عنوان المقال: خاصية الواحدات المتتالية

أ. م. د. رويدة رزاق محسن
جامعة الكوفة \ كلية علوم الحاسوب والرياضيات \ قسم الرياضيات
ruwaida.mohsin@uokufa.edu.iq

خاصية الواحدات المتتالية Consecutive Ones Property (C1P) هي خاصية المصفوفات التي مدخلاتها 0 او 1. مصفوفة الـ(0، 1) تنسب للصفوف عندما تبدل اعمدتها بحيث يظهر في كل صف من صفوفها مجموعة واحدات متتالية. في البرمجة الخطية الصحيحة، خاصية الواحدات المتتالية مهمة حيث تبين ان المسائل التي تعتمد على المصفوفات التي تمتلك هذه الخاصية حلها اسهل من المسألة الاصلية المقترحة. بالفعل، هذه المصفوفات تكون ([1] totally unimodular) . الخاصية C1P تظهر في عدة تطبيقات مثل مسألة تطوير السكك الحديد، أسترداد المعلومات، الجدولة [3]. وتستعمل ايضا في اعادة بناء جينوم الاسلاف، وبناء التطبيقات الفيزياوية من البيانات المهجنة [4 ,2]. في نظرية البيانات هذه الخاصية تساعد في تتبع فترات الرسم البياني [6 ,5].

خاصية الـC1P درست بشكل واسع. كندال [7] عزا اول ذكر لهذه الخاصية الى فلندرس بيترا، عالم اثار من القرن العشرين، [8]. حيث ان عدة طرق من طرق الحدس المهنية أنوجدت لمشكلة ترتيب الودائع الثرية زمنيا قبل اول حل للتعقيد الحسابي[2] الزمني متعدد الحدود polynomial المقترح من قبل فولكرسن و كروس [5].
في 1972، تاكر [1] عرض وصف للمسألة باستخدام مصفوفات جزئية محظورة. بعدها، بوث و لوكر [9] اعطوا خواوزمية ذات وقت خطي لها. حيث وجدوا تبديل لتحويل مصفوفة الـ(0، 1) الى مصفوفة تمتلك خاصية الـ C1P. الخوارزمية التي وجدوها تعتمد على هيكل بيانات الشجرة -PQ. حيث ان مصفوفة الـ (0، 1) تمتلك الـ C1P اذا وفقط اذا كانت الشجرة موجودة. الاساليب التي تعتمد على اشجار-PQ تعتبر معقدة، عموما في 1992، هاس[10] قدم فحص بسيط لخاصية C1P بدون استخدام اشجار-PQ. لقد قدم حل خطي لتعقيد الوقت للمسألة.

بعض التطبيقات

مسألة تسلسل تحديد التاريخ ([7] The Petrie Seriation):

تشمل المسألة ترتيب مواقع القبور زمنيا في مقبرة، حيث ان كل قبر يحتوي (او لا يحتوي) على عدد من الاثار يعود لحقبة زمنية معينة. المسألة تكتب رياضيا بواسطة مصفوفة الـ incidence matrix[2] حيث ان الاعمدة تقابل انواع محددة من الاثار والصفوف تقابل مواقع القبور. اركون و لوري [11] قدما نهج تطوري هجين لحل المشكلة. لقد قدما تبديل للصفوف لتحويل المصفوفة الى مصفوفة تمتلك مجموعة واحدات متعاقبة في كل عمود: وتسمى بمصفوفة بيتري (Petrie matrix). ايجاد هذا التبديل يعادل اظهار ان المصفوفة تمتلك خاصية الـ C1P.

خاصية الأسترجاع المتعاقبة ((Consecutive Retrieval Property (CRP) :

في عام 1972، كوش اقترح خاصية الأسترجاع المتعاقبة لكي تطبق على تنظيم الملفات. على سبيل المثال نظام ملفات الحاسوب. تقسيم المعلومات الى سجلات هي احدى طرق ترتيب المعلومات في نظام الخزن. ليكن R مجموعة السجلات التي تدعى ملف حيث تمتلك تركيب معين يختلف من بيانات الى اخرى. يمكن تخزين هذا الملف على وسيط تخزين S حيث يمثل مجموعة مواقع التخزين. الاسترجاع ذات الصلة أو استرجاع السجلات يعتبر استخدام اساسي لترتيب السجلات في وسيط التخزين. هذه الفايلات تسحب حسب الطلبات Q، و التي قد تمتلك اكثر من بنية اعتمادا على ما هو مطلوب.

1-A matrix is totally unimodular if the determinant of every induced square submatrix is equal to - 1, 0, or +1.

2- في علوم الكمبيوتر ، يعتبر تعقيد الوقت هو التعقيد الحسابي الذي يقيس أو يقدر الوقت المستغرق لتشغيل خوارزمية

3-In mathematics, an incidence matrix is a matrix that shows the relationship between two classes of objects.

تنظيم الملف R في وسيط التخزين بحيث كل طلب في Q يمتلك جواب يسمى تنظيم الملفات. بعض انواع طرق تنظيم الملفات تم تحسينها بالنسبة الى وسيط التخزين الخطي. أفضل هذه الانواع عندما يوجد تنظيم بين R و Q بحيث لكل طلب في Q، السجلات ذات العلاقة تخزن في مواقع خزن متعاقبة في R بدون تكرار. عليه يقال ان مجموعة الطلبات Q تمتلك خاصية الأسترجاع المتعاقبة CRP نسبة الى مجموعة السجلات R وهذا التنظيم يدعى تنظيم الأسترجاع المتعاقب للملفات [12, 13].

حسب هذا الافتراض، يمكن ان ترتبط مصفوفة الـ incident مع $R = \{r_\alpha\}_{\alpha=1}^{m}$ و $Q = \{q_\beta\}_{\beta=1}^{n}$. ليكن A مصفوفة $n \times m$ حيث تقابل الصفوف السجلات m في R والاعمدة تقابل الطلبات n في Q. الـ $a_{\alpha,\beta}$ عنصر في A اذا r_α ذات صلة بـ q_β و 0 غير ذلك. و عليه A تأخذ الشكل

$$A = \begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_m \end{array} \begin{array}{c} \begin{array}{ccccc} q_1 & q_2 & q_3 & \cdots & q_n \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & \dots & 1 \\ 0 & 1 & 1 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix} \end{array}$$

لقد تم دراسة تغيرات هذه المسألة في سياقات مختلفة. على سبيل المثال اعتمادا على فكرة الـ C1P، فولكرسن و كروس [5]، اقترحا طريقة لايجاد فيما اذا التركيب الجيد للجينات يتناسق مع النموذج الخطي للجين [14]. لقد اعطوا حلا عاما لمسألة الـ CRP.

تطبيق فيزياوي للـ DNA:

فولكرسن و كروس [15]، استعملا الـ C1P لدراسة التركيب الجيد للجين اعتمادا على عمل بنزر [16]. التطبيق الفيزياوي لشريط الـ DNA للانسان، لنقل جينوم كامل او كروموسوم تم العثور عليه عن طريق تعيين الترتيب الخطي لمجموعة من العلامات (markers[4]) . واحدة من الطرق المعروفة تعتمد على تطبيقات الاستنساخ (the clone) ، والتي تنشأ من خلال مكتبة أستنساخ. هذه المكتبة تحتوي على عدد كبير من الاستنساخات والتي تمثل قطع قصيرة من الـ DNA. هذه القطع قد تكون فوق بعضها البعض. لفحص فيما اذا كان زوج من القطع فوق بعضها (متقاطعة) أو لا، يتم استخدام تهجين مجسات (probes[5]) قصيرة. يتم تعقب الأستنساخ بواسطة عدد من المجسات و المجس الذي يتم تهجينه الى استنساخ يتم تحديده.

لمسألة تحتوي m من الاستنساخات و n من المجسات، المعلومات للمسألة هي مصفوفة الـ $A = a_{\alpha\beta}$ ذات الابعاد $n \times m$. اذا تم تهجين استنساخ الى مجسة فأن $a_{\alpha\beta} = 1$ و $a_{\alpha\beta} = 0$ للحالات الاخرى. الان، لأيجاد الترتيب الصحيح للمجسات على استنساخ معين، هذه المجسات الـ (1's) يجب ان تكون متعاقبة دون ان يظهر بينها استنساخ (0). عموما اذا ظهر صفر بين الواحدات بيجب ان نتعامل معه، أما من خلال حذف الصفوف التي تحتوي على مجموعة واحدات غير متعاقبة أو نعيد ترتيب الاعمدة للحصول على صفوف تحتوي على مجموعة واحدات متعاقبة. المصفوفات التالية توضح ذلك.

4- A marker is a uniquely recognizable DNA segment; the types of markers are ranged between 1 to 300-400 nucleotide bases in size such as probes, genes, ect.
5- Probes are short segments of DNA and called sequence tagged sits (STSs) which are located in a single sit of the genome.
6- A probe is called hybridize to a clone if it matches a substring in that clone.

$$\begin{array}{cc} & \text{Probes} \\ \text{Clones} & \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}, \end{array} \qquad \begin{pmatrix} \cancel{1} & \cancel{1} & \cancel{0} & \cancel{1} & \cancel{1} \\ 0 & 1 & 1 & 1 & 0 \\ \cancel{1} & \cancel{0} & \cancel{1} & \cancel{0} & \cancel{0} \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

References

[1] A. Tucker, "A structure theorem for the consecutive 1's property," Journal of Combinatorial Theory, Series B, vol. 12, no. 2, pp. 153–162, 1972.

[2] W.-F. Lu and W.-L. Hsu, "A test for the consecutive ones property on noisy data application to physical mapping and sequence assembly," Journal of Computational Biology, vol. 10, no. 5, pp. 709–735, 2003.

[3] D. S. Hochbaumand A. Levin, "Cyclical scheduling and multi-shift scheduling: Complexity and approximation algorithms," Discrete Optimization, vol. 3, no. 4, pp. 327–340, 2006.

[4] F. Alizadeh, R. M. Karp, D. K. Weisser, and G. Zweig, "Physical mapping of chromosomes using unique probes," Journal of Computational Biology, vol. 2, no. 2, pp. 159–184,1995.

[5] D. Fulkerson and O. Gross, "Incidence matrices and interval graphs," Pacific Journal of Mathematics, vol. 15, no. 3, pp. 835–855, 1965.

[6] P. Gilmore and A. Hoffman, "A characterization of comparability graphs and of interval graphs," Canadian Journal of Mathematics, vol. 16, pp. 539–548, 1964.

[7] D. Kendall, "Incidence matrices, interval graphs and seriation in archeology," Pacific Journal of Mathematics, vol. 28, no. 3, pp. 565–570, 1969.

[8] J. Meidanis, O. Porto, and G. P. Telles, "On the consecutive ones property," Discrete Applied Mathematics, vol. 88, no. 1, pp. 325–354, 1998.

[9] K. S. Booth and G. S. Lueker, "Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms," Journal of Computer and System Sciences, vol. 13, no. 3, pp. 335–379, 1976.

[10] W.-L. Hsu, "A simple test for the consecutive ones property," in Algorithms and Computation, pp. 459–468, Springer, 1992.

[11] M. L. Gargano and L. Lurie, "A hybrid evolutionary approach to solving the archaeological seriation problem," Congressus Numerantium, vol. 180, p. 43, 2006.

[12] J. S. Deogun and K. Gopalakrishnan, "Consecutive retrieval property revisited," Information Processing Letters, vol. 69, no. 1, pp. 15–20, 1999.

[13] S. P. Ghosh, "File organization: the consecutive retrieval property," Communications of the ACM, vol. 15, no. 9, pp. 802–808, 1972.

[14] J. S. Deogun, V. V. Raghavan, and T. K. Tsou, "Organization of clustered files for consecutive retrieval," ACM Transactions on Database Systems (TODS), vol. 9, no. 4, pp. 646–671, 1984.

[15] D. Fulkerson, O. Gross, et al., "Incidence matrices with the consecutive 1's property," Bulletin of the American Mathematical Society, vol. 70, no. 5, pp. 681–684, 1964.

[16] S. Benzer, "On the topology of the genetic fine structure," Proceedings of the National Academy of Sciences of the United States of America, vol. 45, no. 11, pp. 1607–1620, 1959.