CS 370: Undergraduate Reading and Research

Spring 2024

General Information

Student: Jose Alvarez Loredo, **EID:** jea3268 enriprivacc@utexas.edu
Supervising Instructor: Siddhartha Chatterjee, sidchat@utexas.edu
*Adjustments to the course format are not final.

Description

370/350C Cossover project: Evaluate the space-time tradeoffs for different hashing algorithms on modern computing systems, measuring architectural effects especially for very large tables.

Class Meetings:

Schedule to be announced (TBA). Additionally, to ensure effective collaboration, we can organize brief meetings aligned with the CS 350C class schedule in the spring semester, either before or after the class sessions.

Initial Information

Recommended Paper: arXiv:2209.07663.

Hash Tables: There are three primary types of hash tables to consider, categorized by their collision resolution methods: Open Addressing, Chaining, Cuckoo Hashing.

Starting Points:

- A good starting point for understanding cuckoo hashing is the Wikipedia entry: Cuckoo Hashing on Wikipedia.
- For practical insights and techniques, the following blog has several pertinent articles: Writing a Damn Fast Hash Table with Tiny Memory Footprints.

Performance Measurements: It is noted that many performance metrics available are outdated. Hence, it is recommended to prioritize performance measurements, beginning with the most efficient implementations currently available.

Topic	Description	Due Date
Preliminary Work	**Review Readings:** Carefully read through the materials provided in the *'Starting Points'* section. **Additional Readings:** Engage in an extensive review of supplementary academic papers relevant to the topic. **Outline Completion:** Based on your understanding from the readings, complete the course outline for CS-370. Cover all necessary topics and key concepts that will be explored throughout the course.	01/16/2024
Hashing	**Implementation Task:** Gain a deeper insight into hash table operations by implementing the Open Addressing, Chaining, and Cuckoo Hashing methods. **Note:** Implementations are for educational purposes only. The performance measurements for the project will utilize the optimized algorithms developed by external experts.	01/19/2024
Research Paper	**Research Paper Template:** Develop a LaTeX template for the research paper, adhering to standard formatting rules. Investigate additional guidelines on structuring the paper, ensuring compliance with common academic standards and best practices. **Paper:** Finish the Introduction section	01/26/2024
Literature Review	**Literature Review Preparation:** Conclude all supplementary readings pertinent to the research paper. Proceed with drafting the literature review section, integrating insights and findings from these readings.	02/02/2024
Checkpoint 1	**Instructor Consultation:** Schedule a progress check meeting with the supervising instructor. Actively incorporate any feedback or suggestions received during the consultation into your work. **Paper:** Finish the Literature Review section	02/09/2024
Methodology	**Simulation/Implementation:** Simulate the optimized hashing algorithms. Ensure the chosen approach effectively demonstrates the algorithms' functionalities and nuances. **Testing Environment Setup:** Establish a stable and consistent testing environment that mirrors modern computing systems. Critical for obtaining accurate performance metrics of the algorithms.	02/16/2024
Data Preparation	**Data Set Preparation:** Acquire or create extensive data sets necessary for evaluating the hash tables. The data sets should be representative of real-world scenarios to ensure the relevance and applicability of the test results. **Paper:** Finish the Methodology section	02/23/2024

Table 1: Project Timeline

Topic	Description	Due Date
Experimentation	**Run Tests:** Execute the hashing algorithms using the prepared data sets. Ensure diverse testing conditions are employed to effectively assess the algorithms' performance under various architectural impacts. **Data Collection:** Meticulously document the performance metrics of each algorithm across different test conditions.	03/01/2024
Results	**Data Presentation:** Devise an effective method for incorporating the collected data into the research paper's results section. Focus on clarity and comprehensibility in presenting the findings. **Continued Testing:** If necessary, continue executing tests throughout the week to gather additional data or validate previous results. **Paper Revision:** Initiate a comprehensive review and revision process for the entire research paper. This stage includes refining content, enhancing clarity, and ensuring coherence across all sections.	03/08/2024
Catch-up week	**March 11 - 16: Spring Break**	03/15/2024
Checkpoint 2	**Instructor Consultation:** Schedule a progress check meeting with the supervising instructor. Actively incorporate any feedback or suggestions received during the consultation into your work. **Paper:** Finish the Results section	03/22/2024
Analysis	**Algorithm Comparison:** Conduct a detailed analysis of each hashing algorithm, focusing on their performance in the context of space-time tradeoffs. Assess how efficiently each algorithm manages memory and processing time. **Data Visualization:** Utilize graphs or charts to effectively illustrate the space-time trade-offs. Ensure that these visualizations are clear, accurate, and aid in understanding the comparative analysis.	03/29/2024
Discussion	**Architectural Impact Assessment:** Examine the effects of various computing architectures on the performance of each algorithm. This evaluation should highlight how different hardware configurations influence algorithm efficiency. **Paper:** Finish the Discussion section	04/05/2024
Final Research	**Conclusion Development:** Begin crafting the conclusion section of the paper. Concurrently, engage in post-results research to deepen your understanding of the observed outcomes. Aim to draw meaningful connections between the results and the broader context of your research topic.	04/12/2024
Conclusion	**Paper:** Finish the Conclusion section	04/19/2024
Wrap-up	**End of Course:** All tasks and assignments pertaining to CS-370 should be fully completed by this time. **Paper:** Finish Revisions	04/26/2024
End	**Final Class Day**	04/29/2024

Table 2: Project Timeline