ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΜΕΤΑΠΤΥΧΙΑΚΟ
ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
«Μαθηματικά και Εφαρμογές τους»

Μεταπτυχιακή Εργασία
«ΠΙΘΑΝΟΘΕΩΡΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΣΕ
ΚΑΤΩ ΦΡΑΓΜΑΤΑ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ»

Χόνδρος Αναστάσιος
Επιβλέπων Καθηγητής: Κολουντζάκης Μιχάλης
Ηράκλειο, Μάρτιος 2011

Επιτροπή Αξιολόγησης

- Θ. Γαρεφαλάκης, Επίκουρος Καθηγητής, Τμήμα Μαθηματικών, Πανεπιστήμιο Κρήτης
- Μ. Ι. Καραβέλας, Επίκουρος Καθηγητής, Τμήμα Εφαρμοσμένων Μαθηματικών, Πανεπιστήμιο Κρήτης
- Μ. Κολουντζάκης (Επιβλέπων Καθηγητής), Καθηγητής, Τμήμα Μαθηματικών, Πανεπιστήμιο Κρήτης

Ευχαριστίες

Θέλω να ευχαριστήσω τον καθηγητή μου κ.Κολουντζάκη Μιχάλη για την καθοδήγησή του και τον χρόνο που μου αφιέρωσε, καθώς και τα άλλα δύο μέλη της επιτροπής, κ.Καραβέλα Μενέλαο και τον κ.Γαρεφαλάκη Θεόδουλο, για την προσεκτική μελέτη της εργασίας και τις χρήσιμες παρατηρήσεις τους. Επίσης, ευχαριστώ τους γονείς και συγγενείς μου, για την οικονομική στήριξη όλα αυτά τα χρόνια καθώς και τους φίλους μου για τις απίστευτες ώρες «βλακείας» που περάσαμε μαζί. Τέλος, ένα μεγάλο «Ευχαριστώ» στην κοπέλα μου Αγγελική για την υπομονή της.

Περίληψη

Σκοπός της εργασίας που ακολουθεί είναι να δείξει την «δύναμη» των πιθανοθεωρητικών μεθόδων, χάρη στις οποίες τελικά μπορούμε να αποφανθούμε εάν τελικά κάτι «υπάρχει» χωρίς αναγκαστικά να το κατασκευάσουμε. Έτσι, στα κεφάλαια 1 και 2, δείχνουμε την ύπαρξη υποσυνόλων του $\mathbb{N}$ τα οποία έχουν κάποιες συγκεκριμένες ιδιότητες. Συγκεκριμένα, στο κεφάλαιο 1, δείχνουμε την ύπαρξη *ασυμπτωτικής προσθετικής βάσης* με συνάρτηση αναπαράστασης μεγέθους $\mathcal{O}(\log(x))$, ενώ στο κεφάλαιο 2, την ύπαρξη και κατασκευή σε πολυωνυμικό χρόνο ενός *sum-free* συνόλου, από ένα δεδομένο υποσύνολο του $\mathbb{N}$. Στα κεφάλαια 3 και 4, ασχολούμαστε με την συνάρτηση *parity* και χρησιμοποιώντας πιθανοθεωρητικά επιχειρήματα πάλι, δείχνουμε ότι δεν μπορεί να υπολογιστεί από λογικά κυκλώματα σταθερού βάθους και πολυωνυμικού μεγέθους καθώς και κάποια άλλα φράγματα.

Περιεχόμενα

Κεφάλαιο 0

Εισαγωγή

Η παρούσα εργασία εκπονήθηκε στο πλαίσιο του διατμηματικού μεταπτυχιακού προγράμματος σπουδών «Μαθηματικά και Εφαρμογές τους», του Μαθηματικού Τμήματος του Πανεπιστημίου Κρήτης. Όπως μαρτυρά και ο τίτλος της εργασίας, θα ασχοληθούμε με πιθανοθεωρητικές μεθόδους και στο μεγαλύτερο μέρος της ασχολούμαστε με την συνάρτηση *parity*, η οποία για είσοδο n μεταβλητών $\in \{0,1\}$, υπολογίζει το άθροισμά τους mod2. Η συγκεκριμένη συνάρτηση, εμφανίζεται σε αρκετούς τομείς στη θεωρία υπολογιστών καθώς σχετίζεται με αρκετές άλλες συναρτήσεις, όπως για παράδειγμα η συνάρτηση *majority*. Δικαιολογημένα λοιπόν υπήρξε ενδιαφέρον για την πολυπλοκότητα της και υπολογισμού κάτω φραγμάτων γι΄ αυτήν. Ο όρος, «πολυπλοκότητα κυκλώματος», εισήχθη από τον Shannon το 1948 [1], με τον οποίο και εννοούσε το μέγεθος του «μικρότερου» κυκλώματος που υπολογίζει την συνάρτηση, αφού πολύ απλά τον ενδιέφερε η ελαχιστοποίηση του hardware για τον υπολογισμό της συνάρτησης. Αργότερα, το 1965, οι Hartmanis & Stearns [2], συνέδεσαν την έννοια της πολυπλοκότητας με τις μηχανές Turing, ενώ το 1972 ο Savage [3], συνέδεσε την πολυπλοκότητα κατά Turing μιας συνάρτησης με την πολυπλοκότητα του κυκλώματός της. Έπρεπε να περιμένουμε μέχρι το 1984, για να αποδειχθούν καλύτερα φράγματα πέρα των γραμμικών (Paul, 1977, [4]) και των τετραγωνικών (Nečiporuk, [5]). Στην εργασία τους, οι Furst, Saxe & Sipser [14], κάνοντας χρήση μιας νέας τότε τεχνικής, ονομαζόμενης «τυχαίοι περιορισμοί» («random restrictions»), κατάφεραν να αποδείξουν υπερπολυωνυμικά φράγματα στην περίπτωση σταθερού βάθους κυκλωμάτων. Το ίδιο αποτέλεσμα, είχε αποδειχθεί και από τον Ajtai [6], ανεξάρτητα ένα χρόνο νωρίτερα, χρησιμοποιώντας διαφορετική μέθοδο, αλλά «συνδυαστικής» φύσεως πάλι. Συγκεκριμένα, η κεντρική ιδέα της μεθόδου, συνοψίζεται στην τιμοδοσία μεταβλητών, τυχαίου υποσυνόλου των μεταβλητών, ώστε να «εξαλειφθούν» κάποιες πύλες του κυκλώματος και συνεπώς να προκύψει κύκλωμα μικρότερου μεγέθους και άρα πιο εύκολο στην ανάλυσή του. Κάνοντας χρήση της ίδιας μεθόδου αλλά με περισσότερη ανάλυση, ο Yao [15], ένα χρόνο αργότερα, κατάφερε να αποδείξει εκθετικά κάτω φράγματα για την συνάρτηση *parity* ενώ ο Håstad [16] το 1989, κάνοντας χρήση μιας ισχυρότερης μορφής της ιδέας των Furst, Saxe & Sipser, κατάφερε να δείξει βέλτιστα, σχεδόν, φράγματα. Να σημειώσουμε, ότι στα ίδια

αποτελέσματα με τους Furst, Saxe & Sipser, κατέληξαν και οι Razborov - Smolensky [18], [19] το 1987, χρησιμοποιώντας όμως διαφορετικό πιθανοθεωρητικό επιχείρημα σε συνδυασμό με αλγεβρικές μεθόδους.
Στο υπόλοιπο κομμάτι της εργασίας, στα κεφάλαια 2 και 3 δηλαδή, αναπαράγουμε αποτελέσματα της δουλειάς του Paul Erdös, όπου με πιθανοθεωρητικά επιχειρήματα, δείχνουμε την ύπαρξη υποσυνόλου του $\mathbb{N}$ τέτοιου ώστε, κάθε «αρκετά» μεγάλο x να μπορεί να γραφεί ως άθροισμα 2 στοιχείων του συνόλου αυτού, ενώ ακόμα δείχνουμε, πως μπορούμε να εξάγουμε από ένα σύνολο, σε πολυωνυμικό χρόνο, ένα υποσύνολό τέτοιο ώστε κανένα στοιχείο του υποσυνόλου να μην μπορεί να γραφεί ως άθροισμα 2 άλλων στοιχείων.

Κεφάλαιο 1

Ασυμπτωτική Προσθετική Βάση

Ένα σύνολο $E \subseteq \mathbb{N}$, θα το ονομάζουμε *ασυμπτωτική προσθετική βάση τάξης 2*, εάν η συνάρτηση αναπαράστασης r με

$$r(x) = r_E(x) = |\{(a,b) : a,b \ \in E \,\text{με}\, a \leq b \,\&\, x = a+b\}|$$

είναι θετική για όλα τα «αρκετά» μεγάλα ακέραια x.

Δύο προφανή παραδείγματα τέτοιων συνόλων, είναι το σύνολο των φυσικών $\mathbb{N}$ και το σύνολο $\{1,2,4,6,\ldots\}$. Προφανώς, στα παραπάνω παραδείγματα, η συνάρτηση $r(x)$ έχει τάξη μεγέθους $\mathcal{O}(x)$. Ενδιαφερόμαστε να δούμε αν υπάρχει ασυμπτωτική βάση της οποίας η συνάρτηση αναπαράστασης είναι μικρότερη. Παρουσιάζονται δύο τρόποι για να απαντήσουμε στο ερώτημά μας: Είτε δείχνουμε με κάποιο τρόπο την ύπαρξη ενός τέτοιου συνόλου E, είτε ακόμη καλύτερα, το κατασκευάζουμε. Προφανώς, η δεύτερη προσέγγιση φαίνεται δυσκολότερη και συνεπώς τείνουμε προς την πρώτη, όπου ο Paul Erdös [7], [8] έχει ήδη απαντήσει για μας. Συγκεκριμένα απέδειξε ότι υπάρχει *ασυμπτωτική προσθετική βάση τάξης* 2 και θετικά c_1, c_2 τέτοια ώστε:

$$c_1 \log(x) \leq r(x) \leq c_2 \log(x)\ ,\ \text{για αρκετά μεγάλα } x.$$

Μάλιστα, ο λόγος c_1/c_2 μπορεί να πάρει τιμές όσο κοντά στο 1 επιθυμούμε.

Για αρχή, ορίζουμε τις πιθανότητες

$$p_x = K\left(\frac{\log(x)}{x}\right)^{1/2},$$

για όλα τα x για τα οποία το δεξί μέλος έχει νόημα, δηλαδή $p_x \in [0,1]$ αλλιώς θέτουμε $p_x = 0$. Η σταθερά K θα προσδιοριστεί αργότερα ανάλογα με τις ανάγκες μας.

Ορίζουμε τώρα ένα σύνολο E, θέτοντας

$$\mathbb{P}[x \in E] = p_x, \ \text{ ανεξάρτητα για όλα τα } x.$$

Μόλις ορίσαμε έμμεσα, ένα σύνολο το οποίο θέλουμε να μελετήσουμε και είμαστε έτοιμοι να δείξουμε ότι με μεγάλη πιθανότητα έχει την απαιτούμενη ιδιότητα:

$$c_1 \log(x) \leq r_E(x) \leq c_2 \log(x), \quad \text{για κάθε φυσικό αριθμό } x.$$

Ορίζουμε εν συνεχεία, τις ανεξάρτητες δείκτριες τυχαίες μεταβλητές X_j

$$X_j = \begin{cases} 1 & , \quad \text{εαν } j \in E \\ 0 & , \quad \text{αλλιώς} \end{cases}$$

με μέση τιμή

$$\mathbb{E}[X_j] = p_j \ .$$

Παρατηρούμε τώρα, σύμφωνα με τον τρόπο που έχουν οριστεί οι X_j, ότι

$$r(x) = \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} X_j X_{x-j} \tag{1.1}$$

και άρα

$$\mathbb{E}[r(x)] = \mathbb{E}\big[\sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} X_j X_{x-j}\big] = \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} \mathbb{E}[X_j X_{x-j}] = \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} p_j p_{x-j} \tag{1.2}$$

όπου στην τελευταία ισότητα εκμεταλλευτήκαμε την ανεξαρτησία των X_j.

Ένας τρόπος για να δει κανείς την ορθότητα της εξίσωσης 1.1, είναι να παρατηρήσει ότι οι μοναδικοί τρόποι ώστε να εκφράσει κάποιος τον αριθμό x ως άθροισμα δύο μη αρνητικών αριθμών a, b είναι:

$$\{(a,b) = (j, x-j), j = 1, \ldots, \lfloor \frac{x}{2} \rfloor\}$$

όπου $j \in \{1, 2, 3, \ldots, \lfloor \frac{x}{2} \rfloor\}$ επειδή ακριβώς απαιτούμε $1 \leq a \leq b$. Συνεπώς μας ενδιαφέρει να εξετάσουμε για αυτά τα j, εάν οι αριθμοί j & $x - j \in E$. Άρα, σύμφωνα με τον τρόπο που έχουν οριστεί οι μεταβλητές X_j,

$$j \in E \Leftrightarrow X_j = 1$$

και άρα

$$j \ \& \ (x-j) \in E \Leftrightarrow X_j X_{x-j} = 1$$

και άρα πράγματι έχουμε την ποσότητα $r(x)$ του δεξιού μέλους.

Αν κάνουμε την αντικατάσταση $p_j = K \left(\frac{\log(j)}{j}\right)^{\frac{1}{2}}$, τότε καταλήγουμε

$$\mathbb{E}[r(x)] = \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} K^2 \left(\frac{\log(j)\log(x-j)}{j(x-j)}\right)^{\frac{1}{2}} \tag{1.3}$$

Παρατηρώντας κανείς, ότι η συνάρτηση $f(t) = \log(t)\log(x-t)$ είναι αύξουσα στο διάστημα $[1, \frac{x}{2}]$ και γνωρίζοντας ότι

$$\frac{x}{2} \geq \frac{x}{\log(x)} \geq 1 \quad (\text{για αρκετά μεγάλο } x,\ \text{π.χ. } x > 10)$$

καταλήγει στο εξής:

$$\Big(\log\big(\frac{x}{\log(x)}\big) \cdot \log\big(x - \frac{x}{\log(x)}\big)\Big)^{1/2} \sum_{j=\frac{x}{\log(x)}}^{\lfloor \frac{x}{2} \rfloor} K^2(j(x-j))^{-1/2}$$

$$\leq \mathbb{E}[r(x)] \leq$$

$$\log(x/2) \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} K^2(j(x-j))^{-1/2}$$

και κάνοντας χρήση της ανισότητας

$$\log\left(x - \frac{x}{\log(x)}\right) \geq \log\left(\frac{x}{\log(x)}\right) \quad (\text{επίσης ισχύει για αρκετά μεγάλο } x)$$

παίρνει σαν αποτέλεσμα

$$\log\left(\frac{x}{\log(x)}\right) \sum_{j=\frac{x}{\log(x)}}^{\lfloor \frac{x}{2} \rfloor} K^2(j(x-j))^{-1/2} \leq \mathbb{E}[r(x)] \leq \log(\frac{x}{2}) \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} K^2\big(j(x-j)\big)^{-1/2} \tag{1.4}$$

Γράφουμε τώρα το δεξί άθροισμα της (1.4) ως εξής:

$$\sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} (j(x-j))^{-1/2} = \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} \left[\Big(\frac{j}{x}(1-\frac{j}{x})\Big)^{-1/2} \cdot x^{-1} \right] = \frac{1}{x} \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} \left[\Big(\frac{j}{x}(1-\frac{j}{x})\Big)^{-1/2} \right]$$

$$= \frac{1}{x}\Big[\big(\frac{1}{x}(1-\frac{1}{x})\big)^{-\frac{1}{2}} + \big(\frac{2}{x}(1-\frac{2}{x})\big)^{-\frac{1}{2}} + \cdots + \big(\frac{x/2}{x}(1-\frac{x/2}{x})\big)^{-\frac{1}{2}} \Big]$$

Αν αντικαταστήσουμε το $\frac{1}{x}$ στον πρώτο όρο με s και θεωρήσουμε τη συνάρτηση f με τύπο

$$f(s) = \big(s(1-s)\big)^{-\frac{1}{2}},$$

τότε η ακολουθία αποτελεί ένα Riemann άθροισμα της f στις θέσεις:

$$\xi_1 = \frac{1}{x}, \xi_2 = \frac{2}{x}, \ldots, \xi_{x/2} = \frac{x/2}{x} (= 1/2) \quad (\xi_1 \to 0, \xi_{\lfloor \frac{x}{2} \rfloor} \to \frac{1}{2}) \text{ όταν } x \to \infty$$

Έτσι, το ζητούμενο άθροισμα, στην περίπτωση όπου $x \to \infty$, συγκλίνει στο ορισμένο ολοκλήρωμα της f στο διάστημα $[0, \frac{1}{2}]$ δηλαδή με:

$$\int_0^{1/2} \big(s(1-s)\big)^{-\frac{1}{2}}\, ds$$

Εργαζόμενοι όμοια και για το άθροισμα που φράσσει από «κάτω» την ποσότητα $\mathbb{E}[r(x)]$ στην σχέση (1.4), καταλήγουμε και πάλι στο ίδιο ακριβώς ολοκλήρωμα. Παίρνοντας τώρα τον λόγο των 2 μεγεθών που φράσσουν την ποσότητα $\mathbb{E}[r(x)]$ βλέπουμε ότι

$$\lim_{x\to\infty} \frac{\log(\frac{x}{2})K^2 \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} (j(x-j))^{-1/2}}{\log(\frac{x}{\log(x)})K^2 \sum_{j=\frac{x}{\log(x)}}^{\lfloor \frac{x}{2} \rfloor} (j(x-j))^{-1/2}} =$$

$$\lim_{x\to\infty} \frac{\log(\frac{x}{2})}{\log(\frac{x}{\log(x)})} = 1$$

και άρα, σε συνδυασμό με την σχέση (1.4) συμπεραίνουμε ότι

$$\lim_{x\to\infty} \frac{\log(\frac{x}{2})K^2 \sum_{j=1}^{\lfloor \frac{x}{2} \rfloor} (j(x-j))^{-1/2}}{\mathbb{E}[r(x)]} = 1$$

Η τελευταία σχέση, μπορεί να γραφεί σε πιο «συμπαγή» μορφή με τη βοήθεια του συμβόλου " $\sim$ " ως εξής:

$$\mathbb{E}[r(x)] \sim IK^2 \log(x), \quad I = \int_0^{1/2} \big(s(1-s)\big)^{-\frac{1}{2}}\, ds \tag{1.5}$$

Ορίζουμε τώρα ως «ανεπιθύμητα» γεγονότα A_x, γεγονότα δηλαδή τα οποία δεν θα θέλαμε να πραγματοποιηθούν, τα εξής:

$$A_x = \{x : |r(x) - \mathbb{E}[r(x)]| > \frac{1}{2}\mathbb{E}[r(x)]\}\}, \quad x = 1, 2, \ldots$$

Θέλουμε να εξασφαλίσουμε ότι οι ποσότητες $\{r(x)\}$ δεν θα «απέχουν πολύ» από την αναμενόμενη τιμή τους. Στην περίπτωση μας, το «πολύ» είναι $1/2$, παρόλο που οποιαδήποτε σταθερά και να επιλέξουμε στη θέση του $1/2$, μπορούμε αν ορίσουμε κατάλληλα τα c_1 και c_2 να επιτύχουμε το ζητούμενο αποτέλεσμα

$$c_1 \log(x) \le r(x) \le c_2 \log(x),$$

αρκεί βέβαια να επιλέξουμε κατάλληλα και τη σταθερά K.

Στο σημείο αυτό, θα αναφέρουμε ένα θεώρημα σχετικά με τα *φράγματα Chernoff* [10], το οποίο και θα εφαρμόσουμε στους υπολογισμούς μας:

Θεώρημα 1 *Έστω X τυχαία μεταβλητή, με $X = X_1 + \cdots + X_k$ και X_j ανεξάρτητες δείκτριες τυχαίες μεταβλητές. Τότε,*

$$\forall \varepsilon > 0 : \quad \mathbb{P}\big[|X - \mathbb{E}[X]| > \varepsilon \mathbb{E}[X]\big] \le 2\exp(-c_\varepsilon \mathbb{E}[X])$$

όπου το c_ε εξαρτάται από το ε με τέτοιο τρόπο ώστε

$$c_\varepsilon = \min\{-\log(e^\varepsilon (1+\varepsilon)^{-(1+\varepsilon)}),\, \varepsilon^2/2\}$$

Παρατηρούμε τώρα, ότι αν σταθεροποιήσουμε το x, η ποσότητα

$$r(x) = \sum_{j=1}^{\lfloor x/2 \rfloor} X_j X_{x-j}$$

ικανοποιεί τις προϋποθέσεις του **Θεωρήματος 1** αφού αν θέσουμε

$$X_j X_{x-j} = X'_j, \quad j = 1, 2, \ldots$$

τότε πράγματι

$$r(x) = X'_1 + \cdots + X'_{\lfloor x/2 \rfloor}$$

δηλαδή άθροισμα δείκτριων μεταβλητών. Και συνεπώς μπορούμε πλέον να εφαρμόσουμε το **Θεώρημα 1** και να πάρουμε ότι:

$$\mathbb{P}[A_x] = \mathbb{P}\Big[\ \big|r(x) - \mathbb{E}[r(x)]\ \big| > \frac{1}{2}\mathbb{E}[r(x)]\Big] \leq$$

$$2e^{-c_{1/2} IK^2 \log(x)} \leq 2e^{-\frac{1}{2} c_{1/2} IK^2 \log(x)} = 2x^{-a}$$

όπου $a = \frac{1}{2} c_{1/2} IK^2$ και $c_{1/2}$ σταθερά που εξαρτάται από το $1/2$ με τρόπο που αναφέρεται στο θεώρημά μας.

Αν τώρα επιλέξουμε τη σταθερά K έτσι ώστε να πάρουμε $a > 1$, θα έχουμε καταλήξει σε κάτι πολύ βολικό: από τη στιγμή που η σειρά $\sum \mathbb{P}[A_x]$ είναι *συγκλίνουσα*, έπεται ότι για οποιαδήποτε σταθερά, όσο κοντά στο 0 επιθυμούμε, ας πούμε εδώ για παράδειγμα $1/2$ τότε

$$\exists\, n_o \in \mathbb{N} \ : \ \sum_{x \geq n_0} \mathbb{P}[A_x] < \frac{1}{2}$$

Άρα, με πιθανότητα τουλάχιστον $1 - \frac{1}{2} = \frac{1}{2}$, κανένα από τα $A_x,\ x \geq n_0$, δεν πραγματοποιείται, όπου ως A_x ορίσαμε τα γεγονότα που θα θέλαμε να αποκλείσουμε.

Συνδυάζοντας τα παραπάνω, καταλήγουμε στην ύπαρξη συνόλου $E \subseteq \mathbb{N}$ τέτοιου ώστε

$$\frac{1}{2} IK^2 \log(x) \leq r_E(x) \leq \frac{3}{2} IK^2 \log(x), \quad \forall x \geq n_0$$

το οποίο ήταν και το ζητούμενο.

Κεφάλαιο 2

Εξαγωγή ενός sum-free υποσυνόλου από ένα σύνολο σε πολυωνυμικό χρόνο

Ένα σύνολο ακεραίων E, θα το ονομάζουμε *sum-free* εάν

$$x + y \neq z \quad \forall\, x, y, z \in \mathrm{E}\ .$$

Για παράδειγμα, κάθε υποσύνολο των περιττών όπως και σύνολα της μορφής $\{N/2, \ldots, N\}$ αποτελούν σύνολα με αυτήν την ιδιότητα.

Δεδομένου ενός συνόλου $\mathrm{A} = \{k_1, \ldots, k_N\}$ όπου $k_i \in \mathbb{Z}\ \forall i$, θα δείξουμε πώς μπορούμε να εξάγουμε ένα *sum-free* $\mathrm{E} \subseteq \mathrm{A}$ με $|\mathrm{E}| > \frac{1}{3}|\mathrm{A}|$. Επίσης, θα δούμε ότι ο αλγόριθμος μας υλοποιείται σε πολυωνυμικό χρόνο ως προς το μέγεθος της εισόδου.

Έχει αποδειχθεί από τους *Erdös, Alon* και *Kleitman* [11], [12] ότι κάθε σύνολο

$$\mathrm{A} = \{k_1, \ldots, k_N\} \text{ με } \mathrm{A} \subseteq \mathbb{Z}\ ,$$

περιέχει ένα *sum-free* υποσύνολο E, όπου $|\mathrm{E}| > \frac{1}{3}N$. Η απόδειξη είναι πιθανοθεωρητική και μάλιστα τέθηκε το ερώτημα από τους *Alon* και *Kleitman*, εάν μπορεί να κατασκευαστεί αλγόριθμος πολυωνυμικού χρόνου ως προς το μέγεθος του προβλήματος τέτοιος ώστε να εξάγει ένα *sum-free* υποσύνολο E. Εύκολα βλέπει κανείς, ότι το μέγεθος του προβλήματος στην περίπτωση μας είναι

$$l = \sum_{j=1}^{N} \log_2(k_j)\ ,$$

αφού τόσα δυαδικά ψηφία απαιτούνται για να αναπαρασταθούν αυτοί οι N αριθμοί.

Θα δείξουμε ότι, κάνοντας κάποιες μικρές τροποποιήσεις, η απόδειξη των *Alon* και *Kleitman* μπορεί να μετατραπεί σε έναν τέτοιο αλγόριθμο. Ένας παρόμοιος αλγόριθμος κατασκευάστηκε ανεξάρτητα και από τους *Alon, Kriz* και *Nesetril* [13].

Μια περίπτωση όπου ένας τέτοιος αλγόριθμος αποδεικνύεται χρήσιμος, είναι στην προσέγγιση των αριθμών *Ramsey*. Πολλά κάτω φράγματα για τους αριθμούς *Ramsey* έχουν βρεθεί, αναλύοντας ένα σύνολο σε *sum-free* υποσύνολα. Να θυμίσουμε εδώ, ότι ο αριθμός *Ramsey*, $R(m,n)$, είναι ο ελάχιστος αριθμός κορυφών έτσι ώστε κάθε γράφημα με σύνολο κορυφών μεγέθους $R(m,n)$, είτε περιέχει κλίκα μεγέθους m, είτε περιέχει αντι-κλίκα μεγέθους n.

Συμβολισμός: Έστω p πρώτος. Τότε $\mathbb{Z}_p = \{0,1,\ldots,p-1\}$ ενώ με $\mathbb{Z}_p^*$ θα συμβολίζουμε την πολλαπλασιαστική ομάδα, δηλαδή $\mathbb{Z}_p^* = \{1,2,\ldots,p-1\}$.

Θεώρημα 2 *Έστω p πρώτος της μορφής $p = 3k+2$ και $w(x)$ μία μη αρνητική συνάρτηση ορισμένη στο $\mathbb{Z}_p^*$. Ορίζουμε*

$$w = \sum_{x \in \mathbb{Z}_p^*} w(x)$$

και υποθέτουμε ακόμα ότι $w > 0$.
Τότε υπάρχει ένα sum-free υποσύνολο E' του $\mathbb{Z}_p^*$ για το οποίο ισχύει ότι

$$\sum_{x \in E'} w(x) > \frac{1}{3} w \, .$$

Απόδειξη: Έστω $S = \{k+1,\ldots,2k+1\}$. Παρατηρούμε τώρα ότι για οποιαδήποτε 2 στοιχεία $x, y \in S$ έχουμε το εξής :

$$\begin{gathered} x + y = (k+j) + (k+l),\ 1 \le j \le l \le k+1 \\ \Rightarrow x + y = 2k + (j+l),\ 2 \le j+l \le 2k+2 \\ \Rightarrow 2k+2 \le x+y \le 4k+2 \\ \Rightarrow x + y \pmod p \in [1,k] \cup [2k+2, 3k+2] \\ \Rightarrow \nexists\ z \in S :\ x + y \equiv z \pmod p \end{gathered}$$

και άρα έχουμε δείξει ότι το σύνολο S είναι *sum-free* και επίσης $|S| > \frac{p-1}{3}$.
Έστω τώρα τυχαία μεταβλητή t, ομοιόμορφα κατανεμημένη στο $\mathbb{Z}_p^*$ και f τ.ω.

$$f(t) = \sum_{(t \cdot x) \in S} w(x)$$

Παίρνουμε μέση τιμή και έχουμε

$$\mathbb{E}[f(t)] = \mathbb{E}\Big[\sum_{(t \cdot x) \in S} w(x) \Big] = \sum_{(t \cdot x) \in S} \mathbb{E}[w(x)] = |S| \frac{w}{p-1} > \frac{w}{3} \qquad (2.1)$$

Για να καταλήξουμε στην τελευταία ισότητα πρέπει πρώτα να παρατηρήσουμε ότι η απεικόνιση T με

$$\begin{gathered} \mathrm{T} : \mathbb{Z}_p^* \to \mathbb{Z}_p^* \\ x \to t \cdot x \, , \quad t \in \mathbb{Z}_p^* \end{gathered}$$

είναι «1-1» και «επί».

- **«1-1»** Έστω x, y τέτοια ώστε $x \neq y$ και έστω ότι $tx = ty$

  $$\Rightarrow t\cdot(x-y) = 0 \Rightarrow t = 0 \text{ ή } x-y = 0 \quad (\mathbb{Z}_p \text{ σώμα } \Rightarrow \nexists \text{ μηδενοδιαιρέτες})$$

  όμως

  $$t \neq 0 \text{ αφού } t \in \mathbb{Z}_p^* \text{ και } x \neq y \text{ εξ' υποθέσεως, άτοπο.}$$

- **«επί»** Έχουμε ήδη δείξει ότι η απεικόνιση T είναι «1-1» και σε συνδυασμό με το γεγονός ότι

  $$tx \in \mathbb{Z}_p^*\,, \quad \forall x$$

  έχουμε και την ιδιότητα του «επί».

Γυρνάμε πάλι πίσω και μπορούμε πλέον να συμπεράνουμε λόγω της (2.1) ότι

$$\exists\, t_0 \in \mathbb{Z}_p^* \text{ τ.ω. } f(t_0) > \frac{w}{3}$$

Σε αντίθετη περίπτωση, αν δηλαδή ίσχυε ότι $f(t) \leq \frac{w}{3}\ \forall t$, τότε θα είχαμε και ότι

$$\mathbb{E}[f(t)] \ \leq \frac{w}{3}\,, \quad \text{το οποίο δεν συμβαίνει.}$$

Ορίζουμε τώρα το σύνολο E', έτσι ώστε $E' = {t_0}^{-1}S$. Συνεπώς και το E' είναι και αυτό *sum-free*, γιατί αλλιώς δεν θα ήταν ούτε το σύνολο S *sum-free* και επιπλέον ικανοποιεί και τη σχέση

$$\sum_{x\in E'} w(x) = \sum_{(t_0\cdot x)\in S} w(x) = f(t_0) > \frac{w}{3}$$

Κάνουμε τώρα την εξής παρατήρηση: Ο μέγιστος αριθμός πρώτων παραγόντων που μπορούν να εμφανιστούν στην ανάλυση ενός αριθμού n είναι το πολύ $\log_2(n)$ το πλήθος. Πράγματι, αν

$$n = p_1^{v_1}\cdots p_k^{v_k}$$

με p_i διαφορετικούς πρώτους παράγοντες και $v_i \geq 1$, τότε $p_i \geq 2$ και άρα

$$n \geq 2^{v_1}\cdots 2^{v_k} = 2^{v_1+\ldots v_k} \geq 2^k$$

και τελικά

$$k \leq \log_2(n).$$

Συνεπώς, το πλήθος των πρώτων που εμφανίζονται στην παραγοντοποίηση όλων των αριθμών που ανήκουν στο A είναι το πολύ l, όπου l να θυμίσουμε ισούται με το μέγεθος της εισόδου. Κάνοντας τώρα χρήση του θεωρήματος πρώτων αριθμών για αριθμητικές προόδους, μπορούμε να ισχυριστούμε ότι ο $(l+1)$-ιοστός πρώτος, είναι μικρότερος από $(l+1)\log(l+1)$, άρα και μικρότερος από $3l\log_2(l)$. Τώρα ξέρουμε, ότι ανάμεσα σε τουλάχιστον $l+1$ πρώτους και από τη στιγμή που οι πρώτοι διαιρέτες του συνόλου A είναι το πολύ l, ότι

υπάρχει ένας πρώτος της μορφής $p = 3k + 2$, ο οποίος δεν διαιρεί κανένα στοιχείο του A.
Ορίζουμε τώρα

$$w(x) = |\{t \in \mathrm{A} : t \equiv x \pmod p\}|$$

και θέλουμε να υπολογίσουμε την ποσότητα

$$w = \sum_{x \in \mathbb{Z}_p^*} w(x) \ .$$

Για ευκολία, ας συμβολίσουμε με A_x το σύνολο $\{t \in \mathrm{A} : t \equiv x \pmod p\}$. Ένας εύκολος τρόπος για να εργαστούμε είναι να δούμε ότι η σχέση $\pmod p$ επί ενός συνόλου, επάγει σχέση ισοδυναμίας και άρα *διαμερίζει*, στην περίπτωσή μας, το σύνολο A στις κλάσεις ισοδυναμίας $\mathrm{A_x}$. Ειδικότερα, το σύνολο $\mathrm{A_p}$ είναι κενό, αφού ο p δεν διαιρεί κανένα στοιχείο του A και συνεπώς

$$|\mathrm{A}| = \sum_{x \in \mathbb{Z}_p} |\mathrm{A_x}| = \sum_{x \in \mathbb{Z}_p^*} |\mathrm{A_x}| = \sum_{x \in \mathbb{Z}_p^*} w(x)$$

Άρα

$$w = \sum_{x \in \mathbb{Z}_p^*} w(x) = |\mathrm{A}| = N$$

και κάνοντας τώρα χρήση του **Θεωρήματος 2**, μπορούμε να συμπεράνουμε ότι

$$\exists\, \mathrm{E}' \subseteq \mathbb{Z}_{\imath}{}^* \text{ τ.ω. } \sum_{x \in E'} w(x) > \frac{1}{3} w \ \text{ με } \ \mathrm{E}' \textit{ sum-free.}$$

Όμως

$$\sum_{x \in E'} w(x) = \sum_{x \in E'} |\{t \in \mathrm{A} : t \equiv x \pmod p\}| = |t \in \mathrm{A} : t \bmod p \in \mathrm{E}'| = |\mathrm{E}|$$

και επίσης

$$w = N$$

Συνδυάζοντας τις 2 τελευταίες σχέσεις συμπεραίνουμε ότι

$$\exists\, \mathrm{E} \subseteq \mathrm{A} \text{ με } |\mathrm{E}| > \frac{1}{3} N \ .$$

Επιπλέον, το σύνολο E είναι και αυτό *sum-free*, και αυτό γιατί αν υποθέσουμε το αντίθετο, τότε θα υπήρχαν $x, y, z \in \mathrm{E}'$ τ.ω. $x + y = z$ και συνεπώς και $x + y \equiv z \pmod p$.
Όμως $x \bmod p$, $y \bmod p$, $z \bmod p \in \mathrm{E}'$ (εξ' ορισμού), που θα σήμαινε πως το σύνολο E' δεν θα ήταν *sum-free*, το οποίο δεν είναι αληθές.

Περιληπτικά τα βήματα του αλγορίθμου είναι τα εξής:

1. Υπολογίζουμε όλους τους πρώτους $\leq 3l \log_2(l)$.

2. Βρίσκουμε έναν πρώτο $p = 3k + 2$ τ.ω. p δεν διαιρεί κανένα x, όπου $x \in \mathrm{A}$.

3. Υπολογίζουμε τη συνάρτηση $w(x)\ \forall x \in \mathbb{Z}_p^*$.

4. Βρίσκουμε με δοκιμές ένα $t_0 \in \mathbb{Z}_p^*$ τ.ω. $f(t_0) > \frac{N}{3}$ και κατασκευάζουμε το σύνολο $E' = t_0^{-1}S$.

5. Κατασκευάζουμε τέλος το σύνολο $E = \{t \in \mathrm{A} : t \bmod p \in E'\}$.

Με μια γρήγορη εκτίμηση, μπορούμε να καταλήξουμε στα εξής φράγματα για τα κόστη των βημάτων 1 έως 5:

- Για το βήμα 1, χρησιμοποιώντας το «Κόσκινο του Ερατοσθένη», βλέπουμε ότι έχουμε μία λίστα από $3l\log_2(l)$ αριθμούς, $\{1, \ldots, 3l\log_2(l)\}$. Επίσης, σε κάθε βήμα του αλγορίθμου αυτού, γίνονται το πολύ $\frac{3l\log_2(l)}{2}$ «διαγραφές» στη λίστα, όσο και το μέγιστο πλήθος διαιρετών του $3l\log_2(l))$ δηλαδή, ενώ τα βήματα είναι $\frac{3l\log_2(l)}{2}$, αυτός είναι και ο μέγιστος δυνατός διαιρέτης του $3l\log_2(l)$ και άρα βλέπουμε ότι υλοποιούνται συνολικά $\mathcal{O}(l^2\log^2(l)) = \mathcal{O}(l^3)$ βήματα.

- Για το βήμα 2, έχοντας ήδη τη λίστα όλων των πρώτων από το βήμα 1, «κρατάμε» όλους τους πρώτους $\equiv 2 \pmod 3$(το οποίο γίνεται σε $\mathcal{O}(l\log(l)) = \mathcal{O}(l^2)$ βήματα, αφού τόσο είναι και το μέγεθος της λίστας) και τώρα για κάθε έναν τέτοιο πρώτο p, οι οποίοι είναι συνολικά πλήθους $\mathcal{O}(l^2)$) κάνουμε $|\mathrm{A}| = N \leq \sum_{j=1}^{N}\log_2(k_j) = l$ διαιρέσεις, για να δούμε αν διαιρεί κάποιο στοιχείο του συνόλου A. Άρα το συνολικό κόστος γι' αυτό το βήμα είναι τάξεως $\mathcal{O}(l^3)$.

- Για το βήμα 3, όπου έχουμε να υπολογίσουμε τη συνάρτηση $w, \forall x \in \mathbb{Z}_p^*$, βλέπουμε ότι έχουμε να υπολογίσουμε $p \leq 3l\log_2(l) \leq 3l^2$, άρα $\mathcal{O}(l^2)$ τιμές της συνάρτησης w. Όμως, για δεδομένη είσοδο x_0 και για κατάλληλο πρώτο p, για να υπολογιστεί η $w(x_0)$, χρειάζεται το πολύ $|\mathrm{A}| = N = \mathcal{O}(l)$ βήματα, αν θεωρήσουμε ότι η πράξη $x_0 \bmod p$ έχει κόστος 1. Επομένως, για το βήμα 3 απαιτούνται συνολικά $\mathcal{O}(l^3)$ βήματα.

- Για το βήμα 4, βλέπουμε ότι έχουμε αρχικά να υπολογίσουμε το πολύ p τιμές της συνάρτησης f, δηλαδή $\mathcal{O}(l^2)$ τιμές. Για να υπολογισθεί όμως μία τιμή της συνάρτησης f, χρειάζονται να αθροιστούν το πολύ $|S| \leq p = \mathcal{O}(l^2)$ τιμές της συνάρτησης w, η οποία υπολογίσαμε στο βήμα 3 ότι έχει κόστος $\mathcal{O}(l)$. Άρα για την εύρεση ενός $t_0 \in \mathbb{Z}_p^* : f(t_0) > \frac{N}{3}$, χρειάζονται το πολύ $\mathcal{O}(l^5)$ βήματα. Κατόπιν για να κατασκεύασουμε το σύνολο $E' = t_0^{-1}S$, δεν έχουμε παρά να κάνουμε $|S| \leq p = \mathcal{O}(l)$ πολλαπλασιασμούς. Τελικά, το κόστος του βήματος 4 είναι της τάξεως του $\mathcal{O}(l^5)$.

- Για το βήμα 5, δεν έχουμε παρά να εξετάσουμε $\forall t \in \mathrm{A}$, δηλαδή $\mathcal{O}(|\mathrm{A}| = \mathcal{O}(l))$ βήματα, αν $t \pmod p \in E'$, δηλαδή $\mathcal{O}(|E'|) = \mathcal{O}(|S|) = \mathcal{O}(p) =$

$\mathcal{O}(l^2)$ συγκρίσεις, άρα συνολικά $\mathcal{O}(l^3)$ βήματα, όπου και πάλι δεχόμαστε ότι η πράξη $a \bmod b$ έχει κόστος 1.

Τελικά, για το συνολικό κόστος του αλγορίθμου μας, δεν έχουμε παρά να αθροίσουμε τα επιμέρους κόστη των βημάτων 1 έως 5 και συγκεκριμένα:

$$\mathcal{O}(l^3) + \mathcal{O}(l^3) + \mathcal{O}(l^3) + \mathcal{O}(l^5) + \mathcal{O}(l^3) = \mathcal{O}(l^5)$$

το οποίο προφανώς μας δίνει ένα πολυωνυμικό φράγμα ως προς το μέγεθος της εισόδου.

Κεφάλαιο 3

Κάτω φράγματα για την συνάρτηση PARITY

3.1 Κυκλώματα AND-OR

Λέγοντας συνάρτηση *parity*, **PARITY**, εννοούμε την άπειρη οικογένεια συναρτήσεων

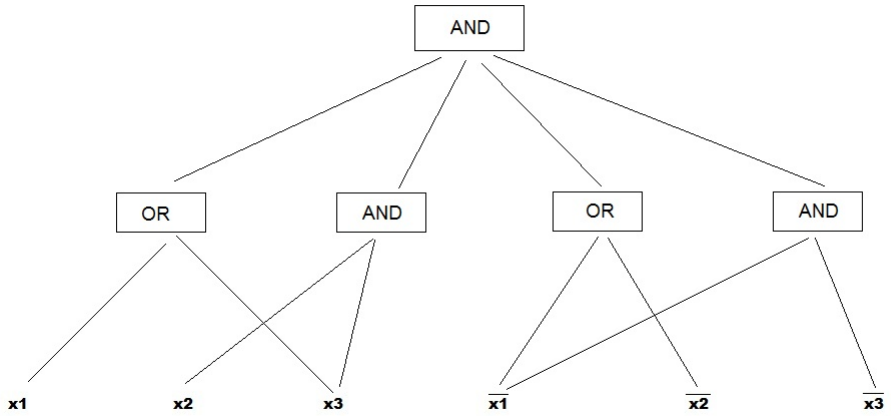
$$par_n : \{0,1\}^n \to \{0,1\},\ n = 1, 2, 3, \ldots$$

με

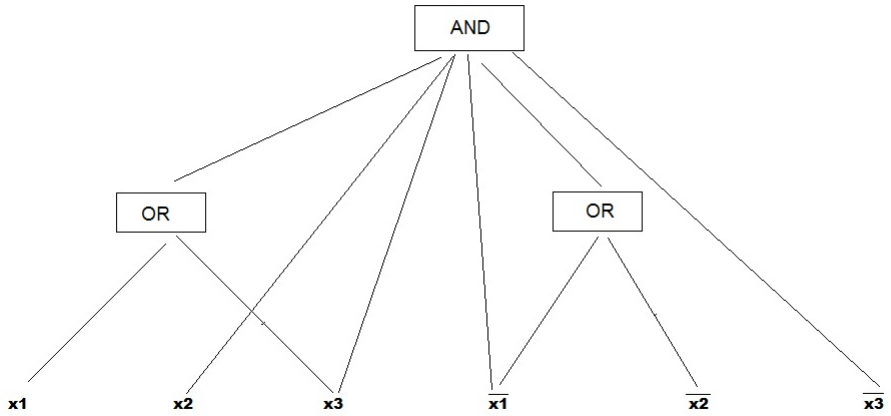
$$par_n(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} x_i\right) \bmod 2$$

Δεδομένου τώρα ότι αρκετές συναρτήσεις μπορούν να εκφραστούν με τη βοήθεια της συνάρτησης *parity*, εύκολα μπορεί κανείς να κατανοήσει το ενδιαφέρον για την κατασκευή όσο το δυνατόν «μικρότερου» συνδυαστικού κυκλώματος το οποίο να υλοποιεί την παραπάνω συνάρτηση.

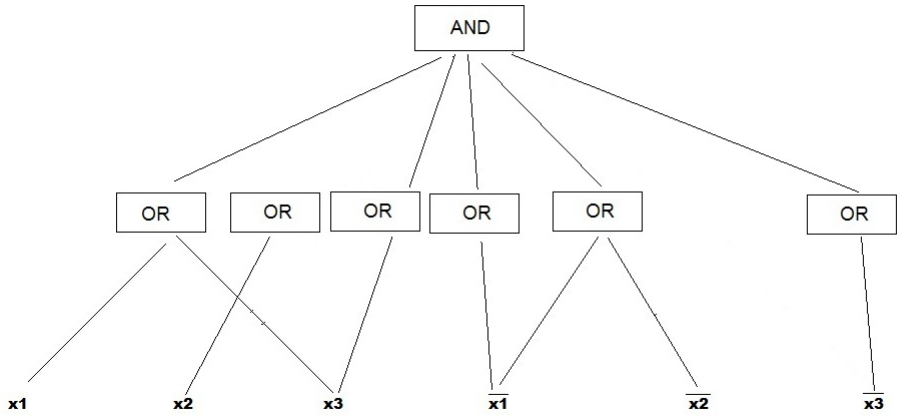
Στην παρούσα φάση, με τον όρο «κύκλωμα», θα αναφερόμαστε σε κυκλώματα τα οποία αποτελούνται μόνο από πύλες **AND** και **OR** με απεριόριστο εύρος εισόδου και όπου τα δεδομένα θα αποτελούνται από τις μεταβλητές x_i ή τις αρνήσεις τους $\overline{x_i}$.



Πύλες του ίδιου τύπου, όπου η μία συνδέεται απευθείας με την άλλη, στο σχήμα μπορούν να συγχωνευθούν σε μία χωρίς να επέλθει αλλαγή στη συνάρτηση που υλοποιείται από το κύκλωμα.



Έτσι λοιπόν, τέτοιου τύπου κυκλώματα, μπορούν να μετασχηματιστούν σε ισοδύναμα κυκλώματα, απλά προσθέτοντας πύλες με εύρος εισόδου 1:



Έχοντας κάνει λοιπόν τις παραπάνω μετατροπές στο κύκλωμά μας, έχουμε καταλήξει σε ένα κύκλωμα το οποίο έχει στο πρώτο επίπεδο μόνο πύλες **OR**, ενώ στο δεύτερο επίπεδο μόνο πύλες **AND**. Μπορούμε να συνεχίσουμε όμοια ώστε τελικά να εναλλασσόμαστε μεταξύ επιπέδων τα οποία αποτελούνται μόνο από ένα είδος πύλης, **OR** ή **AND**.

3.2 Κάποια «εύκολα» φράγματα

Κάνουμε τώρα μερικές παρατηρήσεις σχετικά με τα κυκλώματα μας, στις οποίες θα στηριχτούμε αργότερα για να δείξουμε τα φράγματά μας.

Για αρχή μπορούμε εύκολα να δούμε ότι οποιαδήποτε συνάρτηση *Boole* n μεταβλητών, μπορεί να υλοποιηθεί από ένα κύκλωμα 2 επιπέδων. Και αυτό γιατί μπορούμε να μετατρέψουμε τη δεδομένη συνάρτηση σε ισοδύναμη κανονική

συζευκτική μορφή είτε σε κανονική διαζευκτική μορφή και άρα προκύπτει ένα κύκλωμα $\mathbf{AND-OR}$ βάθους 2 στην πρώτη περίπτωση, είτε ένα κύκλωμα $\mathbf{OR-AND}$ και πάλι βάθους 2 στην δεύτερη περίπτωση. Για παράδειγμα, αν υποθέσουμε ότι είχαμε την εξής συνάρτηση

$$f = (x_1 \wedge x_2) \vee (x_1 \wedge x_3) \wedge (x_4 \vee x_5),$$

θα μπορούσε να μετασχηματιστεί ως εξής:

$f = (x_1 \vee x_1) \wedge (x_2 \vee x_1) \wedge (x_1 \vee x_3) \wedge (x_2 \vee x_3) \wedge (x_4 \vee x_5)$ (κανονική συζευκτική μορφή)

είτε

$f = (x_1 \wedge x_2) \vee (x_1 \wedge x_3 \wedge x_4) \vee (x_1 \wedge x_3 \wedge x_5)$ (κανονική διαζευκτική μορφή)

Γενικά όμως, υπάρχει και περίπτωση εκθετικής αύξησης των εμφανίσεων των μεταβλητών, το οποίο και φαίνεται στο παρακάτω παράδειγμα:

$$f = (x_1 \wedge y_1) \vee (x_2 \wedge y_2) \vee \cdots \vee (x_n \wedge y_n)$$

η οποία μετασχηματίζεται σε κανονική συζευκτική μορφή

$$f = (x_1 \vee \cdots \vee x_{n-1} \vee x_n) \wedge (x_1 \vee \cdots \vee x_{n-1} \vee y_n) \wedge \cdots \wedge (y_1 \vee \cdots \vee y_{n-1} \vee y_n)$$

με συνέπεια οι n διαζεύξεις να αυξηθούν σε 2^n συζεύξεις.

Έστω τώρα, ότι έχουμε ένα κύκλωμα βάθους d και μεγέθους s, το οποίο υλοποιεί την par_n. Τότε, έχοντας κατά νου τους κανόνες $DeMorgan$

$$\overline{x \vee y} = \overline{x} \wedge \overline{y}$$

$$\overline{x \wedge y} = \overline{x} \vee \overline{y}$$

για να πάρουμε ένα κύκλωμα που να υλοποιεί την $\overline{par_n}$ δεν έχουμε παρά να αντικαταστήσουμε τις εισόδους x_i με $\overline{x_i}$ και αντίστροφα και όλες τις πύλες **AND** με πύλες **OR** και αντίστροφα. Παρατηρούμε ότι το προκύπτον κύκλωμα έχει το ίδιο βάθος και μέγεθος με το αρχικό.

Μία ακόμη ιδιότητα που μπορούμε εύκολα να ελέγξουμε την ισχύ της, είναι πως δεδομένου κυκλώματος υπολογισμού της par_n και με το να θέσουμε κάποιες από τις μεταβλητές μας ίσες με 1 και κάποιες άλλες ίσες με 0, το κύκλωμα δεν υπολογίζει τίποτε άλλο παρά την συνάρτηση $parity$ ή την αντίθετή της σε λιγότερες πλέον μεταβλητές. Για να πειστεί κανείς, δεν έχει παρά να θυμηθεί τον ορισμό:

$$par_n(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} x_i \right) \bmod 2$$

και άρα αν υποθέσουμε ότι

- $x_1 = 0$ τότε,

  $$par_n(x_1, \ldots, x_n) = par_n(0, x_2, \ldots, x_n) = par_{n-1}(x_2, \ldots, x_n)$$

  ενώ αν

- $x_1 = 1$ τότε,

$$par_n(x_1, \ldots, x_n) = par_n(1, x_2, \ldots, x_n) = 1 + par_{n-1}(x_2, \ldots, x_n) = \overline{par_{n-1}}$$

Εργαζόμενοι όμοια, για οποιοδήποτε υποσύνολο του $\{x_i\}_{i=1}^n$ και οποιαδήποτε τιμοδοσία $\in \{0, 1\}$, βλέπουμε ότι πράγματι και πάλι υπολογίζουμε είτε την *parity* είτε την $\overline{parity}$, σε λιγότερες όμως μεταβλητές.

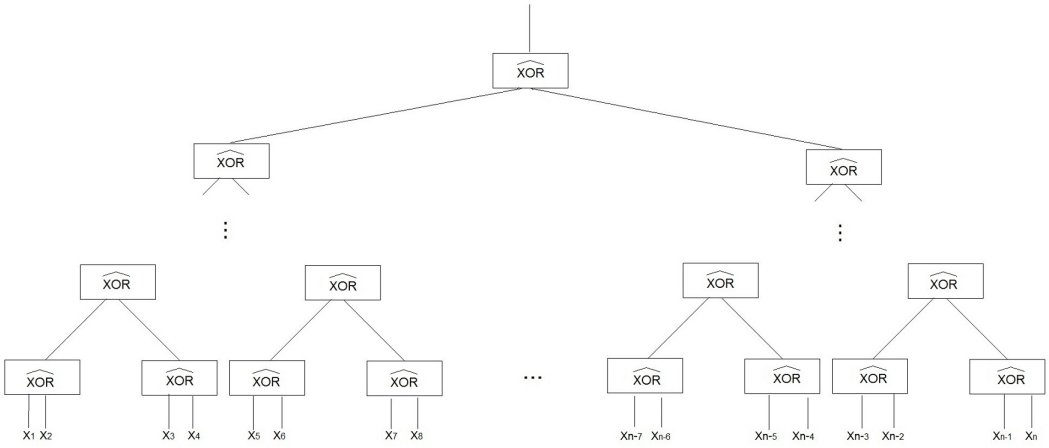
Θέλουμε τώρα να δούμε, εάν η **PARITY** μπορεί να υλοποιηθεί μέσω κυκλωμάτων *σταθερού βάθους* και *πολυωνυμικού μεγέθους*. Θα θέλαμε δηλαδή να εξετάσουμε, εάν υπάρχει σταθερά d και πολυώνυμο p, τέτοια ώστε για όλα τα n η συνάρτηση par_n να μπορεί να υλοποιηθεί μέσω κυκλώματος βάθους d και μεγέθους το πολύ $p(n)$.

Θεώρημα 3 *Η* **PARITY** *δεν μπορεί να υλοποιηθεί μέσω* **OR** $-$ **AND** *κυκλώματος πολυωνυμικού μεγέθους και βάθους* 2.

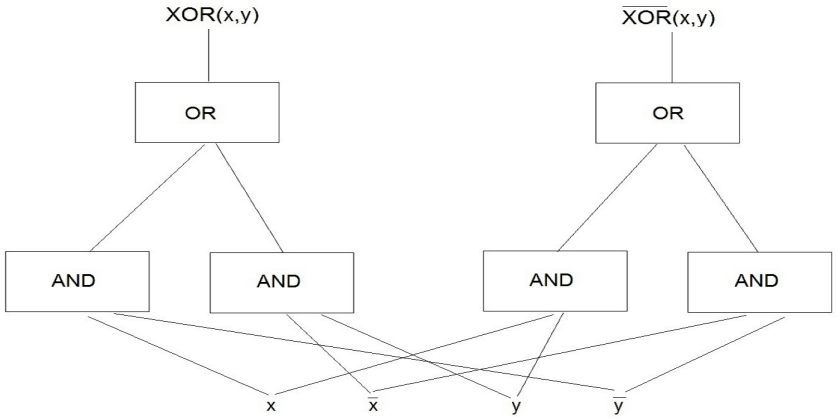
Απόδειξη:

Ας θυμηθούμε για αρχή, ότι αν έχουμε $x_1, \ldots, x_n$ μεταβλητές, τότε στο κύκλωμα μπορούν να εμφανιστούν και τα $\overline{x_1}, \ldots, \overline{x_n}$, άρα συνολικά $2n$ το πλήθος δεδομένα. Ξεκινάμε δείχνοντας ότι κάθε πύλη **AND** στο πρώτο επίπεδο, θα πρέπει αναγκαστικά να έχει *ακριβώς* n δεδομένα ως είσοδο και πιο συγκεκριμένα, σε κάθε πύλη θα εμφανίζεται είτε το x_i είτε το $\overline{x_i}$. Υποθέτουμε, για να καταλήξουμε σε αντίφαση, ότι σε κάποια πύλη *δεν* εμφανίζεται κάποια μεταβλητή x_j αλλά *ούτε* και η άρνησή της $\overline{x_j}$. Παρατηρούμε τώρα, ότι με κατάλληλη ανάθεση τιμών στις μεταβλητές ή στις αρνήσεις τους, που εμφανίζονται ως είσοδοι στην πύλη αυτή, μπορούμε να πάρουμε ως αποτέλεσμα την τιμή 1 από την πύλη **AND** και συνεπώς και την τιμή 1 ως αποτέλεσμα από όλο το κύκλωμα. Τώρα, ότι τιμή και να θέσουμε στη μεταβλητή x_j, το αποτέλεσμα του κυκλώματος παραμένει ανεπηρέαστο, πράγμα το οποίο δείχνει ότι το κύκλωμά μας δεν υπολογίζει την par_n, άτοπο. Επίσης, δεν μπορούν να υπάρξουν πύλες με είσοδο τόσο τη μεταβλητή x_j όσο και τη μεταβλητή $\overline{x_j}$, διότι τότε το αποτέλεσμα της πύλης **AND** θα είναι πάντα 0. Άρα οι πύλες **AND** στο πρώτο επίπεδο, θα έχουν *ακριβώς* n δεδομένα ως είσοδο. Όμως αυτό σημαίνει, πώς κάθε πύλη **AND** μπορεί να συλλάβει μόνο μία γραμμή από τον πίνακα αληθείας της par_n. Από τη στιγμή όμως που υπάρχουν 2^{n-1} γραμμές στον πίνακα αληθείας, με την τιμή 1, συμπεραίνουμε ότι απαιτούνται τελικά 2^{n-1} πύλες **AND** και εδώ ολοκληρώνεται η απόδειξή μας.

Μπορούμε όμως να δείξουμε ότι η **PARITY** μπορεί να υλοποιηθεί μέσω πολυωνυμικού μεγέθους κυκλώματος, βάθους $\mathcal{O}\big(\log(n)\big)$. Για ευκολία στους υπολογισμούς, υποθέτουμε ότι το πλήθος των μεταβλητών είναι δύναμη του 2, δηλαδή $n = 2^k$ για κάποιο $k \in \mathbb{N}$. Χάριν κομψότητας και απλότητας, θα χρησιμοποιήσουμε στο κύκλωμά μας πύλες $\widehat{\textbf{XOR}}$, των οποίων τη λειτουργία θα εξηγήσουμε αμέσως παρακάτω. Σχηματικά, το κύκλωμα μας παρουσιάζει τη μορφή «ισορροπημένου» δυαδικού δένδρου:



Είναι φανερό, ότι το βάθος του κυκλώματος, για είσοδο n μεταβλητών, είναι της τάξεως $\mathcal{O}(\log(n))$, ενώ το μέγεθος είναι της τάξεως $\mathcal{O}(n\log(n)) = \mathcal{O}(n^2)$ δηλαδή πολυωνυμικό, ως προς τις πύλες $\widehat{\mathbf{XOR}}$. Μένει λοιπόν να αναλύσουμε τον τρόπο λειτουργίας των πυλών αυτών. Σχηματικά, παρουσιάζουν την εξής δομή:



Παρατηρούμε, ότι αποτελούνται μόνο από πύλες **OR** και **AND** και μάλιστα με εύρος εισόδου 2. Το γεγονός επίσης ότι κάθε πύλη $\widehat{\mathbf{XOR}}$ αποτελείται από 6 πύλες, 4 πύλες **OR** και 2 πύλες **AND** μας επιτρέπει να πούμε πως τελικά το συνολικό κύκλωμα έχει μέγεθος της τάξης $\mathcal{O}(n^2)$, ως προς πύλες **OR** και **AND**, και βάθος $\mathcal{O}(n\log(n)) = \mathcal{O}(n^2)$. Παρατηρούμε επίσης, πως η κάθε πύλη $\widehat{\mathbf{XOR}}$, προκειμένου να υλοποιήσει την πράξη $\mathbf{XOR}(x, y)$, χρειάζεται τόσο τις μεταβλητές x, y όσο και τις αρνήσεις τους, $\overline{x}$ & $\overline{y}$, και αυτός είναι και ο λόγος που ταυτόχρονα με κάθε πύλη **XOR** υλοποιούμε και μία πύλη $\overline{\mathbf{XOR}}$.

3.3 PARITY $\notin AC^0$

Συμβολισμός: Με $\mathbf{AC^k}$ θα συμβολίζουμε την κλάση όλων των συναρτήσεων *Boole*, οι οποίες μπορούν να υλοποιηθούν μέσω κυκλωμάτων πολυωνυμικού μεγέθους και $\mathcal{O}(\log^k(n))$ βάθους, αποτελούμενα από πύλες **OR** και **AND** απεριόριστου εύρους εισόδου. Μόλις πριν δείξαμε ότι

$$\mathbf{PARITY} \in \mathbf{AC^1}.$$

Σκοπός μας όμως είναι να δείξουμε ότι

$$\mathbf{PARITY} \notin \mathbf{AC^0}.$$

Με το *Θεώρημα (3)*, έχουμε ήδη κάνει το πρώτο μας βήμα, το οποίο θα αποτελέσει και τη βάση της επαγωγής μας. Η απόδειξη αποτελεί έργο των Furst, Saxe & Sipser [14], οι οποίοι έκαναν χρήση μιας νέας τότε τεχνικής, με το όνομα «τυχαίοι περιορισμοί» και η οποία έχει έκτοτε χρησιμοποιηθεί αρκετές φορές. Το αποτέλεσμα αυτό βελτιώθηκε αργότερα, από «μη πολυωνυμικό μέγεθος» σε «τουλάχιστον εκθετικό», αρχικά από τον Yao [15] και αργότερα από τον Håstad [16], του οποίου μάλιστα η απόδειξη θεωρείται πολύ πρωτοποριακή.

Εδώ όμως θα αναφερθούμε στην ασθενέστερη εκδοχή των Furst, Saxe & Sipser η οποία έχει το πλεονέκτημα να είναι πιο απλή και αποτελεί και μια καλή ευκαιρία να εξοικειωθούμε με τους «τυχαίους περιορισμούς». Πριν προχωρήσουμε όμως, ας θυμηθούμε μερικά πράγματα από Θεωρία Πιθανοτήτων:

- *Διωνυμική Κατανομή* Όταν διεξάγουμε ένα πείραμα με 2 πιθανά αποτελέσματα, έστω επιτυχία και αποτυχία, τα οποία συμβαίνουν με πιθανότητες p και $1-p$ αντίστοιχα, τότε η πιθανότητα να έχουμε k ακριβώς επιτυχίες, σε n ανεξάρτητες επαναλήψεις είναι

  $$\binom{n}{k} p^k q^{n-k}$$

  Αν X η τυχαία μεταβλητή που μετράει το πλήθος επιτυχιών σε n δοκιμές, με πιθανότητες επιτυχίας και αποτυχίας όπως και πριν, τότε

  $$\mathbb{E}[X] = np \text{ και } Var[X] = np(1-p).$$

- *Ανισότητα Chebyshev* Ξεκινάμε από την ανισότητα του Markov η οποία λέει πως αν X τυχαία, μη αρνητική μεταβλητή, με πεπερασμένη μέση τιμή και $t > 0,\ t \in \mathbb{R}$ τότε

  $$\mathbb{P}[X \geq t] \leq \frac{\mathbb{E}[X]}{t}$$

  Για να πάρουμε την επιθυμητή ανισότητα δεν έχουμε παρά να θεωρήσουμε ως X την $(X - \mathbb{E}[X])^2$ και ως t τον αριθμό t^2 και άρα παίρνουμε πλέον ως αποτέλεσμα

  $$\mathbb{P}[(X - \mathbb{E}[X])^2 \geq t^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{t^2}$$

Παρατηρούμε όμως ότι

$$(X - \mathbb{E}[X])^2 \geq t^2 \Leftrightarrow |X - \mathbb{E}[X]| \geq t$$

και άρα σε συνδυασμό με το ότι $\mathbb{E}[(X - \mathbb{E}[X])^2] = Var[X]$ καταλήγουμε στο εξής:

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq t] \leq \frac{Var[X]}{t^2}$$

- *Ένα διαφορετικό φράγμα για τη διωνυμική κατανομή:* $\mathbb{P}[X \geq a] \leq p^a 2^n$

$$\mathbb{P}[X \geq a] = \sum_{i=a}^{n} \binom{n}{i} p^i (1-p)^{n-i} \leq \sum_{i=a}^{n} \binom{n}{i} p^i \leq p^a \sum_{i=a}^{n} \binom{n}{i} \leq p^a 2^n$$

Προφανώς η ανισότητα αυτή είναι χρήσιμη μόνο εάν $p^a 2^n \leq 1$.

Επιστρέφουμε πάλι στο κυρίως μέρος της απόδειξης και για να δείξουμε ότι η **PARITY** δεν μπορεί να υλοποιηθεί από κύκλωμα πολυωνυμικού μεγέθους και σταθερού βάθους, είναι αρκετό να δείξουμε τον εξής ισχυρισμό:

Ισχυρισμός 1 *$\forall t$ $\forall c$ $\forall$πολυώνυμο p, η* **PARITY** *δεν μπορεί να υλοποιηθεί μέσω κυκλώματος μεγέθους $p(n)$ και βάθους t το οποίο να έχει εύρος εισόδου $\leq c$, φραγμένο δηλαδή εύρος στο αρχικό επίπεδο.*

Θεώρημα 4 **PARITY** $\notin$ **AC^0**

Το τελευταίο θεώρημα είναι άμεση συνέπεια του *Ισχυρισμού 1*. Για να το δει κανείς, δεν έχει παρά να υποθέσει το αντίθετο, ότι δηλαδή **PARITY** $\in$ **AC^0**. Αυτό θα σήμαινε ότι υπάρχει οικογένεια κυκλωμάτων, πολυωνυμικού μεγέθους, και σταθερού βάθους t. Εάν τώρα ενσωματώσουμε ένα ακόμη επίπεδο, πριν το πρώτο επίπεδο του κυκλώματος, αποτελούμενο από πύλες είτε **OR** είτε **AND**, με τρόπο τέτοιο ώστε κάθε μεταβλητή που εμφανιζόταν ως είσοδο στο αρχικό κύκλωμα, να αποτελεί πλέον τη μία και μοναδική είσοδο σε ακριβώς μία από τις πύλες του «νέου» επιπέδου, και απο κει να εμφανίζεται στις πύλες όπως είχαμε αρχικά υποθέσει, θα έχουμε καταφέρει να κατασκευάσουμε ένα νέο κύκλωμα, βάθους $t + 1$, πολυωνυμικού και πάλι μεγέθους, με σταθερό εύρος εισόδου 1 στο αρχικό επίπεδο, πράγμα το οποίο έρχεται σε αντίθεση με τον *Ισχυρισμό 1*.

Θα πρέπει λοιπόν τώρα να αποδείξουμε τον *Ισχυρισμό 1*. Έχουμε ήδη δείξει νωρίτερα, στο **Θεώρημα 3**, ότι πράγματι στην περίπτωση όπου το βάθος t του κυκλώματος είναι ίσο με 2, τότε δεν υπάρχει κύκλωμα πολυωνυμικού μεγέθους, οποιουδήποτε εύρους εισόδου, το οποίο να υλοποιεί την *parity*.

Για τη γενική περίπτωση, υποθέτουμε ότι ο *Ισχυρισμός 1* είναι αναληθής. Άρα υπάρχει κάποιο $t > 2$, τέτοιο ώστε η **PARITY** να υλοποιείται από κύκλωμα πολυωνυμικού μεγέθους, βάθους t και με φραγμένο, από σταθερά, εύρος εισόδου στο αρχικό επίπεδο. Έστω λοιπόν ο ελάχιστος τέτοιος αριθμός t και έστω k αριθμός, αυστηρά μεγαλύτερος από το βαθμό του πολυωνύμου που

φράσσει το μέγεθος του κυκλώματος και c η σταθερά που φράσσει το εύρος εισόδου στο αρχικό επίπεδο. Θα δείξουμε ότι βάσει του υποθετικού κυκλώματος, με τις παραπάνω ιδιότητες, μπορούμε να κατασκευάσουμε ένα άλλο κύκλωμα, που να υλοποιεί την *parity*, πολυωνυμικού πάλι μεγέθους, με φραγμένο εύρος εισόδου αλλά πλέον με βάθος $t-1$, το οποίο θα έρχεται σε αντίθεση με την ελαχιστότητα του t και άρα θα έχουμε αποδείξει την ορθότητα του *Ισχυρισμού 1*.

Για την κατασκευή των νέων κυκλωμάτων, θα ακολουθήσουμε την εξής στρατηγική: Έστω $S = S_1, S_2, \ldots$ η οικογένεια κυκλωμάτων βάθους t που υπολογίζει την **PARITY**. Το κύκλωμα S_n^r, βάθους πλέον $t-1$, θα κατασκευάζεται παίρνοντας ένα στοιχείο της ακολουθίας S με δείκτη μεγαλύτερο του n, ας πούμε εδώ για παράδειγμα $4n^2$, και δίνοντας τότε τιμές στις «κατάλληλες» $(4n^2 - n)$ μεταβλητές μέσω της τυχαίας τιμοδοσίας r. Τότε το κύκλωμά μας θα εξαρτάται πλέον από n μεταβλητές. Επίσης, θα κατασκευάσουμε με τέτοιο τρόπο το κύκλωμα, ώστε οι πύλες του δεύτερου επιπέδου να έχουν φραγμένο εύρος εισόδου μετά τον περιορισμό, ώστε όταν εφαρμόσουμε τους κανόνες επιμερισμού για να αλλάξουμε τις θέσεις των επιπέδων 1 και 2, να μην προκύψει εκθετική αύξηση στο κύκλωμα, όπως αναμένεται να συμβεί στη γενική περίπτωση. Επίσης το κύκλωμά μας, μετά την εναλλαγή πρώτου και δεύτερου επιπέδου, θα έχει πάλι φραγμένο εύρος εισόδου στο «νέο» αρχικό επίπεδο και επιπλέον το δεύτερο και τρίτο επίπεδο από τη στιγμή που θα αποτελούνται από πύλες της ίδιας λειτουργίας, θα μπορούν να συγχωνευθούν σε ένα επίπεδο, οδηγώντας σε κύκλωμα βάθους $t-1$. Το μέγεθος του S_n^r θα είναι τετραγωνικό ως προς το μέγεθος του αρχικού κυκλώματος, $\mathcal{O}\left((4n^2)^k\right) = \mathcal{O}(n^{2k})$ και συνεπώς ο βαθμός του πολυωνύμου που θα φράσσει το μέγεθος θα διπλασιαστεί.

Υπάρχει όμως το πρόβλημα του πώς θα επιλέξουμε τις «κατάλληλες» μεταβλητές στις οποίες θα αναθέσουμε τιμές, και εδώ βρίσκεται η βασική ιδέα της απόδειξης: Δοκιμάζουμε έναν **«τυχαίο περιορισμό»** ή αλλιώς μία **«τυχαία αντικατάσταση»**. Εάν μπορέσουμε να δείξουμε ότι η πιθανότητα του να πάρουμε έναν κατάλληλο περιορισμό, δηλαδή ένα κατάλληλο κύκλωμα μετά την εφαρμογή του περιορισμού, είναι θετική, τότε μπορούμε να συμπεράνουμε ότι υπάρχει πράγματι ένας τέτοιος. Θα εφαρμόσουμε τον εξής *«τυχαίο περιορισμό»*: Για κάθε μεταβλητή x_i, ανεξάρτητα από τις άλλες, εκτελούμε το ακόλουθο πείραμα, με τα εξής 3 πιθανά αποτελέσματα:

- Με πιθανότητα $\frac{1}{\sqrt{n}}$ η μεταβλητή x_i παραμένει μεταβλητή
- Με πιθανότητα $\frac{1-\frac{1}{\sqrt{n}}}{2}$ η μεταβλητή x_i παίρνει την τιμή 0
- Με πιθανότητα $\frac{1-\frac{1}{\sqrt{n}}}{2}$ η μεταβλητή x_i παίρνει την τιμή 1

Η λογική με την οποία ορίστηκαν οι παραπάνω πιθανότητες είναι η εξής: Για αρχή, οι πιθανότητες σύμφωνα με τις οποίες η μεταβλητή x_i παίρνει κάποια τιμή $\in \{0, 1\}$, αναμένεται να είναι ίσες, ώστε αν αναγκαστούμε στο επιχείρημα που θα ακολουθήσει, να μπορέσουμε να αλλάξουμε τους ρόλους των πυλών

AND και **OR** και x_i με $\overline{x_i}$ καθώς και 1 με 0, χωρίς να χρειαστεί να προβούμε σε νέους υπολογισμούς. Επίσης, αποδεικνύεται χρήσιμο πρακτικά η πιθανότητα μιας μεταβλητής να παραμείνει μεταβλητή, να είναι όσο το δυνατόν μικρότερη και ειδικότερα για λόγους που θα φανερωθούν αργότερα, πολυωνυμικά μικρότερη από το n.

Έστω r, ο συμβολισμός του *«τυχαίου περιορισμού»* και ${x_i}^r$ το αποτέλεσμα του x_i μετά την εφαρμογή του περιορισμού έτσι ώστε ${x_i}^r \in \{0, 1, x_i\}$ και έστω S_n^r το κύκλωμα S_n, μετά τον περιορισμό. Έχουμε ήδη δείξει ότι το S_n^r θα υλοποιεί και αυτό είτε τη συνάρτηση $parity$ είτε τη συνάρτηση $\overline{parity}$ σε λιγότερες πλέον μεταβλητές. Σε κάθε περίπτωση πάντως, έχουμε και πάλι ήδη δείξει, ότι υπάρχει κύκλωμα ίδιου βάθους και μεγέθους, που να υλοποιεί την $parity$. Χρησιμοποιώντας Θεωρία Πιθανοτήτων, το αναμενόμενο πλήθος μεταβλητών X στο κύκλωμα S_n^r είναι:

$$\mathbb{E}[X_{S_n^r}] = n\mathbb{E}[X_{1_{S_n^r}}] = n\frac{1}{\sqrt{n}} = \sqrt{n}$$

και

$$Var[X_{S_n^r}] = nVar[X_{1_{S_n^r}}] = n\frac{1}{\sqrt{n}}\left(1 - \frac{1}{\sqrt{n}}\right) \leq \sqrt{n}$$

Τώρα φαίνεται ότι το πλήθος X των μεταβλητών που παραμένουν μεταβλητές στο κύκλωμα S_n μετά τον περιορισμό, πρέπει να εκφράζεται πολυωνυμικά ως προς το αρχικό πλήθος, ώστε να καταφέρουμε να δώσουμε ένα πολυωνυμικό άνω φράγμα μέσω της ανισότητας Chebyshev όπως και θα κάνουμε τώρα:

$$\begin{aligned}\mathbb{P}[X \leq \frac{\sqrt{n}}{2}] &= \mathbb{P}[X \leq \frac{\mathbb{E}[X]}{2}] \\ &\leq \mathbb{P}\left[|X - \mathbb{E}[X]| \geq \frac{\mathbb{E}[X]}{2}\right] \leq \frac{Var[X]}{\frac{\mathbb{E}^2[X]}{4}} \\ &\leq 4\frac{1}{\sqrt{n}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\end{aligned}$$

Συνεπώς, για n αρκετά μεγάλο, η πιθανότητα αυτή τείνει στο 0, δηλαδή

$$\mathbb{P}[X \leq \frac{\sqrt{n}}{2}] \to 0\ , \quad n \to \infty$$

Με άλλα λόγια, με μεγάλη πιθανότητα, θα παραμείνουν τουλάχιστον $\frac{\sqrt{n}}{2}$ μεταβλητές στο κύκλωμα S_n^r μετά τον περιορισμό.

Κανείς μπορεί να αντιτείνει σε αυτό το σημείο, ότι δεν είναι καθόλου ξεκάθαρο το πώς θα παραχθεί μία ακολουθία κυκλωμάτων $S^r = S_1^r, S_2^r, \ldots$ χωρίς κενά. Αυτό που γνωρίζουμε στα σίγουρα, είναι ότι για δεδομένο n, αν εφαρμόσουμε τον περιορισμό στο κύκλωμα S_{4n^2}, θα πάρουμε ένα κύκλωμα με αναμενόμενο πλήθος μεταβλητών $2n$ και με μεγάλη πιθανότητα πάνω από n μεταβλητές.

$$\mathbb{E}[X_{S_{4n^2}^r}] = 2n\ , \quad Var[X] = 4n^2\frac{1}{\sqrt{n}}\left(1 - \frac{1}{\sqrt{n}}\right) \leq 4n\sqrt{n}$$

και

$$\mathbb{P}[X_{S^r_{4n^2}} \leq \frac{\sqrt{4n^2}}{2} = n] \leq \mathbb{P}[|X - \mathbb{E}[X]| \geq \frac{\mathbb{E}[X]}{2}] \leq \frac{4}{\sqrt{n}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

Άρα, σίγουρα υπάρχει ένα κύκλωμα m μεταβλητών με m: $n \leq m \leq 4n^2$. Θέτουμε λοιπόν τις επιπλέον $(m-n)$ μεταβλητές ίσες με 0 και πλέον έχουμε το επιθυμητό κύκλωμα n ακριβώς μεταβλητών.

Στη συνέχεια θα δώσουμε άνω φράγματα για τα ενδεχόμενα όπου το κύκλωμά μας θα έχει ανεπιθύμητες ιδιότητες. Εάν καταφέρουμε να δείξουμε ότι κάθε ένα από αυτά τα ενδεχόμενα, πεπερασμένα στο πλήθος, συμβαίνουν με πιθανότητα που τείνει στο 0 καθώς $n \to \infty$, τότε για μεγάλα τουλάχιστον n, μπορούμε να ισχυριστούμε ότι στο κύκλωμά μας δεν πραγματοποιείται κανένα απ' αυτά τα ενδεχόμενα. Πρώτα θα δείξουμε ότι με μεγάλη πιθανότητα οι πύλες στο επίπεδο 2, μετά την εφαρμογή του περιορισμού, θα εξαρτώνται μόνο από σταθερό πλήθος μεταβλητών. Για να το πετύχουμε αυτό, μπορούμε να υποθέσουμε ότι στο πρώτο επίπεδο έχουμε πύλες **OR** ενώ στο δεύτερο πύλες **AND**. Εάν η κατάσταση είναι ανεστραμμένη, μπορούμε να επαναλάβουμε το επιχείρημα μας αντιστρέφοντας τους ρόλους των πυλών **AND** και **OR**, 0 και 1, x_i και $\overline{x_i}$.

Για τον υπολογισμό των πιθανοτήτων των ενδεχομένων που θέλουμε να ελέγξουμε, θα θεωρήσουμε μία οποιαδήποτε, προκαθορισμένη πύλη **AND** στο δεύτερο επίπεδο και θα δούμε ότι ο *«τυχαίος περιορισμός»* έχει ανεπιθύμητες ιδιότητες, στην περίπτωση μας εξάρτηση από πολλές μεταβλητές, με πιθανότητα το πολύ $\mathcal{O}(\frac{1}{n^k})$. Όμως από τη στιγμή που υπάρχουν σ' όλο το κύκλωμα το πολύ $\mathcal{O}(n^{k-1})$ πύλες, συμπεραίνουμε ότι η πιθανότητα οποιαδήποτε πύλη **AND** στο δεύτερο επίπεδο, να εξαρτάται από πολλές μεταβλητές, είναι το πολύ $\mathcal{O}(\frac{1}{n^k})\mathcal{O}(n^{k-1}) = \mathcal{O}(\frac{1}{n})$. Άρα με μεγάλη πιθανότητα, **καμία** από τις πύλες **AND** στο δεύτερο επίπεδο δεν θα έχει αυτήν την ανεπιθύμητη ιδιότητα (εξάρτηση από πολλές μεταβλητές) μετά την εφαρμογή του *«τυχαίου περιορισμού»*. Χρησιμοποιώντας επαγωγή θα αποδείξουμε τον παρακάτω ισχυρισμό:

Ισχυρισμός 2 *Για κάθε* **AND** $-$ **OR** *κύκλωμα, το οποίο έχει εύρος εισόδου το πολύ c, και μέγεθος το πολύ $p(n)$, όπου $p(n)$ πολυώνυμο βαθμού $< k$, υπάρχει σταθερά $e = e(c,k)$, εξαρτώμενη δηλαδή μόνο από τα c και k, έτσι ώστε η πιθανότητα μετά τον «τυχαίο περιορισμό», κάποια προκαθορισμένη πύλη* **AND** *να εξαρτάται από περισσότερες από e μεταβλητές, να είναι το πολύ $\mathcal{O}(\frac{1}{n^k})$.*

Απόδειξη Ισχυρισμού 2

Όπως είπαμε και πριν, η απόδειξη θα γίνει με επαγωγή στο c. Η βάση της επαγωγής είναι για $c = 1$. Σε αυτήν την περίπτωση, δεν υπάρχουν πύλες **OR**, αφού τα δεδομένα «περνάνε» απευθείας στο επίπεδο 2 και άρα έχουμε μόνο την πύλη **AND**. Θα εξετάσουμε 2 περιπτώσεις, ανάλογα με το πόσο μεγάλο εύρος εισόδου έχει η πύλη **AND**.

- **Περίπτωση Β1**: Το εύρος της πύλης **AND** είναι τουλάχιστον $4k\ln(n)$.

  Σε αυτήν την περίπτωση είναι πολύ πιθανό, να υπάρχει τουλάχιστον μία μεταβλητή, η οποία μετά την εφαρμογή του «τυχαίου περιορισμού», να

έχει πάρει την τιμή 0 και άρα η πύλη **AND** δεν εξαρτάται από καμία μεταβλητή.

Απόδειξη Περίπτωσης Β1

$$\begin{aligned}
\mathbb{P}[\text{η πύλη } \mathbf{AND} \text{ δεν δίνει } 0] &= \mathbb{P}[\text{καμία μεταβλητή δεν έχει πάρει τιμή } 0] \\
&= \mathbb{P}[\text{μία συγκεκριμένη μεταβλητή δεν έχει πάρει τιμή } 0]^{\#\text{μεταβλητών}} \\
&\leq \left(1 - \frac{1 - 1/\sqrt{n}}{2}\right)^{4k\ln(n)} \\
&\leq \left(\frac{3}{4}\right)^{4k\ln(n)} \qquad (\text{ισχύει για } n \geq 4) \\
&= n^{4k\ln(\frac{3}{4})} \leq n^{4k(-\frac{1}{4})} = n^{-k}
\end{aligned}$$

όπου στην τελευταία γραμμή εφαρμόσαμε

$$a^{\ln(b)} = b^{\ln(a)} \text{ και } \ln(1-x) \leq -x$$

- **Περίπτωση Β2**: Το εύρος εισόδου της πύλης **AND** είναι μικρότερο από $4k\ln(n)$.

  Σε αυτήν την περίπτωση είναι πολύ πιθανό μετά την εφαρμογή του «τυχαίου περιορισμού», να παραμείνει μόνο σταθερό πλήθος μεταβλητών. Την διαίσθησή μας αυτήν, την ενισχύει τουλάχιστον το γεγονός ότι αν X είναι το πλήθος των μεταβλητών που παραμένουν μετά τον περιορισμό, τότε

$$\mathbb{E}[X] \leq 4k\ln(n)\frac{1}{\sqrt{n}} \to 0, \qquad n \to \infty$$

Απόδειξη Περίπτωσης Β2

$\mathbb{P}[\text{η πύλη } \mathbf{AND} \text{ εξαρτάται από περισσότερες από } m \text{ μεταβλητές}]=$

$$\begin{aligned}
\sum_{i=m+1}^{4k\ln(n)} \binom{4k\ln(n)}{i} \left(\frac{1}{\sqrt{n}}\right)^i \left(1 - \frac{1}{\sqrt{n}}\right)^{n-i} &\leq \sum_{i=m}^{4k\ln(n)} \binom{4k\ln(n)}{i} \left(\frac{1}{\sqrt{n}}\right)^i \\
&\leq \left(\frac{1}{\sqrt{n}}\right)^m \sum_{i=m}^{4k\ln(n)} \binom{4k\ln(n)}{i} \\
&\leq \left(\frac{1}{\sqrt{n}}\right)^m 2^{4k\ln(n)} \\
&\leq n^{-\frac{m}{2}} n^{4k\ln(2)} \\
&= n^{4k\ln(2) - \frac{m}{2}}
\end{aligned}$$

Αν τώρα θέσουμε $m = 2k\,[4\ln(2) + 1]$ παίρνουμε τελικά:
$\mathbb{P}[\text{η πύλη } \mathbf{AND} \text{ εξαρτάται από περισσότερες από } m \text{ μεταβλητές}]=\mathcal{O}(n^{-k})$

Έχουμε αποδείξει τη βάση της επαγωγής, όπου $e(1,k) = 2k\big[4\ln(2)+1\big]$. Υποθέτουμε λοιπόν τώρα, ότι ο **Ισχυρισμός (2)** ισχύει για εύρος εισόδου έως και $c-1$, δηλαδή για εύρος εισόδου $i, i \in \{1,2,\ldots,c-1\}$, υπάρχουν σταθερές $\{e_i\}_{i=1}^{c-1}$, τέτοιες ώστε η πιθανότητα μία προκαθορισμένη πύλη **AND** στο δεύτερο επίπεδο να εξαρτάται από περισσότερες από $e(i,k)$ μεταβλητές, είναι $\mathcal{O}(n^{-k})$.

Επαγωγικό Βήμα

- **Περίπτωση 1:** *Πριν τον «τυχαίο περιορισμό», η πύλη* **AND** *στο δεύτερο επίπεδο έχει τουλάχιστον* $4^c k\ln(n)$ *πύλες* **OR** *από κάτω της, με ξένες μεταβλητές.*

  Σε αυτήν την περίπτωση δηλαδή, υπάρχουν περισσότερες από $4^c k\ln(n)$ πύλες **OR** οι οποίες ανα δύο δεν έχουν κοινές μεταβλητές. Θα δείξουμε ότι μετά την εφαρμογή του «τυχαίου περιορισμού», είναι πολύ πιθανό σε μία από τις πύλες **OR**, όλες οι είσοδοι να έχουν πάρει την τιμή 0, το οποίο σημαίνει με τη σειρά του πως και το αποτέλεσμα της πύλης **AND** θα είναι 0. Δηλαδή στην περίπτωση αυτή, η πύλη **AND** δεν εξαρτάται από καμία μεταβλητή, οπότε ο ισχυρισμός μας έχει αποδειχθεί.

  Απόδειξη Περίπτωσης 1

$$\begin{aligned}
\mathbb{P}[\text{η πύλη } \mathbf{AND} \text{ δεν δίνει } 0] &= \mathbb{P}[\text{καμία από τις πύλες } \mathbf{OR} \text{ δεν δίνει } 0] \\
&\leq \mathbb{P}[\text{μία προκαθορισμένη πύλη } \mathbf{OR} \text{ δεν δίνει } 0]^{\#\text{πυλών}} \\
&= \big(1 - \mathbb{P}[\text{μία προκαθορισμένη πύλη } \mathbf{OR} \text{ δίνει } 0]\big)^{\#\text{πυλών}} \\
&\leq \left(1 - \big(\frac{1-\frac{1}{\sqrt{n}}}{2}\big)^c\right)^{4^c k\ln(n)} \\
&\leq \big(1-4^{-c}\big)^{4^c k\ln(n)} \\
&= n^{4^c k\ln(1-4^{-c})} \leq n^{4^c k(-4^{-c})} \\
&= n^{-k}
\end{aligned}$$

  όπου πάλι στην προτελευταία γραμμή, χρησιμοποιήσαμε το γεγονός ότι εαν $n \geq 4$ τότε

$$\frac{1-\frac{1}{\sqrt{n}}}{2} \geq \frac{1}{4} \Rightarrow \left(\frac{1-\frac{1}{\sqrt{n}}}{2}\right)^c \geq \left(\frac{1}{4}\right)^c$$

  και συνεπώς

$$1 - \left(\frac{1-\frac{1}{\sqrt{n}}}{2}\right)^c \leq 1 - \left(\frac{1}{4}\right)^c$$

- **Περίπτωση 2:** *Πριν τον «τυχαίο περιορισμό», η πύλη* **AND** *στο επίπεδο 2 έχει λιγότερες από* $4^c k\ln(n)$ *πύλες* **OR** *από κάτω της, με ξένες μεταβλητές.*

Απόδειξη Περίπτωσης 2

Σε αυτήν την περίπτωση, διαλέγουμε ένα μεγιστικό σύνολο πυλών **OR** με *ξένες μεταβλητές* και έστω H το σύνολο των μεταβλητών που εμφανίζονται στις πύλες αυτές. Από υπόθεση, το εύρος εισόδου φράσσεται από τη σταθερά c και άρα μπορούμε να συμπεράνουμε ότι $|H| \leq c4^c k \ln(n)$ και να πούμε με βεβαιότητα ότι σε κάθε πύλη **OR** του πρώτου επιπέδου, εμφανίζεται τουλάχιστον μία μεταβλητή από το σύνολο H, αλλιώς το σύνολο H δεν θα ήταν μεγιστικό. Υπάρχουν $l = 2^{|H|}$ δυνατά διανύσματα $\in \{0,1\}^{|H|}$, ανάλογα με την τιμή που μπορούμε να αναθέσουμε στην κάθε μεταβλητή. Αν τώρα επιλέξουμε ένα διάνυσμα από τα παραπάνω και το θεωρήσουμε ως είσοδο στο $\mathbf{AND-OR}$ κύκλωμά μας, τότε τουλάχιστον μία μεταβλητή σε κάθε πύλη **OR** παίρνει τιμή και παύει πλέον να αποτελεί μεταβλητή. Δηλαδή, το εύρος εισόδου των πυλών **OR** στο πρώτο επίπεδο, μειώνεται και φράσσεται πλέον από τη σταθερά $(c-1)$. Και τώρα μπορούμε πλέον να εφαρμόσουμε την επαγωγική υπόθεση. Έστω $A_1, A_2, \ldots, A_l$ τα l πιθανά κυκλώματα που μπορούν να προκύψουν με αυτόν τον τρόπο. Κάθε ένα δηλαδή από τα l διανύσματα «δίνει» διαφορετικό κύκλωμα. Σύμφωνα πάλι με την επαγωγική υπόθεση, η πιθανότητα ώστε η συνάρτηση του κυκλώματος A_j^r, δηλαδή το κύκλωμα A_j μετά την εφαρμογή του «τυχαίου περιορισμού», να εξαρτάται από περισσότερες από $e(c-1,k)$ μεταβλητές, είναι της τάξης $\mathcal{O}(n^{-k})$.

Η συνάρτηση $\mathbf{f}$, που υπολογίζει το κύκλωμά μας, μπορεί να εκφραστεί με τη βοήθεια των $\{A_j\}_{j=1}^{l}$. Για παράδειγμα, υποθέστε ότι $\mathrm{H} = \{x_1, x_2\}$. Άρα $l = 4$ και άρα

$$\mathbf{f} = (\overline{x_1} \wedge \overline{x_2} \wedge A_1) \vee (\overline{x_1} \wedge x_2 \wedge A_2) \vee (x_1 \wedge \overline{x_2} \wedge A_3) \vee (x_1 \wedge x_2 \wedge A_4)$$

Τώρα αντί να χρησιμοποιήσουμε το κύκλωμα ώστε να καθορίσουμε την εξάρτηση της $\mathbf{f}$ μετά τον *«τυχαίο περιορισμό»* από τις μεταβλητές, μας βολεύει περισσότερο να εργαστούμε με την ισοδύναμη αναπαράσταση με τη βοήθεια των $\{A_j\}_{j=1}^{l}$.

Με μεγάλη πιθανότητα μετά τον *«τυχαίο περιορισμό»*, το σύνολο H θα αποτελείται μόνο από **σταθερό** πλήθος μεταβλητών και άρα και το πλήθος των όρων (πλήθος διαζεύξεων) στην αναπαράσταση της $\mathbf{f}$, θα φράσσεται και αυτό από μία **σταθερά**. Αυτό συμβαίνει διότι, εάν υποθέσουμε ότι q από τις $|H|$ μεταβλητές πάρουν τιμή, είτε 0 είτε 1, τότε ακριβώς 1 από τις 2^q τιμοδοσίες δίνει τιμή 1 από τη σύζευξη αυτών των q μεταβλητών και άρα μένουν στην αναπαράσταση της $\mathbf{f}$, $2^{\#\text{μεταβλητών}}$ όροι που δεν είναι ταυτοτικά ίσοι με 0. Έστω λοιπόν X_H, η τυχαία μεταβλητή που μετράει το πλήθος των μεταβλητών του συνόλου H μετά την εφαρμογή του *«τυχαίου περιορισμού»*. Γνωρίζοντας ότι η X_H ακολουθεί διωνυμική κατανομή με

$p = \frac{1}{\sqrt{n}}$ και θέτοντας $4^c k = d$ υπολογίζουμε:

$$\begin{aligned}\mathbb{P}[X_H > a] &= \sum_{i=a}^{cd\ln(n)} \binom{cd\ln(n)}{i} \left(\frac{1}{\sqrt{n}}\right)^i \left(1 - \frac{1}{\sqrt{n}}\right)^{cd\ln(n)-i} \\ &\leq \left(\frac{1}{\sqrt{n}}\right)^a 2^{cd\ln(n)} \leq n^{-\frac{a}{2}} \cdot n^{cd} \\ &= n^{-\frac{a}{2}+cd}\end{aligned}$$

Αν θέσουμε $a = 2cd + 2k$ παίρνουμε τελικά

$$\mathbb{P}[X_H > 2cd + 2k] = \mathcal{O}(n^{-k}) \ .$$

Δηλαδή με μεγάλη πιθανότητα έχουμε ότι $X_H \leq 2cd + 2k$ και άρα σε αυτήν την περίπτωση, η αναπαράσταση της $\mathbf{f}$, αποτελείται το πολύ από m όρους, όπου $m := 2^{2cd+2k} \geq 2^{X_H}$.
Θέτουμε τώρα $e(c,k) = (2cd + 2k) + m \cdot e(c-1,k)$ και παίρνουμε:

$$\begin{aligned}&\mathbb{P}[\text{η } \mathbf{f} \text{ εξαρτάται από περισσότερες από } e(c,k) \text{ μεταβλητές}] \\ &\leq \mathbb{P}[X_H > 2cd + 2k] \\ &+ m \cdot \mathbb{P}[\text{ένα προκαθορισμένο } A_j \text{ εξαρτάται από περισσότερες από } e(c-1,k) \text{ μεταβλητές}] \\ &\leq \mathcal{O}(n^{-k}) \ + \ m \cdot \mathcal{O}(n^{-k}) \\ &= \mathcal{O}(n^{-k})\end{aligned}$$

αποτέλεσμα το οποίο ολοκληρώνει την απόδειξη του **Ισχυρισμού 2**.

Τώρα όμως είμαστε σε θέση να αποδείξουμε και τον **Ισχυρισμό 1**. Πρέπει να υπάρχει ένας «περιορισμός», ο οποίος «αφήνει» αρκετές μεταβλητές, αλλά τελικά οι πύλες στα επίπεδα 1 και 2 να εξαρτώνται μόνο από σταθερό πλήθος μεταβλητών. Έτσι, με σταθερό πλέον κόστος, μπορούμε να εναλλάξουμε τις θέσεις του πρώτου και του δεύτερου επιπέδου ώστε να έχουμε στο πρώτο επίπεδο πύλες **AND** και στο δεύτερο πύλες **OR**. Όμως τώρα, τα επίπεδα 2 και 3 αποτελούνται από πύλες **OR** και τα δύο οπότε και μπορούν να συγχωνευθούν σε ένα επίπεδο. τελικά, κατασκευάσαμε κύκλωμα πολυωνυμικού μεγέθους και πάλι, σταθερού εύρους στο επίπεδο 1, αλλά βάθους $t-1$, το οποίο έρχεται σε αντίθεση με την ελαχιστότητα του t.

Κεφάλαιο 4

Η συνάρτηση PARITY ξανά

4.1 PARITY $\notin AC^0$: Μία διαφορετική απόδειξη

Το προηγούμενο αποτέλεσμα των Furst, Saxe & Sipser, δηλαδή ότι η συνάρτηση *parity* δεν μπορεί να υπολογιστεί από κύκλωμα της κλάσης $\mathbf{AC^0}$ (σταθερό βάθος, πολυωνυμικό μέγεθος, απεριόριστο εύρος εισόδου), αποδείχθηκε αργότερα με εντελώς διαφορετικό τρόπο από τους A.Razborov [18] & R.Smolensky [19]. Το παρόν κεφάλαιο ασχολείται με αυτήν την απόδειξη και κάποια αποτελέσματα που πηγάζουν άμεσα από την εργασία τους. Η τεχνική που χρησιμοποίησαν είναι αλγεβρικής φύσεως, αλλά χρησιμοποιεί επίσης και πιθανοθεωρητικό επιχείρημα. Η βασική δομή της απόδειξης είναι:

1. Πρώτα δείχνουμε ότι κάθε συνάρτηση f που μπορεί να υλοποιηθεί από κύκλωμα της κλάσης $\mathbf{AC^0}$, μπορεί τότε να προσεγγιστεί από ένα πολυώνυμο p, μικρού βαθμού. Προσέγγιση στην περίπτωσή μας σημαίνει ότι για «αρκετές» n-άδες $(a_1, \ldots, a_n) \in \{0,1\}^n$ ισχύει:

$$f(a_1, \ldots, a_n) = p(a_1, \ldots, a_n)$$

2. Κατόπιν θα δείξουμε ότι η συνάρτηση *parity* δεν μπορεί να προσεγγιστεί με αυτήν την έννοια από πολυώνυμο μικρού βαθμού.

Ξεκινάμε λοιπόν με το πρώτο βήμα. Προφανώς η συνάρτηση $\mathbf{AND}(x_1, \ldots, x_n)$ μπορεί να αναπαρασταθεί μέσω του πολυωνύμου p

$$p = \prod_{i=1}^{n} x_i \ , \ x_i \in \{0,1\}$$

Η συνάρτηση **OR** είναι λιγότερο προφανής, αλλά αν θυμηθούμε τους κανόνες *DeMorgan* και ότι $\mathbf{NOT}(x_1) = 1 - x_1$, τότε

$$\mathbf{OR}(x_1, \ldots, x_n) = x_1 \vee \cdots \vee x_n = \overline{\overline{x_1} \wedge \cdots \wedge \overline{x_n}} = 1 - \prod_{i=1}^{n} (1 - x_i)$$

Το πρόβλημα τώρα είναι, ότι εν γένει τα πολυώνυμα έχουν βαθμό n και αυτό γιατί περιέχουν μονώνυμα στα οποία εμφανίζονται όλα τα x_i. Μπορούμε όμως να ξεπεράσουμε αυτό το πρόβλημα, χρησιμοποιώντας μία μέθοδο που οφείλεται στους Valiant & Vazirani [17]. Κατασκευάζουμε λοιπόν ένα τυχαίο πολυώνυμο ως εξής:

Έστω $S_0 = \{1, \ldots, n\}$. Επιπλέον, θεωρούμε τα τυχαία κατασκευασμένα σύνολα

$$S_{i+1} \subseteq S_i$$

έτσι ώστε

$$\forall j \in S_i \;\Rightarrow\; j \in S_{i+1} \text{ με πιθανότητα } 1/2.$$

Δηλαδή εάν έχουμε το σύνολο S_i, βρίσκουμε το S_{i+1}, ρίχνοντας από ένα αμερόληπτο νόμισμα για κάθε στοιχείο του S_i, ανεξάρτητα για όλα τα στοιχεία. Θεωρούμε τώρα την ακολουθία $S_0, S_1, \ldots, S_{\log(n)+2}$ κατασκευασμένη με τον τρόπο που μόλις περιγράψαμε και έστω q_i ο συμβολισμός για το τυχαίο πολυώνυμο

$$q_i(x) = \sum_{j \in S_i} x_j$$

το οποίο είναι προφανώς βαθμού 1. Διακρίνουμε 2 περιπτώσεις:

- $\mathbf{OR}(x_1, \ldots, x_n) = 0$ οπότε και $x_i = 0 \quad \forall i\ 1, \ldots, n$
  Άρα και όλα τα q_i έχουν την τιμή 0 και συνεπώς το πολυώνυμο

  $$1 - p(x_1, \ldots, x_n) \tag{4.1}$$

  με

  $$p = \prod_{i=0}^{\log(n)+2} (1 - q_i)$$

  ισούται με 0. Παρατηρήστε ότι το p έχει βαθμό $\mathcal{O}(\log(n))$.

- $\mathbf{OR}(x_1, \ldots, x_n) = 1$
  Δηλαδή υπάρχει ένα τουλάχιστον x_i με την τιμή 1. Θα δείξουμε, ότι σε αυτήν την περίπτωση, η πιθανότητα κάποιο q_i να έχει ακριβώς την τιμή 1 είναι μεγαλύτερη από $\frac{1}{2}$. Άρα

  $$\mathbb{P}[1 - p = 1] = \mathbb{P}\Big[1 - \prod_{i=0}^{\log(n)+2} (1 - q_i) = 1\Big] =$$

  $$\mathbb{P}[\prod_{i=0}^{\log(n)+2} (1 - q_i) = 0] = \mathbb{P}[\exists i : q_i = 1] \geq \frac{1}{2}$$

Μέχρι εδώ, αυτό που έχουμε δείξει είναι πως για κάθε $x = (x_1, \ldots, x_n)$ η πιθανότητα $\mathbb{P}[\mathbf{OR}(x) = 1 - p(x)]$ είναι μεγαλύτερη ή ίση από $\frac{1}{2}$. Βέβαια, το ζητούμενο, το οποίο και δείχνουμε λίγο παρακάτω, δεν είναι ακριβώς αυτό αλλά

πως το πλήθος των σημείων x στα οποία **OR** και $1-p$ συμφωνούν, είναι περισσότερα από τα μισά.

Άρα, το πολυώνυμο $(1-p)$ «προσεγγίζει» την συνάρτηση **OR** κατά μία έννοια. Μπορούμε όμως να το βελτιώσουμε σημαντικά, δημιουργώντας επιπλέον ανεξάρτητα πολυώνυμα p_k, έστω $p_1, \ldots, p_t$ και τότε να σχηματίσουμε το πολυώνυμο

$$p_1 \cdot \ \ldots \ \cdot p_t$$

βαθμού $\mathcal{O}(t\log(n))$ για την προσέγγισή μας. Η πιθανότητα λάθους βέβαια σε αυτήν την περίπτωση είναι $< (\frac{1}{2})^t$. Έτσι, εάν θέλουμε να εξασφαλίσουμε πως το περιθώριο λάθους της προσέγγισης της συνάρτησης

$$\mathbf{OR}(x_1, \ldots, x_n)$$

από το πολυώνυμο

$$\big(1-(p_1 \cdot \ \ldots \ \cdot p_t)\big)(x_1, \ldots, x_n)$$

θα είναι μικρότερο από ένα δεδομένο ϵ, τότε δεν έχουμε παρά να επιλέξουμε το t έτσι ώστε

$$\frac{1}{2^t} \leq \epsilon \Leftrightarrow \frac{1}{\epsilon} \leq 2^t \Leftrightarrow \log_2(1/\epsilon) \leq t. \tag{4.2}$$

Μπορούμε ακόμα να εκφράσουμε την συνάρτηση **AND** μέσω της συνάρτησης **OR**:

$$\mathbf{AND}(x_1, \ldots, x_n) = \mathbf{NOT}\left(\mathbf{OR}\big(\mathbf{NOT}(x_1), \ldots, \mathbf{NOT}(x_n)\big)\right)$$

δηλαδή τελικά

$$\mathbf{AND}(x_1, \ldots, x_n) \approx p(1-x_1, \ldots, 1-x_n) \tag{4.3}$$

Μένει όμως να δείξουμε ότι για κάθε επιλογή μη κενού συνόλου T, που αντιστοιχεί στο σύνολο των μεταβλητών x_j με τιμή 1, $T \subseteq S_0$, η πιθανότητα να υπάρχει ένα τουλάχιστον $i \in \{0, \ldots, \log(n)+2\}$ τέτοιο ώστε $|T \cap S_i| = 1$ είναι μεγαλύτερη από $1/2$.

Για τον υπολογισμό της ζητούμενης πιθανότητας, διαμερίζουμε τον χώρο συμβάντων στα εξής γεγονότα:

$$E^c = \{\forall i \text{ ισχύει ότι } |T \cap S_i| > 1\}, E = \{\exists i \text{ τέτοιο ώστε } |T \cap S_i| \leq 1\}$$

όπου και το γεγονός E διαμερίζεται στα εξής γεγονότα:

$$E_j = \{|T \cap S_j| > 1, |T \cap S_{j+1}| \leq 1\}, \quad j = 0, 1, \ldots, \log(n)+1$$

Αν με K συμβολίσουμε το επιθυμητό γεγονός, δηλαδή

$$K = \{\exists i \text{ τέτοιο ώστε } |T \cap S_i| = 1\}$$

τότε προκύπτει η εξής σχέση

$$\mathbb{P}[K] = \mathbb{P}[K|E^c] \cdot \mathbb{P}[E^c] + \mathbb{P}[K|E] \cdot \mathbb{P}[E] \tag{4.4}$$

Εύκολα βλέπει κανείς ότι τα γεγονότα E^c και K είναι ασυμβίβαστα και άρα $\mathbb{P}[K|E^c] = 0$. Μένει λοιπόν να υπολογίσουμε τις ποσότητες $\mathbb{P}[K|E]$ και $\mathbb{P}[E]$.

- *Υπολογίζουμε την ποσότητα* $\mathbb{P}[E]$:
  $\mathbb{P}[E] = 1 - \mathbb{P}[E^c]$ και $\mathbb{P}[E^c] = \mathbb{P}[\forall i$ ισχύει $|T \cap S_i| > 1]$ πράγμα που συνεπάγεται και ότι $\mathbb{P}[|T \cap S_{\log(n)+2}| > 1]$. Η συνεπαγωγή προκύπτει από την ιδίοτητα των συνόλων $\{S_i\}$ και συγκεκριμένα $S_{i+1} \subseteq S_i$, για κάθε i. Άρα,
  $$\mathbb{P}[|T \cap S_{\log(n)+2}| > 1] \leq \mathbb{P}[|S_{\log(n)+2}| \geq 1]$$
  και κάνοντας χρήση της ανισότητας Markov έχουμε ότι :
  $$\mathbb{P}[|S_{\log(n)+2}| \geq 1] \leq \frac{\mathbb{E}[|S_{\log(n)+2}|]}{1} = \mathbb{E}[|S_{\log(n)+2}|] = n \cdot 2^{-(\log(n)+2)} = \frac{1}{4}$$
  Τελικά, $\mathbb{P}[E] = 1 - \mathbb{P}[E^c] \geq 1 - \frac{1}{4} = \frac{3}{4}$

- *Υπολογίζουμε την ποσότητα* $\mathbb{P}[K|E]$:
  $$\mathbb{P}[K|E] = \frac{1}{\mathbb{P}[E]}\mathbb{P}[K \cap E]$$
  $$= \frac{1}{\mathbb{P}[E]}\Big[\mathbb{P}[(K \cap E_1) \cup \cdots \cup (K \cap E_{\log(n)+1})\Big] = \frac{1}{\mathbb{P}[E]} \sum_{j=1}^{\log(n)+1} \mathbb{P}[K \cap E_j]$$
  Δηλαδή
  $$\mathbb{P}[K|E] = \frac{1}{\mathbb{P}[E]} \sum_{j=1}^{\log(n)+1} \mathbb{P}[K|E_j]\mathbb{P}[E_j]$$
  Για να υπολογίσουμε τώρα την πιθανότητα αυτήν, θα δείξουμε ότι $\mathbb{P}[K|E_j] \geq \frac{2}{3}$ για κάθε j και άρα τελικά
  $$\mathbb{P}[K|E] \geq \frac{2}{3}\frac{1}{\mathbb{P}[E]} \sum_{j=1}^{\log(n)+1} \mathbb{P}[E_j] = \frac{2}{3} \tag{4.5}$$
  αφού όπως έχουμε ήδη πει, τα E_j αποτελούν διαμέριση του E. Υπολογίζουμε λοιπόν, για το τυχόν j, την ποσότητα $\mathbb{P}[K|E_j]$:
  $$\mathbb{P}\Big[|T \cap S_i| = 1\Big||T \cap S_j| > 1, |T \cap S_{j+1}| \leq 1\Big] = \frac{\mathbb{P}\Big[|T \cap S_i| = 1, |T \cap S_j| > 1, |T \cap S_{j+1}| \leq 1\Big]}{\mathbb{P}\Big[|T \cap S_j| = 1, |T \cap S_{j+1}| \leq 1\Big]}$$
  $$= \frac{\mathbb{P}\Big[|T \cap S_{j+1}| = 1, |T \cap S_j| > 1\Big]}{\mathbb{P}\Big[|T \cap S_{j+1}| \leq 1, |T \cap S_j| > 1\Big]}$$
  $$= \frac{\mathbb{P}\Big[|T \cap S_{j+1}| = 1\Big||T \cap S_j| > 1\Big] \cdot \mathbb{P}\Big[|T \cap S_j| > 1\Big]}{\mathbb{P}\Big[|T \cap S_{j+1}| \leq 1\Big||T \cap S_j| > 1\Big] \cdot \mathbb{P}\Big[|T \cap S_j| > 1\Big]}$$

Εάν υποθέσουμε ότι $|T \cap S_j| = t > 1$, αντικαθιστώντας στην παραπάνω σχέση παίρνουμε:

$$\frac{\mathbb{P}\Big[|T \cap S_{j+1}| = 1 \Big| |T \cap S_j| = t\Big]}{\mathbb{P}\Big[|T \cap S_{j+1}| \leq 1 \Big| |T \cap S_j| = t\Big]}$$

Υπολογίζουμε τις πιθανότητες και έχουμε:

$$\frac{\binom{t}{1}(\frac{1}{2})^t}{\binom{t}{0}(\frac{1}{2})^t + \binom{t}{1}(\frac{1}{2})^t} = \frac{t}{t+1} \geq \frac{2}{3}$$

Αντικαθιστούμε στην σχέση (4.4) και παίρνουμε τελικά:

$$\mathbb{P}[K] \geq \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2}$$

Στη συνέχεια θέλουμε να δείξουμε πως υλοποιείται προσεγγιστικά ένα κύκλωμα της κλάσης $\mathbf{AC^0}$, μεγέθους s και βάθους d με τη βοήθεια των πολυωνύμων που κατασκεύασαμε νωρίτερα. Υποθέτουμε, για λόγους ευκολίας ως προς την εποπτεία, ότι το κύκλωμα είναι σε «κανονικοποιημένη» μορφή, σε κάθε επίπεδο δηλαδή εμφανίζονται είτε μόνο πύλες **OR** είτε μόνο πύλες **AND**. Εάν θέλουμε να έχουμε πιθανότητα σφάλματος μικρότερη από ϵ για το πολυώνυμο που περιγράφει ολόκληρο το κύκλωμα, είναι αρκετό σε κάθε μία από τις s πύλες, να έχουμε αντίστοιχη πιθανότητα σφάλματος, μικρότερη από ϵ/s. Σύμφωνα με την σχέση 4.2, αρκεί να σχηματίσουμε $\log_2(s/\epsilon)$ το πλήθος πολυώνυμα για να το εξασφαλίσουμε αυτό. Γνωρίζουμε από τη σχέση 4.3 ότι τα πολυώνυμα που προσεγγίζουν είτε πύλη **OR** είτε πύλη **AND**, έχουν τον ίδιο βαθμό. Επίσης, ένα άνω φράγμα πλέον για το εύρος εισόδου των πυλών του κυκλώματος και συνεπώς για το βαθμό των πολυωνύμων, αποτελεί όχι το πλήθος n των μεταβλητών, αλλά το μέγεθος s του ίδιου του κυκλώματος. Συνδυάζοντας τα παραπάνω, βλέπουμε ότι τα πολυώνυμα θα έχουν βαθμό $\mathcal{O}\big(\log(s/\epsilon) \cdot \log(s)\big)$. Έχουμε όμως παραμελήσει το γεγονός ότι το κύκλωμα έχει βάθος d, αποτελείται δηλαδή από d επίπεδα. Με άλλα λόγια, αφού κάθε επίπεδο περιγράφεται είτε από την συνάρτηση **OR** είτε από την **AND**, η συνάρτηση που περιγράφει ολόκληρο το κύκλωμα, δίνεται από τη σύνθεση των αντίστοιχων πολυωνύμων, δηλαδή πολυώνυμο και πάλι και συγκεκριμένα βαθμού ίσου με το γινόμενο των βαθμών των επιμέρους πολυωνύμων, δηλαδή

$$\mathcal{O}\big(\log^d(s/\epsilon) \cdot \log^d(s)\big)$$

Αν υποθέσουμε για παράδειγμα, ότι θέλουμε να περιγράψουμε ένα κύκλωμα βάθους 2 $\mathbf{OR} - \mathbf{AND}$:

$$\mathbf{OR}\big(\mathbf{AND}(x_1, x_3, \ldots, x_{n-1}, x_n), \ldots, \mathbf{AND}(x_1, \overline{x_2}, \ldots, \overline{x_n})\big)$$

τότε θα καταλήγαμε στο εξής πολυώνυμο P σύμφωνα με τις σχέσεις 4.1 και 4.3:

$$P = 1 - p\Big(p(1-x_1, 1-x_3, \ldots, 1-x_{n-1}, 1-x_n),\ \ldots,\ p(1-x_1, x_2, \ldots, x_n) \Big)$$

Είναι φανερό ότι $deg(P) = deg(1-p) \cdot deg(p) = deg^2(p)$ και πως η ίδια διαδικασία μπορεί να εφαρμοστεί για να περιγράψει κύκλωμα οποιουδήποτε βάθους d. Αφού πειστεί κανείς για τα προηγούμενα, τότε μπορεί αντί για τις προσεγγίσεις $1-p$ και p, να χρησιμποποιήσει τα πολυώνυμα

$$1 - p_1 \cdot \ldots \cdot p_t \quad \text{και} \quad p_1 \cdot \ldots \cdot p_t \ ,$$

ανάλογα με την πιθάνοτητα σφάλματος που επιθυμεί.

Τελικά, μπορούμε να ισχυριστούμε ότι κάθε συνάρτηση f η οποία ανήκει στην κλάση $\mathbf{AC}^0$, μπορεί να προσεγγιστεί από κάποιο τυχαίο πολυώνυμο p, μικρού βαθμού, έτσι ώστε

$$\forall (a_1, \ldots, a_n) \in \{0,1\}^n \text{ έχουμε ότι } \mathbb{P}[f(a_1, \ldots, a_n) = p(a_1, \ldots, a_n)] \geq 1 - \epsilon$$

και ας πούμε εδώ για παράδειγμα ότι $1 - \epsilon = 0,9$. Τότε, αν με X_p συμβολίσουμε το πλήθος των σημείων στα οποία f και p συμφωνούν, εύκολα φαίνεται ότι

$$\mathbb{E}[X_P] \geq 0,9 \cdot 2^n$$

απ' όπου και συμπεραίνουμε πως
υπάρχει κάποιο συγκεκριμένο πολυώνυμο p το οποίο συμφωνεί με την f σε τουλάχιστον $0,9 \cdot 2^n$ σημεία και εδώ ολοκληρώνεται το πρώτο κομμάτι της απόδειξης, όπως προαναφέραμε στην αρχή του κεφαλαίου.

Ξεκινάμε τώρα το δεύτερο κομμάτι της απόδειξης και το πρώτο μας βήμα είναι να ασχοληθούμε με το θέμα της αναπαράστασης των λογικών τιμών **TRUE** και **FALSE** , όπου μέχρι τώρα τις αντιστοιχίζαμε στις τιμές 1 και 0. Από εδώ και στο εξής όμως, θα αποδειχθεί πιο χρήσιμο να χρησιμοποιήσουμε το γραμμικό μετασχηματισμό με τύπο:

$$x \mapsto 1 - 2x$$

και άρα

$$0 \mapsto 1 \quad \& \quad 1 \mapsto -1$$

Αντιστοιχίσαμε δηλαδή την τιμή **TRUE** στην τιμή -1 και την τιμή **FALSE** στην τιμή 1. Προφανώς ο αντίστροφος μετασχηματισμός δίνεται από τον τύπο

$$x \mapsto \frac{1-x}{2}$$

Έτσι λοιπόν, εφαρμόζοντας τον μετασχηματισμό αυτόν στο πολυώνυμο p, παίρνουμε το εξής πολυώνυμο q, ίδιου βαθμού, ακριβώς επειδή ο μετασχηματισμός είναι γραμμικός:

$$q(y_1, \ldots, y_n) = 1 - 2p\left(\frac{1-x_1}{2}, \ldots, \frac{1-x_n}{2}\right),\ x_i \in \{-1, 1\}$$

το οποίο «συμφωνεί» με την f σε $0,9 \cdot 2^n$ στοιχεία του $\{-1,1\}^n$.

Υποθέτουμε τώρα ότι η συνάρτηση *parity* ανήκει στην κλάση $\mathbf{AC}^0$. Συνεπώς, πρέπει να υπάρχει μία συνάρτηση q, πολυώνυμο για την ακρίβεια, όπως παραπάνω, για την *parity*. Δηλαδή, για τουλάχιστον $0,9 \cdot 2^n$ στοιχεία του $\{-1,1\}^n$ πρέπει να ισχύει

$$q(y_1, \ldots, y_n) = par(y_1, \ldots, y_n) = \prod_{i=1}^{n} y_i$$

Και αυτό συμβαίνει διότι, μετά τον μετασχηματισμό, η συνάρτηση *parity* αντιστοιχεί ακριβώς στον πολλαπλασιασμό των μεταβλητών. Με άλλα λόγια, ο μετασχηματισμός αυτός, απεικονίζει το διάνυσμα $(x_1, \ldots, x_n)$ στο $(y_1, \ldots, y_n)$ και άρα εάν έχουμε περιττό(ή άρτιο) πλήθος από «1» στο $(x_1, \ldots, x_n)$ θα έχουμε $par(x_1, \ldots, x_n) = 1$ (0 αντίστοιχα) και επίσης περιττό(άρτιο) και πάλι πλήθος από «-1» στο $(y_1, \ldots, y_n)$ και άρα $\prod_{i=1}^{n} y_i = -1$ (1 αντίστοιχα) δηλαδή

$$par(x_1, \ldots, x_n) \longrightarrow \prod_{i=1}^{n} y_i$$

Προχωράμε τώρα σε ένα λήμμα το οποίο και θα ολοκληρώσει και την απόδειξη μας.

Λήμμα 1 *Δεν υπάρχει πολυώνυμο βαθμού $\frac{\sqrt{n}}{2}$ το οποίο να «συμφωνεί» με την συνάρτηση $\prod_{i=1}^{n} y_i$ σε $0,9 \cdot 2^n$ στοιχεία του $\{-1,1\}^n$.*

Ως άμεση συνέπεια του παραπάνω λήμματος, παίρνουμε το εξής αποτέλεσμα:

$$\mathbf{PARITY} \notin \mathbf{AC}^0.$$

Απόδειξη Λήμματος 1: Υποθέτουμε, για να καταλήξουμε σε αντίφαση, πως η συνάρτηση $parity \in AC^0$ και άρα σύμφωνα με τα προηγούμενα, υπάρχει πολυώνυμο q, βαθμού $\frac{\sqrt{n}}{2}$ και σύνολο S με $|S| \geq 0,9 \cdot 2^n$, τέτοια ώστε q και *parity* να «συμφωνούν» σε κάθε στοιχείο S. Η $par_n(x_1, \ldots, x_n)$ όμως, μέσω του μετασχηματισμού που ορίσαμε πριν, αντιστοιχεί πλέον στην $\prod_{i=1}^{n} y_i$ το οποίο συνεπάγεται πως υπάρχει πολυώνυμο q με:

$$q(y) = \prod_{i=1}^{n} y_i \quad \forall y \in S.$$

Για να καταλήξουμε σε άτοπο, θα συνεχίσουμε ως εξής:

1. Αντιστοιχίζουμε, με τη βοήθεια του πολυωνύμου q, κάθε συνάρτηση από το S στο $\mathbb{R}$ με ένα πολυώνυμο βαθμού $\leq \frac{n+\sqrt{n}}{2}$. Η διάσταση του χώρου που αυτές σχηματίζουν, είναι ακριβώς $|S|$ αφού για να προσδιοριστεί μία τέτοια συνάρτηση, χρείαζονται $|S|$ το πλήθος σημεία.

2. Υπολογίζουμε το πλήθος μονωνύμων n μεταβλητών, βαθμού $\leq \frac{n+\sqrt{n}}{2}$, το οποίο και θα μας δώσει τη διάσταση του χώρου πολυωνύμων n μεταβλητών, βαθμού $\leq \frac{n+\sqrt{n}}{2}$.

- *Δείχνουμε πως πραγματοποιείται η αντιστοίχιση*

  Έστω $f : S \to \mathbb{R}$ και $S \subseteq \{-1, 1\}^n$. Αντιστοιχίζουμε την f με ένα πολυώνυμο p_f έτσι ώστε f και p_f συμφωνούν σε κάθε στοιχείο του S, μέσω της εξής αντιστοίχισης:

$$p_f(x_1, \ldots, x_n) = \sum_{y \in S} f(y) \prod_{i=1}^{n} \frac{y_i x_i + 1}{2} \tag{4.6}$$

  Παρατηρούμε ότι εάν $y \neq x$, τότε υπάρχει τουλάχιστον ένας δείκτης j τέτοιος ώστε $y_j \neq x_j$ και άρα

$$\prod_{i=1}^{n} \frac{y_i x_i + 1}{2} = 0 \quad \text{αφού } y_j, x_j \in \{-1, +1\}$$

  Συνεπώς, από το άθροισμα της σχέσης 4.6, μόνο ένας όρος «επιζεί» και συγκεκριμένα αυτός για τον οποίο ισχύει $x = y$. Παρατηρούμε όμως ότι ο βαθμός του p_f είναι εν γένει n. Μπορούμε όμως, αντί για το πολυώνυμο p_f να αντιστοιχίσουμε την f με ένα πολυώνυμο $\hat{p}_f$ το οποίο είναι βαθμού το πολύ $\frac{n+\sqrt{n}}{2}$ και εδώ θα χρειαστούμε το πολυώνυμο q. Εάν κάποιο πολυώνυμο έχει βαθμό $> \frac{n}{2}$, συνεπάγεται πως θα αποτελείται από κάποια μονώνυμα βαθμού μεγαλύτερου από $\frac{n}{2}$. Κάθε τέτοιο μονώνυμο, θα έχει την εξής μορφή: $\prod_{i \in T} x_i$ με $|T| > \frac{n}{2}$. Τώρα, για κάθε $x \in S$ έχουμε το εξής:

$$\begin{aligned} \prod_{i \in T} x_i &= \prod_{i=1}^{n} x_i \cdot \prod_{i \notin T} x_i \\ &= q(x_1, \ldots, x_n) \cdot \prod_{i \notin T} x_i \end{aligned}$$

  Αντικαθιστούμε δηλαδή κάθε τέτοιο μονώνυμο, από το γινόμενο του πολυωνύμου q επί το «συμπληρωματικό» μονώνυμο. Από τη στιγμή όμως που το q είναι το πολύ βαθμού $\frac{\sqrt{n}}{2}$ και το «συμπληρωματικό» μονώνυμο είναι το πολύ βαθμού $\frac{n}{2}$, μπορούμε λοιπόν να γράψουμε το πολυώνυμο p_f ως $\hat{p}_f$, το οποίο συμφωνεί με την f σε κάθε στοιχείο του S και είναι βαθμού το πολύ $\frac{n+\sqrt{n}}{2}$. Πετύχαμε δηλαδή, να αναπαραστήσουμε την f, πάνω στο σύνολο S, από πολυώνυμα με βαθμό το πολύ $\frac{n+\sqrt{n}}{2}$.

- *Υπολογίζουμε το πλήθος των μονωνύμων n μεταβλητών, βαθμού έως και* $\frac{n+\sqrt{n}}{2}$

Προφανώς, υπάρχουν

$$L = \sum_{i=0}^{\frac{n+\sqrt{n}}{2}} \binom{n}{i}$$

τέτοια μονώνυμα. Για να υπολογίσουμε τώρα το L εργαζόμαστε ως εξής:

$$\begin{aligned} L &= \sum_{i=0}^{\frac{n+\sqrt{n}}{2}} \binom{n}{i} \\ &= \sum_{i=0}^{\frac{n}{2}} \binom{n}{i} + \sum_{i=\frac{n}{2}+1}^{\frac{n+\sqrt{n}}{2}} \binom{n}{i} \\ &< \frac{1}{2} \cdot 2^n + \frac{\sqrt{n}}{2} \binom{n}{n/2} \\ &\leq \frac{1}{2} \cdot 2^n + \frac{\sqrt{n}}{2} \frac{1}{\sqrt{\frac{\pi n}{2}}} \cdot 2^n \\ &\leq \left(\frac{1}{2} + \frac{1}{\sqrt{2\pi}} \right) \cdot 2^n \\ &< 0,9 \cdot 2^n \end{aligned}$$

Συνδυάζοντας τα 2 προηγούμενα αποτελέσματα έχουμε:

$$\begin{aligned} 0,9 \cdot 2^n = |S| &= dim\{\text{χώρος συναρτήσεων } f \text{ με } f : S \Rightarrow \mathbb{R}\} \\ &\leq dim\{\text{χώρος πολυωνύμων βαθμού} \leq \frac{n+\sqrt{n}}{2}\} \\ &< 0,9 \cdot 2^n \end{aligned}$$

Ολοκληρώνεται λοιπόν εδώ η απόδειξη του **Ισχυρισμού 1**.

4.2 MAJORITY $\notin AC^0$

Μπορούμε πλέον να χρησιμοποιήσουμε το προηγούμενο αποτέλεσμα,

$$\textbf{PARITY} \notin \textbf{AC}^0 \,,$$

για να δείξουμε ότι το ίδιο συμβαίνει και με κάποιες άλλες συναρτήσεις Boole, όπως για παράδειγμα η συνάρτηση *majority*. Για τον σκοπό αυτό, πριν προχωρήσουμε, θα εισάγουμε μία έννοια αναγωγής, η οποία συνδέεται με την κλάση $\mathbf{AC^0}$:

Μία οικογένεια συναρτήσεων Boole $F = (f_1, f_2, \dots)$, *όπου*

$$f_n : \{0,1\}^n \to \{0,1\}$$

θα ονομάζεται $\mathbf{AC^0}-$ *αναγώγιμη σε μία οικογένεια* $G = (g_1, g_2, \ldots)$, *εάν* $\exists\, d, p$ *(*d *σταθερά και* p *πολυώνυμο) τέτοια ώστε,*

$$\forall n \;\exists \text{ κυκλώματα υλοποίησης της } f_n, \text{βάθους} \leq d \text{ και μεγέθους} \leq p(n),$$

τα οποία να αποτελούνται από πύλες **AND** *και* **OR** *αλλά και από πύλες* g_i, *όπου* i *τυχαίο.*

Αναφέρουμε τώρα ένα λήμμα το οποίο και θα αποδείξουμε:

Λήμμα 2 *Έστω* F $\mathbf{AC^0}-$ *αναγώγιμη στη* G *και* $G \in \mathbf{AC^0}$. *Τότε,*

$$F \in \mathbf{AC^0}.$$

Απόδειξη

$G \in \mathbf{AC^0} \Leftrightarrow \exists\, d$ και πολυώνυμο p τέτοια ώστε:

$$\forall i \text{ ισχύει ότι βάθος } (g_i) \leq d \text{ και μέγεθος } (g_i) \leq p(i)$$

Όμως, αφού F είναι $\mathbf{AC^0}-$ αναγώγιμη στη $G \Leftrightarrow \exists\, d'$ και πολυώνυμο p' τέτοια ώστε

$$\forall i \quad \text{βάθος}(f_i) \leq d' \text{ και μέγεθος } (f_i) \leq p'(i)$$

που όμως τώρα τα κυκλώματα $\{f_i\}$ αποτελούνται εκτός από πύλες **AND** και **OR** και από κυκλώματα g_i.
Είναι προφανές, ότι ένα άνω φράγμα για το βάθος των $\{f_i\}$ αποτελεί η σταθερά

$$d \cdot d' \;. \tag{4.7}$$

(Θεωρούμε την ακραία περίπτωση για παράδειγμα, όπου όλο το κύκλωμα f_i αποτελείται μόνο από πύλες - κυκλώματα g_k)
Όμοια σκεπτόμενοι λοιπόν, ένα άνω φράγμα για το μέγεθος του f_i προκύπτει ως εξής:
Έστω ότι αποτελείται μόνο από πύλες - κυκλώματα g_k, για κάποιους δείκτες k. Ξέρουμε όμως ότι μπορεί να αποτελείται το πολύ από $p'(i)$ πύλες - κυκλώματα της οικογένειας G και ένα πολυώνυμο που φράσσει το μέγεθος των $\{g_i\}$ είναι το p. Δεδομένου όμως, ότι κάθε πύλη-κύκλωμα g_k που εμφανίζεται στο κύκλωμα f_i, μπορεί να έχει ως είσοδο το πολύ όσο το μέγεθος του f_i, συμπεραίνουμε ότι ένα άνω φράγμα για τα μεγέθη των $\{g_k\}$ που εμφανίζονται στο f_i είναι το

$$p(p'(i)).$$

Και αφού αναφέραμε, ότι μπορεί να υπάρχουν το πολύ $p'(i)$ το πλήθος $\{g_k\}$, τότε το πολυώνυμο που φράσσει το μέγεθος των κυκλωμάτων της F είναι το

$$p' \cdot (p \circ p') \tag{4.8}$$

Συνδυάζοντας τις 4.7 και 4.8 συμπεραίνουμε ότι $F \in \mathbf{AC^0}$.

Δύο οικογένειες τέτοιων κυκλωμάτων, αποτελούν οι

$$PARITY = (par_1, par_2, \dots) \text{ και } MAJORITY = (maj_1, maj_2, \dots)$$

όπου

$$maj_n(x_1, \dots, x_n) = 1 \Leftrightarrow \exists \text{ τουλάχιστον } \frac{n}{2} \text{ δείκτες } i \text{ τ.ω. } x_i = 1.$$

Πράγματι, η αναγωγή της $PARITY$ στη $MAJORITY$ είναι πολύ εύκολη αν προηγουμένως δούμε ότι:

- Η συνάρτηση $majority$ αποτελεί ειδική περίπτωση των συναρτήσεων T_k, όπου

  $T_k(x_1, \dots, x_n) = 1 \Leftrightarrow$ τουλάχιστον k από τα x_i είναι τέτοια ώστε $x_i = 1$.

  Δηλαδή
  $$maj_n(x_1, \dots, x_n) = T_{n/2}(x_1, \dots, x_n)$$
  Με τη βοήθεια της $majority$ μπορούμε να εκφράσουμε οποιαδήποτε T_k:

  $*k < \frac{n}{2}$
  $$T_k(x_1, \dots, x_n) = maj_{2(n-k)}(x_1, \dots, x_n, \underbrace{1, \dots, 1}_{n-2k})$$

  Πράγματι, $maj_{2(n-k)} = 1 \Leftrightarrow$ τουλάχιστον $(n-k)$ $\{x_i\}_{i=1}^{2(n-k)}$ έχουν τιμή $1 \Leftrightarrow$ τουλάχιστον $(n-k)-(n-2k) = k$ από τα $\{x_i\}_{i=1}^{n}$ έχουν τιμή $1 \Leftrightarrow T_k(x_1, \dots, x_n) = 1$

  $*k \geq \frac{n}{2}$
  $$T_k(x_1, \dots, x_n) = maj_{2k}(x_1, \dots, x_n, \underbrace{0, \dots, 0}_{2k-n})$$

  Πράγματι,

  $maj_{2k}(x_1, \dots, x_n, \underbrace{0, \dots, 0}_{2k-n}) = 1 \Leftrightarrow$ τουλάχιστον k από τα $\{x_i\}_{i=1}^{n}$ έχουν τιμή 1

  $$\Leftrightarrow T_k(x_1, \dots, x_n) = 1$$

- Με τη βοήθεια τώρα των T_k μπορούμε να κατασκευάσουμε τις E_k:

  $$E_k(x_1, \dots, x_n) = 1 \Leftrightarrow k \text{ ακριβώς από τα } x_i \text{ έχουν τιμή } 1.$$

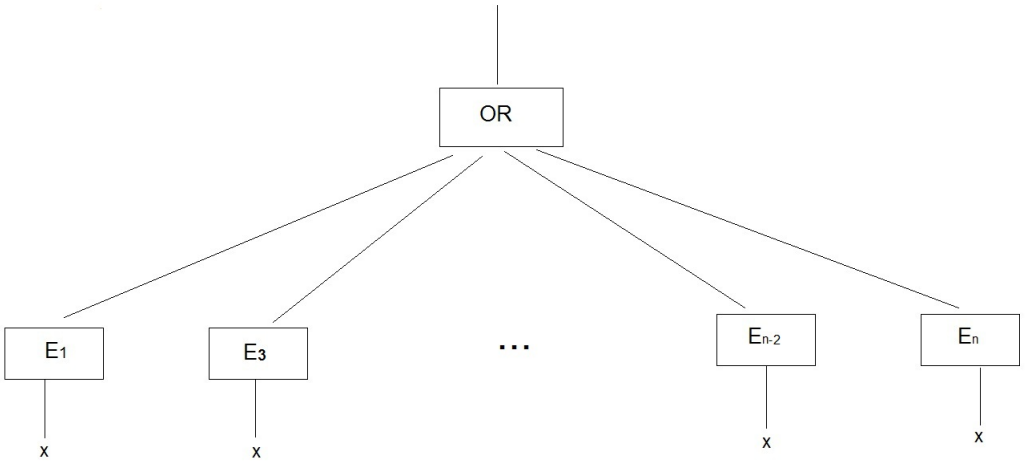
  Η E_k υλοποιείται ως εξής:

  $$E_k(x_1, \dots, x_n) = T_k(x_1, \dots, x_n) \wedge \overline{T_{k+1}(x_1, \dots, x_n)}$$

Επιστρέφουμε πάλι, στην $\mathbf{AC^0}-$αναγωγή της **PARITY** στη **MAJORITY**. Εύκολα βλέπει πλέον κανείς, πώς μπορεί να υλοποιήσει τη συνάρτηση *parity* με τη βοήθεια των T_k, δηλαδή της *majority*:

$$par_n(x_1, \ldots, x_n) = E_1 \vee E_3 \vee \ldots \vee E_{n-2} \vee E_n$$

Στο σχήμα παρακάτω, με x συμβολίζουμε το διάνυσμα $x = (x_1, \ldots, x_n, \overline{x_1}, \ldots, \overline{x_n})$ και θεωρούμε το n περιττό.



Προφανώς το παραπάνω κύκλωμα έχει σταθερό βάθος 2 και αν αναλύσουμε τις πύλες E_i σε πύλες T_i και T_{n-i}, το βάθος αυξάνεται σε 3 και δεν είναι δύσκολο να δούμε, ότι το μέγεθος μετρούμενο σε πύλες **AND** και **OR** και πύλεσ-κυκλώματα *majority* είναι $3\frac{n}{2} + 1$, δηλαδή της τάξης του $\mathcal{O}(n)$.

Ως άμεση συνέπεια λοιπόν των παραπάνω, έχουμε το εξής:

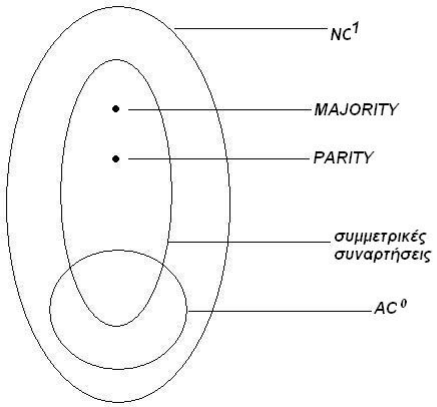
Θεώρημα 5 MAJORITY $\notin \mathbf{AC^0}$

Στην πραγματικότητα, ανάλογα επιχειρήματα μπορούν να χρησιμοποιηθούν για να δείξει κανείς ότι *όλες οι συμμετρικές συναρτήσεις* μπορούν να αναχθούν στη συνάρτηση *majority* μέσω $\mathbf{AC}^0-$αναγωγής. Με τον όρο «*συμμετρική*» χαρακτηρίζουμε μία συνάρτηση της οποίας η τιμή μένει αναλλοίωτη κάτω από οποιαδήποτε μετάθεση των μεταβλητών, δηλαδή η τιμή της εξαρτάται *μόνο* από το *άθροισμα* των μεταβλητών $\sum_{i=1}^{n} x_i$.
Μία τέτοια συνάρτηση f, μπορεί να προσδιοριστεί εξ΄ ολοκλήρου από ένα διάνυσμα τιμών της μορφής

$$(f_0, f_1, \ldots, f_n)$$

όπου κάθε f_k δίνει την τιμή της f όταν $\sum_{i=1}^{n} x_i = k$. Συνεπώς, υπάρχουν ακριβώς 2^{n+1} διαφορετικές *συμμετρικές* συναρτήσεις. Επιπλέον, μπορεί κανείς να δείξει ότι όλες οι *συμμετρικές* συναρτήσεις ανήκουν στην κλάση $\mathbf{NC}^1$(*φραγμένο*

εύρος, βάθος $\mathcal{O}(\log^1(n))$) και ότι η συνάρτηση *majority* δεν ανάγεται μέσω της $\mathbf{AC}^0$ στη συνάρτηση *parity*. Όλα τα παραπάνω απεικονίζονται σχηματικά ως εξής:



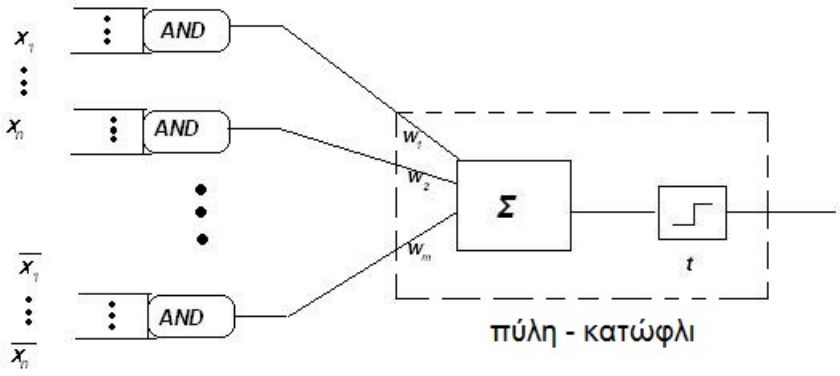
4.3 Κυκλώματα perceptrons και η συνάρτηση parity

Στη συνέχεια, θα αναφερθούμε σε μία διαφορετική τεχνική, για να δείξουμε ότι μία συνάρτηση δεν μπορεί να υλοποιηθεί από ένα κύκλωμα, κάποιων συγκεκριμένων χαρακτηριστικών, το οποίο πρωτοεισήχθη από τους Minsky & Papert [21]. Πιο συγκεκριμένα, θα «περιγράψουμε» ένα συγκεκριμένο είδος αναλογικού κυκλώματος, το *perceptron*, μέσω του πολυωνύμου του και θα συγκρίνουμε το βαθμό αυτό με το βαθμό του ελάχιστου πολυωνύμου που υπολογίζει τη συνάρτηση *parity*. Λέγοντας *perceptron*, εννοούμε ένα κύκλωμα βάθους 2 το οποίο έχει ως πύλη εξόδου μία πύλη-κατώφλι. Δεχόμαστε ότι οι πύλες στο πρώτο επίπεδο μπορούν να υπολογίσουν οποιαδήποτε συνάρτηση Boole. Μπορούμε όμως οποιαδήποτε συνάρτηση *Boole* να την μετασχηματίσουμε σε κανονική διαζευκτική μορφή και άρα τελικά να έχουμε μόνο πύλες **AND** στο πρώτο επίπεδο. Έχοντας δώσει λοιπόν, οι πύλες του πρώτου επιπέδου τα αποτελέσματά τους $a_i \in \{0, 1\}$, βάσει των συναρτήσεων που υπολογίζουν, στη συνέχεια κάθε a_i πολλαπλασιάζεται με ένα βάρος $w_i \in \mathbb{R}$ και προκύπτει ως αποτέλεσμα ένα άθροισμα της μορφής:

$$\sum_i a_i w_i$$

Τέλος, αυτό το άθροισμα συγκρίνεται με την τιμή t της πύλης - κατώφλι και αν το άθροισμα είναι μεγαλύτερο από t, το κύκλωμα δίνει αποτέλεσμα 1, αλλιώς δίνει αποτέλεσμα 0. Με άλλα λόγια

$$perceptron_t(x_1, \ldots, x_n) = \begin{cases} 1, & \text{εάν } \sum_i a_i w_i \geq t \\ 0, & \text{αλλιώς.} \end{cases} \tag{4.9}$$



Ένα perceptron λοιπόν προσδιορίζεται από το εύρος εισόδου n, από τις boolean συναρτήσεις του πρώτου επιπέδου, τα βάρη w_i και το κατώφλι t. Να σημειώσουμε ότι η πύλη-κατώφλι μπορεί να έχει απεριόριστο εύρος εισόδου. Η εφαρμογή τέτοιων κυκλωμάτων στον τομέα της αναγνώρισης προτύπου(pattern recognition) αλλά και ως βιολογικό μοντέλο του νευρώνα, ώθησε κάποιους στο να ερευνήσουν την περίπτωση όπου οι πύλες στο επίπεδο 1, δεν εξαρτώνται από όλα τα δεδομένα εισόδου $x_1, \ldots, x_n$, αλλά από ένα μικρότερο υποσύνολό τους. Υπήρχε η διαίσθηση ότι μεταβάλλοντας κανείς τις τιμές των βαρών w_i, τότε τέτοιου τύπου κυκλώματα θα μπορούσαν να υπολογίσουν («να μάθουν») οποιαδήποτε συνάρτηση Boole. Οι Minsky & Papert έδειξαν όμως, ότι τέτοια κυκλώματα δεν μπορούν να υπολογίσουν τη συνάρτηση *parity*. Το αποτέλεσμα αυτό, το οποίο θα δείξουμε στη συνέχεια, είναι και αυτό στο οποίο χρεώνεται η διακοπή χρηματοδότησης ερευνών πάνω στα νευρωνικά δίκτυα, ώσπου μόλις πρόσφατα η αξία τους αναγνωρίστηκε εκ νέου.

Πριν προχωρήσουμε με την απόδειξη των Minsky και Papert, να θυμίσουμε ότι μία πύλη **AND** μπορεί να περιγραφεί μέσω ενός πολυωνύμου και πιο συγκεκριμένα από το γινόμενο των μεταβλητών που εμφανίζονται ως είσοδοι στην πύλη. Για παράδειγμα, έστω ότι σε μία πύλη **AND**, τα $x_1, x_2, \overline{x_3}$ αποτελούν είσοδο. Τότε, το πολυώνυμο που αντιστοιχεί σε αυτήν την πύλη είναι το $x_1 \cdot x_2 \cdot (1 - x_3)$. Προφανώς για πεδίο τιμών το $\{0, 1\}$, το σύνολο τιμών είναι πάλι το $\{0, 1\}$ και άρα το πολυώνυμο υπολογίζει ακριβώς τη συνάρτηση **AND**. Επιπλέον, ο βαθμός του πολυωνύμου είναι ίσος με το πλήθος των μεταβλητών που εμφανίζονται στην πύλη και άρα μικρότερος από n. Διατυπώνουμε τώρα το θεώρημα των Minsky & Papert:

Θεώρημα 6 *Η συνάρτηση parity δεν μπορεί να υλοποιηθεί από ένα perceptron, όπου κάθε πύλη εισόδου εξαρτάται από λιγότερες από n μεταβλητές.*

Απόδειξη: Έστω f_i, $i = 1, \ldots, m$, το πολυώνυμο βαθμού $s < n$ το οποίο υλοποιεί την $i-$οστή πύλη **AND** του κυκλώματος. Τότε, τα πολυώνυμα αυτά πολλαπλασιάζονται πρώτα με τα αντίστοιχα βάρη w_i και κατόπιν προσθέτονται μεταξύ τους $\sum_{i=1}^{m} w_i f_i$. Προκύπτει λοιπόν, ένα πολυγραμμικό πολυώνυμο βαθμού $< n$. Με την προϋπόθεση ότι ένα τέτοιο κύκλωμα *perceptron* μπορεί να υπολογίσει τη συνάρτηση *parity*, τότε θα υπάρχει μία σταθερά t τέτοια ώστε:

$$\sum_{i=1}^{m} w_i f_i \geq t \Leftrightarrow par(x_1, \ldots, x_n) = 1$$

Δηλαδή το πρόσημο του πολυωνύμου p, με

$$p = \sum_{i=1}^{m} w_i f_i - t$$

καθορίζει το αποτέλεσμα της συνάρτησης *parity* για είσοδο $(x_1, \ldots, x_n)$. Στη συνέχεια, θα κάνουμε χρήση της συμμετρικότητας της συνάρτησης *parity*. Δηλαδή, για οποιαδήποτε μετάθεση π του συνόλου $\{x_1, \ldots, x_n\}$ ισχύει:

$$par_n(x_1, \ldots, x_n) = par_n(\pi(x_1), \ldots, \pi(x_n))$$

Μετά, κατασκεύαζουμε το εξής πολυώνυμο q:

$$q(x_1, \ldots, x_n) = \sum_{\pi} p(\pi(x_1), \ldots, \pi(x_n))$$

Βλέπουμε ότι το q είναι και αυτό βαθμού $< n$ και κατασκευασμένο έτσι ώστε να είναι συμμετρικό. Επίσης, για κάθε διάνυσμα $x \in \{0,1\}^n$ έχουμε ότι $q(x) = n!p(x)$ και συνεπώς εξακολουθεί να ισχύει η σχέση

$$q(x) > 0 \Leftrightarrow par(x) = 1 \tag{4.10}$$

Επομένως το πολυώνυμο q, μπορεί να γραφεί στην εξής μορφή:

$$q(x_1, \ldots, x_n) = \sum_{d=0}^{s} A_d \, t_d(x_1, \ldots, x_n)$$

όπου A_d είναι συντελεστές και t_d είναι το άθροισμα όλων των μονωνύμων βαθμού d, δηλαδή

$$t_d(x_1, \ldots, x_n) = \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=d}} \prod_{j \in S} x_j \tag{4.11}$$

Έχουμε όμως την εξής ιδιότητα:

$$q(\underbrace{1, 0, \ldots, 1, \ldots, 1, 0}_{k \text{ το πλήθος από "1"}}) = q(\underbrace{1, 1, \ldots, 1}_{k}, 0, \ldots, 0)$$

Όμως,

$$q(\underbrace{1,1,\ldots,1}_{k},0,\ldots,0)=\sum_{d=0}^{s}A_d\,t_d(\underbrace{1,\ldots,1}_{k},0,\ldots,0) \tag{4.12}$$

Ξαναγυρνάμε στην 4.11 και βλέπουμε ότι η ποσότητα $t_d(\underbrace{1,\ldots,1}_{k},0,\ldots,0)$ μας δείχνει στην πραγματικότητα με πόσους τρόπους μπορούμε να πάρουμε υποσύνολα μεγέθους d του συνόλου $\{x_1,x_2,\ldots,x_n\}$ στα οποία εμφανίζονται μόνο «1». Άρα,

$$t_d(\underbrace{1,\ldots,1}_{k},0,\ldots,0)=\binom{k}{d}$$

και αντικαθιστώντας στην 4.12, έχουμε:

$$q(\underbrace{1,1,\ldots,1}_{k},0,\ldots,0)=\sum_{d=0}^{s}A_d\binom{k}{d}$$

Εισάγουμε λοιπόν, το πολυώνυμο r μιας μεταβλητής, που ορίζεται από τον τύπο:

$$r(x)=\sum_{d=0}^{s}A_d\binom{x}{d}$$

Το πολυώνυμο r όμως, είναι μιας μεταβλητής, βαθμού $s<n$ με την ιδιότητα

$$r(k)>0\Leftrightarrow q(x_1,\ldots,x_n)>0\Leftrightarrow par(x_1,\ldots,x_n)=1$$

δηλαδή,

$$r(k)>0\Leftrightarrow k\in\{0,1,\ldots,n\}\text{ και }k\text{ περιττός}$$

και ταυτόχρονα

$$r(k)<0\Leftrightarrow k\in\{0,1,\ldots,n\}\text{ και }k\text{ άρτιος}$$

Είναι προφανές, ότι από τη στιγμή που υπάρχουν n διαστήματα στα οποία το πολυώνυμο r αλλάζει πρόσημο, τότε θα έχει και τουλάχιστον n ρίζες. Όμως ένα πολυώνυμο το οποίο έχει n ρίζες, πρέπει να είναι τουλάχιστον βαθμού n ενώ το πολυώνυμο q είναι βαθμού $<n$, άτοπο.

Βιβλιογραφία

[1] C.E. Shannon, *A Mathematical Theory of Communication*, Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948

[2] Hartmanis, J., and Stearns, R. E. *On the computational complexity of algorithms*, Transaction of the American Mathematical Society 117 (1965)

[3] J. E. Savage, *Computational work and time on finite machines*, Journal of the ACM, 19(4):660-674, October 1972

[4] Paul, W. *A 2.5n-lower bound on the combinational complexity of Boolean functions*, SIAM J. Comput. 6, 427–443

[5] E. I. Nečiporuk, *On a Boolean function*, Doklady of the Academy of Sciences of the USSR, 169(4):765-766 (in Russian), 1966

[6] Miklós Ajtai, Σ_1^1-*Formulae on Finite Structures*, Annals of Pure and Applied Logic, 24 (1983) 1-48

[7] P. Erdös, *On a problem of Sidon in additive number theory*, Acta Sci. Math. (Szeged), 15 (1953-54), 255-259

[8] P. Erdös, *Problems and results in additive number theory*, Colloque sur la Théorie des Nombres (CBRM, Bruxelles), 127-137

[9] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley Interscience Series in Discrete Mathematics and Optimization, 1992

[10] H. Chernoff, *A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Stat. 23 (1952), 493-509

[11] N. Alon and D.J. Kleitman, *Sum-free subsets*, in A. Baker, B. Bollobás, and A. Hajnál eds., A Tribute to Paul Erdös, Cambridge University Press (1990), 13-26

[12] P.Erdös, *Extremal Problems in number theory*, Proceedings of the Symp. Pure Math. VIII, AMS (1965), 181-189

[13] N. Alon, I. Kriz and J. Nesetril, *How to color shift hypergraphs*, Studia Scientiarum Mathematicarum Hungarica 30 (1995), 1-11

[14] M.L. Furst, J.B. Saxe, M. Sipser, *Parity, circuits, and the polynomial-time hierarchy*, Mathematical Systems Theory 17 (1984) 13-27

[15] Andrew C. Yao, *Lower bounds by probabilistic arguments*, focs, pp.420-428, 24th Annual Symposium on Foundations of Computer Science (FOCS 1983), 1983

[16] J. Håstad, *Almost optimal lower bounds for small depth circuits*, Proceedings of the 18th Annual Symposium on Theory of Computing, ACM, 1986, 6-12

[17] L. Valiant, V. Vazirani, *NP is as easy as detecting unique solutions*, Theoretical Computer Science 47 (1986, 85-93)

[18] A. Razborov, *Lower bounds on the size of bounded depth networks over a complete basis with logical addition*, Mathematical Notes of the Academy of Sciences of the USSR 41 (1987), 333-338

[19] R. Smolensky, *Algebraic methods in the theory of lower bounds for Boolean circuit complexity*, Proceedings of the 19th Annual Symposium on Theory of Computing, ACM, 1979, 77-82

[20] R. Beigel, *The polynomial method in circuit complexity*, Structure in Complexity Theory Conference, IEEE, 1993

[21] M.L. Minsky, S.A. Papert, *Perceptrons*, MIT Press, 1969

[22] J. Aspnes, R. Beigel, M. Furst, S. Rudich, *The expressive power of voting polynomials*, Combinatorica 14 (1994) 135-148