

原创

基于CYK+PCFG的短语结构句法分析

2018-11-25 23:08:05

不愿透露姓名的广外男子

阅读量 736

☆ 收藏

更多

版权声明：本文为博主原创文章，遵循 [CC 4.0 BY-SA](#) 版权协议，转载请附上原文出处链接和本声明。
本文链接：<https://blog.csdn.net/Chase1998/article/details/84504191>

1 概述

句法分析是自然语言处理中的关键技术之一。其基本任务是确定句子的句法结构(syntatic structure)或句子中词之间的依存关系。形式语法理论的目标是建立一种数学模型(形式语言)来刻画自然语言。句法分析根据形式语法的不同可以分为基于短语结构语法的句法分析、基于依存关系语法的依存句法分析。本文介绍使用CYK(一种自底向上的动态规划算法)+PCFG(概率上下文无关文法)的基于短语结构语法的句法分析。

2 理论描述

2.1 CFG (上下文无关文法)

形式语法 $G = \langle V_N, V_T, R, S \rangle$

- V_N 非终端语符集(辅助词汇的有限集合) , 如 $\{S, NP, VP, ART, N, V\}$
- V_T 终端语符集(基本词汇的有限集合) , 如 $\{the, a, boy, sees, cat, dirty\}$
- R 一组由有限个重写规则组成的规则集, 如 $\{ VP \rightarrow V NP, ART \rightarrow the|a, \dots \}$
- S 表示起始符

句子：由语法 G_0 从起始符 S 可派生出来的终端语符列构成由 G_0 生成的句子。

语言：所有由语法 G_0 从起始符 S 可派生出来的终端语符列构成 G_0 生成的语言。

2.2 PCFG (概率上下文无关文法)

PGFG在CFG的基础上引入了P, 加上了每个规则的概率。

PCFG中定义一棵句法树的概率为所有用到的规则概率的乘积 , 一般来说, 概率值大的更可能是正确的句法树。

我们可以利用句子S的所有可能的句法分析树来统计句子的概率:

$$P(S) = \sum_{t=1}^T P(S, T)$$

2.3 CNF (Chomsky Normal Form)

如果一个上下文无关文法的每个产生式的形式为：

$$A \rightarrow BC \text{ 或 } A \rightarrow a,$$

即规则的右部或者是两个非终结符或者是一个终结符，则它是具有Chomsky范式的CFG。

任何CFG都可以转变成一个弱等价的 Chomsky范式语法。

2.4 CYK算法

给定一个句子 $s = w_1, w_2, \dots, w_n$, 和一个上下文无关文法PCFG, $G=(T, N, S, R, P)$;

定义一个跨越单词 i 到 j 的概率最大的语法成分 π :

$$\pi(i, j, X) \ (i, j \in 1 \dots n, X \in N),$$

目标是找到一个属于 $\pi[1, n, S]$ 的所有树中概率最大的那棵。

CYK算法用于PCFG下的句法分析：

- 基本定义：for all $i=1, \dots, n, X \in N$
$$\pi(i, i, X) = q(X \rightarrow w_i) \text{ (if } X \rightarrow w_i \text{ 没有出现在语法中, 则定义 } q(X \rightarrow w_i) = 0)$$
- 递归定义：for all $i=1, \dots, n, j=(i+1), \dots, n, X \in N$

👍 4

🔗

💬 1

📖

☆

📱

⏪

⏩

👑 赏

🏠 脉

多功能双肩包

活出质感

限时折扣

¥49

关闭

🔊

举报

$$\pi(i, j, X) = \max(q(X \rightarrow YZ) \times \pi(i, k, Y) \times \pi(k + 1, j, Z)) \quad (i \leq k \leq j - 1)$$

3 样例描述

给定以下PCFG，实现句子“fish people fish tanks”最可能的统计句法树。

S → NP VP	0.9	N → <i>people</i>	0.5
S → VP	0.1	N → <i>fish</i>	0.5
VP → V NP	0.5	N → <i>tanks</i>	0.2
VP → V	0.1	N → <i>rods</i>	0.1
VP → V @VP_V	0.3	V → <i>people</i>	0.1
VP → V PP	0.1	V → <i>fish</i>	0.6
@VP_V → NP PP	1.0	V → <i>tanks</i>	0.3
NP → NP NP	0.1	P → <i>with</i>	1.0
NP → NP PP	0.2		
NP → N	0.7		
PP → P NP	1.0		



第一步：构造4*4矩阵。

根据CYK算法，每格Cell[i, j]包含了跨越单词i+1, j+1的所有语法成分（实际计算中下标是从0开始的）。

以Cell[1, 3]为例，Cell[1, 3]格中的成分分别为：(1,1)和(2,3)组成，(1,2)和(3,3)组成，包含了 people fish tanks 所有语法成分。

Fish	people	fish	tanks
score[0][0]	score[0][1]	score[0][2]	score[0][3]
	score[1][1]	score[1][2]	score[1][3]
		score[2][2]	score[2][3]
			score[3][3]

第二步：处理叶子节点中的单词。

Fish	people	fish	tanks
score[0][0] N → fish 0.2 V → fish 0.6	score[0][1]	score[0][2] 4	score[0][3]
	score[1][1] N → people 0.5 V → people 0.1	score[1][2] 1	score[1][3]
		score[2][2] N → fish 0.2 V → fish 0.6	score[2][3]
		 	score[3][3] N → tanks 0.2 V → tanks 0.3 https://blog.csdn.net/Chase1998

第三步：根据叶子节点中单词的词性递归地找一元匹配规则。

以[0][0]中NP → N 0.14为例， $0.14 = 0.7$ (规则集中NP→N) * 0.2 ([0][0]中的N→fish)。

多功能双肩包
活出质感

限时折扣
¥49

关闭



举报

→ VP 0.003

多功能双肩包
活出质感

限时折扣
¥49

关闭

关闭



[illegible]

[1][3]同理。

多功能双肩包
活出质感
限时折扣
¥49
关闭



Fish	people	fish	tanks
score[0][0] N → fish 0.2 V → fish 0.6 NP → N 0.14 VP → V 0.06 S → VP 0.006	score[0][1] NP → NP NP 0.0049 VP → V NP 0.105 S → NP VP 0.00126 S → VP 0.0105	score[0][2] NP → NP NP 6.86e-5 NP → NP NP 6.86e-5 VP → V NP 0.00147 S → NP VP 8.82e-4 S → NP VP 3.087e-5 S → VP 1.47e-4	score[0][3] <div>4</div> <div>1</div>
	score[1][1] N → people 0.5 V → people 0.1 NP → N 0.35 VP → V 0.01 S → VP 0.001	score[1][2] NP → NP NP 0.0049 VP → V NP 0.007 S → NP VP 0.0189 S → VP 0.0007	score[1][3] NP → NP NP 6.86e-5 NP → NP NP 6.86e-5 VP → V NP 9.8e-6 S → NP VP 0.01 S → NP VP 1.32e-5 S → VP 9.8e-6 <div>赏</div>
		score[2][2] N → fish 0.2 V → fish 0.6 NP → N 0.14 VP → V 0.06 S → VP 0.006	score[2][3] <div>脉</div> NP → NP NP 0.00150 VP → V NP 0.042 S → NP VP 0.00378 S → VP 0.0042
			score[3][3] N → tanks 0.2 V → tanks 0.3 NP → N 0.14 VP → V 0.03 S → VP 0.003

第六步：处理根节点。

根据PCYK算法我们分别从[0][0]+[1][3]，[0][1]+[2][3]，[0][2]+[3][3]找对应的匹配规则，再对结果找对应[0][3]的一元规则，这样便可覆盖率所有单
当同一个非终端语符有多条规则时，我们仅保留其最大项。

多功能双肩包
活出质感

限时折扣
¥49

关闭



Fish	people	fish	tanks
score[0][0] N → fish 0.2 V → fish 0.6 NP → N 0.14 VP → V 0.06 S → VP 0.006	score[0][1] NP → NP NP 0.0049 VP → V NP 0.105 S → NP VP 0.00126 S → VP 0.0105	score[0][2] NP→NP NP 6.86e-5 NP→NP NP 6.86e-5 VP → V NP 0.00147 S → NP VP 8.82e-4 S → NP VP 3.087e-5 S → VP 1.47e-4	score[0][3] NP→NP NP 9.60e-6 NP→NP NP 9.60e-6 NP→NP NP 9.60e-6 VP→V NP 2.058e-1 S → NP VP 1.234e-1 S→NP VP 1.852e-1 S → NP VP 1.852e-1 S → VP 2.058e-1
	score[1][1] N → people 0.5 V → people 0.1 NP → N 0.35 VP → V 0.01 S → VP 0.001	score[1][2] NP → NP NP 0.0049 VP → V NP 0.007 S → NP VP 0.0189 S → VP 0.0007	score[1][3] NP→NP NP 6.86e-5 NP→NP NP 6.86e-5 VP → V NP 9.8e-6 S → NP VP 0.011 S → NP VP 1.32e-1 S → VP 9.8e-6
		score[2][2] N → fish 0.2 V → fish 0.6 NP → N 0.14 VP → V 0.06 S → VP 0.006	score[2][3] NP→NP NP 0.00196 VP → V NP 0.042 S → NP VP 0.00378 S → VP 0.0042
			score[3][3] N → tanks 0.2 V → tanks 0.3 NP → N 0.14 VP → V 0.03 S → VP 0.003

第七步：回溯。

从根节点的开始标志S出发，按照之前保留的路径找出概率最大句法树。

多功能双肩包
活出质感

限时折扣
¥49

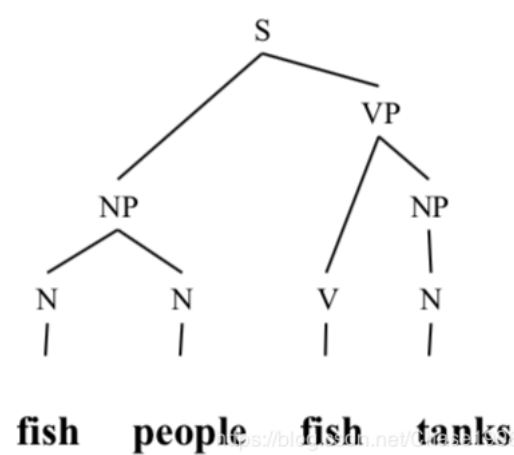
关闭

🔊

举报

Fish	people	fish	tanks
score[0][0] N → fish 0.2 V → fish 0.6 NP → N 0.14 VP → V 0.06 S → VP 0.006	score[0][1] NP → NP NP 0.0049 VP → V NP 0.105 S → VP 0.0105	score[0][2] NP → NP NP 6.86e-5 VP → V NP 0.001 S → NP VP 8.82e-5	score[0][3] NP → NP NP 9.604e-5 VP → V NP 2.058e-5 S → NP VP 1.8522e-5
	score[1][1] N → people 0.5 V → people 0.1 NP → N 0.35 VP → V 0.01 S → VP 0.001	score[1][2] NP → NP NP 0.0007 VP → V NP 0.007 S → NP VP 0.018	score[1][3] NP → NP NP 6.86e-5 VP → V NP 9.8e-5 S → NP VP 0.01323
		score[2][2] N → fish 0.2 V → fish 0.6 NP → N 0.14 VP → V 0.06 S → VP 0.006	score[2][3] NP → NP NP 0.0019 VP → V NP 0.042 S → VP 0.0042
			score[3][3] N → tanks 0.2 V → tanks 0.3 NP → N 0.14 VP → V 0.03 S → VP 0.003

最后，PCFG生成的概率最大句法树结果如图：




4 总结

普通的回溯法（backtracking）在最坏的情况下需要指数时间才能解决这样的问题，而CYK算法只需要多项式时间就够了。CYK算法采用了动态规划的思想。

- 有 $(\frac{n(n+1)}{2}) = O(n_2)$ 格子
- 有 $O(n)$ 个可能的分裂点
- 总的时间复杂度为 $O(n_3)$


上述例子给出的PCFG并不是严格意义上的乔姆斯基范式，如果是CNF的话可以不用对本身格子里的语法再进行规则匹配，所以在实际应用中应首先进行转换，采用乔姆斯基范式来进行CYK算法。

文章最后发布于: 2018-11-2

 **用户画像建立用户洞察**
用户画像数据
5803阅读

 想对作者说点什么



 **hztj2005** 4个月前 #1楼

原理解释得清楚。在Jalaj Thanaki《Python Natural Language Processing》第126页也看到这个表。但该书没有详述S有两条规则符合时，保留概率较大的这一点 VP 0.00126、S ->VP 0.0105中保留后一条。（0,1）格中S ->VP 0.0105 的概率 = 规则S ->VP的概率0.1 * 本格上一条 VP->V NP的概率0.105，这是递归实现的。

句法分析算法PCFG学习

阅读数 1433

在做自然语言处理的时候，我们要用到句法结构分析。找到词与词之间的关系并且分析语义。http://blog.csdn.net/... 博文 来自： baidu_15113429...

基于概率最大化的中文分词算法 Python

阅读数 1225

1 概述汉语自动分词是把没有明显分界标志的字串切分为词串。包括：标点符号、数字、数学符号、各种 标记、人名... 博文 来自： Chase1998的博客

CYK算法简介与实现

阅读数 4298

摘要CYK算法是一个基于“动态规划”算法设计思想，用于测试串w对于一个上下文无关文法L的成员性的一个算法。... 博文 来自： ssaalkjhgf的博客

NLP底层技术之句法分析

阅读数 2万+

句法分析是自然语言处理（naturallanguageprocessing,NLP）中的关键底层技术之一，其基本任务是确定句子的句... 博文 来自： 王琨的博客

 **cdn及cdn加速原理**
海外cdn加速专线
9835阅读

从入门到精通，Java学习路线导航（附学习资源）

阅读数 8802

引言最近也有很多人来向我“请教”，他们大都是一些刚入门的新手，还不了解这个行业，也不知道从何学起，开始的... 博文 来自： java_sha的博客

学会了这些技术，你离BAT大厂不远了

阅读数 2803

每一个程序员都有一个梦想，梦想着能够进入阿里、腾讯、字节跳动、百度等一线互联网公司，由于身边的环境等原... 博文

程序员真是太太太太有趣了！！！！

阅读数 4万+

点击上方“程序员小明”，选择“星标”今晚可以不加班！网络上虽然已经有了很多关于程序员的话题，但大部分人... 博文 来自： 程序员小明

别再翻了，面试二叉树看这 11 个就够了~

阅读数 7万+

写在前边数据结构与算法：不知道你有没有这种困惑，虽然刷了很多算法题，当我去面试的时候，面试官让你手写一... 博文 来自： 一个不甘平凡的码农

60 个让程序员崩溃的瞬间，哈哈哈哈哈哈哈哈

阅读数 3569

阅读本文大概需要 2.3333 分钟。前方高能，每一个程序员看完，你不笑死个人，你来找我，我自己看了好几遍，反... 博文 来自： stormzhang的专栏

句法结构分析:基于PCFG的基本分析方法 - weixin_305220..._CSDN博客

Java环境配置-知其所以然 - Chase1998的博客 - CSDN博客

反转！“只问了1个框架，就给了35K的Python岗”
学Python的程序员建议收藏！

史上最详细的IDEA优雅整合Maven+SSM框架（详细思路+附带源码）

网上很多整合SSM博客文章并不能让初探ssm的同学思路完全的清晰，可以试着关掉整合教程，摇两下头骨，哈一大...

NLP底层技术之句法分析** - 王琨的博客 - CSDN博客**

依存分析 CKY算法 - 扬帆起航 - CSDN博客

对计算机专业来说学历真的重要吗？

我本科学校是渣渣二本，研究生学校是985，现在毕业五年，校招笔试、面试，社招面试参加了两年了，就我个人的...



勿在浮沙筑高台LS
689篇文章
排名:2000+
[关注](#)



ssaalkjhgf
1篇文章
排名:千里之外
[关注](#)



Vico_Men
30篇文章
排名:千里之外
[赏](#)

斐波那契数列引发的思考 Kotlin - Chase1998的博客 - CSDN博客

cyk追捕神系列二 - jiang_1603的博客 - CSDN博客

接班马云的为何是张勇？

上海人、职业经理人、CFO背景，集齐马云三大不喜欢的张勇怎么就成了阿里接班人？作者|王琳本文经授权转载自燃...

我花了一夜用数据结构**给女朋友写个H5走迷宫游戏**

起因又到深夜了，我按照以往在csdn和公众号写着数据结构！这占用了我大量的时间！我的超越妹妹严重缺乏陪伴而...

自然语言处理:基于PCFG**的中文**句法分析**样例 - 程裕强的..._CSDN博客**

基于规则的双向最大匹配中文分词 Java - Chase1998的博客 - CSDN...

Spring高级技术梳理

Spring高级技术梳理 序言正文SpringDate部分Spring全家桶之SpringData——预科阶段Spring全家桶之SpringDat...

Python再次连续卫冕编程语言排行榜，Java和C下降，你怎么看？
Python的火，有目共睹，但是作为程序员的我们应该怎么选择？

句法结构**分析:基于**PCFG**的基本分析方法——给定一个句子..._CSDN博客**

技术人员要拿百万年薪，必须要经历这9个段位

很多人都问，技术人员如何成长，每个阶段又是怎样的，如何才能走出当前的迷茫，实现自我的突破。所以我结合我...

如何在Windows中开启"上帝模式"

原文链接：https://mp.weixin.qq.com/s?__biz=MzIwMjE1MjMyMw==&mid=2650202982&idx=1&...

十大优秀编程项目，让你的简历金光闪闪

全文共3241字，预计学习时长6分钟被问到如何学习编程时，最常听到的问题就是：“有没有什么新项目的好点子？...

面试官：兄弟，说说基本类型和包装类型的区别吧

Java 的每个基本类型都对应了一个包装类型，比如说 int 的包装类型为 Integer，double 的包装类型为 Double。基...

2019诺贝尔经济学奖得主：贫穷的本质是什么？

2019年诺贝尔经济学奖，颁给了来自麻省理工学院的 阿巴希·巴纳吉（ Abhijit Vinayak Banerjee ）、艾丝特·杜芙若...

4

1

脉

多功能双肩包

活出质感

限时折扣

¥49

关闭

那些拿到 60K Offer 的 AI 程序员，后来都怎么样了？

刚刚拿到阿里offer，工作地点杭州。值得去吗？

4

阅读数 2万+

博文来自: Dakshesh的博客

linux：最常见的linux命令（centOS 7.6）

最常见，最频繁使用的20个基础命令如下：皮一下，这都是干货偶，大佬轻喷一、linux关机命令：1.shutdown命令...

1

阅读数 3798

博文来自: vegetable_haker...

NLP学习记录——句法分析

句法分析一般而言，句法分析分为两个主要的方向：一是句法结构分析，或称成分句法分析、短语句法分析，亦即co...

阅读数 1万+

博文来自: qq_37171771的博客

世界上最好的学习法：费曼学习法

你是否曾幻想读一遍书就记住所有的内容？是否想学习完一项技能就马上达到巅峰水平？除非你是天才，不然这是不...

阅读数 5万+

博文来自: 程序新视界

系统学习NLP（五）--句法分析

转自：https://www.jianshu.com/p/fb408b6a0904 真佩服作者的毅力，把基础概念都敲出来了。。。句法分析的...

赏

阅读数 6242

博文来自: 工作笔记

程序员连拿3份Offer，每份工资竟超出原来薪资5万，这几点分享给你！

作为一名老码农！Python 116K 超过 C++，薪酬排行第一

脉

学Linux到底学什么

来源：公众号【编程珠玑】作者：守望先生网站：https://www.yanbinghu.com/2019/09/25/14472.html前言我们...

阅读数 2万+

博文来自: 守望的博客-编程珠玑

深入理解C语言指针

一、指针的概念要知道指针的概念，要先了解变量在内存中如何存储的。在存储时，内存被分为一块一块的。每一块...

阅读数 4万+

博文来自: ZackSock的博客

cyk追楠神系列一(SDUT3703)

cyk追楠神系列一

...

阅读数 120

博文来自: Mercury_Lc的博客

500行代码，教你用python写个微信飞机大战

这几天在重温微信小游戏的飞机大战，玩着玩着就在思考人生了，这飞机大战怎么就可以做的那么好，操作简单，简...

阅读数 6万+

博文来自: Python专栏

焦虑的互联网人及35岁定律

一、35岁定律好像现在整个互联网行业都在贩卖焦虑，从之前的30岁淘汰，到现在的35岁，到最新的39.9岁，你会...

阅读数 5319

博文来自: 十步杀一人-千里不...



他51岁，给卫星写代码：彪悍的人生，从来不怕晚！

如何从零基础进击编程实战？风云卫星数据工程师全部告诉你！

为什么我不建议你在当前购买 5G 手机

首发于公众号：BaronTalk，欢迎关注每年的八九月份都是各大厂商发布年自己旗舰机的时候，最近 vivo、oppo、...

阅读数 2万+

博文来自: BaronTalk

唐僧团队要裁员，你会裁谁？

提问：西游记取经团为了节约成本，唐太宗需要在这个团队里裁掉一名队员，该裁掉哪一位呢，为什么?为了完成西...

阅读数 4万+

博文来自: 王学明

8000字干货：那些很厉害的人是怎么构建知识体系的

本文约8000字，正常阅读需要15~20分钟。读完本文可以获得如下收益：分辨知识和知识体系的差别理解如何用八...

阅读数 4万+

博文来自: 程序视界——聚焦...

第二弹！python爬虫批量下载高清图

文章目录前言下载免费高清图下载带水印的精选图代码与总结前言在上一篇文章没高质量配图？python爬虫绕过...

阅读数 3万+

博文来自: bigsai

MySQL数据库—SQL汇总

一、准备下文整理常见SQL语句的用法，使用MySQL5.7测试，参考了尚硅谷MySQL教程及用例。用例sql：链接: htt...

举报

阅读数 3万+

博文来自: Sirm23333

多功能双肩包

活出质感

限时折扣

¥49

关闭



一本可陪伴一辈子的笔记本

可擦可写，可循环利用，支持OCR识别，让你的笔记本变得智能

十八、神奇算法之句法分析树的生成

把一句话按照句法逻辑组织成一棵树，由人来做这件事是可行的，但是由机器来实现是不可思议的，然而算法世界就...

只因写了一段爬虫，公司200多人被抓！

“一个程序员写了个爬虫程序，整个公司200多人被端了。” “不可能吧！” 刚从朋友听到这个消息的时候，我有点...

动画：用动画给面试官解释 TCP 三次握手过程

作者 | 小鹿来源 | 公众号：小鹿动画学编程写在前边TCP 三次握手过程对于面试是必考的一个，所以不但要掌握 TCP ...

句法分析 (PCFG , Transition-based parsing)

句法分析的基本任务是确定句子的句法结构或者句子中词汇之间的依存关系。 句法分析分为句法结构分析 (synt...

多线程编程是后台开发人员的基本功

这里先给大家分享一个小故事：在我刚开始参加工作的那年，公司安排我开发一款即时通讯软件（IM，类似于QQ聊...



治疗肺积水大概需要多少钱
治疗肺积水要多少钱
3571阅读

CYK算法详解

在计算机科学领域，CYK算法（也称为Cocke-Younger-Kasami算法）是一种用来对上下文无关文法（CFG，Conte...

爬虫小程序 - 爬取王者荣耀全皮肤

王者荣耀全皮肤图片爬取

动态规划入门到熟悉，看不懂来打我啊

2.1斐波那契系列问题 2.2矩阵系列问题 2.3跳跃系列问题 3.1 01背包 3.2 完全背包 3.3多重背包 3.4 一些变形选讲 2....

从入门到精通，Java学习路线导航

引言 最近也有很多人来向我"请教"，他们大都是一些刚入门的新手，还不了解这个行业，也不知道从何学起，开始的...

如何优雅的爬妹子网

from urllib import request import os from user_agents import ua_list import time import random impor...



cdn及cdn加速原理
国外cdn加速
9835阅读

花了20分钟，给女朋友们写了一个web版群聊程序

参考博客 [1]https://www.byteslounge.com/tutorials/java-ee-html5-websocket-example

Java入门学习路线目录索引（持续更新中）

新增：Redis 入门 【Redis缓存】 - 入门——Redis介绍和环境搭建【Redis缓存】 - Redis数据结构、基本命令操作...

八大排序(C语言)

void BubbleSort();//冒泡 void SelectSort();//选择 void InsertSort();//直接插入 void ShellSort();//希...

程序员必须掌握的核心算法有哪些？

由于我之前一直强调数据结构以及算法学习的重要性，所以就有一些读者经常问我，数据结构与算法应该要学习到哪...

python 程序员进阶之路：从新手到高手的100个模块

在知乎和CSDN的圈子里，经常看到、听到一些 python 初学者说，学完基础语法后，不知道该学什么，学了也不知...



多功能双肩包
活出质感
限时折扣
¥49
关闭

<div><div><div><div><div><div></div><div>sci论文终于接收了</div></div></div><div><div><div></div><div>润色sci论文</div></div><div><div></div><div>3283阅读</div></div></div></div></div></div>	4	
<div><div><div><div><div><div></div><div>Python——画一棵漂亮的樱花树（不同种樱花+玫瑰+圣诞树喔）</div></div></div><div><div><div></div><div>最近翻到一篇知乎，上面有不少用Python（大多是turtle库）绘制的树图，感觉很漂亮，我整理了一下，挑了一些我...</div></div></div></div></div></div>	1	阅读数 6万+ 博文
<div><div><div><div><div><div></div><div>这应该是把计算机网络五层模型讲的最好是文章了，看不懂你打我</div></div></div><div><div><div></div><div>帅地：用心写好每一篇文章！前言 天各一方的两台计算机是如何通信的呢？在成千上万的计算机中，为什么一台计...</div></div></div></div></div></div>		阅读数 2万+ 博文
<div><div><div><div><div><div></div><div>HTML CSS整理笔记</div></div></div><div><div><div></div><div>常见字体单位： 1.em 移动端常用的字体尺寸单位，说白em就相当于“倍”，比如设置当前的div的字体大小为1.5e...</div></div></div></div></div></div>		阅读数 2万+ 博文
<div><div><div><div><div><div></div><div>史上最全的mysql基础教程</div></div></div><div><div><div></div><div>启动与停止 启动mysql服务 sudo /usr/local/mysql/support-files/mysql.server start 停止mysql服务 sudo /usr/l...</div></div></div></div></div></div>		阅读数 3万+ 博文
<div><div><div><div><div><div></div><div>为什么你学不会递归？告别递归，谈谈我的经验</div></div></div><div><div><div></div><div>可能很多人在大一的时候，就已经接触了递归了，不过，我敢保证很多人初学者刚开始接触递归的时候，是一脸懵逼...</div></div></div></div></div></div>		阅读数 4万+ 博文
<div><div><div><div><div><div></div><div>有哪些让程序员受益终生的建议</div></div></div><div><div><div></div><div>从业五年多，辗转两个大厂，出过书，创过业，从技术小白成长为基层管理，联合几个业内大牛回答下这个问题，希...</div></div></div></div></div></div>		阅读数 6万+ 博文
<div><div><div><div><div><div></div><div>大学四年自学走来，这些私藏的实用工具/学习网站我贡献出来了</div></div></div><div><div><div></div><div>大学四年，看课本是不可能一直看课本的了，对于学习，特别是自学，善于搜索网上的一些资源来辅助，还是非常...</div></div></div></div></div></div>		阅读数 16万+ 博文
<div><div><div><div><div><div></div><div>linux系列之常用运维命令整理笔录</div></div></div><div><div><div></div><div>本博客记录工作中需要的linux运维命令，大学时候开始接触linux，会一些基本操作，可是都没有整理起来，加上是...</div></div></div></div></div></div>		阅读数 1万+ 博文
<div><div><div><div><div><div></div><div>大学四年，我把私藏的自学「学习网站/实用工具」都贡献出来了</div></div></div><div><div><div></div><div>在分享之前，先说说初学者如何学习编程，这个话题想必非常的重要，要学好编程，给你一些学习网站也好、实用工...</div></div></div></div></div></div>		阅读数 9万+ 博文
<div><div><div><div><div><div></div><div>中国麻将：世界上最早的区块链项目</div></div></div><div><div><div></div><div>中国麻将：世界上最早的区块链项目 最近区块链这个玩意又被市场搞的很是火热，相信大部分人都不太清楚这玩意到...</div></div></div></div></div></div>		阅读数 5万+ 博文
<div><div><div><div><div><div></div><div>比特币原理详解</div></div></div><div><div><div></div><div>一、什么是比特币 比特币是一种电子货币，是一种基于密码学的货币，在2008年11月1日由中本聪发表比特币白皮书...</div></div></div></div></div></div>		阅读数 4万+ 博文
<div><div><div><div><div><div></div><div>Python 基础（一）：入门必备知识</div></div></div><div><div><div></div><div>Python 入门必备知识，你都掌握了吗？</div></div></div></div></div></div>		阅读数 2万+ 博文
<div><div><div><div><div><div></div><div>兼职程序员一般可以从什么平台接私活？</div></div></div><div><div><div></div><div>这个问题我进行了系统性的总结，以下将进行言简意赅的说明和渠道提供，希望对各位小猿/小媛们有帮助~ 根据我...</div></div></div></div></div></div>		阅读数 7万+ 博文
<div><div><div><div><div><div></div><div>Ngrok：超简单的内网穿透，了解一下？</div></div></div><div><div><div></div><div>【1】什么是内网穿透？首先，我们生活中的网络从应用上可以分为内网和外网；内网就是你自己的网络环境，就你...</div></div></div></div></div></div>		阅读数 1万+ 博文
<div><div><div><div><div><div></div><div>反射全解</div></div></div><div><div><div></div><div>反射的概念 反射的引入： Object obj = new Student(); 若程序运行时接收到外部传入的一个对象，该对象的编译类...</div></div></div></div></div></div>		阅读数 1万+ 博文
<div><div><div><div><div><div></div><div>死磕C语言指针</div></div></div><div><div><div></div><div>兜兜转转还是逃不过 C 语言，这该死的缘分。 先看一眼我的西野七濑 学习自：https://zhuanlan.zhihu.com/p/891...</div></div></div></div></div></div>		阅读数 1万+ 博文
<div><div><div><div><div><div></div><div>Python十大装B语法</div></div></div><div><div><div></div><div>Python 是一种代表简单思想的语言，其语法相对简单，很容易上手。不过，如果就此小视 Python 语法的精妙和深...</div></div></div></div></div></div>		阅读数 8万+ 博文
<div><div><div><div><div><div></div><div>数据库优化 - SQL优化</div></div></div><div><div><div></div><div>从一个示例入手，带你一步一步掌握SQL优化的技巧！</div></div></div></div></div></div>		阅读数 1万+ 博文

多功能双肩包
活出质感

限时折扣
¥49

关闭



举报

送给单身猿们的表白神器 问天下男生，有谁想单身？又有谁想单身一辈子？虽然本人也是单身狗，但是也是有一个远大的理想，哈哈，大白天...		阅读数 1万+
python学习目录 这是我学习python的一套流程，从入门到上手 一、Python入门、环境搭建、变量、数据类型 二、Python运算符、...	4	阅读数 2万+ 博文
刷了几千道算法题，这些我私藏的刷题网站都在这里了！ 遥想当年，机缘巧合入了 ACM 的坑，周边巨擘林立，从此过上了"天天被虐似死狗"的生活... 然而我是谁，我可是死...	1	阅读数 4万+ 博文
JavaScript 为什么能活到现在？ 作者 司徒正美 责编 郭芮 出品 CSDN（ID：CSDNnews）JavaScript能发展到现在的程度已经经历不少的坎坷，...		阅读数 9033 博文
别翻了，这篇文章绝对让你深刻理解java类的加载以及ClassLoader源码分析【JVM篇二】 点进文章的盆友不如先来做一道非常常见的面试题，如果你能做出来，可能你早已掌握并理解了java的类加载机制，...		阅读数 1万+ 博文
Nginx 原理和架构 Nginx 是一个免费的，开源的，高性能的 HTTP 服务器和反向代理，以及 IMAP / POP3 代理服务器。Nginx 以其高...		阅读数 7741 博文
致 Python 初学者 欢迎来到“Python进阶”专栏！来到这里的每一位同学，应该大致上学习了很多 Python 的基础知识，正在努力成长...	 	阅读数 5万+ 博文
YouTube排名第一的励志英文演讲《Dream(梦想)》 Idon’ t know what that dream is that you have, I don’t care how disappointing it might have been as y...		阅读数 1万+ 博文
“狗屁不通文章生成器” 登顶GitHub热榜，分分钟写出万字形式主义大作 一、垃圾文字生成器介绍 最近在浏览GitHub的时候，发现了这样一个骨骼清奇的雷人项目，而且热度还特别高。 项...		阅读数 4万+ 博文
程序员：我终于知道post和get的区别 是一个老生常谈的话题，然而随着不断的学习，对于以前的认识有很多误区，所以还是需要不断地总结的，学而时习...		阅读数 5万+ 博文
加快推动区块链技术和产业创新发展，2019可信区块链峰会在京召开 11月8日，由中国信息通信研究院、中国通信标准化协会、中国互联网协会、可信区块链推进计划联合主办，科技行...		阅读数 2145 博文
Java世界最常用的工具类库 Apache Commons Apache Commons有很多子项目 Google Guava 参考博客		阅读数 1万+ 博文
程序员把地府后台管理系统做出来了，还有3.0版本！12月7号最新消息：已在开发中有github地址 第一幕：缘起 听说阎王爷要做个生死簿后台管理系统，我们派去了一个程序员..... 996程序员做的梦： 第一场：团队...		阅读数 3万+ 博文
python json java mysql pycharm android linux json格式 c#网页附加添加到母版页 c# 图 设计器 c# chart示例 c# mysql插入 c# sha256 加密 c#窗体上的又添加事件 c#打印图片文件 c#後台調前台js c#控制台美化 c# 标记 浮点型		



不愿透露姓名的广外男子

TA的个人主页 > 私信 关注

原创	粉丝	获赞	评论	访问
10	9	6	3	6711

等级: 博客 已

周排名: 20万+

积分: 172

总排名: 30万+

勋章: 



举报

多功能双肩包

活出质感

限时折扣

¥49

关闭



腾讯云

学生服务器套餐

10元/月

• 1核2G • 1M宽带

• 50GB存储

立即购买



最新文章

- 上下文无关文法到乔姆斯基范式的转换-CFG2CNF
- 基于特定语料库生成HMM转移概率分布和发射概率分布用于词性标注 Python
- 基于HMM+Viterbi算法的词性标注 Python
- 基于概率最大化的中文分词算法 Python
- 基于规则的双向最大匹配中文分词 Java

分类专栏

-  Kotlin 3篇
-  Android基础
-  Java 2篇
-  自然语言处理 6篇

归档

- 2018年11月 4篇
- 2018年10月 2篇
- 2018年9月 2篇
- 2018年7月 2篇

热门文章

- 上下文无关文法到乔姆斯基范式的转换-CFG2CNF
阅读数 1892
- 基于特定语料库生成HMM转移概率分布和发射概率分布用于词性标注 Python
阅读数 1251
- 基于概率最大化的中文分词算法 Python
阅读数 1225
- 基于HMM+Viterbi算法的词性标注 Python
阅读数 1083
- 基于CYK+PCFG的短语结构句法分析
阅读数 735

最新评论

- 基于CYK+PCFG的短语结构句法...
hztj2005 : 原理解释得清楚。在Jalaj Thanaki 《Python Natural Language Processing》第12 ...
- 基于特定语料库生成HMM转移概率分...
Chase1998 : [reply]qq_41290310[reply] 可以看一下前几篇博客

4

1

赏

脉



举报

多功能双肩包

活出质感

限时折扣

¥49

关闭

基于特定语料库生成HMM转移概率分...
qq_41290310 : 请问这里的 hmm_viterbi和max_
probability_seg能给出代码吗？



腾讯云

学生服务器套餐

10元/月

• 1核2G • 1M宽带

• 50GB存储

立即购买





程序人生



CSDN资讯

 QQ客服

 客服论坛


 kefu@csdn.net

 400-660-0108

工作时间 8:30-22:00

关于我们 招聘 广告服务 网站地图

京ICP备19004658号 经营性网站备案信息

 公安备案号 11010502030143

©1999-2019 北京创新乐知网络技术有限公司

网络110报警服务

北京互联网违法和不良信息举报中心

中国互联网举报中心 家长监护 版权申诉

4

1





多功能双肩包

活出质感

限时折扣

¥49

关闭



举报