

# Genome 540 Discussion

Feb. 3<sup>rd</sup> 2026



# HW4 Intro

# Overview

1. Write a program to find the highest weight path on a WDAG
2. Convert (by hand!) the example WDAG provided, run your program on that.
3. Write another program to read in a genome and scoring scheme, generate a linked list (simple WDAG, edges = nucleotides)
4. Run your program from Step 1 on the linked list

# Formatting WDAGS

V (vname1)

V (vname2)

V (vname3)

....

E (ename1) (vname1) (vname2) (edgeweight)

E (ename2) (vname2) (vname5) (edgeweight)

E (ename3) (vname3) (vname5) (edgeweight)

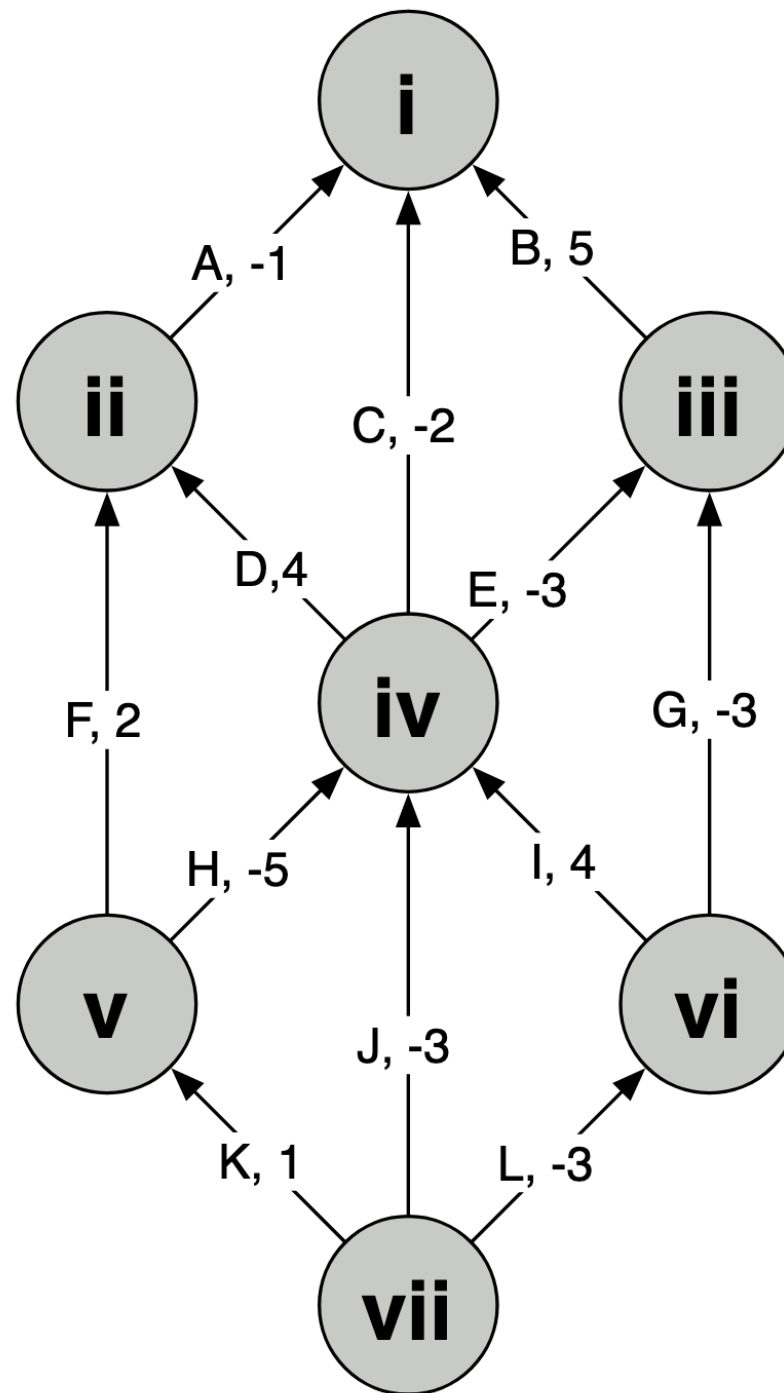
...

# Formatting WDAGS

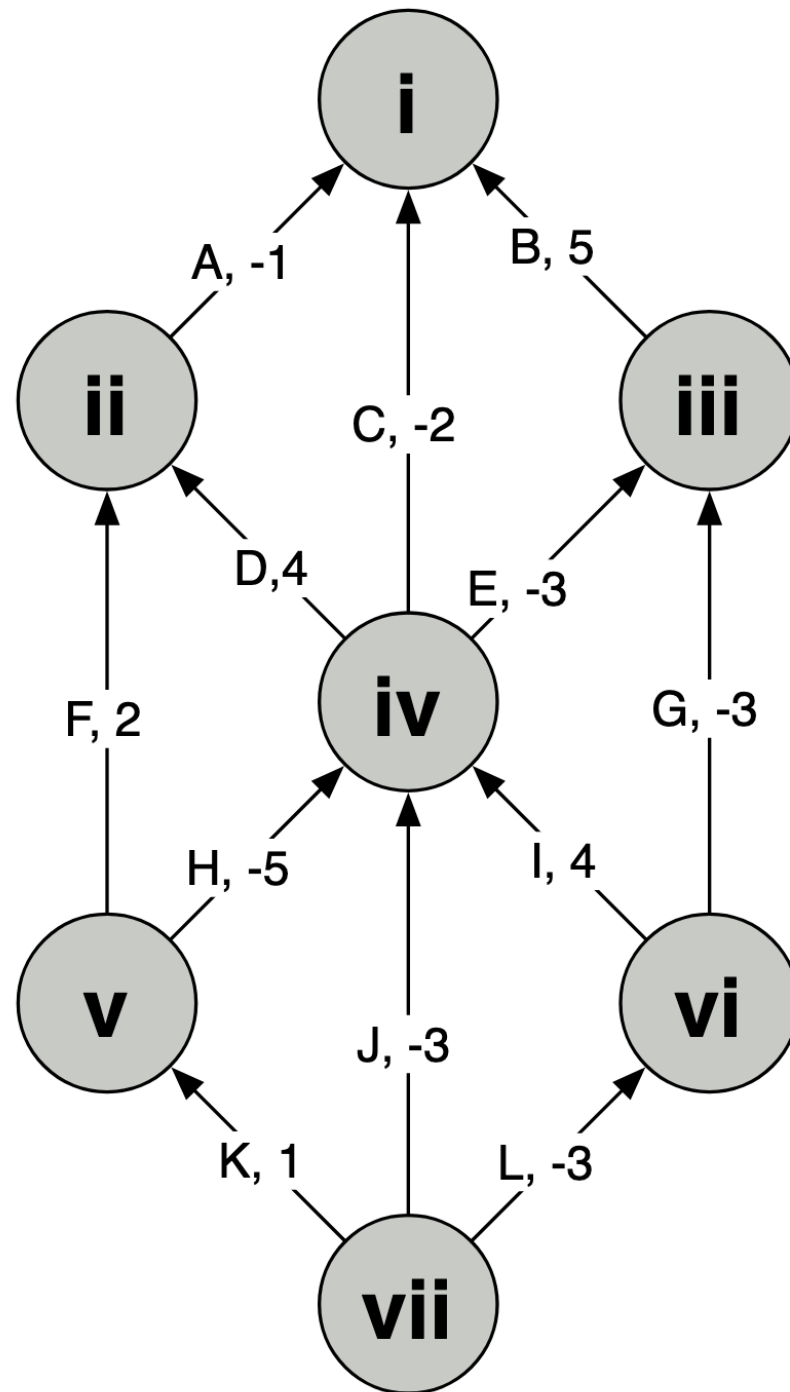
Careful! Are we assuming parent nodes to precede children in the WDAG file?

-> Yes! So make sure this is true when you generate your own WDAG files!

Example!  
Formatting

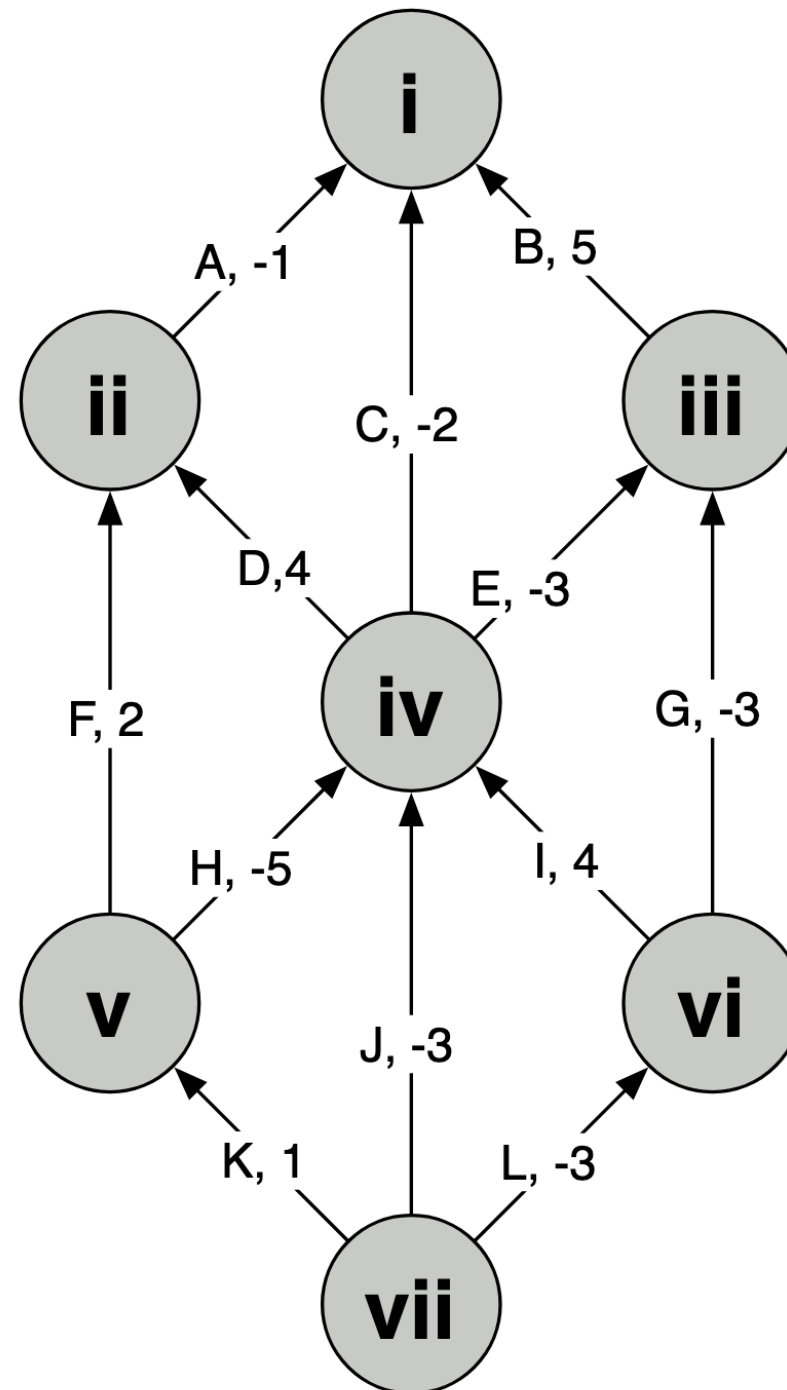


# Example! Dynamic Programming



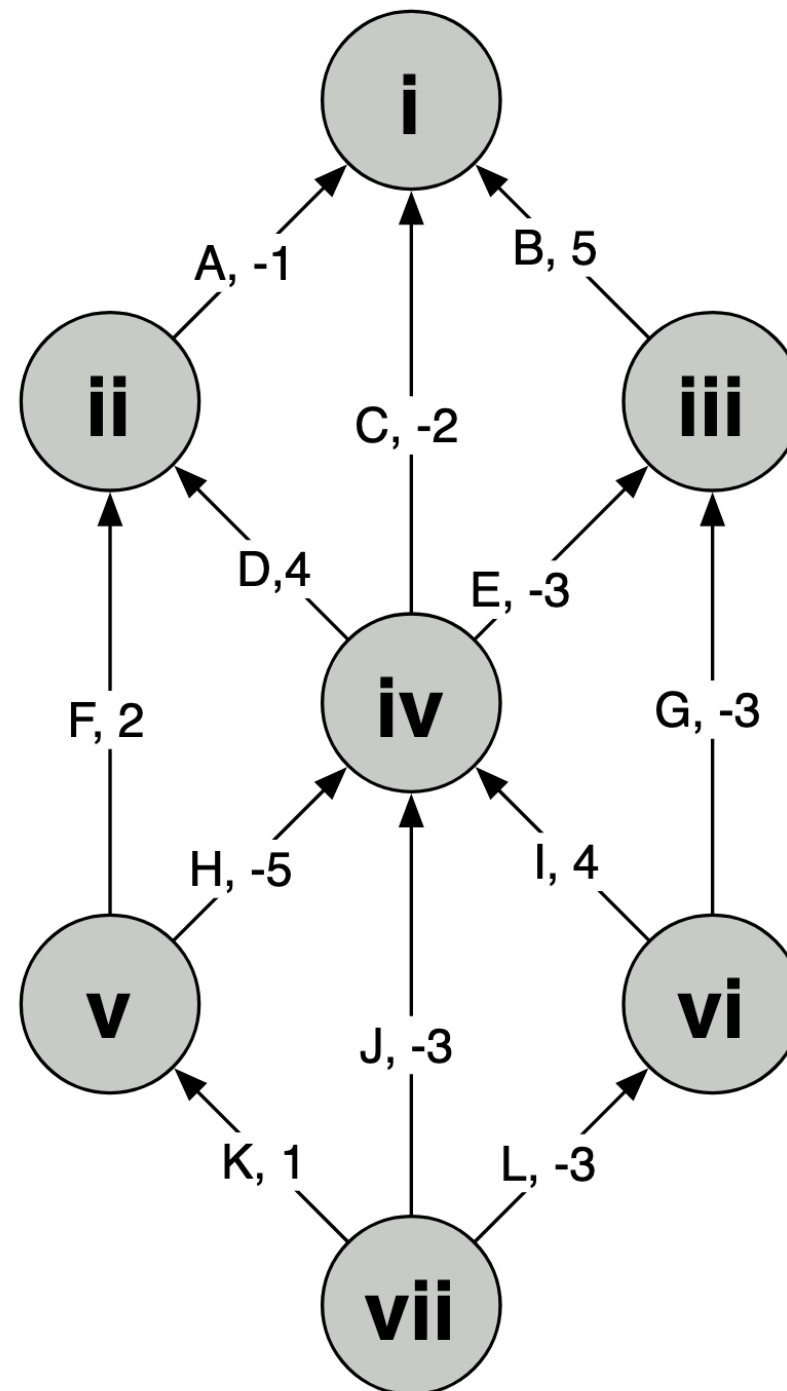
What is the highest weight path if the ends are constrained to (vii) and (i)?

What is the highest weight path if the ends are unconstrained?

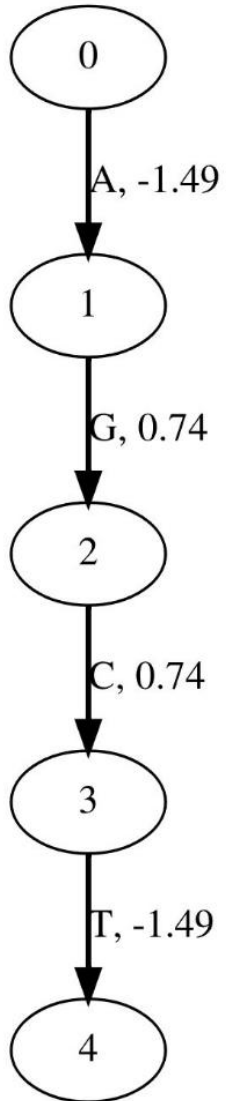


How should you structure your data in your program?

What would be useful?  
An easy way to go from child to parent? Parent to child? From node to incoming edges? Outgoing edges?



# Linked List from DNA



Sequence: AGCT

Scoring File:

A = -1.49

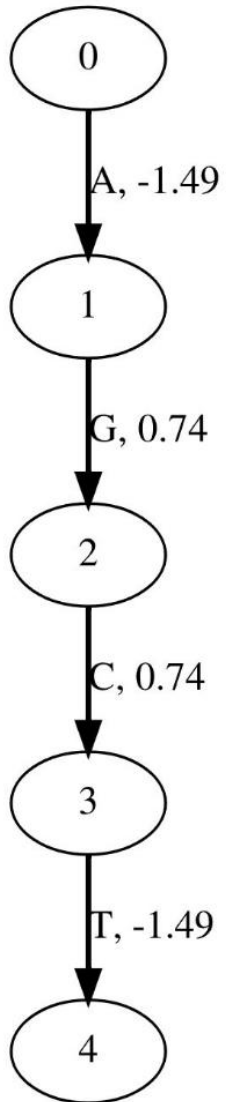
T = -1.49

C = 0.74

G = 0.74

What will the  
corresponding WDAG  
file look like?

# Linked List from DNA



What do arbitrary paths on this graph look like?

What does the scoring file tell us about the kinds of sequences we're looking for?