

## 基于动态词对齐的交互式机器翻译\*

马斌, 蔡东风, 季铎, 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:** 在传统的机器翻译 (Machine Translation, MT) 与计算机辅助翻译 (Computer Aided Translation, CAT) 中, 译员与翻译引擎之间的交互受到很大限制, 于是产生了交互式机器翻译 (Interactive Machine Translation, IMT) 技术。但传统的模型只考虑当前源语与部分翻译的目标语的信息, 没有将用户交互后的对齐信息加入到之后的预测模型中。本文基于词预测交互式机器翻译的研究思路, 将用户交互翻译过程中的鼠标点选行为转化为中间译文的词对齐信息, 进而在翻译交互过程中实现了对译文的动态词对齐标注, 并在词对齐信息和输入译文的约束下提高了传统词预测的准确性。

**关键词:** 交互式机器翻译; 词对齐; 预测模型

**中图分类号:** TP391

**文献标识码:** A

## Improving Interactive Machine Translation by Dynamic Word Alignment

MA Bin, CAI Dongfeng, JI Duo, YE Na, WU Chuang

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

**Abstract:** In traditional machine translation (MT) and computer aided translation (CAT), the interaction between human translators and machines is very limited. So the interactive machine translation (IMT) technology has emerged. The traditional models didn't add the alignment information obtained in the interaction of translators and machine into the subsequent prediction models and only considered the information of the current source language and the partial translation of the target language. In this paper, we propose a method based on the research of the word prediction interactive machine translation, translate the selection behavior of the mouse in the interactive process into the word alignment information of the partial translation, and then label the dynamic word alignment of the translation. The approach has improved the accuracy of prediction under the constraint of the dynamic word alignment and the input translation.

**Key words:** Interactive Machine Translation; word alignment; prediction model

### 1 引言

统计机器翻译 (Statistical Machine Translation, SMT) 技术的提出为机器翻译技术走向实用化打下了基础。但现有的自动翻译的质量还远没有达到人们的需求, 不能完全取代人类<sup>[1][2]</sup>。计算机辅助翻译

(Computer Aided Translation, CAT) 的提出, 为机器与人类译员提供合作的平台, 由计算机完成翻译过程中繁琐而相对简单的任务, 由人类译员完成相对重要的翻译任

务, 有效提高了翻译人员的效率。为进一步提高翻译效率, 增大人类译员在翻译过程中的主导地位, 提出了交互式机器翻译

(Interactive Machine Translation, IMT) 的概念, 允许人类译员在翻译过程中进行人工的干预和指导, 从而获取较高质量的译文<sup>[3]</sup>。近年来, 有研究人员将人机交互过程转变为在翻译记忆的参考译文上进行后编辑<sup>[4][5]</sup> (Post-Editing)。这种方法一方面比较依赖于翻译资料中重复内容, 另一方面也限制了人类译员在翻译过程中的交互。

\* **基金项目:** 国家自然科学基金 (61403262); 国家自然科学基金 (61402299)

**作者简介:** 马斌 (1990—), 男, 硕士研究生, 主要研究方向为自然语言处理, 机器翻译; 蔡东风 (1958—), 男, 博士, 教授, 主要研究方向为自然语言处理, 人工智能, 信息检索; 季铎 (1981—), 男, 副教授, 主要研究方向为自然语言处理, 机器翻译; 叶娜 (1981—), 女, 讲师, 博士, 主要研究方向为自然语言处理, 机器翻译; 吴闯 (1985—), 男, 硕士, 主要研究方向为自然语言处理, 机器翻译。

为此, Foster 等人开展了一个有趣的研究工作, 提出使用词预测 (word prediction) 技术来实现在人机交互中的预测式交互翻译 (interactive-predictive machine translation, IPMT), 该方法将用户的翻译行为看作是一个顺序的选词过程, 而系统需要在每一个翻译选词的过程中通过前缀不断预测下一译文, 每次译员的选择都影响着下一次机器的推荐, 真正实现了系统与用户在翻译过程中的深度交互, 并在欧盟法案的翻译中获得成功应用。后期 Foster<sup>[6]</sup> 进一步又开展了面向短语和用户词典的 IPMT 技术完善工作, Alabau<sup>[7]</sup> 于 2012 年通过改进 IPMT 的交互界面对 IPMT 和 PE 技术进行了比较分析, 在真实的人工评测中, IPMT 在翻译效率和用户满意度两个层面获得了肯定。

IPMT 的交互式翻译过程如图 1 所示<sup>[8]</sup>, 初始时给定输入原语句  $f$ , 系统预测用户的下一步的输入  $e$ , 而后用户通过接受系统的预测或者输入正确译文  $k$  的方式与系统进行交互, 与此同时系统将现有用户的输入  $e_p$  作为约束, 在用户每次操作 (单个字母的输入) 后动态产生更为符合用户预期的预测。在翻译的过程中, 用户与系统不断交互, 直到整个句子翻译完成。

<b>SOURCE (f):</b>	Para encender la impresora:	
<b>REFERENCE (e):</b>	To power on the printer:	
<b>ITER-0</b>	$e_p$ $\hat{e}_s$	( ) <i>To switch on a printer:</i>
<b>ITER-1</b>	a k	To <b>power</b>
	$e_p$ $\hat{e}_s$	To power <i>on a printer:</i>
<b>ITER-2</b>	a k	on <b>the</b>
	$e_p$ $\hat{e}_s$	To power on the <i>printer:</i>
<b>FINAL</b>	a k	printer: #
	$e_p = e$	To power on the printer:

图 1 IPMT 翻译过程。由西班牙语翻译到英语, 系统推荐为斜体字, 用户接受的为正常字体, 用户输入为粗体字。

从上述交互过程中可以看到 IPMT 在人工翻译的基础上, 通过系统的预测自动或半自动的完成用户译文输入实现了翻译效率的提高。因此, 交互过程中用户实际单词输入数量与原句子单词总数的比 (word stroke

rate, WSR) 也自然作为衡量 IPMT 技术的总主要指标。

IPMT 的研究工作主要集中在三个层面, 首先从算法效率方面, Kern 最早提出采用词图 (word Graph) 的方法对尽可能翻译候选进行优化存储, 提高用户每次操作的相应速度, Sanchis-Trilles<sup>[9]</sup> 在 2014 年采用三种策略对词图进行了裁枝, 在保证指标不变的情况下提升了系统的相应效率; 与此同时, 更多研究人员希望 IPMT 在一次操作中能够提供更多的预测内容, 并将主要的预测单元从词扩展到短语和句子, 尽管后期的人工评价对此也给予了积极的评价, 但由于更长的预测不仅使得相关模型更为复杂影响了交互体验, 同时也给用户理解机器的预测带来了较多的困扰; 为了进一步丰富 IPMT 的交互方法, 手写、语音输入等交互手段也被引入到 IPMT 系统中, 在交互式机器学习 (interactive machine learning) 的理论框架中, 既提高了相关输入的识别准确率又保证了用户通过简单的人机交互获得更好的翻译体验。

然而从 IPMT 的技术发展来看, 研究人员主要还是从用户的输入角度讨论人机的翻译交互, 而忽略了其他的用户交互行为对翻译效率的影响, 例如用户利用鼠标完成的译文点选、用户视觉关注内容以及利用电子词典的查词行为等。本文基于词预测交互式机器翻译的研究思路, 将用户交互翻译过程中的鼠标点选行为转化为中间译文的词对齐信息, 进而在翻译交互过程中实现了对译文的动态词对齐标注, 并在对齐信息和输入译文的约束下提高了传统词预测的准确性。

## 2 词预测模型

### 2.1 传统词预测模型

翻译交互过程中的上下文环境主要是由原语  $S = s_1, s_2, \dots, s_{n-1}, s_n$  以及  $p(x|S, T)$  部分的翻译结果  $T = t_1, t_2, \dots, t_{k-2}, t_{k-1}$  构成, 则待预测译文  $x$  的概率可转化为条件概率。但由于  $S, T$  对于当前预测都为常量, 因此 Foster 利用线性插值对其进行了近似求解, 即:

$$p(x|S,T) = \lambda_{(S,T)} * p(x|T) + (1 - \lambda_{(S,T)})p(x|S) \quad (1)$$

由于输入译文是从左向右依次构成的, 随着输入的不断增长, 也导致语言模型  $p(x|T)$  在预测模型中的重要程度不同, 因此  $\lambda_{(S,T)}$  被定义为 S 和 T 的线性插值函数, 用于调节在翻译过程中的不同时期语言模型和翻译模型的比重。

上式中  $p(x|T)$  为语言模型, 由 SRILM 工具训练得到;  $p(x|S)$  为翻译模型, 由词典, 双语对齐语料及词对齐关系训练得到, 其计算方法入下:

$$p(x|S) = \sum_{i=0}^{l_s} p(x|S_i)a(i|k,l_s) \quad (2)$$

其中  $l_s$  表示原语 S 的长度,  $s_i$  代表原语中第 i 个位置的单元, k 代表译文 x 所处的位置。  $p(x|s_i)$  为传统的词翻译模型,  $a(i|k,l_s)$  代表原语第 i 个位置与译文第 k 个位置在原语  $l_s$  条件下的对齐概率。

尽管用户的输入译文 T 作为约束是交互翻译的重要参考, 但是上述模型中该约束仅用于语言模型, 而在翻译模型中未引入上述约束, 并也由此导致其预测准确率不高。

## 2.2 基于词对齐的词预测模型

设 A 为原语 S 以及部分的翻译结果 T 间的对齐关系, 但由于 T 是不完整的译文, 因此从原语到目标语的对齐关系不完整。因此, 本文将对齐关系设定为从目标语到原语, 即  $(t, a_t)$  (以下用  $a_t$  表示),  $a_t \in A$  代表目标语中第 t 个词与原语第  $a_t$  个词存在翻译对齐关系。

待预测译文 x 将被置于翻译结果的第 k 个位置, 其对齐关系为  $a_k$ , 则译文 x 的预测概率可转化为条件概率  $p(x, a_k | S, A, T)$ 。此时译文的预测就变成对上述概率最大值的寻找, 可以通过以下公式给出:

$$(x, a_k) = \arg \max_{x, a_k} p(x, a_k | S, A, T) \quad (3)$$

依据贝叶斯公式上述计算公式 (3) 也可直接推导为:

$$\begin{aligned} (x, a_k) &= \arg \max_{x, a_k} p(x, a_k | S, A, T) \\ &= \arg \max_{x, a_k} p(x|S, A, a_k, T) * p(a_k | S, A, T) \end{aligned} \quad (4)$$

其中,  $p(x|S, A, a_k, T)$  为译文 x 在对齐关系  $a_k \cup A$  的预测概率;  $p(a_k | S, A, T)$  为对齐关系的预测概率。

有了部分对齐结果后, 设  $a_k \cup A = A'$ , 则, 即  $p(x|S, A, a_k, T) = p(x|S, A', T)$ , 进一步利用贝叶斯公式可得到:

$$p(x|S, A', T) \propto p(S, A' | T, x) p(x|T) \quad (5)$$

$p(x|T)$  为语言模型, 而  $p(S, A' | T, x)$  为词对齐关系  $A'$  条件下, 原语 S 以及部分的翻译结果  $T, x = t_1, t_2, \dots, t_{k-2}, t_{k-1}, x$  的翻译概率。本文采用 IBM mode13 方法对翻译概率进行求解, 进一步我们将上述公式分解:

$$\begin{aligned} p(S, A' | T, x) &= \sum_{j=0}^{k-1} n_g(\phi_{a_j} | s_{a_j}) p(s_{a_j} | t_j) d(a_j | a_{j-1}, l_s) \\ &+ n_g(\phi_{a_k} | s_{a_k}) p(s_{a_k} | x) d(a_k | a_l, l_s) \end{aligned} \quad (6)$$

其中  $d(a_j | a_{j-1}, l_s)$  为方向位变概率, 是沿着翻译方向构造的, 根据输入词的位置预测输出词的位置的概率分布;  $n_g(\phi_i | s_i)$  为的繁衍概率, 表示对于每个源语词  $s_i$  通常被翻译成多少个输出单词的概率分布。但与传统的繁衍概率  $n_g(\phi_i | s_i)$  不同, 由于交互翻译中的译文是部分的翻译结果, 导致依据当前对齐获得的繁衍数量  $\phi_i$  不准确, 因此本文采用的翻译概率为原语  $s_i$  繁衍数量大于  $\phi_i$  的概率, 即可表示为:

$$n_g(\phi_i | s_i) = 1 - \sum_{\xi=0}^{\phi_i} n(\xi | s_i) \quad (7)$$

在给定用户输入 T 的前提下

$$\sum_{j=0}^{k-1} n_g(\phi_{a_j} | s_{a_j}) p(s_{a_j} | t_j) d(a_j | a_{j-1}, l_s)$$

的数值可以近似为常量, 因此:

$$\begin{aligned} p(S, A | T, x) &= n_g(\phi_{a_k} | s_{a_k}) p(s_{a_k} | x) d(a_k | a_l, l_s) + C \end{aligned} \quad (8)$$

## 2.3 算法实现

本文在生成预测模型前, 使用 Niutrans 的 GIZA++ 做词对齐处理, 得到词对齐结果:

使用 SRILM 工具得到语言模型。使用训练语料训练得到以下模型：

- 对齐概率模型，源语的某一位置对齐到目标语某一位置的概率模型，使用词对齐结果统计得到；
- 方向位变概率模型，在当前翻译位置的前词所对齐的源语的位置条件下，源语的某一位置与当前翻译位置的对齐概率模型，使用词对齐结果统计得到；
- 翻译概率模型，源语某词翻译到目标与某词的概率模型，利用双语词典和词对齐结果使用公式 (2) 得到；
- 繁衍模型，源语某已经翻译过  $n$  次的词再次被翻译  $m$  次概率模型，使用双语词典和词对齐结果统计得到；

将公式 (4) (5) (8) 整合可以得到公式 (3)，将方向位变概率模型，繁衍概率模型，翻译模型整合到一起得到基于翻译模型的条件概率  $p(x, a_k | S, A, T)$ 。通过公式 (1) 使用 8000 句开发集语料进行迭代，选出不同预测位置时，能使正确答案平均排名最高的参数值，即得到线性插值参数  $\lambda_{(S,T)}$ 。

预测算法的伪代码如下：

Step1: 载入双语对齐的测试语料  
(源语 S 长为  $m$ ，目标语 T 长度为  $n$ ，提供  $h$  个候选)

```

Step2: 总预测数 all=0,
      正确预测数 right=0 翻译正确标记 tag=0
for i=0 to n-1 do
    得到得分从大到小排序的
    翻译候选列表 answer.
    tag=0
    all++
    for j=0 to h do
        if answer[0] == T[i]
            tag=1
            记录对齐位置
            break
        else do
            continue
        if tag do
            right++
    更新预测模型
    
```

Step3: score = right / all

在翻译目标语的每个词后，会记录其对齐关系，以及源语端的词。通过这些信息，更新方向位变概率模型，语言模型，繁衍概

率模型等重新生成预测模型，进行下一位置的预测

### 3 实验

#### 3.1 实验设置

本文实验是在英语到汉语的翻译工作中做的。双语平行语料是 LDC2003E14 的 10 万句平行句对，使用 NiuTrans 的 GIZA++ 做词对齐处理。单语语料为搜狗实验室的 SogouCA 的 150 万句中文新闻语料，并使用 SRILM 工具进行语言模型的训练。所有中文语料使用中科院分词工具对汉语进行分词。

使用人工标注的 1000 句双语对齐语料用于测试。统计计算在不同候选数下的正确率。具体过程为：

本文采用  $n$ -best 的推荐方式，根据提供的原语，系统在当前翻译位置上给出  $n$  ( $n$  取 1, 3, 5, 10) 个翻译候选供用户选择，使用翻译候选的正确率来评价模型：

$$accuracy = \frac{N_c}{N}$$

其中， $N$  表示译文预测的总次数， $N_c$  表示用户使用的翻译候选，即译文预测正确的次数。

本文以 Foster 的线性插值方法作为 Baseline 同现有方法作比较。

#### 3.2 实验结果

图 2 横坐标为预测的候选数，纵坐标为 Baseline 和动态词对齐方法 DWA 的预测正确率。

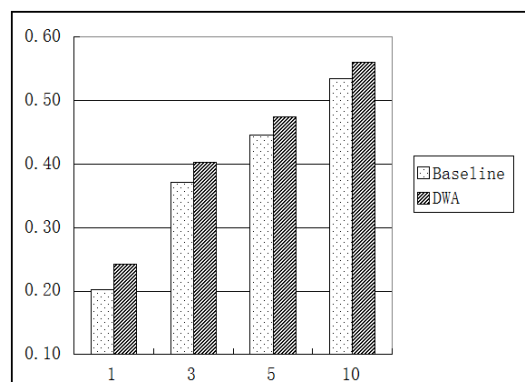


图 2 所有词预测正确率

如图 2，随着候选数的增加，两种方法的正确率均有所提高，而 DWA 方法的正确率

始终高于 Baseline 的方法。在 1 候选时 DWA 预测正确率为 24.25%，Baseline 的预测正确率为 20.24%，此时 DWA 方法提升效果最明显，提高了 4.01 个百分点。

经过实验我们发现，该模型对实词的预测提高较为明显，以下我们从停用词和实词的角度研究了预测效果。

### 3.2.1 不同种类的预测分析

图 3 为两种方法下目标语中停用词的预测正确率。停用词约占目标语中的 20%。

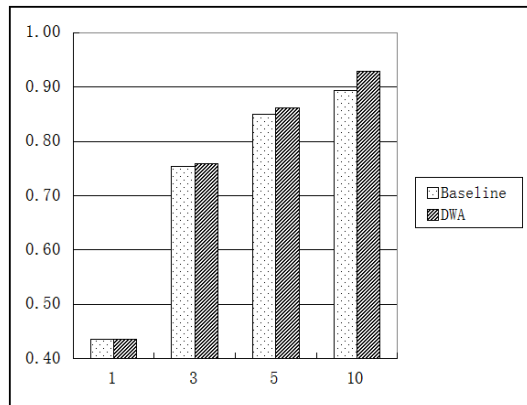


图 3 停用词的预测正确率

如图 3，停用词的预测正确率从第一候选开始就比较高。主要原因是，停用词数目相对实词来说很少，同时比较好预测，所以一开始候选数为 1 的时候停用词的预测正确率就很高。

另外，随着候选数的增大，提升效果有所提高。主要原因是，停用词相对实词，数目少了很多，也较容易预测，随着候选数的增大，停用词在 n-best 预测列表里数目越大，使预测准确率的提升增大。

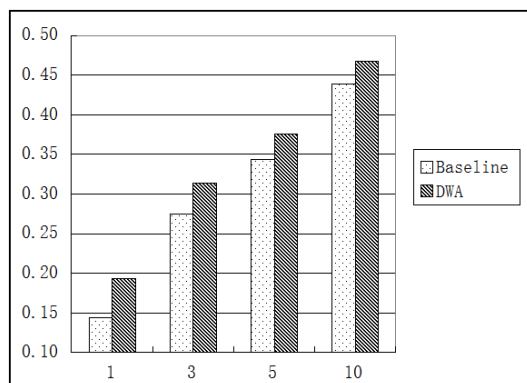


图 4 实词的预测正确率

图 4 为两种方法下目标语中实词的预测正确率，虽然相对于图 2 中正确率下降很多，

但是 DWA 方法相对于 Baseline 的提升也很大。当候选数为 1 时，DWA 方法的正确率为 19.36%，而 Baseline 的预测正确率为 14.35%，高出了 5.01 个百分点。从这一角度也体现出 DWA 模型对词预测的提升有实质性的提高。

另一方面，随着候选数的增大，提升效果有所降低。主要原因是，当候选数增大时，数目少，较好预测的停用词在翻译候选中的比例增大，此时对实词的预测受到影响，使实词的预测准确率提升效果有所减低。

### 4.2.2 词对齐关系对预测正确率的影响

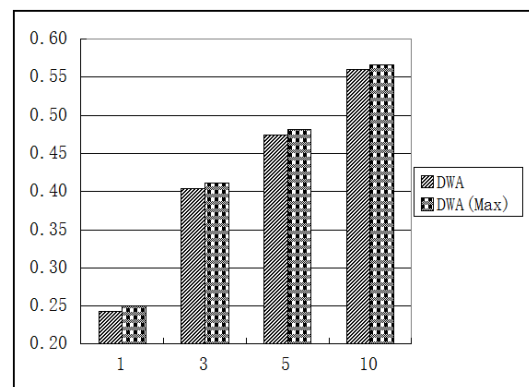


图 5 不同词对齐关系下的预测正确率

为了进一步研究词对齐结果对模型的影响，我们把正确的词对齐结果加入到模型中。如图 5，在当前翻译位置，给定上一翻译的正确对齐位置后，得到 DWA (Max) 的预测正确率，如图 DWA (Max) 的预测正确率相对于 DWA 方法有所提升，主要原因是，在提供了正确的词对齐关系后，可以模型的预测效果会达到最好效果。但是只提升了约 0.7 个百分点，主要原因是，测试语料中的词对齐结果是由 GIZA++ 处理得到的，由于测试语料自身问题，其中有很多对空的情况，无法在每次预测时给出目标语前词的对齐的位置

## 4 总结和展望

本文提出一种动态词对齐的方法，在传统交互式机器翻译的基础上，将用户与翻译引擎交互后的词语对齐信息加入到之后的预测模型中，减少了模型中的搜索空间，提高了模型的效率和预测准确率。为了进一步分析模型的预测效果，加入了正确的词对齐结果，并对最后预测效果进行分析。从实

验结果上可以看出,词对齐信息对预测效果有一定的影响,而且本文的动态词对齐策略已经接近人工指定的效果,相对于传统方法可提高 4.01 个百分点。

下一步的研究工作的重点是优化调序模型,同时在模型中加入短语的预测,减少用户翻译句子时的选择操作,进一步提高模型的预测正确率,提升翻译的工作效率。

### 参考文献

- [1] Koehn P, Haddow B. Interactive assistance to human translators using statistical machine translation methods[J]. MT Summit XII, 2009.
- [2] Och F J, Zens R, Ney H. Efficient search for interactive statistical machine translation[C] //Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1. Association for Computational Linguistics, 2003: 387-393.
- [3] Foster G, Isabelle P, Plamondon P. Target-text mediated interactive machine translation[J]. Machine Translation, 1997, 12(1-2): 175-194.
- [4] Simard M, Ueffing N, Isabelle P, et al. Rule-based translation with statistical phrase-based post-editing[J]. 2007.
- [5] Simard M, Foster G. Pepr: Postedit propagation using phrase-based statistical machine translation[J]. Proceedings of the XIV Machine Translation Summit, 2013: 191-198.
- [6] Foster G, Kuhn R, Johnson J H. Phrasetable smoothing for statistical machine translation[J]. 2006.
- [7] Alabau V, Leiva L A, Ortiz-Martinez D, et al. User evaluation of interactive machine translation systems[C]//Proc. EAMT. 2012: 20-23.
- [8] González-Rubio J, Ortiz-Martínez D, Casacuberta F. On the use of confidence measures within an interactive-predictive machine translation system[C]//Proc. EAMT. 2010.
- [9] Sanchis-Trilles G, Ortiz-Martínez D, Casacuberta F. Efficient wordgraph pruning for interactive translation prediction[C]//Annual Conference of the European Association for Machine Translation (EAMT). 2014.
- [10] Och F J, Ney H. Discriminative training and maximum entropy models for statistical machine translation[C]//Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2002: 295-302.
- [11] Koehn P, Och F J, Marcu D. Statistical phrase-based translation[C]//Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1. Association for Computational Linguistics, 2003: 48-54.
- [12] Arnold D. Why translation is difficult for computers[J]. Computers and Translation: A translator's guide, Amsterdam and Philadelphia: John Benjamins, 2003: 119-42.
- [13] Kay M. The proper place of men and machines in language translation[J]. machine translation, 1997, 12(1-2): 3-23.
- [14] Koehn P, Haddow B. Interactive assistance to human translators using statistical machine translation methods[J]. MT Summit XII, 2009.
- [15] Och F J, Ney H. Improved statistical alignment models[C]//Proceedings of the 38th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2000: 440-447.
- [16] Ortiz-Martinez D, García-Varea I, Casacuberta F. Interactive machine translation based on partial statistical phrase-based alignments[C] //RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING. 2009: 330.
- [17] Langlais P, Foster G, Lapalme G. TransType: a computer-aided translation typing system [C]//Proceedings of the 2000 NAACL-ANLP Workshop on Embedded machine translation systems-Volume 5. Association for Computational Linguistics, 2000: 46-51.