

Μεταγλωττιστές

Γιώργος Δημητρίου

Μάθημα 11^ο

Γεννήτορας Τελικού Κώδικα

-
-
- Ο γεννήτορας τελικού κώδικα είναι το πιο κρίσιμο τμήμα του μεταγλωττιστή και αντιμετωπίζει πολύπλοκα προβλήματα
 - Βέλτιστη χρήση της αρχιτεκτονικής
- Κακός γεννήτορας τελικού κώδικα δίνει κακό τελικό κώδικα, ανεξάρτητα από τον υπόλοιπο μεταγλωττιστή
 - Κακός κώδικας = αργός κώδικας

Τελικό Πρόγραμμα

-
-
- Το πρόγραμμα που παράγει ο γεννήτορας τελικού κώδικα συνοδεύεται από το χώρο στατικών δεδομένων:
- Βοηθητικά δεδομένα, όπως σταθερές συμβολοσειρές, σταθερές που δε μπορούν να αποτελέσουν τελούμενο εντολής τελικής γλώσσας, ή πίνακες διευθύνσεων
- Στατικές μεταβλητές με αρχικές τιμές

Παραγωγή Τελικού Κώδικα

-
-
- Ενδιάμεσος κώδικας εντολών και εκφράσεων δίνει εκτελέσιμο τελικό κώδικα, αλλά και τιμές σταθερών, διευθύνσεων και συμβολοσειρών που υπάρχουν στον αρχικό κώδικα
- Πίνακες δέσμευσης δίνουν διευθύνσεις και μετατοπίσεις στους αντίστοιχους χώρους δεδομένων, αλλά και αρχικές τιμές στατικών μεταβλητών

Χώροι Δεδομένων

-
-
- Αριθμός bytes για κάθε τύπο δεδομένων της ενδιάμεσης γλώσσας
 - Απαιτείται γνώση της ακριβούς αναπαράστασης κάθε τύπου δεδομένων στην τελική γλώσσα και της πιθανής ευθυγράμμισης στην αποθήκευση
- Μετασχηματισμός του σχεδίου δέσμευσης θέσεων σε σχέδιο δέσμευσης διευθύνσεων του χώρου δεδομένων
 - Απόλυτες διευθύνσεις ή μετατοπίσεις

Σχήματα

Παραγωγής Τελικού Κώδικα

- Σχήμα παραγωγής τελικού κώδικα είναι η διαδικασία μετάφρασης κάθε τύπου κόμβου ενδιαμέσου κώδικα σε συγκεκριμένες ακολουθίες εντολών τελικής γλώσσας
- Η μετάφραση γίνεται σε βήματα πηγαίνοντας από το επίπεδο ενδιαμέσης γλώσσας προς τα κάτω στο επίπεδο τελικής γλώσσας
 - Αρχικός ενδιαμέσος κώδικας → τελικός ενδιαμέσος κώδικας
 - Τελικός ενδιαμέσος κώδικας → τελικός κώδικας

Μετάφραση Εκφράσεων

- Τιμές αριστερής και δεξιάς προσπέλασης
- Κλήσεις υποπρογραμμάτων
- Τελεστές
- Προσωρινές μεταβλητές και δέσμευση χωρητήτων
- Επιλογή εντολών τελικού κώδικα

Τιμές Αριστερής Προσπέλασης

- Ανάπτυξη κόμβων ενδιάμεσου κώδικα για ακριβή υπολογισμό διευθύνσεων
 - Υπολογισμός διεύθυνσης στοιχείου πίνακα:
$$\text{address} = \text{base_address} + [(i_1 - L_1) \times N_2 \times \dots \times N_m + (i_2 - L_2) \times N_3 \times \dots \times N_m + \dots + (i_m - L_m)] \times \text{elem_size}$$
- Προσθήκη κόμβων αποδεικτοδότησης για μετατροπή σε τιμές δεξιάς προσπέλασης
 - Ψευδοτελεστής LOAD

Τιμές Δεξιάς Προσπέλασης

- Οι τιμές βαθμωτών μεταβλητών μπορούν να φορτωθούν από τη μνήμη με δύο τρόπους:
 - με τη βοήθεια του τελεστή LOAD
 - απ' ευθείας, αν η διεύθυνση μπορεί να υπολογιστεί στην ίδια εντολή με τη φόρτωση
- Οι τιμές σταθερών παράγονται επίσης με δύο τρόπους:
 - με ενσωμάτωση στον εκτελέσιμο κώδικα
 - με τοποθέτηση στο χώρο στατικών δεδομένων και φόρτωση από τη μνήμη

Κλήσεις Υποπρογραμμάτων

-
-
- Αποτίμηση πραγματικών παραμέτρων και τοποθέτηση των τιμών τους στη στοίβα
- Ενημέρωση συνδέσμου προσπέλασης και δείκτη στοίβας, ανάλογα με την υλοποίηση
- Εκτέλεση άλματος με σύνδεση στον κώδικα του υποπρογράμματος
- Ανάγνωση της τιμής του αποτελέσματος από τη στοίβα

Τελεστές

- Εκτός από τους τελεστές αναφοράς σε στοιχειά πίνακα και κλήσης υποπρογραμμάτων, οι περισσότεροι δεν απαιτούν ιδιαίτερη ανάλυση
- Αν η τελική γλώσσα δεν περιέχει εντολή που να εκτελεί απ' ευθείας τη λειτουργία του τελεστή:
 - μετατροπή σε απλούστερους κόμβους
 - χρήση υποπρογράμματος (πχ. από βιβλιοθήκη)

Προσωρινές Μεταβλητές

Όταν δεν πρόκειται να προσθέσουμε άλλους κόμβους στον ενδιάμεσο κώδικα, αντιστοιχίζουμε μια προσωρινή μεταβλητή σε κάθε κόμβο αποτίμησης:

$$x*y+5*(i+2)$$

temp1: x, temp2: y, temp3: $x*y$, temp4: 5,
temp5: i, temp6: $i+2$, temp7: $5*(i+2)$, temp8:
 $x*y+5*(i+2)$

Δέσμευση Καταχωρητών

- Έχοντας βρει με ανάλυση ροής ελέγχου τη σειρά εκτέλεσης των εντολών και αποτίμησης των κόμβων σε εκφράσεις:
- Υπολογίζουμε τους χρόνους ζωής των προσωρινών μεταβλητών και κατασκευάζουμε το γράφημα αλληλεπιδράσεων
- Δεσμεύουμε καταχωρητές χρωματίζοντας το γράφημα

Χρωματισμός Γραφήματος

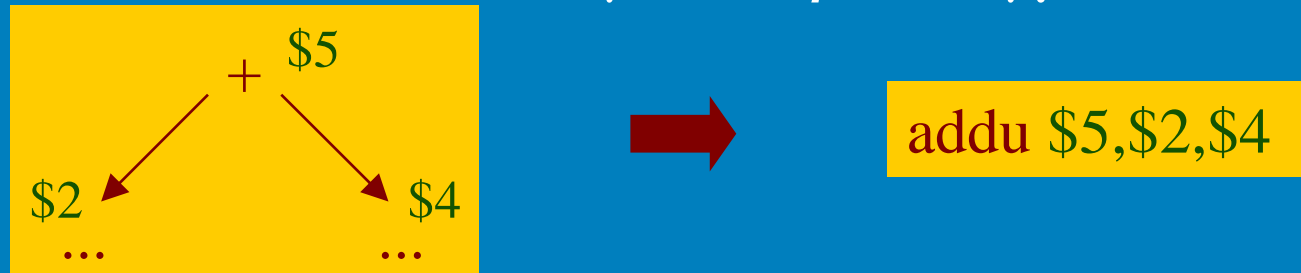
- Έλεγχος χρωματισμού k χρωμάτων για δέσμευση k καταχωρητών
- Ο χρωματικός αριθμός ενός γραφήματος είναι τουλάχιστον ίσος με το μέγεθος της μεγαλύτερης κλίκας του γραφήματος
- Αποδίδουμε χρώμα στους κόμβους του γραφήματος, έχοντας ταξινομήσει αυτούς σε φθίνουσα σειρά αριθμού γειτόνων

Διάχυση και Προδέσμευση

- Αν δεν αρκούν οι διαθέσιμοι καταχωρητές:
 - Διαχέουμε στη στοίβα τον κόμβο με τη μεγαλύτερη διάρκεια ζωής, διαγράφουμε τις αλληλεπιδράσεις του, και επαναλαμβάνουμε!
 - Διάχυση = Αποθήκευση της τιμής μετά τον υπολογισμό της και επαναφόρτωσή της πριν τη χρήση της
- Αν κάποιοι καταχωρητές είναι προδεσμευμένοι (πχ για πέρασμα παραμέτρων):
 - Εφαρμόζουμε την προδέσμευση στο γράφημα και μετά χρωματίζουμε τους υπόλοιπους κόμβους!

Επιλογή Εντολών Τελικού Κώδικα

- Στην απλούστερη περίπτωση αποδίδουμε σε κάθε κόμβο μια εντολή ή μια συγκεκριμένη ακολουθία εντολών, για παράδειγμα:



- Υπάρχουν κι άλλες μέθοδοι βέλτιστης επιλογής εντολών, όπως με ταίριασμα προτύπων και δυναμικό προγραμματισμό

Μετάφραση Εντολών

- Για εντολές ελέγχου ροής (διακλαδώσεις, βρόχοι):
 - Σχήμα παραγωγής κώδικα για κάθε τύπο κόμβου εντολής, με μετατροπή σε απλούς κόμβους ελέγχου ροής που διαθέτει η τελική γλώσσα
 - Στοιχειώδης ανάλυση ροής ελέγχου στον αρχικό ενδιάμεσο κώδικα βοηθάει στη σύνδεση ροής (προορισμοί αλμάτων) στον τελικό κώδικα
- Κατάλληλα σχήματα παραγωγής τελικού κώδικα για όλες τις εντολές

Μορφή Τελικού Κώδικα

-
-
- Γλώσσα μηχανής σε απόλυτη μορφή
 - Άμεση εκτέλεση σε συγκεκριμένες διευθύνσεις
- Γλώσσα μηχανής σε επανατοποθετήσιμη και διασυνδέσιμη μορφή
 - Απαιτείται κλήση του συνδέτη και του φορτωτή πριν την εκτέλεση
 - Η συνηθισμένη μορφή εμπορικών μεταγλωττιστών
- Συμβολική γλώσσα