Paraules k-completes

Enric Gago Hobel, Pedro Francisco Agredano Amengual, Iván Núñez Fuster

Assignatura de Models Matemàtics en l'Educació Secundària, Llicenciatura de Matemàtiques

***Abstract*—Amb aquest article es pretén apropar al lector sobre la utilitat de les paraules k-completes de manera que es pugui fer una idea general de què són i fins i tot com obtenir de forma senzilla una paraula completa de tamany k amb k arbitrari.**

I. Introducció

Qualque vegada ens haurem demanat com podem calcular la distància entre la Terra i un altre planeta, doncs una manera de fer-ho és mitjançant l'emissió d'un senyal en la direcció d'aquest planeta de manera que el que mesuram és el temps que triga aquest senyal a retornar rebotat. Ara bé, hem de ser capaços de detectar-lo a la tornada i no confondre'l, per exemple, amb soroll. Però, de quin tipus de senyals estam parlant? Estam parlant d'una seqüècia d'estímuls que s'emeten amb una certa freqüència, separats per espais de temps. Aquests senyals els podem representar per bits, on l'1 indica que s'ha emès el senyal i el 0 que no, és a dir, un dels espais de temps en que no s'ha emès res. Aquestes representacions per bits són el que anomenam paraules binàries (paraules on les lletres són 0's i 1's). De totes maneres, tampoc no podem emetre una paraula qualsevulla ja que a l'hora de rebre-la de nou pot ocasionar confusió.

II. Petits exemples

Si enviam per exemple el senyal

$$01100101011$$

la idea és rebre exactament el mateix senyal o qualque cosa semblant, el problema és que si rebem

001111**011**1**011**1

a causa del soroll no sabem exactament quan començam a rebre el nostre senyal i per tant no podrem determinar amb precisió quan el rebrem. Per evitar aquest problema i poder rebre el senyal igualment es varen inventar les paraules k-completes. Les paraules k-completes són paraules que no contenten cap subparaula de longitud k repetida. D'aquesta manera se soluciona el problema anterior perquè quan detectem de nou el senyal, a partir del bocí que recuperem podrem saber quant temps fa que va ser emès ja que aquests bocins de longitud k de la paraula són únics, no es repeteixen, i per tant podrem saber quan s'emeteren i a partir d'aquí determinar la distància cercada.

Facem un poc de pràctica, si consideram la paraula

$$1001$$

les seves subparaules de longitud 2 són, per ordre, 10, 00 i 01, que són totes diferents. Ara bé, no és una paraula 2-completa perquè no conté totes les subparaules de longitud dos, faltaria l'11. Podem determinar d'una manera intuïtiva la longitud d'una paraula 2-completa. El nombre màxim de paraules binàries diferents de longitud dos són 4: 00, 01, 10, 11. Si consideram la següent paraula, sí és 2-completa

$$10011$$

Ara bé, de longitud més gran que 5 és impossible ja que si tenim una paraula 2-completa de longitud 6 per exemple, serà de la forma

$$abcdef$$

on totes les subparaules de longitud dos són ab, bc, cd, de, ef i òbviament, com només n'hi ha quatre diferents, alguna de les anteriors es repetiran i per tant no podrà ser una paraula 2-completa. Si no n'hi ha de longitud 6 tampoc no n'hi haurà d'una longitud superior.

Estudiam ara un poquet les paraules 3-completes, un exemple d'una paraula on totes les subparaules de longitud tres són diferents és

$$100110$$

110, 001, 011, 110 són diferents. El nombre màxim de paraules binàries de longitud 3 és 8 (2^3), que són

000	001	010	100
011	101	110	111

Si consideram la paraula 3-completa

$$1110001011$$

veim que té 8 subparaules de longitud 3 diferents, és a dir, el màxim possible, i per tant no hi pot haver una amb més dígits perquè si no es repetirien algunes subparaules i ja no seria 3-completa, és a dir, les paraules 3-completes són de 10 dígits.

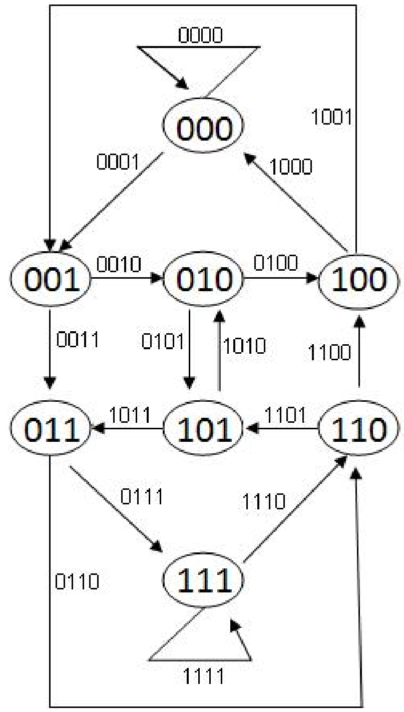
En general, quina és la longitud d'una paraula k-completa? Veient els exemples anteriors el raonament és molt senzill. En total hi ha 2^k paraules binàries diferents de longitud k i per ser paraula k-completa les haurà de contenir totes aquestes de forma seqüencial, ara bé, quants dígits té aquesta paraula? Hem de tenir en compte que aquestes 2^k subparaules estan "incloses" unes dins les altres, és a dir, del primer dígit al k-èssim és una subparaula, del segon al $(k+1)$-èssim una altra diferent i així 2^k vegades, és a dir, a partir de la k-èssima posició només afegim un dígit fins a 2^k vegades, per tant, podem deduir que en total hi ha $2^k + k - 1$ dígits.

D'aquesta manera, la longitud màxima d'una paraula 50-completa és de $2^{50} + 50 - 1 = 1125899906842673$ dígits. Tal vegada no ens hem fixat degut a que els exemples anteriors no són massa grans però se satisfà que a les paraules k-completes els primers $k-1$ i darrers $k-1$ dígits coincideixen. A l'exemple anterior, **11**100010**11**.

III. Construcció de paraules k-completes

Trobar paraules 2-completes i 3-completes no ha estat gaire complicat, ara bé, a mesura que k augmenta no és tan trivial trobar aquestes paraules. Per fer-ho ens ajudarem de la construcció d'un graf dirigit de manera que els nodes representin totes les paraules binàries de longitud $k-1$ i els arcs representin totes les paraules binàries de longitud k, de tal manera que una paraula k-completa correspongui a un circuit *eulerià*, és a dir, que recorre tots els arcs exactament una vegada d'aquest graf. La idea serà la següent, de cada node, que és una paraula de longitud $k-1$, en sortiran dos arcs dirigits, aquests dos arcs seran les paraules que corresponen d'afegir un 0 o un 1 a la dreta del node del qual surten obtenint així una paraula binària de longitud k i es dirigiran a aquells nodes que coincideixin amb els darrers $k-1$ dígits de l'aresta (dit d'una altra manera, llevam el primer dígit de l'aresta i miram a quin node correspon). Per exemple, si estam a $k=3$, aleshores del node 10 sortiran les arestes 100 i 101, la primera es dirigirà al node 00 i la segona al node 01.

Amb la construcció d'aquests grafs es pot construir d'una manera molt visual diferents paraules k-completes de manera molt senzilla, per exemple, ens demanam construir paraules 4-completes i 5-completes. Construirem només el graf per a $k=4$, el graf és el següent



Ara només ens hem de situar a un node i recórrer cada arc una i només una vegada. Per exemple, una paraula 4-completa que comenci amb el node 000 seria

$$0000111101011001000$$

Si començam en el vèrtex 110, obtenim per exemple,

$$1100001001111010110$$

Per a construir una paraula 5-completa s'hauria de fer la construcció del graf corresponent, després, una vegada fet ja és molt fàcil, un exemple de paraula 5-completa és

$$111110000010001101001110110010101111$$

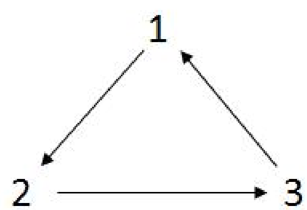
Fins aquí hem donat un mètode per trobar paraules k-completes mitjançant la construcció de grafs com ja s'ha dit abans, ara bé, aquests grafs sempre admenten un circuit eulerià? Ja que en cas contrari no sempre podrem fer ús d'aquest mètode. La resposta és afirmativa, de fet existeix un teorema que ho corrobora, és el següent

Un graf dirigit té un circuit eulerià si i només si és connex i tots els seus nodes tenen el mateix grau de sortida que d'arribada.

Òbviament, els nostres grafs són d'aquest estil ja que si no no tindria sentit parlar-ne d'aquest teorema.

Hem de provar les dues implicacions, la implicació cap a la dreta és més fàcil, de fet ho provarem per contrarecíproc. Vegem-ho. Volem provar que si un graf no és connex o els seus nodes no tenen el mateix grau de sortida que d'arribada aleshores el graf en qüestió no és eulerià. Si el graf no és connex òbviament no serà eulerià ja que existiran dos nodes que no es podran unir mitjançant una trajectòria sense repetir els arcs. Per altra banda, si existeix qualque node tal que el grau de sortida sigui diferent al grau d'entrada aleshores tampoc es podrà recórrer tots els arcs, per exemple, si en un node el grau de sortida és superior al grau d'entrada hi hauran sortides que no es poden recòrrer i per tant no és eulerià.

Vegem ara l'altra implicació, que si un graf és connex i tots els seus nodes tenen el mateix grau de sortida que d'arribada aleshores té un circuit eulerià. Per demostrar-ho farem inducció sobre el nombre d'arcs. Vegem que per a 3 arcs és cert. Un graf dirigit que sigui connex, amb tres arcs i que el grau d'entrada a cada node sigui igual al de sortida ha de ser de la forma



que efectivament té un circuit eulerià, és $C = (1, 2, 3, 1)$. Suposem ara que la propietat també és certa si el nombre de arcs és $3, 4, \dots, k-1$. Vegem que també és certa si el nombre d'arcs és k. La idea és construir un graf eulerià amb la hipòtesi d'inducció. El que feim és cercar un circuit qualsevol en G, suposem que hem començat en un vèrtex v_1 i formem un circuit C_1 que acabi en v_1, si aquest circuit ha passat per tots els vèrtexos aleshores ja hem acabat, en cas contrari el que feim serà suprimir els arcs emprats al circuit C_1. Ara s'hauran creat una o més components connexes en G i cada una satisfà que el nombre d'arcs és menor que k, i com G satisfà les hipòtesis d'inducció aleshores cada component connexa és euleriana. Ara hem d'unir-los de tal manera que aquesta unió sigui un circuit eulerià però això és molt fàcil, només hem de començar en un dels vèrtexos on s'uneixen dues o més de les components connexes i recórrer el graf com la unió de les components.

D'aquesta manera ens asseguram de que donat qualsevol $k \geq 3$ existeixen paraules k-completes, ja que el graf que hem definit abans sempre és construïble i com acabam de demostrar que sempre podrem trobar un circuit eulerià aleshores sempre hi haurà paraules k-completes. Per a $k = 1, 2$ també, ja que les paraules $01, 10$ són 1-completes i de 2-completes ja les hem vistes abans, per tant podem concloure que per tot $k \geq 1$ existeixen paraules k-completes.