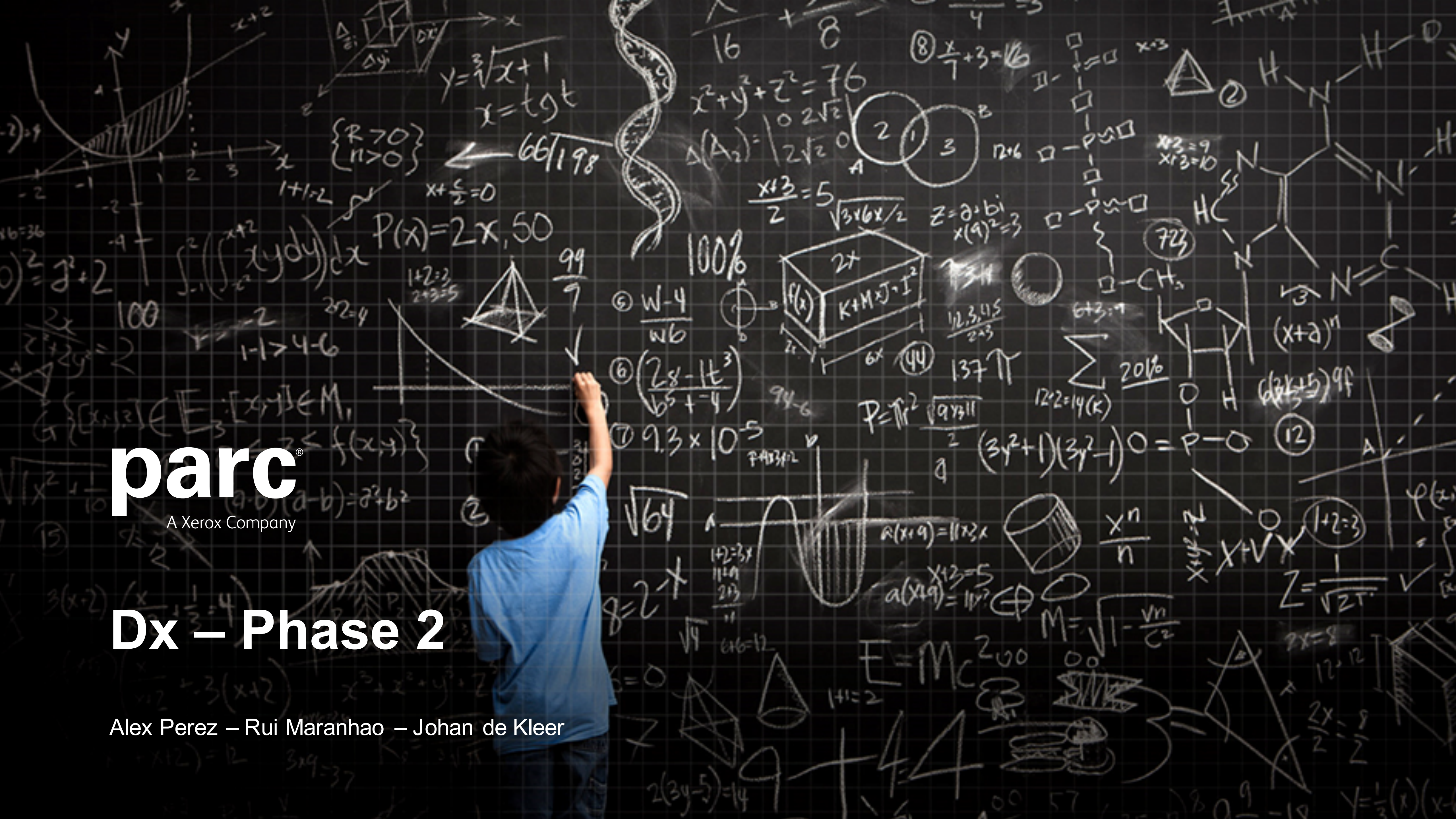


Dx – Phase 2

Alex Perez – Rui Maranhao – Johan de Kleer

Phase 2’s Goals

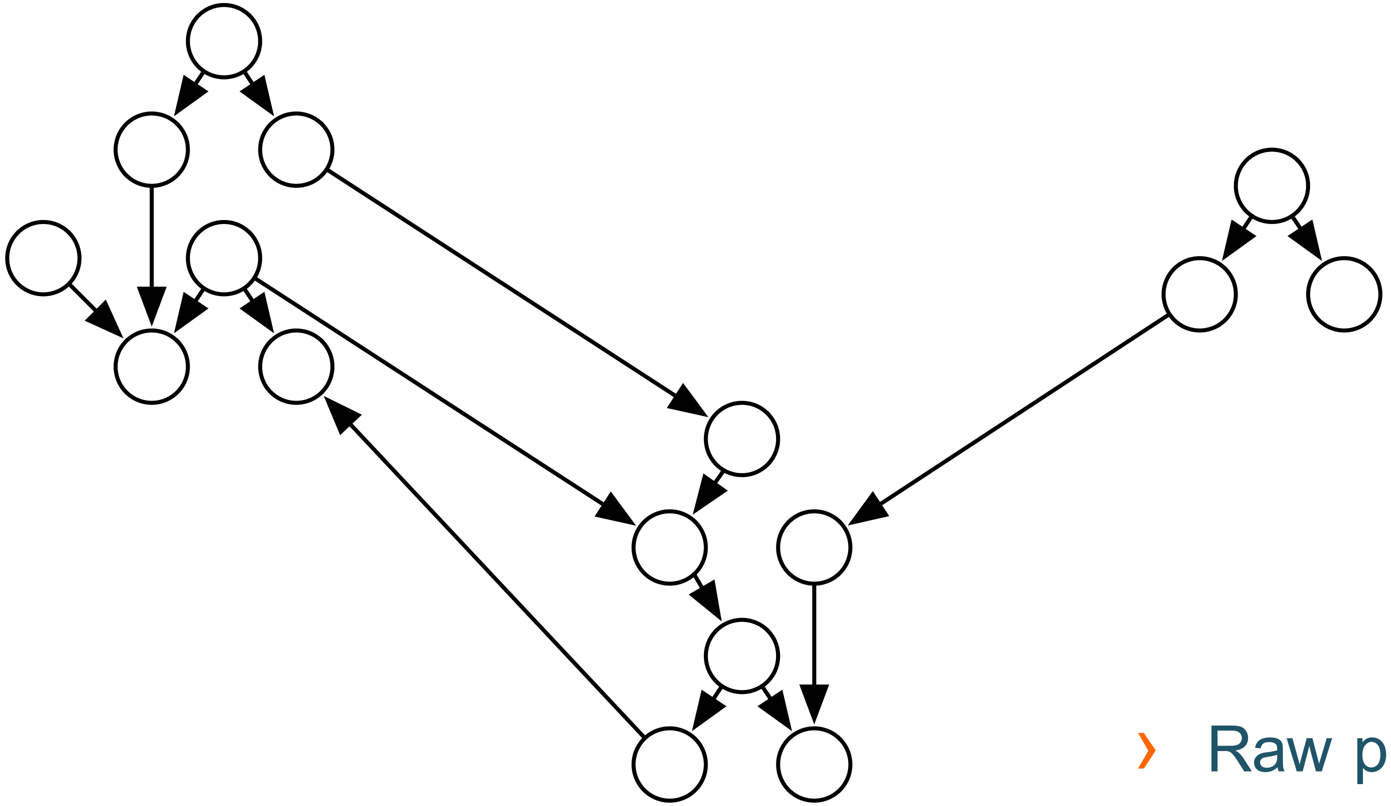
- Focus: Integration
- Run inside TC-in-a-box
  - Consume data from Titan
  - Write output to Titan
  - Input/Output Format
    - specified in the architecture document (*language.md*)
- Module tested with integration and unit tests
- Deadline: 05/06/2016 (tentatively)

ATMS-based APT Campaign Diagnosis

– and its dependencies (in bold) –

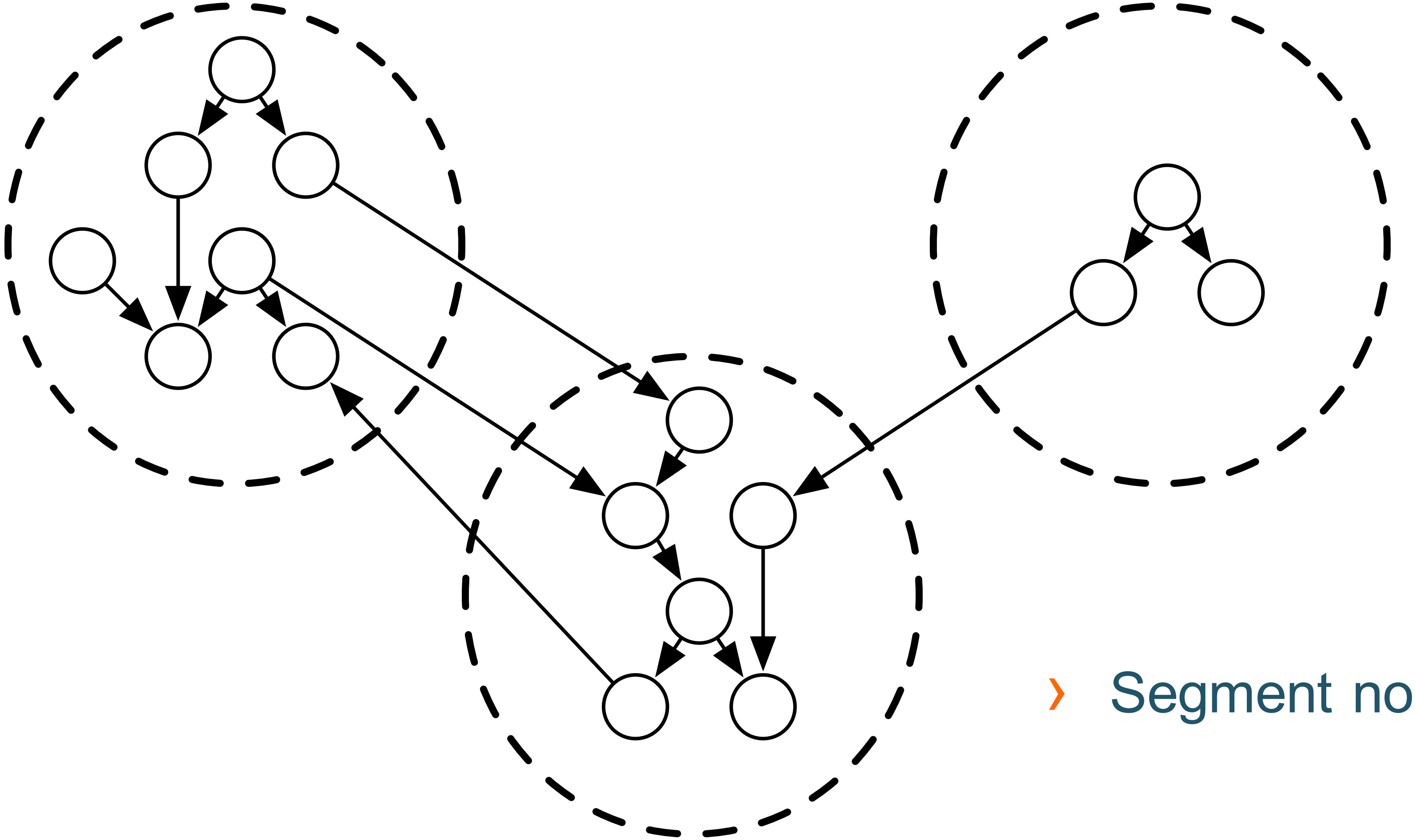
- Dx’s hybrid approach works in two phases
  1. For each segment, an ATMS is used to cache the minimal set of APT activities
     - APT activities are specified in the **APT grammar**, as stored in the **Kb**
     - The **Ac** labels segments using the APT grammar. A segment has a tuple of segments and confidence score
     - This *preprocessing* works both in forensic and real-time mode
     - Low time/space complexity
  2. Using the ATMS and provenance data (i.e., lowest granularity) confirm whether it is an APT campaign or not
     - Why zooming-in into the provenance data?
       - Segment granularity may have lost causal information
       - Graph at the segmentation level may not be a DAG
     - Use **Ac confidence scores** and **Ad scores** to rank APT campaigns

The need to zoom-in



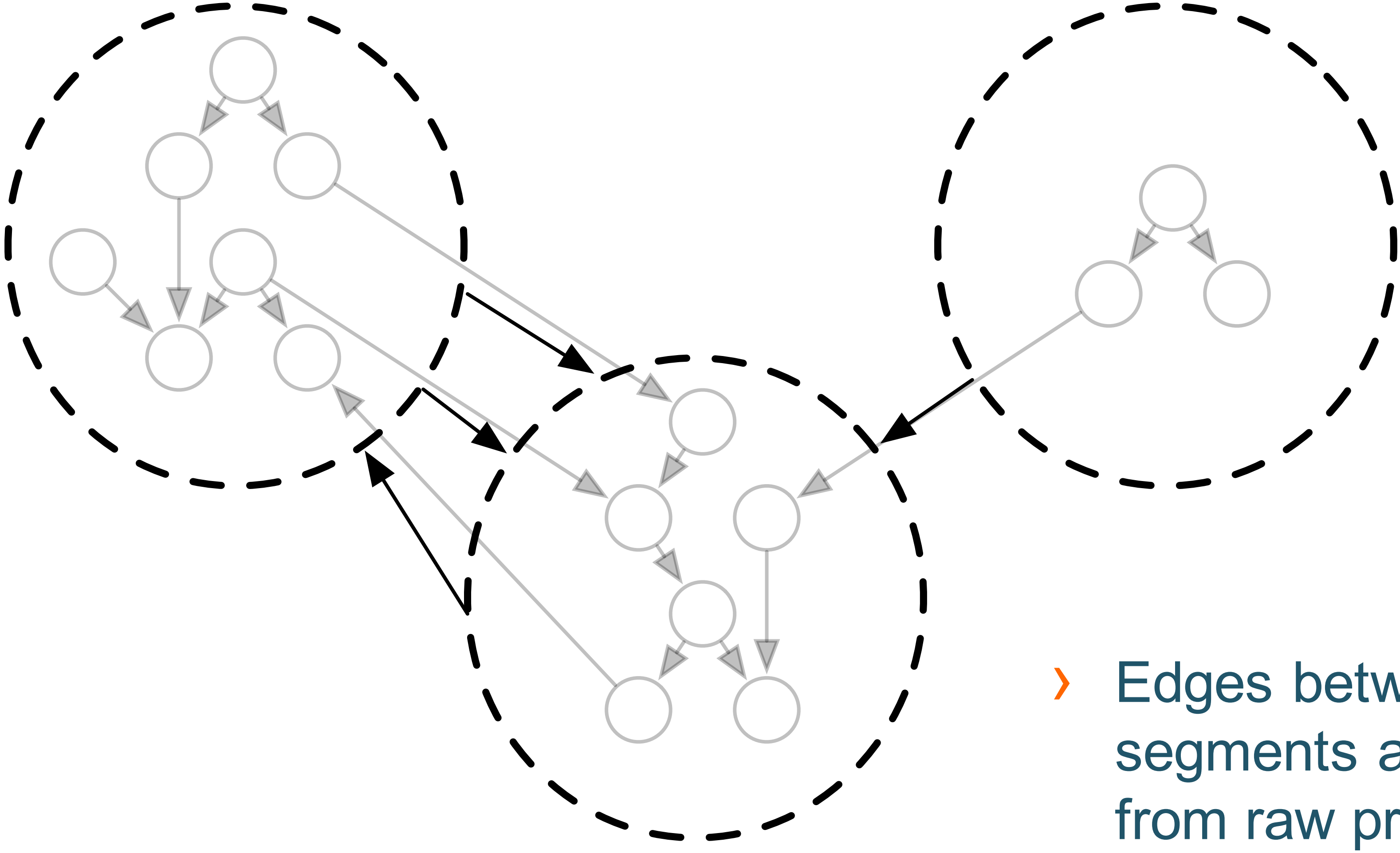
- Raw provenance graph

The need to zoom-in



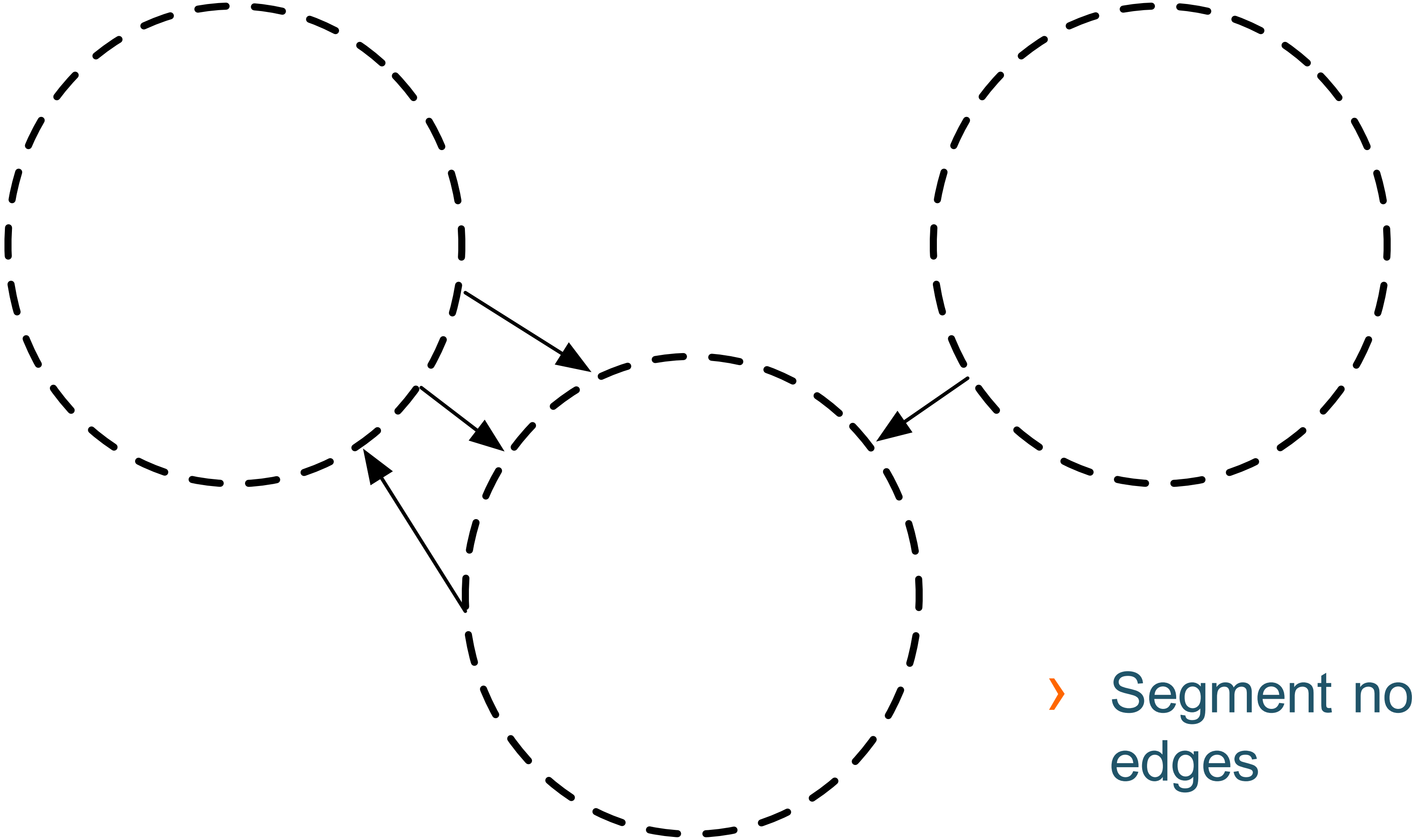
- Segment nodes

The need to zoom-in



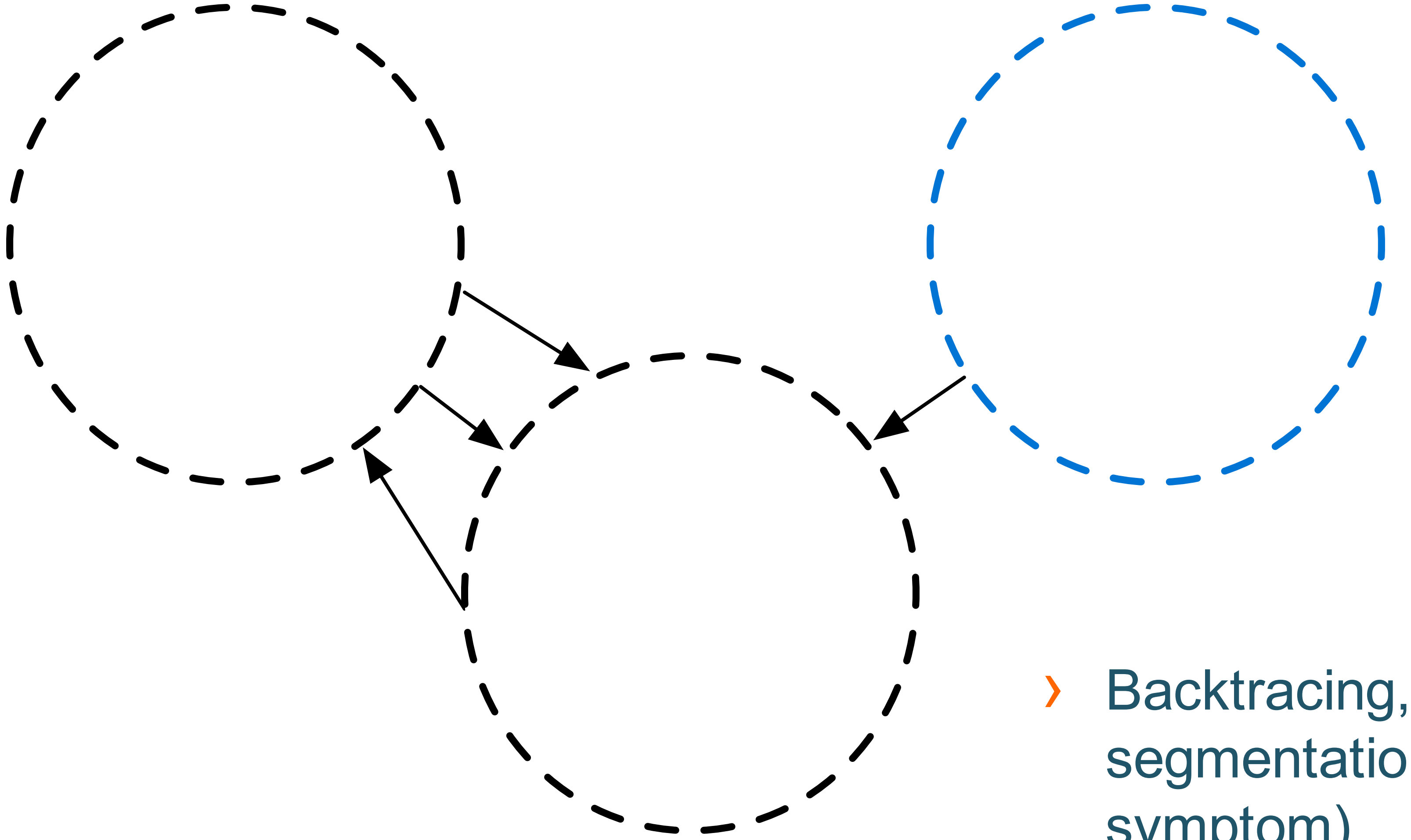
- Edges between segments are inherited from raw provenance
(abstraction may not be a DAG)

The need to zoom-in



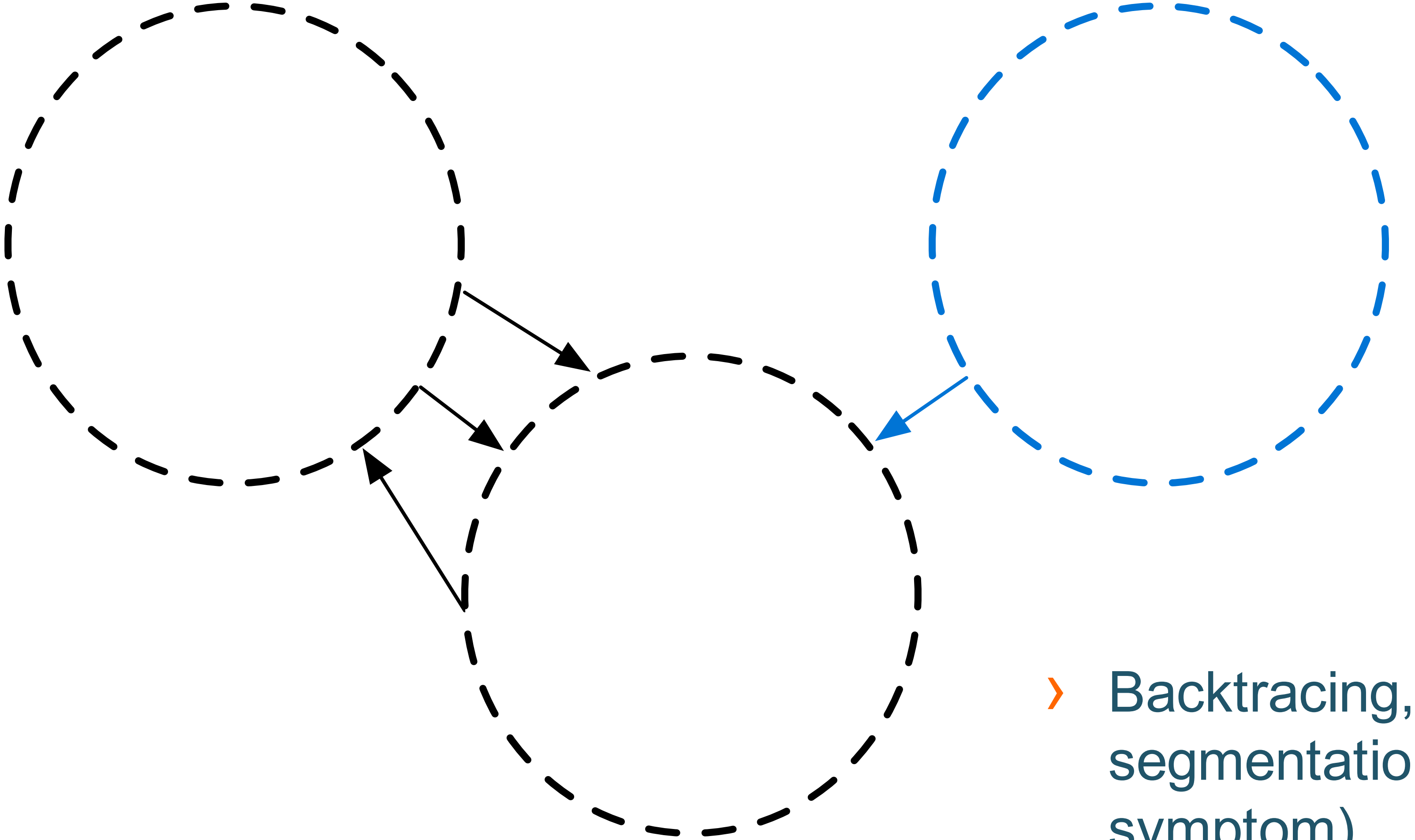
- Segment nodes and edges

The need to zoom-in



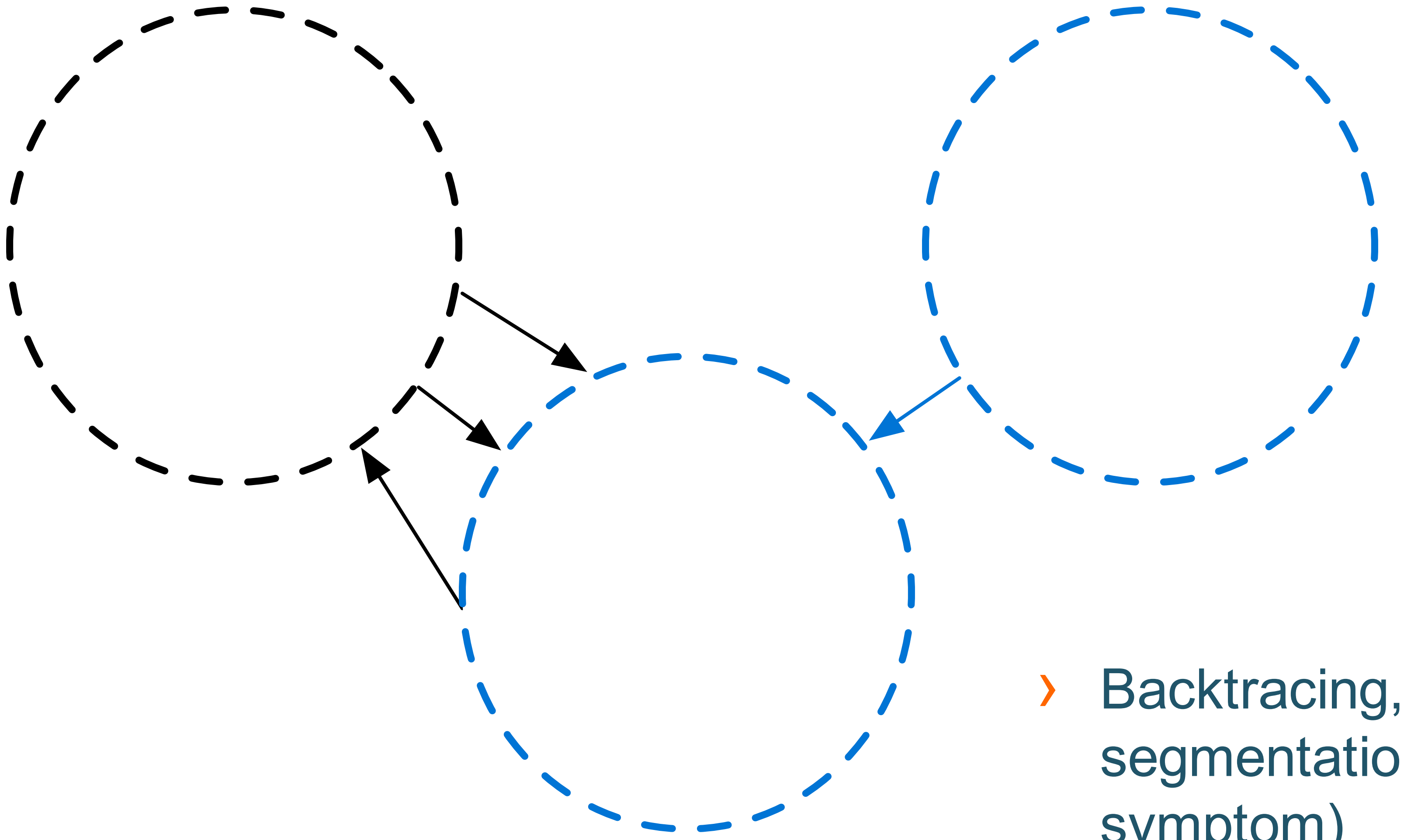
- Backtracing, given a segmentation node (e.g., symptom)

The need to zoom-in



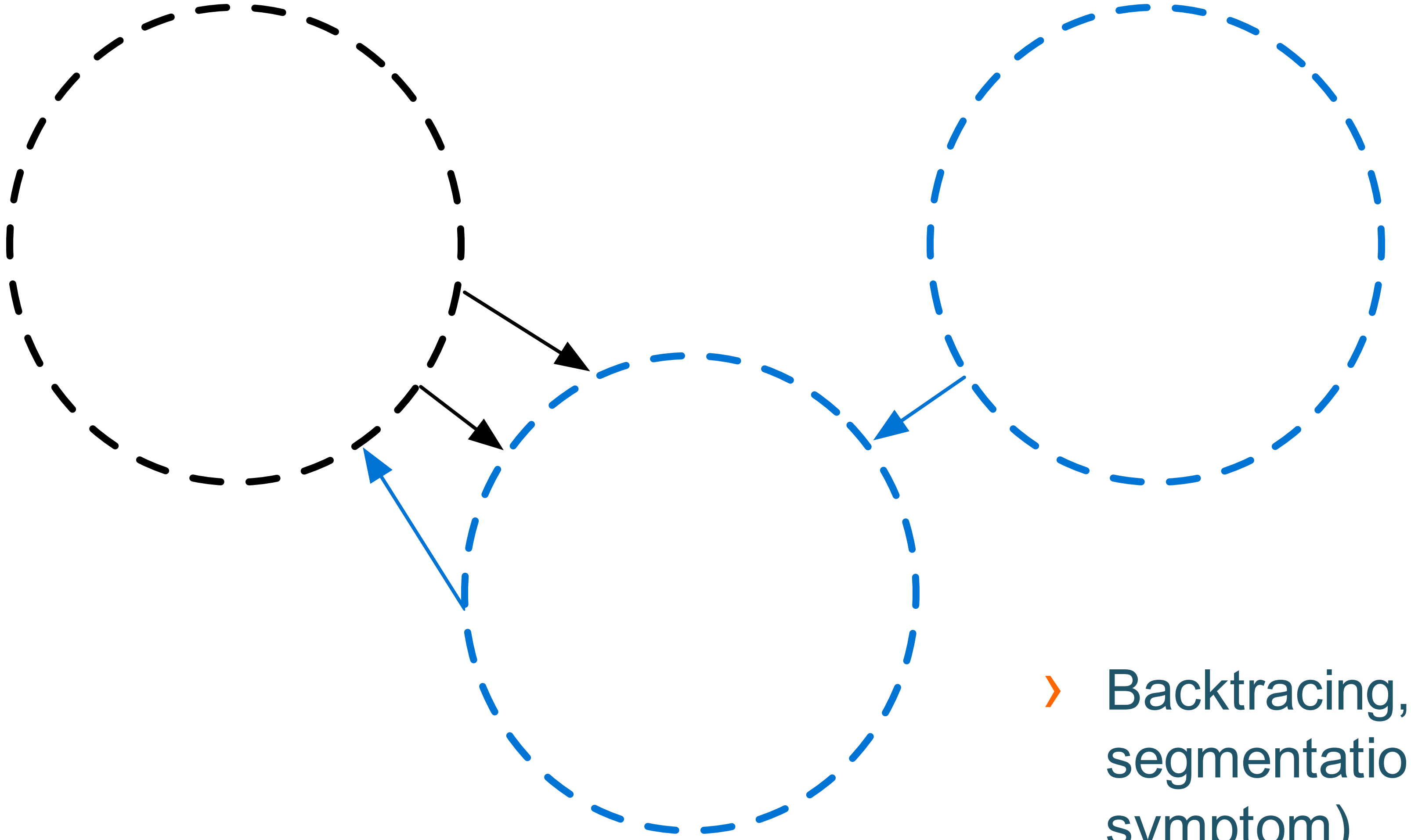
- Backtracing, given a segmentation node (e.g., symptom)

The need to zoom-in



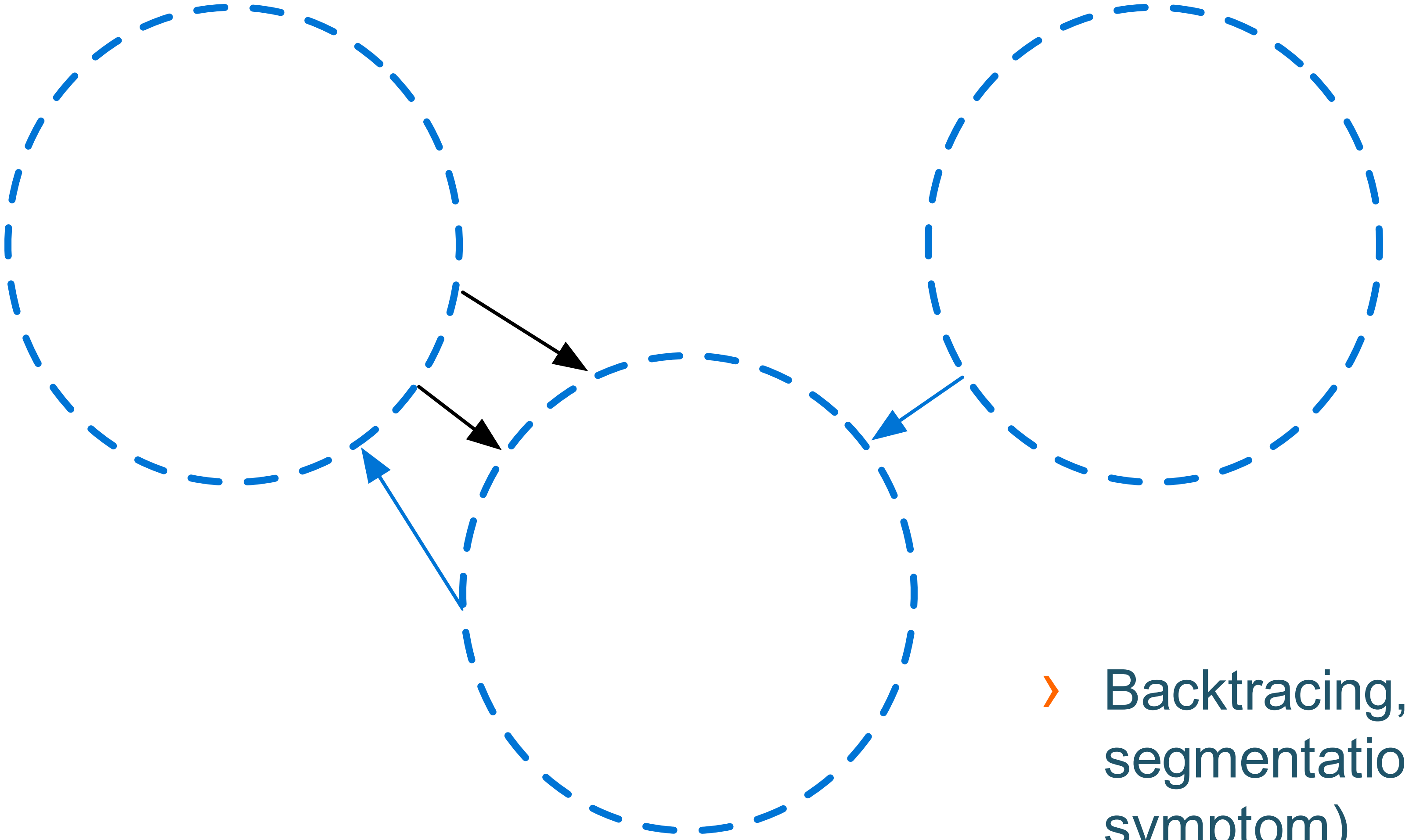
- Backtracing, given a segmentation node (e.g., symptom)

The need to zoom-in



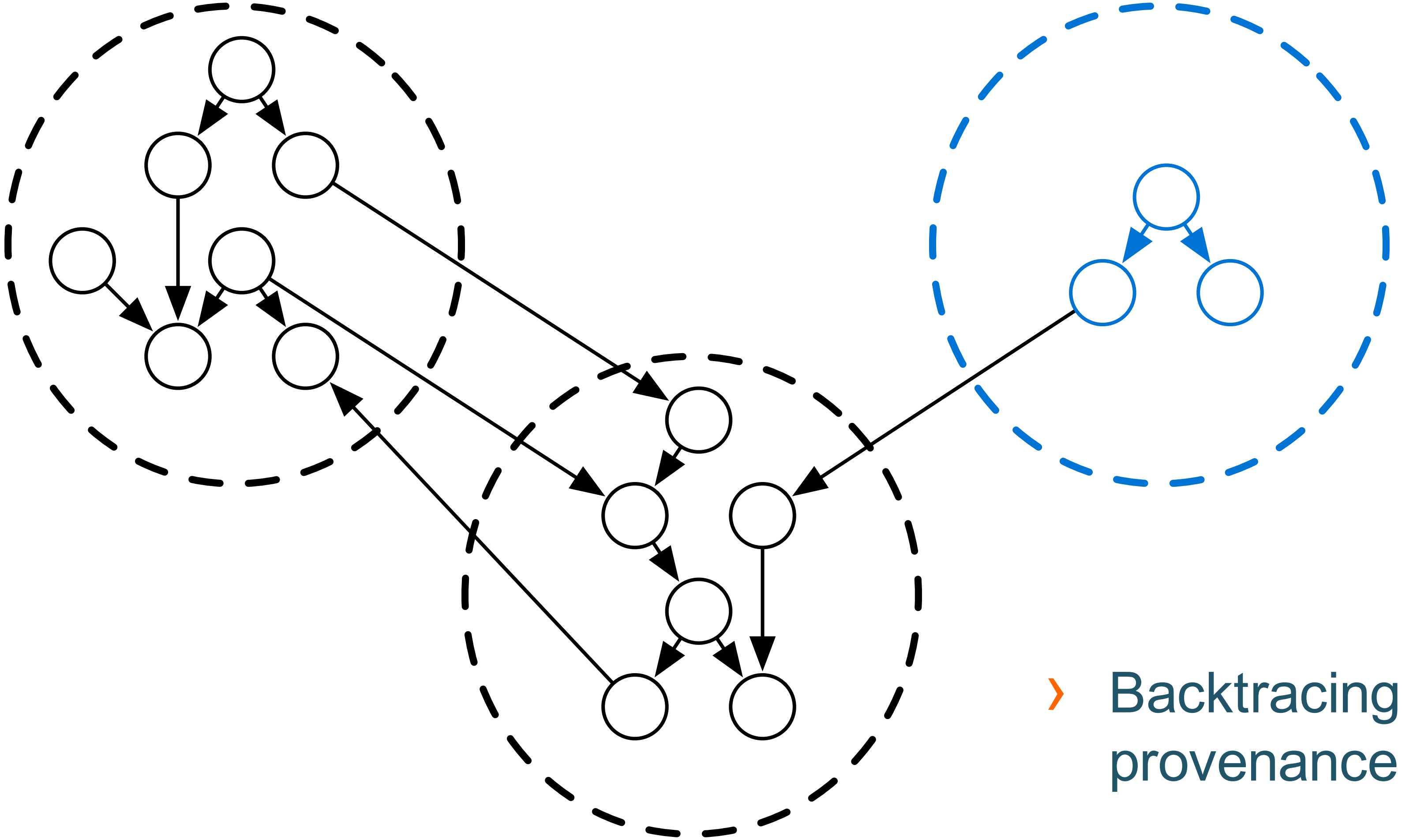
- Backtracing, given a segmentation node (e.g., symptom)

The need to zoom-in



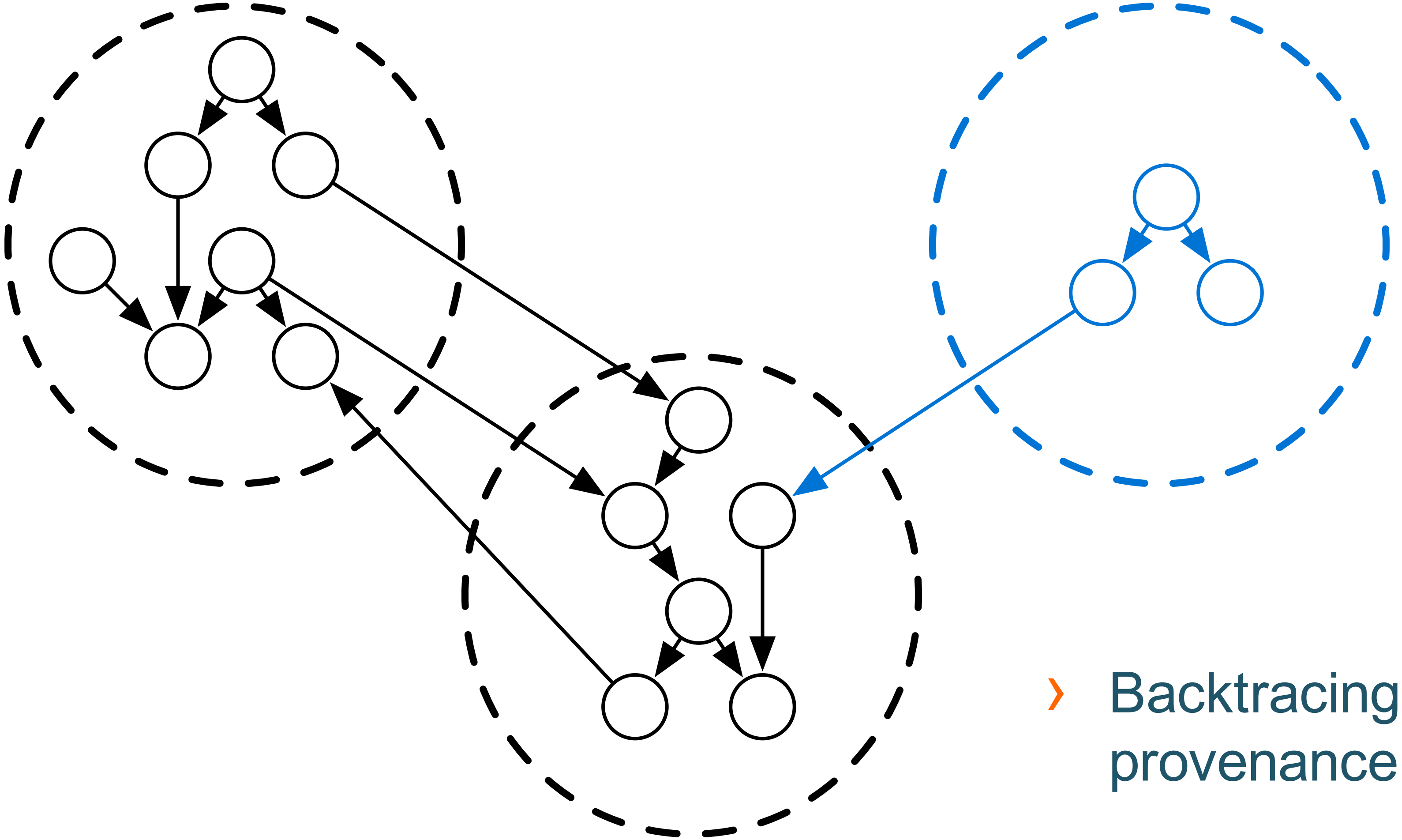
- Backtracing, given a segmentation node (e.g., symptom)

The need to zoom-in



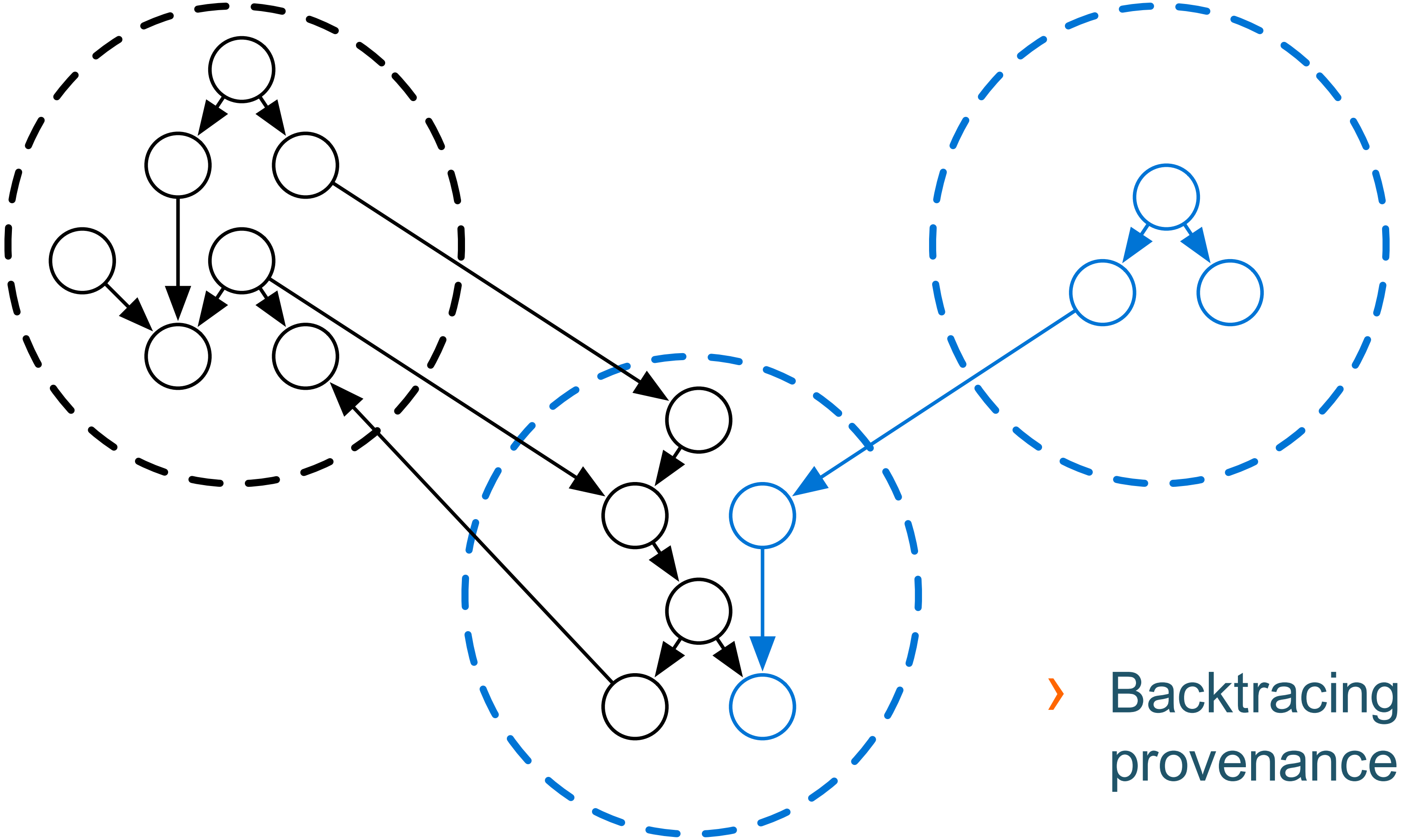
- Backtracing at the provenance level

The need to zoom-in



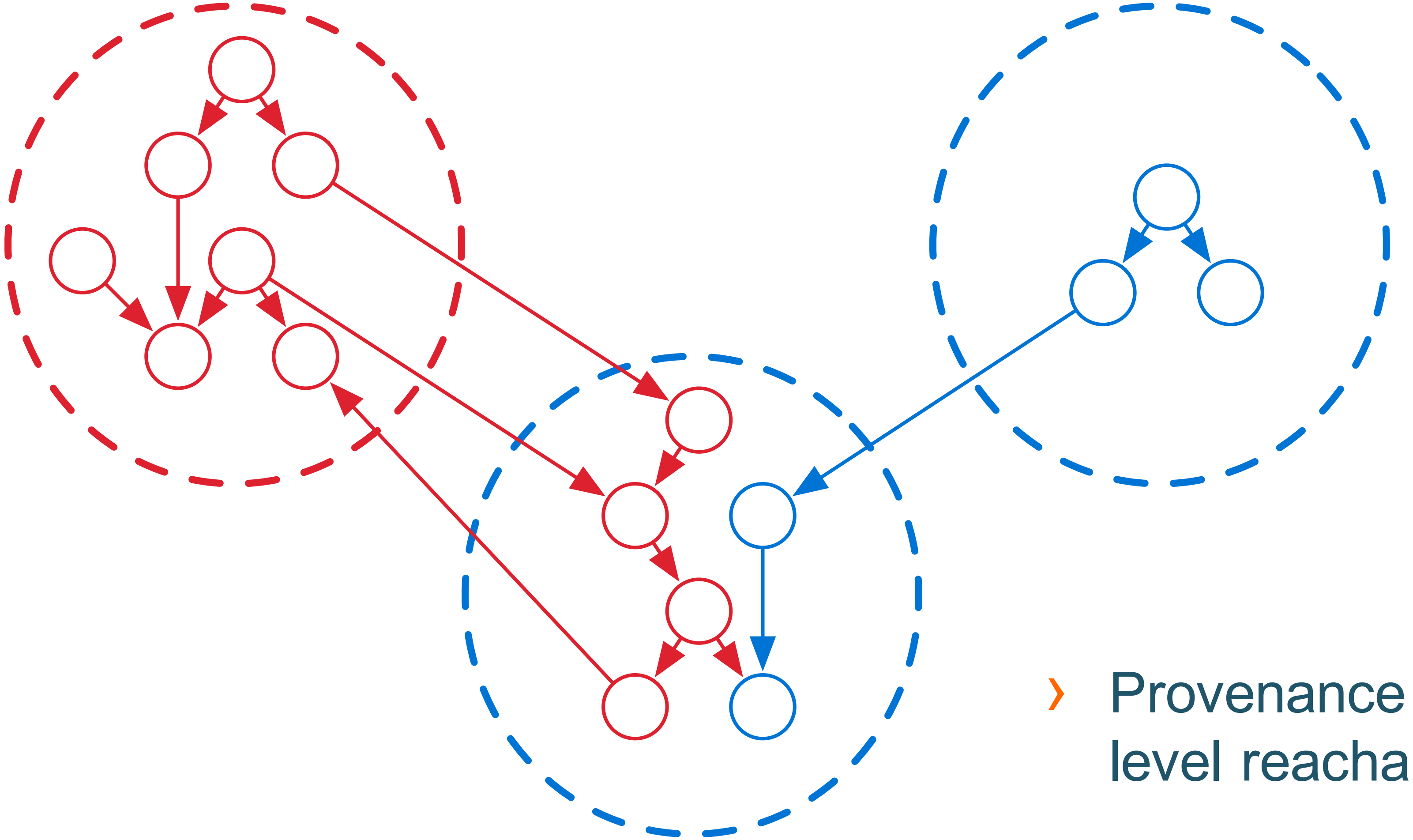
- Backtracing at the provenance level

The need to zoom-in



- Backtracing at the provenance level

The need to zoom-in

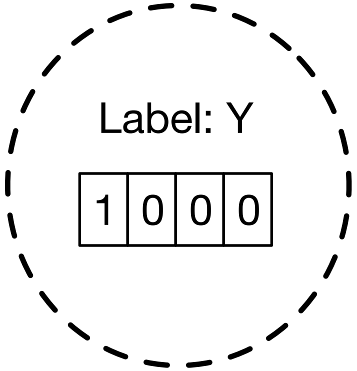


- Provenance vs. segment level reachability

© 2015 PARC, All rights reserved.

parc
A Xerox Company

DX Illustrated

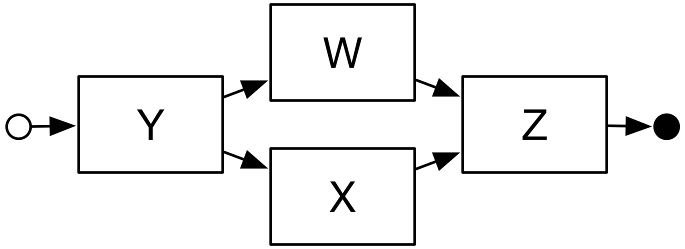


Label: Y

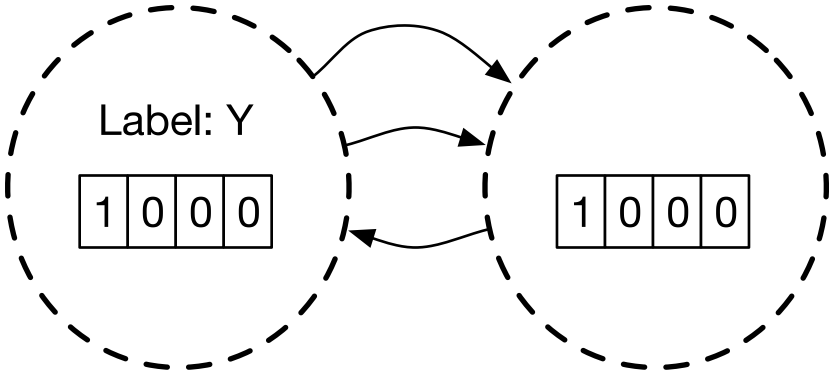
1	0	0	0

APT Grammar Y(W|X)Z encoded as

Y	W	X	Z

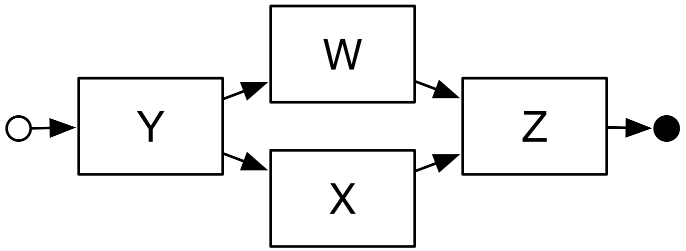


DX Illustrated



APT Grammar Y(W|X)Z encoded as

Y	W	X	Z

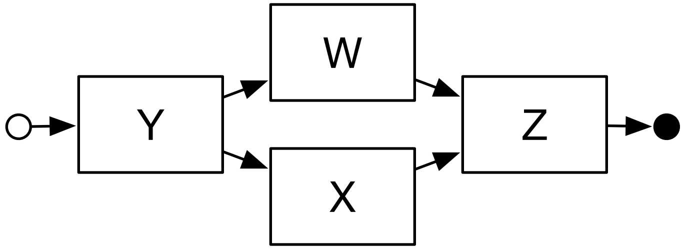


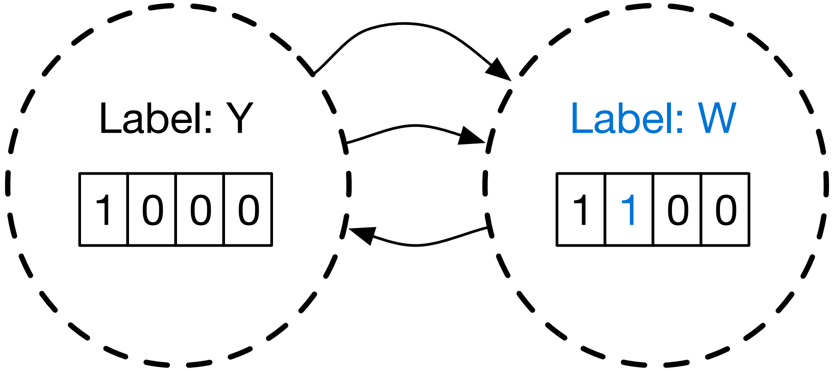
- When a new segment is observed, the parents’ *environment* is copied

DX Illustrated

APT Grammar Y(W|X)Z encoded as

Y	W	X	Z



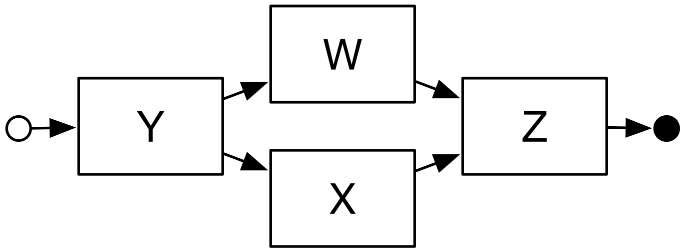


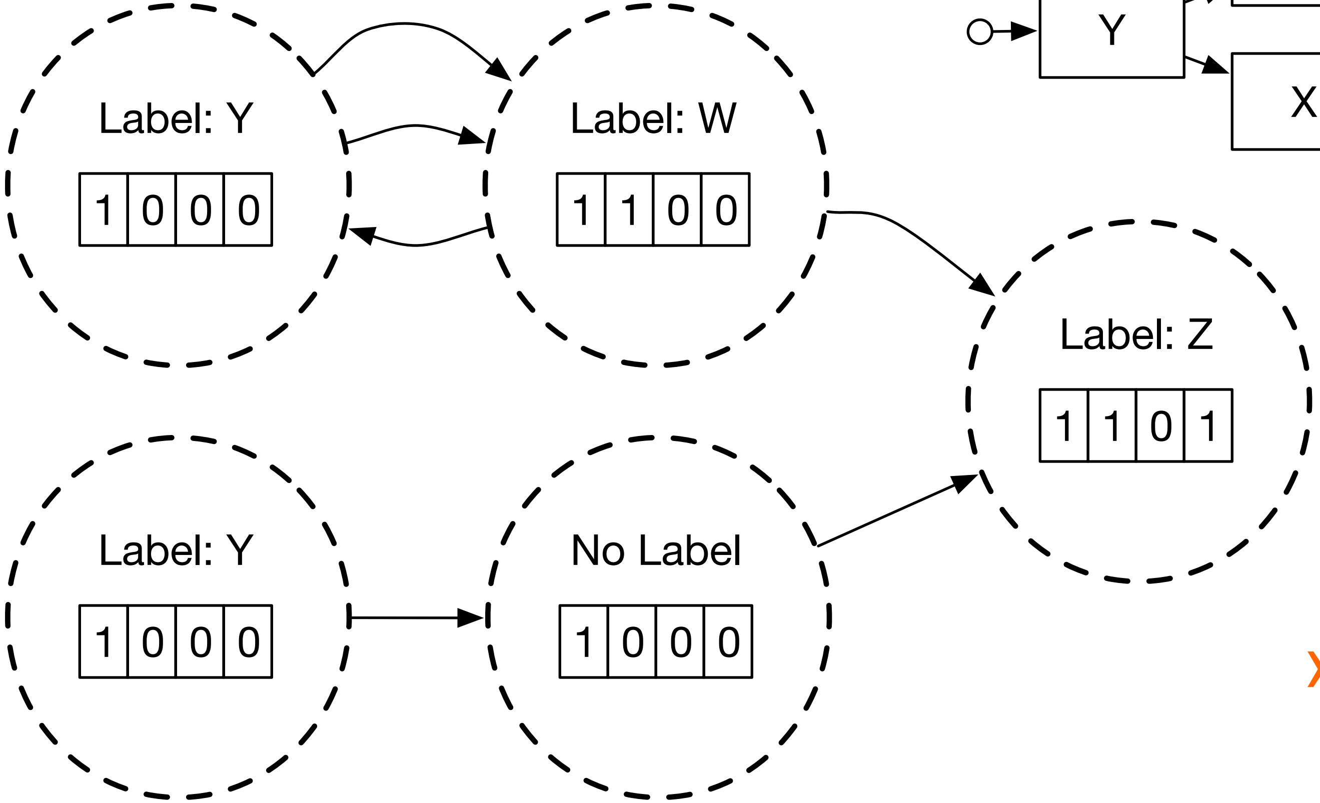
- The environment is updated with the segments' Ac labels

DX Illustrated

APT Grammar Y(W|X)Z encoded as

Y	W	X	Z



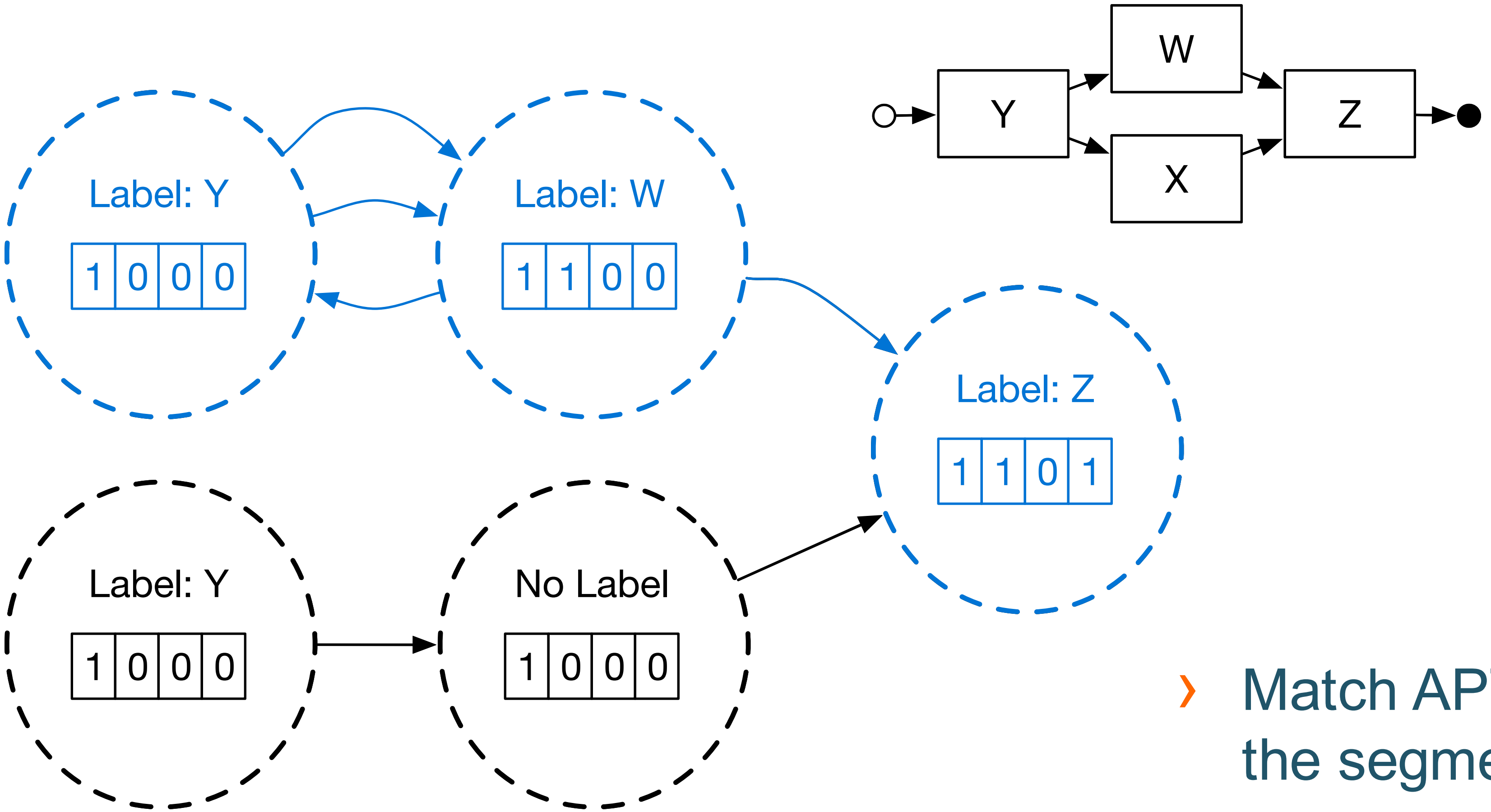


- *Repeated for all segments*

DX Illustrated

APT Grammar Y(W|X)Z encoded as

Y	W	X	Z

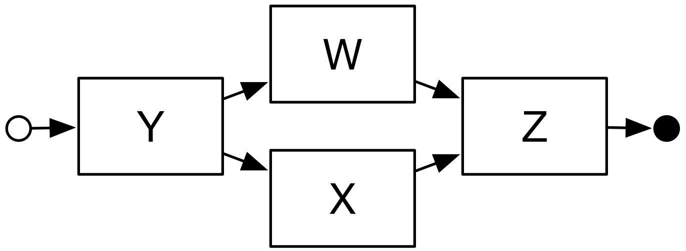


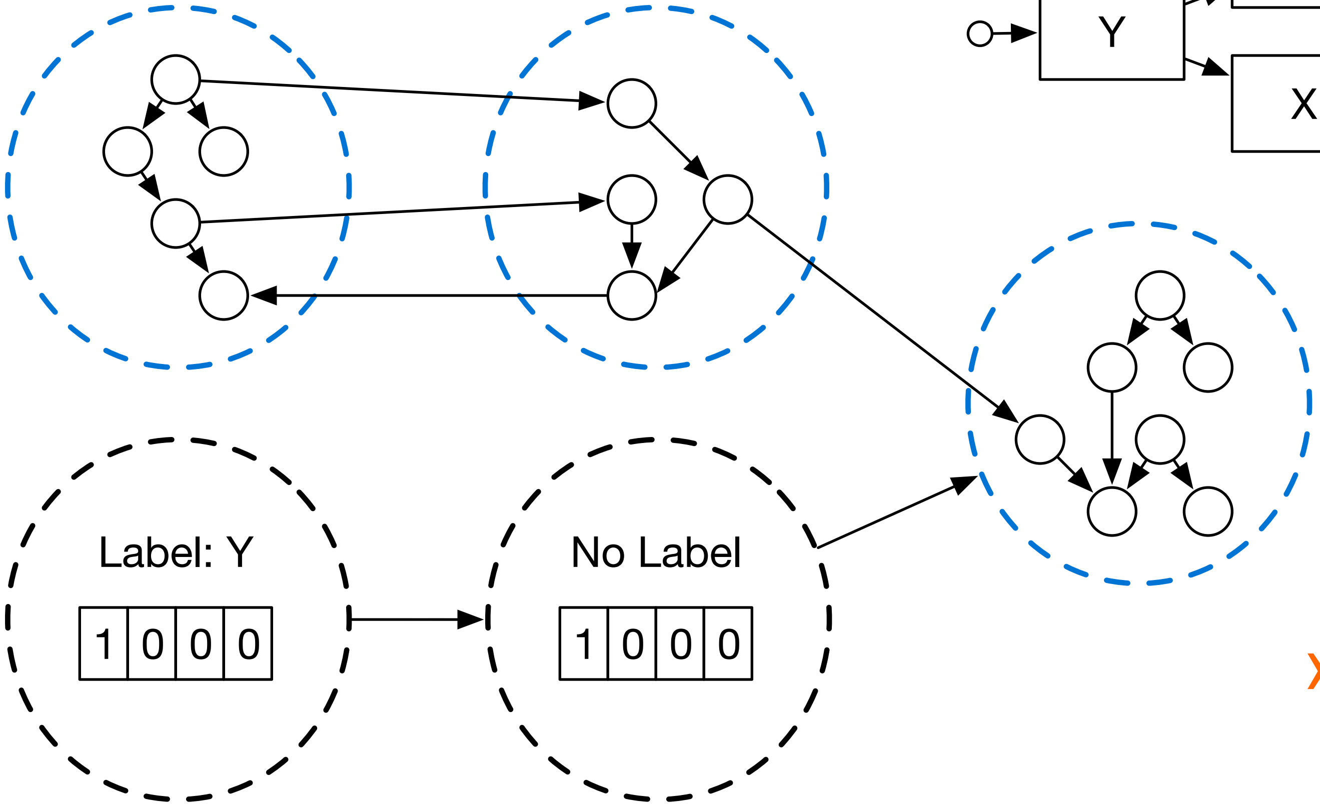
- Match APT Grammar at the segment level

DX Illustrated

APT Grammar Y(W|X)Z encoded as

Y	W	X	Z





Label: Y

1	0	0	0

No Label

1	0	0	0

- Zoom-in to check if there is a causal path at the provenance level

THANK YOU

parc®

A Xerox Company