Higher Order Voronoi Diagrams for Concept Boundaries and Tessellations

Ickjai Lee
School of Math, Physics & IT
James Cook University
Douglas, Townsville, QLD 4811, Australia
Ickjai.Lee@jcu.edu.au

Kyungmi Lee
School of Business & IT
Charles Stuart University
Albury, NSW, Australia
jolee@csu.edu.au

Abstract

Representing concepts and their boundaries is of great importance in concept learning and formation. We introduce a flexible framework for representing concept boundaries and tessellations through the higher order Voronoi diagrams. The framework provides different levels of concept granularity modelling vagueness and indeterminacy of concept boundaries. The proposed approach robustly supports concept learning and formation.

1. Introduction

Representing concepts is a core task in concept formation and learning. Symbolic representation and associationism (in particular connectionism) have been two popular ways of representing concepts in the literature. However, it has been pointed out that connectionism is too fine-grained and symbolic representation is too coarse, and some cognitive tasks require an intermediate representation approach [8]. A conceptual spaces approach is a midway representation to bridge a gap between too detailed raw data from sensory receptors and too coarse symbolic knowledge representations [8]. This conceptual spaces approach provides a robust framework for representing properties and concepts that are core elements for many cognitive tasks including concept formation, concept learning, concept management, inductive inferences, semantics, categorization and prediction.

A conceptual space is seen as a collection of one or more domains and each domain consists of a number of quality dimensions. Objects are identified with multidimensional points (also called knoxel [5]) within conceptual spaces, concepts are denoted by a collection of regions representing properties and their relations in conceptual spaces. Recently, the conceptual spaces based framework has been used in various applications including computer vision [5], cluster reasoning [14, 17], concept categorization [16], representing dynamic actions [4, 6] and conceptual similarities of verbs [12]. In addition, managing concepts and categories in data-rich environments has been studied [9, 15]. Details of conceptual spaces can be found in [8].

Concept learning and formation in a conceptual space is equivalent to a tessellation of the space. The two widely adopted approaches are nearest neighbor learning (well supported by the machine learning community [20]) and prototype learning (well supported by the cognitive science community [8]). The ordinary Voronoi diagram is a robust way of tessellating the space [23] and it has been used for modelling these learning approaches for concept learning. However, main limitations of these approaches include their inability of forming a new concept, rigidity and inflexibility.

This paper investigates approaches for representing concepts and their boundaries using the higher order Voronoi diagrams. We introduce a flexible method of representing different levels of concept boundaries for modelling the vagueness of concept boundaries. The proposed approach robustly supports both concept formation and concept learning. The remainder of this paper is organized as follows. Section 2 surveys traditional approaches for representing concept boundaries within conceptual spaces. Section 3 outlines our approach for modelling concept boundaries using the higher order Voronoi diagrams. Section 4 introduces a geospatial concept formation and learning with concept boundaries derived from the higher order Voronoi diagrams. Section 5 concludes with final remarks.

2. Tessellating Concepts

Objects are represented by multidimensional points in conceptual spaces and a concept is a formulated by a set of exemplars (previously categorized examples of objects in the concept). Thus, representing a concept is equivalent to representing a set of exemplars. Representing a set of multidimensional points is typically achieved by the following two methods in conceptual spaces: prototype *vs.* individual. The former follows the philosophy of learning

6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)
0-7695-2841-4/07 $25.00 © 2007 IEEE

IEEE COMPUTER SOCIETY

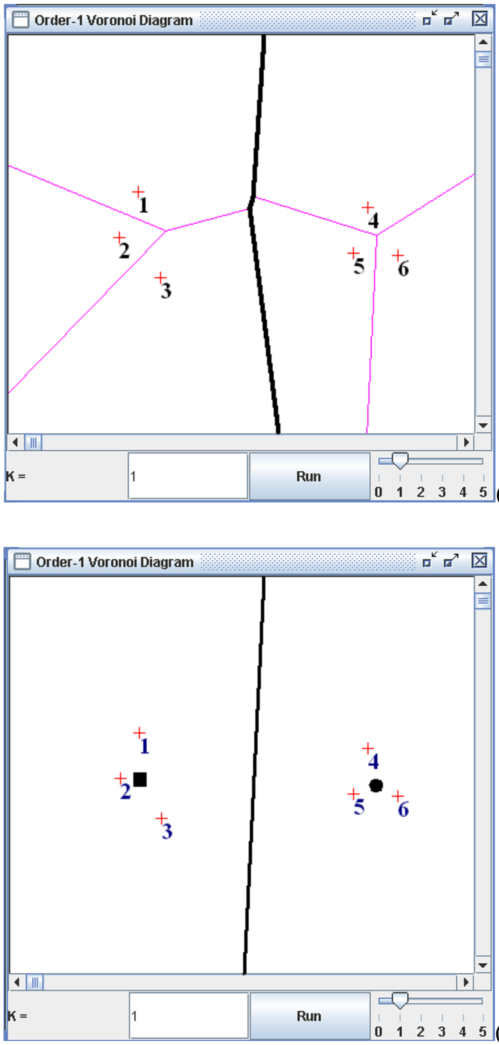

Figure 1. VOI and VOP tessellations of $P = \{p_1, p_2, p_3, p_4, p_5, p_6\}$: (a) VOI($P$); (b) VOP($P$).

economy and widely used in the cognitive science community [8]. Concepts have prototypes and members exhibit central tendency. Therefore, a cloud of points is represented by a representative (prototype). The prototype representation (modelled by the Voronoi diagram of prototypes denoted by VOP) has been the core concept of many learning mechanisms such as k-means [18] and k-medoid [10] clustering approaches. The individual representation has been widely used in the machine learning community [20]. In the instance-based representation (modelled by the Voronoi diagram of individuals denoted by VOI), each member (instance) equally contributes to learning. All members affect the membership of a new object and more similar (closer) members have more (stronger) impact on learning. In this learning, identifying geospatial neighbors (k-nearest neighbors) is an important task.

Figure 1 depicts VOI and VOP tessellations of 6 examples. Let use assume three examples $P_1 = \{p_1, p_2, p_3\}$ form a concept and the other three $P_2 = \{p_4, p_5, p_6\}$ form another concept. Concept boundaries that distinguish these two concepts are shown as thick black edges in this figure. In these approaches, concepts are mutually exclusive and collectively exhaustive. Concept boundaries can be used for concept learning.

3. Higher Order Voronoi Diagrams based Concept Boundaries

3.1. Higher Order Voronoi Diagrams

The ordinary Voronoi diagram tessellates the space in such a way that every location in the space is assigned to the closest generator. A set of locations closest to more than one generator formulates the ordinary Voronoi diagram consisting of Voronoi edges and vertices. Higher order Voronoi diagrams are natural and useful extensions of the ordinary Voronoi diagram for more than one generator [23]. They provide tessellations where each region has the same k (ordered or unordered) closest sites for a given k. The order-k Voronoi diagram $\mathcal{V}^{(k)}$ is a set of all order-k Voronoi regions $\mathcal{V}^{(k)} = \{V(P_1^{(k)}), \ldots, V(P_n^{(k)})\}$, where the order-$k$ Voronoi region $V(P_i^{(k)})$ for a random subset $P_i^{(k)}$ consisting of k points out of P is defined as follows:

$$V(P_i^{(k)}) = \{p | \arg\max_{p_r \in P_i^{(k)}} d(p, p_r) \leq \arg\min_{p_s \in P \setminus P_i^{(k)}} d(p, p_s)\}, \tag{1}$$

where, k generators are not ordered. The order-k Voronoi diagram $\mathcal{V}^{(k)}$ becomes the ordinary Voronoi diagram when k is 1.

When k generators are ordered, the diagram becomes the ordered order-k Voronoi diagram $\mathcal{V}^{<k>}$. It is defined as $\mathcal{V}^{<k>} = \{V(P_1^{<k>}), \ldots, V(P_n^{<k>})\}$, where the ordered order-k Voronoi region $V(P_i^{<k>})$ is defined as

$$V(P_i^{<k>}) = \{p | d(p, p_{i1}) \leq \ldots \leq d(p, p_{ik}) \leq d(p, p_j),$$
$$p_j \in P \setminus \{p_{i1}, \ldots, p_{ik}\}\}. \tag{2}$$

Figure 2 depicts the order-3 Voronoi diagram and ordered order-3 Voronoi diagram of P. A shaded region in Figure 2(a) depicts an order-3 Voronoi region of P_1 where all generators belong to the same concept. Thus, this region is indicative of the concept P_1. Its neighboring regions (that share a Voronoi edge with $V(P_1^{(3)})$) are Voronoi regions of $\{p_1, p_2, p_4\}$, $\{p_1, p_3, p_4\}$, $\{p_1, p_3, p_5\}$ and $\{p_2, p_3, p_5\}$. These neighboring regions have members that belong to

6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)
0-7695-2841-4/07 $25.00 © 2007 IEEE

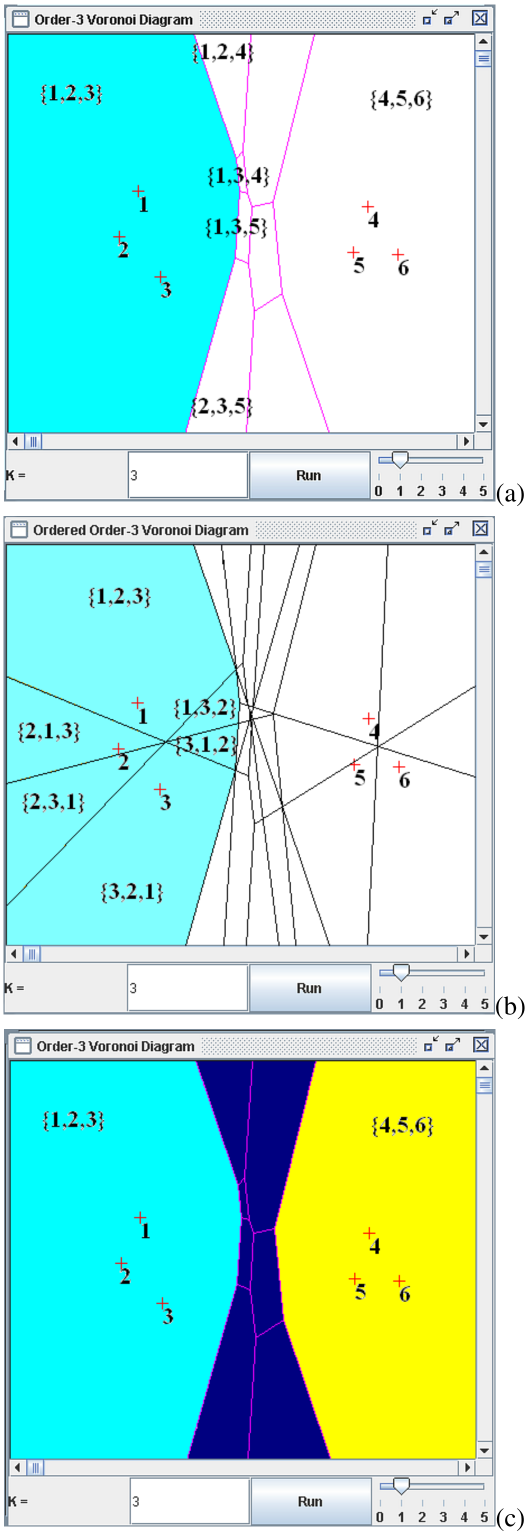

Figure 2. The order-3 Voronoi diagram and ordered order-3 Voronoi diagram of P (P is the same as in Figure 1): (a) $\mathcal{V}^{(3)}(P)$; (b) $\mathcal{V}^{<3>}(P)$; (c) Concept territories of P.

different concepts. For instance, the Voronoi region of $\{p_1, p_2, p_4\}$ has members p_1, p_2 belonging to the concept P_1 and a member p_4 belonging to the concept P_2. Thus, this region is an indicative of a neutral zone that can be used for learning a new concept. The shaded order-3 Voronoi region is further decomposed and modelled in a much finer form by the ordered order-3 Voronoi diagram as shown in Figure 2(b). This region is further decomposed into $3!$ number of regions in the ordered order-3 Voronoi diagram. Figure 2(c) depicts concept territories of P_1 and P_2 shaded in light grey and the neutral zone shaded in dark grey.

Several algorithmic approaches [1, 3, 7, 13] have been proposed to efficiently compute higher order diagrams in the computational geometry community. Dehne [7] proposed an $O(n^4)$ time algorithm that constructs the complete $\mathcal{V}^{(k)}$. Several other attempts [1, 3, 13] have been made to improve the computational time requirement of Dehne's algorithm. The best known algorithm for the order-k Voronoi diagram is $O(k(n-k)\log n + n\log^3 n)$ [2]. Thus, the complete $\mathcal{V}^{(k)}$ requires $O(k^2(n-k)\log n + kn\log^3 n)$ time. Once $\mathcal{V}^{(k)}$ is computed, then $\mathcal{V}^{<k>}$ can be derived by $\bigcup \mathcal{V}^{(k)}$.

3.2. Concept Boundaries and Tessellations

Since an order-k Voronoi region has k generators as the closest (most similar) k points among P, the order-k Voronoi region constitutes a concept when all k generators belong to the same concept. A k-th order concept tessellation can be obtained by merging order-k Voronoi regions those generators all belong to the same cluster. k-th order concept boundaries are a set of order-k Voronoi edges whose two incident order-k Voronoi regions have generators belonging to different concepts.

An order-k Voronoi region whose generators belonging to different concepts constitutes a neutral region that does not belong to any concept. A merge of all these regions forms the k-th order neutral zone that can be used for formulating a new concept. An order-k Voronoi region is either part of one of concept territories or the neutral zone. Thus, we can obtain a space tessellation that is collectively exhaustive and mutually exclusive.

4. A Case Study with Geospatial Concepts

Formulating and learning geospatial concepts are important issues in geospatial data mining and knowledge discovery [19]. Intelligent agents are autonomously building geospatial concepts that can be used for learning and reasoning. In this case study, real geospatial datasets from Brisbane, the capital city of Queensland, Australia are considered. We utilize raw crime and geospatial feature datasets recorded in the year of 1996 and 1998 by the

6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)
0-7695-2841-4/07 $25.00 © 2007 IEEE

Queensland Police Service around 217 central urban suburbs of Brisbane as our study region. Brisbane is continuously experiencing steady population and crime growth [21, 22]. Building crime related concepts in this region provides a valuable resource to city planners, policing agencies and criminologists.

Figure 3 depicts a dataset with the study region. The dataset consists of 4 different geospatial concepts: post office ($\{p_3, p_4, p_5, p_6, p_7\}$), murder ($\{p_8, p_9, p_{10}, p_{11}, p_{12}\}$), caravan park ($\{p_{13}, p_{14}, p_{15}, p_{16}, p_{17}\}$) and unknown crime type ($\{p_{18}, p_{19}, \ldots, p_{36}, p_{37}\}$).

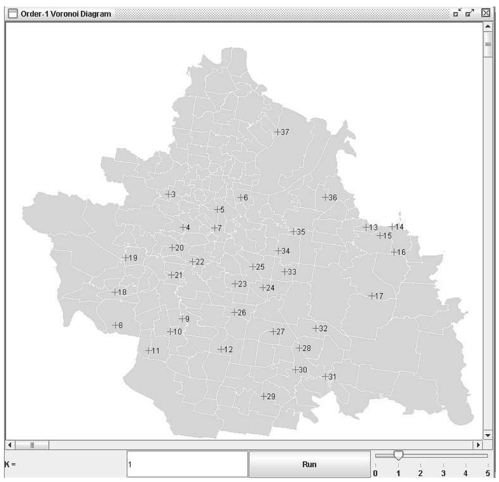

Figure 3. A study region with a dataset $P = \{p_3, p_4, \ldots, p_{36}, p_{37}\}$.

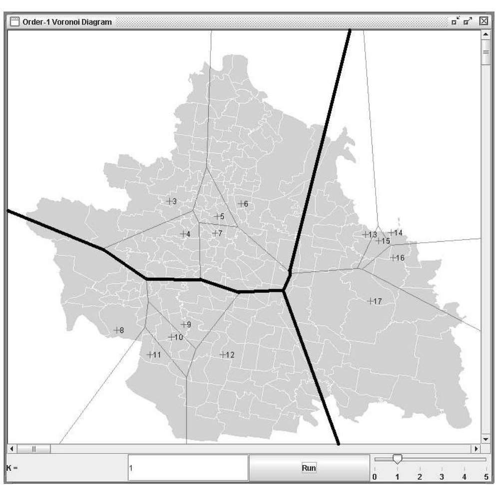

Figure 4. Concept boundaries defined by the ordinary Voronoi diagram.

Each concept is geospatially correlated and aggregated. Locations of post offices, murder incidents and caravan parks are recorded in the year of 1996, and locations of unknown crime type are recorded in the year of 1998. In this scenario, we build the higher order Voronoi diagrams inspired concept boundaries and use them for learning the unknown crime incidents. First, we build concept boundaries of the three known concepts: post office, murder and caravan park. These concept boundaries derived from the higher order Voronoi diagrams are shown from Figure 4 to Figure 8. As k becomes bigger the neutral zone widens.

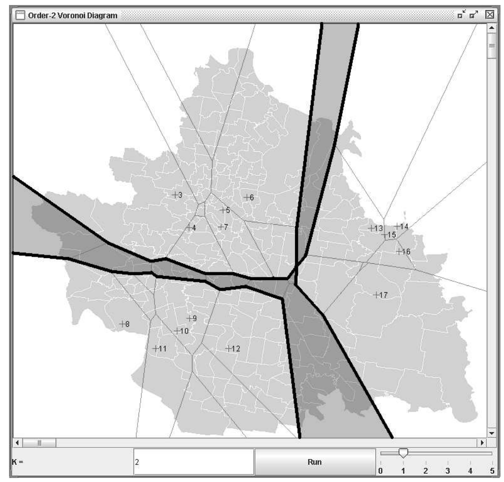

Figure 5. Concept boundaries defined by the 2-order Voronoi diagram.

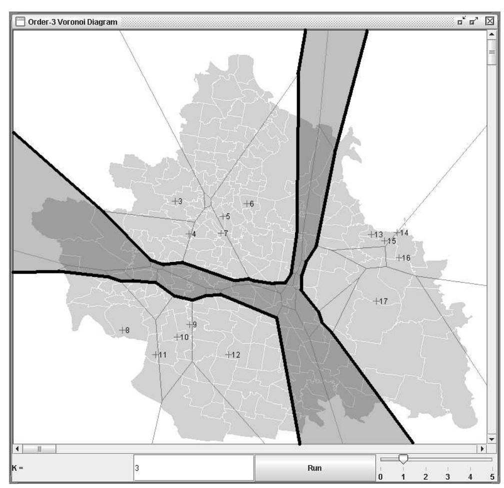

Figure 6. Concept boundaries defined by the 3-order Voronoi diagram.

6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)
0-7695-2841-4/07 $25.00 © 2007 IEEE

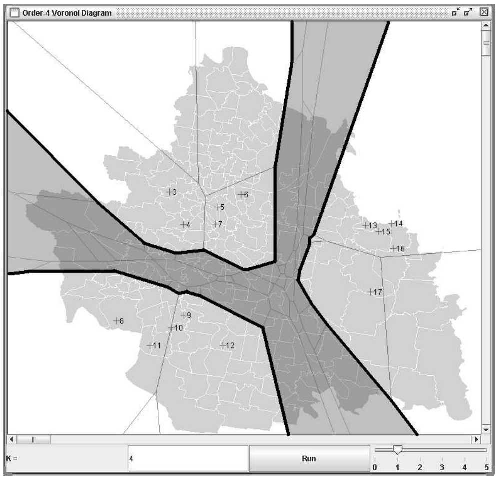

Figure 7. Concept boundaries defined by the 4-order Voronoi diagram.

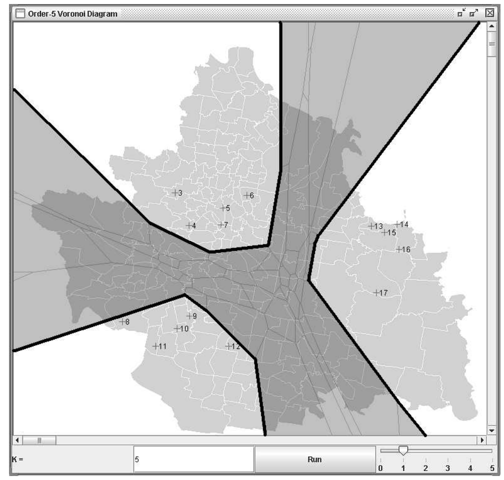

Figure 8. Concept boundaries defined by the 5-order Voronoi diagram.

A series of the higher order Voronoi diagrams defines different levels of concept boundaries that model indeterminacy and vagueness of concepts. Once concept boundaries and tessellations constructed, they can be used for concept learning and formation. Table 1 summarizes learning of 20 unknown objects with the complete higher order Voronoi diagrams. Undoubtedly, VOP (Figure 4) assigns all unknown objects to one of existing concepts, and fails to form a new concept. Concept formation is much improved with the use of the order-2 Voronoi diagram as shown in Figure 5. With the order-3, order-4 and order-5 Voronoi diagrams, more than half of unknown objects fall into the neutral zone that may trigger a new concept formation based on the majority voting [11]. This trend is well depicted in Figure 9. The number of unknown objects misclassified into one of existing concepts is gradually decreasing as k grows whilst the number of objects falling into the neutral zone is constantly increasing. This simple example demonstrates the flexibility and robustness of concept boundaries derived from the higher order Voronoi diagrams.

Table 1. Learning with the complete higher order Voronoi diagrams.

$\mathcal{V}^{(k)}$	$\mathcal{V}^{(1)}$	$\mathcal{V}^{(2)}$	$\mathcal{V}^{(3)}$	$\mathcal{V}^{(4)}$	$\mathcal{V}^{(5)}$
Post office	9	6	5	3	0
Murder	8	4	3	2	0
Caravan park	3	1	1	0	0
Neutral zone	0	9	10	15	20

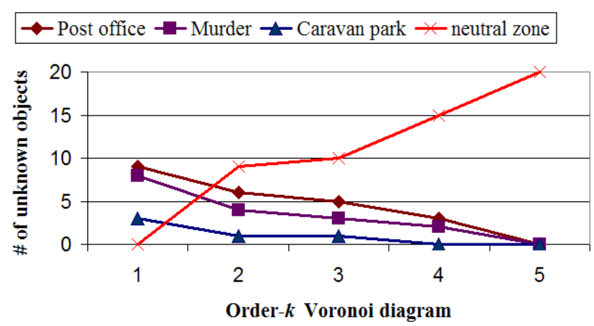

Figure 9. Concept learning with the complete higher order Voronoi diagram.

5. Conclusions

The Voronoi diagram provides a cognitively economical way of representing information about concepts. VOI and VOP have been two popular approaches for representing concepts in conceptual spaces. In this paper, we provide a flexible concept tessellation framework overcoming the rigidity and inflexibility of traditional approaches. It is based on the complete higher order Voronoi diagrams and robustly supports concept learning and formation. Experimental results demonstrate the usefulness and superiority of our approach.

References

[1] F. Aurenhammer and O. Schwarzkopf. A Simple On-Line Randomized Incremental Algorithm for Computing Higher Order Voronoi Diagrams. In *Proceedings of the Symposium on Computational Geometry*, pages 142–151, North Conway, NH, 1991. ACM Press.

[2] M. Berg, M. Kreveld, M. Overmars, and O. Schwarzkoph. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, West Sussex, second edition, 2002.

[3] B. Chazelle and H. Edelsbrunner. An Improved Algorithm for Constructing kth-Order Voronoi Diagrams. *IEEE Transactions on Computers*, 36(11):1349–1354, 1987.

[4] A. Chella, M. Frixione, and S. Gaglio. Towards a Conceptual Representation of Actions. In E. Lamma and P. Mello, editors, *Proceedings of the 5th Congress of the Italian Association for Artificial Intelligence*, Lecture Notes in Computer Science 1792, pages 333–344, Bologna, Italy, 1997. Springer.

[5] A. Chella, M. Frixione, and S. Gaglio. Conceptual Spaces for Computer Vision Representations. *Artificial Intelligence Review*, 16(2):137–152, 2001.

[6] A. Chella, M. Frixione, and S. Gaglio. Symbolic and Conceptual Representation of Dynamic Scenes: Interpreting Situation Calculus on Conceptual Spaces. In F. Esposito, editor, *Proceedings of the 7th Congress of the Italian Association for Artificial Intelligence*, Lecture Notes in Computer Science 2175, pages 333–343, Bari, Italy, 2001. Springer.

[7] F. K. H. Dehne. On $O(N^4)$ Algorithm to Construct all Voronoi Diagrams for K-Nearest Neighbor Searching in the Euclidean plane. In J. Díaz, editor, *Proceedings of the International Colloquium on Automata, Languages and Programming*, Lecture Notes in Artificial Intelligence 154, pages 160–172, Barcelona, Spain, 1983. Springer.

[8] P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. The MIT Press, Cambridge, 2000.

[9] P. Gärdenfors and M.-A. Williams. Reasoning about Categories in Conceptual Spaces. In B. Nebel, editor, *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, pages 385–392, Washington, D.C., 2001. Morgan Kaufmann.

[10] L. Kaufman and P. J. Rousseuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, New York, 1990.

[11] L. I. Kuncheva. *Combining Pattern Classifier: Methods and Algorithms*. Wiley, New Jersey, 2004.

[12] K. Lagus, A. Airola, and M. Creutz. Data Analysis of Conceptual Similarities of Finnish Verbs. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*, pages 566–571, Virginia, 2002.

[13] D.-T. Lee. On k-Nearest Neighbor Voronoi Diagrams in the Plane. *IEEE Transactions on Computers*, 31(6):478–487, 1982.

[14] I. Lee. Fast Qualitative Reasoning about Categories in Conceptual Spaces. In A. Abraham, M. Köppen, and K. Franke, editors, *Proceedings of the Third International Conference on Hybrid Intelligent Systems*, pages 341–350, Melbourne, Australia, 2003. IOS Press.

[15] I. Lee. Efficient Management of Conceptual Spaces Using Data Mining Techniques. In M. Mohammadian, editor, *Proceedings of the 2004 International Conference on Computational Intelligence for Modelling Control and Automation*, pages 492–499, Gold Coast, Australia, 2004.

[16] I. Lee. Hybrid Soft Categorization in Conceptual Spaces. In M. Ishikawa, S. Hashimoto, M. Paprzycki, E. Barakova, K. Yoshida, M. Köppen, D. W. Corne, and A. Abraham, editors, *Proceedings of the Fourth International Conference on Hybrid Intelligent Systems*, pages 74–79. IEEE Computer Society, 2004.

[17] I. Lee and M.-A. Williams. Multi-level Clustering and Reasoning about Its Clusters Using Region Connection Calculus. In K.-Y. Whang, J. Jeon, K. Shim, and J. Srivastava, editors, *Proceedings of the 7th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Lecture Notes in Artificial Intelligence 2637, pages 283–294, Seoul, Korea, 2003. Springer.

[18] J. MacQueen. Some Methods for Classification and Analysis of Multivariate Observations. In *Proceedings of the 5th Berkeley Symposium on Maths and Statistics Problems*, volume 1, pages 281–297, 1967.

[19] H. J. Miller and J. Han. *Geographic Data Mining and Knowledge Discovery: An Overview*. Cambridge University Press, Cambridge, UK, 2001.

[20] T. M. Mitchell. *Machine Learning*. WCB/McGraw-Hill, New York, 1997.

[21] A. T. Murray, I. McGuffog, J. S. Western, and P. Mullins. Exploratory Spatial Data Analysis Techniques for Examining Urban Crime. *British Journal of Criminology*, 41:309–329, 2001.

[22] A. T. Murray and T. Shyy. Integrating Attribute and Space Characteristics in Choropleth Display and Spatial Data Mining. *International Journal of Geographical Information Science*, 14:649–667, 2000.

[23] A. Okabe, B. N. Boots, K. Sugihara, and S. N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, West Sussex, second edition, 2000.

6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)
0-7695-2841-4/07 $25.00 © 2007 IEEE