ΔΥΣΚΟΛΙΑ ΣΤΗΝ ΠΡΟΣΕΓΓΙΣΙΜΟΤΗΤΑ

Επιμέλεια : Γεωργίου Κωστής

Παρουσίαση στα πλαίσια του μαθήματος: “Δίκτυα και πολυπλοκότητα”

Φεβρουάριος 2004

$\mu\Pi\lambda\forall$

Κίνητρα για τη μελέτη της μη προσεγγισιμότητας

- Ο πληρέστερος χαρακτηρισμός της πολυπλοκότητας προβλημάτων

- Η κατανόηση της προσεγγισιμότητας NP-hard προβλημάτων

Υπάρχοντα αποτελέσματα

Παράγοντας δυσκολίας

	Κλάση 1	Κλάση 2	Κλάση 3
Min	Σταθερός (>1)	$\Omega(\log n)$	n^{ε}, για κάποιο ε θετικό
Max	Σταθερός (<1)	$O(1/\log n)$	$1/n^{\varepsilon}$, για κάποιο ε θετικό
	Max 3-SAT vertex cover Steiner tree	set cover	Clique

Η βάση των αποδείξεων και ένα παράδειγμα

- Το PCP (probabilistically checkable proof systems) θεώρημα

- Αποδεικνύεται αναγωγή από το SAT που απεικονίζει στιγμιότυπό του, φ, σε ένα γράφο $G=(E,V)$ τέτοιο που
  - φ ικανοποιήσιμος ➔ ο G έχει vertex cover μεγέθους $\leq \frac{2}{3}|V|$
  - φ μη ικανοποιήσιμος ➔το μικρότερο vertex cover του G έχει μέγεθος $> a \cdot \frac{2}{3}|V|, (a>1)$

Η βάση των αποδείξεων και ένα παράδειγμα (2)

Πόρισμα:

Δεν υπάρχει προσεγγιστικός αλγόριθμος για το vertex cover με λόγο α, εκτός και αν $P = NP$

Απόδειξη

✓ Δυσκολία στη διάκριση γράφων με vertex cover $\leq \frac{2}{3}|V|$, από αυτούς με cover $> a \cdot \frac{2}{3}|V|$

✓ Αν είχαμε προσεγγιστικό αλγόριθμο με λόγο a, θα πετυχαίναμε cover μεγέθους $\leq a \cdot \frac{2}{3}|V|$ για γράφους πρώτης κατηγορίας.

✓ Θα διαχωρίζαμε έτσι τις δύο κλάσεις γράφων ∎

Βασικοί ορισμοί – αναγωγές

- **Ορισμός**: Gap introducing reductions
  - Έστω Π ένα πρόβλημα ελαχιστοποίησης.
  - Υπάρχουν συναρτήσεις $f, a>1$ τέτοιες που αν φ στιγμιότυπο του SAT και $x \in \Pi$ τότε
    - φ ικανοποιήσιμος → $OPT(x) \le f(x)$
    - φ μη ικανοποιήσιμος → $OPT(x) > a(|x|) \cdot f(x)$

➢ Για προβλήματα μεγιστοποίησης υπάρχει αντίστοιχος ορισμός με $\alpha<1$. Οι ανισότητες τότε είναι αντίστροφες

Βασικοί ορισμοί – αναγωγές (2)

- **Ορισμός**: Gap preserving reductions
  - Έστω Π_1, Π_2 προβλήματα ελαχιστοποίησης και μεγιστοποίησης αντίστοιχα.
  - Αν $x \in \Pi_1$, η αναγωγή υπολογίζει στιγμιότυπο $y \in \Pi_2$ τέτοιο ώστε υπάρχουν συναρτήσεις $f_1, a > 1, f_2, b < 1$

$$OPT(x) \le f_1(x) \Rightarrow OPT(y) \le f_2(y)$$
$$OPT(x) > a(|x|) \cdot f_1(x) \Rightarrow OPT(y) > b(|y|) \cdot f_2(y)$$

Η χρήση των δύο αναγωγών

Παρατήρηση: Έστω $\Pi_1 \xrightarrow{\Gamma} \Pi_2$ (Γ gap-preserving αναγωγή),

- Για κατάλληλη gap-introducing αναγωγή Γ' $\mathrm{SAT} \le \Pi_1$
  $\Rightarrow$ δεν υπάρχει $b(|y|)$ προσεγγιστικός αλγόριθμος για το Π_2.

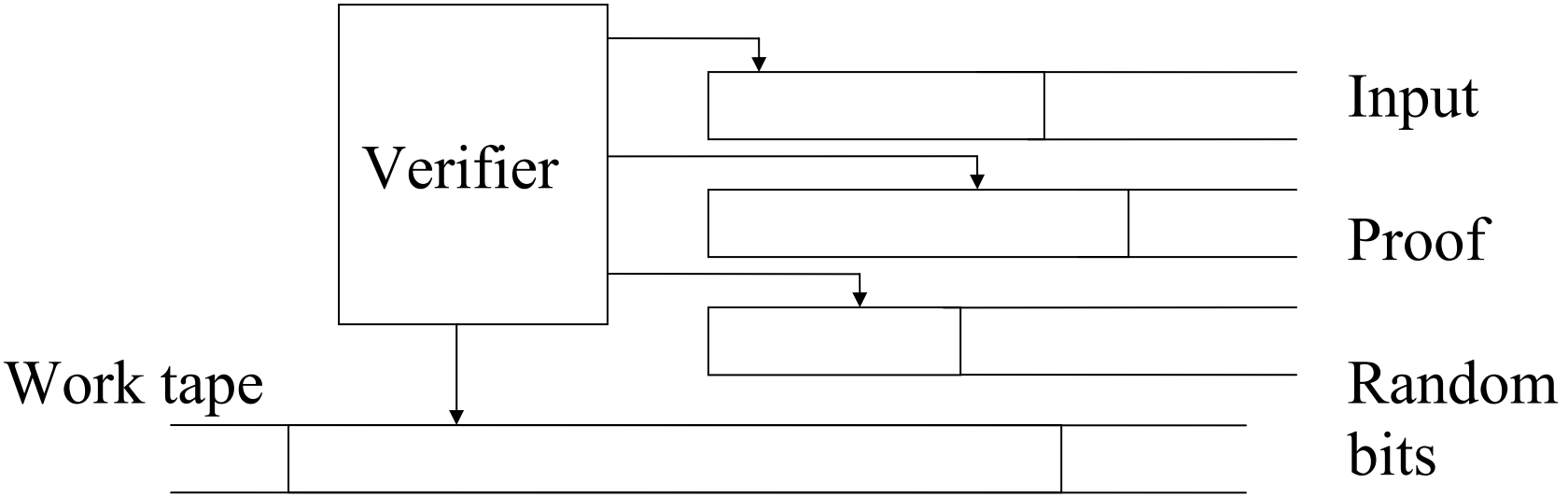
Απόδειξη

✓ Συνθέτουμε τις δύο αναγωγές

Το PCP θεώρημα

- **Ορισμός**: Verifier

  Είναι μια πολυωνυμική μηχανή Turing, που έχει πρόσβαση σε δύο επιπλέον ταινίες, μια με τα τυχαία bits και μια με την απόδειξη.



Το PCP θεώρημα (2)

- **Ορισμός**: $\mathrm{PCP}(f(n), g(n))$

  Η κλάση προβλημάτων που αναγνωρίζονται από verifier που χρησιμοποιεί

  - $O(f(n))$ τυχαία bits
  - $O(g(n))$ bits από την απόδειξη

  $L \in \mathrm{PCP}(f(n), g(n))$:

  - $x \in L$ ➔ υπάρχει απόδειξη που ο V αποδέχεται με πιθανότητα 1
  - $x \notin L$ ➔για κάθε απόδειξη ο V αποδέχεται με πιθανότητα $< \frac{1}{2}$ (γνωστό και ως σφάλμα).

Το PCP θεώρημα (3)

- **Πρόταση:** $\mathrm{NP} = \mathrm{PCP}\left(0, \bigcup_{k \in R^+} \{n^k\}\right)$

- **Θεώρημα:** $\mathrm{NP} = \mathrm{PCP}(\log n, 1)$

Απόδειξη:

✓ $\mathrm{NP} \supseteq \mathrm{PCP}(\log n, 1)$ (απλό)

✓ $\mathrm{NP} \subseteq \mathrm{PCP}(\log n, 1)$: βασίζεται στην κατασκευή verifier για το 3-SAT με σφάλμα $< \frac{1}{2}$.

✓ Παρατήρηση : αν έχουμε m clauses εύκολα φτιάχνουμε verifier με σφάλμα $\leq 1 - \frac{1}{m}$.

Το PCP θεώρημα (4)

Πρόβλημα:

Έστω V ένας $\mathrm{PCP}(\log n, 1)$ verifier του 3-SAT. Να βρεθεί απόδειξη y που να μεγιστοποιεί την πιθανότητα αποδοχής ενός στιγμιότυπου από το V.

Πρόταση:

Δεν υπάρχει ½ προσεγγιστικός αλγόριθμος για το παραπάνω πρόβλημα εκτός αν $\mathrm{P} = \mathrm{NP}$.

Το PCP θεώρημα (5)

Απόδειξη

- ✓ Αν υπήρχε τέτοιος αλγόριθμος θα μας παρείχε απόδειξη για ένα ικανοποιήσιμο τύπο του προβλήματος 3 SAT, όπου ο V θα είχε πιθανότητα αποδοχής $\geq 1/2$.
- ✓ Για ένα τύπο μη ικανοποιήσιμο η πιθανότητα αποδοχής θα είναι $< 1/2$.
- ✓ Ο V μπορεί να προσομοιωθεί για όλα τα $O(\log n)$ τυχαία strings και να υπολογιστεί έτσι η πιθανότητα αποδοχής σε πολυωνυμικό χρόνο.
- ✓ Άρα αποφασίζει το SAT σε πολυωνυμικό χρόνο.

Δυσκολία του MAX-3 SAT

Βοηθητικό πρόβλημα: MAX k-FUNCTION SAT

Έστω n μεταβλητές και m συναρτήσεις με k μεταβλητές η κάθε μια (k σταθερός). Να βρεθεί αποτίμηση των μεταβλητών που ικανοποιεί μέγιστο αριθμό από συναρτήσεις.

Λήμμα:

Υπάρχει σταθερά k και gap-introducing αναγωγή από το SAT στο MAX k-FUNCTION SAT που απεικονίζει τον τύπο ϕ στο στιγμιότυπο $I \in$ MAX k-FUNCTION SAT τέτοιο ώστε

- ϕ ικανοποιήσιμος ➔ $\mathrm{OPT}(I) = m$
- ϕ μη ικανοποιήσιμος ➔ $\mathrm{OPT}(I) < \frac{1}{2}m$.

Δυσκολία του MAX-3 SAT (2)

Απόδειξη

- ✓ Έστω $V \in \text{PCP}(\log n, 1)$ με σταθερές c, q .
- ✓ Υπάρχουν n^c διαφορετικά random strings και για κάθε ένα διαβάζονται q bits.
- ✓ Για κάθε ένα από αυτά τα qn^c bits ορίζω μια νέα μεταβλητή. Το μέρος της απόδειξης που συμβουλευόμαστε μπορεί να ειδωθεί σαν αποτίμηση για τις αντίστοιχες μεταβλητές.
- ✓ Ο verifier δεδομένων (και σταθεροποιημένων) ϕ, r , ρωτά q bits και αποδέχεται ή όχι. Αυτό μπορούμε να το δούμε ως λογική συνάρτηση των q bits. (άρα $k=q$)
- ✓ Για κάθε r παίρνω έτσι μια συνάρτηση f_r. Αυτές είναι οι ζητούμενες. ∎

Δυσκολία του MAX-3 SAT (3)

Θεώρημα:

Υπάρχει σταθερά $\varepsilon_M > 0$ και gap-introducing αναγωγή από το SAT στο MAX-3 SAT που μετατρέπει μια πρόταση ϕ σε μια πρόταση ψ τέτοιες ώστε αν m είναι το πλήθος των clauses της τελευταίας :

- ϕ ικανοποιήσιμος ➔ $\mathrm{OPT}(\psi) = m$
- ϕ μη ικανοποιήσιμος ➔ $\mathrm{OPT}(\psi) < (1 - \varepsilon_M)m$

Απόδειξη:

✓ Για κάθε μια από τις f_r συναρτήσεις του λήμματος μπορεί κανείς να φτιάξει φόρμουλα ψ_r σε μορφή SAT.

Δυσκολία του MAX-3 SAT (4)

- ✓ Παίρνουμε $\psi = \wedge_r \psi_r$.
- ✓ Αν ϕ ικανοποιήσιμη τότε ψ ικανοποιήσιμη
- ✓ Αν ϕ μη ικανοποιήσιμη τότε $> \frac{1}{2} n^c \left(= \frac{1}{2} m \right)$ clauses του ψ είναι ψευδείς
- ✓ Μετατρέπουμε τον ψ σε μορφή 3 SAT εισάγοντας νέες μεταβλητές
- ✓ Το θεώρημα ισχύει για $\varepsilon_M = \frac{1}{2^{q+1}(q-2)}$ ∎

Πόρισμα:

Έστω $\varepsilon_M > 0$ η παραπάνω σταθερά. Δεν υπάρχει $1-\varepsilon_M$ προσεγγιστικός αλγόριθμος για το MAX-3 SAT εκτός και αν $\mathrm{P} = \mathrm{NP}$.

Δυσκολία του φραγμένου MAX-3 SAT

- **Ορισμός:** MAX-3 SAT(k)

Ορίζουμε MAX-3 SAT(k) να είναι ο περιορισμός του MAX-3 SAT όπου κάθε μεταβλητή εμφανίζεται το πολύ k φορές.

- **Θεώρημα:**

Υπάρχει gap-preserving αναγωγή από το MAX-3 SAT στο MAX-3 SAT(29) που μετατρέπει μια πρόταση ϕ σε μια πρόταση ψ τέτοιες ώστε αν m,m' είναι το πλήθος των clauses των ϕ,ψ αντίστοιχα τότε

$$\mathrm{OPT}(\phi) = m \Rightarrow \mathrm{OPT}(\psi) = m'$$

$$\mathrm{OPT}(\phi) < (1-\varepsilon_M)m \Rightarrow \mathrm{OPT}(\psi) < (1-\varepsilon_b)m'$$

με $\varepsilon_b = \varepsilon_M / 43$

Δυσκολία του φραγμένου MAX-3 SAT (2)

- **Πόρισμα:**

Δεν υπάρχει $1-\varepsilon_b$ προσεγγιστικός αλγόριθμος για το MAX-3 SAT(29) εκτός και αν $\mathrm{P} = \mathrm{NP}$.

Δυσκολία του vertex cover

- **Ορισμός:** VC(d)

Ορίζουμε με VC(d) τον περιορισμό του vertex cover που σε κάθε του στιγμιότυπο κάθε κόμβος έχει βαθμό το πολύ d.

- **Θεώρημα:**

Υπάρχει gap-preserving αναγωγή από το MAX-3 SAT(29) στο VC(30) που μετατρέπει ένα τύπο ϕ σε ένα γράφημα $G=(E,V)$ τέτοιο που αν m είναι το πλήθος των clauses του ϕ:

$$\mathrm{OPT}(\phi)=m \Rightarrow \mathrm{OPT}(G)\le \frac{2}{3}|V|$$

$$\mathrm{OPT}(\phi)=(1-\varepsilon_b)m \Rightarrow \mathrm{OPT}(G)>(1+\varepsilon_u)\frac{2}{3}|V|$$

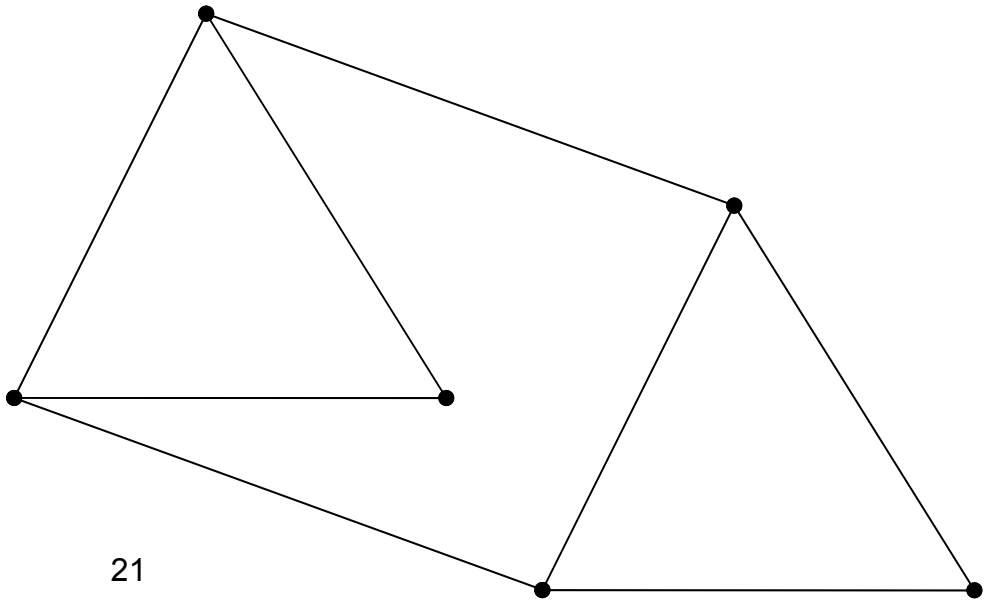
με $\varepsilon_u = \varepsilon_b / 2$.

Δυσκολία του vertex cover (2)

Απόδειξη:

- ✓ Ο γράφος θα έχει $3m$ ($m = \frac{|V|}{3}$) κόμβους και αν σε δύο clauses υπάρχει η ίδια μεταβλητή με την άρνησή της τότε υπάρχει και η αντίστοιχη ακμή στο γράφο, π.χ.

$$(x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee x_3)$$



Δυσκολία του vertex cover (3)

- ✓ Κάθε κόμβος θα έχει ακριβώς 2 ακμές (από κάθε clause) και το πολύ 28 επιπλέον λόγω των πολλαπλών εμφανίσεων των μεταβλητών.
- ✓ Το μέγεθος του maximum independent set του G είναι ακριβώς OPT(φ) διότι:
  - Αν έχουμε αποτίμηση του φ μπορούμε να πάρουμε ένα κόμβο για κάθε ικανοποιήσιμο clause. Αυτό είναι independent set
  - Αν έχουμε ένα independent set I και θέσουμε τα αντίστοιχα literals TRUE, κάθε επέκταση αυτής της αποτίμησης θα ικανοποιεί $|I|$ clauses.
- ✓ Το συμπλήρωμα του maximum independent set είναι minimum vertex cover, άρα :
- ✓ $\mathrm{OPT}(\phi) = m \Rightarrow \mathrm{OPT}(G) = 2m$ και έπεται το θεώρημα ∎

Δυσκολία του Steiner tree

Θεώρημα:

Υπάρχει gap preserving αναγωγή από το VC(30) στο πρόβλημα Steiner tree. Ένα γράφημα $G=(E,V)$ του VC(30) ανάγεται στο $H=(R,S,\text{cost})$ που οι ακμές ικανοποιούν την τριγωνική ανισότητα στο $R \cup S$ έτσι ώστε

$$OPT(G) \le \frac{2}{3}|V| \Rightarrow OPT(H) \le |R| + \frac{2}{3}|S| - 1$$

$$OPT(G) > (1+\varepsilon_u)\frac{2}{3}|V| \Rightarrow OPT(H) > (1+\varepsilon_s)\left(|R| + \frac{2}{3}|S| - 1\right)$$

με $\varepsilon_s = \frac{4}{97}\varepsilon_u$.

Δυσκολία του Steiner tree (2)

Απόδειξη: (σκιαγράφηση)

- ✓ Παίρνουμε ένα required κόμβο για κάθε ακμή του G και ένα Steiner κόμβο για κάθε κόμβο του G.
- ✓ Μια ακμή με Steiner κόμβους έχει κόστος 1
- ✓ Μια ακμή με required κόμβους έχει κόστος 2
- ✓ Κάθε άλλη ακμή έχει κόστος 1 ανν ο αντίστοιχος κόμβος και ακμή του G συσχετίζονται.
- ✓ Ισχύει ότι ο G έχει vertex cover μεγέθους c ανν ο H έχει Steiner tree κόστους $|R|+c-1$.

Δυσκολία του Steiner tree (3)

- ✓ (→) Ας πάρουμε τους Steiner κόμβους που αντιστοιχούν σε ένα vertex cover μεγέθους c.
- ✓ Υπάρχει δένδρο που καλύπτει κόμβους των $R \cup S_c$, που κάθε του ακμή έχει κόστος 1 (λόγω κατασκευής), άρα το ζητούμενο.
- ✓ (←) Αν πάρουμε Steiner tree μεγέθους $|R|+c$-1 Υπάρχει Steiner tree ίδιου μεγέθους, με ακμές κόστους 1.

Πόρισμα:

Δεν υπάρχει $1+\varepsilon_s$ προσεγγιστικός αλγόριθμος για το Steiner tree εκτός και αν $\mathrm{P} = \mathrm{NP}$.

Δυσκολία του clique

- Ανήκει σε κλάση προβλημάτων που είναι πολύ δύσκολα στην προσέγγιση

- Ακόμα και μια τετριμμένη εφικτή λύση δεν είναι πολύ χειρότερη από μια προσεγγιστική λύση

Πρόβλημα: clique

Δεδομένου μη κατευθυνόμενου γράφου με μη αρνητικά βάρη στους κόμβους, να βρεθεί κλίκα μέγιστου βάρους. (Θα ασχοληθούμε με cardinality πρόβλημα)

Δυσκολία του clique (2)

- Υπάρχει $e_q > 0$ τέτοιο ώστε δεν υπάρχει δεν υπάρχει $1/n^{e_q}$ προσεγγιστικός αλγόριθμος για το παραπάνω πρόβλημα.
- Με τη γνώση που έχουμε μέχρι στιγμής μπορούμε μόνο να αποδείξουμε αδυναμία ½ προσέγγισης για το πρόβλημα της κλίκας
- Η απόδειξη γίνεται με gap-preserving αναγωγή από το SAT που χρησιμοποιεί το PCP θεώρημα.
- Ο βαθμός της αδυναμίας της προσέγγισης είναι ακριβώς το πιθανοτικό λάθος της απόδειξης του PCP(log$n$,1) για το SAT.

Δυσκολία του clique (3)

- **Ορισμός:** Γενίκευση του PCP

$L \in \mathrm{PCP}_{c(n),s(n)}(r(n),q(n))$ ανν υπάρχει verifier που στην είσοδο x με n bits δέχεται μια τυχαία ταινία μήκους το πολύ $O(r(n))$, ρωτά $O(q(n))$ bits από την απόδειξη και

$x \in L$ ➔ ο verifier αποδέχεται με πιθανότητα $\geq c(n)$ (completeness)

$x \notin L$ ➔ ο verifier αποδέχεται με πιθανότητα $< s(n)$ (soundness)

- **Πόρισμα:** $\mathrm{PCP}(r(n),q(n)) = \mathrm{PCP}_{1,\frac{1}{2}}(r(n),q(n))$

Δυσκολία του clique (4)

- Προσομοίωση του PCP(log n,1) k φορές δίνει soundness $1/2^k$
- Απαιτεί O(k log n) τυχαία bits και O(k) ερωτήσεις στην απόδειξη.
- Για να έχω αντίστροφα πολυωνυμική δυσκολία προσεγγισιμότητας θέλουμε το k να είναι $\Omega(\log n)$
- Τότε όμως θα απαιτούνται $O(\log^2 n)$ τυχαία bits. (πρόβλημα)

Ορισμός: Expander γράφοι:
Όλοι οι κόμβοι έχουν τον ίδιο βαθμό και $|E(S,S^c)| > \min(|S|,|S^c|)$, όπου $E(S,S^c)$ είναι οι ακμές στο cut (S,S^c).

Δυσκολία του clique (5)

- Σταθερού βαθμού expander γράφος μπορεί να χρησιμοποιηθεί για να παράγουμε O(log n) strings με b log n bits το κάθε ένα Κάθε κόμβος από τους n^b θα έχει ένα μοναδικό label από b log n bits.

- Αν και όχι τυχαία αυτά τα bits είναι τουλάχιστον "ικανοποιητικά τυχαία"

- Για να παράγω $O(\log^2 n)$ τυχαία bits χρειάζομαι μόνο O(log n) bits: b log n για να επιλέξω τυχαία ένα κόμβο και μια σταθερά (ο βαθμός του γράφου) για κάθε βήμα του τυχαίου περιπάτου.

Δυσκολία του clique (6)

Θεώρημα:

Έστω H expander γράφος με n^b κόμβους και S ένα υποσύνολο των κόμβων αυτού με μέγεθος $< n^b/2$. Υπάρχει σταθερά k τέτοια ώστε

$\boldsymbol{Pr}[$*σε ένα τυχαίο περίπατο μήκους k log n όλοι οι κόμβοι είναι μέσα στο S*$] < \frac{1}{n}$

Θεώρημα: $\mathrm{NP} = \mathrm{PCP}_{1,\frac{1}{2}}(\log n, 1) = \mathrm{PCP}_{1,\frac{1}{n}}(\log n, \log n)$

Απόδειξη: Για το ότι $\mathrm{PCP}_{1,\frac{1}{2}}(\log n, 1) \subseteq \mathrm{PCP}_{1,\frac{1}{n}}(\log n, \log n)$:

✓ Αν μια γλώσσα ανήκει στην πρώτη κλάση θα βρούμε verifier που να την αποδέχεται με τις προδιαγραφές της δεύτερης.

Δυσκολία του clique (7)

- ✓ Ο νέος verifier θα κατασκευάζει το expander γράφο και θα δημιουργεί ένα τυχαίο περίπατο (για τα απαιτούμενα $O(\log^2 n)$ τυχαία bits).
- ✓ Θα προσομοιώνεται τότε ο verifier F της πρώτης κλάσης log n φορές και θα αποδεχόμαστε ανν ο F αποδέχεται όλες τις φορές.
- ✓ *Completeness* = 1
- ✓ *Soundness* : αν ένα στοιχείο δεν ανήκει στη γλώσσα τότε ο F λέει ναι σε $< n^b/2$ τυχαία strings από την απόδειξη
- ✓ Παίρνουμε για S το αντίστοιχο σύνολο κόμβων των παραπάνω strings
- ✓ Ο νέος verifier θα αποδέχεται αν ο περίπατος μένει στο S

Δυσκολία του clique (8)

Θεώρημα:

Για σταθερές b και q υπάρχει gap-preserving αναγωγή από το SAT στο πρόβλημα clique που μετατρέπει ένα τύπο φ μεγέθους n σε ένα γράφο $G=(E,V)$ με $|V| = n^{b+q}$ έτσι ώστε

- φ ικανοποιήσιμος ➔ $\mathrm{OPT}(G) \geq n^b$
- φ μη ικανοποιήσιμος ➔ $\mathrm{OPT}(G) < n^{b-1}$

Απόδειξη (σκιαγράφηση)

✓ Θα κατασκευάσουμε γράφο με τη βοήθεια του $\mathrm{PCP}_{1,\frac{1}{n}}(\log n, \log n)$ verifier του SAT.

Δυσκολία του clique (9)

- ✓ Έστω ότι ο verifier χρησιμοποιεί $b \log n$ τυχαία bits και συμβουλεύεται $q \log n$ bits από την απόδειξη.
- ✓ Αν φ τύπος μεγέθους n, τότε για κάθε string r μεγέθους $b \log n$ και κάθε αποτίμηση τ των $q \log n$ μεταβλητών του τύπου, υπάρχει από ένας κόμβος $v_{r,\tau}$.
- ✓ Δύο κόμβοι θα λέγονται *συνεπείς* μεταξύ τους ανν οι αντίστοιχες αποτιμήσεις τους συμφωνούν στα αντίστοιχα τυχαία bits (που έχουν κοινά) που ρωτά ο verifier.
- ✓ Δύο κόμβοι θα είναι γειτονικοί ανν είναι συνεπείς μεταξύ τους και αποδέχονται την είσοδο.

Δυσκολία του clique (10)

Πόρισμα:

Δεν υπάρχει $1/n^{e_q}$ προσεγγιστικός αλγόριθμος για το πρόβλημα clique ($\varepsilon_q = \frac{1}{b+q}$) εκτός και αν $P = NP$.

Δυσκολία του set cover

- **Ορισμός:** two-prover-one round $2\text{P1R}_{c,s}(r(n))$

Ο verifier κάνει από μία ερώτηση σε δύο provers P_1, P_2 που δεν επικοινωνούν μεταξύ τους, έχοντας στη διάθεσή του $r(n)$ τυχαία bits

- $x \in L$ ➔ υπάρχει ζεύγος αποδείξεων που παρέχουν οι provers έτσι που ο verifier αποδέχεται με πιθανότητα $\geq c$ (completeness)
- $x \notin L$ ➔ για κάθε ζεύγος αποδείξεων που παρέχουν οι provers ο verifier αποδέχεται με πιθανότητα $< s$ (soundness)

Ο verifier αποφασίζει σε πολυωνυμικό χρόνο

- **Θεώρημα:**

Υπάρχει σταθερά $\varepsilon_P > 0$ τέτοια που $\text{NP} = 2\text{P1R}_{1,1-\varepsilon_P}(\log n)$

Δυσκολία του set cover (2)

Απόδειξη:

- ✓ Για το δύσκολο εγκλεισμό $\mathrm{NP} \subseteq \mathrm{2P1R}_{1,1-\varepsilon_P}(\log n)$ αρκεί να δείξει κανείς ότι το SAT βρίσκεται στη δεύτερη κλάση
- ✓ Χρησιμοποιείται gap-introducing αναγωγή από το SAT στο MAX-3 SAT(5).
- ✓ Ο verifier ρωτά τους δύο provers για τον τύπο του SAT και την απεικόνισή του στο MAX-3 SAT(5)

Δυσκολία του set cover (3)

- Αποδεικνύεται ότι η gap-introducing αναγωγή από το SAT στο MAX-3 SAT(5) μπορεί να ικανοποιεί τα εξής:
  - Κάθε μεταβλητή του SAT εμφανίζεται ακριβώς 5 φορές.
  - Κάθε clause έχει τρεις διακριτές μεταβλητές
  - Ως συνέπεια αυτών αποδεικνύεται ότι οι αποδείξεις των δύο provers μπορούν να έχουν το ίδιο μήκος

Δυσκολία του set cover (4)

- **Ορισμός:** (Ένα σύστημα συνόλων)

ο Έστω U σύνολο που θα το λέμε το σύμπαν και το σύστημα συνόλων $\left(U, A_1, \ldots, A_m, A_1^c, \ldots, A_m^c\right)$ με $A_i \subseteq U$.

ο Good cover θα ονομάζουμε ένα cover που θα αποτελείται από ένα σύνολο και το συμπλήρωμά του.

ο Ένα cover που δεν περιέχει ένα σύνολο και το συμπλήρωμά του θα λέγεται bad cover.

Δυσκολία του set cover (5)

- **Θεώρημα:**

Υπάρχει πολυώνυμο $p(.,.)$ για το οποία υπάρχει randomized αλγόριθμος ο οποίος για κάθε m,l παράγει ένα σύστημα συνόλων $\left(U,A_1,\ldots,A_m,A_1^c,\ldots,A_m^c\right)$ με $|U| = p\left(m,2^l\right)$. Με πιθανότητα >½ κάθε bad cover έχει μέγεθος $>l$. Επιπλέον ο αλγόριθμος είναι πολυωνυμικός στον μέγεθος του U.

Δυσκολία του set cover (6)

- Μείωση της πιθανότητας λάθους του $2\text{P1R}_{1,1-\varepsilon_P}(\log n)$
- Κανείς Θα περίμενε ότι μετατρέποντας το $2\text{P1R}_{1,1-\varepsilon_P}(\log n)$ αλγόριθμο σε έναν με k ερωτήσεις (με μια απάντηση) προς τους provers, θα πετύχαινε σφάλμα $(1-\varepsilon_P)^k$
- Αυτό δεν είναι αλήθεια
- Οι provers θα μπορούσαν να κοιτάξουν και τις k ερωτήσεις και να συνδυάσουν τις απαντήσεις ώστε να μεγαλώσουν την πιθανότητα εξαπάτησης
- k ανεξάρτητες ερωταποκρίσεις δεν αφορούν συστήματα $2\text{P1R}_{1,1-\varepsilon_P}(\log n)$

Δυσκολία του set cover (7)

- **Θεώρημα:**

Έστω ότι το σφάλμα ενός 2P1R αλγορίθμου είναι $\delta<1$. Η πιθανότητα παραπλάνησης με k ερωτήσεις παραμένει μικρότερη από δ^{dk} όπου d εξαρτάται μόνο από το μέγεθος των απαντήσεων του αρχικού αποδεικτικού συστήματος.

Δυσκολία του set cover (8)

Θεώρημα:

Υπάρχει σταθερά $c>0$ για την οποία υπάρχει randomized gap-introducing αναγωγή Γ που απαιτεί χρόνο $n^{O(\log\log n)}$, από το SAT στο set cover που μετατρέπει ένα τύπο φ σε ένα σύστημα συνόλων S με σύμπαν μεγέθους $n^{O(\log\log n)}$ έτσι ώστε

- φ ικανοποιήσιμος → $\mathrm{OPT}(S)= 2n^k$
- φ μη ικανοποιήσιμος → $\Pr\left[\mathrm{OPT}(S) > cn^k k\log n\right] > 1/2$

όπου n είναι το μέγεθος των αποδείξεων των δύο provers του 2P1R (πολυωνυμικά μεγάλο στο μέγεθος του φ) και $k \in O(\log\log n)$

Δυσκολία του set cover (9)

Απόδειξη: (Σκιαγράφηση)

- ✓ Θυμόμαστε τον 2P1R verifier του SAT, έστω V
- ✓ Εφαρμόζουμε $k = O(\log\log n)$ φορές παράλληλα τον V.
- ✓ Για τον νέο verifier υπάρχουν $(3n)^k$ τυχαία strings
- ✓ Θέτουμε $m = 2^k, l = O(k \log n)$ και κατασκευάζουμε ένα σύστημα συνόλων
- ✓ Για κάθε ένα τυχαίο string φτιάχνουμε και ένα τέτοιο σύστημα συνόλων $\left(U_r, A_r^1, \ldots, A_r^{2^k}, \left(A_r^1\right)^C, \ldots, \left(A_r^{2^k}\right)^C\right)$
- ✓ Η αναγωγή του θεωρήματος απεικονίζει ένα τύπο φ σε ένα σύστημα με σύμπαν το $\mathbf{Y} = \bigcup_r U_r$. Ανάλογα με τις ερωτήσεις του verifier καθορίζεται και η οικογένεια υποσυνόλων του.

Δυσκολία του set cover (10)

- **Πόρισμα:**
  Υπάρχει σταθερά b τέτοια ώστε δεν υπάρχει $b \log n$ προσεγγιστικός αλγόριθμος για το πρόβλημα set cover εκτός και αν $\mathrm{NP} \subseteq \mathrm{ZTIME}\left(n^{O(\log\log n)}\right)$[1]

- Για πρώτη φορά οι Lund και Γιαννακάκης (1994) δείχνουν ότι δεν υπάρχει $\log n /2$ προσεγγιστικός αλγόριθμος για το set cover εκτός και αν $\mathrm{NP} \subseteq \mathrm{ZTIME}\left(n^{O(\log\log n)}\right)$

- Οι Naor, Schulman και Srinivasan αποδεικνύουν την ίδια προσεγγιστική δυσκολία με παραδοχή περί του $D\mathrm{TIME}\left(n^{O(\log\log n)}\right)$

[1] $\mathrm{ZTIME}\left(f(n)\right)$ είναι η κλάση προβλημάτων που λύνονται από randomized αλγόριθμους σε χρόνο $f(n)$.

Τέλος παρουσίασης