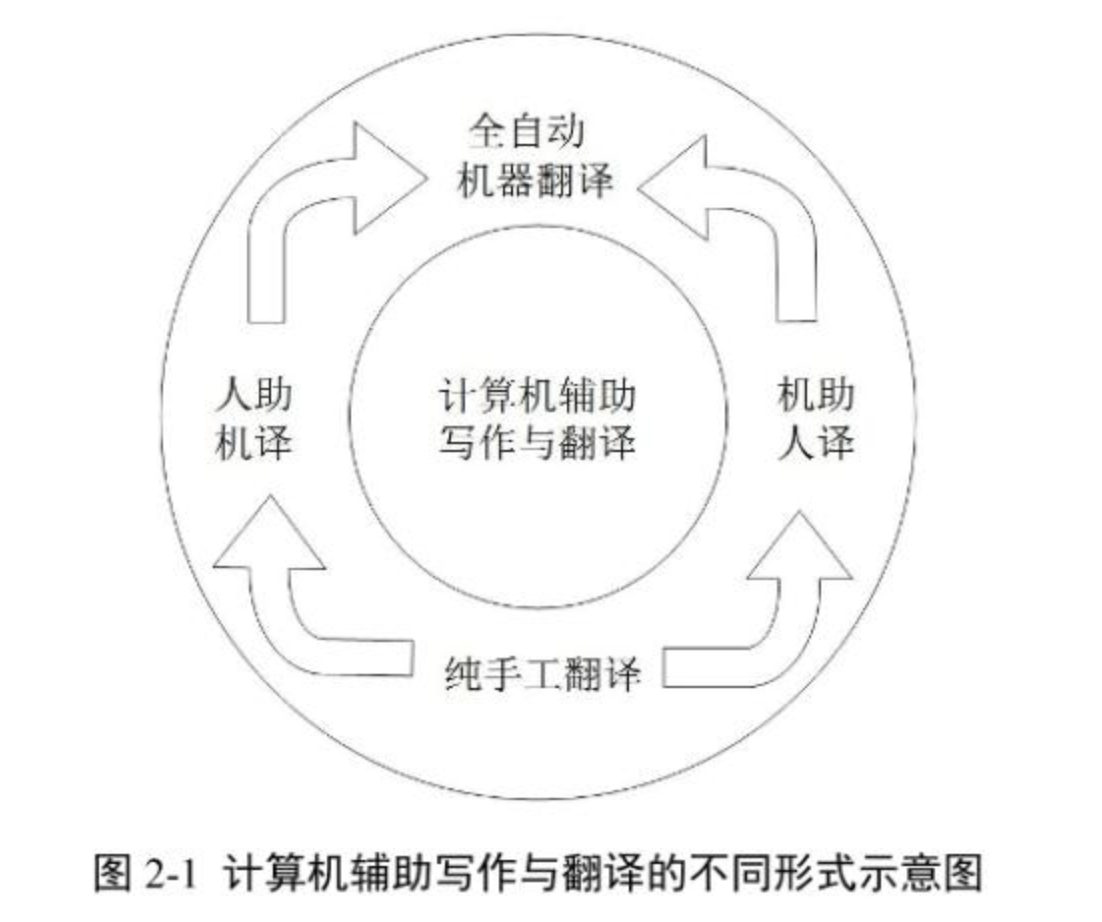
1.技术原理（李老师部分）

1计算机辅助翻译的不同形式

人们研究机器翻译的初衷就是为了让它帮助翻译工作者提高工作效率和翻译质量，其最高的期望就是实现全自动的高质量的机器翻译FAHQMT(FullAutomatedHighQualityMachineTranslation)“，但基于目前的发展现状，这一目标在短期内是难以实现的。

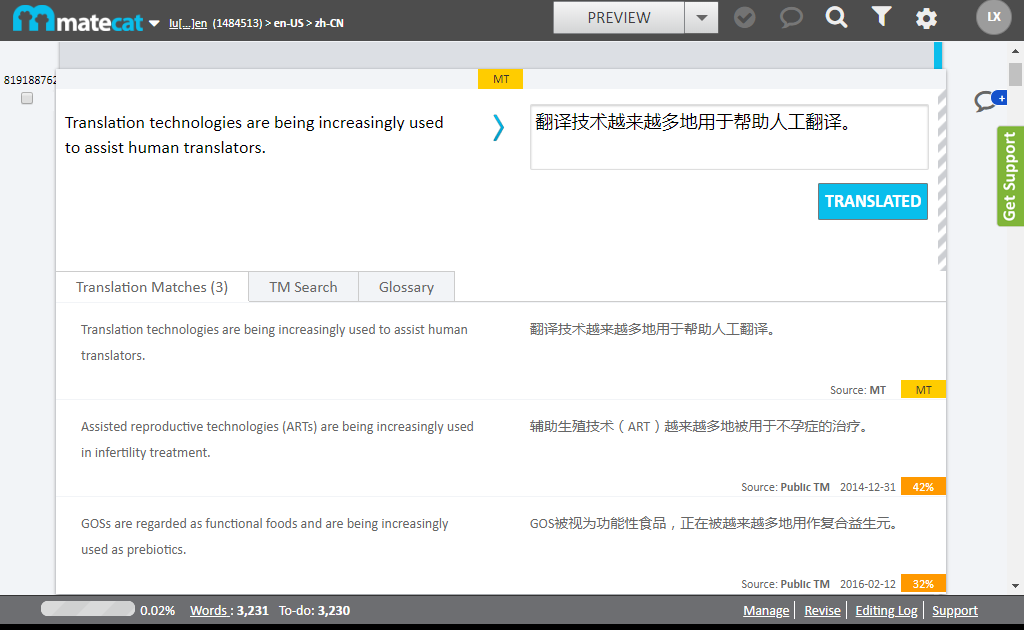


“机助人译”形式核心是计算机的预翻译和辅助译后编辑这两个部分，计算机预翻译可以通过基于统计机器翻译、基于翻译记忆库、基于规则辅助引擎获得支持。“机助人译”这种形式的重要一点在于为用户提供一个具有多格式兼容的译后编辑平台。这也是多数商业化计算机辅助翻译工具的卖点所在。“人助机译”形式的计算机辅助翻译的则是重在交互和“人助”的过程。交互是指用户在写作与翻译的时候，通过一个友好的人机交互接口，与计算机互动，指导计算机输出更高质量或者说满足人类写作和翻译质量的结果；如果计算机能记录用户工作的交互过程的日志，通过对它们的学习，改进自己的工作模式和输出质量，也是很重要的研究内容。在整个人助机译的交互过程中，从计算机的角度看，是在帮助用户写作和翻译，从用户的角度看，是用户在指导计算机，帮助计算机作出更准确的判断与预测。

**译后编辑**

给出机器翻译，由译员在此基础上进行修改，这叫做译后编辑模式。机器翻译可以让译员快速地了解原文的大致意思，如果结果不错，译员只需进行一点润色便可以完成翻译。

大多数CAT软件都支持调用机器翻译。在线CAT平台MateCAT，以译后编辑为主，并且除了机器翻译，还支持调用在线翻译记忆服务MyMemory进行辅助翻译。在创建项目时还需要选择领域，以利用针对性更强的资源。



MateCAT还可以根据用户译后编辑的结果修改背后的机器翻译的模型参数，做增量式学习，这样一个机器翻译的错误被纠正后，下次就不用再纠正了。SDL Trados 2017的Adaptive MT功能和此类似。和MateCAT类似的国内CAT平台有yeekit。

但是机器翻译有时不能输出一个很好的句子结果，另外这种在机器翻译上修改的模式显得没有创造性。而译后编辑的模式也并不是一个很好的人机交互模式[2](http://blog.xulihang.me/use-machine-translation-to-help-human-translation/" \l "fn:atanet)。于是就有了另一种模式。

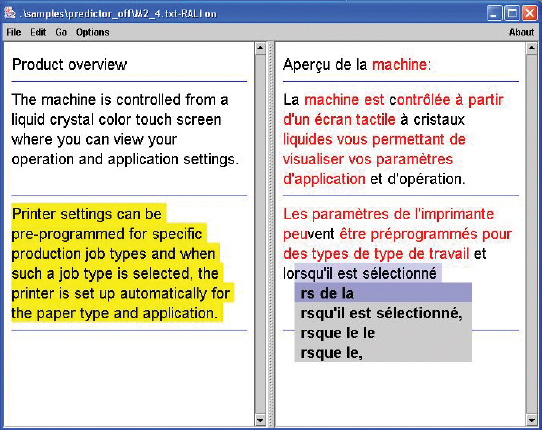
**交互式机器翻译**

Kay最早研究了交互式机器翻译(InteractiveMachineTranslation，IMT)技术，其中用户的工作被限制为消除原语句子中的句法歧义，规则系统负责对经过消歧处理的原文进行翻译，用户不必对目标语进行处理，甚至不需要具有任何关于目标语的语言知识。这种方法在某些特定应用中具有一定优势，但同时也存在着严重的不足，特别是用户不能直接控制最终译文的质量，以及对于熟练的翻译人员来说消除歧义是否真的比直接翻译更容易也存在疑问。

为此，Foster等开展了一项有趣的研究工作，提出使用词预测(wordprediction)技术来实现在人机交互中的预测式交互翻译(interactive—predictivemachinetranslation，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，

## TransType

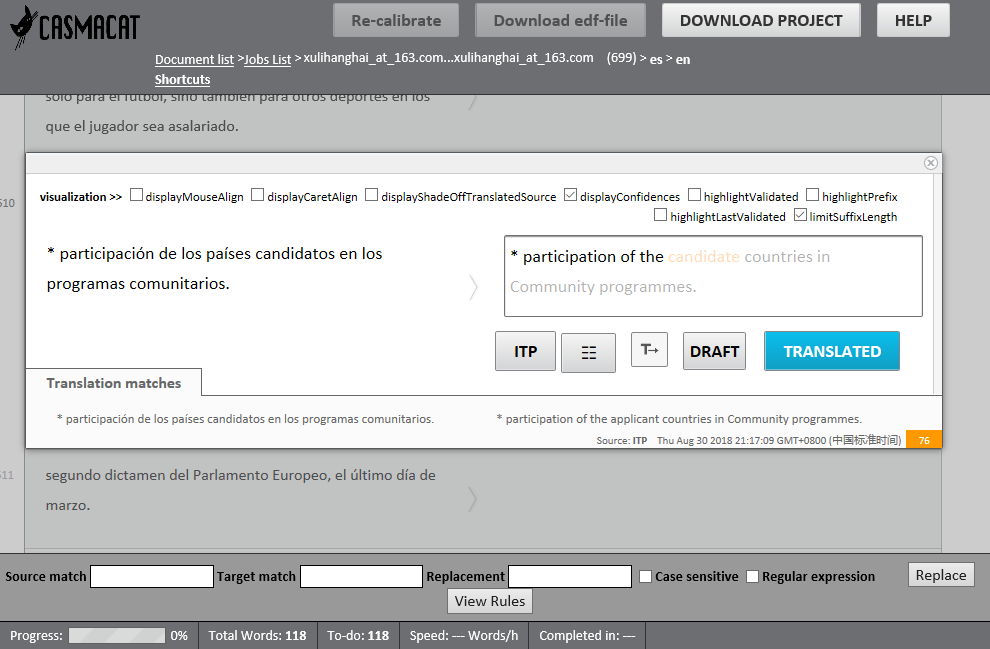
使用在线机器翻译的用户一般不知道统计式机器翻译系统可以生成一串备选翻译，TransType根据用户的输入对备选翻译列表进行筛选，得到合适的翻译建议。结果会以下拉列表的形式呈现。



之后的几年当中，又有很多的研究人员对TransType系统进行了改进。Langlais等人在2000年对系统的用户界面和词的预测提出了改进[2j。2002年，由许多欧盟研究机构共同参与的TransType2项目，创新性的把一个完全的基于数据驱动的机器翻译系统嵌入到交互式翻译框架中，并且在每一次的交互过程中，翻译系统都会根据翻译人员给出的译文前缀，预测出一个或者多个最好的后缀补全译文，供翻译人员选择。在TransType2项目中，很多的研究人员对系统进行分析，并且提出很多种方法来解决这些问题。TransType的这两个项目极大的推动了交互式机器翻译技术的发展。

## CASMACAT

CASMACAT（Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation）[3](http://blog.xulihang.me/use-machine-translation-to-help-human-translation/" \l "fn:casmacat)基于MateCAT开发，它的一个特点是机器翻译可以根据用户的操作不断学习完善。它使用了一个专门的统计式机器翻译系统，和翻译界面的耦合程度高。用户输入翻译后（前缀），系统会给出一个机器翻译结果（后缀），单词还可以设置用颜色来表示置信程度。



CASMACAT可以记录用户的翻译过程，以便进行相关研究。

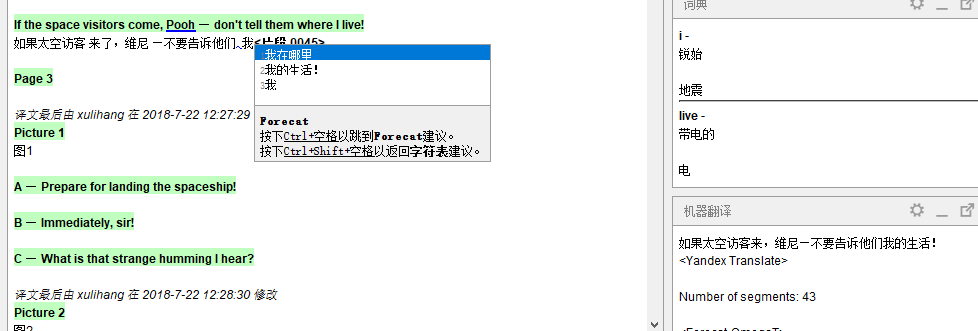
## Forecat

在2014年，Pérez-Ortiz等人提出了一个黑盒的方法。它可以使用任意机器翻译系统、翻译记忆和双语词典提供的双语资源，为翻译提供参考。该方法将原文以n-gram的办法分为多个片段，然后每个片段都利用双语资源进行翻译，然后根据一定的排序方法（目前是简单的距离方法），给用户提供有限的几个建议。用户输入一个前缀，就会做进一步筛选。因为使用了各种双语资源，再叫交互式机器翻译就不合适了，一般叫做交互式翻译预测。黑箱的方法更容易理解，运行成本低，可利用资源多，试验结果显示比CASMACAT这样的白箱方法效果更好。[4](http://blog.xulihang.me/use-machine-translation-to-help-human-translation/" \l "fn:compare)

Forecat将原文拆分为若干片段，片段的最大长度可以定义，默认为4个词。然后，它使用机器翻译或者翻译记忆翻译这些片段，然后根据用户输入匹配译文片段。匹配译文开头时，如果开头一样，还要根据输入框中的位置和对应原文的位置进行排序。这样得到的片段会很多，很多距离较远的片段其实提供的参考价值更大，所以一般采用所处编辑位置最长和最短的片段。

建议以下拉列表的方式提供，默认显示4个建议。

Forecat使用Java编写，除了提供Web应用外，还可以集成到开源CAT工具OmegaT中使用。



2010年，Ortiz和Casacuberta等人，将在线学习的思想加入到了交互式机器翻译技术当中。其主要思想是利用用户的反馈信息来不断的完善系统的底层模型[a3。Gonzdlez-Rubio和Or—tiz等人，将机器译文的置信度评价作为其是否需要和翻译人员进行交互的衡量，从而有效地平衡了翻译人员的工作量和系统翻译结果的准确率[4]。2012年，Gonzdlez—Rubio和Ortiz等人[5]，将动态学习的方法引入到交互式机器翻译系统当中，使系统可以增量式的从已经翻译完的句子中学习，从而明显地提高后续句子的翻译准确率，有效减少了翻译人员的工作量。2013年，JesflsGonzdlez-Rubio和Dan—ielOrtlz—Martinez等人L6]，将基于层次短语的翻译模型应用到了交互式机器翻译当中，并且采用了超图作为机器和用户之间的交互接口。

**TranSmart原理（zz部分）**

Source：腾讯AI Lab：深度解读AI辅助翻译的研究及应用

<https://cloud.tencent.com/developer/article/1363107>

Outline：

1. 简单介绍机器翻译及其困难，引出为什么要有交互式机器翻译
2. 全自动输出翻译结果 vs 人机交互输出翻译结果
3. 翻译技术及其原理

3.1 “一评估三交互”

3.2.1 交互一：译后编辑

3.2.2 交互二：交互式机器翻译

3.2.3 交互三：翻译输入法

3.2 在线学习方法

3.3 翻译记忆

**机器翻译的困难**

机器翻译领域特有的一些困难。

我们一般认为，之所以目前的机器翻译质量还不够好，下边这些原因占了很大一部分。

首先，自然语言中普遍存在的歧义和未知现象始终难以被已有训练数据覆盖并被模型及时学习到，即便能覆盖也可能有长尾问题；其次，翻译本身并不仅仅是字符串的转换，比如我们并不能直接简单通过字面就能正确翻译“青梅竹马”这个词；再者，翻译的解并不唯一，始终存在人为的标准，这样直接做训练模型的损失函数失效；最后，有的翻译问题实在太难了，可能需要翻译大家穷其一生，更遑论机器翻译模型。比如在见到“最是那一低头的温柔，像一朵水莲花不胜凉风的娇羞”的时候，机器翻译几乎只能选择启动自杀程序（如果有的话）。

在这种背景下，除了极少一部分机器翻译能完成得很好的翻译任务以外，如果要正式使用自动译文，几乎都需要不同程度的人工介入。

这是我们为什么投入人机交互式机器翻译的最初动机。

**全自动输出翻译结果 vs. 人机交互输出翻译结果：**

**全自动输出翻译结果：**

无法保证译文质量。

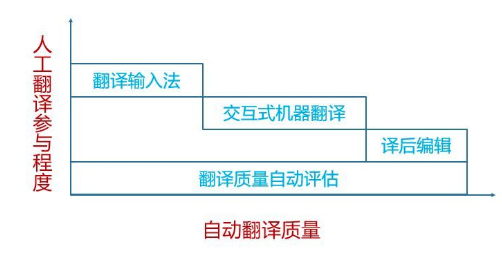
**人机交互输出翻译结果：**

1. 接受用户提供译文干预
2. 及时学习用户修改反馈
3. 实时提供翻译辅助信息

**人机交互式机器翻译技术及其原理（以TranSmart为例）**

**“一评估三交互”**

腾讯AI Lab根据深度调研的结果，总结出“一评估三交互”的人机交互翻译范式。



**“三交互”**

三种人机交互形式：翻译输入法、交互式机器翻译和译后编辑，分别适用于自动译文质量比较差、质量尚可和质量较优的情形。

“**一评估**”

一评估指贯穿全过程的机器翻译自动译文质量评估技术，在具体的应用场景中，通过自动翻译质量评估，向用户推荐合适的交互手段。

在人机交互翻译范式中，**机器翻译需要完成的任务**可以分成六类：

1. 整句更新，理解用户干预，输入当前句子更好的自动译文供人工译员参考和决策；

2. 优质翻译片断提示，以便用户快速采用机器翻译建议；

3. 翻译输入法，方便用户在翻译过程中精准组词，以快速录入人工译文；

4. 在线学习，及时学习用户修改反馈；

5. 语义理解，实时提供翻译辅助信息；

6. 快速解码， 所有交互过程必须实时响应。

**“三交互”中的技术原理**

1. **译后编辑**



当自动译文质量比较好的情况下，通过译后编辑这种模式就可以快速完成翻译任务。

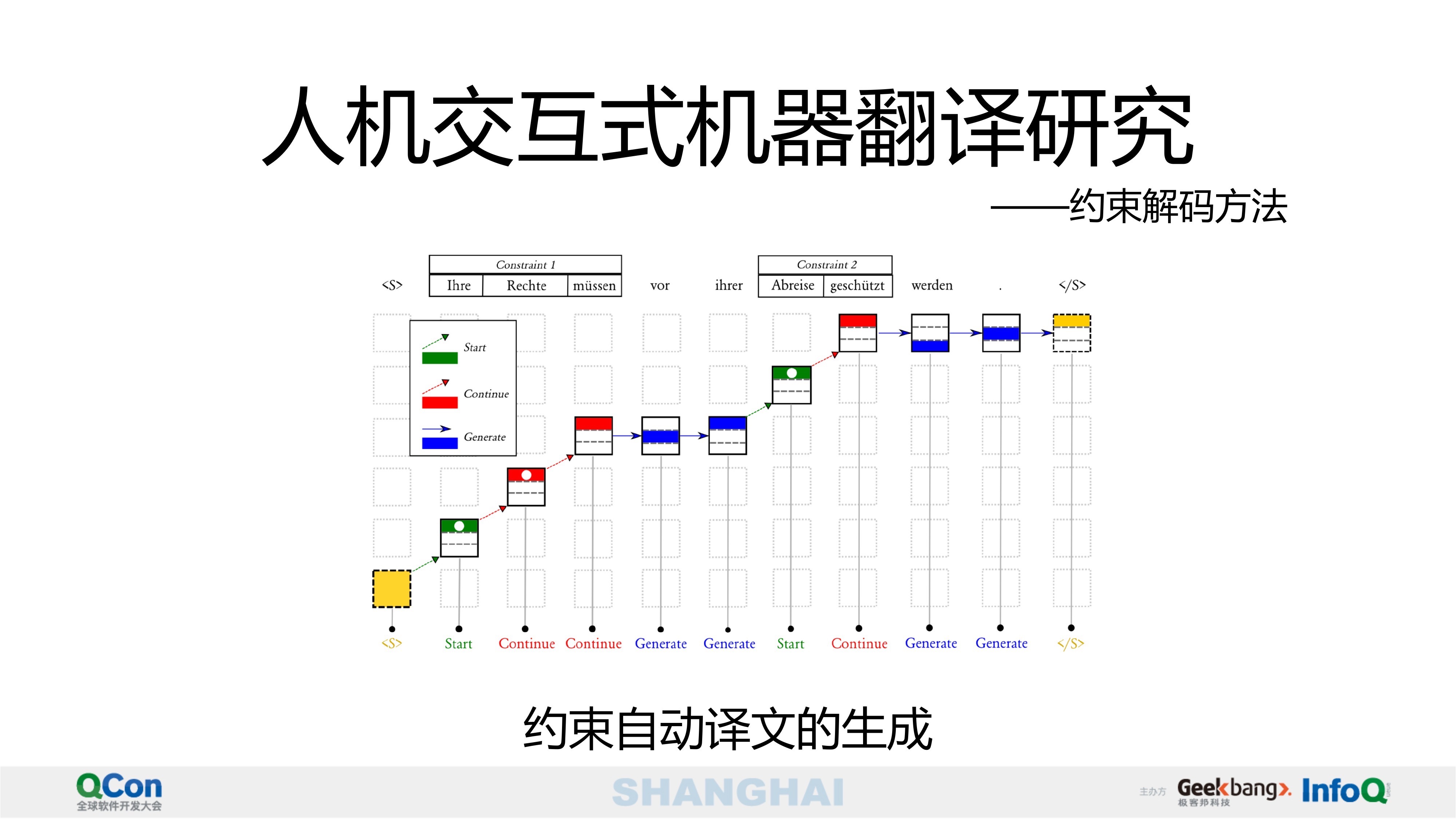
所谓译后编辑，简单而言就是指通过人工修改自动译文来完成翻译。主要的技术难点是如何准确自动推荐优质译文，业界通用的技术为Quality Estimation，即无参考译文的机器翻译质量评估。实现途径一般是特征工程，同时辅以人工翻译规则。

1. **交互式机器翻译**



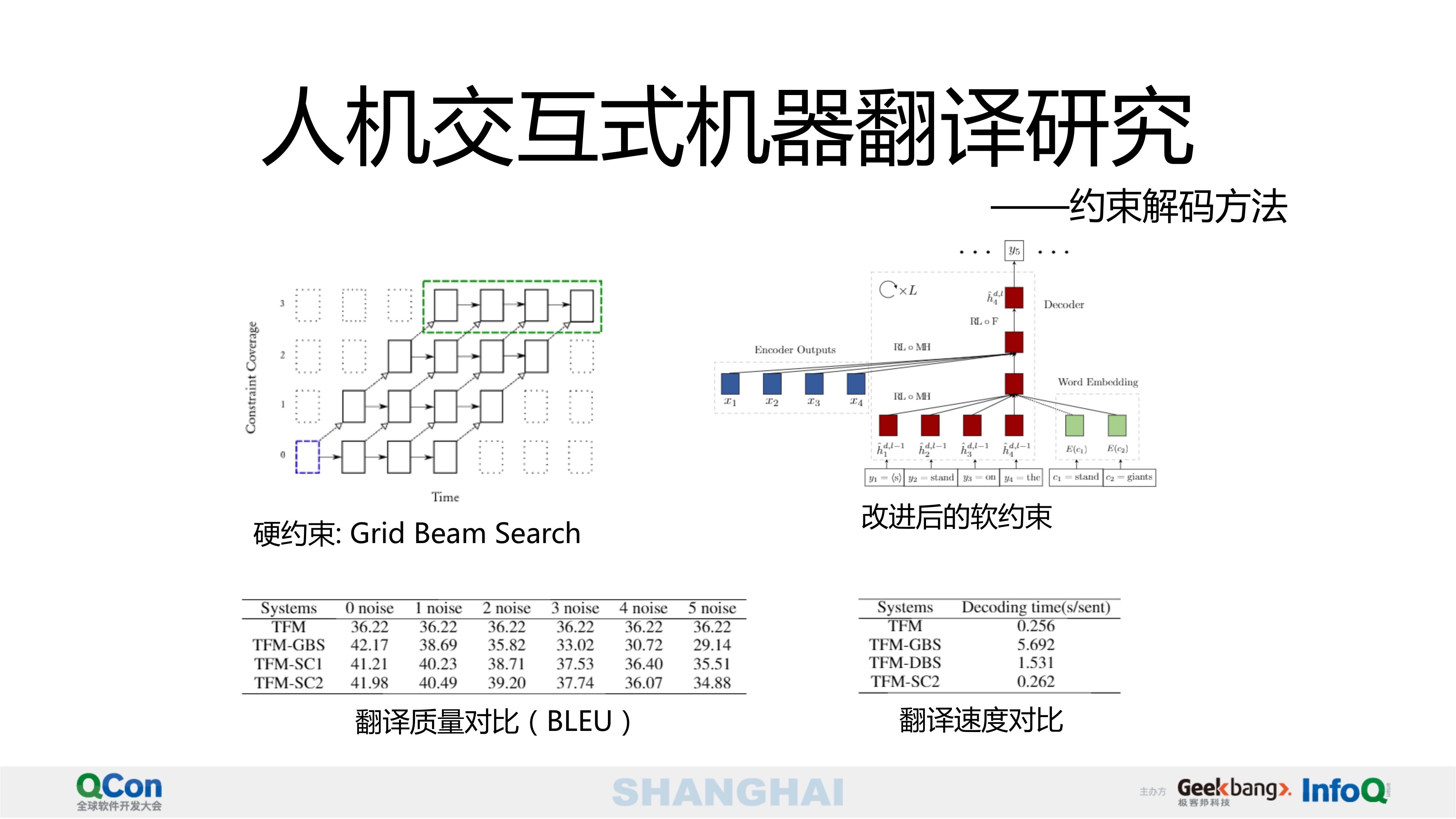
当机器译文质量尚可，但缺乏译后编辑价值时，就需要使用交互式机器翻译技术。交互式机器翻译的中心思想在于，机器翻译根据用户已输入片断，提供片断级的翻译建议。优化目标是通过尽可能少的人工干预，生成尽可能准确的长翻译片断。主要的技术难点在于如何理解用户给定的干预信息，干预信息包括但不限于译文片断、术语以及不希望出现的词。

**交互式机器翻译的技术核心：约束解码方法**



在通用的机器翻译系统中，模型的推理阶段一般采用标准的Beam Search算法。但在人机交互式机器翻译系统中，生成的自动译文往往需要满足一定约束，比如前缀约束、后缀约束、译文中必须出现或者必须不出现哪些片断。

比较简单且成熟的解码方法称为Grid Beam Search，即在推理的每一个时刻都试图扩展约束，最后输出满足所有约束的翻译假设。优点是代码简单直接，缺点是解码速度下降比较明显，且对约束的噪声比较敏感。



为了改进GBS的缺陷，将约束引入了训练过程，即将约束作为额外特征引入模型，经实验证明，能达到与GBS基本一致的解码质量，能容忍最多60%的约束噪声，且翻译速度与标准Transformer模型基本一致。

1. **翻译输入法**



译后编辑与交互式机器翻译的一个共同点都是需要用户阅读理解自动译文，如果自动译文质量比较差的情况下，这是不被用户所接受的。在比较严肃的领域，如医疗、医药、法律等领域是经常发生的事。

在这种情形下，机器翻译就不能辅助人工翻译提高翻译效率吗？

答案是完全可行的，途径就是翻译输入法，在用户不阅读理解自动译文的前提下，通过机器翻译提升人工翻译效率。

例如，用户希望录入的最终译文是“中国考虑改革公务员福利制度”，一般在录入时，我们都希望一次能多输入几个字。搜狗、谷歌、微软输入法虽然都很优秀，但是很遗憾，由于缺乏对翻译上下文的感知，如果输入的拼音串太长，则很容易解码不成功。

2.工具测试

3.痛点难点

交互式机器翻译动机与定位：除了极少一部分机器翻译能完成得很好的翻译任务以外，如果要正式使用自动译文，几乎都需要不同程度的人工介入

延迟感 实用性低

韩林涛 tranSMART

术语未收录

快捷键

术语库收录不对

机器翻译不够好

黄国平：

站在人工翻译侧，机器翻译需要能够接受用户提供的译文干预，及时学习用户的修改反馈，同时实时提供翻译辅助信息

一评估指贯穿全过程的机器翻译自动译文质量评估技术，在具体的应用场景中，通过自动翻译质量评估，向用户推荐合适的交互手段；

三交互指三种人机交互手段，分别是翻译输入法、交互式机器翻译和译后编辑，分别适用于自动译文质量比较差、质量尚可和质量较优的情形。

在“一评估三交互”的人机交互翻译范式中，机器翻译系统承担助手角色，辅助人工翻译在不降低译文质量的前提下显著加快翻译效率。

因此，机器翻译需要承担的任务可以分为六类：

1. 整句更新，理解用户干预，输入当前句子更好的自动译文供人工译员参考和决策；

2. 优质翻译片断提示，以便用户快速采用机器翻译建议；

3. 翻译输入法，方便用户在翻译过程中精准组词，以快速录入人工译文；

4. 在线学习，及时学习用户修改反馈；

5. 语义理解，实时提供翻译辅助信息；

6. 快速解码， 所有交互过程必须实时响应。

可以看到，人机交互式机器翻译既有工程技术的难题，也有目前难以找到有效解决方案的学术问题。

首先要定义明确的需求，即我们需要什么语种、什么领域、用于什么场景，以及如何现有业务系统打通；然后准备足量的双语平行句对作用训练语料；接着集成已经可以应用的最新技术，比如输入法、术语抽取、翻译片断挖掘等；再然后就是调试GPU集群、并行加速训练等；最后部署上线并迭代更新。

往往会出现很多现实问题。比如最终部署的系统并不能完全预定任务，主要可能是对自动译文质量过高的期望，或者发现开源系统不是那么容易填坑，另外就是如何甄别出切实可用的技术，并拉通目标场景的技术链条。

机器翻译观察：

各行各业的“非专业”译者可以凭借自己的专业知识，快速扩大这个原本不足百万的用户群体。

:第一, 传统IMT方法通过逐字确认前缀的方式自左向右补全句子, 一个短语需要解码多次才能被翻译正确;第二, 翻译正确的译文后缀会随着前缀的更改重新解码, 有时原本正确的译文在增加了前缀约束后会被翻译需要译员重新确认, 影响交互效率;第三, 限制了人机交互方式, 译员无法从其他方式给予更充分的指导;第四, 一些研究人员对传统交互方式进行扩展, 允许译员在译文中选择单语片段，但是从错误的译文里识别正确的译文片段十分困难, 增加了译员的认知负担。

摘自：基于双语短语约束的交互式机器翻译方法

此论文中的方法：1.译员可以选择术语的解释

2.译员可以拆分短语

3.可选短语的排序

本次目标：减少译员工作量和认知负担、应用于专业领域、

解决的问题：减少系统反应时间和键盘敲击率，译文后缀预测的精确度，收集语料的使用（译员标记正确，怎么学习），只有前缀约束

建议：源语与目标语言对齐展示，后缀译文最大化，减少每个字符都有无用提示。与译前编辑结合？加入HMM对齐模型，使用户反馈信息不断完善翻译系统；词图作为交互接口进行交互，搜索不到合适译文时再解码。动态学习，

交互：图形界面、鼠标

问题：只能从左至右，一个短语需要输入多次前缀；译员无法对其进行其他指导，

有时候后缀正确，但是输入前缀后后缀改变，→基于片段的

版本控制

不能翻译文档 文档的交互？

4.优化