

**Вопросы к государственному экзамену
Магистерская программа «Суперкомпьютерные системы и приложения»**

1. Поколения архитектур компьютеров и парадигмы программирования.
2. Архитектурные особенности современных микропроцессоров. Особенности организации подсистемы памяти и динамической памяти, их влияние на производительность суперкомпьютеров при решении различных задач, в том числе задач обработки больших графов.
3. Высокоскоростные коммуникационные сети. Основные характеристики: задержка, пропускная способность, темп выдачи сообщений, бисекционная пропускная способность. Технология RDMA. Топологии сетей: тор, жирное дерево, Flattened Butterfly, Dragonfly.
4. Последовательная и параллельная сложность алгоритмов, информационный граф и ресурс параллелизма алгоритмов.
5. Проблемы и методы эффективной организации параллельных вычислений при суперкомпьютерном решении задач обработки больших графов.
6. Архитектурные особенности графических процессоров, направленные на массивно-параллельные вычисления. Методы эффективной организации параллельных вычислений на графических процессорах.
7. Методы статической и динамической балансировки загрузки процессоров: сдваивания, геометрического параллелизма, коллективного решения, конвейерного параллелизма, диффузной балансировки загрузки.
8. Декомпозиция расчетных сеток. Критерии декомпозиции. Иерархические методы разбиения графов большого размера. Методы геометрической и спектральной бисекции, метод инкрементного роста, метод локального уточнения.
9. Параллельные алгоритмы сортировки данных.
10. Лямбда-архитектура обработки больших данных.
11. Консенсус в распределённых системах.
12. Технология распределённого хранения информации блокчейн.
13. Клеточные автоматы: определение, элементарные клеточные автоматы, классификация Вольфрама, двумерные клеточные автоматы, типы окрестностей, игра "Жизнь", параллельная реализация.
14. Сети Петри: определение, примеры, вариации (сети с приоритетами, ингибиторные сети, цветные сети), моделирование параллельных процессов.
15. Генетические алгоритмы: операторы генетических алгоритмов, особенности кодирования (двоичное, целочисленное, непрерывное, перестановками), сходимость генетических алгоритмов (теория схем), островная модель, клеточные генетические алгоритмы.
16. Информационная безопасность. Шифрование данных. Криптографическая стойкость. Симметричная криптография. Блочный шифр (DES) и его режимы. Ассиметричные схемы (RSA и Диффи-Хеллмана). Код аутентификации (MAC). Цифровая подпись (DSA).
17. Односторонняя и двусторонняя модели передачи сообщений в библиотеке MPI, принципы программирования и синхронизации параллельных процессов. Протоколы Eager и Rendezvous для реализации двусторонней модели.
18. Параллельные алгоритмы обработки графов. Алгоритмы поиска вширь, Беллмана-Форда, дельта-сетепинга, Боровки, лувенский, распространения меток, Брандеса.
19. Математические основы GraphBLAS: данные, операции.
20. Псевдополиномиальные алгоритмы решения задач: разбиение, рюкзак, расписание для многопроцессорной системы (число процессоров фиксировано).
21. Метод ветвей и границ на примере минимаксной задачи теории расписаний. Приближенные алгоритмы решения NP-трудных задач: упаковка в контейнеры, рюкзак, коммивояжер, расписание для многопроцессорной системы, вершинное покрытие. Оценки их сложности и погрешности.
22. Средние и эмпирические операционные характеристики стратегий распознавания (классификаторов, регрессий). Проблема недообучения и переобучения. Проблема устойчивости решений. Роль обучающей, валидационной и контрольной выборок при построении распознающей системы. Скользящий контроль (кросс-валидация). Регуляризация на примере линейной регрессии, её роль.
23. Проблема смещения-дисперсии. Ансамбли классификаторов. Основные этапы работы типичного базового классификатора, возможность коррекции на разных этапах. Бэггинг. Случайные подпространства. Бустинг. Случайный лес как композиция основных подходов к построению ансамбля.
24. Задачи кластеризации, сопоставление с операцией группирования и задачей классификации. Различные постановки: разбиение, стохастическая, нечёткая, иерархическая, упорядочивание, однокластерная (последовательная). Задача и процедура K-средних. Метод нечеткой кластеризации C-средних. Агломеративная кластеризация.

Список рекомендованной литературы

1. Воеводин В.В., Воеводин Вл.В. [Параллельные вычисления](#). - СПб.: БХВ-Петербург, 2002. - 608 с.
2. Яковлевский М.В. Введение в параллельные методы решения задач: Учебное пособие. - М.: Издательство Московского университета, 2012. - 328 с.
3. Антонов А.С. Технологии параллельного программирования MPI и OpenMP: Учеб. пособие. - М.: Издательство Московского университета, 2012. - 344 с. (Серия "Суперкомпьютерное образование"). ISBN 978-5-211-06343
4. А. В. Боресков и др. Параллельные вычисления на GPU. Архитектура и программная модель CUDA: Учебное пособие. - Издательство Московского университета, 2012, 336 стр.
5. Интернетресурсы: <http://parallel.ru>, <http://AlgoWiki-Project.org>,
6. www.theochem.ruhr-uni-bochum.de/research/marx/marx.pdf
7. Ершов Н.М., Попова Н.Н. Естественные модели параллельных вычислений: Учебное пособие. - М., Макс-пресс, 2016
8. Жуматий С.А., Дацюк О.В. "Администрирование суперкомпьютеров и кластерных систем". Москва, Изд-во МГУ, 2014.
9. Давыдов Э.Г. *Исследование операций*. М.: Высшая школа, 1990.
10. Морозов В.В. *Основы теории игр*. М.: МГУ, 2002.
11. Кормент Т., Лейзерсон Ч., Ривест Р., Штайн К. *Алгоритмы. Построение и анализ*. М.: МЦНМО, 2005.
12. Пападимитриу Х., Стайглиц К. *Комбинаторная оптимизация. Алгоритмы и сложность*. М.: Мир, 1985.
13. Гэри М., Джонсон Д. *Вычислительные машины и труднорешаемые задачи*. М.: Мир, 1982.
14. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. - М.: ДМК Пресс. - 2015. - 400 с. ISBN 978-5-97060-273-7 (Flach P. Machine learning: the art and science of algorithms that makes sense of data. - Cambridge University Press, 2012)
Bishop С.М. *Pattern recognition and machine learning*. - Springer, 2006.
Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию // Киев: Наукова думка. - 2004.
15. А. Антонопулос Осваиваем биткойн. Программирование блокчейна - М.: ДМК Пресс, 2018. - 428 с.
16. Э. Таненбаум, М. ван Стеен *Распределенные системы*. - М.: ДМК Пресс, 2021. - 584 с.

Дополнительная литература

17. Коэльо Л. П., Ричарт В. Построение систем машинного обучения на языке Python. - М.: ДМК Пресс. - 2016. (Coelho L. P., Rich-ert W. Building machine learning systems with Python. - 2nd ed. - Packt Publishing Ltd, 2015.)
18. Max Kuhn, Kjell Johnson. *Applied Predictive Modeling*. - Springer, 2013.
19. Hastie, T., Tibshirani R., Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. - 2nd ed. - Springer-Verlag, 2009. - 746 p. - ISBN 978-0-387-84857-0.
20. Журавлев Ю. И., Рязанов В. В., Сенько О. В. «Расознавание». Математические методы. Программная система. Практические применения. - М.: Фазис, 2006. ISBN 5-7036-0108-8.
21. I.H. Witten, E. Frank *Data Mining: Practical Machine Learning Tools and Techniques*. - 2nd ed. - Morgan Kaufmann, 2005 ISBN 0-12-088407-0

Ресурсы интернета

www.machinelearning.ru