(به نام خداوند بخشنده‌ی مهربان)

گزارش فاز اول پروژه‌ی درس کامپایلر

OpenUnderstand

گروه 7:

دیار حامدی

سید عماد موسوی

افشین زنگنه

مرتضی شهرابی فراهانی

فهرست مطالب

[مقدمه: 3](#_Toc102188045)

[نتایج حاصل شده: 3](#_Toc102188046)

[خلاصه‌ای از کارهای انجام شده: 3](#_Toc102188047)

[ساختار ادامه‌ی گزارش: 3](#_Toc102188048)

[نحوه‌ی تقسیم کار: 3](#_Toc102188049)

[روش پیشنهادی: 5](#_Toc102188050)

[فایل main.py: 5](#_Toc102188051)

[تابع getListOfFiles: 6](#_Toc102188052)

[تابه getFileEntity: 6](#_Toc102188053)

[تابع Parse: 7](#_Toc102188054)

[Define\_defineby.py: 7](#_Toc102188055)

[ارزیابی: 8](#_Toc102188056)

[مشکلات و چالش‌ها: 15](#_Toc102188057)

[نتیجه‌گیری و کارهای آتی: 16](#_Toc102188058)

# مقدمه:

## نتایج حاصل شده:

Understand ابزاری قدرتمند برای تجزیه و تحلیل کدها است. در واقع در این نرم‌افزار، تمامی موجودیت‌ها که شامل فایل‌ها، پکیج‌ها، توابع، متغیرها و... است را در پروژه بررسی می‌کند و همچنین نوع رابطه‌ی بین هر کدام از این موجودیت‌ها وجود دارد، خط و نام فایلی که این رابطه در آن رخ داده و سایر موارد اینچنینی را بررسی می‌کند. موارد گفته شده توسط برنامه‌ی Understand انجام می‌شود. اما هدف از این پروژه پیاده‌سازی متن‌باز از Understand Python API برای تجزیه و تحلیل کدهای منبع است. این کار برای کدهای جاوا با استفاده از ابزارهای Python و Antlr-4 انجام می‌شود. با اجرای این پروژه، کارکردهای برنامه‌ی Understand به صورت رایگان و متن باز در اختیار برنامه‌نویسان قرار می‌گیرد.

## خلاصه‌ای از کارهای انجام شده:

در فاز اول پروژه، وظیفه‌ی گروه 7 پیاده‌سازی API مربوط به بخش define / defineIn پروژه‌ی OpenUnderstand بوده‌است. در ابتدا مکان‌ پروژه‌ای که قصد بررسی آن را داریم به برنامه داده می‌شود. سپس برای هرکدام از فایل‌ها و بخش‌های مختلف آن پروژه، بخش‌های define/ defineIn که کد آن‌ها در این بخش زده شده‌است، پیدا می‌شود و در دیتابیس در نظر گرفته شده برای پروژه ریخته می‌شود.

## ساختار ادامه‌ی گزارش:

در ادامه‌ی گزارش، بخش‌های مختلفی از پروژه که در این فاز توسط گروه 7 تغییر داده شده با جزئیات بیشتر و توضیح بیشتر به ترتیب بررسی می‌شود و توضیحات هر بخش به صورت جداگانه داده می‌شود.

## نحوه‌ی تقسیم کار:

بیشتر کار پروژه به صورت تیمی و مشورتی انجام شد. در ابتدا اعضای گروه در چند جلسه‌ی حضوری در سایت دانشکده باهم به بررسی پروژه پرداختند و کدهای هر بخش ما مشارکت همه‌ی اعضای گروه زده شده است. اما اگر بخواهیم به صورت ریزتر، برای هر فرد مشخص کنیم که روی چه بخشی تمرکز بیشتری داشته‌است، نحوه‌ی تقسیم کار بدین صورت می‌شود.

دیار حامدی و سید عماد موسوی: بیشتر بررسی کد define / defineIn

افشین زنگنه: نوشتن داک پروژه

مرتضی شهرابی فراهانی: بخش تست و مقایسه‌ی نتایج با نرم افزار understand

البته تمامی اعضای گروه در تمام بخش‌های پروژه همکاری داشتند و تقسیم‌بندی بالا تنها بیانگر بخشی است که هریک از افراد در آن بخش تاثیر بیشتری داشته‌اند.

# روش پیشنهادی:

در ابتدا برنامه‌ی Understand توسط هریک از اعضای گروه نصب شد و سعی شد تا با بررسی آن با بخش‌های مختلف آن آشنا شویم. پس از به دست آمدن آشنایی نسبی با برنامه و نحوه‌ی کار آن، چند پروژه‌ی نمونه به برنامه داده شد و خروجی آن‌ها برای define, defineby بررسی شد و حالت‌های مختلفی که define, defineby رخ می‌داد یادداشت شد تا در بخش کدزنی به آن‌ها پرداخته شود. در ادامه به سراغ پروژه‌ی OpenUnderstand رفتیم.

برای توضیح روش پیشنهادی، ابتدا فایل main.py که فایل ابتدایی پروژه است بررسی می‌شود و عملکرد آن توضیح داده می‌شود. این فایل شامل توابع مختلفی است که در این بخش توابع مرتبط با پروژه‌ی گروه ما هم بررسی می‌شوند. سپس در ادامه سایر فایل‌هایی که توسط گروه ما به پروژه اضافه شده‌اند هم بررسی می‌شوند و نقش هریک از آن‌ها توضیح داده می‌شود.

## فایل main.py:

این فایل، فایل اصلی و ابتدایی پروژه است. در این فایل ابتدا در بخش main آن، یک شی از کلاس اصلی این فایل که کلاس Project است ساخته می‌شود. سپس دیتابیس مرتبط با این اجرا ساخته می‌شود و در ادامه توسط تابع getListOfFiles که به عنوان ورودی آدرس پروژه را دریافت می‌کند لیست آدرس تمام فایل‌های جاوای پروژه برگردانده می‌شود. سپس به ازای هر کدام از آدرس‌های فایل‌های جاوا، یکبار تابع getFileEntity را صدا می‌زنیم. این تابع به عنوان ورودی آدرس یک فایل جاوا را دریافت می‌کند و در ادامه Entity مربوط به آن فایل را تشخیص می‌دهد و آن را بر می‌گرداند.

در ادامه برای همین فایل تابع Parse را صدا می‌زنیم و با استفاده از این فایل درخت تجزیه‌ی این فایل ساخته می‌شود و برگردانده می‌شود.(توضیحات بیشتر مربوط به هریک از این توابع گفته شده در خطوط بالا در صفحات جلوتر آمده است. جهت دسترسی سریع‌تر می‌توانید به فهرست مطالب مراجعه کنید.)

پس از ساختن درخت مربوط به این فایل، باید listener های مربوط به این فایل یک به یک صدا زده‌شوند. در این بخش listener های مختلفی صدا زده می‌شوند که هریک وظیفه‌ی خاص خود را دارند و برای تکمیل اطلاعات یک بخش از دیتابیس کلی پروژه به کار برده می‌شوند. توضیح همه‌ی این موارد خارج از اهداف تسک محول شده به گروه 7 است. بنابراین در این بخش تنها به توضیح listener مربوط به بخش خودمان یعنی بخش define / defineIn اکتفا می‌کنیم. توضیح مربوط به listener ای که برای define / defineby ساخته شده است در بخش بعدی به صورت کامل آمده است.

### تابع getListOfFiles:

این تابع به عنوان ورودی، آدرس پروژه را دریافت می‌کند. یک لیست به نام allFiles ساخته می‌شود و در ادامه فایل‌های مختلف موجود در آدرس داده شده بررسی می‌شوند و آدرس تمامی فایل‌هایی که به پسوند .java ختم می‌شوند(که همان فایل‌هایی هستند که ما قصد بررسی آن‌ها را داریم) به لیست allFiles اضافه می‌شوند.

### تابه getFileEntity:

در این تابع ابتدا روی آدرسی که به عنوان ورودی به تابع داده می‌شود، یک سری اصلاحات انجام می‌شود تا به فرمت مورد نظرمان در بیاید. در ادامه این فایل را به صورت یک Entity ذخیره می‌کنیم به طوری که برای فیلد kind مقدار 1، برای فیلد name نام فایل که از روی آدرس آن به دست می‌آید، برای path آدرسی که در تابع به فرمت درست ساخته بودیم و در نهایت در contents هم محتوای فایل ذخیره می‌شود و Entity مورد نظر ساخته می‌شود و در نهایت به عنوان خروجی توسط تابع برگردانده می‌شود.

### تابع Parse:

هدف نهایی این تابع ساختن درخت تجزیه‌ی فایلی است که آدرس آن به عنوان ورودی به این تابع داده می‌شود. برای این مهم، لازم است مراحل زیر طی شود. این مراحل به ترتیب عبارتند از:

* باز کردن فایل با استفاده از آدرس فایل که به عنوان ورودی به تابع داده شده است.
* فراخوانی تابع JavaLexer مربوط به فایل ساخته شده و ذخیره خروجی آن در متغیر lexer
* دریافت توکن‌های مربوط به lexer ای که در بخش قبل برگردانده شده بود توسط تابع CommonTokenStream
* فراخوانی parser مربوط به زبان جاوا توسط تابع JavaParserLabeled
* و در نهایت ساخت درخت مربوطه توسط فراخوانی parser.compilationUnit()
* برگرداندن درخت ساخته شده به عنوان خروجی تابع

## Define\_defineby.py:

# ارزیابی:

در ابتدای امر، برای ارزیابی برنامه چند فایل کوچک را به Understand و OpenUnderstand

دادیم و نتایج آن‌ها را با بررسی یک به یک نتایج این دو برنامه، مقایسه کردیم. سپس بعد از اینکه ایرادات مربوط به این فایل‌های کوچک برطرف شدند، برای بررسی فایل‌ها و پروژه‌های بزرگتر نیاز بود تا یک فایل برای بررسی دقیق‌تر نتایج Understand ایجاد شود. برای این کار از api هایی که خود نرم‌افزار understand در اختیار ما قرار می‌داد استفاده کردیم.

ابتدا لازم است تا پروژه را به نرم افزار understand بدهیم تا فایلی با پسوند .udb توسط نرم افزار ساخته شود. سپس با استفاده از دستور زیر می‌توانیم به محتوای تولید شده توسط نرم‌افزار understand دسترسی داشته باشیم

try:

import understand

except ImportError:

print("Can not import understand")

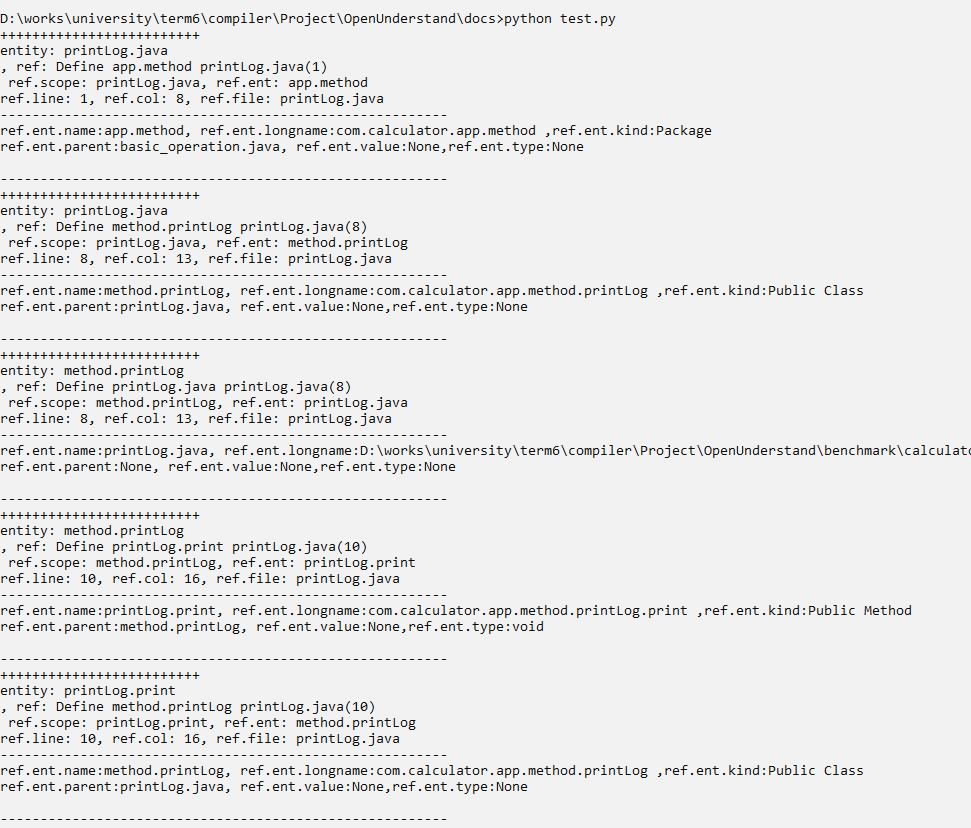
db=understand.open(r"D:\works\university\term6\compiler\Project\OpenUnderstand\benchmark\JSON\JSON.udb")

در ادامه می‌توانیم به نتایج به شیوه‌ی دلخواه دسترسی داشته باشیم. برای این منظور گروه ما نتایج را در کنسول پرینت می‌کرد. همچنین یک متغیر counter هم برای محاسبه‌ی تعداد فیلدهای ساخته شده تعریف شده است. محتوای کامل فایل بالا با نام test.py در زیر آورده شده است.

from pprint import pprint  
  
try:  
 import understand  
except ImportError:  
 print("Can not import understand")  
  
db = understand.open(r"D:\works\university\term6\compiler\Project\OpenUnderstand\benchmark\calculator\_app\calculator\_app\calculator\_app1.udb")  
  
counter = 0  
  
for ent in db.ents():  
 for ref in ent.refs():  
  
 if ref.kindname() == "Define":  
 counter += 1  
   
 print("+++++++++++++++++++++++++")  
  
 print(f"entity: {ent}\n, ref: {ref}\n ref.scope: {ref.scope()}, ref.ent: {ref.ent()}\n"  
 f"ref.line: {ref.line()}, ref.col: {ref.column()}, ref.file: {ref.file().name()}")  
 print("--------------------------------------------------------")  
 print(f"ref.ent.name:{ref.ent().name()}, ref.ent.longname:{ref.ent().longname()} ,ref.ent.kind:{ref.ent().kind()}\n"  
 f"ref.ent.parent:{ref.ent().parent()}, ref.ent.value:{ref.ent().value()},ref.ent.type:{ref.ent().type()}\n")  
 print("--------------------------------------------------------")  
  
print(counter)

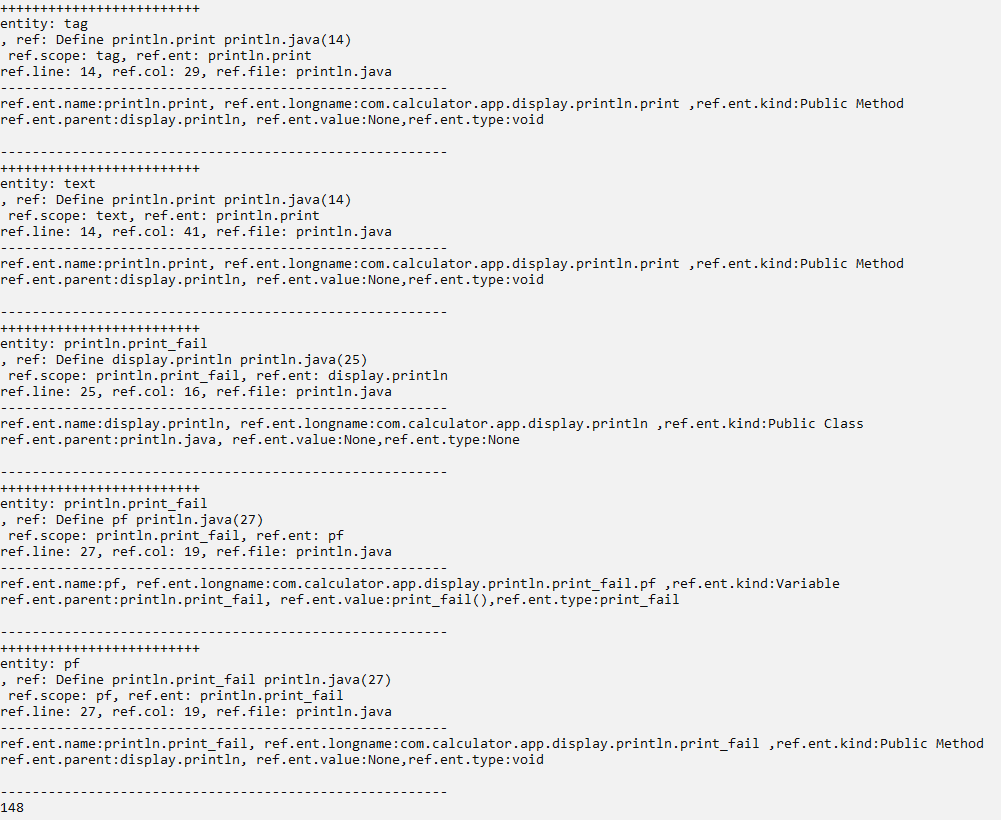
کد1 فایل test.py

در ادامه نتایج بدست آمده از اجرای برنامه‌ی calculator\_app توسط برنامه‌ی understand و همچنین توسط کد نوشته شده‌ی ما آورده شده است.



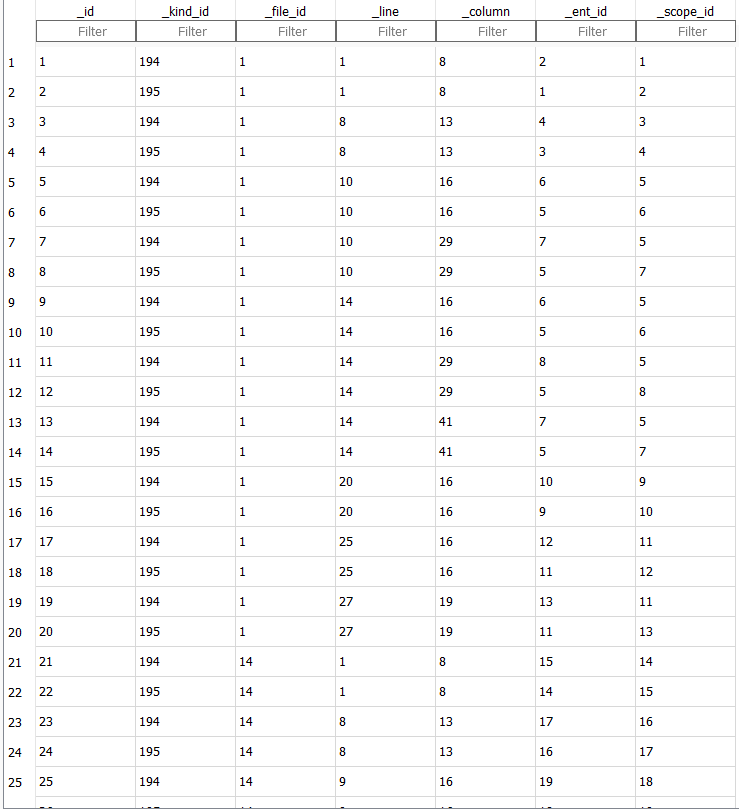
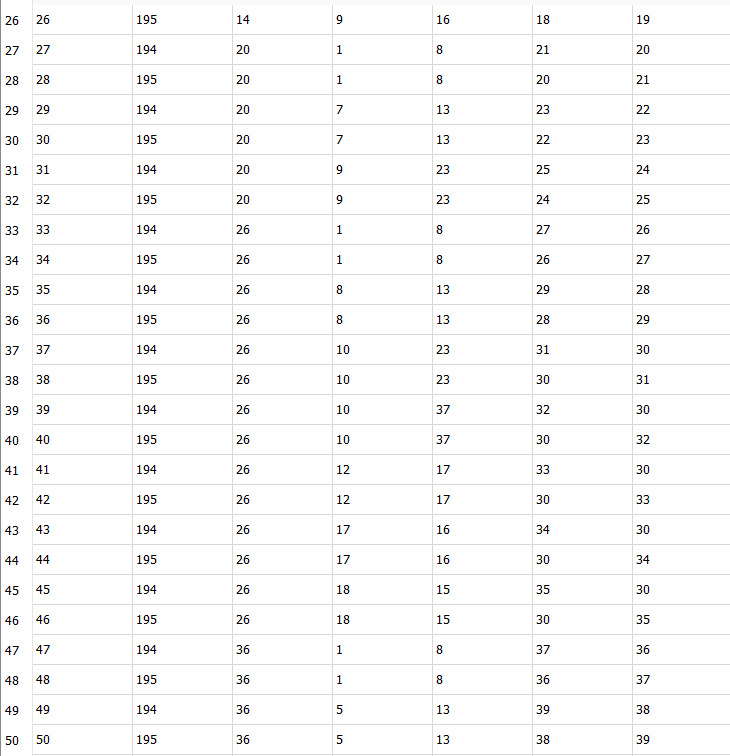
در این پروژه در مجموع 148 مورد define, defineby پیدا شد. به علت اینکه پوشش دادن تمام این موارد در فایل باعث زیاد شدن بیهوده‌ی حجم داکیومنت می‌شود و شاید از کیفیت لازم هم برخوردا نباشد، بنابراین عکس سایر موارد به طور کامل آورده نشده است و در عکس بعدی تعداد define, defineby ها آورده شده که بتوانیم با بخش بعدی که نتایج openUnderstand است بهتر مقایسه انجام بدهیم.

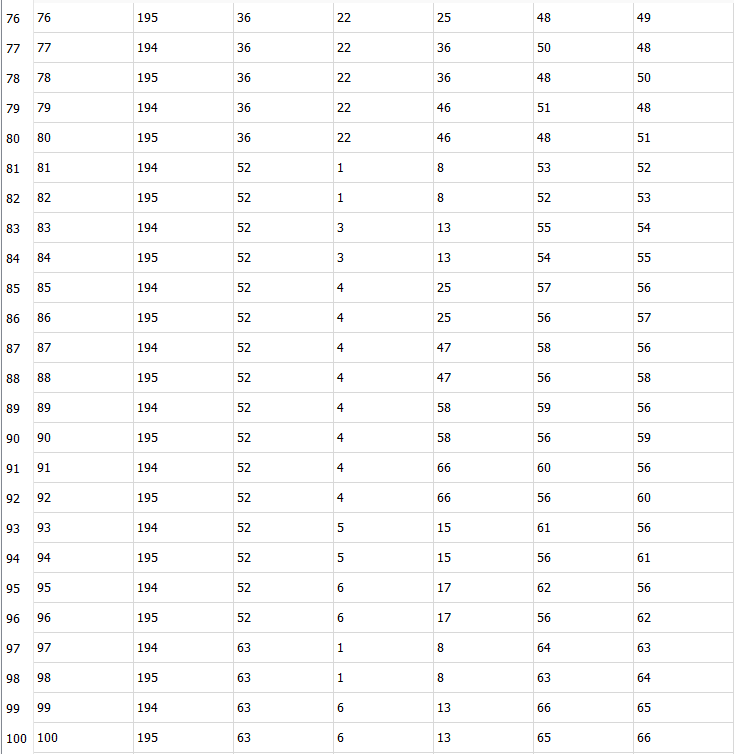
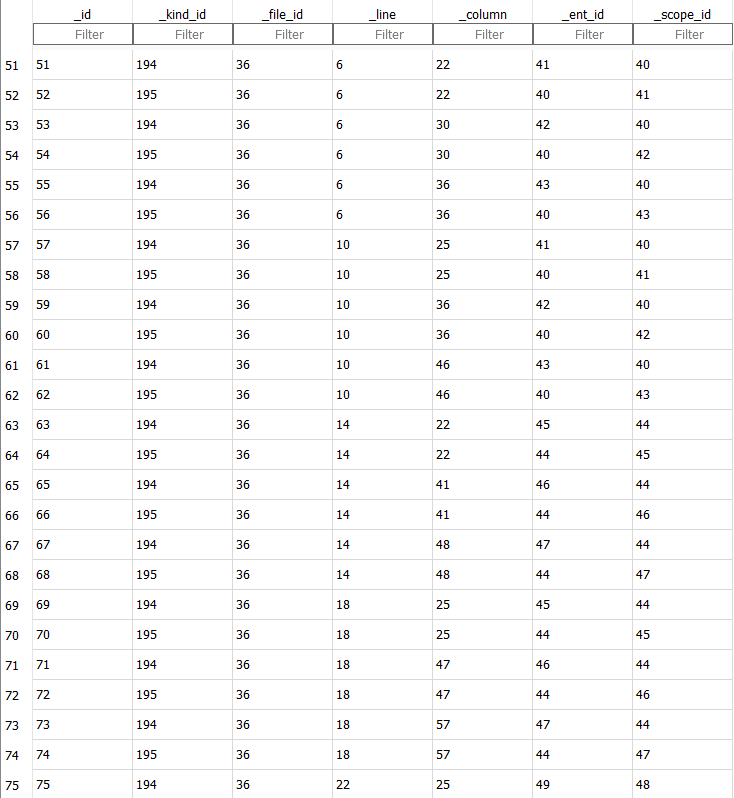
شکل 1نتایج مربوط به دیتابیس understand

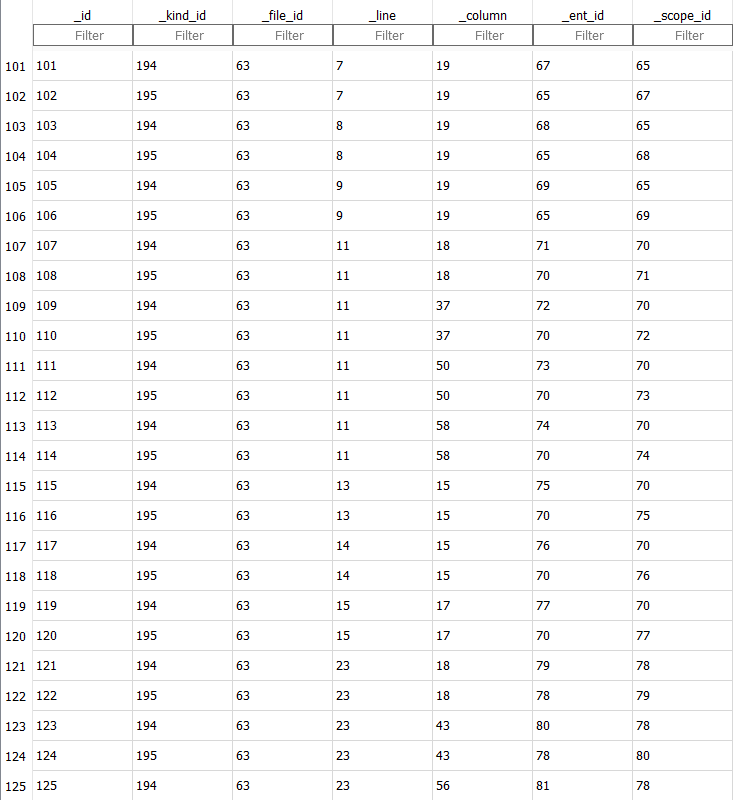
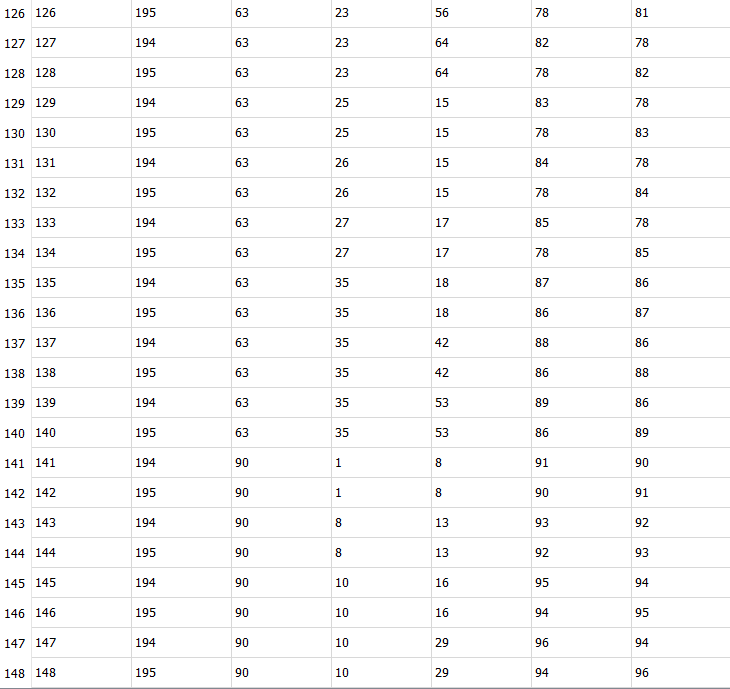


شکل 2 ادامه‌ی نتایح مربوط به دیتابیس Understand

در عکس‌های بعدی، نتایج مربوط به دیتابیس OpenUnderstand آورده شده است.





شکل 3 مجموعه شکل‌های نتایج دیتابیس OpenUnderstand

با بررسی نتایج متوجه می‌شویم که هم دیتابیس OpenUnderstand به ما 148 خروجی می‌دهد و هم خروجی فایل test.py. همچنین در صورتی که دقیق‌تر بررسی کنیم، هر یک از سطرهای این دو خروجی باهم مطابقت دارند. بنابراین برنامه به درستی کار می‌کند.

# مشکلات و چالش‌ها:

یکی از مشکلات بوجود آمده در هنگام تلاش برای ساخت فایل تست بود. در ابتدا ما فکر می‌‎کردیم که امکان دسترسی به دیتابیس Understand وجود ندارد و به همین علت زمان زیادی را صرف بررسی دستی نتایج برنامه‌ی خودمان با Understand کردیم.

پس از آنکه متوجه شدیم امکان بررسی بهتر نتایج وجود دارد، با بررسی api های سایت Understand متوجه شدیم که اگر آدرس فایل udb ساخته شده را به متد open از کتابخانه‌ی understand بدهیم، می‌توانیم به داده‌ها دسترسی داشته باشیم. اما در ابتدا با دادن پروژه‌های مختلف به برنامه‌ی understand فایل udb توسط برنامه ساخته نمیشد. در اینجا متوجه شدیم که باید نسخه‌های قدیمی‌تر از understand را نصب می‌کردیم و ورژن کنونی با کرک، همچنین فایلی نمی‌سازد.

پس از نصب دوباره‌ی نسخه‌های قدیمی‌تر برنامه، وقتی تلاش کردیم تا فایل test.py را اجرا کنیم، این بار برنامه متد open را شناسایی نمی‌کرد. پس از پرس و جو از سایر دانشجویان و همچنین دستیاران آموزشی، متوجه شدیم که بعد از نصب برنامه‌ی understand لازم بود تا یک فایل از پوشه‌ی Scripts به نام api\_install\_test.py را اجرا کنیم و پس از رفع ارورهای آن امکان شناسایی متد open وجود دارد.

پس از اجرای مراحل بالا، درنهایت توانستیم فایل test.py را اجرا کنیم.

# نتیجه‌گیری و کارهای آتی:

با استفاده از بخش پیاده‌سازی شده می‌توان رابطه define/definein را در کدهای جاوا پیدا و تجزیه و تحلیل کرد. در آینده می‌توان راهی برای مقایسه نتایج ابزار Open Understand و ابزار Understand ارائه داد تا بتوان به صورت اتوماتیک نتایج را مقایسه کرد و از صحت نتایج اطمینان یافت.