

به نام خدا

فاز اول پروژه طراحی کامپایلر

عبارات منظم

اعداد: $[-+]?([1-9][0-9]^*|0)$

شناسه ها: $[a-zA-Z_]([0-9]| [a-zA-Z_])^*$

رشته ها: $"(\backslash|.)^*"$

کاراکتر: $'(\backslash|.)'$

توضیحات: $\backslash \backslash .^* \backslash n$

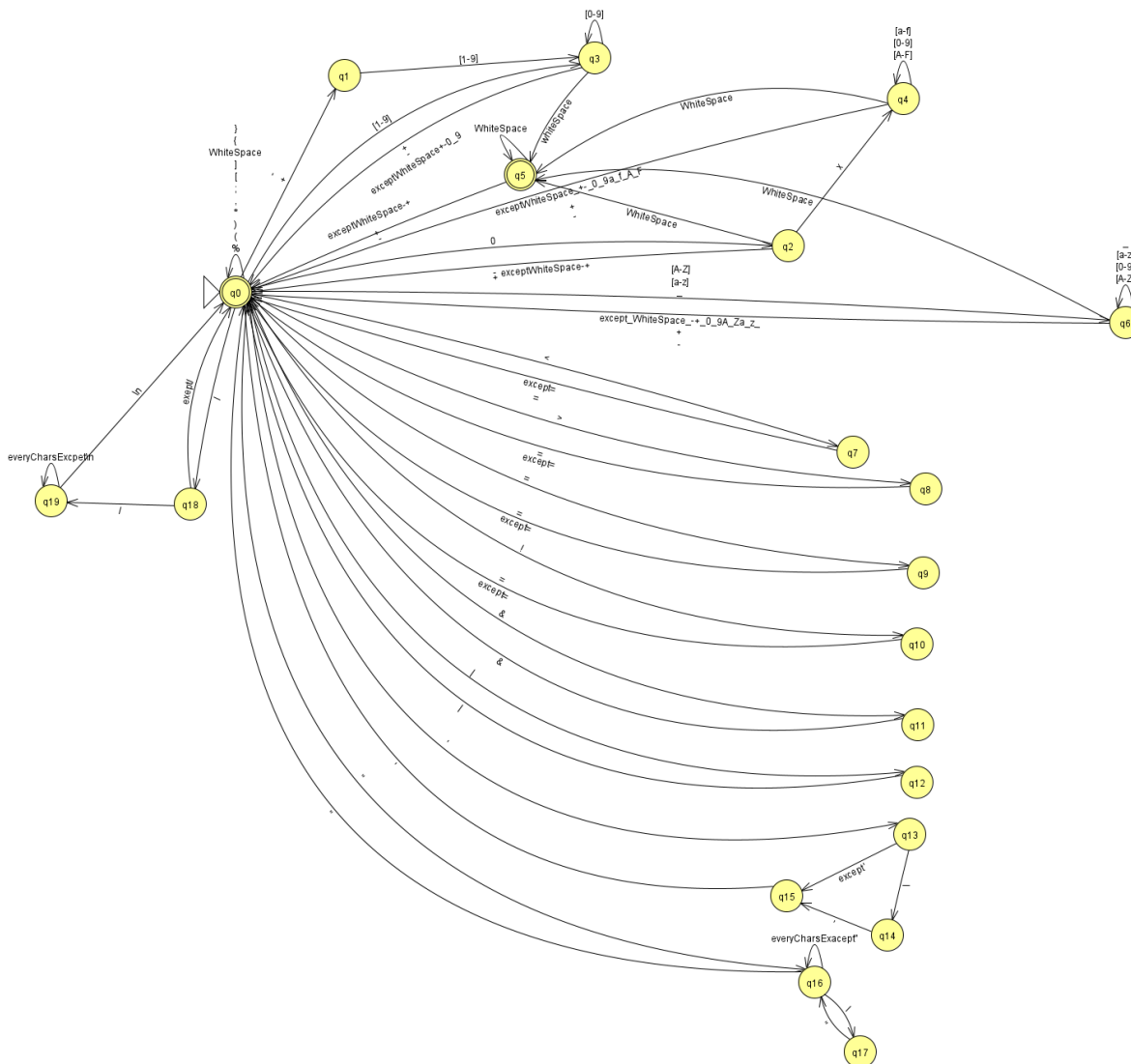
کلمات کلیدی:

`(bool | break | char | continue | else | false | for | if | int | print | return | true)`

با استفاده از این عبارت ها ماشین NFA رسم شده و سپس آن ماشین NFA تبدیل به DFA می شود. نمودار DFA را نتوانستم به دلیل زیاد بودن حالت ها کامل رسم کنم و حالت های تشکیل کلمات کلیدی را در نمودار آورده نشده و صرفا در کد پیاده سازی شده شناسایی می شوند. در فایل زیپ ارسالی عکس آن را می توان به صورت واضح دید. در پیاده سازی کد برای انتقال هایی که از واژه `except` برای آن ها استفاده شده ، در حالت بعدی ادامه بررسی باید از حرف قبلی شروع شود. در غیر این صورت بعضی از توکن ها را نمی توان تشخیص داد و در کد پیاده سازی شده این موضوع لحاظ شده است.

کد با زبان `java script` پیاده سازی شده و در آن برای هر حالت یک آبجکت در نظر گرفته شده. `Property` های این آبجکت ها یال های خروجی از هر حالت هستند و `value` های این `property` ها، آبجکتی هستند که شامل حالت بعدی پس از انتقال از طریق آن یال و توکن های استخراج شده در اثر انتقال از آن یال می باشد که حداکثر ۲ توکن های بعد از هر انتقال در این پیاده سازی تولید می شود

برای بررسی همه حالت های انتقال و بررسی توکن های استخراج شده برای هر حالت یک تابع پیاده سازی شده سعی کردم که کلمات کلیدی را در دیاگرام DFA در نظر بگیرم ولی با کشیدن آن ها نمودار بسیار بزرگ و پیچیده می شد و همچنین کد پیاده سازی شده بسیار طولانی می شد و در نوشتن آن مشکل داشتم. راه حلی که به ذهنم رسید این بود که وقتی توکن مربوط به شناسه تشخیص می شود ، ابتدا آن را با تمام کلمات کلیدی زبان مقایسه کند اگر آن توکن در کلمات کلیدی پیدا شد به جای این که توکن شناسه استخراج شود، توکن مربوط به آن کلمه کلید استخراج شود و اگر در کلمات کلیدی نبود، توکن شناسه استخراج شود.



برای اجرا کردن کد، فایل index.html را در مرورگر اجرا کنید. یک text box وجود دارد که در آن کدی به زبان PL نوشته می شود و با کلیک کردن روی دکمه compile عمل تحلیل لغوری و استخراج توکن های از کد نوشته شده انجام می شود.