Justified Terminological Reasoning

Thomas Studer

Institut für Informatik und angewandte Mathematik
Universität Bern, Swizerland
tstuder@iam.unibe.ch
http://www.iam.unibe.ch/∼tstuder

Abstract. Justification logics are epistemic logics that include explicit justifications for an agent's knowledge. In the present paper, we introduce a justification logic $\mathcal{JALC}$ over the description logic $\mathcal{ALC}$. We provide a deductive system and a semantics for our logic and we establish soundness and completeness results. Moreover, we show that our logic satisfies the so-called internalization property stating that it internalizes its own notion of proof. We then sketch two applications of $\mathcal{JALC}$: (i) the justification terms can be used to generate natural language explanations why an $\mathcal{ALC}$ statement holds and (ii) the terms can be used to study data privacy issues for description logic knowledge bases.

Keywords: Justification logic, description logic, inference tracking, explanations, data privacy

1 Introduction

Description logics [7] are a variant of modal logic that is used in knowledge representation to model the universe of discourse of an application domain and to reason about it. In the present paper we study the basic logic $\mathcal{ALC}$ which is the minimal description logic that is closed under boolean connectives. Our aim is to extend $\mathcal{ALC}$ with so-called justification terms yielding a justification logic over $\mathcal{ALC}$.

Justification logics [4] are epistemic logics that feature explicit justifications for an agent's knowledge and they allow to reason with and about these justifications. The first logic of this kind, the *logic of proofs* LP, has been developed by Artemov [2, 3] to solve the problem of a provability semantics for S4. Since then many applications of justification logics have been studied. For instance, these logics have been used to create a new approach to the logical omniscience problem [6], to explore self-referential proofs [18], to study evidence tracking [5], and to investigate the role of the announcement as a justification in public announcement logics [10, 11].

Instead of the simple statement A *is known*, denoted $\Box A$, justification logics reason about justifications for knowledge by using the construct $[t]A$ to formalize t *is a justification for* A, where the evidence term t can be viewed as an informal justification or a formal mathematical proof depending on the application.

Evidence terms are built by means of operations that correspond to the axioms of S4 as Fig. 1 shows.

S4 axioms	LP axioms	
$\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$	$[t](\phi \to \psi) \to ([s]\phi \to [t \cdot s]\psi)$	(application)
$\Box\phi \to \phi$	$[t]\phi \to \phi$	(reflexivity)
$\Box\phi \to \Box\Box\phi$	$[t]\phi \to [!t][t]\phi$	(inspection)
	$[t]\phi \vee [s]\phi \to [t + s]\phi$	(sum)

Fig. 1. Axioms of S4 and LP

Internalization is a key property for any justification logic. It states that for each derivation $\mathcal{D}$ of a theorem A of the logic in question, there is a step-by-step construction that transforms $\mathcal{D}$ into a term $t_{\mathcal{D}}$ such that $[t_{\mathcal{D}}]A$ is also a theorem of the logic. Therefore, the term $t_{\mathcal{D}}$ describes why, according to the logic, A must hold.

In this paper, we introduce a new logic $\mathcal{JALC}$ of justified $\mathcal{ALC}$ - that is we extend $\mathcal{ALC}$ by justification terms - and study its main features. We start with a brief introduction to the description logic $\mathcal{ALC}$. In Section 3, we introduce the language of $\mathcal{JALC}$ and present a deductive system for it. We then prove the so-called Lifting lemma saying that $\mathcal{JALC}$ internalizes its own notion of proof. We define a semantics for $\mathcal{JALC}$ and establish soundness and completeness of the deductive system in Section 4. Then a section about applications follows where we give a detailed example of internalization. We make use of this example to illustrate how internalization can be applied to

1. the problem of generating natural language explanations and
2. the problem of data privacy for $\mathcal{ALC}$ knowledge bases.

In Section 6, we present related work. Finally we conclude the paper and mention some further research directions.

2 The description logic $\mathcal{ALC}$

In this section, we briefly recall the main definitions concerning $\mathcal{ALC}$. We will not only study subsumption but also introduce formulas for $\mathcal{ALC}$, which will be useful when we add justification terms. Also compactness of $\mathcal{ALC}$ will play an important role later.

Definition 1 (Concept). *We start with countably many* concept names *and* role names. *The set of* concepts *is then defined inductively as follows:*

1. *Every concept name is a concept.*
2. *If C and D are concepts, R is a role name, then the following expressions are concepts:*

$$\neg C, \quad C \sqcap D, \quad \forall R.C.$$

As usual, we define $C \sqcup D := \neg(\neg C \sqcap \neg D)$, $\exists R.C := \neg\forall R.\neg C$, and $\top := A \sqcup \neg A$ for some fixed concept name A.

Definition 2 ($\mathcal{L}_A$ formula).

1. *If C and D are concepts, then $C \sqsubseteq D$ is an (atomic) $\mathcal{L}_A$ formula.*
2. *If ϕ and ψ are $\mathcal{L}_A$ formulas, then the following expressions are $\mathcal{L}_A$ formulas:*

$$\neg\phi, \quad \phi \wedge \psi.$$

Definition 3 ($\mathcal{ALC}$ interpretation). *An $\mathcal{ALC}$* interpretation *is a pair $\mathcal{I} = (\Delta_\mathcal{I}, \cdot^\mathcal{I})$ where $\Delta_\mathcal{I}$ is a non-empty set called the* domain *of $\mathcal{I}$ and $\cdot^\mathcal{I}$ maps each concept name A to a subset $A^\mathcal{I} \subseteq \Delta_\mathcal{I}$ and each role name R to a binary relation $R^\mathcal{I}$ on $\Delta_\mathcal{I}$. An interpretation is extended to non-atomic concepts as follows.*

1. $(\neg C)^\mathcal{I} = \Delta_I \setminus C^\mathcal{I}$
2. $(C \sqcap D)^\mathcal{I} = C^\mathcal{I} \cap D^\mathcal{I}$
3. $(\forall R.C)^\mathcal{I} = \{d \in \Delta_I \ : \ \forall d' \in \Delta_I (R^\mathcal{I}(d, d') \to C^\mathcal{I}(d'))\}$

Definition 4 ($\mathcal{ALC}$ satisfiability). *We inductively define when an $\mathcal{L}_A$ formula is satisfied in an $\mathcal{ALC}$ interpretation $\mathcal{I}$.*

1. *$\mathcal{I} \models C \sqsubseteq D$ iff $C^\mathcal{I} \subseteq D^\mathcal{I}$*
2. *$\mathcal{I} \models \neg\phi$ iff not $\mathcal{I} \models \phi$*
3. *$\mathcal{I} \models \phi \wedge \psi$ iff $\mathcal{I} \models \phi$ and $\mathcal{I} \models \psi$*

We say an $\mathcal{L}_A$ formula ϕ is $\mathcal{ALC}$ valid *(and write $\models_{\mathcal{ALC}} \phi$) if for all interpretations $\mathcal{I}$ we have $\mathcal{I} \models \phi$. For a set of $\mathcal{L}_A$ formulas Φ, we write $\models_{\mathcal{ALC}} \Phi$ if for all $\phi \in \Phi$ we have $\models_{\mathcal{ALC}} \phi$*

By the work of Schild [22], we know that $\mathcal{ALC}$ can be seen as a notational variant of the multi-modal logic K_n. Thus we can transfer results from K_n to $\mathcal{ALC}$. In particular, we immediately get the following lemma about compactness of $\mathcal{ALC}$ from compactness of K_n.

Lemma 1 ($\mathcal{ALC}$ compactness). *$\mathcal{ALC}$ is compact: for any set Φ of $\mathcal{L}_A$ formulas we have*

$$\models_{\mathcal{ALC}} \Phi \text{ if and only if for all finite subsets } \Phi' \subseteq \Phi \text{ we have } \models_{\mathcal{ALC}} \Phi'.$$

3 Syntax of $\mathcal{JALC}$

The aim of this section is to introduce the language of $\mathcal{JALC}$, the logic of justified $\mathcal{ALC}$. Then we present a deductive system for $\mathcal{JALC}$ and show that it satisfies internalization.

Definition 5 (Terms). *We fix countable sets of constants* Con *and variables* Var*, respectively. Terms t are now built according to the following grammar*

$$t ::= x \mid c \mid t \cdot t \mid t + t \mid !t$$

where x is a variable and c is a constant. Tm *denotes the set of terms.*

Definition 6 ($\mathcal{L}_J$ Formula).

1. *If C and D are concepts, then $C \sqsubseteq D$ is an (atomic) $\mathcal{L}_J$ formula.*
2. *If ϕ and ψ are $\mathcal{L}_J$ formulas and t is a term, then the following expressions are $\mathcal{L}_J$ formulas:*

$$\neg\phi, \quad \phi \wedge \psi, \quad [t]\phi.$$

We denote the set of $\mathcal{L}_J$ formulas by Fml_J. As usual, we define

$$\phi \vee \psi := \neg(\neg\phi \wedge \neg\psi) \quad \text{and} \quad \phi \to \psi := \neg\phi \vee \psi.$$

Note that every $\mathcal{L}_A$ formula also is an $\mathcal{L}_J$ formula.

Definition 7 ($\mathcal{JALC}$ deductive system). *The* axioms of $\mathcal{JALC}$ *consist of all* Fml_J *instances of the following schemes:*

1. *All valid $\mathcal{L}_A$ formulas ϕ, i.e. for which $\models_{\mathcal{ALC}} \phi$ holds*
2. $[t](\phi \to \psi) \to ([s]\phi \to [t \cdot s]\psi)$ *(application)*
3. $[t]\phi \vee [s]\phi \to [t + s]\phi$ *(sum)*
4. $[t]\phi \to \phi$ *(reflexivity)*
5. $[t]\phi \to [!t][t]\phi$ *(introspection)*

A constant specification $\mathcal{CS}$ *is any subset*

$$\mathcal{CS} \subseteq \{[c]\phi \ : \ c \text{ is a constant and } \phi \text{ is an axiom of } \mathcal{JALC}\}.$$

A constant specification $\mathcal{CS}$ is called axiomatically appropriate *if for every axiom ϕ of $\mathcal{JALC}$, there is a constant c such $[c]\phi \in \mathcal{CS}$.*

The deductive system $\mathcal{JALC}(\mathcal{CS})$ *is the Hilbert system that consists of the above axioms of $\mathcal{JALC}$ and the following rules of* modus ponens *and* axiom necessitation*:*

$$\frac{\phi \quad \phi \to \psi}{\psi} \ , \qquad \frac{}{[c]\phi} \ \text{ where } [c]\phi \in \mathcal{CS}.$$

For a set of $\mathcal{L}_J$ formulas Φ we write $\Phi \vdash_{\mathcal{CS}} \phi$ to state that ϕ is derivable from Φ in $\mathcal{JALC}(\mathcal{CS})$. When the constant specification $\mathcal{CS}$ is clear from the context we will write only $\vdash$ instead of $\vdash_{\mathcal{CS}}$.

We say a set Φ of $\mathcal{L}_J$ formulas is $\mathcal{CS}$ *consistent* if there exists a formula ϕ such that $\Phi \nvdash_{\mathcal{CS}} \phi$. The set Φ is called *maximal $\mathcal{CS}$ consistent* if it is $\mathcal{CS}$ consistent but has no proper extension that is $\mathcal{CS}$ consistent.

The Lifting Lemma states the $\mathcal{JALC}$ internalizes its own notion of proof. This is a standard property that any justification logic should have.

Lemma 2 (Lifting lemma). *Let $\mathcal{CS}$ be an axiomatically appropriate constant specification. If*

$$[x_1]\phi_1, \ldots, [x_n]\phi_n, \psi_1, \ldots \psi_m \vdash_{\mathcal{CS}} \chi,$$

then there is a term $t(x_1, \ldots, x_n, y_1, \ldots, y_m)$ such that

$$[x_1]\phi_1, \ldots, [x_n]\phi_n, [y_1]\psi_1, \ldots, [y_m]\psi_m \vdash_{\mathcal{CS}} [t(x_1, \ldots, x_n, y_1, \ldots, y_m)]\chi.$$

Proof. Let Φ be the set $\{[x_1]\phi_1, \ldots, [x_n]\phi_n, [y_1]\psi_1, \ldots, [y_m]\psi_m\}$. We proceed by induction on the length of the derivation of χ and distinguish the following cases.

1. χ is an axiom of $\mathcal{JALC}$. Since $\mathcal{CS}$ is axiomatically appropriate, there is a constant c such that $[c]\chi \in \mathcal{CS}$. Thus $\vdash [c]\chi$ follows by axiom necessitation.
2. χ is $[x_i]\phi_i$ for some i. We find $[x_i]\phi_i \vdash [!x_i][x_i]\phi_i$ by *(introspection)* and modus ponens.
3. χ is ψ_i for some i. We immediately have $[y_i]\psi_i \vdash [y_i]\psi_i$.
4. χ follows from $\psi \to \chi$ and ψ by modus ponens. By the induction hypothesis there are terms $t_1(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $t_2(x_1, \ldots, x_n, y_1, \ldots, y_m)$ such that
$$\Phi \vdash [t_1(x_1, \ldots, x_n, y_1, \ldots, y_m)](\psi \to \chi)$$
and
$$\Phi \vdash [t_2(x_1, \ldots, x_n, y_1, \ldots, y_m)]\psi.$$
Thus
$$\Phi \vdash [t_1(x_1, \ldots, x_n, y_1, \ldots, y_m) \cdot t_2(x_1, \ldots, x_n, y_1, \ldots, y_m)]\chi$$
follows from *(application)* and applying modus ponens twice.
5. χ is the conclusion of axiom necessitation. Then χ has the form $[c]\chi'$. Thus we find $\vdash [!c][c]\chi'$ by *(introspection)* and modus ponens. □

4 Semantics of $\mathcal{JALC}$

The semantics of $\mathcal{JALC}$ is based on so-called F-models [16] for justification logics. These models consist of a Kripke frame and an evidence function specifying for each state which terms are admissible evidence for which formulas. This evidence function has to satisfy certain closure conditions matching the axioms of $\mathcal{JALC}$. Finally, we assign to each state an $\mathcal{ALC}$ interpretation that gives meaning to concept and role names.

Definition 8 ($\mathcal{JALC}$ model). *A $\mathcal{JALC}$ model meeting a constant specification $\mathcal{CS}$ is a tuple $\mathcal{M} = (W, \lhd, \mathcal{E}, I)$ where*

1. *W is a non-empty set (of* states*)*
2. *$\lhd$ is a binary relation on W that is transitive and reflexive*
3. *$\mathcal{E}$ is an evidence function $\mathcal{E} : W \times \mathsf{Tm} \to \mathcal{P}(\mathsf{Fml}_J)$ that satisfies the following closure conditions for any states $w, v \in W$:*
   (a) if $[c]\phi \in \mathcal{CS}$, then $\phi \in \mathcal{E}(w, c)$
   (b) if $\lhd(w, v)$, then $\mathcal{E}(w, t) \subseteq \mathcal{E}(v, t)$
   (c) if $(\phi \to \psi) \in \mathcal{E}(w, t)$ and $\phi \in \mathcal{E}(w, s)$, then $\psi \in \mathcal{E}(w, t \cdot s)$
   (d) $\mathcal{E}(w, s) \cup \mathcal{E}(w, t) \subseteq \mathcal{E}(w, s + t)$
   (e) if $\phi \in \mathcal{E}(w, t)$, then $[t]\phi \in \mathcal{E}(w, !t)$
4. *I associates with each $w \in W$ an $\mathcal{ALC}$ interpretation $I(w) = (\Delta_w, \cdot^{\mathcal{I}(w)})$.*

We use the standard notion of satisfiability for F-models. A formula $[t]\phi$ holds at a state w if ϕ holds at all states reachable from w and the term t is admissible evidence for ϕ at w.

Definition 9 (Satisfiability). *We inductively define when a formula is satisfied in a model $\mathcal{M} = (W, \lhd, \mathcal{E}, I)$ at a world $w \in W$.*

1. $\mathcal{M}, w \models C \sqsubseteq D$ *iff* $C^{\mathcal{I}(w)} \subseteq D^{\mathcal{I}(w)}$
2. $\mathcal{M}, w \models \neg\phi$ *iff not* $\mathcal{M}, w \models \phi$
3. $\mathcal{M}, w \models \phi \wedge \psi$ *iff* $\mathcal{M}, w \models \phi$ *and* $\mathcal{M}, w \models \psi$
4. $\mathcal{M}, w \models [t]\phi$ *iff* $\mathcal{M}, w' \models \phi$ *for all* $w' \in W$ *such that* $w \lhd w'$ *and* $\phi \in \mathcal{E}(w, t)$.

We write $\models_{\mathcal{CS}} \phi$ and say that the formula ϕ is *valid with respect to the constant specification* $\mathcal{CS}$ if for all models $\mathcal{M} = (W, \lhd, \mathcal{E}, I)$ that meet $\mathcal{CS}$ and all $w \in W$ we have $\mathcal{M}, w \models \phi$.

As usual, soundness follows by a straightforward induction on the length of $\mathcal{JALC}(\mathcal{CS})$ derivations.

Theorem 1 (Soundness). *Let ϕ be an $\mathcal{L}_J$ formula and $\mathcal{CS}$ a constant specification. We have*

$$\vdash_{\mathcal{CS}} \phi \quad \textit{implies} \quad \models_{\mathcal{CS}} \phi.$$

In the remainder of this section, we will establish completeness of the deductive system $\mathcal{JALC}(\mathcal{CS})$. Our aim is to construct a canonical model. To do so, we first need to show that certain $\mathcal{ALC}$ interpretations exist.

Definition 10. *Let Φ be a set of $\mathcal{L}_J$ formulas. We set*

$$G_\Phi := \{\phi \in \Phi \ : \ \phi \textit{ is an } \mathcal{L}_A \textit{ formula}\}.$$

Lemma 3 (Existence of $\mathcal{ALC}$ interpretations). *There exists a function I_G that maps any consistent set of formulas Φ to an $\mathcal{ALC}$ interpretation $I_G(\Phi)$ such that $I_G(\Phi) \models G_\Phi$.*

Proof. We show that there exists an $\mathcal{ALC}$ interpretation $\mathcal{I}$ such that $\mathcal{I} \models G_\Phi$. We suppose G_Φ is not $\mathcal{ALC}$ satisfiable and aim at a contradiction. By compactness of $\mathcal{ALC}$ there exists a finite subset $\Phi'_G \subseteq \Phi_G$ that is not $\mathcal{ALC}$ satisfiable. That means $\bigwedge \Phi'_G$ is not satisfiable which implies $\models_{\mathcal{ALC}} \neg \bigwedge \Phi'_G$. Therefore, $\neg \bigwedge \Phi'_G$ is an axiom of $\mathcal{JALC}$ and thus

$$\vdash \neg \bigwedge \Phi'_G. \tag{1}$$

Since $\Phi'_G \subseteq \Phi_G$, we also obtain

$$\Phi \vdash \bigwedge \Phi'_G. \tag{2}$$

From (1) and (2) we conclude $\Phi \vdash A$ for any formula A, which contradicts the assumption that Φ is consistent.

Thus for any $\mathcal{JALC}$ consistent set of formulas Φ there exists an $\mathcal{ALC}$ interpretation $\mathcal{I}$ with $\mathcal{I} \models G_\Phi$. We let I_G be a function that chooses for each consistent Φ such an interpretation $\mathcal{I}$. □

Definition 11 (Canonical model). *We define the* canonical model $\mathcal{M} = (W, \lhd, \mathcal{E}, I)$ *meeting a constant specification* $\mathcal{CS}$ *as follows.*

1. *W is the set of all maximal $\mathcal{CS}$ consistent subsets of Fml_J*
2. *$w \lhd v$ if and only if for all $t \in \mathsf{Tm}$, we have $[t]A \in w$ implies $A \in v$*
3. *$\mathcal{E}(w,t) := \{A \in \mathsf{Fml}_J \ : \ [t]A \in w\}$*
4. *$I := I_G$.*

It is standard to show that the canonical model is indeed a $\mathcal{JALC}$ model, meaning that $\lhd$ satisfies the frame conditions and $\mathcal{E}$ satisfies the closure conditions of evidence functions. For details we refer to Fitting [16]. Thus we have the following lemma.

Lemma 4. *The canonical model meeting a constant specification $\mathcal{CS}$ is a $\mathcal{JALC}$ model meeting $\mathcal{CS}$.*

Lemma 5 (Truth lemma). *Let $\mathcal{M}$ be the canonical model meeting a constant specification $\mathcal{CS}$. For all $\mathcal{L}_J$ formulas ϕ and all states w in $\mathcal{M}$, we have*

$$\phi \in w \text{ if and only if } \mathcal{M}, w \models_{\mathcal{CS}} \phi.$$

Proof. Proof by induction on the structure of ϕ. If ϕ is atomic, then we have by Lemma 3

$$\phi \in w \text{ iff } \phi \in G_w \text{ iff } I_G(w) \models \phi \text{ iff } M, w \models_{\mathcal{CS}} \phi.$$

The cases where ϕ is not atomic are standard and follow easily from the closure conditions on the evidence function, again see [16]. □

As usual, the Truth lemma implies completeness of the corresponding deductive system.

Theorem 2 (Completeness). *Let ϕ be an $\mathcal{L}_J$ formula and $\mathcal{CS}$ be a constant specification. We have*

$$\models_{\mathcal{CS}} \phi \quad \text{implies} \quad \vdash_{\mathcal{CS}} \phi.$$

5 Applications

Inference tracking. One distinguished feature of justification terms is that they keep track of the inferences made in a logical derivation. Let us illustrate this with the following example about a business information system storing information about managers and their salaries.

Let Φ be a knowledge base containing the three statements:

1. If a person gets a high salary, then she handles key accounts only

$$\texttt{high} \sqsubseteq \forall \texttt{handles.keyAcc} \quad . \qquad (\phi_1)$$

2. Everyone who handles something gets a high or a low salary

$$\exists \texttt{handles}.\top \sqsubseteq \texttt{high} \sqcup \texttt{low} \quad . \tag{ϕ_2}$$

3. Person 1 handles something that is not a key account

$$\texttt{P1} \sqsubseteq \exists \texttt{handles}.\neg \texttt{keyAcc} \quad . \tag{ϕ_3}$$

From this knowledge base we can derive that Person 1 gets a low salary. That is we have

$$\phi_1, \phi_2, \phi_3 \vdash \texttt{P1} \sqsubseteq \texttt{low}. \tag{3}$$

However, this does not give us any information on how the derivation looks like. We can change this situation by applying the Lifting lemma to (3) which results in

$$[v]\phi_1, [w]\phi_2, [x]\phi_3 \vdash_{\mathcal{CS}} [t]\texttt{P1} \sqsubseteq \texttt{low}.$$

Now the term t will provide explicit information about a derivation of $\texttt{P1} \sqsubseteq \texttt{low}$.

To apply the Lifting lemma we need an axiomatically appropriate constant specification $\mathcal{CS}$. We assume $\mathcal{CS}$ is such that for all concepts A, B, C and role names R, the following are elements of $\mathcal{CS}$:

$$\begin{gathered}
[a](A \sqsubseteq B \to \exists R.A \sqsubseteq \exists R.B), \\
[b](A \sqsubseteq B \to (C \sqsubseteq A \to C \sqsubseteq B)), \\
[c](A \sqsubseteq B \to \neg B \sqsubseteq \neg A), \\
[d](A \sqsubseteq B \sqcup C \to (A \sqsubseteq \neg B \to A \sqsubseteq C)), \\
[e]\neg \texttt{keyAcc} \sqsubseteq \top.
\end{gathered}$$

We now find that from $[v]\phi_1, [w]\phi_2, [x]\phi_3$ the following statements are derivable in $\mathcal{JALC}(\mathcal{CS})$:

$$\begin{gathered}
[a \cdot e]\exists \texttt{handles}.\neg \texttt{keyAcc} \sqsubseteq \exists \texttt{handles}.\top \\
[b \cdot (a \cdot e)](\texttt{P1} \sqsubseteq \exists \texttt{handles}.\neg \texttt{keyAcc} \to \texttt{P1} \sqsubseteq \exists \texttt{handles}.\top) \\
[(b \cdot (a \cdot e)) \cdot x]\texttt{P1} \sqsubseteq \exists \texttt{handles}.\top \\
[b \cdot w](\texttt{P1} \sqsubseteq \exists \texttt{handles}.\top \to \texttt{P1} \sqsubseteq \texttt{high} \sqcup \texttt{low}) \\
[(b \cdot w) \cdot ((b \cdot (a \cdot e)) \cdot x)](\texttt{P1} \sqsubseteq \texttt{high} \sqcup \texttt{low}) \\
[c \cdot v]\exists \texttt{handles}.\neg \texttt{keyAcc} \sqsubseteq \neg \texttt{high} \\
[(b \cdot (c \cdot v)) \cdot x]\texttt{P1} \sqsubseteq \neg \texttt{high} \\
[(d \cdot ((b \cdot w) \cdot ((b \cdot (a \cdot e)) \cdot x))) \cdot ((b \cdot (c \cdot v)) \cdot x)]\texttt{P1} \sqsubseteq \texttt{low}
\end{gathered}$$

In the last line, the justification term

$$(d \cdot ((b \cdot w) \cdot ((b \cdot (a \cdot e)) \cdot x))) \cdot ((b \cdot (c \cdot v)) \cdot x) \tag{4}$$

represents the logical steps that led to the conclusion $\texttt{P1} \sqsubseteq \texttt{low}$. We can now use this justification term for two purposes: to give explanations and to study data privacy.

Explanations. The justification terms can be employed to give a natural language description of the reasoning steps performed in the proof. For instance, since the term c justifies $A \sqsubseteq B \rightarrow \neg B \sqsubseteq \neg A$, we can translate, for instance, $[c \cdot v]\psi$ into

taking the contrapositive of the statement justified by v results in ψ.

Using v from our example, we get

taking the contrapositive of $\texttt{high} \sqsubseteq \forall \texttt{handles}.\texttt{keyAcc}$
gives us $\exists \texttt{handles}.\neg \texttt{keyAcc} \sqsubseteq \neg \texttt{high}$.

Of course, for practical applications it is important to find the right level of abstraction. In a long proof, we do not want to mention every single axiom and every single application of an inference rule that is used. That means we do not give an explanation for every single proof constant and every single application occurring in a proof term. Instead, terms of a certain complexity should be regarded as one unit representing one step in the proof. For example, because the variable x justifies $\texttt{P1} \sqsubseteq \exists \texttt{handles}.\neg \texttt{keyAcc}$ and a, b, e are constants and thus justify logical axioms, we can read

$$[(b \cdot (a \cdot e)) \cdot x]\texttt{P1} \sqsubseteq \exists \texttt{handles}.\top$$

in a more abstract way as

$\texttt{P1} \sqsubseteq \exists \texttt{handles}.\neg \texttt{keyAcc}$ implies
$\texttt{P1} \sqsubseteq \exists \texttt{handles}.\top$ by simple logical reasoning in $\mathcal{ALC}$.

Data Privacy. Inference tracking is also important for applications in the area of data privacy. In privacy aware applications only a part of a given knowledge base is publicly accessible (say via views or via aggregation) and other parts (say containing personally identifiable information) should be kept secret. A violation of privacy occurs if it is possible for an agent to infer some secret information from the public part of the knowledge base.

There are basically two possibilities to prevent such privacy violations: (i) to refuse an answer to a query, that is make the public part of the knowledge base smaller, or (ii) to lie about the answer, that is distort the knowledge base. In both cases it is important to understand what led to the privacy breach. That means to understand how it was possible a secret could be inferred from the public knowledge. Again, if we model this situation in a justification logic, then we can apply the Lifting lemma to obtain a term that tracks the inferences leading to the leaked secret. This term is essentially a blueprint of a derivation of the secret. Thus it contains information about which elements of the published part of the knowledge base are responsible for the privacy violation and this information can be used to alter the knowledge base such that it does no longer leak private data.

Consider our example above about the knowledge base Φ. We assume that $\texttt{P1} \sqsubseteq \texttt{low}$ should be kept secret since it contains information that is related to a specific person. As we have seen before, there is a violation of privacy since we have $\Phi \vdash \texttt{P1} \sqsubseteq \texttt{low}$ and now the question is what part of Φ is responsible for this. The justification term (4) constructed by the Lifting lemma contains the variables v, w, x. This tells us that $\{\phi_1, \phi_2, \phi_3\}$ is a subset of Φ from which the secret can be inferred. Thus to prevent this privacy breach it is a good strategy to restrict access to at least on these three elements. Of course, this does not guarantee privacy since there may be other derivations of the secret that start from a different subset of Φ. Still the justification term provides valuable information for a heuristic to construct a privacy preserving knowledge base.

6 Related work

Modalized description logics. The study of multi-agent epistemic description logics started with the investigations by Laux, Gräber, and Bürckert [19, 17]. In those papers, like in $\mathcal{JALC}$, the modal operators apply only to axioms, but not to concepts. A similar approach for temporalizing (instead of modalizing) logics had earlier been provided by Finger and Gabbay [15]. Baader and Laux [8] present a description logic in which modal operators can be applied to both axioms and concepts. Modalized description logics of this kind have then been investigated in detail, see for example [20].

Explanations. For some reasoning services offered by an ontological information system, users will not only be interested in the result but also in an explanation of it. That is, users will need to understand how deductions were made and what manipulations were done. There are many studies on how to generate explanations. We confine ourselves to mentioning two of them. McGuiness and Pinheiro da Silva [21] give an overview about requirements for answer explanation components of such a system. A very promising approach to provide explanations is based on meta-inferencing [1]. While processing a query, the reasoning engine produces a proof tree for any given answer. This proof tree acts then as input for a second inference run which returns answers that are explaining the proof tree in natural language. In $\mathcal{JALC}$ we can employ the justification terms as input to this second inference run.

Data privacy. Although knowledge base systems enter more and more application domains, privacy issues in the context of description logics are not yet well studied. Notable exceptions are the following: Calvanese et al. [12] address the problem of privacy aware access to ontologies. They show how view based query answering is able to conceal from the user information that are not logical consequences of the associated authorization views. Grau and Horrocks [13] study different notions of privacy for logic based information systems. They look at privacy preserving query answering as reasoning problem and establish a connection between such reasoning problems and probabilistic privacy guarantees such as perfect privacy. Safe reasoning strategies for very expressive description

logics and for hierarchical ontologies are studied in [9, 24]. The approach used there is based on the principles of locality and conservative extensions for description logics. A decision procedure for provable data privacy in the context of $\mathcal{ALC}$ as well as a sufficient condition for $\mathcal{ALC}$ data privacy that can be checked efficiently is presented in [23].

7 Conclusion

We extended the description logic $\mathcal{ALC}$ with justifications. This results in an epistemic logic $\mathcal{JALC}$ over $\mathcal{ALC}$ where not only an agent's knowledge can be expressed but one also has explicit justifications for that knowledge. We presented a deductive system as well as a semantics for $\mathcal{JALC}$ and proved soundness and completeness. Moreover, we showed that the justification terms of $\mathcal{JALC}$ reflect its provability notion and allow thus to internalize proofs of $\mathcal{JALC}$. We finally explored two applications of this property: generating explanations and data privacy.

It is worth noticing that our approach for adding justifications is very general and does not rely on the particular choice of $\mathcal{ALC}$. This is due to the fact that there is no deep interaction between the justification logic part and the description logic part of $\mathcal{JALC}$ (this is very similar to Finger and Gabbay's [15] way of temporalizing logics). Basically, we only needed the compactness property of $\mathcal{ALC}$ to establish completeness. Thus we could add justifications also to other description logics in the same way as presented here for $\mathcal{ALC}$ and our applications would still be possible.

Further work starts with investigating more deeply the basic properties of $\mathcal{JALC}$. We need decision procedures for it and their complexities have to be determined. On a more practical level, we have to elaborate on the applications of $\mathcal{JALC}$. In particular, we would like to fully develop the *justification terms as explanations* approach and we think it is also worthwhile to further investigate justification terms in the context of data privacy.

There is also a second direction of future work: namely, to combine justifications and terminological reasoning by integrating justification terms in concept descriptions. The definition of a concept then includes a clause of the form

$$\text{if } t \text{ is a term and } C \text{ is a concept, then } [t]C \text{ is a concept.}$$

On the semantic side, the concept $[t]C$ includes all individuals a for which t justifies that a belongs to C. A similar approach was explored for pure modal logic where concepts of the form $\Box C$ were included into the language of description logic, see for instance [8]. Concepts of this modalized form turned out to be important for several applications including procedural extension of description logics [14]. We believe that considering justifications in concept descriptions also is very promising.

References

1. J. Angele, E. Moench, H. Oppermann, S. Staab, and D. Wenke. Ontology-based query and answering in chemistry: OntoNova @ project halo. In *Proc. of the 2nd Int. Semantic Web Conference ISWC*, 1993.
2. S. N. Artemov. Operational modal logic. Technical Report MSI 95–29, Cornell University, Dec. 1995.
3. S. N. Artemov. Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, 7(1):1–36, Mar. 2001.
4. S. N. Artemov. The logic of justification. *The Review of Symbolic Logic*, 1(4):477–513, Dec. 2008.
5. S. N. Artemov. Tracking evidence. In A. Blass, N. Dershowitz, and W. Reisig, editors, *Fields of Logic and Computation, Essays Dedicated to Yuri Gurevich on the Occasion of His 70th Birthday*, volume 6300 of *Lecture Notes in Computer Science*, pages 61–74. Springer, 2010.
6. S. N. Artemov and R. Kuznets. Logical omniscience as a computational complexity problem. In A. Heifetz, editor, *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Twelfth Conference (TARK 2009)*, pages 14–23, Stanford University, California, July 6–8, 2009. ACM.
7. F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, 2nd edition, 2007.
8. F. Baader and A. Laux. Terminological logics with modal operators. In *Proc. of IJCAI '95*, pages 808–814. Morgan Kaufmann, 1995.
9. J. Bao, G. Slutzki, and V. Honavar. Privacy-preserving reasoning on the semantic web. In *Web Intelligence 2007*, pages 791–797, 2007.
10. S. Bucheli, R. Kuznets, B. Renne, J. Sack, and T. Studer. Justified belief change. In X. Arrazola and M. Ponte, editors, *LogKCA-10, Proceedings of the Second ILCLI International Workshop on Logic and Philosophy of Knowledge, Communication and Action*, pages 135–155. University of the Basque Country Press, 2010.
11. S. Bucheli, R. Kuznets, and T. Studer. Partial realization in dynamic justification logic. In L. Beklemishev and R. de Queiroz, editors, *Proc. of Wollic 2011*, volume 6642 of *LNCS*, pages 35–51. Springer, 2011.
12. D. Calvanese, G. De Giacomo, M. Lenzerini, and R. Rosati. View-based query answering over description logic ontologies. In *Principles of Knowledge Representation and Reasoning*, pages 242–251, 2008.
13. B. Cuenca Grau and I. Horrocks. Privacy-preserving query answering in logic-based information systems. In *ECAI 2008*, 2008.
14. F. Donini, M. Lenzerini, D. Nardi, W. Nutt, and A. Schaerf. An epistemic operator for escription logics. *Artificial Intelligence*, 100(1–2):225–274, 1998.
15. M. Finger and D. Gabbay. Adding a temporal dimension to a logic system. *Journal of Logic Language and Information*, 1:203–233, 1992.
16. M. Fitting. The logic of proofs, semantically. *Annals of Pure and Applied Logic*, 132(1):1–25, Feb. 2005.
17. A. Gräber, H. Bürckert, and A. Laux. Terminological reasoning with knowledge and belief. In *Knowledge and Belief in Philosophy and Artificial Intelligence*, pages 29–61. Akademie Verlag, 1995.
18. R. Kuznets. Self-referential justifications in epistemic logic. *Theory of Computing Systems*, 46(4):636–661, May 2010.

19. A. Laux. Beliefs in multi-agent worlds: a terminological approach. In *Proc. of ECAI '94*, pages 299–303, 1994.
20. C. Lutz, H. Sturm, F. Wolter, and M. Zakharyaschev. A tableau decision algorithm for modalized ALC with constant domains. *Studia Logica*, 72:199–232, 2002.
21. D. McGuiness and P. Pinheiro da Silva. Inference web: Portable and shareable explanations for question answering. In *Proc. of AAAI Workshop on New Directions for Question Answering*, 1993.
22. K. Schild. A correspondence theory for terminological logics: Preliminary report. In *Proc. of IJCAI*, pages 466–471, 1991.
23. P. Stouppa and T. Studer. Data privacy for $\mathcal{ALC}$ knowledge bases. In S. Artemov and A. Nerode, editors, *LFCS 2009*, volume 5407 of *LNCS*, pages 409–421. Springer, 2009.
24. T. Studer. Privacy preserving modules for ontologies. In A. Pnueli, I. Virbitskaite, and A. Voronkov, editors, *PSI'09*, volume 5947 of *LNCS*, pages 380–387. Springer, 2010.