

机器翻译

维基百科，自由的百科全书

機器翻譯（英語：Machine Translation，經常簡寫為MT，簡稱**机译**或**机翻**）屬於計算語言學的範疇，其研究藉由计算机程序將文字或演說從一種自然語言翻譯成另一種自然語言。簡單來說，機器翻譯是通过將一個自然語言的字辭取代成另一個自然語言的字辭。藉由使用語料庫的技術，可達成更加複雜的自動翻譯，包含可更佳的處理不同的文法結構、辭彙辨識、慣用語的對應等。

目前的機器翻譯軟體通常可允許針對特定領域或是專業領域（例如天氣預報）來加以客製化，目的在於將辭彙的取代縮小於該特定領域的專有名詞上，以藉此改進翻譯的結果。這樣的技術适合針對一些使用較正规或是較制式化陳述方式的領域。例如政府機關公文或是法律相關文件，這類型的文句通常比一般的文句更加正式與制式化，其機器翻譯的結果通常比日常对话等非正式场合所使用语言的翻译结果更加符合语法。

目前的翻译机器，有時可以得到可以理解的翻译结果，但是想要得到較有意義的翻譯結果，往往需要在輸入語句時適當地編輯，以利電腦程式分析。

但是，機器翻譯的結果好壞，往往取決於譯入及譯出語之間的詞彙、文法結構、語系甚至文化上的差異，例如英语與荷兰语同為印歐語系日耳曼語族，這兩種語言間的機器翻譯結果通常比汉语與英语間機器翻译的結果好。

因此，要改善機器翻譯的結果，人為的介入仍顯相當重要。

一般而言，大眾使用機器翻譯的目的只是為了获知原文句子或段落的要旨，而不是精確的翻譯。總的来说，機器翻譯的效果并没有达到可以取代人工翻译的程度，所以無法成為正式的翻譯。

不過現在已有越來越多的公司嘗試以機器翻譯的技術來提供其公司網站多語系支援的服務。例如微軟公司試將其 MSDN 以機器翻譯來自動翻譯成多國語言，如上文所说，知识库作为专业领域，其文法较为制式化，翻译结果亦更加符合自然语言。



302	云南菜爆松茸	Sautéed tricholoma matsutake with coriander and garlic	158元
303	白油爆鸡枞	Stir-fried wikipedia	158元
	云南辣鸡枞	Stir-fried wikipedia with pimentos	158元
304	香油鸡枞蒸水蛋	Steam eggs with wikipedia	48元
305	寸金蒜片油鸡枞	Fried special wikipedia	128元
306	双椒牛肝菌	Sautéed king truffle with mixed pimentos	88元

機器翻譯有時會得出這樣令人難以理解又啼笑皆非的翻譯結果

目录

歷史

翻譯流程

方法

[規則法](#)

[轉化法](#)

[辭典法](#)

[知識翻譯](#)

[範例法](#)

[統計法](#)

参见

参考来源

外部連結



巴厘岛景区上的 no entry（禁止進入）被機器翻譯成「沒有進入」



巴厘岛景区上的 forbidden to stand on（禁止站在上方）被機器翻譯成了「嚴禁站在」

歷史

機器翻譯的概念最早可追溯到17世紀。1629年，哲學家笛卡兒提出了世界語言的概念，即將不同語言相同含義的詞匯以統一符號表示。笛卡兒、萊布尼茲、貝克（Cave Beck）、基爾施（Athanasius Kircher）以及貝希爾（Johann Joachim Becher）等人曾試圖編寫類似世界語言的辭典。直到近代，藉由機械的輔助，機器翻譯的可行性大為提升。20世紀初期便有多位科學家與發明家陸續提出機器翻譯的理論與實作計畫或想法。沃伦·韦弗被誉为机器翻译的鼻祖。他拋却了俄语文本的含义，转而视为一堆“密码”。在美国和欧洲，他的团队和继任者在工作时都遵循着一个常理：“任何语言都是由一堆词汇和一套语法规则组成。只要把两种词汇放到机器里，按照人类组合这两种词汇的方式，为之建立一套完整的规则，机器就能破译“密码”。”

1954年乔治城-IBM实验中，成功將約60句的俄文自動翻譯成英文，被視為機器翻譯可行的開端。自此開始的十年間，政府與企業相繼投入相當的資金，用於機器翻譯的研究上。然而，自動語音處理顧問委員會（Automatic Language Processing Advisory Committee，簡稱ALPAC）在1966年提出的一項報告中表明十年來的機器翻譯研究進度緩慢，未達預期。該項報告使得之後的研究資金大為減縮，直到近1980年代，由於電腦運算科技的進步，以及演算成本相對降低，才使政府與企業對機器翻譯再次提起興趣，特別是在統計法機器翻譯的領域上。

翻譯流程

從人為的翻譯來看機器翻譯，翻譯的過程可被細分如下：

1. 解譯來源文字的文意
2. 重新編譯此解析後所得的文意至目標語言。

在這看似簡單的步驟之後其實是複雜的認知操作。要能解譯來源文字的完整意義，一個譯者必須能夠分析與詮釋整段文章的所有特徵，必須能夠深度的了解其文法、語義、語法、成語等等，相當於了解來源語言的文化背景。譯者同時也必須兼備目標語言相同深度的知識。

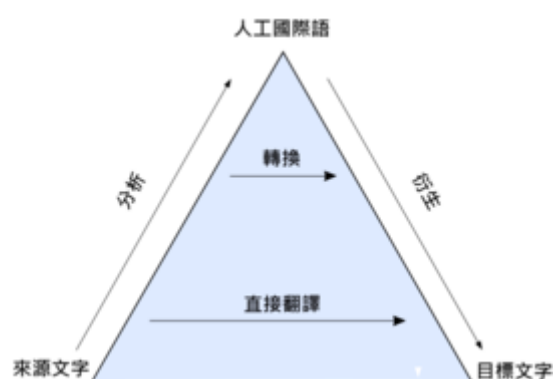
於是，這對機器翻譯便是一項挑戰，即：要如何設計一個程式使其能夠如同真人一樣的「了解（認知）」一段文字，並且能夠「創造」一段好似真人實際寫作出來的目標語言的文字。

方法

機器翻譯可以使用一種基於語言規則的語法，文字將會依語言學的方式來進行翻譯，即一個最合適的目標語言的字詞將會被用來取代來源語言的字詞。

能夠優先解決對自然語言的正確認知與辨識，被視為機器翻譯是否能夠成功的最主要關鍵。

一般而言，用規則法（rule-based method）分析一段文字，通常會先建立目標語言中介的、象徵性的表義字詞。再根據這中介的表義字詞來決定使用人工國際語言（interlingual）化的機器翻譯，或是使用轉化原則法的機器翻譯（transfer-based machine translation）。這些方法都必須擁有具備足夠形態學的、語句學的、以及語義學的資訊以及大量的字詞規則所建構的辭彙。



常見機器翻譯的難處在於無法給於適當且足夠龐大的資訊，來滿足不同領域或是不同法則的機器翻譯法。舉例來說，對於一個需要統計學法則的翻譯法，給予它大量的多語言素材是必要的，但對於文法式法則的翻譯法便顯得沒有太大意義。

規則法

規則法機器翻譯的範例包含了轉化法（transfer-based）、中間語法（interlingual）、以及辭典法（dictionary-based）機器翻譯。

轉化法

辭典法

機器翻譯可利用辭典的詞彙作翻譯。因為這種翻譯是「字對字」的，所以通常各字之間在意思上都沒有任何關聯。這種機器翻譯法最適用於具有冗長的詞語列表（意即非完整的句子）。例如產品型錄的翻譯。

知識翻譯

範例法

所谓范例法，即基于实例的翻译方法。基本思路是电脑模拟大量翻译实例（翻译语料库），进行有效替换的翻译策略。因此该方法依赖于翻译语料库的质量、规模和覆盖面。如果有完全一样的例句，则直接采用范例的译文；如果有多个相似的例句，则自动模拟相似度最高的译文，只需翻译不同部分即可；如果没有相似的译文，则必须进行基于统计或规则的方法进行翻译。根据乔姆斯基的转换生成语法而言，这种方法永远也无法赶上人的语言的变化。因此，这种方法算是比较笨的方法，类似于字典，我们可以从中查到有用的字词，甚至短语，但写出什么东西，却是字典无法实现的。因此这种方法有一定的实用性，但局限性也显而易见。

統計法

统计机器翻译：是目前非限定领域机器翻译中，性能较佳的一种方法。统计机器翻译的基本思想是通过对大量的平行语料进行统计分析，构建统计翻译模型，进而使用此模型进行翻译。从早期基于词的机器翻译已经过渡到基于短语的翻译，并正在融合句法信息，以进一步提高翻译的精确性。

统计机器翻译的首要任务是为语言的产生构造某种合理的统计模型，并在此统计模型基础上，定义要估计的模型参数，并设计参数估计算法。早期的基于词的统计机器翻译采用的是噪声信道模型，采用最大似然准则进行无监督训练，而近年来常用的基于短语的统计机器翻译则采用区分性训练方法，一般来说需要参考语料进行有监督训练。貝氏模型（Bayesian Model）也是一種機器翻譯方法。

近年來在語言服務產業掀起波瀾的神經機器翻譯就是利用巨大的人工神經網絡計算一連串字詞的機率以產生文意精確的翻譯。將在未來數年持續改變翻譯及語言在地化產業。在投入大量翻譯資料集（data sets）以訓練人工智慧和機器學習模組後，神經機器翻譯的品質已大幅改善。更重要的是，當神經機器翻譯與人工編修搭配，無論在技術還是文化層面，更能達到一流的譯文品質。因此，神經翻譯與人工編修搭配需求龐大。^[1]

参见

- [電腦輔助翻譯](#)

参考来源

1. 優譯堂 Ulatus. [MTPE再進化：神經機器翻譯](#). 2019-06-18 [2019-06-24]. （原始内容存档于2020-08-03） （中文（臺灣）） .

外部連結

- [The Advantages and Disadvantages of Machine Translation \(http://www.omniglot.com/language/articles/machinetranslation.htm\)](http://www.omniglot.com/language/articles/machinetranslation.htm) （[页面存档备份 \(https://web.archive.org/web/20201028053352/http://www.omniglot.com/language/articles/machinetranslation.htm\)](https://web.archive.org/web/20201028053352/http://www.omniglot.com/language/articles/machinetranslation.htm)，存于互联网档案馆）
- [International Association for Machine Translation \(IAMT\) \(http://www.eamt.org/iamt.php\)](http://www.eamt.org/iamt.php) （[页面存档备份 \(https://web.archive.org/web/20100624162302/http://www.eamt.org/iamt.php\)](https://web.archive.org/web/20100624162302/http://www.eamt.org/iamt.php)，存于互联网档案馆）

- [Machine Translation Archive \(http://www.mt-archive.info\)](http://www.mt-archive.info) ([页面存档备份 \(https://web.archive.org/web/20190401232615/http://www.mt-archive.info/\)](https://web.archive.org/web/20190401232615/http://www.mt-archive.info/)), 存于[互联网档案馆](#)) by [John Hutchins](#). An electronic repository (and bibliography) of articles, books and papers in the field of machine translation and computer-based translation technology
 - [Machine translation \(computer-based translation\) \(http://www.hutchinsweb.me.uk/\)](http://www.hutchinsweb.me.uk/) ([页面存档备份 \(https://web.archive.org/web/20080414052156/http://www.hutchinsweb.me.uk/\)](https://web.archive.org/web/20080414052156/http://www.hutchinsweb.me.uk/)), 存于[互联网档案馆](#)) – Publications by John Hutchins (includes [PDFs](#) of several books on machine translation)
 - [Machine Translation and Minority Languages \(https://web.archive.org/web/20080120192259/http://bowland-files.lancs.ac.uk/monkey/ihe/mille/paper2.htm\)](https://web.archive.org/web/20080120192259/http://bowland-files.lancs.ac.uk/monkey/ihe/mille/paper2.htm)
 - [John Hutchins 1999 \(http://www.foreignword.com/Technology/art/Hutchins/hutchins99.htm\)](http://www.foreignword.com/Technology/art/Hutchins/hutchins99.htm) ([页面存档备份 \(https://web.archive.org/web/20070907182416/http://www.foreignword.com/Technology/art/Hutchins/hutchins99.htm\)](https://web.archive.org/web/20070907182416/http://www.foreignword.com/Technology/art/Hutchins/hutchins99.htm)), 存于[互联网档案馆](#))
-

取自“<https://zh.wikipedia.org/w/index.php?title=机器翻译&oldid=72430248>”

本页面最后修订于2022年6月30日 (星期四) 03:46。

本站的全部文字在知识共享 署名-相同方式共享 3.0协议之条款下提供，附加条款亦可能应用。（请参阅使用条款）
Wikipedia®和维基百科标志是维基媒体基金会的注册商标；维基™是维基媒体基金会的商标。
维基媒体基金会是按美国国内稅收法501(c)(3)登记的非营利慈善机构。