عنوان پروژه :

پیاده سازی تحلیلگر لغوی و نحوی

استاد راهنما:

دکتر سکینه اسدی

گردآورنده:

مجتبی صفایی(۹۵۱۲۲۹۱۰۱۴)

پوریا نادری (۹۵۱۲۲۹۱۰۲۴)

سوال :

برنامه ای بنویسید که با دریافت قطعه کدی ابتدا آن را به لیستی از توکن های تبدیل کند و سپس طبق گرامر داده شده و با استفاده از پارسر LL(K) توکن ها را پارس کند.

توضیحات کد:

ابتدا برنامه باید از روی یک فایل متن مورد نظر کاربر را بخواند و آن را به صورت یک متن بدون کاراکتر "\n" ذخیره کند.

پس از آن تابع lex صدا زده می شود که این تابع یک رشته را به عنوان ورودی دریافت می کند.

تابع Lex:

در این تابع ما دو لیست مهم داریم که یکی از آن ها لیست توکن ها و یکی لیست کلمات است.

لیست کلمات شامل همه ی کلمات موجود در کد است. این کلمات شامل کامنت ها نیز می شود.

برای جدا سازی کلمات ما از یک حلقه استفاده می کنیم که این حلقه هر بار یک کاراکتر را خوانده و اگر آن کاراکتر برابر با فاصله نبود این کاراکتر را به کاراکتر های خوانده شده قبلی اضافه می کند و اگر برابر با فاصله بود کاراکتر اضافه نمی شود.

در ادامه یک شرط داریم که اگر کاراکتر بعدی یکی از کلمات کلیدی یا یکی از symbol ها بود باید کلمه ای که تاکنون خوانده ایم را به لیست کلمات اضافه کند.

پس از آن یک حلقه داریم که تک تک کلمات را جست و جو می کند و کلماتی مانند "==" را تولید می کند. برای این کار اگر کلمه ای برابر با "=" بود و کلمه بعدی نیز "=" بود این دو را تبدیل به یک کلمه کرده و دومی را از لیست حذف می کند.

پس از آن یک حلقه داریم که برای تولید توکن هاست این حلقه به ترتیب هر کلمه را چک کرده و به آن ها یک نوع اختصاص می دهد و در آخر این لیست توکن ها را بر می گرداند.

در ادامه و تابع اصلی لیست توکن را به یک لیست از کلمات تبدیل می کنیم که اینجا به جای اعداد نوع آن اعداد و به جای نام متغیر ها id را جایگزین کرده و یک لیست تولید می شود که تحلیلگر نحوی می تواند با آن کار کند.

تحلیلگر نحوی براساس گرامر زیر کار می کند.

هر متغیر در گرامر زیر یک تابع در کد ما است.

یک تابع به اسم eat اضافه شده است که با دریافت توکنی که باید دیده شود چک می کند که آیا توکن کنونی با توکن مورد نظر برابر است یا خیر و اگر برابر بود توکن را رد کرده و توکن های بعدی در ادامه توابع خود چک می شوند.