Transaction Processing

CPS 116
Introduction to Database Systems

Announcements (November 27)

- Homework #4 due this Thursday
- Project demo period starts on December 7
  - Each project gets a 30-minute with me and Yi
  - Watch for an email this weekend scheduling demo slots
- Final exam on December 15

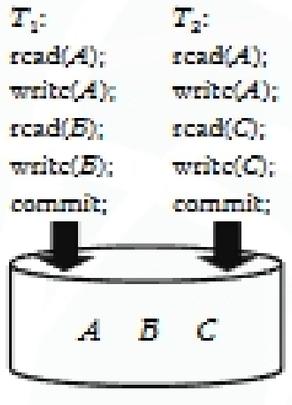
Review

- ACID
  - Atomicity: TX's are either completely done or not done at all
  - Consistency: TX's should leave the database in a consistent state
  - Isolation: TX's must behave as if they are executed in isolation
  - Durability: Effects of committed TX's are resilient against failures
- SQL transactions

```
-- Begins implicitly
SELECT …;
UPDATE …;
ROLLBACK | COMMIT;
```

Concurrency control

- Goal: ensure the "I" (isolation) in ACID



Good versus bad schedules

Good!

T_1	T_2
r(A)	
w(A)	
r(B)	
w(B)	
	r(A)
	w(A)
	r(C)
	w(C)

T_1	T_2
r(A)	
	r(A)
w(A)	
	w(A)
r(B)	
	r(C)
w(B)	
	w(C)

T_1	T_2
r(A)	
w(A)	
	r(A)
	w(A)
r(B)	
	r(C)
w(B)	
	w(C)

Serial schedule

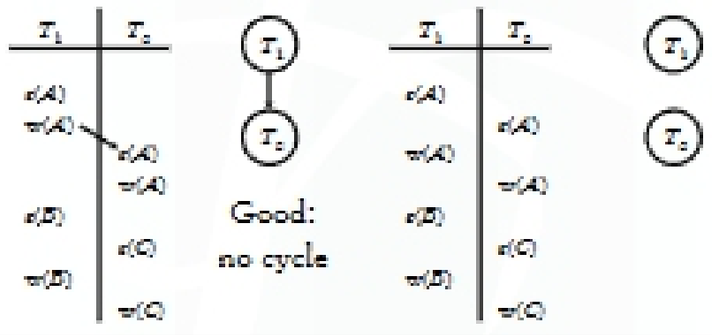
- Execute transactions in order, with no interleaving of operations
  - T_1.r(A), T_1.w(A), T_1.r(B), T_1.w(B), T_2.r(A), T_2.w(A), T_2.r(C), T_2.w(C)
  - T_2.r(A), T_2.w(A), T_2.r(C), T_2.w(C), T_1.r(A), T_1.w(A), T_1.r(B), T_1.w(B)
  - ☞Isolation achieved by definition!
- Problem: no concurrency at all
- Question: how to reorder operations to allow more concurrency

Conflicting operations

- Two operations on the same data item conflict if at least one of the operations is a write
  - $r(X)$ and $w(X)$ conflict
  - $w(X)$ and $r(X)$ conflict
  - $w(X)$ and $w(X)$ conflict
  - $r(X)$ and $r(X)$ do not
  - $r/w(X)$ and $r/w(Y)$ do not
- Order of conflicting operations matters
  - E.g., if $T_1.r(A)$ precedes $T_2.w(A)$, then conceptually, T_1 should precede T_2

Precedence graph

- A node for each transaction
- A directed edge from T_i to T_j if an operation of T_i precedes and conflicts with an operation of T_j in the schedule



Conflict-serializable schedule

- A schedule is conflict-serializable iff its precedence graph has no cycles
- A conflict-serializable schedule is equivalent to some serial schedule (and therefore is "good")
  - In that serial schedule, transactions are executed in the topological order of the precedence graph
  - You can get to that serial schedule by repeatedly swapping adjacent, non-conflicting operations from different transactions