



# CSCD43: Database Systems Technology

# Lecture 12

Wael Aboulsaadat

Acknowledgment: these slides are based on Prof. Garcia-Molina & Prof. Ullman slides accompanying the course's textbook.



## DBMS Architecture





# Component for Concurrency Control



# Transactions

- Historical note:
  - Turing Award for Transaction concept
  - Jim Gray (1998)
- Interesting reading:

Transaction Concept: Virtues and Limitations by Jim Gray http://www.hpl.hp.com/techreports/tandem/TR-81.3.pdf



# Issues with Transactions: Example Bank database: 3 Accounts



# Property: A + B + C = 1500

Money does not leave the system

Issues with Transactions: Example Transaction T1: Transfer 100 from A to B A = 500, B = 500, C = 500Read (A, t) t = t - 100Write (A, t) Read (B, t) t = t + 100Write (B, t) A = 400, B = 600, C = 500

Issues with Transactions: Example Transaction T2: Transfer 100 from A to C A = 500, B = 500, C = 500Read (A, s) s = s - 100Write (A, s) Read (C, s) S = S + 100Write (C, s) A = 400, B = 500, C = 600

<b>— —</b>	<b>— —</b>		-	*
Iransaction 11	Iransaction 12	A	В	
Read (A, t)		500	500	500
t = t - 100 Write (A, t)		400	500	500
Read (B, t) t = t + 100 Write (B, t)		400	600	500
	Read (A, s) s = s - 100 Write (A, s) Read (C, s)	300	600	500
	s = s + 100 Write (C, s)	300 + 600	600 ) + 600	600 = 1500

Transaction T1	Transaction T2	A	В	C 🖉
Read (A, t)		500	500	500
t = t - 100				
Write (A, t)		400	500	500
	Read (A, s)			
	s = s - 100			
	Write (A, s)	300	500	500
Read (B, t)				
t = t + 100		200	(00	F00
Write (B, t)		300	600	500
	Read (C, s)			
	s = s + 100			
	Write (C, s)	300	600	600
	30	0 + 60	0 + 600	) = 1500

Transaction T1	Transaction T2	A	В	С
Read (A, t)		500	500	500
t = t - 100				
	Read (A, s)			
	s = s - 100			
	Write (A, s)	400	500	500
Write (A, t)		400	500	500
Read (B, t)				
t = t + 100 Write (B. t)		400	600	500
	Read (C, s)			
	s = s + 100			
	Write (C, s)	400	600	600
	40	00 + 60	0 + 600	) = 1600



# Terminology

- Schedule:
  - The exact sequence of (relevant) actions of one or more transactions



# Problems

- Which schedules are "correct"?
   Mathematical characterization
- How to build a system that allows only "correct" schedules?
  - Efficient procedure to enforce correctness



CSCD43: Database System	s Technology	A	В	C 🛓
Serial Schedule Read (A, t)		500	500	500
t = t - 100				
T1 Write (A, t)				
Read (B, t)				
t = t + 100				
Write (B, t)		400	600	500
	Read (A, s)			
	s = s - 100			
	Write (A, s)			
T2	Read (C, s)			
	s = s + 100			
	Write (C, s)	300	600	600
		300 + 600	0 + 600	) = 1500

CSCD43: Database Sys	tems Technology	A	В	C 👻
Serial Schedu	e Read (A, s)	500	500	500
Τ2	s = s - 100 Write (A, s) Read (C, s) s = s + 100			
Read (A, t) t = t - 100 Write (A, t) T1 Read (B, t) t = t + 100	Write (C, s)	400	500	600
Write (B, t)		300	600	600
	30	00 + 60	0 + 600	) = 1500



# Serial Schedule





# Serial Schedule

- If any action of transaction T<sub>1</sub> precedes any action of T<sub>2</sub>, then all action of T<sub>1</sub> precede all action of T<sub>2</sub>
- The correctness principle tells us that every serial schedule will preserve consistency of the database state



• What's the problem with a Serial Schedule?



# Serializability

- A schedule is called *serializable* if its final effect is the same as that of a *serial schedule*
- Serializability → schedule is fine and does not result in inconsistent database
  - Since serial schedules are fine
- Non-serializable schedules are unlikely to result in consistent databases
- Scheduler ensures serializability



# Serializability

- Not possible to look at all *n!* serial schedules to check if the effect is the same
  - Instead we ensure serializability by allowing or not allowing certain schedules



# Conflict Serializability

- Weaker notion of serializability
- Depends only on reads and writes
- Which steps can be interleaved and which cannot





# Conflict Serializability

- Recall from OS course:
  - Multitasking
  - context switch

В

50

105



<u></u> T1	T2	<u></u>	T2
read(A) A = A -50 write(A)	read(A) tmp = A*0.1 A = A – tmp write(A)	read(A) A = A -50 write(A)	read(A) tmp = A*0.1 A = A – tmp
read(B)		read(B)	write(A)
B=B+50 write(B)	read(B) B = B+ tmp write(B)	B=B+50 write(B)	read(B) B = B+ tmp write(B)
Effect: <u>Befo</u> A 10	re <u>After</u> 00 45	Effect: <u>Be</u>	fore <u>After</u> 100 45

University of Toronto Scarborough

В

50

105

= =



<u>T1</u>	T2	<u></u>	T2
read(A)		read(A)	
A = A -50		A = A - 50	
write(A)		write(A)	
	read(A)		read(A)
	$tmp = A^*0.1$		tmp = A*0.1
	A = A - tmp		A = A - tmp
	write(A)		write(A)
read(B)		read(B)	
B=B+50		B = B + 50	
write(B)			read(B)
	read(B)	write(B)	
	B = B + tmp		B = B + tmp
	write(B)		write(B)
Effect: <u>Befor</u>	<u>e Atter</u>	Effect: <u>Be</u>	tore <u>After</u>
B 50	) 105	:== A B	50 55

#### CSCD43: Database Systems Technology



T1	T2	<u></u>	T2
read(A)		read(A)	
A = A - 50		A = A - 50	
write(A)		write(A)	
	read(A)		read(A)
	$tmp = A^*0.1$		tmp = A*0.1
	A = A - tmp		A = A - tmp
	write(A)		I
		read(B)	
read(B)		B=B+50	
B = B + 50			write(A)
write(B)		write(B)	
	read(B)		read(B)
	B = B + tmp		B = B + tmp
	write(B)		write(B)
Effect: Befor	re <u>After</u>	Effect: Be	efore After
A 10	00 45	== A	100 45
B 5	0 105	В	50 105



	Τ2	T1	T2
read(A)		read(A)	
A = A - 50		A = A - 50	
write(A)		write(A)	
	read(A)		
	tmp = A*0.1	read(B)	
	A = A - tmp	B=B+50	
	write(A)	write(B)	
			read(A)
read(B)			tmp = A*0.1
B = B + 50			A = A - tmp
write(B)			write(A)
	read(B)		read(B)
	B = B + tmp		B = B + tmp
	write(B)		write(B)
Effect: <u>Befor</u>	<u>e After</u>	Effect: <u>Ber</u>	fore <u>After</u>
A 10	U 45	== A	100 45 50 105
D 3(	J 100	D	50 105



# Simpler Notation

# $r_{T}(X)$ Transaction T reads X

#### $w_{T}(X)$ Transaction T writes X



# What is X in r (X)?

- X could be any component of a database:
  - Attribute of a tuple
  - Tuple
  - Block in which a tuple resides
  - A relation
  - ...



# Non-Conflicting Steps

- Two Reads
  - E.g., r<sub>i</sub>(X); r<sub>i</sub>(Y)
- Read and write of different database element
  E.g., r<sub>i</sub>(X); w<sub>j</sub>(Y)
- Two writes of different database elements
  E.g., w<sub>i</sub>(X); w<sub>j</sub>(Y)



# Conflicting Steps

- Two actions of the same transaction
  E.g., r<sub>i</sub>(X); w<sub>i</sub>(Y)
- Two writes of the same database element
  E.g., w<sub>i</sub>(X); w<sub>j</sub>(X)
- A read and a write of the same database element
  E.g., r<sub>i</sub>(X); w<sub>j</sub>(X)



# Conflict Serializability

- Conflict-equivalent schedules:
  - If S can be transformed into S' through a series of swaps, S and S' are called *conflict-equivalent*
  - conflict-equivalent guarantees same final effect on the database
- A schedule S is conflict-serializable if it is conflictequivalent to a serial schedule



# Conflict-Serializability

- Commercial systems generally support *conflict-serializability* 
  - Stronger notion than serializability
- Turn a given schedule to a serial one by make as many nonconflicting swaps as we wish



# Testing for conflict-serializability

- Given a schedule, determine if it is conflict-serializable
- Construct a *precedence-graph* over the transactions
  - A directed edge from T1 and T2, if they have conflicting instructions, and T1's conflicting instruction comes first
- If there is a cycle in the graph, not conflict-serializable
  - Can be checked in at most O(n+e) time, where *n* is the number of vertices, and *e* is the number of edges
- If there is none, conflict-serializable

# Precedence Graph

- Precedence graph for schedule S:
  - Nodes: Transactions in S
  - Edges:  $Ti \rightarrow Tj$  whenever
    - S: ... ri (X) ... wi(X) ...
    - S: ... wi (X) ... wj (X) ...
    - S: ... ri(X) ... wj (X) ...

#### Note: not necessarily consecutive



# Enforcing Serializability





# Graph Theory 101 Directed Graph:





# Graph Theory 101 Edges Directed Graph:

### 

# Graph Theory 101 Directed Graph:








# Graph Theory 101

#### Acyclic Graph: A graph with no cycles



# Graph Theory 101 Acyclic Graph:





- $T_i \rightarrow T_i$  whenever:
  - There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

 $r_i(X); w_i(Y)$  $w_i(X); w_j(X)$  $r_i(X); w_i(X)$ 



•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

S<sub>1</sub>: r<sub>2</sub>(A); r<sub>1</sub>(B); w<sub>2</sub>(A); r<sub>3</sub>(A); w<sub>1</sub>(B); w<sub>3</sub>(A); r<sub>2</sub>(B); w<sub>2</sub>(B);

 $r_i(X); w_i(Y)$  $w_i(X); w_j(X)$  $r_i(X); w_i(X)$ 



 $w_i(X); w_j(X)$  $r_i(X); w_i(X)$ 

## Precedence Graph – Example 1

•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

$$S_1: r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B);$$





•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .



University of Toronto Scarborough



•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

 $S_1: r_2(A); r_1(B); w_2(A); r_3(A); w_1(B); w_3(A); r_2(B); w_2(B);$ 





•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

S<sub>2</sub>: r<sub>2</sub>(A); r<sub>1</sub>(B); w<sub>2</sub>(A); r<sub>2</sub>(B); r<sub>3</sub>(A); w<sub>1</sub>(B); w<sub>3</sub>(A); w<sub>2</sub>(B);

 $r_i(X); w_i(Y)$  $w_i(X); w_j(X)$  $r_i(X); w_i(X)$ 



 $w_i(X); w_j(X)$  $r_i(X); w_i(X)$ 

## Precedence Graph – Example 2

•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

 $S_2: r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B);$ 





 $w_i(X); w_i(X)$ 

 $r_i(X)$ ;  $w_i(X)$ 

#### Precedence Graph – Example 2

•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

 $S_2: r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B);$ 





 $w_i(X); w_i(X)$ 

 $r_i(X); W_i(X)$ 

#### Precedence Graph – Example 2

•  $T_i \rightarrow T_i$  whenever:

– There is an action of  $T_i$  that occurs before a conflicting action of  $T_j$ .

 $S_2: r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B);$ 



CSCD43: Database Systems Technology





University of Toronto Scarborough



# Conflict-serializable

- Two schedules are *conflict-equivalent* if they can be turned one into the other by a sequence of nonconflicting swaps of adjacent actions
- A schedule is *conflict-serializable* if it is conflict-equivalent to a serial schedule



#### Conflict Serializable Schedule

Transformations: swap non-conflicting actions





#### Transformation



#### r1(A) w1(A) r2(A) r1(B) w2(A) w1(B) r2(B) w2(B)

University of Toronto Scarborough





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





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





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





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





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





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





 $r_{1}(A); w_{1}(A); r_{2}(A); w_{2}(A); r_{1}(B); w_{1}(B); r_{2}(B); w_{2}(B)$  $r_{1}(A); w_{1}(A); r_{2}(A); r_{1}(B); w_{2}(A); w_{1}(B); r_{2}(B); w_{2}(B)$  $r_{1}(A); w_{1}(A); r_{1}(B); r_{2}(A); w_{2}(A); w_{1}(B); r_{2}(B); w_{2}(B)$  $r_{1}(A); w_{1}(A); r_{1}(B); r_{2}(A); w_{1}(B); w_{2}(A); r_{2}(B); w_{2}(B)$ 





 $r_{1}(A); w_{1}(A); r_{2}(A); w_{2}(A); r_{1}(B); w_{1}(B); r_{2}(B); w_{2}(B)$  $r_{1}(A); w_{1}(A); r_{2}(A); r_{1}(B); w_{2}(A); w_{1}(B); r_{2}(B); w_{2}(B)$  $r_{1}(A); w_{1}(A); r_{1}(B); r_{2}(A); w_{2}(A); w_{1}(B); r_{2}(B); w_{2}(B)$  $r_{1}(A); w_{1}(A); r_{1}(B); r_{2}(A); w_{1}(B); w_{2}(A); r_{2}(B); w_{2}(B)$ 





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



# Enforcing Serializability

University of Toronto Scarborough



## How to enforce serializable schedules?

*Option 1:* run system, recording P(S); at end of day, check for P(S) cycles and declare if execution was good! Unrealistic...!



# How to enforce serializable schedules?



But how ???

2.A) Buffer transactions during n seconds, stop DBMS, make schedule, execute schedule, repeat... unrealistic...!

# 2.B) Use a locking protocol!







# <u>Rule #1:</u> Well-formed transactions

Ti: ... Ii(A) ... pi(A) ... ui(A) ...



# Rule #2 Legal scheduler

# $S = \dots Ii(A) \dots Ui(A) \dots Ii(A)$

University of Toronto Scarborough



 What schedules are legal? What transactions are well-formed?
S1 = l1(A)l1(B)r1(A)w1(B)l2(B)u1(A)u1(B) r2(B)w2(B)u2(B)l3(B)r3(B)u3(B)



 What schedules are legal? What transactions are well-formed?
S1 = I1(A)I1(B)r1(A)w1(B)I2(B)u1(A)u1(B) r2(B)w2(B)u2(B)I3(B)r3(B)u3(B)



 What schedules are legal? What transactions are well-formed?
S1 = l1(A)l1(B)r1(A)w1(B)l2(B)u1(A)u1(B) r2(B)w2(B)u2(B)l3(B)r3(B)u3(B)

 $S2 = I_1(A)r_1(A)w_1(B)u_1(A)u_1(B)$  $I_2(B)r_2(B)w_2(B)I_3(B)r_3(B)u_3(B)$ 



• What schedules are legal? What transactions are well-formed?  $S1 = I_1(A)I_1(B)r_1(A)w_1(B)I_2(B)u_1(A)u_1(B)$  $r_2(B)w_2(B)u_2(B)I_3(B)r_3(B)u_3(B)$  $S2 = I_1(A)r_1(A)w_1(B)u_1(A)u_1(B)$  $I_2(B)r_2(B)w_2(B)I_3(B)r_3(B)u_3(B)$ 



 What schedules are legal? What transactions are well-formed?
S1 = l1(A)l1(B)r1(A)w1(B)l2(B)u1(A)u1(B) r2(B)w2(B)u2(B)l3(B)r3(B)u3(B)

 $S2 = I_1(A)r_1(A)w_1(B)u_1(A)u_1(B)$  $I_2(B)r_2(B)w_2(B)I_3(B)r_3(B)u_3(B)$ 

 $S3 = I_1(A)r_1(A)u_1(A)I_1(B)w_1(B)u_1(B)$  $I_2(B)r_2(B)w_2(B)u_2(B)I_3(B)r_3(B)u_3(B)$


# Exercise:

 What schedules are legal? What transactions are well-formed?
S1 = l1(A)l1(B)r1(A)w1(B)l2(B)u1(A)u1(B) r2(B)w2(B)u2(B)l3(B)r3(B)u3(B)

 $S2 = I_1(A)r_1(A)w_1(B)u_1(A)u_1(B)$ I\_2(B)r\_2(B)w\_2(B)I\_3(B)r\_3(B)u\_3(B)

 $S3 = I_1(A)r_1(A)u_1(A)I_1(B)w_1(B)u_1(B)$  $I_2(B)r_2(B)w_2(B)u_2(B)I_3(B)r_3(B)u_3(B)$ 

# Locking Example

- T1: Read(A); A =A+100; Write(A); Read(B); B=B+100; Write(B);
- T2: Read(A); A = A\*2; Write(A); Read(B); B=B\*2; Write(B);

# Serial Schedule





# Schedule A

T1	T2
l1(A);Read(A)	
A←A+100;Write(A);u1(A)	
	I2(A);Read(A)
	A←Ax2;Write(A);u <sub>2</sub> (A)
	I2(B);Read(B)
	B←Bx2;Write(B);u <sub>2</sub> (B)
I1(B);Read(B)	
B←B+100;Write(B);u <sub>1</sub> (B)	



### Schedule A 25 T2 25 Τ1 I1(A);Read(A) A←A+100;Write(A);u1(A) 125 l<sub>2</sub>(A);Read(A) $A \leftarrow Ax2; Write(A); u_2(A)$ 250 $I_2(B)$ ; Read(B) B←Bx2;Write(B);u<sub>2</sub>(B) 50 I1(B);Read(B) $B \leftarrow B + 100$ ; Write(B); u<sub>1</sub>(B) 150 250 150



## <u>Rule #3</u> Two phase locking (2PL) for transactions





## 

# Schedule B





# Schedule B





#### Schedule B T1 T2 25 25 I<sub>1</sub>(A);Read(A) $A \leftarrow A + 100$ ; Write(A) $I_1(B); u_1(A)$ 125 delayed $I_2(A)$ ; Read(A) A←Ax2;Write(A) 250 Read(B);B ← B+100 125 Write(B); u<sub>1</sub>(B) $I_2(B)$ ; $u_2(A)$ ; Read(B) $B \leftarrow Bx2; Write(B); u_2(B);$ 250 250 250