

#### MONASH INFORMATION TECHNOLOGY

# FIT2004 Algorithms and Data Structures

Ian Wern Han Lim lim.wern.han@monash.edu

Referencing materials by Nathan Companez, Aamir Cheema, Arun Konagurthu and Lloyd Allison





# Faculty of Information Technology, Monash University

#### COMMONWEALTH OF AUSTRALIA

Copyright Regulations 1969

This material has been reproduced and communicated to you by or on behalf of Monash University pursuant to Part VB of the Copyright Act 1968 (the Act). The material in this communication may be subject to copyright under the Act. Any further reproduction or communication of this material by you may be the subject of copyright protection under the Act. Do not remove this notice



Ready?

# **Agenda**

Directed Acyclic Graph (DAG)



#### Agenda

- Directed Acyclic Graph (DAG)
- Topological Sort
  - Kahn's algorithm
  - Depth-First Search (DFS) modification





Let us begin...



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG
- What is a DAG?



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG
- What is a DAG?
  - Directed



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG
- What is a DAG?
  - Directed
  - Acyclic



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG
- What is a DAG?
  - Directed directed edeg
  - Acyclic no cyclic
  - and of course it is a Graph...



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG
  - Can you give me an example of a real world DAG?
- What is a DAG?
  - Directed
  - Acyclic
  - and of course it is a Graph...

#### What is it?



- Graphs are very commonly used to model real world scenario
  - One of which is a DAG
  - Can you give me an example of a real world DAG?
  - Can you give me an example of a real world non-DAG?

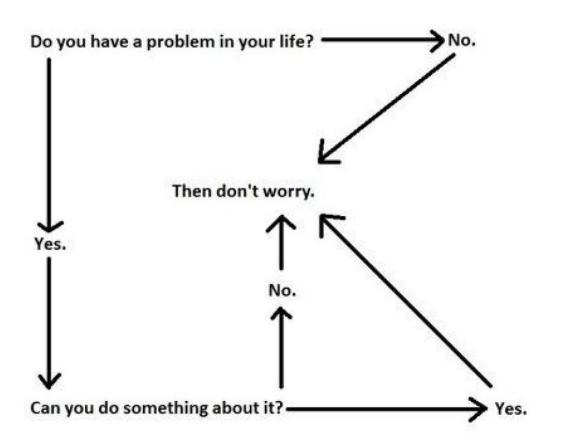
have cycle

- What is a DAG?
  - Directed
  - Acyclic
  - and of course it is a Graph...

#### What is it?



#### A real world DAG



What is it?

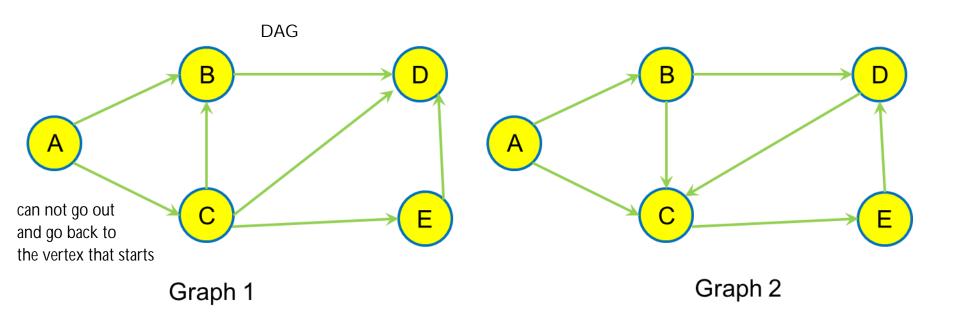
A real world not DAG







- Which graph is a DAG?
  - And where is the cycle?





# Questions?

#### What is it for?



Prerequisite mapping

#### What is it for?



- Prerequisite mapping
  - If I have a directed edge from A to B, this means I need A before B

#### What is it for?



- Prerequisite mapping
  - If I have a directed edge from A to B, this means I need A before B
    - Common in project management
    - Common in talent/skill trees!



#### What is it for?



#### Prerequisite mapping

- If I have a directed edge from A to B, this means I need A before B
  - Common in project management
  - Common in talent/skill trees!
  - Your university units!
    - Pass 1045
    - Pass 1008
    - Pass 2004
    - Pass 3155

#### What is it for?



So for an edge <A,B>

#### What is it for?



- So for an edge <A,B>
  - A is a prerequisite for B
  - A is an ancestor of B
  - A is the subset of B
  - A is ordered before B

#### What is it for?



- So for an edge <A,B>
  - A is a prerequisite for B
  - A is an ancestor of B
  - A is the subset of B
  - A is ordered before B
    - This enable us to sort a DAG



# Questions?

# Ordering of Vertices



- A topological sort
  - 排列
  - Permutation of vertices in a DAG
  - Vertex U will appear before vertex V if we have edge <U,V>

order the one unit in front that is pre-requisite of the most of the units

#### Ordering of Vertices



#### A topological sort

- Permutation of vertices in a DAG
- Vertex U will appear before vertex V if we have edge <U,V>
- Vertex U will appear before vertex W is we have edge <U,W>

#### Ordering of Vertices



#### A topological sort

- Permutation of vertices in a DAG
- Vertex U will appear before vertex V if we have edge <U,V>
- Vertex U will appear before vertex W is we have edge <U,W>
- But if we don't have edge <V,W> then V and W are of the same order

#### Ordering of Vertices



#### A topological sort

- Permutation of vertices in a DAG
- Vertex U will appear before vertex V if we have edge <U,V>
  - U<V</li>
- Vertex U will appear before vertex W is we have edge <U,W>
  - U<W</li>
- But if we don't have edge <V,W> then V and W are of the same order
  - **V==W** U V W or U W V

# Ordering of Vertices

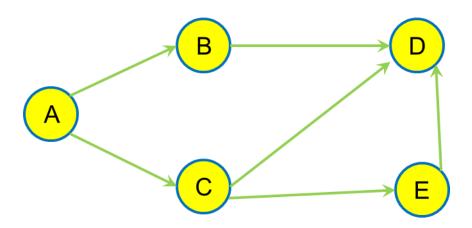


- A topological sort
  - Permutation of vertices in a DAG
  - Vertex U will appear before vertex V if we have edge <U,V>
    - U<V</li>
  - Vertex U will appear before vertex W is we have edge <U,W>
    - U<W</li>
  - But if we don't have edge <V,W> then V and W are of the same order
    - V==W
- So we have a DAG of your units
- Topological sort of this DAG gives you the order of units to take!

# Ordering of Vertices



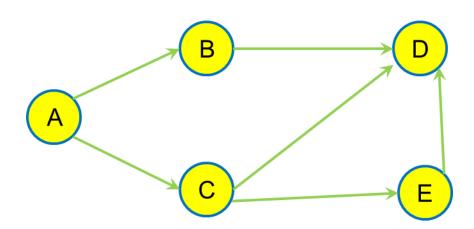
Which one of these are not a valid topological sort of the DAG?



# Ordering of Vertices



- Which one of these are not a valid topological sort of the DAG?
  - 1. A, B, C, E, D
  - 2. A ,C, B, E, D
  - 3. A, C, E, B, D E can be before B, as long as B before D
  - 4. A, B, **E**, C, D



# Ordering of Vertices



- Topological sort can be done via
  - Kahn's algorithm
  - A modified DFS



# Questions?

# Kahn's Algorithm

# For topological sort



What is the concept like?

# Kahn's Algorithm

# For topological sort



- What is the concept like?
  - Start with vertices without incoming edges



- What is the concept like?
  - Start with vertices without incoming edges
  - Delete all outgoing edges from the vertex



- What is the concept like?
  - Start with vertices without incoming edges
  - Delete all outgoing edges from the vertex
  - Add in vertices without incoming edges



- What is the concept like?
  - Start with vertices without incoming edges
  - Delete all outgoing edges from the vertex
  - Add in vertices without incoming edges
  - Repeat!

# For topological sort



#### Algorithm as follow

```
sorted_list = []
  process = []
  add all vertices without incoming edges to process
□ while len(process) > 0:
      vertex u = process.pop()
      sorted list.append(vertex u)
      for edge in vertex_u.edges:
          remove edge from graph
          if edge.vertex v has no incoming edges:
              process.append(vertex v)
⊡ if graph still has edges:
      print("Error. Not a cycle")
⊟ else:
      print(sorted list)
```

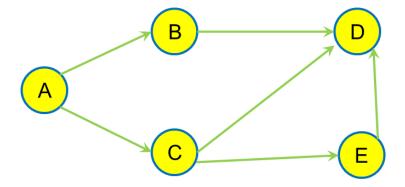
remove outgoing edge only add vertex that has no incoming edge



- Algorithm as follow
  - Let us try it out

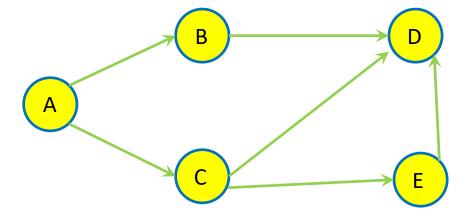
```
sorted_list = []
  process = []
  add all vertices without incoming edges to process
□ while len(process) > 0:
      vertex u = process.pop()
      sorted list.append(vertex u)
      for edge in vertex_u.edges:
          remove edge from graph
          if edge.vertex_v has no incoming edges:
              process.append(vertex v)
∃ if graph still has edges:
      print("Error. Not a cycle")
⊟ else:
      print(sorted_list)
```

- Algorithm as follow
  - Let us try it out



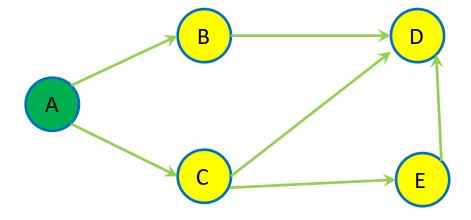
```
process process
```

- Algorithm as follow
  - Let us try it out



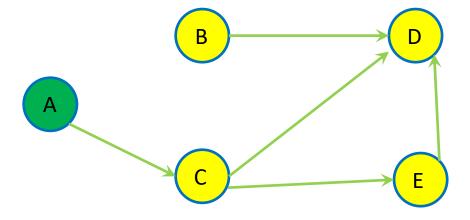
sorted			
process	А		

- Algorithm as follow
  - Let us try it out



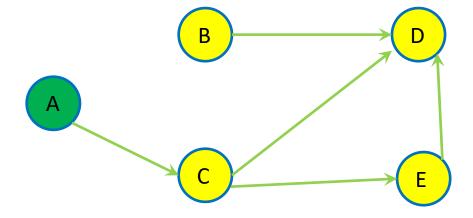
sorted	A		
process			

- Algorithm as follow
  - Let us try it out



```
sorted A process
```

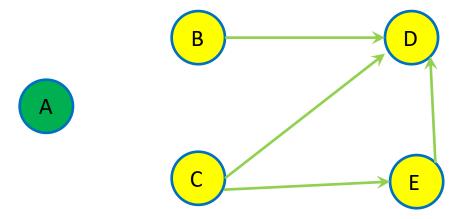
- Algorithm as follow
  - Let us try it out



```
sortedAprocessB
```

# For topological sort

- Algorithm as follow
  - Let us try it out

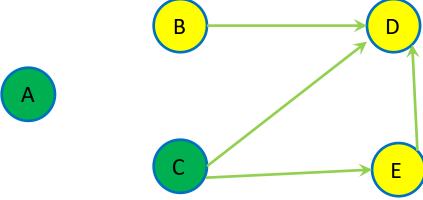


sorted	Α			
process	В	С		

if queue, B serve first if stact, C serve first

# For topological sort

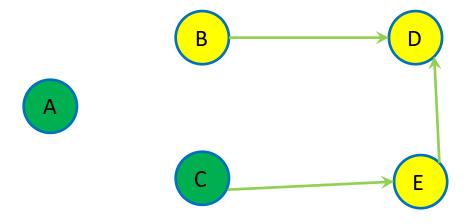
- Algorithm as follow
  - Let us try it out



delete outgoing edge from sorted element: C

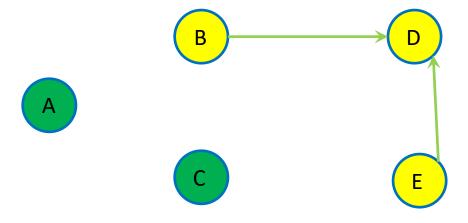
sorted	A	С		
process	В			

- Algorithm as follow
  - Let us try it out



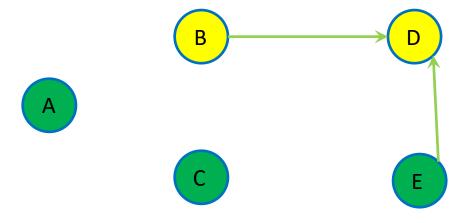
```
sortedACprocessB
```

- Algorithm as follow
  - Let us try it out



```
sortedACprocessBE
```

- Algorithm as follow
  - Let us try it out



```
sortedACEprocessB
```

- Algorithm as follow
  - Let us try it out









sorted	A	С	E	
process	В			

- Algorithm as follow
  - Let us try it out









sorted	A	С	E	В	
process					

- Algorithm as follow
  - Let us try it out











sorted	Α	С	E	В	
process	D				

# For topological sort

- Algorithm as follow
  - Let us try it out











not unique, can have different order of elements size of process > 1, then not unique

sorted	A	С	E	В	D
process					

choose either B or C to be with either queue or stack to be unique



- Algorithm as follow
  - Seemed simple right?
  - Let us zoomed in to the algorithm more



- Algorithm as follow
  - Seemed simple right?
  - Let us zoomed in to the algorithm more

```
# for output
  sorted list = []
 # tracks number of incoming edges
  incoming edges = [0] * len(vertices)

        ∃ for edge in graph:

                               # edge <u,v>
      incoming edges[vertex_v] += 1
  # process queue or stack
  process = []

    □ for vertex v in incoming edges:
      if incoming edges[vertex v] == 0:
          process.append(vertex v)
  # kahn's
⊟ while len(process) > 0:
      vertex u = process.pop()
      sorted list.append(vertex u)
      for edge in vertex u.edges:
          incoming edges[edge.vertex v] -= 1
          if incoming edges[vertex_v] == 0:
              process.append(vertex v)
  # results
☐ if graph still has edges:
      print("Error. Not a cycle")
∃ else:
      print(sorted list)
```



- Algorithm as follow
  - Seemed simple right?
  - Let us zoomed in to the algorithm more
- Complexity?

```
# for output
  sorted list = []
# tracks number of incoming edges
  incoming edges = [0] * len(vertices)

        ∃ for edge in graph:

                               # edge <u,v>
      incoming edges[vertex v] += 1
  # process queue or stack
  process = []

    □ for vertex v in incoming edges:
      if incoming edges[vertex v] == 0:
          process.append(vertex v)
  # kahn's
⊟ while len(process) > 0:
      vertex u = process.pop()
      sorted list.append(vertex u)
      for edge in vertex u.edges:
          incoming edges[edge.vertex v] -= 1
          if incoming edges[vertex_v] == 0:
               process.append(vertex v)
  # results

☐ if graph still has edges:
      print("Error. Not a cycle")
∃ else:
      print(sorted list)
```

#### Kahn's Algorithm ~= BFS

## For topological sort



# Algorithm as follow

- Seemed simple right?
- Let us zoomed in to the algorithm more

O(E)

O(E)

# Complexity?

O(V+E) time space

incoming edge don't care about u (start of edge) only cares v (end of edge)

one vertex in process only has small part of all edges in graph all these edges combine = O(E) since process eventually run through all vertices so O(V+E)

|incoming edges| > 0 still has edges there was a cycle

```
# for output
  sorted list = []
  # tracks number of incoming edges
  incoming edges = [0] * len(vertices)

    ∃ for edge in graph:

                              # edge <u,v>
      incoming edges[vertex_v] += 1
  # process queue or stack
  process = []

        ∃ for vertex v in incoming edges:

      if incoming edges[vertex v] == 0:
          process.append(vertex v)
  # kahn's
vertex u = process.pop()
      sorted list.append(vertex u)
      for edge in vertex u.edges:
          incoming edges[edge.vertex v] -= 1
          if incoming_edges[vertex_v] == 0:
              process.append(vertex v)
  # results
☐ if graph still has edges:
      print("Error. Not a cycle")
∃ else:
      print(sorted list)
```



# Questions?

# Modified for topological sorting



We saw the complexity of Kahn's being O(V+E)



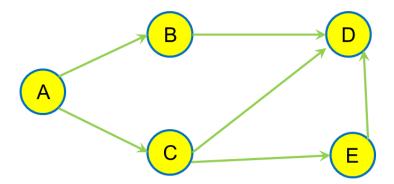
- We saw the complexity of Kahn's being O(V+E)
- Can we modify DFS to do the same?

```
28 ☐ def dfs_topological(vertex_u):
29     vertex_u.visited = True
30 ☐ for edge in vertex_u.edges:
31 ☐ if edge.vertex_v.visited == False:
32     dfs_topological(vertex_v)
```



- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see

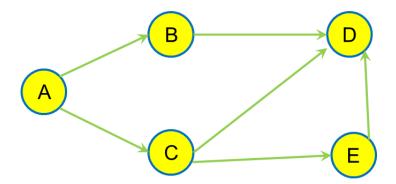
```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  fedge.vertex_v.visited == False:
32     dfs_topological(vertex_v)
```





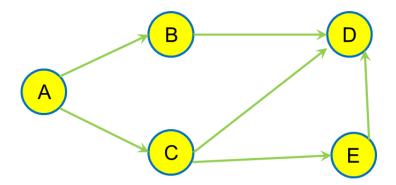
- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A

```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```



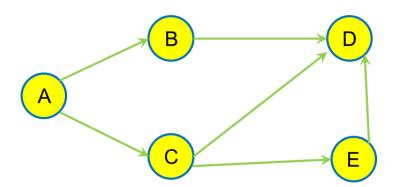


- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B





- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D



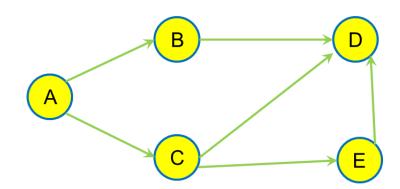
# Modified for topological sorting



- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C backtrack to A then go to C

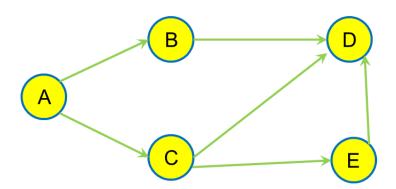
recursion

```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```



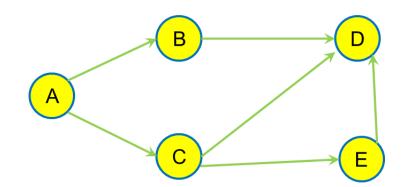


- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C
    - Go to E



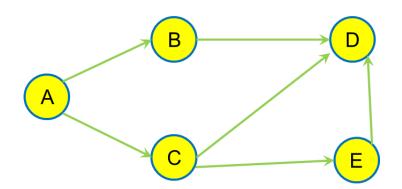


- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C
    - Go to E
    - So we have A, B, D, C, E





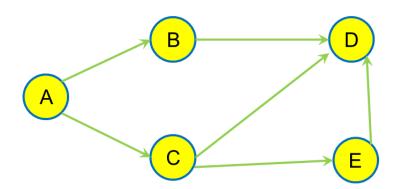
- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C
    - Go to E
    - So we have A, B, D, C, E
  - Any other DFS order?





- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C
    - Go to E
    - So we have A, B, D, C, E
  - Any other DFS order?
    - A, C, D, E, B

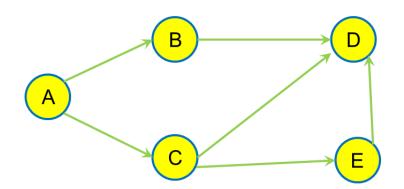
```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```





- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C
    - Go to E
    - So we have A, B, D, C, E
  - Any other DFS order?
    - A, C, D, E, B
    - A, C, E, D, B

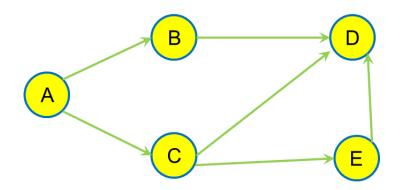
```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```





- Can we modify DFS to do the same?
  - Let us run DFS on the following graph and see
    - Start from A
    - Go to B
    - Go to D
    - Go to C
    - Go to E
    - So we have A, B, D, C, E
  - Any other DFS order?
    - A, C, D, E, B
    - A, C, E, D, B
    - A, C, E, B, D?

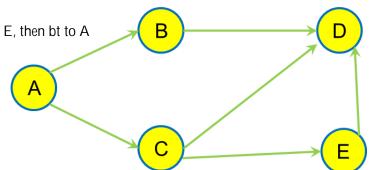
```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```





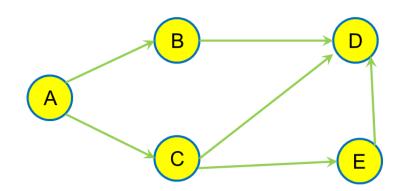
- Can we modify DFS to do the same?
  - Any other DFS order?
    - A, B, D, C, E
    - A, C, D, E, B backtrack to C from D, then go to E, then bt to A then go to B
    - A, C, E, D, B

```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```





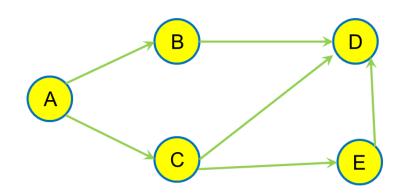
- Can we modify DFS to do the same?
  - Any other DFS order?
    - A, B, D, C, E
    - A, C, D, E, B
    - A, C, E, D, B
  - A possible topological sort
    - A, B, C, E, D
    - A, C, B, E, D





- Can we modify DFS to do the same?
  - Any other DFS order?
    - A, B, D, C, E
    - A, C, D, E, B
    - A, C, E, D, B
  - A possible topological sort
    - A, B, C, E, D
    - A, C, B, E, D
  - Notice something?

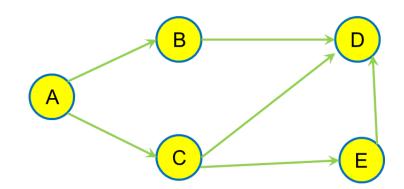
```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```





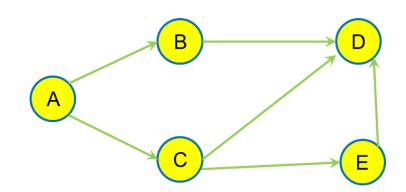
- Can we modify DFS to do the same?
  - Any other DFS order?
    - A, B, D, C, E
    - A, C, D, E, B
    - A, C, E, D, B
  - A possible topological sort
    - A, B, C, E, D
    - A, C, B, E, D
  - Notice something?
    - When we reach the end of the DFS, we go back to an earlier vertex but this vertex should be early in topological sort (such as vertex B or C)

```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```



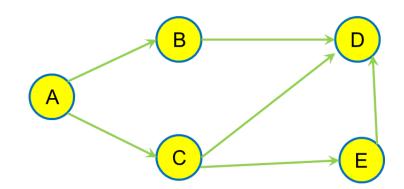


- Can we modify DFS to do the same?
  - Any other DFS order?
    - A, B, D, C, E
    - A, C, D, E, B
    - A, C, E, D, B
  - A possible topological sort
    - A, B, C, E, D
    - A, C, B, E, D
  - Notice something?
    - When we reach the end of the DFS, we go back to an earlier vertex but this vertex should be early in topological sort (such as vertex B or C)





- Can we modify DFS to do the same?
  - Any other DFS order?
    - A, B, D, C, E
    - A, C, D, E, B
    - A, C, E, D, B
  - A possible topological sort
    - A, B, C, E, D
    - A, C, B, E, D
  - Notice something?
    - When we reach the end of the DFS, we go back to an earlier vertex but this vertex should be early in topological sort (such as vertex B or C)



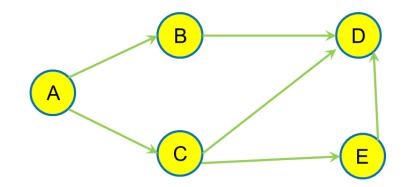


- Can we modify DFS to do the same?
  - Any other DFS order?

```
A, B, D, C, E push D, B then Push E, C
A, C, D, E, B finally A
```

- A, C, E, D, B original : A,B,C,D,E
- A possible topological sort
  - A, B, C, E, D for modified DFS
  - A, C, B, E, D for BFS
- Notice something?
  - When we reach the end of the DFS, we go back to an earlier vertex but this vertex should be early in topological sort (such as vertex B or C)
  - So how do we arrange it?

```
28  def dfs_topological(vertex_u):
29     vertex_u.visited = True
30  for edge in vertex_u.edges:
31  def dfs_topological(vertex_v)
```



## Modified for topological sorting



Here's how we modify with a stack!
last in first out

```
□ def dfs topological(vertex u):
      # start for result
      stack = []
                                                         separate into two, make sure A(root) to be pushed to stack
      # run DFS
      vertex u.visited = True
      for edge in vertex u.edges:
           if edge.vertex v.visited == False:
               dfs topological aux(vertex v,stack)
      # output
                                                             DFS
      print(stack)
□ def dfs topological aux(vertex u,stack):
                                                         edge.vertex_v = D == True
      vertex u.visited = True
                                                         only do it on A, C,E,D, B
      for edge in vertex u.edges:
                                                         push D, E, C, B, A
           if edge.vertex v.visited == False:
                                                         serve A. B. C. E. D
               dfs topological aux(vertex v)
      # add to stack
      stack.push(vertex u)
```

backtrack to here to push to stack

## Modified for topological sorting



#### Complexity?

```
□ def dfs topological(vertex u):
      # start for result
      stack = []
      # run DFS
      vertex u.visited = True
      for edge in vertex u.edges:
          if edge.vertex v.visited == False:
              dfs topological aux(vertex v,stack)
      # output
      print(stack)
□ def dfs topological aux(vertex u,stack):
      vertex u.visited = True
      for edge in vertex u.edges:
          if edge.vertex v.visited == False:
                                                      serve
              dfs topological aux(vertex v)
      # add to stack
      stack.push(vertex u)
```

not next: just mark edge.vertex\_v = TRUE

reverse the stack on purpose since it first in last out

push back in reverse order since visit: A -> B -> D -> C -> E -> push D -> B -> E -> C -> A



- Complexity?
  - O(V+E) since we only added a stack

```
□ def dfs topological(vertex u):
      # start for result
      stack = []
      # run DFS
      vertex u.visited = True
      for edge in vertex u.edges:
          if edge.vertex v.visited == False:
              dfs topological aux(vertex v,stack)
      # output
      print(stack)
□ def dfs topological aux(vertex u,stack):
      vertex u.visited = True
      for edge in vertex_u.edges:
          if edge.vertex v.visited == False:
              dfs topological aux(vertex v)
      # add to stack
      stack.push(vertex u)
```



# Questions?



# Thank You