

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

1. 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.
2. Высокоскоростные коммуникационные сети. Основные характеристики: задержка, пропускная способность, темп выдачи сообщений, бисекционная пропускная способность. Технология RDMA. Топологии сетей: тор, жирное дерево, Flattened Butterfly, Dragonfly.
3. Образцы (паттерны) проектирования, их классификация и способ описания. Пример образца проектирования.
4. Проблемы и методы эффективной организации параллельных вычислений при суперкомпьютерном решении задач обработки больших графов.
5. Архитектурные особенности графических процессоров, направленные на массивно-параллельные вычисления. Методы эффективной организации параллельных вычислений на графических процессорах.
6. Методы статической и динамической балансировки загрузки процессоров: сдваивания, геометрического параллелизма, коллективного решения, конвейерного параллелизма, диффузной балансировки загрузки.
7. Декомпозиция расчетных сеток. Критерии декомпозиции. Иерархические методы разбиения графов большого размера. Методы геометрической и спектральной бисекции, метод инкрементного роста, метод локального уточнения.
8. Параллельные алгоритмы сортировки данных.
9. Лямбда-архитектура обработки больших данных.
10. Консенсус в распределённых системах.
11. Корпоративные хранилища данных (datawarehouse), озёра данных.
12. Клеточные автоматы: определение, элементарные клеточные автоматы, классификация Вольфрама, двумерные клеточные автоматы, типы окрестностей, игра "Жизнь", параллельная реализация.
13. Сети Петри: определение, примеры, вариации (сети с приоритетами, ингибиторные сети, цветные сети), моделирование параллельных процессов.
14. Генетические алгоритмы: операторы генетических алгоритмов, особенности кодирования (двоичное, целочисленное, непрерывное, перестановками), сходимость генетических алгоритмов (теория схем), островная модель, клеточные генетические алгоритмы.
15. 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).
16. Односторонняя и двусторонняя модели передачи сообщений в библиотеке MPI, принципы программирования и синхронизации параллельных процессов. Протоколы Eager и Rendezvous для реализации двусторонней модели.
17. Параллельные алгоритмы обработки графов. Алгоритмы поиска вширь, Беллмана-Форда, дельта-сетеппинга, Боровки, лувенский, распространения меток, Брандеса.
18. Математические основы GraphBLAS: данные, операции.
19. Позиционные игры с полной информацией и игры общего вида. Смешанные стратегии и стратегии поведения. Алгоритм Куна определения совершенного подигрового равновесия.
20. Задачи распределения ресурсов. Теорема Гермейера. Дискретная задача распределения ресурсов. Лемма Гиббса.
21. 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.
22. 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.
23. Задачи кластеризации, сопоставление с операцией группирования и задачей классификации. Различные постановки: разбиение, стохастическая, нечёткая, иерархическая, упорядочивание, однокластерная (последовательная). Задача и процедура К-средних. Метод нечеткой кластеризации С-средних. Агломеративная кластеризация.

Список рекомендованной литературы

1. Воеводин В.В., Воеводин Вл.В. [Параллельные вычисления](#). -СПб.: БХВ-Петербург, 2002. -608с.
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"). ISBN978-5-211-06343
4. А. В. Боресков и др. Параллельные вычисления на GPU. Архитектура и программная модель CUDA: Учебное пособие. -Издательство Московского университета, 2012, 336 стр.
5. Интернет ресурсы: <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. Давыдов Э.Г. *Исследование операций*. М.: Высшая школа, 1990.
9. Морозов В.В. *Основы теории игр*. М.: МГУ, 2002.
10. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. *Алгоритмы. Построение и анализ*. М.: МЦНМО, 2005.
11. Пападимитриу Х., Стайн К. *Комбинаторная оптимизация. Алгоритмы и сложность*. М.: Мир, 1985.
12. Гэри М., Джонсон Д. *Вычислительные машины и трудно решаемые задачи*. М.: Мир, 1982.
13. 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 make sense of data. –Cambridge University Press, 2012) Bishop C.M. Pattern recognition and machine learning. – Springer, 2006.
14. 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. —Spring-er-Verlag, 2009. — 746 p.— ISBN978-0-387-84857-0.
20. Журавлев Ю. И., Рязанов В. В., Сенько О. В. «Распознавание». Математические методы. Программная система. Практические применения. —М.: Фазис, 2006. ISBN5-7036-0108-8.
21. I.H. Witten, E. Frank Data Mining: Practical Machine Learning Tools and Techniques. — 2nd ed. — Morgan Kaufmann, 2005 ISBN0-12-088407-0
22. A. Lumsdaine, D. Gregor. Challenges in parallel graph processing // Parallel Process. Lett. 17, 5, 2007.
23. Blondel, Vincent D., et al. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment 2008.10 (2008).
24. Что каждый программист должен знать о памяти, перевод Капустин С.В., М.Ульянов, Н.Ромоданов, 2009-2012. Оригинал U.Drepper, 2007.
25. U. Meyer, P. Sanders. Delta-Stepping: A Parallel Single Source Shortest Path Algorithm. In Proceedings of the 6th Annual European Symposium on Algorithms (ESA-98), Venice, Italy, August, 24 - 26, 1998, 393-404.
26. U. Brandes. A Faster Algorithm for Betweenness Centrality. Journal of Mathematical Sociology, vol. 25, no. 2, pp. 163–177, 2001.
27. U. Raghavan, R. Albert, S. Kumara. Near linear time algorithm to detect community structures in large-scale networks.

28. Головина Е.А., Семенов А.С., Фролов А.С. Исследование производительности задачи поиска вширь в графе на сопроцессорах семейства Intel Xeon Phi // Вычислительные методы и программирование. – 2014. – Т. 15. – С. 49-58.
29. J. Kepner et al. Mathematical foundations of the GraphBLAS. 2016 IEEE High Performance Extreme Computing Conference (HPEC). IEEE, 2016.

Ресурсы интернета

www.machinelearning.ru