

新的规则抽取约束方法改进基于句法翻译模型的性能

贲国生, 谢军, 刘群

智能信息处理重点实验室

计算技术研究所

中国科学院

{benguosheng, xiejun, liuqun}@ict.ac.cn

摘要

当前的基于句法模型的机器翻译中, 翻译规则的抽取都要受到严格的约束。在翻译规则自动获取算法设计中, 我们发掘了更细粒度的语言学知识, 提出了在抽取翻译规则的过程中, 使用更宽松的子树一致性跨度约束来抽取翻译规则, 并在基于句法的两大模型: 成分树到串和依存树到串中做了验证。在成分树到串系统(树到串系统) BLEU 值提高了 2.62 个点, 在依存树到串系统中 BLEU 提高了 0.37 个点。

1 概述

近年来, 基于句法的翻译模型成为统计机器翻译领域研究热点, 并取得了较大的进展。基于语言学句法的翻译模型以句法分析器输出的句法树结构为基础, 利用句法树所包含的语言学知识来指导翻译过程。基于句法的翻译模型能取得很好的翻译性能, 现已显著超过了基于短语的翻译模型性能。

目前, 统计机器翻译领域常用的文法主要分为两种: 短语结构文法(phrase structure grammar) (Xavier and Collins, 2009; DeNeefe and Knight, 2009; Galley et al., 2004; Liu, 2006; Mi 2008; Yamada and Knight, 2001) 依存文法(dependency grammar) (Ding and Palmer, 2004; Lin, 2004; Quirk, 2005; Shen et al., 2008;

Xiong et al., 2007)。其中, 短语结构文法也被称为成分文法(constituency grammar)。两种文法均以树结构来描述句子中各个成分(constituents)之间的关系, 分别称被称为短语结构树(或成分树)和依存树。

基于成分树的翻译模型主要以树替换文法 TSG (Tree Substitution Grammar) 和树粘接文法 TAG(Tree Adjoining Grammar) (Joshi et al., 1975) 这两类文法为基础构建。根据翻译规则的源端和目标端利用的语言学知识的差异, 基于成分树的模型大致可以分为三种类型: 1) 串到树模型(Galley et al., 2006; Galley et al., 2004; Yamada and Knight, 2001), 仅利用目标语言句法树的语言学知识; 2) 树到串模型(Liu et al., 2006; Huang et al., 2006; Mi et al., 2008), 仅利用源语言的句法树的语言学知识; 3) 树到树模型(Mi et al., 2010; Zhang et al., 2008), 同时利用源语言句法树和目标语言句法树的语言学知识。

基于依存树的翻译模型旨在描述句子中各个成分之间的语法关系, 兼有句法知识和一定的语义知识。在基于依存树的模型研究中, 研究人员从不同的角度进行了探索, 提出了多种不同的树分解方法或形式化文法, 构建了几种不同类型的翻译模型。类似于基于成分树的翻译模型分类方法, 这些模型也可以大致分为三类: 1) 串到依存树模型(Shen and Weischedel, 2010; Shen et al., 2008), 仅利用目标语言句法树的语言学知识; 2) 依存树到串模型(Xiong et al., 2007; Xie et al., 2011), 仅利用源语言句法树的语

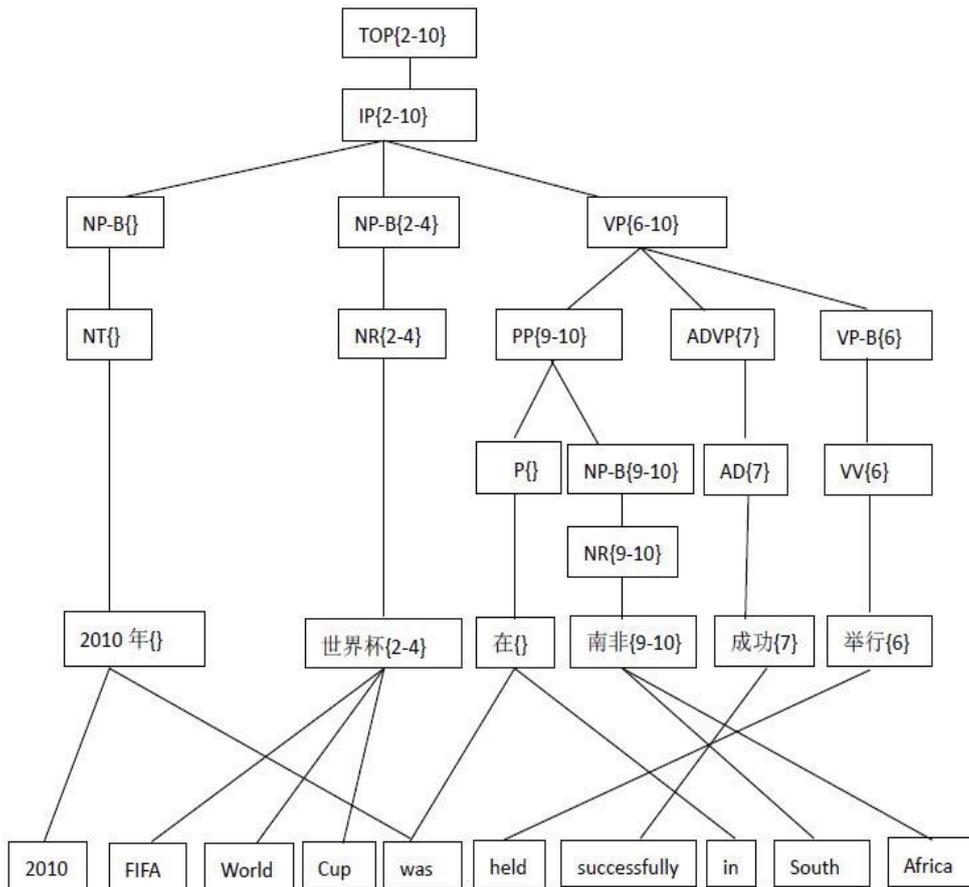


图 1：子树一致性跨度的成分树

语言学知识； 3) 依存树到依存树模型(Ding and Palmer, 2004; Ding and Palmer, 2005; Lin, 2004; Quirk et al., 2005), 同时利用源语言句法树和目标语言句法树的语言学知识。

构建基于句法的翻译模型时, 主要涉及三个任务 (Xie et al., 2011): 1) 翻译模型的设计: 根据定义的基本结构单元和基本操作设计翻译模型, 确定翻译规则表示方法; 2) 翻译规则的自动获取: 根据翻译规则表示方法和基本操作, 设计合适的翻译规则自动获取方法; 3) 解码算法: 根据定义的翻译规则表示方法和基本操作设计适当的解码算法。其中, 在基于句法的翻译模型中, 翻译规则的自动获取都要受到一定的约束, 通常, 我们都使用短语一致性跨度约束来抽取翻译规则。在本文中, 我们提出在抽取翻译规则的过程中, 使用更为宽松的子树一致性跨度约束来代替传统的短语一致性跨度约束抽取翻译规则。实验结果表明, 在抽取

翻译规则的过程中我们能抽取更多的有效规则且在成分树到串和依存树到串两个句法模型中都能有效地提高机器翻译的质量。

2 子树一致性跨度约束 VS 短语一致性跨度约束

2002年(Och et al., 2002)提出在基于短语的模型中, 双语短语需要满足对齐一致性。短语对齐一致性定义为: 给定三元组 $\langle f, e, A \rangle$, 其中, f 为源语言句子, e 为目标语言句子, A 为 f 和 e 的对齐关系:

$$(f_j^{j+m}, e_i^{i+n}) : \forall (i', j') \in A:$$

$$j \leq j' \leq j+m \leftrightarrow i \leq i' \leq i+n$$

而在基于句法的模型中, 句法树所覆盖的短语倾向于作为整体翻译为目标短语。句法树中每个短语都是固定的, 每个子树都代表着

一个语义单元。这与 Och 在 2002 年提出任何连续的词语都可被认为是一个短语有着本质的差异。因此，在基于句法的模型中，短语的一致性约束是非常强的。从语言学的角度来说，源端短语和目标端短语中任何词的对齐关系都可代表源端短语和目标端短语的对齐关系，因此，在基于句法树的翻译规则抽取中我们就可以使用更为宽松的约束来抽取翻译规则。例如，2002 年 (Fox 2002) 就曾经提出依存树具有更好的跨语言短语聚合性质。类似地，我们在成分树上也发现类似性质。因此，在基于句法的翻译模型的规则抽取中，我们可以尝试利用更为宽松的子树一致性跨度约束来抽取翻译规则。

2.1 子树一致性跨度

在这里我们采用 2011 年 (Xie et al. 2011) 提出的相关定义来解释在基于句法的模型中采取更为宽松的约束方法。

定义 1 节点 n 的对齐 $align(n)$ 定义为与节点 n 存在对齐关系的所有目标语言词语的位置的集合。

例如，在图 1 中，节点“2010 年”的对齐 $align(2010 年) = \{1,5\}$

定义 2 节点 n 的对齐 $align(n)$ 的闭包 $closure(align(n))$ 定义为包含集合 $align(n)$ 中的所有元素的最小连续集合。

我们将 $closure(align(n))$ 简记为 $closure(n)$ 。例如，在图 1 中，节点“2010 年”的对齐 $align(2010 年)$ 的闭包 $closure(2010 年) = closure(\{1,5\}) = \{1,2,3,4,5\}$ ，为了方便起见，将集合 $\{i, \dots, j\}$ 简记为 $\{i-j\}$ 。例如 $\{1,2,3,4,5\}$ 可以简记为 $\{1-5\}$

定义 3 节点 n 的词跨度 $wsp(n)$ 定义为该节点对齐的闭包 $closure(n)$ 。

例如，在图 1 中，节点“2010 年”的词跨度 $wsp(2010 年) = closure(2010 年) = \{1-5\}$

定义 4 若节点 n 的词跨度 $wsp(n)$ 满足如下性质，则称节点 n 是对齐一致的：

$$\forall_{n' \neq n} wsp(n') \cap wsp(n) = \emptyset$$

例如，在图 1 中，节点“2010 年”和节点“在”的词跨度存在交集， $wsp(2010 年)$

$\cap wsp(在) = \{5\}$ ，因此，在这两个节点不是对齐一致的，其他节点都是对齐一致的。

定义 5 以节点 n 为根的子树 T' ，节点 n 的短语跨度 $psp(n)$ 定义为：

$$psp(n) = closure(\bigcup_{n' \in T'} wsp(n'))$$

即节点 n 的短语跨度 $psp(n)$ 为子树 T' 中所有节点的词跨度的并集的闭包。

定义 6 若节点 n 的短语跨度 $psp(n)$ 满足如下性质，则称节点 n 的短语跨度是对齐一致的：

$$\forall_{n' \in T \text{ 且 } n' \neq n} psp(n') \cap psp(n) = \emptyset$$

例如，在图 1 中，节点 PP 的短语跨度 $psp(PP) = closure(psp(P) \cup psp(NP-B)) = \emptyset$ 。这里因为节点 P 不满足短语对齐一致性，所以节点 PP 的短语跨度为空。

定义 7 以节点 n 为根的子树 T' ，节点 n 的子树一致性跨度 $dsp(n)$ 定义为：

$$dsp(n) = closure(\bigcup_{n' \in T' \text{ 且 } n' \text{ 是对齐一致的}} wsp(n'))$$

即节点 n 的子树一致性跨度 $dsp(n)$ 为子树 T' 中所有对齐一致节点的词跨度的并集的闭包。

例如，在图 1 中，节点 PP 的子树一致性跨度 $dsp(PP) = closure(dsp(P) \cup dsp(NP-B)) = \{9-10\}$ 。

子树一致性跨度约束不同于短语一致性跨度约束，是一种更为宽松的约束。

2.2 树标记算法

给定三元组 $\langle T, S, A \rangle$ ，树标记的过程可以通过简单地对句法树 T 进行一次后序遍历完成。遍历过程中，对于访问到的每个节点 n ，根据对齐关系 A 设置词跨度 $wsp(n)$ ，然后计算节点 n 的子树一致性跨度。完成树标记后，即可根据每个节点 n 的子树一致性跨度抽取翻译规则。

3 满足子树一致性跨度约束的规则抽取

在基于句法的翻译模型中，主要分为两

大类：成分树模型和依存树模型。在成分树模型和依存树模型的翻译规则抽取中都需要满足短语一致性跨度约束。在成分树模型中，我们以 2006 年(Liu et al., 2006)提出的树到串的翻译模型为例；在依存树模型中，我们以 2011 年 (Xie et al., 2011) 提出的依存树到串的翻译模型为例，分别介绍满足子树一致性跨度约束的规则抽取。

3.1 成分树到串模型中满足子树一致性跨度约束的规则抽取

2006 年(Liu et al., 2006)提出的树到串模型是一种典型的成分树到串翻译模型。树到串翻译模型描述了源语言句法树和目标语言串之间的对齐关系。这个模型是语言学基于句法的(linguistically syntax-based)。在翻译规则的抽取中使用的是短语一致性跨度约束。

2006 年(Liu et al., 2006)提出的树到串翻译模型的规则抽取算法如下：

输入：源语言句法树 $T(f_1^J)$ ，目标语言句子 e_1^I ，词语对齐 A

- [1]后序遍历源语言句法树 $T(f_1^J)$ ，对每个节点进行编号。
- [2]for id :=1 to nodeCount($T(f_1^J)$)
- [3]确定于编号 id 对应的 (树, 串, 对齐) 三元组 $(T,S,A)=identify(id,T(f_1^J),e_1^I,A)$
- [4]如果三元组不满足对齐一致性，则继续(continue)
- [5]否则
- [6]抽取树到串翻译规则
- [7]end for

图 2：满足短语一致性跨度约束树到串模型翻译规则抽取算法

在这个抽取算法中，我们认为可以使用

更为宽松子树一致性跨度约束来代替短语一致性跨度约束抽取翻译规则：

输入：源语言句法树 $T(f_1^J)$ ，目标语言句子 e_1^I ，词语对齐 A

- [1]后序遍历源语言句法树 $T(f_1^J)$ ，对每个节点进行编号。
- [2]for id :=1 to nodeCount($T(f_1^J)$)
- [3]确定于编号 id 对应的 (树, 串, 对齐) 三元组 $(T,S,A)=identify(id,T(f_1^J),e_1^I,A)$
- [4] 如果节点 n 是叶子节点：
判断是否满足短语一致性跨度，如果不满足则继续；
如果节点是非叶子节点：
找出节点 n 的子树一致性跨度。
- [5]抽取树到串翻译规则
- [6]end for

图 3：满足子树一致性跨度约束树到串翻译规则抽取算法

通过以上更为放松的子树一致性跨度约束，我们会抽取更多的有效的翻译规则。例如，在图 1 中，如果我们使用短语一致性跨度约束，我们可以抽取以下规则：

- <" (AD 成功)";"successfully",{(1,1)}>
- <" (ADVP (AD 成功)";"successfully",{(1,1)}>
- <" (NP-B (NR 南非)";" South Africa",{(1,1),(1,2)}>
- <" (NP-B (NR 世界杯)";"FIFA World Cup",{(1,1),(1,2),(1,3)}>
- <" (NR 南非)";"South Africa",{(1,1),(1,2)}>
- <" (NR 世界杯)";"FIFA World Cup",{(1,1),(1,2),(1,3)}>
- <" (VP-B (VV 举行)";"held",{(1,1)}>
- <" (VV 举行)";"held",{(1,1)}>

如果我们使用子树一致性跨度约束则可以抽取以下规则：

```

<" ( AD 成功 )" ,"successfully",{(1,1)}>
<" ( ADVP ( AD 成功 ) )" ,"successfully",{(1,1)}>
<" ( NP-B ( NR 南非 ) )" ,"South Africa",{(1,1),(1,2)}>
<" ( NP-B ( NR 世界杯 ) )" ,"FIFA World Cup",{(1,1),(1,2),(1,3)}>
<" ( NR 南非 )" ,"South ca ",{(1,1),(1,2)}>
<" ( NR 世界杯 )" ,"FIFA World Cup",{(1,1),(1,2),(1,3)}>
<" ( VP ( PP ) ( ADVP ( AD 成功 ) ) ( VP-B ( VV 举行 ) ) )" ,"held successfully",{(1,3),(2,2),(3,1)}>
<" ( VP ( PP ) ( ADVP ( AD 成功 ) ) ( VP-B ) )" ,"X successfully X",{(1,3),(2,2),(3,1)}>
<" ( VP ( PP ) ( ADVP ) ( VP-B ( VV 举行 ) ) )" ,"held X X",{(1,3),(2,2),(3,1)}>
<" ( VP ( PP ) ( ADVP ) ( VP-B ) )" ,"X X X",{(1,3),(2,2),(3,1)}>
<" ( VP-B ( VV 举行 ) )" ,"held",{(1,1)}>
<" ( VV 举行 )" ,"held",{(1,1)}>

```

我们可以发现如果使用短语一致性跨度约束的话，对于规则

```

<" ( VP ( PP ) ( ADVP ( AD 成功 ) ) ( VP-B ( VV 举行 ) ) )" ,"held successfully",{(1,3),(2,2),(3,1)}>
<" ( VP ( PP ) ( ADVP ( AD 成功 ) ) ( VP-B ) )" ,"X successfully X",{(1,3),(2,2),(3,1)}>
<" ( VP ( PP ) ( ADVP ) ( VP-B ( VV 举行 ) ) )" ,"held X X",{(1,3),(2,2),(3,1)}>

```

```

<" ( VP ( PP ) ( ADVP ) ( VP-B ) )" ,"X X X",{(1,3),(2,2),(3,1)}>

```

是无法抽取的，但是如果我们采用子树一致性跨度约束，这些规则是可以抽取的，在实际中，我们发现这种规则多是正确有效的规则。

3.2 依存树到串模型中满足子树一致性跨度约束的规则抽取

在依存树到串的翻译模型中，我们采用2011年(Xie et al., 2011)提出的依存树到串翻译模型。依存树到串的翻译模型的规则抽取算法如下：

-
- [1] 后序遍历源语言句法树，对每个节点进行编号。
 - [2] for 每个可访问到的节点 n
 - [3] if(wsp(n) 是对齐一致性的且对于 $\forall n' \in \text{child}(n) \text{ dsp}(n') \neq \emptyset$ 且 $\text{psp}(n') \cap \text{wsp}(n) = \emptyset$ $n' \in \text{child}(n)$)
 - [4] 根据当前中心词-依存节点及对齐关系生成翻译词汇化规则 R_{lex} 和非词汇化规则 R_{unlex}
 - R_{unlex}
 - [5] 将 R_{lex} 和 R_{unlex} 加到规则表 RT.
 - [6] end for
-

图 4: 满足短语一致性跨度约束依存树到串模型规则抽取算法

如果采用子树一致性跨度约束抽取翻译规则，则抽取算法为：

[1] 后序遍历源语言句法树，对每个节点进行编号。

[2] for 每个可访问到的节点 n

[3] if ($\text{hsp}(n)$ 是对齐一致的且对于 $\forall n' \in$

$\text{child}(n) \text{ dsp}(n') \neq \emptyset$ 且

$\text{dsp}(n') \cap \text{wsp}(n) = \emptyset$
 $n' \in \text{child}(n)$)

[4] 根据当前中心词-依存节点及对齐关系生成翻译词汇化规则 R_{lex} 和非词汇化规则

R_{unlex}

[5] 将 R_{lex} 和 R_{unlex} 加到规则表 RT.

[6] end for

图 5: 满足子树一致性跨度约束依存树到串模型规则抽取算法

在这里， $\text{psp}(n')$ 表示短语一致性跨度约束， $\text{dsp}(n)$ 表示子树一致性跨度约束。

例如，对于“2010年 世界杯 在 南非 成功 举行”的依存分析和对齐结果如图6所示。对于每个节点标出了相应的子树一致性跨度。例如，如果我们使用短语一致性跨度约束则可以抽取规则：

$\langle \text{"(成功)", "successfully"} \rangle$
 $\langle \text{"(南非)", "South Africa"} \rangle$

如果我们使用子树一致性跨度约束，我们可以抽取以下规则：

$\langle \text{"((x1:世界杯)(x2:在)(成功)举行)", "x1 was held successfully x2"} \rangle$
 $\langle \text{"((x1:世界杯)(x2:在)(x3:AD)举行)", "x1 was held x3 x2"} \rangle$
 $\langle \text{"((x1:NR)(x2:P)(成功)举行)", "x1 was held successfully x2"} \rangle$
 $\langle \text{"((x1:NR)(x2:P)(x3:AD)举行)", "x1 held x3 x2"} \rangle$
 $\langle \text{"((x1:世界杯)(x2:在)(x3:AD)x4:VV)", "x1 x2 x3 x4"} \rangle$
 $\langle \text{"((x1:世界杯)(x2:在)(成功)x4:VV)", "x1 x4"} \rangle$

$\text{successfully x2} \rangle$
 $\langle \text{"((x1:NR)(x2:P)(成功)x4:VV)", "x1 x4 successfully x2"} \rangle$
 $\langle \text{"(x1:NR)(x2:P)(x3:AD)x4:VV)", "x1 x4 x3 x2"} \rangle$
 $\langle \text{"成功", "successfully"} \rangle$
 $\langle \text{"南非", "South Africa"} \rangle$

对比发现对于规则：

$\langle \text{"((x1:世界杯)(x2:在)(成功)举行)", "x1 was held successfully x2"} \rangle$
 $\langle \text{"((x1:世界杯)(x2:在)(x3:AD)举行)", "x1 was held x3 x2"} \rangle$
 $\langle \text{"((x1:NR)(x2:P)(成功)举行)", "x1 was held successfully x2"} \rangle$
 $\langle \text{"((x1:NR)(x2:P)(x3:AD)举行)", "x1 held x3 x2"} \rangle$
 $\langle \text{"((x1:世界杯)(x2:在)(x3:AD)x4:VV)", "x1 x2 x3 x4"} \rangle$
 $\langle \text{"((x1:世界杯)(x2:在)(成功)x4:VV)", "x1 x4 successfully x2"} \rangle$
 $\langle \text{"((x1:NR)(x2:P)(成功)x4:VV)", "x1 x4 successfully x2"} \rangle$
 $\langle \text{"(x1:NR)(x2:P)(x3:AD)x4:VV)", "x1 x4 x3 x2"} \rangle$

如果使用短语一致性跨度约束这些规则是无法抽取的，但是如果使用子树一致性跨度约束则是可以抽取的。实际中，这些规则都是有效的规则，在依存树到串的翻译规则的抽取中，如果使用子树一致性跨度约束能抽取更多的有效的翻译规则，有助于提高翻译质量。

System	NIST05(%)
cons2str(baseline)	22.99
cons2str(looser)	25.61(+2.62)
dep2str(baseline)	33.06
dep2str(looser)	33.43(+0.37)

表一: 实验结果

4 实验

我们的实验主要在两个模型上测试，在成分树到串的模型中，我们选用的是刘洋2006在ACL中提出的树到串的系统(Lynx)；

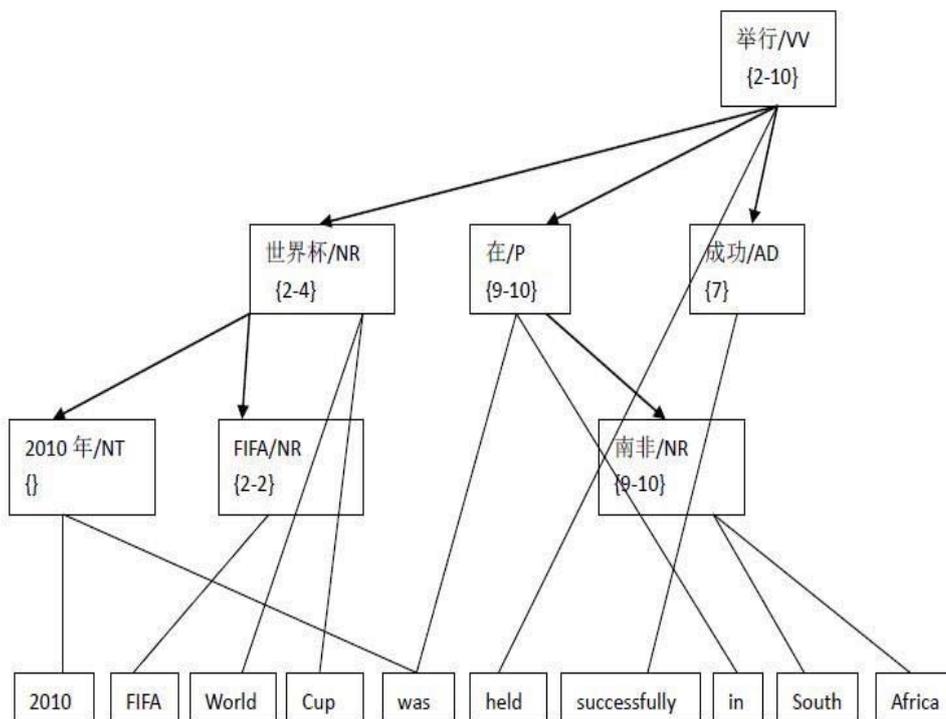


图 6: 子树一致性跨度的依存树

在依存树到串模型中，我们选用的是谢军 2011 年在 EMNLP 中提出的依存树到串的系统。

4.1 成分树到串系统

在成分树到串系统的 Lynx 中，训练语料是 FBIS 语料，语言模型我们选用的是 SRI 语言模型工具 (Stolcke, 2002) 训练的经过 KN 平滑的 4 元语言模型，开发集我们选用是 NIST2002 的 878 句，测试集我们选取的是 NIST2005 的 1082 句。翻译质量的评价，我们选用是 BLEU，测试工具是大小写敏感的 mteval-v11b.pl。对于汉语端的句法分析我们用的是 2005 年熊得意 (Xiong et al., 2005) 开发的汉语句法分析器。对于 Lynx 的配置，我们选择和 Liu2006 一样的配置， $h=3$ 和 $c=5$ ，在开发集和测试集的上，解码器的配置，我们设置 $tatTable-limit=20, tatTable-threshold =$

$0, stack-limit = 50, stack-threshold = 0.0001$. 我们将 Liu 提出的系统作为 baseline 系统，Lynx 测得的 BLEU 值是 22.99，我们通过使用更为宽松的约束方法测的 BLEU 值是 25.61，提高了 2.62 个点。

4.2 依存树到串系统

在依存树到串系统中，我们选用的是谢军 2011 年 EMNLP 中提出的依存树到串的模型，我们采用的训练语料是 1.5M 的 LDC 的语料，我们用的是 Stanford Parser (Klein and Manning, 2003) 进行句法映射成依存结构，节点被标记成 POS 标记，边被标记成 typed 依存。我们的词语对齐用的是 Och 提出的 GIZA++ (Och and Ney, 2003) 获得，语言模型是 SRI 工具 (Stolcke, 2002) 获得的经过 KN 平滑的 4 元语言模型。我们使用 NIST02 作为开发集，NIST05 作为测试集。翻译质量

评价是用大小写敏感的 NIST BLEU-4 评价。我们以谢军的依存树到串系统作为 baseline 系统,测得 BLEU 值是 33.06,经过放松约束后,我们测的 BLEU 值是 33.43,提高了 0.37 个点。

通过在成分树到串的 Lynx 系统和依存树到串系统上,我们发现经过放松约束抽取翻译规则后,翻译的质量都有了较大的提高,表明翻译规则抽取算法中,使用更为宽松的子树一致性跨度约束抽取翻译规则有助于提高机器翻译的质量。

5 结论

在本文中,我们从更细粒度的语言学知识的角度,提出了在基于句法翻译模型的翻译规则自动获取算法设计中,使用更为宽松的子树一致性约束来代替传统的短语一致性跨度约束抽取翻译规则,并在基于句法的两大模型成分树到串翻译模型和依存树到串翻译模型中得到了验证。因此在基于句法的翻译模型的抽取规则部分,使用更为宽松的子树一致性约束能够有效地提高机器翻译的质量。

参考文献

- Aravind K Johsi, L Levy, and M Takahashi. Tree adjunct grammars. *Journal of Computer and System Sciences*,10(1):136-163,1975
- Michel Galley, Jonathan Graehl, Kevin Knight, Daniel Marcu, Steve DeNeefe, Wei Wang, and Ignacio Thayer. Scalable inference and training of context-rich syntactic translation models. In *Proceeding of ACL 2006*, pages 961-968, Sydney, Australia, July 2006. Association for Computational Linguistics
- Michel Galley, Mark Hopkins, Kevin Knight, and Daniel Marcu. 2004. What's in a translation rule? In *Proceedings of NAACL-HLT 2004*, pages 273-280
- Kenji Yamada and Kevin Knight. A syntax-based statistical translation model. In *Proceedings of the Conference of the Association for Computational Linguistics*.
- L.Huang, K.Knight, and A. Joshi. A syntax-directed translator with extended domain of locality. In *Proceedings of the Workshop on Computationally Hard Problems and Joint Inference in Speech and Language Processing*, pages 1-8. Association for Computational Linguistics.2006.
- Yang Liu, Qun Liu, and Shouxun Lin. 2006. Tree-tostring alignment template for statistical machine translation. In *Proceedings of ACL 2006*, pages 609-616, Sydney, Australia, July.
- Haitao Mi and Liang Huang. Forest-based translation rule extraction. In *Proceedings of EMNLP2008*, pages 206-214, Honolulu,Hawaii, October 2008.
- Min Zhang, Hongfei Jiang, Aiti Aw, Haizhou Li, Chew Lim Tan, and Sheng Li. A tree sequence alignment-based tree-to-tree translation model. In *Proceeding of ACL-08:HLT*,pages 559-567, Columbus, Ohio, June 2008. Association for Computational Linguistics.
- Libin Shen and Ralph Weischedel. String-to-dependency statistical machine translation. *Computational Linguistics*.
- Libin Shen, Jinxi Xu, and Ralph Weischedel. A new string-to-dependency machine translation algorithm with a target dependency language model. In *Proceedings of ACL 2008: HLT*,pages 577-585, Columbus, Ohio, June 2008. Association for Computational Linguistics.
- Deyi Xiong, Qun Liu, and Shouxun Lin. A dependency treelet string correspondence model for statistical machine translation. In *Proceedings of the Second Workshop on Statistical Machine Translation*, pages 40-47, Prague, Czech Republic, June. 2007.

Jun Xie, Haitao Mi, Qun Liu. A novel dependency-to-string model for statistical machine translation. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pages 216-226, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.

Yuan Ding and Martha Palmer. Synchronous dependency insertion grammars: A grammar formalism for syntax based statistical mt. In COLING 2004 Recent Advances in Dependency Grammar. In Proceedings of ACL 2005.

Research on Dependency-to-String Statistical Translation Models. PhD thesis, Institute of Computing Technology, Chinese Academy of Sciences, June, 2012.

Franz J. Och. 2002. Statistical machine translation: From single-word models to alignment templates. Ph.D. thesis, Computer Science Department, RWTH Aachen, Germany, October.

Heidi J. Fox. Phrasal cohesion and statistical machine translation. In Proceedings of EMNLP2002, pages 304-311, 2002.

Deyi Xiong, Qun Liu, and Shouxun Lin. A dependency treelet string correspondence model for statistical machine translation. In Proceedings of the Second Workshop on Statistical Machine Translation, pages 40-47, Prague, Czech Republic, June 2007.