

smtfeatures

July 17, 2019

```
[80]: import pandas as pd
import seaborn as sns
from sklearn.preprocessing import MultiLabelBinarizer

filename='features.csv'

[99]: sorts = set()
df = pd.read_csv(filename, delimiter=';')
df['hasArrays'] = df['hasArrays'].replace({False: 0, True: 1})

for i,s in df.iterrows():
    sorts_array = s['sorts'].strip('[]').split(', ')
    df.at[i,'sorts'] = list(sort_array)
    sorts = sorts.union(set(sort_array))

feature_names = ['#functions', '#quantifiers', 'DAGSize', 'sorts', 'hasArrays', '
    ↪ solvertime']
df_filtered = df[feature_names]
sns.pairplot(df_filtered)

[99]: <seaborn.axisgrid.PairGrid at 0x7f92f2371fd0>
```

