ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΙΚΑ ΜΑΘΗΜΑΤΑ

Θεωρία Υπολογισμού

Ενότητα 13: Ελαχιστοποίηση αυτομάτων

Επ. Καθ. Π. Κατσαρός
Τμήμα Πληροφορικής

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδεια χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

1. Ισοδυναμία συμβολοσειρών

2. Ελαχιστοποίηση προσδιοριστικού πεπερασμένου αυτόματου
   - Διαίρεση καταστάσεων σε κλάσεις
   - Περιγραφή αλγορίθμου ελαχιστοποίησης καταστάσεων
   - Εφαρμογή της ελαχιστοποίησης καταστάσεων
   - Βήματα αλγορίθμου ελαχιστοποίησης προσδιοριστικού αυτόματου

Ισοδυναμία συμβολοσειρών

Ορισμός 1 (Ισοδυναμία ως προς γλώσσα)

Έστω $L \in \Sigma^*$ μία γλώσσα και έστω $x, y \in \Sigma^*$. Λέμε ότι τα x και y είναι **ισοδύναμα ως προς την** L, συμβολίζουμε $x \approx_L y$, αν για όλα τα $z \in \Sigma^*$, αληθεύει ότι: $xz \in L$ αν και μόνο αν $yz \in L$. Παρατηρούμε ότι η $\approx_L$είναι σχέση ισοδυναμίας (ανακλαστική, συμμετρική και μεταβατική).

Ορισμός 2 (Ισοδυναμία ως προς ΠΑ)

Έστω $M = (K, \Sigma, \delta, s, F)$ ένα ντεντερμινιστικό πεπερασμένο αυτόματο. Λέμε ότι δύο συμβολοσειρές είναι **ισοδύναμες ως προς το** M ($x \sim_M y$), αν και οι δύο οδηγούν το M από την s στην ίδια κατάσταση. Δηλαδή, $x \sim_M y$, αν υπάρχει κατάσταση q τέτοια ώστε $(s, x) \vdash_M^* (q, e)$ και $(s, y) \vdash_M^* (q, e)$.

Θεώρημα 3

Για κάθε ντεντερμινιστικό πεπερασμένο αυτόματο $M = (K, \Sigma, \delta, s, F)$ και για όλες τις συμβολοσειρές $x, y \in \Sigma^*$, αν $x \sim_M y$, τότε $x \approx_{L(M)} y$.

Θεώρημα 4 (Myhill-Nerode)

Έστω $L \subseteq \Sigma^*$ μία κανονική γλώσσα. Τότε υπάρχει ένα ντεντερμινιστικό πεπερασμένο αυτόματο, με ακριβώς τόσες καταστάσεις όσες κλάσεις ισοδυναμίας στην $\approx_L$, που δέχεται την L.

Ελαχιστοποίηση προσδιοριστικού πεπερασμένου αυτόματου

- Η μετατροπή ενός μη προσδιοριστικού αυτόματου, σε προσδιοριστικό, μπορεί να προκαλέσει εκθετική αύξηση του αριθμού των καταστάσεων, που στη συνέχεια πρέπει να περιοριστεί με την εφαρμογή ενός αλγορίθμου ελαχιστοποίησης
- Έχει θεμελιωθεί και θεωρητικά, ότι για δοθέν προσδιοριστικό αυτόματο, υπάρχει ένα μοναδικό ισοδύναμο αυτόματο με τον ελάχιστο αριθμό καταστάσεων.
- Ο αλγόριθμος των Hopcroft& Ullman, που περιγράφεται στη συνέχεια, παράγει το ελάχιστο αυτόματο, διαιρώντας τις καταστάσεις σε κλάσεις ισοδυναμίας, έτσι ώστε όλες οι καταστάσεις της ίδιας κλάσης να αναγνωρίζουν τις ίδιες συμβολοσειρές.

Διαίρεση καταστάσεων σε κλάσεις

Ορισμός 5 (Κατάσταση κατάληξη)

Λέμε ότι, σε ένα αυτόματο, η συμβολοσειρά w διακρίνει την κατάσταση s από την κατάσταση t αν ξεκινώντας από την s και διοχετεύοντας την w, οδηγούμαστε σε μία κατάληξη, ενώ ξεκινώντας από την t και διοχετεύοντας την w, οδηγούμαστε σε κατάσταση, που δεν είναι κατάληξη.

- Θεωρούμε ότι η συμβολοσειρά ϵ διακρίνει τις καταλήξεις από τις υπόλοιπες καταστάσεις του αυτόματου.
- Ο αλγόριθμος λειτουργεί ως μία συνεχής βελτίωση της αρχικής διαίρεσης του χώρου των καταστάσεων, σε καταλήξεις και μη καταλήξεις.

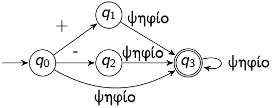
Περιγραφή αλγορίθμου ελαχιστοποίησης καταστάσεων

- Οι κλάσεις ισοδυναμίας συγκροτούνται σταδιακά από καταστάσεις, ώστε
  - κάθε ζεύγος καταστάσεων από *διαφορετικές κλάσεις* να έχει διακριθεί από κάποια συμβολοσειρά εισόδου,
  - για κάθε ζεύγος καταστάσεων της *ίδιας κλάσης*, να μην έχει συμβεί ακόμη κάτι τέτοιο.

  Βήματα:
- Για κάθε κλάση ισοδυναμίας και για κάθε είσοδο a, ελέγχονται οι μεταβάσεις από όλες τις καταστάσεις.
  - Αν οι μεταβάσεις οδηγούν σε καταστάσεις διαφορετικών κλάσεων,
    - τότε δημιουργούνται δύο κλάσεις ισοδυναμίας, έτσι ώστε όλες οι καταστάσεις της κάθε κλάσης ισοδυναμίας να οδηγούν με την είσοδο a στην ίδια κλάση καταστάσεων.
- Η διαδικασία αυτή επαναλαμβάνεται μέχρι οι κλάσεις ισοδυναμίας να μη μπορούν να διαιρεθούν άλλο.

Εφαρμογή της ελαχιστοποίησης καταστάσεων (1/3)

Παρουσιάζουμε την ελαχιστοποίηση του αυτόματου του που αντιστοιχεί στην κανονική έκφραση $(+|-)?\{\text{ψηφίο}\}\{\text{ψηφίο}\}^*$



- Κατά την ελαχιστοποίηση του αυτόματου θα ληφθεί υπόψη και η κατάσταση «λάθος».
- Καταγράφουμε σε πίνακα για κάθε ζεύγος καταστάσεων το αν αυτές διακρίνονται από κάποια συμβολοσειρά εισόδου.
  - Αρχικά βρίσκουμε τα ζεύγη καταστάσεων που διακρίνονται από την ϵ, δηλαδή που αποτελούνται από μία κατάληξη και μία μη κατάληξη.

0				
	1			
		2		
x	x	x	3	
			x	«λάθος»

Εφαρμογή της ελαχιστοποίησης καταστάσεων (2/3)

- Για τις επόμενες συμβολοσειρές εισόδου
  - {ψηφίο}
    - υπάρχουν οι μεταβάσεις (0,3),(1,3),(2,3),(3,3) και «(λάθος»,«λάθος»).
    - Αφού για τις μεταβάσεις (0,3) και «(λάθος»,«λάθος») οι τελικές καταστάσεις 3 και «λάθος» έχουν διακριθεί στον πίνακα, συνάγεται το ίδιο και για τις 0 και «λάθος».

0				
	1			
		2		
x	x	x	3	
x			x	«λάθος»

    - Ομοίως, διακρίνεται η «λάθος» από τις 1 και 2.

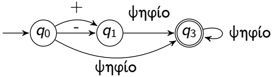
0				
	1			
		2		
x	x	x	3	
x	x	x	x	«λάθος»

Εφαρμογή της ελαχιστοποίησης καταστάσεων (3/3)

- Για τις επόμενες συμβολοσειρές εισόδου
  - $+$ και $-$
    - Αφού για τις μεταβάσεις με $+$ (0,1) και (1,«λάθος») οι τελικές καταστάσεις 1 και «λάθος» έχουν διακριθεί στον πίνακα, συνάγεται το ίδιο και για τις 0 και 1. Ομοίως με το $-$ η 0 διακρίνεται από την 2.

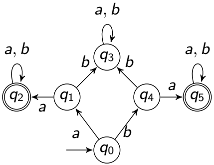
0				
x	1			
x		2		
x	x	x	3	
x	x	x	x	«λάθος»

- Αφού τελικά μόνο οι καταστάσεις 1 και 2 ανήκουν στην ίδια κλάση ισοδυναμίας, τις ενώνουμε σε μία κατάσταση 1 και τελικά παίρνουμε το ελάχιστο αυτόματο.



Εφαρμογή της ελαχιστοποίησης καταστάσεων (1/3)

Παρουσιάζουμε την ελαχιστοποίηση ενός αυτόματου με περισσότερες από μία καταλήξεις.



- Οι καταστάσεις 2 και 5 είναι τελικές, άρα η συμβολοσειρά ϵ τις διακρίνει από τις υπόλοιπες.

0					
	1				
x	x	2			
		x	3		
		x		4	
x	x		x	x	5

Εφαρμογή της ελαχιστοποίησης καταστάσεων (2/3)

- Παρατηρούμε ότι με a έχουμε (0,1) και (1, 2), αλλά οι 1 και 2 διακρίνονται. Συμπεραίνουμε ότι διακρίνεται και η 0 από την 1. Ομοίως διακρίνονται η 0 από την 4, η 4 από την 3 και η 3 από την 1.
- Οι 0 και 3 δε διακρίνονται μεταξύ τους γιατί με μετάβαση a έχουμε τις καταστάσεις 1 και 3 και με μετάβαση b τις καταστάσεις 4 και 3,που και στις δύο περιπτώσεις δεν έχουν βρεθεί να διακρίνονται μεταξύ τους.

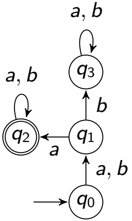
0					
x	1				
x	x	2			
	x	x	3		
x		x	x	4	
x	x		x	x	5

Εφαρμογή της ελαχιστοποίησης καταστάσεων (3/3)

- Παρατηρούμε ότι από τις 0 και 3 με a υπάρχουν μεταβάσεις αντίστοιχα στις 1 και 3, που όμως τώρα διακρίνονται μεταξύ τους και άρα διακρίνουμε τις 0 και 3.

0					
x	1				
x	x	2			
x	x	x	3		
x		x	x	4	
x	x		x	x	5

- Τελικά, οι 1-4 και 2-5 είναι ισοδύναμες μεταξύ τους.



Βήματα αλγορίθμου (1/2)

Αλγόριθμος: Ελαχιστοποίηση προσδιοριστικού αυτόματου

Είσοδος: Προσδιοριστικό αυτόματο M, με χώρο καταστάσεων S, αλφάβητο Σ, αρχική κατάσταση s και σύνολο καταλήξεων A

Έξοδος: Ελάχιστο προσδιοριστικό αυτόματο M' με γλώσσα $L(M)$

1. Θεωρούμε την διαίρεση, Δ_0, του χώρου των καταστάσεων, στην κλάση των καταλήξεων A και την κλάση $S - A$.
2. Για (κάθε κλάση ισοδυναμίας G της Δ_i) επανέλαβε

   διαίρεσε το G σε κλάσεις, έτσι ώστε δύο καταστάσεις s και t του G να είναι στην ίδια κλάση, αν και μόνο αν, για όλους τους χαρακτήρες του αλφαβήτου οι μεταβάσεις οδηγούν σε καταστάσεις της ίδιας κλάσης της Δ_i και αντικατέστησε στη Δ_i, την κλάση G, με τις νέες κλάσεις.

Βήματα αλγορίθμου (2/2)

Αλγόριθμος: Ελαχιστοποίηση προσδιοριστικού αυτόματου (Συν.)

3. Αν $\Delta_i = \Delta_j$ τότε $\Delta_{final} = \Delta_i$, διαφορετικά επανέλαβε το βήμα 2.
4. Επιλέγεται μία αντιπροσωπευτική κατάσταση από κάθε κλάση της Δ_{final}. Οι καταστάσεις αυτές θα αποτελέσουν τις καταστάσεις του ελάχιστου αυτόματου M'.
5. Διαγράφεται από το M' κάθε κατάσταση (και οι μεταβάσεις της) που
   - δεν είναι κατάληξη και έχει μεταβάσεις στον εαυτό της για όλους τους χαρακτήρες του αλφαβήτου (*νεκρή* κατάσταση)
   - που δεν είναι προσεγγίσιμες από την αρχική κατάσταση

Τέλος ενότητας

Επεξεργασία: Εμμανουέλα Στάχτιαρη
Θεσσαλονίκη, 24/07/2014