Heuristics for Licenses Composition

Guido GOVERNATORI\textsuperscript{a,1}, Ho-Pun LAM\textsuperscript{a}, Antonino ROTOLO\textsuperscript{b}, Serena VILLATA\textsuperscript{c} and Fabien GANDON\textsuperscript{c}

\textsuperscript{a}NICTA Queensland Research Laboratory
\textsuperscript{b}University of Bologna
\textsuperscript{c}INRIA Sophia Antipolis

Abstract. The Web of Data is assisting to a growth of interest with respect to the open challenge of representing and reasoning in an automated way over licenses and copyright. In this paper, we deal with the problem of checking the compatibility of a set of licenses associated to a single query result returned on the Web of Data, and subsequently compose them into a so called composite license. More precisely, we analyze two composition heuristics, \textit{AND-composition} and \textit{OR-composition}, showing how they can be used to combine the deontic components specified by the licenses, i.e., permissions, obligations, and prohibitions, and which are the most suitable combinations depending on the starting licenses. Such heuristics are evaluated using the SPINdle logic reasoner.

Introduction

In the Web of Data [11], the problem of handling the licensing terms associated to the data in an automated way is becoming more and more important. Several challenges arise to represent the licensing information in a machine-readable format (i.e., from the definition of lightweight vocabularies like ORDL\textsuperscript{2} and Creative Commons\textsuperscript{3} up to the definition of specific ontology design patterns for licenses applied to Linked Data resources\textsuperscript{4}), and to reason over such information to achieve more complex goals like checking the compatibility of a set of licenses and compose them in a compliant way. In particular, the problem of combining the set of terms belonging to heterogeneous licenses or contracts has been studied in different contexts [4,3,18,17]. However, a deeper analysis of the possible composition heuristics w.r.t. the specific deontic component (i.e., permissions, obligations, and prohibitions) expressed through such licensing terms is needed.

In this paper, we address the following research question: how to reason over the composition of a set of licenses in such a way that their deontic component guides the choice of the heuristic to be adopted? We answer this research question by adopting the defeasible deontic logic semantics presented in [17,10] and the SPINdle reasoning engine [13] to actually compute the composite licenses following two composition heuristics, namely the \textit{AND-composition} and the \textit{OR-composition}.

In particular, the overall workflow of the licenses composition framework we propose is visualized in Figure 1. Our application scenario consists in a data consumer who queries

\textsuperscript{1}Corresponding Author: guido.governatori@nicta.com.au
\textsuperscript{2}http://u3.org/ns/ordl/2/
\textsuperscript{3}http://creativecommons.org/ns
\textsuperscript{4}http://bit.ly/DesignPatternsLicenses
the Web of Data through a SPARQL endpoint in order to obtain some data (step 1, Fig. 1). The aim of our licenses composition framework is to retrieve the license(s) associated to the query result (steps 1-2, Fig. 1), and, if there is more than one license associated to the triples, we verify the compatibility of the different licensing terms composing them into a unique composite license, which is the only license returned to the consumer. In particular, we first translate the retrieved licenses from RDF to the SPINdle syntax (step 4, Fig. 1) and then the whole theory, containing all the licenses to be composed, is loaded in SPINdle. SPINdle manages two kinds of composition heuristics (step 5, Fig. 1): the AND-composition (i.e., the composite license entails a deontic effect if all the licenses composing it entail such deontic effect), and the OR-composition (i.e., the composite license entails a deontic effect if there is at least one licenses that entails such effect, and no license prevents it). These two heuristics can be combined together to produce the composite licenses (step 6, Fig. 1), allowing in such a way a different treatment for each deontic component. Finally, we return to the consumer the query result together with the URI of the composite license expressed in machine-readable format, i.e., RDF (step 7, Fig. 1).

In this paper, we concentrate only on the AND-composition and OR-composition heuristics and their possible combination. Other composition heuristics like the Constraining Value [3,17] and quantitative heuristics are not considered, and we left their definition in our framework as future work.

The remainder of this paper is as follows. In Section 1 we formally define our defeasible deontic logic for licenses composition and the two composition heuristics. In Section 2 we evaluate the heuristics using the SPINdle reasoner, analyzing how they can be combined together to better suit the data publisher’s needs. Section 3 we discuss the related literature comparing it with the proposed approach.

1. Composition heuristics for data licensing

In this section, we introduce the defeasible deontic logic we rely on to automatically generate the composite licenses ensuring its compliance w.r.t. the single licenses composing it.

We propose an extension of Defeasible Logic, extending earlier works [9,8], to handle license composition. Previous versions of this logic were proposed in [17,10]. The current
version, as will see, is much more compact than the one in [17] and proposes a new and
more intuitive reading of AND-composition and OR-composition than the one in [10].
Both improvements allow us to easily generate composite licenses. Dealing with license
composition requires reasoning about two components:

**Factual and ontology component**: the first component is meant to describe the facts
with respect to which Web of Data licenses are applied as well as the ontology of
concepts involved by licenses (thus modeling, e.g., concept inclusion);

**Deontic component**: the second component aims at capturing the deontic aspects of
Web of Data licenses, thus offering mechanisms for reasoning about obligations,
prohibitions, and permissions in force in each license, and in their composition.

We focus on the deontic component, even though, for the sake of completeness, we
illustrate the proposed method by also handling, in standard Defeasible Logic, the factual
and ontology component, as done in [2,17]. Standard Defeasible Logic is just an option,
and the factual and ontology component can be handled in any other suitable logic and by
resorting to a separate reasoner. Also, notice that we assume that all licenses share a same
ontology, or that the ontologies are aligned.

The formal language of the logic is rule-based. Literals can be plain, such as \( p, q, r, \ldots \),
or modal, such \( Op \) (obligatory), \( Pp \) (permitted), and \( Fp \) (forbidden/prohibited). Ontology
rules work as regular Defeasible Logic rules for deriving plain literals, while the logic of
deontic rules provide a constructive account of the basic deontic modalities (obligation,
prohibition, and permission). However, while we assume that all licenses share a same
ontology, the purpose of the formalism is mainly to establish the conditions to derive dif-
ferent deontic conclusions from different licenses, and check whether they are compatible
so that they can be attributed to a composite license. Hence, we need to keep track of how
these deontic conclusions are obtained. To this purpose, deontic rules (and, as we will see,
their conclusions) are parametrized by labels referring to licenses.

An ontology rule such as \( a_1, \ldots, a_n \Rightarrow b \) supports the conclusion of \( b \), given \( a_1, \ldots, a_n \),
and so it states that, from the viewpoint of any license any instance enjoying \( a_1, \ldots, a_n \) is
also an instance of \( b \). On the contrary, rules as \( a, Ob \Rightarrow b \) state that, if \( a \) is the case and \( b \)
is obligatory, then \( Op \) holds in the perspective of license \( l_2 \), i.e., \( p \) is obligatory for \( l_2 \).

The proof theory we propose aims at offering an efficient method for reasoning about
the deontic component of each license and, given that method, for combining different
licenses, checking their compatibility, and establishing what deontic conclusions can be
drawn from the composite license. In other words, if \( l_c = l_1 \odot \cdots \odot l_n \) is the composite
license obtained from \( l_1, \ldots, l_n \), the conclusions derived in the logic for \( l_1, \ldots, l_n \) are also
used to establish those that hold in \( l_c \).

**Formal language and basic concepts** The basic language is defined as follows. Let
\( \text{Lic} = \{ l_1, l_2, \ldots, l_n \} \) be a finite set of licenses. Given a set \( \text{PROP} \) of propositional atoms,
the set of literals \( \text{Lit} \) is the set of such atoms and their negation; as a convention, if \( q \)
is a literal, \( \lnot q \) denotes the complementary literal (if \( q \) is a positive literal \( p \) then \( \lnot q = \lnot p \); and if \( q = \lnot p \), then \( \lnot q \) is \( p \)). Let us denote with \( \text{MOD} = \{ O, P, F \} \) the set of basic
deontic modalities. The set \( \text{ModLit} \) of modal literals is defined as follows: i) if \( X \in \text{MOD} \)
and \( l \in \text{Lit} \), then \( Xl \) and \( \lnot Xl \) are modal literals, ii) nothing else is a modal literal.

Every rule is of the type \( r : A(r) \Rightarrow C(r) \), where: \( r \) is a unique identifier for the rule;
\( A(r) = \{ a_1, \ldots, a_n \} \), the antecedent is a set literal if \( r \) is an ontology rule, and a set of
modal literals and literals if \( r \) is a deontic rule; \( C(r) \) the consequent is a literal; if \( r \) is a
deontic rule \( Y \in \text{MOD} \) represents the type of conclusion obtained\(^5\) and \( x \in \text{Lic} \) indicates to which license the rule refers to; \( Y \) and \( x \) are not used for ontology rules.

The intuition behind the different arrows is the following. **Strict rules** have the form \( a_1, \ldots, a_n \Rightarrow_b \beta \). **Defeasible rules** have the form \( a_1, \ldots, a_n \Rightarrow_b^\dagger \beta \). A rule of the form \( a_1, \ldots, a_n \Rightarrow_b^\dagger \beta \) is a defeater. Analogously, for ontology rules, where arrows do not have superscripts and subscripts. The three types of rules establish the strength of the relationship. Strict rules provide the strongest connection between a set of premises and their conclusion: whenever the premises are deemed as indisputable so is the conclusion. Defeasible rules allow to derive the conclusion unless there is evidence for its contrary. Finally, defeaters suggest that there is a connection between its premises and the conclusion not strong enough to warrant the conclusion on its own, but such that it can be used to defeat rules for the opposite conclusion.

A multi-license theory is the knowledge base which is used to reason about the applicability of license rules under consideration.

**Definition 1** A multi-license theory is a structure \( D = (F, L, R^c, \{R^O_l\}_{l \in \text{Lic}}, \succ) \), where \( F \subseteq \text{Lit} \cup \text{ModLit} \) is a finite set of facts; \( L \subseteq \text{Lic} \) is a finite set of licenses; \( R^c \subseteq \text{Lic} \) is a finite set of ontology rules; \( \{R^O_l\}_{l \in \text{Lic}} \) is finite family of sets of obligation rules; \( \succ \) is an acyclic relation (called superiority relation) defined over \( (R^c \times R^c) \cup (R^O \times R^O) \), where \( R^O, R^O_l \subseteq \{R^O_l\}_{l \in \text{Lic}} \).

\( R[b] \) and \( R^X[b] \) with \( X \in \{c, O\} \) denote the set of all rules whose consequent is \( b \) and of all rules (of type \( X \)). Given a set of rules \( R \) the sets \( R_s, R_d, \) and \( R_{dft} \) denote, respectively, the subsets of \( R \) of strict rules, defeasible rules, and defeaters.

**Proof theory** A proof \( P \) of length \( n \) is a finite sequence \( P(1), \ldots, P(n) \) of tagged literals of the type \( +\Delta^X q, -\Delta^X q, +\partial^X q, \) and \( -\partial^X q \), where \( X \in \{ c, Y \} \). The proof conditions below define the logical meaning of such tagged literals. As a conventional notation, \( P(1..i) \) denotes the initial part of the sequence \( P \) of length \( i \). Given a multi-license theory \( D, +\Delta^X q \) means that literal \( q \) is provable in \( D \) with the mode \( X \) using only facts and strict rules, \(-\Delta^X q \) that it has been proved in \( D \) that \( q \) is not definitely provable in \( D \) with the mode \( X \), \( +\partial^X q \) that is defeasibly provable in \( D \) with the mode \( X \), and \(-\partial^X q \) that \( q \) is not defeasibly provable in \( D \) with the mode \( X^7 \).

Given \( \# \in \{ \Delta, \partial \} \), \( P = P(1), \ldots, P(n) \) is a proof for \( p \) in \( D \) for the license \( l \) iff \( P(n) = +\# p \) when \( p \in \text{Lit} \), \( P(n) = +\#Y/q \) when \( p = Xq \in \text{ModLit} \) and \( P(n) = -\#Y/q \) when \( p = Yq \in \text{ModLit} \).

The proof conditions aim at determining what conclusions can be obtained within composite licenses by using the source licenses.

We concentrate here on deontic effects of licenses, thus working on the obligations, prohibitions, permissions entailed by the composition of a given set of licenses (instead of the composition of the clauses). In [17,10], OR- and AND-compositions were basically characterized as follows:

- **OR-composition**: \( l_c \) entails a deontic effect if there is at least one license that entails such effect (and no license prevents it).
- **AND-composition**: \( l_c \) entails a deontic effect if all licenses entail it.

---

\(^5\)We will see why we do not need rules for prohibitions and permissions.

\(^6\)Notice that we may have that \( l = l' \).

\(^7\)As we will see, we shall adopt a reading of permissions according to which they can only be defeasible. Hence, we will not define the cases \( \pm\Delta^Y q \) where \( Y = P \).
In this paper, we adopt another approach and associate the different heuristics to the derivation of different deontic effects. In particular, OR-composition allows to establish what obligations hold in the composite license: something is obligatory if there is at least one license supporting it. For permissions, instead, AND-composition is the case, as to prove that something is permitted in the composite license we have either to prove that this is the case in all licenses or to exclude that the opposite is obligatory in some license.

Some notational conventions and concepts that we will use throughout the remainder of this section: i) let \( L_c = L_1 \odot \cdots \odot L_n \) be any composite license that can be obtained from the set of licenses \( L_c = \{L_1, \ldots, L_n\} \subseteq L \); ii) let \( X, Y \in \text{MOD} \).

As usual with Defeasible Logic, we have proof conditions for the monotonic part of the theory (proofs for the tagged literals \( \pm \Delta^F p \)) and for the non-monotonic part (proofs for the tagged literals \( \pm \Delta^D p \)). To check licenses’ compatibility and compose them means to apply the proof conditions of the logic to a multi-license where the set of licenses is \( L = L_c \). Since the proof theory for the ontology component (\( \pm \Delta^F p \) and \( \pm \Delta^D p \)) is the one for standard Defeasible Logic we will omit it and refer the reader to [1]. For \# \in \{\Delta, \partial\} and \( Y \in \{O, P, F\} \), notice that conditions governing conclusions for the composite license \( L_c \) and for each license \( L_i \) interplay recursively: indeed, we may use a conclusion for \( L_i \) to fire a rule in \( L_c \).

**OR-composition and Obligations** Let us first define the condition for monotonic derivations of the obligations in each license \( L_i \) and the condition for the monotonic derivations of the obligations in the composite license \( L_c \): this second case is a first illustration of the OR-composition heuristics. Assume \( x \in \{c, i\} \).

\begin{align*}
+\Delta^O_c: \text{If } P(n + 1) = +\Delta^O_c q \text{ then,} \\
(1) & \exists q \in F \text{ or} \\
(2) & \forall a, X, b, Yd \in A(r): \exists q \in F \text{ or} \\
+\Delta^A a, +\Delta^X b, -\Delta^V d \in P(1..n) & +\Delta^D a, +\Delta^X b, -\Delta^V d \in P(1..n)
\end{align*}

Definite proof conditions for prohibitions can be simply obtained:

\begin{align*}
+\Delta^O_c^P: \text{If } P(n + 1) = +\Delta^O_c^P q \text{ then,} \\
+\Delta^A a, +\Delta^X b, -\Delta^V d \in P(1..n).
\end{align*}

A second illustration of the OR-composition is offered in the defeasible derivations of the obligations in \( L_c \):

\begin{align*}
+\Delta^O_d: \text{If } P(n + 1) = +\Delta^O_d q \text{ then,} \\
(1) & \exists q \in F \text{ or} \\
(2) & \forall a, X, b, Yd \in A(r): +\Delta^A a, +\Delta^X b, -\Delta^V d \in P(1..n) \text{ and} \\
(2.2) & \exists q \in F \text{ such that} \\
(2.2.1) & \exists r \in R^O_d \text{ such that} \\
(2.2.2) & \forall a, X, b, Yd \in A(r): +\Delta^A a, +\Delta^X b, -\Delta^V d \in P(1..n) \text{ and} \\
(2.2.3) & \forall s \in R^O_d[\sim:\sim q], \text{ either} \\
(2.2.3.1) & \exists a \in A(s) \text{ or } Xb \in A(s) \text{ or } Y \in A(s): \\
-\Delta^A a \in P(1..n), \text{ or } -\Delta^X b \in P(1..n), \text{ or } +\Delta^V d \in P(1..n); \text{ or} \\
(2.2.3.2) & \exists q \in F \text{ such that} \\
+\Delta^A a, +\Delta^X b, -\Delta^V d \in P(1..n), \text{ and } t > s.
\end{align*}

\(^8\) For space reasons, we will omit in the remainder the all the other proof conditions for the deontic effects in each license, defeasible conditions for prohibitions, and all the negative proof conditions, i.e., for \( -\Delta^O_d, -\Delta^O_c, -\Delta^O_c^P, -\Delta^O_d^P, -\Delta^P, -\Delta^P_c, -\Delta^P_c^P, -\Delta^P_d, -\Delta^P_d^P, \) negative conditions that can all be obtained from positive conditions applying the so-called Principle of Strong Negation [7].
As usual in standard Defeasible Logic, to show that a literal \( q \) is defeasibly provable we have two choices: (1) we show that \( q \) is already definitely provable; or (2) we need to argue using the defeasible part of a multi-license theory \( D \). For this second case, some (sub)conditions must be satisfied. First, we need to consider possible reasoning chains in support of \( \sim q \) with the modes \( l_{\Delta} \) and \( X^\Delta \), and show that \( \sim q \) is not definitely provable with that mode (2.1 below). Second, we require that there must be a strict or defeasible rule with mode at hand for \( q \) which can apply (2.2 below). Third, we must consider the set of all rules which are not known to be inapplicable and which permit to get \( \sim q \) with the mode under consideration (2.3 below). Essentially, each rule \( s \) of this kind attacks the conclusion \( q \). To prove \( q, s \) must be counterattacked by a rule \( t \) for \( q \) with the following properties: i) \( t \) must be applicable, and ii) \( t \) must prevail over \( s \). Thus each attack on the conclusion \( q \) must be counterattacked by a stronger rule. In other words, \( r \) and the rules \( t \) form a team (for \( q \) that defeats the rules \( s \).

**AND-composition and Permissions** The concept of permission is much more elusive (for a discussion, see, e.g., [14]). Here, we minimize complexities by adopting perhaps the two simplest options among those discussed in [7]. Such options model permissions either as obtained

1. when it is possible to show that the opposite obligations are not provable; or
2. from permissive norms with defeaters for obligations; a defeater like \( a_1, \ldots, a_n \sim^{l_{\Delta}} q \) states that some \( q \) is permitted (Pq) in the license \( l_{\Delta} \), since it is meant to block deontic defeasible rules for \( \sim q \), i.e., rules supporting \( O \sim q \).

The first type of permissions corresponds to the so-called weak permissions, according to which some \( q \) is permitted (Pq) because it can be obtained from the fact that \( \sim q \) is not provable as mandatory [19]. The second type of permissions is just one way for modeling explicit permissive clauses for proving Pq (strong permissions of \( q \)): for an extensive treatment of defeasible permissions, see [6]. This reading suggests that permissions are essentially defeasible.

**Permission, version I (Weak Permission)**

\[
\Delta^\mathcal{P}_t: \text{If } P(n + 1) = +\Delta^\mathcal{P}_t q \text{ then } (1) -\Delta^\mathcal{P}_t \sim q \in P(1..n).
\]

The first type of permission might be useful for combination for ‘public domain’ type of license, meaning, that unless explicitly obliged or forbidden data can be used freely.

**Permission, version II (Strong Permission)**

\[
\Delta^\mathcal{P}_t: \text{If } P(n + 1) = +\Delta^\mathcal{P}_t q \text{ then } (1) (1.1) -\Delta^\mathcal{P}_t \sim q \in P(1..n) \text{ and }
\]

(1.2) \( \forall r \in \text{Lic,} \)

(1.2.1) \( \exists r \in \mathcal{R}_{\mathcal{L}_A} [q]: \forall a, Xb, -Yd \in A(r): +\Delta^\mathcal{P}_t a, +\Delta^\mathcal{P}_t b, -\Delta^\mathcal{P}_t d \in P(1..n) \text{ and } (\text{either})
\]

(1.2.2) \( \forall l \in \text{Lic,} \forall r \in \mathcal{R}_{\mathcal{L}_A} [q] \)

(1.2.3.1) \( \exists a \in A(s) \) or \( Xb \in A(s) \) or \( \sim Y \in A(s) \):

\( -\Delta^\mathcal{P}_t a \in P(1..n), \) or \( -\Delta^\mathcal{P}_t b \in P(1..n), \) or \( +\Delta^\mathcal{P}_t d \in P(1..n); \) or

(1.2.3.2) \( \forall l \in \text{Lic,} \exists r \in \mathcal{R}_{\mathcal{L}_A} [q]: \forall a, Xb, -Yd \in A(r),
\]

\( +\Delta^\mathcal{P}_t a, +\Delta^\mathcal{P}_t b, -\Delta^\mathcal{P}_t d \in P(1..n), \) and \( t \succ s. \)

Notice that the logic presented here is a variant of the one developed in [9,8]. On account of this fact, results of soundness and linear computational complexity can be directly imported here [17,10].

Let us consider now an example that illustrate some aspects of the proof theory, and how the heuristics are used.
Example 2 Consider two datasets published on the LOD cloud\textsuperscript{9} associated to licenses \( l_1 \) and \( l_2 \), respectively. License \( l_1 \) permits Derivative and obliges for Share-Alike, while license \( l_2 \) prohibits Derivative, permits Reproduction, and obliges for Notice.

\[
R^{O_1} = \{ r_1 : \Rightarrow l_1 \text{ Share-Alike}, \quad r_2 : \Rightarrow l_1 \text{ Derivative} \} \\
R^{O_2} = \{ r_3 : \Rightarrow l_2 \sim \text{Derivative}, \quad r_4 : \Rightarrow l_2 \text{ Notice}, \quad r_5 : \Rightarrow l_2 \text{ Reproduction} \}
\]

The data publisher has to decide which heuristics better suits her own needs such that the composite license protects as desired the reuse of the released data. First, she needs to include the obligations present in each single license (Share-Alike, Notice) to be compliant with their normative semantics. Thus OR-composition is used to compose obligations. Concerning permissions (Derivative, Reproduction), she has to check that every single license includes the specific permission, thus adopting AND-composition. Given that license \( l_2 \) obliges for \( \sim \text{Derivative} \) (i.e., prohibits Derivative), we cannot include such permission in the composite license. Hence, \( + \partial^{O_1} \text{ Share-Alike}, + \partial^{O_1} \text{ Notice}, \) and \( + \partial^{P_1} \text{ Reproduction} \).

2. Heuristics’ implementation

In this section, we analyze the composition heuristics we have previously defined, further evaluating them using the SPINdle defeasible reasoner.

We show now how the heuristics we propose can be used to check the compatibility and combine four real world licenses, widely adopted in the Linked Open Data scenario, using the SPINdle reasoner. The licenses we consider are the Creative Commons Public Domain Mark 1.0\textsuperscript{10}, the OS Open Data license\textsuperscript{11}, the Creative Commons Attribution-NoDerivs license\textsuperscript{12}, and the Creative Commons Attribution-NonCommercial-ShareAlike\textsuperscript{13}. The deontic component of such licenses is as follows:

- **CC PDM**
  - Permissions: Reproduction, Distribution, Derivative Works.
  - OS OpenData
  - Permissions: Reproduction, Distribution, Derivative Works.
  - Obligations: Notice, Attribution.
- **CC-BY-ND**
  - Permissions: Reproduction, Distribution, Derivative Works.
  - Obligations: Notice, Attribution, Share Alike.
  - Prohibitions: Non Commercial.
- **CC-BY-NC-SA**
  - Permissions: Reproduction, Distribution.
  - Obligations: Notice, Attribution.

Notice that licenses CC PDM, OS OpenData and CC-BY-ND allow to make commercial use of the work, i.e., the permission is not explicitly stated but it is ensured by

\begin{itemize}
  \item \text{http://lod-cloud.net/}
  \item \text{http://creativecommons.org/publicdomain/mark/1.0/}
  \item \text{http://www.ordnancesurvey.co.uk/docs/licences/os-opendata-licence.pdf}
  \item \text{http://creativecommons.org/licenses/by-nd/3.0/}
  \item \text{http://creativecommons.org/licenses/by-nc-sa/3.0/}
\end{itemize}
the absence of the obligation for Non Commercial, and that license CC-BY-ND does not permit Derivative Works even if such prohibition is not mentioned.

As mentioned before, we have to verify the compatibility of different licensing terms by composing them into a unique composed theory. The implementation is based on the two transformations proposed in [10] for AND- and OR-compositions.

\[
\text{tor}(r) = \begin{cases} 
  r : A(r) \leftrightarrow p & r \in R_c \\
  r : A(r) \rightarrow \alpha & r \in R^{O_i}_c, l_i \in \text{Lic} \\
  r : A(r) \Rightarrow \alpha & r \in R^{O}_c, l_i \in \text{Lic} \\
  r : A(r) \Rightarrow -\alpha \sim p & r \in R^{P}_d, l_i \in \text{Lic}
\end{cases}
\]

For the OR-heuristic we can use the \text{tor} transformation as it is. For weak permission SPINdle generates the conclusion +\partial^P q as soon as a conclusion \neg\partial^{O_i} \sim q is generated (alternatively we could add the following set of rules \{\neg\partial^{O_i} \sim q \mid l_i \in \text{Lic}, q \in \text{Lit}\} but this transformation could lead to many additional rules, with the consequent degrade in the performance).

For AND-composition, we consider the following transformations:

\[
\text{tando}(r) = \{ r_{ij} : A(r) \sim \alpha_i C(r) \mid r \in R^{O_i}_c, \} \cup \{ r : A(r) \Rightarrow -\alpha \sim C(r) \mid r \in R^{O}_d, \} \cup \{ r \mid r \in R^{P}_d \} \cup \{ q_i : O^{i} q \Rightarrow \partial^P q, q_i : -O^{i} q \Rightarrow \partial^P q, q_i : -O^{i} q \Rightarrow \partial^P q \mid l_i \in \text{Lic}, c \in \text{Lit} \}
\]

\[
\text{tandsp}(r) = \{ p_q : P^p q, P^p q, P^p q \Rightarrow \partial^P q, q_i : -O^{i} q \Rightarrow \partial^P q, q_i : -O^{i} q \Rightarrow \partial^P q \mid q \mid l_i \in \text{Lic}, c \in \text{Lit} \}
\]

The transformations \text{tor, tando, tandsp} and \text{tandwp} are used to map rules in different licenses to the composed theory. The combination of \text{tando} and \text{tandsp} gives us the transformation given in [10] to handle the AND-composition. However, we can use \text{tando} and \text{tandwp} to compute the AND-composition with weak permission instead of strong permission. In addition it is possible to use the \text{tor} and \text{tandsp} (or \text{tandwp}) to model the hybrid composition P-AND/OR-O discussed in Example 2. In fact, for the OR-heuristic, to prove +\partial^P p we only need to prove that O^p \Rightarrow p is provable by a license; while in the AND-heuristic we have to show that the literal is provable in all licenses. However, the case for permission is a bit different as, in addition to the condition above, it also requires least one license permits p. To this end, we have implemented a \text{theory composer} to apply the transformations above to different licenses and compose them into a single defeasible theory, before passing it to SPINdle for reasoning.

Even though one may argue that modifying the inference engine or devise a new reasoning algorithm to solve our problem can achieve a better performance. However, our approach can give us, at least, three advantages: (1) the implementation of the licenses composer is a lot simpler when comparing with modifying the reasoning engine or implementing a new inference algorithm; (2) instead of computing the conclusions for permission (AND-composition) and obligation (OR-composition) with two separate inferences (as described in [10]), the proposed approach enable us to compute all conclusions with one single inference, which is simpler and can reduce the time required for initializing the reasoning engine; (3) we can utilize the features provided by SPINdle to capture different intuitions, i.e., ambiguity propagation, well-founded semantics, or their combinations, in the future without additional work.
Table 1 shows the composition and reasoning time used for composing the four licenses mentioned in the previous section. As expected, there is not much different in the time required to generate the composed theory (and reasoning) with or without the OR-composition, and there is a minimal overhead with the P-AND/OR-O-composition. However, the cases with AND-composition (the latter two cases) is a bit worse (but still in acceptable range) as plethora of propositions are added to the composed theory to ensure consistency among different licenses.

### 3. Related Work

The closest set of related work is in the area of contract compatibility and composition in the context of services composition. Comerio [3] analyses which kind of qualitative and quantitative heuristics to be used for contact composition in the context of service composition. Qualitative heuristics include AND- and OR-composition heuristics plus the Constraining Value one where the most constraining value among the ones offered by the contracts of the services to be composed is included in the composite service. Quantitative heuristics, instead, include MIN, MAX, AVG and SUM (the composite contract offers the minimum (resp. maximum, mean, sum) among the values offered by the contracts of the single services involved in the composition. In this paper, we do not consider quantitative heuristics, and we propose a fine grained analysis of the ADN- and OR-composition heuristics and how to combine them with respect to the deontic component of the single licenses to compose. Gangadharan et al. [4] address the issue of service license composition and compatibility analysis, specifying a matchmaking algorithm which verifies whether two service licenses are compatible. In case of a positive answer, the services can be composed and the framework determines the license of the composite service. Truong et al. [18] address a similar problem concerning data contracts: in contracts composition, first the comparable contractual terms from the different contracts are retrieved, and second an evaluation of the new contractual terms for the data mash-up is addressed. In this paper, we concentrate on the evaluation of the composition heuristics and how they can be composed to better suit the data publisher’s needs.

Other related work concerns reasoning about licenses [16,5] or licensing issues in the Semantic Web scenario [15,12]. However, they do not address the issue of licenses composition which is the goal of this paper.

### 4. Concluding remarks

The development of new models and tools for the advanced management of legal information and knowledge in the Web of Data raises open challenges. In this paper, we...
have formally defined and evaluated two heuristics for combining the licensing terms of a set of licenses in a compliant way, using our defeasible deontic logic. The OR- and AND-composition heuristics have been coded into the SPINdle defeasible logic reasoner, and the system’s evaluation show the applicability of the proposed formal approach in the Web of Data scenario. Several future directions will be considered. First, we will enlarge the set of composition heuristics taking into account also qualitative ones and the Constraining-value heuristic. Second, we are planning the development of a standalone licensing module able to check the compatibility of a set of licenses in an automated way, and to generate the machine-readable composite license. Finally, even if our framework allows to reason about certain characteristics of licenses, e.g., whether attribution is required or commercial usages are permitted, it is still an open problem the fact that there is no uniform, cross-national definition of essential legal terms. We will further investigate this open issue.

References


