Separabilidad Geométrica aplicada a las Búsquedas por Rangos

Edilma Olinda Gagliardi
LIDIC[1]
Departamento de Informática
Facultad de Ciencias Físico, Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
oli@unsl.edu.ar
Fax: 54-2652-430224

Gregorio Hernández Peñalver
Departamento de Matemática Aplicada
Facultad de Informática
Universidad Politécnica de Madrid, España
gregorio@fi.upm.es
Fax: 34-91-3367426

Resumen

El objetivo de esta propuesta es mostrar un trabajo de investigación, en el que presentamos los aspectos teóricos y prácticos relevantes para las búsquedas por rangos y separabilidad de objetos geométricos. Nuestra pretensión es poder articular ambas temáticas, proponiendo nuevas formas de búsquedas en el espacio, a través de particionamientos obtenidos por la aplicación de criterios de separabilidad.

Palabras Claves: Búsquedas por rangos, Separabilidad geométrica, Rangos simpliciales.

1. Introducción

La Geometría Computacional es una disciplina que brinda un marco teórico y formal para el diseño de estructuras y análisis de algoritmos requeridos para dar soluciones a problemas de tipo geométrico que surgen en las más diversas áreas de la Informática. [Abe00], [BKOS97], [BY98], [GO97], [SU00], [Tou85], [Tou92].

Nuestra introducción al estudio de la disciplina se basa en un problema que se presenta a menudo en bases de datos: es el estudio de los rangos y las consultas por rangos, denominado *Búsqueda por Rangos*. Esta temática vista desde una perspectiva geométrica, nos abre una puerta a una línea de investigación vigente y con resultados recientes en Geometría: la *Separabilidad de objetos geométricos* [Sea02]. Los criterios de separabilidad tienen aplicaciones interesantes, como por ejemplo el análisis de imágenes, clasificación de datos, entre otros. Siempre que sea necesario discriminar objetos en un espacio de trabajo, por algún atributo del mismo, los criterios de separabilidad juegan un papel importante.

Nuestra propuesta consiste en vincular las Búsquedas por rangos con la Separabilidad geométrica, por medio de cuñas, bandas y combinaciones de ellas. Las regiones se determinan en base a atributos propios de los objetos geométricos y de su ubicación en el espacio. Buscamos generar subregiones del espacio que se distingan por las propias características de sus individuos. Las consultas, básicamente, se reducen a casos particulares de las búsquedas por rangos.

En el presente trabajo, primeramente, introducimos las Búsquedas por rangos, luego las nociones de Separabilidad y la vinculación propuesta entre ambas temáticas. Por último, brindamos nuestras conclusiones y perspectivas de trabajos futuros.

2. Búsquedas por Rangos

La Búsqueda por rangos es uno de los problemas centrales en Geometría Computacional, tanto por la variedad de aplicaciones que posee, como por la gran cantidad de problemas geométricos pueden resolverse observándolos como problemas relacionados a las búsquedas por rango [AE98], [Aga97],[Mat94].

---

[1] Laboratorio de Investigación y Desarrollo en Inteligencia Computacional. Director: Dr. Raúl Gallard.

• Parcialmente subvencionado por *Proyecto de la UPM de Al2003-1010-2.55 Geometría Computacional*.

Un *Espacio de Rangos* es un sistema $\Omega = (U, F)$, donde U es un conjunto de objetos geométricos y F es una familia de subconjuntos de U. Los elementos de F son llamados *Rangos* de Ω. El sistema Ω es llamado *Espacio de Rango Finito* si el conjunto U es finito.

Algunos ejemplos de espacios de rangos generales son:

Ω_1 = (R^d, { h/ h es un semiespacio en R^d }), Ω_2 = (R^d, { h/ h es una bola en R^d })

Un problema de búsqueda por rangos puede expresarse del siguiente modo: Dados un espacio de rangos $\Omega = (U, F)$, S un subconjunto de objetos de U y R un rango de F, consultar los objetos geométricos que están en $S \cap R$. En este caso, a R se lo llama *rango de consulta* (query range).

La Búsqueda por rangos, se clasifican en cuatro tipos de búsqueda por rangos: *Ortogonales* (*BRO*), *SemiAlgebraicos* (*BRSa*), *SemiEspaciales* (*BRSp*) y *Simpliciales* (*BRSx*). En cuanto a las consultas, dado un rango de consulta puede interesarnos efectuar alguna de las siguientes clases sobre él: *Consultas de Recuento* (*range counting*), *Consultas de Reporte* (*range reporting*), *Consultas Booleanas* (*range emptiness*) y *Consultas Extremales*. Como es típico en Geometría Computacional, usamos el modelo de computación Real RAM (*Real Random Access Machine*). La mayoría de las estructuras de datos para búsquedas o consultas por rango, son construidas en forma recursiva, dividiendo el espacio de objetos geométricos en varias regiones, con propiedades geométricas deseables sobre ellas. Estas estructuras de datos son referidas como esquemas de descomposición jerárquicos.

Las búsquedas por rangos simpliciales son de nuestro interés. Las consultas en este contexto son consultas en polígonos; es decir, los rangos son polígonos. Todo polígono puede descomponerse en piezas más pequeñas, realizando una triangulación del mismo. De esta forma, la consulta en el polígono se transforma en consultas sobre triángulos de la triangulación. El resultado de la consulta es la unión de las respuestas de cada triángulo. Este concepto puede extenderse al espacio d-dimensional, donde tenemos símplices. Un símplice es el cierre convexo de d+1 puntos independientes [HW87], [Mat92a], [Mat92c], [Sar98], [Wel88], [Wil82].

Dado que el tamaño y el tiempo de consulta de cualquier estructura de datos tiene al menos complejidad lineal y logarítmica respectivamente, nosotros consideramos estos extremos: i) ¿Cuán rápido puede ser contestada una consulta de rango simplicial usando estructura de datos de tamaño lineal? ii) ¿Qué espacio se requiere para una estructura de datos que conteste una consulta en tiempo logarítmico?. La combinación de estos dos extremos nos lleva a una competencia espacio-tiempo.

La mayoría de las estructuras de datos de tamaño orden lineal para BRSx, se basan en los *Árboles de Partición* (*Partition Trees*). Dados n puntos en el plano, dividimos el plano en varias regiones, de modo que todas las regiones contengan una cantidad semejante de puntos y que cualquier recta dada corte una o varias de las regiones. La eficiencia del árbol de partición se determina por el esquema de partición utilizado [Wil82].

Podemos mejorar los tiempos de consulta cambiando de $O(\sqrt{n})$ a $O(\log n)$. La idea subyacente es la misma que en los árboles de partición con particiones simpliciales disjuntas, sólo que nos trasladamos al espacio dual.

3. Separabilidad geométrica

Los trabajos de separabilidad están orientados a dos o más conjuntos disjuntos de objetos geométricos, básicamente para puntos, bajo diversos criterios de separabilidad (*bandas*, *cuñas*, *sectores*, etc.). Los principales resultados se deben a Seara [Sea02], y los resultados de su tesis doctoral fueron publicados previamente [AHMS01] [DHMS01] [HNRS01] [AHMS00] [HMRS99] [HNRS98], por mencionar sólo algunas referencias.

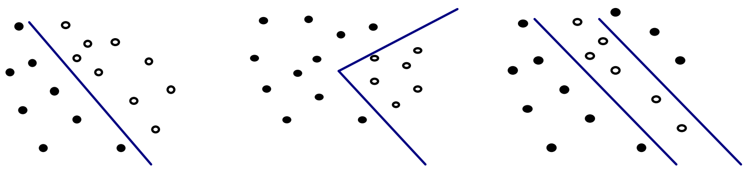
Se dan dos conjuntos disjuntos *P* y *Q* de objetos en el plano, clasificados como dos clases de objetos: *rojos* y *azules*, respectivamente. Eventualmente, los objetos pueden ser: puntos, segmentos, polígonos o círculos. En el caso de los polígonos, *N* y *M* representan el número total de segmentos de los polígonos de *P* y de *Q*. En otro caso, *N* y *M* representan la cardinalidad de *P* y *Q*, respectivamente. En cualquiera de los casos, *n* es el máximo de *N* y *M*.

Dada una familia C de curvas en el plano, decimos que los conjuntos P y Q son C-separables si existe una curva $C_i \in C$, llamada separador, tal que cada componente conectada de $R^2 - C_i$ contiene solamente objetos, o bien de P, o bien de Q. Atendiendo a la familia C, se tienen diferentes tipos de separabilidad:

Si *C* es la familia de rectas en el plano, tenemos la *separabilidad lineal*. Dos conjuntos *P* y *Q* son *linealmente separables* si y sólo si sus cierres convexos no se intersecan.

Dos conjuntos disjuntos *P* y *Q* de objetos en el plano son *separables por una cuña* si existe una cuña que contiene solamente todos los objetos de uno de los conjuntos. Aquí se estudia el problema de decidir la separabilidad por cuñas calculando la ubicación de los posibles ápices (vértices) de cuñas; como una extensión al problema, se estudia hallar la cuña de ángulo mínimo (máximo).

Dos conjuntos disjuntos *P* y *Q* de objetos en el plano son *separables por banda* si existe una banda, determinada por dos rectas paralelas, que contiene solamente todos los objetos de uno de los conjuntos. Aquí se estudia el problema de decidir la separabilidad por bandas calculando el conjunto de intervalos de pendientes de bandas separadoras; como una extensión al problema, se estudia hallar la banda mínima (máxima).



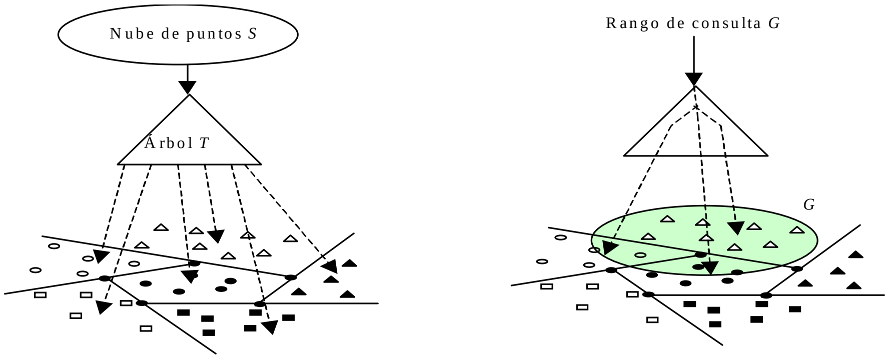
Separabilidad líneal   Separabilidad por Cuña   Separabilidad por Banda

La separabilidad de conjuntos de objetos geométricos está demostrada para separar dos conjuntos de puntos, segmentos, polígonos o círculos en el plano con cuñas y bandas, pudiendo encontrarse cuñas de máximo (mínimo) ángulo y bandas de máximo (mínimo) ancho. Los algoritmos para resolver los problemas de decisión y optimización propuestos anteriormente se ejecutan en $O(n \log n)$ tiempo; salvo en el caso de la separación lineal, que toma $O(n)$ tiempo.

Finalmente, algunos de los resultados expuestos pueden ser extendidos en el plano al uso de otros objetos geométricos que no sean puntos (segmento, círculos, etc.), y al espacio tridimensional [HSS03] en donde sólo existen algunos algoritmos para resolver algunos de los problemas vistos. En este marco, también queda abierto el estudio sobre nuevos criterios de separabilidad; es decir, no sólo usar planos o semiplanos, sino otros objetos como pirámides dobles (bipirámides), cilindros, cónicas, cónicas dobles, etc.

4. Separabilidad Geométrica aplica Búsquedas por Rangos

En esta sección queremos unificar las nociones de Búsquedas por rangos con las de Separabilidad geométrica. Nos interesan particularmente las descripciones de las curvas que determinan las regiones que contienen tales conjuntos disjuntos, dado que ellas constituyen los separadores geométricos. Con esta información podemos obtener un esquema de partición, en cuyo caso estamos en condiciones de crear una estructura de datos apropiada para acceder a los puntos. En la siguiente figura, podemos visualizar estas ideas:



Nube de puntos, dividida en regiones por alguna característica de los mismos

Regiones seleccionadas para la búsqueda

Con respecto a las consultas, podemos realizar la búsqueda por las características dadas de los puntos, o bien, por rangos. En el primer caso, por ejemplo, podemos solicitar las regiones correspondientes a los puntos rojos, pudiendo obtener como respuesta una o varias regiones. O bien, consultar cuáles son los diferentes colores existentes. También, podemos utilizar un rango de consulta, que puede abarcar parte de una región o varias regiones. Por ejemplo, un rango de consulta dado por un semiplano *H*, definido por una recta *h*, puede corresponderse con partes de regiones, con regiones completas, y posiblemente (es deseable), dejar excluidas regiones íntegras.

Los criterios de separabilidad utilizados para obtener las regiones poligonales pueden ser: lineal, por cuñas, por bandas, por doble cuña, etc. Es posible utilizar herramientas inteligentes[2] que nos brinden información acerca de cuál criterio de separabilidad es el más apropiado.

Es interesante y deseable pensar que las nubes de puntos acepten criterios de separabilidad acotados a complejidades bajas, aunque no es relevante, porque la principal presunción que podemos hacer es que la estructura se genera una única vez. Como hemos visto, puede ser prioritario crear estructuras con bajo costo de almacenamiento o bien crear estructuras que privilegien los tiempos de respuestas. Por asumir, justamente, que la estructura se crea una vez y que es estática, es que los costos de construcción son amortizados en base al criterio de ahorro elegido. Además, el objetivo principal es realizar muchas consultas por diferentes regiones sobre la misma nube de puntos.

5. Conclusiones y visión de futuro

El grupo de trabajo en la UNSL, conjuntamente con docentes de la UPM, han dado inicio en el presente año a un proyecto de investigación, Proyecto de la UPM de AL2002-1010-2.43 y AL2003-1010-2.55 Geometría Computacional, cuyo objetivo principal es la consolidación de la línea de trabajo en Geometría Computacional en la UNSL, aportando nuevos enfoques y técnicas algorítmicas a las líneas de investigación ya establecidas en su Departamento de Informática dentro de LIDIC (Laboratorio de Investigación y Desarrollo en Inteligencia Computacional). Esto constituye una ocasión inmejorable para impulsar investigadores en período de formación, la realización de tesis doctorales y de maestría en el campo de estudio.

6. Referencias bibliográficas

[Abe00] Abellanas Oar, M. *Descubriendo la Geometría Algorítmica, 2000*. http://www.dma.fi.upm.es/mabellanas/divulgación/GeometriaAlgoritmica.html

[AE98] Agarwal, P.; Erickson, J. *Geometric range searching and its relatives*; Advances in Discrete and

[2] Algunas herramientas de Inteligencia Computacional incluyen Metaheurísticas, Redes Neuronales, Algoritmos Evolutivos, Inteligencia Artificial Distribuida, Agentes, etc.

Comput. Geom. (B. Chazelle,J. Goodman, and R. Pollack, eds.), American Mathematical Society, Providence, 1998.

[Aga97] Agarwal, P.; *Range searching* CRC Handbook of Computational Geometry (J. Goodmand and J. O'Rourke, eds.).

[AHMS00] Arkin, E.M.; Hurtado, F.; Mitchell,S.B.; Seara, C.; Skiena, S.S.; *Some Separability Problems in the Plane*; 16th European Workshop on Computational Geometry, Eilat, Israel, 2000.

[AHMS01] Arkin, E.M.; Hurtado, F.; Mitchell,S.B.; Seara, C.; *Some Lower Bounds on Geometric Separability Problems*; 11th Fall Workshop on Computational Geometry, Polytechnic University, Brooklyn, NY, 2001.

[BKOS97] de Berg, M; Kreveld, Overmars, M; Schwarzkopf. *Computational Geometry: algorithms and applications*, Springer Verlag, BH 1997

[BY98] Boissonnat, J.D.; Yvinec, M. *Algorithmic Geometry*, Cambridge University Press, 1998.

[DHMS01] Devillers, O; Hurtado, F.; Mitchell,S.B.; Seara, C.; *Separating Several Point Sets in the Plane*; 13th Canadian Conference on Computational Geometry, 2001.

[GO97] Goodman, J.; O'Rourke, J. *Handbook of Discrete and Computational Geometry*. CRC Press 1997.

[HMRS99] Hurtado, F.; Mora, M.; Ramos, P.A.; Seara, C.; *Separación de Objetos en el Plano por Doble Cuña y por Θ-Poligonal*; VIII Encuentros de Geometría Computacional, Castelló, España, 1999.

[HNRS01] Hurtado, F.; Noy, M.; Ramos, P.A.; Seara, C.; *Separating Objects in the Plane with Wedges and Strips*; Discrete Applied mathematics, Vol 109, pp 109-138, 2001.

[HNRS98] Hurtado, F.; Noy, M.; Ramos, P.A.; Seara, C.; *Separating Objects in the Plane with Wedges and Strips*; 110th Canadian Conference on computational Geometry, Montreal, Canadá, 1998.

[HSS03] Hurtado, F.; Seara, C.; Sethia, S.; *Red-Blue separability problems ein 3D;* to appear in the 19th European Workshop on Computational Geometry, Bonn, Germany, March 24-26, 2003.

[HW87] Haussler, D.; Welzl,E,; *Epsilon nets and simplex range queries*; Discrete Comput. Geom., 2:127-151;1987.

[Mat92a] Matousek, J. *Efficient partition tree*; Discrette Comput. Geom., 8:315-334, 1992.

[Mat92c] Matousek, J.*Range searching with efficient hierarchical cuttings*; Proc. 8 ACM Symposium on Computational Geometry, 276-285, 1992.

[Mat94] Matousek,J.; *Geometric Range searching*; ACM Comput. Survey, 26:421-461;1994.

[Sar98] Sariel, Har-Peled; *Constructing planar cuttings in theory and practice*;

[Sea02] Seara, C; *On geometric separability*; Tesis Doctoral 2002, Univ. Politécnica de Cataluña.

[SU00] Sack, J.R.; Urrutia, J.. *Handbook of Computational Geometry*. Elsevier Science B.V. 2000.

[Tou85] Toussaint, G.T. *Computational Geometry* Edited by North-Holland, Amsterdam, 1985 .

[Tou92] Toussaint, G.T. *What is computational geometry?* Proc. IEEE, vol. 80, No. 9, pp. September,1992,1347-1363.

[Wel88] Welzl, E.; *Partition Tree for triangule counting and other range searching problems*; In Proc. 4 ACM Sympossim on Computational Geometry, 23-33,1988.

[Wil82] Willard, D.E.; *Polygon retrieval;* SIAM J. Comput., 11;149-165;1982.

[WL85] Willard, D.E.; Lueker, G.S*: kadding range restriction capability to dinamic data structures*. J.ACM, 32:597-617;1985.

[Y83] Yao,F.F.; *A 3-space partition and its applications.* Proc. 15th Annu. ACM Sympos. Theory Comput. 258-263, 1983.

[YDEP89] Yao,F.F.;Dobkin,D.P.;Edelbrusnner,H; Paterson,M.S.; *Partitioning space for range queries*; SIAM L Compuy. 18:371-384;1989.

[YY85] Yao, A.C.;Yao,F.F.; *A general approach to D-dimensional geometric queries*; Proc. 17th Annu. ACM Sympos. Theory Comput. 163-168, 1985.