Αλγόριθμοι Δικτύων και Πολυπλοκότητα

Εισαγωγή σε βασικές έννοιες αλγορίθμων και πολυπλοκότητας και γραφοθεωρητικών προβλημάτων

Άρης Παγουρτζής

ΕΜΠ — ΑΛΜΑ

Ευχαριστίες: μέρος των διαφανειών αυτών προέρχεται από σημειώσεις του Ε. Ζάχου για διάφορα μαθήματα της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών Ε.Μ.Π.

Επισκόπηση μαθήματος

- Υπολογιστική πολυπλοκότητα και προσεγγισιμότητα γραφοθεωρητικών προβλημάτων με εφαρμογές σε δίκτυα: Matching, Shortest Paths, Vertex Cover, Traveling Salesman Problem, Steiner tree, Maximum Flow, Maximum Edge-Disjoint Paths, Multicommodity Flow, Facility Location, Multicut, k-Center, Scheduling, Clustering.
- Κατανεμημένα πρωτόκολλα: εκλογή αρχηγού, broadcasting, gossiping, secret sharing. Ασύρματα ad hoc δίκτυα. Συγχρονισμένοι και ασύγχρονοι αλγόριθμοι. Fault tolerance, Byzantine adversaries.
- Προβλήματα αυτόνομων οντοτήτων, εξερεύνηση δικτύων, προβλήματα συνάντησης (rendezvous), εντοπισμός βλαβών σε δίκτυα. Πρωτόκολλα δρομολόγησης, compact routing, geometric routing.

Επισκόπηση μαθήματος (συν.)

- Ειδικά θέματα: χρονοδρομολόγηση (scheduling), δρομολόγηση και ανάθεση συχνοτήτων σε οπτικά δίκτυα, αλγόριθμοι πλοήγησης, προγραμματισμός δρομολογίων οχημάτων.
- Στοιχεία θεωρίας παιγνίων: σημεία ισορροπίας Nash, "κόστος της αναρχίας". Εγωιστική δρομολόγηση σε κλασικά, ασύρματα και οπτικά δίκτυα. Παίγνια συμφόρησης. Σύγκλιση σε ισορροπίες Nash και σχεδίαση μηχανισμών.

Βιβλιογραφία: αλγόριθμοι

1. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest: **Introduction to algorithms**. The MIT Press/ McGraw-Hill Book Company, 1989. (or 2nd edition.)
2. S. Dasgupta, C. H. Papadimitriou, and U. V. Vazirani: **Algorithms**. MacGraw-Hill, 2006. (Διατίθεται ελεύθερα στο διαδίκτυο).
3. David P. Williamson, David B. Shmoys. **The Design of Approximation Algorithms**. Cambridge University Press, 2010. (available free online)
4. Vijay V. Vazirani. **Approximation Algorithms**. Springer, 2001.
5. Dorit S. Hochbaum. **Approximation Algorithms for NP-Hard Problems**. PWS Publishing Company, 1997.

Βιβλιογραφία: δίκτυα και κατανεμημένοι υπολογισμοί

1. Roger Wattenhofer. **Principles of Distributed Computing**. ETH Zuerich course notes, 2011.
2. Nancy A. Lynch. **Distributed Algorithms**. Morgan Kaufmann Publishers, San Mateo, CA, 1996.
3. Robert E. Tarjan. **Data Structures and Network Algorithms**. SIAM, 1983.
4. Ravindra K. Ahuja, Thomas L. Magnanti, James B. Orlin. **Network flows: Theory, Algorithms, and Applications**. Prentice-Hall, 1993.
5. Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. **Algorithmic Game Theory**. Cambridge University Press, New York, NY, USA.

Εισαγωγή

- Προβλήματα, Αλγόριθμοι, Πολυπλοκότητα
- Γράφοι
- Ευεπίλυτα προβλήματα γράφων (στην κλάση **P**): κύκλος Euler, διάσχιση γράφων / προσβασιμότητα (Reachability), συντομότερα μονοπάτια (Shortest Paths), ελάχιστο συνδετικό δένδρο (Minimum Spanning Tree), μέγιστη ροή (Max Flow), ταίριασμα (Matching), ευσταθές ταίριασμα (Stable Matching), χρωματισμός ακμών σε διμερείς γράφους (Bipartite Edge Coloring.
- **NP**-πλήρη προβλήματα γράφων: Vertex Cover, Clique, Hamilton Circuit/Cycle (HC), Traveling Salesman Problem (TSP), 3-Colorability, Subgraph Isomorphism, 3-Dimensional Matching (3DM).
- Graph Isomorphism?

Υπολογιστικά Προβλήματα

Υπολογιστικό πρόβλημα: καθορισμός αντιστοίχισης έγκυρων *δεδομένων εισόδου (στιγμιοτύπου)* σε *δεδομένα εξόδου (απαντήσεις / λύσεις)*.

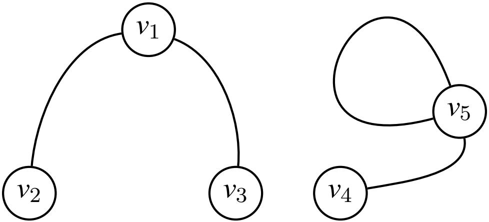
Μαθηματική περιγραφή: σχέση (relation) μεταξύ συμβολοσειρών.

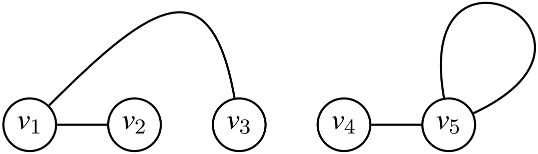
Παράδειγμα. Το πρόβλημα Satisfiability (SAT)

Προβλήματα απόφασης, προβλήματα βελτιστοποίησης.

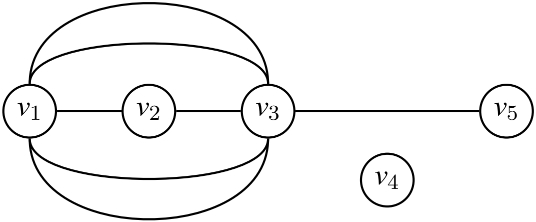
Γράφοι (ή γραφήματα)

$$E = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_4, v_5\}, \{v_5, v_5\}\}$$





Πολυγράφημα



Ορολογία γραφημάτων

Συνδεδεμένες κορυφές, παραγόμενος (induced) υπογράφος,

Συνεκτικότητα (connectivity), συνεκτικές συνιστώσες (connected components).

Κατευθυνόμενος γράφος: ισχυρή και ασθενής συνεκτικότητα.

Δένδρο (tree): συνεκτικός ακυκλικός γράφος.

Πλήρης γράφος (K_n), διμερής γράφος (πλήρης διμερής: $K_{n,m}$).

Επίπεδος γράφος (ανν δεν περιέχει K_5, $K_{3,3}$ ως ελάσσονα γραφήματα)

Χαρακτηρισμοί με αποκλεισμό ελασσόνων (minors)

Έλασσον υπογράφημα (minor): κάθε γράφημα που προκύπτει από το αρχικό με διαγραφή κόμβου, διαγραφή ακμής, ή σύνθλιψη ακμής.

Σύνολο παρεμπόδισης ή αποκλεισμού: ένα σύνολο ελασσόνων υπογραφημάτων που ο αποκλεισμός τους χαρακτηρίζει ακριβώς μια οικογένεια γραφημάτων.

$\{K_5, K_{3,3}\}$: σύνολο αποκλεισμού για τους επίπεδους γράφους.

Θεώρημα Robertson-Seymour: κάθε οικογένεια γραφημάτων που είναι κλειστή ως προς minors χαρακτηρίζεται από πεπερασμένο σύνολο αποκλεισμού.

Πόρισμα: κάθε ιδιότητα που είναι κλειστή ως προς minors μπορεί να ελεγχθεί σε πολυωνυμικό χρόνο (αλλά συνήθως δεν ξέρουμε πώς)!

Παράσταση Γράφου

Πίνακας γειτνίασης (adjacency matrix)

Πίνακας πρόσπτωσης (incidence matrix)

Λίστες γειτνίασης (adjacency lists): αποδοτική παράσταση σε αραιούς γράφους.

Ένας ενδιαφέρων πίνακας

- Laplacian matrix: $Q(G) = D(G) - A(G)$
- $Q(G) = E'(G) \cdot E'^T(G)$ (E': προσανατολισμένος πίνακας πρόσπτωσης)
- Χρησιμεύει, μεταξύ άλλων, στον υπολογισμό του πλήθους των spanning trees (Kirchhoff's matrix tree theorem):

$$t(G) = \frac{1}{n}\lambda_1\lambda_2 \ldots \lambda_{n-1} = \det(Q_{11}(G))$$

- Επίσης, το πλήθος των συνεκτικών συνιστωσών ενός γράφου ισούται με το πλήθος των μηδενικών ιδιοτιμών του $Q(G)$.

Περαιτέρω εφαρμογές: spectral graph theory (βλ. και Unique Games Conjecture).

Unique Games Conjecture [Khot'02]

- Έστω γράφος με περιορισμούς στις ακμές: π.χ. χρωματικούς.
- Κάθε ανάθεση τιμής σε κόμβο καθορίζει μοναδικά τιμές γειτόνων.
- Το πρόβλημα ικανοποίησης όλων των περιορισμών είναι στο **P**.
- Το πρόβλημα μεγιστοποίησης δεν είναι!
- UCG: Για κάθε $\delta > 0$ και $\varepsilon > 0$ υπάρχει Unique Game όπου είναι **NP**-hard να διακρίνει κανείς αν $OPT > 1 - \varepsilon$ ή αν $OPT < \delta$.
- Ακόμη και αν δίνεται στιγμιότυπο με 99% των περιορισμών να μπορούν να ικανοποιηθούν, είναι **NP**-hard να ικανοποιηθεί έστω και το 1%.
- Πλούσια έρευνα: expander graphs, spectral graph theory.

Κλάσεις Πολυπλοκότητας

P: προβλήματα απόφασης επιλύσιμα σε πολυωνυμικό χρόνο από κάποιον ντετερμινιστικό αλγόριθμο.

NP: προβλήματα απόφασης επιλύσιμα σε πολυωνυμικό χρόνο από κάποιον μη ντετερμινιστικό αλγόριθμο. Πιθανές λύσεις (πιστοποιητικά, αποδείξεις, μάρτυρες) ελέγξιμες σε πολυωνυμικό χρόνο.

Το μεγάλο ανοιχτό ερώτημα: $\mathbf{P} \stackrel{?}{=} \mathbf{NP}$

NP-completeness, αναγωγές.

NP-πλήρη προβλήματα γράφων

Vertex Cover

Clique

Hamilton Circuit/Cycle (HC)

Traveling Salesman Problem (TSP)

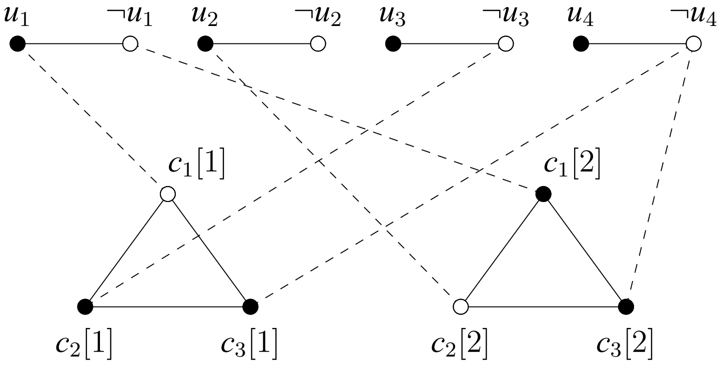
3-Colorability

Subgraph Isomorphism

3-Dimensional Matching (3DM)

Αναγωγή 3SAT $\leq$ VERTEX COVER

$$: (u_1 \vee \neg u_3 \vee \neg u_4) \wedge (\neg u_1 \vee u_2 \vee \neg u_4)$$



Η είναι ικανοποιήσιμη ανν υπάρχει vertex cover μεγέθους $\leq k = n + 2m = 8$ στον γράφο που κατασκευάσαμε.

Αλλες Αναγωγές

$$
\begin{tikzcd}
 & & & & & \text{HC} \arrow[r] & \text{TSP} \\
 & & & & \text{VC} \arrow[ur] \arrow[r] & \text{Clique} \arrow[r] & \text{Subgraph Isomorphism} \\
\text{Any NP problem} \arrow[r, "Cook"] & \text{Sat} \arrow[r] & \text{3Sat} \arrow[urr] \arrow[r] \arrow[ddrr] & \text{3DM} \arrow[dr] & & & \\
 & & & & \text{Partition} \arrow[r] & \text{DKnapsack} & \\
 & & & & \text{3-Graph Colorability} & &
\end{tikzcd}
$$

Προβλήματα γράφων στην κλάση **P**

Κύκλος Euler.

Reachability - Διάσχιση Γράφων: DFS, BFS, D-Search.

Συντομότερα μονοπάτια. Συνεκτικές συνιστώσες.

Ελάχιστο συνδετικό δένδρο (minimum spanning tree).

Μέγιστη ροή. Perfect matching. Stable matching.

Χρωματισμός ακμών σε διμερή γράφο (bipartite edge coloring).

Accessibility problems - Διάσχιση γράφων

Αναζήτηση κατά βάθος (Depth First Search - DFS).

Αναζήτηση κατά πλάτος (Breadth First Search - BFS).

D-search: όμοιο με BFS, αλλά με στοίβα αντί για ουρά.

Διάσχιση γράφων: DFS

procedure dfs(v:vertex);

begin

visited[v] $\leftarrow$ `True`

for all vertices u adjacent to v do

if not visited[u] then dfs(u)

end

Πολυπλοκότητα: $O(|V| + |E|)$.

Διάσχιση γράφων: BFS

```
procedure bfs(v:vertex);
begin
   initialize queue with v; visited[v]:=true
   repeat
      dequeue(u)
      for all vertices w adjacent to u do
         if not visited[w] then
            begin visited[w] ← true; enqueue(w) end
   until queue is empty
end
```

Πολυπλοκότητα: $O(|V| + |E|)$.

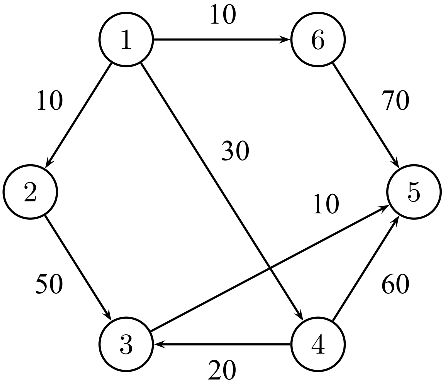
Συντομότερα μονοπάτια: Dijkstra

procedure Dijkstra;

begin (* Αρχικοποίηση *)
    $S \leftarrow \{1\}$
    for $i \leftarrow 2$ to n do
        $D[i] \leftarrow C[1, i]$; $P[i] \leftarrow 1$
    for $i \leftarrow 2$ to n-1 do
  begin
      select w from $V - S$ such that $D[w]$ is minimum;
      $S := S \cup \{w\}$;
      for all v in $V - S$ do
          if $D[v] > D[w] + C[w, v]$ then
              $P[v] \leftarrow w$;
              $D[v] \leftarrow D[w] + C[w, v]$ ▷ update
  end
end

Πολυπλοκότητα: $O(|V|^2)$
All-pairs shortest paths: $O(|V|^3)$

Αλγόριθμος Dijkstra: παράδειγμα



Βήμα	S	w	D					P				
			2	3	4	5	6	2	3	4	5	6
-	{1}	-	10	∞	30	∞	10	1	1	1	1	1
2	{1,2}	2		60	30	∞	10		2			
3	{1,2,6}	6		60	30	80					6	

Αλγόριθμος Bellman-Ford

Εκτελείται σε $|V| - 1$ στάδια.

Στο στάδιο i ενημερώνεται κάθε κόμβος v (που βρίσκεται σε απόσταση το πολύ i ακμών από τον αρχικό) με το συντομότερο μονοπάτι από τον s στον v που έχει το πολύ i ακμές.

Αυτό επιτυγχάνεται με εκτέλεση για κάθε ακμή $(w, v) \in E$ της εντολής:

if $D[v] > D[w] + C[w, v]$ then
  $P[v] \leftarrow w$
  $D[v] \leftarrow D[w] + C[w, v]$

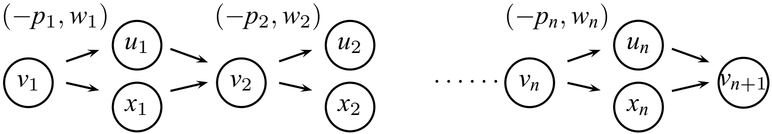
Πολυπλοκότητα: $(|V||E|)$

Παρατήρηση: δεν δουλεύει αν υπάρχουν κύκλοι αρνητικού βάρους. Μπορεί όμως να τους εντοπίζει με κατάλληλη τροποποίηση (*Ασκηση: βρείτε πώς*).

Restricted (Constrained) Shortest Path: **NP**-hard

Το πρόβλημα ορίζεται σε γράφους με δύο συναρτήσεις κόστους στις ακμές (π.χ. κόστος σε χρήμα c_{ij} και σε χρόνο t_{ij}). Ζητείται να ελαχιστοποιηθεί ο συνολικός χρόνος χωρίς να ξοδευτούν πάνω από ένα ποσό K χρημάτων (ή, ισοδύναμα, να ελαχιστοποιηθεί το κόστος χωρίς ο χρόνος να ξεπεράσει κάποιο όριο).

Αναγωγή από το D-Knapsack:



Good news: admits an FPTAS [Warburton, 1987].

Ελάχιστο Συνδετικό Δένδρο (Minimum Spanning Tree - MST)

Αλγόριθμος Prim: Διαλέγουμε κάθε φορά ακμή ελαχίστου κόστους έτσι ώστε ο νέος υπογράφος να παραμένει δέντρο.

Αλγόριθμος Kruskal: Διαλέγουμε κάθε φορά ακμή ελαχίστου κόστους έτσι ώστε ο νέος υπογράφος να μην έχει κύκλους.

Κοινή ιδέα των δύο αλγορίθμων: ξεκινώντας από τον γράφο χωρίς ακμές, και ενώνοντας επαναληπτικά δύο οποιαδήποτε συμπληρωματικά υποσύνολα κόμβων S και $V \setminus S$ που ακόμη δεν έχουν ακμή μεταξύ τους με ελαχίστου βάρους ακμή καταλήγουμε σε ελάχιστο συνδετικό δένδρο.

Ελάχιστο Συνδετικό Δένδρο (συν.)

Γιατί δουλεύει η ιδέα; (ένωση ενός οποιουδήποτε υποσυνόλου κόμβων S με υπόλοιπο γράφο $V \setminus S$ με την ελαφρύτερη ακμή)

Λήμμα. Ένα σύνολο ακμών που είναι υποσχόμενο (υποσύνολο ενός MST) παραμένει υποσχόμενο αν του προσθέσουμε ελάχιστη ακμή μεταξύ *οποιουδήποτε* συνόλου συνεκτικών συνιστωσών του γράφου (που ορίζεται από τις ακμές του συνόλου) και του υπόλοιπου γράφου.

Απόδειξη. Στον πίνακα.

Άλλη εφαρμογή της ιδέας: Αλγόριθμος Borůvka, προσφέρεται για παραλληλοποίηση. Κάθε συνεκτική συνιστώσα (connected component) συνδέεται με την ελαφρύτερη δυνατή ακμή με κάποια από τις υπόλοιπες συνιστώσες. Αρχικά κάθε κόμβος είναι συνιστώσα. Σε κάθε ‘γύρο’ ο αριθμός των συνιστωσών μειώνεται στο μισό. Χρειάζεται διαφορετικά βάρη στις ακμές, ή τρόπο επίλυσης ‘ισοπαλιών’.

Πολυπλοκότητα: $(|E| \log |V|)$.

Αλγόριθμος Prim: υλοποίηση

Κάθε φορά επιλέγεται ο κόμβος με την ελάχιστη απόσταση από το μέχρι στιγμής κατασκευασμένο δένδρο και προστίθεται στο δένδρο.

Πολυπλοκότητα: $O(|V|^2)$ (απλή υλοποίηση), $O(|E| \log |V|)$ με binary heap, $O(|V| \log |V| + |E|)$ με fibonacci heap,

Αλγόριθμος Kruskal: υλοποίηση

Κάθε φορά επιλέγεται ακμή ελαχίστου κόστους και εάν δεν δημιουργεί κύκλο στο μέχρι στιγμής δάσος προστίθεται σε αυτό, αλλιώς απορρίπτεται.

Πολυπλοκότητα: $O(|E| \log |V|)$ (υλοποίηση με Union-Find, Union by Rank)

Μέγιστη ροή (Max Flow)

Δοθέντος γράφου με βάρη που αντιπροσωπεύουν *χωρητικότητες* (network) και δύο κόμβων s, t, ζητείται να δρομολογηθεί όσο το δυνατόν μεγαλύτερη ροή από τον s στον t.

Θεώρημα

(Max Flow – Min Cut) Η μέγιστη ροή ισούται με την ελάχιστη (ως προς χωρητικότητα) τομή (σύνολο ακμών) που διαχωρίζει τον s από τον t.

Αλγόριθμος Ford-Fulkerson

Επιλογή μονοπατιού από τον s στον t. Δρομολόγηση ροής ίσης με την ελάχιστη χωρητικότητα ακμής στο μονοπάτι.

Επανάληψη της διαδικασίας στο υπολειμματικό δίκτυο (residual network) ώσπου να μην υπάρχει πλέον μονοπάτι από τον s στον t.

Ορολογία: Τα μονοπάτια που χρησιμοποιεί ο αλγόριθμος λέγονται συνήθως μονοπάτια επαύξησης (augmenting paths).

Πολυπλοκότητα: $O(|f^*||E|)$, f^* η μέγιστη ροή.

Βελτιώσεις: Αλγόριθμος Edmonds-Karp $O(|V||E|^2)$ (shortest paths), αλγόριθμος Goldberg $O(|V|^2|E|)$ και $O(|V|^3)$ (preflow-push).

Τέλειο ταίριασμα (Perfect Matching)

Σε διμερείς γράφους:

Ανάγεται στο πρόβλημα της μέγιστης ροής.

Μονοπάτια επαύξησης: στην περίπτωση αυτή είναι ουσιαστικά μονοπάτια όπου εναλλάσσονται ακμές εκτός του τρέχοντος matching με ακμές εντός του τρέχοντος matching (alternating paths) και όπου η πρώτη και τελευταία ακμή είναι εκτός matching.

Πολυπλοκότητα: $O(|V||E|)$ (επειδή $|f^*| \leq |V|/2$). Βελτίωση Hopcroft-Karp: $O(|V|^{5/2})$.

Σχετικό πρόβλημα: Stable Marriage.

Πρόβλημα Stable Marriage/Matching

Δίνεται πλήρης διμερής γράφος, όπου οι κόμβοι κάθε συνόλου έχουν σειρά προτίμησης για τους κόμβους του άλλου συνόλου (ή απλά λίστες/πίνακας προτίμησης).

Ζητείται matching M όπου να μην υπάρχει ζευγάρι εκτός ταιριάσματος, όπου και οι δύο κόμβοι να επιθυμούν τον άλλο περισσότερο από το ‘ταίρι’ τους στο M.

Αλγόριθμος Gale-Shapley (1962): οι άντρες προτείνουν, οι γυναίκες αποδέχονται ή όχι (ή και αντίστροφα). Ευνοεί τους προτείνοντες.

Πολυπλοκότητα: $O(n^2)$.

Παραλλαγές (εκτός γάμου): ανάθεση ειδικευόμενων γιατρών σε νοσοκομεία (hospital/residents problem), επιλογή φοιτητών σε σχολές, Stable Roommates. Μερικές παραλλαγές είναι NP-complete.

Χρωματισμός ακμών

Σε γενικούς γράφους: αρκούν $\Delta + 1$ χρώματα (απλοί γράφοι) και $\Delta + \mu$ χρώματα (πολυγραφήματα) [Vizing, 1964].

Το ερώτημα εάν αρκούν Δ χρώματα είναι NP-complete.

Καλύτερος μέχρι στιγμής προσεγγιστικός αλγόριθμος για πολυγραφήματα: $(1 + \frac{3}{\sqrt{2OPT}})$-apx. (APTAS) [Sanders and Steurer, 2005]

Σε διμερείς γράφους (και πολυγραφήματα): αρκούν πάντοτε Δ χρώματα (König, 1916, αλγόριθμος $(|E||V|)$).

Καλύτερος μέχρι στιγμής αλγόριθμος: $(|E| \log \Delta)$ [Cole, Ost, and Schirra, 2001].

Χρωματισμός ακμών σε διμερείς γράφους

Θεώρημα. [König, 1916] Σε διμερείς γράφους (και πολυγραφήματα) αρκούν πάντοτε Δ χρώματα.

Απόδειξη (ιδέα): θεωρώντας maximal matching με Δ χρώματα, αν υποθέσουμε ότι υπάρχει αχρωμάτιστη ακμή μπορούμε να βρούμε augmenting path, φτάνοντας σε αντίφαση.

Από το θεώρημα προκύπτουν 2 αλγόριθμοι:

(α) Επαναληπτική εύρεση και αφαίρεση perfect matching (σε Δ-κανονικό γράφο): $O(\Delta n^{\frac{5}{2}})$.

(β) Επαναληπτικός χρωματισμός ακμών βασισμένος στην απόδειξη: $O(|V||E|) = O(\Delta n^2)$.

Γραφοθεωρητικά προβλήματα (επισκόπηση)

- Ευεπίλυτα προβλήματα γράφων (στην κλάση **P**): κύκλος Euler, διάσχιση γράφων / προσβασιμότητα (Reachability), συντομότερα μονοπάτια (Shortest Paths), ελάχιστο συνδετικό δένδρο (Minimum Spanning Tree), μέγιστη ροή (Max Flow), ταίριασμα (Matching), ευσταθές ταίριασμα (Stable Matching), χρωματισμός ακμών σε διμερείς γράφους (Bipartite Edge Coloring.
- **NP**-πλήρη προβλήματα γράφων: Vertex Cover, Clique, Hamilton Circuit/Cycle (HC), Traveling Salesman Problem (TSP), 3-Colorability, Subgraph Isomorphism, 3-Dimensional Matching (3DM). Restricted Shortest Path.
- Ενδιάμεση πολυπλοκότητα: Graph Isomorphism?

Αντιμετώπιση **NP**-πληρότητας: προσεγγιστικοί αλγόριθμοι, περιορισμός στιγμιοτύπων, παραμετρική πολυπλοκότητα, βελτιωμένοι εκθετικοί αλγόριθμοι.