

任何两个字母如果不一样，都算错一个，但是我们没有考虑读音的相似性，在键盘上字母位置的接近程度等等。比如在键盘上j和h都是用右手食指打出来的，可能会打错，因此shangjai很可能是上海。但是shangbai可能不是上海，因为b和h是用不同的手打字的。

4. 对于本身比较长的单词，我们应该允许编辑距离差得比较大，但是这样一来，计算量就会增加，如何保证纠错覆盖率和使用资源的多少，就是工程的艺术了。

最后，总结一下今天的内容：

1. 很多人问我什么是一通百通，对于一种知识和技能的熟练运用，就是一通百通。假如我们学习了维特比算法，那么不仅通信的问题会解决，计算机软件中设计拼写纠错的程序也会写了，就是灵活运用，一通百通。从某种程度上讲，拼写纠错和通信的纠缠是等价问题。

2. 科学和工程还是有区别的，从今天的例子中你应该有更深的体会。在科学上，采用一个好的和不好的方法，差距会非常大。在工程上，很多实际问题不考虑清楚，做出来的东西就不好用。

3. 很多时候，改进是无止尽的，有的人做事情只做了50%-60%，就觉得全做好了，就满足了，其实如果要往深里做，就会发现还有很多问题有待解决。编写一个拼写纠错程序，看似简单，如果要把各种问题讲清楚，只要有足够的篇幅，我可能能讲整整一个星期。精品和粗制滥造的产品差别就在于此。

希望这周的内容对你有所启发。

祝近安

吴军

2018年8月15日

吴军的谷歌方法论

一份智能时代的行动指南

版权归得到App所有，未经许可不得转载

