Warm-up and Terminology (30 points)
1. (5 points) Write your name and student id on your blue book. Shahram Ghandeharizadeh, 777-77-7777
2. (5 points) True or False: Atomicity means repeating an operation multiple times produces the same result. False
3. (5 points) True or False: A transaction may abort after executing its last statement and prior to generating its <commit> log record. True
4. (5 points) True or False: Time stamp based protocols are free from deadlocks. True
5. (5 points) True or False: With two phase locking, the serial order of transactions is the same as their order of arrival. False
6. (5 points) True or False: IS and IX lock modes are compatible. True

Transaction Processing and Concurrency Control (20 points)

Crash Recovery (20 points)
8. (10 points) Assume logging with immediate database modification, what operations are performed during recovery from a system crash? (To simplify discussion, assume no check pointing.) Read the log file and (1) redo all those transactions with a start and commit log entries, (2) undo all those transactions with a start log record and no commit log entry.
9. (10 points) How frequently should a log-based protocol issue checkpoints and why? The frequency of checkpoints should depend on how quickly one desires the system to recover from a failure. For example, if the worst case recovery time should be less than 1 minute then the system should checkpoint every minute.

Normal Forms (20 points)
10. (10 points) When is the fourth normal form violated? Give an example. Fourth normal form is violated when two independent multi-valued attributes are kept in the same record. Example is an company that keeps track of its employees skills and the languages they can use fluently. Storing these two in the same table violates the fourth normal form.
11. (10 points) Violating the 3rd normal form may result in loss of information. Show this with an example. Imagine keeping track of an employee, his department and the department’s address in one table. When the department becomes empty (no employees) then it might be deleted all together.

Continuous Media Servers (10 points)
12. (10 points) A large block size (say 1 Megabyte) enhances useful utilization of a disk drive for continuous media servers. This is because a disk performs useful work
when transferring data and waste-full work when incurring seeks and rotational latency. Why do continuous media servers assume large block sizes while record-based database management systems (DBMSs) use small block sizes (8 Kilobytes)? This is because record-based DBMSs use a small fraction of a block, namely, a record. With large blocks, reading many records is a huge waste of time because most of the content of the page is not used. Moreover, it reduces the number of buffer pool hits because large frames reduce the number of buffer pool frames. In the worst case scenario where the number of frames might be limited, the buffer pool might exhibit a thrashing behavior.