

【超星尔雅学习通】大数据算法网课章节答案



超星尔雅学习通大数据算法(哈尔滨工业大学)网课章节测试答案

1.1 大数据的定义与特点

1、【单选题】以下关于大数据的特点，叙述错误的是)。

- A、速度慢
- B、多元、异构
- C、数据规模大
- D、基于高度分析的新价值

我的答案： A

2、【单选题】在《法华经》中，“那由他”描写的“大”的数量级是()。

- A、 10^7
- B、 10^{14}
- C、 10^{28}
- D、 10^{56}

我的答案： C

3、【多选题】以下选项中，大数据涉及的领域中包括()。

- A、社交网络
- B、医疗数据
- C、计算机艺术

D、 医疗数据

我的答案： ABC(D)

4、 【多选题】大数据的应用包括)。

A、 预测

B、 推荐

C、 商业情报分析

D、 科学研究

我的答案： ABCD

5、 【判断题】目前，关于大数据已有公认的确定义。

我的答案： X

6、 【判断题】大数据种类繁多，在编码方式、数据格式、应用特征等方面都存在差异。)

我的答案： √

1.2 大数据算法(1)

1、 【单选题】大数据求解计算问题过程的第三步一般是)。

A、 判断可计算否

B、 判断能行可计算否

C、 算法设计与分析

D、 用计算机语言实现算法

我的答案： C

2、 【多选题】在大数据求解计算问题中，判断是否为能行可计算的因素包括)。

- A、数据量
- B、资源约束
- C、速度约束
- D、时间约束

我的答案： ABD

3、【判断题】大数据求解计算问题过程的第一步是确定该问题是否可计算。

我的答案： √

4、【判断题】大数据计算模型与一般小规模计算模型一样，都使用的是图灵机模型。

我的答案： √

1.3 大数据算法(2)

1、【多选题】资源约束包括()。

- A、CPU
- B、网络带宽
- C、内存
- D、外存

我的答案： ABCD

2、【多选题】大数据算法可以不是()。

- A、云计算
- B、精确算法
- C、内存算法

D、 串行算法

我的答案： BCD

3、【判断题】大数据算法是在给定的时间约束下，以大数据为输入，在给定资源约束内可以生成满足给定约束结果的算法。

我的答案： X

4、【判断题】MapReduce是一种比较好实现大数据算法的编程架构，在生产中得到广泛应用。

我的答案： √

5、【判断题】大数据算法是仅在电子计算机上运行的算法。

我的答案： X

1.4大数据的特点与大数据算法

1、【单选题】众包算法是用来解决)。

A、访问全部数据时间过长

B、数据难于放入内存计算

C、单个计算机难以保存全部数据，计算需要整体数据

D、计算机计算能力不足或知识不足，需要人来帮忙

我的答案： D

2、【单选题】大数据算法存在很多难题，对于访问全部数据时间过长的的问题，采用的解决方案是)。

A、将数据存储到磁盘上

B、仅基于少量数据进行计算

C、读取部分数据

D、并行处理

我的答案： C

3、【多选题】大数据算法的()特点，使其与大数据算法密切相关的。

A、数据量大

B、基于高度分析的新价值

C、速度快

D、多样性、复杂性

我的答案： AC

4、【判断题】为解决单个计算机难以保存全部数据的问题，通常会采用并行处理的技术，此技术会涉及到时间亚线性算法。

我的答案： X

1.5大数据算法设计与分析

1、【单选题】大数据算法涉及到外存的时候，通常要分析)。

A、时间空间复杂性

B、IO 复杂性

C、结果质量

D、通讯复杂性

我的答案： B

2、【多选题】下列选项中，属于智能仿生算法的是)。

A、遗传算法

B、近似算法

C、模拟退火算法

D、数据流算法

我的答案： AC

3、【判断题】在线算法/数据流算法是面向大数据速度快的特点提出的。))

我的答案： ✓

4、【判断题】对于数据流算法或在线算法，经常要分析结果的近似比。 ()

我的答案： X

5、【判断题】随机算法是利用随机化的方法来进行大数据处理，是大数据算法设计技术之一。))

我的答案： ✓

2.1 亚线性算法的定义

1、【单选题】计算在一个给定社交网络中平均每人的朋友个数，在不访问所有顶点的情况下，进行精确计算最少需要访问)个顶点。

A、 n

B、 n+1

C、 2n

D、 n-1

我的答案： D

2、【多选题】亚线性是指)等的消耗是输入规模o。))

A、 时间

B、 空间

C、IO

D、通讯

我的答案： ABCD

3、【判断题】性质检测算法属于亚线性空间算法的一类。()

我的答案： X

2.2 水库抽样——空间亚线性算法

1、【判断题】在经典的水库抽样中，要求空间复杂性为 $O(k)$ ，是指与抽样大小有关，而与整个数据的数据量无关。

我的答案： √

2、【判断题】水库抽样算法的采样是均匀的。

我的答案： √

2.3 平面图直径——时间亚线性计算算法

1、【单选题】以下()不是衡量分析近似解代价与优化解代价差距的方法。

A、Ratio Bound

B、相对误差

C、绝对误差

D、 $(1-)$ -近似

我的答案： C

2、【单选题】利用平面图的直径近似算法得到的解，在最坏情况下，也不会小于最优解的()。

A、二分之一

B、三分之一

C、四分之一

D、十分之一

我的答案： A

3、【单选题】采用平面图的直径近似算法的动机是无法在要求的时间内得到()。

A、相似解

B、完整解

C、精确解

D、近似解

我的答案： C

4、【判断题】在平面图的直径近似算法中，要求点之间的距离满足三角不等式是指在 i 、 j 、 k 三个点中， i 到 j 的距离加上 j 到 k 的距离小于 i 到 k 的距离。

我的答案： X

5、【判断题】近似算法能给出一个优化问题的优化解。

我的答案： X

6、【判断题】Ratio Bound 越大，则近似解越坏。()

我的答案： √

2.4全0数组判定——时间亚线性判定算法

1、【多选题】在判定问题的近似中，对于近似解需要区分的是)。

A、是

B、否

C、差不离

D、差得很远

我的答案： AD

2、【判断题】全0数组判定的近似算法的证据引理是：如果一次测试以大于等于 p 的概率获得一个证据，那么 $s=2/p$ 轮测试得到证据的概率大于等于 $1/2$ 。

我的答案： X

3、【判断题】判定问题的近似解是指：输入满足某种性质或近似满足某种性质。

我的答案： X

3.1 数据流中频繁元素

1、【判断题】Zipf 原则是指典型的频率分布是高度偏斜的，存在很多频繁元素。)

我的答案： X

2、【判断题】在数据流模型中，从数据流中可以计算简单的函数，如最大值、最小值、求和等，且处理这些函数时通常使用单个寄存器 s 。()

我的答案： √

3、【判断题】数据流模型中，数据流是指来自某个域中的元素序列。()

我的答案： √

4、【判断题】在数据流模型中，内存远远大于数据的规模。（）

我的答案： X

3.2 频繁元素计算算法

1、【单选题】频繁元素计算算法又称为)算法。

A、MM

B、MG

C、GM

D、MP

我的答案： B

2、【单选题】频繁元素计算算法有效的原因是源于)。

A、取近似解

B、证据引理

C、错误界限和k 成反比

D、Zipf 原则

我的答案： D

3、【判断题】在频繁元素计算算法中，计数器x 减少的次数依赖于有几个减少计数器的步骤。（）

我的答案： √

4、【判断题】在频繁元素计算算法中，当数据流中元素的总个数远大于估计值与真实值相差的最多值时，可以得到频繁项一个好的估计。

我的答案： √

3.3 最小生成树

1、【单选题】时间亚线性算法的思想是：利用特定子图联通分量的数量估计最小生成树的()。

- A、近似值
- B、精确值
- C、权重
- D、界限

我的答案： C

2、【单选题】求最小生成树是一个贪心法，可以用()算法来解决。

- A、Prime
- B、并行
- C、MG
- D、内存

我的答案： A

3、【判断题】对联通分量个数的估计可以利用随机化方法。

我的答案： √

3.4序列有序的判定

1、【单选题】对于输入 n 个数的数组 $(x_1, x_2, x_3, \dots, x_n)$ ，输出：这个数组是否有序。远离意味着必须删除大于 0 个元素才能保证剩下的元素有序。

- A、 $-n$
- B、 n
- C、 $/n$

D、n/

我的答案： B

2、【判断题】如果一次测试以大于等于 p 的概率获得一个证据，那么 $s=2/p$ 轮测试得到证据的概率大于等于 $3/4$ 。

我的答案： X

3、【判断题】关于证明如果输入远离有序，则存在大于 n 个“坏索引”的问题，可以采用证明其逆否命题的方法。

我的答案： √

4.1外存存储结构与外存算法

1、【单选题】下列选项中叙述正确的是)。

A、磁盘的访问可以随机读、随机取

B、磁盘系统传输大规模连续的数据块的范围是18-32k

C、磁盘访问比主存访问的速度快

D、大多数程序在RAM模型上运行

我的答案： D

2、【判断题】对于大数据而言，标准计算理论模型失效的原因之一是内存是有限的，无法存储所有的内存。)

我的答案： √

3、【判断题】当内存不够或者算法设计不好时，如果数据量达到一定规模以上，运行时间会急剧增加。

我的答案： √

4、【判断题】现代计算机有复杂的存储层次，存储单元的访问是以块

为单位的数据移动。()

我的答案： ✓

4.2.1 外存算法示例：外存排序算法(1)

1、【单选题】外存归并排序，以)为单位进行调度。

A、比特

B、兆

C、块

D、层

我的答案： C

2、【判断题】排序分为内部排序和外部排序，外部排序是因排序的数据很大，一次不能容纳全部的排序内容，在排序过程中需要访问外存。()

我的答案： ✓

4.2.2 外存算法示例：外存排序算法(2)

1、【单选题】在外排序的快速排序中，分割元素的选择非常重要。

我的答案： ✓

4.3 外存数据结构示例：外存查找树

1、【单选题】在内存中的二分搜索树中，通常使用()来维护树的平衡。()

A、置换

B、分裂

C、旋转

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。
。如要下载或阅读全文，请访问：<https://d.book118.com/986240053124010144>