sciendo
Control and Cybernetics
vol. **51** (2022) No. 2
pages: 267-283
DOI: 10.2478/candc-2022-0017

A novel approach to hierarchical contextual bipolar queries: A winnow operator approach*

by

Janusz Kacprzyk[*,**,***] **and Sławomir Zadrożny**[*,**]

*Systems Research Institute, Polish Academy of Sciences
Newelska 6, 01–447 Warsaw, Poland
{kacprzyk,zadrozny}@ibspan.waw.pl
**WIT – Warsaw School of Information Technology
Newelska 6, 01–447 Warsaw, Poland
***Faculty of Computer Science, Electronics and Telecommunications
AGH University of Science and Technology
Al. Mickiewicza 30 30-059 Kraków, Poland

Abstract: We propose a new approach to the bipolar database queries, which involve a necessary (required) and optional (desired) conditions, connected with a non-conventional aggregation operator "and possibly", combined with a context, exemplified by "find houses which are cheap and – with respect to other houses in town – possibly close to a railroad station". We use our winnow operator based interpretation of the bipolar queries. We assume that the query, posed by the human user, involves terms, which do not directly relate to attributes, and which are then to be *decoded* using a concept of a query hierarchy, leading to the queries, which involve terms directly related to attribute values. The original query is considered to be of level 0, at the bottom of the precisiation hierarchy, then its required and optional parts are assumed to be bipolar queries themselves, both accounting for context. The precisiation proceeds further, to level 1 queries, level 2, etc. A real estate related example is provided as illustration.

Keywords: database query, bipolar query, context, fuzzy logic, user intention, user preference

1. Introduction

The paper is concerned with database querying, as well as with information retrieval, Web search, etc., in which we face a serious problem with a discrepancy

*Submitted: April 2022; Accepted: May 2022.

(gap) between the human user and the computer, that is, a system providing access to data/information such as, e.g., a database management system (DBMS). Basically, for a human being, natural language is the only fully natural way of articulation and communication, and the humans tend to pose questions, queries, requests, etc. primarily in natural language, which is clearly strange to the computer. Therefore, such questions, queries, etc. should first be transformed into another form, suitable for the computer. One of the challenges of such a transformation is the fact that the user often phrases his or her queries using imprecise terms and – what is even more relevant for this paper – using terms which do not directly correspond to the features of the queried data. In our paper we will use examples from real estate databases and their related querying, where a query such as "find all houses that are comfortable and well located" may very well express human intentions and preferences. However, "comfortable" and "well located" are inherently imprecise and general terms that may not have directly corresponding numerical data features if taken literally. Thus, such a query should be further precisiated by *decoding* these "comfortable" and "well located" terms into some set of characteristics, which are directly related to the data features exemplified by "with a couple of large rooms" for the case of "comfortable", or "close to public transportation" for the case of "well located". Notice that in this case we have the number of rooms and their size, and the distance to, for instance, a bus stop, so that, first, the values are numerical and can be directly included in the database, and the fuzzy conditions "a couple", "large" and "close" constitute an imprecise formulation of the query terms with flexible constraints on values of these numerical attributes.

Our aim is therefore to provide a database user with a possibly high flexibility in forming a query by making it possible to use, first, imprecise conditions, which can be of different importance. The point of departure for our work is a new direction in fuzzy querying, which tries to further extend some sophisticated approaches, based on fuzzy logic with linguistic quantifiers, to be more specific in their version due to Kacprzyk and Ziółkowski (1986), and Kacprzyk, Zadrożny and Ziółkowski (1989). Then, we further assume that we explicitly account for *bipolarity* in human judgments and intentions and/or preferences, notably using the approach by Zadrożny and Kacprzyk (2012), see also Kacprzyk and Zadrożny (1999). We also add to the bipolar queries our approach to the involvement of *context* in bipolar queries, as proposed by Zadrożny, Kacprzyk and Dziedzic (2015), see also Zadrożny et al. (2014). And, finally, we combine these constituents with our concept of hierarchical contextual bipolar queries, proposed by Kacprzyk and Zadrożny (2018) in its basic form. Now, we extend this approach by using a winnow operator based representation of bipolar queries which is a new element.

In Section 2 we briefly remind the essence of one of possible approaches to bipolarity in querying. Next, in Section 3 the concept of the context, as used in our approach to bipolar queries, is briefly discussed. Then, in Section 4 we will present the new hierarchical contextual bipolar queries.

2. Bipolar fuzzy queries

The term *bipolar query*, coined in our context by Dubois and Prade (2002, 2008,a,b, 2009), is basically related to having two types of query conditions, which may be viewed as expressing negative and positive user preferences. This is implied by the results of psychological research, which finds that a human being in his/her assessments or evaluations usually prefers to use some sort of a *bipolar scale* through:

- some degree of being *negative*, i.e., to be rejected,
- some degree of being *positive*, i.e., to be accepted.

Fuzzy logic has been found to constitute a very powerful tool for formalizing and processing such bipolar queries, and more information on the traditional approach to the formalization of bipolarity in judgments and evidence via fuzzy logic and possibility theory can be found in, for instance, the works of Dubois and Prade and their collaborators, e.g.: Benferhat et al. (2008), Dubois and Prade (2002, 2008a,b, 2009), cf. also Dubois and Prade (2014), Dziedzic, Kacprzyk and Zadrożny (2014), Hadjali, Kaci and Prade (2011), Lietard and Rocacher (2009), Lietard, Tamani and Rocacher (2011), Matthé et al. (2011), etc.

In practice, two bipolar scales are used (cf. Grabisch, Greco and Pirlot, 2008):

- *bipolar univariate* and
- *unipolar bivariate*,

with the former assuming one scale with three main levels of, respectively, negative, neutral and positive evaluation, gradually changing from one end of the scale to another, usually represented by $[-1, 1]$, while the latter assuming two independent scales for a positive and negative evaluation, usually represented by $[0, 1]$. The latter is more convenient for our purposes, because it makes it possible to separately deal with the negative and positive evaluations.

The crucial issue is to assume a proper semantics of the negative and positive evaluations. In our context, we assume that the objects (tuples in the database context) with a negative evaluation are rejected and a positive evaluation contributes to the overall evaluation of an object (tuple) only if it is not rejected. Here, we additionally assume that these positive/negative evaluations are gradual, with values from the unit interval $[0, 1]$. Even if the positive evaluations play a weaker role, they are equally important as negative evaluations in the case when there exist multiple non-rejected objects with positive evaluations. This semantics is usually formalized by introducing a special aggregation operator "and possibly".

A prototypical example of such a bipolar query is:

$$C \text{ and possibly } P \tag{1}$$

exemplified by "find a house which is inexpensive (C) and possibly close to public transportation (P)", meaning, in this case, that the above query is satisfied by a tuple t only if either one of the two conditions holds:

1. it satisfies (possibly to a high degree) both conditions C and P, or
2. it satisfies only C and there is no tuple in the whole database which satisfies both conditions.

Thus, we adopt our approach (cf. Zadrożny, 2005; Zadrożny and Kacprzyk, 2006, 2012; Zadrożny, De Tré and Kacprzyk, 2010) that is basically an extension of the classic Lacroix and Lavency (2010) (nonfuzzy) approach, to the so-called queries with preferences, cf. Zadrożny and Kacprzyk (2012) for details. Some other interpretations of the "and possibly" operator can be found in, for instance, Bosc and Pivert (2012), Lietard, Tamani and Rocacher (2011), or Dujmovic (1979), see also De Tré, Zadrożny and Bronselaer (2010).

The concept of such a bipolar query first appeared in the seminal paper of Lacroix and Lavency (1987) who proposed the use of a query (C, P) with two categories of conditions: C which is required (mandatory), and P, which expresses just mere preferences (desires). The semantics is provided by the "and possibly" operator, that is – if at least one tuple in the given database satisfies both mandatory and desired condition, then the "and possibly" operator is interpreted as the standard conjunction, otherwise only the mandatory condition is taken into account.

Such an aggregation operator has been later proposed in a fuzzy logic setting independently by Dubois and Prade (1988) in default reasoning, Yager (1992, 1996) in multicriteria decision making, and Bordogna and Pasi (1995) in information retrieval.

The bipolar queries with the "and possibly" operator may be also viewed as a special case of Chomicki's (2002, 2003) *queries with preferences*, which are based on an extra relational algebra operator, the *winnow* (cf. Zadrożny and Kacprzyk, 2012). The winnow is associated with a *preference relation* over the universe of tuples, and returns as a result those tuples, which are *non-dominated* with respect to this preference relation.

Let us denote with T a set of tuples and with R a preference relation, defined over this set, $R \subseteq T \times T$. If two tuples $s, t \in T$ are in relation R, i.e., $R(s,t)$, then we will say that tuple s *dominates* tuple t with respect to relation R.

Then the *winnow* operator ω_R is defined as

$$\omega_R(T) = \{t \in T : \neg\exists_{s\in T}\ R(s,t)\} \tag{2}$$

and returns a subset of the non-dominated tuples with respect to R.

Let us briefly illustrate this concept on a set of tuples T, representing real estate properties described with, among other, the attributes `city` and `price`. Suppose that we look for the *cheapest* houses in each city. Then the following preference relation will be of use

$$R(s,t) \Leftrightarrow (s.\mathtt{city} = t.\mathtt{city}) \wedge (s.\mathtt{price} < t.\mathtt{price}) \tag{3}$$

where $t.A$ denotes the value of attribute A (e.g., `price`) at tuple t. Then, the *winnow* operator $\omega_R(T)$ will select the houses sought, since we will get as an answer a set of houses, which are non-dominated with respect to R, i.e., for which there is no other house in the same city which has a lower price.

The winnow operator may be adopted for the fuzzy setting (a fuzzy set of tuples T and a fuzzy preference relation R) in several ways.

In Zadrożny and Kacprzyk (2006) we proposed the following fuzzy counterpart of the original *winnow* operator

$$\mu_{\omega_R(T)}(t) = \text{truth}(T(t) \wedge \forall_s \; (T(s) \rightarrow \neg R(s,t))) \tag{4}$$

where $\mu_{\omega_R(T)}(t)$ denotes the value of the membership degree of tuple t to the fuzzy set of tuples, defined by $\omega_R(T)$, while $T(\cdot)$ and $R(\cdot,\cdot)$ denote the membership functions of the respective fuzzy set and fuzzy relation.

The bipolar query (C, P) with the "and possibly" operator may be expressed using the fuzzy *winnow* operator as follows (Chomicki, 2002; Zadrożny and Kacprzyk, 2006). Let R be a fuzzy preference relation defined as

$$R(s,t) \Leftrightarrow P(s) \wedge \neg P(t) \tag{5}$$

where, as previously, $R(\cdot,\cdot)$ and $P(\cdot)$ denote the membership functions of the corresponding fuzzy relation and fuzzy set.

Then the bipolar query with the "and possibly" operator may be expressed as the combination of the selection and the fuzzy *winnow* operators $\omega_R(\sigma_C(T))$, i.e.,

$$\mu_{\omega_R(\sigma_C(T))}(t) = \text{truth}(C(t) \wedge \forall_s \; (C(s) \rightarrow (\neg P(s) \vee P(t)))) \tag{6}$$

where $\sigma_C(T)$ is a usual "fuzzy" extension of the standard relational algebra selection operator, i.e., $\mu_{\sigma_C(T)}(t) = C(t)$.

The crucial issue of bipolar query semantics assumed in this work can be again summarized as follows. The unipolar bivariate scale is assumed, and a special interpretation, in which the negative and positive assessments are considered to correspond to the *required* and *desired* conditions, i.e. the negative assessment is identified with the degree, to which the required condition is *not satisfied* as, e.g., if a house sought has to be cheap (the required condition), then its negative assessment corresponds to the degree to which it is not cheap. The desired condition directly corresponds to the positive assessment.

In the approach of Lacroix and Lavency (1987), the crisp (nonfuzzy) conditions C and P are used. Then, a bipolar query (C, P) can be processed via the "first select using C then order using P" strategy, i.e., by finding tuples satisfying C and, second, choosing from among them those satisfying P, if any. This strategy may be easily generalised to the case in which C is crisp and P is fuzzy, then the second step consists in a non-increasing *ordering* of the tuples satisfying C according to their degree of satisfaction of P.

A fuzzification of the original Lacroix and Lavency's (1987) approach, when both C and P are fuzzy, was proposed by Zadrożny (2005), and Zadrożny and Kacprzyk (2006, 2012). We will present this approach in some detail, because it will provide us with a good conceptual point of departure. However, we will actually use in the paper a different representation of bipolar queries, using the above mentioned winnow operator (cf. Zadrożny and Kacprzyk, 2012). For some other approaches, see Bosc et al. (2010), Lietard and Rocacher (2009), Lietard, Rocacher and Bosc (2009), etc.

We consider the general form of the bipolar query (1), with C – the complement of the negative assessment (e.g., "price is cheap"), and P – the positive assessment (e.g., located "near a railroad station").

Then, the semantics of the bipolar query (1) may be formally characterized as follows:

- a tuple t belongs to the answer set of the query (C, P), given by (1), if it satisfies:

$$C(t) \text{ and possibly } P(t) \equiv C(t) \wedge \exists s(C(s) \wedge P(s)) \Rightarrow P(t) \quad (7)$$

  (in the original Lacroix and Lavency's, 1987, approach $P(t)$ and $C(t)$ are clearly the binary predicates),
- and if there are tuples satisfying both P and C, then (7) boils down to $C \wedge P$, while otherwise it boils down to C alone.

Basically, the fuzzification of the above mentioned concepts of a bipolar query, can be done in the following ways (cf. Zadrożny and Kacprzyk, 2012, for a comprehensive analysis):

- by a direct fuzzification of (7) (the predicates $P(t)$ and $C(t)$ are now fuzzy):

$$C(t) \text{ and possibly } P(t) \equiv C(t) \wedge \exists s \; (C(s) \wedge P(s)) \Rightarrow P(t); \quad (8)$$

- by a direct fuzzification of the winnow operator (cf. Chomicki, 2003) and applying it with a preference relation based on $P(\cdot)$, i.e., s is preferred to t if and only if $P(s)$ and $\neg P(t)$, combined with the selection operator referring to the condition C:

$$C(t) \text{ and possibly } P(t) \equiv C(t) \wedge \neg\exists s \; ((C(s) \wedge P(s) \wedge \neg P(t))); \quad (9)$$

- by using our fuzzy version of the winnow operator (cf. Zadrożny and Kacprzyk, 2012) and applying it with a preference relation based on $P(\cdot)$ as above:

$$C(t) \text{ and possibly } P(t) \equiv C(t) \wedge \forall s \; (C(s) \Rightarrow (\neg P(s) \vee P(t))); \quad (10)$$

and, clearly, these forms are equivalent in the classic Boolean logic; in this paper, (10) will be used.

The above logic formulas can then be precisiated in the sense that one can choose from among the available specific forms a specific form of the conjunction and disjunction, i.e. a t-norm and t-conorm (often called an s-norm), and the negation. They form the so-called De Morgan Triples $(\wedge, \vee, \neg)$ that comprise a t-norm operator $\wedge$, a t-conorm (s-norm) operator $\vee$ and a negation operator such that $\neg(x \vee y) = \neg x \wedge \neg y$, which then give rise to specific S-implication and R-implication operators modeling the implication in the formulas (8)-(10), cf. Zadrożny and Kacprzyk (2012).

3. Contextual bipolar queries

The concept of *context* is crucial for many areas, notably all those involving knowledge representation and discovery, intention modeling, preference representation, etc. In the area of bipolar queries the inclusion of context implied a new type of such queries, *contextual bipolar query*, proposed by Zadrożny. Kacprzyk and Dziedzic (2014, 2015); cf. also Kacprzyk and Zadrożny (2018), as well as Dziedzic, Zadrożny and Kacprzyk (2012).

As already mentioned, a bipolar query is meant to follow the required/desired semantics here and, interpreted as (1), the query "C and possibly P", given by (7), is satisfied by a tuple t only if either of two conditions holds:

1. it satisfies (of course, possibly to a high degree) both conditions C and P, or
2. it satisfies C and there is no tuple in the *whole database* which satisfies both conditions.

In a practical setting, it may be worthwhile to understand the "and possibly" in (1) in such a way that the satisfaction of both conditions C and P is meant in a certain *context*. For instance, while looking for inexpensive *and possibly* comfortable hotels in several cities we can have a serious problem in case of some very expensive regions or cities.

Namely, if in just one city, say Cheapcomfytown, there exists a cheap and comfortable hotel, then the regular bipolar query "cheap and possibly comfortable hotels" will return only the hotel(s) from Cheapcomfytown and no hotels, whatever cheap they are, from all other cities will be returned, since in order to be included to the answer set they are required to be both cheap and comfortable. The inclusion of a *context* can help to better account for the possibility of satisfying both conditions regarding a suitable subset of tuples, belonging to the *context* of a given tuple. A contextual bipolar query may be exemplified by:

Find *cheap* and possibly – *with respect to the hotels located in the same city* – *comfortable* hotels (11) to be meant as to be satisfied by a hotel if:

1. it is cheap (to a high degree) and is comfortable (to a high degree), or
2. it is cheap (to a high degree) and there is no other hotel *located in the same city* which is both cheap and comfortable. (12)

The new "and possibly + context" operator may be formalized as follows. The context of a tuple t is identified with a part of the database comprising tuples s somehow related to t, as defined by an additional binary predicate W, i.e.,

$$Context(t) = \{s \in T : W(t, s)\}, \tag{13}$$

where T denotes the whole database (relation).

The "and possibly + context" operator has three arguments:

$$C \text{ and possibly } P \text{ with respect to } W \tag{14}$$

where the predicates C and P should be interpreted, as previously, as representing the required and desired conditions, respectively, while the predicate W denotes the context.

Then, the formula (14) is interpreted as:

$$C(t) \;\; \textit{and possibly} \;\; P(t) \;\; \textit{with respect to} \;\; W \equiv \\ C(t) \wedge (\exists s(W(t, s) \wedge C(s) \wedge P(s)) \Rightarrow P(t)). \tag{15}$$

As we have already mentioned, the above traditional representation of the contextual bipolar query is shown for illustration, and in this paper we use a winnow operator based representation of the bipolar queries (10) so that (15) becomes

$$C(t) \;\; \textit{and possibly} \;\; P(t) \;\; \textit{with respect to} \;\; W \equiv \\ C(t) \wedge \forall s \; ((W(t, s) \wedge C(s)) \Rightarrow (\neg P(s) \vee P(t))). \tag{16}$$

In our example (11), C and P represent the properties of "cheap" and "comfortable", respectively, while W denotes the relation of being "located in the same city", i.e., $W(t, s)$ is true if both tuples represent hotels located in the same city; W can also be fuzzy.

The use of the winnow operator in modelling of the contextual bipolar queries, as shown by (16), provides for some further interesting interpretations and generalizations. First of all, notice that the right hand side of the implication symbol in (16) is the negation of the formula, representing dominance of tuple s over tuple t, as defined by (5). Thus, we can replace it in (16) with some other, more general, preference relation $R(\cdot,\cdot)$, giving rise to such a more general form of the contextual bipolar query (this generalization applies, in fact, also to bipolar query as defined by (10))

$$C(t) \;\; \textit{and possibly} \;\; P(t) \;\; \textit{with respect to} \;\; W \equiv \\ C(t) \wedge \forall s \; ((W(t,s) \wedge C(s)) \Rightarrow \neg R(s,t)) \quad (17)$$

where $P(t)$ refers to the notion of non-dominance with respect to the relation R(.,.), exemplified by

> Find hotels which are *cheap* and possibly – *with respect to the hotels located in the same city* – *not much worse, simultaneously, in terms of comfort and attractiveness of the neighborhood* (18)

where "much worse" (denoted below as $MuchWorse$) is a fuzzy relation, which is used here to define a preference relation R such that

$$R(s,t) = MuchWorse(t.comf, s.comf) \wedge MuchWorse(t.neigh, s.neigh) \quad (19)$$

where `comf` and `neigh` correspond to the attributes representing comfort and neighborhood attractiveness, respectively. In this way the user may express his or her preferences in a more faithful way. The interpretation of such an extended query in terms of the bipolarity is more subtle: hotels which are not cheap are still rejected and among those which are cheap the ones non-dominated with respect to the comfort and neighborhood attractiveness are preferred.

The second opportunity for interesting interpretations, provided by the *winnow* based model of the contextual bipolar queries, corresponds to the concept of context. The definition of context may be a part of the preference relation – cf. formula (3), where preference for cheaper houses is restricted to the houses located in the same city. Moreover, a clear separation of particular predicates – concerning the required condition C, the context W and the preference relation R – may be obtained by observing the following logical equivalence (the Law of Importation):

$$((\phi \wedge \psi) \Rightarrow \rho) \equiv (\phi \Rightarrow (\psi \Rightarrow \rho)), \quad (20)$$

which is valid in classical logic, but may be also easily verified for the most popular choices of the fuzzy connectives, cf., e.g., Jayaram (2008). Thus, (17) may be presented in the following equivalent form

$$C(t) \;\; \textit{and possibly} \;\; P(t) \;\; \textit{with respect to} \;\; W \equiv \\ C(t) \wedge \forall s \; (C(s) \Rightarrow (W(t,s) \Rightarrow \neg R(s,t))). \quad (21)$$

Such a separation makes it possible to get a deeper insight into the interplay of particular components of the contextual bipolar queries.

It is worth noting that the relation modeling the context, i.e. predicate $W(t,s)$, defines basically a partition (in a broad sense) of the set of tuples, crisp or fuzzy. This can be formalized in various ways, for instance by specifying an equivalence, similarity, etc. relation, an ordering, or even a modal logic based interpretation, cf. Zadrożny, Kacprzyk and Dziedzic (2014, 2015).

4. Hierarchical contextual bipolar queries

Now, by following the argument for the *hierarchical bipolar queries*, proposed by Kacprzyk and Zadrożny (2013a) (see also a related concept of a compound bipolar query, cf. Kacprzyk and Zadrożny, 2013b, 2017, 2018, 2019), we will present how the degrees of truth for particular tuples are calculated for the *contextual bipolar queries* in a hierarchical context. We will follow, for clarity and intuitive appeal, a very illustrative real estate example. The concept of such a hierarchical contextual bipolar query was proposed by Kacprzyk and Zadrożny (2018), but we extend it here by using a novel approach through the winnow operator representation of contextual bipolar queries. Basically, a hierarchical query is meant here in the sense that a query posed by the human user may involve terms, which do not directly relate to attributes as they concern some more complex terms, exemplified by *financially advantageous* for a house. This is then to be decoded by decomposing it into subterms that either explicitly relate to some attribute values in the database, for instance a *low price* and a *low interest rate of a bank loan*, or require further decoding in the same vein. The original query is therefore meant to be of level 0 at the bottom of the precisiation hierarchy, its required and optional parts are assumed to be bipolar queries themselves, both accounting for context. Then, the precisiation proceeds further, to level 1 queries, level 2, etc.

To start with, suppose that a customer of a real estate agency looks for houses that are "*financially advantageous* and possibly *well located*". However, he/she is interested in houses in different parts of the town and is fully aware that in some of these parts the satisfaction of both conditions is perfectly possible, because these parts are not expensive, but the locations are good. However, in some other parts of the town this may not be the case, so that a contextual bipolar query should be more appropriate as a tool to express the real intention and preference of the customer. His/her initial query, extended with context, can be:

> find a house that is *financially advantageous* (C_0) *and possibly* – *with respect to other houses in the same part of town* (W) – is *well located* (P_0) (22)

to be meant that a house satisfies this query if:

1. it is financially advantageous (C_0) and is well located (to a high degree), (P_0) or

2. it is financially advantageous (to a high degree) (C_0) and there is no other house *located in the same part of town*, which is both financially advantageous and well located,

and notice that the lower index 0, associated with the predicates C and P, indicates that these correspond to the initial ("time zero") level of the query, and we also assume – for the clarity of presentation – that the context W is the same for all levels. Clearly, one can also assume different contexts for different levels, which does not change the essence of the approach.

Then, using the interpretation of bipolar queries due to (16), i.e. via the winnow operator, the query (22) is interpreted as:

$$C_0(t) \textit{ and possibly } P_0(t) \text{ with respect to } W \equiv \\ C_0(t) \wedge \forall s((W(t,s) \wedge C_0(s)) \Rightarrow ((\neg P_0(s)) \vee P_0(t))), \quad (23)$$

where t and s, as previously, denote tuples (which here represent houses).

Assuming that the terms in the above initial query are not related directly to the attributes in the real estate database, corresponding to the features of the real estate properties, implies that the real estate agent has to "decode" the intentions/preferences of the customer, expressed through the required condition involving "financially advantageous", and the desired condition involving "well located", in context W. The very essence of "financially advantageous" and "well located" is clearly subjective, but suppose that these conditions may be "decoded" as (notice that the upper indexes of the C and P conditions will correspond to the level of concept hierarchy, i.e. generality, at which a concept in question is considered):

- the extent of the predicate C_0, representing the required condition "a financially advantageous house", may be equated with the answer set of the following query:

  find a house that is *inexpensive* (${C_1}^{C_0}$) *and possibly – with respect to other houses located in the same part of the town* (W) – in a *modern building* (${P_1}^{C_0}$) (24)

  which, similarly as for (23), yields

$${C_1}^{C_0}(t) \textit{ and possibly } {P_1}^{C_0}(t) \textit{ with respect to } W \equiv \\ {C_1}^{C_0}(t) \wedge \forall s((W(t,s) \wedge {C_1}^{C_0}(s)) \Rightarrow \\ (\neg {P_1}^{C_0}(s) \vee {P_1}^{C_0}(t))); \quad (25)$$

- the extent of the predicate P_0, representing the desired condition "well located", may be equated with the answer set of the following query:

  find a house that is *in an affluent part of the town* (${C_1}^{P_0}$) *and possibly – with respect to other houses located in the same part of the town* (W) – is *close to a recreational area* (${P_1}^{P_0}$) (26)

which, similarly as for (23), yields

$$C_1{}^{P_0}(t) \textit{ and possibly } P_1{}^{P_0}(t) \textit{ with respect to } W \equiv$$
$$C_1{}^{P_0}(t) \wedge \forall s((W(t,s) \wedge C_1{}^{P_0}(s)) \Rightarrow$$
$$(\neg P_1{}^{P_0}(s) \vee P_1{}^{P_0}(t))) \quad (27)$$

and these can be viewed as the first level query formulations (intentions/preferences).

Assuming that these conditions still do not refer to the database real attributes, i.e., among the database attributes there are no direct counterparts of the concepts of *inexpensive*, *modern building*, *affluent part of the town* nor *close to recreational area*, this calls for a further *decoding*:

- for the second level formulation of the first level required condition $C_1{}^{C_0}$, i.e. *inexpensive*, the following decoding may be possible:

  find a house that has a *low price* $(C_2{}^{C_1{}^{C_0}})$ *and possibly – with respect to other houses located in the same part of the town* (W) – has a *good bank loan offer* $(P_2{}^{C_1{}^{C_0}})$ (28)

  which, similarly as for (25), yields

$$C_2{}^{C_1{}^{C_0}}(t) \textit{ and possibly } P_2{}^{C_1{}^{C_0}}(t) \text{ with respect to } W \equiv$$
$$C_2{}^{C_1{}^{C_0}}(t) \wedge \forall s((W(t,s) \wedge C_2{}^{C_1{}^{C_0}}(s))$$
$$\Rightarrow (\neg P_2{}^{C_1{}^{C_0}}(s) \vee P_2{}^{C_1{}^{C_0}}(t))); \quad (29)$$

- for the second level formulation of the first level desired condition $P_1{}^{C_0}$, i.e. “modern building”, the following decoding may be possible:

  find a house that has an *intelligent energy management* $(C_2{}^{P_1{}^{C_0}})$ *and possibly – with respect to other houses located in the same part of the town* (W) – has *fast elevators* $(P_2{}^{P_1{}^{C_0}})$ (30)

  which, similarly as for (27), yields

$$C_2{}^{P_1{}^{C_0}}(t) \textit{ and possibly } P_2{}^{P_1{}^{C_0}}(t) \text{ with respect to } W \equiv$$
$$C_2{}^{P_1{}^{C_0}}(t) \wedge \forall s((W(t,s) \wedge C_2{}^{P_1{}^{C_0}}(s)) \Rightarrow$$
$$(\neg P_2{}^{P_1{}^{C_0}}(s) \vee P_2{}^{P_1{}^{C_0}}(t))); \quad (31)$$

- for the second level formulation of the first level required condition $C_1{}^{P_0}$, i.e. “an affluent part of town”:

  find a house that is in a *quiet zone* $(C_2{}^{C_1{}^{P_0}})$ *and possibly – with respect to other houses located in the same part of town* (W) – is *close to the business district* $(P_2{}^{C_1{}^{P_0}})$ (32)

which, similarly as for (31), yields

$$\begin{aligned}&{C_2^{C_1}}^{P_0}(t) \textit{ and possibly } {P_2^{C_1}}^{P_0}(t) \text{ with respect to } W \equiv\\&{C_2^{C_1}}^{P_0}(t) \wedge \forall s((W(t,s) \wedge {C_2^{C_1}}^{P_0}(s)) \Rightarrow\\&\qquad\qquad(\neg {P_2^{C_1}}^{P_0}(s) \vee {P_2^{C_1}}^{P_0}(t)));\end{aligned} \tag{33}$$

- for the second level formulation of the first level desired condition ${P_1}^{P_0}$, i.e. "close to a recreational area":

find a house that is *close to a park* (${C_2^{P_1}}^{P_0}$) *and possibly* – *with respect to other houses located in the same part of the town* (W) – is *close to a lake* (${P_2^{P_1}}^{P_0}$) (34)

which, similarly as for (33), yields

$$\begin{aligned}&{C_2^{P_1}}^{P_0}(t) \textit{ and possibly } {P_2^{P_1}}^{P_0}(t) \text{ with respect to } W \equiv\\&{C_2^{P_1}}^{P_0}(t) \wedge \forall s((W(t,s) \wedge {C_2^{P_1}}^{P_0}(s)) \Rightarrow\\&\qquad\qquad(\neg {P_2^{P_1}}^{P_0}(s) \vee {P_2^{P_1}}^{P_0}(t)))\end{aligned} \tag{35}$$

and, if necessary, one can continue until all the conditions involve attributes present in the database.

Notice that the context condition in the above queries (i.e. "with respect to other houses located in the same part of the town") is assumed the same for all queries and conditions but, in general, we can also employ local contexts, if needed.

5. Concluding remarks

We have further extended our two novel classes of bipolar database queries, the contextual bipolar queries and hierarchical bipolar queries, which has yielded a new class of hierarchical contextual bipolar queries. In this work, we have employed as the representation of bipolar queries our extension of the winnow operator based representation as a viable alternative to our logic based approach. The use of the winnow operator provides a comparable effectiveness and efficiency as the above mentioned logic based approach, but can be viewed to be preferable by people from the database community who are more interested or active in a broader area of preference queries. The discussion was illustrated on a real estate example, in which the customer has started with very general linguistic terms in the query, which are not directly related to the database attributes, and then has step by step precisiated the semantics of the terms used to finally arrive at a form, in which the real database attributes may be used.

Acknowledgments

Partially supported by National Science Centre under grant 2020/39/I/ST7/02285 "Parametric optimization of models and systems with time delays using metaheuristics".

References

Benferhat, S., Dubois, D., Kaci, S. and Prade, H. (2008) Modeling positive and negative information in possibility theory. *International Journal of Intelligent Systems* 23, 1094–1118.

Bordogna, G. and Pasi, G. (1995) Linguistic aggregation operators of selection criteria in fuzzy information retrieval. *International Journal of Intelligent Systems* **10**(2) 233–248.

Bosc, P. and Pivert, O. (1995) SQLf: A relational database language for fuzzy querying. *IEEE Transactions on Fuzzy Systems* **3**(1) 1–17.

Bosc, P. and Pivert, O. (2012) On four noncommutative fuzzy connectives and their axiomatization. *Fuzzy Sets and Systems*, 202, 42–60.

Bosc, P., Pivert, O., Mokhtari, A. and Lietard, L. (2010) Extending relational algebra to handle bipolarity. In: S.Y. Shin, S. Ossowski, M. Schumacher, M.J. Palakal, Ch.-Ch. Hung (eds.): *Proceedings of the 2010 ACM Symposium on Applied Computing (SAC)*, Sierre, Switzerland, ACM, 2010, 1718–1722.

Chomicki, J. (2002) Querying with intrinsic preferences. *Lecture Notes in Computer Science* **2287**, 34–51.

Chomicki, J. (2003) Preference formulas in relational queries. *ACM Transactions on Database Systems*, **28**(4), 427–466.

De Tré, G., Zadrożny, S. and Bronselaer, A. (2010) Handling bipolarity in elementary queries to possibilistic databases. *IEEE Transactions on Fuzzy Sets* **18**(3) 599–612.

De Tré, G., Zadrożny, S., Matthe, T., Kacprzyk, J. and Bronselaer, A. (2009) Dealing with positive and negative query criteria in fuzzy database querying. *Lecture Notes in Computer Science* (5822), 593–604.

Dubois, D. and Prade, H. (1988) Default reasoning and possibility theory. *Artificial Intelligence* **25**(2) 243–257.

Dubois, D. and Prade, H. (2002) Bipolarity in flexible querying. In: T. Andreasen, A. Motro, H. Christiansen, H. Larsen (Eds.), *Flexible Query Answering Systems, FQAS 2002, Lecture Notes in Computer Science*, **2522**. Springer, Berlin, Heidelberg, 174–182.

Dubois, D. and Prade, H. (2008a) Handling bipolar queries in fuzzy information processing. In: J. Galindo, ed., *Handbook of Research on Fuzzy Information Processing in Databases.* Information Science Reference, New York, USA, 97–114.

Dubois, D. and Prade, H. (2008b) An introduction to bipolar representations of information and preference. *Int. J. Intell. Syst.* **23**, 866–877.

Dubois, D. and Prade, H. (2009) An overview of the asymmetric bipolar representation of positive and negative information in possibility theory. *Fuzzy Sets and Systems* **160** (10), 1355–1366.

Dubois, D. and Prade, H. (2014) Modeling "and if possible" and "or at least": Different forms of bipolarity in flexible querying. In: O. Pivert, S. Zadrożny, eds., *Flexible Approaches to Data Management, a volume dedicated to Patrick Bosc*, Springer, 3–19.

Dujmović, J. (1979) Partial absorption function. *Journal of the University of Belgrade, EE Dept.* **659**, 156–163.

Dziedzic, M., Zadrożny, S. and Kacprzyk, J. (2012) Towards bipolar linguistic summaries: a novel fuzzy bipolar querying based approach. *FUZZ-IEEE 2012, International Conference on Fuzzy Systems*, 1–8.

Dziedzic, M., Kacprzyk, J. and Zadrożny, S. (2014) Contextual bipolarity and its quality criteria in bipolar linguistic summaries. *Technical Transactions: Automatic Control* 4-AC, 117–127.

Fodor, J. and Roubens, M. (1994) *Fuzzy Preference Modelling and Multicriteria Decision Support. Series D: System Theory, Knowledge Engineering and Problem Solving*, Kluwer Academic Publishers.

Grabisch, M., Greco, S. and Pirlot, M. (2008) Bipolar and bivariate models in multicriteria decision analysis: Descriptive and constructive approaches. *Int. J. Intell. Syst.* 23, 930–969.

Hadjali, A., Kaci, S. and Prade, H. (2011) Database preference queries – A possibilistic logic approach with symbolic priorities. *Annals of Mathematics and Artificial Intelligence*, **63** (3-4), 357–383.

Jayaram, B. (2008) On the law of importation $(x \wedge y) \longrightarrow z \equiv (x \longrightarrow (y \longrightarrow z))$ in fuzzy logic. *IEEE Transactions on Fuzzy Systems*, **16** (1), 130–144.

Kacprzyk, J. and Zadrożny, S. (1999) The paradigm of computing with words in intelligent database querying. In: L. A. Zadeh and J. Kacprzyk (eds.), *Computing with Words in Information/Intelligent Systems 2. Applications.* Physica, Heidelberg, 383–398.

Kacprzyk, J. and Zadrożny, S. (2013a) Hierarchical bipolar fuzzy queries: Towards more human consistent flexible queries. *FUZZ-IEEE 2013, International Conference on Fuzzy Systems*, 1–8.

Kacprzyk, J. and Zadrożny, S. (2013b) Compound bipolar queries: combining bipolar queries and queries with fuzzy linguistic quantifiers. In: *Proceedings of the 8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-13)*, University of Milano-Bicocca, Milan, Italy, September 11-13, Atlantis Press, 848–855.

Kacprzyk, J. and Zadrożny, S. (2016) Compound Bipolar Queries: A Step Towards an Enhanced Human Consistency and Human Friendliness. In: Stan Matwin and Jan Mielniczuk (Eds.): *Challenges in Computational Statistics and Data Mining.* Springer International Publishing, Cham, Heidelberg, New York, Dordrecht, 93–111.

Kacprzyk, J. and Zadrożny, S. (2017) Compound bipolar queries: The case of data with a variable quality. *Proc. FUZZ-IEEE 2017, IEEE*

International Conference on Fuzzy Systems, Naples, Italy, July 9-12, IEEE Press, 1–6.

Kacprzyk, J. and Zadrożny, S. (2018) Towards a Hierarchical Extension of Contextual Bipolar Queries. *Proc. IPMU 2018, Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*, Cadiz, Spain, June 11-15, Springer, 63–74.

Kacprzyk, J. and Zadrożny, S. (2019) Compound Bipolar Queries: The Case of Data with a Variable Quality. *Proc. FUZZ-IEEE 2019, IEEE International Conference on Fuzzy Systems*, New Orleans, USA, June 23–26, 1–6.

Kacprzyk, J., Zadrożny, S. and De Tré, G. (2015) Fuzziness in database management systems: half a century of developments and future prospects. *Fuzzy Sets and Systems* 281, 300–307.

Kacprzyk, J., Zadrożny, S. and Dziedzic, M. (2014) A novel view of bipolarity in linguistic data summaries. In: L. Koczy, C.R. Pozna and J. Kacprzyk (Eds.): *Issues and Challenges of Intelligent Systems and Computational Intelligence*. Springer Verlag, Berlin, Heidelberg, 215–229.

Kacprzyk, J. and Ziółkowski, A. (1986) Database queries with fuzzy linguistic quantifiers. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC 16**, 474–479.

Kacprzyk, J., Zadrożny, S. and Ziółkowski, A. (1989) FQUERY III+: a "human consistent" database querying system based on fuzzy logic with linguistic quantifiers. *Information Systems* **14**(6), 443–453.

Lacroix, M. and Lavency, P. (1987) Preferences: Putting more knowledge into queries. In: *Proc. of the 13 Int. Conf. on Very Large Databases*, 217–225.

Lietard, L. and Rocacher, D. (2009) On the Definition of Extended Norms and Co-norms to Aggregate Fuzzy Bipolar Conditions. In: J.P. Carvalho, D. Dubois, U. Kaymak and J.M. da Costa Sousa (eds.): *Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference* (Lisbon, Portugal, July 20-24, 2009), 513–518.

Lietard, L., Rocacher, D. and Bosc, P. (2009) On the Extension of SQL to Fuzzy Bipolar Conditions. *Proceedings of NAFIPS-2009 Conference*, 2009, 1–6.

Lietard, L., Tamani, N. and Rocacher, D. (2011) Fuzzy bipolar conditions of type "or else". In: *FUZZ-IEEE*, IEEE, 2546–2551.

Matthé, T., De Tré, G., Zadrożny, S., Kacprzyk, J. and Bronselaer, A. (2011) Bipolar database querying using bipolar satisfaction degrees. *Int. J. Intell. Syst.* **26**(10), 890–910.

Matthé, T., Nielandt, J., Zadrożny, S. and De Tré, G. (2014) Constraint-wish and satisfied-dissatisfied: an overview of two approaches for dealing with bipolar querying. *Flexible Approaches in Data, Information and Knowledge Management*, 21–44.

YAGER, R. (1992) Higher structures in multi-criteria decision making. *International Journal of Man-Machine Studies* **36**, 553–570.

YAGER, R. (1996) Fuzzy logic in the formulation of decision functions from linguistic specifications, *Kybernetes* **25**(4), 119–130.

ZADROŻNY, S. (2005) Bipolar queries revisited. In: V. Torra, Y. Narukawa, and S. Miyamoto (Eds.), *Modelling Decisions for Artificial Intelligence (MDAI 2005)*, LNAI, **3558**, Springer-Verlag, Berlin, Heidelberg, 387–398.

ZADROŻNY, S., DE TRÉ, G. AND KACPRZYK J. (2010) Remarks on various aspects of bipolarity in database querying. In: *Proc. of DEXA'10*, Bilbao, Spain, IEEE Computer Society, 323–327.

ZADROŻNY, S. AND KACPRZYK, J. (2006) Bipolar queries and queries with preferences. In: S. Bressan, J. Kung and R. Wagner, eds., *Database and Expert System Applications: 17th International Conference, DEXA 2006*, Kraków, Poland. Springer Verlag, 415–419.

ZADROŻNY, S. AND KACPRZYK, J. (2012) Bipolar queries: An aggregation operator focused perspective. *Fuzzy Sets and Systems* 196, 69–81.

ZADROŻNY, S., KACPRZYK, J. AND DE TRÉ, G. (2012) Bipolar queries in textual information retrieval: A new perspective. *Information Processing & Management* **48** (3), 390-398.

ZADROŻNY, S., KACPRZYK, J. AND DZIEDZIC, M. (2015) Contextual bipolar queries. In: *Proc. IFSA/EUSFLAT 2015*, Gijon, Spain, Atlantis Press, 1266–1273.

ZADROŻNY, S., KACPRZYK, J., DZIEDZIC, M. AND DE TRÉ, G. (2014) Contextual bipolar queries. In: Mo Jamshidi, Vladik Kreinovich, Janusz Kacprzyk (Eds.): *Advance Trends in Soft Computing.* Springer, 421–428.