Παραμετρικοί Αλγόριθμοι και Μητροειδή: η χρήση των συνόλων αντιπροσώπευσης

Πετροπαναγιωτάκη Μαρία

Τριμελής Επιτροπή
Θηλυκός Δημήτριος, Τμ. Μαθηματικών, Ε.Κ.Π.Α.
Κολλιόπουλος Σταύρος, Τμ. Πληροφορικής και Τηλ/νιών, Ε.Κ.Π.Α.
Κυρούσης Ελευθέριος, Τμ. Μαθηματικών, Ε.Κ.Π.Α.

Περιεχόμενα

Εισαγωγή

Η θεωρία των μητροειδών παρέχει χρήσιμα εργαλεία και τεχνικές για την επίλυση διαφόρων αλγοριθμικών προβλημάτων. Η αλγεβρική δομή τους καθιστά την χρήση τους ως βάσεις δεδομένων εξαιρετικά αποτελεσματική στην κατασκευή άπληστων, προσεγγιστικών και παραμετρικών αλγορίθμων.

Ο Monien [14] όρισε και χρησιμοποίησε την έννοια των αντιπροσωπευτικών οικογενειών, για την κατασκευή αλγορίθμου που βρίσκει μονοπάτι σε γράφημα. Σε αυτόν τον ορισμό, παρ'ότι δεν γίνεται ξεκάθαρη αναφορά σε αυτό, υπονοείται ένα ομοιόμορφο μητροειδές. Ο Marx [13], γενικεύοντας αυτόν τον ορισμό σε μητροειδή, έγινε ο πρώτος που τα χρησιμοποίησε για την κατασκευή παραμετρικών αλγορίθμων.

Έστω $\mathcal{M} = (E, \mathcal{I})$ μητροειδές και $\mathcal{S} = \{S_1, \ldots, S_t\}$ οικογένεια υποσυνόλων του E με $|S_i| = p$ $\forall i \in \{1, \ldots, t\}$. Μια οικογένεια $\mathcal{S}' \subseteq \mathcal{S}$ είναι q-αντιπροσωπευτική της $\mathcal{S}$ αν για κάθε σύνολο $Y \subseteq E$ με το πολύ q στοιχεία ισχύει το εξής: αν υπάρχει σύνολο $X \in \mathcal{S}$ τέτοιο ώστε $X \cap Y = \emptyset$ και $X \cup Y \in \mathcal{I}$ τότε υπάρχει και σύνολο $X' \in \mathcal{S}'$ τέτοιο ώστε $X' \cap Y = \emptyset$ και $X' \cup Y \in \mathcal{I}$. Με άλλα λόγια, αν κάποιο σύνολο της οικογένειας $\mathcal{S}$ μπορεί να επεκταθεί σε μεγαλύτερο ανεξάρτητο σύνολο προσθέτοντάς του q νέα στοιχεία, τότε υπάρχει και κάποιο σύνολο της $\mathcal{S}'$, το οποίο μπορεί να επεκταθεί σε μεγαλύτερο ανεξάρτητο σύνολο προσθέτοντάς του τα ίδια q στοιχεία.

Από το Θεώρημα των Δύο-Οικογενειων του Bollobás [3] και τη γενικευσή του σε διανυσματικούς χώρους από τον Lovász [10], προκύπτει ότι υπάρχει q-αντιπροσωπευτική οικογένεια με το πολύ $\binom{p+q}{p}$ σύνολα. Πώς μπορεί, όμως, να κατασκευαστεί αυτή η οικογένεια;

Όσον αφορά τα ομοιόμορφα μητροειδή, ο Monien [14] σχεδιάζει αλγόριθμο που υπολογίζει μια q-αντιπροσωπευτική οικογένεια της $\mathcal{S}$ μεγέθους το πολύ $\sum_{i=0}^{q} p^i$ σε $\mathcal{O}\big(|\mathcal{S}|pq \cdot \sum_{i=0}^{q} p^i\big)$ χρόνο. Ένας ακόμη αλγόριθμος για την περίπτωση των ομοιόμορφων μητροειδών δίνεται από τον Marx [12]. Ο αλγόριθμος αυτός υπολογίζει σε $\mathcal{O}(|\mathcal{S}|^2 p^q)$ χρόνο q-αντιπροσωπευτική οικογένεια μεγέθους $\binom{p+q}{p}$. Μερικά χρόνια αργότερα, ο Marx [13] «αλγοριθμοποιεί» την απόδειξη του Lovász και κατασκευάζει q-αντιπροσωπευτική οικογένεια, όταν το μητροειδές είναι γραμμικό. Το μέγεθος της οικογένειας είναι και σε αυτή την περίπτωση ίσο με $\binom{p+q}{p}$, ενώ ο χρόνος του αλγορίθμου ισούται με $f(p,q)(||A_{\mathcal{M}}|| \cdot |\mathcal{S}|)^{\mathcal{O}(1)}$, όπου $f(p,q)$ πολυωνυμική συνάρτηση ως προς $(p+q)^p$ και $\binom{p+q}{p}$ και $A_{\mathcal{M}}$ ο πίνακας αναπαράστασης του μητροειδούς $\mathcal{M}$. Άρα, όταν το p είναι σταθερό, $f(p,q) = (p+q)^{\mathcal{O}(1)}$. Σε αντίθετη περίπτωση, ο χρόνος που χρειάζεται ο αλγόριθμος φράσσεται από $2^{\mathcal{O}(k \log k)}(||A_{\mathcal{M}}|| \cdot |\mathcal{S}|)^{\mathcal{O}(1)}$.

Στο πλαίσιο αυτής της εργασίας δίνουμε δύο πιο γρήγορους αλγορίθμους για τον υπολογισμό αντιπροσωπευτικών οικογενειών τόσο στην περίπτωση των γραμμικών όσο και στην περίπτωση των ομοιόμορφων μητροειδών. Οι αλγόριθμοι αυτοί, σχεδιάστηκαν από τους Fomin, Lokshtanov, Panolan και Saurabh στο [21]. Συγκεκριμένα, για γραμμικά μητροειδή αποδεικνύουμε το παρακάτω θεώρημα.

Θεώρημα 0.1. *Έστω $\mathcal{M} = (E, \mathcal{I})$ γραμμικό μητροειδές τάξης $p+q = k$, $A_{\mathcal{M}}$ πίνακας αναπαράστασής του επί του σώματος $\mathbb{F}$ και $\mathcal{S} = \{S_1, \ldots, S_t\}$ οικογένεια ανεξάρτητων υποσυνόλων του E με $|S_i| = p$ $\forall i \in \{1, \ldots, t\}$. Τότε υπάρχει αλγόριθμος που υπολογίζει μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της*

$\mathcal{S}$ μεγέθους το πολύ $\binom{p+q}{p}$, σε $\mathcal{O}\Big(\binom{p+q}{p}tp^{\omega}+t\binom{p+q}{q}^{\omega-1}\Big)$ χρόνο, όπου ως μονάδα χρόνου θεωρούμε μια πράξη στο $\mathbb{F}$ και $\omega<2.373$ ο εκθέτης για τον πολλαπλασιασμό πινάκων.

Η απόδειξη του Θεωρήματος 0.1 βασίζεται στην σχετική απόδειξη του Lovász [10] και εκμεταλλεύεται το γεγονός ότι η συνάρτηση της ορίζουσας είναι πολυγραμμική.

Στη περίπτωση των ομοιόμορφων μητροειδών μπορούμε να βελτιώσουμε περαιτέρω τον χρόνο.

Θεώρημα 0.2. *Έστω $\mathcal{S}=\{S_1,\ldots,S_t\}$ οικογένεια υποσυνόλων ενός συνόλου με n στοιχεία τέτοια ώστε $|S_i|=p\ \forall\{1,\ldots,t\}$, q ακέραιος και $x\in(0,1)$. Τότε υπάρχει αλγόριθμος, που υπολογίζει σε χρόνο $\mathcal{O}\big((1-x)^{-q}\cdot 2^{o(p+q)}\cdot t\cdot\log n\big)$ μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ τέτοια ώστε $|\mathcal{S}'|\leq x^{-p}(1-x)^{-q}\cdot 2^{o(p+q)}$.*

Για την απόδειξη του Θεωρήματος 0.2, ορίζουμε τις συλλογές διαχωρισμού και στη συνέχεια δίνουμε αλγόριθμο για την κατασκευή τους. Μια *n-p-q-συλλογή διαχωρισμού* $\mathcal{C}$ είναι ένα ζεύγος $(\mathcal{F},\chi)$, όπου $\mathcal{F}$ οικογένεια υποσυνόλων ενός συνόλου U με n στοιχεία και χ συνάρτηση από το $\binom{U}{p}$ στο $2^{\mathcal{F}}$ τέτοια ώστε: (α) για κάθε $A\in\binom{U}{p}$ και κάθε $F\in\chi(A)$, $A\subseteq F$ και (β) για κάθε $A\in\binom{U}{p}$ και $B\in\binom{U\setminus A}{q}$, υπάρχει $F\in\chi(A)$ τέτοιο ώστε $A\subseteq F$ και $F\cap B=\emptyset$. Το *μέγεθος* της $(\mathcal{F},\chi)$ ισούται με $|\mathcal{F}|$, ενώ ο *μέγιστος βαθμός* ορίζεται ως $\max\limits_{A\in\binom{U}{p}}|\chi(A)|$. Για την κατασκευή των συλλογών διαχωρισμού, σχεδιάζουμε αλγόριθμο ο οποίος για κάθε n, p και q υπολογίζει την οικογένεια $\mathcal{F}$ της συλλογής διαχωρισμού $(\mathcal{F},\chi)$ και έπειτα επιστρέφει την οικογένεια $\chi(A)$ για είσοδο κάποιο σύνολο $A\in\binom{U}{p}$.

Τέλος, δείχνουμε πως υπολογίζοντας αποδοτικά αντιπροσωπευτικές οικογένειες, μπορούμε να κατασκευάσουμε γρηγορότερους παραμετρικούς αλγορίθμους.

ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ. Στο πρόβλημα *ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ* δίνονται μητροειδή $\mathcal{M}_1,\ldots,\mathcal{M}_\ell$ μαζί με τους πίνακες αναπαράστασής τους και ακέραιος k και διερωτόμαστε αν υπάρχει κοινό ανεξάρτητο σύνολο για όλα τα μητροειδή μεγέθους τουλάχιστον k. Το πρόβλημα αυτό για $\ell=2$ ανήκει στην κλάση P, δηλαδή υπάρχει αλγόριθμος πολυωνυμικού χρόνου. Επομένως, η εύρεση παραμετρικού αλγορίθμου έχει νόημα για $\ell\geq 3$.

Ο Marx στο [13] σχεδιάζει αλγόριθμο για το παραπάνω πρόβλημα, στα πλαίσια του οποίου ορίζεται μια οικογένεια $\mathcal{S}$ και υπολογίζεται αντιπροσωπευτική οικογένεια της σε $2^{\mathcal{O}(k\log k)}(||A_{\mathcal{M}}||\cdot|\mathcal{S}|)^{\mathcal{O}(1)}$ χρόνο. Αντικαθιστώντας τον υπολογισμό αυτόν με τον αλγόριθμο του Θεωρήματος 0.1, προκύπτει γρηγορότερος αλγόριθμος για το πρόβλημά μας.

ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ. Σε αυτό πρόβλημα ενδιαφερόμαστε να βρούμε κύκλο μήκους τουλάχιστον k σε κατευθυνόμενο γράφημα με n κορυφές και m ακμές.

Ο πρώτος FPT αλγόριθμος για το παραπάνω πρόβλημα δόθηκε από τους Gabow και Nie στο [17]. Ο αναμενόμενος χρόνος του ισούται με $k^{2k}2^{\mathcal{O}(k)}nm$, ενώ ο χρόνος χειρότερης περίπτωσης με $\mathcal{O}(k^{2k}2^{\mathcal{O}(k)}nm\log n)$ ή $\mathcal{O}(k^{3k}nm)$. Έπεται ότι αν στο πρόβλημα αναζητούμε κύκλο μήκους τουλάχιστον $\log n/\log\log n$, τότε ο αλγόριθμος αναμένεται να τον βρει, εφόσον υπάρχει, σε πολυωνυμικό χρόνο.

Οι Fomin, Lokshtanov, Panolan και Saurabh στο [21] δίνουν τον αλγόριθμο, τον οποίο θα περιγράψουμε παρακάτω. Ο αλγόριθμος αυτός τρέχει σε $\mathcal{O}(6.75^{k+o(k)}mn^2\log n)$ χρόνο, ενώ αν αναζητούμε κύκλο μήκους τουλάχιστον $\log n$ τρέχει σε πολυωνυμικό χρόνο. Οι Bjorklund, Husfeldt, και Khanna στο [8] απέδειξαν ότι αν ισχύει η Υπόθεση Εκθετικού Χρόνου (Exponential Time Hypothesis) των Impagliazzo, Paturi και Zane [30], δεν υπάρχει αλγόριθμος πολυωνυμικού χρόνου που βρίσκει κατευθυνόμενο κύκλο μήκους $\Omega\big((f(n)\log n\big)$, όπου f μη-φθίνουσα, μη-φραγμένη, υπολογίσιμη συνάρτηση. Επομένως, ο αλγόριθμος αυτός μειώνει το χάσμα μεταξύ άνω και κάτω φράγματος.

k-***ΜΟΝΟΠΑΤΙ.*** Σε αυτό το πρόβλημα, το οποίο μελετήθηκε στο [5] ως το πρώτο παράδειγμα παραμετρικού αλγορίθμου, μας δίνεται γράφημα G με n κορυφές και m ακμές και ακέραιος k. Σκοπός είναι να βρεθεί αν υπάρχει μονοπάτι στο G μήκους ακριβώς k.

Οι Monien [14] και Bodlaender [2] έδειξαν ότι το πρόβλημα είναι αποδοτικώς επιλύσιμο με σταθερή παράμετρο. Συγκεκριμένα, ο αλγόριθμος του πρώτου τρέχει σε $\mathcal{O}(k!nm)$ χρόνο και του δευτέρου σε $\mathcal{O}(k!2^k n)$ χρόνο. Αρκετά χρόνια αργότερα οι Alon, Yuster και Zwick [37] εισαγάγουν μια καινούρια τεχνική για την κατασκευή FPT αλγορίθμων, την τεχνική χρωματικής κωδικοποίησης (color coding), και με αυτόν τον τρόπο κατασκευάζουν ντετερμινιστικό αλγόριθμο που τρέχει σε $\mathcal{O}(c^k n \log n)$ (c μεγάλο) χρόνο. Το 2007, στο [35], κατασκευάζεται αλγόριθμος που τρέχει σε $\mathcal{O}(4^{k+\mathcal{O}(\log^3 k)}nm)$ χρόνο. Τέλος, το 2016, κατασκευάζεται ο αλγόριθμος που θα παρουσιάσουμε στη συνέχεια, ο οποίος βελτιώνει περαιτέρω τον χρόνο σε $\mathcal{O}(2.619^k n \log n)$.

Κεφάλαιο 1

Αλγόριθμοι και Πολυπλοκότητα

1.1 Σύντομη Ιστορία της Πολυπλοκότητας

Όλα ξεκίνησαν από μια μηχανή...

Το 1936, ο Turing επινόησε ένα θεωρητικό μοντέλο υπολογισμού, τη μηχανή Turing, βασισμένο στη μαθηματική σκέψη. Πρόκειται, για ένα πολύ απλό μοντέλο υπολογισμού, το οποίο όμως δεν μπορεί να κατασκευαστεί αφού διαθέτει άπειρη μνήμη. Παρ' όλα αυτά μια μηχανή Turing μπορεί να κάνει ότι και ένας πραγματικός υπολογιστής.

Ο Turing μέσω της μηχανής του δίνει σαφή ορισμό στην έννοια του αλγορίθμου, μια έννοια με μακρά ιστορία στα μαθηματικά. Σε πειργραφικό επίπεδο αλγόριθμος είναι ένα πεπερασμένο σύνολο σαφώς ορισμένων κανόνων, οι οποίοι περιγράφουν τα βήματα για να λυθεί ένα συγκεκριμένο πρόβλημα. Υπάρχει, όμως, αλγόριθμος για κάθε πρόβλημα; Όπως απέδειξε και ο Turing τα προβλήματα χωρίζονται σε αυτά που μπορούν να επιλυθούν αλγοριθμικά και ορισμένα άλλα που δεν μπορούν. Η κατηγοριοποίηση αυτή, καθώς και ο βαθμός ανεπιλυσιμότητας των διαφόρων προβλημάτων αποτελεί αντικείμενο της Θεωρίας Υπολογισιμότητας.

Ακόμη και όταν ένα πρόβλημα είναι επιλύσιμο, πιθανόν να είναι πρακτικά ανεπίλυτο εάν η επίλυση του απαιτεί εξωπραγματικό χρόνο ή υπερβολικά πολλή μνήμη. Με ποιό τρόπο μπορούμε να μετρήσουμε τους πόρους που απαιτούνται για την επίλυση των αλγοριθμικών προβλημάτων; Η απάντηση σε αυτό το ερώτημα δόθηκε στις αρχές της δεκαετίας του '60 από τους Hartmanis και Stearns. Πρότειναν την μέτρηση τόσο του χρόνου όσο και του χώρου σε συνάρτηση με το μήκος της εισόδου, εγκαινιάζοντας έτσι έναν νέο κλάδο της Θεωρίας των Υπολογιστών, την Υπολογιστική Πολυπλοκότητα.

Στη συνέχεια, έγινε προσπάθεια ταξινόμησης των προβλημάτων ανάλογα με το χρόνο ή το χώρο που απαιτείται για την επίλυσή τους. Με αυτό τον τρόπο ορίστηκαν οι κλάσεις πολυπλοκότητας, οι πιο γνωστές από τις οποίες είναι οι P και NP. Και οι δύο αυτές κλάσεις αφορούν την μέτρηση του χρόνου. Η κλάση P περιέχει όλα τα προβλήματα τα οποία μπορούν να επιλυθούν αποδοτικά ενώ στην NP ανήκουν όλα αυτά για τα οποία αν μας δοθεί μια λύση μπορούμε να την επαληθεύσουμε αποδοτικά. Με τον όρο αποδοτικά εννοούμε την ύπαρξη ενός ντετερμινιστικού αλγορίθμου, ο οποίος επιλύει το πρόβλημα σε χρόνο το πολύ πολυωνυμικό ως προς το μέγεθος της εισόδου του.

Προφανώς όλα τα προβλήματα της κλάσης P ανήκουν και στην NP. Ισχύει όμως και το αντίστροφο; Το ερώτημα $P \stackrel{?}{=} NP$ αποτελεί το σημαντικότερο ερώτημα της Υπολογιστικής Πολυπλοκότητας και παραμένει, μέχρις στιγμής, αναπάντητο, αν και οι περισσότεροι επιστήμονες πιστεύουν ότι $P \neq NP$.

Η δουλειά των Cook και Karp, στις αρχές της δεκαετίας του '70, έδειξε ότι υπάρχουν κάποια προβλήματα στην κλάση NP, τα οποία αποκαλούνται NP-πλήρη και είναι πιο δύσκολα. Αν για κάποιο από αυτά βρεθεί αποδοτικός αλγόριθμος τότε το πρόβλημα $P \stackrel{?}{=} NP$ έχει θετική απάντηση. Μέχρις στιγμής, οι μόνοι αλγόριθμοι για τα NP-πλήρη προβλήματα είναι εκθετικοί ή (στην καλύτερη περίπτωση)

υποεκθετικοί, ενώ δεν έχει αποδειχθεί η μη-ύπαρξη πολυωνυμικού αλγορίθμου για κάποιο από αυτά.

Μπορούμε όμως να επιλύσουμε NP-πλήρη ή και ακόμα πιο δύσκολα προβλήματα αποδοτικά ή σχεδόν αποδοτικά; Υπάρχουν διάφορες τεχνικές τις οποίες μπορούμε να δοκιμάσουμε προκειμένου να κατασκευάσουμε γρηγορότερους αλγορίθμους, κάποιες από τις οποίες είναι:

Προσεγγιστικοί Αλγόριθμοι για προβλήματα βελτιστοποίησης. Αντί να ψάχνουμε για την βέλτιστη λύση, ψάχνουμε για μια «σχεδόν» βέλτιστη.

Πιθανοτικοί Αλγόριθμοι κάνουν τυχαίες επιλογές σε κάποιο ή κάποια βήματα με πολύ μικρή πιθανότητα αποτυχίας.

Παραμετρικοί Αλγόριθμοι, των οποίων η πολυπλοκότητα μετριέται συναρτήσει όχι μόνο του μήκους εισόδου αλλά και κάποιας παραμέτρου.

Οι παραπάνω τεχνικές αναπτύχθηκαν τις τελευταίες δεκαετίες και κάθε μία από αυτές αποτελεί έναν ξεχωριστό κλάδο της Θεωρίας των Υπολογιστών. Η πιο «σύγχρονη» είναι η τελευταία, η οποία άρχισε να αναπτύσσεται μόλις τη δεκαετία του '90 από τους Downey και Fellows και θα μας απασχολήσει στη συνέχεια.

1.2 Παραμετρικοί Αλγόριθμοι

Στην παράγραφο αυτή, θα δώσουμε κάποιους τυπικούς ορισμούς σχετικά με τους Παραμετρικούς Αλγορίθμους.

Έστω Σ πεπερασμένο αλφάβητο (π.χ. το δυαδικό αλφάβητο $\{0,1\}$). Θα συμβολίζουμε με Σ^* το σύνολο όλων των πεπερασμένων ακολουθιών που αποτελούνται από στοιχεία του Σ.

Ορισμός 1.1. *Παραμετροποιημένο πρόβλημα ονομάζεται κάθε υποσύνολο L του $\Sigma^* \times \mathbb{N}$.*
Για ένα στιγμιότυπο $(x,k) \in L$, το k ονομάζεται παράμετρος του προβλήματος, ενώ το $|x|+k$ ορίζεται ως το μέγεθος του στιγμιοτύπου.

Για παράδειγμα, στο πρόβλημα 3-*ΙΚΑΝΟΠΟΙΗΣΙΜΟΤΗΤΑ* ή αλλιώς 3-*SAT* δίνεται μια Boolean φόρμουλα ϕ σε κανονική διαζευκτική μορφή, όπου κάθε φράση (clause) της περιέχει το πολύ 3 λεξιγράμματα (literals) και το ζητούμενο είναι να βρεθεί μια απονομή αληθοτιμών ώστε η φόρμουλα να είναι αληθής. Ένα παραμετροποιημένο στιγμιότυπο, λοιπόν, του 3-*SAT* είναι ένα ζεύγος (ϕ,k). Τι συμβολίζει όμως η παράμετρος k; Το k μπορεί να είναι το πλήθος των μεταβλητών της ϕ ή το πλήθος των φράσεων της ϕ ή ακόμα και κάτι άλλο λιγότερο προφανές.

Επομένως, χρησιμοποιώντας κανείς την φαντασία του, μπορεί να σκεφτεί πολλές παραμέτρους για το ίδιο πρόβλημα. Ωστόσο, επιτυχημένη θεωρείται η παραμετροποίηση που ικανοποιεί δύο ιδιότητες. Ο αλγόριθμος, πρώτον, να τρέχει αρκετά γρήγορα για μικρές τιμές της παραμέτρου και δεύτερον, να τρέχει σε χρόνο $f(k)$ φορές ένα πολυώνυμο του μεγέθους της εισόδου.

Ορισμός 1.2. *Λέμε ότι ένα πρόβλημα $L \subseteq \Sigma^* \times \mathbb{N}$ είναι αποδοτικώς επιλύσιμο με σταθερή παράμετρο (fixed parameter tractable) αν υπάρχουν αλγόριθμος $\mathcal{A}$, υπολογίσιμη συνάρτηση $f : \mathbb{N} \to \mathbb{N}$ και σταθερά c τέτοια ώστε*

για δεδομένο στιγμιότυπο $(x, k) \in \Sigma^* \times \mathbb{N}$ *ο αλγόριθμος* $\mathcal{A}$ *αποφασίζει αν* $(x, k) \in L$ *σε χρόνο* $\mathcal{O}\big(f(k) \cdot |(x, k)|^c\big)$.

Η κλάση των ανωτέρων προβλημάτων συμβολίζεται με FPT και ο αλγόριθμος $\mathcal{A}$ *καλείται αλγόριθμος σταθερής παραμέτρου (fixed parameter algorithm) ή FPT αλγόριθμος.*

Για περισσότερα, παραπέμπουμε τον αναγνώστη στα [22], [5] και [18].

Κεφάλαιο 2

Μητροειδή

Θα ξεκινήσουμε το κεφάλαιο αυτό με δύο παραδείγματα.

Παράδειγμα 2.1. *Θεωρούμε τα ακόλουθα διανύσματα του* $\mathbb{R}^3$*:*

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_5 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

$$v_6 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \quad \textit{και} \quad v_7 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

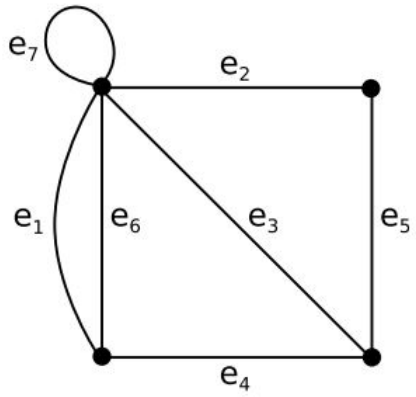
Είναι εύκολο να βρούμε ποιά από αυτά είναι γραμμικά ανεξάρτητα. Υπενθυμίζουμε ότι τα διανύσματα $v_1, \ldots, v_n$ *ονομάζονται γραμμικά ανεξάρτητα αν κανένα από αυτά δεν μπορεί να γραφεί ως γραμμικός συνδυασμός των υπολοίπων, δηλαδή αν η διανυσματική εξίσωση* $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ *έχει μοναδική λύση, την* $\lambda_1 = \cdots = \lambda_n = 0$*.*
π.χ. τα v_1, v_2, v_3 *καθώς και τα* v_2, v_4 *είναι γραμμικά ανεξάρτητα, ενώ τα* v_2, v_3, v_5 *και* v_1, v_2, v_3, v_5 *όχι.*
Δεν είναι δύσκολο να αποδείξουμε ότι ισχύουν οι ακόλουθες ιδιότητες:

- *Αν από ένα σύνολο γραμμικώς ανεξάρτητων διανυσμάτων αφαιρέσουμε ένα ή περισσότερα διανύσματα, τότε το σύνολο παραμένει γραμμικά ανεξάρτητο.*
- *Για δύο σύνολα* A *και* B *γραμμικώς ανεξάρτητων διανυσμάτων τέτοια ώστε το* B *να έχει περισσότερα διανύσματα από το* A*, υπάρχει ένα διάνυσμα* v *το οποίο ανήκει αποκλειστικά στο* B *και προσθέτοντάς το στο* A *προκύπτει ένα σύνολο γραμμικώς ανεξάρτητων διανυσμάτων. π.χ. για* $A = \{v_1, v_2, v_3\}$ *και* $B = \{v_2, v_4\}$*, το* $B \cup \{v_1\}$ *είναι προφανώς γραμμικά ανεξάρτητο.*

Το επόμενο παράδειγμα προέρχεται μια τελείως διαφορετική περιοχή των Μαθηματικών, τη Θεωρία Γραφημάτων.

Γράφημα G είναι ένα διατεταγμένο ζεύγος (V, E), όπου V είναι ένα σύνολο κορυφών ή κόμβων και E ένα σύνολο ακμών, δηλαδή γραμμών που συνδέουν δύο κορυφές. Κάθε ακμή μπορεί να θεωρηθεί ως μια διμελής σχέση μεταξύ των κόμβων. Η σχέση αυτή μπορεί να είναι συμμετρική. Αν δεν είναι, το γράφημα ονομάζεται κατευθυνόμενο.

Παράδειγμα 2.2. *Θεωρούμε το μη-κατευθυνόμενο γράφημα του παρακάτω σχήματος.*



Ένα σύνολο S ακμών του γραφήματος G ονομάζεται ανεξάρτητο αν το υπογράφημα $(V(G), S)$ είναι δάσος, δηλαδή δεν περιέχει κανέναν κύκλο. Για παράδειγμα, τα σύνολα $\{e_1, e_2, e_3\}$ και $\{e_2, e_4\}$ είναι ανεξάρτητα, ενώ τα $\{e_1, e_6\}$ και $\{e_7\}$ όχι. Όπως και στο προηγούμενο παράδειγμα, έτσι και σε αυτό, ισχύουν τα εξής:

- *Κάθε υποσύνολο ενός ανεξάρτητου συνόλου είναι κι αυτό ανεξάρτητο, αφού αφαιρώντας ακμές αποκλείεται να δημιουργηθεί κύκλος.*
- *Για δύο ανεξάρτητα σύνολα A και B, όπου το A περιλαμβάνει περισσότερες ακμές, υπάρχει τουλάχιστον μια ακμή που ανήκει μόνο στο A και προσθέτοντάς τη στο B δεν δημιουργείται κύκλος.*
  π.χ. για $A = \{e_2, e_3, e_6\}$ και $B = \{e_1, e_5\}$, το σύνολο $B \cup \{e_2\}$ είναι δένδρο, δηλαδή ανεξάρτητο.

Αυτές τις ομοιότητες παρατήρησαν τη δεκαετία του ΄30, οι Whitney [28], Birkhoff [1], Maclane [11] και Waerden [27] και αναρωτήθηκαν αν υπάρχουν και σε άλλες μαθηματικές δομές. Έτσι προέκυψε η έννοια του μητροειδούς.

Ορισμός 2.1. *Μητροειδές καλείται ένα διατεταγμένο ζεύγος $\mathcal{M} = (E, \mathcal{I})$, όπου E είναι ένα πεπερασμένο σύνολο και $\mathcal{I}$ είναι μια οικογένεια από υποσύνολα του E που ικανοποιούν τις εξής συνθήκες:*

1. *$\emptyset \in \mathcal{I}$.*
2. *Αν $A' \subseteq A$ και $A \in \mathcal{I}$, τότε $A' \in \mathcal{I}$.*
3. *Αν $A, B \in \mathcal{I}$ και $|A| < |B|$, τότε υπάρχει $e \in (B \setminus A)$ τέτοιο ώστε $A \cup \{e\} \in \mathcal{I}$.*

Τα σύνολα της οικογένειας $\mathcal{I}$ καλούνται ανεξάρτητα και το σύνολο E σύμπαν του $\mathcal{M}$.

Η 2η ιδιότητα συχνά αναφέρεται ως κληρονομική ιδιότητα (hereditary property) και η 3η ως ιδιότητα της αύξησης (augmentation property).

Τα υποσύνολα τα οποία δεν είναι ανεξάρτητα ονομάζονται εξαρτημένα. Ένα μεγιστικό σύνολο της οικογένειας $\mathcal{I}$ -δηλαδή, ένα ανεξάρτητο σύνολο το οποίο γίνεται εξαρτημένο αν του προσθέσουμε οποιοδήποτε στοιχείο- ονομάζεται **βάση** του μητροειδούς. Από τον ορισμό των βάσεων, έπεται ότι καμία βάση δεν είναι υποσύνολο κάποιας άλλης βάσης. Επίσης, για B_1, B_2 βάσεις του μητροειδούς και $e_1 \in B_1$, υπάρχει $e_2 \in B_2$ τέτοιο ώστε το σύνολο $\big(B_1 \setminus \{e_1\}\big) \cup \{e_2\}$ να είναι κι αυτό βάση. Τέλος, αποδεικνύεται ότι όλες οι βάσεις έχουν το ίδιο πλήθος στοιχείων, το οποίο ονομάζεται **τάξη** του μητροειδούς $\mathcal{M}$ και συμβολίζεται με $rank(\mathcal{M})$.

Πόρισμα 2.1. *Όλες οι βάσεις ενός μητροειδούς έχουν το ίδιο πλήθος στοιχείων.*

Απόδειξη. Έστω $\mathcal{M} = (E, \mathcal{I})$ μητροειδές και έστω B_1, B_2 δύο βάσεις του με διαφορετικό πλήθος στοιχείων, έστω $|B_1| < |B_2|$. Από το αξίωμα της ανταλλαγής προκύπτει ότι υπάρχει στοιχείο $e \in B_2 \setminus B_1$ τέτοιο ώστε $B_1 \cup \{e\} \in \mathcal{I}$, δηλαδή το B_1 δεν είναι μεγιστικό σύνολο του $\mathcal{I}$. Άτοπο. □

Παρατηρούμε ότι τόσο το μητροειδές του Παραδείγματος 2.1 όσο και αυτό του Παραδείγματος 2.2 έχουν τάξη ίση με 3.

2.1 Κλάσεις Μητροειδών

Στη παράγραφο αυτή, παρουσιάζουμε ορισμένες κλάσεις μητροειδών.

2.1.1 Γραφικά Μητροειδή

Για ένα μη-κατευθυνόμενο γράφημα $G = (V, E)$, ορίζουμε το *γραφικό (graphic)* μητροειδές $\mathcal{M}_G$. Το σύμπαν του μητροειδούς $\mathcal{M}_G$ είναι το σύνολο των ακμών $E(G)$ του γραφήματος και ένα σύνολο $A \subseteq E(G)$ είναι ανεξάρτητο αν το υπογράφημα $(V(G), A)$ είναι ακυκλικό.

Στο Παράδειγμα 2.2 ορίσαμε ένα γραφικό μητροειδές. Μπορεί, όμως, να οριστεί μητροειδές με αυτό τον τρόπο για οποιοδήποτε γράφημα;

Προφανως, οι δύο πρώτες συνθήκες του Ορισμού 2.1 ικανοποιούνται για κάθε γράφημα. Για την ιδιότητα της αύξησης, θεωρούμε ανεξάρτητα σύνολα A, $B \subseteq E(G)$ με $|A| < |B|$. Έστω μια συνεκτική συνιστώσα C του $(V(G), A)$, δηλαδή ένα υπογράφημα του $(V(G), A)$ στο οποίο κάθε κορυφή συνδέεται με κάθε άλλη μέσω ενός μονοπατιού. Έστω ότι το σύνολο B έχει πιο πολλές ακμές στη συνεκτική συνιστώσα C από ότι το σύνολο A, τότε υπάρχει κύκλος στο B. Άτοπο, αφού το B είναι ακυκλικό. Άρα σε κάθε συνεκτική συνιστώσα του $(V(G), A)$, το γράφημα $(V(G), B)$ έχει το πολύ $|A|$ ακμές. Αφού $|A| < |B|$, έπεται ότι υπάρχει μια ακμή $e \in B$ με άκρα σε διαφορετικές συνεκτικές συνιστώσες του $(V(G), A)$. Άρα, το σύνολο $A \cup \{e\}$ είναι ανεξάρτητο, και επομένως το $\mathcal{M}_G$ είναι μητροειδές για κάθε γράφημα G.

2.1.2 Ομοιόμορφα Μητροειδή

Έστω σύνολο E με n στοιχεία και η οικογένεια υποσυνόλων του E

$$\mathcal{I} = \{A \subseteq E \mid |A| \leq k\}.$$

Παρατηρούμε ότι, $\emptyset \in \mathcal{I}$ και αν $A \in \mathcal{I}$ τότε και $A' \in \mathcal{I}$, για κάθε $A' \subseteq A$. Επιπλέον, για A, $B \in \mathcal{I}$ με $|A| < |B| \leq k$ ισχύει ότι $A \cup \{e\} \in \mathcal{I}$ για κάθε $e \in B \setminus A$, αφού $|A \cup \{e\}| < k + 1$. Επομένως, το διατεταγμένο ζεύγος $(E, \mathcal{I})$ είναι μητροειδές, το οποίο καλείται *ομοιόμορφο (uniform)* και συμβολίζεται με $U_{k,n}$.

Καθώς κάθε ανεξάρτητο σύνολο του $U_{k,n}$ έχει το πολύ k στοιχεία, έπεται ότι η τάξη του είναι ίση με k. Γενικότερα, ένα ομοιόμορφο μητροειδές μπορεί να οριστεί πλήρως μόνο από το σύμπαν του E και την τάξη του k.

2.1.3 Μητροειδή Διαμέρισης

Για τον ορισμό των μητροειδών διαμέρισης, υπενθυμίζουμε τι είναι η διαμέριση ενός συνόλου.

Ορισμός 2.2. *Διαμέριση (partition) ενός συνόλου U καλείται μια οικογένεια $\mathcal{U}_P = \{U_1, \ldots, U_t\}$ υποσυνόλων του U τέτοια ώστε $\bigcup_{i=1}^{t} U_i = U$ και $U_i \neq U_j$ για κάθε $i \neq j$.*

Τα σύνολα U_i καλούνται μέρη της διαμέρισης.

Έστω σύνολο E, μια διαμέριση αυτού σε t μέρη $E_1, \ldots, E_t$ και t μη-αρνητικοί ακέραιοι $k_1, \ldots, k_t$. Ορίζουμε την οικογένεια $\mathcal{I}$ υποσυνόλων του E,

$$\mathcal{I} = \big\{ A \subseteq E \mid |A \cap E_i| \leq k_i \quad \forall i \in \{1, \ldots, t\} \big\}.$$

Θα αποδείξουμε ότι το διατεταγμένο ζεύγος $\mathcal{M} = (E, \mathcal{I})$ είναι μητροειδές:

1. Αφού $|\emptyset \cap E_i| = 0 \leq k_i \quad \forall i \in \{1, \ldots t\}$, έπεται ότι $\emptyset \in \mathcal{I}$.
2. Έστω $A \in \mathcal{I}$ και $A' \subseteq A$. Αφού $A \in \mathcal{I}$, $|A \cap E_i| \leq k_i \ \forall i \in \{1, \ldots t\}$. Άρα $|A' \cap E_i| \leq |A \cap E_i| \leq k$, δηλαδή $A' \in \mathcal{I}$.
3. Έστω $A, B \in \mathcal{I}$ με $|A| < |B|$. Τότε υπάρχει κάποιο $i \in \{1, \ldots, t\}$ τέτοιο ώστε $|A \cap E_i| < |B \cap E_i|$ και επομένως για κάθε στοιχείο $e \in (B \setminus A) \cap E_i$ ισχύει $A \cup \{e\} \in \mathcal{I}$.

Το $\mathcal{M} = (E, \mathcal{I})$ ονομάζεται *μητροειδές διαμέρισης (partition matroid)*. Παρατηρούμε ότι για τον ορισμό του χρειαζόμαστε μόνο τη διαμέριση του συνόλου E και τους ακεραίους $k_1, \ldots, k_t$.

2.1.4 Γραμμικά Μητροειδή και Αναπαράσταση

Τα μητροειδή, σαν κι αυτό του Παραδείγματος 2.1, τα οποία χρησιμοποιούν την έννοια της γραμμικής ανεξαρτησίας ονομάζονται *γραμμικά (linear)*. Θεωρούμε ένα σύνολο E και σε κάθε στοιχείο του αντιστοιχούμε ένα διάνυσμα v_e επί του σώματος $\mathbb{F}$. Τα διανύσματα των διαφορετικών στοιχείων του E ανήκουν όλα στο ίδιο σώμα και έχουν την ίδια διάσταση. Η οικογένεια $\mathcal{I}$ των ανεξάρτητων συνόλων του μητροειδούς αποτελείται από όλα τα σύνολα $A \subseteq E$ για τα οποία το $\{v_e \mid e \in A\}$ είναι σύνολο γραμμικά ανεξάρτητων διανυσμάτων.

Προφανώς, για το διατεταγμένο ζεύγος $(E, \mathcal{I})$ ισχύουν οι δύο πρώτες συνθήκες του Ορισμού 2.1. Για την ιδιότητα της αύξησης, θεωρούμε A, $B \in \mathcal{I}$ με $|A| < |B|$. Τότε το σύνολο $\{v_e \mid e \in A\}$ παράγει χώρο διάστασης $|A|$, ενώ το $\{v_e \mid e \in B\}$ παράγει χώρο διάστασης $|B|$. Έπεται ότι υπάρχει $b \in B$ τέτοιο ώστε το διάνυσμα v_b να μην ανήκει στον γραμμικό χώρο του $\{v_e \mid e \in A\}$ και άρα $A \cup \{b\} \in \mathcal{I}$. Επομένως, το διατεταγμένο ζεύγος $(E, \mathcal{I})$ είναι μητροειδές.

Ορίζουμε πίνακα M με μια στήλη για κάθε $e \in E$ τέτοιο ώστε το διάνυσμα-στήλη που αντιστοιχεί στο στοιχείο e να είναι το v_e. Τότε λέμε ότι ο πίνακας M *αναπαριστά* το μητροειδές $\mathcal{M} = (E, \mathcal{I})$ επί του σώματος $\mathbb{F}$ και το μητροειδές $\mathcal{M} = (E, \mathcal{I})$ καλείται *αναπαριστάσιμο επί του* $\mathbb{F}$. Στο σημείο αυτό σημειώνουμε ότι δεν είναι όλα τα μητροειδή αναπαραστάσιμα, αφού τότε η έννοια του μητροειδούς θα ήταν ισοδύναμη με την γραμμική ανεξαρτησία και άρα άνευ νοήματος. Ωστόσο, τα διάφορα μητροειδή που έχουμε ορίσει είναι όλα αναπαραστάσιμα επί κάποιου σώματος.

Πρόταση 2.1. *Ένα γραφικό μητροειδές είναι αναπαραστάσιμο σε οποιοδήποτε σώμα με τουλάχιστον 2 στοιχεία.*

Απόδειξη. Θεωρούμε τον πίνακα $M_{\mathcal{M}}$, όπου οι γραμμές του αντιστοιχούν στις κορυφές του γραφήματος και οι στήλες του στις ακμές. Στη στήλη που αντιστοιχεί στην ακμή $e = (i, j)$, όλες οι τιμές είναι ίσες με 0 εκτός από αυτές στην θέση i που είναι ίση με 1 και στην θέση j που είναι ίση με -1. Αυτή είναι μια αναπαράσταση του μητροειδούς στους πραγματικούς αριθμούς. Για αναπαράσταση πάνω σε οποιοδήποτε σώμα $\mathbb{F}$, αρκεί να αντικαταστήσουμε το -1 με τον αντίθετο του 1 στο $\mathbb{F}$. □

Πρόταση 2.2. *Ένα ομοιόμορφο μητροειδές $U_{k,n}$ είναι αναπαραστάσιμο επί του πεπερασμένου σώματος $GF(s)$, για κάθε $s > n$.*

Απόδειξη. Θεωρούμε $E = \{e_1, e_2, \ldots, e_n\}$, το σύμπαν του $U_{k,n}$. Σε κάθε e_i αντιστοιχούμε ένα μη-μηδενικό στοιχείο α_i του σώματος $GF(s)$. Θέτουμε $v_i = (1, \alpha_i^1, \alpha_i^2, \ldots, \alpha_i^{k-1})$ να είναι το διάνυσμα που αντιστοιχεί στο e_i.

Αφού έχουμε k-διάστατα διανύσματα, έπεται ότι για κάθε $A \subseteq E$ με $|A| > k$ το σύνολο $\{v_i \mid e_i \in A\}$ είναι γραμμικά εξαρτημένο. Θέλουμε να δείξουμε ότι για κάθε $A \subseteq E$ με $|A| = k$, το σύνολο $\{v_i \mid e_i \in A\}$ είναι γραμμικά ανεξάρτητο. Προφανώς, το ίδιο ισχύει και για κάθε σύνολο με λιγότερα στοιχεία. Άρα, αρκεί να δείξουμε ότι κάθε $k \times k$ υποπίνακας M_A του πίνακα

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & \cdots & n \\ 1 & 2^2 & 3^2 & \cdots & n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2^{k-1} & 3^{k-1} & \ldots & n^{k-1} \end{pmatrix}$$

έχει μη-μηδενική ορίζουσα. Γνωρίζουμε, όμως, ότι η ορίζουσα του πίνακα M_A είναι ίση με

$$\prod_{\substack{0 \le i < j \le n \\ e_i \in A, e_j \in A}} \alpha_j - \alpha_i,$$

δηλαδή είναι γινόμενο μη-μηδενικών στοιχείων του σώματος $GF(s)$. Άρα, η ορίζουσα του πίνακα M_A, καθώς και κάθε υποπίνακά του, είναι μη-μηδενική. Έπεται ότι το σύνολο $\{v_i \mid e_i \in A\}$ είναι γραμμικά ανεξάρτητο. □

Πρόταση 2.3. *Ένα μητροειδές διαμέρισης είναι αναπαραστάσιμο επί του πεπερασμένου σώματος $GF(s)$, για κάθε $s > \max_{i \le k} (|E_i| + 1)$.*

Η απόδειξη είναι άμεση συνέπεια της προηγούμενης Πρότασης, του Παραδείγματος 2.3 και της Πρότασης 2.4.

2.2 Πράξεις σε Μητροειδή

2.2.1 Ευθύ Άθροισμα Μητροειδών

Έστω τα ξένα ανά δύο σύνολα $E_1, \ldots, E_t$ και τα μητροειδή

$$\mathcal{M}_1 = (E_1, \mathcal{I}_1),\ \mathcal{M}_2 = (E_2, \mathcal{I}_2),\ \ldots,\ \mathcal{M}_t = (E_t, \mathcal{I}_t).$$

Ορίζουμε το *ευθύ άθροισμα (direct sum)* $\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_t = (E, \mathcal{I})$, θέτοντας $E = \bigcup_{i=1}^{t} E_i$ και

$$\mathcal{I} = \big\{X \subseteq E \mid X \cap E_i \in \mathcal{I}_i \quad \forall i \in \{1, \ldots, t\}\big\}.$$

Το $\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_t$ είναι μητροειδές. Προφανώς, $\emptyset \in \mathcal{I}$ και αν $A \in \mathcal{I}$ τότε και $A' \in \mathcal{I}$ για κάθε $A' \subseteq A$. Για την ιδιότητα της αύξησης, θεωρούμε A, $B \in \mathcal{I}$ με $|A| < |B|$. Επειδή $|A| < |B|$, υπάρχει $j \in \{1, \ldots, t\}$ τέτοιο ώστε $|A \cap E_j| < |B \cap E_j|$. Επίσης, $A \cap E_j \in \mathcal{I}_j$ και $B \cap E_j \in \mathcal{I}_j$. Άρα υπάρχει $e \in (B \cap E_j) \setminus (A \cap E_j)$ τέτοιο ώστε $\big(A \cup \{e\}\big) \cap E_j \in \mathcal{I}_j$ και $\big(A \cup \{e\}\big) \cap E_i \in \mathcal{I}_i$, $\forall i \neq j$. Έπεται ότι $A \cup \{e\} \in \mathcal{I}$.

Παράδειγμα 2.3. *Στο ευθύ άθροισμα των ομοιόμορφων μητροειδών $U_{|E_1|,k_1}, U_{|E_2|,k_2}, \ldots, U_{|E_t|,k_t}$, το σύμπαν είναι ίσο με $E = \bigcup_{i=1}^{t} E_i$ και ένα σύνολο A είναι ανεξάρτητο αν για κάθε $i \in \{1, \ldots, t\}$, $A \cap E_i \in \mathcal{I}_i$, δηλαδή $|A \cap E_i| \leq k_i$. Άρα, ένα μητροειδές διαμέρισης μπορεί να οριστεί ως ευθύ άθροισμα ομοιόμορφων μητροειδών.*

Έστω ότι τα μητροειδή είναι αναπαραστάσιμα όλα επί του ίδιου σώματος $\mathbb{F}$ και έστω M_i ο πίνακας αναπαράστασης του $\mathcal{M}_i = (E, \mathcal{I}_i)$ για κάθε $i \in \{1, \ldots, t\}$. Τότε, ο πίνακας

$$M = \begin{pmatrix} M_1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & M_t \end{pmatrix}$$

είναι ο πίνακας αναπαράστασης του $\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_t$.
Για την ορθότητα της κατασκευής, παραπέμπουμε τον αναγνώστη στο [13].

Πρόταση 2.4. *Έστω $\mathcal{M}_1, \ldots, \mathcal{M}_t$ μητροειδή αναπαραστάσιμα επί του σώματος $\mathbb{F}$. Τότε μια αναπαράσταση του $\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_t$ μπορεί να βρεθεί σε πολυωνυμικό χρόνο.*

2.2.2 Περιορισμός Μητροειδούς

Ως *t-περιορισμός (t-truncation)* ενός μητροειδούς $\mathcal{M} = (E, \mathcal{I})$ ορίζεται το μητροειδές $\mathcal{M}' = (E, \mathcal{I}')$ όπου ένα σύνολο $A \subseteq E$ είναι ανεξάρτητο στο $\mathcal{M}'$ αν και μόνο αν το A είναι ανεξάρτητο στο $\mathcal{M}$ και $|A| \leq t$.

Το $\mathcal{M}'$ είναι μητροειδές.

1. $\emptyset \in \mathcal{I}$ και $|\emptyset| = 0 \leq t$. Άρα $\emptyset \in \mathcal{M}'$.
2. Έστω $A \in \mathcal{I}'$ και $A' \subseteq A$. Τότε $A \in \mathcal{I}$ και $|A| \leq t$, δηλαδή $A' \in \mathcal{I}$ και $|A'| \leq |A| \leq t$, δηλαδή $A' \in \mathcal{I}'$.
3. Έστω A, $B \in \mathcal{I}'$ με $|A| < |B|$. Έπεται ότι A, $B \in \mathcal{I}$ και $|A| < |B| \leq t$. Άρα υπάρχει $e \in B \setminus A$ τέτοιο ώστε $A \cup \{e\} \in \mathcal{I}$ και $|A \cup \{e\}| \leq t$, αφού $|A| < t$. Άρα $A \cup \{e\} \in \mathcal{I}'$.

Πρόταση 2.5. *Έστω $\mathcal{M}$ μητροειδές, M ο πίνακας αναπαράστασής του επί του πεπερασμένου σώματος $GF(s)$ και ακέραιος t. Τότε υπάρχει πιθανοτικός πολυωνυμικός αλγόριθμος, ο οποίος υπολογίζει μια αναπαράσταση του t-περιορισμού $\mathcal{M}'$ του $\mathcal{M}$.*

Για την απόδειξη, θα χρειαστούμε το παρακάτω λήμμα, το οποίο αποδείχθηκε ανεξάρτητα από τους DeMillo και Lipton στο [9], Zippel στο [36] και Schwartz στο [25].

Λήμμα 2.1 (Λήμμα Zippel-Schwartz). *Έστω $\mathbb{F}$ σώμα και $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ πολυώνυμο βαθμού το πολύ d, όχι ταυτοτικά μηδέν. Έστω, επίσης, ένα πεπερασμένο υποσύνολο S του $\mathbb{F}$. Διαλέγουμε τιμές $\alpha_1, \ldots \alpha_n$ του S, την κάθε μια με πιθανότητα $\frac{1}{|S|}$. Τότε*

$$\Pr\big(p(\alpha_1, \ldots, \alpha_n) = 0\big) \leq \frac{d}{|S|}.$$

Απόδειξη της Πρότασης 2.5. Έστω $s \geq xt$ για κάποιο $x \geq 2$. Θεωρούμε ότι το μητροειδές $\mathcal{M}$ έχει τάξη r, δηλαδή ο πίνακας M έχει διαστάσεις $r \times n$. Επίσης, θεωρούμε τον πίνακα R διαστάσεων $t \times r$, όπου κάθε στοιχείο του διαλέγεται με πιθανότητα $\frac{1}{s}$ από το σώμα $GF(s)$.

Ισχυριζόμαστε ότι με μεγάλη πιθανότητα, το μητροειδές $\mathcal{M}'$ με πίνακα αναπαράστασης τον RM είναι t-περιορισμός του μητροειδούς $\mathcal{M}$. Αφού ο $t \times n$ πίνακας RM δεν μπορεί να έχει περισσότερες από t ανεξάρτητες στήλες, αρκεί να δείξουμε ότι ένα σύνολο με t στοιχεία είναι ανεξάρτητο στο $\mathcal{M}'$ αν και μόνο αν είναι ανεξάρτητο στο $\mathcal{M}$. Πράγματι, έστω $X \subseteq E$ με $|X| = t$ και M_0, ο $r \times t$ υποπίνακας του M που αποτελείται από τις στήλες που αντιστοιχούν στα στοιχεία του X. Θέτουμε $B_0 = RM_0$. Αν το X δεν είναι ανεξάρτητο στο $\mathcal{M}$ τότε οι στήλες του M_0 δεν είναι ανεξάρτητες και επομένως ούτε και οι στήλες του B_0. Άρα το X δεν είναι ανεξάρτητο στο $\mathcal{M}'$. Αντίστροφα, έστω X ανεξάρτητο στο $\mathcal{M}$. Τότε οι στήλες του M_0 είναι ανεξάρτητες και επομένως $\det(M_0) \neq 0$. Έπεται ότι $\Pr[\det(RM_0) \neq 0] > 0$. Αφού ο πίνακας R είναι τυχαίος μπορούμε να θεωρήσουμε τα tr στοιχεία του ως μεταβλητές και επομένως την ορίζουσα $det(RM_0)$ ως πολυώνυμο t βαθμού. Από το Λήμμα 2.1 έχουμε ότι

$$\Pr[\det(RM_0) = 0] \leq \frac{t}{s} \leq \frac{1}{x}.$$

Έπεται ότι με πιθανότητα τουλάχιστον $1 - \frac{1}{x}$ ο πίνακας RM αναπαριστά τον t-περιορισμό του $\mathcal{M}$. □

2.2.3 Ένωση και Τομή Μητροειδών

Έστω τα μητροειδή $\mathcal{M}_1 = (E_1, \mathcal{I}_1)$, ..., $\mathcal{M}_t = (E_t, \mathcal{I}_t)$. Ορίζουμε την *ένωση* τους ως

$$\mathcal{M}_1 \vee \cdots \vee \mathcal{M}_t = (E, \mathcal{I}),$$

όπου $E = \bigcup_{i=1}^{t} E_i$ και $\mathcal{I} = \{I_1 \cup \cdots \cup I_t \mid I_i \in \mathcal{I}_i \quad \forall i \in \{1, \ldots, t\}\}$.

Στη συνέχεια, διακρίνουμε δύο περιπτώσεις και δείχνουμε ότι και στις δύο το διατεταγμένο ζεύγος $(E, \mathcal{I})$ είναι μητροειδές.

Έστω ότι τα σύνολα $E_1, \ldots, E_t$ είναι ξένα ανά δύο. Θα δείξουμε ότι ισχύουν οι τρεις συνθήκες του Ορισμού 2.1. Προφανώς, $\emptyset \in \mathcal{I}$ και αν $A \in \mathcal{I}$ τότε και $A' \in \mathcal{I}$, για κάθε $A' \subseteq A$. Για την ιδιότητα της αύξησης, θεωρούμε σύνολα $A, B \in \mathcal{I}$ με $|A| < |B|$. Τότε $A = \bigcup_{i=1}^{t} A_i$ και $B = \bigcup_{i=1}^{t} B_i$, όπου $A_i, B_i \in \mathcal{I}_i$, $\forall i \in \{1, \ldots, t\}$. Τα $A_1, \ldots, A_t$ καθώς και τα $B_1, \ldots, B_t$ είναι ξένα ανά δύο και επειδή $|A| < |B|$, υπάρχει κάποιο $j \in \{1, \ldots, t\}$ τέτοιο ώστε $|A_j| < |B_j|$. Άρα υπάρχει $e \in B_j \setminus A_j$ τέτοιο ώστε $A_j \cup \{e\} \in \mathcal{I}_j$ και επειδή $e \notin \mathcal{I}_i$ $\forall i \neq j$, έπεται ότι $A \cup \{e\} \in \mathcal{I}$.

Έστω ότι τα σύνολα $E_1, \ldots, E_t$ δεν είναι ξένα ανά δύο. Θα χρειαστούμε το ακόλουθο λήμμα.

Λήμμα 2.2. *Έστω $\mathcal{M}' = (E', \mathcal{I}')$ μητροειδές, σύνολο E' και αμφιμονοσήμαντη συνάρτηση $f : E' \to E$. Ορίζουμε την οικογένεια $\mathcal{I} = \{f(I') \mid I' \in \mathcal{I}'\}$. Τότε το διατεταγμένο ζεύγος $\mathcal{M} = (E, \mathcal{I})$ είναι μητροειδές.*

Απόδειξη. Αποδεικνύουμε τις τρεις ιδιότητες του Ορισμού 2.1.

1. $f(\emptyset) = \emptyset$, και άρα $\emptyset \in \mathcal{I}$.

2. Έστω $A \in \mathcal{I}$ και $B \subseteq A$. Τότε, αφού $A \in \mathcal{I}$, υπάρχει $A' \in \mathcal{I}'$ τέτοιο ώστε $f(A') = A$. Έπεται ότι για κάθε $e \in A$, $f^{-1}(e) \cap A' \neq \emptyset$. Θέτουμε $B' = \{e' \in A' \mid f(e') \in B\}$. Τότε $B = f(B')$ και αφού $B' \subseteq A'$, $B' \in \mathcal{I}'$ και άρα $B \in \mathcal{I}$.

3. Έστω $A, B \in \mathcal{I}$ με $|A| < |B|$. Έστω A' το ελάχιστο σύνολο για το οποίο ισχύει $f(A') = A$. Αντίστοιχα, έστω B' το ελάχιστο σύνολο τέτοιο ώστε $f(B') = B$. Έπεται ότι $|A'| = |A|$ και $|B'| = |B|$. Αφού $A', B' \in \mathcal{I}'$ και $|A'| < |B'|$, υπάρχει $e' \in B' \setminus A'$ τέτοιο ώστε $A' \cup \{e'\} \in \mathcal{I}'$. Οπότε, $A \cup f(e') \in \mathcal{I}$ και $f(e') \in B \setminus A$.

Έπεται ότι το διατεταγμένο ζεύγος $\mathcal{M} = (E, \mathcal{I})$ είναι μητροειδές. □

Οπότε, αν τα σύνολα $E_1, \ldots, E_t$ δεν είναι ξένα ανά δύο, θεωρούμε τα αντίγραφα τους $E'_1, \ldots, E'_t$, τα οποία είναι ξένα ανά δύο. Έστω τα μητροειδή $\mathcal{M}'_i = (E'_i, \mathcal{I}')$, όπου $\mathcal{I}'_i = \mathcal{I}_i$. Θέτουμε $E' = \bigcup_{i=1}^{t} E'_i$ και ορίζουμε το διατεταγμένο ζεύγος $\mathcal{M}' = (E', \mathcal{I}')$, όπου $\mathcal{I}' = \{I'_1 \cup \cdots \cup I'_k \mid I'_i \in \mathcal{I}_i\}$. Προφανώς, το $\mathcal{M}'$ είναι μητροειδές. Τέλος, ορίζουμε συνάρτηση $f : E' \to E$, όπου $E = \bigcup_{i=1}^{t} E_i$ και $f(e') = e$, αν το e' είναι αντίγραφο του e. Από το παραπάνω λήμμα, έπεται ότι το $\mathcal{M}$ είναι μητροειδές.

Σε αντίθεση με την ένωση, η τομή των μητροειδών δεν είναι μητροειδές, ακόμα και όταν τα μητροειδή έχουν κοινό σύμπαν E. Για παράδειγμα, έστω τα μητροειδή $\mathcal{M}_1 = (E, \mathcal{I}_1)$ και $\mathcal{M}_2 = (E, \mathcal{I}_2)$ με σύμπαν $E = \{a, b, c, d\}$,

$$\mathcal{I}_1 = \big\{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{b, d\}, \{c, d\}\big\}$$
$$\text{και } \mathcal{I}_2 = \big\{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, d\}, \{b, c\}, \{c, d\}\big\}$$

Παρατηρούμε ότι η τομή τους δεν είναι μητροειδές. Πράγματι, το $\mathcal{M} = (E, \mathcal{I}_1 \cap \mathcal{I}_2)$, όπου

$$\mathcal{I}_1 \cap \mathcal{I}_2 = \big\{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{c, d\}\big\}$$

δεν ικανοποιεί το αξίωμα της ανταλλαγής, αφού $\{c\} \in \mathcal{I}_1 \cap \mathcal{I}_2$ και $\{a, b\} \in \mathcal{I}_1 \cap \mathcal{I}_2$, ενώ $\{a, c\} \notin \mathcal{I}_1 \cap \mathcal{I}_2$ και $\{b, c\} \notin \mathcal{I}_1 \cap \mathcal{I}_2$.

Κεφάλαιο 3

Σύνολα Αντιπροσώπευσης

Έστω $\mathcal{S} = \{S_1, S_2, \cdots S_t\}$ οικογένεια υποσυνόλων ενός συνόλου E. Θα λέμε ότι η $\mathcal{S}$ είναι p-οικογένεια αν κάθε σύνολο της έχει p στοιχεία, δηλαδή $|S_i| = p \ \forall i \in \{1, \ldots, t\}$.

Η έννοια των συνόλων αντιπροσώπευσης ορίστηκε για πρώτη φορά από τον Monien στο [14]. Ο αρχικός ορισμός ήταν ο εξής:

Ορισμός 3.1. *Έστω σύνολο E και $\mathcal{S}$ p-οικογένεια υποσυνόλων του E. Λέμε ότι η οικογένεια $\mathcal{S}' \subseteq \mathcal{S}$ q-αντιπροσωπεύει την $\mathcal{S}$ αν ισχύει το εξής:*

> *για κάθε σύνολο $Y \subseteq E$ μεγέθους q για το οποίο υπάρχει σύνολο $X \in \mathcal{S}$ με $X \cap Y = \emptyset$, υπάρχει σύνολο $X' \in \mathcal{S}'$ με $X' \cap Y = \emptyset$.*

Για παράδειγμα, έστω $\mathcal{S}$ η οικογένεια που αποτελείται από τα δισύνολα

$$\{2,4\}, \{1,5\}, \{1,6\}, \{1,7\}, \{3,6\}, \{3,8\}, \{4,8\} \text{ και } \{4,7\}.$$

Η υποοικογένεια $\mathcal{S}' = \big\{\{2,4\}, \{1,5\}\big\}$ 1-αντιπροσωπεύει την $\mathcal{S}$. Πράγματι, αφού η $\mathcal{S}'$ περιέχει δύο ξένα σύνολα, έπεται ότι για κάθε σύνολο $Y \subseteq [8]$ με $|Y| = 1$, είτε $\{2,4\} \cap Y = \emptyset$ είτε $\{1,5\} \cap Y = \emptyset$. Με τον ίδιο τρόπο, μπορούμε να δείξουμε ότι η $\mathcal{S}'' = \big\{\{2,4\}, \{1,5\}, \{3,6\}\big\}$ 2-αντιπροσωπεύει την $\mathcal{S}$ και η $\mathcal{S}''' = \big\{\{2,4\}, \{1,5\}, \{3,6\}, \{1,7\}, \{3,8\}, \{4,8\}\big\}$ την 3-αντιπροσωπεύει.

Αρκετά χρόνια αργότερα, ο Marx στο [13] έδωσε έναν πιο γενικευμένο ορισμό των συνόλων αντιπροσώπευσης, ο οποίος χρησιμοποιεί και την έννοια του μητροειδούς.

Ορισμός 3.2. *Έστω μητροειδές $\mathcal{M} = (E, \mathcal{I})$ και $\mathcal{S}$ p-οικογένεια υποσυνόλων του E. Λέμε ότι η οικογένεια $\mathcal{S}' \subseteq \mathcal{S}$ είναι q-αντιπροσωπευτική (q-representative) της $\mathcal{S}$ αν ισχύει το εξής:*

> *για κάθε σύνολο $Y \subseteq E$ μεγέθους το πολύ q για το οποίο υπάρχει σύνολο $X \in \mathcal{S}$ με $X \cap Y = \emptyset$ και $X \cup Y \in \mathcal{I}$, υπάρχει σύνολο $X' \in \mathcal{S}'$ με $X' \cap Y = \emptyset$ και $X' \cup Y \in \mathcal{I}$.*

Αν η $\mathcal{S}'$ είναι q-αντιπροσωπευτική της $\mathcal{S}$ γράφουμε $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$.

Παρατηρήστε ότι, αν το μητροειδές $\mathcal{M}$ στον Ορισμό 3.2 είναι το ομοιόμορφο μητροειδές τάξης τουλάχιστον $p + q$ τότε οι δύο ορισμοί ταυτίζονται. Από εδώ και στο εξής, θα χρησιμοποιούμε τον Ορισμό 3.1 όταν αναφερόμαστε σε ομοιόμορφα μητροειδή και τον Ορισμό 3.2 σε όλες τις υπόλοιπες περιπτώσεις.

Στη συνέχεια θα διατυπώσουμε και θα αποδείξουμε τρία βασικά λήμματα για τα σύνολα αντιπροσώπευσης.

Λήμμα 3.1. *Έστω $\mathcal{M} = (E, \mathcal{I})$ μητροειδές και $\mathcal{S}$ p-οικογένεια υποσυνόλων του E. Αν $\widehat{\mathcal{S}} \subseteq^q_{rep} \mathcal{S}$ και $\mathcal{S}' \subseteq^q_{rep} \widehat{\mathcal{S}}$, τότε $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$.*

Απόδειξη. Έστω $Y \subseteq E$ μεγέθους το πολύ q τέτοιο ώστε να υπάρχει $X \in \mathcal{S}$ με $X \cap Y = \emptyset$ και $X \cup Y \in \mathcal{I}$. Επειδή $\widehat{\mathcal{S}} \subseteq^q_{rep} \mathcal{S}$, υπάρχει $\widehat{X} \in \widehat{\mathcal{S}}$ με $\widehat{X} \cap Y = \emptyset$ και $\widehat{X} \cup Y \in \mathcal{I}$. Επίσης, επειδή $\mathcal{S}' \subseteq^q_{rep} \widehat{\mathcal{S}}$, υπάρχει $X' \in \mathcal{S}'$ με $X' \cap Y$ και $X' \cup Y \in \mathcal{I}$. Άρα $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$. □

Λήμμα 3.2. *Έστω $\mathcal{M} = (E, \mathcal{I})$ μητροειδές και $\mathcal{S}$ p-οικογένεια υποσυνόλων του E. Αν $\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_k$ και $\mathcal{S}'_i \subseteq^q_{rep} \mathcal{S}_i$ για κάθε $1 \leq i \leq k$, τότε $\bigcup_{i=1}^{k} \mathcal{S}'_i \subseteq^q_{rep} \mathcal{S}$.*

Απόδειξη. Έστω $Y \subseteq E$ μεγέθους το πολύ q τέτοιο ώστε να υπάρχει $X \in \mathcal{S}$ με $X \cap Y = \emptyset$ και $X \cup Y \in \mathcal{I}$. Αφού $\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_k$, υπάρχει ένα i τέτοιο ώστε $X \in \mathcal{S}_i$. Έπεται ότι υπάρχει $X' \in \mathcal{S}'_i \subseteq \bigcup_{i=1}^{k} \mathcal{S}'_i$ με $X' \cap Y = \emptyset$ και $X' \cup Y \in \mathcal{I}$. Άρα $\bigcup_{i=1}^{k} \mathcal{S}'_i \subseteq^q_{rep} \mathcal{S}$. □

Το τρίτο λήμμα αναφέρεται στην παρακάτω πράξη, η οποία ορίζεται για οικογένειες συνόλων $\mathcal{S}_1$ και $\mathcal{S}_2$:

$$\mathcal{S}_1 \bullet \mathcal{S}_2 = \{S_1 \cup S_2 \mid S_1 \in \mathcal{S}_1 \text{ και } S_2 \in \mathcal{S}_2 \text{ και } S_1 \cap S_2 = \emptyset\}$$

Λήμμα 3.3. *Έστω $\mathcal{M} = (E, \mathcal{I})$ μητροειδές τάξης k και $\mathcal{S}_1, \mathcal{S}_2$ οικογένειες υποσυνόλων του E με p_1 και p_2 στοιχεία αντίστοιχα. Αν $\mathcal{S}'_1 \subseteq^{k-p_1}_{rep} \mathcal{S}_1$ και $\mathcal{S}'_2 \subseteq^{k-p_2}_{rep} \mathcal{S}_2$, τότε $\mathcal{S}'_1 \bullet \mathcal{S}'_2 \subseteq^{k-p_1-p_2}_{rep} \mathcal{S}_1 \bullet \mathcal{S}_2$.*

Απόδειξη. Έστω $Y \subseteq E$ μεγέθους το πολύ $q = k - p_1 - p_2$ τέτοιο ώστε να υπάρχει $X \in \mathcal{S}_1 \bullet \mathcal{S}_2$ τέτοιο ώστε $X \cap Y = \emptyset$ και $X \cup Y \in \mathcal{I}$. Έπεται ότι υπάρχουν $S_1 \in \mathcal{S}_1$ και $S_2 \in \mathcal{S}_2$ τέτοια ώστε $S_1 \cup S_2 = X$ και $S_1 \cap S_2 = \emptyset$. Αφού $\mathcal{S}'_1 \subseteq^{k-p_1}_{rep} \mathcal{S}_1$, υπάρχει $S'_1 \in \mathcal{S}'_1$ τέτοιο ώστε $S'_1 \cap (S_2 \cup Y) = \emptyset$ και $S'_1 \cup (S_2 \cup Y) \in \mathcal{I}$. Επίσης, αφού $\mathcal{S}'_2 \subseteq^{k-p_2}_{rep} \mathcal{S}_2$, υπάρχει $S'_2 \in \mathcal{S}'_2$ τέτοιο ώστε $S'_1 \cap (S'_2 \cup Y) = \emptyset$ και $S'_1 \cup S'_2 \cup Y \in \mathcal{I}$. Άρα $S'_1 \cup S'_2 \in \mathcal{S}'_1 \bullet \mathcal{S}'_2$ και $(S'_1 \cup S'_2) \cup Y \in \mathcal{I}$, οπότε $\mathcal{S}'_1 \bullet \mathcal{S}'_2 \subseteq^{k-p_1-p_2}_{rep} \mathcal{S}_1 \bullet \mathcal{S}_2$. □

3.1 Γρήγορος Υπολογισμός Συνόλων Αντιπροσώπευσης

Σε αυτή την παράγραφο, παρουσιάζουμε έναν αλγόριθμο, ο οποίος με είσοδο μια οικογένεια συνόλων υπολογίζει μια q-ανιπροσωπευτική οικογένεια αυτής. Ο αλγόριθμος βασίζεται σε σχετική απόδειξη του Lovász. Ο Lovász στο [10] απέδειξε ότι, σε οποιοδήποτε γραμμικό μητροειδές, κάθε p-οικογένεια έχει μια q-αντιπροσωπευτική οικογένεια με το πολύ $\binom{p+q}{p}$ σύνολα. Οι Fomin, Lokshtanov και Saurabh στο [21] έδωσαν το παρακάτω θεώρημα, το οποίο μετατρέπει την απόδειξη του Lovász σε αλγόριθμο.

Ως ω συμβολίζουμε τον εκθέτη για τον πολλαπλασιασμό πινάκων. Μέχρις στιγμής, το καλύτερο άνω φράγμα που γνωρίζουμε για αυτό είναι μικρότερο του 2.373 [29].

Θεώρημα 3.1. *Έστω $\mathcal{M} = (E, \mathcal{I})$ γραμμικό μητροειδές τάξης $p+q = k$, $A_{\mathcal{M}}$ πίνακας αναπαράστασής του επί του σώματος $\mathbb{F}$ και $\mathcal{S} = \{S_1, \ldots, S_t\}$ p-οικογένεια ανεξάρτητων υποσυνόλων του E. Τότε υπάρχει αλγόριθμος που υπολογίζει μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ μεγέθους το πολύ $\binom{p+q}{p}$, σε $\mathcal{O}\Big(\binom{p+q}{p} tp^{\omega} + t\binom{p+q}{q}^{\omega-1}\Big)$ χρόνο, όπου ως μονάδα χρόνου θεωρούμε μια πράξη στο $\mathbb{F}$.*

Με άλλα λόγια, η q-αντιπροσωπευτική οικογένεια δεν χρειάζεται να έχει ποτέ πάνω από $\binom{p+q}{p}$ σύνολα, ενώ μπορεί να υπολογιστεί σε χρόνο πολυωνυμικό ως προς το μέγεθος της αρχικής οικογένειας.

Για την απόδειξη θα χρειαστούμε την γενικευμένη επέκταση του Laplace (generalized Laplace expansion) για ορίζουσες. Συμβολίζουμε με $R(A)$ και $C(A)$ το σύνολο των γραμμών και των στηλών του πίνακα $A = (\alpha_{i,j})$, αντίστοιχα. Για $I \subseteq R(A)$ και $J \subseteq C(A)$, συμβολίζουμε με $A[I, J] = \big(\alpha_{i,j} \mid i \in I, j \in J\big)$ τον υποπίνακα του A με γραμμές τις γραμμές του συνόλου I και στήλες αυτές του συνόλου J. Τέλος, για $I \subseteq [n]$ θέτουμε $\bar{I} = [n] \setminus I$ και $\sum I = \sum_{i \in I} i$.

Πρόταση 3.1. *Για $n \times n$ πίνακα A και $J \subseteq C(A) = [n]$, ισχύει*

$$\det(A) = \sum_{I \subseteq [n], |I| = |J|} (-1)^{\sum I + \sum J} \det(A[I, J]) \cdot \det(A[\bar{I}, \bar{J}]).$$

Για την απόδειξη της παραπάνω ταυτότητας, ο αναγνώστης παραπέμπεται στο [15].

Επίσης, θεωρούμε ότι ο πίνακας αναπαράστασης $A_{\mathcal{M}}$ του μητροειδούς $\mathcal{M}$ έχει $rank(\mathcal{M})$ γραμμές, καθώς αν δεν έχει εφαρμόζουμε απαλοιφή Gauss και παίρνουμε πίνακα με $rank(\mathcal{M})$ γραμμές σε πολυωνυμικό χρόνο. Τώρα, είμαστε έτοιμοι να δώσουμε την απόδειξη του Θεωρήματος 3.1.

Απόδειξη του Θεωρήματος 3.1. Αν $t \leq \binom{k}{p}$, τότε $\mathcal{S}' = \mathcal{S}$.

Αν $t > \binom{k}{p}$, για κάθε $e \in E$ γνωρίζουμε ότι υπάρχει μια στήλη $x_e \in \mathbb{F}^k$ στον πίνακα $A_{\mathcal{M}}$ που αντιστοιχεί στο στοιχείο αυτό. Επίσης για κάθε $S_i \in \mathcal{S}$, θεωρούμε το διάνυσμα

$$\vec{s_i} = (s_i[I])_{I \in \binom{[k]}{p}}$$

όπου $s_i[I] = \det(A_{\mathcal{M}}[I, S_i])$. Προφανώς, $\vec{s_i} \in \mathbb{F}^{\binom{k}{p}}$.

Ορίζουμε τον $\binom{k}{p} \times t$ πίνακα $H_{\mathcal{S}} = (\vec{s_1}, \ldots, \vec{s_t})$, με στήλες τα διανύσματα $\vec{s_i}$, και θεωρούμε $\mathcal{W}$ ένα σύνολο με $rank(H_{\mathcal{S}})$ το πλήθος γραμμικά ανεξάρτητες στήλες του πίνακα $H_{\mathcal{S}}$. Με άλλα λόγια, το $\mathcal{W}$ αποτελεί βάση της κυρτής θήκης των διανυσμάτων $\vec{s_1}, \ldots, \vec{s_t}$. Αφού το πλήθος των γραμμικά ανεξάρτητων γραμμών ενός πίνακα ισούται με το πλήθος των γραμμικά ανεξάρτητων στηλών, έπεται ότι $|\mathcal{W}| = rank(H_{\mathcal{S}}) \leq \binom{k}{p}$.

Θέτουμε $\mathcal{S}' = \{S_\alpha \mid \vec{s_\alpha} \in \mathcal{W}\}$ και έστω $|\mathcal{S}'| = \ell$. Επειδή $|\mathcal{W}| = |\mathcal{S}'|$, έχουμε ότι $\ell \leq \binom{k}{p}$. Χωρίς βλάβη της γενικότητας, μπορούμε να θεωρήσουμε ότι $\mathcal{S}' = \{S_i \mid 1 \leq i \leq \ell\}$ και $\mathcal{W} = \{\vec{s_1}, \ldots, \vec{s_\ell}\}$.

Μένει να δείξουμε ότι $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$. Έστω $S \in \mathcal{S}$ και $S \notin \mathcal{S}'$. Θα δείξουμε ότι αν υπάρχει $Y \subseteq E$ με $|Y| \leq q$ τέτοιο ώστε $S \cap Y = \emptyset$ και $S \cup Y \in \mathcal{I}$, τότε υπάρχει $S' \in \mathcal{S}'$ με $S' \cap Y = \emptyset$ και $S' \cup Y \in \mathcal{I}$. Διακρίνουμε 2 περιπτώσεις.

1η Περίπτωση: $|Y| = q$. Αφού $S \cap Y = \emptyset$, έπεται ότι $|S \cup Y| = p+q = k$ και αφού $S \cup Y \in \mathcal{I}$, οι αντίστοιχες στήλες του πίνακα $A_{\mathcal{M}}$ είναι γραμμικά ανεξάρτητες, και επομένως $\det(A_{\mathcal{M}}[R(A_{\mathcal{M}}), S \cup Y]) \neq 0$.

Υπενθυμίζουμε ότι $\vec{s} = (s[I])_{I \in \binom{[k]}{p}}$, όπου $s[I] = \det(A_{\mathcal{M}}[I, S])$ και αντίστοιχα ορίζουμε

$$\vec{y} = (y[L])_{L \in \binom{[k]}{q}}$$

όπου $y[L] = \det(A_{\mathcal{M}}[L, Y])$. Επίσης, ορίζουμε

$$\gamma(\vec{s}, \vec{y}) = \sum_{I \in \binom{[k]}{p}} (-1)^{\sum I + \sum J} s[I] \cdot y[\bar{I}].$$

Από Πρόταση 3.1 και επειδή $\binom{k}{p} = \binom{k}{k-p} = \binom{k}{q}$,

$$\gamma(\vec{s}, \vec{y}) = \det(A_{\mathcal{M}}[R(A_{\mathcal{M}}), S \cup Y]) \neq 0.$$

Γνωρίζουμε ότι το διάνυσμα $\vec{s}$ μπορεί να γραφτεί ως γραμμικός συνδυασμός των διανυσμάτων του $\mathcal{W}$, δηλαδή $\vec{s} = \sum_{i=1}^{\ell} \lambda_i \vec{s_i}$, με $\lambda_i \in \mathbb{F}$ και όχι όλα ίσα με 0. Επομένως,

$$\begin{aligned}
\gamma(\vec{s}, \vec{y}) &= \sum_{I} (-1)^{\sum I + \sum J} s[I] \cdot y[\bar{I}] \\
&= \sum_{I} (-1)^{\sum I + \sum J} \Big(\sum_{i=1}^{\ell} \lambda_i s_i[I] \Big) y[\bar{I}] \\
&= \sum_{i=1}^{\ell} \lambda_i \Big(\sum_{I} (-1)^{\sum I + \sum J} s_i[I] y[\bar{I}] \Big) \\
&= \sum_{i=1}^{\ell} \lambda_i \det(A_{\mathcal{M}}[R(A_{\mathcal{M}}), S_i \cup Y])
\end{aligned}$$

Τέλος, ορίζουμε

$$\sup(S) = \{S_i \mid S_i \in \mathcal{S}' \text{ και } \lambda_i \det(A_{\mathcal{M}}[R(A_{\mathcal{M}}), S_i \cup Y]) \neq 0\}.$$

Αφού $\gamma(\vec{s}, \vec{y}) \neq 0$, έπεται ότι $\sum_{i=1}^{\ell} \lambda_i \det(A_{\mathcal{M}}[R(A_{\mathcal{M}}), S \cup Y]) \neq 0$, δηλαδή $\sup(S) \neq \emptyset$. Οπότε για κάθε $S' \in \sup(S) \subseteq \mathcal{S}'$,

$$\det(A_{\mathcal{M}}[R(A_{\mathcal{M}}), S' \cup Y]) \neq 0.$$

Άρα $S' \cap Y \neq \emptyset$ και $S' \cup Y \in \mathcal{I}$, δηλαδή $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$.

2η Περίπτωση: $|Y| = q' < q$. Αφού το μητροειδές $\mathcal{M}$ έχει τάξη ίση με $k = p + q$, υπάρχει ανεξάρτητο υπερσύνολο Y' του Y με ακριβώς q στοιχεία τέτοιο ώστε $S \cap Y' = \emptyset$ και $S \cup Y' \in \mathcal{I}$. Έπεται ότι υπάρχει $S' \in \mathcal{S}'$ με $S' \cap Y' = \emptyset$ και $S' \cup Y' \in \mathcal{I}$. Αφαιρώντας τα στοιχεία του συνόλου $Y' \setminus Y$, προκύπτει ότι $S' \cap Y' = \emptyset$ και $S' \cup Y \in \mathcal{I}$. Άρα $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$.

Για να κατασκευάσουμε την οικογένεια $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$, κατασκευάζουμε τον πίνακα $H_{\mathcal{S}}$. Αρχικά, για κάθε σύνολο S_i υπολογίζουμε το διανύσμα $\vec{s_i}$, δηλαδή υπολογίζουμε τις ορίζουσες $\det(A_{\mathcal{M}}[I, S_i])$, για κάθε $I \in \binom{[k]}{p}$. Για την ορίζουσα ενός $p \times p$ πίνακα χρειαζόμαστε $\mathcal{O}(p^\omega)$ χρόνο, σύμφωνα με το [7]. Άρα συνολικά για τον πίνακα $H_{\mathcal{S}}$, χρειαζόμαστε $\mathcal{O}\Big(\binom{p+q}{p} t p^\omega\Big)$ χρόνο. Στη συνέχεια, θέλουμε να βρούμε βάση $\mathcal{W}$ της γραμμικής θήκης των διανυσμάτων-στηλών του $\binom{k}{p} \times t$ πίνακα $H_{\mathcal{S}}$. Σύμφωνα με το [16], ο υπολογισμός αυτός πραγματοποιείται σε $\mathcal{O}\Big(t\binom{p+q}{p}^{\omega-1}\Big)$ χρόνο. Τέλος, από το $\mathcal{W}$, μπορούμε εύκολα να υπολογίσουμε την οικογένεια $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$. Άρα, συνολικά χρειαζόμαστε $\mathcal{O}\Big(\binom{p+q}{p} t p^\omega + t\binom{p+q}{q}^{\omega-1}\Big)$ χρόνο. □

Στο Θεώρημα 3.1, υποθέσαμε ότι $rank(\mathcal{M}) = p + q$. Τι συμβαίνει, όμως, για $rank(\mathcal{M}) > p + q$; Στην περίπτωση αυτή, υπολογίζουμε πρώτα τον πίνακα αναπαράστασης ενός $(p + q)$-περιορισμού του μητροειδούς $\mathcal{M}$ χρησιμοποιώντας την Πρόταση 2.5. Έπειτα, εφαρμόζουμε το Θεώρημα 3.1 στο καινούριο μητροειδές και τελικά, προκύπτει το ακόλουθο αποτέλεσμα.

Θεώρημα 3.2. *Έστω $\mathcal{M} = (E, \mathcal{I})$ γραμμικό μητροειδές με $rank(\mathcal{M}) > p + q$, $A_{\mathcal{M}}$ πίνακας αναπαράστασης του επί του σώματος $\mathbb{F}$ και $\mathcal{S} = \{S_1, \ldots, S_t\}$ p-οικογένεια ανεξάρτητων υποσυνόλων του E. Τότε υπάρχει πιθανοτικός αλγόριθμος που υπολογίζει μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ μεγέθους το πολύ $\binom{p+q}{p}$, σε $\mathcal{O}\Big(\binom{p+q}{p}tp^{\omega} + t\binom{p+q}{q}^{\omega-1}\Big)$ χρόνο, όπου ως μονάδα χρόνου θεωρούμε μια πράξη στο $\mathbb{F}$.*

3.2 Γρήγορος Υπολογισμός Συνόλων Αντιπροσώπευσης για Ομοιόμορφα Μητροειδή

Σε αυτή την παράγραφο, θα δείξουμε ότι όταν το μητροειδές είναι ομοιόμορφο, μπορούμε να κατασκευάσουμε q-αντιπροσωπευτική οικογένεια, αποφεύγοντας τον υπολογισμό οριζουσών καθώς και την εύρεση βάσης. Σαν αποτέλεσμα, ο αλγόριθμος που προκύπτει είναι πιο γρήγορος από αυτόν του Θεωρήματος 3.1. Υπενθυμίζουμε ότι, αφού το μητροειδές είναι ομοιόμορφο, χρησιμοποιούμε τον Ορισμό 3.1.

Ορισμός 3.3. *Λέμε ότι μια οικογένεια $\mathcal{F}$ διαχωρίζει το σύνολο A από το σύνολο B όταν υπάρχει σύνολο $F \in \mathcal{F}$ τέτοιο ώστε $A \subseteq F$ και $B \cap F = \emptyset$.*

Ορισμός 3.4. *Έστω σύνολο U με n στοιχεία. Μια n-p-q-συλλογή διαχωρισμού (n-p-q-separating collection) $\mathcal{C}$ είναι ένα διατεταγμένο ζεύγος $(\mathcal{F}, \chi)$, όπου $\mathcal{F}$ οικογένεια υποσυνόλων του U και χ συνάρτηση από το $\binom{U}{p}$ στο $2^{\mathcal{F}}$ τέτοια ώστε:*

1. *Για κάθε $A \in \binom{U}{p}$ και $F \in \chi(A)$, $A \subseteq F$.*
2. *Για κάθε $A \in \binom{U}{p}$ και $B \in \binom{U \setminus A}{q}$, η οικογένεια $\chi(A)$ διαχωρίζει το A από το B.*

Το μέγεθος της $\mathcal{C}$ είναι ίσο με $|\mathcal{F}|$, ενώ ο βαθμός της με $\max_{A \in \binom{U}{p}} |\chi(A)|$.

Η *κατασκευή* μιας συλλογής διαχωρισμού είναι μια δομή δεδομένων, η οποία για κάθε n, p και q κατασκευάζει και εξάγει μια οικογένεια $\mathcal{F}$ υποσυνόλων ενός σύμπαντος U με n στοιχεία. Μετά την κατασκευή, μπορούμε να δώσουμε σαν είσοδο στην δομή δεδομένων ένα σύνολο A και να μας επιστρέψει την οικογένεια $\chi(A)$. Το ζεύγος $\mathcal{C} = (\mathcal{F}, \chi)$ πρέπει να είναι μια n-p-q-συλλογή διαχωρισμού.

Ο χρόνος που χρειάζεται η δομή δεδομένων μέχρι να κατασκευάσει και να εξάγει την οικογένεια $\mathcal{F}$ ονομάζεται *χρόνος αρχικοποίησης (initialization time)* και συμβολίζεται με $\tau(n, p, q)$. Ο *χρόνος απάντησης (query time)* συμβολίζεται με $\tau_Q(n, p, q)$ και είναι ο μέγιστος χρόνος που χρειάζεται η δομή δεδομένων για να υπολογίσει το $\chi(A)$ ανάμεσα σε όλα τα $A \in \binom{U}{p}$. Τέλος, το μέγεθος της $\mathcal{C}$ συμβολίζεται με $\zeta(n, p, q)$ και ο βαθμός με $\Delta(n, p, q)$.

Τώρα, είμαστε έτοιμοι να δώσουμε το βασικό εργαλείο για την κατασκευή του αλγορίθμου μας.

Λήμμα 3.4. *Για κάθε $x \in (0,1)$ υπάρχει κατασκευή n-p-q-συλλογής διαχωρισμού με τις εξής παραμέτρους:*

- *μέγεθος*, $\zeta(n,p,q) \leq 2^{\mathcal{O}(\frac{p+q}{\log\log(p+q)})} \cdot \frac{1}{x^p(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$
- *χρόνος αρχικοποίησης*, $\tau_I(n,p,q) \leq 2^{\mathcal{O}(\frac{p+q}{\log\log(p+q)})} \cdot \frac{1}{x^p(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot n \log n$
- *βαθμός*, $\Delta(n,p,q) \leq 2^{\mathcal{O}(\frac{p+q}{\log\log(p+q)})} \cdot \frac{1}{(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$
- *χρόνος ανταπόκρισης*, $\tau_Q(n,p,q) \leq 2^{\mathcal{O}(\frac{p+q}{\log\log(p+q)})} \cdot \frac{1}{(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$.

Πριν αποδείξουμε το Λήμμα 3.4, θα δούμε πώς μπορούμε να το χρησιμοποιήσουμε για τον υπολογισμό συνόλων αντιπροσώπευσης ομοιόμορφων μητροειδών, ενώ στην υπόλοιπη παράγραφο θα ασχοληθούμε με την απόδειξη του.

Θεώρημα 3.3. *Έστω σύνολο U με n στοιχεία, $\mathcal{S}$ p-οικογένεια υποσυνόλων του U, ακέραιος q και $x \in (0,1)$. Τότε υπάρχει αλγόριθμος, που υπολογίζει σε χρόνο*

$$\mathcal{O}\left(\frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot n \log n + |\mathcal{S}| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\right),$$

μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ τέτοια ώστε $|\mathcal{S}'| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n$.

Απόδειξη. Αν $|\mathcal{S}| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n$, ο αλγόριθμος θέτει $\mathcal{S}' = \mathcal{S}$ και τερματίζει. Οπότε, υποθέτουμε ότι $|\mathcal{S}| > \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n$. Ο αλγόριθμος, χρησιμοποιώντας το Λήμμα 3.4, κατασκευάζει $\mathcal{C} = (\mathcal{F}, \chi)$, μια n-p-q-συλλογή διαχωρισμού. Αν $|\mathcal{S}| \leq |\mathcal{F}|$, θέτει $\mathcal{S}' = \mathcal{S}$ και τερματίζει. Διαφορετικά, κατασκευάζει την οικογένεια $\mathcal{S}'$ ως εξής:

> Αρχικά, θέτει $\mathcal{S}' = \emptyset$ και χαρακτηρίζει «αχρησιμοποίητο» κάθε σύνολο στην οικογένεια $\mathcal{F}$. Στη συνέχεια, για κάθε $A \in \mathcal{S}$, υπολογίζει το σύνολο $\chi(A)$ και ψάχνει για «αχρησιμοποίητο» σύνολο $F \in \chi(A)$. Το πρώτο τέτοιο σύνολο που θα βρεθεί, χαρακτηρίζεται «χρησιμοποιημένο», ενώ το σύνολο A τοποθετείται στην οικογένεια $\mathcal{S}'$. Αν δεν βρεθεί τέτοιο σύνολο, τότε ο αλγόριθμος συνεχίζει με το επόμενο σύνολο της $\mathcal{S}$, χωρίς να εισάγει το A στην $\mathcal{S}'$.

Το μέγεθος της $\mathcal{S}'$ είναι μικρότερο ή ίσο από

$$|\mathcal{F}| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n,$$

αφού κάθε φορά που ο αλγόριθμος προσθέτει ένα σύνολο A στην $\mathcal{S}'$ ένα «αχρησιμοποίητο» σύνολο της $\mathcal{F}$ γίνεται «χρησιμοποιημένο». Όσον αφορά τον χρόνο υπολογισμού της οικογένειας $\mathcal{S}'$, η αρχικοποίηση της $\mathcal{C} = (\mathcal{F}, \chi)$ χρειάζεται $\frac{1}{x^p(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot 2^{o(p+q)} \cdot n \log n$ χρόνο. Για κάθε σύνολο $A \in \mathcal{S}$, ο αλγόριθμος υπολογίζει το $\chi(A)$ σε $\frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$ χρόνο. Στη συνέχεια, για κάθε σύνολο του $\chi(A)$ ελέγχει αν έχει χαρακτηρισθεί ως «χρησιμοποιημένο» σε $\frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$ χρόνο. Οπότε, συνολικά ο χρόνος που χρειάζεται ο αλγόριθμος είναι ίσος με

$$\mathcal{O}\left(\frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot n \log n + |\mathcal{S}| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\right).$$

Τέλος, πρέπει να δείξουμε ότι $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$. Έστω $A \in \mathcal{S}$ και $B \subseteq U$ τέτοιο ώστε $|B| = q$ και $A \cap B = \emptyset$. Αν $A \in \mathcal{S}'$, δεν έχουμε τίποτα να δείξουμε. Υποθέτουμε, λοιπόν, ότι $A \notin \mathcal{S}'$ και θέλουμε να δείξουμε

ότι υπάρχει $A' \in \mathcal{S}'$ τέτοιο ώστε $A' \cap B = \emptyset$. Το ζεύγος $\mathcal{C} = (\mathcal{F}, \chi)$ είναι n-p-q-συλλογή διαχωρισμού, άρα η οικογένεια $\chi(A)$ διαχωρίζει το A από το B. Επειδή $A \notin \mathcal{S}'$, το F είχε ήδη χαρακτηριστεί ως «χρησιμοποιημένο» σε προηγούμενο στάδιο του αλγορίθμου. Επομένως, όταν ο αλγόριθμος χαρακτήρισε το F «χρησιμοποιημένο», πρόσθεσε ένα σύνολο $A' \neq A$ στην οικογένεια $\mathcal{S}'$ τέτοιο ώστε $F \in \chi(A')$, δηλαδή $A' \subseteq F$. Αφού $F \cap B = \emptyset$ και $A' \subseteq F$, έπεται ότι $A' \cap B = \emptyset$, και αφού $A' \in \mathcal{S}'$, προκύπτει το ζητούμενο. □

Για την απόδειξη του Λήμματος 3.4 χρειαζόμαστε τρία βοηθητικά λήμματα. Το πρώτο από αυτά, αποδεικνύει ότι υπάρχει μια n-p-q-συλλογή διαχωρισμού. Για την κατασκευή της, μαντεύουμε μια οικογένεια $\mathcal{F}$ με το επιθυμητό μέγεθος, ορίζουμε μια συνάρτηση χ και ελέγχουμε αν το ζεύγος $(\mathcal{F}, \chi)$ είναι όντως n-p-q-συλλογή διαχωρισμού.

Λήμμα 3.5. *Για $x \in (0,1)$, υπάρχει κατασκευή n-p-q-συλλογής διαχωρισμού με τις εξής παραμέτρους:*

- *μέγεθος*, $\zeta(n,p,q) = \mathcal{O}\Big(\frac{1}{x^p(1-x)^q} \cdot (p+q+1) \cdot \log n\Big)$
- *χρόνος αρχικοποίησης*, $\tau_I(n,p,q) = \mathcal{O}\Big(\binom{2^n}{\zeta(n,p,q)} \cdot \frac{1}{x^p(1-x)^q} \cdot n^{\mathcal{O}(p+q)}\Big)$
- *βαθμός*, $\Delta(n,p,q) = \mathcal{O}\Big(\frac{1}{(1-x)^q} \cdot (p+q+1) \cdot \log n\Big)$
- *χρόνος ανταπόκρισης*, $\tau_Q(n,p,q) = \mathcal{O}\Big(\frac{1}{x^p(1-x)^q} \cdot n^{\mathcal{O}(1)}\Big)$.

Απόδειξη. Αρχικά θα δώσουμε πιθανοτικό αλγόριθμο για την κατασκευή μιας n-p-q-συλλογής διαχωρισμού $\mathcal{C} = (\mathcal{F}, \chi)$ με το ζητούμενο μέγεθος και βαθμό, και στη συνέχεια θα εφαρμόσουμε τη διαδικασία της αποτυχαιοποίησης (derandomization) ώστε να «λάβουμε» ντετερμινιστικό αλγόριθμο.

Θέτουμε $t = \frac{1}{x^p(1-x)^q} \cdot (p+q+1) \cdot \log n$ και κατασκευάζουμε την οικογένεια $\mathcal{F} = \{F_1, \ldots, F_t\}$ υποσυνόλων ενός συνόλου U με n στοιχεία, ως εξής:

> Κάθε σύνολο F_i είναι τυχαίο υποσύνολο του U, αφού ένα στοιχείο του U ανήκει στο F_i με πιθανότητα ίση με x. Κάθε στοιχείο ανήκει ή όχι στο F_i ανεξάρτητα από τα υπόλοιπα στοιχεία του U, και επομένως η κατασκευή των διαφορετικών συνόλων της $\mathcal{F}$ γίνεται ανεξάρτητα. Επίσης, για κάθε $A \in \binom{U}{p}$, θέτουμε $\chi(A) = \{F \in \mathcal{F} \mid A \subseteq F\}$.

Προφανώς, το μέγεθος της $\mathcal{C} = (\mathcal{F}, \chi)$ είναι το ζητούμενο.

Στη συνέχεια, θα δείξουμε ότι η $\mathcal{C}$ είναι n-p-q-συλλογή διαχωρισμού με πιθανότητα κοντά στο 1. Προφανώς, η 1η συνθήκη του Ορισμού 3.4 ικανοποιείται. Για την 2η συνθήκη, σταθεροποιούμε σύνολα $A \in \binom{U}{p}$, $B \in \binom{U \setminus A}{q}$ και ακέραιο $i \in \{1, \ldots, t\}$. Τότε

$$\Pr\big(A \subseteq F_i \text{ και } F_i \cap B = \emptyset\big) = x^p(1-x)^q.$$

Κάθε σύνολο της $\mathcal{F}$ κατασκευάζεται ανεξάρτητα από τα υπόλοιπα. Επομένως η πιθανότητα να μην υπάρχει F_i τέτοιο ώστε $A \subseteq F_i$ και $F_i \cap B = \emptyset$ είναι ίση με

$$\big(1 - x^p(1-x)^q\big)^t \leq e^{-(p+q+1)\log n} = \frac{1}{n^{p+q+1}}.$$

Υπάρχουν $\binom{n}{p}$ και $\binom{n-p}{q}$ επιλογές για τα σύνολα $A \in \binom{U}{p}$ και $B \in \binom{U \setminus A}{q}$, αντίστοιχα. Αφού η $\chi(A)$ περιλαμβάνει όλα τα σύνολα της $\mathcal{F}$ τα οποία είναι υπερσύνολα του A, η $\chi(A)$ διαχωρίζει το A από το

B όταν το ίδιο ισχύει και για την $\mathcal{F}$. Επομένως, από γνωστή ανισότητα πιθανοτήτων (union bound) έχουμε ότι η 2η συνθήκη του Ορισμού 3.4 αποτυγχάνει με πιθανότητα μικρότερη ή ίση του

$$\begin{aligned}\frac{1}{n^{p+q+1}} \cdot \binom{n}{p} \cdot \binom{n-p}{q} &= \frac{1}{n^{p+q+1}} \cdot \frac{(n-p-q+1)\dots n}{p!q!} \\ &\leq \frac{(n-p-q+1)\dots n}{n^{p+q+1}} \\ &\leq \frac{n^{p+q}}{n^{p+q+1}} = \frac{1}{n}.\end{aligned}$$

Για να βρεθεί άνω φράγμα του βαθμού της $\mathcal{C} = (\mathcal{F}, \chi)$, θεωρούμε σύνολο $A \in \binom{U}{p}$ και ακέραιο $i \in \{1, \dots, t\}$. Τότε $\Pr(A \subseteq F_i) = x^p$. Άρα μπορούμε να θεωρήσουμε ότι για την κατασκευή της $\chi(A)$ εφαρμόζουμε t πειράματα επιτυχίας-αποτυχίας, όπου η πιθανότητα επιτυχίας είναι ίση με x^p, δηλαδή ο πληθάριθμος $|\chi(A)|$ ακολουθεί διωνυμική κατανομή με παραμέτρους t και x^p. Η αναμενόμενη τιμή του $|\chi(A)|$ είναι ίση με

$$E\big[|\chi(A)|\big] = t \cdot x^p = \frac{1}{(1-x)^q} \cdot (p+q+1) \cdot \log n.$$

Από όριο Chernoff [26, Θεώρημα 4.4(3)], έχουμε

$$\Pr\Big(|\chi(A)| \geq 6E[|\chi(A)|]\Big) \leq 2^{-6E[|\chi(A)|]} \leq \frac{1}{n^{p+q+1}}$$

και από union bound

$$\Pr\Big(\exists A \text{ τ.ω. } |\chi(A)| \geq 6E\big[|\chi(A)|\big]\Big) \leq \frac{1}{n^{q+1}} \leq \frac{1}{n}.$$

Οπότε, η πιθανότητα η $\mathcal{C} = (\mathcal{F}, \chi)$ να είναι n-p-q-συλλογή διαχωρισμού με το επιθυμητό μέγεθος και βαθμό είναι τουλάχιστον ίση με $1 - \frac{2}{n} > 0$. Στην περίπτωση όπου $1 - \frac{2}{n} \leq 0$, η $\mathcal{F}$ περιλαμβάνει όλα (το πολύ 4) υποσύνολα του U.

Για να κατασκευάσουμε ντετερμινιστικά την $\mathcal{F}$, αρκεί να «δοκιμάσουμε» όλες τις πιθανές οικογένειες $\mathcal{F}$ μεγέθους t και να ελέγξουμε για καθεμιά αν είναι n-p-q-συλλογή διαχωρισμού. Έχουμε $\binom{2^n}{\zeta(n,p,q)}$ πιθανές οικογένειες και για καθεμιά χρειαζόμαστε $\mathcal{O}\big(t \cdot n^{\mathcal{O}(p+q)}\big) = \mathcal{O}\left(\frac{1}{x^p(1-x)^q} \cdot n^{\mathcal{O}(p+q)}\right)$ χρόνο. Έπεται ο ζητούμενος χρόνος αρχικοποίησης.

Τέλος, όσον αφορά τον χρόνο ανταπόκρισης, για κάθε $F \in \mathcal{F}$ ελέγχουμε αν $A \subseteq F$ και αν ισχύει τότε εισάγουμε το F στην οικογένεια $\chi(A)$. Δεδομένου ότι η $\mathcal{F}$ έχει ήδη υπολογιστεί, η όλη διαδικασία γίνεται στον ζητούμενο χρόνο. □

Το μόνο μειονέκτημα της κατασκευής του Λήμματος 3.5 είναι ο χρόνος αρχικοποίησης, τον οποίο και θα βελτιώσουμε με το ακόλουθο Λήμμα. Για τον σκοπό αυτό θα χρειαστούμε την κατασκευή k-τέλειων συναρτήσεων κατακερματισμού, η οποία δόθηκε από τους Alon, Yuster και Zwick στο [37].

Ορισμός 3.5. *Έστω U και V σύνολα με n και r στοιχεία αντίστοιχα.*
Μια οικογένεια συναρτήσεων $\{f_1, \dots, f_t\}$ με $f_i : U \to V$, $\forall i \in \{1, \dots, t\}$ ονομάζεται k-τέλεια οικογένεια συναρτήσεων κατακερματισμού (k-perfect family of hash functions) αν για κάθε $S \subseteq U$ με $|S| = k$ υπάρχει κάποιο i τέτοιο ώστε ο περιορισμός της f_i στο S να είναι 1-1.

Πρόταση 3.2. *Για κάθε σύνολο U με n στοιχεία, υπάρχει k-τέλεια οικογένεια $\{f_1, \dots, f_t\}$ συναρτήσεων κατακερματισμού με $f_i : U \to [k^2]$, $\forall i \in \{1, \dots, t\}$ και $t = \mathcal{O}(k^{\mathcal{O}(1)} \log n)$, η οποία μπορεί να κατασκευαστεί σε $\mathcal{O}(k^{\mathcal{O}(1)} n \log n)$ χρόνο.*

Λήμμα 3.6. *Αν υπάρχει κατασκευή n-p-q-συλλογής διαχωρισμού $(\widehat{\mathcal{F}}, \widehat{\chi})$ με χρόνο αρχικοποίησης $\tau_I(n,p,q)$, μέγεθος $\zeta(n,p,q)$, βαθμό $\Delta(n,p,q)$ και χρόνο ανταπόκρισης $\tau_Q(n,p,q)$, τότε υπάρχει κατασκευή n-p-q-συλλογής διαχωρισμού με τις ακόλουθες παραμέτρους:*

- *μέγεθος*, $\zeta'(n,p,q) \leq \zeta\big((p+q)^2,p,q\big) \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$
- *χρόνος αρχικοποίησης*, $\tau'_I(n,p,q) = \mathcal{O}\Big(\tau_I\big((p+q)^2,p,q\big) + \zeta\big((p+q)^2,p,q\big) \cdot (p+q)^{\mathcal{O}(1)} \cdot n \log n\Big)$
- *βαθμός*, $\Delta'(n,p,q) \leq \Delta\big((p+q)^2,p,q\big) \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$
- *χρόνος ανταπόκρισης*, $\tau'_Q(n,p,q) = \mathcal{O}\bigg(\Big(\tau_Q\big((p+q)^2,p,q\big) + \Delta\big((p+q)^2,p,q\big)\Big) \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n\bigg)$.

Απόδειξη. Έστω U σύνολο με n στοιχεία και ακέραιοι p, q. Χρησιμοποιώντας την Πρόταση 3.2, κατασκευάζουμε $(p+q)$-τέλεια οικογένεια $\{f_1,\ldots,f_t\}$ συναρτήσεων κατακερματισμού με $f_i : U \to \big[(p+q)^2\big]$, $\forall i \in \{1,\ldots,t\}$. Επίσης, κατασκευάζουμε $(p+q)^2$-p-q-συλλογή διαχωρισμού $(\widehat{\mathcal{F}}, \widehat{\chi})$, όπου $\widehat{\mathcal{F}}$ οικογένεια υποσυνόλων του συνόλου $\big[(p+q)^2\big]$. Λόγω υπόθεσης, θεωρούμε ότι η κατασκευή αυτή γίνεται «αυτόματα».

Πριν προχωρήσουμε, θα χρειαστούμε μερικούς συμβολισμούς:

Για $S \subseteq U$, $f_i(S) = \{f_i(s) \mid s \in S\}$.

Για $T \subseteq \big[(p+q)^2\big]$, $f_i^{-1}(T) = \{s \in U \mid f(s) \in T\}$.

Για οικογένεια $\mathcal{Z}$ υποσυνόλων του U, $f_i(\mathcal{Z}) = \{f_i(S) \mid S \in \mathcal{Z}\}$.

Για οικογένεια $\mathcal{W}$ υποσυνόλων του $\big[(p+q)^2\big]$, $f_i^{-1}(\mathcal{W}) = \{f_i^{-1}(T) \mid T \in \mathcal{W}\}$.

Τώρα, είμαστε έτοιμοι να κατασκευάσουμε n-p-q-συλλογή διαχωρισμού με τις επιθυμητές παραμέτρους. Θέτουμε

$$\mathcal{F} = \bigcup_{i \in \{1,\ldots,t\}} f_i^{-1}(\widehat{\mathcal{F}})$$

και για κάθε $A \in \binom{U}{p}$,

$$\chi(A) = \bigcup_{\substack{i \in \{1,\ldots,t\} \\ |f_i(A)| = |A|}} f_i^{-1}\Big(\widehat{\chi}\big(f_i(A)\big)\Big).$$

Το ζεύγος $\mathcal{C} = (\mathcal{F}, \chi)$ είναι n-p-q-συλλογή διαχωρισμού. Αφού για κάθε $\widehat{F} \in \widehat{\chi}\big(f_i(A)\big)$, $f_i(A) \subseteq \widehat{F}$, έπεται ότι $A \subseteq F$ για κάθε $F \in \chi(A)$. Οπότε, ισχύει η 1η συνθήκη του Ορισμού 3.4. Όσον αφορά τη 2η συνθήκη, θεωρούμε σύνολα $A \in \binom{U}{p}$ και $B \in \binom{U \setminus A}{q}$ και θέλουμε να δείξουμε ότι το $\chi(A)$ διαχωρίζει το A από το B. Αφού $\{f_1,\ldots,f_t\}$ $(p+q)$-τέλεια οικογένεια συναρτήσεων κατακερματισμού, υπάρχει κάποιο $i \in \{1,\ldots,t\}$ τέτοιο ώστε η συνάρτηση f_i περιορισμένη στο $A \cup B$ να είναι 1-1. Επίσης, το ζεύγος $(\widehat{\mathcal{F}}, \widehat{\chi})$ είναι $(p+q)^2$-p-q-συλλογή διαχωρισμού, και επομένως η οικογένεια $\widehat{\chi}\big(f_i(A)\big)$ διαχωρίζει το $f_i(A)$ από το $f_i(B)$. Άρα, η οικογένεια $f_i^{-1}\Big(\widehat{\chi}\big(f_i(A)\big)\Big)$ διαχωρίζει το A από το B. Η f_i περιορισμένη στο σύνολο A είναι 1-1. Άρα $f_i^{-1}\Big(\widehat{\chi}\big(f_i(A)\big)\Big) \in \chi(A)$, δηλαδή η οικογένεια $\chi(A)$ διαχωρίζει το A από το B.

Χρειαζόμαστε $\mathcal{O}\big((p+q)^{\mathcal{O}(1)} n \log n\big)$ χρόνο για την κατασκευή $(p+q)$-τέλειας οικογένειας συναρτήσεων κατακερματισμού και $\mathcal{O}\Big(\tau_I\big((p+q)^2,p,q\big)\Big)$ χρόνο για να αρχικοποιήσουμε την οικογένεια $\widehat{\mathcal{F}}$ μεγέθους $\zeta\big((p+q)^2,p,q\big)$. Άρα, συνολικά, χρειαζόμαστε $\mathcal{O}\Big(\zeta\big((p+q)^2,p,q\big) \cdot (p+q)^{\mathcal{O}(1)} \cdot n \log n\Big)$ χρόνο για την

κατασκευή της $\mathcal{F}$ από την $\widehat{\mathcal{F}}$ και την τέλεια οικογένεια συναρτήσεων κατακερματισμού. Έπεται η ζητούμενη τιμή για τον χρόνο αρχικοποίησης $\tau'_I(n,p,q)$. Επίσης, $|\mathcal{F}| \leq |\widehat{\mathcal{F}}| \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n$, δηλαδή ισχύει το ζητούμενο άνω φράγμα για το μέγεθος $\zeta'(n,p,q)$ της δομής δεδομένων.

Για τον βαθμό $\Delta'(n,p,q)$, ισχύει:

$$|\chi(A)| \leq \sum_{\substack{i \leq t \\ |f_i(A)|=|A|}} |\widehat{\chi}\big(f_i(A)\big)| \leq \Delta\big((p+q)^2,p,q\big)(p+q)^{\mathcal{O}(1)} \log n,$$

δηλαδή το ζητούμενο άνω φράγμα.

Για τον υπολογισμό της $\chi(A)$ ελέγχουμε για κάθε $i \in \{1,\ldots,t\}$, αν η f_i περιορισμένη στο A είναι 1-1 σε $\mathcal{O}\big((p+q)^{\mathcal{O}(1)} \cdot \log n\big)$ χρόνο. Για κάθε τέτοιο i, υπολογίζουμε το σύνολο $f_i(A)$ και έπειτα την οικογένεια $\widehat{\chi}\big(f_i(A)\big)$ σε $\mathcal{O}\Big(\tau_Q\big((p+q)^2,p,q\big)\Big)$ χρόνο. Στη συνέχεια, υπολογίζουμε την οικογένεια $f_i^{-1}\Big(\widehat{\chi}\big(f_i(A)\big)\Big)$ σε $\mathcal{O}\Big(|\widehat{\chi}\big(f_i(A)\big)| \cdot (p+q)^{\mathcal{O}(1)}\Big) = \mathcal{O}\Big(\Delta\big((p+q)^2,p,q\big) \cdot (p+q)^{\mathcal{O}(1)}\Big)$ χρόνο και προσθέτουμε τα σύνολα αυτά στην οικογένεια $\chi(A)$. Επαναλαμβάνουμε την διαδικασία $\mathcal{O}\big((p+q)^{\mathcal{O}(1)} \cdot \log n\big)$ φορές. Άρα για τον υπολογισμό της $\chi(A)$ χρειαζόμαστε συνολικά

$$\mathcal{O}\bigg(\Big(\tau_Q\big((p+q)^2,p,q\big) + \Delta\big((p+q)^2,p,q\big)\Big) \cdot (p+q)^{\mathcal{O}(1)} \log n\bigg)$$

χρόνο, δηλαδή τον ζητούμενο χρόνο ανταπόκρισης $\tau'_Q(n,p,q)$. □

Στο επόμενο λήμμα, ανάγουμε το πρόβλημα της εύρεσης n-p-q-συλλογής διαχωρισμού, στο ίδιο πρόβλημα αλλά με μικρότερες τιμές για τα p και q.

Ορισμός 3.6. *Διαδοχική διαμέριση (consecutive partition) του συνόλου* $[n]$ *ονομάζεται μια διαμέριση* $\mathcal{U}_P = \{U_1,\ldots,U_t\}$ *του* $[n]$ *τέτοια ώστε, για κάθε* $i \in \{1,\ldots,t\}$ *και ακεραίους* $1 \leq x \leq y \leq z$*, αν* $x \in U_i$ *και* $z \in U_i$ *τότε και* $y \in U_i$*.*

Συμβολίζουμε με $\mathscr{P}_t^n$ την οικογένεια που αποτελείται από όλες τις διαδοχικές διαμερίσεις του $[n]$ με t μέρη, κάποια από τα οποία ενδεχομένως κενά. Είναι εύκολο να δείξουμε ότι για κάθε t,

$$|\mathscr{P}_t^n| = \binom{n+t-1}{t-1}.$$

Επίσης, συμβολίζουμε με $\mathcal{Z}_{s,t}^p$ το σύνολο των t-άδων $(p_1,p_2,\ldots,p_t) \in \mathbb{Z}^t$, για τις οποίες $\sum_{i=1}^{t} p_i = p$ και $0 \leq p_i \leq s$, $\forall i \in \{1,\ldots,t\}$. Για κάθε s, t ισχύει $|\mathcal{Z}_{s,t}^p| \leq \binom{p+t-1}{t-1}$. Πράγματι, $\mathcal{Z}_{s,t}^p \subseteq \mathscr{P}_t^p$, αφού υπάρχει αμφιμονοσήμαντη απεικόνιση από το σύνολο των διαδοχικών διαμερίσεων του $[p]$ με t μέρη στο σύνολο των τρόπων με τους οποίους μπορεί να γραφτεί ο αριθμός p ως άθροισμα t μη-αρνητικών ακεραίων $p_1,p_2,\ldots,p_t$, χωρίς να λάβουμε υπόψιν το άνω φράγμα των p_i.

Λήμμα 3.7. *Για ακεραίους* p *και* q *θέτουμε* $s = \big\lfloor \log^2(p+q) \big\rfloor$ *και* $t = \big\lceil \frac{p+q}{s} \big\rceil$*. Αν υπάρχει κατασκευή* n-p-q*-συλλογής διαχωρισμού* $(\widehat{\mathcal{F}},\widehat{\chi})$ *με χρόνο αρχικοποίησης* $\tau_I(n,p,q)$*, μέγεθος* $\zeta(n,p,q)$*, βαθμό* $\Delta(n,p,q)$ *και χρόνο ανταπόκρισης* $\tau_Q(n,p,q)$*, τότε υπάρχει κατασκευή* n-p-q*-συλλογής διαχωρισμού με τις ακόλουθες παραμέτρους:*

- *μέγεθος,* $\zeta'(n,p,q) \leq |\mathscr{P}_t^n| \cdot \sum_{(p_1,\ldots,p_t) \in \mathcal{Z}_{s,t}^p} \prod_{i=1}^{t} \zeta(n,p_i,s-p_i)$

- *χρόνος αρχικοποίησης*, $\tau'_I(n,p,q) = \mathcal{O}\Big(\Big(\sum_{\substack{\hat{p}\le s,p\\ s-\hat{p}\le q}} \tau_I(n,\hat{p},s-\hat{p})\Big) + \zeta'(n,p,q)\cdot n^{\mathcal{O}(1)}\Big)$
- *βαθμός*, $\Delta'(n,p,q) \le \Delta^*(n,p,q) = |\mathscr{P}^n_t|\cdot|\mathcal{Z}^p_{s,t}|\cdot \max_{(p_1,\ldots,p_t)\in\mathcal{Z}^p_{s,t}} \prod_{i=1}^{t}\Delta(n,p_i,s-p_i)$
- *χρόνος ανταπόκρισης*, $\tau'_Q(n,p,q) = \mathcal{O}\Big(\Delta^*(n,p,q)\cdot n^{\mathcal{O}(1)} + |\mathscr{P}^n_t|\cdot|\mathcal{Z}^p_{s,t}|\cdot t\cdot\big(\max_{\substack{\hat{p}\le s,p\\ s-\hat{p}\le q}} \tau_Q(n,\hat{p},s-\hat{p})\big)\Big)$.

Απόδειξη. Θέτουμε $s = \lfloor \log^2(p+q)\rfloor$ και $t = \lceil \frac{p+q}{s}\rceil$. Χωρίς βλάβη της γενικότητας, θεωρούμε $U=[n]$. Αρχικά, ο αλγόριθμος κατασκευάζει n-$\hat{p}$-$(s-\hat{p})$-συλλογές διαχωρισμού $(\widehat{\mathcal{F}}_{\hat{p}},\chi_{\hat{p}})$, για κάθε $\hat{p}$ τέτοιο ώστε $0\le\hat{p}\le s$, $\hat{p}\le p$ και $s-\hat{p}\le q$.

Πριν συνεχίσουμε την περιγραφή του αλγορίθμου, θα χρειαστούμε τις παρακάτω πράξεις. Για σύνολο $U'\subseteq U$ και οικογένειες $\mathcal{A}$ και $\mathcal{B}$ ορίζουμε:

$$\mathcal{A}\sqcap U' = \{A\cap U' \mid A\in\mathcal{A}\}$$

$$\mathcal{A}\circ\mathcal{B} = \{A\cup B \mid A\in\mathcal{A}\wedge B\in\mathcal{B}\}.$$

Τώρα, είμαστε έτοιμοι να ορίσουμε n-p-q-συλλογή διαχωρισμού $(\mathcal{F},\chi)$.
Θέτουμε

$$\mathcal{F} = \bigcup_{\substack{(U_1,\ldots,U_t)\in\mathscr{P}^n_t\\ (p_1,\ldots,p_t)\in\mathcal{Z}^p_{s,t}\ \text{τ.ω.}\\ \forall i: s-p_i\le q}} (\widehat{\mathcal{F}}_{p_1}\sqcap U_1)\circ(\widehat{\mathcal{F}}_{p_2}\sqcap U_2)\circ\cdots\circ(\widehat{\mathcal{F}}_{p_t}\sqcap U_t) \tag{3.1}$$

και για κάθε $A\in\binom{U}{p}$,

$$\chi(A) = \bigcup_{\substack{\{U_1,\ldots,U_t\}\in\mathscr{P}^n_t\\ (p_1,\ldots,p_t)\in\mathcal{Z}^p_{s,t}\ \text{τ.ω.}\\ \forall U_i: |U_i\cap A|=p_i, s-p_i\le q}} \Big[\big(\widehat{\chi}_{p_1}(A\cap U_1)\sqcap U_1\big)\circ\big(\widehat{\chi}_{p_2}(A\cap U_2)\sqcap U_2\big)\circ \cdots\circ\big(\widehat{\chi}_{p_1}(A\cap U_1)\sqcap U_1\big)\Big]. \tag{3.2}$$

Το διατεταγμένο ζεύγος $(\mathcal{F},\chi)$ είναι n-p-q-συλλογή διαχωρισμού. Έστω $A\in\binom{U}{p}$ και $B\in\binom{U\setminus A}{q}$. Υπάρχει τουλάχιστον μια διαδοχική διαμέριση $\{U_1,\ldots,U_t\}$ του U τέτοια ώστε για κάθε $i\in\{1,\ldots,t\}$, $|(A\cup B)\cap U_i|\le\lceil\frac{p+q}{t}\rceil = s$. Για κάθε $i\in\{1,\ldots,t\}$, θέτουμε $p_i = |A\cap U_i|$ και $q_i = |B\cap U_i| = s-p_i$. Παρατηρούμε ότι $p_i\le p$, $q_i\le q$ και ότι το $(\widehat{\mathcal{F}}_{p_i},\widehat{\chi}_{p_i})$ είναι n-p_i-q_i-συλλογή διαχωρισμού, για κάθε i. Επομένως, για κάθε i, η οικογένεια $\widehat{\chi}_{p_i}(A\cap U_i)$ διαχωρίζει το $A\cap U_i$ από το $B\cap U_i$. Έστω $F_i\in\widehat{\chi}_{p_i}(A\cap U_i)$ τέτοιο ώστε $A\cap U_i\subseteq F_i$ και $F_i\cap(B\cap U_i)=\emptyset$. Θέτουμε $F=\bigcup_{i=1}^{t}(F_i\cap U_i)$ και παρατηρούμε ότι $A\subseteq F$ και $B\cap F=\emptyset$. Από την κατασκευή της $\chi(A)$ έχουμε ότι $F\in\chi(A)$ και επομένως η $\chi(A)$ διαχωρίζει το A από το B. Για την απόδειξη της 1ης συνθήκης του Ορισμού 3.4, ακολουθείται η ίδια διαδικασία.

Για τον χρόνο αρχικοποίησης, ο αλγόριθμος χρειάζεται $\mathcal{O}\Big(\sum_{\substack{\hat{p}\le s,p\\ s-\hat{p}\le q}}\tau_I(n,\hat{p},s-\hat{p})\Big)$ χρόνο για την αρχικοποίηση των n-$\hat{p}$-$(s-\hat{p})$-συλλογών διαχωρισμού για κάθε $\hat{p}\le s,p$ και $s-\hat{p}\le q$. Στη συνέχεια, χρησιμοποιεί την σχέση (3.1) και υπολογίζει τα σύνολα της οικογένειας $\mathcal{F}$, το καθένα σε $n^{\mathcal{O}(1)}$ χρόνο. Έπεται η ζητούμενη τιμή για το $\tau'_I(n,p,q)$.

Άμεση συνέπεια της σχέσης (3.1) είναι το άνω φράγμα του μεγέθους $\zeta'(n,p,q)$, ενώ της σχέσης (3.2) το άνω φράγμα του βαθμού $\Delta'(n,p,q)$.

Τέλος, για τον υπολογισμό της $\chi(A)$ χρησιμοποιούμε την σχέση 3.2. Για κάθε $\{U_1,\dots,U_t\} \in \mathscr{P}_t^n$ και $(p_1,\dots,p_t) \in \mathcal{Z}_{s,t}^p$ τέτοια ώστε $p_i = |U_i \cap A| \leq p$ και $s - p_i \leq q$ για κάθε $i \in \{1,\dots,t\}$ ακολουθούμε την εξής διαδικασία: Αρχικά, για κάθε $i \in \{1,\dots,t\}$ υπολογίζουμε το $\widehat{\chi}_{p_i}(A \cap U_i)$ σε $\mathcal{O}\big(\tau_Q(n,p_i,s-p_i)\big)$ χρόνο. Χρησιμοποιώντας την σχέση (3.2), υπολογίζουμε τα σύνολα της $\chi(A)$, το καθένα σε $n^{\mathcal{O}(1)}$ χρόνο. Συνολικά, έχουμε:

$$\begin{aligned}
\tau'_Q(n,p,q) &\leq \mathcal{O}\Bigg(\Delta^*(n,p,q)\cdot n^{\mathcal{O}(1)} + \sum_{\substack{\{U_1,\dots,U_t\}\in\mathscr{P}_t^n \\ (p_1,\dots,p_t)\in\mathcal{Z}_{s,t}^p \text{ τ.ω.} \\ \forall U_i : |U_i\cap A|=p_i, s-p_i\leq q}} \Big[\sum_{i=1}^{t}\tau_Q(n,p_i,s-p_i)\Big]\Bigg) \\
&\leq \mathcal{O}\Bigg(\Delta^*(n,p,q)\cdot n^{\mathcal{O}(1)} + |\mathscr{P}_t^n|\cdot|\mathcal{Z}_{s,t}^p|\cdot \max_{\substack{(p_1,\dots,p_t)\in\mathcal{Z}_{s,t}^p \text{ τ.ω.} \\ \forall i\,:\, s-p_i\leq q}} \Big[\sum_{i=1}^{t}\tau_Q(n,p_i,s-p_i)\Big]\Bigg) \\
&\leq \mathcal{O}\Bigg(\Delta^*(n,p,q)\cdot n^{\mathcal{O}(1)} + |\mathscr{P}_t^n|\cdot|\mathcal{Z}_{s,t}^p|\cdot t\cdot \max_{\substack{(p_1,\dots,p_t)\in\mathcal{Z}_{s,t}^p \text{ τ.ω.} \\ \forall i\,:\, s-p_i\leq q}} \big(\tau_Q(n,p_i,s-p_i)\big)\Bigg) \\
&\leq \mathcal{O}\Big(\Delta^*(n,p,q)\cdot n^{\mathcal{O}(1)} + |\mathscr{P}_t^n|\cdot|\mathcal{Z}_{s,t}^p|\cdot t\cdot \max_{\substack{\hat{p}\leq s,p \\ s-\hat{p}\leq q}} \tau_Q(n,\hat{p},s-\hat{p})\Big),
\end{aligned}$$

δηλαδή τον ζητούμενο χρόνο ανταπόκρισης. □

Τώρα, έχουμε στην διαθεσή μας όλα τα απαραίτητα εργαλεία για την απόδειξη του Λήμματος 3.4.

Απόδειξη του Λήμματος 3.4. Έστω U σύνολο με n στοιχεία. Διακρίνουμε τρεις περιπτώσεις, ανάλογα με την τιμή του x.

1η Περίπτωση: $x \leq \frac{1}{n}$. Σε αυτή την περίπτωση, θέτουμε

$$\mathcal{F} = \{F \subseteq U \mid |F| = p\}$$

και για κάθε $A \in \binom{U}{p}$,

$$\chi(A) = \{F \in \mathcal{F} \mid A \subseteq F\}.$$

Προφανώς, οι δύο ιδιότητες του Ορισμού 3.4 ικανοποιούνται, η 1η κατευθείαν από τον ορισμό της $\chi(A)$ και η 2η γιατί $\chi(A) = \{A\}$. Άρα το διατεταγμένο ζεύγος $(\mathcal{F},\chi)$ είναι n-p-q-συλλογή διαχωρισμού.

Παρατηρούμε ότι $|\mathcal{F}| = \binom{n}{p} \leq n^p$. Αφού $n \leq \frac{1}{x}$, το μέγεθος της n-p-q-συλλογής διαχωρισμού φράσσεται από το ζητούμενο. Αφού μπορούμε να παραθέσουμε όλα τα στοιχεία της $\mathcal{F}$ σε n^p χρόνο, έπεται το ζητούμενο άνω φράγμα του χρόνου αρχικοποίησης. Για κάθε $A \in \binom{U}{p}$, $|\chi(A)| = 1$, οπότε έπονται τα ζητούμενα φράγματα για τον βαθμό και τον χρόνο ανταπόκρισης.

2η Περίπτωση: $1 - x \leq \frac{1}{n}$. Σε αυτή την περίπτωση, θέτουμε

$$\mathcal{F} = \{F \subseteq U \mid |F| = n - q\}$$

και για κάθε $A \in \binom{U}{p}$,

$$\chi(A) = \{F \in \mathcal{F} \mid A \subseteq F\}.$$

Προφανώς, οι δύο ιδιότητες του Ορισμού 3.4 ικανοποιούνται, η 1η κατευθείαν από τον ορισμό της $\chi(A)$ και η 2η γιατί $\chi(A) = \{U \setminus B\}$. Άρα το διατεταγμένο ζεύγος $(\mathcal{F}, \chi)$ είναι n-p-q-συλλογή διαχωρισμού.

Παρατηρούμε ότι $|\mathcal{F}| = \binom{n}{n-q} \leq n^q$. Αφού $n \leq \frac{1}{1-x}$, το μέγεθος της n-p-q-συλλογής διαχωρισμού φράσσεται από το ζητούμενο. Αφού μπορούμε να παραθέσουμε όλα τα στοιχεία της $\mathcal{F}$ σε n^q χρόνο, έπεται το ζητούμενο άνω φράγμα του χρόνου αρχικοποίησης. Για κάθε $A \in \binom{U}{p}$,

$$|\chi(A)| \leq |\mathcal{F}| \leq \frac{1}{(1-x)^q},$$

οπότε έπονται τα ζητούμενα φράγματα για τον βαθμό και τον χρόνο ανταπόκρισης.

3η Περίπτωση: $x, 1-x > \frac{1}{n}$. Αρχικά, κατασκευάζουμε n-p-q-συλλογή διαχωρισμού, χρησιμοποιώντας το Λήμμα 3.5, με παραμέτρους:

- μέγεθος, $\zeta^1(n,p,q) = \mathcal{O}\Big(\frac{1}{x^p(1-x)^q} \cdot (p+q+1)\log n\Big)$
- χρόνος αρχικοποίησης, $\tau_I^1(n,p,q) = \mathcal{O}\Big(\binom{2^n}{\zeta^1(n,p,q)} \cdot \frac{1}{x^p(1-x)^q} \cdot n^{\mathcal{O}(p+q)}\Big)$
- βαθμός, $\Delta^1(n,p,q) = \mathcal{O}\Big(\frac{p+q+1}{(1-x)^q}\log n\Big)$
- χρόνος ανταπόκρισης, $\tau_Q^1(n,p,q) = \mathcal{O}\Big(\frac{1}{x^p(1-x)^q} n^{\mathcal{O}(1)}\Big) = \mathcal{O}(2^n n^{\mathcal{O}(1)})$.

Εφαρμόζουμε το Λήμμα 3.6 στην παραπάνω κατασκευή και προκύπτει νέα συλλογή διαχωρισμού με τις ακόλουθες παραμέτρους:

- μέγεθος, $\zeta^2(n,p,q) = \mathcal{O}\Big(\frac{1}{x^p(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n\Big)$
- χρόνος αρχικοποίησης,

$$\begin{aligned}
\tau_I^2(n,p,q) &= \mathcal{O}\Big(\tau_I^1\big((p+q)^2,p,q\big) + \zeta^1\big((p+q)^2,p,q\big) \cdot (p+q)^{\mathcal{O}(1)} \cdot n\log n\Big) \\
&= \mathcal{O}\left(\frac{2^{2^{(p+q)^2}}}{x^p(1-x)^q} \cdot (p+q)^{\mathcal{O}(p+q)} + \left(\frac{(p+q)^{\mathcal{O}(1)}}{x^p(1-x)^q} \cdot n\log n\right)\right) \\
&= \mathcal{O}\left(\frac{(p+q)^{\mathcal{O}(p+q)}}{x^p(1-x)^q}\left(2^{2^{(p+q)^2}} + n\log n\right)\right)
\end{aligned}$$

- βαθμός, $\Delta^2(n,p,q) = \mathcal{O}\Big(\frac{1}{(1-x)^q} \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n\Big)$ και
- χρόνος ανταπόκρισης, $\tau_Q^2(n,p,q) = \mathcal{O}\Big(\big(2^{(p+q)^2} + \frac{1}{(1-x)^q}\big)(p+q)^{\mathcal{O}(1)} \cdot \log n\Big)$.

Εφαρμόζουμε το Λήμμα 3.7 στην παραπάνω κατασκευή. Υπενθυμίζουμε ότι $s = \lfloor \log^2(p+q) \rfloor$ και $t = \lceil \frac{p+q}{s} \rceil$. Με αυτό τον τρόπο, προκύπτει μια νέα συλλογή διαχωρισμού με:

- μέγεθος,

$$\begin{aligned}\zeta^3(n,p,q) &\le |\mathcal{P}_t^n| \cdot \sum_{(p_1,\ldots,p_t)\in\mathcal{Z}_{s,t}^p} \prod_{i=1}^{t} \zeta^2(n,p_i,s-p_i) \\ &\le n^{\mathcal{O}(t)} \cdot |\mathcal{Z}_{s,t}^p| \cdot \max_{(p_1,\ldots,p_t)\in\mathcal{Z}_{s,t}^p} \prod_{i=1}^{t} \zeta^2(n,p_i,s-p_i) \\ &\le n^{\mathcal{O}(t)} \cdot (p+q)^{\mathcal{O}(t)} \cdot \frac{1}{x^p(1-x)^{q+s}} \cdot s^{\mathcal{O}(t)} \cdot (\log n)^{\mathcal{O}(t)} \\ &\le n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)} \cdot \frac{1}{x^p(1-x)^q}\end{aligned}$$

επειδή $\left(\frac{1}{1-x}\right)^s \le n^s \le n^{\mathcal{O}(t)}$.

- χρόνος αρχικοποίησης,

$$\begin{aligned}\tau_I^3(n,p,q) &= \mathcal{O}\left(\left(\sum_{\substack{\widehat{p}\le s,p\\ s-\widehat{p}\le q}} \tau_I^2(n,\widehat{p},s-\widehat{p})\right) + \zeta^3(n,p,q)\cdot n^{\mathcal{O}(1)}\right) \\ &= \mathcal{O}\left(\left(\sum_{\substack{\widehat{p}\le s,p\\ s-\widehat{p}\le q}} \frac{s^{\mathcal{O}(s)}}{x^{\widehat{p}}(1-x)^{s-\widehat{p}}}\left(2^{2^{s^2}} + n\log n\right)\right) + \zeta^3(n,p,q)\cdot n^{\mathcal{O}(1)}\right) \\ &= \mathcal{O}\left(\frac{(\log(p+q))^{\mathcal{O}(\log^2(p+q))}}{x^p(1-x)^q}\left(2^{\log^4(p+q)} + n\log n\right) + \right. \\ &\qquad\qquad \left. + n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)} \cdot \frac{1}{x^p(1-x)^q}\right)\end{aligned}$$

- βαθμός,

$$\begin{aligned}\Delta^3(n,p,q) &= \Delta^{*3}(n,p,q) \\ &\le |\mathcal{P}_t^n| \cdot |\mathcal{Z}_{s,t}^p| \cdot \max_{(p_1,\ldots,p_t)\in\mathcal{Z}_{s,t}^p} \prod_{i=1}^{t} \Delta^2(n,p_i,s-p_i) \\ &\le n^{\mathcal{O}(t)} \cdot (p+q)^{\mathcal{O}(t)} \cdot \frac{1}{(1-x)^{q+s}} \cdot s^{\mathcal{O}(t)} \cdot (\log n)^{\mathcal{O}(t)} \\ &\le n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)} \cdot \frac{1}{(1-x)^q}\end{aligned}$$

επειδή $\left(\frac{1}{1-x}\right)^s \le n^{\mathcal{O}(t)}$.

- χρόνος ανταπόκρισης,

$$\begin{aligned}\tau_Q^3(n,p,q) &\leq \mathcal{O}\Big(\Delta^{*3}(n,p,q)\cdot n^{\mathcal{O}(1)} + |\mathcal{P}_t^n|\cdot|\mathcal{Z}_{s,t}^p|\cdot t\cdot \max_{\substack{\widehat{p}\leq p,s\\ s-\widehat{p}\leq q}} \tau_Q^2(n,\widehat{p},s-\widehat{p})\Big)\\ &\leq \mathcal{O}\left(\Delta^{*3}(n,p,q)\cdot n^{\mathcal{O}(1)} + n^{\mathcal{O}(t)}\cdot \max_{\substack{\widehat{p}\leq p,s\\ s-\widehat{p}\leq q}} \left(2^{s^2}+\frac{1}{(1-x)^{s-\widehat{p}}}\right) s^{\mathcal{O}(1)}\log n\right)\\ &\leq \mathcal{O}\left(\frac{n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}}{(1-x)^q} + n^{\mathcal{O}(t)}\cdot s^{\mathcal{O}(1)}\cdot\log n\left(2^{s^2}+\frac{1}{(1-x)^q}\right)\right)\\ &\leq \mathcal{O}\left(\frac{n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}}{(1-x)^q}\right)\end{aligned}$$

Στη συνέχεια, εφαρμόζοντας το Λήμμα 3.6 στην 3η κατασκευή, προκύπτει η 4η κατασκευή με τις ακόλουθες παραμέτρους:

- βαθμός, $\zeta^4(n,p,q)\leq 2^{\mathcal{O}\left(\frac{p+q}{\log(p+q)}\right)}\cdot\frac{1}{x^p(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n$,
- χρόνος αρχικοποίησης,

$$\begin{aligned}\tau_I^4(n,p,q) &\leq \mathcal{O}\Big(\tau_I^3\big((p+q)^2,p,q\big)+\zeta^3\big((p+q)^2,p,q\big)\cdot(p+q)^{\mathcal{O}(1)}\cdot n\log n\Big)\\ &\leq 2^{\log^4(p+q)}\cdot\frac{\big(\log(p+q)\big)^{\mathcal{O}(\log^2(p+q))}}{x^p(1-x)^q}+\frac{2^{\mathcal{O}\left(\frac{p+q}{\log(p+q)}\right)}}{x^p(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}n\log n\end{aligned}$$

- βαθμός,

$$\begin{aligned}\Delta^4(n,p,q) &\leq \Delta^3\big((p+q)^2,p,q\big)\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\\ &\leq \frac{2^{\mathcal{O}\left(\frac{p+q}{\log(p+q)}\right)}}{x^p(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\end{aligned}$$

- χρόνος ανταπόκρισης,

$$\begin{aligned}\tau_Q^4(n,p,q) &\leq \mathcal{O}\Big(\Big(\tau_Q^3\big((p+q)^2,p,q\big)+\Delta^3\big((p+q)^2,p,q\big)\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\Big)\\ &\leq \frac{2^{\mathcal{O}\left(\frac{p+q}{\log(p+q)}\right)}}{(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}\log n\end{aligned}$$

Για την 5η κατασκευή, εφαρμόζουμε το Λήμμα 3.7 στην 4η κατασκευή. Όπως και στην 3η κατασκευή, $s=\lfloor\log^2(p+q)\rfloor$ και $t=\lceil\frac{p+q}{s}\rceil$. Οι παράμετροί της φράσσονται ως εξής:

- μέγεθος,

$$\begin{aligned}\zeta^5(n,p,q) &\leq |\mathcal{P}_t^n|\cdot\sum_{(p_1,\ldots,p_t)\in\mathcal{Z}_{s,t}^p}\prod_{i=1}^t\zeta^4(n,p_i,s-p_i)\\ &\leq n^{\mathcal{O}(t)}\cdot(p+q)^{\mathcal{O}(t)}\cdot s^{\mathcal{O}(t)}\cdot 2^{\mathcal{O}\left(\frac{st}{\log s}\right)}\cdot(\log n)^{\mathcal{O}(t)}\cdot\frac{1}{x^p(1-x)^{q+s}}\\ &\leq n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot 2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{x^p(1-x)^q}\end{aligned}$$

επειδή $\left(\frac{1}{1-x}\right)^s \leq n^{\mathcal{O}(t)}$.

- χρόνος αρχικοποίησης,

$$\begin{aligned}
\tau_I^5(n,p,q) &\leq \mathcal{O}\Bigg(\Big(\sum_{\substack{\widehat{p}\leq s,p\\ s-\widehat{p}\leq q}} \tau_I^4(n,\widehat{p},s-\widehat{p})\Big) + \zeta^5(n,p,q)\cdot n^{\mathcal{O}(1)}\Bigg)\\
&\leq \mathcal{O}\Bigg(s\frac{2^{2^{\log^4 s}}\cdot(\log s)^{\mathcal{O}(\log^2 s)}}{x^p(1-x)^q} + \frac{2^{\mathcal{O}\left(\frac{s}{\log s}\right)}}{x^p(1-x)^q}\cdot n\log n +\\
&\qquad + n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot\frac{2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}}{x^p(1-x)^q}\Bigg)\\
&\leq \mathcal{O}\Bigg(s\frac{2^{2^{\log^4 s}}\cdot(\log s)^{\mathcal{O}(\log^2 s)}}{x^p(1-x)^q} + n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot\frac{2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}}{x^p(1-x)^q}\Bigg)\\
&\leq \mathcal{O}\Bigg(\frac{2^{2^{\log^4 s}}\cdot s^{\mathcal{O}(s)}}{x^p(1-x)^q} + n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot\frac{2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}}{x^p(1-x)^q}\Bigg)\\
&\leq \mathcal{O}\Bigg(\frac{2^{2^{(2\log\log(p+q))^4}}\cdot\big(\log(p+q)\big)^{\mathcal{O}(\log^2(p+q))}}{x^p(1-x)^q} +\\
&\qquad + n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot\frac{2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}}{x^p(1-x)^q}\Bigg)\\
&\leq \mathcal{O}\Bigg(n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot\frac{2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}}{x^p(1-x)^q}\Bigg)
\end{aligned}$$

επειδή $2^{2^{(2\log\log(p+q))^4}}, \big(\log(p+q)\big)^{\mathcal{O}(\log^2(p+q))} \leq 2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}$.

- βαθμός,

$$\begin{aligned}
\Delta^5(n,p,q) &\leq \Delta^{*5}(n,p,q)\\
&= |\mathcal{P}_t^n|\cdot|\mathcal{Z}_{s,t}^p|\cdot\max_{(p_1,\ldots,p_t)\in\mathcal{Z}_{s,t}^p}\prod_{i=1}^{t}\Delta^4(n,p_i,s-p_i)\\
&\leq n^{\mathcal{O}(t)}\cdot(p+q)^{\mathcal{O}(t)}\cdot\frac{2^{\mathcal{O}\left(\frac{st}{\log s}\right)}}{(1-x)^{q+s}}\cdot s^{\mathcal{O}(t)}\cdot(\log n)^{\mathcal{O}(t)}\\
&\leq n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot 2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{(1-x)^q}
\end{aligned}$$

επειδή $\left(\frac{1}{1-x}\right)^s \leq n^{\mathcal{O}(t)}$.

- χρόνος ανταπόκρισης,

$$\begin{aligned}
\tau_Q^5(n,p,q) &\leq \mathcal{O}\Big(\Delta^{*5}(n,p,q)\cdot n^{\mathcal{O}(1)} + |\mathcal{P}_t^n|\cdot|\mathcal{Z}_{s,t}^p|\cdot\max_{\substack{\widehat{p}\leq s\\ s-\widehat{p}\leq q}}\tau_Q^4(n,\widehat{p},s-\widehat{p})\Big)\\
&\leq n^{\mathcal{O}\left(\frac{p+q}{\log^2(p+q)}\right)}\cdot 2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{(1-x)^q}
\end{aligned}$$

Τέλος, εφαρμόζουμε το Λήμμα 3.6 στην 5η κατασκευή και παίρνουμε την 6η κατασκευή.

- μέγεθος,

$$\begin{aligned}\zeta(n,p,q) &\leq \zeta^5\big((p+q)^2,p,q\big)\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\\ &\leq 2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{x^p(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}\log n\end{aligned}$$

- χρόνος αρχικοποίησης,

$$\begin{aligned}\tau_I(n,p,q) &\leq \mathcal{O}\Big(\tau_I^5\big((p+q)^2,p,q\big)+\zeta^5\big((p+q)^2,p,q\big)\cdot(p+q)^{\mathcal{O}(1)}\cdot n\log n\Big)\\ &= \mathcal{O}\Big(2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{x^p(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}n\log n\Big)\end{aligned}$$

- βαθμός,

$$\begin{aligned}\Delta(n,p,q) &\leq \Delta^5\big((p+q)^2,p,q\big)\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\\ &\leq \mathcal{O}\Big(2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\Big)\end{aligned}$$

- χρόνος ανταπόκρισης,

$$\begin{aligned}\tau_Q(n,p,q) &\leq \mathcal{O}\Big(\Big(\tau_Q^5\big((p+q)^2,p,q\big)+\Delta^5\big((p+q)^2,p,q\big)\Big)\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\Big)\\ &\leq \mathcal{O}\Big(2^{\mathcal{O}\left(\frac{p+q}{\log\log(p+q)}\right)}\cdot\frac{1}{(1-x)^q}\cdot(p+q)^{\mathcal{O}(1)}\cdot\log n\Big)\end{aligned}$$

Η 6η κατασκευή έχει τις ζητούμενες παραμέτρους.

□

3.2.1 Μπορούμε καλύτερα;

Στη παράγραφο αυτή, θα αποδείξουμε μια ακόμα πιο γρήγορη εκδοχή του Θεωρήματος 3.3.

Λήμμα 3.8. *Έστω U σύνολο με n στοιχεία, $\mathcal{S}$ p-οικογένεια υποσυνόλων του U, q ακέραιος και $x\in(0,1)$. Τότε υπάρχει αλγόριθμος, που υπολογίζει σε χρόνο*

$$\mathcal{O}\Big((p+q)^{\mathcal{O}(1)}n\log n+|\mathcal{S}|\cdot\frac{1}{(1-x)^q}\cdot 2^{o(p+q)}\cdot\log n\Big)$$

μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$, τέτοια ώστε $|\mathcal{S}'|\leq\frac{1}{x^p(1-x)^q}\cdot 2^{o(p+q)}\cdot\log n$.

Απόδειξη. Αν $|\mathcal{S}|\leq\frac{1}{x^p(1-x)^q}\cdot 2^{o(p+q)}\cdot\log n$, τότε ο αλγόριθμος θέτει $\mathcal{S}'=\mathcal{S}$ και τερματίζει. Οπότε, υποθέτουμε ότι $|\mathcal{S}|>\frac{1}{x^p(1-x)^q}\cdot 2^{o(p+q)}\cdot\log n$.

Ο αλγόριθμος, χρησιμοποιώντας την Πρόταση 3.2, κατασκευάζει $(p+q)$-τέλεια οικογένεια $\{f_1,\ldots,f_t\}$ συναρτήσεων κατακερματισμού με $f_i:U\to\big[(p+q)^2\big]$, $\forall i\in\{1,\ldots,t\}$ και $t=\mathcal{O}\big((p+q)^{\mathcal{O}(1)}\log n\big)$ σε χρόνο $\mathcal{O}\big((p+q)^{\mathcal{O}(1)}n\log n\big)$. Στη συνέχεια για κάθε f_i με $i\in\{1,\ldots,t\}$, κατασκευάζει μια οικογένεια $\mathcal{S}_i'$ ως εξής:

Αρχικά, ο αλγόριθμος κατασκευάζει $[(p+q)^2]$-p-q-συλλογή διαχωρισμού $(\mathcal{F}_i, \chi_i)$, χρησιμοποιώντας το Λήμμα 3.4, θέτει $\mathcal{S}'_i = \emptyset$ και χαρακτηρίζει «αχρησιμοποίητο» κάθε σύνολο στην οικογένεια $\mathcal{F}_i$. Για κάθε $A \in \mathcal{S}$, ο αλγόριθμος ελέγχει αν η συνάρτηση f_i περιορισμένη στο A είναι 1-1. Αν όχι, τότε ο αλγόριθμος συνεχίζει στο επόμενο σύνολο της $\mathcal{S}$, χωρίς να εισάγει το A στην $\mathcal{S}'_i$. Σε αντίθετη περίπτωση, υπολογίζει την οικογένεια $\chi_i(A)$ και ψάχνει «αχρησιμοποίητο» σύνολο $F \in \chi_i(A)$. Το πρώτο τέτοιο σύνολο F που θα βρεθεί χαρακτηρίζεται «χρησιμοποιημένο», ενώ το σύνολο A τοποθετείται στην οικογένεια $\mathcal{S}_i$. Αν δεν βρεθεί κανένα τέτοιο σύνολο F τότε ο αλγόριθμος συνεχίζει στο επόμενο σύνολο της $\mathcal{S}$, χωρίς να εισάγει το A στην $\mathcal{S}'_i$.

Τέλος, θέτει

$$\mathcal{S}' = \bigcup_{i=1}^{t} \mathcal{S}'_i.$$

Το μέγεθος της $\mathcal{S}'_i$ είναι μικρότερο ή ίσο από $|\mathcal{F}| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log(p+q)$, αφού κάθε φορά που ο αλγόριθμος προσθέτει ένα σύνολο A στην $\mathcal{S}'$ ένα «αχρησιμοποίητο» σύνολο της $\mathcal{F}$ γίνεται «χρησιμοποιημένο». Άρα, το μέγεθος της $\mathcal{S}'$ είναι μικρότερο ή ίσο από

$$|\mathcal{F}| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log(p+q) \cdot (p+q)^{\mathcal{O}(1)} \cdot \log n \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n.$$

Όσον αφορά τον υπολογισμό του χρόνου, η διαδικασία που ακολουθείται είναι ανάλογη με αυτή του Θεωρήματος 3.3.

Τέλος, πρέπει να δείξουμε ότι $\mathcal{S}' \subseteq^q_{rep} \mathcal{S}$. Έστω $A \in \mathcal{S}$ και $B \subseteq U$ τέτοιο ώστε $|B| = q$ και $A \cap B = \emptyset$. Αν $A \in \mathcal{S}'$, δεν έχουμε τίποτα να δείξουμε. Υποθέτουμε, λοιπόν, ότι $A \notin \mathcal{S}'$ και θέλουμε να δείξουμε ότι υπάρχει $A' \in \mathcal{S}'$ τέτοιο ώστε $A' \cap B = \emptyset$. Αφού $\{f_1, \ldots, f_t\}$ $(p+q)$-τέλεια οικογένεια συναρτήσεων κατακερματισμού, υπάρχει $i \in \{1, \ldots, t\}$ τέτοιο ώστε ο περιορισμός της f_i στο $A \cup B$ να είναι 1-1. Επίσης, το $(\mathcal{F}_i, \chi_i)$ είναι $[(p+q)^2]$-p-q-συλλογή διαχωρισμού, οπότε υπάρχει $F \in \chi_i(A)$ τέτοιο ώστε $A \subseteq F$ και $F \cap B = \emptyset$. Αφού $A \notin \mathcal{S}'$ έπεται ότι $A \notin \mathcal{S}'_i$ και επομένως το σύνολο F είχε χαρακτηρισθεί «χρησιμοποιημένο» σε προηγούμενο στάδιο του αλγορίθμου. Επομένως, όταν ο αλγόριθμος χαρακτήρισε το F «χρησιμοποιημένο», πρόσθεσε ένα σύνολο $A' \neq A$ τέτοιο ώστε $F \in \chi_i(A')$, δηλαδή $A' \subseteq F$. Αφού $F \cap B = \emptyset$ και $A' \subseteq F$, έπεται $A' \cap B = \emptyset$ και επειδή $A' \in \mathcal{S}'_j \subseteq \mathcal{S}'$, προκύπτει το ζητούμενο. □

Παρατηρούμε ότι στο Λήμμα 3.8 το μέγεθος της οικογένειας $\mathcal{S}'$ εξαρτάται από το μέγεθος n του συνόλου U. Μπορούμε όμως να κατασκευάσουμε ένα καινούριο σύνολο U' με μέγεθος το πολύ $|\mathcal{S}|p+q$.

Πράγματι, αν $n \leq |\mathcal{S}|p+q$, τότε θέτουμε $U' = U$. Διαφορετικά, το σύνολο U' αποτελείται από τα στοιχεία του U, τα οποία ανήκουν σε κάποιο σύνολο της οικογένειας $\mathcal{S}$, και q νέα στοιχεία. Το U' κατασκευάζεται σε $\mathcal{O}(|\mathcal{S}|p+q)$ χρόνο και παρατηρούμε ότι $|U'| \leq |\mathcal{S}|p+q$ και $|U'| \leq n$.

Θα δείξουμε ότι μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$, στην οποία κάθε σύνολο αποτελείται από στοιχεία του U', παραμένει q-αντιπροσωπευτική και όσον αφορά το σύνολο U. Θεωρούμε $X \in \mathcal{S}$ και $Y \subseteq U$, $|Y| \leq q$ τέτοιο ώστε $X \cap Y = \emptyset$. Έστω $Y' = Y \setminus U'$ και Y'' ένα τυχαίο υποσύνολο του $U' \setminus U$ μεγέθους $|Y'|$. Θέτουμε $Z = (Y \setminus Y') \cup Y''$. Είναι εύκολο να δούμε ότι $|Z| = |Y|$ και $X \cap Z = \emptyset$. Από τον ορισμό της q-αντιπροσωπευτικής οικογένειας, υπάρχει $X' \in \mathcal{S}'$ τέτοιο ώστε $X' \cap Z = \emptyset$. Αφού $Y' \cap X' = \emptyset$, έπεται $X' \cap Y = \emptyset$.

Στη συνέχεια εφαρμόζουμε το Λήμμα 3.8 και παίρνουμε το ακόλουθο αποτέλεσμα.

Λήμμα 3.9. *Έστω U σύνολο με n στοιχεία, $\mathcal{S}$ p-οικογένεια υποσυνόλων του U, q ακέραιος και $x \in (0,1)$. Τότε υπάρχει αλγόριθμος, που υπολογίζει σε χρόνο*

$$\mathcal{O}\Big(|\mathcal{S}| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\Big)$$

μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ τέτοια ώστε $|\mathcal{S}'| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log|\mathcal{S}|$.

Τέλος, είμαστε έτοιμοι να διατυπώσουμε έναν ακόμα πιο γρήγορο αλγόριθμο για τον υπολογισμό αντιπροσωπευτικής οικογένειας σε ομοιόμορφα μητροειδή. Ο αλγόριθμος αυτός αποτελεί βελτίωση του αλγορίθμου του Θεωρήματος 3.3.

Θεώρημα 3.4. *Έστω U σύνολο με n στοιχεία, $\mathcal{S}$ p-οικογένεια υποσυνόλων του U, q ακέραιος και $x \in (0,1)$. Τότε υπάρχει αλγόριθμος, που υπολογίζει σε χρόνο*

$$\mathcal{O}\Big(|\mathcal{S}| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\Big)$$

μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ τέτοια ώστε $|\widehat{\mathcal{S}}| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)}$.

Απόδειξη. Θέτουμε $\mathcal{S} = \mathcal{S}_1$ και χρησιμοποιώντας το Λήμμα 3.9 υπολογίζουμε μια σειρά από αντιπροσωπευτικές οικογένειες

$$\mathcal{S}_2 \subseteq^q_{rep} \mathcal{S}_1, \ldots, \mathcal{S}_m \subseteq^q_{rep} \mathcal{S}_{m-1},$$

όπου m ο ελάχιστος ακέραιος για τον οποίο $|\mathcal{S}_m| \geq \frac{|\mathcal{S}_{m-1}|}{2}$. Με άλλα λόγια, $|\mathcal{S}_i| \leq \frac{|\mathcal{S}_i|}{2}\ \forall i \in \{1, \ldots, m-1\}$ και $|\mathcal{S}_m| \geq \frac{|\mathcal{S}_{m-1}|}{2}$. Από το Λήμμα 3.1 προκύπτει ότι η $\mathcal{S}_m$ είναι q-αντιπροσωπευτική οικογένεια της $\mathcal{S}$. Επίσης, από το Λήμμα 3.9 έχουμε

$$\begin{aligned} |\mathcal{S}_m| &\leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log|\mathcal{S}_{m-1}| \\ &\leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \cdot \log 2|\mathcal{S}_m| \\ \text{Άρα, } \frac{|\mathcal{S}_m|}{\log|\mathcal{S}_m|} &\leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)} \end{aligned}$$

Γνωρίζουμε ότι για αριθμούς a και b, αν $a \leq b$ τότε $a\log^2 a \leq b\log^2 b$. Οπότε,

$$\frac{|\mathcal{S}_m|}{\log|\mathcal{S}_m|} \log^2\Big(\frac{|\mathcal{S}_m|}{\log|\mathcal{S}_m|}\Big) \leq x^{-p}(1-x)^{-q} \cdot 2^{o(p+q)}$$

και από την παραπάνω ανισότητα προκύπτει

$$|\mathcal{S}_m| \leq \frac{|\mathcal{S}_m|}{\log|\mathcal{S}_m|} \log^2\Big(\frac{|\mathcal{S}_m|}{\log|\mathcal{S}_m|}\Big) \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)}$$

και άρα $|\mathcal{S}_m| \leq \frac{1}{x^p(1-x)^q} \cdot 2^{o(p+q)}$. Επίσης, από το Λήμμα 3.9, ο συνολικός χρόνος T υπολογισμού της $\mathcal{S}_m$ είναι

$$\begin{aligned} T &= \sum_{i=1}^{m-1} |\mathcal{S}_i| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n \\ &= \sum_{i=1}^{m-1} \mathcal{O}\Big(\frac{|\mathcal{S}|}{2^{i-1}} \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\Big) \\ &= \mathcal{O}\Big(|\mathcal{S}| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\Big) \end{aligned}$$

□

Θέτοντας $x = \frac{p}{p+q}$ στο Θεώρημα 3.4, ελαχιστοποιούμε το μέγεθος της αντιπροσωπευτικής οικογένειας και προκύπτει το ακόλουθο Πόρισμα.

Πόρισμα 3.1. *Έστω U σύνολο με n στοιχεία, $\mathcal{S}$ p-οικογένεια υποσυνόλων του U και ακέραιος q. Τότε υπάρχει αλγόριθμος, που υπολογίζει σε χρόνο*

$$\mathcal{O}\left(|\mathcal{S}| \cdot \left(\frac{p+q}{q}\right)^q \cdot 2^{o(p+q)} \cdot \log n\right)$$

μια q-αντιπροσωπευτική οικογένεια $\mathcal{S}'$ της $\mathcal{S}$ τέτοια ώστε $|\widehat{\mathcal{S}}| \leq \binom{p+q}{p} \cdot 2^{o(p+q)}$.

Κεφάλαιο 4

Εφαρμογές

Στο κεφάλαιο αυτό θα δούμε πως η «γρήγορη» κατασκευή αντιπροσωπευτικών οικογενειών μπορεί να χρησιμοποιηθεί για τον σχεδιασμό FPT αλγορίθμων.

4.1 ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ

Το πρώτο πρόβλημα που θα μελετήσουμε είναι το εξής:

ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ

Είσοδος: πίνακες αναπαράστασης $A_{\mathcal{M}_1}, \ldots, A_{\mathcal{M}_\ell}$ επί του σώματος $GF(s)$ των μητροειδών $\mathcal{M}_1 = (E, \mathcal{I}_1), \ldots, \mathcal{M}_\ell = (E, \mathcal{I}_\ell)$ και ακέραιος k

Ερώτηση: Υπάρχει $S \subseteq E$ με $|S| \geq k$ τέτοιο ώστε $S \in \mathcal{I}_i$ για κάθε $i \in \{1, \ldots, \ell\}$;

Προφανώς, για $\ell = 1$ η λύση του προβλήματος είναι τετριμμένη, αφού ψάχνουμε για ένα ανεξάρτητο σύνολο του μητροειδούς. Τι συμβαίνει, όμως, για $\ell \geq 2$; Αν η τομή μητροειδών παρέμενε μητροειδές, τότε το πρόβλημα θα είχε πάντα προφανή λύση και επομένως θα ανήκε στην κλάση P. Όπως δείξαμε στην παράγραφο 2.2.3, κάτι τέτοιο δεν ισχύει. Ωστόσο, και για $\ell = 2$ το πρόβλημα ανήκει στην κλάση P. Για την απόδειξη, παραπέμπουμε τον αναγνώστη στο [24], όπου δίνεται αλγόριθμος πολυωνυμικού χρόνου για το πρόβλημα.

Για $\ell \geq 3$, το πρόβλημα δυσκολεύει. Όχι μόνο ανήκει στην κλάση NP, αλλά είναι και NP-πλήρες. Θα αποδείξουμε την NP-πληρότητα για $\ell = 3$ και προφανώς το ίδιο θα ισχύει και για κάθε $\ell \geq 3$.

Θεώρημα 4.1. *Το πρόβλημα ΤΟΜΗ 3-ΜΗΤΡΟΕΙΔΩΝ είναι NP-πλήρες.*

Απόδειξη. Προφανώς, το πρόβλημα ανήκει στην κλάση NP. Οπότε, αρκεί να δείξουμε ότι υπάρχει αναγωγή από το πρόβλημα *ΧΑΜΙΛΤΟΝΙΑΝΗ ΔΙΑΔΡΟΜΗ* σε διμερή γραφήματα στο πρόβλημά μας. Στο πρόβλημα *ΧΑΜΙΛΤΟΝΙΑΝΗ ΔΙΑΔΡΟΜΗ* σε διμερή γραφήματα, η είσοδος είναι ένα διμερές γράφημα $G = (V, E)$ με μέρη V_L και V_R. Ο στόχος είναι να αποφασίσουμε αν υπάρχει μονοπάτι που επισκέπτεται κάθε κορυφή του G ακριβώς μια φορά. Γνωρίζουμε ότι το πρόβλημα *ΧΑΜΙΛΤΟΝΙΑΝΗ ΔΙΑΔΡΟΜΗ* παραμένει NP-πλήρες ακόμα και όταν περιορίσουμε την είσοδο σε διμερή γραφήματα.

Κατασκευάζουμε τρία μητροειδή $\mathcal{M}_L$, $\mathcal{M}_R$ και $\mathcal{M}_G$. Τα σύνολα E_L, E_R και E_G αποτελούνται από το σύνολο των ακμών E του G. Για το $\mathcal{M}_L$ ένα σύνολο $X \subseteq E$ είναι ανεξάρτητο αν κάθε κορυφή του V_L είναι άκρο σε δύο το πολύ από τις ακμές του X. Αντίστοιχα, για το $\mathcal{M}_R$ ένα σύνολο $X \subseteq E$ είναι

ανεξάρτητο αν κάθε κορυφή του V_R είναι άκρο σε δύο το πολύ από τις ακμές του X. Παρατηρούμε ότι τα $\mathcal{M}_L$ και $\mathcal{M}_R$ είναι όντως μητροειδή, και μάλιστα μητροειδή διαμέρισης. Τέλος, το $\mathcal{M}_G$ ορίζεται ως το γραφικό μητροειδές του γραφήματος G.

Ισχυριζόμαστε ότι ένα σύνολο X τουλάχιστον $n-1$ ακμών είναι ανεξάρτητο στα $\mathcal{M}_L$, $\mathcal{M}_R$ και $\mathcal{M}_G$ αν και μόνο αν το X είναι το σύνολο ακμών ενός μονοπατιού στο G που επισκέπτεται κάθε κορυφή του G ακριβώς μια φορά. Πράγματι, αφού X ανεξάρτητο στο μητροειδές $\mathcal{M}_G$, έπεται ότι το σύνολο X είναι ακυκλικό. Το X έχει μέγεθος τουλάχιστον $n-1$, άρα το γράφημα (V, X) είναι συνδετικό δένδρο (spanning tree) T του G, δηλαδή δένδρο που περιλαμβάνει όλες τις κορυφές του G. Επίσης, το X είναι ανεξάρτητο στα μητροειδή $\mathcal{M}_L$ και $\mathcal{M}_R$, οπότε ο μέγιστος βαθμός του T είναι ίσος με 2. Στο σημείο αυτό ολοκληρώνεται η απόδειξη, αφού ένα συνδετικό δέντρο με μέγιστο βαθμό 2 είναι μονοπάτι που επισκέπτεται κάθε κορυφή του G ακριβώς μια φορά. □

Στη συνέχεια, θα δώσουμε πιθανοτικό FPT αλγόριθμο για κάθε $\ell \geq 3$, με παραμέτρους το πλήθος ℓ των μητροειδών και το μέγεθος k του ανεξάρτητου συνόλου που ψάχνουμε. Πρώτα, όμως θα ορίσουμε ένα ενδιάμεσο πρόβλημα, για το οποίο θα δώσουμε ντετερμινιστικό FPT αλγόριθμο.

ΠΑΚΕΤΑΡΙΣΜΑ d-ΣΥΝΟΛΩΝ ΜΗΤΡΟΕΙΔΟΥΣ

Είσοδος: πίνακας αναπαράστασης $A_{\mathcal{M}}$ επί του σώματος $GF(s)$ του μητροειδούς $\mathcal{M} = (E, \mathcal{I})$ με $rank(\mathcal{M}) = kd$ και d-οικογένεια $\mathcal{S}$ υποσυνόλων του E

Ερώτηση: Υπάρχει $\mathcal{S}' \subseteq \mathcal{S}$ με $|\mathcal{S}'| = k$ τέτοια ώστε τα σύνολα της $\mathcal{S}'$ να είναι ξένα ανά δύο και $\bigcup_{S \in \mathcal{S}'} S \in \mathcal{I}$;

Θεώρημα 4.2. *Υπάρχει αλγόριθμος που επιλύει το πρόβλημα ΠΑΚΕΤΑΡΙΣΜΑ d-ΣΥΝΟΛΩΝ ΜΗΤΡΟΕΙΔΟΥΣ σε $\mathcal{O}\big(|\mathcal{S}|^{\mathcal{O}(1)} + k^{\mathcal{O}(dk)}\big)$ χρόνο, όπου ως μονάδα χρόνου θεωρούμε μια πράξη στο $GF(s)$.*

Απόδειξη. Χρησιμοποιώντας το Θεώρημα 3.1, υπολογίζουμε $d(k-1)$-αντιπροσωπευτική οικογένεια $\mathcal{T}$ της $\mathcal{S}$, μεγέθους το πολύ $\binom{dk}{d} \leq (ek)^d$. Για το σκοπό αυτό χρειαζόμαστε το πολύ $|\mathcal{S}|^{\mathcal{O}(1)}$ πράξεις επί του σώματος $GF(s)$.

Ισχυριζόμαστε ότι, υπάρχει οικογένεια $\mathcal{S}' \subseteq \mathcal{S}$ με $|\mathcal{S}'| = k$ τέτοια ώστε τα σύνολά της να είναι ανά δύο ξένα και $\bigcup_{X \in \mathcal{S}'} X \in \mathcal{I}$ αν και μόνο αν υπάρχει οικογένεια $\mathcal{T}' \subseteq \mathcal{T}$ με τις ίδιες ιδιότητες. Προφανώς, κάθε οικογένεια $\mathcal{T}' \subseteq \mathcal{T}$, η οποία έχει τις ζητούμενες ιδιότητες είναι και υποοικογένεια της $\mathcal{S}$. Για την ευθεία κατεύθυνση, θεωρούμε την οικογένεια $\mathcal{S}'$ η οποία αποτελεί λύση του προβλήματος και έχει τα περισσότερα κοινά σύνολα με την $\mathcal{T}$. Έστω, προς άτοπο, ότι η $\mathcal{S}'$ περιέχει κάποιο σύνολο $A \in \mathcal{S} \setminus \mathcal{T}$. Ορίζουμε

$$B = \bigcup_{X \in \mathcal{S}' \setminus \{A\}} X.$$

Τα σύνολα της $\mathcal{S}'$ είναι ξένα ανά δύο, άρα $A \cap B = \emptyset$ και

$$A \cup B = A \cup \left(\bigcup_{X \in \mathcal{S}' \setminus \{A\}} X \right) = \bigcup_{X \in \mathcal{S}'} X \in \mathcal{I}.$$

Επειδή $\mathcal{T} \subseteq^{d(k-1)}_{rep} \mathcal{S}$, έπεται ότι υπάρχει $A' \in \mathcal{T}$ τέτοιο ώστε $A' \cap B = \emptyset$ και $A' \cup B \in \mathcal{I}$. Αφού $A' \cap B = \emptyset$, το σύνολο A' είναι ξένο από κάθε σύνολο του $\mathcal{S}' \setminus \{A\}$. Συνεπώς, τα σύνολα της οικογένειας

$\mathcal{T}' = \big(\mathcal{S}' \setminus \{A\}\big) \cup \{A'\}$ είναι ξένα ανά δύο. Επιπλέον,

$$\bigcup_{X \in \mathcal{T}'} X = \Big(\bigcup_{X \in \mathcal{S}' \setminus \{A\}} X \Big) \cup A' = B \cup A' \in \mathcal{I}.$$

Άρα, η οικογένεια $\mathcal{T}'$ αποτελεί λύση του προβλήματος, η οποία περιλαμβάνει περισσότερα κοινά σύνολα με την οικογένεια $\mathcal{T}$ από ότι η $\mathcal{S}'$. Άτοπο, λόγω της επιλογής της $\mathcal{S}'$.

Πώς θα βρούμε, όμως, την οικογένεια $\mathcal{T}' \subseteq \mathcal{T}$ με τις ζητούμενες ιδιότητες; Θα ελέγξουμε κάθε υποοικογένεια της $\mathcal{T}$ με ακριβώς k σύνολα. Συνολικά, έχουμε

$$\binom{(ek)^d}{k} \leq e^{(d+1)k} \cdot k^{(d-1)k} \leq k^{(d+1)k} \cdot k^{(d-1)k} = k^{\mathcal{O}(kd)}$$

επιλογές. Για κάθε επιλογή της $\mathcal{T}'$, πρέπει να ελέγξουμε αν είναι όντως λύση του προβλήματος. Αυτό μπορεί να επιτευχθεί σε $(k+d)^{\mathcal{O}(1)}$ χρόνο. Άρα, για την εύρεση της $\mathcal{T}'$ χρειαζόμαστε

$$k^{\mathcal{O}(dk)} \cdot (k+d)^{\mathcal{O}(1)} = k^{\mathcal{O}(dk)}$$

χρόνο. Προσθέτοντας και τον χρόνο υπολογισμού της αντιπροσωπευτικής οικογένειας $\mathcal{T}$, προκύπτει το ζητούμενο. □

Στη συνέχεια, θα δώσουμε αλγόριθμο για μια ειδική περίπτωση του προβλήματος. Στην περίπτωση αυτή, τα μητροειδή, άρα και οι αντίστοιχοι πίνακες αναπαράστασης, έχουν τάξη ίση με k.

Λήμμα 4.1. *Θεωρούμε την ειδική περίπτωση του προβλήματος ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ στην οποία τα μητροειδή έχουν όλα την ίδια τάξη, έστω k. Τότε υπάρχει αλγόριθμος που επιλύει το πρόβλημα σε $\mathcal{O}\big(|E|^{\mathcal{O}(1)} + k^{\mathcal{O}(k\ell)}\big)$ χρόνο, όπου ως μονάδα χρόνου θεωρούμε μια πράξη στο $GF(s)$.*

Απόδειξη. Ανάγουμε το πρόβλημα *ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ* στην ειδική περίπτωση του προβλήματος *ΠΑΚΕΤΑΡΙΣΜΑ d-ΣΥΝΟΛΩΝ ΜΗΤΡΟΕΙΔΟΥΣ*. Στο πρώτο πρόβλημα, η είσοδος είναι ℓ πίνακες αναπαράστασης $A_{\mathcal{M}_1}, \ldots, A_{\mathcal{M}_\ell}$ των μητροειδών $\mathcal{M}_1 = (E, \mathcal{I}_1), \ldots, \mathcal{M}_\ell = (E, \mathcal{I})$ και ακέραιος k. Αρχικά, κατασκευάζουμε ℓ αντίγραφα $E_1, \ldots, E_\ell$ του συνόλου E και για κάθε $i \in \{1, \ldots, \ell\}$, θεωρούμε το μητροειδές $\mathcal{M}_i^* = (E_i, \mathcal{I}_i)$. Προφανώς, ένας πίνακας αναπαράστασης του $\mathcal{M}_i^*$ είναι ο $A_{\mathcal{M}_i}$. Θέτουμε

$$\mathcal{M}^* = \mathcal{M}_1^* \oplus \cdots \oplus \mathcal{M}_\ell^*.$$

Από την Πρόταση 2.4 μπορούμε να βρούμε έναν πίνακα αναπαράστασης $A_{\mathcal{M}^*}$ του μητροειδούς $\mathcal{M}^*$ σε πολυωνυμικό χρόνο με $rank(\mathcal{M}^*) = k\ell$. Το $\mathcal{M}^*$ είναι το μητροειδές εισόδου του προβλήματος *ΠΑΚΕΤΑΡΙΣΜΑ d-ΣΥΝΟΛΩΝ ΜΗΤΡΟΕΙΔΟΥΣ* και $d = \ell$.

Για κάθε στοιχείο $x \in E$, θέτουμε $S_x = \{x_1, \ldots, x_\ell\}$, όπου x_i το αντίγραφο του x στο E_i. Επίσης, θέτουμε $\mathcal{S} = \{S_x \mid x \in E\}$. Προφανώς, όλα τα σύνολα στην $\mathcal{S}$ είναι ανά δύο ξένα. Επιπλέον, ένα σύνολο $X \subseteq E$ είναι ανεξάρτητο στα $\mathcal{M}_1, \ldots, \mathcal{M}_\ell$ αν και μόνο αν $\bigcup_{x \in X} S_x$ είναι ανεξάρτητο στο $\mathcal{M}^*$.

Άρα το στιγμιότυπο $\langle \mathcal{M}_1, \ldots, \mathcal{M}_\ell, k \rangle$ ανήκει στο πρόβλημα *ΠΑΚΕΤΑΡΙΣΜΑ ℓ-ΣΥΝΟΛΩΝ ΜΗΤΡΟΕΙΔΟΥΣ* αν και μόνο αν υπάρχει $\mathcal{S}' \subseteq \mathcal{S}$ με $|\mathcal{S}'| = k$ τέτοια ώστε όλα τα σύνολα της $\mathcal{S}'$ να είναι ξένα ανά δύο και το σύνολο $\bigcup_{S_x \in \mathcal{S}'} S_x$ να είναι ανεξάρτητο στο $\mathcal{M}^*$.

Επομένως, αρκεί να εφαρμόσουμε τον αλγόριθμο του Θεωρήματος 4.2 για το στιγμιότυπο του προβλήματος *ΠΑΚΕΤΑΡΙΣΜΑ ℓ-ΣΥΝΟΛΩΝ ΜΗΤΡΟΕΙΔΟΥΣ* που περιγράψαμε παραπάνω. Η λύση που θα προκύψει, μετά από κατάλληλο μετασχηματισμό θα αποτελεί και λύση του αρχικού μας προβλήματος. Άρα, χρειαζόμαστε $\mathcal{O}\big(|E|^{\mathcal{O}(1)} + k^{\mathcal{O}(k\ell)}\big)$ χρόνο. □

Στο σημείο αυτό θα θέλαμε να παραβλέψουμε την υπόθεση του Λήμματος 4.1 ότι τα μητροειδή $\mathcal{M}_1, \ldots, \mathcal{M}_\ell$ έχουν όλα την ίδια τάξη k. Αν το πλήθος στοιχείων του σώματος $GF(s)$ είναι αρκετά μεγάλο, τότε χρησιμοποιώντας την Πρόταση 2.5 μετατρέπουμε τους πίνακες αναπαράστασης των $\mathcal{M}_1, \ldots, \mathcal{M}_\ell$ σε πίνακες τάξης το πολύ k. Διαφορετικά, μετατρέπουμε τους πίνακες αναπαράστασης των $\mathcal{M}_1, \ldots, \mathcal{M}_\ell$ σε πίνακες με στοιχεία από σώμα $GF(s')$, όπου το s' είναι αρκετά μεγάλο για να εφαρμόσουμε την Πρόταση 2.5. Λαμβάνοντας υπόψιν τα παραπάνω, καθώς και την απόδειξη της Πρότασης 2.5 προκύπτει το ακόλουθο θεώρημα.

Θεώρημα 4.3. *Υπάρχει πιθανοτικός αλγόριθμος που επιλύει το πρόβλημα ΤΟΜΗ ℓ-ΜΗΤΡΟΕΙΔΩΝ σε $\mathcal{O}\big(|E|^{\mathcal{O}(1)} + k^{\mathcal{O}(\ell k)}\big)$ χρόνο, όπου ως μονάδα χρόνου θεωρούμε μια πράξη στο $GF(s)$.*

4.2 *ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ*

Το πρόβλημα k-*ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ*, στο οποίο διερωτόμαστε αν σε ένα κατευθυνόμενο γράφημα υπάρχει κύκλος μήκους ακριβώς k, είναι γνωστό ότι ανήκει στην κλάση *NP*. Για το πρόβλημα αυτό υπάρχει αλγόριθμος, ο οποίος χρησιμοποιεί την τεχνική χρωματικής κωδικοποίησης (color coding) που αναπτύχθηκε από τους Alon, Yuster και Zwick στο [37] και τρέχει σε χρόνο $2^{\mathcal{O}(k)}n^{\mathcal{O}(1)}$. Στο πρόβλημα *ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ*, από την άλλη, παρά την όποια φαινομενική ομοιότητα δεν μπορεί να εφαρμοστεί.

ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ

Είσοδος: κατευθυνόμενο γράφημα $D = (V, A)$ με $|V| = n$, $|A| = m$ και θετικός ακέραιος k

Ερώτηση: Υπάρχει κατευθυνόμενος κύκλος μήκους τουλάχιστον k στο D;

Τα δύο προβλήματα διαφέρουν, αφού είναι πιθανόν ένας κύκλος μήκους τουλάχιστον k να είναι πολύ μεγαλύτερος από k (θα μπορούσε να είναι ακόμα και Χαμιλτονιανός). Οπότε οι τεχνικές για την επίλυση του πρώτου προβλήματος δεν μπορούν να εφαρμοστούν και στο δεύτερο.

Θεωρούμε κατευθυνόμενο γράφημα $D = (V, A)$, θετικό ακέραιο k και το ομοιόμορφο μητροειδές $U_{n,2k} = (E, \mathcal{I})$ με $E = V$ και $\mathcal{I} = \{S \subseteq V \mid |S| \leq 2k\}$. Για κορυφές $u, v \in V$ ορίζουμε την οικογένεια

$$\mathcal{P}^i_{uv} = \big\{X \subseteq V \mid u, v \in X, |X| = i \text{ και υπάρχει } uv\text{-μονοπάτι στο } D \text{ το οποίο επισκέπτεται όλες τις κορυφές του } X\big\}$$

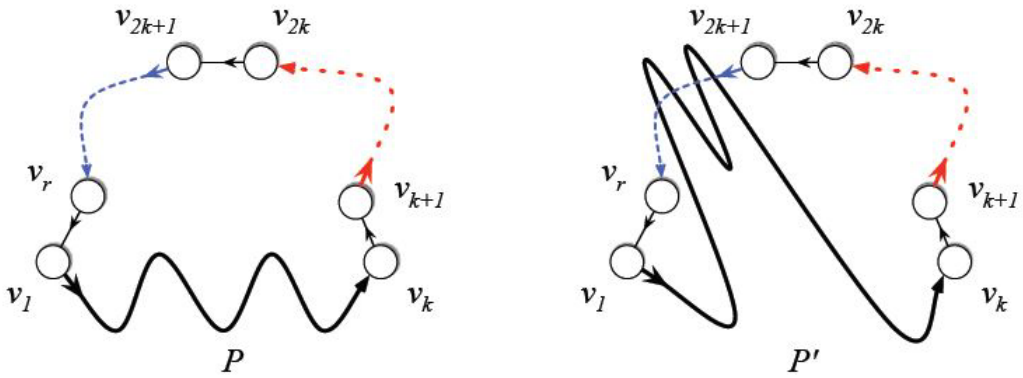
Ορίζουμε ως εναγόμενο υπογράφημα το γράφημα που προκύπτει με την αφαίρεση κάποιων κορυφών του αρχικού γραφήματος και παρατηρούμε το εξής:

Λήμμα 4.2. *Έστω D ένα κατευθυνόμενο γράφημα. Το D έχει κατευθυνόμενο κύκλο μήκους τουλάχιστον k αν και μόνο αν υπάρχει ένα ζεύγος κορυφών $u, v \in V$ και σύνολο $X \in \widehat{\mathcal{P}}^k_{uv} \subseteq^k_{rep} \mathcal{P}^k_{uv}$ τέτοια ώστε το D να έχει κατευθυνόμενο κύκλο C, στον οποίο οι κορυφές του X ενάγουν κατευθυνόμενο μονοπάτι.*

Απόδειξη. Η αντίστροφη κατεύθυνση είναι προφανής, αφού αν ο κύκλος C περιλαμβάνει μονοπάτι μήκους τουλάχιστον k, τότε προφανώς το μήκος του C είναι τουλάχιστον k.

Για την ευθεία κατεύθυνση, θεωρούμε $C^* = v_1v_2 \dots v_rv_1$ τον μικρότερο κατευθυνόμενο κύκλο στο D μήκους τουλάχιστον k. Με άλλα λόγια, $r \geq k$ και δεν υπάρχει κύκλος μήκους r' με $k \leq r' < r$. Διακρίνουμε δύο περιπτώσεις:

- **1η Περίπτωση** Για $r \leq 2k$, θέτουμε $u = v_1$, $v = v_k$ και μονοπάτια $P = v_1v_2 \dots v_k$ και $Q = v_{k+1} \dots v_r$. Επειδή το μονοπάτι Q έχει το πολύ k κορυφές, $V(P) \cap V(Q) = \emptyset$ και $\widehat{\mathcal{P}}^k_{uv} \subseteq^k_{rep} \mathcal{P}^k_{uv}$, υπάρχει κατευθυνόμενο μονοπάτι P' τέτοιο ώστε $X = V(P') \in \widehat{\mathcal{P}}^k_{uv}$ και $X \cap V(Q) = \emptyset$. Αντικαθιστώντας το P με το P' στον κύκλο C^*, προκύπτει κατευθυνόμενος κύκλος C μήκους τουλάχιστον k, στον οποίο οι κορυφές του συνόλου X ενάγουν μονοπάτι.
- **2η Περίπτωση** Για $r > 2k$, θέτουμε $u = v_1$, $v = v_k$ και μονοπάτια $P = v_1 \dots v_k$, $Q = v_{k+1} \dots v_{2k}$ και $R = v_{2k+1} \dots v_r$. Επειδή το μονοπάτι Q έχει το πολύ k κορυφές, $V(P) \cap V(Q) = \emptyset$ και $\widehat{\mathcal{P}}^k_{uv} \subseteq^k_{rep} \mathcal{P}^k_{uv}$, υπάρχει κατευθυνόμενο μονοπάτι P' τέτοιο ώστε $X = V(P') \in \widehat{\mathcal{P}}^k_{uv}$ και $X \cap V(Q) = \emptyset$. Το P' δεν είναι απαραίτητα ξένο με το μονοπάτι R, οπότε αντικαθιστώντας το P με το P' μπορεί να προκύψουν κοινές κορυφές μεταξύ των μονοπατιών P' και R (βλέπε παρακάτω σχήμα).



Θέτουμε C' τον κύκλο που προκύπτει αντικαθιστώντας το P με το P' και ισχυριζόμαστε ότι $X \cap V(R) = \emptyset$.

Έστω $X \cap V(R) \neq \emptyset$ και έστω v_α η τελευταία κορυφή που «επισκεπτόμαστε» στο μονοπάτι P', η οποία ανήκει και στο μονοπάτι R. Θέτουμε $P'[v_\alpha, v_k]$ το υπομονοπάτι του P' από την v_α στην v_k και

$$R' = \begin{cases} \emptyset & \text{για } v_\alpha = v_{2k+1} \\ R[v_{2k+1}, v_{\alpha-1}] & \text{διαφορετικά} \end{cases}$$

όπου $R[v_{2k+1}, v_{\alpha-1}]$ το υπομονοπάτι του R από την v_{2k+1} στην $v_{\alpha-1}$. Παρατηρούμε ότι το γράφημα $\overline{C} = P'[v_\alpha, v_k]QR'$ είναι μια τρύπα στο γράφημα D, δηλαδή ένα εναγόμενο υπογράφημα του D που είναι κύκλος. Προφανώς $|\overline{C}| \geq |Q| = k$. Επειδή $v_1 \notin P'[v_\alpha, v_k]$, $|P'[v_\alpha, v_k]| < |P'| = |P|$ και επειδή $v_\alpha \notin R'$, $|R'| < |R|$. Τελικά έχουμε

$$k \leq |\overline{C}| = |P'[v_\alpha, v_k]| + |Q| + |R'| < |P| + |Q| + |R| = |C^*|$$

Δηλαδή, ο $\overline{C}$ είναι κατευθυνόμενος κύκλος μήκους τουλάχιστον k, μικρότερος από τον C^*. Άτοπο, λόγω της επιλογής του C^*. Άρα, αντικαθιστώντας το P με το P' στον C^*, προκύπτει κατευθυνόμενος κύκλος C μήκους τουλάχιστον k, στον οποίο οι κορυφές του συνόλου X ενάγουν μονοπάτι.

□

Για τον υπολογισμό q-αντιπροσωπευτικής οικογένειας $\widehat{\mathcal{P}}^k_{uv}$ της $\mathcal{P}^k_{uv}$ κατασκευάζουμε τον παρακάτω αλγόριθμο ωμής βίας.

Λήμμα 4.3. *Έστω D κατευθυνόμενο γράφημα με n κορυφές και m ακμές, $u \in V(D)$ και το ομοιόμορφο μητροειδές $U_{n,\ell} = (E, \mathcal{I})$ με $E = V(D)$ και $\mathcal{I} = \{S \subseteq V(D) \mid |S| \leq \ell\}$. Τότε για κάθε $p \in \{2, \ldots, \ell\}$ και $v \in V(D) \setminus \{u\}$, υπάρχει αλγόριθμος που υπολογίζει μια $(\ell - p)$-αντιπροσωπευτική οικογένεια $\widehat{\mathcal{P}}^p_{uv}$ της $\mathcal{P}^p_{uv}$, μεγέθους το πολύ*

$$\binom{\ell}{p} \cdot 2^{o(\ell)}$$

σε χρόνο

$$\mathcal{O}\left(2^{o(\ell)} m \log n \max_{i \in [p]} \left\{ \binom{\ell}{i-1} \left(\frac{\ell}{\ell - i}\right)^{\ell - i} \right\}\right).$$

Επιπλέον, στον ίδιο χρόνο υπολογίζει και μια διάταξη των κορυφών κάθε συνόλου της οικογένειας $\widehat{\mathcal{P}}^p_{uv}$ τέτοια ώστε να αντιστοιχεί σε ένα μονοπάτι του D.

Απόδειξη. Έστω $V(D) = \{u, v_1, \ldots, v_{n-1}\}$ οι κορυφές του γραφήματος $D = (V, A)$ και ο $(p-1) \times (n-1)$ πίνακας $\mathcal{D}$. Στις γραμμές του αντιστοιχίζουμε τους ακεραίους $2, \ldots, p$ και στις στήλες του τις κορυφές $v_1, \ldots, v_{n-1}$. Στη θέση $\mathcal{D}[i, v]$ θα αποθηκεύσουμε την οικογένεια $\widehat{\mathcal{P}}^i_{uv} \subseteq^{\ell-i}_{rep} \mathcal{P}^i_{uv}$. Συμπληρώνουμε τον πίνακα $\mathcal{D}$ κατά αύξουσα σειρά γραμμών. Για $i = 2$, $\mathcal{D}[2, v] = \big\{\{u, v\}\big\}$ αν $uv \in A(D)$.
Έστω ότι έχουμε συμπληρώσει τον πίνακα μέχρι και την i-οστή γραμμή. Θέτουμε

$$\mathcal{N}^{i+1}_{uv} = \bigcup_{w \in N^-(v)} \widehat{\mathcal{P}}^i_{uw} \bullet \{v\},$$

όπου με $N^-(v)$ συμβολίζουμε το σύνολο των κορυφών w για τις οποίες $wv \in A(D)$.

Για να συμπληρώσουμε τη θέση $\mathcal{D}[i+1, v]$ του πίνακα $\mathcal{D}$ εργαζόμαστε ως εξής:
Αρχικά, παρατηρούμε ότι

$$\mathcal{N}^{i+1}_{uv} = \bigcup_{w \in N^-(v)} \mathcal{D}[i, w] \bullet \{v\}$$

Έχουμε ήδη υπολογίσει την οικογένεια στη θέση $\mathcal{D}[i, w]$ για κάθε $w \in N^-(v)$. Από το Πόρισμα 3.1, έχουμε $|\widehat{\mathcal{P}}^i_{uw}| \leq \binom{\ell}{i} 2^{o(\ell)}$ και άρα $|\mathcal{N}^{i+1}_{uv}| \leq \deg^-(v) \binom{\ell}{i} 2^{o(\ell)}$, όπου με $\deg^-(v)$ συμβολίζουμε το πλήθος των κορυφών w για τις οποίες $wv \in A(D)$. Επιπλέον, μπορούμε να υπολογίσουμε την οικογένεια $\widehat{\mathcal{N}}^{i+1}_{uv}$ σε $\mathcal{O}\big(\deg^-(v) \binom{\ell}{i} 2^{o(\ell)}\big)$ χρόνο. Στη συνέχεια, χρησιμοποιώντας ξανά το Πόρισμα 3.1, υπολογίζουμε οικογένεια $\widehat{\mathcal{N}}^{i+1}_{uv} \subseteq^{\ell-(i+1)}_{rep} \mathcal{N}^{i+1}_{uv}$ σε

$$\mathcal{O}\left(t \cdot \left(\frac{\ell}{\ell - (i+1)}\right)^{\ell-(i+1)} \cdot \log n\right)$$

χρόνο, όπου $t = \deg^-(v) \binom{\ell}{i} 2^{o(\ell)}$.

Ισχυριζόμαστε ότι $\mathcal{N}^{i+1}_{uv} \subseteq^{\ell-(i+1)}_{rep} \mathcal{P}^{i+1}_{uv}$.
Πράγματι, έστω $S \in \mathcal{P}^{i+1}_{uv}$ και σύνολο Y με $|Y| = \ell - (i+1)$ και $S \cap Y = \emptyset$. Επειδή $S \in \mathcal{P}^{i+1}_{uv}$, υπάρχει ένα κατευθυνόμενο μονοπάτι $P = u\alpha_1 \ldots \alpha_{i-1} v$ στο D τέτοιο ώστε $S = \{u, \alpha_1, \ldots, \alpha_{i-1}, v\}$ και $\alpha_{i-1} \in N^-(v)$. Το υπομονοπάτι $P[u, \alpha_{i-1}]$ του P από την u στην α_{i-1} ανήκει στην οικογένεια $\mathcal{P}^i_{u\alpha_{i-1}}$, δηλαδή $X^* = S \setminus \{u\} \in \mathcal{P}^i_{u\alpha_{i-1}}$. Θέτουμε $Y^* = Y \cup \{v\}$ και παρατηρούμε ότι $X^* \cap Y^* = \emptyset$ και $|Y^*| = \ell - i$. Στον πίνακα $\mathcal{D}$ έχουμε ήδη υπολογίσει την οικογένεια $\widehat{\mathcal{P}}^i_{u\alpha_{i-1}} \subseteq^{\ell-i}_{rep} \mathcal{P}^i_{u\alpha_{i-1}}$, οπότε γνωρίζουμε

ότι υπάρχει σύνολο $\widehat{X}^* \in \widehat{\mathcal{P}}^i_{u\alpha_{i-1}}$ τέτοιο ώστε $\widehat{X}^* \cap Y^* = \emptyset$. Επειδή $\alpha_{i-1} \in N^-(v)$ και $\widehat{X}^* \cap \{v\} = \emptyset$, $\widehat{X}^* \bullet \{v\} = \widehat{X}^* \cup \{v\} \in \mathcal{N}^{i+1}_{uv}$. Άρα, για $S' = \widehat{X}^* \cup \{v\} \in \mathcal{N}^{i+1}_{uv}$ ισχύει $S' \cap Y = \emptyset$ και έπεται το ζητούμενο.

Από το Λήμμα 3.3, έπεται $\widehat{\mathcal{N}}^{i+1}_{uv} = \widehat{\mathcal{P}}^{i+1}_{uv} \subseteq^{\ell-(i+1)}_{rep} \mathcal{P}^{i+1}_{uv}$. Τελικά, στη θέση $\mathcal{D}[i+1, v]$ του πίνακα $\mathcal{D}$ αποθηκεύουμε την οικογένεια $\widehat{\mathcal{N}}^{i+1}_{uv}$. Διατάσσουμε τις κορυφές ενός συνόλου της οικογένειας $\widehat{\mathcal{P}}^p_{uv}$ με βάση τη σειρά με την οποία κατασκευάστηκε χρησιμοποιώντας την πράξη $\bullet$. Τότε κάθε διατεταγμένο σύνολο στην οικογένεια $\widehat{\mathcal{P}}^p_{uv}$ αναπαριστά ένα μονοπάτι στο γράφημα.

Όσον αφορά τον χρόνο του αλγορίθμου, φράσσεται από

$$\begin{aligned} & \mathcal{O}\left(\sum_{i=2}^{p}\sum_{j=1}^{n-1} \deg^-(v_j)\binom{\ell}{i-1}\left(\frac{\ell}{\ell-i}\right)^{\ell-i} 2^{o(\ell)}\log n\right) \\ = \; & \mathcal{O}\left(2^{o(\ell)}\log n \sum_{i=2}^{p}\sum_{j=1}^{n-1} \deg^-(v_j)\binom{\ell}{i-1}\left(\frac{\ell}{\ell-i}\right)^{\ell-i}\right) \\ = \; & \mathcal{O}\left(2^{o(\ell)} m \log n \max_{i\in[p]}\left\{\binom{\ell}{i-1}\left(\frac{\ell}{\ell-i}\right)^{\ell-i}\right\}\right). \end{aligned}$$

□

Στο σημείο αυτό, είμαστε έτοιμοι να δώσουμε FPT αλγόριθμο για το πρόβλημα *ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ*.

Θεώρημα 4.4. *Υπάρχει αλγόριθμος που επιλύει το πρόβλημα ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ σε $\mathcal{O}(8^{k+o(k)}mn^2)$ χρόνο.*

Απόδειξη. Έστω D κατευθυνόμενο γράφημα. Από το Λήμμα 4.2, το D έχει κατευθυνόμενο κύκλο μήκους τουλάχιστον k αν και μόνο αν υπάρχουν κορυφές $u, v \in V(D)$ και μονοπάτι P' με $V(P') \in \widehat{\mathcal{P}}^k_{uv} \subseteq^k_{rep} \mathcal{P}^k_{uv}$ τέτοιο ώστε το D να έχει κατευθυνόμενο κύκλο C, στον οποίο οι κορυφές του P' ενάγουν κατευθυνόμενο μονοπάτι. Αρχικά, υπολογίζουμε $\widehat{\mathcal{P}}^k_{uv} \subseteq^k_{rep} \mathcal{P}^k_{uv}$ για κάθε $u, v \in V(D)$. Για τον σκοπό αυτό, χρησιμοποιούμε το Λήμμα 4.3 για κάθε $u \in V(D)$ με $\ell = 2k$ και $p = k$ και χρειαζόμαστε $\mathcal{O}(8^{k+o(k)}mn\log n)$ χρόνο. Επιπλέον, για κάθε $X \in \widehat{\mathcal{P}}^k_{uv}$ υπολογίζουμε ένα κατευθυνόμενο uv-μονοπάτι P_X χρησιμοποιώντας τις κορυφές του X. Θέτουμε

$$\mathcal{Q} = \bigcup_{u,v\in V(D)} \widehat{\mathcal{P}}^k_{uv}.$$

Για κάθε $X \in \mathcal{Q}$ και το αντίστοιχο uv-μονοπάτι P_X, ελέγχουμε αν υπάρχει vu-μονοπάτι στο D αποφεύγοντας όλες τις κορυφές του X εκτός από τις u και v. Ο έλεγχος αυτός μπορεί να πραγματοποιηθεί χρησιμοποιώντας κάποιον από τους αλγορίθμους BFS ή DFS σε $\mathcal{O}(m+n)$ χρόνο. Αν όντως βρεθεί τέτοιο μονοπάτι, τότε η απάντηση στο πρόβλημα είναι θετική και ο ζητούμενος κατευθυνόμενος κύκλος είναι αυτός που προκύπτει από την ένωση του P_X με το vu-μονοπάτι που βρέθηκε. Διαφορετικά, αν δεν βρεθεί τέτοιο μονοπάτι για κανένα $X \in \mathcal{Q}$, τότε δεν υπάρχει κατευθυνόμενος κύκλος μήκους τουλάχιστον k στο D. Η ορθότητα του αλγορίθμου προκύπτει από το Λήμμα 4.2.

Από το Πόρισμα 3.1, το μέγεθος μιας οικογένειας $\widehat{\mathcal{P}}^k_{uv}$ είναι το πολύ ίσο με $\binom{2k}{k}2^{o(k)}$ και υπάρχουν το πολύ n^2 επιλογές για τα πιθανά ζεύγη κορυφών u, v. Έπεται ότι το μέγεθος της $\mathcal{Q}$ είναι το πολύ ίσο με $n^2\binom{2k}{k}2^{o(k)} \leq n^2 4^{k+o(k)}$. Τελικά, ο αλγόριθμος χρειάζεται το πολύ

$$\mathcal{O}\big(8^{k+o(k)}mn\log n + 4^{k+o(k)}(n^2m + n^3)\big) = \mathcal{O}\big(8^{k+o(k)}mn^2\big).$$

□

Στην πραγματικότητα, υπάρχει ακόμα πιο γρήγορος αλγόριθμος για το πρόβλημα *ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ*. Στον αλγόριθμο που δώσαμε παραπάνω ο χρόνος εξαρτάται κυρίως από τον υπολογισμό των αντιπροσωπευτικών οικογενειών $\widehat{\mathcal{P}}^p_{uv} \subseteq^q_{rep} \mathcal{P}^p_{uv}$ για $2 \leq p \leq k$ και $q = 2k - p$. Για τον σκοπό αυτό, χρησιμοποιήσαμε το Θεώρημα 3.4 με το x να παίρνει την τιμή που ελαχιστοποιεί το μέγεθος της αντιπροσωπευτικής οικογένειας, δηλαδή $x = \frac{p}{p+q}$. Μπορούμε, όμως, να διαλέξουμε x το οποίο να ελαχιστοποιεί τον χρόνο υπολογισμού της αντιπροσωπευτικής οικογένειας και με αυτό τον τρόπο να βελτιώσουμε τον χρόνο επίλυσης του προβλήματός μας.

Ποιά είναι, λοιπόν, η τιμή του x που ελαχιστοποιεί χρόνο υπολογισμού της $\widehat{\mathcal{P}}^p_{uv} \subseteq^q_{rep} \mathcal{P}^p_{uv}$, για $2 \leq p \leq k$ και $q = 2k - p$; Συμβολίζουμε με $s_{p,q}$ το μέγεθος της $\widehat{\mathcal{P}}^p_{uv}$. Επειδή $\widehat{\mathcal{P}}^p_{uv} \subseteq^q_{rep} \mathcal{N}^p_{uv} \subseteq^q_{rep} \mathcal{P}^p_{uv}$, ο υπολογισμός της $\widehat{\mathcal{P}}^p_{uv}$ εξαρτάται από το μέγεθος της $\mathcal{N}^p_{uv}$, το οποίο εξαρτάται με τη σειρά του από το μέγεθος της $\widehat{\mathcal{P}}^{p-1}_{uw}$. Άρα, $|\mathcal{N}^p_{uv}| \leq s_{p-1,q+1} \cdot n$. Οπότε οι τιμές των $s_{p-1,q+1}$ και $s_{p,q}$ είναι σχεδόν ίσες, δηλαδή $s_{p-1,q+1} \approx s_{p,q}$ και από Θεώρημα 3.4, ο χρόνος υπολογισμού της $\widehat{\mathcal{P}}^p_{uv} \subseteq^q_{rep} \mathcal{N}^p_{uv} \subseteq^q_{rep} \mathcal{P}^p_{uv}$ είναι ίσος με

$$\begin{aligned}
&\mathcal{O}\left(|\mathcal{N}^p_{uv}| \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot \log n\right) \\
=&\mathcal{O}\left(s_{p,q} \cdot \frac{1}{(1-x)^q} \cdot 2^{o(p+q)} \cdot n \log n\right) \\
=&\mathcal{O}\left(\frac{1}{x^p(1-x)^{2q}} \cdot 2^{o(p+q)} \cdot n \log n\right)
\end{aligned}$$

Για να ελαχιστοποιήσουμε τον παραπάνω χρόνο υπολογισμού, αρκεί να ελαχιστοποιήσουμε την συνάρτηση $f(x) = x^{-p} \cdot (1-x)^{-2q}$. Γνωρίζουμε ότι, αν υπάρχει τιμή x^* για την οποία $f'(x^*) = 0$ και $f''(x^*) > 0$, τότε η τιμή αυτή θα ελαχιστοποιεί και την $f(x)$. Οπότε, υπολογίζουμε

$$\begin{aligned}
f'(x) &= 0 \\
-p \cdot x^{-p-1}(1-x)^{-2q} + 2q \cdot x^{-p}(1-x)^{-2q-1} &= 0 \\
-p \cdot (1-x) + 2q \cdot x &= 0 \\
x &= \frac{p}{p+2q}
\end{aligned}$$

και θέτουμε $x^* = \frac{p}{p+2q}$. Αν ισχύει και $f''(x^*) > 0$, τότε, όντως, η $f(x)$ ελαχιστοποιείται για x^*.

$$\begin{aligned}
f'(x) &= x^{-p}(1-x)^{-2q}(-p \cdot x^{-1} + 2q \cdot (1-x)^{-1}) \\
&= f(x) \cdot (-p \cdot x^{-1} + 2q \cdot (1-x)^{-1}) \\
f''(x) &= f(x) \cdot (p \cdot x^{-2} + 2q \cdot (1-x)^{-2}) + f'(x) \cdot (-p \cdot x^{-1} + 2q \cdot (1-x)^{-1}) \\
f''(x^*) &= f(x^*) \cdot \left(p \cdot (x^*)^{-2} + 2q \cdot (1-x^*)^{-2}\right) > 0
\end{aligned}$$

Τελικά, ο χρόνος τον οποίο χρειαζόμαστε για να υπολογίσουμε $\widehat{\mathcal{P}}^p_{uv} \subseteq^q_{rep} \mathcal{P}^p_{uv}$ ελαχιστοποιείται για $x = \frac{p}{p+2q}$.

Λήμμα 4.4. *Έστω D ένα κατευθυνόμενο γράφημα με n κορυφές και m ακμές, $u \in V(D)$ και το ομοιόμορφο μητροειδές $U_{n,\ell} = (E, \mathcal{I})$ με $E = V(D)$ και $\mathcal{I} = \{S \subseteq V(D) \mid |S| \leq \ell\}$. Τότε για κάθε $p \in \{2, \ldots, \ell\}$ και $v \in V(D) \setminus \{u\}$, υπάρχει αλγόριθμος που υπολογίζει μια $(\ell - p)$-αντιπροσωπευτική οικογένεια $\widehat{\mathcal{P}}^p_{uv}$ της $\mathcal{P}^p_{uv}$ μεγέθους το πολύ*

$$\left(\frac{2\ell - p}{p}\right)^p \left(\frac{2\ell - p}{2\ell - 2p}\right)^{\ell - p} \cdot 2^{o(\ell)}$$

σε χρόνο

$$\mathcal{O}\left(2^{o(\ell)} m \log n \max_{i\in[p]}\left\{\left(\frac{2\ell-i}{i}\right)^{i}\left(\frac{2\ell-i}{2\ell-2i}\right)^{2\ell-2i}\right\}\right).$$

Απόδειξη. Η απόδειξη είναι ίδια με αυτή του Λήμματος 4.3, με την μόνη διαφορά ότι για τον υπολογισμό των οικογενειών $\widehat{\mathcal{N}}^j_{uv} = \widehat{\mathcal{P}}^j_{uv}$, χρησιμοποιούμε το Θεώρημα 3.4 με την τιμή του x να είναι ίση με

$$x_j = \frac{j}{j+2(\ell-j)} = \frac{j}{2\ell-j}.$$

Έστω $s_{j,\ell-j}$ το μέγεθος της αντιπροσωπευτικής οικογένειας $\widehat{\mathcal{N}}^j_{uv} = \widehat{\mathcal{P}}^j_{uv}$, δηλαδή

$$s_{j,\ell-j} = x_j^{-j}(1-x_j)^{\ell-j}\cdot 2^{o(\ell)}.$$

Έστω ότι έχουμε υπολογίσει $\widehat{\mathcal{P}}^j_{uw}$ μεγέθους $s_{j,\ell-j}$ και την έχουμε αποθηκεύσει στην θέση $\mathcal{D}[j,w]$ του πίνακα $\mathcal{D}$, για κάθε $j \le i$ και $w \in \{v_1,\ldots,v_{n-1}\}$. Επειδή έχουμε ορίσει $\mathcal{N}^{i+1}_{uv} = \bigcup_{w\in N^-(v)} \widehat{\mathcal{P}}^i_{uw} \bullet \{v\}$, ισχύει

$$\begin{aligned}|\mathcal{N}^{i+1}_{uv}| &\le s_{i,\ell-i}\cdot \deg^-(v)\\ &\le x_i^{-i}(1-x_i)^{\ell-i}\cdot 2^{o(\ell)}\deg^-(v)\end{aligned}$$

Από το Θεώρημα 3.4, ο χρόνος υπολογισμού της $\widehat{\mathcal{N}}^{i+1}_{uv}$ είναι ίσος με

$$s_{i,\ell-i}\cdot(1-x_{i+1})^{\ell-(i+1)}\cdot 2^{o(\ell)}\cdot\deg^-(v)\cdot\log n. \tag{4.1}$$

Επιπλέον, για κάθε $3 < i < p$, ισχύει

$$s_{i,\ell-i} \le e^2\cdot(i+1)\cdot s_{i+1,\ell-(i+1)}. \tag{4.2}$$

Πράγματι, από τους ορισμούς των $s_{i,\ell-i}$ και x_{i+1} έχουμε

$$\begin{aligned}\frac{s_{i,\ell-i}}{s_{i+1,\ell-(i+1)}} &= \frac{x_i^{-i}(1-x_i)^{-\ell+i}}{x_{i+1}^{-(i+1)}(1-x_{i+1})^{-\ell+(i+1)}}\\ &= \left(\frac{2\ell-i}{2\ell-(i+1)}\right)^{\ell}\cdot\frac{(i+1)^{i+1}}{i^i}\cdot\frac{\big(2\ell-2(i+1)\big)^{\ell-(i+1)}}{(2\ell-2i)^{\ell-i}}\\ &\le \left(1+\frac{1}{2\ell-(i+1)}\right)^{2\ell-(i+1)}\cdot(i+1)\cdot\left(1+\frac{1}{i}\right)^{i}\\ &\le e^2\cdot(i+1)\end{aligned}$$

Από τις σχέσεις (4.1) και (4.2), προκύπτει ότι η οικογένεια $\widehat{\mathcal{P}}^p_{uv}$ υπολογίζεται σε χρόνο ίσο με

$$\begin{aligned}&\mathcal{O}\left(\sum_{i=2}^{p}\sum_{j=1}^{n-1} s_{i,\ell-i}\cdot\deg^-(v_j)\cdot(1-x_i)^{-\ell+i}\cdot 2^{o(\ell)}\cdot\log n\right)\\ =&\mathcal{O}\left(2^{o(\ell)}\cdot m\log n\cdot\max_{i\in[p]}\left\{\left(\frac{2\ell-i}{i}\right)^{i}\left(\frac{2\ell-i}{2\ell-2i}\right)^{2\ell-2i}\right\}\right)\end{aligned}$$

και το μέγεθος της είναι ίσο με

$$s_{p,\ell-p} = (x_p)^{-p}(1-x_p)^{-\ell+p}\cdot 2^{o(\ell)} = \left(\frac{2\ell-p}{p}\right)^{p}\left(\frac{2\ell-p}{2\ell-2p}\right)^{\ell-p}\cdot 2^{o(\ell)}.$$

□

Τώρα, έχουμε στη διάθεση μας έναν πιο γρήγορο αλγόριθμο υπολογισμού της αντιπροσωπευτικής οικογένειας $\widehat{\mathcal{P}}^k_{uv} \subseteq^k_{rep} \mathcal{P}^k_{uv}$. Χρησιμοποιώντας, το Λήμμα 4.4 για $\ell = 2k$ και $p = k$, υπολογίζουμε $\widehat{\mathcal{P}}^k_{uv}$ για κάθε $v \in V(D) \setminus \{u\}$ σε χρόνο

$$\mathcal{O}\left(2^{o(k)} m \log n \cdot \max_{i \in [p]} \left\{ \left(\frac{4k-i}{i}\right)^i \left(\frac{4k-i}{4k-2i}\right)^{4k-2i} \right\}\right)$$

Για $i = k$, προκύπτει η μέγιστη τιμή της συνάρτησης $f(k) = \left(\frac{4k-i}{i}\right)^i \left(\frac{4k-i}{4k-2i}\right)^{4k-2i}$. Επομένως, ο χρόνος υπολογισμού της $\widehat{\mathcal{P}}^k_{uv}$ για κάθε u, $v \in V(D)$ είναι το πολύ ίσος με $\mathcal{O}(6.75^{k+o(k)} nm \log n)$. Από το Θεώρημα 3.4, το μέγεθος της $\widehat{\mathcal{P}}^k_{uv}$ για κάθε u, $v \in V(D)$, είναι το πολύ ίσο με $\mathcal{O}(4.5^{k+o(k)})$. Τελικά, ακολουθώντας την ίδια διαδικασία με αυτήν του Θεωρήματος 4.4, προκύπτει το ακόλουθο θεώρημα.

Θεώρημα 4.5. *Υπάρχει αλγόριθμος που επιλύει το πρόβλημα ΜΑΚΡΥΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΣ ΚΥΚΛΟΣ σε $\mathcal{O}(6.75^{k+o(k)} mn^2)$ χρόνο.*

4.3 k-ΜΟΝΟΠΑΤΙ

Σε αυτή την παράγραφο, μελετάμε το ακόλουθο πρόβλημα:

k-ΜΟΝΟΠΑΤΙ

Είσοδος: μη-κατευθυνόμενο γράφημα G με n κορυφές και m ακμές και θετικός ακέραιος k

Ερώτηση: Υπάρχει μονοπάτι μήκους k στο G;

Αρχικά, προσθέτουμε στο γράφημα G μια καινούρια κορυφή s, την οποία συνδέουμε με κάθε κορυφή στο $V(G)$. Έστω G' το καινούριο γράφημα. Προφανώς, το G έχει μονοπάτι μήκους k αν και μόνο αν το G' έχει μονοπάτι μήκους $k+1$ με αφετηρία την κορυφή s. Θεωρούμε το ομοιόμορφο μητροειδές $U_{n,k+1} = (E, \mathcal{I})$, με $E = V(G)$ και $\mathcal{I} = \{S \subseteq V(G) \mid |S| \leq k+1\}$. Για κορυφές s, $v \in V(G')$, υπενθυμίζουμε ότι έχουμε ορίσει την οικογένεια

$$\mathcal{P}^i_{sv} = \big\{X \subseteq V(G') \mid s, v \in X, |X| = i \text{ και υπάρχει } sv\text{-μονοπάτι στο } G' \\ \text{το οποίο επισκέπτεται όλες τις κορυφές του } X\big\}.$$

Οπότε για την επίλυση του προβλήματος, αρκεί να βρεθεί κορυφή $v \in V(G)$ τέτοια ώστε η οικογένεια $\mathcal{P}^{k+1}_{sv}$ να είναι μη-κενή.

Ο αλγόριθμος ελέγχει αν η $\mathcal{P}^{k+1}_{sv}$ είναι μη-κενή, υπολογίζοντας $\widehat{\mathcal{P}}^{k+1}_{sv} \subseteq^0_{rep} \mathcal{P}^{k+1}_{sv}$ και ελέγχοντας αν η $\widehat{\mathcal{P}}^{k+1}_{sv}$ είναι μη-κενή. Πράγματι, έστω ότι η $\mathcal{P}^{k+1}_{sv}$ είναι μη-κενή, τότε υπάρχει σύνολο $A \in \mathcal{P}^{k+1}_{sv}$, για το οποίο (προφανώς) $A \cap \emptyset = \emptyset$. Επειδή $\widehat{\mathcal{P}}^{k+1}_{sv} \subseteq^0_{rep} \mathcal{P}^{k+1}_{sv}$, υπάρχει σύνολο $A' \in \widehat{\mathcal{P}}^{k+1}_{sv}$ για το οποίο $A \cap \emptyset = \emptyset$. Έπεται ότι η οικογένεια $\widehat{\mathcal{P}}^{k+1}_{sv}$ είναι μη-κενή.

Χρησιμοποιώντας το Λήμμα 4.4 για $\ell = p = k+1$, ο αλγόριθμος υπολογίζει $\widehat{\mathcal{P}}_{sv}^{k+1} \subseteq_{rep}^{0} \mathcal{P}_{sv}^{k+1}\ \forall v \in V(G)$ σε

$$2^{o(k)} \cdot m \log n \cdot \max_{i \in [k+1]} \left\{ \left(\frac{2(k+1)-i}{i} \right)^{i} \left(\frac{2(k+1)-i}{2(k+1)-2i} \right)^{2(k+1)-2i} \right\}$$

χρόνο. Ο χρόνος αυτός μεγιστοποιείται για $i = \left(1 - \frac{1}{\sqrt{5}}\right)(k+1)$. Συνολικά, ο αλγόριθμος για να υπολογίσει $\widehat{\mathcal{P}}_{sv}^{k+1} \subseteq_{rep}^{0} \mathcal{P}_{sv}^{k+1}\ \forall v \in V(G)$, χρειάζεται το πολύ $\phi^{2k+o(k)} m \log^2 n = \mathcal{O}(2.619^k m \log n)$ χρόνο, όπου ϕ είναι η χρυσή τομή $\phi = \frac{1+\sqrt{5}}{2}$. Στον ίδιο χρόνο, ο αλγόριθμος υπολογίζει και μια διάταξη των κορυφών κάθε συνόλου της οικογένειας $\widehat{\mathcal{P}}_{sv}^{k+1}$, τέτοια ώστε να αντιστοιχεί σε ένα μονοπάτι του G. Άρα συνολικά, ο αλγόριθμος χρειάζεται $\mathcal{O}(2.619^k m \log n)$ χρόνο, για να βρει αν υπάρχει $v \in V(G)$ με $\widehat{\mathcal{P}}_{sv}^{k+1} \neq \emptyset$, δηλαδή αν στο γράφημα G' υπάρχει μονοπάτι μήκους $k+1$ με αφετηρία την κορυφή s.

Με ανάλογο τρόπο, μπορούμε να βρούμε αλγόριθμο ίδιου χρόνου στην περίπτωση που η είσοδος μας είναι κατευθυνόμενο γράφημα. Ωστόσο, όταν το γράφημα είναι μη-κατευθυνόμενο, μπορούμε να βελτιώσουμε τον αλγόριθμο, χρησιμοποιώντας το παρακάτω αποτέλεσμα.

Πρόταση 4.1. *[2] Υπάρχει αλγόριθμος, ο οποίος για γράφημα G και ακέραιο k, είτε υπολογίζει ένα μονοπάτι μήκους τουλάχιστον k είτε χρησιμοποιεί τον αλγόριθμο DFS και υπολογίζει δένδρο με ρίζα κάποια κορυφή του G και βάθος το πολύ k. Ο αλγόριθμος αυτός χρειάζεται $\mathcal{O}(k^2 n)$ χρόνο.*

Χρησιμοποιούμε τον αλγόριθμο της Πρότασης 4.1 και σε χρόνο $\mathcal{O}(k^2 n)$ υπολογίζουμε είτε ένα μονοπάτι μήκους τουλάχιστον k είτε ένα δένδρο με βάθος το πολύ k. Στην 1η περίπτωση, έχουμε βρει μονοπάτι μήκους k. Στη 2η περίπτωση, επειδή στο δένδρο του DFS όλα τα μονοπάτια από τη ρίζα στα φύλλα έχουν μήκος το πολύ k και δεν υπάρχουν ακμές διασταύρωσης (cross edges), προκύπτει ότι το δένδρο έχει το πολύ $\mathcal{O}(k^2 n)$ ακμές. Στη συνέχεια, εφαρμόζουμε στο δένδρο τον αλγόριθμο για τα σύνολα αντιπροσώπευσης που περιγράψαμε παραπάνω. Τελικά, προκύπτει το παρακάτω αποτέλεσμα.

Θεώρημα 4.6. *Υπάρχει αλγόριθμος που επιλύει το πρόβλημα k-ΜΟΝΟΠΑΤΙ σε $\mathcal{O}(2.619^k n \log n)$ χρόνο.*

Επίλογος

Σε αυτήν την εργασία παρουσιάσαμε έναν αλγόριθμο για τον υπολογισμό αντιπροσωπευτικών οικογενειών σε γραμμικά μητροειδή και αποδείξαμε ότι στην περίπτωση των ομοιόμορφων μητροειδών, μπορεί να βρεθεί ακόμη πιο γρήγορος αλγόριθμος. Στην συνέχεια δείξαμε πως ο αποδοτικός υπολογισμός αντιπροσωπευτικών οικογενειών, μπορεί να χρησιμοποιηθεί για τον σχεδιασμό παραμετρικών αλγορίθμων, τόσο ντετερμινιστικών όσο και πιθανοτικών.

Οι αλγόριθμοι και τα αποτελέσματα σε σχέση με τον υπολογισμό αντιπροσωπευτικών οικογενειών έχουν χρησιμοποιηθεί σε διάφορες εφαρμογές, τα τελευταία χρόνια. Κατ΄ αρχάς, παραπέμπουμε τον αναγνώστη στα [21], [6], [19], [23], [34], [31] και [32], όπου δίνονται ντετερμινιστικοί παραμετρικοί αλγόριθμοι για διάφορα προβλήματα. Από την άλλη, στο [4] έχουν δοθεί γραμμικού χρόνου κατασκευές στα ακόλουθα προβλήματα: perfect hash family, cover-free family και separating hash family. Οι Lokshtanov, Misra, Panolan και Saurabh στο [20] έδωσαν ντετερμινιστικό αλγόριθμο για τον υπολογισμό του πίνακα αναπαράστασης του t-περιορισμού ενός γραμμικού μητροειδούς. Χρησιμοποιώντας τον αλγόριθμο αυτό, προκύπτει ντετερμινιστική εκδοχή του Θεωρήματος 3.2 και αποτυχαιοποίηση διαφόρων γνωστών παραμετρικών αλγορίθμων. Τέλος, ο Zehavi στο [33] δίνει πιο γρήγορο αλγόριθμο για το πρόβλημα k-*ΜΟΝΟΠΑΤΙ*, ο οποίος τρέχει σε $\mathcal{O}(2.597^k n^{\mathcal{O}(1)})$.

Κατά την άποψή μας, στα παραπάνω διαφαίνονται δύο πράγματα. Από την μία πλευρά, η αξία των μητροειδών στον σχεδιασμό αλγορίθμων, η οποία προκύπτει από την αλγεβρική δομή τους, δίνοντας έτσι άλλο ένα παράδειγμα όπου μία κλασσική περιοχή των μαθηματικών μπορεί να δώσει ενδιαφέροντα αποτελέσματα στην υπολογιστική πολυπλοκότητα. Από την άλλη, εστιάζοντας στην παραμετρική πολυπλοκότητα, βλέπουμε πως η χρήση μίας απλής σχετικά έννοιας, αυτής των αντιπροσωπευτικών οικογενειών, μπορεί να χρησιμοποιηθεί για την βελτίωση της απόδοσης αλγορίθμων, τόσο ως προς το πλήθος βημάτων που θα χρειαστούν, όσο και ως προς την παραμετρική τους εξάρτηση.

Κλείνουμε με δύο παρατηρήσεις, σε σχέση με κάποια δουλειά που θα μπορούσε να ακολουθήσει. Καταρχάς, σε πολλά προβλήματα που μελετούμε, εμφανίζονται (ή και «κρύβονται») γραφικά και ομοιόμορφα μητροειδή. Ως εκ τούτου, η βελτίωση του χρόνου υπολογισμού αντιπροσωπευτικών οικογενειών για τις συγκεκριμένες κλάσεις μητροειδών θα είχε άμεσες συνέπειες στην βελτίωση του χρόνου επίλυσης μιας ευρείας κλάσης προβλημάτων. Τέλος, όλες οι εφαρμογές που έχουμε δει μέχρις στιγμής αφορούν ομοιόμορφα μητροειδή, γραμμικά μητροειδή και μητροειδή διαμέρισης. Πιστεύουμε πως θα ήταν και ενδιαφέρουσα και χρήσιμη η αναζήτηση εφαρμογών που θα βασίζονται και σε άλλους τύπους (κλάσεις) μητροειδών.

Βιβλιογραφία

[1] G. Birkhoff. Abstract linear dependence and lattices. *American Journal of Mathematics*, 57:800–804, 1935.

[2] H. L. Bodlaender. On linear time minor tests with depth-first search. *J. Algorithms*, 14:1–23, 1980.

[3] B. Bollobás. On generalized graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 16:447–452, 1965.

[4] N. H. Bshouty. Linear time constructions of some d -restriction problems. In *Algorithms and Complexity - 9th International Conference, CIAC 2015, Paris, France, May 20-22, 2015. Proceedings*, pages 74–88, 2015.

[5] R. G. Downey, M. R. Fellows. *Fundamentals of Parameterized Complexity*. Springer-Verlag, 1999.

[6] F. V. Fomin, P. A. Golovach. Long circuits and large euler subgraphs. *SIAM J. Discrete Math.*, 28(2):878–892, 2014.

[7] J. Bunch, J. Hopcroft. Triangular factorization and inversion by fast matrix multiplication. *Mathematics of Computation*, pages 231–236, 1974.

[8] A. Björklund, T. Husfeldt, S. Khanna. Approximating longest directed paths and cycles. In *Automata, Languages and Programming: 31st International Colloquium, ICALP 2004, Turku, Finland, July 12-16, 2004. Proceedings*, pages 222–233, 2004.

[9] R. A. DeMillo, R. J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7:193–195, 1978.

[10] L. Lovász. Flats in matroids and geometric graphs. *In Combinatorial Surveys (Proc. 6th British Combinatorial Conf.), Academic Press*, pages 45–86, 1977.

[11] S. MacLane. Some interpretations of abstract linear dependence in terms of projective geometry. *American Journal of Mathematics*, 58:236–240, 1936.

[12] D. Marx. Parameterized coloring problems on chordal graphs. In *Parameterized and Exact Computation, First International Workshop, IWPEC 2004, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 83–95, 2004.

[13] D. Marx. A parameterized view on matroid optimization problems. *Theor. Comput. Sci.*, 410(44):4471–4479, 2009.

[14] B. Monien. How to find long paths efficiently. *Analysis and Design of Algorithms for Combinatorial Problems*, 109:239–254, 1985.

[15] K. Murota. *Matrices and matroids for systems analysis*. Springer, 2000.

[16] L. Bodlaender, M. Cygan, S. Kratsch, J. Nederlof. Solving weighted and counting variants of connectivity problems parameterized by treewidth deterministically in single exponential time. *CoRR*, abs/1211.1505, 2012.

[17] H. N. Gabow, S. Nie. Finding a long directed cycle. *ACM Trans. Algorithms*, 4(1), 2008.

[18] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford, 2006.

[19] P. Goyal, N. Misra, F. Panolan. Faster deterministic algorithms for r-dimensional matching using representative sets. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2013, December 12-14, 2013, Guwahati, India*, pages 237–248, 2013.

[20] D. Lokshtanov, P. Misra, F. Panolan, S. Saurabh. Deterministic truncation of linear matroids. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, pages 922–934, 2015.

[21] F. V. Fomin, D. Lokshtanov, S. Saurabh. Efficient computation of representatives sets with applications in parameterized and exact algorithms. *CoRR*, abs/1304.4626, 2013.

[22] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, S. Saurabh. *Parameterized Algorithms*. Springer, 2015.

[23] P. Goyal, P. Misra, F. Panolan, G. Philip, S. Saurabh. Finding even subgraphs even faster. *CoRR*, abs/1409.4935, 2014.

[24] A. Schrijver. *Combinatorial Optimization. Polyhedra and Efficiency*, volume A. Springer, 2003.

[25] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27:701–717, 1980.

[26] M. Mitzenmacher, E. Upfal. *Probability and computing: Randomized algorithms and probabilistic analysis*. Cambridge University Press, 2005.

[27] B. L. van der Waerden. *Moderne Algebra*. Springer, 1937.

[28] H. Whitney. On the abstract properties of linear dependence. *American Journal of Mathematics*, 57:509–533, 1935.

[29] V. V. Williams. Multiplying matrices faster than coppersmith-winograd. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 887–898, 2012.

[30] R. Impagliazzo, R. Paturi, F. Zane. Which problems have strongly exponential complexity. *J. Comput. Syst. Sci.*, 63(4):512–530, 2001.

[31] H. Shachnai, M. Zehavi. Parameterized algorithms for graph partitioning problems. *CoRR*, abs/1403.0099, 2014.

[32] H. Shachnai, M. Zehavi. Representative families: A unified tradeoff-based approach. *J. Comput. Syst. Sci.*, 82(3):488–502, 2016.

[33] M. Zehavi. Mixing color coding-related techniques. In *Algorithms - ESA 2015 - 23rd Annual European Symposium, Patras, Greece, September 14-16, 2015, Proceedings*, pages 1037–1049, 2015.

[34] R. Y. Pinter, H. Shachnai, M. Zehavi. Deterministic parameterized algorithms for the graph motif problem. In *Mathematical Foundations of Computer Science 2014 - 39th International Symposium, MFCS 2014, Budapest, Hungary, August 25-29, 2014. Proceedings, Part II*, pages 589–600, 2014.

[35] J. Chen, J. Kneis, S. Lu, D. Molle, S. Richter, P. Rossmanith, S.-H. Sze, F. Zhang. Randomized divide-and-conquer: improved path, matching, and packing algorithms. *SIAM J. Comput.*, 38:2526–2547, 2009.

[36] R. Zippel. Probabilistic algorithms for sparse polynomials. *International Symposium on Symbolic and Algebraic Computation*, pages 216–226, 1979.

[37] N. Alon, R. Yuster, U. Zwick. Color-coding. *J. ACM*, 42:844–856, 1995.