Αλγόριθμοι και Πολυπλοκότητα

Ν. Μ. Μισυρλής

Τμήμα Πληροφορικής και Τηλεπικοινωνιών,

Πανεπιστήμιο Αθηνών

NP-Πληρότητα

Εισαγωγή

- Η θεωρία πολυπλοκότητας ασχολείται με την ιεράρχηση των υπολογιστικών προβλημάτων με βάση τη δυσκολία επίλυσής τους.
- Το πρόβλημα στη θεωρητική πληροφορική ορίζεται ανεξάρτητα από τα δεδομένα εισόδου.
- Δεδομένων n αντικειμένων με βάρη $w_j, j = 1 \dots n$ και αξίες $p_j, j = 1 \dots n$ ζητάμε ένα υποσύνολο αντικειμένων συνολικού βάρους το πολύ W που έχουν τη μέγιστη συνολική αξία.
- Ένα στιγμιότυπο του προβλήματος είναι το ακόλουθο:
- Δεδομένων 5 αντικειμένων με βάρη $1, 4, 5, 2, 3$ και αξίες $3, 4, 7, 2, 5$ αντίστοιχα, ζητάμε ένα υποσύνολο αντικειμένων συνολικού βάρους το πολύ 7 που έχουν τη μέγιστη συνολική αξία.

NP-Πληρότητα

Εισαγωγή

- Η δυσκολία επίλυσης ενός προβλήματος ορίζεται σε σχέση με τους υπολογιστικούς πόρους που καταναλώνονται από την καλύτερη γνωστή μέθοδο επίλυσης.
- Οι υπολογιστικοί πόροι κατά κύριο λόγο συνίστανται από το μέγεθος απαιτούμενης μνήμης, καθώς και από τον απαιτούμενο ασυμπτωτικό χρόνο επίλυσης.
- Για το λόγο αυτό μας ενδιαφέρει η οριακή (ασυμπτωτική) συμπεριφορά της κατανάλωσης μνήμης και του χρόνου επίλυσης.
- Μας ενδιαφέρει η συμπεριφορά ενός αλγόριθμου σε ό,τι αφορά την επίλυση των «μεγάλων» στιγμιοτύπων ενός προβλήματος.

NP-Πληρότητα

Εισαγωγή

- Μελετάμε τη δυσκολία των προβλημάτων από την άποψη του χρόνου εκτέλεσης του ταχύτερου γνωστού αλγόριθμου.
- Σαν «δύσκολο» θα χαρακτηρίζεται ένα πρόβλημα για το οποίο ο καλύτερος γνωστός αλγόριθμος έχει χρόνο εκτέλεσης εκθετικό ως προς το μέγεθος της εισόδου (π.χ. $O(2^n)$).
- Σαν «εύκολο» χαρακτηρίζουμε ένα πρόβλημα το οποίο έχει αλγόριθμο πολυωνυμικού χρόνου.
- Το πρόβλημα του πλανώδιου πωλητή είναι δύσκολο, ενώ το πρόβλημα του μονοπατιού ελάχιστου κόστους είναι εύκολο.

NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*

- Η θεωρία πολυπλοκότητας ταξινομεί τα προβλήματα σε κλάσεις (σύνολα) ισοδυναμίας που ορίζουν ότι τα προβλήματα στην ίδια κλάση έχουν την ίδια δυσκολία.
- P (Deterministic Polynomial Time) και NP (Non deterministic Polynomial Time).
- Οι κλάσεις *P* και *NP* ορίζονται ως προς **προβλήματα απόφασης**, δηλαδή προβλήματα στα οποία καλούμαστε να απαντήσουμε μια συγκεκριμένη ερώτηση με ΝΑΙ ή ΟΧΙ.

NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*

- Η κλάση *P* περιλαμβάνει όλα εκείνα τα προβλήματα απόφασης, τα οποία επιλύονται από ένα ντετερμινιστικό αυτόματο σε πολυωνυμικό χρόνο.
- Η κλάση *NP* περιλαμβάνει όλα τα προβλήματα απόφασης που επιλύονται από ένα μη ντετερμινιστικό αυτόματο σε πολυωνυμικό χρόνο.
- Ένας ισοδύναμος ορισμός της κλάσης *NP* έχει ως εξής:
- Η κλάση *NP* περιλαμβάνει όλα τα προβλήματα απόφασης για τα οποία αν μας δοθεί ένα πιστοποιητικό της απάντησης ΝΑΙ, μπορούμε να επαληθεύσουμε σε πολυωνυμικό χρόνο ότι είναι σωστή.

NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*

- Τα ντετερμινιστικά αυτόματα αντιστοιχούν στις μηχανές και στους αλγορίθμους.
- κάθε βήμα υπολογισμού καθορίζεται (αποφασίζεται) πλήρως από το προηγούμενο:
- διαισθητικά αυτός είναι και ο μόνος τρόπος με τον οποίο γνωρίζουμε πως να υπολογίζουμε.
- Τα μη ντετερμινιστικά αυτόματα έχουν τη δυνατότητα να επιλέγουν το επόμενο βήμα τους μέσα από ένα σύνολο δυνατών υπολογιστικών βημάτων, και ανεξάρτητα από το ποιο ήταν το προηγούμενο βήμα.

NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*

- Δύο διαφορετικές εκτελέσεις ενός μη ντετερμινιστικού αλγορίθμου δεν αντιστοιχούν αναγκαστικά στην ίδια ακολουθία υπολογιστικών βημάτων.
- Τα μη ντετερμινιστικά αυτόματα είναι ένα θεωρητικό κατασκεύασμα, που γενικεύει τα ντετερμινιστικά, υπό την έννοια ότι μια διαδοχή βημάτων ενός ντετερμινιστικού αυτομάτου μπορεί να ανήκει και σε ένα μη ντετερμινιστικό, αλλά το αντίθετο δεν ισχύει.
- Είναι φανερό ότι ένα πιστοποιητικό μπορεί να βρεθεί σε πολυωνυμικό χρόνο για όλα τα προβλήματα της *P* αφού αυτά επιλύονται σε πολυωνυμικό χρόνο.

NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*

- Από τα προηγούμενα προκύπτει ότι $P \subseteq NP$.
- Κάποια προβλήματα είναι πλήρη για μια κλάση
- Δηλαδή την περιγράφουν πλήρως υπό την ακόλουθη έννοια
- η αποδοτική επίλυση ενός πλήρους προβλήματος για μια κλάση πολυπλοκότητας θα επιφέρει την επίλυση όλων των ισοδύναμων προβλημάτων που ανήκουν στην κλάση αυτή.

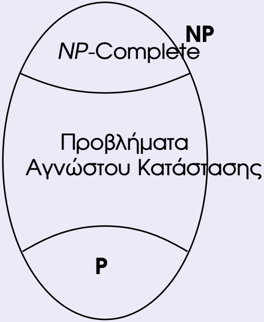
NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*

- Τα *NP*-πλήρη προβλήματα είναι «δύσκολα» σύμφωνα με όσα γνωρίζουμε.
- Είναι προβλήματα για τα οποία γνωρίζουμε ότι ανήκουν στην *NP*, αλλά δεν γνωρίζουμε αν ανήκουν στην *P*
- Ο καλύτερος αλγόριθμος που έχουμε για τα προβλήματα αυτά είναι εκθετικού χρόνου.
- Η αποδοτική επίλυση ενός από αυτά θα επιφέρει την αποδοτική επίλυση όλων.

NP-Πληρότητα

Οι κλάσεις πολυπλοκότητας *P* και *NP*



Σχήμα: Η κλάση *NP*: περιλαμβάνει εύκολα προβλήματα ($P \subseteq NP$), προβλήματα πλήρη για την κλάση (*NP*-πλήρη), καθώς και προβλήματα για τα οποία δε γνωρίζουμε πολυωνυμικό αλγόριθμο, αλλά δεν έχουμε αποδείξει ότι είναι πλήρη για την *NP*.

NP-Πληρότητα

Προβλήματα Απόφασης και NP−πληρότητα

- Ένα πρόβλημα απόφασης ορίζεται από ένα πεπερασμένο σύνολο S, μέσα στο οποίο αναζητείται ένα στοιχείο του που ικανοποιεί μια ιδιότητα P.
- Το ερώτημα στο οποίο καλούμαστε να απαντήσουμε είναι αν υπάρχει ένα τέτοιο στοιχείο $s \in S$.

NP-Πληρότητα

Προβλήματα Απόφασης και $NP-$πληρότητα

Ένα πρόβλημα απόφασης X είναι NP-πλήρες αν:

1. $X \in NP$.
2. Η επίλυση κάθε άλλου προβλήματος απόφασης $\Pi \in NP$ ανάγεται σε πολυωνυμικό χρόνο στην επίλυση του X. Αυτό γράφεται σαν: $\Pi \leq_P X$ για κάθε $\Pi \in NP$.

NP-Πληρότητα

Προβλήματα Απόφασης και NP−πληρότητα

- Αν μπορούσαμε να επιλύσουμε το X σε πολυωνυμικό χρόνο, θα μπορούσαμε να επιλύσουμε και το Π σε πολυωνυμικό χρόνο, με την προϋπόθεση ότι η αναγωγή είναι επίσης πολυωνυμικού χρόνου.
- Η αναγωγή ενός προβλήματος Π σε ένα πρόβλημα X σημαίνει ότι κάθε στιγμιότυπο του Π μπορεί να διατυπωθεί σαν στιγμιότυπο του X με συγκεκριμένους μετασχηματισμούς (αντιστοιχίες).

NP-Πληρότητα

Το πρόβλημα της ικανοποιησιμότητας

- Ίσως το πιο διάσημο πρόβλημα απόφασης στην θεωρία πολυπλοκότητας είναι το πρόβλημα της ικανοποιησιμότητας (Satisfiability Problem - SAT) ένος τύπου λογικής πρώτης τάξης σε κανονική συζευκτική μορφή (Conjunctive Normal Form - CNF).
- Ένα παράδειγμα CNF λογικού τύπου:

$$F = (x_1 \vee \bar{x}_2 \vee x_5) \wedge (x_2 \vee \bar{x}_4 \vee x_3) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_5)$$

NP-Πληρότητα

Το πρόβλημα της ικανοποιησιμότητας

- Η ιδιότητα P για κάθε ανάθεση $s \in \{true, false\}^5 = S$ ορίζεται σαν η αλήθεια της πρότασης $F(s) = true$.
- Αν βρούμε $s^\star \in S$ τέτοιο ώστε $F(s^\star) = true$, τότε αποφασίζουμε ότι η F είναι ικανοποιήσιμη.
- Αν $F(s) = false$ για κάθε $s \in S$ τότε αποφασίζουμε ότι η F είναι μη ικανοποιήσιμη.
- Το SAT έιναι το πρώτο πρόβλημα που αποδείχθηκε πλήρες για την κλάση NP.

NP-Πληρότητα

Το πρόβλημα της ικανοποιησιμότητας

- Η εύρεση ενός αλγόριθμου πολυωνυμικού χρόνου για το SAT θα σημάνει την επίλυση σε πολυωνυμικό χρόνο όλων των προβλημάτων της κλάσης NP, $P = NP$.
- Δε γνωρίζουμε αν υπάρχει πολυωνυμικού χρόνου αλγόριθμος για το SAT.

$$P = NP?$$

NP-Πληρότητα

Προβλήματα Συνδυαστικής Βελτιστοποίησης

- Ένα στιγμιότυπο προβλήματος βελτιστοποίησης ορίζεται από ένα πεπερασμένο χώρο εφικτών λύσεων S και από μια συνάρτηση αποτίμησης των λύσεων $f : S \rightarrow \Re$.
- Μια εφικτή λύση $s \in S$ του προβλήματος ικανοποιεί τους περιορισμούς του προβλήματος.

NP-Πληρότητα

Παραδείγματα Συνδυαστικής Βελτιστοποίησης

- Στο πρόβλημα του Δέντρου Επικάλυψης Ελάχιστου Κόστους ενός γράφου $G(V, E, W)$ με κόστη στις ακμές, το σύνολο S περιέχει όλα τα υποσύνολα ακμών του γράφου που κατασκευάζουν ένα δέντρο.
- Στο πρόβλημα του σακιδίου Knapsack το σύνολο S αποτελείται από τα υποσύνολα αντικειμένων που «χωράνε» μέσα στο σακίδιο.
- Στο πρόβλημα του πλανώδιου πωλητή (Travelling Salesman Problem - TSP) το σύνολο S ορίζεται από όλους τους δυνατούς κύκλους του γράφου, οι οποίοι περνούν μία φορά από όλους τους κόμβους του γράφου.

NP-Πληρότητα

Παραδείγματα Συνδυαστικής Βελτιστοποίησης

- Η συνάρτηση f ονομάζεται αντικειμενική συνάρτηση (objective function), και προσφέρει μια αποτίμηση της κάθε εφικτής λύσης $s \in S$. Στα προβλήματα που αναφέρθηκαν η f δίνει αντίστοιχα:
  - το κόστος του δέντρου επικάλυψης
  - την αξία του υποσυνόλου αντικειμένων
  - το κόστος του κάθε κύκλου.
- Ζητάμε δηλαδή εφικτή λύση $s^\star \in S$ τέτοια ώστε:

$$f(s^\star) = opt_{s \in S} f(s)$$

'Οπου το *opt* αντικαθίσταται από min ή max ανάλογα με τον ορισμό του προβλήματος.

NP-Πληρότητα

Προβλήματα $NP-$δύσκολα

- Οι κλάσεις P και NP ορίστηκαν σε σχέση με προβλήματα απόφασης (ΠΑ).
- Σε κάθε ΠΣΒ μπορεί να αντιστοιχηθεί ένα ΠΑ.
- Αντίστροφα, συμβαίνει ένας αλγόριθμος που επιλύει ένα ΠΑ να μπορεί να χρησιμοποιηθεί για την επίλυση του αντίστοιχου ΠΣΒ.
- Με την εισαγωγή ενός αριθμού k στα δεδομένα ενός ΠΣΒ εύκολα το μετατρέπουμε σε ΠΑ ζητώντας $s \in S$ τέτοιο ώστε $f(s) \leq k$.

NP-Πληρότητα

Προβλήματα $NP-$δύσκολα

- Το ΠΣΒ είναι τουλάχιστον τόσο δύσκολο όσο το αντίστοιχό του ΠΑ.
- Αντίστροφα, αν έχουμε έναν αλγόριθμο για την επίλυση του ΠΑ, τότε μπορούμε να τον χρησιμοποιήσουμε για την επίλυση του αντίστοιχου ΠΣΒ, χρησιμοποιώντας διχοτομική αναζήτηση πάνω στο πεδίο τιμών του k.
- **NP-hardness**:
  Ένα πρόβλημα X είναι NP-δύσκολο αν είναι τουλάχιστον τόσο δύσκολο όσο ένα NP-πλήρες πρόβλημα.
- Ο ορισμός αυτός διαφέρει από την NP-πληρότητα στο ότι αφορά προβλήματα τα οποία δεν ανήκουν στην κλάση NP.
- Τα προβλήματα βελτιστοποίησης δεν ανήκουν στην NP διότι δεν είναι προβλήματα απόφασης.
- Ένα πρόβλημα X είναι NP-δύσκολο αν για κάθε $\Pi \in NP$ ισχύει $\Pi \leq_P X$ και $X \notin NP$. Αν επιπλέον $X \in NP$ τότε το X είναι NP-πλήρες.

NP-Πληρότητα

Αναγωγές

- Τα *NP*-πλήρη και *NP*-δύσκολα προβλήματα συνεχίζουν να αντιστέκονται σε οποιαδήποτε προσπάθεια επίλυσης σε πολυωνυμικό χρόνο.
- Για τα περισσότερα προβλήματα, για τα οποία δε γνωρίζουμε πολυωνυμικό αλγόριθμο, συμβαίνει να έχει αποδειχθεί η *NP*-πληρότητά τους.
- Είναι καλή πρακτική να προσπαθούμε να αποδείξουμε την *NP*-πληρότητα ενός προβλήματος όταν έχουμε αποτύχει να βρούμε πολυωνυμικού χρόνου αλγόριθμο μετά από εκτεταμένη προσπάθεια (και το αντίστροφο).

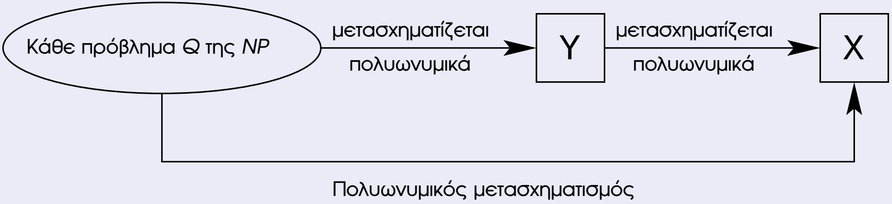
NP-Πληρότητα

Αναγωγές

- Για να αποδείξουμε ότι ένα πρόβλημα $X \in NP$ είναι NP-πλήρες προσπαθούμε να το ανάγουμε σε πολυωνυμικό χρόνο σε ένα άλλο πρόβλημα Y, για το οποίο γνωρίζουμε ήδη ότι είναι NP-πλήρες.
- Σχεδιάζουμε δηλαδή έναν αλγοριθμικό μετασχηματισμό πολυωνυμικού χρόνου των στιγμιοτύπων του Y σε στιγμιότυπα του X.
- Οι αναγωγές πολυωνυμικού χρόνου έχουν μεταβατική ιδιότητα.

NP-Πληρότητα

Αναγωγές



NP-Πληρότητα

Αγνώστου Κατάστασης Προβλήματα

- Υπάρχουν προβλήματα της κλάσης NP για τα οποία δε γνωρίζουμε πολυωνυμικού χρόνου αλγόριθμο που τα επιλύει, ενώ δεν έχουμε επίσης απόδειξη της NP-πληρότητάς τους.
- Ίσως το πιο διάσημο παράδειγμα από αυτά είναι ο ισομορφισμός δύο γράφων.
- Είναι εύκολο να δει κανείς ότι το πρόβλημα αυτό ανήκει στο NP: στην περίπτωση θετικής απάντησης στο ερώτημα αυτό με μια απεικόνιση f σαν πιστοποιητικό μπορούμε να ελέγξουμε την ορθότητα της απάντησης σε πολυωνυμικό $O(|V(G)|^2)$ χρόνο.
- Ένας δεδομένος αριθμός n είναι πρώτος ή όχι;
- Το πρόβλημα αυτό ανήκει στο P όπως αποδείχθηκε πρόσφατα (2002) μετά από δεκαετίες αναζήτησης.