Chapter 17

Introduction to Transaction Processing Concepts and Theory
Chapter Outline

1 Introduction to Transaction Processing
2 Transaction and System Concepts
3 Desirable Properties of Transactions
4 Characterizing Schedules based on Recoverability
5 Characterizing Schedules based on Serializability
6 Transaction Support in SQL
Transaction Processing: Introduction

➲ Single-User System:
• At most one user at a time can use the system.

➲ Multiuser System:
• Many users can access the system concurrently.

➲ Concurrency
• Interleaved processing:
  • Concurrent execution of processes is interleaved in a single CPU
• Parallel processing:
  • Processes are concurrently executed in multiple CPUs.
A **Transaction**: 
- Logical unit of database processing that includes one or more access operations (read - retrieval, write - insert or update, delete).

A transaction (set of operations) may be stand-alone specified in a high level language like SQL submitted interactively, or may be embedded within a program.

**Transaction boundaries**: 
- Begin and End transaction.

An **application program** may contain several transactions separated by the Begin and End transaction boundaries.
SIMPLE MODEL OF A DATABASE (for purposes of discussing transactions):

- **A database** is a collection of named data items
- **Granularity** of data - a field, a record, or a whole disk block (Concepts are independent of granularity)
- Basic operations are **read** and **write**
  - **read_item(X)**: Reads a database item named X into a program variable. To simplify our notation, we assume that the program variable is also named X.
  - **write_item(X)**: Writes the value of program variable X into the database item named X.
READ AND WRITE OPERATIONS:
➲ Basic unit of data transfer from the disk to the computer main memory is one block. In general, a data item (what is read or written) will be the field of some record in the database, although it may be a larger unit such as a record or even a whole block.
➲ read_item(X) command includes the following steps:
- Find the address of the disk block that contains item X.
- Copy that disk block into a buffer in main memory (if that disk block is not already in some main memory buffer).
- Copy item X from the buffer to the program variable named X.
READ AND WRITE OPERATIONS (contd.):

- `write_item(X)` command includes the following steps:
  - Find the address of the disk block that contains item X.
  - Copy that disk block into a buffer in main memory (if that disk block is not already in some main memory buffer).
  - Copy item X from the program variable named X into its correct location in the buffer.
  - Store the updated block from the buffer back to disk (either immediately or at some later point in time).
Two sample transactions

FIGURE 17.2 Two sample transactions:

- (a) Transaction T1
  - read_item (X);
  - X:=X-N;
  - write_item (X);
  - read_item (Y);
  - Y:=Y+N;
  - write_item (Y);
- (b) Transaction T2
  - read_item (X);
  - X:=X+M;
  - write_item (X);
Why Concurrency Control is needed:

❖ **The Lost Update Problem**
  - This occurs when two transactions that access the same database items have their operations interleaved in a way that makes the value of some database item incorrect.

❖ **The Temporary Update (or Dirty Read) Problem**
  - This occurs when one transaction updates a database item and then the transaction fails for some reason (see Section 17.1.4).
  - The updated item is accessed by another transaction before it is changed back to its original value.

❖ **The Incorrect Summary Problem**
  - If one transaction is calculating an aggregate summary function on a number of records while other transactions are updating some of these records, the aggregate function may calculate some values before they are updated and others after they are updated.
Concurrent execution is uncontrolled:
(a) The lost update problem.

Figure 17.3
Some problems that occur when concurrent execution is uncontrolled. (a) The lost update problem. (b) The temporary update problem. (c) The incorrect summary problem.

<table>
<thead>
<tr>
<th>Time</th>
<th>$T_1$</th>
<th>$T_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>read_item(X);</td>
<td>read_item(X);</td>
</tr>
<tr>
<td></td>
<td>$X := X - N$;</td>
<td>$X := X + M_i$;</td>
</tr>
<tr>
<td></td>
<td>write_item(X);</td>
<td>write_item(X);</td>
</tr>
<tr>
<td></td>
<td>read_item(Y);</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Y := Y + N$;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>write_item(Y);</td>
<td></td>
</tr>
</tbody>
</table>

Item $X$ has an incorrect value because its update by $T_1$ is lost (overwritten).
Concurrent execution is uncontrolled: (b) The temporary update problem.

Figure 17.3
Some problems that occur when concurrent execution is uncontrolled. (a) The lost update problem. (b) The temporary update problem. (c) The incorrect summary problem.

(b)

<table>
<thead>
<tr>
<th>Time</th>
<th>$T_1$</th>
<th>$T_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>read_item($X$); $X := X - N$; read_item($X$);</td>
<td>read_item($X$); $X := X + M$; write_item($X$);</td>
</tr>
<tr>
<td></td>
<td>read_item($Y$);</td>
<td></td>
</tr>
</tbody>
</table>

Transaction $T_1$ fails and must change the value of $X$ back to its old value; meanwhile $T_2$ has read the temporary incorrect value of $X$. 
**Concurrent execution is uncontrolled:**

(c) The incorrect summary problem.

**Figure 17.3**
Some problems that occur when concurrent execution is uncontrolled. (a) The lost update problem. (b) The temporary update problem. (c) The incorrect summary problem.

<table>
<thead>
<tr>
<th>(c)</th>
<th>$T_1$</th>
<th>$T_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>read_item($X$); $X := X - N$;</td>
<td>$sum := 0;$</td>
</tr>
<tr>
<td></td>
<td>write_item($X$);</td>
<td>read_item($A$);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$sum := sum + A;$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\vdots$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\vdots$</td>
</tr>
<tr>
<td></td>
<td>read_item($Y$); $Y := Y + N$;</td>
<td>$read_item($X$);</td>
</tr>
<tr>
<td></td>
<td>write_item($Y$);</td>
<td>$sum := sum + X;$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$read_item($Y$);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$sum := sum + Y;$</td>
</tr>
</tbody>
</table>

$T_3$ reads $X$ after $N$ is subtracted and reads $Y$ before $N$ is added; a wrong summary is the result (off by $N$).
Why recovery is needed?

Transaction may fail because of,

1. A computer failure (system crash):
   A hardware or software error occurs in the computer system during transaction execution. If the hardware crashes, the contents of the computer’s internal memory may be lost.

2. A transaction or system error:
   Some operation in the transaction may cause it to fail, such as integer overflow or division by zero. Transaction failure may also occur because of erroneous parameter values or because of a logical programming error. In addition, the user may interrupt the transaction during its execution.
3. Local errors or exception conditions detected by the transaction:

Certain conditions necessitate cancellation of the transaction. For example, data for the transaction may not be found. A condition, such as insufficient account balance in a banking database, may cause a transaction, such as a fund withdrawal from that account, to be canceled.

A programmed abort in the transaction causes it to fail.

4. Concurrency control enforcement:

The concurrency control method may decide to abort the transaction, to be restarted later, because it violates serializability or because several transactions are in a state of deadlock (see Chapter 18).
5. Disk failure:
   Some disk blocks may lose their data because of a read or write malfunction or because of a disk read/write head crash. This may happen during a read or a write operation of the transaction.

6. Physical problems and catastrophes:
   This refers to an endless list of problems that includes power or air-conditioning failure, fire, theft, sabotage, overwriting disks or tapes by mistake, and mounting of a wrong tape by the operator.
A **transaction** is an atomic unit of work that is either completed in its entirety or not done at all.

- For recovery purposes, the system needs to keep track of when the transaction starts, terminates, and commits or aborts.

**Transaction states:**
- Active state
- Partially committed state
- Committed state
- Failed state
- Terminated State
Recovery manager keeps track of the following operations:

- **begin_transaction**: This marks the beginning of transaction execution.
- **read** or **write**: These specify read or write operations on the database items that are executed as part of a transaction.
- **end_transaction**: This specifies that read and write transaction operations have ended and marks the end limit of transaction execution.
- At this point it may be necessary to check whether the changes introduced by the transaction can be permanently applied to the database or whether the transaction has to be aborted because it violates concurrency control or for some other reason.
- **commit_transaction**: This signals a successful end of the transaction so that any changes (updates) executed by the transaction can be safely committed to the database and will not be undone.

- **rollback (or abort)**: This signals that the transaction has ended unsuccessfully, so that any changes or effects that the transaction may have applied to the database must be undone.
Recovery techniques use the following operators:

- **undo**: Similar to rollback except that it applies to a single operation rather than to a whole transaction.
- **redo**: This specifies that certain transaction operations must be redone to ensure that all the operations of a committed transaction have been applied successfully to the database.
Figure 17.4
State transition diagram illustrating the states for transaction execution.
The System Log

- **Log** or **Journal**: The log keeps track of all transaction operations that affect the values of database items.
  - This information may be needed to permit recovery from transaction failures.
  - The log is kept on disk, so it is not affected by any type of failure except for disk or catastrophic failure.
  - In addition, the log is periodically backed up to archival storage (tape) to guard against such catastrophic failures.
• T in the following discussion refers to a unique **transaction-id** that is generated automatically by the system and is used to identify each transaction:

• Types of log record:
  • `[start_transaction,T]`: Records that transaction T has started execution.
  • `[write_item,T,X,old_value,new_value]`: Records that transaction T has changed the value of database item X from old_value to new_value.
  • `[read_item,T,X]`: Records that transaction T has read the value of database item X.
  • `[commit,T]`: Records that transaction T has completed successfully, and affirms that its effect can be committed (recorded permanently) to the database.
  • `[abort,T]`: Records that transaction T has been aborted.
Desirable Properties of Transactions

ACID properties:

➲ **Atomicity**: A transaction is an atomic unit of processing; it is either performed in its entirety or not performed at all.

➲ **Consistency preservation**: A correct execution of the transaction must take the database from one consistent state to another.

➲ **Isolation**: A transaction should not make its updates visible to other transactions until it is committed; this property, when enforced strictly, solves the temporary update problem and makes cascading rollbacks of transactions unnecessary (see Chapter 21).

➲ **Durability or permanency**: Once a transaction changes the database and the changes are committed, these changes must never be lost because of subsequent failure.
Transaction Schedules and Characterization

➲ Transaction schedule or history:

- When transactions are executing concurrently in an interleaved fashion, the order of execution of operations from the various transactions forms what is known as a transaction schedule (or history).

➲ A schedule (or history) $S$ of $n$ transactions $T_1, T_2, \ldots, T_n$:

- It is an ordering of the operations of the transactions subject to the constraint that, for each transaction $T_i$ that participates in $S$, the operations of $T_1$ in $S$ must appear in the same order in which they occur in $T_1$.

- Note, however, that operations from other transactions $T_j$ can be interleaved with the operations of $T_i$ in $S$. 
**Operation Conflict**

- Two operations in a schedule are said to be conflict IF
  - They belong to different transactions
  - They access the same data item $X$
  - At least one of the operation is write($X$)
    (ex: $S : r1(X), r2(X), w1(X)$)

- **Committed Projection**
  - Transactions are submitted continuously to the system, so it is difficult to encounter a complete schedule
  - CP includes only the operations of the committed transactions
Characterizing Schedules based on Recoverability

Schedules classified on recoverability:

- **Recoverable schedule:**
  - One where no transaction needs to be rolled back.
  - A schedule $S$ is recoverable if no transaction $T$ in $S$ commits until all transactions $T'$ that have written an item that $T$ reads have committed.

- **Cascadless schedule:**
  - One where every transaction reads only the items that are written by committed transactions.
Schedules requiring cascaded rollback:
- A schedule in which uncommitted transactions that read an item from a failed transaction must be rolled back.

Strict Schedules:
- A schedule in which a transaction can neither read or write an item X until the last transaction that wrote X has committed.
Example

- Sa': r1(X), r2(X), w1(X), r1(Y), w2(X), c2, w1(Y), c1

- Sc : r1(X), w1(X), r2(X), r1(Y), w2(X), c2, a1 (non recoverable)

- Sd : r1(X), w1(X), r2(X), r1(Y), w2(X), w1(Y) c1, c2

- Se : r1(X), w1(X), r2(X), r1(Y), w2(X), w1(Y) a1, a2
Characterizing Schedules based on Serializability

➲ Serial schedule:

• A schedule $S$ is serial if, for every transaction $T$ participating in the schedule, all the operations of $T$ are executed consecutively in the schedule.
  • Otherwise, the schedule is called nonserial schedule.

➲ Serializable schedule:

• A schedule $S$ of $n$ transactions is serializable if it is equivalent to some serial schedule of the same $n$ transactions.
Example

- r1(X), w1(x), r1(Y), w1(Y), c1, r2(X), w2(X), c2
- r2(X), w2(X), c2; r1(X), w1(X), r1(Y), w1(Y), c1
- r1(X), r2(X), w1(X), r1(Y) w2(X), w1(Y), c2, c1
- r1(X), w1(X), r2(X), w2(X), r1(Y), W1(Y),
Different forms of Schedule equivalence

➲ Conflict equivalent:
  • Two schedules are said to be conflict equivalent if the order of any two conflicting operations is the same in both schedules.

➲ Conflict serializable:
  • A schedule $S$ is said to be conflict serializable if it is conflict equivalent to some serial schedule $S'$. 
Characterizing Schedules based on Serializability (3)

➲ Being serializable is not the same as being serial
➲ Being serializable implies that the schedule is a correct schedule.

• It will leave the database in a consistent state.
• The interleaving is appropriate and will result in a state as if the transactions were serially executed, yet will achieve efficiency due to concurrent execution.
Testing for conflict serializability: Algorithm 17.1:

- Looks at only read_Item (X) and write_Item (X) operations
- Constructs a precedence graph (serialization graph) - a graph with directed edges
- An edge is created from Ti to Tj if one of the operations in Ti appears before a conflicting operation in Tj
- The schedule is serializable if and only if the precedence graph has no cycles.
Constructing the Precedence Graphs

FIGURE 17.7 Constructing the precedence graphs for schedules A and D from Figure 17.5 to test for conflict serializability.

- (a) Precedence graph for serial schedule A.
- (b) Precedence graph for serial schedule B.
- (c) Precedence graph for schedule C (not serializable).
- (d) Precedence graph for schedule D (serializable, equivalent to schedule A).
Another example of serializability

Testing

**Figure 17.8**

Another example of serializability testing.

(a) The read and write operations of three transactions $T_1$, $T_2$, and $T_3$. (b) Schedule $E$. (c) Schedule $F$. 

<table>
<thead>
<tr>
<th>Transaction $T_1$</th>
<th>Transaction $T_2$</th>
<th>Transaction $T_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>read_item(X);</td>
<td>read_item(Z);</td>
<td>read_item(Y);</td>
</tr>
<tr>
<td>write_item(X);</td>
<td>read_item(Y);</td>
<td>read_item(Z);</td>
</tr>
<tr>
<td>read_item(Y);</td>
<td>write_item(Y);</td>
<td>read_item(Y);</td>
</tr>
<tr>
<td>write_item(Y);</td>
<td>read_item(X);</td>
<td>write_item(Y);</td>
</tr>
<tr>
<td></td>
<td>write_item(X);</td>
<td>write_item(Z);</td>
</tr>
</tbody>
</table>
Another Example of Serializability Testing

Figure 17.8
Another example of serializability testing. (a) The read and write operations of three transactions $T_1$, $T_2$, and $T_3$. (b) Schedule E. (c) Schedule F.

<table>
<thead>
<tr>
<th>Time</th>
<th>Transaction $T_1$</th>
<th>Transaction $T_2$</th>
<th>Transaction $T_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>read_item(X);</td>
<td>read_item(Z);</td>
<td>read_item(Y);</td>
</tr>
<tr>
<td></td>
<td>write_item(X);</td>
<td>read_item(Y);</td>
<td>read_item(Z);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>write_item(Y);</td>
<td></td>
</tr>
<tr>
<td></td>
<td>read_item(Y);</td>
<td></td>
<td>write_item(Y);</td>
</tr>
<tr>
<td></td>
<td>write_item(Y);</td>
<td></td>
<td>write_item(X);</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>read_item(X);</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>write_item(Z);</td>
</tr>
</tbody>
</table>

Schedule E
Another Example of Serializability Testing

**Figure 17.8**
Another example of serializability testing.
(a) The read and write operations of three transactions $T_1$, $T_2$, and $T_3$. (b) Schedule E. (c) Schedule F.

<table>
<thead>
<tr>
<th>Time</th>
<th>Transaction $T_1$</th>
<th>Transaction $T_2$</th>
<th>Transaction $T_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>read_item($X$);</td>
<td></td>
<td>read_item($Y$);</td>
</tr>
<tr>
<td></td>
<td>write_item($X$);</td>
<td></td>
<td>read_item($Z$);</td>
</tr>
<tr>
<td></td>
<td>read_item($Y$);</td>
<td></td>
<td>write_item($Y$);</td>
</tr>
<tr>
<td></td>
<td>write_item($Y$);</td>
<td></td>
<td>write_item($Y$);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>read_item($Z$);</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>write_item($X$);</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>read_item($X$);</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>write_item($X$);</td>
<td></td>
</tr>
</tbody>
</table>

Schedule F
View equivalence:
- A less restrictive definition of equivalence of schedules.

View serializability:
- Definition of serializability based on view equivalence.
- A schedule is view serializable if it is view equivalent to a serial schedule.
Two schedules are said to be view equivalent if the following three conditions hold:

- The same set of transactions participates in S and S’, and S and S’ include the same operations of those transactions.
- For any operation Ri(X) of Ti in S, if the value of X read by the operation has been written by an operation Wj(X) of Tj (or if it is the original value of X before the schedule started), the same condition must hold for the value of X read by operation Ri(X) of Ti in S’.
- If the operation Wk(Y) of Tk is the last operation to write item Y in S, then Wk(Y) of Tk must also be the last operation to write item Y in S’.
The premise behind view equivalence:

- As long as each read operation of a transaction reads the result of *the same write operation* in both schedules, the write operations of each transaction must produce the same results.

- "The view": the read operations are said to see *the same view* in both schedules.
Relationship between view and conflict equivalence:

- The two are same under **constrained write assumption** which assumes that if T writes X, it is constrained by the value of X it read; i.e., new X = f(old X)
- Conflict serializability is **strict**er than view serializability. With unconstrained write (or blind write), a schedule that is view serializable is not necessarily conflict serializable.
- Any conflict serializable schedule is also view serializable, but not vice versa.
Relationship between view and conflict equivalence (cont):

- Consider the following schedule of three transactions
  - T1: r1(X), w1(X);  T2: w2(X);  and  T3: w3(X):
  - Schedule Sa: r1(X); w2(X); w1(X); w3(X); c1; c2; c3;

- In Sa, the operations w2(X) and w3(X) are blind writes, since T1 and T3 do not read the value of X.
  - Sa is view serializable, since it is view equivalent to the serial schedule T1, T2, T3.
  - However, Sa is not conflict serializable, since it is not conflict equivalent to any serial schedule.
Other Types of Equivalence of Schedules

- Under special **semantic constraints**, schedules that are otherwise not conflict serializable may work correctly.

  - Using commutative operations of addition and subtraction (which can be done in any order) certain non-serializable transactions may work correctly.
Other Types of Equivalence of Schedules (contd.)

Example: bank credit / debit transactions on a given item are **separable and commutative**.

- Consider the following schedule $S$ for the two transactions:
  - $Sh: r_1(X); w_1(X); r_2(Y); w_2(Y); r_1(Y); w_1(Y); r_2(X); w_2(X)$;
  - Using conflict serializability, it is **not serializable**.
  - However, if it came from a (read, update, write) sequence as follows:
    - $r_1(X); X := X - 10; w_1(X); r_2(Y); Y := Y - 20; r_1(Y);
    - $Y := Y + 10; w_1(Y); r_2(X); X := X + 20; (X)$;
  - Sequence explanation: debit, debit, credit, credit.
  - It is a **correct schedule for the given semantics**.
6 Transaction Support in SQL2 (1)

- A single SQL statement is always considered to be **atomic**.
  - Either the statement completes execution without error or it fails and leaves the database unchanged.

- With SQL, there is **no explicit Begin Transaction statement**.
  - Transaction initiation is done implicitly when particular SQL statements are encountered.

- Every transaction **must have an explicit end statement**, which is either a COMMIT or ROLLBACK.
Characteristics specified by a SET TRANSACTION statement in SQL2:

✏ **Access mode:**

- READ ONLY or READ WRITE.
- The default is READ WRITE unless the isolation level of READ UNCOMMITTED is specified, in which case READ ONLY is assumed.

✏ **Diagnostic size** $n$, specifies an integer value $n$, indicating the number of conditions that can be held simultaneously in the diagnostic area.
Characteristics specified by a SET TRANSACTION statement in SQL2 (contd.):

- **Isolation level** `<isolation>`, where `<isolation>` can be READ UNCOMMITTED, READ COMMITTED, REPEATABLE READ or SERIALIZABLE. The default is SERIALIZABLE.

  - With SERIALIZABLE: the interleaved execution of transactions will adhere to our notion of serializability.
  - However, if any transaction executes at a lower level, then serializability may be violated.
Potential problem with lower isolation levels:

- **Dirty Read:**
  - Reading a value that was written by a transaction which failed.

- **Nonrepeatable Read:**
  - Allowing another transaction to write a new value between multiple reads of one transaction.
  - A transaction T1 may read a given value from a table. If another transaction T2 later updates that value and T1 reads that value again, T1 will see a different value.
  - Consider that T1 reads the employee salary for Smith. Next, T2 updates the salary for Smith. If T1 reads Smith's salary again, then it will see a different value for Smith's salary.
Potential problem with lower isolation levels (contd.):

- **Phantoms:**
  - New rows being read using the same read with a condition.
    - A transaction T1 may read a set of rows from a table, perhaps based on some condition specified in the SQL WHERE clause.
    - Now suppose that a transaction T2 inserts a new row that also satisfies the WHERE clause condition of T1, into the table used by T1.
    - If T1 is repeated, then T1 will see a row that previously did not exist, called a phantom.
Sample SQL transaction:
EXEC SQL whenever sqlerror go to UNDO;
EXEC SQL SET TRANSACTION
  READ WRITE
  DIAGNOSTICS SIZE 5
  ISOLATION LEVEL SERIALIZABLE;
EXEC SQL INSERT
  INTO EMPLOYEE (FNAME, LNAME, SSN, DNO, SALARY)
  VALUES ('Robert','Smith','991004321',2,35000);
EXEC SQL UPDATE EMPLOYEE
  SET SALARY = SALARY * 1.1
  WHERE DNO = 2;
EXEC SQL COMMIT;
  GOTO THE_END;
UNDO: EXEC SQL ROLLBACK;
THE_END: ...
Possible violation of serializability:

<table>
<thead>
<tr>
<th>Isolation level</th>
<th>Dirty read</th>
<th>nonrepeatable read</th>
<th>phantom</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ UNCOMMITTED</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>READ COMMITTED</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>REPEATABLABLE READ</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>SERIALIZABLE</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
Summary

- Transaction and System Concepts
- Desirable Properties of Transactions
- Characterizing Schedules based on Recoverability
- Characterizing Schedules based on Serializability
- Transaction Support in SQL