Introduction to Data Management
CSE 344

Unit 7: Transactions
Schedules
Implementation
Two-phase Locking

(3-4 lectures)
Class Overview

• Unit 1: Intro
• Unit 2: Relational Data Models and Query Languages
• Unit 3: Non-relational data
• Unit 4: RDMBS internals and query optimization
• Unit 5: Parallel query processing
• Unit 6: DBMS usability, conceptual design

• Unit 7: Transactions
  – Writing DB applications
  – Locking and schedules
Data Management Pipeline

Transactions
- Application programmer

Schema designer
- name
- product
- price

Conceptual Schema

Transactions
- Database administrator

Physical Schema
Transactions

• We use database transactions everyday
  – Bank $$$ transfers
  – Online shopping
  – Signing up for classes

• Applications that talk to a DB **must** use transactions in order to keep the database consistent.
What’s the big deal?
Challenges

• Suppose we only serve one app at a time
  – No problem…

• Suppose we execute apps concurrently
  – What’s the problem?

• Want: multiple operations to be executed *atomically* over the same DBMS
What can go wrong?

• Manager: balance budgets among projects
  – Remove $10k from project A
  – Add $7k to project B
  – Add $3k to project C

• CEO: check company’s total balance
  – SELECT SUM(money) FROM budget;

• This is called a dirty / inconsistent read aka a WRITE-READ conflict
What can go wrong?

- App 1:  
  ```sql
  SELECT inventory FROM products
  WHERE pid = 1
  ```

- App 2:  
  ```sql
  UPDATE products SET inventory = 0
  WHERE pid = 1
  ```

- App 1:  
  ```sql
  SELECT inventory * price FROM products
  WHERE pid = 1
  ```

- This is known as an unrepeatable read aka READ-WRITE conflict
What can go wrong?

Account 1 = $100
Account 2 = $100
Total = $200

• App 1:
  – Set Account 1 = $200
  – Set Account 2 = $0

• App 2:
  – Set Account 2 = $200
  – Set Account 1 = $0

• At the end:
  – Total = $200

• App 1: Set Account 1 = $200

• App 2: Set Account 2 = $200

• App 1: Set Account 2 = $0

• App 2: Set Account 1 = $0

• At the end:
  – Total = $0

This is called the lost update aka WRITE-WRITE conflict
What can go wrong?

• Buying tickets to the next Bieber concert:
  – Fill up form with your mailing address
  – Put in debit card number
  – Click submit
  – Screen shows money deducted from your account
  – [Your browser crashes]

Lesson:
Changes to the database should be **ALL or NOTHING**
Transactions

• Collection of statements that are executed atomically (logically speaking)

BEGIN TRANSACTION
  [SQL statements]
COMMIT  or  ROLLBACK (=ABORT)

If BEGIN... missing, then TXN consists of a single instruction
Know your chemistry transactions: ACID

- **Atomic**
  - State shows either all the effects of txn, or none of them

- **Consistent**
  - Txn moves from a DBMS state where integrity holds, to another where integrity holds
    - remember integrity constraints?

- **Isolated**
  - Effect of txns is the same as txns running one after another (i.e., looks like batch mode)

- **Durable**
  - Once a txn has committed, its effects remain in the database
Atomic

- **Definition**: A transaction is ATOMIC if all its updates must happen or not at all.

```sql
-- Example: move $100 from A to B:
BEGIN TRANSACTION;
  UPDATE accounts SET bal = bal - 100 WHERE acct = A;
  UPDATE accounts SET bal = bal + 100 WHERE acct = B;
COMMIT;
```
**Isolated**

- **Definition** An execution ensures that txns are isolated, if the effect of each txn is as if it were the only txn running on the system.

---

**App 1:**
```
BEGIN TRANSACTION;

SELECT inventory
FROM products
WHERE pid = 1;

SELECT inventory * price
FROM products
WHERE pid = 1;

COMMIT
```

**App 2:**
```
BEGIN TRANSACTION;

UPDATE products
SET inventory = 0
WHERE pid = 1;

COMMIT;
```
Consistent

• Recall: integrity constraints govern how values in tables are related to each other
  – Can be enforced by the DBMS, or ensured by the app

• How consistency is achieved by the app:
  – App programmer ensures that txns only takes a consistent DB state to another consistent state
  – DB makes sure that txns are atomic and isolated

• Can defer checking the validity of constraints until the end of a transaction
Durable

• A transaction is durable if its effects continue to exist after the transaction and even after the program has terminated

• How?
  – By writing to disk!
  – More in 444
Rollback transactions

• If the app gets to a state where it cannot complete the transaction successfully, execute ROLLBACK

• The DB returns to the state prior to the transaction

• What are examples of such program states?
ACID

• Atomic
• Consistent
• Isolated
• Durable

• Enjoy this in HW8!

• Again: by default each statement is its own tx
  – Unless auto-commit is off then each statement starts a new tx
Implementing Transactions

Need to address two problems:

• ”I” – Isolation:
  – Means concurrency control
  – We will discuss this

• “A” – Atomicity:
  – Means recover from crash
  – We will not discuss this (see 444)
Transactions Demo
Transaction Schedules
Modeling a Transaction

• Database = a collection of elements
  – An element can be a record (logical elements)
  – Or can be a disc block (physical element)

Database: A B C D ...

• Transaction = sequence of read/writes of elements
Schedules

A **schedule** is a sequence of interleaved actions from all transactions.
Serial Schedule

• A *serial schedule* is one in which transactions are executed one after the other, in some sequential order.

• **Fact:** nothing can go wrong if the system executes transactions serially.

• But DBMS don’t do that because we want better overall system performance.
### Example

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ(A, t)</td>
<td>READ(A, s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(A, t)</td>
<td>WRITE(A,s)</td>
</tr>
<tr>
<td>READ(B, t)</td>
<td>READ(B,s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(B,t)</td>
<td>WRITE(B,s)</td>
</tr>
</tbody>
</table>

A and B are elements in the database, t and s are variables in txn source code.
### Example of a (Serial) Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>READ(A, t)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( t := t + 100 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WRITE(A, t)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>READ(B, t)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( t := t + 100 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WRITE(B, t)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>READ(A, s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( s := s \times 2 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WRITE(A, s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>READ(B, s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( s := s \times 2 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WRITE(B, s)</td>
<td></td>
</tr>
</tbody>
</table>
Another Serial Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ(A, s)</td>
<td>READ(A, s)</td>
</tr>
<tr>
<td>s := s*2</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(A, s)</td>
<td>WRITE(A, s)</td>
</tr>
<tr>
<td>READ(B, s)</td>
<td>READ(B, s)</td>
</tr>
<tr>
<td>s := s*2</td>
<td>WRITE(B, s)</td>
</tr>
<tr>
<td>READ(A, t)</td>
<td></td>
</tr>
<tr>
<td>t := t+100</td>
<td></td>
</tr>
<tr>
<td>WRITE(A, t)</td>
<td></td>
</tr>
<tr>
<td>READ(B, t)</td>
<td></td>
</tr>
<tr>
<td>t := t+100</td>
<td></td>
</tr>
<tr>
<td>WRITE(B, t)</td>
<td></td>
</tr>
</tbody>
</table>

Time
A schedule is *serializable* if it is equivalent to a serial schedule.
A Serializable Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ(A, t)</td>
<td>READ(A, s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(A, t)</td>
<td>WRITE(A, s)</td>
</tr>
<tr>
<td>READ(B, t)</td>
<td>READ(B, s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(B, t)</td>
<td>WRITE(B, s)</td>
</tr>
</tbody>
</table>

This is a **serializable** schedule.

This is **NOT** a serial schedule.
A Non-Serializable Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ(A, t)</td>
<td>READ(A, s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(A, t)</td>
<td>WRITE(A, s)</td>
</tr>
<tr>
<td>READ(B, t)</td>
<td>READ(B, s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s*2</td>
</tr>
<tr>
<td>WRITE(B, t)</td>
<td>WRITE(B, s)</td>
</tr>
</tbody>
</table>
How do We Know if a Schedule is Serializable?

Notation:

$T_1: r_1(A); w_1(A); r_1(B); w_1(B)$

$T_2: r_2(A); w_2(A); r_2(B); w_2(B)$

Key Idea: Focus on conflicting operations
Conflicts

- Write-Read – WR
- Read-Write – RW
- Write-Write – WW
Conflict Serializability

Conflicts: (i.e., swapping will change program behavior)

Two actions by same transaction $T_i$:

Two writes by $T_i$, $T_j$ to same element

Read/write by $T_i$, $T_j$ to same element
Conflict Serializability

- A schedule is *conflict serializable* if it can be transformed into a serial schedule by a series of swappings of adjacent non-conflicting actions.

- Every conflict-serializable schedule is serializable.
- The converse is not true (why?)
Conflict Serializability

Example:

\[ r_1(A); w_1(A); r_2(A); w_2(A); r_1(B); w_1(B); r_2(B); w_2(B) \]
Conflict Serializability

Example:

\[
\begin{align*}
& r_1(A); w_1(A); r_2(A); w_2(A); r_1(B); w_1(B); r_2(B); w_2(B) \\
\end{align*}
\]
Conflict Serializability

Example:

\[ r_1(A); w_1(A); r_2(A); w_2(A); r_1(B); w_1(B); r_2(B); w_2(B) \]
Conflict Serializability

Example:

\[
\begin{align*}
\text{r}_1(\text{A}); \text{w}_1(\text{A}); \text{r}_2(\text{A}); &\quad \text{w}_2(\text{A}); \text{r}_1(\text{B}); &\quad \text{w}_1(\text{B}); \text{r}_2(\text{B}); \text{w}_2(\text{B}) \\
\text{r}_1(\text{A}); \text{w}_1(\text{A}); \text{r}_2(\text{A}); &\quad \text{r}_1(\text{B}); &\quad \text{w}_2(\text{A}); \text{w}_1(\text{B}); \text{r}_2(\text{B}); \text{w}_2(\text{B}) \\
\text{r}_1(\text{A}); \text{w}_1(\text{A}); \text{r}_1(\text{B}); &\quad \text{w}_1(\text{B}); \text{r}_2(\text{A}); &\quad \text{w}_2(\text{A}); \text{r}_2(\text{B}); \text{w}_2(\text{B})
\end{align*}
\]
Conflict Serializability

Example:

```
r_1(A); w_1(A); r_2(A); w_2(A); r_1(B); w_1(B); r_2(B); w_2(B)
```

```
r_1(A); w_1(A); r_2(A); r_1(B); w_2(A); w_1(B); r_2(B); w_2(B)
```

```
r_1(A); w_1(A); r_1(B); r_2(A); w_2(A); w_1(B); r_2(B); w_2(B)
```

```
r_1(A); w_1(A); r_1(B); w_1(B); r_2(A); w_2(A); r_2(B); w_2(B)
```

....
Serializable, Not Conflict-Serializable

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ(A, t)</td>
<td>READ(A,s)</td>
</tr>
<tr>
<td>t := t+100</td>
<td>s := s + 200</td>
</tr>
<tr>
<td>WRITE(A, t)</td>
<td>WRITE(A,s)</td>
</tr>
<tr>
<td></td>
<td>READ(B,s)</td>
</tr>
<tr>
<td></td>
<td>s := s + 200</td>
</tr>
<tr>
<td></td>
<td>WRITE(B,s)</td>
</tr>
<tr>
<td></td>
<td>READ(B, t)</td>
</tr>
<tr>
<td></td>
<td>t := t+100</td>
</tr>
<tr>
<td></td>
<td>WRITE(B,t)</td>
</tr>
</tbody>
</table>
Testing for Conflict-Serializability

Precedence graph:
• A node for each transaction $T_i$,
• An edge from $T_i$ to $T_j$ whenever an action in $T_i$ conflicts with, and comes before an action in $T_j$

• The schedule is conflict-serializable iff the precedence graph is acyclic
Example 1

\[ r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B) \]
Example 1

\[ r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B) \]
Example 1

\[ r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B) \]
Example 1

\[ r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B) \]
Example 1

\[ r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B) \]
Example 1

This schedule is conflict-serializable
Example 2

\[ r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B) \]
Example 2

\[ r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B) \]
Example 2

$\text{r}_2(\text{A}); \text{r}_1(\text{B}); \text{w}_2(\text{A}); \text{r}_2(\text{B}); \text{r}_3(\text{A}); \text{w}_1(\text{B}); \text{w}_3(\text{A}); \text{w}_2(\text{B})$
Example 2

\[ r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B) \]
Example 2

r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B)
Example 2

\( r_2(A); \ r_1(B); \ w_2(A); \ r_2(B); \ r_3(A); \ w_1(B); \ w_3(A); \ w_2(B) \)
Example 2

$r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B)$
Example 2

This schedule is NOT conflict-serializable
Implementing Transactions
Scheduler

• **Scheduler a.k.a. Concurrency Control Manager**
  – The module that schedules the transaction’s actions
  – Goal: ensure the schedule is serializable

• We discuss next how a scheduler may be implemented
Implementing a Scheduler

Two major approaches:

- **Locking Scheduler**
  - Aka “pessimistic concurrency control”
  - SQLite, SQL Server, DB2

- **Multiversion Concurrency Control (MVCC)**
  - Aka “optimistic concurrency control”
  - Postgres, Oracle: Snapshot Isolation (SI)

We discuss only locking schedulers in this class
Lock-based Implementation of Transactions
Locking Scheduler

Simple idea:
- Each element has a unique lock
- Each transaction must first acquire the lock before reading/writing that element
- If the lock is taken by another transaction, then wait
- The transaction must release the lock(s)

By using locks scheduler ensures conflict-serializability
What Data Elements are Locked?

Major differences between vendors:

• Lock on the entire database
  – SQLite

• Lock on individual records (“elements”)
  – SQL Server, DB2, etc
Actions on Locks

$L_i(A)$ = transaction $T_i$ acquires lock for element A

$U_i(A)$ = transaction $T_i$ releases lock for element A

Let’s see this in action…
A Non-Serializable Schedule

T1
READ(A)
A := A + 100
WRITE(A)

T2
READ(A)
A := A * 2
WRITE(A)
READ(B)
B := B * 2
WRITE(B)

READ(B)
B := B + 100
WRITE(B)
Example

T1

L₁(A); READ(A)
A := A + 100
WRITE(A); U₁(A); L₁(B)

READ(B)
B := B + 100
WRITE(B); U₁(B);

T2

L₂(A); READ(A)
A := A * 2
WRITE(A); U₂(A);
L₂(B); BLOCKED…

…GRANTED; READ(B)
B := B * 2
WRITE(B); U₂(B);

Scheduler has ensured a conflict-serializable schedule
But…

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₁(A); READ(A)</td>
<td>L₂(A); READ(A)</td>
</tr>
<tr>
<td>A := A+100</td>
<td>A := A*2</td>
</tr>
<tr>
<td>WRITE(A); U₁(A);</td>
<td>WRITE(A); U₂(A);</td>
</tr>
<tr>
<td></td>
<td>L₂(B); READ(B)</td>
</tr>
<tr>
<td></td>
<td>B := B*2</td>
</tr>
<tr>
<td></td>
<td>WRITE(B); U₂(B);</td>
</tr>
<tr>
<td>L₁(B); READ(B)</td>
<td></td>
</tr>
<tr>
<td>B := B+100</td>
<td></td>
</tr>
<tr>
<td>WRITE(B); U₁(B);</td>
<td></td>
</tr>
</tbody>
</table>

Locks did not enforce conflict-serializability !!! What’s wrong ?
Two Phase Locking (2PL)

The 2PL rule:

In every transaction, all lock requests must precede all unlock requests
### Example: 2PL transactions

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1(A); L_1(B); \text{READ}(A)$</td>
<td></td>
</tr>
<tr>
<td>$A := A + 100$</td>
<td>$L_2(A); \text{READ}(A)$</td>
</tr>
<tr>
<td>WRITE($A$); $U_1(A)$</td>
<td></td>
</tr>
<tr>
<td>READ($B$)</td>
<td>$A := A \times 2$</td>
</tr>
<tr>
<td>$B := B + 100$</td>
<td>WRITE($A$);</td>
</tr>
<tr>
<td>WRITE($B$); $U_1(B)$</td>
<td>$L_2(B); \text{BLOCKED...}$</td>
</tr>
<tr>
<td></td>
<td>$\ldots\text{GRANTED;} \text{READ}(B)$</td>
</tr>
<tr>
<td></td>
<td>$B := B \times 2$</td>
</tr>
<tr>
<td></td>
<td>WRITE($B$); $U_2(A); U_2(B);$</td>
</tr>
</tbody>
</table>

Now it is conflict-serializable
Two Phase Locking (2PL)

**Theorem**: 2PL ensures conflict serializability
Two Phase Locking (2PL)

**Theorem:** 2PL ensures conflict serializability

**Proof.** Suppose not: then there exists a cycle in the precedence graph.
Two Phase Locking (2PL)

**Theorem:** 2PL ensures conflict serializability

**Proof.** Suppose not: then there exists a cycle in the precedence graph.

Then there is the following **temporal** cycle in the schedule:

![Diagram](image)
Theorem: 2PL ensures conflict serializability

Proof. Suppose not: then there exists a cycle in the precedence graph.

Then there is the following **temporal** cycle in the schedule: $U_1(A) \Rightarrow L_2(A)$  why?

$U_1(A)$ happened strictly before $L_2(A)$.
Two Phase Locking (2PL)

**Theorem:** 2PL ensures conflict serializability

**Proof.** Suppose not: then there exists a cycle in the precedence graph.

Then there is the following *temporal* cycle in the schedule: $U_1(A) \rightarrow L_2(A)$why?
Two Phase Locking (2PL)

**Theorem:** 2PL ensures conflict serializability

**Proof.** Suppose not: then there exists a cycle in the precedence graph.

Then there is the following **temporal** cycle in the schedule:

\[ U_1(A) \rightarrow L_2(A) \]
\[ L_2(A) \rightarrow U_2(B) \]

why?

L₂(A) happened strictly before U₁(A)
Two Phase Locking (2PL)

**Theorem:** 2PL ensures conflict serializability

**Proof.** Suppose not: then there exists a cycle in the precedence graph.

![Diagram](image.png)

Then there is the following **temporal** cycle in the schedule:

\[ U_1(A) \rightarrow L_2(A) \]
\[ L_2(A) \rightarrow U_2(B) \]

why?
Two Phase Locking (2PL)

**Theorem**: 2PL ensures conflict serializability

**Proof**: Suppose not: then there exists a cycle in the precedence graph. Then there is the following **temporal** cycle in the schedule:

\[ U_1(A) \rightarrow L_2(A) \]
\[ L_2(A) \rightarrow U_2(B) \]
\[ U_2(B) \rightarrow L_3(B) \]

why?
Two Phase Locking (2PL)

**Theorem:** 2PL ensures conflict serializability

**Proof.** Suppose not: then there exists a cycle in the precedence graph.

Then there is the following **temporal** cycle in the schedule:

- $U_1(A) \rightarrow L_2(A)$
- $L_2(A) \rightarrow U_2(B)$
- $U_2(B) \rightarrow L_3(B)$

......etc......
Theorem: 2PL ensures conflict serializability

Proof. Suppose not: then there exists a cycle in the precedence graph.

Then there is the following temporal cycle in the schedule:

\[ \begin{align*}
U_1(A) & \rightarrow L_2(A) \\
L_2(A) & \rightarrow U_2(B) \\
U_2(B) & \rightarrow L_3(B) \\
L_3(B) & \rightarrow U_3(C) \\
U_3(C) & \rightarrow L_1(C) \\
L_1(C) & \rightarrow U_1(A)
\end{align*} \]

Cycle in time: Contradiction
## A New Problem: Non-recoverable Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₁(A); L₁(B); READ(A)</td>
<td>L₂(A); READ(A)</td>
</tr>
<tr>
<td>A := A + 100</td>
<td>A := A * 2</td>
</tr>
<tr>
<td>WRITE(A); U₁(A)</td>
<td>WRITE(A);</td>
</tr>
<tr>
<td></td>
<td>L₂(B); BLOCKED…</td>
</tr>
<tr>
<td>READ(B)</td>
<td>…GRANTED; READ(B)</td>
</tr>
<tr>
<td>B := B + 100</td>
<td>B := B * 2</td>
</tr>
<tr>
<td>WRITE(B); U₁(B);</td>
<td>WRITE(B); U₂(A); U₂(B);</td>
</tr>
<tr>
<td></td>
<td>Commit</td>
</tr>
</tbody>
</table>

Rollback
## A New Problem: Non-recoverable Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₁(A); L₁(B); READ(A)</td>
<td>L₂(A); READ(A)</td>
</tr>
<tr>
<td>A := A + 100</td>
<td>A := A * 2</td>
</tr>
<tr>
<td>WRITE(A); U₁(A)</td>
<td>WRITE(A);</td>
</tr>
<tr>
<td>READ(B)</td>
<td>L₂(B); BLOCKED…</td>
</tr>
<tr>
<td>B := B + 100</td>
<td>…GRANTED; READ(B)</td>
</tr>
<tr>
<td>WRITE(B); U₁(B);</td>
<td>B := B * 2</td>
</tr>
<tr>
<td></td>
<td>WRITE(B); U₂(A); U₂(B);</td>
</tr>
<tr>
<td></td>
<td>Commit</td>
</tr>
</tbody>
</table>

Elements A, B written by T1 are restored to their original value.
# A New Problem: Non-recoverable Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_1(A); L_1(B); ) ( \text{READ}(A) )</td>
<td>( L_2(A); ) ( \text{READ}(A) )</td>
</tr>
<tr>
<td>( A := A + 100 )</td>
<td>( A := A \times 2 )</td>
</tr>
<tr>
<td>( \text{WRITE}(A); ) ( U_1(A) )</td>
<td>( \text{WRITE}(A); )</td>
</tr>
<tr>
<td>( \text{READ}(B) )</td>
<td>( L_2(B); ) BLOCKED…</td>
</tr>
<tr>
<td>( B := B + 100 )</td>
<td>( \ldots \text{GRANTED}; ) ( \text{READ}(B) )</td>
</tr>
<tr>
<td>( \text{WRITE}(B); ) ( U_1(B) )</td>
<td>( B := B \times 2 )</td>
</tr>
<tr>
<td>( \text{Rollback} )</td>
<td>( \text{WRITE}(B); ) ( U_2(A); ) ( U_2(B) ); Commit</td>
</tr>
</tbody>
</table>

Elements A, B written by T1 are restored to their original value.

Dirty reads of A, B lead to incorrect writes.
## A New Problem: Non-recoverable Schedule

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₁(A); L₁(B); READ(A)</td>
<td>L₂(A); READ(A)</td>
</tr>
<tr>
<td>A := A + 100</td>
<td>A := A * 2</td>
</tr>
<tr>
<td>WRITE(A); U₁(A)</td>
<td>WRITE(A);</td>
</tr>
<tr>
<td>READ(B)</td>
<td>L₂(B); BLOCKED…</td>
</tr>
<tr>
<td>B := B + 100</td>
<td>…GRANTED; READ(B)</td>
</tr>
<tr>
<td>WRITE(B); U₁(B)</td>
<td>B := B * 2</td>
</tr>
<tr>
<td></td>
<td>WRITE(B); U₂(A); U₂(B);</td>
</tr>
</tbody>
</table>

Elements A, B written by T1 are restored to their original value.

Can no longer undo!

Dirty reads of A, B lead to incorrect writes.
Strict 2PL

The Strict 2PL rule:

All locks are held until commit/abort:
All unlocks are done together with commit/abort.

With strict 2PL, we will get schedules that are both conflict-serializable and recoverable.
## Strict 2PL

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₁(A); READ(A)</td>
<td></td>
</tr>
<tr>
<td>A := A + 100</td>
<td></td>
</tr>
<tr>
<td>WRITE(A);</td>
<td></td>
</tr>
<tr>
<td>L₁(B); READ(B)</td>
<td></td>
</tr>
<tr>
<td>B := B + 100</td>
<td></td>
</tr>
<tr>
<td>WRITE(B);</td>
<td>L₂(A); BLOCKED…</td>
</tr>
<tr>
<td>Rollback &amp; U₁(A); U₁(B);</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>…GRANTED; READ(A)</td>
</tr>
<tr>
<td>A := A * 2</td>
<td></td>
</tr>
<tr>
<td>WRITE(A);</td>
<td></td>
</tr>
<tr>
<td>L₂(B); READ(B)</td>
<td></td>
</tr>
<tr>
<td>B := B * 2</td>
<td></td>
</tr>
<tr>
<td>WRITE(B);</td>
<td>Commit &amp; U₂(A); U₂(B);</td>
</tr>
</tbody>
</table>
Strict 2PL

• Lock-based systems always use strict 2PL

• Easy to implement:
  – Before a transaction reads or writes an element A, insert an L(A)
  – When the transaction commits/aborts, then release all locks

• Ensures both conflict serializability and recoverability
Another problem: Deadlocks

• $T_1$: $R(A), W(B)$
• $T_2$: $R(B), W(A)$

• $T_1$ holds the lock on $A$, waits for $B$
• $T_2$ holds the lock on $B$, waits for $A$

This is a deadlock!
Another problem: Deadlocks

To detect a deadlocks, search for a cycle in the waits-for graph:

- $T_1$ waits for a lock held by $T_2$;
- $T_2$ waits for a lock held by $T_3$;
- \ldots
- $T_n$ waits for a lock held by $T_1$

Relatively expensive: check periodically, if deadlock is found, then abort one TXN; re-check for deadlock more often (why?)
Lock Modes

- **S** = shared lock (for READ)
- **X** = exclusive lock (for WRITE)

Lock compatibility matrix:

```
None       S       X
None       None   None   None
S          None   None   None
X          None   None   None
```
## Lock Modes

- **S** = shared lock (for READ)
- **X** = exclusive lock (for WRITE)

### Lock Compatibility Matrix:

<table>
<thead>
<tr>
<th></th>
<th>None</th>
<th>S</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Lock Granularity

• **Fine granularity locking** (e.g., tuples)
  – High concurrency
  – High overhead in managing locks
  – E.g., SQL Server

• **Coarse grain locking** (e.g., tables, entire database)
  – Many false conflicts
  – Less overhead in managing locks
  – E.g., SQL Lite

• **Solution:** lock escalation changes granularity as needed
Throughput (TPS) vs. # Active Transactions

Lock Performance

TPS = Transactions per second

To avoid thrashing, use admission control

Why?
Announcement

Final review

• Saturday, 2pm,

• GWN 301
Phantom Problem

- So far we have assumed the database to be a static collection of elements (=tuples)

- If tuples are inserted/deleted then the *phantom problem* appears
Suppose there are two blue products, A1, A2:

**Phantom Problem**

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELECT *</td>
<td>INSERT INTO Product(name, color)</td>
</tr>
<tr>
<td>FROM Product</td>
<td>VALUES (‘A3’,’blue’)</td>
</tr>
<tr>
<td>WHERE color=‘blue’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SELECT *</td>
</tr>
<tr>
<td></td>
<td>FROM Product</td>
</tr>
<tr>
<td></td>
<td>WHERE color=‘blue’</td>
</tr>
</tbody>
</table>

Is this schedule serializable?
Suppose there are two blue products, A1, A2:

Phantom Problem

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELECT *</td>
<td>INSERT INTO Product(name, color) VALUES ('A3', 'blue')</td>
</tr>
<tr>
<td>FROM Product</td>
<td></td>
</tr>
<tr>
<td>WHERE color='blue'</td>
<td></td>
</tr>
<tr>
<td>SELECT *</td>
<td></td>
</tr>
<tr>
<td>FROM Product</td>
<td></td>
</tr>
<tr>
<td>WHERE color='blue'</td>
<td></td>
</tr>
</tbody>
</table>

Is this schedule serializable?

No: T1 sees a “phantom” product A3
Suppose there are two blue products, A1, A2:

Phantom Problem

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELECT *</td>
<td>INSERT INTO Product(name, color) VALUES (‘A3’,’blue’)</td>
</tr>
<tr>
<td>FROM Product</td>
<td></td>
</tr>
<tr>
<td>WHERE color=‘blue’</td>
<td></td>
</tr>
</tbody>
</table>

R₁(A₁); R₁(A₂); W₂(A₃); R₁(A₁); R₁(A₂); R₁(A₃)
Suppose there are two blue products, A1, A2:

**Phantom Problem**

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>**SELECT ** *</td>
<td><strong>INSERT INTO</strong> Product(name, color)</td>
</tr>
<tr>
<td><strong>FROM Product</strong></td>
<td><strong>VALUES</strong> (‘A3’, ’blue’)</td>
</tr>
<tr>
<td><strong>WHERE color=’blue’</strong></td>
<td></td>
</tr>
<tr>
<td>**SELECT ** *</td>
<td></td>
</tr>
<tr>
<td><strong>FROM Product</strong></td>
<td></td>
</tr>
<tr>
<td><strong>WHERE color=’blue’</strong></td>
<td></td>
</tr>
</tbody>
</table>

*R_1(A1); R_1(A2); W_2(A3); R_1(A1); R_1(A2); R_1(A3)*

*W_2(A3); R_1(A1); R_1(A2); R_1(A1); R_1(A2); R_1(A3)*
Suppose there are two blue products, A1, A2:

**Phantom Problem**

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELECT *</td>
<td>INSERT INTO Product(name, color) VALUES ('A3','blue')</td>
</tr>
<tr>
<td>FROM Product</td>
<td></td>
</tr>
<tr>
<td>WHERE color='blue'</td>
<td></td>
</tr>
<tr>
<td>SELECT *</td>
<td></td>
</tr>
<tr>
<td>FROM Product</td>
<td></td>
</tr>
<tr>
<td>WHERE color='blue'</td>
<td></td>
</tr>
</tbody>
</table>

But this is conflict-serializable!

\[ R_1(A1); R_1(A2); W_2(A3); R_1(A1); R_1(A2); R_1(A3) \]

\[ W_2(A3); R_1(A1); R_1(A2); R_1(A1); R_1(A2); R_1(A3) \]
Phantom Problem

• A “phantom” is a tuple that is invisible during part of a transaction execution but not invisible during the entire execution.

• In our example:
  – T1: reads list of products
  – T2: inserts a new product
  – T1: re-reads: a new product appears!

• Conflict-serializability assumes DB is static
• When DB is dynamic then c-s is not serializable.
Dealing With Phantoms

- Lock the entire table
- Lock the index entry for ‘blue’
  - If index is available
- Or use predicate locks
  - A lock on an arbitrary predicate

Dealing with phantoms is expensive!
Summary of Serializability

• Serializable schedule = equivalent to a serial schedule
• (strict) 2PL guarantees conflict serializability
  – What is the difference?
• Static database:
  – Conflict serializability implies serializability
• Dynamic database:
  – Conflict serializability plus phantom management implies serializability
Weaker Isolation Levels

• Serializable are expensive to implement

• SQL allows the application to choose a more efficient implementation, which is not always serializable: weak isolation levels
Isolation Levels in SQL

1. “Dirty reads”
   SET TRANSACTION ISOLATION LEVEL READ UNCOMMITTED

2. “Committed reads”
   SET TRANSACTION ISOLATION LEVEL READ COMMITTED

3. “Repeatable reads”
   SET TRANSACTION ISOLATION LEVEL REPEATABLE READ

4. Serializable transactions
   SET TRANSACTION ISOLATION LEVEL SERIALIZABLE

ACID
Lost Update

Write-Write Conflict

\[ T_1: \text{READ}(A) \]
\[ T_1: A := A+5 \]
\[ T_1: \text{WRITE}(A) \]

\[ T_2: \text{READ}(A) \]
\[ T_2: A := A*1.3 \]
\[ T_2: \text{WRITE}(A) \]

Never allowed at any level
1. Isolation Level: Dirty Reads

- “Long duration” WRITE locks
  - Strict 2PL

- No READ locks
  - Read-only transactions are never delayed

Possible problems: dirty and inconsistent reads
1. Isolation Level: Dirty Reads

Write-Read Conflict

$T_1$: WRITE($A$)

$T_2$: READ($A$)

$T_1$: ABORT
1. Isolation Level: Dirty Reads

Write-Read Conflict

$T_1$: $A := 20; \ B := 20$

$T_1$: WRITE($A$)

$T_1$: WRITE($B$)

$T_2$: READ($A$);

$T_2$: READ($B$);

Inconsistent read
2. Isolation Level: Read Committed

- “Long duration” WRITE locks
  - Strict 2PL
- “Short duration” READ locks
  - Only acquire lock while reading (not 2PL)

Unrepeatable reads:
When reading same element twice, may get two different values
2. Isolation Level: Read Committed

Read-Write Conflict

$T_1$: WRITE(A)

$T_2$: READ(A);  

Unrepeatable read
3. Isolation Level: Repeatable Read

- "Long duration" WRITE locks
  - Strict 2PL
- "Long duration" READ locks
  - Strict 2PL

This is not serializable yet !!!
4. Isolation Level Serializable

- “Long duration” WRITE locks
  - Strict 2PL
- “Long duration” READ locks
  - Strict 2PL
- Predicate locking
  - To deal with phantoms
Beware!

In commercial DBMSs:
• Default level may not be serializable
• Default level differs between DBMSs
• Some engines support subset of levels!
• Also, some DBMSs do NOT use locking and different isolation levels can lead to different pbs

Bottom line: Read the doc for your DBMS!
Case Study: SQLite

- SQLite is very simple

- Lock types
  - READ LOCK (to read)
  - RESERVED LOCK (to write)
  - PENDING LOCK (wants to commit)
  - EXCLUSIVE LOCK (to commit)
Step 1: when a transaction begins

- Acquire a READ LOCK (aka "SHARED" lock)
- All these transactions may read happily
- They all read data from the database file
- If the transaction commits without writing anything, then it simply releases the lock
SQLite

Step 2: when one transaction wants to write

- Acquire a **RESERVED LOCK**
- May coexists with many READ LOCKs
- Writer TXN may write; these updates are only in main memory; others don't see the updates
- Reader TXN continue to read from the file
- New readers accepted
- No other TXN is allowed a RESERVED LOCK
SQLite

Step 3: when writer transaction wants to commit, it needs *exclusive lock*, which can’t coexists with *read locks*

- Acquire a **PENDING LOCK**
- May coexists with old READ LOCKs
- No new READ LOCKS are accepted
- Wait for all read locks to be released

Why not write to disk right now?
SQLite

Step 4: when all read locks have been released
- Acquire the **EXCLUSIVE LOCK**
- Nobody can touch the database now
- All updates are written permanently to the database file

- Release the lock and **COMMIT**
SQLite

begin transaction

None

READ LOCK

first write

RESERVED LOCK

commit requested

PENDING LOCK

no more read locks

EXCLUSIVE LOCK

commit

commit executed
SQLite Demo

create table r(a int, b int);
insert into r values (1,10);
insert into r values (2,20);
insert into r values (3,30);
Demonstrating Locking in SQLite

T1:
begin transaction;
select * from r;
-- T1 has a READ LOCK

T2:
begin transaction;
select * from r;
-- T2 has a READ LOCK
Demonstrating Locking in SQLite

T1:
  update r set b=11 where a=1;
  -- T1 has a RESERVED LOCK

T2:
  update r set b=21 where a=2;
  -- T2 asked for a RESERVED LOCK: DENIED
Demonstrating Locking in SQLite

T3:

begin transaction;
select * from r;
commit;

-- everything works fine, could obtain READ LOCK
Demonstrating Locking in SQLite

T1:

commit;

-- SQL error: database is locked
-- T1 asked for PENDING LOCK -- GRANTED
-- T1 asked for EXCLUSIVE LOCK -- DENIED
Demonstrating Locking in SQLite

T3:

begin transaction;
select * from r;
-- T3 asked for READ LOCK-- DENIED (due to T1)

T2:

commit;
-- releases the last READ LOCK; T1 can commit
How do anomalies show up in schedules?

• What could go wrong if we didn’t have concurrency control:
  – Dirty reads (including inconsistent reads)
  – Unrepeatable reads
  – Lost updates

Many other things can go wrong too
Demonstration with SQL Server

**Application 1:**
create table R(a int);
insert into R values(1);
set transaction isolation level *serializable*;
begin transaction;
select * from R; -- get a shared lock

**Application 2:**
set transaction isolation level *serializable*;
begin transaction;
select * from R; -- get a shared lock
insert into R values(2); -- blocked waiting on exclusive lock
-- App 2 unblocks and executes insert after app 1 commits/aborts
Demonstration with SQL Server

Application 1:
create table R(a int);
insert into R values(1);
set transaction isolation level repeatable read;
begi
select * from R; -- get a shared lock

Application 2:
set transaction isolation level repeatable read;
begi
select * from R; -- get a shared lock
insert into R values(3); -- gets an exclusive lock on new tuple
  -- If app 1 reads now, it blocks because read dirty
  -- If app 1 reads after app 2 commits, app 1 sees new value
Final Exam

1. Relational data model (SQL)
2. Semistructured data model (SQL++)
3. Datalog
4. RDBMS Internals (execution, optimization)
5. Parallel Query Processing
6. Conceptual Design
7. Transactions