Relation Web

- Let D be the set of descriptors.
- Let R be the set of possible relations.
- Let $G = (V, E)$ be a directed graph where $V = D$ and E is a set of edges such that each $v \in V$ has $|R| \times |D|$ edges, one per relation per descriptor (exhaustively listing all possible relations between this descriptor and all other descriptors).
- Goal: assign each edge $e \in E$ a weight $P(e) \in [0, 1]$ that represents the probability of that being a valid relation.
- Idea: as triples are processed, adjust the weight of not only the "explicit" relation between the relevant descriptors that is directly stated with that triple, but also the "implicit" relation(s).

Algorithm

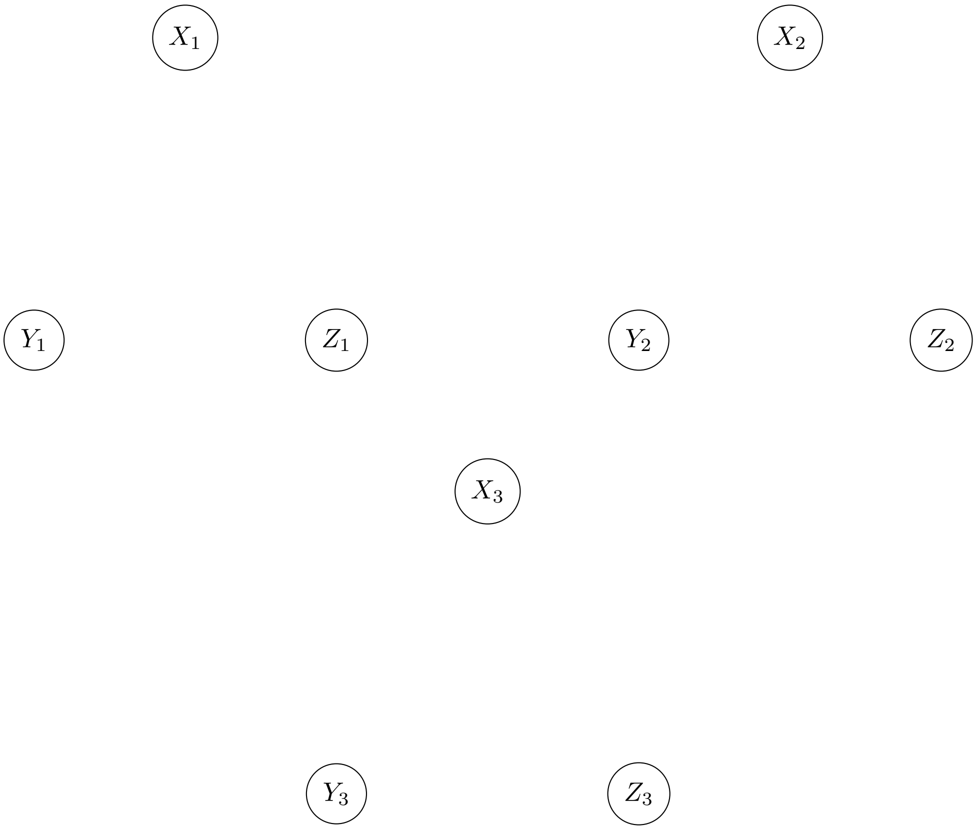
- Assume that we have a stream S of verified confidence measures for specific relations between two descriptors.
- We want to sequentially construct a relation web, processing one $s \in S$ at a time, updating the web for each new piece of information.
- Let L be the set of labels.
- Let D be the set of descriptors.
- Let L_p be the set of all unique ordered pairwise combinations of $l \in L$.
- Let $C(l_p)$, where $l_p = (l_1, l_2) \in L_p$, be the set of all possible paths between l_1 and l_2, where each path is generalized to the labels on the vertices.
- Let $C_p(l_p)$, be the set of all unique unordered pairwise combinations of possible paths between l_1 and l_2, where each path is generalized to the labels on the vertices.
- Let $L(d)$, where $d \in D$, be the label of d.
- Let $L(c^*)$, where c^* is a path in the graph, be a path with the descriptors replaced with their labels.
- Let $Corr_{l_p}(c_p)$, where $c_p = (c_1, c_2) \in C_p(l_p)$, be an assignment of a correlation score between paths c_1 and c_2.
- When processing a new measure $s \in S$:
  Extrapolation Stage:
  - Assign edge e specified by s.
  - For each path $d_1 \xrightarrow{r_{1,2}} \ldots \xrightarrow{r_{n-1,n}} d_n = c_1^*$ that e is a part of:
    - For each path c_2^* starting at d_1 and ending at d_n:
      - (Recursively?) update weights along c_2^* according to $Corr_{(L(d_1),L(d_2))}((L(c_1^*), L(c_2^*)))$.

  Correlation Reassessment Stage:
  - Reassess correlations ($Corr_{l_p}(c_p)\ \forall l_p, c_p$) based on current state of graph.
  - For each pair of labels (l_1, l_2):
    - Initialize list of vectors V.
    - For each pair of descriptors (d_1, d_2) with labels (l_1, l_2):
      - Construct a vector v where each entry corresponds to a label-generalized path between the descriptors, and store the sum of the edge weights along that path.
      - Add v to V.
    - Using V, determine pairwise correlations between all label-generalized paths c_1, c_2, ignoring paths with uninitialized values, storing results in $Corr_{(l_1,l_2)}(c_1, c_2)$.

Questions:

- How to convert OpenIE classifications to descriptor relation confidence measures?
- Math for σ-assignment.
  - Divide by path length.
- Potentially recursive σ-assignment.
- How to get correlation scores?
- Directed correlations? E.g. if this score is high, how often is the other one high, and if this score is low, how often is the other one low? And in both directions.

Event: Initialization



Event: Recieved from S: Confidence of 1.0 for relation 1 between X_1 and Y_1.

Extrapolation Stage:

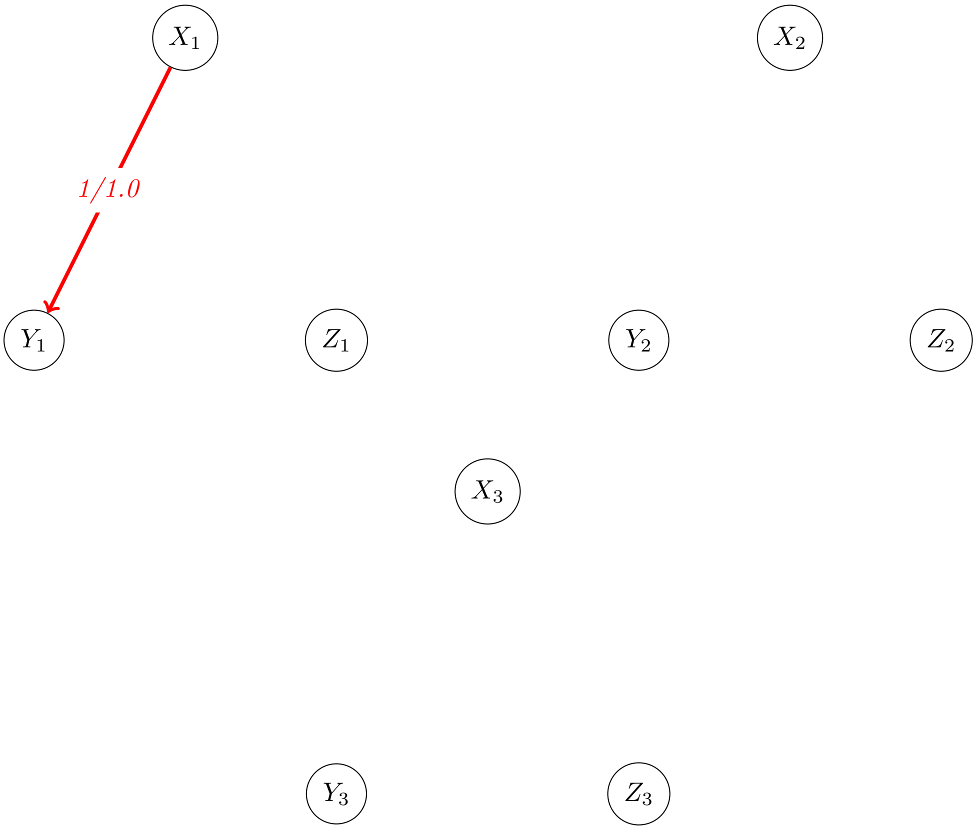
- Because there are not any correlations yet, no other edges are updated.

Correlation Reassessment Stage:

- Correlation vectors:

	$X \to Y$			$Y \to Z$			$X \to Z$		
	$X_1 \to Y_1$	$X_2 \to Y_2$	$X_3 \to Y_3$	$Y_1 \to Z_1$	$Y_2 \to Z_2$	$Y_3 \to Z_3$	$X_1 \to Z_1$	$X_2 \to Z_2$	$X_3 \to Z_3$
$X \xrightarrow{1} Y$	1								

- No correlations can be drawn.



Event: Recieved from S: Confidence of 1.0 for relation 2 between Y_1 and Z_1.

Extrapolation Stage:

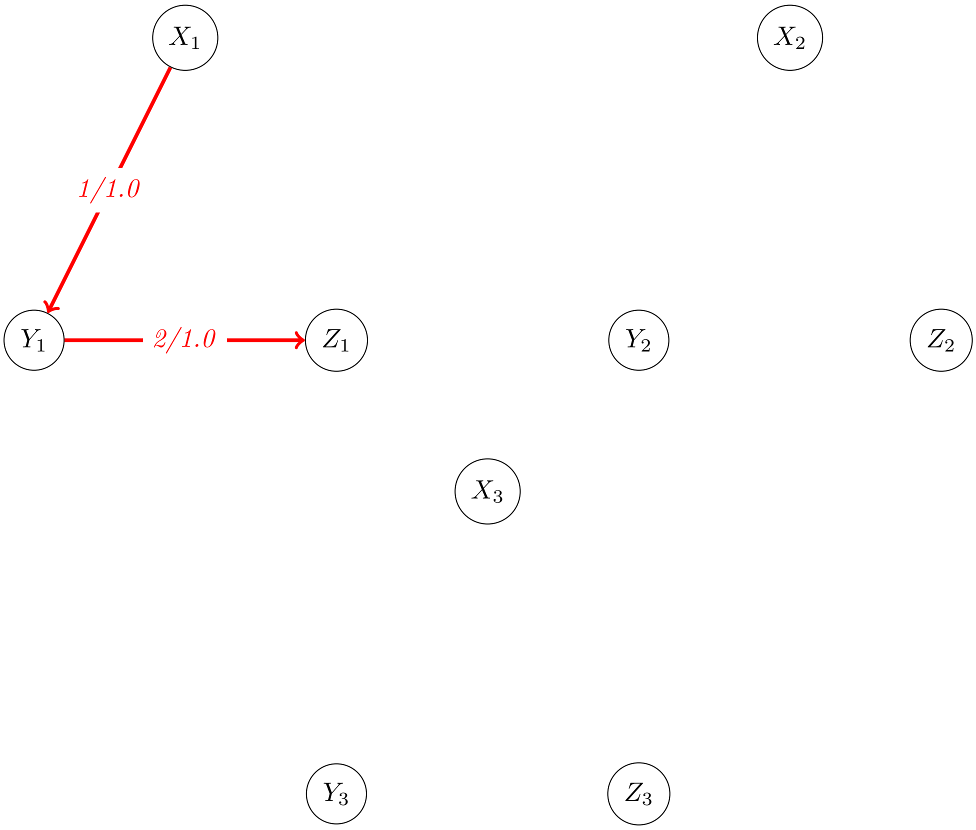
- Because there are not any correlations yet, no other edges are updated.

Correlation Reassessment Stage:

- Correlation vectors:

$X \to Y$				$Y \to Z$				$X \to Z$			
	$X_1 \to Y_1$	$X_2 \to Y_2$	$X_3 \to Y_3$		$Y_1 \to Z_1$	$Y_2 \to Z_2$	$Y_3 \to Z_3$		$X_1 \to Z_1$	$X_2 \to Z_2$	$X_3 \to Z_3$
$X \xrightarrow{1} Y$	1			$Y \xrightarrow{2} Z$	1			$X \xrightarrow{1} Y \xrightarrow{2} Z$	2		

- No correlations can be drawn.



Event: Recieved from S: Confidence of 1.0 for relation 3 between X_1 and Z_1.

Extrapolation Stage:

- Because there are not any correlations yet, no other edges are updated.

Correlation Reassessment Stage:

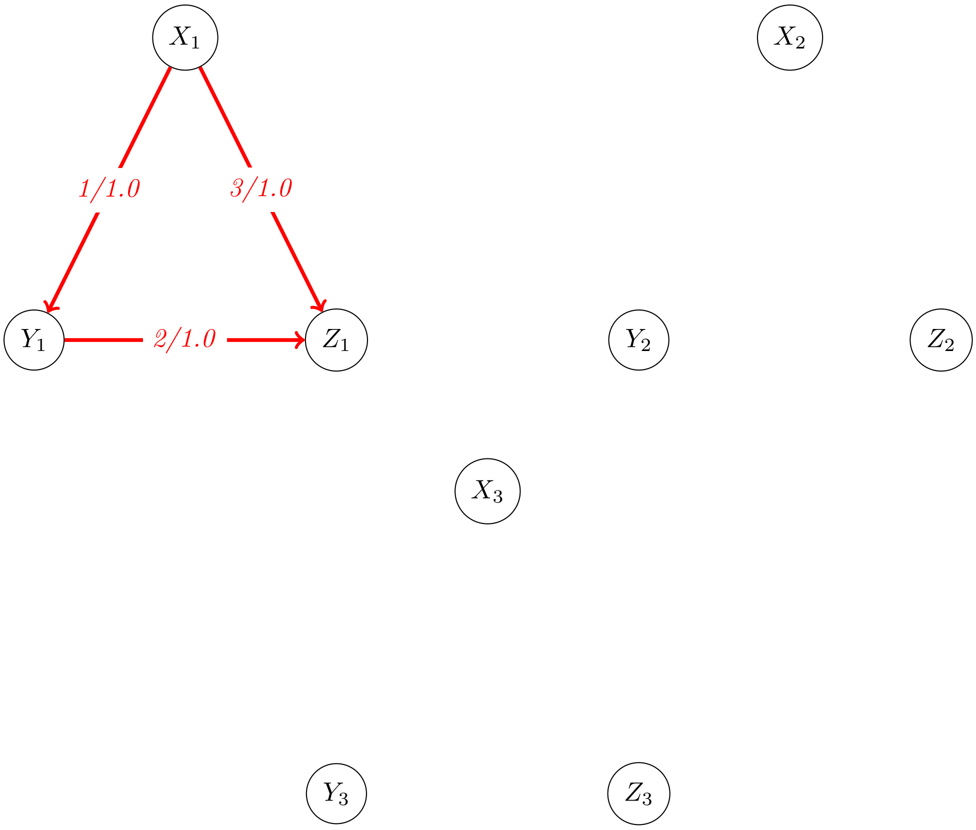
- Correlation vectors:

$X \rightarrow Y$	$X_1 \rightarrow Y_1$	$X_2 \rightarrow Y_2$	$X_3 \rightarrow Y_3$
$X \xrightarrow{1} Y$	1		

$Y \rightarrow Z$	$Y_1 \rightarrow Z_1$	$Y_2 \rightarrow Z_2$	$Y_3 \rightarrow Z_3$
$Y \xrightarrow{2} Z$	1		

$X \rightarrow Z$	$X_1 \rightarrow Z_1$	$X_2 \rightarrow Z_2$	$X_3 \rightarrow Z_3$
$X \xrightarrow{3} Z$	1		
$X \xrightarrow{1} Y \xrightarrow{2} Z$	2		

- No correlations can be drawn.



Event: Recieved from S: Confidence of 1.0 for relation 1 between X_2 and Y_2, confidence of 1.0 for relation 3 between X_2 and Y_2.

Extrapolation Stage:

- Because there are not any correlations yet, no other edges are updated.

Correlation Reassessment Stage:

- Correlation vectors:

$X \rightarrow Y$

	$X_1 \rightarrow Y_1$	$X_2 \rightarrow Y_2$	$X_3 \rightarrow Y_3$
$X \xrightarrow{1} Y$	1	1	

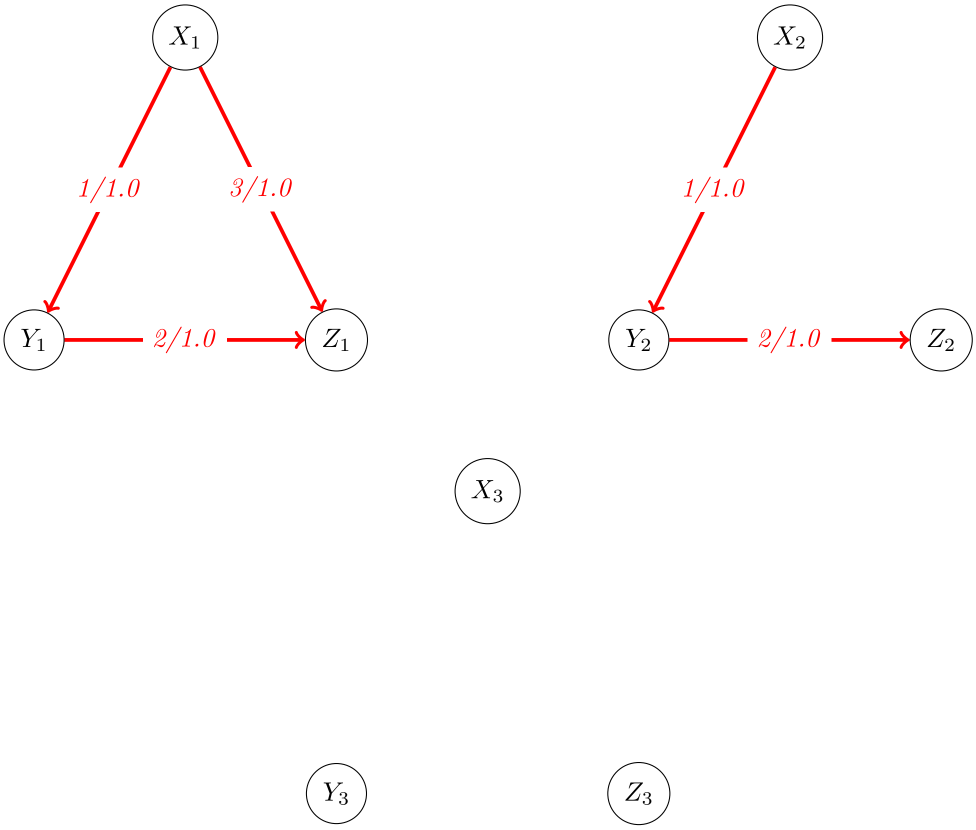
$Y \rightarrow Z$

	$Y_1 \rightarrow Z_1$	$Y_2 \rightarrow Z_2$	$Y_3 \rightarrow Z_3$
$Y \xrightarrow{2} Z$	1	1	

$X \rightarrow Z$

	$X_1 \rightarrow Z_1$	$X_2 \rightarrow Z_2$	$X_3 \rightarrow Z_3$
$X \xrightarrow{3} Z$	1		
$X \xrightarrow{1} Y \xrightarrow{2} Z$	2	2	

- No correlations can be drawn.



Event: Recieved from S: Confidence of 1.0 for relation 1 between X_2 and Z_2.

Extrapolation Stage:

- Because there are not any correlations yet, no other edges are updated.

Correlation Reassessment Stage:

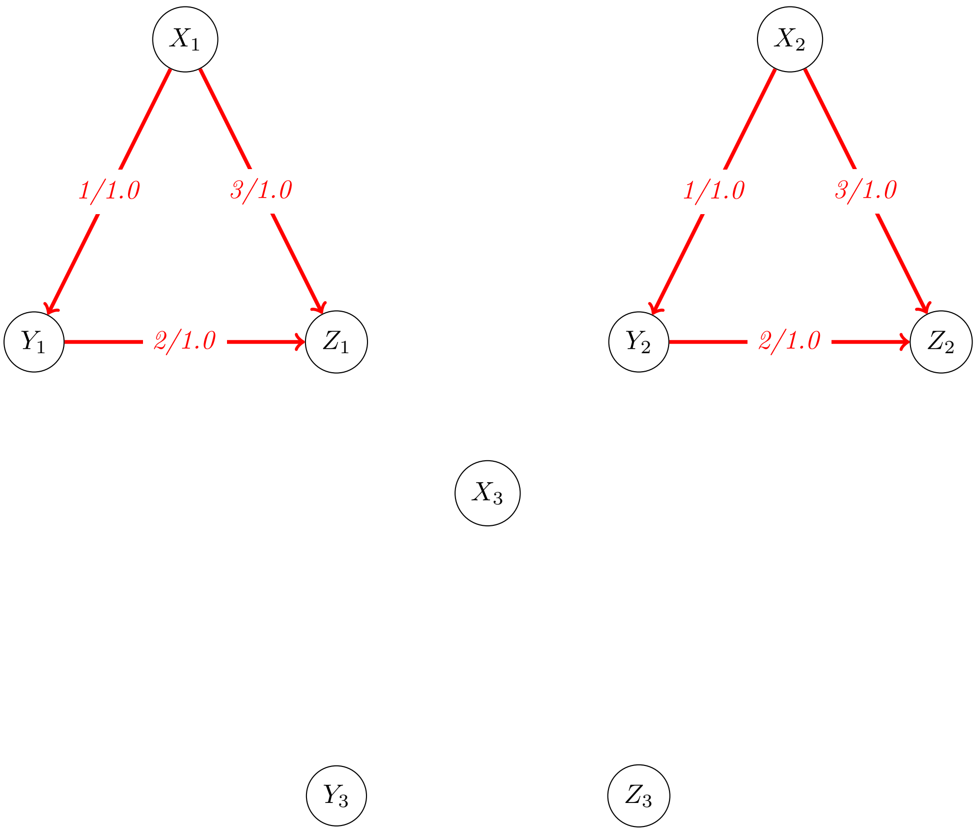
- Correlation vectors:

$X \to Y$	$X_1 \to Y_1$	$X_2 \to Y_2$	$X_3 \to Y_3$
$X \xrightarrow{1} Y$	1	1	

$Y \to Z$	$Y_1 \to Z_1$	$Y_2 \to Z_2$	$Y_3 \to Z_3$
$Y \xrightarrow{2} Z$	1	1	

$X \to Z$	$X_1 \to Z_1$	$X_2 \to Z_2$	$X_3 \to Z_3$
$X \xrightarrow{3} Z$	1	1	
$X \xrightarrow{1} Y \xrightarrow{2} Z$	2	2	

- A positive correlation now exists between $X \xrightarrow{3} Z$ and $X \xrightarrow{1} Y \xrightarrow{2} Z$. That is, $Corr_{(X,Z)}(X \xrightarrow{3} Z, X \xrightarrow{1} Y \xrightarrow{2} Z) > 0$.



Event: Recieved from S: Confidence of 1.0 for relation 1 between X_3 and Y_3.

Extrapolation Stage:

- Because the path $X_3 \xrightarrow{1} Y_3 \xrightarrow{2} Z$ is incomplete, no extrapolation occurrs.

Correlation Reassessment Stage:

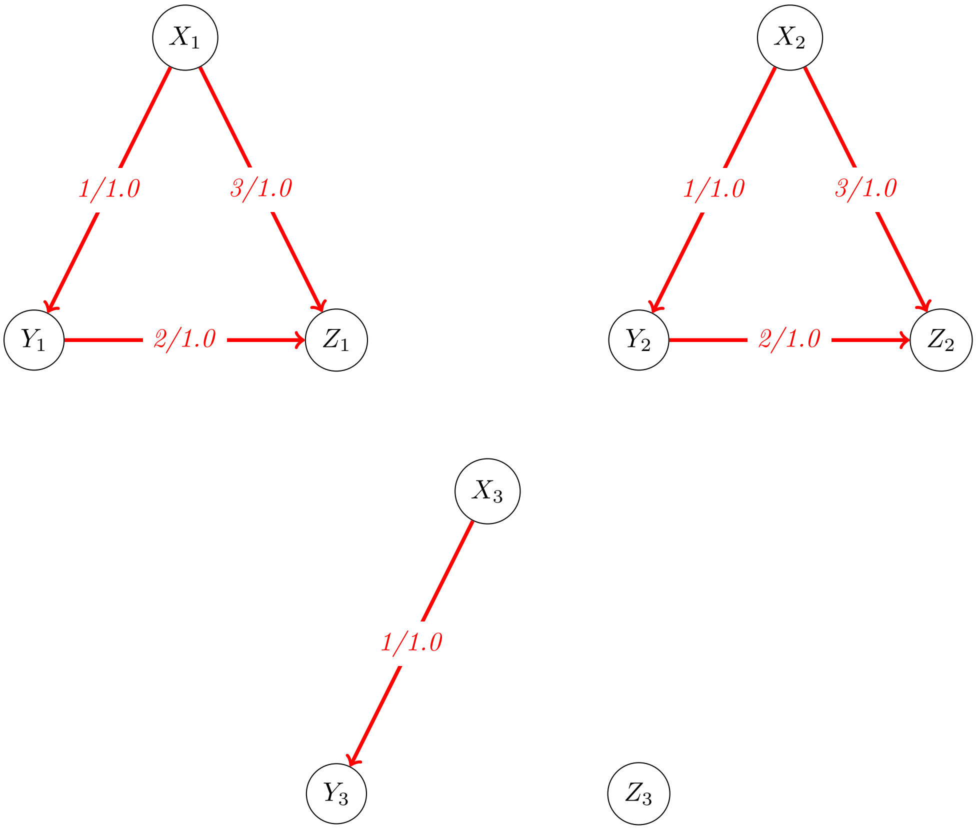
- Correlation vectors:

$X \to Y$	$X_1 \to Y_1$	$X_2 \to Y_2$	$X_3 \to Y_3$
$X \xrightarrow{1} Y$	1	1	1

$Y \to Z$	$Y_1 \to Z_1$	$Y_2 \to Z_2$	$Y_3 \to Z_3$
$Y \xrightarrow{2} Z$	1	1	

$X \to Z$	$X_1 \to Z_1$	$X_2 \to Z_2$	$X_3 \to Z_3$
$X \xrightarrow{3} Z$	1	1	
$X \xrightarrow{1} Y \xrightarrow{2} Z$	2	2	

- $Corr_{(X,Z)}(X \xrightarrow{3} Z, X \xrightarrow{1} Y \xrightarrow{2} Z) > 0$.
- No new correlations can be drawn.



Event: Recieved from S: Confidence of 1.0 for relation 2 between X_3 and Y_3.

Extrapolation Stage:

- Because the path $X_3 \xrightarrow{1} Y_3 \xrightarrow{2} Z_3$ is complete, extrapolation occurs between it and the path $X_3 \xrightarrow{3} Z_3$, using $Corr_{(X,Z)}(X \xrightarrow{3} Z, X \xrightarrow{1} Y \xrightarrow{2} Z)$. A weight σ is applied to the correlated path $X_3 \xrightarrow{3} Z_3$.

Correlation Reassessment Stage:

- Correlation vectors:

$X \to Y$	$X_1 \to Y_1$	$X_2 \to Y_2$	$X_3 \to Y_3$
$X \xrightarrow{1} Y$	1	1	1

$Y \to Z$	$Y_1 \to Z_1$	$Y_2 \to Z_2$	$Y_3 \to Z_3$
$Y \xrightarrow{2} Z$	1	1	1

$X \to Z$	$X_1 \to Z_1$	$X_2 \to Z_2$	$X_3 \to Z_3$
$X \xrightarrow{3} Z$	1	1	σ
$X \xrightarrow{1} Y \xrightarrow{2} Z$	2	2	2

- $Corr_{(X,Z)}(X \xrightarrow{3} Z, X \xrightarrow{1} Y \xrightarrow{2} Z) > 0$.
- $Corr_{(X,Z)}(X \xrightarrow{3} Z, X \xrightarrow{1} Y \xrightarrow{2} Z)$ is adjusted.