

2021-2022 秋 期末考试

'''

Author: XiangxinZhong

Date: 2021年12月29日17:14:06

LastEditors: XiangxinZhong

Description: 2021秋季期末考试总结

'''

说在前面：本着造福下一代的宗旨，于每场考试结束写下题目回忆，并给出个人复习建议。回忆题与真实考题很可能有出入，各位且看且珍惜，如果觉得不错的话给个star吧hhh ☺。

一、操作系统

1. 题型

仅包含能够回忆起的题目

选择题

1. openEuler;
2. OS: 何老师重点

填空题

简答题

1. 线程与进程的区别。
2. 微内核内核结构定义、优势、应用的OS? (最后一个我没写出来)
3. 简述死锁的4个必要条件。
4. 如何看待openEuler的新特性，简述4点。

综合题

1. 信号量的应用：前驱图，课本P113;
2. 分页存储管理：TLB时间计算，课本191例题;
3. 磁盘调度算法：FCFS, SSTF, SCAN移动磁道计算;
4. 进程周转时间：计算非抢占式和抢占式SJF的周转时间;
5. 综合题：TLB的综合使用，据说是考研真题;

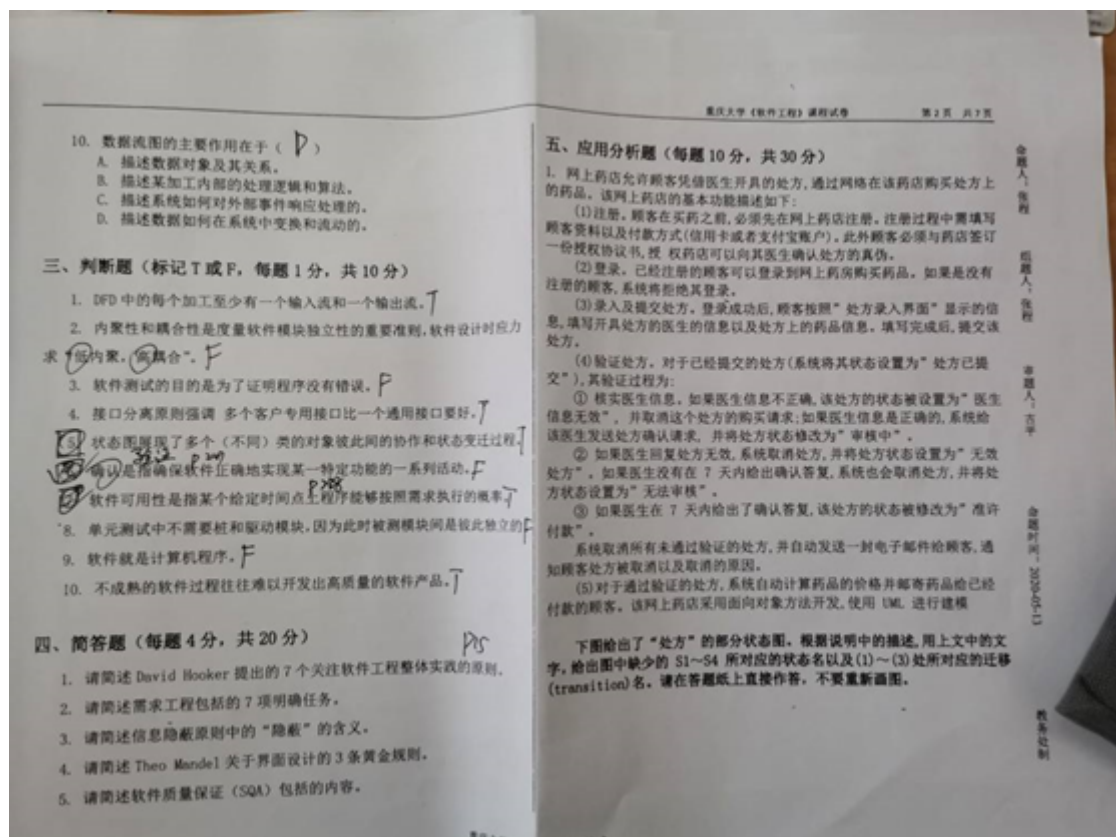
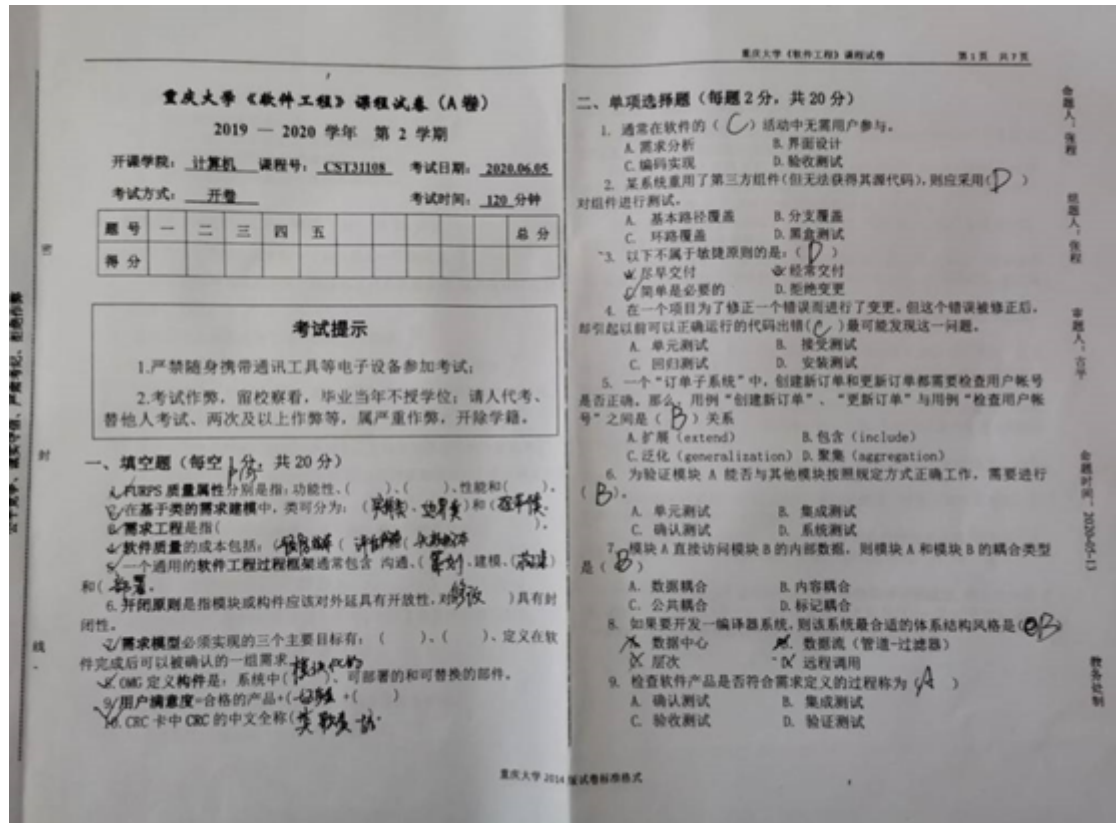
2. 复习建议

1. 课本+PPT: 根据何老师画的重点复习;
2. 有时间可以做点计算题

二、软件工程

1. 题型

参考2020年期末



2. 某高校欲开发一个成绩管理系统，记录并管理所有选修课程的学生的平时成绩和考试成绩，其主要功能描述如下：

(1) 每门课程都有3到6个单元构成，每个单元结束后会进行一次测试，其成绩作为这门课程的平时成绩。

(2) 学生的平时成绩和考试成绩均由每门课程的主讲教师上传给成绩管理系统。

(3) 在记录学生成绩之前，系统需要验证这些成绩是否有效，首先，根据学生信息文件来确认该学生是否选修这门课程，若没有，那么这些成绩是无效的；如果他的确选修了这门课程，再根据课程信息文件和课程单元信息文件来验证平时成绩是否与这门课程所包含的单元相对应，如果是，那么这些成绩是有效的，否则无效。

(4) 对于有效成绩，系统将其保存在课程成绩文件中，对于无效成绩，系统会单独将其保存在无效成绩文件中，并将详细情况提交给教务处。在教务处没有给出具体处理意见之前，系统不会处理这些成绩。

(5) 若一门课程的所有有效的平时成绩和考试成绩都被系统记录，系统会发送课程完成通知给教务处，告知该门课程的成绩已经齐全。教务处根据需要，请求系统生成相应的成绩列表，用来提交考试委员会审查。

重庆大学《软件工程》课程试卷 第3页 共7页

(6) 在生成成绩列表之前，系统会生成一份成绩报告给主讲教师，以便核对是否存在错误。主讲教师须将核对之后的成绩报告返还系统。

(7) 根据主讲教师核对后的成绩报告，系统生成相应的成绩列表，提交考试委员会进行审查。考试委员会在审查之后，上交一份成绩单给结果系统，对于所有通过审查的成绩，系统生成最终的成绩单，并通知每个选课学生。

现采用结构化方法对这个系统进行分析与设计，得到如下图2-1所示的顶层数据流图和图2-2所示的1层数据流图。

请根据说明中的描述，用上文中的文字，给出图中外部实体 E1-E4 的名称、数据存储 D1-D4 的名称。另外，图2-2中缺失了两个数据流，请补充（起点和终点请采用数据流图2-2中的符号或名称，数据流请用上文中的文字）。请在答题卡上直接作答，不要重新画图。

图2-1 顶层数据流图

重庆大学 2014 级试卷标准格式

合题人：张程 命题人：张程 审题人：吉平 命题时间：2015-05-11 教务处理

图2-2 1层数据流图

重庆大学《软件工程》课程试卷 第4页

3. 请采用基本路径测试法设计下列伪码程序的测试用例。

要求：

(1) 画出流程图

(2) 给出基本路径

(3) 给出满足基本路径测试标准的最小测试用例集。提示：测试用例=测试输入+预期输出

```

1: Start Input(a, b, c, d)
2: If (a>0)
3:   and (b>0)
4:   Then x=a+b
5:   Else x=a-b
6: End if
7: If (c>a)
8:   or (d>b)
9:   Then y=c-d
10: Else y=c+d
11: End if
12: Print(x, y) Stop

```

重庆大学 2014 级试卷标准格式

2. 试卷总结

2.1 填空题

知识点在课本上几乎都能找到，考前一定要认真过一遍课本，建议用荧光笔标记每页出现的名词，方便查阅。

2.2 选择题

相比于填空题需要对课本知识点有一定理解。

2.3 简答题

不难，抄书就完事了。

2.4 应用题

1. 数据流图，填空：外部实体、数据存储、流；
2. 类图：给定背景和类图，填写图中类名；
3. 白盒测试
 - 计算环路复杂度
 - 写出独立路径
 - 写出测试用例

3. 复习建议

课本很重要，像软工这种完形填空考试答案课本上都有。

三、计算机网络

1. 题型

选择题&填空题

只记得我没写出来的几个点，感觉有点偏。

1. ADSL是**什么复用**
2. PCM：采样、**量化**、编码
3. 海明码

简答题

1. 什么是协议？协议三要素？
2. 物理接口四个特性及具体实例
3. 数据链路层如何实现速度匹配和可靠传输？
4. 计算成功传输时间：两地距离1000m, 传输速率10Mbps, 数据量1500B, 传播速度 $2 \times 10^8 m/s$ 。
5. UDP和IP都是无连接的不可靠协议，能否直接用IP为应用层提供服务？为什么？
6. 电子邮件的4个协议是什么？作用分别是什么？

综合题

1. RIP协议
 - 简述RIP原理
 - R5与R1-R4相邻，给了R1-R4的路由表，计算R5的路由表。
2. TCP
 - 简述快重传，快恢复原理
 - 1-20次传输的拥塞窗口大小变化及原因分别是什么

2. 复习建议

可以结合谢希仁教材和李老师PPT复习，有时间做点王道考研的题。IP和TCP是重点，大题一般也会考。
ps 应用层协议也要记下。

四、自然语言处理

1. 题型

参考: <https://github.com/VayneDuan>

简单说一下变化:

- 计算题
 1. CNN计算: 1) 算卷积 2) padding后算卷积 (**步长改变了, 我没看见被坑了**) 3) 算average-pooling 4) 算RELU值
 2. 维特比算法: 1) 维特比变量 2) 给出隐状态序列
- 简答题记不得了, 难度不大, 就是抄书
- 综合题2021年是3个, 也是抄书, 不过很多, 记得抓紧时间抄书。

今年NLP难度还行, 题量很恐怖, 抄书抄的手快断了, 有的同学都没写完。

五、机器学习

说在前面: 今年ML计算题改成了算法题, 所以更像文科了, 主要就是看懂西瓜书和PPT上的算法, 背下重要的知识点就行。

题目回忆可能与真实题目有出入, 大致意思是一样的。

1. 题型

简答题 (顺序不一定对)

1. 简述假设空间和版本空间。
2. 简述训练误差、测试误差、泛化误差。是否训练误差小, 泛化误差一定小?
3. 朴素贝叶斯解决了什么困难? 提出的假设是什么?
4. 对数几率回归解决的是回归问题还是二分类问题? 训练方法是什么?

算法题

1. 补充注释: k-means聚类, 补充几条语句的注释

输入: 样本集 $D = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$;
 聚类簇数 k .

过程:

- 1: 从 D 中随机选择 k 个样本作为初始均值向量 $\{\mu_1, \mu_2, \dots, \mu_k\}$
- 2: **repeat**
- 3: 令 $C_i = \emptyset$ ($1 \leq i \leq k$)
- 4: **for** $j = 1, \dots, m$ **do**
- 5: 计算样本 \mathbf{x}_j 与各均值向量 μ_i ($1 \leq i \leq k$) 的距离: $d_{ji} = \|\mathbf{x}_j - \mu_i\|_2$;
- 6: 根据距离最近的均值向量确定 \mathbf{x}_j 的簇标记: $\lambda_j = \arg \min_{i \in \{1, 2, \dots, k\}} d_{ji}$;
- 7: 将样本 \mathbf{x}_j 划入相应的簇: $C_{\lambda_j} = C_{\lambda_j} \cup \{\mathbf{x}_j\}$;
- 8: **end for**
- 9: **for** $i = 1, \dots, k$ **do**
- 10: 计算新均值向量: $\mu'_i = \frac{1}{|C_i|} \sum_{\mathbf{x} \in C_i} \mathbf{x}$;
- 11: **if** $\mu'_i \neq \mu_i$ **then**
- 12: 将当前均值向量 μ_i 更新为 μ'_i
- 13: **else**
- 14: 保持当前均值向量不变
- 15: **end if**
- 16: **end for**
- 17: **until** 当前均值向量均未更新
- 18: **return** 簇划分结果

输出: 簇划分 $\mathcal{C} = \{C_1, C_2, \dots, C_k\}$

上面这段代码中的中文部分。

2. 补充代码：补充AdaBoost的代码

输入: 训练集 $D = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$;
 基学习算法 \mathcal{L} ;
 训练轮数 T .

过程:

- 1: $\mathcal{D}_1(\mathbf{x}) = 1/m$.
- 2: **for** $t = 1, 2, \dots, T$ **do**
- 3: $h_t = \mathcal{L}(D, \mathcal{D}_t)$;
- 4: $\epsilon_t = P_{\mathbf{x} \sim \mathcal{D}_t}(h_t(\mathbf{x}) \neq f(\mathbf{x}))$;
- 5: **if** $\epsilon_t > 0.5$ **then break**
- 6: $\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right)$;
- 7:
$$\begin{aligned} \mathcal{D}_{t+1}(\mathbf{x}) &= \frac{\mathcal{D}_t(\mathbf{x})}{Z_t} \times \begin{cases} \exp(-\alpha_t), & \text{if } h_t(\mathbf{x}) = f(\mathbf{x}) \\ \exp(\alpha_t), & \text{if } h_t(\mathbf{x}) \neq f(\mathbf{x}) \end{cases} \\ &= \frac{\mathcal{D}_t(\mathbf{x}) \exp(-\alpha_t f(\mathbf{x}) h_t(\mathbf{x}))}{Z_t} \end{aligned}$$
- 8: **end for**

输出: $H(\mathbf{x}) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(\mathbf{x}) \right)$

综合题

1. SVM

- SVM分类还是回归?
- 不能正确分类的问题还能用基本型SVM解决吗?
- 上一问的解决办法: 我回答的是软间隔和核方法

2. 集成学习

- 根据图片判断集成学习是否提升性能

| 测试例1 测试例2 测试例3 | | | | 测试例1 测试例2 测试例3 | | | | 测试例1 测试例2 测试例3 | | | |
|----------------|---|---|---|----------------|---|---|---|----------------|---|---|---|
| h_1 | ✓ | ✓ | × | h_1 | ✓ | ✓ | × | h_1 | ✓ | × | × |
| h_2 | × | ✓ | ✓ | h_2 | ✓ | ✓ | × | h_2 | × | ✓ | × |
| h_3 | ✓ | × | ✓ | h_3 | ✓ | ✓ | × | h_3 | × | × | ✓ |
| 集群 | ✓ | ✓ | ✓ | 集群 | ✓ | ✓ | × | 集群 | × | × | × |
| (a) 集群提升性能 | | | | (b) 集群不起作用 | | | | (c) 集群起负作用 | | | |

- 简述影响集成学习性能的因素

$$E = \overline{E} - \overline{A}$$

- 这个漂亮的式子显示:个体学习器精确性越高、多样性越大,则集成效果越好。称为误差-分歧分解

- 枚举提升集成学习效果的方法

2. 复习建议

考前看下西瓜书+陈老师的PPT就行了,主要是概念要理解和记忆,尽量全面一点(防止考到没看过的概念)。像常考的点,比如:过拟合vs欠拟合,分类vs回归,监督学习vs非监督学习.....,最好先总结答案考试直接默写。

六、系统结构

本想把题目写在草稿纸上撕下来带出考场,没想到被老师发现了qaq,系统结构纯纯的“回忆”版了。

说在前面:题型和去年差不多,但是难度陡增,考了很多动态优化的内容,我印象中去年几乎没怎么考优化的题目。

1. 题型

判断题

都比较简单,具体题目忘记了,会考一些计组的内容,大多数还是系统结构的知识点,看书和PPT理解概念就行了。

简答题

- 简述数据依赖并举例:WAW, WAR, RAW
- 简述Scoreboard和Tomasulo的区别,二者的优缺点分别是什么?
- 给定双重循环简述时间局部性和空间局部性,就是下面这个:

Example

```
for ( i= 0 ; i< 20 ; i++)  
    for( j=0; j<10; j++)  
        a[i]=a[i]*j;
```

- Spatial locality: for the outer loop, it accesses memory $a[0], a[1], \dots, a[20]$ sequentially.
- Temporal locality: for the inner loop, it accesses memory $a[i]$ repeatedly in each iteration.

4. cache的优化技术：各简述一种降低miss rate和miss penalty的方法及它们的原理（其实就是PPT上的这一页）

课本P421 Cache Optimization Techniques

Average memory access time =
Hit time + Miss rate \times Miss penalty

- **Reducing the miss rate**（重点）
larger block size, larger cache size, and higher associativity
- **Reducing the miss penalty**（重点）
multilevel caches and giving reads priority over writes
- **Reducing the time to hit in the cache**（了解）
avoiding address translation when indexing the cache

5. 2-bit 饱和预测器，题型和PPT上这一页差不多

Example

A snapshot of the taken/not-taken behavior of a branch is:

... T T T T T T T T T T N N T T N N T N N T

If the branch predictor used is a 2-bit saturating counter, how many of the last ten branches are predicted correctly?

Answer :

According to the branch predictor in textbook, the prediction for the last ten branches are:

ST, WT, SN, WN, ST, WT, SN, **WN**, **SN**, SN

According to the 2-bit saturating counter:

ST, WT, WN, **WT**, ST, WT, WN, WT, **WN**, SN

No matter which one you use, the answer is 2 branches are predicted correctly.

综合题

1. 和半期考试这道题目非常类似（可以说题型几乎一样，只是改了指令）

6. Pipeline Timing Chart (3*10%=30%)

```
Loop: LD R1, 0(R2) ; load R1 from address 0+R2
      LD R5, 2(R2) ; load R5 from address 2+R2
      DADDI R5, R1, R5 ; R5=R1+R5
      SD R5, 0(R2) ; store R5 at address 0+R2
      DADDI R2, R2, #4 ; R2=R2+4
      DSUB R4, R3, R2 ; R4=R3-R2
      BNEZ R4, Loop ; branch to Loop if R4!=0
```

Assume branch instruction finishes at ID stage, and new PC available at the next cycle.

(1) Data hazards are caused by data dependences in the code. Whether a dependency causes a hazard depends on the machine implementation (i.e., number of pipeline stages). List all of the data dependences in the code above. Record the register, source instruction, and destination instruction.

(2) Use a pipeline timing chart (no more than 20 cycles) to show the timing of this instruction sequence for the 5-stage MIPS pipeline without any forwarding or bypassing hardware but **assuming that a register read and a write in the same clock cycle “forwards” through the register file**. Assume that the branch is handled by flushing the pipeline.

(3) Use a pipeline timing chart (no more than 15 cycles) to show the timing of this instruction sequence for the 5-stage MIPS pipeline with full forwarding and bypassing hardware. Assume that the branch is handled by predicting it as not taken.

2. loop unrolling, 和PPT上的题型大致相同

Example: a Simple Parallel Loop

```
for (i=1000; i>0; i=i-1) {
    x[i] = x[i] + s;
}
```



compiled into MIPS
assembly language

```
loop:  L.D    F0, 0(R1)    ; R1 = highest address of array element
      ADD.D  F4, F0, F2    ; add scalar s (stored in F2)
      S.D    F4, 0(R1)    ; store result
      DADDUI R1,R1,#-8     ; decrement pointer by 8 bytes (DW)
      BNE    R1,R2, loop  ; 8(R2) points to x[1], last to process
```

3. Scoreboard题目

- 简述动态调度的优缺点
- 给一段代码，简述里面的WAW, WAR, RAW
- 推导scoreboard的执行过程，在表格中填写每条指令issue、read operands、execution complete、commit的周期，跟PPT上面的这道题很相似（只是没有FU和寄存器状态表，只有指令状态表）

Scoreboard Example: Clock Cycle = 1

| Instruction | Issue | Read Operands | Exec. Complete | Write Result |
|-------------------|-------|---------------|----------------|--------------|
| L.D F6, 34(R2) | 1 | | | |
| L.D F2, 45(R3) | | | | |
| MUL.D F0, F2, F4 | | | | |
| SUB.D F8, F6, F2 | | | | |
| DIV.D F10, F0, F6 | | | | |
| ADD.D F6, F8, F2 | | | | |

| Name | Busy | Op | Fi | Fj | Fk | Qj | Qk | Rj | Rk |
|---------|------|-----|----|----|----|----|----|----|------|
| Integer | yes | L.D | F6 | | R2 | | | | true |
| Mult-1 | no | | | | | | | | |
| Mult-2 | no | | | | | | | | |
| Add | no | | | | | | | | |
| Divide | no | | | | | | | | |

| | F0 | F2 | F4 | F6 | F8 | F10 | ... | F30 |
|----|----|----|----|---------|----|-----|-----|-----|
| FU | | | | Integer | | | | |

• issue first instruction

2. 复习建议

今年的难度比去年大了很多，动态优化那四个方法考了很多（scoreboard、tomasulo、分支预测、循环展开）。建议提前一周开始复习，看书和肖老师的PPT就差不多了。btw强烈建议听下肖老师的复习课。