因卓教育答题卡生成和识别方案

作者: 张文虎

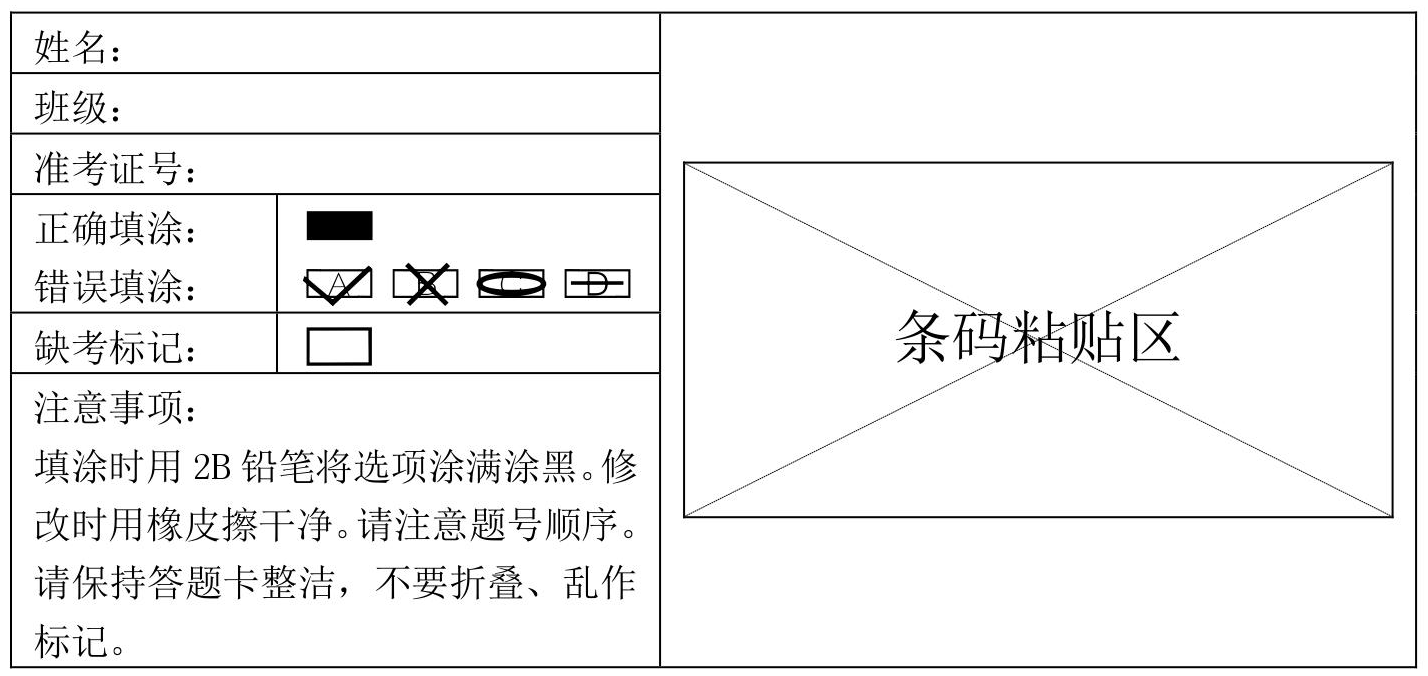
# **答题卡生成**

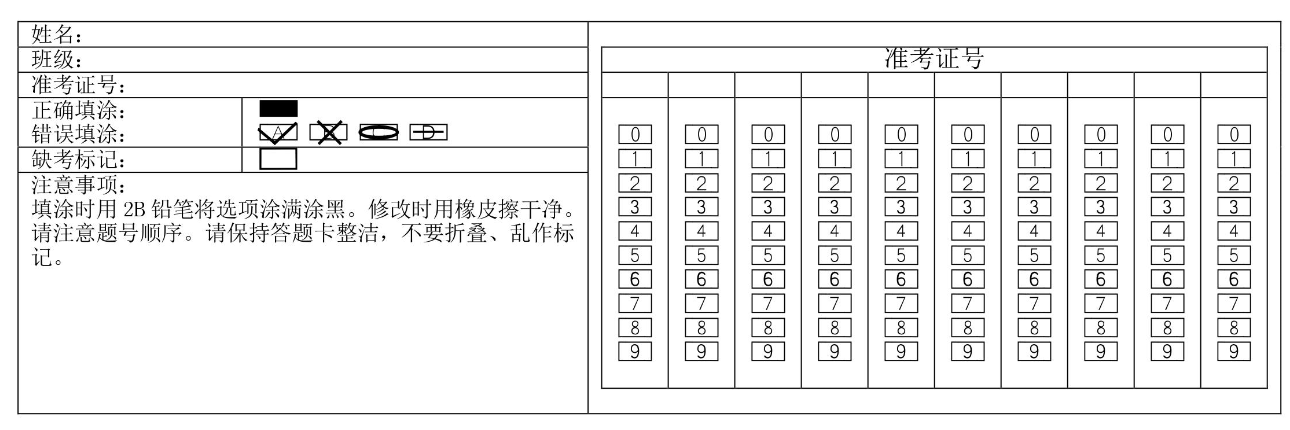
## 答题卡信息

**答题卡大小**：A4一栏、A3二栏、A3三栏

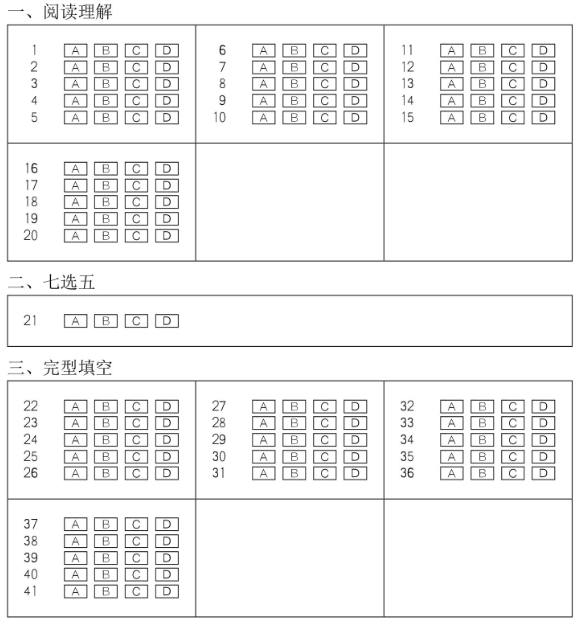
**答题卡组成元素：**

1. 答题卡标题，比如：xxx答题卡
2. 答题卡头，分条形码和填涂两种样式

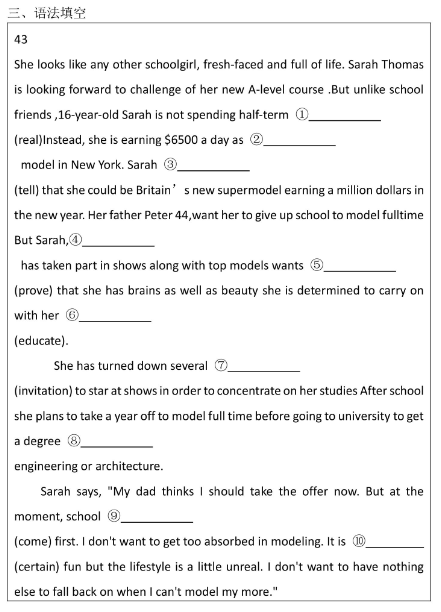




1. 客观题

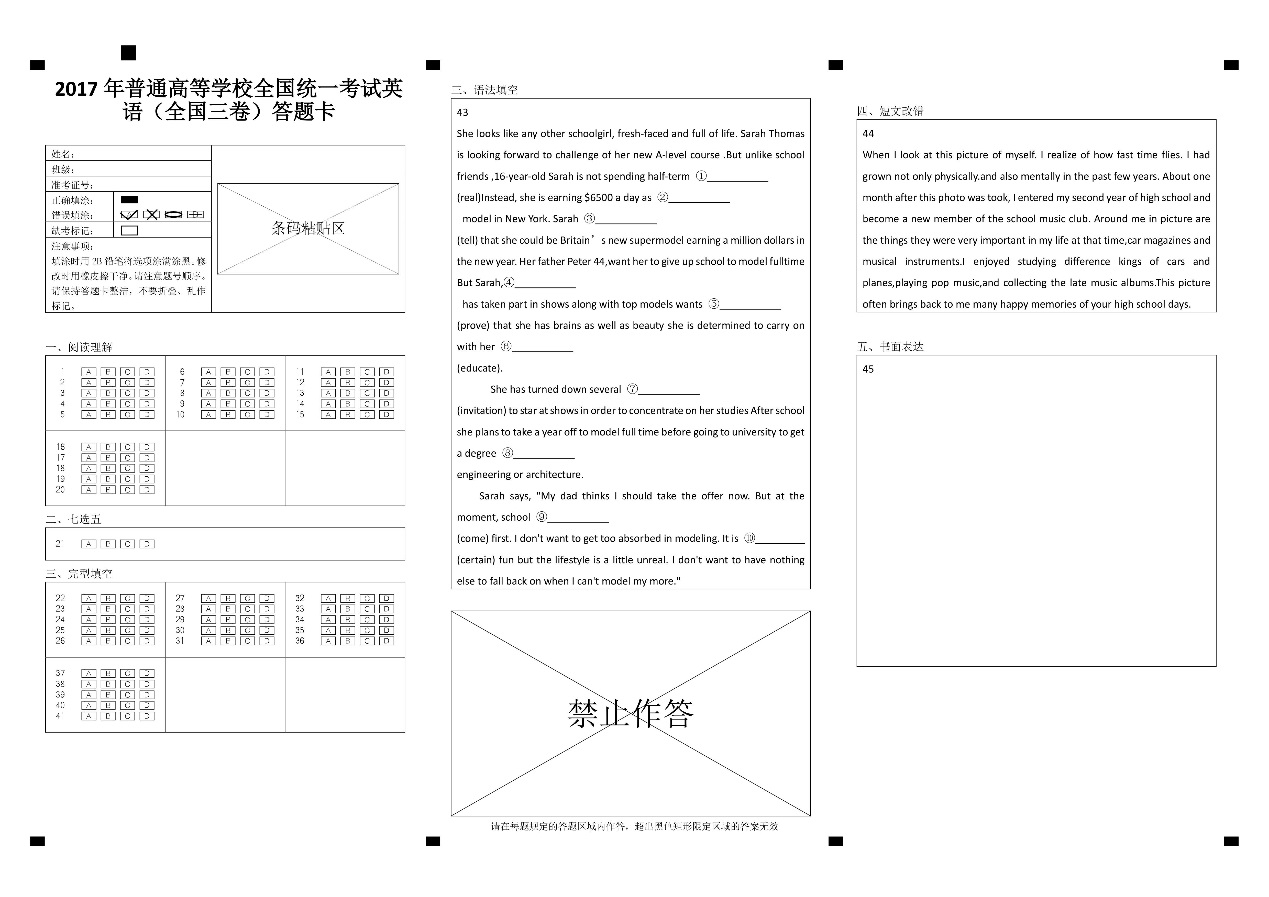


1. 主观题





1. 答题卡定位符，用于定位答题卡每一栏的区域，以及纠正放反的答题卡



## 如何生成答题卡

* 答题卡标题使用对应试卷的名称



RE：答题卡标题为两行，首行是**考试名称**（不是~~试卷名称~~，便于统一）；第二行是“科目+答题卡”。

例：2017年高三期末考试，语文科目的答题卡标题如下

2017年高三期末考试

语文答题卡

**确认：**如果考试里面同时存在多个年级的话，同科目的可能会存在多张试卷，比如高一语文；高二语文。这种情况怎么显示？

* 答题卡卡头使用固定的两种样式：条形码、填涂
* 答题卡答题区域根据题型生成，配置答题卡的试卷需要遵循以下原则
  + 同一题型内不允许同时出现主观题和客观题

RE：允许出现。

如语文的阅读理解题型，会同时存在选择、填空、解答。但当一个大题题型内同时存在主客观题的小问时，执行切图后人工批改，而不是机器阅卷。

**确认：**

* + 如果同一题型都是客观题，客观题必须属于统一类型，比如：都是选择题，并且选项数量一致。  
    RE：允许不统一类型（单多选），且选项数量也不一致（上限为4个选项）。
  + 小题内的小问不允许同时存在客观题和主观题

RE:允许同时存在。但当一个大题题型内同时存在主客观题的小问时，执行切图后人工批改，而不是机器阅卷。

* + 整个试卷的布局，要求先是客观题然后是主观题

RE：OK。出现在客观题答题区后的，统一以主观题形式批改。

* 试题的答题区域不允许跨栏，因此同一栏剩余的区域不足以容纳一题的答题区的时候，需要设置为禁止答题区。

RE：允许跨栏，相当于一个题目引用2个答题框切图。（是否能自动匹配，如果不能允许人工设置映射关系）

* 所有的答题区域都必须要使用线框框起来，用于答题卡识别
* 使用定位符来定位栏
* 答题卡在后台使用HTML格式存储，下载的时候直接下载为图片

**举例说明：**

**试卷：**

2017年普通高等学校招生全国统一考试数学(理)

**试卷结构：**

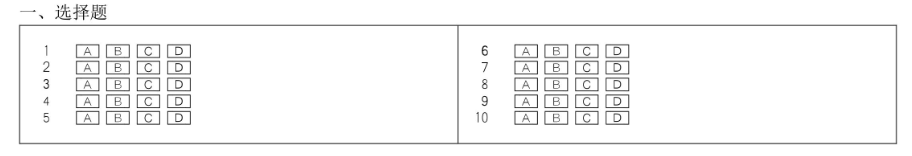
1. 选择题（共10题）
2. 填空题（共5题）
3. 解答题（共6题）

**生成步骤：**

* 检查题型下面所有的小问是否都为同一类型，如果不是同一类型，报错，如果是，进入下一步。

RE：允许存在，根据试题结构引用而定。

* 如果所有小问都为选择题，根据小问数量生成客观题答题区域，这里为10



* 如果小题内所有小问都是主观题，针对每一个小题生成一个答题区域。

RE：如果小题内，存在小问是主观题，则针对该小题生成一个答题区域。

例：语文阅读理解中，第一小问单选，第二小问解答，则生成这个小题的答题区域，此时第1小问的选择也通过切图后人工批改。



* 所有题型的答题区域生成完后，用户根据实际答题区所需要的大小去调整答题区域。

RE：同时可对答题卡进行编辑，如插入文字、图片、下划线。

# 答题卡识别

**输入**

* 考试信息
* 答题卡对应的试卷的试题结构
* 答题卡扫描图片

**识别步骤**

* 使用图像识别算法根据定位点识别出答题卡有效区域
* 使用图像识别算法检测出所有的矩形区域，是否需要增加更多的定位点来帮助识别矩形区域，需要后续开发人员确认
* 答题卡卡头：识别出学生的考号，如果考号无法识别需要记录异常。
* 客观题答题区域：根据试题结构和识别出的矩形框的顺序来确认是否为客观题答题区域，如果是的话，识别出学生客观题的答题，并上报给后台服务器进行学生答案的存储。
* 主观题答题区域：根据试题结构和识别出的矩形框的顺序来确认是否为客观题答题区域，如果是的话，切割并上传到后台服务器进行学生答案的存储。
* 根据试题结构和识别出来的矩形框的顺序来确定答题区域和试卷中试题的对应关系

**异常处理**

* 考号无法识别
* 考号重复
* 白卷
* 答题卡定位符缺失

**特殊说明**

* 允许扫描答题卡张数超过考试人数。（避免报错）
* 允许处理异常结束后保存的成绩条目数超过考试人数。
* 处理异常时，读取的学生、班级数据是发布考试时的学生、班级数据。
* 答题卡上统一按照**小问**进行编号；切图图框按照每个**小题**进行切图。