

H1 开发功能记录

H2 2、功能实现

H3 2.1 实现英文查询功能

```
收到的消息 : ["数字",1,"根据",1,"根据",1,"路",1]  
timeout  
发送: timeout  
收到的消息 : ["tumour",3,"times",3,"without",3,"time",3]  
time  
发送: time  
收到的消息 : ["lime",1,"tide",1,"times",1,"time",0]  
out  
发送: out  
收到的消息 : ["put",1,"our",1,"but",1,"out",0]  
█
```

H3 2.2、实现了缓存系统

- 能够进行数据的缓存、轮询更新缓存，缓存数据读取，数据持久化，能够进行LRU缓存淘汰

```
数据写入缓存  
hello["hallo",1,"hell",1,"hullo",1,"hello",0]  
消息发送完毕.....处理前收到的大小6  
处理后收到的大小5  
morey  
172.17.42.252:8989 --> 172.17.42.252:34546>> 发送过来: morey  
2019-06-21 14:40:02,567: INFO mylogger : [SpellcorrectServer.  
cpp:onMessage:50]info, H 发送了: , "мЖ  
获取0:号线程缓存  
数据写入缓存  
morey["morley",1,"morel",1,"money",1,"more",1]  
消息发送完毕.....>> poll timeout
```

```

数据写入缓存
ebokk["evoke",2,"ebooks",2,"book",2,"ebook",1]
进行了淘汰
列表大小为 = 10
capacity大小为 = 10
消息发送完毕.....定时更新所有Cache
更新单个缓存数据
更新单个缓存数据
更新单个缓存数据
更新单个缓存数据
更新单个缓存数据
更新单个缓存数据
更新单个缓存数据
处理前收到的大小6
处理后收到的大小5
haeld
172.17.42.252:8989 --> 172.17.42.252:34548>> 发送过来: h
2019-06-21 14:40:35.283: INFO mylogger : [SpellcorrectS
AutoRelease()
缓存大小 = 10
读取缓存内容
缓存文件名 = /root/project/oP/online/data/cache.dat
打印文件缓存内容
10
10
key = ebokk
value = ["evoke",2,"ebooks",2,"book",2,"ebook",1]
key = book
value = ["ebook",1,"look",1,"took",1,"book",0]
key = ling
value = ["lying",1,"king",1,"line",1,"long",1]
key = rean
value = ["mean",1,"real",1,"read",1,"ran",1]
key = mdsa
value = ["miss",2,"idea",2,"most",2,"must",2]
key = msg
value = ["sg",1,"meg",1,"mug",1,"ms",1]
key = hullo
value = ["follow",2,"fellow",2,"hullo",1,"hollow",1]
key = look
value = ["looks",1,"book",1,"took",1,"look",0]

```

H3 2.3 实现中文分词和索引的建立

- 词典

896	网步走	1
897	两段	1
898	两汉	1
899	两派	2
900	两点	5
901	两用	1
902	两百	2
903	两百多	2
904	两百多年	1
905	两种	29
906	两立	2
907	两章	1
908	两端	3
909	两篇	3

- 索引

```

2753 黄 12306 18417 18418 18419 18420 18421 18422 18423 18424
2754 黏 18437
2755 黑 11724 18438 18439 18440 18441 18442 18443 18444 18445
2756 默 2248 7604 7605 11603 12234 16767 18448 18451 18452 184
2757 黛 18457
2758 黜 18458
2759 黠 12550 18459
2760 鼎 18460 18461 18462 18463
2761 鼓 5047 5048 9996 9997 10635 17421 17462 18464 18465 1846
2762 鼻 13074 13845 18470 18471 18472 18473 18474
2763 齐 2535 4373 7587 9717 9718 13335 14025 18475 18476 18477
2764 齿 12439 17858 18484
2765 龙 3224 4176 4598 7150 9771 11612 12582 12583 12627 14668
2766 癸 18489

```

H3 2.4 实现中文查询

```

王道在线
发送: 王道在线
收到的消息 : ["王",3,"道",3,"存在",3,"在",3]
在线
发送: 在线
收到的消息 : ["在内",1,"曲线",1,"在于",1,"在",1]
更新
发送: 更新
收到的消息 : ["更多",1,"新",1,"更",1,"更新",0]
数据
发送: 数据
收到的消息 : ["数学",1,"依据",1,"根据",1,"据",1]

```

H2 3 总结

- 在写的过程中遇到了很多弯路，比如说在选择数据结构上面不够合理，导致算法复杂度很大，建立索引的时候特别慢。
- 在获取配置文件的时候也设计的不合理，循环写很麻烦。
- 功能上面大致实现了中文查询和英文查询，以及缓存和缓存淘汰。
- 缓存和缓存淘汰用的是最基础的LRU，存在问题是list的大小和Map大小不能同步，就是list大小可以控制，但是Map会一直变大。不能进行淘汰。
- 客户端的编写过于简单，没有加上输入检查，应该在客户端也加入输入检查，服务器也判断一下输入是否合法，目前如果发送了空字符，服务器会发生段错误，需要检查，客户端接收服务器的显示也没有进行优化，知识单独的收到json文件直接打印而且。
- 反正协议传输上面设计的不够合理。
- 服务器写分词，分词在客户端，如果，没有给服务器添加上。后期优化需要添加
- 分词设置的不够合理。没有使用jiaba最好的那个分词模式，后期可以加上优化分词
- 词典树功能没有完成。
- 没有使用内存数据库，有时间可以直接扩展到缓存数据库中