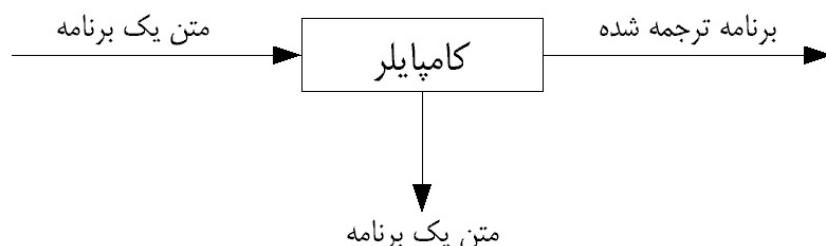


مبانی کامپایلر

۱-۱ تعریف کامپایلر

کامپایلر برنامه‌ای است که در ورودی خود، متن یک برنامه را که طبق قوانین و دستور زبان تدوین شده برای آن کامپایلر مشخص شده است پذیرفته و در خروجی، برنامه‌ای به زبان دوم ایجاد می‌نماید. این زبان دوم می‌تواند ماشین یا برنامه‌ای به زبان دیگر باشد.



برنامه ترجمه شده در بعضی محیط‌ها مثل Unix معمولاً به زبان C می‌باشد. برای مثال کامپایلر راتفورد که نوع پیشرفته‌تر از فرترن می‌باشد در خروجی خود برنامه به زبان C را ایجاد می‌کند که توسط کامپایلر C تبدیل به کد ماشین می‌گردد.

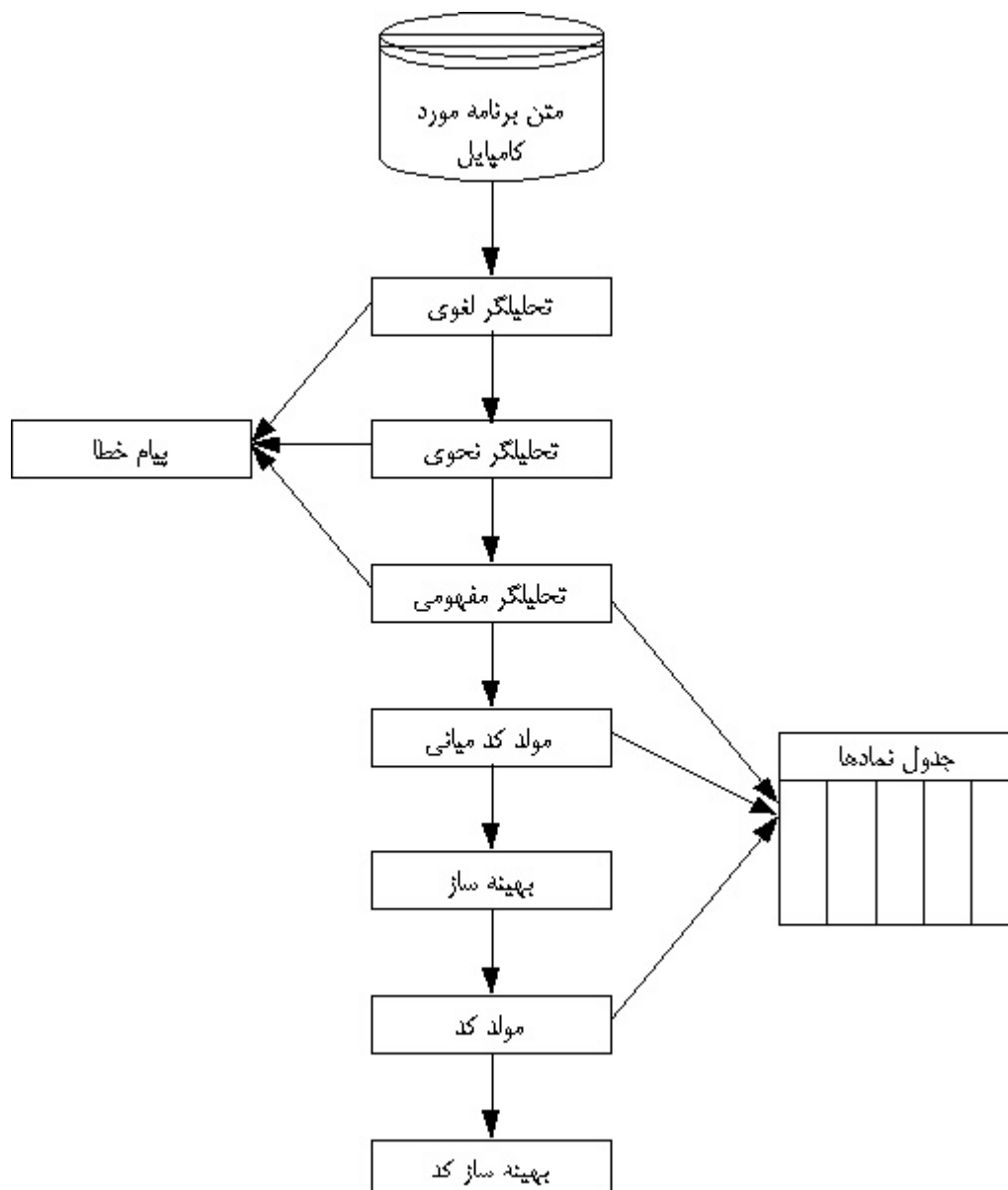
بعضی از کامپایلرها که در اصطلاح مفسر یا Interpreter نامیده می‌شوند همگام با ترجمه هر جمله اجرایی، آن را به اجرا در می‌آورند. برای نمونه نرم‌افزار Dbase را می‌توان نام برد. البته این ترجمه همراه با اجرا، زمان اجرا را بسیار طولانی می‌کند اما امکاناتی در اختیار برنامه‌نویس قرار می‌دهد که کامپایلرهای عادی قادر به ایجاد آن نیستند. برای مثال در زبان Dbase دستورالعمل ماکرو که با عملگر & مشخص می‌شود، این امکان را فراهم می‌کند که بتوان در زمان اجرا محتوی یک رشته را به اجرا درآورد. به مثال زیر توجه نمایید:

```
A = B - 10
&A
a, 2.3 get A
&A
```

در مثال فوق با اجرای اولین دستورالعمل A و محتوای A به عنوان یک دستورالعمل تلقی می‌شود و به اجرا درمی‌آید و به این ترتیب محتوای B برابر ۱۰ قرار می‌گیرد. عمل کامپایل در طی مراحل مشخص که در بخش بعدی در مورد آن بحث خواهد شد، انجام می‌گردد. این مراحل می‌توانند به طور همزمان انجام شوند.

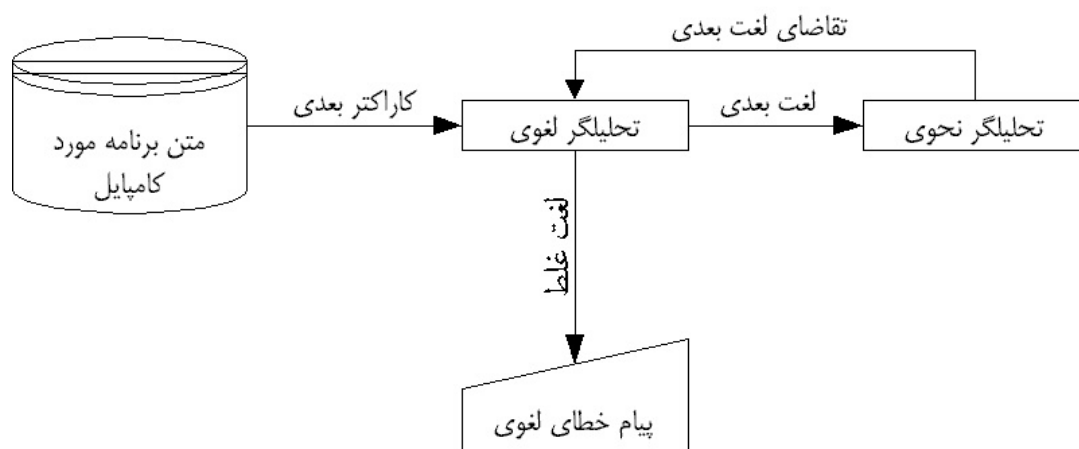
۲-۱ اجزای کامپایلر

همانطور که در شکل زیر مشخص شده است، کامپایلرها در حالت کلی از هفت بخش اصلی تشکیل می‌شوند. در این بخش به طور خلاصه بخش‌های متفاوت کامپایلر تشریح می‌شود.



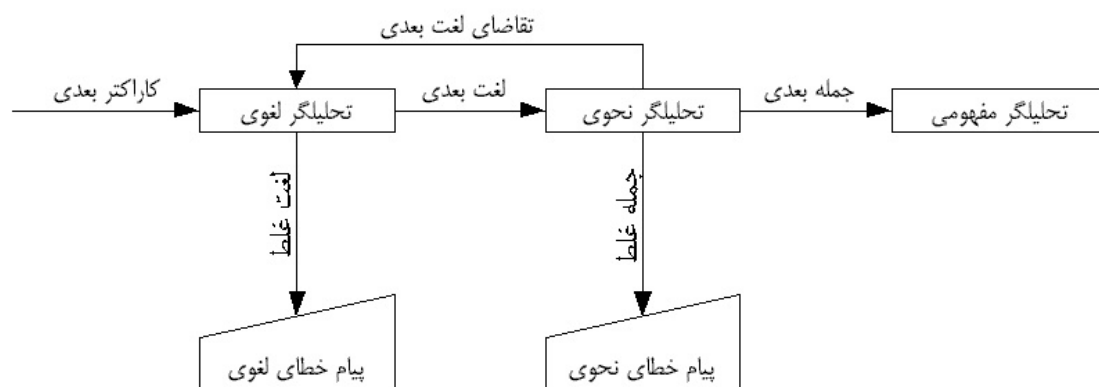
الف - تحلیلگر لغوی

وظیفه تابع تحلیلگر لغوی یا در اصطلاح خارجی Lexical Analyser، تشخیص لغات در متن برنامه مورد کامپایل است. توسط بخش دیگری از کامپایلر تحت عنوان تحلیلگر نحوی مورد فراخوانی قرار می‌گیرد. با هر بار فراخوانی این تابع لغت بعدی را از متن برنامه تشخیص می‌دهد و اطلاعات لازم در مورد لغت را برای تحلیلگر نحوی ارسال می‌دارد. هر زبان برنامه سازی قوانین لغوی مربوط به خود را دارد. قوانین لغوی بیانگر فرم کلی انواع لغات در زبان برنامه سازی است. تحلیلگر لغوی در صورت وجود خطا در قالب بندی لغات استفاده شده در متن برنامه مورد کامپایل اعلام خطای لغوی می‌نماید. تابع تحلیلگر لغوی در فصل دوم از این کتاب مورد بررسی قرار خواهد گرفت.



ب - تحلیلگر نحوی

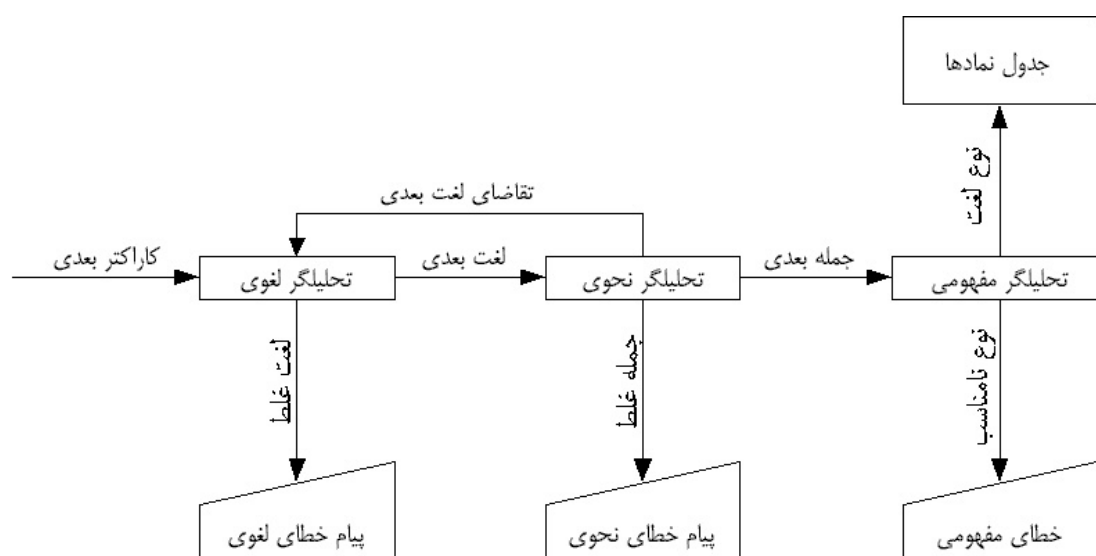
وظیفه تحلیلگر نحوی یا در اصطلاح Syntax Analyser، تشخیص صحت فرم ظاهری متن برنامه از لحاظ دستورالعمل زبان برنامه سازی مربوطه است. تحلیلگر نحوی با فراخوانی تحلیلگر لغوی، لغات را از متن برنامه مورد کامپایل دریافت و صحت قرار گرفتن آنها در کنار یکدیگر بر اساس دستورالعمل زبان مربوطه مورد آزمون و تحلیل نحوی قرار می‌دهد. در صورتی که از لحاظ دستورالعمل زبان، برنامه مورد کامپایل دارای خطا باشد، پیام خطا توسط تحلیلگر نحوی صادر می‌گردد.



انواع روش های تحلیل نحوی و تشخیص صحت برنامه ها بر اساس دستورالعمل و گرامر زبان های برنامه سازی در فصل های ۳ و ۴ و ۵ مورد بررسی واقع خواهد شد.

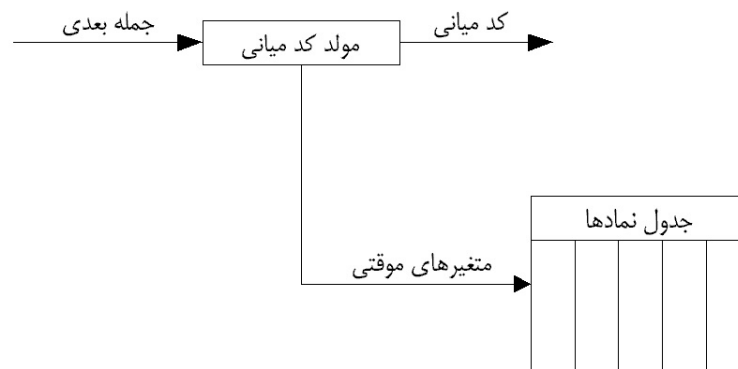
ج- تحلیلگر مفهومی

کار تحلیلگر مفهومی یا در اصطلاح Semantic Analyser تعیین صحت مفهوم جملات است. ممکن است یک جمله از نظر نحوی صحیح ولی از لحاظ مفهومی دارای خطا باشد. برای مثال جمله مهرداد/آتش را خورد، علیرغم اینکه در دستور زبان فارسی از نظر نحوی صحیح می باشد از نظر مفهومی نادرست است. در کامپایلرها به طور معمول عمل تحلیل مفهومی محدود به آزمون نوع یا Type Checking است. تحلیلگر مفهومی وابسته به نوع اسامی و متغیرها که در جدول نمادها مشخص شده، صحت استفاده آنها را مورد آزمون قرار می دهد. برای مثال اگر متغیری از نوع رشته تعریف شده باشد، نمی توان آن را با متغیری از نوع صحیح یا اعشاری جمع یا مقایسه نمود. فصل ۷ در ارتباط با آزمون نوع است.



د- مولد کد میانی

مولد کد میانی بخشی از کامپایلر است که در ورودی خود جملات تشخیص داده شده توسط تحلیلگر نحوی را پذیرفته و در خروجی کد واسطه یا میانی تولید می کند. کد میانی به سادگی قابل تبدیل و نزدیک به زبان ماشین است اما مستقل از ساختار و جزئیات هر گونه ماشین خاص می باشد.



یکی از مزیت های ایجاد کد میانی افزایش قابلیت حمل کامپایلر بر روی سخت افزار با کد متفاوت و ساختار متفاوت است. به این ترتیب که برنامه کامپایل شده را بتوان بر روی کامپیوترها با کد ماشین و اسمبلی متفاوت کامپایل نموده، با اجرای آن متن برنامه های مورد کامپایل را به کد میانی تبدیل کرد و سپس با نوشتن یک برنامه کوچک این کد واسطه را به کد ماشین مورد نظر تبدیل نمود. مزیت دیگر ایجاد کد میانی، فراهم نمودن شرایط خوب برای بهینه سازی کد برنامه هاست. کد میانی در فصل ۶ مطرح شده است.

۵- بهینه ساز کد میانی

هدف از بهینه سازی کد میانی می تواند تقلیل حجم و افزایش سرعت اجرایی کد ماشین حاصل از حاصل کار کامپایلر باشد. برای این منظور بهینه ساز، کد میانی حاصل از مولد کد میانی را مورد تحلیل قرار می دهد. به طور معمول تحلیل کد میانی با آزمون نمادی برنامه ها تحقق می یابد، به این ترتیب که مفهوم برنامه ها در زمان کامپایل، با تبدیل کد میانی به یک گراف به نام گراف جریان مورد بررسی و تحلیل واقع شده، سعی می کنند تا حجم کد حاصل را تقلیل داده، در صورت امکان کد زائد را مشخص و حذف نمایند. بهینه سازی موضوع فصل ۸ این کتاب است.

۶- بهینه ساز کد

هدف از بهینه سازی کد تقلیل حجم اجرایی برنامه است. در این مرحله برای انجام عمل بهینه سازی و رسیدن به اهداف آن ساختار اسمبلی و کد ماشین و امکانات سخت افزاری آن در نظر گرفته می شود و از دستورالعمل هایی که سریعتر و با حجم کمتر به اجرا در می آیند، استفاده می شود. بهینه ساز کد در صورت امکان دستورالعمل های سریع و کم حجم را با دستورالعمل های به کار برده شده در خروجی مولد کد که همان کد اجرایی برنامه است جایگزین می کند.

البته مراحل فوق الذکر بعضاً ممکن است در ایجاد کامپایلر حذف شود. برای مثال قسمت بهینه سازی و یا تحلیل مفهومی ممکن است در یک کامپایلر وجود نداشته باشد. و نیز ممکن است کلیه مراحل در یک مرحله

و به صورت همزمان اجرا شود. معمولاً عملیات کامپایل برنامه ها در دو و یا در اصطلاح در دو **pass** انجام می شود.

۳-۱ کامپایلر کامپایلرها

کامپایلر کامپایلر در واقع یک مولد کامپایلر میباشد و یا برنامه ای است که در ورودی خود قوانین لغوی و گرامر کدی که برای هر جمله عبارت و عنصری از زبان را که باید توسط مولد کد ایجاد شود را در ورودی گرفته ، در خروجی خود یک کامپایلر ایجاد می کند.

در طی فصلهای بعدی خواهیم دید که چگونه می توان یک مولد تحلیلگر لغوی و تحلیلگر نحوی را به سادگی تولید نمود. از کامپایلر کامپایلرهای رایج یکی **YACC** را می توان نام برد. این مولد کامپایلر قواعد گرامری و کد مورد نظر برای تبدیل و در واقع کامپایل جملات متن برنامه کامپایلر را در ورودی خود دریافت می کند. خروجی **YACC** برنامه کامپایلر به زبان **C** است. عمل تحلیل لغوی برای این کامپایلر کامپایلر توسط یک مولد تحلیلگر لغوی **lex** انجام می شود.