Relatedness of Fuzzy Sets

S. Schockaert*, M. De Cock, C. Cornelis

Dept. of Applied Mathematics and Computer Science

Ghent University, Belgium

Abstract

Comparison measures to assess similarity, inclusion, and overlap of fuzzy sets are well known and widely used. This paper goes one step further by bootstrapping a more general concept of relatedness of fuzzy sets based on the relatedness of their elements. We study properties of 6 relatedness measures of this kind, and we illustrate their use in interesting applications such as information retrieval, temporal reasoning and recommender systems.

1 Introduction

Fuzzy relations have proven to be very useful tools for the representation of relatedness, which is more often than not a matter of degree. An abundance

*Corresponding author. Krijgslaan 281 (S9), 9000 Gent, Belgium, tel: +3292644772, fax: +3292644995, e-mail: Steven.Schockaert@UGent.be

of applications based on this principle is described in the literature. As a well known example from information retrieval (see e.g. [4]), we recall that a fuzzy relation Q between a set of documents D and a set of terms T can be used to represent the importance $Q(d,t)$ of term t for document d. Likewise a fuzzy relation R in T is often used such that $R(t_1,t_2)$ denotes the degree of relatedness of terms t_1 and t_2, e.g. synonymy.

On the other hand, comparison measures to assess the similarity, inclusion and overlap of fuzzy sets have been around for many years now. Formally they are fuzzy relations over universes of fuzzy sets. The overlap of fuzzy sets A and B in T can for instance be measured by

$$\sup_{t\in T} \mathcal{T}(A(t), B(t)) \tag{1}$$

in which $\mathcal{T}$ is a triangular norm. If A corresponds to a weighted query q of keywords and B to a document d such that $B(t) = Q(d,t)$ for all t in T, then (1) can be used as a score to rank d in the retrieval results for q. This score measures the extent to which document d has at least one keyword in common with the query q. Documents that do not contain any of the query keywords get a score of 0, even if they contain content that might be relevant to the query. A helpful solution to this problem might be query expansion, the process of adding related terms. This gives rise to the following, alternative scoring formula

$$\sup_{t'\in T} \sup_{t\in T} \mathcal{T}(A(t'), \mathcal{T}(R(t',t), B(t))) \tag{2}$$

which is a direct fuzzification of

$$(\exists t' \in T)(\exists t \in T)(t' \in A \wedge (t',t) \in R \wedge t \in B) \tag{3}$$

Formula (2) constructs a relatedness measure between fuzzy sets based on the relatedness of their elements. As such it offers a more nuanced view than (1) where the membership value $B(t)$ is compared only to the membership value $A(t)$ of t in A; in (2) it is compared to the membership values of every element t' of A that is related to some degree $R(t', t)$ to t. The resulting formula is slightly more complicated but offers many advantages in applications.

The aim of this paper is to introduce a new class of relatedness measures in the spirit of (2). Formally, we will start from 3 different kinds of compositions of fuzzy relations. After recalling these compositions as well as the necessary background on triangular norms and their residual implicators in Section 2, we introduce 6 relatedness measures for fuzzy sets in Section 3. We show how they are ordered (namely which is the most strict and the most loose measure, and how the others fit in between). We show that they exhibit transitive behaviour, which allows to deduct information about the degree of relatedness of A and C from the relatedness of A and B and B and C respectively. In this context, fuzzy sets A, B and C can belong to different universes, which makes our measures clearly more general than traditional ones such as (1). Finally we illustrate their practical relevance for temporal reasoning and recommender systems in Section 4.

2 Preliminaries

Throughout this paper, let $\mathcal{T}$ denote a left–continuous triangular norm (t-norm for short). Recall that by definition a t-norm is an increasing $[0,1]^2 - [0,1]$ mapping that satisfies the following conditions, for all a, b and c in $[0,1]$

$$\mathcal{T}(1,a) = a \tag{4}$$

$$\mathcal{T}(a,b) = \mathcal{T}(b,a) \tag{5}$$

$$\mathcal{T}(\mathcal{T}(a,b),c) = \mathcal{T}(a,\mathcal{T}(b,c)) \tag{6}$$

Furthermore let $\mathcal{I}$ denote its residual implicator, which is defined as

$$\mathcal{I}(a,b) = \sup\{\lambda|\lambda \in [0,1] \text{ and } \mathcal{T}(a,\lambda) \leq b\} \tag{7}$$

$\mathcal{I}$ is decreasing in the first and increasing in the second argument. Furthermore

$$\mathcal{I}(1,b) = b \tag{8}$$

for all b in $[0,1]$, and $\mathcal{I}(0,0) = \mathcal{I}(0,1) = 1$. $\mathcal{T}$ and $\mathcal{I}$ satisfy the residuation principle, i.e. for all a, b and c in $[0,1]$ it holds that

$$\mathcal{T}(a,b) \leq c \Leftrightarrow a \leq \mathcal{I}(b,c) \tag{9}$$

In other words, $\mathcal{T}$ and $\mathcal{I}$ form a Galois connection. It can be shown that for all a, b, c and d in $[0,1]$ it holds that (see e.g. [6])

$$\begin{aligned}
a \leq b &\Leftrightarrow \mathcal{I}(a,b) = 1 && (10)\\
\mathcal{T}(\mathcal{I}(a,b),c) &\leq \mathcal{I}(a,\mathcal{T}(b,c)) && (11)\\
\mathcal{T}(a,\mathcal{I}(a,b)) &\leq b && (12)\\
\mathcal{I}(\mathcal{T}(a,b),c) &= \mathcal{I}(a,\mathcal{I}(b,c)) && (13)\\
\mathcal{I}(a,\mathcal{I}(b,c)) &= \mathcal{I}(b,\mathcal{I}(a,c)) && (14)\\
\mathcal{T}(\mathcal{I}(a,b),\mathcal{I}(c,d)) &\leq \mathcal{I}(\mathcal{T}(a,c),\mathcal{T}(b,d)) && (15)
\end{aligned}$$

If J is an arbitrary index set and if $(a_j)_{j\in J}$ and $(b_j)_{j\in J}$ are families in $[0,1]$, it holds that

$$\begin{aligned}
\mathcal{T}(\sup_{j\in J} a_j, b) &= \sup_{j\in J} \mathcal{T}(a_j, b) && (16)\\
\mathcal{I}(\sup_{j\in J} a_j, b) &= \inf_{j\in J} \mathcal{I}(a_j, b) && (17)\\
\mathcal{I}(a, \inf_{j\in J} b_j) &= \inf_{j\in J} \mathcal{I}(a, b_j) && (18)\\
\mathcal{T}(\inf_{j\in J} a_j, b) &\leq \inf_{j\in J} \mathcal{T}(a_j, b) && (19)\\
\mathcal{I}(\inf_{j\in J} a_j, b) &\geq \sup_{j\in J} \mathcal{I}(a_j, b) && (20)\\
\mathcal{I}(a, \sup_{j\in J} b_j) &\geq \sup_{j\in J} \mathcal{I}(a, b_j) && (21)
\end{aligned}$$

Let R be a fuzzy relation from U to V (i.e. a fuzzy set in $U \times V$) and S a fuzzy relation from V to W. The well known sup-$\mathcal{T}$-composition of R and S is the fuzzy relation $R \circ_\mathcal{T} S$ from U to W defined by, for all u in U and w in W,

$$R \circ_\mathcal{T} S(u,w) = \sup_{v\in V} \mathcal{T}(R(u,v), S(v,w)) \tag{22}$$

The subproduct $R \triangleleft_{\mathcal{I}} S$ and the superproduct $R \triangleright_{\mathcal{I}} S$ of R and S are defined by [2]

$$R \triangleleft_{\mathcal{I}} S(u,w) = \inf_{v \in V} \mathcal{I}(R(u,v), S(v,w)) \tag{23}$$

$$R \triangleright_{\mathcal{I}} S(u,w) = \inf_{v \in V} \mathcal{I}(S(v,w), R(u,v)) \tag{24}$$

A fuzzy relation R from U to U (also called a fuzzy relation in U) is said to be reflexive iff $R(u,u) = 1$ for all u in U; R is called irreflexive iff $R(u,u) = 0$ for all u in U; R is called symmetric iff $R(u,v) = R(v,u)$ for all u and v in U. R is called $\mathcal{T}$- transitive iff

$$\mathcal{T}(R(u,v), R(v,w)) \leq R(u,w) \tag{25}$$

for all u, v and w in U.

3 Relatedness Measures

Let A be a fuzzy set in U, R a fuzzy relation from U to V and B a fuzzy set in V. We would like to express the degree to which A and B are related w.r.t. R. Formally we will first transform A and B into fuzzy relations and then compose them with R. The left extension $\overleftarrow{A}$ of A is the fuzzy relation in U defined for all u_1 and u_2 in U by

$$\overleftarrow{A}(u_1, u_2) = A(u_2) \tag{26}$$

Likewise, the right extension $\overrightarrow{B}$ of B is the fuzzy relation in V defined for all v_1 and v_2 in V by

$$\overrightarrow{B}(v_1, v_2) = B(v_1) \tag{27}$$

Since $\overleftarrow{A}$ is a fuzzy relation from U to U, the composition $\overleftarrow{A} \circ_{\mathcal{T}} R$ is a fuzzy relation from U to V. Moreover, since $\overrightarrow{B}$ is a fuzzy relation from V to V, the composition $(\overleftarrow{A} \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} \overrightarrow{B}$ is a fuzzy relation from U to V. It is easy to see that $(\overleftarrow{A} \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} \overrightarrow{B}$ is a constant mapping whose value in an arbitrary point (u, v) expresses the degree to which there exists an element of B which is related to an element of A. In [3] $(\overleftarrow{A} \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} \overrightarrow{B}$ is called the (fuzzy relational) double image of A and B. For convenience, we will denote the value of the mapping $(\overleftarrow{A} \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} \overrightarrow{B}$ in an arbitrary point (u, v) of $U \times V$ by $(A \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} B$. By using other compositions than sup-$\mathcal{T}$–composition, we obtain alternative relatedness measures. In this paper, we will focus on the following measures:

- The degree to which an element of B is related to an element of A:

$$(A \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} B = \sup_{v \in V} \mathcal{T}(\sup_{u \in U} \mathcal{T}(A(u), R(u, v)), B(v))$$

- The degree to which an element of A is related to an element of B:

$$A \circ_{\mathcal{T}} (R \circ_{\mathcal{T}} B) = \sup_{u \in U} \mathcal{T}(A(u), \sup_{v \in V} \mathcal{T}(R(u, v), B(v)))$$

- The degree to which every element of B is related to every element of A:

$$(A \lhd_{\mathcal{I}} R) \rhd_{\mathcal{I}} B = \inf_{v \in V} \mathcal{I}(B(v), \inf_{u \in U} \mathcal{I}(A(u), R(u, v)))$$

- The degree to which every element of A is related to every element of B:

$$A \lhd_{\mathcal{I}} (R \rhd_{\mathcal{I}} B) = \inf_{u \in U} \mathcal{I}(A(u), \inf_{v \in V} \mathcal{I}(B(v), R(u, v)))$$

- The degree to which an element of B is related to every element of A:

$$(A \lhd_{\mathcal{I}} R) \circ_{\mathcal{T}} B = \sup_{v \in V} \mathcal{T}(\inf_{u \in U} \mathcal{I}(A(u), R(u, v)), B(v))$$

- The degree to which every element of A is related to an element of B:

$$A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B) = \inf_{u \in U} \mathcal{I}(A(u), \sup_{v \in V} \mathcal{T}(R(u,v), B(v)))$$

- The degree to which every element of B is related to an element of A:

$$(A \circ_{\mathcal{T}} R) \rhd_{\mathcal{I}} B = \inf_{v \in V} \mathcal{I}(B(v), \sup_{u \in U} \mathcal{T}(A(u), R(u,v)))$$

- The degree to which an element of A is related to every element of B:

$$A \circ_{\mathcal{T}} (R \rhd_{\mathcal{I}} B) = \sup_{u \in U} \mathcal{T}(A(u), \inf_{v \in V} \mathcal{I}(B(v), R(u,v)))$$

Proposition 1 (Coincidence)**.**

$$(A \circ_{\mathcal{T}} R) \circ_{\mathcal{T}} B \quad = \quad A \circ_{\mathcal{T}} (R \circ_{\mathcal{T}} B) \tag{28}$$

$$(A \lhd_{\mathcal{I}} R) \rhd_{\mathcal{I}} B \quad = \quad A \lhd_{\mathcal{I}} (R \rhd_{\mathcal{I}} B) \tag{29}$$

Proof. A proof for (28) follows from (16) and the associativity of $\mathcal{T}$, i.e. (6). Likewise a proof for (29) follows from (18) and (14). □

As a consequence of Proposition 1, we will omit the parentheses and write $A \circ_{\mathcal{T}} R \circ_{\mathcal{T}} B$ and $A \lhd_{\mathcal{I}} R \rhd_{\mathcal{I}} B$. Computing the relatedness with the empty set using the measures above can lead to results that are not very intuitive, though mathematically correct. This phenomenon is related to the fact that quantification over the empty set always evaluates to true. It becomes apparent when computing e.g. the degree to which every element of $\emptyset$ is related to every element of B, namely $\emptyset \lhd_{\mathcal{I}} R \rhd_{\mathcal{I}} B = 1$. To avoid this case, in the remainder of this paper we will assume that the fuzzy sets under consideration are normalized, i.e. that they contain at least one element to membership degree 1. This requirement is usually easy to satisfy in applications.

$$
\begin{tikzcd}
& A \circ_{\mathcal{T}} R \circ_{\mathcal{T}} B & \\
A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B) \arrow[ur] & & (A \circ_{\mathcal{T}} R) \rhd_{\mathcal{T}} B \arrow[ul] \\
(A \lhd_{\mathcal{T}} R) \circ_{\mathcal{T}} B \arrow[u] & & A \circ_{\mathcal{T}} (R \rhd_{\mathcal{T}} B) \arrow[u] \\
& A \lhd_{\mathcal{T}} R \rhd_{\mathcal{T}} B \arrow[ul] \arrow[ur] &
\end{tikzcd}
$$

Figure 1: Ordering of relatedness measures

Proposition 2. The relatedness measures obey the ordering displayed in Figure 1 i.e.

$$
\begin{aligned}
A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B) &\leq A \circ_{\mathcal{T}} R \circ_{\mathcal{T}} B && (30)\\
(A \circ_{\mathcal{T}} R) \rhd_{\mathcal{I}} B &\leq A \circ_{\mathcal{T}} R \circ_{\mathcal{T}} B && (31)\\
A \lhd_{\mathcal{I}} R \rhd_{\mathcal{I}} B &\leq (A \lhd_{\mathcal{I}} R) \circ_{\mathcal{T}} B && (32)\\
A \lhd_{\mathcal{I}} R \rhd_{\mathcal{I}} B &\leq A \circ_{\mathcal{T}} (R \rhd_{\mathcal{I}} B) && (33)\\
(A \lhd_{\mathcal{I}} R) \circ_{\mathcal{T}} B &\leq A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B) && (34)\\
A \circ_{\mathcal{T}} (R \rhd_{\mathcal{I}} B) &\leq (A \circ_{\mathcal{T}} R) \rhd_{\mathcal{I}} B && (35)
\end{aligned}
$$

Proof. Since A is normalized there exists an m in U such that $A(m) = 1$. Furthermore using (8) and (4) we obtain

$$
\begin{aligned}
A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B) &\leq \mathcal{I}(A(m), \sup_{v \in V} \mathcal{T}(R(m,v), B(v)))\\
&= \mathcal{T}(A(m), \sup_{v \in V} \mathcal{T}(R(m,v), B(v)))\\
&\leq A \circ_{\mathcal{T}} (R \circ_{\mathcal{T}} B)
\end{aligned}
$$

The proofs of (31)-(33) are analogous. By (19), we have

$$
\begin{aligned}
& (A \lhd_{\mathcal{I}} R) \circ_{\mathcal{T}} B \\
\leq \quad & \sup_{v \in V} \inf_{u \in U} \mathcal{T}(\mathcal{I}(A(u), R(u,v)), B(v))
\end{aligned}
$$

and by (11) and (21)

$$
\begin{aligned}
\leq \quad & \sup_{v \in V} \inf_{u \in U} \mathcal{I}(A(u), \mathcal{T}(R(u,v), B(v))) \\
\leq \quad & \inf_{u \in U} \sup_{v \in V} \mathcal{I}(A(u), \mathcal{T}(R(u,v), B(v))) \\
\leq \quad & A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B)
\end{aligned}
$$

which proves (34); the proof of (35) is analogous. □

The following proposition also involves a fuzzy relation S from V to W and a fuzzy set C in W.

Proposition 3 (Transitivity)**.**

$$
\begin{gathered}
\mathcal{T}(A \lhd_{\mathcal{I}} R \rhd_{\mathcal{I}} B, B \lhd_{\mathcal{I}} S \rhd_{\mathcal{I}} C) \\
\leq A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} S) \rhd_{\mathcal{I}} C
\end{gathered} \tag{36}
$$

$$
\begin{gathered}
\mathcal{T}((A \lhd_{\mathcal{I}} R) \circ_{\mathcal{T}} B, (B \lhd_{\mathcal{I}} S) \circ_{\mathcal{T}} C) \\
\leq (A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} S)) \circ_{\mathcal{T}} C
\end{gathered} \tag{37}
$$

$$
\begin{gathered}
\mathcal{T}(A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} B), B \lhd_{\mathcal{I}} (S \circ_{\mathcal{T}} C)) \\
\leq A \lhd_{\mathcal{I}} ((R \circ_{\mathcal{T}} S) \circ_{\mathcal{T}} C)
\end{gathered} \tag{38}
$$

$$
\begin{gathered}
\mathcal{T}((A \circ_{\mathcal{T}} R) \rhd_{\mathcal{I}} B, (B \circ_{\mathcal{T}} S) \rhd_{\mathcal{I}} C) \\
\leq (A \circ_{\mathcal{T}} (R \circ_{\mathcal{T}} S)) \rhd_{\mathcal{I}} C
\end{gathered} \tag{39}
$$

$$\mathcal{T}(A \circ_{\mathcal{T}} (R \rhd_{\mathcal{I}} B), B \circ_{\mathcal{T}} (S \rhd_{\mathcal{I}} C))$$
$$\leq A \circ_{\mathcal{T}} ((R \circ_{\mathcal{T}} S) \rhd_{\mathcal{I}} C) \tag{40}$$

Proof. As an example we prove (36). Since B is normalized there exists an m in V such that $B(m) = 1$.

$$\begin{aligned}
& \mathcal{T}(A \lhd_{\mathcal{I}} R \rhd_{\mathcal{I}} B, B \lhd_{\mathcal{I}} S \rhd_{\mathcal{I}} C) \\
= \; & \mathcal{T}(\inf_{v \in V} \mathcal{I}(B(v), \inf_{u \in U} \mathcal{I}(A(u), R(u,v))), \\
& \inf_{w \in W} \mathcal{I}(C(w), \inf_{v \in V} \mathcal{I}(B(v), S(v,w)))) \\
\leq \; & \mathcal{T}(\mathcal{I}(B(m), \inf_{u \in U} \mathcal{I}(A(u), R(u,m))), \\
& \inf_{w \in W} \mathcal{I}(C(w), \mathcal{I}(B(m), S(m,w))))
\end{aligned}$$

By (8) we obtain

$$\begin{aligned}
= \; & \mathcal{T}(\inf_{u \in U} \mathcal{I}(A(u), R(u,m)), \\
& \inf_{w \in W} \mathcal{I}(C(w), S(m,w)))
\end{aligned}$$

By (19), we have

$$\begin{aligned}
\leq \; & \inf_{u \in U} \inf_{w \in W} \mathcal{T}(\mathcal{I}(A(u), R(u,m)), \\
& \mathcal{I}(C(w), S(m,w)))
\end{aligned}$$

and by (15)

$$
\begin{aligned}
&\leq \inf_{u\in U} \inf_{w\in W} \mathcal{I}(\mathcal{T}(A(u), C(w)), \\
&\qquad\qquad \mathcal{T}(R(u,m), S(m,w))) \\
&\leq \inf_{u\in U} \inf_{w\in W} \mathcal{I}(\mathcal{T}(A(u), C(w)), \\
&\qquad\qquad \sup_{v\in V} \mathcal{T}(R(u,v), S(v,w))) \\
&= \inf_{u\in U} \inf_{w\in W} \mathcal{I}(\mathcal{T}(A(u), C(w)), (R \circ_{\mathcal{T}} S)(u,w))
\end{aligned}
$$

Finally, by (13) and (18)

$$
\begin{aligned}
&= \inf_{u\in U} \inf_{w\in W} \mathcal{I}(C(w), \mathcal{I}(A(u), (R \circ_{\mathcal{T}} S)(u,w))) \\
&= \inf_{w\in W} \mathcal{I}(C(w), \inf_{u\in U} \mathcal{I}(A(u), (R \circ_{\mathcal{T}} S)(u,w))) \\
&= A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} S) \rhd_{\mathcal{I}} C
\end{aligned}
$$

$\square$

Proposition 4 (Reflexivity)**.** *If R is a reflexive fuzzy relation in U then*

$$A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} A) = 1 \tag{41}$$

$$(A \circ_{\mathcal{T}} R) \rhd_{\mathcal{I}} A = 1 \tag{42}$$

Proof. As an example we prove (41).

$$\begin{aligned} & A \lhd_{\mathcal{I}} (R \circ_{\mathcal{T}} A) \\ = \; & \inf_{u \in U} \mathcal{I}(A(u), \sup_{v \in U} \mathcal{T}(R(u,v), A(v))) \\ \geq \; & \inf_{u \in U} \mathcal{I}(A(u), \mathcal{T}(R(u,u), A(u))) \\ = \; & \inf_{u \in U} \mathcal{I}(A(u), A(u)) \end{aligned}$$

By (10) we verify that the latter is 1. □

Proposition 5 (Irreflexivity)**.** *If R is an irreflexive fuzzy relation in U then*

$$A \circ_{\mathcal{T}} (R \rhd_{\mathcal{I}} A) = 0 \tag{43}$$

$$(A \lhd_{\mathcal{I}} R) \circ_{\mathcal{T}} A = 0 \tag{44}$$

Proof. As an example we prove (43).

$$\begin{aligned} & A \circ_{\mathcal{T}} (R \rhd_{\mathcal{I}} A) \\ = \; & \sup_{u \in U} \mathcal{T}(A(u), \inf_{v \in U} \mathcal{I}(A(v), R(u,v))) \\ \leq \; & \sup_{u \in U} \mathcal{T}(A(u), \mathcal{I}(A(u), R(u,u))) \end{aligned}$$

By (12) we verify that the latter equals 0. □

4 Applications

The measures introduced in Section 3 can be used to compute virtually any kind of relatedness between fuzzy sets (defined in the same or in different universes).

The semantics of the relatedness is brought into the measure through a fuzzy relation between the universes of the fuzzy sets. We demonstrate this process with 2 applications. Firstly, using fuzzy ordering relations between time points such as "December 1, 1930 came long before January 31, 1992", we propose expressive representations of qualitative relations between time events such as "The *late* 1930's came *long* before the *early* 1990's". Note that both the events as well as the relation between them is imprecise!

The second application example concerns recommender systems. Our approach resembles the information retrieval techniques touched upon in Section 1. Like a document is represented as a fuzzy set of terms, a customer is represented by the items he bought and/or liked in the past. We use a similarity relation between items to identify similarities between users, even if their past shopping baskets do not have any items in common. We also show how this helps to overcome the "new item problem", which is how to decide on recommending an item that previously has not been bought by anyone, e.g. because it is newly introduced into the market.

Temporal Reasoning. Temporal information plays an important role in our knowledge about the world. In practice, available temporal information is often qualitative, rather than quantitative, in nature. Several qualitative temporal reasoning models have been developed (see [10] for an overview). One option is to model the time span of an event as an interval in $\mathbb{R}$; Allen [1] defined a set of 13 jointly exhaustive and pairwise disjoint interval relations for this purpose. For example, $before([a, b], [c, d])$ holds iff $b < c$, $meets([a, b], [c, d])$ holds iff $b = c$ and $overlaps([a, b], [c, d])$ holds iff $a < c$, $b < d$ and $c < b$. All 13 interval relations can be expressed by referring to the relative positioning of the endpoints of the intervals.

However, many events are inherently imprecise, i.e. they are characterized by a gradual beginning and/or ending. The time span of this kind of events can be represented by a fuzzy time interval, i.e. a normalized, convex fuzzy set in $\mathbb{R}$. This however requires a suitable generalization of Allen's interval relations. For example, a fuzzy relation $meets$ has to be defined, such that for fuzzy time intervals A and B, $meets(A, B)$ expresses the degree to which the end of A is equal to the beginning of B. In [5] such an interpretation has been introduced. However, this approach suffers from a number of important disadvantages. For example, if A and B are continuous mappings, then $meets(A, B) \leq before(A, B)$, i.e. this approach fails to properly differentiate between $meets$ and $before$. In the alternative approach introduced in [7], many important properties, such as $\mathcal{T}$-transitivity, antisymmetry and (ir)reflexivity, are lost.

By using the relatedness measures introduced in this paper, we obtain an interpretation which allows to model imprecise qualitative relations, without any of the aforementioned shortcomings of the existing approaches. Indeed, let $L^{\ll}$ be an irreflexive, $\mathcal{T}$–transitive fuzzy relation in $\mathbb{R}$ such that for u and v in $\mathbb{R}$, $L^{\ll}(u, v)$ expresses the extent to which u is much smaller than v. Moreover, let $L^{\preccurlyeq}$ be a reflexive, $\mathcal{T}$–transitive fuzzy relation in $\mathbb{R}$ such that $L^{\preccurlyeq}(u, v)$ expresses the extent to which u is smaller than or approximately equal to v. Then $A \lhd_{\mathcal{I}}$

$L^{\lll} \triangleright_{\mathcal{I}} B$ expresses the degree to which the end of A is much smaller than the beginning of B. As as consequence, we define the fuzzy relation $before$ by

$$before(A, B) = A \triangleleft_{\mathcal{I}} L^{\lll} \triangleright_{\mathcal{I}} B \tag{45}$$

In the same way, we have

$$meets(A, B) = \min(A \triangleleft_{\mathcal{I}} L^{\preccurlyeq} \triangleright_{\mathcal{I}} B, B \circ_{\mathcal{T}} L^{\preccurlyeq} \circ_{\mathcal{T}} A) \tag{46}$$

where $A \triangleleft_{\mathcal{I}} L^{\preccurlyeq} \triangleright_{\mathcal{I}} B$ expresses the degree to which the end of A is smaller than or approximately equal to the beginning of B, and $B \circ_{\mathcal{T}} L^{\preccurlyeq} \circ_{\mathcal{T}} A$ expresses the degree to which the beginning of B is smaller than or approximately equal to the end of A. Similar interpretations can be given for the other qualitative interval relations. The properties established in Section 3 support temporal reasoning.

Recommender systems. Recommender systems are applications that anticipate the needs of e-service users by referring them to information items (or products, in the context of e-commerce) that are likely to interest them. One popular strategy is collaborative filtering (CF) [9], which recommends an item to someone if it was purchased and/or rated positively by others who share similar characteristics with them. Another approach is called content-based (CB), or reclusive [11], and identifies objects that sufficiently resemble those the user has liked in the past. Fuzzy set theory lends itself very well to reflect the uncertainties and gradedness inherent to these processes; in [8], Perny and Zucker developed a hybrid CF-CB approach to predict the degree to which a user will like an item, given fuzzy relations R and S that represent user, resp. item similarity. A few important problems remain; we briefly address how the relatedness measures introduced in this paper can aid.

Traditionally, in the absence of explicit profile information, users' preference judgments are compared to assess the similarity of users. Let U and I be the universe of users, resp. items, and let A and B be the fuzzy sets of items liked in the past by user u, resp. user v (the degree of "liking" an item could be established based on a rating of the product expressed by the users). The formula

$$R(u,v) = \inf_{i \in I} \mathcal{I}(A(i), B(i)) \tag{47}$$

evaluates to what extent all items that u likes are also popular with v. Such an approach however quickly incurs practical problems, as it is hardly realistic to expect that every user has rated every item, and so $A(i)$ and $B(i)$ are not always available. A patch is therefore sometimes provided, replacing I by the set of items that u and v have commonly rated, see e.g. [8], but this is not a satisfying nor an elegant solution either. Suppose for instance that I is a book collection, and that u has positively evaluated novels by Edgar Allan Poe and Mary Shelley, and that v likes "Wuthering heights" by Emily Bronte and "Dracula" by Bram Stoker (but did not rate any of the books u read), hence both of them seem to be fond of gothic literature, and u might make a very good advisor to v. However, since they have no rated items in common, the above approach cannot discover this shared interest. However, when

$$R(u,v) = A \triangleleft_{\mathcal{I}} (S \circ_{\mathcal{T}} B) \tag{48}$$

the fuzzy relation S (which in general is relatively easy to establish based on the internal representation of the items) comes into play, exploiting the similarity between the considered items by evaluating to what extent for each item u likes, there exists a *similar* item that v likes. Under these assumptions, $R(u, v)$ can indeed assume a high value in our example.

Another problem concerns recommendation itself. In the traditional setting, i is recommended to user u only when similar users approve of i, i.e.

$$\sup_{v \in U} \mathcal{T}(R(u, v), P(v, i)) \tag{49}$$

is used to predict the preference of user u towards item i, based on the revealed preference $P(v, i)$ of every other user v towards i. This approach is in particular susceptible to the so-called new item problem: items not yet rated/purchased by anyone cannot be recommended, which is a significant drawback of CF. To overcome this problem, as a predictive value, we propose to use the relatedness of the fuzzy set of users similar to u with the fuzzy set of products similar to i. These fuzzy sets Ru and Si are respectively defined by $Ru(v) = R(u, v)$ and $Si(j) = S(i, j)$ for all v in U and all j in I, giving rise to the final recommendation score $Ru \circ_{\mathcal{T}} P \circ_{\mathcal{T}} Si$.

5 Conclusion

We introduced 6 different measures to assess the relatedness of fuzzy sets, based not only on the degree to which they contain elements (as is traditionally done) but also on the relatedness of these elements. The measures are partially ordered. They inherit the (ir)reflexive and transitive behaviour of the underlying fuzzy relation on the level of the elements. We used the measures to express imprecise qualitative relations between time events, as well as a solution to the new item problem in recommender system. In general, problems that require the comparison between fuzzy sets in any way, even if in different universes, are possible candidates for the application of our measures.

Acknowledgements The authors would like to thank the Fund for Scientific Research–Flanders for funding the research reported on in this paper. The Fuzziness and Uncertainty Modelling Research Unit is a participant of the European COST Action 274 TARSKI: Theory and Applications of Relational Structures as Knowledge Instruments.

References

[1] J.F. Allen, Maintaining knowledge about temporal intervals, Communications of the ACM 26(11), 1983, 832–843.

[2] W. Bandler and L.J. Kohout, Fuzzy relational products as a tool for analysis and synthesis of the behaviour of complex natural and artificial systems, in: Fuzzy Sets: Theory and Application to Policy Analysis and Information Systems, S.K. Wang and P.P. Chang, eds., Plenum Press, New York and London, 1980, pp. 341-367.

[3] J.A. Goguen, L-fuzzy sets, Journal of Mathematical Analysis and Applications 18, 1967, 145-174.

[4] S. Miyamoto, T. Miyake, and K. Nakayama, Generation of a pseudothesaurus for information retrieval based on co-occurrences and fuzzy set operations, IEEE Transactions on Systems, Man, and Cybernetics 13(1), 1983, 62-70.

[5] G. Nagypál and B. Motik, A fuzzy model for representing uncertain, subjective and vague temporal knowledge in ontologies, Lecture Notes in Computer Science 2888, 2003, 906–923.

[6] V. Novák, I. Perfilieva, and J. Močkoř, Mathematical principles of fuzzy logic, Kluwer Academic Publishers, 1999.

[7] H. J. Ohlbach, Relations between fuzzy time intervals, in: Proceedings of the 11th International Symposium on Temporal Representation and Reasoning, 2004, pp. 44–51.

[8] P. Perny and J.D. Zucker, Preference-based search and machine learning for collaborative filtering: the "Film-Conseil" movie recommender system, Revue I3 1(1), 2001, 1–40.

[9] P. Resnick, N. Iacovou, M. Sushak, P. Bergstrom, and J. Riedl, GroupLens: an open architecture for collaborative filtering of netnews, in: Proceedings of the 1994 Computer Supported Collaborative Work Conference, 1994, pp. 175–186.

[10] L. Vila, A survey on temporal reasoning in artificial intelligence, AI Communications 7(1), 1994, 4–28.

[11] R. Yager, Fuzzy logic methods in recommender systems, Fuzzy Sets and Systems 136, 2003, 133–149.