

Vagueness, a multifacet concept - A case study on *Ambrosia artemisiifolia* predictive cartography

Patrick Maupin and Anne-Laure Jousselme

Defence R & D Canada - Valcartier

Decision Support Systems

Situation Analysis Support Systems

2459, Pie XI Blvd North

Val-Bélair, QC, G3J 1X5, CANADA

Email: {Patrick.Maupin, Anne-Laure.Jousselme}@drdc-rddc.gc.ca

Abstract—This paper proposes a reflexion on the concept of vagueness. We relate the different kinds of vagueness to relevant mathematical frameworks with particular emphasis on recent works on neutrosophic logic, Dezert-Smarandache theory, and rough sets. The different facets of vagueness are illustrated through a case study on *Ambrosia artemisiifolia* predictive cartography.

I. INTRODUCTION

The presence of *Ambrosia artemisiifolia* in urban areas is becoming a major problem of public health in the area of Montreal (Canada) where approximately 10% of the population exhibit moderate to severe allergy symptoms causing a heavy burden in health costs but also an important economic stress if productivity and absenteeism are considered. Typically, *Ambrosia artemisiifolia* grows in large colonies on recently or continuously disturbed land such as vacant lots, industrial areas, road and railway shoulders. But in many cases the direct detection is simply not possible since the plant might not have emerged yet from ground, waiting for, sometimes for more than forty years, appropriate germination conditions. Another source of difficulty is the fact that the precise ecological laws governing *Ambrosia artemisiifolia*'s niches are not very well known. The information available comes from annual census made by the city's authorities, expert judgments and testimonies, and ecological observations found in various monographs. To this date, the only maps available are derived from pollen counts obtained via a limited number of gravimetric and volumetric captors, as well as from partial surveys made by municipal inspectors. Such maps are too coarse and incomplete, thus of limited use for the implementation of public health and environmental policies. Remote sensing combined with typical ancillary data sources such as land use maps and population census seem to offer a very interesting mean for the cartography of contamination sources, being both synaptic in scope and spatially precise.

Answering the question "What areas are at risk of being invaded by *Ambrosia artemisiifolia* colonies?" requires dealing with uncertainty. Among the different types of uncertainty, vagueness is ubiquitous in geoscience. However, there are different kinds of vagueness: (1) Ontological, (2) linguistic and (3) epistemic. This taxonomy is detailed in section III.

From a formal point of view, vagueness produces borderline cases, that can be represented in different manners: Using a symbolic value for indetermination, using degrees that correspond for example to ill-determinate membership or truth, defining granulations or classes more or less coarse for the representation of indiscernible objects, and defining a richer universe of discourse to refine vague terms. Some mathematical theories use one or more of these representation schemes for vagueness. Among these, fuzzy sets theory and Łukasiewicz many-valued logics use only the notion of degrees (section IV-A), rough sets theory use equivalence relations to represent the indiscernibility among objects, but degrees can also be used to define rough membership functions (section IV-B), Dezert-Smarandache theory is based on the hyperpowerset structure allowing to build more precise concepts upon the vague concepts composing the basic universe of discourse (section IV-C).

The work presented in this paper besides illustrating the different representation schemes for vagueness, explores the composition of mathematical functions for a broader representation of vagueness. Indeed, because different kinds of vagueness can coexist (a blue object with ill-defined boundaries can belong to a given reference class with a certain degree), we should adopt an eclectic point of view or envisage a general framework for representing vagueness. Over a territory of roughly 500 square kilometers ground truth consisting in 3883 survey sites, expert judgments, multitemporal Landsat 7 data, a land use map, as well as information on the socio-economical status of the population will be used to illustrate the outlined concepts (see section II).

II. *Ambrosia artemisiifolia* PREDICTIVE CARTOGRAPHY

Among the principal objectives of managers coping with an invasive plant such as *Ambrosia artemisiifolia* are the spatial delimitation of the infected areas and the estimation of the severity and intensity of the phenomenon over the delimited territory. A precise cartography of the infected areas will allow to estimate the future extension of the plant, measure the effectiveness of control policies as well as identify ecological behavior of the plant. The information derived from this

cartography will to estimate the costs of control or eradication of the targeted plant.

The Island of Montreal is an urbanized area located within the St-Lawrence Lowlands ecoregion, in the St. Lawrence River valley, Canada. Mount Royal, the highest elevation of the island, is located roughly at $-73^{\circ}35'18''$ longitude West, $45^{\circ}30'24''$ latitude North. Warm summers (average of 16.5°C) and cold snowy winters (average of -7°C) are characteristic of the region. The mean annual precipitation in the ecozone ranges between 800-1000 mm. In 2001, the population of the island was approximately 1 800 000 within 27 administrative districts.

The use of ancillary data such as census and land-use maps offers numerous advantages such as low prices, availability and familiarity. Arguments for the use of the city block aggregation level are the facts that it is a convenient management unit and it lightens the computational burden. City blocks are in a first step obtained by the polygonization of street segments. The second step of block formation consists in cutting the previously obtained blocks using the railway network. This step is made necessary in order to avoid the formation of oversized blocks and to respect the natural boundaries between neighborhoods induced by the railways. The same process is applied using the Lachine Canal and Montreal's aqueduct boundaries.

Distance maps are obtained for each land-use class, railways, highways and contour of the Island of Montreal (water). This data is averaged at the block level. For each block the proportion of the different land-use classes is recorded, allowing the distinction between homogeneous and heterogeneous blocks but also given in a certain sense a richer topology to the land-use map. Various features are extracted from a Landsat 7 image taken on September 25th 2000 at 10 am. Among the most useful features are trapezoidal membership functions applied to extract information on ground cover. Simple classes such as bare soil, mineralized cover, nonvascular vegetation are defined and further combined in order to identify spaces typically infested by *Ambrosia artemisiifolia*. See figure 1.

III. VAGUENESS

Uncertainty is a broad concept, that should be avoided without explicit reference to specific contexts. In [1], we discuss and clarify the different facets of uncertainty. Klir and Yuan's taxonomy [2] distinguishes three different kinds of uncertainty-based information (further referred to as uncertainty) that are *Vagueness*, *Nonspecificity* and *Discord*. Hence, vagueness is a kind of uncertainty in the sense that this imperfection of information gives rise to uncertainty as far as the agent's mind is concerned. Black [3] rather distinguishes between *Vagueness*, *Generality* and *Ambiguity*. Also in [4], the authors mention the frequent confusion of vagueness with *Generality*, *Ambiguity* and *Open-texture*. According to Black [3], "The vagueness of a term is shown by producing "borderline cases", i. e. , individuals for which it seems impossible either to apply or not the term." This definition seems to be the more standard one. How are obtained these borderline

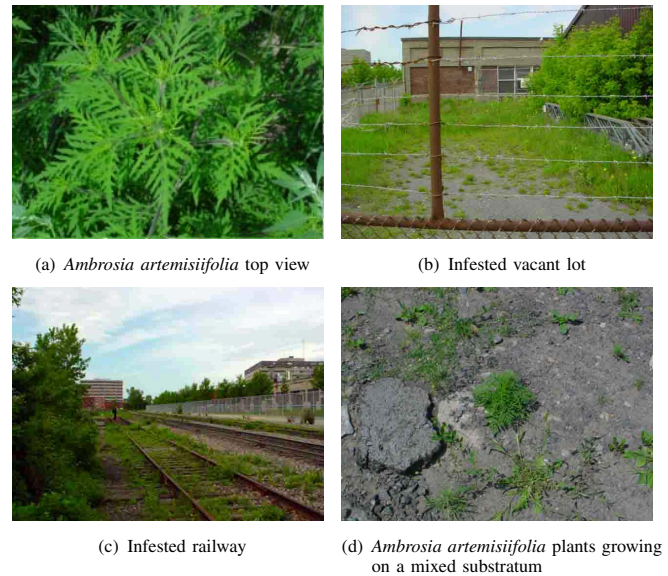


Fig. 1. *Ambrosia artemisiifolia* and its typical occupation of the territory on the Island of Montreal.

cases? The different possible answers lead to different kinds of vagueness.

For our purpose, we base the discussion on the taxonomy presented by figure 2. We consider three different kinds of

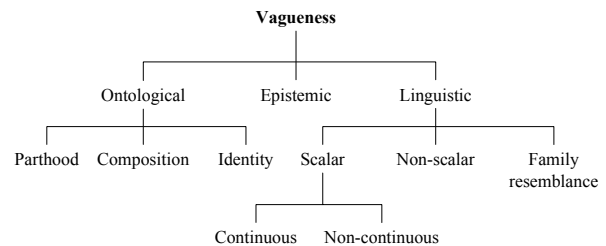


Fig. 2. Proposed taxonomy of vagueness.

vagueness: (1) **Ontological vagueness**, which is about the physical nature of objects. Although some contrary arguments exist, we assume that vague objects exist. For example, the limit of an object can be difficultly determined. In our application, that could correspond to city blocks or contaminated areas. This branch of the typology of vagueness can be enriched by the subclasses of Markosian's taxonomy, i. e. parthood, composition and identity [5]. (2) **Linguistic vagueness**, due to the limitations of natural language. The use of a vague predicate such as *big* for example, does not allow to clearly distinguish two objects. Among our data, such vagueness corresponds to statements such as "close to or on railways", "on vacant lots", "in open spaces". This branch of the typology of vagueness can be enriched by the subclasses described by Post and Berg [4]. (3) **Epistemic vagueness**, caused by the limitations of sensorial apparatus, lack of knowledge or computational limitations. Epistemic vagueness disappears if more information is brought into the situation [4].

IV. THEORETICAL FRAMEWORKS FOR DEALING WITH VAGUENESS

We discuss below three theories developed to deal with vagueness, that must be rather seen as complementary than competitors, since these theories address different kinds of vagueness. Formally, let Θ be a set of all possible outcomes of an experiment, the set of all hypotheses or objects of interest, of all considered situations, $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$.

A. Fuzzy sets theory

Fuzzy sets [6] are means to represent vague (fuzzy) concepts by allowing elements to belong more or less to a given set (*i. e.* class). A fuzzy subset A of Θ is then defined by a function μ_A that assigns to each θ of Θ a value $\mu_A(\theta)$ of the $[0; 1]$ real interval. $\mu_A(\theta)$ is the grade of membership of θ in A . Fuzzy sets theory and its logical counterpart fuzzy logic are adequate frameworks for vagueness, especially linguistic vagueness. Figure 3 shows an example of a membership function for the linguistic variable “close to railways”, which is built on a Euclidian distance (see also figure 1(c)). However, fuzzy sets

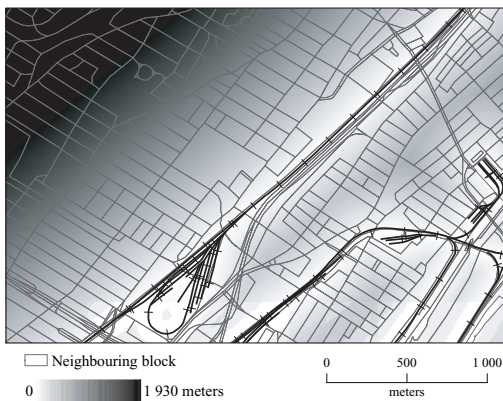


Fig. 3. Membership function of the linguistic variable “close to railways” built on a Euclidian distance [7].

theory seems also to suit well ontological vagueness where objects properties can be modeled by membership functions. For example, fuzzy boundaries at the interface of bare soil or mineralized surfaces and vegetation cover where *Ambrosia artemisiifolia* is very often found (figure 1(b)).

Often, the membership to a given class of objects is evaluated using a similarity criterion with respect to a reference object, group of objects or geometrical feature such as the barycenter of a class in a feature space. This reference feature from which the membership is calculated is called an *end-member* and is considered as a “pure” representative of the class of objects under study. In practice such end members are either reference signatures obtained in field experiments or in laboratory conditions, or as depicted by figure 3 using a mapped physical structure and building a linguistic variable on a proximity function. Some argue that this notion of endmember is not sufficient to describe objects of the real world, and that one must also consider features such as the

boundaries of the concepts. That is the line of view proposed by rough sets theory.

B. Rough sets theory

Rough set theory [8] appears as an alternative to fuzzy set theory for dealing with vagueness. Rough sets concern then a different kind of vagueness than fuzzy sets: indiscernibility (roughness) for the former and fuzziness for the latter. These two aspects of vagueness being often present together, other kinds of sets have been proposed, namely rough fuzzy sets and fuzzy rough sets [9]. On a finite frame of discernment Θ (a list of exclusive and exhaustive hypotheses), an equivalence relation I is defined, indicating which elements of Θ are indiscernible from each other. These elements form an equivalence class $[\theta]_I$, and come to defining a partition of indiscernible objects representing our limited knowledge. Vagueness is then represented by the fact that indiscernible elements both belong and not belong to a given set. A vague concept can then be represented by its lower and upper bounds being unions of elements of the partition. For any subset (concept) A of Θ , a lower approximation $I_*(A) = \{[\theta]_I | [\theta]_I \subseteq A\}$ and an upper approximation $I^*(A) = \{[\theta]_I | [\theta]_I \cap A \neq \emptyset\}$ are defined with respect to I (see figure 4). On figure 5, different levels of

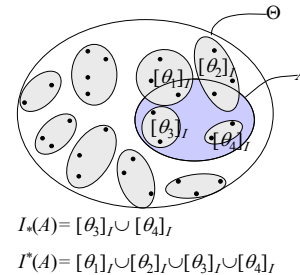
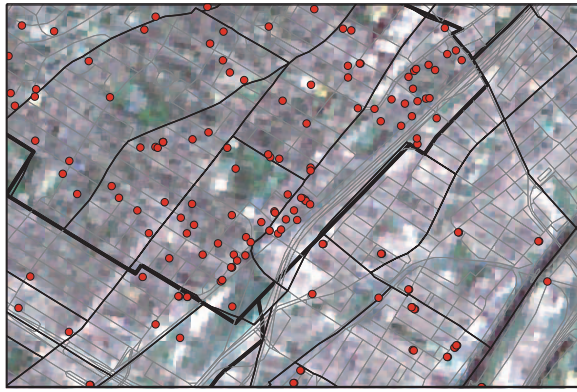


Fig. 4. Upper and lower approximations for rough sets.

roughness are represented: (1) the pixel level (resampled at 7.125 m for the purpose of the application), (2) the neighboring blocks level (3) the census tract level and (4) the district level. The different partitions involved are not necessary refinements of the upper levels. Indeed, a city block for example, is not necessary a union of pixels nor a census tract correspond exactly to a set of complete city blocks. Hence, to define a census tract or city blocks from pixels, the best we can do is to define upper and lower approximations. Similarly, infested areas can be defined using ground observations of *Ambrosia artemisiifolia* or estimated presence with a predictive model [7] for the valuation of the elements of the frame of discernment and blocks, census tracts or district limits using lower and upper approximations. Rough sets theory is thus an interesting framework for dealing with ontological (ex. infested areas) and epistemic vagueness (ex. pixels).

C. Dezert-Smarandache theory

Dezert-Samrandache theory (DSmT) [10] is an extension of the evidence theory [11]. Often interpreted as an extension of the Bayesian theory of probabilities, the theory of evidence



Census tract
 Neighbouring block • *Ambrosia artemisiifolia* observation
 District

Fig. 5. *Ambrosia artemisiifolia* observations and the different levels of spatial analysis [7].

offers the main advantage of better representing uncertainty because the measures are defined on the power set of the frame of discernment, instead of the universe itself as the probability theory does. Indeed, contrary to the probability measure, the belief measure of the DST is non-additive. The main consequence is that **every** subset of Θ is measurable (instead of only the singletons in probability theory). That is why it is often said that DST is built on the power set of Θ whereas probability theory is built on Θ itself. We note the power set, $2^\Theta = \{\emptyset, \theta_1, \dots, \theta_n, (\theta_1 \cup \theta_2), \dots, \Theta\}$. A belief function Bel is thus defined from 2^Θ to $[0, 1]$ to quantify the degree of belief one has in the occurrence of a given A , subset of Θ . Bel is often described using a *basic probability assignment* (or basic belief assignment) m from 2^Θ to $[0, 1]$ that must satisfy the following conditions: (1) $m(\emptyset) = 0$ and (2) $\sum_{A \in 2^\Theta} m(A) = 1$. Then we have $\text{Bel}(A) = \sum_{B \subseteq A, B \in 2^\Theta} m(B)$.

One of the main basic hypotheses of DST is that the frame of discernment Θ is supposed to be a list of mutually exclusive objects, an hypothesis also shared by rough sets theory. Consequently, vague concepts are not allowed to form Θ . DSMT is an extension of DST relaxing this hypothesis of mutually exclusivity, and thus proposes a wider framework that besides the same features of DST for dealing with uncertainty, allows the representation of vague concepts. DSMT is thus built on the hyperpower set $D^\Theta = \{\emptyset, \theta_1, \dots, \theta_n, (\theta_1 \cup \theta_2), \dots, \Theta, (\theta_1 \cap \theta_2), \dots, (\theta_1 \cap \theta_2) \cup \theta_3, \dots\}$. The hyperpower set is constituted of all the combinations of the union and intersection operators applied to the elements of Θ . This structure allows one to account for vague concepts as their intersection is considered. Then a *general basic belief mass* is defined from D^Θ to $[0, 1]$, satisfying the following conditions (1) $m(\emptyset) = 0$ and (2) $\sum_{A \in D^\Theta} m(A) = 1$. Hence, for example elements of the type of $\theta_i \cap \theta_j$, $i \neq j$ are allowed to be measured. The *general belief function* is then defined from the general basic belief assignment. DSMT is thus a more general framework that deals

with both ontological and epistemic uncertainty [12].

For our case study, this framework allows then to represent the three kinds of vagueness (ontological, epistemic and linguistic) previously defined since any vague concept can constitute the frame of discernment. Besides this, DSMT is a natural framework for the construction of mixed concepts. Consider for example $\Theta = \{\text{bare soil, mineralized surfaces, vegetation cover}\}$. DSMT permits to built other elements such as $(\text{bare soil} \cap \text{vegetation cover})$, which indeed corresponds to a privileged area for the installation of *Ambrosia artemisiifolia* colonies. The elements of the hyperpower set of Θ describe then a wide variety of mixed soils, to which belief values can further be assigned.

V. CONCLUSION

In this paper, we proposed a reflexion on the concept of vagueness. We described the different facets of vagueness through three main types of vagueness that are ontological vagueness, linguistic vagueness and epistemic vagueness, and illustrated this taxonomy on a case study on *Ambrosia artemisiifolia* predictive cartography. In remote sensing application, vagueness appears as a multifacet concept: Basically, vagueness is *epistemic* in nature (measures) but it is also *ontological* due to the nature of objects in presence (fuzzy boundaries of areas). Furthermore, vagueness is *linguistic* when human knowledge is a data source (domain knowledge and expert opinions). These different aspects of vagueness should be considered and modeled adequately by the different theories for dealing with vagueness. Prediction and classification results will be proposed in an upcoming paper.

REFERENCES

- [1] A.-L. Jousselme, P. Maupin, and E. Bossé, "Uncertainty in a situation analysis perspective," in *Proceedings of 6th Annual Conference on Information Fusion*, (Cairns, Australia), pp. 1207–1214, July 2003.
- [2] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice Hall International, 1995.
- [3] M. Black, "Vagueness: An exercise in logical analysis," *Philosophy of science*, vol. 4, pp. 427–455, 1937.
- [4] J. F. Post and J. Berg, "Vagueness," in *The Cambridge Dictionary of Philosophy* (R. Audi, ed.), pp. 945–947, Cambridge, UK: Cambridge University Press, 2 ed., 1999.
- [5] N. Markosian, "Brutal Composition," *Philosophical Studies*, vol. 92, pp. 211–249, 1998.
- [6] L. A. Zadeh, "Fuzzy Sets," *Information and control*, vol. 8, pp. 338–353, 1965.
- [7] P. Maupin and P. Apparicio, "Relationships between *Ambrosia artemisiifolia* sites and the physical and social environments of Montreal (Canada)," in *IGARSS*, (Anchorage, Alaska), 2004. To be published.
- [8] Z. Pawlak, *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publisher, 1991.
- [9] D. Dubois and H. Prade, "Rough Fuzzy Sets and Fuzzy Rough Sets," *International Journal of General Systems*, vol. 17, no. 2-3, pp. 191–209, 1990.
- [10] J. Dezert, "Foundations for a new theory of plausible and paradoxical reasoning," *Information & Security: An International Journal*, vol. 9, pp. 90–95, 2002.
- [11] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [12] A.-L. Jousselme and P. Maupin, "Neutrosophic frameworks for situation analysis," in *Advances and Applications of Dezert-Smarandache Theory (DSMT) of Plausible and Paradoxical Reasoning for Data Fusion* (F. Smarandache and J. Dezert, eds.), 2004.