O, and F_O with the predicates $pred(C_O) = \mathit{hotel}$, $pred(R_O) = \mathit{reviewer}$, and $pred(F_O) = \mathit{review}$, and the domains $Dom(C_O) = \{\mathit{hotel}(\mathbf{t}_1), \mathit{hotel}(\mathbf{t}_2), \mathit{hotel}(\mathbf{t}_3)\}$, $Dom(R_O) = \{\mathit{reviewer}(\mathbf{t}_{10}), \mathit{reviewer}(\mathbf{t}_{11})\}$, and $Dom(F_O) = \{\mathit{review}(\mathbf{t}_7), \mathit{review}(\mathbf{t}_8), \mathit{review}(\mathbf{t}_9)\}$ (see Fig. 1), respectively.

Note that the notion of outcome for OCP-theories is inherited from CP-theories (Section 2.2), i.e., a mapping that associates with every variable $X \in \mathcal{X}$ a value $x \in Dom(X)$. Observe that every outcome o of an OCP-theory can be seen as a conjunction of ground atoms over $\Delta \cup \Delta_N$ (e.g., in the above example, the outcome o with $o(C_O) = \mathit{hotel}(\mathbf{t}_1)$, $o(R_O) = \mathit{reviewer}(\mathbf{t}_{10})$, and $o(F_O) = \mathit{review}(\mathbf{t}_9)$ can be seen as the conjunction $\mathit{hotel}(\mathbf{t}_1) \wedge \mathit{reviewer}(\mathbf{t}_{10}) \wedge \mathit{review}(\mathbf{t}_9)$).

As a consequence of the underlying ontology, some of these outcomes may be inconsistent, and some other outcomes may be equivalent. We thus have to ensure that the preference relation encoded in an OCP-theory is well-defined, which is expressed in the notion of consistency of an OCP-theory. To define it, we need some preparatory definitions as follows.

An outcome o of (O, Γ) is *consistent* if $O \cup \{o(X) \mid X \in \mathcal{X}\} \not\models \bot$. Two outcomes o and o' of (O, Γ) are *equivalent*, denoted $o \sim o'$, if $O \cup \{o(X) \mid X \in \mathcal{X}\} \equiv O \cup \{o'(X) \mid X \in \mathcal{X}\}$. An *interpretation* π for (O, Γ) is a total order over the outcomes of (O, Γ). We say π *satisfies* (or *is a model of*) O, denoted $\pi \models O$, if (i) o and o' are consistent, for all $(o, o') \in \pi$, and (ii) $(o, o') \in \pi$ for any two equivalent outcomes o and o'. We say π *satisfies* (or *is a model of*) a ground conditional preference φ, denoted $\pi \models \varphi$, if $\pi_s \supseteq \varphi^\star$, where π_s denotes the strict part of π. We say π *satisfies* (or *is a model of*) a set of conditional preferences Γ, denoted $\pi \models \Gamma$, if it satisfies all ground instances of $\varphi \in \Gamma$. We say π *satisfies* (or *is a model of*) an OCP-theory (O, Γ), denoted $\pi \models (O, \Gamma)$, if it satisfies both O and Γ. The notion of consistency of OCP-theories is then defined as the existence of a model.

Definition 2 (Consistency) An OCP-theory (O, Γ) is *consistent* if it has a model.

We next give an alternative characterization of this notion of consistency. Given an OCP-theory (O, Γ), let $\Gamma^\star_\sim$ denote the transitive closure of the set of all $([a]_\sim, [b]_\sim)$ such that (i) $(a, b) \in \Gamma^\star$ and (ii) a and b are both consistent, where $[s]_\sim$ denotes the $\sim$-equivalence class of s. Then, the consistency of OCP-theories describes the acyclicity of $\Gamma^\star_\sim$.

Theorem 1 *OCP-theory* (O,Γ) *is consistent iff* $\Gamma^\star_\sim$ *is acyclic.*

The notions of dominance between consistent outcomes and of undominance of consistent outcomes are as follows.

Definition 3 (Dominance) Let (O, Γ) be a consistent OCP-theory, and o and o' be consistent outcomes. Then, o *dominates* o' in (O, Γ), denoted $o \succ o'$, if $(o, o') \in \pi$ for all models π of (O, Γ). A consistent outcome o is *undominated* in (O, Γ), if no consistent outcome o' exists with $o' \succ o$.

The next theorem shows that the notion of dominance between consistent outcomes is exactly expressed by $\Gamma^\star_\sim$.

Theorem 2 *Let* (O,Γ) *be a consistent OCP-theory, and o and* o' *be consistent outcomes. Then,* $o \succ o'$ *iff* $([o]_\sim, [o']_\sim) \in \Gamma^\star_\sim$.

4 CQ Answering for OCP-Theories

In this section, we define skyline and k-rank answers for CQs to consistent OCP-theories. We also provide an algorithm for computing their answers.

4.1 Semantics

The following definition formalizes answers for CQs $q(\mathbf{X}) = \exists \mathbf{Y}\, \Phi(\mathbf{X}, \mathbf{Y})$ to O as usual, and relates them to the consistent outcomes of (O, Γ). Here, we assume a bijection β from a set $\Phi_\beta(\mathbf{X}, \mathbf{Y}) \subseteq \Phi(\mathbf{X}, \mathbf{Y})$ of atoms in q to a set of variables of (O, Γ) (and so their values in consistent outcomes) such that every atom $a \in \Phi_\beta(\mathbf{X}, \mathbf{Y})$ has the same predicate as the variable $\beta(a)$. Intuitively, β relates query atoms in q to preference atoms in (O, Γ). Note that when β is empty, the answers for q to O are unrelated to the consistent outcomes of (O, Γ), i.e., they are (unordered) standard answers to CQs.

Definition 4 (Answer) Let (O, Γ), with $O = (D, \Sigma)$, be a consistent OCP-theory, and let $q(\mathbf{X}) = \exists \mathbf{Y}\, \Phi(\mathbf{X}, \mathbf{Y})$ be a CQ. Then, the set of all *answers* for q to (O, Γ) under the consistent outcome o of (O, Γ), denoted $ans(q, \Gamma, O, o)$, is the set of all tuples $\mathbf{a}$ over $\Delta \cup \Delta_N$ for which a homomorphism $\mu\colon \mathbf{X} \cup \mathbf{Y} \to \Delta \cup \Delta_N$ exists such that (i) $D \cup \Sigma \models \mu(\Phi(\mathbf{X}, \mathbf{Y}))$, (ii) $\mu(\mathbf{X}) = \mathbf{a}$, and (iii) $\mu(a) = o(\beta(a))$ for all $a \in \Phi_\beta(\mathbf{X}, \mathbf{Y})$. An *answer* for q to (O, Γ) is an answer for q to (O, Γ) under some consistent outcome o of (O, Γ).

Note that nulls are included in answers, since they actually represent prototypical constants. The following example illustrates the notion of answer for CQs to OCP-theories.

Example 4 Consider the consistent OCP-theory (O, Γ) of Example 3 and the CQ $q(A, B, C, D) = \exists Y\, \mathit{hotel}(A, rome, C, Y) \wedge \mathit{review}(A, B, D)$. Then, $\langle h_2, b, w, p\rangle$ is an answer for q under the consistent outcome $o = \mathit{hotel}(h_2, rome, w, l)$ $\mathit{review}(h_2, b, p)$ $\mathit{reviewer}(b, 20)$. Let $a = \mathit{hotel}(A, rome, C, Y)$ and $\beta(a) = C_o$; once we fix the outcome o, then we have that $o(\beta(a)) = \mathit{hotel}(h_2, rome, w, l)$.

We next sort these answers based on the preferences of the user. We do this via skyline answers [Börzsönyi *et al.*, 2001], a well-known class of answers for preference-based formalisms, and the iterated computation of skyline answers that allows us to assign a *rank* to every answer (via the CP-theory); we refer to these as k-rank answers. We first define skyline answers; note that they are not unique. Indeed, we may have more than one undominated consistent outcome o, and also more than one homomorphism μ that satisfies the conditions (i) and (ii) of Definition 4 for the same undominated consistent outcome o.

Definition 5 (Skyline Answer) Let (O, Γ) be a consistent OCP-theory, and let $q(\mathbf{X}) = \exists \mathbf{Y}\, \Phi(\mathbf{X}, \mathbf{Y})$ be a CQ. A *skyline answer* for q to (O, Γ) is any tuple $\mathbf{a} \in ans(q, \Gamma, O, o)$ for some consistent outcome o such that there is no consistent outcome o' with (i) $o' \succ o$ and (ii) $ans(q, \Gamma, O, o') \neq \emptyset$.

We next define k-rank answers for CQs to OCP-theories via iteratively applying the notion of skyline answer. The definition is technically involved, as answers are obtained by projecting away existentially quantified arguments, thus distinct outcomes may be associated with the same answer, and so distinct levels of the skyline iteration may contain the same answers. We handle this in each iteration step by removing all answers computed in previous iteration steps.

Definition 6 (k-Rank Answer) Let (O, Γ) be a consistent OCP-theory, and $q(\mathbf{X}) = \exists \mathbf{Y}\, \Phi(\mathbf{X}, \mathbf{Y})$ be a CQ. A *k-rank answer* for q to (O, Γ) *outside* a set of ground atoms S is a sequence $\langle \mathbf{a_1}, \ldots, \mathbf{a_k}\rangle$ such that either:

(a) $\mathbf{a_1}, \ldots, \mathbf{a_k}$ are k different skyline answers for q to (O, Γ) not belonging to S, if at least k such answers exist; or

(b) (1) $\mathbf{a_1}, \ldots, \mathbf{a_i}$ are all i different skyline answers for q to (O, Γ) that do not belong to S, and (2) $\langle \mathbf{a_{i+1}}, \ldots, \mathbf{a_k}\rangle$ is a $(k{-}i)$-rank answer for q to $(O, \Gamma{-}\{o\})$ outside $S \cup \{\mathbf{a_1}, \ldots, \mathbf{a_i}\}$, where o is an undominated outcome relative to $\succ$, otherwise.

A *k-rank answer* for q to (O, Γ) is a k-rank answer for q to (O, Γ) outside $\emptyset$.

Note that when no answer for q to (O, Γ) exists, then $\langle\rangle$ is its unique k-rank answer. Informally, a k-rank answer is a sequence of k answers for a CQ to an OCP-theory ranked by following the order among consistent outcomes induced by the OCP-theory. Clearly, as skyline answers are in general not unique, also k-rank answers are in general not unique.

Input: consistent OCP-theory (O, Γ), CQ q, and $k \geqslant 0$.
Output: a k-rank answer $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$ for q to (O, Γ).

1 *Result* $\leftarrow \langle\rangle$; *Outcomes* $\leftarrow \emptyset$; *Checked* $\leftarrow \emptyset$;
2 *k-reached* $\leftarrow$ *false*;
3 **while** *k-reached* $=$ *false* **do**
4   **foreach** *undominated outcome o relative to the partial order obtained from $\succ$ by removing all pairs $o_1 \succ o_2$ such that $o_1 \in$ Checked* **do**
5   *Outcomes* $\leftarrow$ *Outcomes* $\cup \{o\}$;
6   **while** *k-reached* $=$ *false and Outcomes* $\neq \emptyset$ **do**
7   choose $o \in$ *Outcomes* ;
8   *Outcomes* $\leftarrow$ *Outcomes* $- \{o\}$;
9   *Checked* $\leftarrow$ *Checked* $\cup \{o\}$;
10   **foreach** μ *such that* $\mu(\mathbf{X}) \in ans(q, \Gamma, O, o)$ **do**
11   **if** *k-reached* $=$ *false and* $\mu(\mathbf{X}) \notin$ *Result* **then**
12   *Result* $\leftarrow$ *Result* $\circ \langle \mu(\mathbf{X}) \rangle$;
13   **if** *length*(*Result*) $= k$ **then**
14   *k-reached* $\leftarrow$ *true*;
15 **return** *Result*.

ALGORITHM 1: k-Rank-Prefs (O, Γ, Q, k)

4.2 Algorithm

Algorithm 1 computes, given a consistent OCP-theory (O, Γ), a CQ q, and $k \geqslant 0$ as input, a k-rank answer for q to (O, Γ). It exploits the preference order of consistent outcomes to iteratively compute skyline answers for q to (O, Γ). It stops when it reaches k different answers to q. The most preferred answers are the ones related to the undominated consistent outcomes o; thus, the computation starts by adding such o to *Outcomes* (line 4). There are two sources of non-determinism: in line 7, we may choose arbitrarily among different incomparable outcomes; then, in line 10, we may have multiple equivalent μ's, and after selecting some of them, we reach *length*(*Result*) $= k$. Note that the set of all skyline answers for q to (O, Γ) can be computed by setting *k-reached* to true in line 2 and by stopping the algorithm after the set of all answers associated with all undominated outcomes in *Outcomes* are computed, once this set of answers is non-empty.

5 Computational Complexity

We now analyze the computational complexity of deciding consistency, dominance, and CQ skyline membership for OCP-theories. We also delineate some tractable special cases. We assume some familiarity with complexity classes PSPACE (resp.,P, EXP, 2EXP) containing all decision problems that can be solved in polynomial space (resp., polynomial, exponential, double exponential time) on a deterministic Turing machine; see [Johnson, 1990; Papadimitriou, 1994].

5.1 Decidability Paradigms

The main (syntactic) conditions on TGDs that guarantee the decidability of CQ answering are guardedness [Calì *et al.*, 2013], stickiness [Calì *et al.*, 2012b], and acyclicity. Interestingly, each of them has its "weak" counterpart: weak guardedness [Calì *et al.*, 2013], weak stickiness [Calì *et al.*, 2012b] and weak acyclicity [Fagin *et al.*, 2005], respectively.

A TGD σ is *guarded* if an atom $\mathbf{a} \in body(\sigma)$ exists that contains (or "guards") all the body variables of σ. The class of guarded TGDs, denoted G, is defined as the family of all possible sets of guarded TGDs. A key subclass of guarded TGDs are linear TGDs with just one body atom (which is automatically a guard), and the corresponding class is denoted L. *Weakly guarded* TGDs extend guarded TGDs by requiring only "harmful" body variables to appear in the guard, and the associated class is denoted WG. Notice that $\mathsf{L} \subset \mathsf{G} \subset \mathsf{WG}$.

Stickiness is inherently different from guardedness, and its central property is as follows: variables that appear more than once in a body (i.e., join variables) are always propagated (or "stick") to the inferred atoms. A set of TGDs that enjoys the above property is called *sticky*, and the corresponding class is denoted S. Weak stickiness is a relaxation of stickiness where only "harmful" variables are taken into account. A set of TGDs that enjoys weak stickiness is *weakly sticky*, and the associated class is denoted WS. Observe that $\mathsf{S} \subset \mathsf{WS}$.

A set Σ of TGDs is *acyclic* if its predicate graph is acyclic, and the underlying class is denoted A. In fact, an acyclic set of TGDs can be seen as nonrecursive. We say Σ is *weakly acyclic* if its dependency graph enjoys a certain acyclicity condition, which guarantees the existence of a finite canonical model; the associated class is denoted WA. Clearly, $\mathsf{A} \subset \mathsf{WA}$.

Another key fragment of TGDs are *full* TGDs, i.e., TGDs without existentially quantified variables, and the corresponding class is denoted F. If we further assume that full TGDs enjoy linearity, guardedness, stickiness, or acyclicity, then we obtain the classes LF, GF, SF, and AF, respectively.

5.2 Combined Complexity

For complexity classes $\mathcal{C} \subseteq$ PSPACE (resp., deterministic complexity classes $\mathcal{C} \supseteq$ EXP), the following theorem shows the generic results that (i) deciding consistency of OCP-theories (O, Γ) and (ii) deciding dominance between two outcomes of OCP-theories (O, Γ) are both complete for PSPACE (resp., $\mathcal{C}$) when BCQ^N answering in O is complete for $\mathcal{C}$, and hardness holds even for ground atomic BCQs. Here, BCQ^{N}s denote BCQs with elements from $\Delta \cup \Delta_N$ (where Δ_N is as defined in Section 3) and variables as arguments in atoms.

Theorem 3 *Let $\mathcal{T}$ be a class of OCP-theories (O, Γ) where BCQ^N answering in O is complete for a complexity class $\mathcal{C} \subseteq$ PSPACE (resp., deterministic complexity class $\mathcal{C} \supseteq$ EXP), and hardness holds even for ground atomic BCQs. Then, deciding whether (a) some given $(O, \Gamma) \in \mathcal{T}$ is consistent and (b) o dominates o' in (O, Γ), given $(O, \Gamma) \in \mathcal{T}$ and two outcomes o and o', are both complete for PSPACE (resp., $\mathcal{C}$).*

Proof (sketch). (a) For membership, by Theorem 1 it is sufficient to decide whether $\Gamma^\star_\sim$ is acyclic. This can be done by deciding whether $[o]_\sim \succ [o]_\sim$ for some consistent outcome o, which can be done (despite the number of outcomes of (O, Γ) being exponential) by storing at most two outcomes, while exploring all the paths in $\Gamma^\star_\sim$. This requires both deciding BCQ^{N}s to O (for deciding if an outcome is consistent, and whether two consistent outcomes are equivalent), which is in $\mathcal{C}$, and deciding dominance between two outcomes in a standard CP-theory, which is in PSPACE [Goldsmith *et al.*, 2008]. Overall, this is possible in PSPACE (resp., $\mathcal{C}$).

Hardness for PSPACE follows from the more specialized problem of deciding consistency in a standard CP-theory being PSPACE-hard. Hardness for $\mathcal{C}$ holds by a reduction from the $\mathcal{C}$-hard problem of answering ground atomic BCQs, since $O \models \mathbf{a}$ iff $(O \cup \{\mathbf{a}'\}, \{\mathbf{a}' \succ \mathbf{a}\})$ over $\mathcal{X} = \{X\}$ with $Dom(X) = \{\mathbf{a}, \mathbf{a}'\}$ is inconsistent.

(b) For membership, by Theorem 2 it is sufficient to decide whether o and o' are consistent and whether $[o]_\sim \succ [o']_\sim$. As argued in (a), this is overall possible in PSPACE (resp., $\mathcal{C}$).

PSPACE-hardness follows from the PSPACE-hardness of the more specialized problem of deciding dominance between two outcomes in standard CP-theories [Goldsmith *et al.*, 2008]. Hardness for $\mathcal{C}$ holds by a reduction from the same problem as in (a), since $O \models \mathbf{a}'$ iff $\mathbf{a}'' \succ \mathbf{a}$ holds in $(O \cup \{\mathbf{a}''\}, \{\mathbf{a}' \succ \mathbf{a}\})$ over $\mathcal{X} = \{X\}$ with $Dom(X) = \{\mathbf{a}, \mathbf{a}', \mathbf{a}''\}$. □

As a corollary, by the known complexity results for Datalog+/– (see, e.g., [Lukasiewicz *et al.*, 2015]), deciding consistency and dominance for OCP-theories when the underlying ontology is in L, LF, AF, G, WG, S, F, GF, SF, WF, or WA, is complete for the complexity classes shown in Table 1 in the combined, *ba*-combined, and *fp*-combined complexity.

Corollary 4 *Given an OCP-theory (O, Γ), where O is defined in* L, LF, AF, G, WG, S, F, GF, SF, WS, *or* WA, *deciding whether (a) (O, Γ) is consistent and (b) o dominates o', given (O, Γ) and two outcomes o and o', is both complete for the complexity classes shown in Table 1 in the combined, ba-combined, and fp-combined complexity.*

We next show that the complexity of deciding whether a tuple over Δ is in the skyline answer for a CQ to an OCP-theory (O, Γ) is complete for PSPACE (resp., $\mathcal{C}$) when BCQ^N answering in O is complete for the complexity class $\mathcal{C} \subseteq$ PSPACE (resp., deterministic complexity class $\mathcal{C} \supseteq$ EXP).

Theorem 5 *Let $\mathcal{T}$ be a class of consistent OCP-theories (O, Γ) where BCQ^N answering in O is complete for a complexity class $\mathcal{C} \subseteq$ PSPACE (resp., deterministic complexity class $\mathcal{C} \supseteq$ EXP). Then, given $(O, \Gamma) \in \mathcal{T}$, a CQ q, and a tuple $\mathbf{a}$ over $\Delta \cup \Delta_N$, deciding whether $\mathbf{a}$ is in the skyline answer for q to (O, Γ) is complete for PSPACE (resp., $\mathcal{C}$).*

Proof (sketch). For membership, it is sufficient to decide if $\mathbf{a} \in ans(q, \Gamma, O, o)$ for some consistent outcome o such that no consistent outcome o' exists with (i) $o' \succ o$ and (ii) $ans(q, \Gamma, O, o') \neq \emptyset$. By a similar argument as in the proof of Theorem 3, this is overall possible in PSPACE (resp., $\mathcal{C}$).

Hardness for $\mathcal{C}$ follows from the more specialized problem of BCQ^N answering being $\mathcal{C}$-hard. Hardness for PSPACE follows by a reduction from the PSPACE-hard problem of deciding dominance between two outcomes $o = o_1 \dots o_n$ and $o' = o'_1 \dots o'_n$ in standard CP-theories Γ, as $o \succ o'$ in Γ iff $(o_1, \dots, o_n)$ is in the skyline answer for $\bigcup_{i=1}^n \{q_i(X_i)\}$ to $(\bigcup_{i=1}^n \{q_i(o_i)\} \cup \bigcup_{i=1}^n \{q_i(o'_i)\}, \Gamma')$, where Γ' is obtained from Γ by consistently replacing values v_i by $q_i(v_i)$. □

By the known complexity results for Datalog+/– (see, e.g., [Lukasiewicz *et al.*, 2015]), we obtain as a corollary that deciding CQ skyline membership for OCP-theories when the underlying ontology is in L, LF, AF, G, WG, S, F, GF, SF, WF, or WA, is complete for the complexity classes in Table 1 in the combined, *ba*-combined, and *fp*-combined complexity.

Language	Data	Comb.	*ba*-comb.	*fp*-comb.
L, LF, AF	in P	PSPACE	PSPACE	PSPACE
G	in P	2EXP	EXP	PSPACE
WG	EXP	2EXP	EXP	EXP
S, F, GF, SF	in P	EXP	PSPACE	PSPACE
WS, WA	in P	2EXP	2EXP	PSPACE

Table 1: Data, combined, *ba*-combined, and *fp*-combined complexity of deciding consistency, dominance, and CQ skyline membership for OCP-theories in different languages.

Corollary 6 *Given a consistent OCP-theory (O, Γ), where O is defined in* L, LF, AF, G, WG, S, F, GF, SF, WS, *or* WA, *a CQ, and a tuple $\mathbf{a}$ over $\Delta \cup \Delta_N$, deciding if $\mathbf{a}$ is in the skyline answer for q to (O, Γ) is complete for the classes in Table 1 in the combined, ba-combined, and fp-combined complexity.*

5.3 Data Complexity

We now delineate special cases where deciding consistency, dominance, and CQ skyline membership for OCP-theories are all tractable in the data complexity. The following result shows that deciding consistency and dominance for OCP-theories (O, Γ), with $O = (D, \Sigma)$, is data tractable when BCQ^N answering in O is data tractable. Here, data complexity means that Σ and the variables and preferences of Γ are fixed, while D and the domains of Γ are part of the input.

Theorem 7 *Let $\mathcal{T}$ be a class of OCP-theories (O, Γ) where BCQ^N answering in O is possible in polynomial time in the data complexity. Then, deciding whether (a) some given $(O, \Gamma) \in \mathcal{T}$ is consistent and (b) o dominates o', given $(O, \Gamma) \in \mathcal{T}$ and outcomes o and o', are both possible in polynomial time in the data complexity.*

Proof (sketch). (a) In the data complexity, the number of all outcomes of (O, Γ) is polynomial in the number of domain values. As deciding BCQs in O is possible in data polynomial time, deciding if an outcome is consistent, and whether two consistent outcomes are equivalent as well. So, computing $\Gamma^\star_\sim$ and deciding whether $\Gamma^\star_\sim$ is acyclic can also be done in data polynomial time. Overall, by Theorem 1, deciding consistency of (O, Γ) is possible in data polynomial time.

(b) By Theorem 2, it is sufficient to decide whether o and o' are consistent and whether $[o]_\sim \succ [o']_\sim$. As argued in (a), this is overall possible in data polynomial time. □

As a corollary, deciding consistency and dominance for OCP-theories is data tractable when the underlying ontology is formulated in L, LF, AF, G, S, F, GF, SF, WF, or WA, which all allow for data tractable BCQ^N answering.

Corollary 8 *Given an OCP-theory (O, Γ), where O is defined in* L, LF, AF, G, S, F, GF, SF, WF, *or* WA, *deciding whether (a) (O, Γ) is consistent and (b) o dominates o', given (O, Γ) and outcomes o and o', are all possible in polynomial time in the data complexity.*

The next result shows that deciding membership to skyline answers for CQs to consistent OCP-theories is data tractable when BCQ^N answering in O is data tractable.

Theorem 9 *Let $k \geqslant 0$ be fixed. Let $\mathcal{T}$ be a class of consistent OCP-theories (O, Γ) where BCQ^N answering in O can be done in polynomial time in the data complexity. Then, given $(O, \Gamma) \in \mathcal{T}$, a CQ q, and a tuple $\mathbf{a}$ over $\Delta \cup \Delta_N$, deciding whether $\mathbf{a}$ is in the skyline answer for q to (O, Γ) is possible in polynomial time in the data complexity.*

Proof (sketch). As argued in the proof of Theorem 7, computing $\Gamma^{\star}_{\sim}$ can be done in data polynomial time. Hence, by Theorem 1, computing iteratively undominated outcomes can also be done in data polynomial time. Since BCQ^N answering in O is data tractable, deciding whether $\mathbf{a}$ is in the skyline answer for q to (O, Γ) is possible in data polynomial time. □

A corollary is that membership to skyline answers for CQs to consistent OCP-theories is data tractable when the ontology is formulated in L, LF, AF, G, S, F, GF, SF, WF, or WA, which all allow for data tractable BCQ^N answering.

Corollary 10 *Given a consistent OCP-theory (O, Γ), a CQ q, where O is defined in* L, LF, AF, G, S, F, GF, SF, WF, *or* WA, *and a tuple $\mathbf{a}$ over $\Delta \cup \Delta_N$, deciding whether $\mathbf{a}$ is in the skyline answer for q to (O, Γ) is possible in polynomial time in the data complexity.*

The above data tractability results do not carry over to WG, where BCQ^N answering in O is data complete for EXP, and data hardness holds even for ground atomic BCQs: so, data completeness for EXP can be proved similarly to Theorem 5.

6 Related Work

Modeling and dealing with preferences in databases has been studied for almost three decades, since the seminal work of Lacroix and Lavency [1987]; see [Stefanidis *et al.*, 2011] for a survey of notable works in this line. Work has also been carried out in the intersection with databases and knowledge representation and reasoning, such as preference logic programs [Govindarajan *et al.*, 1995], incorporation of preferences into formalisms such as answer set programs [Brewka, 2007], and answering k-rank queries in ontological languages [Lukasiewicz *et al.*, 2013].

In the philosophical tradition, preferences are usually expressed over mutually exclusive "worlds", such as truth assignments to formulas. Bienvenu *et al.*'s work [2010] is framed in this interpretation of preferences, aiming at bridging the gap between several formalisms in AI, such as CP-nets and those studied traditionally in philosophy. In this regard, CP-nets [Boutilier *et al.*, 2004] are one of the most widely known formalisms. Most work on CP-nets has focused on the computation of optimal outcomes and dominance testing, i.e., to check if one outcome of the CP-net is preferred to another. More recently, Wang *et al.* [2012] propose an efficient algorithm and indexing scheme for top-k retrieval in CP-nets.

Recently, there has been some interest in the combination of Semantic Web technologies with preference representation and reasoning. A combination of conditional preferences (very different from CP-theories) with DL reasoning for ranking objects is presented in [Lukasiewicz and Schellhase, 2007]. There, conditional preferences are exploited in the definition of a ranking function that allows to perform a semantic personalized search and ranking over a set of resources annotated via an ontological description. In [Lukasiewicz *et al.*, 2013], Datalog+/– is extended with preference management formalisms closely related to those previously studied for relational databases. The paper also considers skyline and k-rank queries, but the allowed preference statements are more general first-order statements, which are not related to CP-theories. This also comes at the price of higher complexity and data tractability results that hold only for disjunctions of atomic queries and not for CQs like here.

Closest in spirit to this paper is perhaps the preference formalism that combines CP-nets and DLs in [Di Noia *et al.*, 2013], where variable values of CP-nets are satisfiable DL formulas. The main difference with this paper lies in the use of CP-theories here rather than CP-nets, and (more importantly) in the relationship between the ontology and the CP-net/theory. While Di Noia *et al.* [2013] use ontological axioms to restrict CP-net outcomes, here we use the preference information contained in a CP-theory to inform how answers to queries over the ontology should be ranked. Finally, in the context of information retrieval, in [Boubekeur *et al.*, 2007] Wordnet is used to add semantics to CP-net variables. Another interesting approach to mixing qualitative preferences with Semantic Web technology is presented in [Siberski *et al.*, 2006], where an extension of SPARQL is studied that can encode user preferences in the query. Santhanam *et al.* [2010] reduce dominance testing in CP-nets to reachability analysis in a graph of outcomes. Mindolin and Chomicki [2011] explore the p-skyline framework, which extends skylines with the notion of attribute importance in preference relations.

7 Summary and Outlook

We have introduced ontological CP-theories (OCP-theories), which are a novel combination of Datalog+/– ontologies with CP-theories. We have defined skyline and k-rank answers for CQs to OCP-theories. We have also provided an algorithm for computing such skyline and k-rank answers. Furthermore, we have provided a host of precise complexity (including several data tractability) results for deciding consistency, dominance, and CQ skyline membership for OCP-theories.

Interesting topics of ongoing and future research include the implementation and experimental evaluation of the presented approach, as well as the investigation and development of optimized special-case and/or approximation algorithms, and the exploration of further tractable cases of OCP-theories.

Acknowledgments This work was supported by the UK EPSRC grant EP/J008346/1 ("PrOQAW"), by a Google European Doctoral Fellowship, by the EU (FP7/2007-2013) ERC grant 246858 ("DIADEM"), by a Yahoo! Research Fellowship, by the grant PON02_00563_3470993 ("VINCENTE"), and funds provided by CONICET and UNSur.

References

[Bienvenu *et al.*, 2010] M. Bienvenu, J. Lang, and N. Wilson. From preference logics to preference languages, and back. In *Proc. of KR*, pp. 214–224, 2010.

[Börzsönyi *et al.*, 2001] S. Börzsönyi, D. Kossmann, and K. Stocker. The skyline operator. In *Proc. of ICDE*, pp. 421–430, 2001.

[Boubekeur *et al.*, 2007] F. Boubekeur, M. Boughanem, and L. Tamine-Lechani. Semantic information retrieval based on CP-Nets. In *Proc.of FUZZ-IEEE*, pp. 1 –7, 2007.

[Boutilier *et al.*, 1999] C. Boutilier, R. I. Brafman, H. H. Hoos, and D. Poole. Reasoning with conditional ceteris paribus preference statements. In *Proc. of UAI*, pp. 71–80, 1999.

[Boutilier *et al.*, 2004] C. Boutilier, R. I. Brafman, C. Domshlak, H. H. Hoos, and D. Poole. CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *J. Artif. Intell. Res.*, 21:135–191, 2004.

[Brafman *et al.*, 2006] R. I. Brafman, C. Domshlak, and S.E. Shimony. On graphical modeling of preference and importance. *J. Artif. Intell. Res.*, 25(1):389–424, 2006.

[Brewka, 2007] G. Brewka. Preferences, contexts and answer sets. In *Proc. of ICLP*, p. 22, 2007.

[Calì *et al.*, 2012a] A. Calì, G. Gottlob, and T. Lukasiewicz. A general Datalog-based framework for tractable query answering over ontologies. *J. Web Sem.*, 14:57–83, 2012.

[Calì *et al.*, 2012b] A. Calì, G. Gottlob, and A. Pieris. Towards more expressive ontology languages: The query answering problem. *Artif. Intell.*, 193:87–128, 2012.

[Calì *et al.*, 2013] A. Calì, G. Gottlob, and M. Kifer. Taming the infinite chase: Query answering under expressive relational constraints. *J. Artif. Intell. Res.*, 48:115–174, 2013.

[Calvanese *et al.*, 2007] D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. Autom. Reasoning*, 39(3):385–429, 2007.

[Di Noia *et al.*, 2013] T. Di Noia, T. Lukasiewicz, and G. I. Simari. Reasoning with semantic-enabled qualitative preferences. In *Proc. of SUM*, pp. 374–386, 2013.

[Fagin *et al.*, 2005] R. Fagin, P. G. Kolaitis, R. J. Miller, and L. Popa. Data exchange: Semantics and query answering. *Theor. Comput. Sci.*, 336(1):89–124, 2005.

[Goldsmith *et al.*, 2008] J. Goldsmith, J. Lang, M. Truszczynski, and N. Wilson. The computational complexity of dominance and consistency in CP-nets. *J. Artif. Intell. Res.*, 33:403–432, 2008.

[Govindarajan *et al.*, 1995] K. Govindarajan, B. Jayaraman, and S. Mantha. Preference logic programming. In *Proc. of ICLP*, pp. 731–745, 1995.

[Johnson, 1990] D. S. Johnson. A catalog of complexity classes. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume A, chapter 2, pp. 67–161. MIT Press, 1990.

[Lacroix and Lavency, 1987] M. Lacroix and P. Lavency. Preferences: Putting more knowledge into queries. In *Proc. of VLDB*, pp. 1–4, 1987.

[Lukasiewicz and Schellhase, 2007] T. Lukasiewicz and J. Schellhase. Variable-strength conditional preferences for ranking objects in ontologies. *J. Web Sem.*, 5(3):180–194, 2007.

[Lukasiewicz *et al.*, 2013] T. Lukasiewicz, M. V. Martinez, and G. I. Simari. Preference-based query answering in Datalog+/– ontologies. In *Proc. of IJCAI*, pp. 501–518, 2013.

[Lukasiewicz *et al.*, 2015] T. Lukasiewicz, M. V. Martinez, A. Pieris, and G. I. Simari. From classical to consistent query answering under existential rules. In *Proc. of AAAI*, pp. 1546–1552, 2015.

[Mindolin and Chomicki, 2011] D. Mindolin and J. Chomicki. Preference elicitation in prioritized skyline queries. *VLDB J.*, 20(2):157–182, 2011.

[Papadimitriou, 1994] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.

[Santhanam *et al.*, 2010] G. R. Santhanam, S. Basu, and V. Honavar. Dominance testing via model checking. In *Proc. of AAAI*, pp. 357–362, 2010.

[Siberski *et al.*, 2006] W. Siberski, J. Z. Pan, and U. Thaden. Querying the Semantic Web with preferences. In *Proc. of ISWC*, pp. 612–624, 2006.

[Stefanidis *et al.*, 2011] K. Stefanidis, G. Koutrika, and E. Pitoura. A survey on representation, composition and application of preferences in database systems. *ACM Trans. Database Syst.*, 36, 2011.

[Vardi, 1982] M. Y. Vardi. The complexity of relational query languages (extended abstract). In *Proc. of STOC*, pp. 137–146, 1982.

[Wang *et al.*, 2012] H. Wang, X. Zhou, W. Chen, and P. Ma. Top-k retrieval using conditional preference networks. In *Proc. of CIKM*, pp. 2075–2079, 2012.

[Wilson, 2004] N. Wilson. Extending CP-nets with stronger conditional preference statements. In *Proc. of AAAI*, pp. 735–741, 2004.