



華東師範大學
EAST CHINA NORMAL UNIVERSITY

学期汇报

李楷达

51205901093



華東師範大學
EAST CHINA NORMAL UNIVERSITY

项目

上下文实体对齐



NER -> 上下文实体对齐

- 4月7日，格力地产股份有限公司(下称“**格力地产**”)公告称，收到**中国证券登记结算有限责任公司上海分公司**通知，公司控股股东**珠海投资控股有限公司**(下称“**珠海投资**”)所持有公司的股份被冻结。根据公告，本次股权被冻结股份数量8.47亿股，占**珠海投资**所持股份比例100%，占**格力地产**股本41.11%，冻结期限为三年。冻结申请人为**广州市玄元投资管理有限公司**(下称“**玄元投资**”)，冻结原因系合同纠纷。财报显示，截至2019年9月30日，**玄元投资**为格力地产第三大股东。
- 全称 – 简称



思路

一个不太强的假设——对一篇新闻中出现的某个公司简称，至少满足下列条件之一：

- ①该公司属于大公司，较为知名；
- ②该公司的全称在该篇新闻中出现过；
- ③该篇新闻规范度较差，价值较小，忽略其造成的影响有限。

基于上述假设，对新闻中出现的公司简称做如下处理：

- ①尝试在上下文中匹配该简称对应的全称；
- ②若匹配失败，推测该公司属于知名公司，在字典中进行查找。



匹配逻辑

- 公司全称组成：行政区划+字号+行业或经营特点+组织形式
- 广州市玄元投资管理有限公司
- 关键部分：字号+行业，其中前>后
- '安徽省皖能股份有限公司'-'皖能电力','中水集团远洋股份有限公司'-'中水渔业'
- 准确率：98.39%



華東師範大學
EAST CHINA NORMAL UNIVERSITY

比賽

SemEval Task9

表格事实验证



Caption: 2019年世界主要国家经济状况

Header:

| 国家 text | GDP (万亿美元) real | 对外贸易额 (万亿美元) real |
|---------|-----------------|-------------------|
|---------|-----------------|-------------------|

Cells:

| | | |
|----|-------|------|
| 美国 | 21.37 | 5.1 |
| 中国 | 14.34 | 3.95 |
| 德国 | 3.85 | 2.69 |
| 日本 | 5.08 | 1.35 |

Query: 2019年美国国内生产总值比中国高多少?



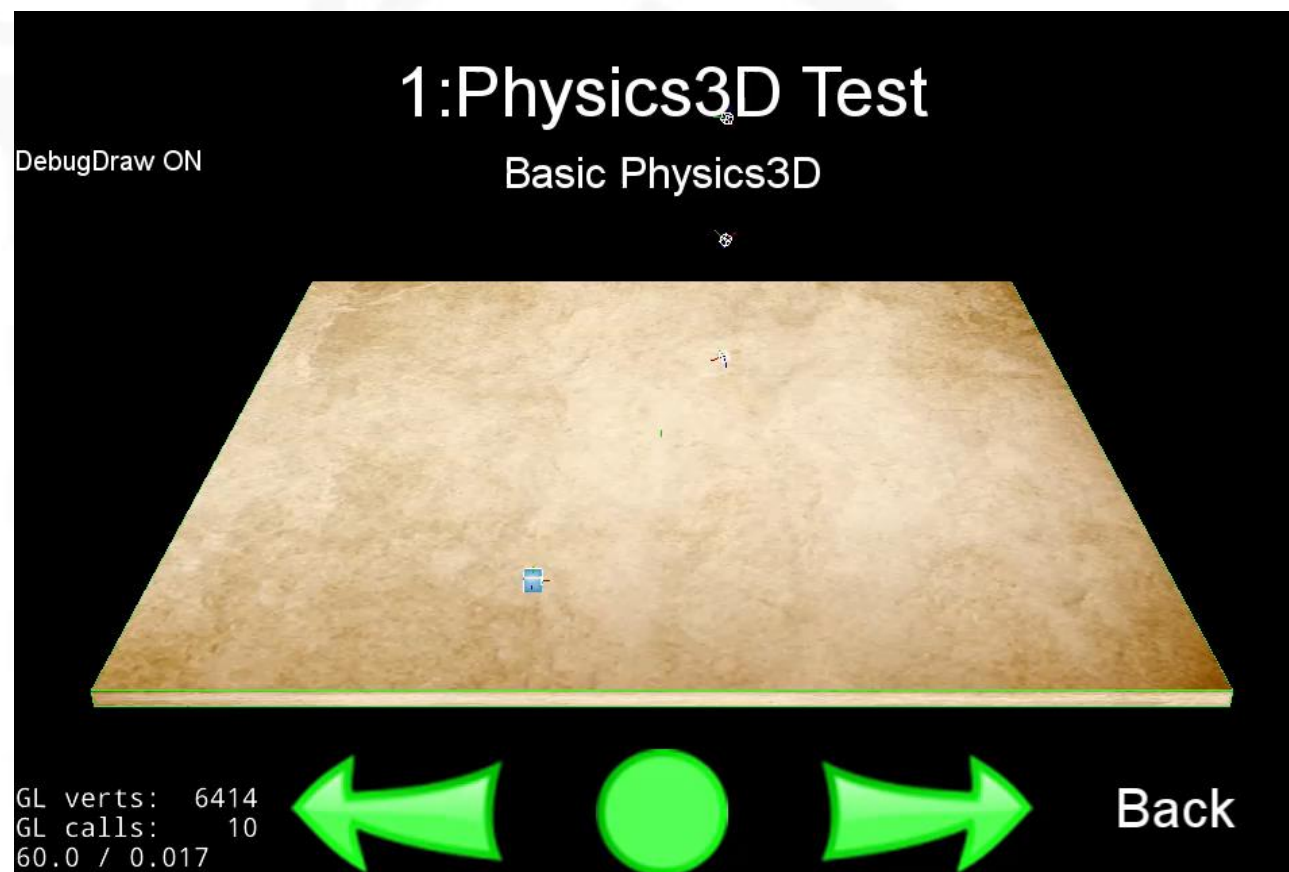
華東師範大學
EAST CHINA NORMAL UNIVERSITY

未来工作

常识获取



可行性分析



在游戏中获取物理常识



- ①原思路存在的问题：1.设置的场景基于已经想到的常识2.为什么不直接从游戏源码中寻找
- ②短期计划：
 - 阅读论文，学习常识获取基础知识（ddl：4.15，阅读至少10篇相关论文）
 - 改进从游戏中获取常识的方法，解决现存问题（ddl：4.30）
 - 探索获取常识的新思路