

Előszó

A fordítóprogramokat általában sejtelmes, nehezen átlátható, titokzatos homály veszi körül. Minden informatikus használja ezeket a programokat, és gyakran megdöbbenve nézik, amikor a nagy gonddal megírt programjuk fordítása végén a fordítóprogram kiírja: „23 errors, 12 warnings”.

Ebben a könyvben azzal foglalkozunk, hogyan működik egy fordítóprogram, hogyan fedezi fel a forrásprogram hibáit, és hogyan készíti el a forrásprogramból a futtatható tárgykódot. Ezzel nemcsak az a célunk, hogy a fordítóprogramok belső világát megvilágítsuk, hanem az, hogy a programozó, ismelve a fordítás algoritmusait, ezt az információt felhasználva jobb, hatékonyabb programot tudjon írni.

A könyvben az *imperatív* nyelvek fordítási módszereit vizsgáljuk. Először megadjuk a fordítóprogramok szerkezetét, majd részletesen tanulmányozzuk a fordítóprogramok részegységeit, amelyek a következők:

- lexikális elemzés,
- szintaktikus elemzés,
 - felülről-lefelé haladó szintaktikus elemzések,
 - alulról-felfelé haladó szintaktikus elemzések,
- szimbólumtábla-kezelés,
- szemantikus elemzés,
- hibakezelés,
- kódgenerálás,
- kódoptimalizálás.

Ez a könyv az Eötvös Loránd Tudományegyetem Informatikai Karán tartott *Fordítóprogramok* tantárgy előadásainak anyagára épül. A fordítóprogramok a számítástudomány és az informatika különböző területeiről származó tudást ötvözik, illetve adatszerkezetek, formális nyelvek, automaták, magasszintű prog-

ramozási nyelvek, assembly nyelvek és assemblerek alapvető ismereteit koncentrálják.

Olyan fordítási módszerekkel foglalkozunk, amelyeket a gyakorlatban használt fordítóprogramokban is alkalmaztak és jelenleg is alkalmaznak. Néhány klasszikus algoritmus után a jelenleg is használt fordítóprogramok működését tanulmányozzuk, és módszereket, eljárásokat adunk arra, hogy hogyan kell egy fordítóprogramot létrehozni.

Mivel elsősorban a gyakorlatra, elvekre, a fordítás módszereire koncentrálnunk, a tételek bizonyítását nem írjuk le, a bizonyítások az Irodalomjegyzékben megadott könyvekben és egyetemi jegyzetekben megtalálhatók. A téma könnyebb megértése és elsajátítása érdekében a könyvben nagyon sok példát adunk, a példák végét a □ jellel jelöljük. A programokat egy könnyen olvasható egyedi pszeudokódban írjuk meg, ha valamely konkrét programnyelv példáját adjuk meg, akkor a példában a nyelvet megnevezzük.

Könyvünkkel elsősorban a programnyelvekkel foglalkozó egyetemi, főiskolai hallgatók és az informatikus szakemberek munkáját szeretnénk segíteni.

* * *

Köszönetemet fejezem ki Kása Zoltán egyetemi tanárnak, a könyv lektorának hasznos tanácsaiért, rendkívüli gondossággal végzett alapos munkájáért.

Köszönöm a munkatársaknak és egyetemi hallgatóknak az észrevételeiket, ezek a könyv megírásához nagy segítséget nyújtottak.

Külön köszönet illeti a Typotex Kiadót, amiért lehetővé tette ennek a könyvnek a kiadását.

Kérem, hogy észrevételeiket a csz@inf.elte.hu címre írják meg!

Budapest, 2005. november 28.

Csörnyei Zoltán

A könyvben a következő jelöléseket használjuk:

G	grammatika,
L	nyelv,
$L(G)$	a G grammatika által meghatározott nyelv,
(N, T, P, S)	grammatika, ahol
	N a nemterminális szimbólumok halmaza,
	T a terminális szimbólumok halmaza,
	P a helyettesítési szabályok halmaza,
	S a kezdőszimbólum,
\rightarrow	helyettesítés,
$A \rightarrow \alpha \mid \beta$	két azonos bal oldalú helyettesítési szabály,
\implies	levezetés,
$\xrightarrow{+}$	egy vagy több alkalmazás,
$\xrightarrow{*}$	nulla, egy vagy több alkalmazás,
a, b, c, \dots	terminális szimbólum (T),
A, B, C, \dots	nemterminális szimbólum (N),
X, Y, Z, \dots	terminális vagy nemterminális szimbólum ($N \cup T$),
$\alpha, \beta, \gamma, \dots$	terminális vagy nemterminális szimbólumok sorozata $((T \cup N)^*)$,
x, y, z, \dots	terminális szimbólumok sorozata (T^*),
ε	az üres szimbólumsorozat.

A fenti jelek indexei a jelek értelmezését nem változtatják meg. Ha egy olyan fogalomtípusra van szükségünk, amelyiknek a fentiekben nem adtunk meg jelkészletet, akkor ezt a hivatkozáskor tesszük meg. Például a nemterminális szimbólumok sorozatát így jelöljük: α , ahol $\alpha \in N^*$. Felhívjuk a figyelmet arra, hogy a könyvben, ha a fenti jeleket használjuk, a jelek értelmezését már nem ismételjük meg.