sim^{π}: A Concept Similarity Measure under an Agent's Preferences in Description Logic $\mathcal{ELH}$

Teeradaj Racharak[1,2], Boontawee Suntisrivaraporn[1] and Satoshi Tojo[2]
[1]School of Information, Computer and Communication Technology, Sirindhorn International Institute of Technology, Thammasat University, Pathumthani, Thailand
[2]Graduate School of Information Science, Japan Advanced Institute of Science and Technology, Nomi, Japan

Keywords: Concept Similarity Measures, Non-standard Reasoning Services, Preference Profile, Description Logics.

Abstract: In Description Logics (DLs), concept similarity measures (CSMs) aim at identifying a degree of commonality between two given concepts and are often regarded as a generalization of the classical reasoning problem of equivalence. That is, any two concepts are equivalent if their similarity degree is one, and vice versa. When two concepts are not equivalent, the level of similarity varies depending not only on the objective factors (i.e. the concept descriptions) but also on the subjective factors (i.e. the agent's preferences). This work presents the notion of a general preference profile to be used in existing similarity measures and exemplifies its applicability with the similarity measure for the DL $\mathcal{ELH}$, called sim^{π}. We show that our measure is expressible for all aspects of preference profile and prove that sim^{π} is preference-invariant w.r.t. equivalence, i.e. similarity between two equivalent concepts is always one regardless of agents' preferences.

1 INTRODUCTION

Agents' preferences are used in a variety of related, but not identical, ways in their daily life: to express what they like and dislike, to express their desired goals when choosing routes for travelling (Son et al., 2003), etc. In psychology, preferences may be conceived of as an agent's attitude towards a set of objects when making decisions (Lichtenstein and Slovic, 2006). Alternatively, preferences can be interpreted as a judgment in a sense of liking or disliking an object (Scherer, 2005).

In Description Logics (DLs), concept similarity measures (CSMs) aim at identifying a degree of commonality between two given concept names and are often regarded as a generalization of the classical reasoning problem of equivalence. That is, any two concepts are equivalent if their similarity degree is one, and vice versa. To date, many semantic CSMs have been developed (cf. Section 4). These developments can induce efficient similarity-oriented DL reasoning services, i.e., to measure if two concepts are similar, to check if a given instance is a relaxed instance of a concept, and to retrieve those instances similar to a given instance. However, relatively limited efforts have been placed on addressing real-world similarity services executed by a user agent, i.e., *finding similarity w.r.t. the needs and preferences of an agent*. These issues can be illustrated with the following example:

Example 1.1. Suppose that Bob, a Ph.D student, wants to visit a place for active activities, and he feels like a place where he can enjoy walking. According to his world, a terminology might have been modeled in DL as follows:

$$\begin{aligned}
\mathsf{ActivePlace} &\sqsubseteq \mathsf{Place} \sqcap \exists \mathsf{canWalk.Trekking} \\
&\quad \sqcap \exists \mathsf{canSail.Kayaking} \\
\mathsf{Mangrove} &\sqsubseteq \mathsf{Place} \sqcap \exists \mathsf{canWalk.Trekking} \\
\mathsf{Beach} &\sqsubseteq \mathsf{Place} \sqcap \exists \mathsf{canSail.Kayaking} \\
\mathsf{canWalk} &\sqsubseteq \mathsf{canDo} \\
\mathsf{canSail} &\sqsubseteq \mathsf{canDo}
\end{aligned}$$

Considering merely the objective aspects of the world, it is reasonable to conclude that both Mangrove and Beach are equally similar to the notion of ActivePlace. Taking into account also Bob's preferences, however, Mangrove appears more suitable to his perception of ActivePlace.

□

The example shows that preferences of an agent play a decisive role in the choice of alternatives. Thus, we need to be able to fine-tune the degree of similarity by employing aspects apart from the objective factors (i.e. the concept descriptions themselves). It is worth observing that, with a few exceptions like

Racharak, T., Suntisrivaraporn, B. and Tojo, S.
sim^{π}: A Concept Similarity Measure under an Agent's Preferences in Description Logic $\mathcal{ELH}$.
DOI: 10.5220/0005813404800487
In *Proceedings of the 8th International Conference on Agents and Artificial Intelligence (ICAART 2016) - Volume 2*, pages 480-487
ISBN: 978-989-758-172-4

sim and *simi* (cf. Section 4), most CSMs do not allow user agents to specify their preferences and use them to identify a degree of similarity between two concepts. The responsibility of finding similar concepts w.r.t. the needs and preferences of an agent rests solely on that agent.

In this work, we exemplify the applicability of the so-called *preference profile* (Racharak et al., 2015), which is a design guideline for the development of concept similarity measures under an agent's preferences, to the similarity measure sim, in symbols sim^{π}. We also exhibit that sim^{π} is expressible for all aspects of preference profile and prove that sim^{π} is preference-invariant w.r.t. equivalence, i.e. similarity between two equivalent concepts is always one regardless of agents' preferences (cf. Section 3).

2 PRELIMINARIES

In Description Logics (DLs), *concept descriptions* are inductively defined by the help of a set of *constructors*, a set of concept names CN, and a set of role names RN. The set of concept descriptions, or simply concepts, for a specific DL $\mathcal{L}$ is denoted by $\mathsf{Con}(\mathcal{L})$. The set $\mathsf{Con}(\mathcal{ELH})$ of all $\mathcal{ELH}$ concepts can be inductively defined by the following grammar,

$$C,\, D \longrightarrow A \mid \top \mid C \sqcap D \mid \exists r.C$$

where $\top$ denotes the *top concept*, $C, D \in \mathsf{Con}(\mathcal{ELH})$, $A \in \mathsf{CN}$ and $r \in \mathsf{RN}$. Conventionally, concept names are denoted by A and B, concept descriptions are denoted by C and D, and role names are denoted by r and s.

A *terminology* or *TBox* O is a finite set of (possibly primitive) concept definitions and role hierarchy axioms, whose syntax is an expression of the form ($A \sqsubseteq D$) $A \equiv D$, and $r \sqsubseteq s$, respectively. A TBox is called *unfoldable* if it contains at most one concept definition for each concept name in CN and does not contain cyclic dependencies. Concept names occurring on the left-hand side of a concept definition are called defined concept names (denoted by $\mathsf{CN}^{\mathsf{def}}$), all other concept names are primitive concept names (denoted by $\mathsf{CN}^{\mathsf{pri}}$). A primitive definition $A \sqsubseteq D$ can easily be transformed into a semantically equivalent full definitions $A \equiv X \sqcap D$ where X is a fresh concept name. When a TBox O is unfoldable, concept names can be expanded by exhaustively replacing all defined concept names by their definitions until only primitive concept names remain. Such concept names are called *fully expanded concept names*. In what follows, we assume that concepts are fully expanded, and as such the TBox can be omitted. Like primitive definitions, a role hierarchy axiom $r \sqsubseteq s$ can be transformed in to a semantically equivalent role definition $r \equiv t \sqcap s$ where t is a fresh role name. Role names occurring on the left-hand side of a role definition are called defined role names, denoted by $\mathsf{RN}^{\mathsf{def}}$. All others are primitive role names, collectively denoted by $\mathsf{RN}^{\mathsf{pri}}$. We also denote a set of all r's super roles by $\mathcal{R}_r = \{s \in \mathsf{RN} \mid r = s$ or $r_i \sqsubseteq r_{i+1} \in O$ where $1 \leq i \leq n, r_1 = r, r_n = s\}$.

In order to define a formal semantics for a specific DL $\mathcal{L}$, we consider an *interpretation* $I = \langle \Delta^I, \cdot^I \rangle$, which consists of a nonempty set Δ^I as the domain of the interpretation and an interpretation function $\cdot^I$ which assigns to every concept name A a set $A^I \subseteq \Delta^I$ and to every role name r a binary relation $r^I \subseteq \Delta^I \times \Delta^I$. The interpretation function $\cdot^I$ is inductively extended to $\mathcal{ELH}$ concepts in the usual manner:

$$\top^I = \Delta; \qquad (C \sqcap D)^I = C^I \cap D^I;$$
$$(\exists r.C)^I = \{a \in \Delta^I \mid \exists b \in \Delta^I : (a,b) \in r^I \wedge b \in C^I\}.$$

An interpretation I is said to be a *model* of a TBox O (in symbols, $I \models O$) if it satisfies all axioms in O. I satisfies axioms $A \sqsubseteq$, $A \equiv C$, and $r \sqsubseteq s$, respectively, if $A^I \subseteq C^I$, $A^I = C^I$, and $r^I \subseteq s^I$. One of the main classical reasoning problems is the *subsumption problem*. That is, given two concept descriptions C and D and a TBox O, C is subsumed by D w.r.t. a TBox O (written as $C \sqsubseteq_O D$) if $C^I \subseteq D^I$ in every model I of O. Furthermore, C and D are equivalent w.r.t. O (written as $C \equiv_O D$) if $C \sqsubseteq_O D$ and $D \sqsubseteq_O C$. When a TBox O is empty or is clear from the context, we omit to denote O, i.e. $C \sqsubseteq D$ and $C \equiv D$.

Concept Similarity Measure (CSM). is one of non-standard DL reasoning services. It determines how similar two concepts are. Formally, given two concept descriptions $C, D \in \mathsf{Con}(\mathcal{L})$ for a specific DL $\mathcal{L}$. Then, a *concept similarity measure* w.r.t. a TBox O is a function $\sim_O : \mathsf{Con}(\mathcal{L}) \times \mathsf{Con}(\mathcal{L}) \to [0,1]$ such that $C \sim_O D = 1$ iff $C \equiv_O D$ (total similarity) and $C \sim_O D = 0$ indicates total dissimilarity between C and D. When a TBox O is clear from the context, we simply write $C \sim D$.

Since we present an extension to sim (Suntisrivaraporn, 2013; Tongphu and Suntisrivaraporn, 2015) for taking into account an agent's preferences, the original definitions of homomorphism degree and sim are included here for self-containment. Let $C \in \mathsf{Con}(\mathcal{ELH})$ be a fully expanded concept to the form:

$$P_1 \sqcap \cdots \sqcap P_m \sqcap \exists r_1.C_1 \sqcap \cdots \sqcap \exists r_n.C_n$$

where $P_i \in \mathsf{CN}^{\mathsf{pri}}$, $r_j \in \mathsf{RN}$, $C_j \in \mathsf{Con}(\mathcal{ELH})$ in the same format, $1 \leq i \leq m$, and $1 \leq j \leq n$. The set $P_1, \ldots, P_m$ and the set $\exists r_1.C_1, \ldots, \exists r_n.C_n$ are denoted by $\mathcal{P}_C$ and $\mathcal{E}_C$, respectively. An $\mathcal{ELH}$ concept description can be structurally transformed into the corresponding $\mathcal{ELH}$ description tree. The root v_0 of the $\mathcal{ELH}$ description tree $\mathcal{T}_C$ has $\{P_1, \ldots, P_m\}$ as its label and has n outgoing edges, each labeled with r_j to a vertex v_j for $1 \leq j \leq n$. Then, a subtree with the root v_j is defined recursively relative to the concept C_j.

Definition 2.1 (Homomorphism Degree (Tongphu and Suntisrivaraporn, 2015))**.** Let $\mathbf{T}^{\mathcal{ELH}}$ be a set of all $\mathcal{ELH}$ description trees and $\mathcal{T}_C, \mathcal{T}_D \in \mathbf{T}^{\mathcal{ELH}}$ corresponds to two $\mathcal{ELH}$ concept names C and D, respectively. The *homomorphism degree* function $\mathsf{hd} : \mathbf{T}^{\mathcal{ELH}} \times \mathbf{T}^{\mathcal{ELH}} \to [0,1]$ is inductively defined as follows:

$$\begin{aligned}\mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C) = {} & \mu \cdot \mathsf{p\text{-}hd}(\mathcal{P}_D, \mathcal{P}_C) \\ & + (1-\mu) \cdot \mathsf{e\text{-}set\text{-}hd}(\mathcal{E}_D, \mathcal{E}_C), \end{aligned} \quad (1)$$

where $|\cdot|$ represents the set cardinality, $\mu = \frac{|\mathcal{P}_D|}{|\mathcal{P}_D \cup \mathcal{E}_D|}$ and $0 \leq \mu \leq 1$;

$$\mathsf{p\text{-}hd}(\mathcal{P}_D, \mathcal{P}_C) = \begin{cases} 1 & \text{if } \mathcal{P}_D = \emptyset \\ \frac{|\mathcal{P}_D \cap \mathcal{P}_C|}{|\mathcal{P}_D|} & \text{otherwise,} \end{cases} \quad (2)$$

$$\mathsf{e\text{-}set\text{-}hd}(\mathcal{E}_D, \mathcal{E}_C) = \begin{cases} 1 & \text{if } \mathcal{E}_D = \emptyset \\ 0 & \text{if } \mathcal{E}_D \neq \emptyset \text{ and } \mathcal{E}_C = \emptyset \\ \mathsf{e}^*(\mathcal{E}_D, \mathcal{E}_C) & \text{otherwise,} \end{cases} \quad (3)$$

where

$$\mathsf{e}^*(\mathcal{E}_D, \mathcal{E}_C) = \sum_{\varepsilon_i \in \mathcal{E}_D} \frac{\max\{\mathsf{e\text{-}hd}(\varepsilon_i, \varepsilon_j) : \varepsilon_j \in \mathcal{E}_C\}}{|\mathcal{E}_D|} \quad (4)$$

with $\varepsilon_i, \varepsilon_j$ existential restrictions; and

$$\mathsf{e\text{-}hd}(\exists r.X, \exists s.Y) = \gamma(\nu + (1-\nu) \cdot \mathsf{hd}(\mathcal{T}_X, \mathcal{T}_Y)) \quad (5)$$

where $\gamma = \frac{|\mathcal{R}_r \cap \mathcal{R}_s|}{|\mathcal{R}_r|}$ and $0 \leq \nu < 1$.

Definition 2.2 ($\mathcal{ELH}$ Similarity Degree (Tongphu and Suntisrivaraporn, 2015))**.** Let C and D be $\mathcal{ELH}$ concept names and $\mathcal{T}_C, \mathcal{T}_D$ be the corresponding description trees. Then, the $\mathcal{ELH}$ *similarity degree* between C and D (in symbols, $\mathsf{sim}(C,D)$) is defined as follows:

$$\mathsf{sim}(C,D) = \frac{\mathsf{hd}(\mathcal{T}_C, \mathcal{T}_D) + \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)}{2} \quad (6)$$

Example 2.1 (Continuation of Example 1.1)**.** Each primitive definition can be transformed to a corresponding equivalent full definition as shown in the following.

$$\begin{aligned} \mathsf{ActivePlace} & \equiv X \sqcap \mathsf{Place} \\ & \quad \sqcap \exists \mathsf{canWalk}.\mathsf{Trekking} \\ & \quad \sqcap \exists \mathsf{canSail}.\mathsf{Kayaking} \\ \mathsf{Mangrove} & \equiv Y \sqcap \mathsf{Place} \sqcap \\ & \quad \sqcap \exists \mathsf{canWalk}.\mathsf{Trekking} \\ \mathsf{Beach} & \equiv Z \sqcap \mathsf{Place} \\ & \quad \sqcap \exists \mathsf{canSail}.\mathsf{Kayaking} \end{aligned}$$

where X, Y and Z are fresh primitive concept names. Furthermore, $\mathcal{R}_{\mathsf{canWalk}} = \{t, \mathsf{canDo}\}$ and $\mathcal{R}_{\mathsf{canSail}} = \{u, \mathsf{canDo}\}$ where t and u are fresh primitive role names. For brevity, let ActivePlace, Mangrove, Place, Trekking, Kayaking, canWalk, and canSail be abbreviated as AP, M, P, T, K, cW, and cS, respectively. Using Definition 2.1, the homomorphism degree from ActivePlace to Mangrove, or $\mathsf{hd}(\mathcal{T}_{\mathsf{AP}}, \mathcal{T}_{\mathsf{M}})$

$$\begin{aligned} & = (\tfrac{2}{4})(\tfrac{1}{2}) + (\tfrac{2}{4}) \\ & (\tfrac{\max\{\mathsf{e\text{-}hd}(\exists \mathsf{cW.T}, \exists \mathsf{cW.T})\}}{2} + \tfrac{\max\{\mathsf{e\text{-}hd}(\exists \mathsf{cS.K}, \exists \mathsf{cW.T})\}}{2}) \\ & = (\tfrac{2}{4})(\tfrac{1}{2}) + (\tfrac{2}{4})(\tfrac{0.5+0.1}{2}) = 0.55 \end{aligned}$$

Similarly, $\mathsf{hd}(\mathcal{T}_{\mathsf{M}}, \mathcal{T}_{\mathsf{AP}}) = 0.67$, $\mathsf{hd}(\mathcal{T}_{\mathsf{AP}}, \mathcal{T}_{\mathsf{B}}) = 0.55$, and $\mathsf{hd}(\mathcal{T}_{\mathsf{M}}, \mathcal{T}_{\mathsf{AP}}) = 0.67$. Thus, $\mathsf{sim}(\mathsf{M}, \mathsf{AP}) = 0.61$ and $\mathsf{sim}(\mathsf{B}, \mathsf{AP}) = 0.61$ □

2.1 Preference Profile

Preference profile is first proposed in (Racharak et al., 2015) as a guideline for developing *CSMs under preferences*. It is a quintuple of preference functions which exhibit five aspects for preference expressions. It can be adopted into the development of arbitrary CSMs and thereby influencing the calculation of CSMs. The syntax and semantics of each aspect are given in term of partial functions since different agents can have different perspectives of preferences. Any CSMs that expose those syntactic forms and satisfy their corresponding semantics will infer a similarity value w.r.t. the needs and preferences of an agent. Each syntax and semantic is presented formally as follows:

Definition 2.3 (Primitive Concept Importance)**.** Let $\mathsf{CN}^{\mathsf{pri}}(O)$ be a set of primitive concept names occurring in O. Then, a *primitive concept importance* is a partial function $\mathfrak{i}^{\mathfrak{c}} : \mathsf{CN} \to \mathbb{R}_{\geq 0}$, where $\mathsf{CN} \subseteq \mathsf{CN}^{\mathsf{pri}}(O)$.

For any $A \in \mathsf{CN}^{\mathsf{pri}}(O)$, $\mathfrak{i}^{\mathfrak{c}}(A) = 1$ captures an expression of normal importance for A, $\mathfrak{i}^{\mathfrak{c}}(A) > 1$ (and $\mathfrak{i}^{\mathfrak{c}}(A) < 1$) indicates that A has higher (and lower, respectively) importance, and $\mathfrak{i}^{\mathfrak{c}}(A) = 0$ indicates that A is entirely ignored by an agent. For example, suppose Bob is keenly interested to visit places. Therefore, he can express as $\mathfrak{i}^{\mathfrak{c}}(\mathsf{Place}) = 2$ for his preference profile.

Definition 2.4 (Role Importance). Let $\mathsf{RN}(\mathcal{O})$ be a set of role names occurring in $\mathcal{O}$. Then, a *role importance* is a partial function $\mathfrak{i}^{\mathfrak{r}} : \mathsf{RN} \to \mathbb{R}_{\geq 0}$, where $\mathsf{RN} \subseteq \mathsf{RN}(\mathcal{O})$.

For any $r \in \mathsf{RN}(\mathcal{O})$, $\mathfrak{i}^{\mathfrak{r}}(r) = 1$ captures an expression of normal importance for r, $\mathfrak{i}^{\mathfrak{r}}(r) > 1$ (and $\mathfrak{i}^{\mathfrak{r}}(r) < 1$) indicates that r has higher (and lower, respectively) importance, and $\mathfrak{i}^{\mathfrak{r}}(r) = 0$ indicates that r is entirely ignored by an agent. For example, Bob is interested to visit places where he can enjoy walking. Therefore, he can also express as $\mathfrak{i}^{\mathfrak{r}}(\mathsf{canWalk}) = 2$ for his preference profile.

Definition 2.5 (Primitive Concepts Similarity). Let $\mathsf{CN}^{\mathsf{pri}}(\mathcal{O})$ be a set of primitive concept names occurring in $\mathcal{O}$. For $A, B \in \mathsf{CN}^{\mathsf{pri}}(\mathcal{O})$, a *primitive concepts similarity* is a partial function $\mathfrak{s}^{\mathfrak{c}} : \mathsf{CN} \times \mathsf{CN} \to [0,1]$, where $\mathsf{CN} \subseteq \mathsf{CN}^{\mathsf{pri}}(\mathcal{O})$, such that $\mathfrak{s}^{\mathfrak{c}}(A,B) = \mathfrak{s}^{\mathfrak{c}}(B,A)$ and $\mathfrak{s}^{\mathfrak{c}}(A,A) = 1$.

For $A, B \in \mathsf{CN}^{\mathsf{pri}}(\mathcal{O})$, $\mathfrak{s}^{\mathfrak{c}}(A,B) = 1$ captures an expression of total similarity between A and B and $\mathfrak{s}^{\mathfrak{c}}(A,B) = 0$ captures an expression of total dissimilarity between A and B. For example, Bob believes that trekking and kayaking are a bit similar in some sense. Hence, he can express $\mathfrak{s}^{\mathfrak{c}}(\mathsf{Trekking}, \mathsf{Kayaking}) = 0.1$ for his preference profile.

Definition 2.6 (Primitive Roles Similarity). Let $\mathsf{RN}^{\mathsf{pri}}(\mathcal{O})$ be a set of primitive role names occurring in $\mathcal{O}$. For $r, s \in \mathsf{RN}^{\mathsf{pri}}(\mathcal{O})$, a *primitive roles similarity* is a partial function $\mathfrak{s}^{\mathfrak{r}} : \mathsf{RN} \times \mathsf{RN} \to [0,1]$, where $\mathsf{RN} \subseteq \mathsf{RN}^{\mathsf{pri}}(\mathcal{O})$, such that $\mathfrak{s}^{\mathfrak{r}}(r,s) = \mathfrak{s}^{\mathfrak{r}}(s,r)$ and $\mathfrak{s}^{\mathfrak{r}}(r,r) = 1$.

For $r, s \in \mathsf{RN}(\mathcal{O})$, $\mathfrak{s}^{\mathfrak{r}}(r,s) = 1$ captures an expression of total similarity between r and s and $\mathfrak{s}^{\mathfrak{r}}(r,s) = 0$ captures an expression of total dissimilarity between r and s. For example, Bob believes that walking is a bit similar to sailing. Hence, he can also express $\mathfrak{s}^{\mathfrak{r}}(t,u) = 0.1$ for his preference profile.

Definition 2.7 (Role Discount Factor). Let $\mathsf{RN}(\mathcal{O})$ be a set of role names occurring in $\mathcal{O}$. Then, a *role discount factor* is a partial function $\mathfrak{d} : \mathsf{RN} \to [0,1]$, where $\mathsf{RN} \subseteq \mathsf{RN}(\mathcal{O})$.

For any $r \in \mathsf{RN}(\mathcal{O})$, $\mathfrak{d}(r) = 1$ captures an expression of total importance on a role (over a corresponding nested concept) and $\mathfrak{d}(r) = 0$ captures an expression of total importance on a nested concept (over a corresponding role). For example, Bob does not concern much if places permit to either walk or to sail. He would rather consider on actual activities which he can perform. Thus, he may express $\mathfrak{d}(\mathsf{canWalk}) = 0.3$ and $\mathfrak{d}(\mathsf{canSail}) = 0.3$ for his preference profile.

Definition 2.8 (Preference Profile). (Racharak et al., 2015) A *preference profile*, in symbol π, is a quintuple $\langle \mathfrak{i}^{\mathfrak{c}}, \mathfrak{i}^{\mathfrak{r}}, \mathfrak{s}^{\mathfrak{c}}, \mathfrak{s}^{\mathfrak{r}}, \mathfrak{d} \rangle$ where $\mathfrak{i}^{\mathfrak{c}}, \mathfrak{i}^{\mathfrak{r}}, \mathfrak{s}^{\mathfrak{c}}, \mathfrak{s}^{\mathfrak{r}}$, and $\mathfrak{d}$ are as defined above and the *default preference profile*, in symbol π_0, is the quintuple $\langle \mathfrak{i}^{\mathfrak{c}}_0, \mathfrak{i}^{\mathfrak{r}}_0, \mathfrak{s}^{\mathfrak{c}}_0, \mathfrak{s}^{\mathfrak{r}}_0, \mathfrak{d}_0 \rangle$ where

$$
\begin{aligned}
\mathfrak{i}^{\mathfrak{c}}_0(A) &= 1 \text{ for all } A \in \mathsf{CN}^{\mathsf{pri}}(\mathcal{O}),\\
\mathfrak{i}^{\mathfrak{r}}_0(r) &= 1 \text{ for all } r \in \mathsf{RN}(\mathcal{O}),\\
\mathfrak{s}^{\mathfrak{c}}_0(A,B) &= 0 \text{ for all } (A,B) \in \mathsf{CN}^{\mathsf{pri}}(\mathcal{O}) \times \mathsf{CN}^{\mathsf{pri}}(\mathcal{O}),\\
\mathfrak{s}^{\mathfrak{r}}_0(r,s) &= 0 \text{ for all } (r,s) \in \mathsf{RN}(\mathcal{O}) \times \mathsf{RN}(\mathcal{O}), \text{ and}\\
\mathfrak{d}_0(r) &= 0.4 \text{ for all } r \in \mathsf{RN}(\mathcal{O}).
\end{aligned}
$$

3 CSM UNDER AGENT'S PREFERENCES

A numerical value obtained by CSMs indicates the similarity between two concept descriptions. For instance, sim(ActivePlace, Mangrove) = 0.61 and sim(ActivePlace, Beach) = 0.61 indicates that the similarity between ActivePlace and Mangrove, and that between ActivePlace and Beach are equivalently 61%. This means both Mangrove and Beach match equally to the general notion of ActivePlace. Unfortunately, this is not true because it does not correspond with his needs and his preferences. Indeed, the similarity degree between ActivePlace and Mangrove should be greater than the similarity degree between ActivePlace and Beach in order to consistent with Bob's perspectives (as exhibited by Figure 1).

In this section, we adopt those aspects of preference profile into our development of concept similarity measure under an agent's preferences for DL $\mathcal{ELH}$. In the following, we have presented formal definition of *concept similarity measures under preference profile*.

Definition 3.1. Given a CSM $\sim$, a preference profile π, and two concepts $C, D \in \mathsf{Con}(\mathcal{L})$. Then, a *concept similarity measure under preference profile* π is a function $\overset{\pi}{\sim} : \mathsf{Con}(\mathcal{L}) \times \mathsf{Con}(\mathcal{L}) \to [0,1]$. A CSM $\sim$ is called *preference invariant w.r.t. equivalence* if

$$C \equiv D \text{ iff } C \overset{\pi}{\sim} D = 1 \text{ for any } \pi$$

By developing a concept similarity measure under preference profile for $\mathcal{ELH}$, we have generalized the measure sim. To avoid confusion, we write $\overset{\pi}{\sim}$ when referring to an arbitrary CSM in a generic sense, whereas specific function symbols, e.g. sim^{π} or hd^{π}, are used when talking about specific CSMs or functions.

In order to consider those aspects of preference profile, we have presented a *total importance function* as $\hat{\mathfrak{i}} : \mathsf{CN}^{\mathsf{pri}} \cup \mathsf{RN} \to \mathbb{R}_{\geq 0}$ based on a concept im-

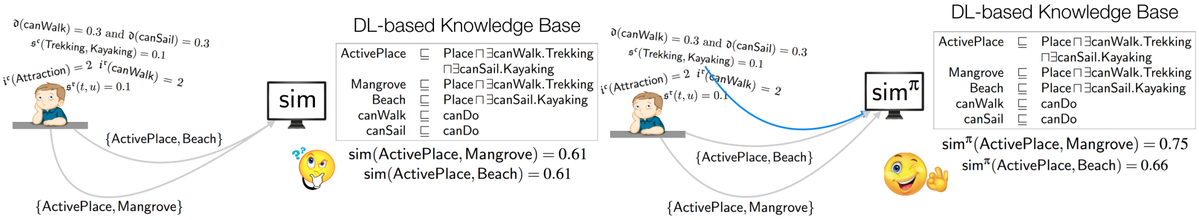

Figure 1: Similarity value with (and without) respect to Bob's preferences.

portance and a role importance.

$$\hat{\mathfrak{i}}(x) = \begin{cases} \mathfrak{i}^c(x) & \text{if } x \in \mathsf{CN}^{\mathsf{pri}} \text{and } \mathfrak{i}^c \text{ is defined on } x \\ \mathfrak{i}^r(x) & \text{if } x \in \mathsf{RN} \text{ and } \mathfrak{i}^r \text{ is defined on } x \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

A *total similarity function* is also presented as $\hat{\mathfrak{s}} : (\mathsf{CN}^{\mathsf{pri}} \times \mathsf{CN}^{\mathsf{pri}}) \cup (\mathsf{RN}^{\mathsf{pri}} \times \mathsf{RN}^{\mathsf{pri}}) \to [0,1]$ using a primitive concept similarity and a primitive role similarity.

$$\hat{\mathfrak{s}}(x,y) = \begin{cases} 1 & \text{if } x = y \\ \mathfrak{s}^c(x,y) & \text{if } (x,y) \in \mathsf{CN}^{\mathsf{pri}} \times \mathsf{CN}^{\mathsf{pri}} \\ & \text{and } \mathfrak{s}^c \text{ is defined on } (x,y) \\ \mathfrak{s}^r(x,y) & \text{if } (x,y) \in \mathsf{RN}^{\mathsf{pri}} \times \mathsf{RN}^{\mathsf{pri}} \\ & \text{and } \mathfrak{s}^r \text{ is defined on } (x,y) \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Similarly, a *total role discount factor function* is presented in the following in term of a function $\hat{\mathfrak{d}} : \mathsf{RN} \to [0,1]$ based on a role discount factor.

$$\hat{\mathfrak{d}}(x) = \begin{cases} \mathfrak{d}(x) & \text{if } \mathfrak{d} \text{ is defined on } x \\ 0.4 & \text{otherwise} \end{cases} \tag{9}$$

Let C and D be $\mathcal{ELH}$ concept names and r and s be role names. Let $\mathcal{T}_C, \mathcal{T}_D, \mathcal{P}_C, \mathcal{P}_D, \mathcal{E}_C, \mathcal{E}_D, \mathcal{R}_r$, and $\mathcal{R}_s$ are as defined in Definition 2.1. Let $\mathbf{T}^{\mathcal{ELH}}$ be a set of all $\mathcal{ELH}$ description trees and $\pi = \langle \mathfrak{i}^c, \mathfrak{i}^r, \mathfrak{s}^c, \mathfrak{s}^r, \mathfrak{d} \rangle$ be a preference profile. The homomorphism degree under preference profile π can be formally defined as follows:

Definition 3.2. The *homomorphism degree under preference profile* π is a function $\mathsf{hd}^\pi : \mathbf{T}^{\mathcal{ELH}} \times \mathbf{T}^{\mathcal{ELH}} \to [0,1]$ defined inductively as follows:

$$\begin{aligned} \mathsf{hd}^\pi(\mathcal{T}_D, \mathcal{T}_C) &= \mu^\pi \cdot \mathsf{p\text{-}hd}^\pi(\mathcal{P}_D, \mathcal{P}_C) \\ &\quad + (1 - \mu^\pi) \cdot \mathsf{e\text{-}set\text{-}hd}^\pi(\mathcal{E}_D, \mathcal{E}_C), \end{aligned} \tag{10}$$

$$\text{where } \mu^\pi = \frac{\sum\limits_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A)}{\sum\limits_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A) + \sum\limits_{\exists r.X \in \mathcal{E}_D} \hat{\mathfrak{i}}(r)}; \tag{11}$$

$$\mathsf{p\text{-}hd}^\pi(\mathcal{P}_D, \mathcal{P}_C) = \begin{cases} 1 & \text{if } \sum\limits_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A) = 0 \\ 0 & \text{if } \sum\limits_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A) \neq 0 \text{ and } \sum\limits_{B \in \mathcal{P}_C} \hat{\mathfrak{i}}(B) = 0 \\ \mathsf{p}^{\pi*}(\mathcal{P}_D, \mathcal{P}_C) & \text{otherwise,} \end{cases} \tag{12}$$

where

$$\mathsf{p}^{\pi*}(\mathcal{P}_D, \mathcal{P}_C) = \frac{\sum\limits_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A) \cdot \max\{\hat{\mathfrak{s}}(A,B) : B \in \mathcal{P}_C\}}{\sum\limits_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A)}; \tag{13}$$

$$\mathsf{e\text{-}set\text{-}hd}^\pi(\mathcal{E}_D, \mathcal{E}_C) = \begin{cases} 1 & \text{if } \sum\limits_{\exists r.X \in \mathcal{E}_D} \hat{\mathfrak{i}}(r) = 0 \\ 0 & \text{if } \sum\limits_{\exists r.X \in \mathcal{E}_D} \hat{\mathfrak{i}}(r) \neq 0 \\ & \text{and } \sum\limits_{\exists s.Y \in \mathcal{E}_C} \hat{\mathfrak{i}}(s) = 0 \\ \mathsf{e}^{\pi*}(\mathcal{E}_D, \mathcal{E}_C) & \text{otherwise,} \end{cases} \tag{14}$$

where

$$\mathsf{e}^{\pi*}(\mathcal{E}_D, \mathcal{E}_C) = \frac{\sum\limits_{\exists r.X \in \mathcal{E}_D} \hat{\mathfrak{i}}(r) \cdot \max\{\mathsf{e\text{-}hd}^\pi(\exists r.X, \varepsilon_j) : \varepsilon_j \in \mathcal{E}_C\}}{\sum\limits_{\exists r.X \in \mathcal{E}_D} \hat{\mathfrak{i}}(r)} \tag{15}$$

with ε_j existential restriction; and

$$\mathsf{e\text{-}hd}^\pi(\exists r.X, \exists s.Y) = \gamma^\pi(\hat{\mathfrak{d}}(r) + (1 - \hat{\mathfrak{d}}(r)) \cdot \mathsf{hd}^\pi(\mathcal{T}_X, \mathcal{T}_Y)) \tag{16}$$

$$\text{where } \gamma^\pi = \begin{cases} 1 & \text{if } \sum\limits_{r' \in \mathcal{R}_r} \hat{\mathfrak{i}}(r') = 0 \\ \dfrac{\sum\limits_{r' \in \mathcal{R}_r} \hat{\mathfrak{i}}(r') \cdot \max\{\hat{\mathfrak{s}}(r', s') : s' \in \mathcal{R}_s\}}{\sum\limits_{r' \in \mathcal{R}_r} \hat{\mathfrak{i}}(r')}, & \text{otherwise.} \end{cases} \tag{17}$$

It is obvious to see that Definition 3.2 exposes all elements of preference profile, viz. $\mathfrak{i}^c, \mathfrak{i}^r, \mathfrak{s}^c, \mathfrak{s}^r$, and $\mathfrak{d}$ since it was generalized alongside the use of the functions $\hat{\mathfrak{i}}$, $\hat{\mathfrak{s}}$, and $\hat{\mathfrak{d}}$.

Intuitively, Equation 10 is defined as the weighted sum of the degree under π of primitive concepts and the degree under π of matching edges. Equation 11 indicates the weight of primitive concept names w.r.t. the importance function. Equation 12 calculates the proportion of best similarity between primitive concept names. Similarly, Equation 14 calculates the proportion of best similarity between existential information from Equation 16 and Equation 17. Equation 16 calculates the degree of similarity between matching edges. Finally, Equation 17 calculates the proportion of best similarity between role names.

Let C and D be $\mathcal{ELH}$ concept names, $\mathcal{T}_C$ and $\mathcal{T}_D$ be the corresponding description trees, and $\pi = \langle \mathfrak{i}^c, \mathfrak{i}^r, \mathfrak{s}^c, \mathfrak{s}^r, \mathfrak{d} \rangle$ be a preference profile. The following definition formally describes the *$\mathcal{ELH}$ similarity degree under preference profile* π.

Definition 3.3. The *$\mathcal{ELH}$ similarity degree under preference profile* π between C and D (denoted by $\mathsf{sim}^\pi(C,D)$) is defined as follows:

$$\mathsf{sim}^\pi(C,D) = \frac{\mathsf{hd}^\pi(\mathcal{T}_C, \mathcal{T}_D) + \mathsf{hd}^\pi(\mathcal{T}_D, \mathcal{T}_C)}{2} \tag{18}$$

Lemma 3.1. For $\mathcal{T}_D, \mathcal{T}_C \in \mathbf{T}^{\mathcal{ELH}}$, $\mathsf{hd}^{\pi_0}(\mathcal{T}_D, \mathcal{T}_C) = \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)$.

Proof

Recall by Definition 2.8 that the default preference profile π_0 is the quintuple $\langle \mathfrak{i}_0^c, \mathfrak{i}_0^r, \mathfrak{s}_0^c, \mathfrak{s}_0^r, \mathfrak{d}_0 \rangle$. Also, suppose a concept name D is of the form:

$$P_1 \sqcap \cdots \sqcap P_m \sqcap \exists r_1.D_1 \sqcap \cdots \sqcap \exists r_n.D_n$$

where $P_i \in \mathsf{CN}^{\mathsf{pri}}$, $r_j \in \mathsf{CN}^{\mathsf{pri}}$, $D_j \in \mathsf{Con}(\mathcal{ELH})$, $1 \le i \le m$, $1 \le j \le n$, $P_1 \sqcap \cdots \sqcap P_m$ is denoted by $\mathcal{P}_D$, and $\exists r_1.D_1 \sqcap \cdots \sqcap \exists r_n.D_n$ is denoted by $\mathcal{E}_D$. Let d be the depth of $\mathcal{T}_D$. We prove that, for any $d \in \mathbb{N}$, $\mathsf{hd}^{\pi_0}(\mathcal{T}_D, \mathcal{T}_C) = \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)$ with mathematical induction.

When $d = 0$, we know that $D = P_1 \sqcap \cdots \sqcap P_m$. To show that $\mathsf{hd}^{\pi_0}(\mathcal{T}_D, \mathcal{T}_C) = \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)$, we need to show that $\mu^{\pi_0} = \mu$ and $\mathsf{p\text{-}hd}^{\pi_0}(\mathcal{P}_D, \mathcal{P}_C) = \mathsf{p\text{-}hd}(\mathcal{P}_D, \mathcal{P}_C)$. Let us derive as follows:

$$\mu^{\pi_0} = \frac{\sum_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A)}{\sum_{A \in \mathcal{P}_D} \hat{\mathfrak{i}}(A) + \sum_{\exists r.X \in \mathcal{E}_D} \hat{\mathfrak{i}}(r)} = \frac{\sum_{i=1}^{m} 1}{\sum_{i=1}^{m} 1 + 0} = \frac{m}{m+0} = \mu.$$

Furthermore, we only need to show $\sum_{A \in \mathcal{P}_D} \max\{\hat{\mathfrak{s}}(A,B) : B \in \mathcal{P}_C\} = |\mathcal{P}_D \cap \mathcal{P}_C|$ in order to show $\mathsf{p\text{-}hd}^{\pi_0}(\mathcal{P}_D, \mathcal{P}_C) = \mathsf{p\text{-}hd}(\mathcal{P}_D, \mathcal{P}_C)$. We know that $\mathfrak{s}_0^c$ maps name identity to 1 and otherwise to 0. Thus, $\sum_{A \in \mathcal{P}_D} \max\{\hat{\mathfrak{s}}(A,B) : B \in \mathcal{P}_C\} = |\{x : x \in \mathcal{P}_D \text{ and } x \in \mathcal{P}_C\}| = |\mathcal{P}_D \cap \mathcal{P}_C|$.

We must now prove that if $\mathsf{hd}^{\pi_0}(\mathcal{T}_D, \mathcal{T}_C) = \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)$ holds for $d = h-1$ where $h > 1$ and $D = P_1 \sqcap \cdots \sqcap P_m \sqcap \exists r_1.D_1 \sqcap \cdots \sqcap \exists r_n.D_n$ then $\mathsf{hd}^{\pi_0}(\mathcal{T}_D, \mathcal{T}_C) = \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)$ also holds for $d = h$. To do that, we have to show $\mathsf{e\text{-}set\text{-}hd}^{\pi_0}(\mathcal{E}_D, \mathcal{E}_C) = \mathsf{e\text{-}set\text{-}hd}(\mathcal{E}_D, \mathcal{E}_C)$. This can be done by showing in the similar manner that $\gamma^{\pi_0} = \gamma$ and $\mathsf{hd}^{\pi_0}(\mathcal{T}_X, \mathcal{T}_Y) = \mathsf{hd}(\mathcal{T}_X, \mathcal{T}_Y)$ from $\mathsf{e\text{-}hd}^{\pi_0}(\exists r.X, \exists s.Y) = \mathsf{e\text{-}hd}(\exists r.X, \exists s.Y)$, where $\exists r.X \in \mathcal{E}_D$ and $\exists s.Y \in \mathcal{E}_C$. Consequently, it follows by induction that, for $\mathcal{T}_D, \mathcal{T}_C \in \mathbf{T}^{\mathcal{ELH}}$, $\mathsf{hd}^{\pi_0}(\mathcal{T}_D, \mathcal{T}_C) = \mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C)$. □

Theorem 3.1. For $C, D \in \mathsf{Con}(\mathcal{ELH})$, $\mathsf{sim}^{\pi_0}(C,D) = \mathsf{sim}(C,D)$.

The above theorem follows from Lemma 3.1, Definition 2.2, and Definition 3.3.

Lemma 3.2. For $\mathcal{T}_D, \mathcal{T}_C \in \mathbf{T}^{\mathcal{ELH}}$, $\mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C) = 1$ iff $\mathsf{hd}^\pi(\mathcal{T}_D, \mathcal{T}_C) = 1$ for any π.

Proof

(Sketch) Let $\pi = \langle \mathfrak{i}^c, \mathfrak{i}^r, \mathfrak{s}^c, \mathfrak{s}^r, \mathfrak{d} \rangle$ be an arbitrary preference profile and $\pi_0 = \langle \mathfrak{i}_0^c, \mathfrak{i}_0^r, \mathfrak{s}_0^c, \mathfrak{s}_0^r, \mathfrak{d}_0 \rangle$ be the default preference profile. This lemma can be shown by mathematical induction on the depth of $\mathcal{T}_D$.

($\Rightarrow$) $\mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C) = 1$ implies that there exists a homomorphism mapping from the root of $\mathcal{T}_D$ to the root of $\mathcal{T}_C$. Consequently, any setting on π does not influence the calculation on $\mathsf{hd}^\pi(\mathcal{T}_D, \mathcal{T}_C)$.

($\Leftarrow$) $\mathsf{hd}^\pi(\mathcal{T}_D, \mathcal{T}_C) = 1$ for any π. In particular, this holds for the default preference profile π_0. By Lemma 3.1, it is the case that $\mathsf{hd}(\mathcal{T}_D, \mathcal{T}_C) = 1$. □

Theorem 3.2. sim^π is *preference invariant w.r.t. equivalence*.

Proof

(Sketch) Given two concepts C and D and an arbitrary preference profile π, we have to show $C \equiv D$ iff $\mathsf{sim}^\pi(C,D) = 1$.

($\Rightarrow$) By Proposition 7 in (Tongphu and Suntisrivaraporn, 2015), we can derive that $\mathsf{sim}(C,D) = 1$. With the usage of Lemma 3.2, Definition 2.2, and Definition 3.3, we can derive that sim^π is *preference invariant w.r.t. equivalence*.

($\Leftarrow$) This can be shown similarly as in the forward direction. □

Example 3.1. (Continuation of Example 1.1) Let enrich the example by assuming Bob's preference profile is expressed as follows: (i) $\mathfrak{i}^c(\mathsf{Place}) = 2$; (ii) $\mathfrak{i}^r(\mathsf{canWalk}) = 2$; (iii) $\mathfrak{s}^c(\mathsf{Trekking}, \mathsf{Kayaking}) = 0.1$; (iv) $\mathfrak{s}^r(t,u) = 0.1$; (v) $\mathfrak{d}(\mathsf{canWalk}) = 0.3$ and $\mathfrak{d}(\mathsf{canSail}) = 0.3$. Let $\mathsf{ActivePlace}$, $\mathsf{Mangrove}$, Place, Trekking, Kayaking, canWalk, and canSail are rewritten shortly as AP, M, P, T, K, cW, and cS, respectively. Using Definition 3.2, $\mathsf{hd}^{\pi}(\mathcal{T}_{\mathsf{AP}}, \mathcal{T}_{\mathsf{M}})$

$$
\begin{aligned}
&= (\tfrac{3}{6}) \cdot \mathsf{p\text{-}hd}^{\pi}(\mathcal{P}_{\mathsf{AP}}, \mathcal{P}_{\mathsf{M}}) + (\tfrac{3}{6}) \cdot \mathsf{e\text{-}set\text{-}hd}^{\pi}(\mathcal{E}_{\mathsf{AP}}, \mathcal{E}_{\mathsf{M}}) \\
&= (\tfrac{3}{6}) \cdot (\tfrac{\mathfrak{i}(X)\cdot\max\{\mathfrak{s}(X,Y),\mathfrak{s}(X,\mathsf{P})\}+\mathfrak{i}(\mathsf{P})\cdot\max\{\mathfrak{s}(\mathsf{P},\mathsf{Y}),\mathfrak{s}(\mathsf{P},\mathsf{P})\})}{\mathfrak{i}(X)+\mathfrak{i}(\mathsf{P})}) \\
&+(\tfrac{3}{6}) \cdot \mathsf{e\text{-}set\text{-}hd}^{\pi}(\mathcal{E}_{\mathsf{AP}}, \mathcal{E}_{\mathsf{M}}) \\
&= (\tfrac{3}{6})(\tfrac{1\cdot\max\{0,0\}+2\cdot\max\{0,1\}}{1+2}) + (\tfrac{3}{6}) \cdot \mathsf{e\text{-}set\text{-}hd}^{\pi}(\mathcal{E}_{\mathsf{AP}}, \mathcal{E}_{\mathsf{M}}) \\
&= (\tfrac{3}{6})(\tfrac{2}{3}) \\
&+(\tfrac{3}{6}) \left[\tfrac{\mathfrak{i}(\mathsf{cW})\cdot\max\{\mathsf{e\text{-}hd}^{\pi}(\exists \mathsf{cW.T},\exists \mathsf{cW,T})\}+1\cdot\max\{0.2035\}}{\mathfrak{i}(\mathsf{cW})+\mathfrak{i}(\mathsf{cS})}\right] \\
&= (\tfrac{3}{6})(\tfrac{2}{3}) + (\tfrac{3}{6}) \left[\tfrac{2\cdot\max\{(1)(0.3+0.7(1))\}+1\cdot\max\{0.2035\}}{\mathfrak{i}(\mathsf{cW})+\mathfrak{i}(\mathsf{cS})}\right] \\
&= (\tfrac{3}{6})(\tfrac{2}{3}) + (\tfrac{3}{6}) \left[\tfrac{(2)(1)+(1)(0.2035)}{2+1}\right] \approx 0.70
\end{aligned}
$$

Following the same step, we obtain $\mathsf{hd}^{\pi}(\mathcal{T}_{\mathsf{M}}, \mathcal{T}_{\mathsf{AP}}) = 0.75$. Hence, $\mathsf{sim}^{\pi}(\mathsf{M}, \mathsf{AP}) \approx 0.73$ by using Definition 3.3.

Furthermore, using Definition 3.2, $\mathsf{hd}^{\pi}(\mathcal{T}_{\mathsf{AP}}, \mathcal{T}_{\mathsf{B}})$

$$
\begin{aligned}
&= (\tfrac{3}{6}) \cdot \mathsf{p\text{-}hd}^{\pi}(\mathcal{P}_{\mathsf{AP}}, \mathcal{P}_{\mathsf{B}}) + (\tfrac{3}{6}) \cdot \mathsf{e\text{-}set\text{-}hd}^{\pi}(\mathcal{E}_{\mathsf{AP}}, \mathcal{E}_{\mathsf{B}}) \\
&= (\tfrac{3}{6}) \cdot (\tfrac{\mathfrak{i}(X)\cdot\max\{\mathfrak{s}(X,Z),\mathfrak{s}(X,\mathsf{P})\}+\mathfrak{i}(\mathsf{P})\cdot\max\{\mathfrak{s}(\mathsf{P},\mathsf{Z}),\mathfrak{s}(\mathsf{P},\mathsf{P})\})}{\mathfrak{i}(X)+\mathfrak{i}(\mathsf{P})}) \\
&+(\tfrac{3}{6}) \cdot \mathsf{e\text{-}set\text{-}hd}^{\pi}(\mathcal{E}_{\mathsf{AP}}, \mathcal{E}_{\mathsf{B}}) \\
&= (\tfrac{3}{6})(\tfrac{1\cdot\max\{0,0\}+2\cdot\max\{0,1\}}{1+2}) + (\tfrac{3}{6}) \cdot \mathsf{e\text{-}set\text{-}hd}^{\pi}(\mathcal{E}_{\mathsf{AP}}, \mathcal{E}_{\mathsf{B}}) \\
&= (\tfrac{3}{6})(\tfrac{2}{3}) \\
&+(\tfrac{3}{6}) \left[\tfrac{2\cdot\max\{0.2035\}+\mathfrak{i}(\mathsf{cS})\cdot\max\{\mathsf{e\text{-}hd}^{\pi}(\exists \mathsf{cS.K},\exists \mathsf{cS,K})\}}{\mathfrak{i}(\mathsf{cW})+\mathfrak{i}(\mathsf{cS})}\right] \\
&= (\tfrac{3}{6})(\tfrac{2}{3}) + (\tfrac{3}{6}) \left[\tfrac{2\cdot\max\{0.2035\}+1\cdot\max\{(1)(0.3+(0.7)(1))\}}{\mathfrak{i}(\mathsf{cW})+\mathfrak{i}(\mathsf{cS})}\right] \\
&= (\tfrac{3}{6})(\tfrac{2}{3}) + (\tfrac{3}{6}) \left[\tfrac{(2)(0.2035)+(1)(1)}{2+1}\right] \approx 0.57
\end{aligned}
$$

Following the same step, we obtain $\mathsf{hd}^{\pi}(\mathcal{T}_{\mathsf{B}}, \mathcal{T}_{\mathsf{AP}}) = 0.75$. Hence, $\mathsf{sim}^{\pi}(\mathsf{B}, \mathsf{AP}) \approx 0.66$ by using Definition 3.3.

These results, i.e. $\mathsf{sim}^{\pi}(\mathsf{M}, \mathsf{AP}) > \mathsf{sim}^{\pi}(\mathsf{B}, \mathsf{AP})$, corresponds with Bob's needs and preferences. □

4 RELATED WORK

Several CSMs abound, but here we investigate those CSMs that exhibit preferential elements and contrast them with aspects of our preference profile. Except the following two works, most CSMs do not consider any of preferential elements. Hence, we omit discussions thereof and merely refer interested readers to their references for further details, namely (Janowicz and Wilkes, 2009; Racharak and Suntisrivaraporn, 2015; D'Amato et al., 2006; Fanizzi and D'Amato,) for structural-based similarity measures and (D'Amato et al., 2009; D'Amato et al., 2008) for interpretation-based measures.

In an extended work of sim, a range of number for discount factor (ν) is used in the similarity application of SNOMED CT. For instance, when the roleGroup is found, the value of ν is set to 0. That approach can be viewed as a specific application of $\mathfrak{d}$ function of preference profile. In *simi*, pairs of primitive concept names and pairs of role names are permitted to impose the similarity values via the function *pm*. For instance, given two primitive concept names A and B, we can establish the 50% similarity between A and B by defining $pm(A,B) = 0.5$. In the similar manner, given two role names r and s, we can establish the 50% similarity between r and s by defining $pm(r,s) = 0.5$. The former is identical to $\mathfrak{s}^{\mathfrak{c}}$ of preference profile; however, the latter differs from $\mathfrak{s}^{\mathfrak{r}}$ of preference profile in a sense that *pm* does not consider primitive role names which contribute to similarity between two arbitrary role names. Furthermore, each primitive concept name and each existential restriction atoms (i.e., those concepts of the form $\exists r.C$) is permitted to be weighted w.r.t. a positive real number via the function g. However, we believe that the imposition on existential restriction atoms will be impractical to use. After all, there can be infinitely many existential restriction atoms. Thus, our sim^{π} is developed according to preference profile by allowing to define an importance over each role name instead. Table 1 shows a summary of existing CSMs which expose elements of preference profile, together with our proposed sim^{π}, where ✔ denotes totally identical to the specified function whereas ✓ denotes partially identical to the specified function.

Table 1: State-of-the-art CSMs embeded preference profile.

CSM	$\mathfrak{i}^{\mathfrak{c}}$	$\mathfrak{i}^{\mathfrak{r}}$	$\mathfrak{s}^{\mathfrak{c}}$	$\mathfrak{s}^{\mathfrak{r}}$	$\mathfrak{d}$
sim^{π}	✔	✔	✔	✔	✔
sim					✔
simi	✓		✔		

5 CONCLUDING REMARKS

In this work, the applicability of the so-called *preference profile* is first exemplified by generalizing the mechanism of the similarity measure sim for the DL $\mathcal{ELH}$, called sim^{π}. Our sim^{π} can nicely utilize preferences of an agent, which are represented in form of a preference profile, for influencing the calculation. We also prove that sim^{π} is backward compatible in the sense that sim^{π} under the default preference profile coincides with sim. This finding together with Proposition 7 in (Tongphu and Suntisrivaraporn, 2015) are important to show that sim^{π} is preference-invariant w.r.t. equivalence, i.e. similarity between two equivalent concepts is always one regardless of agents' preferences. We also investigate existing CSMs and find that none of them, to the best of our knowledge, entirely comply with the preference profile.

There are some directions for our future work. Firstly, it appears to be a natural next step to verify desirable properties in which any CSMs under a preference profile must have. Secondly, we are going to investigate deeply on the possibility of other reasonable aspects to be included in the preference profile, especially when considering more expressive DLs. Thirdly, we intend to carry out an implementation of sim^{π} and perform experiments on realistic ontologies. Finally, it is interesting to explore the possibility to extend preference profile beyond other kinds of similarity-based reasoning services. i.e., relaxed instance checking and relaxed instance retrieval.

ACKNOWLEDGEMENTS

This research is partially supported by Thammasart University Research Fund under the TU Research Scholar, Contract No. TOR POR 1/13/2558; the Center of Excellence in Intelligent Informatics, Speech and Language Technology, and Service Innovation (CILS), Thammasat University; and the JAIST-NECTEC-SIIT dual doctoral degree program.

REFERENCES

D'Amato, C., Fanizzi, N., and Esposito, F. (2006). A dissimilarity measure for alc concept descriptions. *In Proceedings of the 2006 ACM Symposium on Applied Computing*, pages 1695–1699.

D'Amato, C., Fanizzi, N., and Esposito, F. (2009). A semantic similarity measure for expressive description logics. *In CoRR*, abs/0911.5043.

D'Amato, C., Staab, S., and Fanizzi, N. (2008). On the influence of description logics ontologies on conceptual similarity. *In Proceedings of Knowledge Engineering: Practice and Patterns*, pages 48–63.

Fanizzi, N. and D'Amato, C. A similarity measure for the aln description logic. *In Proceedings of CILC 2006 - Italian Conference on Computational Logic*, pages 26–27.

Janowicz, K. and Wilkes, M. (2009). Sim-dla: A novel semantic similarity measure for description logics reducing inter-concept to inter-instance similarity. *In Proceedings of the 6th European Semantic Web Conference on The Semantic Web: Research and Applications*, pages 353–367.

Lehmann, K. and Turhan, A.-Y. (2012). A framework for semantic-based similarity measures for elh-concepts. In del Cerro, L. F., Herzig, A., and Mengin, J., editors, *JELIA*, volume 7519 of *Lecture Notes in Computer Science*, pages 307–319. Springer.

Lichtenstein, S. and Slovic, P., editors (2006). *The construction of preference*. Cambridge University Press, New York.

Racharak, T. and Suntisrivaraporn, B. (2015). Similarity measures for fl0 concept descriptions from an automata-theoretic point of view. In *Information and Communication Technology for Embedded Systems (IC-ICTES), 2015 6th International Conference of*, pages 1–6.

Racharak, T., Suntisrivaraporn, B., and Tojo, S. (2015). Identifying an Agent's Preferences Toward Similarity Measures in Description Logics. *In Proceedings of The 5th Joint International Semantic Technology (JIST 2015)*.

Scherer, K. (2005). What are emotions? and how can they be measured? *Social Science Information*.

Son, T. C., Cao, T., and Pontelli, E. (2003). Planning with preferences using logic programming. In *In Proc. LPNMR'04*, pages 247–260. Springer.

Suntisrivaraporn, B. (2013). A similarity measure for the description logic el with unfoldable terminologies. *In INCoS*, pages 408–413.

Tongphu, S. and Suntisrivaraporn, B. (2015). Algorithms for measuring similarity between elh concept descriptions: A case study on snomed ct. *Journal of Computing and Informatics (accepted on May 7; to appear)*.