**Pages 65–69** mainly discuss the basic concepts of pipelining, like the material covered in class.

Instruction execution is divided into five stages: fetch, decode, operand fetch, execute, and write-back. By fetching instructions in advance, memory access latency can be reduced.

A superscalar architecture uses dual or single pipelines with multiple functional units to execute multiple instructions simultaneously.

**Pages 73–78** mainly discuss the concepts of bits, bytes, and memory storage methods.

The basic binary storage unit in memory is the bit, and the smallest addressable unit composed of multiple bits is called a cell. The address of a cell requires a certain number of bits for identification; an address length of bits can identify up to cells. Modern computers typically use 8 bits as one unit called a byte. In an x-bit computer, one word equals bytes. For example, in a 32-bit computer, one word is bytes. Most instructions operate on whole words.

The order of bytes within a word can vary and is categorised as either big-endian or little-endian. In big-endian format, the most significant byte is stored at the lowest memory address (e.g., address 0—Byte0, address 1—Byte1, address 2—Byte2, address 3—Byte3, etc.). In little-endian format, the order is reversed (e.g., address 0 - Byte3, address 1 - Byte2, etc.). When machines with different byte orders exchange data, byte-order conversion must be performed. Otherwise, integers and strings may become disordered.