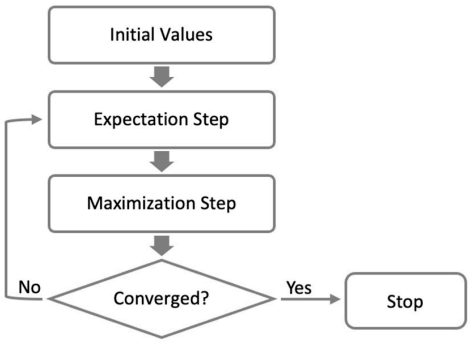
Cbb752b25 Quiz #1 Study Guide

The quiz will be closed-book and last for the whole lecture time.

There are **some topics** from the lectures that we **drilled down** into detail on. Here is a list of them that you should know for the quiz:

- Genomics:
  - high-throughput sequencing & bias (understanding sources of sequencing bias)
  - short-read versus long read sequencing
  - key sequencing techniques and applications (Hi-C, RNA-seq etc.)
  - Somatic versus germline
- Proteomics:
  - key challenges compared to dna sequencing
  - proteomic techniques
- Variant calling:
  - differences between SNV and SV
  - challenges in variant calling and how to reduce errors
- Multi-omics:
  - Chip-seq; RNA-seq; Hi-C
- Database
  - understand database normalization concepts
  - identify normal form violations
- Personal Genome
  - Understand the prevalence of SNPs in the human genome.
  - Comprehend the basic concept of Polygenic Risk Scores (PRS) and their role in personal genomics. And limitations and applications in predicting complex diseases.
- How to do a dynamic program sequence alignment (i.e. Smith Waterman)
  - key conceptual differences between global and local alignment algorithms.
  - Apply a dynamic program sequence alignment sequence alignment, including matrix construction, scoring, and traceback.
- Multiple sequence alignment
  - Understand main steps of the sampling algorithm (EM, Gibbs)
  - EM workflow:



- FAST Alignment
    - Understand the the sequence alignment (main steps, time complexity, Speed v Sensitivity Tradeoff)

For **all other topics**, you should understand at a **high level**, i.e., the contents of the lecture.