****

****

**信息学院软件工程系**

**《JAVA程序设计》实验报告**

实验9

**姓名：宋泽涛**

**学号：25120222201292**

**学院：信息学院**

**专业：软件工程**

**完成时间：2024/4/30**

**一、实验目的及要求**

* 熟悉字符串及正则表达式

**二、实验题目及实现过程**

题目1：

编写程序完成：

1. 输出以下新闻片段中出现的单词（每个单词只输出一次）。
2. 输出以下新闻片段中包含*the*的句子。

PITTSBURGH (AP) — Carnegie Mellon University will hire a researcher from the Library of Congress to help it decode a collection that includes two WWII German Enigma machines.

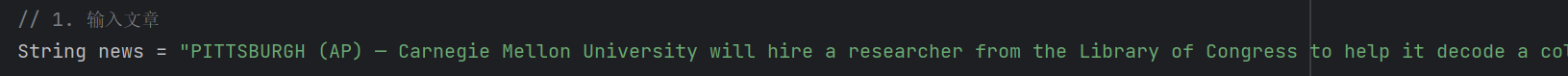
The university wants to encourage the study of 19th and 20th century computers, calculators, encryption machines and other materials related to the history of computer science.

“When we look back and we see this, we see who we remember,” Andrew Moore, dean of CMU’s School of Computer Science, said, adding his students are increasingly asking for courses about the history of the field. “We see people who took technology to save lives and save the world.”

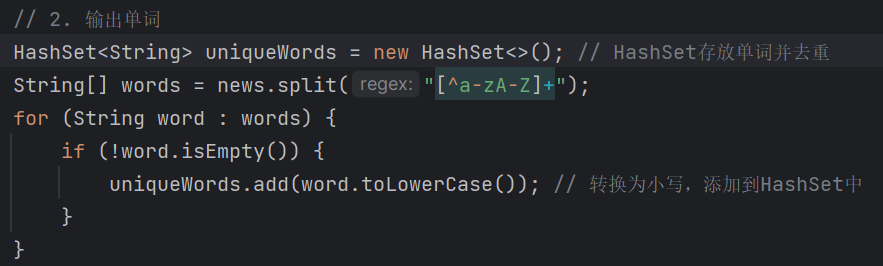
Pamela McCorduck, a prolific author on the history and future of artificial intelligence and the widow of Joseph Traub, a renowned computer scientist and the former head of CMU’s Computer Science Department, permanently loaned to the university a collection of early computers, books and letters. The collection, anchored by a three-rotor and four-rotor Enigma machine, is on display in the Fine and Rare Book Room in CMU’s Hunt Library in Oakland. The gift makes CMU one of a few institutions in the United States with Enigma machines. Even fewer display them.

**实现过程**

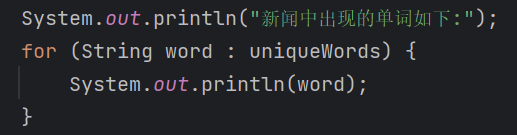
1. **首先创建字符串对象存放待读入的文章**



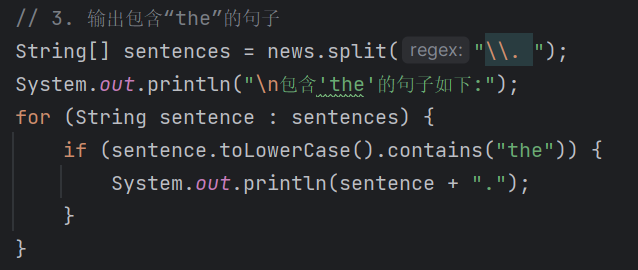
1. **为完成第一个统计单词数目的需求，我们可以创建一个HashSet来存放单词，并用其性质进行去重。至于如何判断一段字符串是单词，可以用正则表达式进行判断：”[^a-zA-Z]+”作为分割标志，表示用任何非字母字符作为分隔符，将字符串news分隔成一个单词数组**



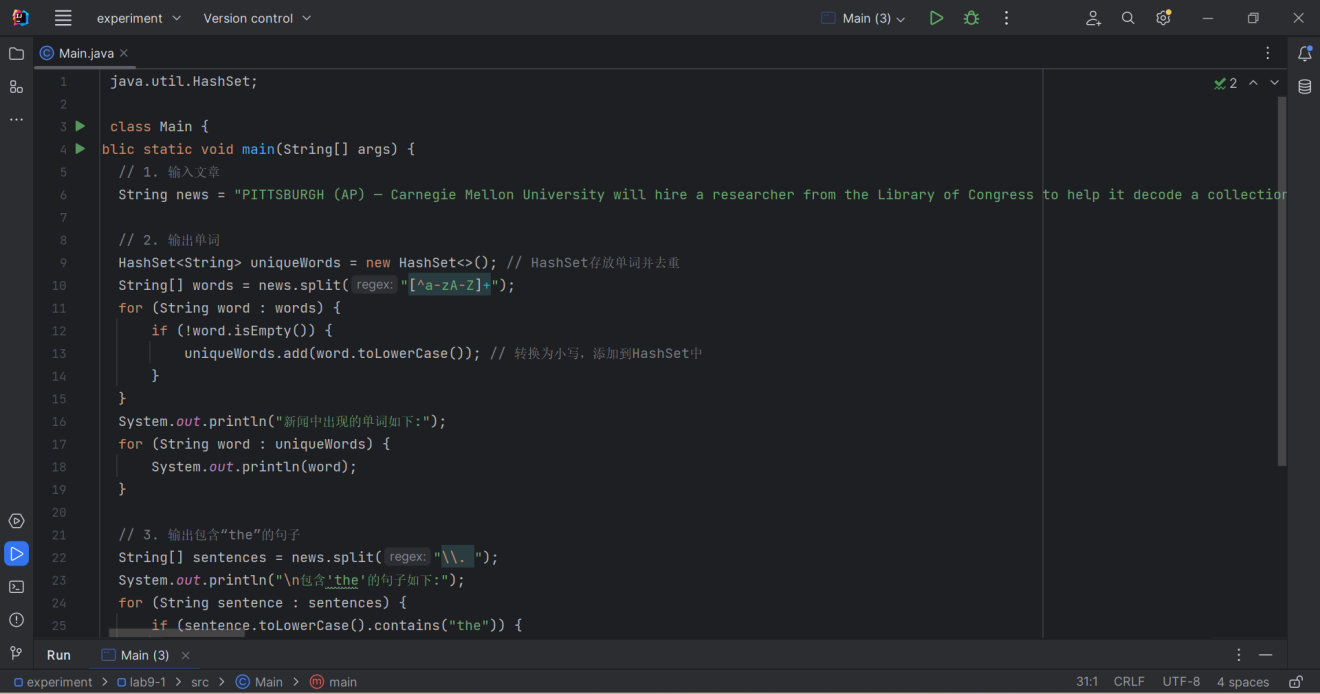
1. **遍历数组输出单词**

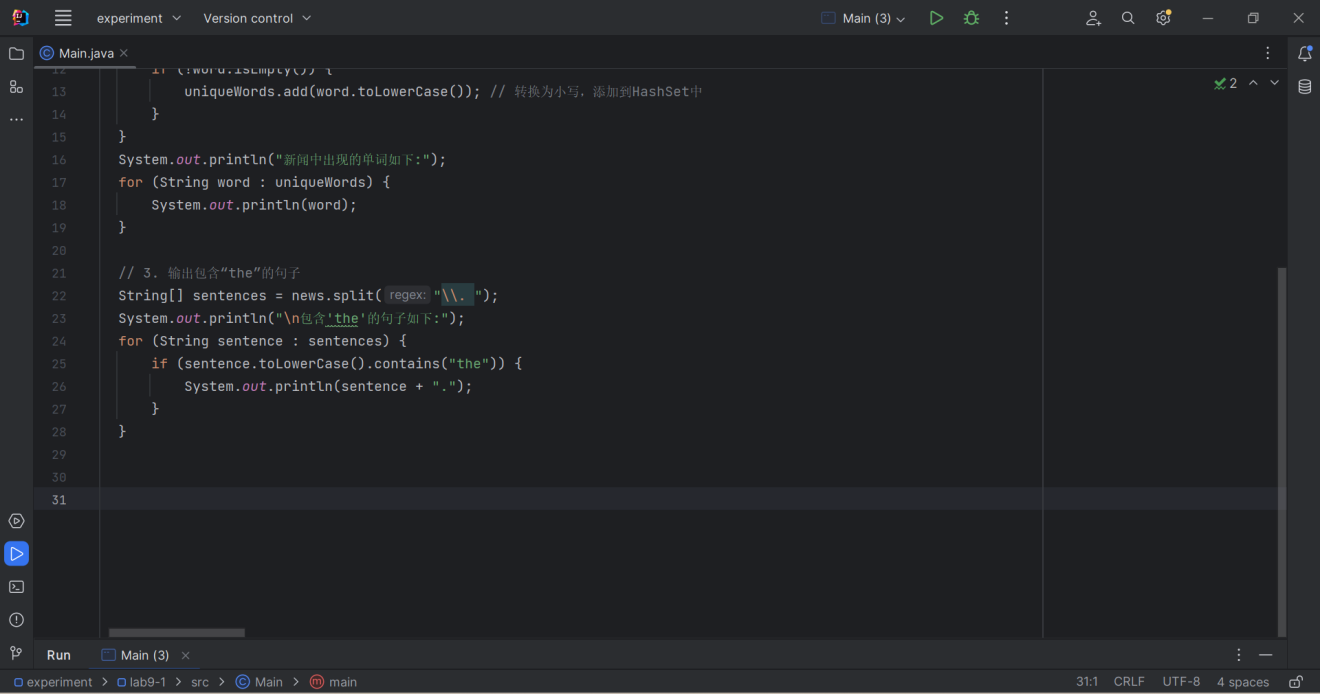


1. **统计句子时，我们需要以句号为分隔符，所以需要通过正则表达式实现。而句号字符’.’在正则表达式中有特殊的含义（匹配除了换行符’\n’之外的任何单个字符），所以需要反斜杠进行转义**

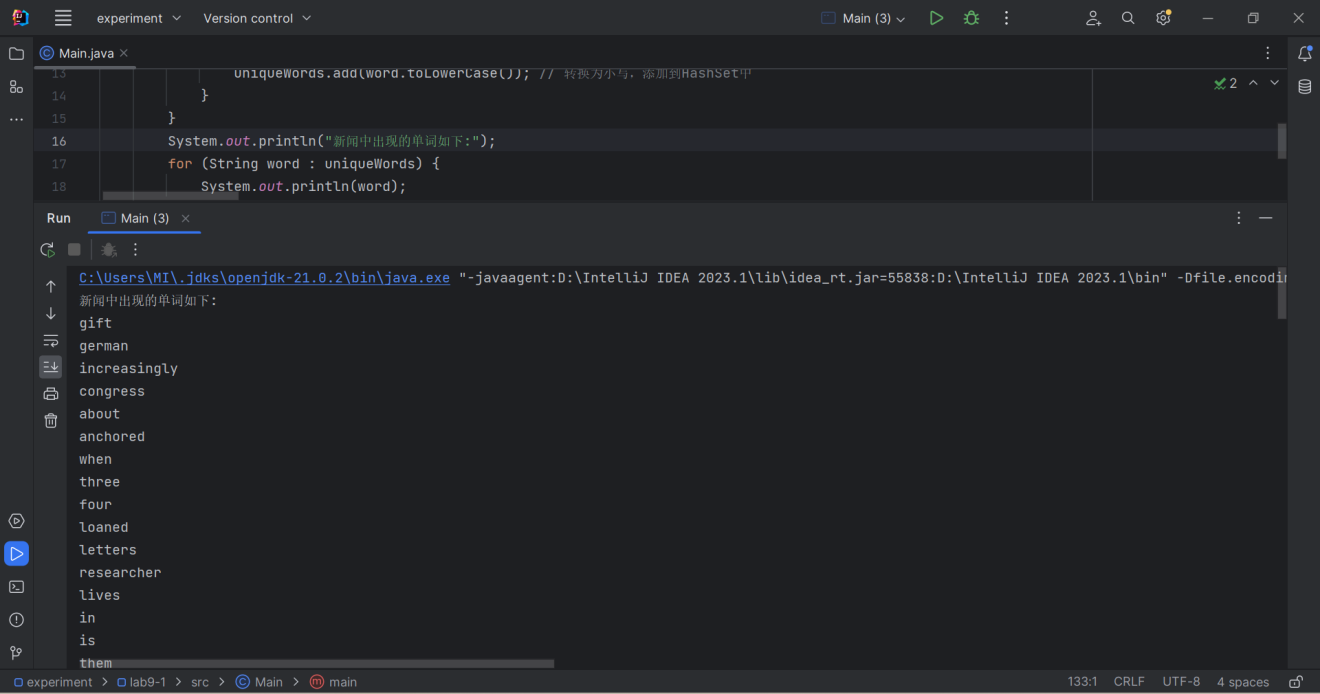


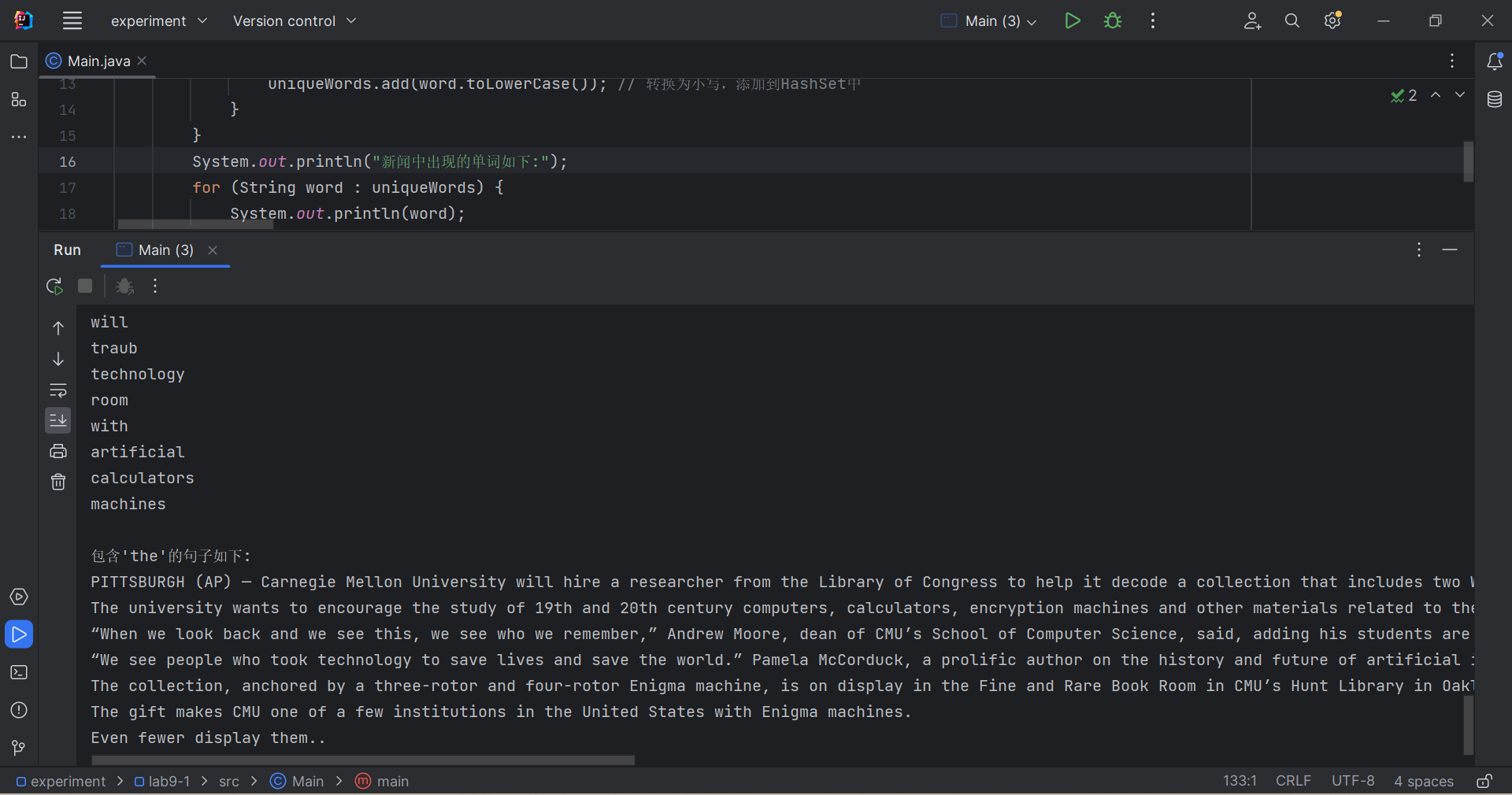
**程序截图**





**运行结果**

****

****

题目2：

用正则表达式对用户输入的用户名、密码、邮箱进行判断，若不满足输入要求则提示出错类型。

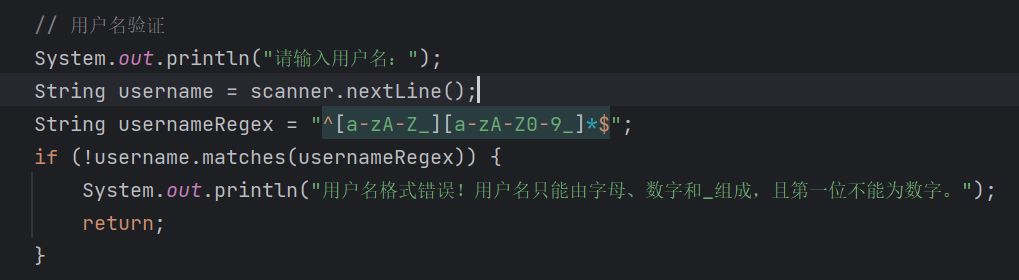
用户名要求：不能为空，只能由字母、数字和\_组成，第一位不能为数字。

密码要求：不能为空，密码长度至少8位，由字母、数字、下划线组成。

邮箱要求：不能为空，需包含”@”符号。”@”符号后需要出现多个由”. ”分割的词。

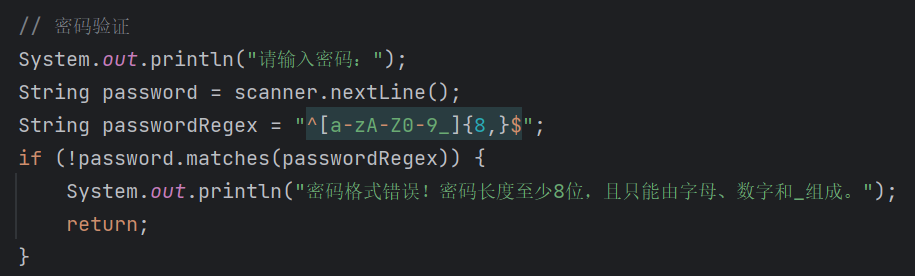
**实现过程**

**1.首先提示用户输入用户名，等待输入后对用户的输入用正则表达式进行判断**



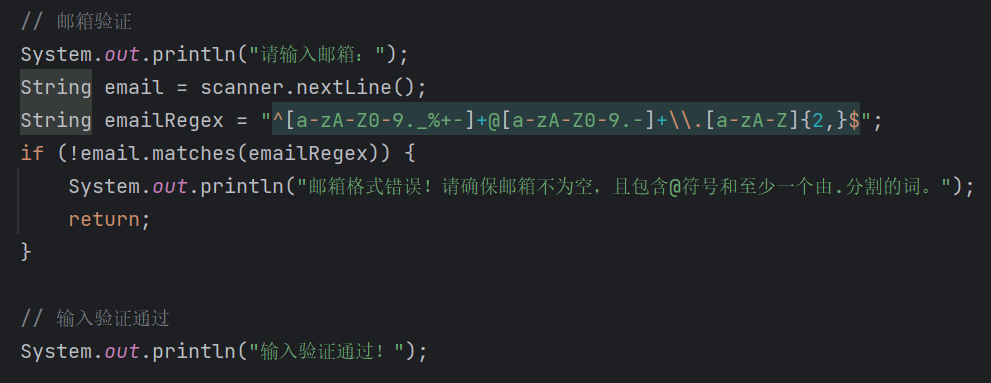
**对于’^[a-zA-Z\_][a-zA-Z0-9\_]\*$’这个正则表达式，’^’表示匹配字符串开始的位置；’[a-zA-Z\_]’是一个字符类，用来匹配大小写字母和下划线中的任何一个字符；’[a-zA-Z0-9\_]\*’这是一个字符类后跟一个量词，匹配零个或多个大小写字母、数字或下划线字符；’$’表示字符串结束的位置**

**2.若用户名合法，则提示用户输入密码，否则提示用户输入的用户名不合法**



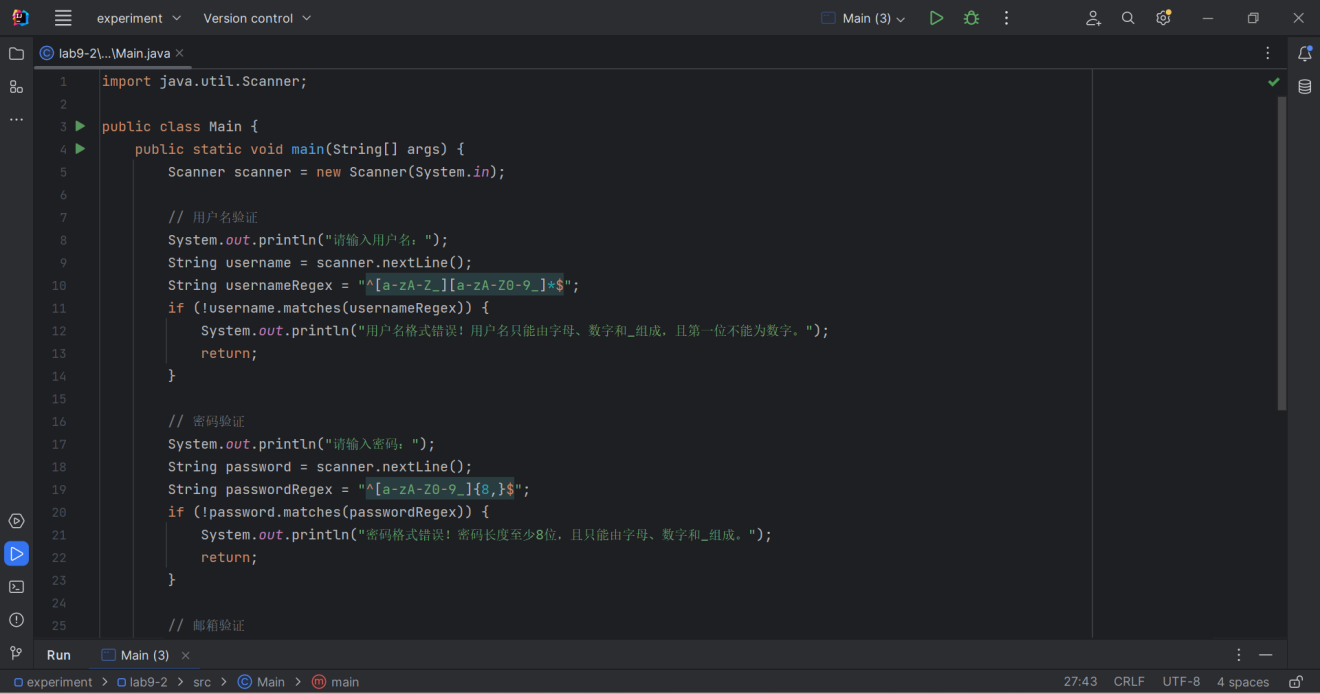
**同样的’[a-zA-Z0-9\_]{8,}$’这个正则表达式，’{8,}’是一个量词，表示前面的字符类’[a-zA-Z0-9\_]’至少要匹配8次，实现“密码长度至少8位”这个需求。**

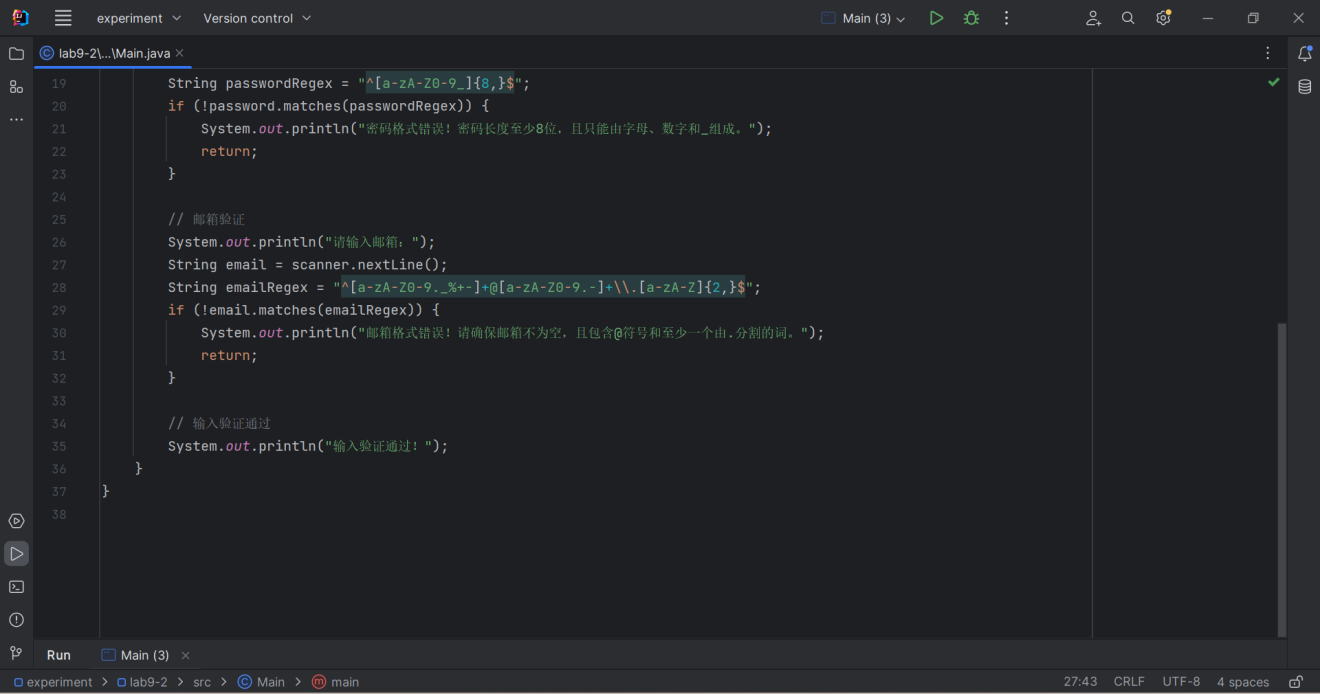
**3.若用户名和密码均合法，则提示用户输入邮箱，否则提示用户输入不合法**



**[这个正则表达式’[a-zA-Z0-9.\_%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}$’中，’[a-zA-Z0-9.\_%+-]’匹配至少一个大小写字母、数字、点号、下划线、百分号、加号或减号；’@’匹配’@’字符；’[a-zA-Z0-9.-]+’匹配至少一个大小写字母、数字、点号或减号；’\\.’匹配实际的点号字符；’[a-zA-Z]{2,}’匹配至少两个大小写字母。](mailto:这个正则表达式’[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+/.[a-zA-Z]{2,}$’中，’[a-zA-Z0-9._%+-]’匹配至少一个大小写字母、数字、点号、下划线、百分号、加号或减号；’@’匹配’@’字符；’[a-zA-Z0-9.-]+’匹配至少一个大小写字母、数字、点号或减号；’/.’匹配实际的点号字符；’[a-zA-Z]{2,}’匹配至少两个大小写字母。)**

**程序截图**

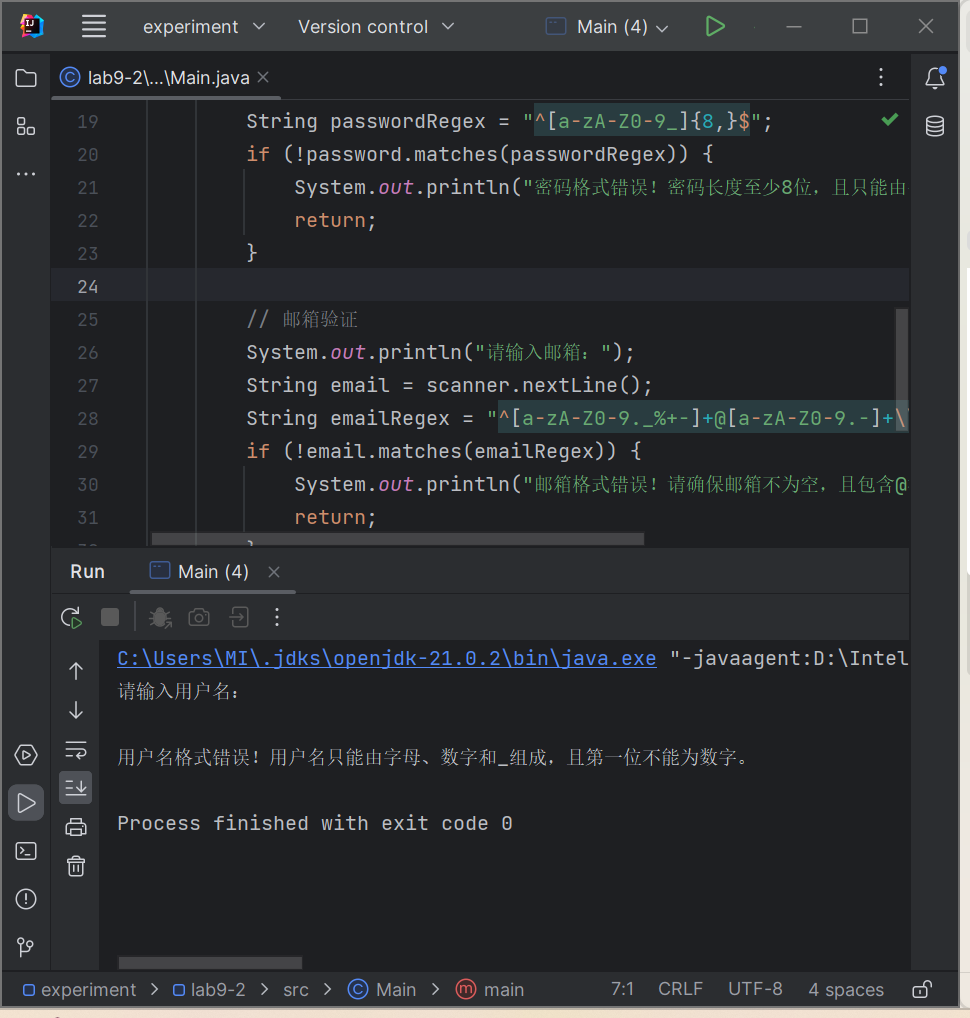
****

****

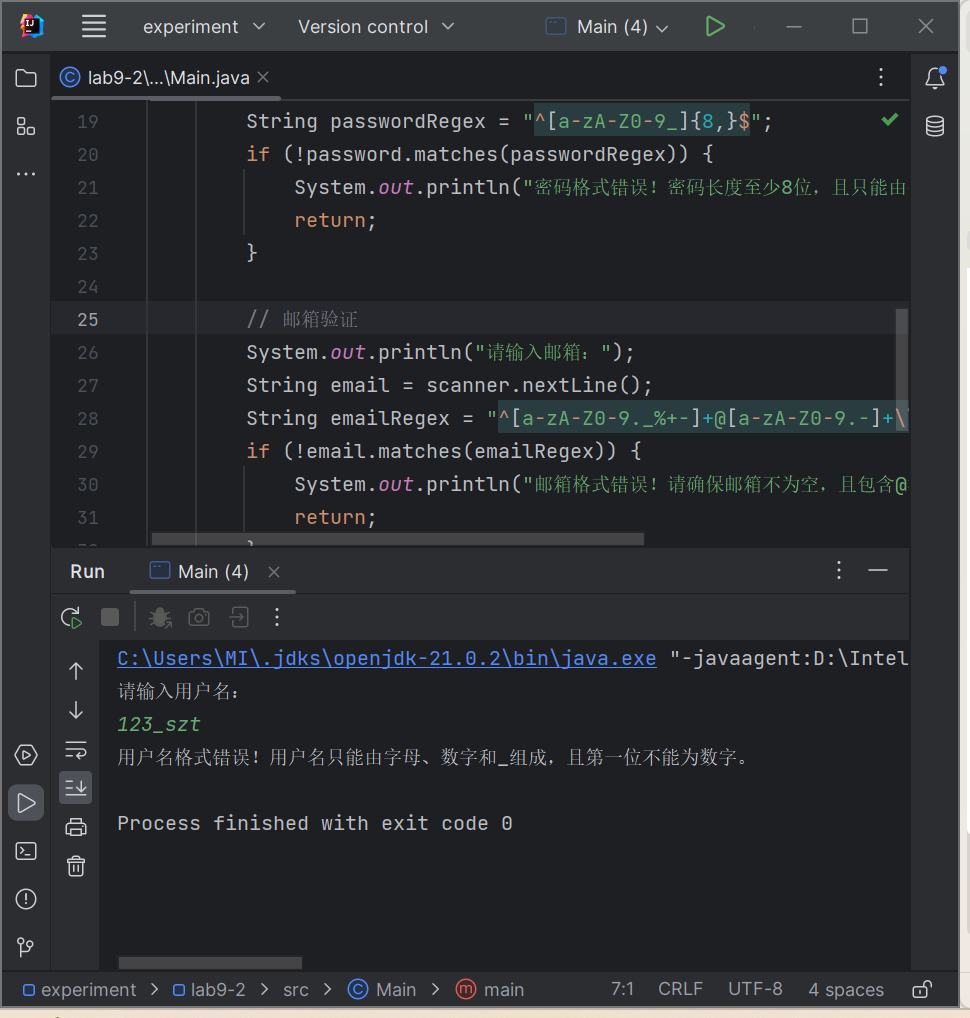
**执行结果截图**

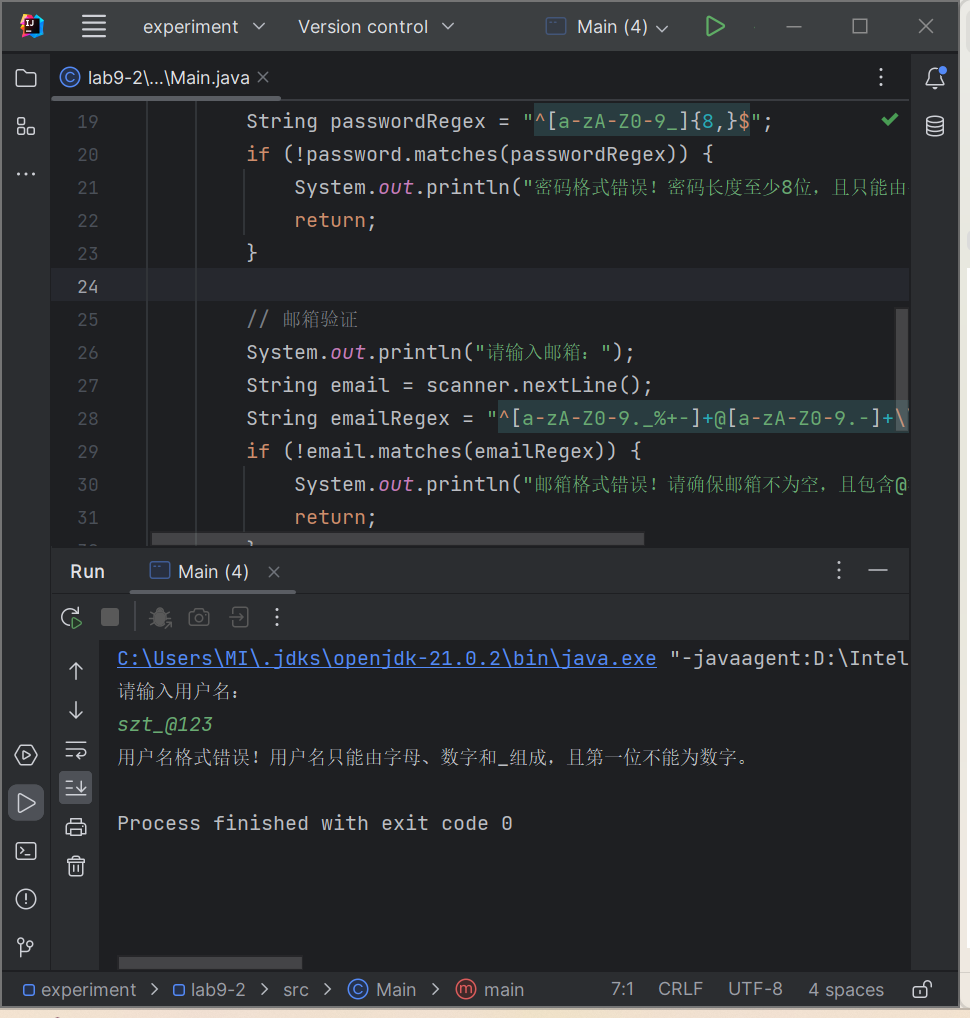
**一、用户名格式验证**

**1.空用户名报错**

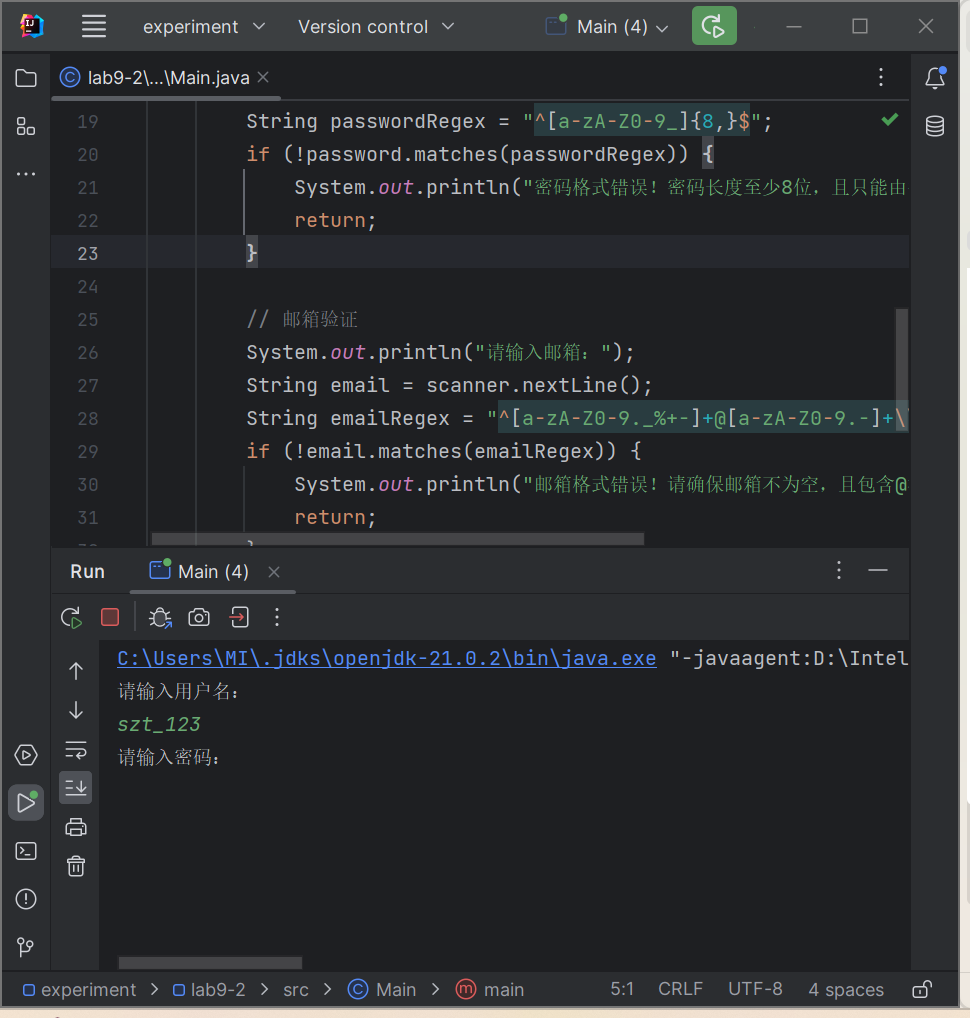


**2.用户名格式错误**



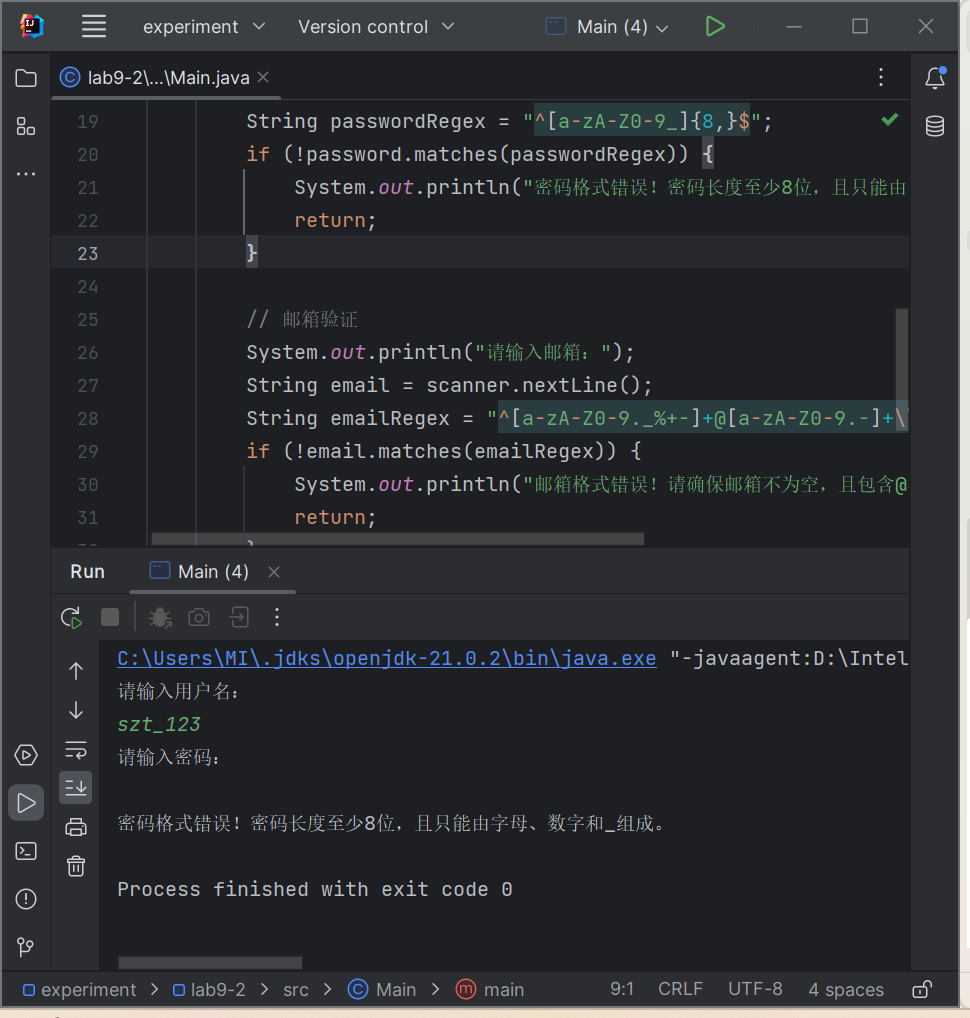


**3.用户名正确**

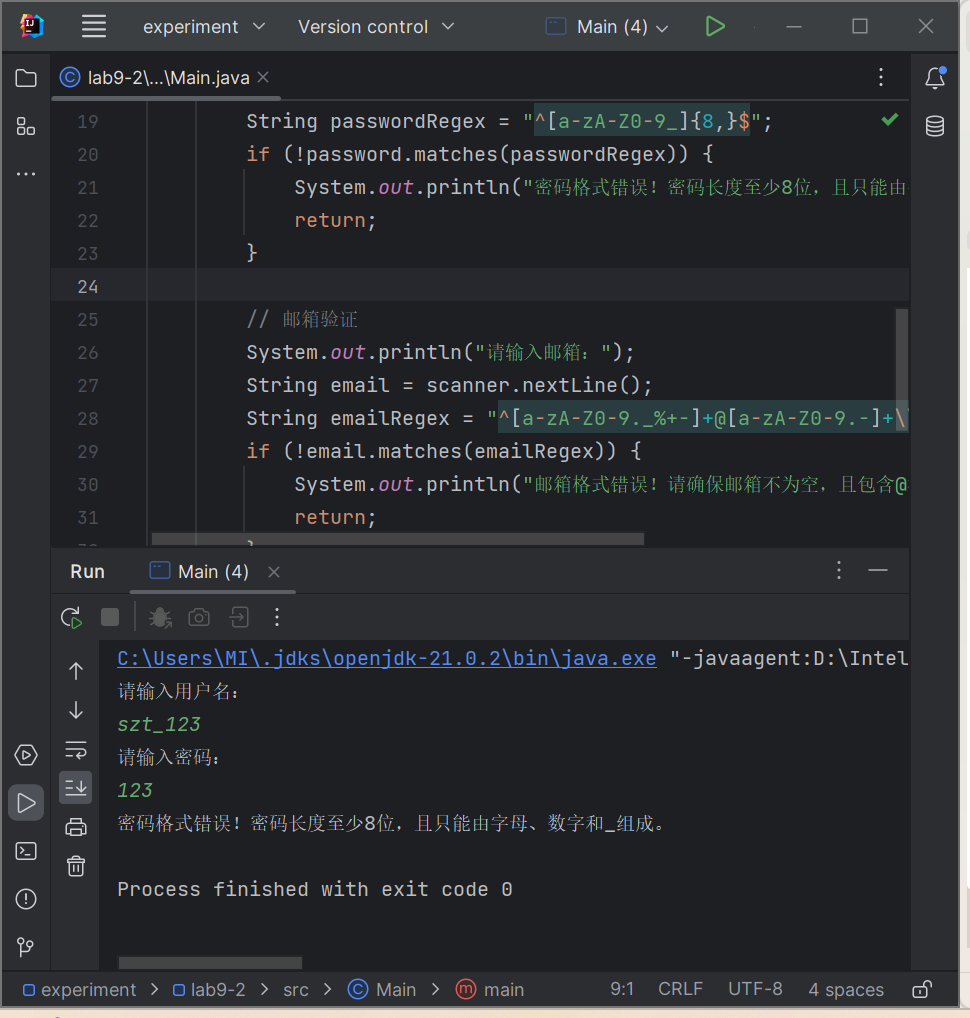


**二、密码格式验证**

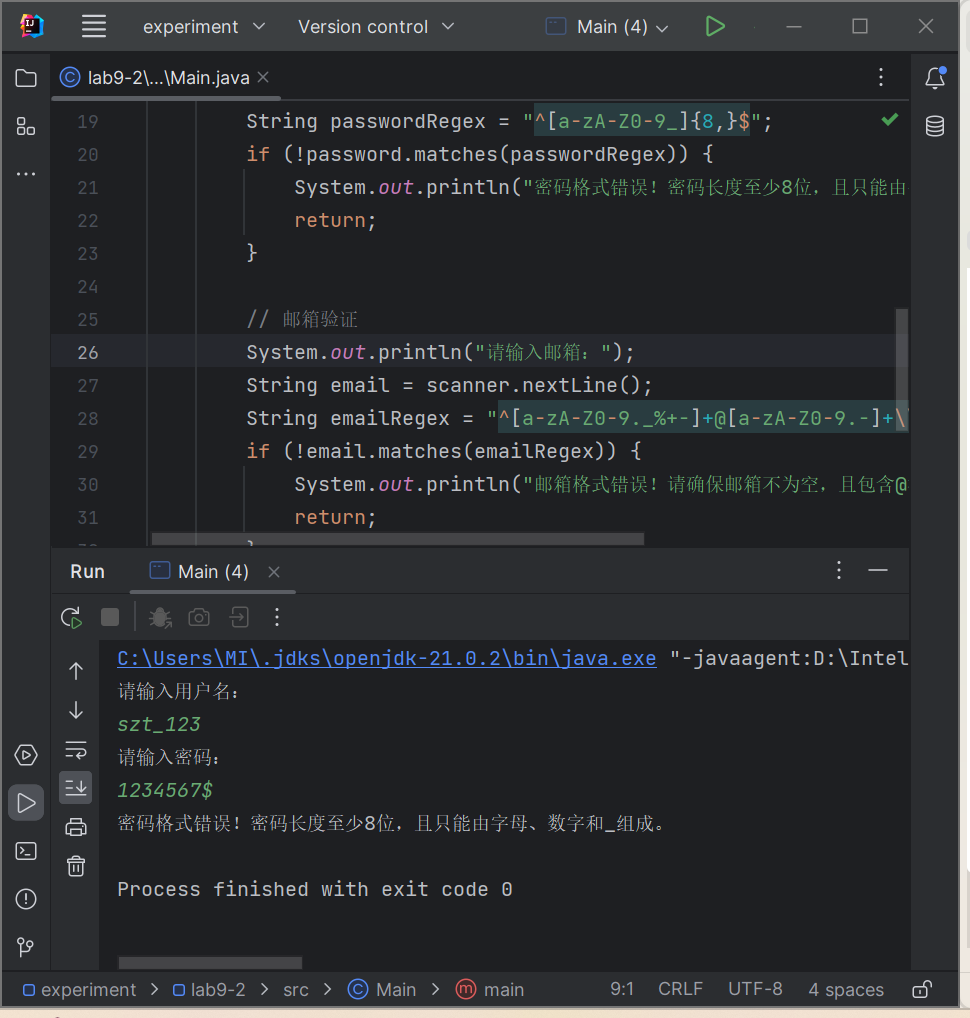
**1.空密码报错**



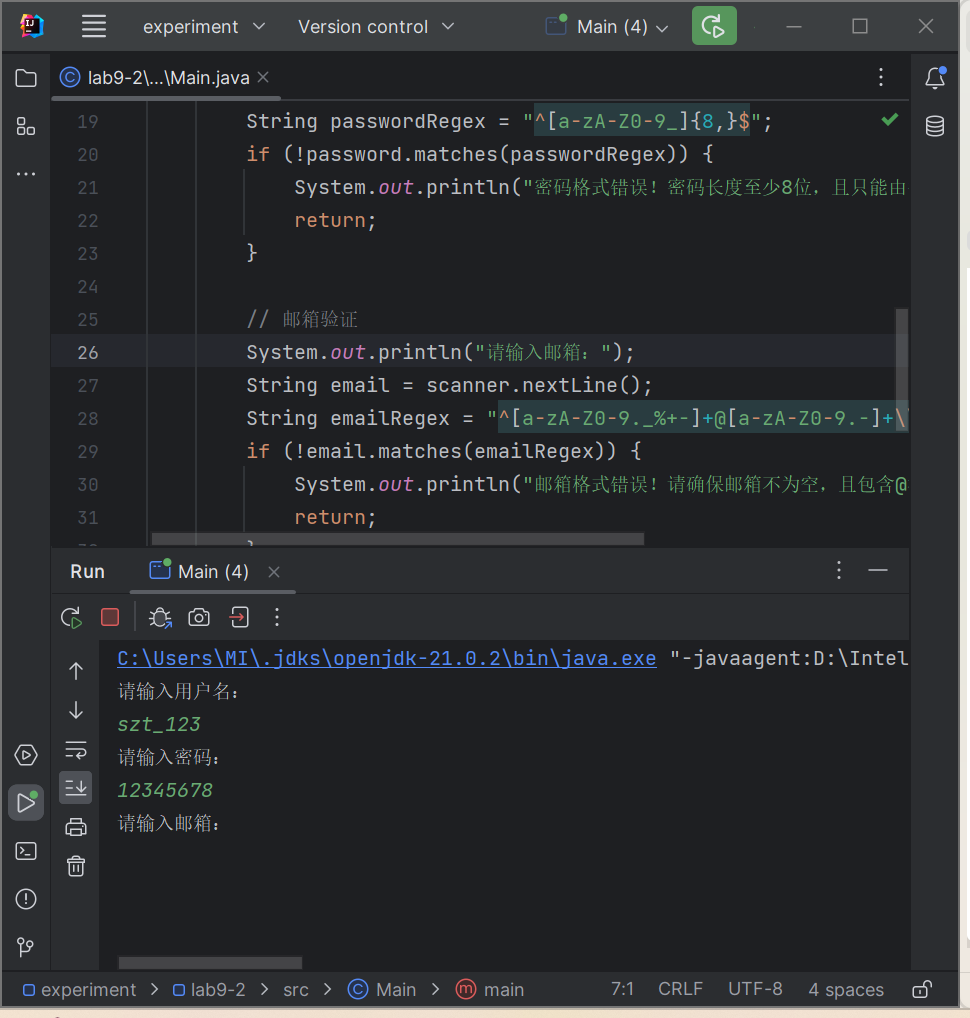
**2.密码长度不够8位**



**3.密码中有其他字符**

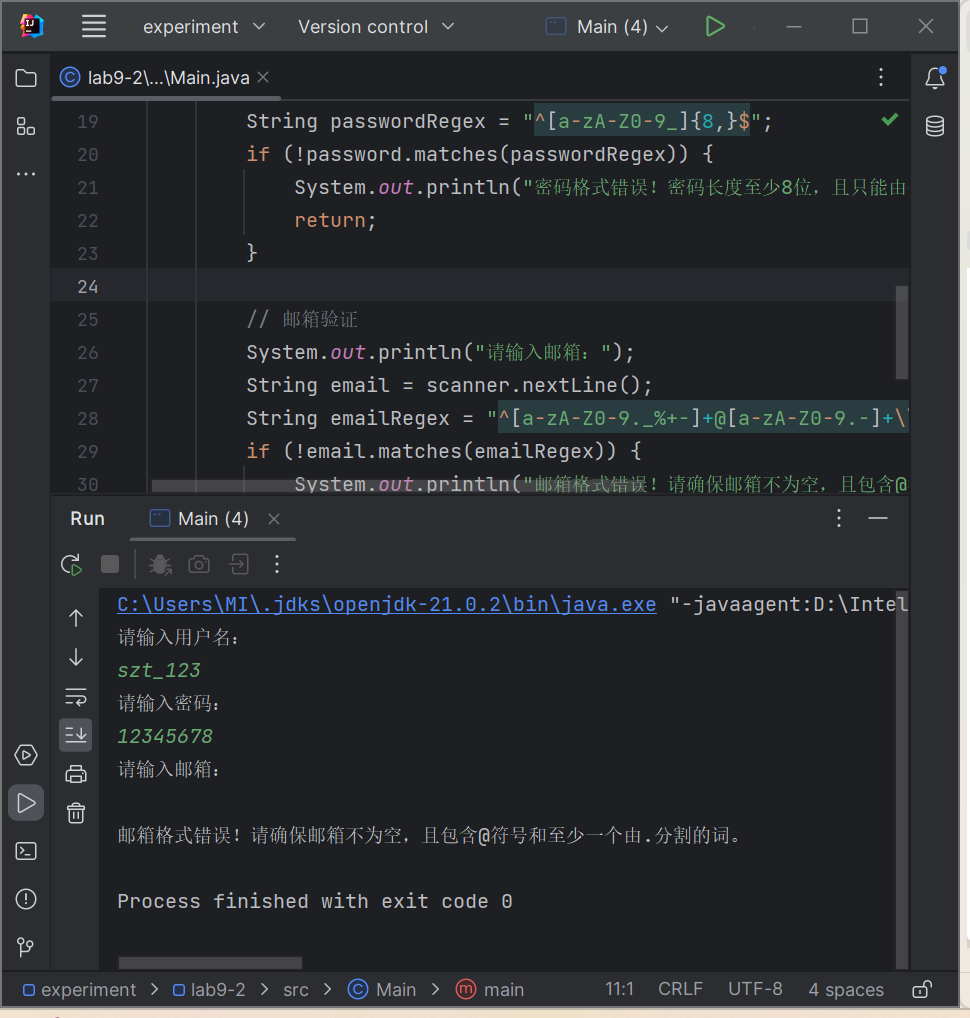


**4.密码格式正确**

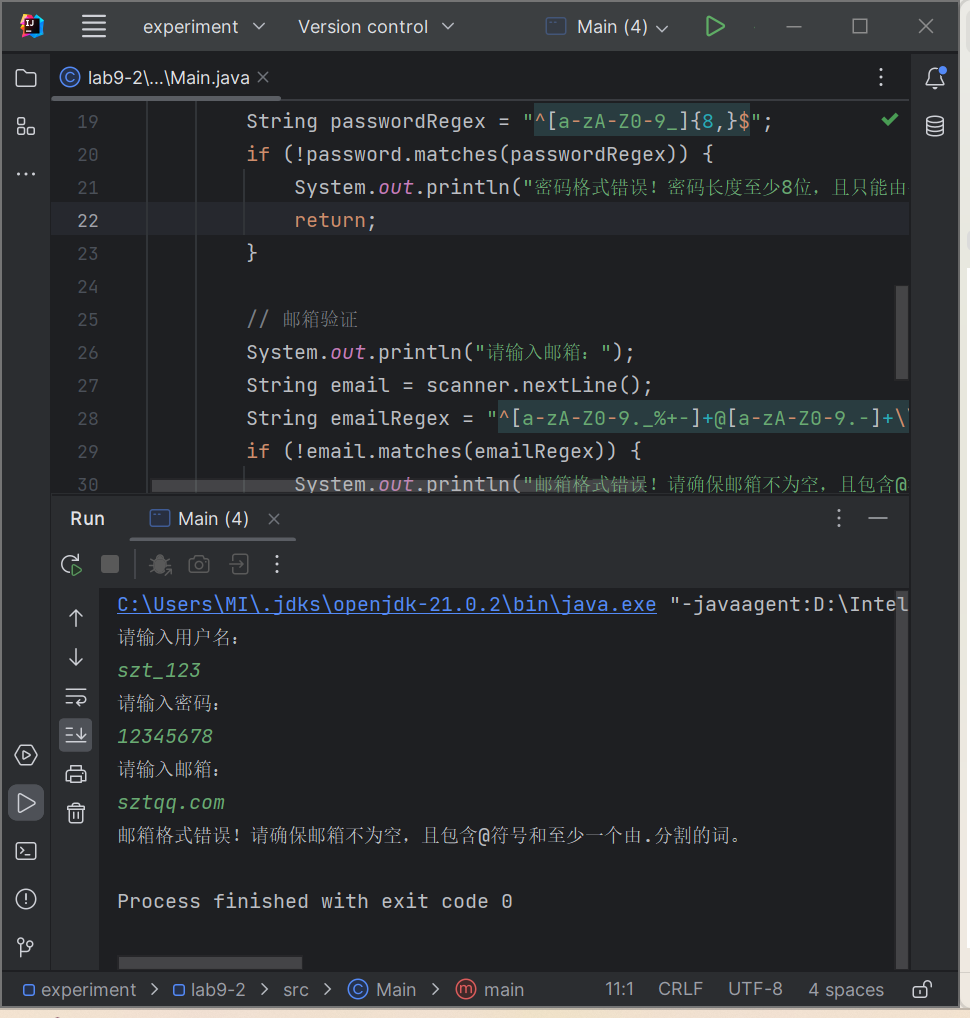


**三、邮箱格式验证**

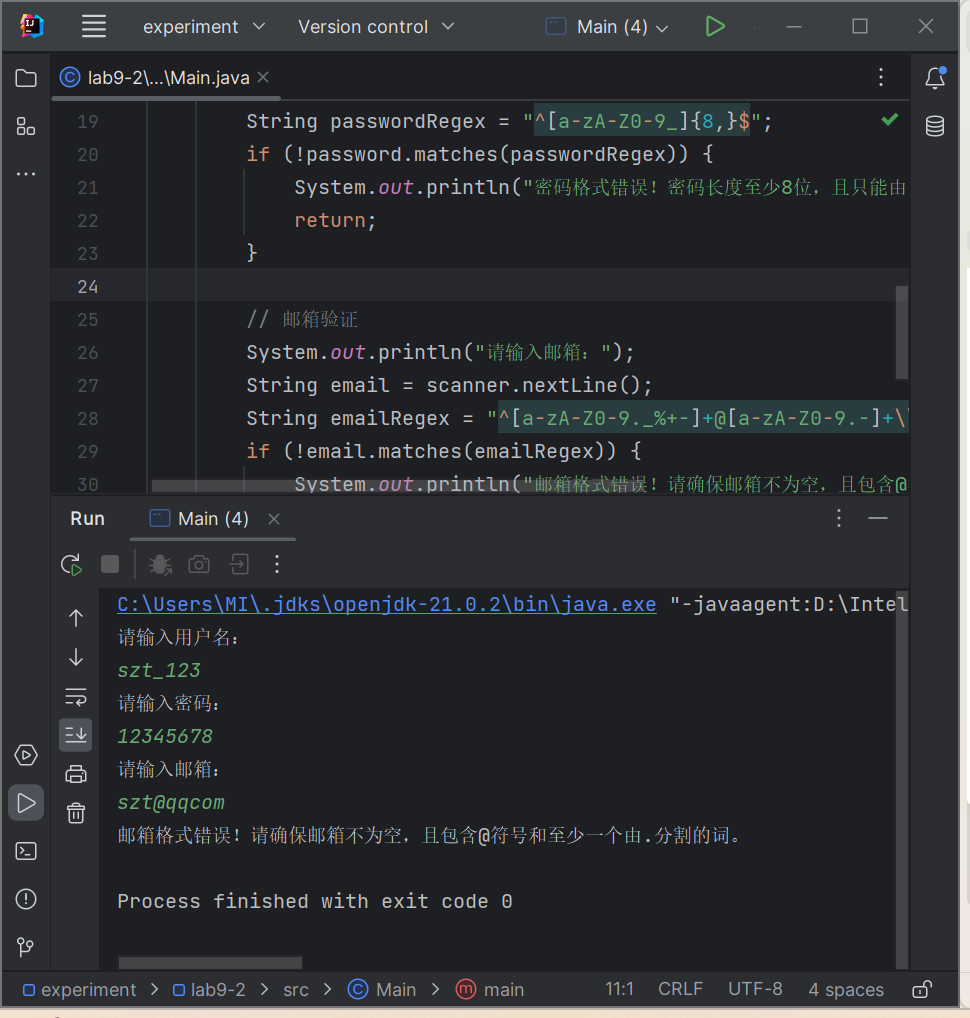
**1.邮箱为空**



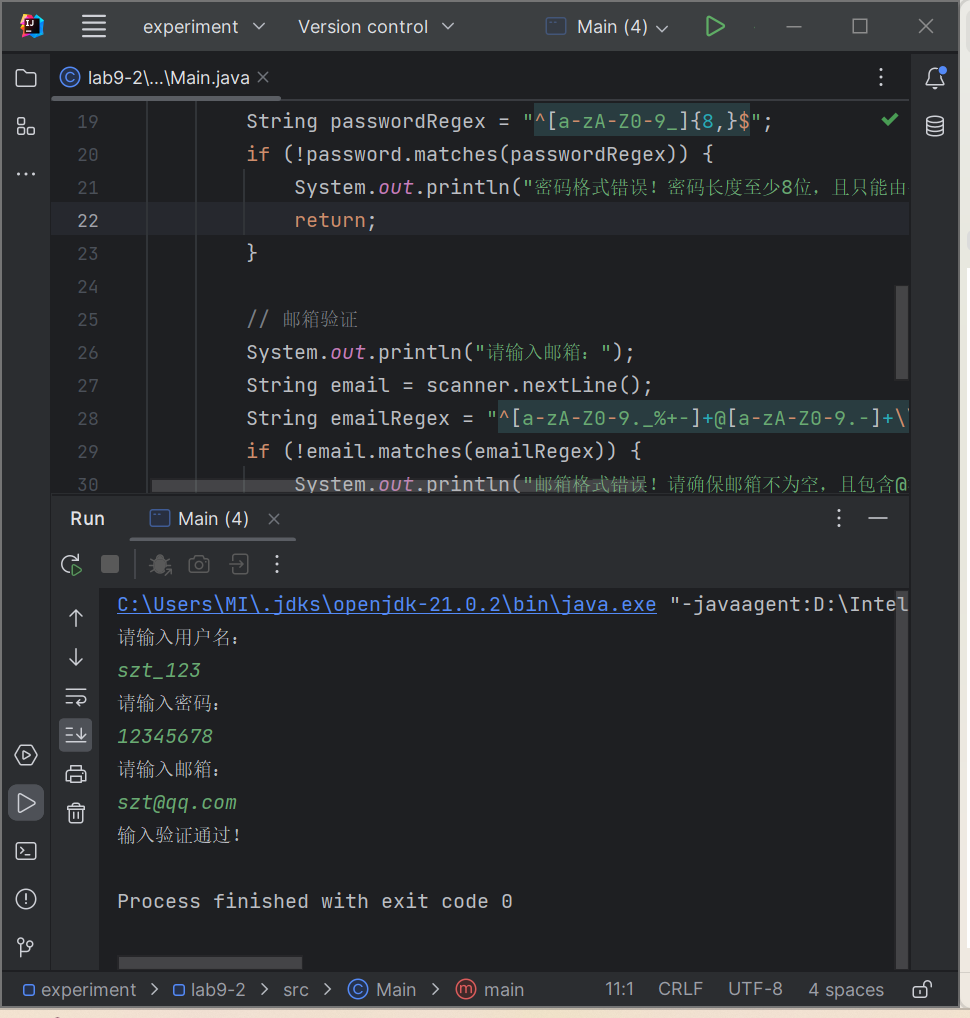
**2.邮箱不包含@**



**3.邮箱不包含’.’**



**4.邮箱正确**



1. **实验总结与心得记录**
2. **学习掌握了通过正则表达式对字符进行分割的方法**
3. **通过实现用户名、密码、邮箱验证和新闻单词统计的实验，增加了使用正则表达式的熟练度**