标准化试卷标记语言

(写另外一篇博文的时候发现跑题太多,所以现在把它独立出来另外写一篇。)

我本科的时候给我爸设计了一种"标准化试卷标记语言"(他是中学英语老师)。当时我写了一个1000来行的 Perl 脚本,可以把这种简单的标记语言转换成美观的 LaTeX 格式文档,并且带有友好的 Tk 图形界面。题目可以包括选择题,填空题,改错题,…… 这种语言的特点是,题目和答案都放在一起,所以出题的时候很直观。经过处理之后,答案与题目分离,并且被放到答案表里正确的位置上。这样出题的人就不会把答案放错位置,也不用担心如何排版。

比如,选择题的格式是这样:

#	0ur	teacher	t	old	us	that	there		no	end	to	learning
	Α.	was B. is *		*	C. has			D. had				

每道题开头都是一个 #, 正确答案后面加一个 *。在处理的时候这些 # 都会被变成题目编号,而有 * 标记的答案选项,会被放到答案表里面。

最有意思的是改错题。因为改错题是一个英语段落,某些行有错,但每行最多只能有一个错。所以我的设计是,在普通的段落里插入这样的记号:

This is an |extraordinary|extrordinary| sentence...

用以"引入错误"。左边是正确的方式,右边是错误的。这里使用 | 是因为这个符号不会在普通的英语文章里出现。| 的左右两边都可以是空白,用以表示"插入"与"删除"。比如:

I don't know how to play |the|| piano.

这样的题目显示出来之后是这样:

I don't know how to play piano.

然后答案里会显示:

在 play 和 piano 之间加 the

如果两个都不是空白,那就是"修改",就像上面的例子。最后我会根据这些标记的位置把段落排版,让每行上面最多只有一个错。为了让段落的行看起来均匀,我使用了一种类似 TeX 的动态规划断行算法。它先算出多种断行方案的"难看程度",然后从中选出最好看的一个。

现在回想起来,我那时候的设计其实是相当先进的。怎么就没想过把它做成产品卖给教育部呢,也许因为觉得这种技术会制造出更大量的题海,祸害更多的中学生:-P 跟我的语言相比,现在一些 blog 系统用的 markup 语言(比如markdown)真是小巫见大巫了。