

双重否定结构自动识别研究

王昱

北京大学中文系/ 北京100871
wangyustu@pku.edu.cn

袁毓林

澳门大学人文学院中文系/ 澳门
北京大学中文系/中国语言学研究
/教育部计算语言学重点实验室/ 北京100871
yulinyuan@um.edu.mo/yuanyl@pku.edu.cn

摘要

双重否定结构是一种“通过两次否定表示肯定意义”的特殊结构，其存在会对自然语言处理中的语义判断与情感分类产生重要影响。本文以“ $\neg \neg P \Rightarrow P$ ”为标准，对现代汉语中所有的“否定词+否定词”结构进行了遍历研究，将双重否定结构按照格式分为了3大类，25小类，常用双重否定结构或构式132个。结合动词的叙实性、否定焦点、语义否定与语用否定等相关理论，本文归纳了双重否定结构的三大成立条件，并据此设计实现了基于规则的双重否定结构自动识别程序。程序实验的精确率为98.85%，召回率为98.90%，F1值为98.85%。同时，程序还从96281句语料中获得了8640句精确率约为99%的含有双重否定结构的句子，为后续基于统计的深度学习模型提供了语料支持的可能。

关键词： 双重否定；自动识别程序；语义分析

The Research on Automatic Recognition of the Double Negation Structure

Wang Yu

Department of Chinese Language
and Literature, Peking University
/ Beijing 100871
wangyustu@pku.edu.cn

Yuan Yulin

Department of Chinese Language and
Literature, University of Macau / Macau
Department of Chinese Language and
Literature, Center for Chinese Linguistics,
Key Laboratory of Computational Linguistics,
Peking University / Beijing 100871
yulinyuan@um.edu.mo/yuanyl@pku.edu.cn

Abstract

The double negation structure is a special structure of "expressing positive meaning through two negations", in which the two negations have an important impact on the semantic analysis and emotional classification in natural language processing. Taking " $\neg \neg P \Rightarrow P$ " as the standard, this paper makes an ergodic study on the "negation word + negation word" structures in modern Chinese; According to the formal features and semantic expressions, the double negation structure is divided into three categories, 25 sub-categories and 132 common structures or constructions. Then, based on the theories of factuality, negation focus, descriptive truth-functional negation and non-truth-functional negation, this paper investigates the double negation structure, summarizes the three conditions for the establishment of the double negation structure, and designs the program of automatically recognizing the double negation structure. The accuracy rate of the recognition of the double negation structure is 98.80%, the recall rate is 98.90%, and the F1 value is 98.85%. Meanwhile, the program also obtains 8640 sentences which 99% contains double negation structure from 96281 sentences, which provides corpus for the subsequent statistical based deep learning model.

Keywords: Double negation, Automatic recognition program, Semantic Analysis

1 引言

在否定用法中，有一种特殊的用法——双重否定。丁声树先生在《现代汉语语法讲话》中将其概括为：“一句话先后用两个否定词，如‘不能不去’，‘没有人不去’，‘非去不可’之类，都是双重否定的句法。双重否定意思是肯定的，不过跟单纯肯定不全一样”。例如，“我不得不喜欢他”指“我得喜欢他”，“我不一定不同意这个观点”指“我可能同意这个观点”⁰。虽然句子使用的是否定形式，但其表达的却是肯定语义，且句子的语气还与单纯肯定有一定的区别（例如，“不得不喜欢”的语气比“喜欢”强，而“不一定不同意”的语气比“同意”弱，等等）。可见，双重否定结构的存在会对句子的语义真值与语气强度产生重要影响。

对于自然语言处理来说，双重否定是处理语料时必须考虑的内容。例如，下图展示了目前常用的自然语言处理应用对双重否定结构的处理情况。左图为微软小冰聊天机器人。如图所示，小冰无法有效识别以否定形式的“不是不”表示肯定的句子的语义与情感。即使是对于“我不是不喜欢你，而是超级喜欢你”这种明显表示积极情感的句子，小冰也仍然会将其误判为消极。右图是百度情感分析应用。我们的输入是“我不认为这件事情不值得称赞”，这句话应该是中性或者偏积极。但是由于无法识别双重否定结构，程序将其误判为了百分之百消极。可见，目前自然语言处理应用尚无法有效识别双重否定结构，且双重否定结构的存在会对程序的语义识别与情感分析产生重要影响。**综上所述，无论是从汉语本体研究的理论意义考虑，还是从自然语言处理的实用价值出发，我们都有必要对双重否定结构进行更全面更深入的研究与探索。**



Figure 1: 双重否定结构处理情况示例图

1.1 文献综述

目前国内关于双重否定结构的理论研究主要集中在双重否定结构的定义、范围、格式和语义等领域。关于双重否定结构的定义与范围，学界存在不少的争论，学者们争议较多的问题有：含有否定词的反诘句是否为双重否定，紧缩复句是否为双重否定，双重否定是否包含复句，等等。关于双重否定结构的格式，已有研究都是从分类与举例的角度进行讨论，尚未有一篇论文从形式的角度对双重否定格式进行具体详细的遍历分析。目前对双重否定格式概括最为全面的是芜崧（1987）所划分的八大类型，25个格式，然而，仍有许多语料中出现的双重否定结构未被芜崧收入其中，如“不可能不”“不应该不”“不是...不...”，等等。关于双重否定结构的语义，叶文曦（2013）、方绪军（2017）、何爱晶（2019）等引入了Ladusaw（1997）的形式语义学，Horn（1985）的元语否定等理论，对一些典型的结构进行了分析，得出了具有解释

©2022 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

基金项目：本课题得到国家科技创新2030“新一代人工智能”重大项目《以自然语言为核心的语义理解理论、模型与方法》（编号：2020AAA0106701）和国家社科基金重大项目《基于“互联网+”的国际汉语教学资源与智慧教育平台研究》（项目编号：18ZDA295）的资助，谨此致谢。

⁰丁声树等.现代汉语语法讲话[M].北京:商务印书馆,2004.200-202.

力的成果。然而，由于双重否定的范围、格式还未确定，目前学者只集中分析了几个典型的结构，覆盖面十分有限，缺乏系统性地总结与梳理。

目前关于双重否定结构的应用研究主要集中在情感分析领域，具体根据研究方法可以分为以下两种。

一、通过搜集典型的双重否定结构（如图2所示），构建双重否定词典，以服务相关的情感分析，如王勇等（2014），吴杰胜、陆奎（2019），等等。这种方法的准确率很高，但是覆盖面不足。

双重否定词典	绝非不、并非不、不是不、不能不、不会不、不可不、不要不、不得不、没有不、无不、不无
--------	---

Figure 2: 王勇等（2014）双重否定词典

二、在否定词的基础上，通过统计修饰每一个情绪词的否定词个数来判断双重否定，并以系数的形式将双重否定的语气功能加入到情感分析的结果当中，如封洋（2016），等等。这种方法的涵盖范围广，但是错误率很高，因为任何含有两次否定的结构都会被判断为表示肯定的双重否定结构。

1.2 本文选题及目标

综上所述，目前学界对汉语双重否定结构的研究成果颇丰但仍然存在一些不足之处，例如，双重否定结构的格式与范围尚不完整；双重否定语料资源匮乏；系统化、全面化的双重否定结构自动识别尚未实现，等等。鉴于此，本研究将“双重否定结构”作为研究对象，试图通过遍历分析与语料考察相结合的方法，对双重否定结构进行以下探索：

1. 梳理双重否定结构格式，使其能够全面覆盖CCL语料库；
2. 总结双重否定结构成立条件，并据此提出相应的计算机识别策略；
3. 建立高F1值的双重否定结构自动识别程序；
4. 进一步验证语言学知识在双重否定结构自动识别过程中的贡献，通过程序测试上述成立条件在双重否定结构识别过程中的作用；
5. 搜集双重否定语料资源，为基于统计的双重否定识别深度学习模型提供支持。

2 双重否定结构的定义标准与考察范围

鉴于语义真值识别和情感极值判断是计算机对否定结构进行语义识别时所面临的主要问题，本文借鉴形式语义学，为双重否定拟定了一个工作定义：只要两次否定与肯定在语义真值上相同，即“ $\neg \neg P \implies P$ ”，即属于双重否定。

目前我们的考察范围为所有“否定词 (+...) + 否定词”中双重否定表肯定的结构。暂不考虑下列情况：

1. 否定词为隐性否定词（即本身语义内含有否定意思的动词，如“讨厌”“拒绝”等。）
2. 否定类型为语用否定的结构（比如“我不是不喜欢你，而是恨你”中的“不是不喜欢”。）
3. “反问句+否定词”结构（比如“难道...不...”等，具体参见刘彬、袁毓林（2017）。）

我们结合吕叔湘（1956/1944）、王力（1985/1943）、朱德熙（1982）等前贤研究，梳理出了十三个否定词，前十个为否定副词，后三个为否定动词，具体如下：

“非、不、别、甬、未、莫、勿、没、没有、休、无、没、没有”

结合语料，我们对“否定词+否定词”组合中符合要求的双重否定结构的格式进行了遍历梳理，整理出了25种“否定词+否定词”可表肯定的结构，具体如表1所示（其中三角形表示该组合不出现/极少出现于实际语料中）。

下面，我们将对上述25种“否定词+否定词”格式展开具体分析，梳理每一种格式中双重否定结构的成立条件，并设计与之相应的计算机识别策略，在此基础上总结所有现代汉语中常用的双重否定结构。

	不	没	没有	无	非	莫	别
不	不…不	不…没…	不…没有…	不无	▲	▲	▲
没	没不 没…不…	没…没…	没…没有…	▲	▲	▲	▲
没有	没有不 没…不…	没有…没	没有…没有…	▲	▲	▲	▲
无	无不 无…不…	▲	▲	▲	▲	▲	▲
非	非不 非…不…	非没 非…没…	非没有 非…没有…	非无	▲	非…莫…	▲
莫	莫不	▲	▲	▲	▲	▲	▲
别	别不	别没	▲	▲	▲	▲	▲

Table 1: 双重否定结构格式调查情况表

3 双重否定结构的成立条件与识别策略

通过初步考察，我们发现，双重否定结构“不V1...不V2”的成立条件最为复杂，需要同时满足以下三个条件：

- ①“不V1”与“不V2”构成述宾关系；
- ②“V1”动词有限制（为非叙实动词）；
- ③“不V1”的否定焦点在“不V2”上。

我们发现，除了“不是...不...”需要区别语义否定与语用否定外，其他双重否定结构的成立条件都已囊括在了上述三个条件中，只是部分细节存在差异。因此，我们先对“不V1...不V2”与“不是...不...”的成立条件与识别策略进行详细分析，再在此基础上，对其他双重否定结构进行讨论。

3.1 “不V1...不V2”双重否定结构的成立条件与识别策略

3.1.1 第一个条件：“不V1”与“不V2”构成述宾关系

“不V1...不V2”的结构类型有并列、主谓、紧缩、述宾等。在各类结构类型中，只有述宾结构的“不+V1+ (...) +不+VP”存在表示双重否定的可能。具体讨论如下：

并列结构的“不V1...不V2”，指“不哭不闹”、“不高不低”这一类表达。袁毓林（1999）指出，并列结构“通常不能通过直接在这种谓词性并列结构的前面加上‘不、没有’等否定词来构成否定式，而是要在这种并列结构的各个直接成分之前分别加上‘不、没有’等否定词。”因此，“不V1...不V2”只是“V1V2”并列结构的单重否定结构，不属于双重否定结构。如“不哭不闹”不等于“哭闹”。

主谓结构的“不V1...不V2”语料数量很少，指“不隐藏不代表泄露”、“不买票不是我的决定”这一类表达。在该类结构中，“不V1”只是一个命题陈述，是交流中的旧信息。“不V1”的“不”与“不V2”的“不”，并没有语义上的关联，并不构成“双重”否定的结构。“不V1...不V2”只是“不V1V2”结构的否定，不属于双重否定结构。如“不隐藏不代表泄露”不等于“隐藏代表泄露”，“不买票不是我的决定”不等于“买票是我的决定”。

紧缩结构的“不V1...不V2”指“不给钱不办事”、“不买票就不让进”这一类表达。紧缩结构虽然在语义上有条件性，但在句法上仍是并列关系，前后并不构成从属结构。关于紧缩条件类的结构是否为双重否定未有定论。本文主要从形式语义学的角度对其进行讨论。

以“不给钱不办事”为例。“给钱办事”语义为“如果给钱，那么办事”。P命题可以分解为q1“给钱”，q2“办事”，逻辑式为蕴含式 $q1 \rightarrow q2$ ，它的等值式为 $\neg q2 \rightarrow \neg q1$ 。而“不给钱不办事”语义为“如果不给钱，那么不办事”，逻辑式应为蕴含式 $\neg q1 \rightarrow \neg q2$ 。从下列真值表本文可以看出， $q1 \rightarrow q2$ 与 $\neg q1 \rightarrow \neg q2$ 的语义真值不一致，不符合“ $\neg \neg P \Rightarrow P$ ”的标准，因此从形式语

义学来看，紧缩语义结构不是双重否定结构。

q1	q2	¬q1	¬q2	q1→q2	¬q2→¬q1	¬q1→¬q2
T	T	F	F	T	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	F
F	F	T	T	T	T	T

Table 2: 紧缩结构语义真值表

最后述宾结构的“不V1...不V2”指“不觉得不好”、“不知道你不来”等V1为述语，“不V2”为宾语的结构。该类结构中的“不...不...”可以理解为肯定。例如，“不认为他明天不会来”可以理解为“认为他明天会来”，“不觉得这件事不妥”可以理解为“觉得这件事妥”，等等。由此，我们可以得出“不V1...不V2”构成双重否定结构的第一个条件：“不V1”与“不V2”构成述宾关系。

3.1.2 第二个条件：V1为非叙实动词

述宾结构的“不V1...不V2”中只有一部分成员，其“不V1”对“不V2”有管辖作用，属于双重否定结构，其他成员仍只表示单纯的否定。试看下列：

- | | |
|------------|-----------|
| 1a.我不知道他不来 | 1b.*我知道他来 |
| 2a.我不幻想他不来 | 2b.*我幻想他来 |
| 3a.我不认为他不来 | 3b.我认为他来 |

通过例句，可以发现，当V1为“认为”时，“不V1...不V2”可以理解为“V1...V2”，而当V1为“知道”、“幻想”时却不能。同样是动词，“知道”、“幻想”、“认为”却存在着区别。本文认为，“不V1”对“不VP”是否有管辖作用与V1的语义有关，具体来说与V1的叙实性有关。

李新良（2015）将叙实性定义为“叙实性是动词的一种语义功能，即动词预设其宾语小句真值的能力。具体来说，肯定式和否定式都预设其宾语小句为真的动词是叙实动词.....肯定式和否定式都不预设其宾语小句为真，也不预设其宾语小句为假的动词是非叙实动词.....肯定式和否定式都预设其宾语小句为假的动词是反叙实动词”。对于叙实动词和反叙实动词来说，由于其预设固定，无论主句有无否定，宾语小句的真值都不变，因此主句无法影响宾语小句的真值，不构成“ $\neg \neg P \Rightarrow P$ ”。而对于非叙实动词（如：认为）来说，由于非叙实动词对宾语小句并没有预设，在述宾结构中，主句中的V1可以对宾语的真值造成影响，存在“ $\neg \neg P \Rightarrow P$ ”的可能。因此，我们可以得出“不V1...不V2”构成双重否定结构的第二个条件：V1为非叙实动词。

3.1.3 第三个条件：“不V1”的否定焦点包含V2

除上述两个条件外，结构中否定焦点的情况也会对“不V1...不V2”是否为双重否定造成影响。试看下列：

- | | |
|--------------------------|--------------|
| 4a.我不认为他不来。 | 4b.我认为他来。 |
| 5a.我不认为他故意不来。 | 5b.*我认为他故意来。 |
| 6a.我不相信他不喜欢我。 | 6b.我相信他喜欢我。 |
| 7a.我不相信他不喜欢我到了看见我就恶心的地步。 | |
| 7b.*我相信他喜欢我到了看见我就恶心的地步。 | |

在例句中，4a、6a可以转换为4b、6b，而5a、7a却不能转换为5b、7b。本文认为这主要与否定的焦点有关。袁毓林(2000)指出“有的成分表达的是句子的预设意义，属于旧信息，事实上它们的意义在否定的情况下仍然得以保持；有的成分表达的是句子的焦点意义，属于新信息，它们是真正被否定的。”当“不V1”的否定焦点不落在“不V2”上时，“不V1”对“V2”不

造成否定，不能构成“ $\neg \neg P$ ”结构，因此无法满足“ $\neg \neg P = P$ ”的条件，不属于双重否定。如“我不认为他故意不来”中的“不V1”否定的是“故意”而不是“不来”，其中“不来”是预设成分，属于旧信息。无论是“认为他故意不来”还是“不认为他故意不来”，语义都是“他不来”。“V1认为”的否定无法影响到“不来”的真值，无法构成“否定+否定”的语义结构。因此，我们可以得出“不V1...不V2”构成双重否定结构的第三个条件：“不V1”的否定焦点落在“不V2”上。

综上所述，“不V1...不V2”需要同时满足①“不V1”与“不V2”构成述宾关系、②V1为非叙实动词、③“不V1”的否定焦点落在“不VP”，三大成立条件，才可构成双重否定结构，表示肯定语义。

3.2 “不是...不V2”双重否定结构的成立条件与识别策略

如章节开头所述，“不是...不V2”若要表示双重否定，除需满足上述“不V1...不V2”的成立条件外，还需保证“不是”为描述性真值否定（descriptive truth-functional negation，又称语义否定），而不是元语否定（non-truth-functional negation，又称语用否定）。“所谓元语否定，就是用元语言对对象语言所描述的非真值语义的否定，...是一种非真值意义否定；与之相对应的是真值否定，否定的是句子的真值条件（truth condition）”¹这种否定常常是引述性否定，是对之前对话中已出现的内容的否定。例句如下：

- 8a. “可现在杀他不容易啊。”有人说。不是不容易，是根本不可能。
8b. * “可现在杀他不容易啊。”有人说。是容易，是根本不可能。

例中否定形式的8a“不是不容易，是根本不可能”不能理解为相应的肯定形式8b“是容易，是根本不可能”。这是因为例8中的“不是”是语用否定，并不对语义真值产生影响，不构成“否定+否定”的语义结构，无法通过两次否定表示肯定。因此，为了识别“不是...不V2”中的双重否定结构，首先需要区分“不是”是语义否定还是语用否定。为此，我们对大量语料进行了考察。我们发现，“不是”的否定情况具体有以下三种：

1. 当“不是”的上文中没有出现“不是”否定的观点时，“不是”只能是语义否定，而不能是语用否定。例如：

- 9a. 甲：你明天来不来参加生日聚会？
9b. 乙：不是我不乐意，而是我这边实在没时间。
语义否定✓ → 我是乐意，（但）是我这边实在没时间。
语用否定× → * “不乐意”这个表述不恰当，是我这边实在没时间。

2. 当“不是”的上文中出现了“不是”否定的观点，且“不是”所在句的后文与被否定观点的真值一致，则“不是”只能是语用否定，而不能是语义否定。例如：

- 10a. 甲：你不乐意明天来参加生日聚会？
10b. 乙：我不是不乐意，而是超级反感。
语义否定× → *我是乐意，是超级反感。
语用否定✓ → “不乐意”这个表述不恰当，（应该）是超级反感。

3. 当“不是”的上文中出现了“不是”否定的观点，且“不是”所在句的后文与被否定观点的真值不一致，则“不是”既可看作语义否定，又可看作语用否定。对于这种情况，从计算机处理的角度出发，我们可以统一将其处理为“不是”表示语义否定。

- 11a. 甲：你不乐意明天来参加生日聚会？
11b. 乙：不是我不乐意，而是我这边实在没时间。
语义否定✓ → 我是乐意，（但）是我这边实在没时间。
语用否定✓ → “不乐意”这个表述不对，（应该）是我这边实在没时间。

综上所述，只有当“不是”的上文中出现了“不是”否定的观点，且“不是”所在句的后文与被否定观点的真值一致时，“不是”才会是语用否定，而其余情况，皆可被计算机视为语义否定。

根据上述条件，我们可以通过计算情感真值的方法，来约束“不是”为语义否定。在文本里，绝大多数表示语用否定的“不是”都只出现在“不+是+不+x，（而）是+y”的结构中。因此对于“不是”是语义否定还是语用否定，我们可以拟定识别策略如下：

¹何爱晶.反叙的非真值义否定和真值义肯定[J].外语研究,2019,36(04):第25页.

4 双重否定自动识别程序的建立

4.1 词库的建立

为了使计算机能够识别助动词、非叙实动词、情感真值，本文对助动词、非叙实动词与情感词进行了梳理，在常用的基础词表中补充了助动词词表、非叙实动词词表与情感词表。助动词方面，本文以郑贵友（1989）整理的助动词范围为基本，结合鲁晓琨（2004）等前人的研究以及现代汉语的使用情况，选取了23个助动词，构成常用助动词词表。具体如下：

能、能够、可能、会、可以、应该、应、应当、一定、要、得、愿意、愿、肯、可、想、要、敢、该、当、准、许、容

非叙实动词方面，结合袁毓林（2014）、李新良（2015）等人对非叙实动词的研究，本文认为非叙实动词多为心理动词。因此本文对心理动词进行了考察。若一个心理动词的宾语的真值无法确定，则该心理动词为非叙实动词。以此为标准，本文对心理动词进行了考察，筛选出了11个非叙实动词³。整理如下：

认为、说、感到、觉得、允许、同意、相信、愿意、希望、考虑、打算

情感词方面，我们结合知网hownet情感词典、台湾大学NTUSD简体中文情感词典与清华大学李军中文褒贬义词典，设计了情感词表，共收纳正面词语10323个，负面词语9411个。

根据上述双重否定结构的识别策略，我们设计编写了双重否定结构自动识别程序。通过该程序，对含有两个否定词的语料文件，进行自动识别实验，输出其中存在的双重否定句以及相应的双重否定结构。程序输出结果示例如下（在程序中，我们还做了双重否定结构的语气识别，由于篇幅原因，本文暂不对其进行讨论。）：

```
此句话为双重否定句，双重否定结构为“无不”，语气为--语气强，加强肯定语气--：*****19...尽的义务。世界各国的义务教育年限有长有短，培养目标的提法各有不同，但是无一不是为了培养
此句话为双重否定句，双重否定结构为“不+助动词+不”，语气为--语气强，加强肯定语气--：*****22...切社会的生存和发展的基础，究竟应当把年轻一代培养成怎样的人，不能不现实社会生
此句话为双重否定句，双重否定结构为“不+助动词+不”，语气为--语气强，加强肯定语气--：*****38...由于基层财政困难，一些农村学校的校长为了保证学校的正常运转，不得不四处筹钱，被
此句话为双重否定句，双重否定结构为“不+助动词+不”，语气为--语气强，加强肯定语气--：*****39...门的练习题、测验卷、考试宝典等，教师逼着买，家长主动买，学生不得不买，于是学生
此句话为双重否定句，双重否定结构为“不+助动词+不”，语气为--语气强，加强肯定语气--：*****40：虽然教育管理者口头上不得不要大搞素质教育，但是他们的内心深处依然深深热恋着
此句话为双重否定句，双重否定结构为“不+助动词+非叙实动词+不”，语气为--语气强，加强肯定语气--：*****60...样轻轻扶去”在如此艰难的情况下，乌克思对自己情绪的把握和管理不能说不
```

Figure 4: 程序输出结果示例图

5 双重否定自动识别实验

5.1 实验语料来源

我们从CCL语料库中，按照各类结构的情况，进行了同等分布提取（即根据各结构在CCL语料库中的语料数量比例进行提取），收集了100000条初始语料。然后，通过程序，对上述语料进行筛选，排除了所有未被成功分句，长度超过150字的句子，最终获得测试语料96281句。

5.2 初始实验

我们使用双重否定结构自动识别程序对96281句语料进行识别，获得了8640句计算机认为含有双重否定结构的句子。由于人力有限，我们无法对8640个句子都进行人工检校。因此，为了计算精确率，我们从程序识别出的句子中随机抽取了1000句进行检测，经过人工检校，以上1000个句子中判断正确的句子为992句，由此我们可以计算，程序识别的精确率约为99.20%；在CCL“否定词+否定词”语料中，双重否定句的比例约为8.9%。

为了进一步验证，我们从96281句语料中随机抽取了1000语料进行检测。通过人工校验，上述1000句测试语料中存在92句含有双重否定结构的句子。然后，我们将上述1000句测试语料输入到双重否定自动识别程序中，程序从中识别出了90句含有双重否定结构的句子。根据计

³非叙实动词的界限并不是完全清晰的。非叙实动词与叙实动词、反叙实动词之间还存在一定的渗透性。由于其情况较为复杂，且对本文研究的影响较小，故暂不讨论。

算，程序的召回率约为97.83%。根据F1值公式，该程序的F1值约为98.51%。CCL“否定词+否定词”语料中双重否定句的比例仍在9%左右。

5.3 正式实验

初始实验表明，上述96281句语料中双重否定结构与非双重否定结构的比例差距很大（约为9: 91），因此为了更好地计算程序的精确率与召回率，我们人工构建了2000句双重否定语料，其中1000句为含有双重否定结构的正例，1000句为不含双重否定结构的负例。我们将语料输入双重否定结构程序中，具体结果如下表：

	1000 正例	1000 负例	总精确率	总召回率	F1 值
	识别出的双重否定句数				
本文程序	989	12	98.80%	98.90%	98.85%
王勇等 (2014)	777	10	98.72%	77.70%	86.96%
封洋 (2016)	1000	1000	50.00%	100%	66.67%

Table 4: 实验结果数据表

本文程序识别的精确率约为98.80%，召回率约为98.90%，F1值约为98.85%，实验结果较王勇等 (2014)、封洋 (2016) 的结果有较明显提升。需要说明的是，本文的实验为封闭测试，检验数据的方式是抽样，且文章对于双重否定结构的判断均来自于作者本人，因此结果会存在一定的偏差。后续我们还会投入更多的时间与人力，来获取更为准确的数据。

5.3.1 实验结果分析

无论是精确率还是召回率，实验的准确率都是在百分之九十多，未达到百分之百。通过分析，我们发现程序识别与召回错误的主要原因与句子的分词与句法分析错误有关。由于分词与句法分析等基础自然语言处理工具的问题，程序对一些句子的句法判断错误，导致一些原本应被判为并列关系、因果关系的成分，被误判为述宾关系，从而使整个双重否定结构的判断错误。示例如下：

34743:...不大紧；有的急于求成把将来要办的事情，拿到今天来办，由于条件不允许迟迟开展不了。

(程序识别结果：双重否定结构为“不+非叙实动词+...不”)

(实际情况：“条件不允许”与“迟迟开展不来”是因果并列关系)

6893:...要的网络公司均未能达到阿尔诺的预期。为此，去年6月底，阿尔诺不得不刹车。他说，他要考虑”战略调整”。

(程序分词：阿尔诺不得不/刹车 实际分词：阿尔诺/不得不/刹车)

当我们输入人工修改后的分词与句法分析结果后，程序的错误得到纠正，精确率与召回率皆可达到100%。

6 双重否定结构成立条件的测试实验

为了测试上述三个条件在双重否定结构识别过程中的作用，进一步验证语言学知识在双重否定结构自动识别过程中的贡献，我们将人工构建的2000句双重否定语料作为输入，测试在取消某一条件后，双重否定结构识别程序的识别情况与召回情况。具体结果如下表：

可以看到，“构成述宾结构”与“动词为非叙实动词”对整个双重否定结构的识别造成较大影响。尤其在召回率方面，相较应我们提供的标准正确数据，没有“构成述宾结构”条件约束的程序较原始结果将额外召回约37倍的错误结构（460: 12），而没有“动词为非叙实动词”条件约束的程序也将额外召回1.4倍的错误结构（29: 12）。

相较来说，“否定焦点”条件会对双重否定结构识别的影响最为轻微。我们认为，这是因为人们在实际语言交流中很少会使用非常复杂的句子（例如：“我不相信他不喜欢我到了看见我就恶心的地步”）。当我们扩大检测数据，用程序对96281句原始语料进行测试时，缺少否定焦点

	1000 正例	1000 负例
	识别出的双重否定句数	
原始数据	989	12
取消“述宾结构”条件限制	1078	460
取消“非叙实动词”条件限制	1031	29
取消“否定角度”条件限制	989	13

Table 5: 双重否定结构成立条件测试实验结果

的程序将会比标准程序额外召回140句双重否定结构，进一步说明否定焦点会对双重否定结构识别程序造成影响，只是由于语料中复杂的句子很少，因此影响较轻微。这种情况也符合我们日常表达的经济性原则。

7 结语

本文以“ $\neg \neg P \Rightarrow P$ ”为标准，借助计算机程序与CCL语料库，对现代汉语中所有的“否定词+否定词”结构进行了遍历研究，实现了以下目标：

1. 将双重否定结构按照格式分为了3大类，25小类，常用双重否定结构或构式132个，进一步地揭露了双重否定结构的全貌；
2. 总结得出了双重否定结构成立的三大条件；并对其进行了实验测试分析，进一步验证了语言学知识在双重否定结构中的作用；
3. 补充了助动词表、非叙实动词表、情感词表等基础词表，编写实现了双重否定结构自动识别程序。程序的识别的精确率约为98.80%，召回率约为98.90%，F1值约为98.85%；
4. 获得了8640句精确率约为99%的标明双重否定结构的句子，为后续建立双重否定语料库提供了支持。具体文件烦请参见脚注链接⁴。

本文还有许多不足未尽之处，例如：本文的识别程序是在规则的基础上建立的，而人为编写的规则未免有不尽之处；本文对双重否定结构的判断皆基于作者个人的语感，未免有疏漏之处；对于一些已经固化的双重否定结构，如“非...不...”、“无非”、“莫非”等，本文的处理还较为粗糙，有待进一步分析与探索，等等。

未来我们拟在本文的基础上，继续展开与深入，具体计划有：

1. 建立双重否定语料库，在程序的辅助下构建数万句级的双重否定语料库，工程量应能控制在数十小时；
2. 探究双重否定结构自动识别的深度学习模型。
3. 对双重否定结构的语用方面进行进一步研究与探索。
4. 对双重否定结构中的构式，例如“非...不可”、“非...莫属”等等，进行进一步的研究与探索。

参考文献

- Horn, L. 1989. *A Nature History of Negation*. University of Chicago Press, Chicago, US, 311-312.
- Ladusaw, W. A. 1997. *Negation and polarity items*. In S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*. Blackwell Publishing Ltd, Oxford, UK, 321-341.
- 丁声树等. 2004. 现代汉语语法讲话. 商务印书馆, 北京, 200-202.
- 方绪军. 2017. “不是不X”、“不是没(有)X”和“没(有)不X”. 语言科学, 16(05):511-521.
- 封洋. 2016. 中文微博情绪分析. 上海交通大学, DOI:10.27307/d.cnki.gsjt.2016.000526.
- 符达维. 1986. 对双重否定的几点探讨. 福建论坛(文史哲版), (6):78-81.

⁴<https://download.csdn.net/download/qq-43342081/86399912>

- 何爱晶. 2019. 反叙的非真值义否定和真值义肯定. 外语研究, 36(04):24-29.
- 郎桂青. 1989. 双重否定句表示肯定的条件. 语文研究, (1) :28.
- 李新良, 王明华. 2015. 汉语动词的叙实性研究的应用前景. 对外汉语研究, (02):120-129.
- 鲁晓琨. 2004. 现代汉语基本助动词语义研究. 中国社会科学出版社, 北京.
- 吕叔湘. 1956. 中国语法要略. 商务印书馆, 北京.
- 王力. 1943. 中国现代语法. 商务印书馆, 北京.
- 王勇, 吕学强, 姬连春. 2014. 基于极性词典的中文微博客情感分类. 计算机应用与软件,(01):40-43+132.
- 芜崧. 1987. 双重否定句的种类与功能. 荆州师专学报(哲社版), (3): 52-57.
- 叶文曦. 2013. 否定和双重否定的多维度研究. 语言学研究,(2):20-31.
- 袁毓林. 1999. 并列结构的否定表达. 语言文字应用,(03):42-46.
- 袁毓林. 2000. 论否定句的焦点、预设和辖域歧义. 中国语文,(02):99-108+189.
- 袁毓林. 2014. 隐性否定动词的叙实性和极项允准功能. 语言科学,13(6):575-586.
- 郑贵友. 1989. 汉语“助动词”的研究刍议. 汉语学习,(06):23-27.