

文章编号: 1003-0077(2018)03-0042-07

## 基于最长名词短语分治策略的神经机器翻译

张学强, 蔡东风, 叶 娜, 吴 闯

(沈阳航空航天大学 人机智能研究中心, 辽宁 沈阳 110136)

**摘 要:** 神经机器翻译自兴起以来, 不断给机器翻译领域带来振奋人心的消息。但神经机器翻译没有显式地利用语言学知识对句子结构进行分析, 因此对结构复杂的长句翻译效果不佳。该文基于分治法的思想, 识别并抽取句子中的最长名词短语, 保留特殊标识或核心词, 与其余部分组成句子框架。通过神经机器翻译系统分别翻译最长名词短语和句子框架, 再将译文重新组合, 缓解了神经机器翻译对句子长度敏感的问题。实验结果表明, 该方法获得的译文与基线系统相比, BLEU 分值提升了 0.89。

**关键词:** 神经机器翻译; 最长名词短语; 分治策略

**中图分类号:** TP391

**文献标识码:** A

### Divide-and-conquer Strategy of Maximal NP for NMT

ZHANG Xueqiang, CAI Dongfeng, YE Na, WU Chuang

(Human-Computer Intelligence Research Center, Shenyang Aerospace University,  
Shenyang, Liaoning 110136, China)

**Abstract:** Neural Machine Translation (NMT) is defected in long sentences with complex structure owing to its neglect of linguistic knowledge of sentence structure. Adopting the idea of divide-and-conquer strategy, this paper proposes to identifying and extracting the Maximal Noun Phrases in a sentence, and retaining special marks or head words and the rest component to form the sentence framework. Then the Maximal Noun Phrases and sentence frames are translated by NMT, respectively. Experimental results show that the method proposed yields 0.89 improvements in terms of BLEU score compared with the baseline system.

**Key words:** neural machine translation; maximal-length noun phrase; divide-and-conquer strategy

## 0 引言

神经机器翻译 (neural machine translation, NMT) 作为一种全新的机器翻译方法, 近年来获得迅速发展。然而, 神经机器翻译仅仅使用一个非线性的神经网络实现自然语言之间的转换<sup>[1]</sup>, 相比统计机器翻译, 译文质量对句子长度更为敏感<sup>[2]</sup>。如何在神经机器翻译中将一个句子在尽量不损失语义信息的前提下, 进行长度上的缩减和结构上的简化是一个值得探究的方向。

一般认为, 自然语言中语义的基本单位是短语。

因此, 将句子级别的对齐和翻译进行到亚句子 (sub-sentence) 的短语一级显得尤为重要。句子中的实体和概念通常可由名词短语 (noun phrase, NP) 来描述, 其捆绑了一个相对完整的语义信息, 具有丰富的句法功能, 可在句中充当主语和宾语等成分。最长名词短语<sup>[3]</sup> (maximal-length noun phrase, MNP) 指不被其他任何名词短语嵌套的名词短语。与一般名词短语相比, MNP 具有更大的粒度, 边界特征较为明显, 有利于句子的整体结构分析。采用分治策略处理 MNP, 既能在亚句子一级上获得更精准的翻译, 也在一定程度上将句子缩短为包含主干信息的句子框架。因此, 准确识别和翻译 MNP, 是利用分治策略提升机器翻

译性能的一个有力手段。

针对神经机器翻译在长句翻译任务上的不足,考虑到 MNP 的处理可以在一定程度上简化句子结构,本文提出一种基于 MNP 分治策略的神经机器翻译方法。该方法基于一个“抽取—翻译—重组”的 MNP 处理框架,旨在将 MNP 独立处理带来更高质量的 MNP 和句子框架译文的优势,与神经机器翻译学习能力强、译文具有较高准确度和流畅度等优势相结合,以达到提升译文整体质量的目的。

## 1 相关研究

### 1.1 短语知识在机器翻译中的应用

在自然语言中,短语作为语义的基本单位,具有重要的意义。将双语短语等语言学知识融入机器翻译中,一直是研究人员孜孜追求的目标。

针对基于短语的统计机器翻译方法未充分利用语言学知识、长距离调序效果不好的问题,丁鹏<sup>[4]</sup>等提出一种基于双语句法短语的统计机器翻译方法。首先,采用一种基于期望最大化(expectation maximization, EM)的算法来抽取双语句法短语。然后,通过三种方法将短语应用到统计机器翻译系统中:(1)将双语句法短语加入训练语料中,训练翻译模型;(2)将其加入短语表中,计算短语的特征值;(3)增加一个句法短语特征到短语表中,表征其是否为句法短语。实验结果表明,这三种方法得到的译文 BLEU 分值分别比基线系统提升了 0.23、0.41 和 0.64。丁鹏等人的方法尽管利用了双语句法短语,但整体框架仍然是基于短语的统计机器翻译方法,长距离调序效果不佳。

针对上述问题,Ren X 等<sup>[5]</sup>提出一种简化专利句子结构以提高翻译性能和后处理效率的方法。首先,采用一种基于统计方法的识别器,对句中的 MNP 进行识别。在中文树库 CTB 5.1 的专利语料上识别结果的 F 值达到 62.28%。然后,对 MNP 进行分析,在识别正确与错误的 MNP 中,分别有 97.92% 和 38.94%,有利于后续的翻译过程。最后,在统计机器翻译方法上分别使用自动方法和人工方法对系统进行评价。与基线系统相比,该系统得到的译文 BLEU 分值提升了 0.62;语义准确度和流畅度分别提升 0.18 和 0.17,翻译效率提升了约 100 字/小时。该方法的

不足在于,没有使用双语 MNP 扩展语料,以训练短语表、翻译模型和调序模型。MNP 作为句子的一部分,翻译规则却与句子不尽相同。导致训练得到的模型能较好地翻译简化后的句子,却不能准确翻译 MNP。

尽管神经机器翻译在英文—法文等语言对翻译任务上表现出最佳的翻译性能<sup>[6-8]</sup>,但其面临着难以融入词典等语言学知识的挑战。Zhang J 等<sup>[9]</sup>提出一种能够桥接神经机器翻译与双语词典的框架,将包含训练数据中的低频词与未登录词等信息的双语词典应用到神经机器翻译中的方法,从而提升翻译性能。为此,Zhang 等人针对该框架采用了两种模型:(1)混合词/字模型(mixed word/character model),将语料中频率超过一定阈值的词保留,而低频词和集外词则使用其字符序列进行替换;(2)伪句对合成模型(pseudo sentence pair synthesis model),针对低频词对或未登录词对( $D_{xi}, D_{yi}$ ),构造  $J$  组包含( $D_{xi}, D_{yi}$ )的句对 $\{(X_{ij}, Y_{ij})\}_{j=1}^J$ 。实验结果表明,与基线系统相比,混合词/字模型与伪句对合成模型的译文 BLEU 分值分别提升了 1.0 和 1.62。Zhang 等人的方法缓解了神经机器翻译方法的未登录词问题,但未能使用短语知识缓解译文质量对句子长度敏感的问题。

### 1.2 神经机器翻译

统计机器翻译(statistical machine translation, SMT)主要存在三个挑战<sup>[10]①</sup>:(1)线性不可分;(2)缺乏合适的语义表示;(3)难以设计特征。而深度学习可以较好地缓解上述问题,因此完全基于深度学习的端到端神经机器翻译应运而生,并获得迅速发展。

研究人员通过将现有的方法和策略引入端到端的神经网络,以实现翻译性能的不断提升。Sutskever 等<sup>[11]</sup>首次将长短期记忆<sup>[12]</sup>(long short-term memory, LSTM)引入到神经机器翻译,以缓解递归神经网络(recurrent neural network, RNN)训练时“梯度消失”的问题,并且在“编码—解码”(encoder-decoder)框架两端同时采用递归神经网络。图 1 给出了 Sutskever 等人提出的神经机器翻译模型。

① 实际上,引文中作者给出了六个挑战,这里只列举其三。

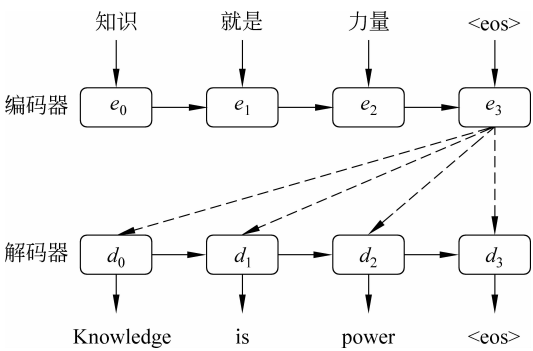


图1 端到端神经机器翻译,隐状态  $e_3$  作为句子向量  $c$

在源端,对于句子  $X=\{x_0, x_1, x_2, x_3\}$ ,编码器递归地依据前一时刻隐状态  $e_{t-1}$  和词  $x_t$  计算当前时刻隐状态  $e_t$ 。直到扫描尾词  $x_n$  随即完成了编码过程,并将最后一个隐状态  $e_n$  作为表示源语言句子的向量  $c$ ,指导并约束后续解码过程。 $e_t$  的计算如式(1)所示。

$$e_t = g(e_{t-1}, x_t) \tag{1}$$

在目标端,解码器递归地依据向量  $c$  和已生成的目标词  $y_{t-1}$  以及上一时刻隐状态  $d_{t-1}$  共同作用于当前时刻隐状态  $d_t$ ,如式(2)所示。

$$d_t = h(d_{t-1}, y_{t-1}, c) \tag{2}$$

得到解码器隐状态  $d_t$  后,目标词  $y_t$  的概率分布可由式(3)得到。

$$p(y_t | y_{<t}, X) = \text{softmax}(f(d_t, y_{t-1}, c)) \tag{3}$$

其中, $g$ 、 $h$  和  $f$  为非线性函数。通过解码器递归地从左至右逐一生成目标词,最终得到完整译文  $Y=\{y_0, y_1, y_2, y_3\}$ 。尽管引入长短期记忆的神经机器翻译在性能上获得大幅提升,却面临着实现准确编码的挑战。因为不论句子长短,编码器都要将其映射为一个固定维度的向量。

针对上述问题,Bengio 等<sup>[13]</sup>提出了基于注意力(attention)的神经机器翻译。解码器在生成目标词  $y_t$  时,动态地注意源语言句中与之相关的上下文  $c_t$ ,而不再关注整个源语言句子。图2给出了引入注意力机制的神经机器翻译模型。

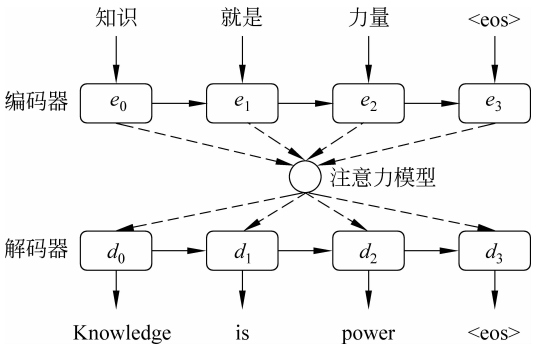


图2 基于注意力机制的神经机器翻译,动态生成上下文向量  $c$

引入注意力的神经机器翻译的关键在于基于注意力的上下文向量  $c$  的生成。当前时刻待生成词  $y_t$  在源端对应的上下文向量  $c_t$  由源语言隐状态序列  $e=\{e_0, e_1, e_2, e_3\}$  和注意力权重  $a_t$  加权求和得到,而注意力权重  $a_t$  由上一时刻解码器隐状态  $d_{t-1}$  和源端隐状态  $e_j$  共同作用产生。如式(4)~(6)所示。

$$c_t = \sum_{j=1}^n a_{t,j} e_j \tag{4}$$

$$a_{t,j} = \frac{\exp(b_{t,j})}{\sum_{k=1}^n \exp(b_{t,k})} \tag{5}$$

$$b_{t,j} = m(d_{t-1}, e_j) \tag{6}$$

其中, $m$  为非线性函数。得到当前时刻上下文向量  $c_t$  后,当前时刻解码器隐状态  $d_t$  与待生成词  $y_t$  的条件概率分布分别可由式(2)和式(3)求解。

尽管长短期记忆和注意力机制的引入能够更好地处理长距离依赖,从而提升神经机器翻译的性能。然而,自然语言中句子长短不一、结构复杂,通过单一神经网络学习翻译知识的方法受到限制。如何利用语言学知识结合分治策略对句子的各部分进行分治与整合,是一个值得研究的问题。

## 2 基于 MNP 分治策略的神经机器翻译

尽管神经机器翻译近年来获得了迅速发展,但目前的方法主要是从数据中自动学习翻译知识,没有充分利用语言学知识显式地指导翻译过程。并且,神经机器翻译使用固定维度的向量表示变化长度的词句,造成结构复杂的长句翻译效果不佳。

针对上述问题,本文提出一种基于 MNP 分治策略的神经机器翻译。该方法主要基于分治法的思想,采用一个“抽取—翻译—重组”的 MNP 处理框架,将单个复杂长句的翻译问题,转化为一个或多个携带子句信息的 MNP 和维系主干信息的句子框架的翻译问题,以实现翻译性能的整体提升。

### 2.1 “抽取—翻译—重组”框架

在分治策略中,通常将单个复杂问题转化为多个相对简单的问题,并分而治之。鉴于 MNP 在句中使用频率高、句法功能丰富以及边界易于识别等事实,本文主要基于“抽取—翻译—重组”的 MNP 处理框架以实现分治策略的神经机器翻译。表1给出了该方法的完整示例。

表 1 “抽取—翻译—重组”框架示例

抽取	原句	儿童 基金会 为 难民营 中 成千上万 个 流离失所 家庭 发放 了 紧急 现金 援助 。
	短语结构句法分析	(( (IP (IP (NP (NN 儿童) (NN 基金会)) (VP (PP (P 为) (NP (LCP (NP (NN 难民营)) (LC 中)) (QP (CD 成千上万) (CLP (M 个))) (NP (NN 流离失所) (NN 家庭)))) (VP (VV 发放) (AS 了) (NP (ADJP (JJ 紧急)) (NP (NN 现金) (NN 援助)))))) (PU。) ) ) )
	句子框架与 MNP	主干：MNP <sub>1</sub> 为 MNP <sub>2</sub> 发放 了 MNP <sub>3</sub> 。 MNP <sub>1</sub> ：儿童 基金会 MNP <sub>2</sub> ：难民营 中 成千上万 个 流离失所 家庭 MNP <sub>3</sub> ：紧急 现金 援助
翻译	句子框架与 MNP 译文	主干：MNP <sub>1</sub> has provided MNP <sub>3</sub> to MNP <sub>2</sub> , and UNHCR has provided MNP <sub>4</sub> to MNP <sub>5</sub> . MNP <sub>1</sub> ：Children’s Fund MNP <sub>2</sub> ：thousands of internally displaced families in camps MNP <sub>3</sub> ：emergency cash assistance
	完整译文	Children’s Fund has provided emergency cash assistance to thousands of internally displaced families in camps.
重组	完整译文	Children’s Fund has provided emergency cash assistance to thousands of internally displaced families in camps.

在示例中，抽取 MNP 时在句子框架中保留特殊标识“MNP<sub>*i*</sub>”( *i* = 1, 2, ... )。作为对比，本文还使用了在句子框架中保留 MNP 核心词的方法。将在 2.3 节、2.4 节和 2.5 节中逐一说明“抽取—翻译—重组”框架的三个步骤，并对抽取 MNP 时保留特殊标识或 MNP 核心词的方法作出详细论述。

2.2 双语 MNP 语料库的构建

本文采用神经机器翻译系统分别对 MNP 和句子框架进行翻译。因此，双语 MNP 语料库的构建是其中重要的一个环节。为保证训练和测试过程中 MNP 的抽取规则一致，本文没有采用双语 MNP 对齐算法进行抽取，而是采用一个“抽取+查表”的方法。步骤描述如下：

(1) 使用分析器对源语言句子进行短语结构句法分析，依据标记匹配和括号对齐等规则抽取 MNP。

(2) 训练并查找短语表，匹配其中与源语言 MNP 对齐分值最高的目标语言 MNP。

上述方法的优势在于每一步都可以加入规则条件，以获得较高质量的双语 MNP。本文在抽取源语言 MNP 以及查找短语表匹配其对应的目标语言 MNP 时，过滤掉长度小于 2 或包含符号、标点等特殊字符的 MNP。得到双语 MNP 后，神经机器翻译系统的训练和测试过程如下：

首先，将双语 MNP 分别加入训练数据集和开发数据集中，利用扩展后的数据集训练神经机器翻译模型。这一做法旨在得到能同时翻译句子和 MNP 的神经机器翻译模型。

其次，对测试数据集进行同样的短语结构句法

分析，抽取 MNP 的同时在句子框架中保留特殊标识或 MNP 核心词。

最后，分别对句子框架和 MNP 进行翻译，将译文重新组合以得到原句的完整翻译。

图 3 给出了基于“抽取—翻译—重组”框架的神经机器翻译系统翻译的过程。考虑到短句子译文质量原本较高，本文只对长度超过阈值 *L* 且可成功抽取 MNP 的句子采用基于“抽取—翻译—重组”框架的分治策略进行处理。

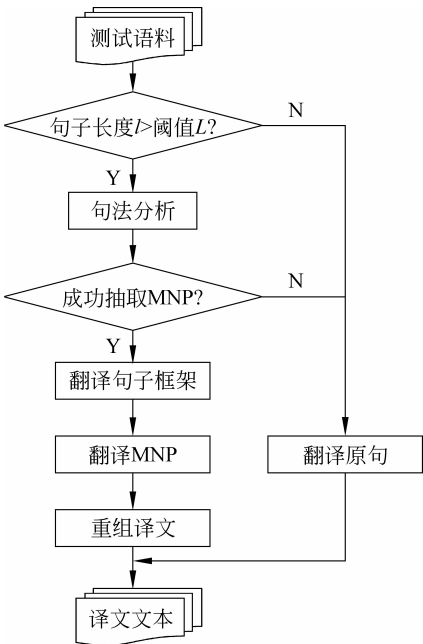


图 3 神经机器翻译系统的“抽取—翻译—重组”过程

2.3 抽取

抽取过程的核心任务是对句子进行短语结构句



给出了测试语料的相关信息。

表 3 测试语料信息

信息	Test
句子数	4 000
成功抽取 MNP 的句子数	1 924
句子框架数	1 924
MNP 数	5 001
句子平均长度	25.86
成功抽取 MNP 的句子平均长度	39.24
句子框架平均长度	19.60
MNP 平均长度	12.14

从表 3 可以看出,相比于成功抽取出 MNP 的句子平均长度,MNP 和句子框架的平均长度分别缩短了 19.64 和 27.10。

3.2 参数设置

本文主要在深度学习框架 Theano 上采用 DL4MT<sup>①</sup> 开源代码,搭建基于注意力机制的神经机器翻译系统。表 4 给出了实验中神经网络的主要参数设置及部分说明。

表 4 网络参数设置及说明

参数	值	说明
网络层数	3	—
隐藏层节点数	1 024	—
词向量维度	600	—
Mini-batch 大小	64	—
学习率	0.5	由损失控制衰减
学习率衰减因子	0.99	—
源语言词表大小	80 000	包含 eos 和 UNK
目标语言词表大小	60 000	包含 eos 和 UNK
节点类型	GRU	—

表 4 中,eos 和 UNK 是置于词表首位的特殊词。将 eos 追加在句尾,表示句子结束。当编码器扫描到 eos 时结束编码,同样地,当解码器生成目标词 eos 时,终止解码过程。由于网络训练过程中 softmax 函数的计算复杂度较高,而其与词表规模成正相关,因此词表大小受到限制。考虑到集外词对神经机器翻译系统的性能影响较大<sup>[14]</sup>,本文将集外词统一替换为特殊词 UNK。

在网络训练过程中,采用随机梯度下降(stochastic gradient descent,SGD)算法进行参数更新。模型测试时,本文采用束搜索(beam search)算法生成最终译文,束大小设置为 10。

3.3 结果与分析

3.3.1 MNP 抽取

本文采用一种基于 MNP 分治策略的神经机器翻译方法,因此,能否准确识别 MNP 直接影响到系统的翻译性能。本文从成功抽取 MNP 的 1 924 个句子中随机抽取 200 句,并对句中的 MNP 进行人工标注。通过比对系统的 MNP 抽取结果和人工标注结果,可计算得到系统 MNP 识别的准确率、召回率、F 值,如表 5 所示。

表 5 MNP 识别结果

准确率/%	召回率/%	F 值/%
73.59	72.42	73.01

由表 5 可以看出,约 27% 的 MNP 识别存在错误。但边界错误的 MNP 并不全都给后续的翻译过程造成消极影响<sup>[5]</sup>。

3.3.2 句长敏感度

为验证句子长度对于译文质量的影响,本文分别在基线系统和 MNP 分治系统上,对测试数据集中的句子按照不同的长度分布进行测试。其中,基线系统指未采用“抽取—翻译—重组”的 MNP 处理框架的神经机器翻译系统。MNP 分治系统包含两种方法,即抽取 MNP 时在句子框架中保留特殊标识“MNP<sub>*i*</sub>”与保留 MNP 核心词。

本文采用 NiuTrans<sup>①</sup> 开源系统中集成的大小写不敏感的 4-gram BLEU 方法对译文质量进行自动评价。如图 5 所示,横坐标表示不同句长分布,纵坐标表示译文 BLEU 分值。

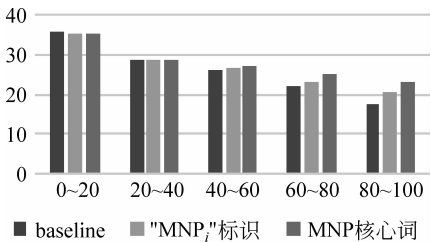


图 5 系统在不同句长分布上的翻译性能

① <https://github.com/nyu-dl/dl4mt-tutorial/>

由图 5 可以看出,随着句子长度的增加,译文质量呈明显下降趋势。特别地,当句子长度超过 20 后译文质量显著下降,基线系统的译文 BLEU 分值下降了 7.23,保留特殊标识“MNP<sub>i</sub>”方法和 MNP 核心词方法的译文 BLEU 分值分别下降了 6.55 和 6.31。

具体来看,主要有三点结论:(1)当句长小于 20 时,基线系统略优于 MNP 分治系统。原因分析如下:首先,神经机器翻译方法原本在短句上翻译性能较好。其次,MNP 分治系统在“抽取—翻译—重组”框架的三个步骤中都存在一定的损失,当这种损失与分治方法带来的提升持平时,分治系统的优势表现得并不明显。(2)当句长超过 20 后,随着句子长度的增大,MNP 分治系统越来越表现出更优的翻译性能。尤其当句长在 80 和 100 之间时,相比于基线系统,保留特殊标识“MNP<sub>i</sub>”和保留 MNP 核心词的方法,译文 BLEU 分值分别提升了 3.10 和 5.75。(3)保留 MNP 核心词的方法在翻译性能上优于保留特殊标识“MNP<sub>i</sub>”的方法,且随着句长的增大,优势愈发明显。

3.3.3 翻译性能

本文采用“抽取—翻译—重组”的 MNP 处理框架,对句子进行短语结构句法分析后抽取 MNP,并保留特殊标识或 MNP 核心词与其他部分组成句子框架。表 6 给出了基线系统、保留特殊标识“MNP<sub>i</sub>”以及保留 MNP 核心词的三种神经机器翻译系统的译文质量。

表 6 译文质量对比	
	BLEU/%
基线系统	28.87
保留“MNP <sub>i</sub> ”	29.23
保留 MNP 核心词	29.76

由表 6 可以看出,基于“抽取—翻译—重组”的 MNP 处理框架,抽取 MNP 时保留特殊标识“MNP<sub>i</sub>”和保留 MNP 核心词的方法在基线系统的基础上,都获得一定的提升。相比于基线系统,保留“MNP<sub>i</sub>”的方法 BLEU 分值提升了 0.36,保留 MNP 核心词的方法 BLEU 分值提升了 0.89。

在分治系统中,由于抽取 MNP 时在句子框架中保留了 MNP 的核心词,在一定程度上提高了句子框架的流畅度和语义完整性,从而相比于保留“MNP<sub>i</sub>”,表现出更好的性能,译文的 BLEU 分值提升了 0.53。

4 总结与展望

本文针对当前神经机器翻译方法的译文质量对句子长度敏感的问题,提出一种基于 MNP 分治策略的神经机器翻译方法。依据组块分析和分治法的思想,对长句进行 MNP 识别和抽取,进一步对 MNP 和句子框架进行独立翻译,从而在一定程度上缓解了神经机器翻译对句子长度敏感的问题。

实验结果表明,该方法通过对训练数据的扩展、翻译前对 MNP 的识别和抽取、翻译中对 MNP 和句子框架的分而治之、翻译后对译文的重组等策略给神经机器翻译带来积极的影响。相对基线系统的方法,BLEU 分值提升了 0.89。

然而,该方法在 MNP 抽取,句子框架与 MNP 的译文重组等方面都存在一定的损失,并且,诸如目标语言 MNP 的单复数等问题尚待解决。下一步研究工作的重心拟定在以下两个方面:首先,将该方法泛化到其他类型的短语结构,以对目前方法做进一步扩充;其次,因为过程中涉及对句子的拆分与整合,应更多地从语言学角度重新思考“抽取—翻译—重组”的分治策略,以采取更优的方法。

参考文献

[1] Zhang J, Zong C. Deep neural networks in machine translation: An overview[J]. IEEE Intelligent Systems, 2015, 30(5): 16-25.

[2] Cho K, Merriënboer B V, Bahdanau D, et al. On the properties of neural machine translation: Encoder-decoder approaches[J]//arXiv: 1409.1259. 2014.

[3] 蔡东风,赵奇猛,饶齐,等. 基于马尔科夫逻辑网的中文专利最大名词短语识别[J]. 中文信息学报, 2016, 30(4): 21-28.

[4] 丁鹏. 基于双语句法短语的统计机器翻译研究[D]. 大连: 大连理工大学硕士学位论文, 2013.

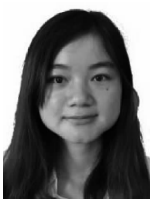
[5] Ren X, Wei Y, Hu R. Simplify sentence structure for improving human post-editing efficiency on Chinese-to-English patent machine translation[C]//Proceedings of 6th Workshop on Patent and Scientific Literature Translation (PSLT6) Miami, 2015: 33-43.

[6] Luong M T, Pham H, Manning C D. Effective Approaches to Attention-based Neural Machine Translation[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 2015: 1412-1421.

(下转第 63 页)

mation retrieval based on LSI with multiple word spaces[C]//Proceedings of the 2nd NTCIR Workshop Meeting on Evaluation of Chinese & Japanese Text Retrieval and Text Summarization, 2001.

[16] Bradford R, Pozniak J. Combining modern machine translation software with LSI for cross-lingual infor-



刘娇(1994—), 硕士, 主要研究领域为多语种自然语言文本信息处理。  
E-mail: mykongkongliu@foxmail.com



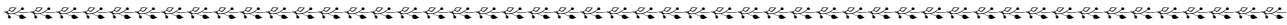
赵亚慧(1974—), 通信作者, 副教授, 主要研究领域为自然语言文本处理。  
E-mail: yhzhaoy@ybu.edu.cn

mation processing[C]//Proceedings of the Eleventh International Conference on Information Technology: New Generations, 2014: 65-72.

[17] 袁松翔. 基于译文特征与译文内容的中英文跨语种抄袭识别技术研究[D]. 上海: 上海交通大学硕士学位论文, 2011.



崔荣一(1962—), 博士, 教授, 主要研究领域为智能计算、模式识别、机器学习、自然语言处理。  
E-mail: cuirongyi@ybu.edu.cn



(上接第 48 页)

[7] Sennrich R, Haddow B, Birch A. Neural Machine Translation of Rare Words with Subword Units[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 2016: 1715-1725.

[8] Wu Y, Schuster M, Chen Z, et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation[J]. arXiv: 1609.08144

[9] Zhang J, Zong C. Bridging Neural Machine Translation and Bilingual Dictionaries[J]. arXiv: 1610.07272

[10] 刘洋. 基于深度学习的机器翻译研究进展[J]. 中国人工智能学会通讯, 2015: 28-32.

[11] Sutskever I, Vinyals O, Le Q V. Sequence to se-

quence learning with neural networks[J]. Advances in Neural Information Processing Systems, 2014(4): 3104-3112.

[12] Graves A. Long short-term memory[M]. Supervised Sequence Labelling with Recurrent Neural Networks. Springer Berlin Heidelberg, 2012: 1735-1780.

[13] Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. arXiv: 1409.0473

[14] Li X, Zhang J, Zong C. Towards zero unknown word in neural machine translation[C]//Proceedings of the International Joint Conference on Artificial Intelligence. AAAI Press, 2016: 2852-2858.



张学强(1992—), 硕士研究生, 主要研究领域为神经机器翻译、自然语言处理。  
E-mail: zxqfightingsau@163.com



叶娜(1981—), 博士, 讲师, 主要研究领域为机器翻译、自然语言处理。  
E-mail: yena\_1@126.com



蔡东风(1958—), 博士, 教授, 主要研究领域为人工智能、自然语言处理、信息检索。  
E-mail: caidf@vip.163.com