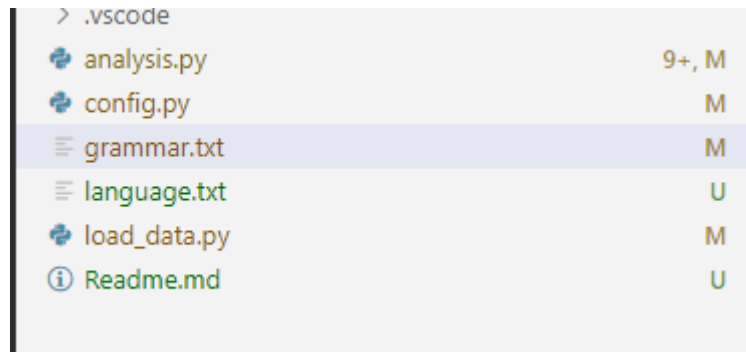


Motivation/Aim

模拟一个SLR (1) 语法分析器

Content Description

代码目录截图



输入输出结果截图

```

c:\python\python\Projects\compiler homework2\analysis.py'
-----读取CFGs成功-----
S->BBB
B->aB | b
-----FIRST函数-----
{'S': ['a', 'b'], 'S': ['a', 'b'], 'B': ['a', 'b']}
-----FOLLOW函数-----
{'S': ['#'], 'S': ['#'], 'B': ['a', 'b', '#']}
-----生成项目集-----
[{'key': 'S', 'value': 'S', 'dot': 0}, {'key': 'S', 'value': 'S', 'dot': 1}, {'key': 'S', 'value': 'BBB', 'dot': 0}, {'key': 'S', 'value': 'BBB', 'dot': 1}, {'key': 'S', 'value': 'BBB', 'dot': 2}, {'key': 'S', 'value': 'BBB', 'dot': 3}, {'key': 'B', 'value': 'aB', 'dot': 0}, {'key': 'B', 'value': 'aB', 'dot': 1}, {'key': 'B', 'value': 'aB', 'dot': 2}, {'key': 'B', 'value': 'b', 'dot': 0}, {'key': 'B', 'value': 'b', 'dot': 1}]
-----项目集规范族-----
[{'key': 'S', 'value': 'S', 'dot': 0}, {'key': 'S', 'value': 'BBB', 'dot': 0}, {'key': 'B', 'value': 'aB', 'dot': 0}, {'key': 'B', 'value': 'b', 'dot': 0}], [{'key': 'S', 'value': 'S', 'dot': 1}, {'key': 'S', 'value': 'BBB', 'dot': 1}, {'key': 'B', 'value': 'aB', 'dot': 1}, {'key': 'B', 'value': 'b', 'dot': 1}], [{'key': 'B', 'value': 'aB', 'dot': 0}, {'key': 'B', 'value': 'b', 'dot': 0}], [{'key': 'S', 'value': 'BBB', 'dot': 2}, {'key': 'B', 'value': 'aB', 'dot': 0}, {'key': 'B', 'value': 'b', 'dot': 0}], [{'key': 'B', 'value': 'aB', 'dot': 2}], [{'key': 'S', 'value': 'BBB', 'dot': 3}]]
-----ACTION-----
[{'a': 's3', 'b': 's4'}, {'#': 'acc'}, {'a': 's3', 'b': 's4'}, {'a': 's3', 'b': 's4'}, {'a': 'r3', 'b': 'r3', '#': 'r3'}, {'a': 's3', 'b': 's4'}, {'a': 'r2', 'b': 'r2', '#': 'r2'}, {'#': 'r1'}]]
-----GOTO-----
[{'S': 1, 'B': 2}, {}, {'B': 5}, {'B': 6}, {}, {'B': 7}, {}, {}]
-----读取language成功-----
baabaaab
-----开始分析-----
Step 1 ACTION(0, a)状态3入栈
Step 2 ACTION(3, b)状态4入栈
Step 3 r3:用B->b规约且GOTO(3, B)状态6入栈
Step 4 r2:用B->aB规约且GOTO(0, B)状态2入栈
Step 5 ACTION(2, a)状态3入栈
Step 6 ACTION(3, a)状态3入栈
Step 7 ACTION(3, b)状态4入栈
Step 8 r3:用B->b规约且GOTO(3, B)状态6入栈
Step 9 r2:用B->aB规约且GOTO(3, B)状态6入栈
Step 10 r2:用B->aB规约且GOTO(2, B)状态5入栈
Step 11 ACTION(5, a)状态3入栈
Step 12 ACTION(3, a)状态3入栈
Step 13 ACTION(3, a)状态3入栈
Step 14 ACTION(3, b)状态4入栈
Step 15 r3:用B->b规约且GOTO(3, B)状态6入栈
Step 16 r2:用B->aB规约且GOTO(3, B)状态6入栈
Step 17 r2:用B->aB规约且GOTO(3, B)状态6入栈
Step 18 r2:用B->aB规约且GOTO(5, B)状态7入栈
Step 19 r1:用S->BBB规约且GOTO(0, S)状态1入栈
acc: 分析成功
PS C:\Users\txcjh\Desktop\Projects\compiler homework2>

```

资源文件截图



Ideas/Methods

```
# 读入CFGs
grammar = load_data.load_CFG()

# 添加S'
grammar.insert(0, {'S\'': 'S'})

collectSign()      # 收集所有的终结符和非终结符
handleFirst()      # FIRST函数
handleFollow()      # FOLLOW函数
buildItemSet()      # 生成项目集
buildItemFamily()   # 生成项目集规范族
buildACTIONGOTO()   # 生成分析表

# 读入语言
lang = load_data.load_lang()
process()           # 分析语言
```

思路如上图

重要的数据结构描述

数据结构以及运行过程截图如下

```
grammar = []          # 文法
lang = ''             # 语言
itemSet = []          # 项目集
itemfamily = []       # 项目规范族

terminalSymbols = set() # 终结符
unterminalSymbols = set() # 非终结符
first = {}            # FIRST集
follow = {}           # FOLLOW集
goMap = []            # Go函数
ACTION = []
GOTO = []
```

```

✓ grammar: [{ 'S': 'S' }, { 'S': 'BBB' }, { 'B': 'aB' }, { 'B': 'b' } ]
  > 0: { 'S': 'S' }
  > 1: { 'S': 'BBB' }
  > 2: { 'B': 'aB' }
  > 3: { 'B': 'b' }
  __len__: 4
✓ itemSet: [{ 'dot': 0, 'key': 'S', 'value': 'S' }, { 'dot': 1, 'key': 'S', 'value': 'S' }, { 'dot': 0, 'key': 'S', 'value': 'BBB' }, { 'dot': 1, 'key': 'S', 'value': 'BBB' }, { 'dot': 2, 'key': 'S', 'value': 'BBB' }, { 'dot': 3, 'key': 'S', 'value': 'BBB' }, { 'dot': 0, 'key': 'B', 'value': 'aB' }, { 'dot': 1, 'key': 'B', 'value': 'aB' }, { 'dot': 2, 'key': 'B', 'value': 'aB' }, { 'dot': 0, 'key': 'B', 'value': 'b' }, { 'dot': 1, 'key': 'B', 'value': 'b' } ]
  __len__: 11
✓ itemfamily: [[{...}, {...}, {...}, {...}], [{...}, {...}, {...}, {...}], [...], [...]]
  > 0: [{ 'dot': 0, 'key': 'S', 'value': 'S' }, { 'dot': 1, 'key': 'S', 'value': 'S' }, { 'dot': 0, 'key': 'S', 'value': 'BBB' }, { 'dot': 1, 'key': 'S', 'value': 'BBB' }, { 'dot': 2, 'key': 'S', 'value': 'BBB' }, { 'dot': 3, 'key': 'S', 'value': 'BBB' }, { 'dot': 0, 'key': 'B', 'value': 'aB' }, { 'dot': 1, 'key': 'B', 'value': 'aB' }, { 'dot': 2, 'key': 'B', 'value': 'aB' }, { 'dot': 0, 'key': 'B', 'value': 'b' }, { 'dot': 1, 'key': 'B', 'value': 'b' } ]
  > 1: [{ 'dot': 1, 'key': 'S', 'value': 'S' }, { 'dot': 1, 'key': 'S', 'value': 'BBB' }, { 'dot': 1, 'key': 'B', 'value': 'aB' }, { 'dot': 1, 'key': 'B', 'value': 'b' } ]
  > 2: [{ 'dot': 1, 'key': 'S', 'value': 'BBB' }, { 'dot': 2, 'key': 'S', 'value': 'BBB' }, { 'dot': 2, 'key': 'B', 'value': 'aB' }, { 'dot': 2, 'key': 'B', 'value': 'b' } ]
  > 3: [{ 'dot': 1, 'key': 'B', 'value': 'aB' }, { 'dot': 2, 'key': 'B', 'value': 'aB' }, { 'dot': 3, 'key': 'S', 'value': 'BBB' } ]
  > 4: [{ 'dot': 1, 'key': 'B', 'value': 'b' }, { 'dot': 2, 'key': 'B', 'value': 'b' } ]
  > 5: [{ 'dot': 2, 'key': 'S', 'value': 'BBB' }, { 'dot': 3, 'key': 'S', 'value': 'BBB' } ]
  > 6: [{ 'dot': 2, 'key': 'B', 'value': 'aB' }, { 'dot': 3, 'key': 'B', 'value': 'aB' } ]
  > 7: [{ 'dot': 3, 'key': 'S', 'value': 'BBB' } ]
  > 8: []
  > 9: []
  > 10: []
  > 11: []
  > 12: []
  > 13: []
  > 14: []
  > 15: []
  > 16: []
  > 17: []
  > 18: []
  > 19: []
  > 20: []
  > 21: []
  > 22: []
  > 23: []
  > 24: []
  > 25: []
  > 26: []
  > 27: []
  > 28: []
  > 29: []
  > 30: []
  > 31: []
  > 32: []
  > 33: []
  > 34: []
  > 35: []
  > 36: []
  > 37: []
  > 38: []
  > 39: []
  > 40: []
  > 41: []
  > 42: []
  > 43: []
  > 44: []
  > 45: []
  > 46: []
  > 47: []
  > 48: []
  > 49: []
  > 50: []
  > 51: []
  > 52: []
  > 53: []
  > 54: []
  > 55: []
  > 56: []
  > 57: []
  > 58: []
  > 59: []
  > 60: []
  > 61: []
  > 62: []
  > 63: []
  > 64: []
  > 65: []
  > 66: []
  > 67: []
  > 68: []
  > 69: []
  > 70: []
  > 71: []
  > 72: []
  > 73: []
  > 74: []
  > 75: []
  > 76: []
  > 77: []
  > 78: []
  > 79: []
  > 80: []
  > 81: []
  > 82: []
  > 83: []
  > 84: []
  > 85: []
  > 86: []
  > 87: []
  > 88: []
  > 89: []
  > 90: []
  > 91: []
  > 92: []
  > 93: []
  > 94: []
  > 95: []
  > 96: []
  > 97: []
  > 98: []
  > 99: []
  > 100: []
  > 101: []
  > 102: []
  > 103: []
  > 104: []
  > 105: []
  > 106: []
  > 107: []
  > 108: []
  > 109: []
  > 110: []
  > 111: []
  > 112: []
  > 113: []
  > 114: []
  > 115: []
  > 116: []
  > 117: []
  > 118: []
  > 119: []
  > 120: []
  > 121: []
  > 122: []
  > 123: []
  > 124: []
  > 125: []
  > 126: []
  > 127: []
  > 128: []
  > 129: []
  > 130: []
  > 131: []
  > 132: []
  > 133: []
  > 134: []
  > 135: []
  > 136: []
  > 137: []
  > 138: []
  > 139: []
  > 140: []
  > 141: []
  > 142: []
  > 143: []
  > 144: []
  > 145: []
  > 146: []
  > 147: []
  > 148: []
  > 149: []
  > 150: []
  > 151: []
  > 152: []
  > 153: []
  > 154: []
  > 155: []
  > 156: []
  > 157: []
  > 158: []
  > 159: []
  > 160: []
  > 161: []
  > 162: []
  > 163: []
  > 164: []
  > 165: []
  > 166: []
  > 167: []
  > 168: []
  > 169: []
  > 170: []
  > 171: []
  > 172: []
  > 173: []
  > 174: []
  > 175: []
  > 176: []
  > 177: []
  > 178: []
  > 179: []
  > 180: []
  > 181: []
  > 182: []
  > 183: []
  > 184: []
  > 185: []
  > 186: []
  > 187: []
  > 188: []
  > 189: []
  > 190: []
  > 191: []
  > 192: []
  > 193: []
  > 194: []
  > 195: []
  > 196: []
  > 197: []
  > 198: []
  > 199: []
  > 200: []
  > 201: []
  > 202: []
  > 203: []
  > 204: []
  > 205: []
  > 206: []
  > 207: []
  > 208: []
  > 209: []
  > 210: []
  > 211: []
  > 212: []
  > 213: []
  > 214: []
  > 215: []
  > 216: []
  > 217: []
  > 218: []
  > 219: []
  > 220: []
  > 221: []
  > 222: []
  > 223: []
  > 224: []
  > 225: []
  > 226: []
  > 227: []
  > 228: []
  > 229: []
  > 230: []
  > 231: []
  > 232: []
  > 233: []
  > 234: []
  > 235: []
  > 236: []
  > 237: []
  > 238: []
  > 239: []
  > 240: []
  > 241: []
  > 242: []
  > 243: []
  > 244: []
  > 245: []
  > 246: []
  > 247: []
  > 248: []
  > 249: []
  > 250: []
  > 251: []
  > 252: []
  > 253: []
  > 254: []
  > 255: []
  > 256: []
  > 257: []
  > 258: []
  > 259: []
  > 260: []
  > 261: []
  > 262: []
  > 263: []
  > 264: []
  > 265: []
  > 266: []
  > 267: []
  > 268: []
  > 269: []
  > 270: []
  > 271: []
  > 272: []
  > 273: []
  > 274: []
  > 275: []
  > 276: []
  > 277: []
  > 278: []
  > 279: []
  > 280: []
  > 281: []
  > 282: []
  > 283: []
  > 284: []
  > 285: []
  > 286: []
  > 287: []
  > 288: []
  > 289: []
  > 290: []
  > 291: []
  > 292: []
```

```
> terminalSymbols: {'a', 'b'}
> unterminalSymbols: {'S', 'B'}
✓ first: {'B': ['a', 'b'], 'S': ['a', 'b'], "S'"
| > "S'": ['a', 'b']
| > 'B': ['a', 'b']
| > 'S': ['a', 'b']
|   __len__: 3
✓ follow: {'B': ['a', 'b', '#'], 'S': ['#'], "S'"
| > "S'": ['#']
| > 'B': ['a', 'b', '#']
| > 'S': ['#']
|   __len__: 3
✓ goMap: [[-1, 'S', 'B', 'a', 'b', -1, -1, -1, -1, -1],
| > 00: [-1, 'S', 'B', 'a', 'b', -1, -1, -1, -1, -1],
| > 01: [-1, -1, -1, -1, -1, -1, -1, -1, -1, -1],
| > 02: [-1, -1, -1, 'a', 'b', 'B', -1, -1, -1, -1],
| > 03: [-1, -1, -1, 'a', 'b', -1, 'B', -1, -1, -1],
| > 04: [-1, -1, -1, -1, -1, -1, -1, -1, -1, -1],
| > 05: [-1, -1, -1, 'a', 'b', -1, -1, 'B', -1, -1],
| > 06: [-1, -1, -1, -1, -1, -1, -1, -1, -1, -1],
| > 07: [-1, -1, -1, -1, -1, -1, -1, -1, -1, -1],
|   __len__: 8
✓ ACTION: [{'a': 's3', 'b': 's4'}, {'#': 'acc'}, ...]
| > 0: {'a': 's3', 'b': 's4'}
| > 1: {'#': 'acc'}
| > 2: {'a': 's3', 'b': 's4'}
| > 3: {'a': 's3', 'b': 's4'}
| > 4: {'#': 'r3', 'a': 'r3', 'b': 'r3'}
| > 5: {'a': 's3', 'b': 's4'}
| > 6: {'#': 'r2', 'a': 'r2', 'b': 'r2'}
| > 7: {'#': 'r1'}
|   __len__: 8
✓ GOTO: [{'B': 2, 'S': 1}, {}, {'B': 5}, {'B': 6...}]
| > 0: {'B': 2, 'S': 1}
| > 1: {}
| > 2: {'B': 5}
| > 3: {'B': 6}
| > 4: {}
| > 5: {'B': 7}
| > 6: {}
| > 7: {}
|   len : 8
```

核心算法

- FIRST函数

- FOLLOW函数
- 求闭包
- 生成项目集
- 生成项目集规范族
- 生成ACTION、GOTO分析表
- 根据SLR (1) 分析表分析语法

Example

在Content Description部分已给出

Problems occurred and related solutions

代码冗余，难以debug——拆分函数后analyse()函数简洁了许多

Your feelings and comments

时间不够，学期课程比较多，期末之前搞完这个实属紧张。。应效仿其他几门课，最后一门作业留到期末完之后交。。^_^