

文章编号: 1006-4710(2013)01-0032-06

# 基于多特征融合的统计机器翻译译文错误检测

王莎, 杜金华, 刘丁

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

**摘要:** 抽取了3种典型的单词后验概率特征(基于固定位置的词后验概率、基于目标位置窗的词后验概率、基于词对齐的词后验概率)和3种语言学特征(词、词性、句法分析器抽取的句法特征),并在此基础上抽取了一个来自源端的单词特征,然后基于中英NIST数据集,采用最大熵分类器来验证不同单词后验概率特征(WPP)独立使用及与其它特征组合后使用时对错误检测性能的影响。实验结果表明,采用不同方法计算得到的单词后验概率特征对分类错误率的影响是显著的,并且在单词后验概率和语言学特征组合基础上加入源端单词特征,可以显著降低分类错误率(CER),提高译文错误检测能力。

**关键词:** 最大熵分类器; 单词后验概率; 语言学特征; 源端单词特征; 错误检测

中图分类号: TP391

文献标志码: A

## Error Detection in Translation Version by Statistical Machine Translation Based on Feature Integration

WANG Sha, DU Jinhua, LIU Ding

(Faculty of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** Three kinds of typical word posterior probability (WPP) features (based on the fixed position, sliding window-WPP and alignment-based WPP) and three kinds of linguistic features (word, POS and LG parsing knowledge) are extracted to detect errors, on the basis of which a source-side word feature is extracted, and then based on NIST data setting the maximum entropy classifier is adopted to test the different word posterior probability features as well as the effect upon the independent use and error detection performances of other feature integration in use. The experimental results show that the WPP features obtained using the different calculation methods have the obvious effects upon the classification error rate and also that the source-side word feature incorporated on the basis of integration of the WPP and linguistic features can significantly reduce the CER values and improve the detection performances in translation version errors.

**Key words:** maximum entropy classifier; word posterior probability; linguistic features; source-side word feature; error detection

近年来,基于统计机器翻译方法的发展,涌现出了多种不同类型的统计机器翻译(SMT)系统,如基于短语、层次短语及句法的机器翻译模型等,并且翻译性能得到了显著提高<sup>[1-3]</sup>。本领域内一些研究人员对于如何将SMT系统应用到实际中做了大量工作。然而,目前来看,统计机器翻译译文质量仍然无法满足工业化的实际需要。因此,为了提高机器翻

译译文质量,自动对译文进行错误判断、分类和分析在SMT技术发展和应用中有着重要作用。

目前,译文错误检测方法多依据如单词后验概率(word posterior probability, WPP)的系统特征进行译文置信度估计,一些研究人员已经提出了多种根据N-best或词图等计算WPP的方法,并且这些方法在机器翻译译文错误检测和置信度估计中得到了广

收稿日期: 2012-11-15

基金项目: 国家自然科学基金资助项目(61100085); 陕西省教育厅专项科研计划基金资助项目(11JK1029)。

作者简介: 王莎,女,硕士生,研究方向为模式识别与智能系统。E-mail: hebenwangsha@126.com。

杜金华,男,副教授,研究方向为模式识别与智能系统。E-mail: jhdu@xaut.edu.cn。

泛应用<sup>[4-7]</sup>。之后,一些研究人员开始尝试利用其他知识源作为特征,并与 WPP 置信特征相结合来进行错误检测,如深层次句法或深层次语义特征等<sup>[8-9]</sup>。然而,由于多数外部特征的复杂性、与具体语言相关等原因,有效的并且通用的特征并不容易被抽取。因此,目前来看,单词后验概率特征、词以及词性等词汇化特征仍然是人们依据的主要特征。

本文主要采用最大熵分类器结合不同方法计算得到的 WPP 特征及三种 WPP 特征与其他特征的组合结果,研究其对译文错误检测能力的影响,主要有以下三点:①验证三种不同方法得到的 WPP 特征对译文错误检测性能的有效性;②分别将三种不同方法得到的 WPP 特征与其他三个语言学特征结合,验证语言学特征对具体语言翻译错误检测任务的预测能力;③在以上特征组合的基础上分别加入抽取的源端单词特征,验证源端单词特征对分类错误率的影响。

## 1 置信度特征

### 1.1 词后验概率特征

词后验概率是统计机器翻译译文错误检测中常用的置信度特征,目标单词的置信度可以用其后验概率即 WPP 来表示<sup>[5-7]</sup>。

用  $S$  表示一个 SMT 系统,给定源语言输入句子  $f_1^J$  和对应的  $N$ -best 列表输出  $e_{n,1}^{n,I_n}$ ,其中  $n = 1, \dots, N$ ;  $e_n$  代表  $N$ -best 列表中第  $n$  个翻译假设,每个翻译假设的翻译概率记为  $p(f_1^J, e_{n,1}^{n,I_n})$ ,错误检测所用的 WPP 即 1-best 列表中每个目标词的后验概率  $p(e_i^J/f_1^J)$ 。下面将详细介绍三种典型的用来计算 WPP 的方法。

#### 1.1.1 基于固定位置的词后验概率

给定源语言输入  $f_1^J$ ,位于最优翻译假设  $e_1^J$  中位置  $i$  的目标单词  $e$  的后验概率即是对  $N$ -best 列表中其他翻译假设在对应的固定位置  $i$  出现单词  $e$  的概率求和,如公式(1)所示:

$$p_i(e/f_1^J) = \frac{\sum_{n=1}^N \delta(e_{n,i}, e) \cdot p(f_1^J, e_{n,1}^{n,I_n})}{\sum_{e'} \sum_{n=1}^N \delta(e_{n,i}, e') \cdot p(f_1^J, e_{n,1}^{n,I_n})} \quad (1)$$

#### 1.1.2 基于目标位置窗的词后验概率

在  $N$ -best 列表中,翻译假设的句长是在某一范围内动态变化的,使得由于不同的翻译假设因句长不同等因素的影响,导致在同一位置  $i$  上的目标词  $e_i$  有可能是不同的,即对应不同的源语言词,因此,难以保证单词  $e$  出现的位置  $i$  是固定不变的,但有可

能出现在位置  $i$  附近,即上下文中。因此,若将初始的固定位置  $i$  变为一个动态值,使其能在初始值的某一范围内滑动,当目标单词出现在限定范围内时,其参与后验概率计算。这就是在固定位置基础上改进的基于位置窗的方法思想。

记位置窗为  $i \pm t$ ,  $t$  为窗口大小,为自然数。如果单词在位置窗范围内出现,同样认为单词出现在当前翻译假设。因此,单词的后验概率可以由此窗范围内的位置上单词出现的后验概率之和来决定,计算公式如(2)式所示:

$$p_{i,t}(e/f_1^J) = \sum_{k=i-t}^{i+t} p_k(e/f_1^J) \quad (2)$$

#### 1.1.3 基于词对齐的词后验概率

另一种直接且有效的计算方法是利用  $N$ -best 列表中生成的最优目标译文  $e_1^J$  与  $N$ -best 列表中其他句子的词对齐信息得到词后验概率。本文实验利用的是基于 Levenshtein 的对齐,根据此对齐信息,最优译文中每个单词  $e$  的后验概率  $p(e/f_1^J, e_1^J)$  可以通过对 Levenshtein 对齐结果中位置  $i$  上包含  $e$  的所有翻译假设的概率之和得到。此方法由 Ueffing 等人最早提出,并通过实验验证了其有效性<sup>[5,7]</sup>。

用  $L(e_1^J, e_{n,1}^{n,I_n})$  表示最优译文  $e_1^J$  和  $N$ -best 列表中其他翻译假设  $e_{n,1}^{n,I_n}$  之间的 Levenshtein 对齐关系,则对于位置  $i$  的单词  $e$ ,其具有 Levenshtein 对齐关系的句子后验概率可由下式给出:

$$p_{lev}(e/f_1^J, e_1^J) = \frac{p_{lev}(e, f_1^J, e_1^J)}{\sum_{e'} p_{lev}(e', f_1^J, e_1^J)} \quad (3)$$

其中:

$$p_{lev}(e/f_1^J, e_1^J) = \sum_{n=1}^N \delta(e, L_i(e_1^J, e_{n,1}^{n,I_n})) \cdot p(f_1^J, e_{n,1}^{n,I_n})$$

式中  $p(f_1^J, e_{n,1}^{n,I_n})$  表示  $N$ -best 列表中给出的翻译假设的后验概率。 $\delta(x, y)$  代表克罗奈克函数,即:

$$\delta(x, y) = \begin{cases} 1 & x = y \\ 0 & \text{其他} \end{cases}$$

## 1.2 语言学特征

近年来,本领域的一些研究人员开始尝试利用外部知识源作为特征,如深层次句法或深层次语义特征等。1.1 节中的单词后验概率特征是基于系统的特征,无法提供足够的语法或句法知识,因此,考虑抽取如句法、语义等语言学特征来提高检测准确率。本文主要用到了两类常用的语言学特征:词汇化特征和句法特征。下面对这两类特征做一个详细的描述。

### 1.2.1 词汇化特征

对于翻译假设中的大多数单词,通常认为出现频率高的单词序列和出现频率较高的词性标注序列为正确的概率要大于出现频率较少的单词序列和词性标注序列。这里,考虑单词及词性标注的上下文信息,对于每个单词和词性标注,取其前2个和后2个及其本身构成特征向量,具体形式表示如下。

单词特征:

$$Word: (w_{-2}, w_{-1}, w, w_1, w_2)$$

词性标注特征:

$$Pos: (pos_{-2}, pos_{-1}, pos, pos_1, pos_2)$$

### 1.2.2 语法特征

句法知识一般是通过句法分析器对源语言或目标语言进行句法分析后得到的。当分析器无法对整个句子进行分析时,则忽略问题单词以找到剩余单词的联系从而完成句法分析。被忽略的单词就成为与句子中其他单词没有连接的单词,称为 null-link 单词。一般认为,这些与其他单词没有联系的单词存在语法错误的可能性更大,利用此语法信息来定义二值句法特征,如下式所示:

$$link(e) = \begin{cases} \text{yes} & e \text{ 与其他词存在连接关系} \\ \text{no} & \text{其他} \end{cases}$$

### 1.3 源端单词特征

统计机器翻译译文错误检测大部分研究中所用的置信度特征均是关于目标端的特征,如 WPP 特征、语言学特征和其他常用的句法语义特征等。在此基础上本文以源端为对象,抽取源端特征。从 MT 解码过程中可以跟踪生成译文中目标短语所对应的源语言短语,并且利用简单的对齐模型能够很快找到对应短语间源语言单词和目标语言单词之间的对齐关系,即提取源端单词特征<sup>[11]</sup>。这些特征用来得到给定错误标注中源语言单词和目标单词出现的相似度。可以认为:如果源语言单词和目标单词同时出现在同一类别的概率很大,那么可以认为生成的目标单词在将来翻译结果中很可能属于此标注类别。

实验中所用的源端单词特征模板如图1所示,第一行为源语言中文序列,第二行是利用统计机器翻译系统得到对应的英文输出结果。

源端单词特征的定义如公式(4)所示:

$$f_{(1,3)}(\text{nearly}) = \begin{cases} 1 & \text{if source - word} = \text{"近"} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

式中,  $f_{(1,3)}$  中的 1 和 3 分别表示目标单词在翻译假设中的位置和对应的源语言单词在源语言句子中的位置。

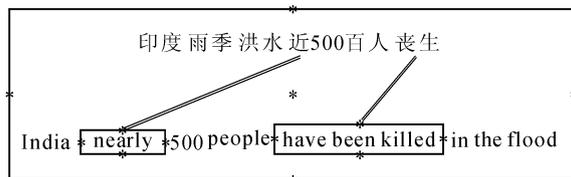


图1 源端单词特征

Fig.1 Source word feature

如果源语言单词对应到目标端语言的是一个短语,如图1中源语言中文词语“丧生”对应的目标端句子中的成分为短语“have been killed”,则需要返回到短语模型训练过程中生成的短语表中,从表中的对齐信息找到中文单词与英文短语中每个单词的一一对应关系,最终得到源端单词对齐信息。

## 2 分类器

机器翻译中的译文错误检测问题通常是作为一个二值分类问题来进行研究,如监督机器学习问题。自然语言处理领域中最常用的分类器包括最大熵分类器<sup>[10-13]</sup>、SVM 分类器等。实验中选择最大熵分类器作为其译文错误检测分类工具,结合抽取的多个不同特征分析比较不同特征组合的分类性能。

最大熵分类器是朴素贝叶斯分类器的一个泛化模型,其本质思想是不考虑任何未知因素,对所有已知因素建立一个连续的模型,使用最大熵方法建模的一个主要优点是不同的特征可以被方便地添加到模型中。将二值分类器样例用  $(x_1, y_1), \dots, (x_n, y_n)$  表示,其中  $x_i$  表示特征样例,本实验中  $y_i \in \{c, i\}$ 。

给定了单词的特征向量后,采用最大熵模型预测单词正确与错误的公式如式(5)所示。

$$p(y/x) = \frac{\exp(\sum_i (\lambda_i f_i(x, y)))}{\sum_y (\exp(\sum_i \lambda_i f_i(x, y)))} \quad (5)$$

式中,  $f_i$  代表二值模型特征函数,其值由单词分类类别  $y(c, i)$  和特征向量  $x$  来决定。  $\lambda_i$  代表对应  $f_i$  的权重。

利用最大熵模型作为译文错误检测的分类器,需要将特征向量转换为分类器所能接受的文本格式。具体格式如下所示:

```
line1:label index1:value1 index2:value2...
line2:c wpp_{-2}:1 ... pos_{-2}:1...word_{-2}:1...
      link_{-2}:1...source_{-2}:1...
line3:c wpp_{-2}:0...pos_{-2}_JJ...
      word_{-2}_white...link_{-2}:1...
      white 白宫...
```

在此特征向量公式中, *line1* 表示最大熵分类器的标准数据模板, 包括标注结果(分类类别)、特征和特征值。*line2* 用来说明对应于 *line1* 的特征及其特征值的具体文本数据格式, *line3* 则是将特征用具体指代内容代替后的数据格式, 为最大熵模型训练时所使用的模板。其中, *line1* 中的 *label* 对应于 *line2* 和 *line3* 中的第一行, 即单词分类结果  $\{c, i\}$  中的值, *line1* 中的 *index* 为抽取的特征, 对应 *line2* 和 *line3* 中的特征, 而 *value* 则表示每个特征对应的函数值, 对应 *line2* 中特征之后的“1”, 因为在 *line2* 中只将特征值为 1 的特征罗列了出来, 故在 *line3* 中对应 *value* 的函数值均为 1。根据最大熵模型的特性训练文本和测试文本特征值省略不写。

### 3 实验和设计分析

#### 3.1 中文-英文统计机器翻译系统

本文实验中的统计机器翻译系统选用基于短语的机器翻译系统: Moses 系统。利用此系统在测试集解码过程中, 为每个源语言句子输出 10 000 个最好翻译假设, 即  $N = 10\ 000$ 。

翻译语言对的方向为中文到英文的翻译。翻译模型训练语料为 LDC 提供的数据集, 共计 3 397 538 句对(主要包括香港新闻、FBIS、ISI 汉英网络数据、新华新闻等)。语言模型采用五元模型, 其训练语料包括以上双语语料的英文部分和英文 Gigaword 的新华部分, 共计约 1 000 万句。

机器翻译系统的开发集为 NIST 机器翻译评测任务 2006 年 current 集, 共计 1664 句, 每个源语言输入对应 4 个参考答案。测试集为 NIST 2005 current 数据集 1 082 句和 NIST 2008 current 数据 1 357 句, 每句对应 4 个参考答案。

#### 3.2 译文错误检测任务

数据标注: 本文利用 TER 工具包<sup>[14]</sup>中的 WER 准则来确定翻译假设中每个单词的真实分类结果。首先找到四个参考译文中与翻译假设编辑距离最小的参考译文作为基准, 利用 WER 准则将假设译文与基准参考译文做词对齐, 如果翻译假设中的单词与参考译文中相同位置上的单词一致则标记为 *c*, 反之标记为 *i*。

SMT 系统生成的 NIST MT 2008 翻译结果中, 1-best 翻译假设包含 38 587 个单词, 其中标为 *c* 的为 14 658 个, 标为 *i* 的 23 929 个, 其样本为“正确”的类别比例(RCW)为 37.99%。NIST MT 2005 翻译结果的 1-best 翻译假设包含 36 497 个单词, 其中标为 *c* 的 15 179 个, 标为 *i* 的 21 318 个, 其样本为“正

确”的类别比例(RCW)为 41.59%。

开发集和测试集: 在译文错误检测任务中使用 NIST MT 2008 数据集(1 357 句)作为最大熵模型参数训练的开发集, 将 NIST MT 2005(1 082 句)数据集作为分类任务的测试集。

#### 3.3 评价准则

译文错误检测任务评价准则有: 分类错误率(CER)、准确率(*P*)、召回率(*R*)和 *F* 准则。

$$CER = \frac{\text{分类类别为错的单词个数}}{\text{单词总个数}} \quad (6)$$

准确率定义为分类器将真实类别为 *i* 的单词准确分类的个数  $n_m$  与分类器标记为 *i* 的单词总数  $n_i$  之比:

$$P = \frac{n_m}{n_i} \quad (7)$$

召回率定义为分类器将真实为类别 *i* 的单词准确分类的个数  $n_m$  与真实类别为 *i* 的单词总数  $n_g$  之比:

$$R = \frac{n_m}{n_g} \quad (8)$$

*F* 准则定义为准确率和召回率的权衡, 即:

$$F = \frac{2 \times P \times R}{P + R} \quad (9)$$

CER 基线水平: 由于作为测试集的 MT-05 数据中  $RCW = 41.59\%$ , 低于 50%, 即标为“incorrect”的单词多于标为“correct”单词, 所以将“correct”的单词全部标为“incorrect”时所得到的评价准则得分, 也就是  $RCW$  作为分类错误率的基线水平。

#### 3.4 译文错误检测实验

##### 3.4.1 基于独立特征的分类实验

利用最大熵分类器分别对三种典型的 WPP 特征和三个语言学特征及抽取的源端单词特征进行独立特征分类实验并分析比较其结果。分类实验结果如表 1 所示。

表 1 独立特征的错误检测结果

Tab. 1 Results of individual feature for Translation

特征	Error Detection			
	CER/%	P/%	R/%	F/%
基线	41.59	-	-	-
WPP_Dir	40.48	63.44	72.46	67.65
WPP_Win	39.70	63.82	73.95	68.51
WPP_Lev	40.12	60.24	92.07	72.83
Wd	39.11	64.20	76.67	69.04
Pos	39.50	61.52	86.46	71.89
Link	40.89	59.55	93.55	72.77
Source	39.68	63.36	76.04	69.12

表中, WPP\_Dir 表示基于固定位置的 WPP, WPP\_Lev 表示基于 Levenshtein 对齐的 WPP, WPP\_Win 表示基于目标位置窗的 WPP。单词特征用 Wd 表示, 源端单词特征用 Source 表示。

其中基线系统的 CER 为 41.59%。

从表中的数据可以得到以下几点结论。

1) 三种典型的 WPP 特征独立使用时都能在一定程度上降低 CER 值, 且基于目标位置窗的计算方法因其灵活性得到了较其他两种方法更低的 CER 结果。基于 Levenshtein 对齐的 WPP 就 F 值而言性能最好。

2) 语言学特征性能要比独立的 WPP 特征好 (Link 除外)。单词特征可以帮助系统更准确地找到错误所在。语法特征 Link 召回率最高, 预示着此特征可以帮助系统找到更多错误。表中的数据同时也揭示了语言学特征对于错误检测的贡献, 表明了语言学特征可以有效降低分类错误率, 提高错误预测能力。

3) 就 CER 而言, 源端单词特征降低了 4.59%。证明了源端单词特征的有效性, 可以考虑提取更丰富的源端信息作为特征加入到分类器中来提高错误检测分类性能。

### 3.4.2 基于组合特征的分类实验

关于组合特征的分类实验做了两组。首先是分别将三种典型的 WPP 特征和抽取的三个语言学特征 (词、词性、由 LG 句法分析器抽取的句法特征) 进行组合, 利用最大熵分类器进行错误分类实验。

表2 第一组组合特征错误检测结果

Tab.2 Results of the first combined features for Translation Error Detection

特征	CER/%	P/%	R/%	F/%
基线	41.59	-	-	-
WPP_Dir + Wd + Pos + Link	35.93	63.93	88.30	74.17
WPP_Win + Wd + Pos + Link	35.55	64.77	85.83	73.83
WPP_Lev + Wd + Pos + Link	35.62	65.31	83.22	73.15

从表2可看出, 就 CER 而言, 与基线系统相比, 三种组合特征的特征分别降低了 13.61%、14.52% 和 14.35% (相对值), 而且 F 值也得到了显著提高。三组实验中, 特征组合“WPP\_Win + Wd + Pos + Link”的分类错误率最低, 而组合“WPP\_Dir + Wd + Pos + Link”的 F 值最高。同时也揭示了语言学特征对于错误检测的贡献。

在以上组合特征实验的基础上加入源端单词特征, 同样利用最大熵分类器进行分类实验, 分析比较

组合特征的预测能力。错误分类实验结果如表3所示。

表3 第二组组合特征错误检测结果

Tab.3 Results of the second combined features for Translation Error Detection

特征	CER/%	P/%	R/%	F/%
基线	41.59	-	-	-
WPP_Dir + Wd + Pos + Link + Source	35.54	64.68	86.24	73.92
WPP_Win + Wd + Pos + Link + Source	35.12	65.75	83.21	73.46
WPP_Lev + Wd + Pos + Link + Source	34.93	67.02	79.16	72.59

从表3和表2可以得出如下结论。

1) 组合特征的分类能力要比单独的特征分类能力强, 且语言学特征和源端单词特征都能在一定程度上提高统计机器翻译译文错误检测能力。

2) 对于加入源端单词特征的组合特征实验, 就 CER 而言, 加入源端单词特征后的组合特征, CER 分别降低了 14.55%、15.56%、16.01% (相对值)。同时, 在加入源端单词特征后, 其 F 值也有一定程度的降低。

3) 表3所用的三种不同方法计算得到的 WPP 中, 基于词对齐的 WPP 与其他四个特征的组合得到了最好的结果, 说明了位置信息与丰富的词对齐信息对错误检测分类的重要性。

## 4 结论和展望

本文基于中英 NIST 数据集, 利用最大熵分类器对三种典型的 WPP 特征、语言学特征、源端单词特征及三种不同的 WPP 特征与其他几个特征的特征组合进行了译文错误分类实验。实验结果表明了语言学特征和源端单词特征对译文错误检测能力的有效性。同时, 可以看出基于词对齐的 WPP 与语言学特征、源端单词特征的组合得到的分类性能最好, 说明了位置信息的重要性。

接下来的工作将从以下几个方向做进一步的研究: 考虑实际情况中单词同义不同形的情况、考虑引进“复述”等技术对本文中仅仅使用 WER 准则进行 c 与 i 的标注方法做出改进; 提取更丰富的源端信息特征, 如利用源端 POS 标注结果代替源端单词特征, 以解决数据稀疏的问题, 提高分类器的分类性能。

## 参考文献:

- [1] Kenji Yamada, Kevin Knight. A syntax-based statistical translation model; Proceedings of ACL-EACL'01, 2001 [C]. Toulouse, France, 2001.
- [2] Philipp Koehn, Franz Josef Och, Daniel Marcu. Statistical phrase-based translation; Proceedings of HLT-NAACL, 2003 [C]. Edmonton, Canada, 2003.
- [3] David Chiang. A hierarchical phrase-based model for statistical machine translation; Proceedings of ACL'05, 2005 [C]. Ann Arbor, MI, 2005.
- [4] Simona Gandrabur, George Foster. Confidence estimation for text prediction; Proceedings of CoNLL, 2003 [C]. Edmonton, Canada, 2003.
- [5] Nicola Ueffing, Klaus Macherey, Hermann Ney. Confidence measures for statistical machine translation; Proceedings of MT Summit IX, 2003 [C]. Nouvelle-Orleans, USA, 2003.
- [6] Blatz John, Erin Fitzgerald, George Foster, et al. Confidence estimation for machine translation; Proceedings of COLING, 2004 [C]. Geneva, Switzerland, 2004.
- [7] Nicola Ueffing, Hermann Ney. Word-Level confidence estimation for machine translation [J]. Computational Linguistics, 2007, 33(1), 9-40.
- [8] Lucia Specia, Nicola Cancedda, Marc Dymetman, et al. Estimating the sentence-level quality of machine translation systems; The 13th Annual Conference of the European Association for Machine Translation (EAMT-2009) [C]. Barcelona, Spain, 2009.
- [9] Lucia Specia, Craig Saunders, Marco Turchi, et al. Improving the confidence of machine translation quality estimates; Proceedings of the 12th Machine Translation Summit, 2009 [C]. Ottawa, Canada, 2009.
- [10] Xiong Deyi, Zhang Min, Li Haizhou. Error detection for statistical machine translation using linguistic features; Proceedings of the 48th ACL, 2010 [C]. Uppsala, Sweden, 2010.
- [11] Nguyen Bach, Fei Huang, Yaser Al-Onaizan. Goodness: A Method for Measuring Machine Translation Confidence; Proceedings of ACL, 2011 [C], Portland, Oregon, USA, 2011.
- [12] Ratnaparkhi A. A maximum entropy for part-of-speech tagging; Proceedings of the empirical methods in natural language processing conference (EMMLP), 1996 [C]. Philadelphia, 1996.
- [13] Zhang Le. Maximum entropy modeling toolkit for Python and C++ [EB/OL]. [http://homepages.inf.ed.ac.uk/s0450736/maxent\\_toolkit.htm](http://homepages.inf.ed.ac.uk/s0450736/maxent_toolkit.htm), 2012-02-10.
- [14] Matthew Snover, Bonnie Dorr, Richard Schwartz, et al. A study of translation edit rate with targeted human annotation; Proceedings of the AMTA'06, 2006 [C]. Boston, MA, USA, 2006.

(责任编辑 王卫勋)

## 简讯

## 西安理工大学两项科技成果获国家科技奖

2013年1月18日上午,2012年度国家科学技术奖励大会在北京隆重举行,我校有两项科技成果获国家科技奖励:由自动化与信息工程学院刘丁教授主持完成的“大尺寸电子级硅单晶炉关键技术及其应用”获得国家技术发明奖二等奖;水利水电学院张建丰教授参与完成的“都市型现代农业高效用水原理与集成技术研究”获得国家科学技术进步奖二等奖。

(摘自理工大新闻网 2013-01-18)