使用指南

感谢您选择 The Prime Machine。我们希望此系统对您有所帮助。我们希望通过使用它,您会发现一些关于语言及语言表达的有趣的例子和使用方法。

什么是语料库 (Corpus)?

从本质上讲,语料库是一个提供单词和词组使用实例的文本集合。这些文本可能来 自书面或口头语言。他们可能是完整的文本,也可能是节选自某文本。语料库 *Corpus*的复数形式是 *Corpora*。

什么是语词检索程式(concordancer)?

Concordancer 是一个计算机程序,它集合了特定单词或词组的实例。使学生或研究 人员可以通过整个语料库来查找词语的所有使用方式。它的工作方式与搜索引擎或 文字处理器的查找功能类似,但其目的是为如何使用单词和短语提供实际示例。

我能用 The Prime Machine 做什么?

*The Prime Machine*在开发过程中充分考虑了语言学习者和老师的需求。您可以使用它来查找语言使用的示例。

它无法取代其他形式的语言学习支持,但是可以用它来补充信息。下表列出了您可以获得语言学习支持的资源,同时还对比了其他资源与*The Prime Machine*中提供的信息的差异。

资源	The Prime Machine
词典	• 更多的实例;
	• 特定领域的实例;
	• 关于典型语言环境的信息(标签及关联)
搜索引擎	• 结果能从用于文本分析的原文而不是网站中得出
	• 关于常用搭配的信息;
	• 所获结果的排序不会受到商业用途的影响
单词&语法书	• 得到的结果可以剖析其实际使用;
	• 得到的结果可以突出典型并显示差异化
	• 可以看到不同体裁及文本类型的对比

一些可能的用途是:

- 通过查找单词以找到有用的短语或搭配;
- •比较两个或多个同义词的使用情况;
- •查看不同体裁或文本类型中单词的使用方式;
- •从其他文本中查找示例,以比较您自己的语言选择。

系统要求

The Prime Machine 是一款客户端-服务器应用程序,这意味着如果需要使用它,您 将需要用互联网访问。它是一个 Windows 应用程序,可以在 Windows Vista 及以上 版本使用。本系统已在 Windows 7 和 Windows 10 系统上进行了测试。

目前,它还不可用于其他平台,但在未来,希望它也可以在 Mac, iOS 和 Android 系统 上使用。

安装

*The Prime Machine*应用程序文件是单个的 EXE 可执行文件。如果要从 ZIP 文件复制此应用程序,应在运行之前将其解压。

只需将应用程序文件放在桌面上(或系统中的其他位置),然后双击它即可运行。

卸载

*The Prime Machine*不会在系统注册表中存储非常多的数据,但是如果需要完全清除信息,可以到 "选项" 标签中,选择 "加载/保存设置" 页,然后单击"清除 tPM 文件关联"和 "清除家庭网络设置" 旁边的 **清除键**。

您也可以直接删除应用程序的 EXE 文件。

🔛 The Prime Machine V3.03									
Q		ſ		College Barb		Clark			
Search	Lines	Frequenci	es	Collocations	Labels	Associates	Options		
Settings			Loa	ad/save setti	ngs				
Search sugg	estions		Load	d settings from file		Load			
Ranking me	thods		Save	e settings to file					
Results			Include DIY Corpus Settings						
Tables and	Clouds		Registry settings						
Collocations			Clear	r tPM file associati					
Charts			Clear	r EULA and "Home					
Special Access									
Load/save s	ettings	Reset							

基本功能

本节将介绍如何使用 *The Prime Machine* 获取语言使用实例,以及如何运用某些主要功能。

打开应用程序

当您第一次使用 *The Prime Machine*时,您需要连接到服务器以访问预准备的语料 库和资源。

<u>西交利物浦大学的学生和工作人员</u>在校园中或使用 西交利物浦大学的虚拟网络时, 请记得勾选"尝试 tPM 内部网络(Try tPM Home Network)"以获得额外的资源。



单击 "连接"(connect), 应用程序将连接到服务器, 并检索可用的语料库的信息。 (如 果您已经保存了结果, 或者您想脱机创建 DIY 语料库, 则可以单击 "脱机工作" work offline)

功能概览

连接到服务器后,将显示主界面。在屏幕顶部有许多选项卡。这些能帮助您查到有 关单词和短语使用的不同类型的信息。

下图展示的生命环(Life-ring)按钮为您提供了有关当前选项卡的信息。

III The Prime Machine V3.03	- <u> </u>
Search Image: Sear	ates Options
🔐 🛱 🔶 🤿 🏂 🧍 🏲 BNC: Academi	c 🗸 🚽
Words and phrases Sentence breakdown Compare with another corpus Research tools	Search History
Search	8 Compare

如果您不确定某一按钮或搜索框的用途,您可以通过用鼠标指向它来获取提示消息。



以下为主选项卡列表及其用途:

Q Search	捜索选项卡	输入要搜索的单词、搭配或短语。
Lines	索引行与卡片 选项卡*	"索引行"搜到的结果与 "卡片" 选项卡相同,但只是 索引行搜索到的结果是以一个表格的形式展现。 且每一行只显示一个结果。在其他软件程序中, 这些称为题内关键词(上下文中的关键字)。如果 选定一个索引行,它也可以同时显示为一张卡片 (请参阅下面的索引行部分)。
ر Frequencies	频率选项卡*	"频率"选项卡可以帮助您查看您所查找的单词和 搭配是否常见。当看两个不同大小的语料库的结 果时,每百万字所计算的频率会帮助您比较结果。
Graphs	图形选项卡*	"图形"选项卡允许您查看所查单词的典型上下文 信息。包括位置(标题、文本位置等)、语法结构 (情态动词、时态、定冠词)等。除了查看显示当 前结果之外,还可以看到所有存储在整个语料库中 的结果的总体比例。
Collocations	搭配选项卡*	"搭配"选项卡根据不同种类数据统计度量,为您 展示了经常被一起使用的词语,即常用搭配。
Labels	标签选项卡*	"标签"选项卡为您提供了所查单词较为重要的标签信息。包括对整个文本的标记,包含所查单词, 作者或发言者的特定部分。
Associates	关联选项卡*	"关联"选项卡显示了所查单词的最常用文本类别。它还显示了某类文本中经常出现的关键词, 以及你所查单词为关键字的文本。
Options	选项	在不同选项卡的顶部会出现一些选项,但要保存或 加载设置并更改为高级设置,您需要使用 "选项" 这一选项卡。

注意:标记为星号(*)的选项卡仅显示通过服务器查询并下载的信息的结果。

"主页" 按钮

'搜索选项卡'中有一些子页和一系列不同的语料库可供选择。如果进入错误页面,请单击 "主页" 按钮来清除搜索框并返回默认语料库。

If you get lost, click the Home Button to clear the search boxes and return to the default corpus.

简单搜索

您可以在每个搜索框中输入1到5个单词。当您开始输入一个单词时,该程序将显示一系列自动生成的单词列表。

如果您在列表中看到所需的单词,只需单击它即可。

Words and phrases 17 Sentence brea	akdown Y Compare with an
cons	🛞 Search
cons (25)	
considered (4437)	
consider (3636)	
considerable (3031)	
consent (2050)	
consideration (2017)	

当您输入了与语料库中的完整单词匹配的单词后,系统将给您进一步的建议,显示为:

- •用同一串字母开头的较长单词;
- ·常用的搭配;
- ・同根词
- 具有类似含义或类似翻译的词语

Words and phrases Sentence breakdown Compare with an	another corpus Research tools Search History
consequence Search	h Compare
consequence (1406)	consequences (1701)
consequences (1701)	
as a consequence (337)	sufferings (35)
consequence of (866)	importance (3623)
inevitable consequence (24)	upshot (42)
direct consequence (33)	aftermath (139)
a consequence of (388)	event (2364)

如果输入的一串字母在当前选定的语料库中找不到,将出现红色下画线来表示错误 输入。这并不意味着拼写错误,只是在当前语料库中不能找到任何结果。另一方面, 如果您有**拼写错误**,可以右键单击该单词以获得拼写建议。

Words and phrases Sent	tence breakdown Co	ompare with an	ot
consegence	(🗱 Search)
Words and phrases Sen	tence breakdown C	compare with a	not
consegence		🗱 Search	
consequences (1701)			
changing (1500)			
consequence (1406)			
consciousness (1355)			
consensus (721)			

当您右键单击标记有错误行的单词时,该程序将检查所有语料库,以查看它们是否包含该单词。例如,如果您在**英国国家语料库的口语语料库**中查找 polymorph 这个单词,它是找不到的并且将被标记为错误。然而,这一**专业术语**在一些学术文本上是

确实存在的,并且其使用频率将会被展示。如果您双击列出的一个语料库,本软件将开始使用该语料库。



什么是单词?

这似乎是一个简单的问题,但是对于 The Prime Machine 系统处理您**输入的字符**时, 有几个要点需要解释。

基于一串字母的搜索中:

• 不同的单词形式在展示结果时不会被组合在一起;

如果你搜索原因(cause)和其复数形式(causes)你会得到不同的结果,同时得到的结果也和搜索其动名词形式(causing)以及过去分词形式(caused)不同。

• 拼写相同但含义不同的单词将混合在一起;

如果你搜索银行(bank),你可能会发现河岸和银行混在一起的例子。

大写字母和小写字母的处理方式相同。这意味着瓷器(china)和中国(China)的结果将混合在一起。

如果一个单词中间的部分是大写,那么则会出现大写字母在建议框的情况。(如:博士 PhD)。

如果要同时查看两种单词形式,可以使用 "比较" 按钮 (见下文)。

搜索搭配和短语

当您在搜索框中输入多个单词时,你要知道某一个词的**常用搭配**及其与其他单词组成的短语的区别。如果你搜索的词组被存储为常用搭配,该软件的数据库也会存储关于此搭配的**语言知识点(例如语义,语用等)**。

在 The Prime Machine 中,

如果您输入或选择一个符合要求的短语,您将能够看到匹配的图形,标签和关联的搭配,以及索引行信息。

对于搭配,单词的顺序很重要。您应该按他们的使用顺序输入单词,如果允许空隙, 在单词之间使用两个点。例如:

- 造成困难 (cause difficulties)
- 根本原因 (root cause)
- 导致..行动 (cause .. action)
- 可能..因为(likely..cause)

The Prime Machine 还存储三, 四和五词的搭配。

如果您在搜索框中输入的单词组合可以在数据库中找到,并作为常用搭配出现,服 务器则会为您进行查询并下载结果。

但是,如果该短语未存储为常用搭配,则系统将检查是否有该组合:

- 连续不断(如: 美国总统 President of the United States)
- 意味着每个单词之间不会出现其他单词,它们必须是按照规定顺序的。

President (总统)_of(的)_the_united(联合)_states(国家)

• 按顺序,但有空隙(如: 总统_美国 President _____ United States)

*意味着单词之间可能会出现其他单词,但它们必须按相同的顺序排列,并且第一到最后一个单词的距离不能超过4个单词。

president (总统) * united (联合的) * states (国家)

• 5 字窗口内的任何命令 (例如, 联合的总统国家 United President States)

"|"符号表示单词可能以任何顺序出现,但第一到最后一个单词的距离不能超过4 个单词。

联合 | 总裁 | 美国 united | president | states

这些搜索查询非常复杂,但是如果输入的短语没有被存储为常用搭配,系统则会给出 建议。如果您单击 "检查其他语料库" 按钮,该程序将检查所有其他语料库,以查看 该短语是否在其中为常用搭配。



R

当您学习和使用英语时,您需要特别注意词形。

在使用 The Prime Machine 语料库检索软件时,您应该非常谨慎的使用其屈折 语,即您应该在查询单词时完全按照它们在句子当中所出现的形式输入。

如果您查询的短语显示 "No priming summary available" 无相关语言知识点 (语义,语用等)信息,也许是因为您输入的短语没有被此软件存储为常用搭 配。即使是这样的情况,您也可根据此软件建议的使用 |,_or*符号加在所查询 短语当中,以便查询到一些索引行。

选择一个语料库

对于学术工作, BNC: 学术语料库是一个很好的选择。这是本系统的默认语料库, 并且在每次打开程序时都被默认选中。

如果您单击下拉菜单,您将能够看到其他可用的语料库。如果您是连接在 the Prime Machine's Home Network (XJTLU),相比没有连接西交利物浦大学网络或虚拟网络的用户,您将能看到并使用更多的语料库。

🔛 The Prime M	achine V3.03							- 0	×
Ø		f		Cland Cland	L'EAL A		R	\bigcirc	
Search	Lines	Frequencies	Graphs	Collocations	Labels	Associates	Options		
$ \begin{array}{c} {} \\ \hline \\ $									
Words an	d phrases Se	ntence breakdow	n Compare	e with another co	rpus Resea	arch tools Searc	h History		
Search Search									

英国国家语料库的每个部分都是单独提供的,但您也可以视其为一个完整的语料库进行搜索。



其他选项可能包括某些专业领域的语料库,如以下的 Hindawi 学术期刊。

Hinday	vi A	cademic Corpora
		Hindawi Mathematics
	Ł	Hindawi Chemistry
(\$	Hindawi Physics
		Hindawi Earth Science and Environment
l	À	Hindawi Engineering
9		Hindawi Computer Science
1	M	Hindawi Social Sciences
(Ð	Hindawi Biological Sciences

每个语料库都有自己的搭配和其统计数据的方法,所以尽量将您搜索的词语与合适的语料库相匹配。

For more information about the latest corproa available check tPM Help 003a and 003b from <u>help.theprimemachine.com</u>.

索引行 (索引行与卡片选项卡)

最重要的结果将展示在 "行" 或 "卡片" 选项卡上。这两个选项卡得到的是完全相同的结果,只是用两种方法展示。

在搜索框中输入要查询的单词或短语,然后单击 "搜索"(Search)。索引行和其他数据信息都是从**服务器**直接下载的,所以可能需要几分钟时间。

Words and phrases Sentence I	breakdown Compare with an
sophisticated	😹 Search

e

当它准备就绪时,程序将跳转到"行"或"卡"选项卡。您可以在查看索引行和索引卡 之间切换。(可以点击选项卡去选择其他默认结果展示方法)。

🔛 Tł	ne Prime Machine V3.0	3								- 0	×
Se	R Lines	f Frequencies		Graphs	Collocations	Labels	Associate	5	Options	\bigcirc	
A ⁺	🛾 💽 🚺	Collocations	•							÷	W
BN	C: Academic: so j	phisticated									200
		Text to the left of	node	Node	Text to the rig	ht of node					
1	nstances amounts to a r	misrepresentation of the	more	sophisticated	falsificationist's p	osition. More than	a hint of this is	cc	more sop	histicated	
2	19)/ Dearlove (1979) se	eeks to develop a rather	more	sophisticated	analysis in his dis	scussion of reorgar	nization, but his	co	falsifica	ationist	
3	of industrial land market	ts in comparison with the	more	sophisticated	understanding o	flabour and capita	l markets./Preli	nii	Humanities	and Arts	
4	is century by Robert Wa	alcott, who argued for a	more	sophisticated	understanding o	f the nature of the	party structure	u	science? Chair	ners, A F	
5	t process of accumulation	on. This is clearly a much	more	sophisticated	analysis of the p	henomena than th	e empirically inv	alie	Open Universi Milton Keynes	ty Press 1992 1-120	
6	from the decades of dec	dicated work, using ever	more	sophisticated	technology, is th	at the creature ski	ills develop no fi	irt			
7	ot be cost-effective to p	oursue without access to	more	50-	technoloav./Sor	tina Options		_	The self-av	owed	
8	state's role has led son	ne Marxists to develo	Tł	he word	vou looke	d for is	operating upon	th	as much in the	passage	, ,
9	omes more experienced	l, her/his needs becor					he more special	ist	quoted on p.4 italics are his.	4, where the However,	e
10	:./ Technical Documenta	ition: As aircraft beco	С	entred ir	i the table	e, with	he proliferation	of	exclusive atte	ntion to	
11	: Diagnosis/Isolation: Air	craft have become e	ea	ch exam	ple in one	row in	in a single airfra	am	amounts to a		
12	gh the tasks to be perfo	ormed may have becol					understanding	ab	more sophistic	ated	
13	it the environment, as r	market demands become	more	sophisticated	so there is a nee	d for operator flex	cibility in product	ior	falsificationist More than a h	s position. int of this is	
14	ion of 'rational choice' in	troduces us to another,	more	sophisticated	conception of th	e rationality of raci	ism, and one wh	ict	contained in the	ne example	
15	that of 'ordinary prude	nt men', need to have a	more	sophisticated	grasp of the fac	tors that should inf	fluence their dec	isi	section conclu	ded	
16	this respect, his pluralis	sm can be regarded as a	more	sophisticated	version of monis	m. The flaw of mor	hism is that it ter	nd:			
17	ing a Marxist, has to pro	ovide an alternative and	more	sophisticated	account of why i	military expenditur	e remains hie?				
18	mentioned earlier, the	HMIs took an altogether	more	sophisticated	view of the relat	ionship of educatio	on to socie	If yo	ou choos	e Lines	s + Viev
19	ill see later, subsequent	experiments have used	more	sopnisticated	measures of rec	all.) Knowing or no	t knowing	the	e card for	r the c	urrently
				PoS		1	:	sele	cted line	appea	rs on tl

🔛 The Prime Machine V3.03



The View Options button allows you to change the way concordance lines are displayed. Click to access the drop-down menu.



关于索引行的更多信息

如果您想快速查看许多示例,则这些索引行很有用。

符号"/"用于分隔段落。

如果将鼠标悬停在行号上,类别和源信息将会出现。

		Text to the left of node		Node	Tex
	1	g instances amounts to a misrepresentation of the more	50	phisticated	falsi
	2	n statements that are inconsistent with them. The more	50	phisticated	falsi
	3	es from the decades of dedicated work, using ever more	50	phisticated	tech
	4	amont polico officera from civilian support staff to more	60	phisticated	tech
	0	Humanities and Arts		ohisticated	tech
		What is this thing called science? Chalmers, A F Ope	n	ohisticated	ana
Ļ		University Press Willon Reynes 1992 1-120		phisticated	indu
	8	to adhere to that position. None of the modern, more	50	phisticated	indu

与所有其他结果一样,单击右键或双击 🖪 🕼 🕼 🕼 可以选择复制或保存结果。 program, the more MetaObjects 允许您在另一台计算机程 序中重新调整结果显示的大小,而不会 影响图像质量。 + Copy as MetaObject 结果最终以图片保存为 JPEG 图像后, 因此无法编辑。如果放大,可能会出现 Ð Copy as Picture 图像不清晰的情况。 5 Save as MetaObject... Save as Picture... 对于 "索引行" 选项卡 (以及其他结果), E) Save as Sheet... 您可以选择复制或保存为电子数据表。 Ø View Search History 保存为工作表的默认文件格式为 î Show info... Microsoft Excel 97-2003 工作簿 (.XLS 文

件)。 但是,您也可以选择另存为 CSV 文件格式(逗号分隔文件)。



"索引行"选项卡将尽可能多地显示文本的左侧和右侧,最多可以显示其 之前的一整句话到之后的一整句话。但是,如果只想看到完整的句子,可 以使用 "视图设置" 下拉菜单中的 "**仅限句子**"(Sentence Only)选项。

更广泛的背景:

	Text to the left of node	Node	Text to the right of node
1	to falsifying instances amounts to a misrepresentation of the more	sophisticated	falsificationist's position. More than a hint of this is contained in th
2	together'(p. 19)/ Dearlove (1979) seeks to develop a rather more	sophisticated	analysis in his discussion of reorganization, but his conclusion is sin
3	he workings of industrial land markets in comparison with the more	sophisticated	understanding of labour and capital markets./Preliminary research
4	middle of this century by Robert Walcott, who argued for a more	sophisticated	understanding of the nature of the party structure under William a
5	the capitalist process of accumulation. This is clearly a much more	sophisticated	analysis of the phenomena than the empirically invalidated instrum
6	1at emerges from the decades of dedicated work, using ever more	sophisticated	technology, is that the creature skills develop no further. At a con
7	e, it would not be cost-effective to pursue without access to more	sophisticated	technology./ Sorting Options

限于同一句:

	Text to the left of node	Node	Text to the right of node
1	to falsifying instances amounts to a misrepresentation of the more	sophisticated	falsificationist's position.
2	/ Dearlove (1979) seeks to develop a rather more	sophisticated	analysis in his discussion of reorganization, but his conclusion is sir
3	he workings of industrial land markets in comparison with the more	sophisticated	understanding of labour and capital markets./
4	middle of this century by Robert Walcott, who argued for a more	sophisticated	understanding of the nature of the party structure under William a
5	This is clearly a much more	sophisticated	analysis of the phenomena than the empirically invalidated instrum
6	nat emerges from the decades of dedicated work, using ever more	sophisticated	technology, is that the creature skills develop no further.
7	e, it would not be cost-effective to pursue without access to more	sophisticated	technology./

关于索引卡的更多信息

如果您希望看到某一单词使用的更大范围的上下文,或 是您想看到它在**段落或标题中的位置**,则这些索引卡片 很有用。

您可以单击右键或双击来复制或保存索引卡。

作为长列表查看的卡可以复制或保存为图片或电子表 格。

B	Copy as MetaObject
Þ	Copy as Picture
5	Save as MetaObject
	Save as Picture
E)	Save as Sheet
0	View Search History
ŝ	Show info



Save all as MetaObjects...

Save all as Pictures... 这张索引卡或

如果您从索引行+屏幕处单击右键,您可以选择保存这张索引卡或所有的索引卡片作为图片。



文本类别和**源信息**可能很有用,但它会使卡片变大。如果您想要隐藏这些信息,可以通过 "视图设置" 下拉菜单来隐藏这些信息。

索引行及索引卡排序

搜索时, The Prime Machine 服务器将发送固定的随机选择的索引行。如果查询的结果比您设定的限制少(或者如果您使用 DIY 语料库),您将收到所有的索引行。

能以不同的方式对这些行和卡进行排序,是本系统的亮 点。这可以帮助您在搜索的结果中看到不同的使用模 式。

对结果进行排序有 12 种不同的方法,如果选择 "按字 母顺序 左 1 (L1) " 或 "按字母顺序右 1 (R1) ",则可 以单击图示按钮以更改排序顺序。





下面列出了为索引行进行排序的	不同方法。
----------------	-------

	随机顺序	固定的随机顺序-如果语料库被重新处理(即在一
T-J_233	Random Order	个主要版本升级后,或如果您再次导入文本到 DIY
		语料库)时,这个顺序才会改变。
	文本顺序	按照文本和句子加载到语料库中的顺序排序。
	Text Order	
	常用搭配	按照词语搭配强度排序-包含很强关系的词组搭配
	Collocations	的索引线将被置顶。(索引线将根据跨文本链接的 排序进一步排序)。
	労田塔配(左)	和党田塔配一样 伯日息云左边的单词
い	市市背配(江) Collocations (L)	和市用指配 件, 但八亚小工边的手问。
	常用搭配(右)	和常用搭配一样,但只显示右边的单词。
1	Collocations (R)	
	跨文本链接	跨文本链接是基于您搜索的单词的左右两边常用
	Links Across Texts	搭配词性来排序的。通常,您可以查看所查询单
0		词的左边 4 个和右边 4 个常用搭配的词性。满足
		这些条件的索引行将被置顶。(其它索引行将按
		照常用搭配强度来排序。)
	跨文本链接(左)	与跨文本链接相同,但仅限于向左文字。
1/2	Links Across Texts (L)	
	跨文本链接(右)	与跨文本链接相同,但仅限于向右文字。
	Links Across Texts (R)	
	文本内链接	根据同一文本中句子和其他句子之间的链接排列
	Links Within Texts	的顺序在每个文本中,某一单词在其他句子中也
		重复出现的索引行将被排在最前面。
		这个可用于 DIY 语料库。
	· · · · · · · · · · · · ·	一个基于短可及较少的低频间的排序力法一家引
	Dictionary Style	行中有许多付亏, 个常使用的问或名子等将做排 在最下面
AZ	按字母排序	在搜索单词左侧的1、2、3或4个单词中按字母
	(左 1- 左 4)	顺序排列。
	Alphabetical L1-L4	
AZ	按字母排序	在搜索单词右侧的 1、2、3 或 4 个单词按字母顺
	(右1-右4)	序排列。
	Alphabetical R1-R4	

怎样获取更多的结果

如果您要查看**更多的索引结果**(更多的索引卡片和索引行),请单击显示结果的窗口 右上角的**加号按钮**。



还可以通过更改选项卡上的设置(Options Tab)来增加每次获取的结果数量。

筛选结果

通过单击 "索引行/卡" 选项卡右上角的图标, 可访问 "筛选" 的下拉菜单。

有许多不同的方法来选择和筛选一些索引行。

- 您可以通过隐藏/显示复选框(Hide/Show checkboxes)来 加亮索引线。
- 您可以使用 "内部搜索..." (Search Inside...) 按钮来标亮 显示 (或清除) 包含想要搜索的特定文本的索引行。
- 您可以使用 "图表"选项卡 (Graphs Tab)上的按钮取消选择某些语言知识点的展示,然后突出显示符合您要求的索引行 (见 22 页)。

突出显示一行或多行后,您可以通过**"筛选左侧"(Filter Left)** (或是**筛选右侧 Filter Right** 等)来显示选定的索引行。

如果要将突出显示的索引行与未突出显示的进行比较,单击" **筛选和比较"(Filter and Compare)**后显示结果将被拆分。左 侧仅显示突出显示的索引行,而其余部分则显示在右侧。



常用搭配

"搭配"选项卡(Collocations tab)显示经常被一起使用的单词。您可以通过单词云 (Word Cloud)或表格来查看结果。



	Collocation	Frequency	Log-likelihood $ abla$	Bayes Factor
1	as a consequence	337	464.42	Very strong evidence
2	consequence of	866	425.16	Very strong evidence
3	inevitable consequence	24	160.23	Very strong evidence
4	direct consequence	33	130.06	Very strong evidence
5	a consequence of	388	62.94	Very strong evidence
6	unintended consequence	7	59.62	Very strong evidence
7	partly consequence	13	44.31	Very strong evidence
8	natural consequence	17	43.89	Very strong evidence
9	consequence this	106	38.95	Very strong evidence
10	logical consequence	9	34.57	Very strong evidence
11	consequence instability	6	34.16	Very strong evidence
12	virtue consequence	8	31.99	Very strong evidence
13	necessary consequence	16	27.89	Very strong evidence
14	consequence unsolicited	3	26.36	Strong evidence
15	consequence breach	9	23.10	Strong evidence
16	probable consequence	5	21.18	Some evidence

计算常用搭配强度的主要算法包括:

对数似然比 (*log-likelihood*)考虑了单词的位置,并且提供了长达 5 字或更长的搭配。

得尔塔 P (*Delta P*)考虑了顺序 (A 然后 B),并且显示 2 词的搭配。

其他三个(互信息 3 值(*MI3*), T 值(*T-Score*)和戴斯系数 (*Dice*)) 是在其他语词检索程序中用于计算常用搭配 的衡量标准。

关于这些算法的表格显示了您所搜索的单词与其左 右两边的4个单词同时出现的比例。

集群不是基于一个数字的衡量标准;它们是在最小频 率之上重复出现的字串,通常是按照长度排序。

Log-Likelihood Delta P MI3 T-Score Dice Clusters

图表选项卡

系统主界面的底部的一组图标用于显示一些典型的上下文。你可以浏览 "图表"选项卡(Graphs Tab)上的信息。要跳转到 "图表"选项卡(Graphs Tab)上的相关页面,只需单击其中一个图标即可。



刚开始您可能会觉得**"图表"选项卡(Graphs Tab)**上的菜单和子菜单非常复杂,但 它们提供了重要且具有总结性的功能信息。**"图表"选项卡(Graphs Tab)**的每个子 菜单都有其自己的**生命环(Life-ring)**信息。

使用指南

🔣 The Prime Machine V3.03			
Search Image: Select all features	cies Graphs C	Collocations Labels	The figures for current set are for the concordance lines which have been downloaded. The figures for all are for
Text Position	First fifth 🔽	consequence (BNC: Academic)	all the concordance lines in the entire
Text Position	First third	90	You can choose to highlight lines on the Lines/Cards Tab by unticking some
Paragraph Position	Last third 🗹 🗌	80	features and clicking Mark left lines
Brep Sentence Position	Last fifth 🗹	60	
Theme/Rheme	Not First or Last third	40	Nonn-36 Magan 33
Quotations		30	Norm 23
		0 First fifth First thir	rd Last third Last fifth Not First or Last third

例如,下面的图表显示了结果(consequence)一词在 BNC 语料库中是如何使用的: 在 BNC 学术语料库中,它经常在句子的前五分之一或三分之一出现。规范(Norm)则显示了这个词在这些语境中,和在所有语料库中的比例,所以很容易看出使用趋势有何不同。



第23页,共26页

并排比较结果

用 The Prime Machine 同时比较两个词或词组非常容易。

只需要简单地在每个搜索框中输入一个搜索字词,然后单击**"比较"(Compare)按** 钮。

建议用于:

- 查看同一个单词的不同单词形式的使用方式 例如:名词分析(*analysis*)与动词分析(*analyse*)
- 查看单词在作为常用搭配时的使用方式
 例如:有效(effective)与成本效益(cost effective)
- 查看具有类似含义或翻译的单词的使用方式 例如:结果(consequence)与结果(result)

比较两个不同语料库的结果

除了看一个语料库的结果,用 The Prime Machine 你也可以并排看到两个不同语料库的搜索结果。这将使您看到在不同体裁或文本类型中单词和词组不同的使用方式。

要使用此功能,只需单击 "**搜索" (search)**选项卡中的 "**与另一个语料库**"

(compare with another corpus)选项卡进行比较即可。在框中输入单词或短语, 然后选择要用于比较的语料库即可。

🔛 The Prime	Machine V3.03							_		×
Search	Lines	f Frequencies	Graphs	Collocations	Labels	Associates	Options		\bigcirc	
6) E	Ç	₽ \$		ย	BN	IC: Academic			•	
Words a	nd phrases Se	entence breakdov	wr Compare	e with another co	rpus Rese	arch tools Sear	ch History			
				*	dF)	BNC: Newspa	apers		-	·
								Co	ompare	e

更高级的设置

标签选项卡

语料库中的句子通常附有标签,以提供有关文本、章节或原创作者的详细信息。数 字统计的标准可以确保一个词或词组出现在特定类型的文本中的机会比预期更多。

文本标签包括文本的主要类别 (如小说或学术文章)。 它们还可能包括有关出版商、 来源或体裁的信息;

章节标签是文本的子标题 (如摘要、导言或结论);

原创作者标签提供关于作者或演讲者的信息(例如他们的名字、年龄、性别等)。

邻里标签基于对当前卡片中单词和词组的重复语义标记。

框上方的线条和箭头显示了在数据库中可见的标签所占的单词的匹配比例。如果 百分比非常高,则意味着云中的标签会对大多数可使用的索引行进行说明。如果百 分比很低,则表示在许多其他文本环境下也会出现该单词。

<u>详细说明</u>:由于可能出现一个单词和多个标签关联的情况,因此这些数字显示为一系列值,而不是单个百分比。较高的百分比显示了在云中(*the cloud*)可以表示的最大发生比例,因为它给出了标记的所有标记的组合频率。

关联选项卡

搭配主要是研究某一关键词左右两边经常一起使用的词,但你也可以通过看到单词 或词组经常出现在同一文本中,来了解一个词是如何使用的。系统中的中的单词和 词组是与主搜索词相同的文字中的**关键词**。

此选项卡还显示搜索的顶级文本类别以及属于该类别的结果百分比。

箭头用来表示如果所有单词在所有文本类别之间平均分配,是高于还是低于预期的比例。

如果您的单词从未出现在类别中,则它将在 "空类别" 框中(Empty Categories box) 列出。这可能表明该词通常不用于这种文本。

如果您有兴趣查看特定类别或标记的关键字列表,则可以在 "主搜索"选项卡 (Search Tab window)窗口中使用 "标记搜索"(Tags)。

选项卡

您可以通过自定义您的搜索来进行很多的更改。

最有用的是**排序方法(Ranking methods)**和结果 (Results)选项卡。

选项卡的每个子节都有其自己的**生命环** Life-ring 信息,因此您可以单击**生命环(Life-ring)**信息查看详细信息。

Settings
Search suggestions
Ranking methods
Results
Tables and Clouds
Collocations
Charts
Special Access
Load/save settings Reset

更多支持和帮助

The Prime Machine 仍在不断升级改进中。

浏览 http://help.theprimemachine.com 以获取更多信息。

Last Updated: 10 May 2018