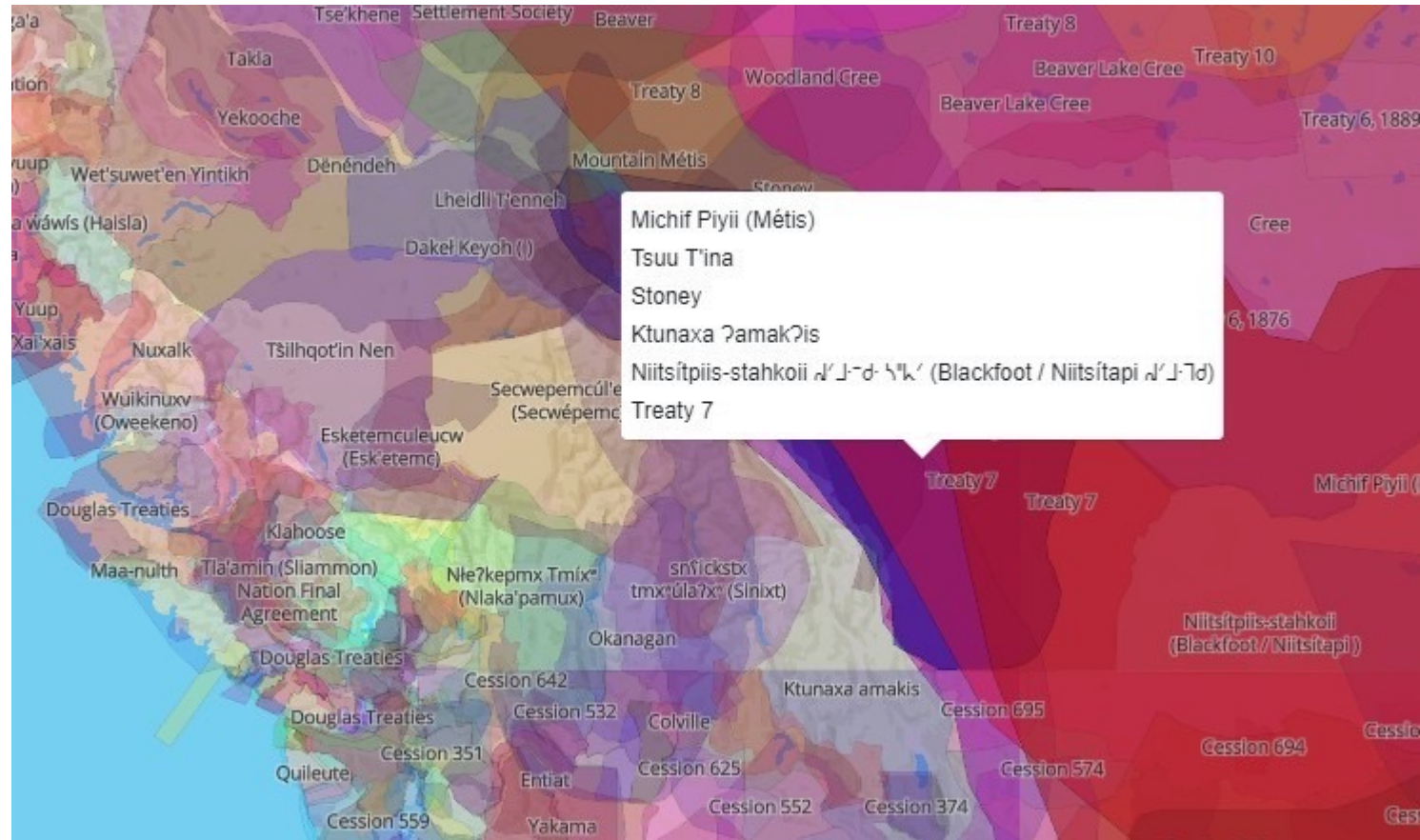


COMP 4522: Advanced DBs

Winter 2024

Dr. Orestes Appel

Land Acknowledgement



Source: <https://native-land.ca>

Who am I? Or better, what do I do?

Name: Orestes Appel (worked for IBM Canada until Summer 2023)

Background: Math & Computing

- PhD Artificial Intelligence, De Montfort-Leicester University, UK
- MSc. Intelligent Systems, De Montfort-Leicester University, UK
- M.S. Computer Science, Indiana University – Bloomington, USA
- BSc. Mathematics & Computing, UCV, Venezuela

Pronouns: He/Him

Call me: Professor, or Orestes

Office: B175 C5

Email: oappel@mtroyal.ca

Who are you?

- A. What is your major?
- B. Are you SQL-fluent?
- C. Experience with programming language(s):
 - **Python**
 - Java
 - C/C++
 - Lisp/Scheme
 - Haskell
 - Other(s)
- D. What do you expect from this class?
- E. Have you worked with Database Systems before?
 - Do you know how to create a DB in a RDBMS?
 - Do you know how to query the data in the database?

Why database systems came to be?

Needed a mechanism to efficiently store, retrieve and manipulate data. That implied:

- Inventing models
- Inventing access techniques
- Translating the above into proper physical access using the best data structures available
- Ensure data integrity and consistency
- High availability

“Magic's just science that we don't understand yet.”

— Arthur C. Clarke

Goals of this course

- Relational Databases: Relational Algebra, Set Theory & Normalization
- RDBMS Transactions Integrity
- Data Warehousing
- Fundamentals of Data Mining
- Deductive Databases (Databases based on a subset of First-Order Logic)
- Distributed Database: issues and models
- RDBMSs limitations and alternatives
- Columnar and in-memory databases
- Non-Relational Database Models: No-SQL (*Key:Value*, etc.), Big Data Map/Reduce
- Extended relational model and its strengths by preserving ACID properties.
- Case Study: MongoDB – A cross-platform document-oriented DBMS
- Physical Storage Models for Persistence
- RDBMS Query Optimization

Challenges and tips

Challenges

- Lots of new terminology!
- Lots of new tools – Datalog, Prolog, Mongo, etc
- **Low-level** programming
- Probably new languages (Python?)

Tips for success

- Take notes - writing in your own words helps to reinforce concepts
- Tell me when something doesn't make sense, you probably aren't alone
- **Start your assignments early**
- Read material in a timely fashion and come to class **prepared**

Course Evaluation

Grade breakdown

Midterm	15%
Final Exam	20%
Assignments (A1=20, A2=15, A3=20)	55%
Readings	10%
Total	100%

Assignments
Individual or in
Group of up to 4

Covid stuff and more

- No food in classes, please
- If you are sick, **stay home**
- If I get sick, sessions may be conducted on-line

For more info, see [Current Health Measures on Campus](#)

Academic Integrity

- As deadlines start piling up, it can be tempting to copy an assignment
- Both **copying** and **allowing your work to be copied** are considered academic misconduct and will be penalized
- If you use an internet resource (e.g. [Stack Overflow](#)), **cite it**
 - just drop the link in your code as a comment
- ChatGPT may not be your best friend after all

Textbook(s): None

- Jeff Hoffer, Ramesh Venkataraman and Heikki Topi. **Modern Database Management**
- We will provide chapters from books and /or articles, as needed
- Notes for this class will be distributed as the semester progresses
- Other books:
 - Database Systems: The Complete Book by Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom.
 - Fundamentals of Database Management Systems by Mark L. Gillenson.

Other resources

Course Resources

- [D2L](#): Grades, announcements [Course Website](#) -
- Lecture notes, lab instructions, tutorial problems, etc.

Computer stuff

- A computer running Windows (XP+), macOS (10.15+), or Linux Tablets or
 - pre-2010 computers will not work
 - Administrative rights
- High speed internet and an audio connection (video nice to have)
- Installed recent version of Google Chrome, MS Edge, Safari or Firefox browser
- A [GitHub](#) account is recommended
- Please [email me](#) describing the problem if any of this is an issue

Third-party Software

GitHub is a third-party service where you will be asked to enter your name and email address. When using a third-party software tool you are **consenting to the disclosure of your personal information** to the external third-party vendor; which falls outside the custody and control of Mount Royal University; therefore, limited identifiable personal information should be entered into the software to protect personal privacy.

Alternatively, students have the right to **opt out of using the software** and may choose another form of academic participation subject to the agreement of the instructor.

Students must **notify their instructor of their intention to opt-out** within the first **two weeks** of the semester.

Next Class

Lecture:

- Relational Data Model – Summary
- Relational Algebra