Relation Web

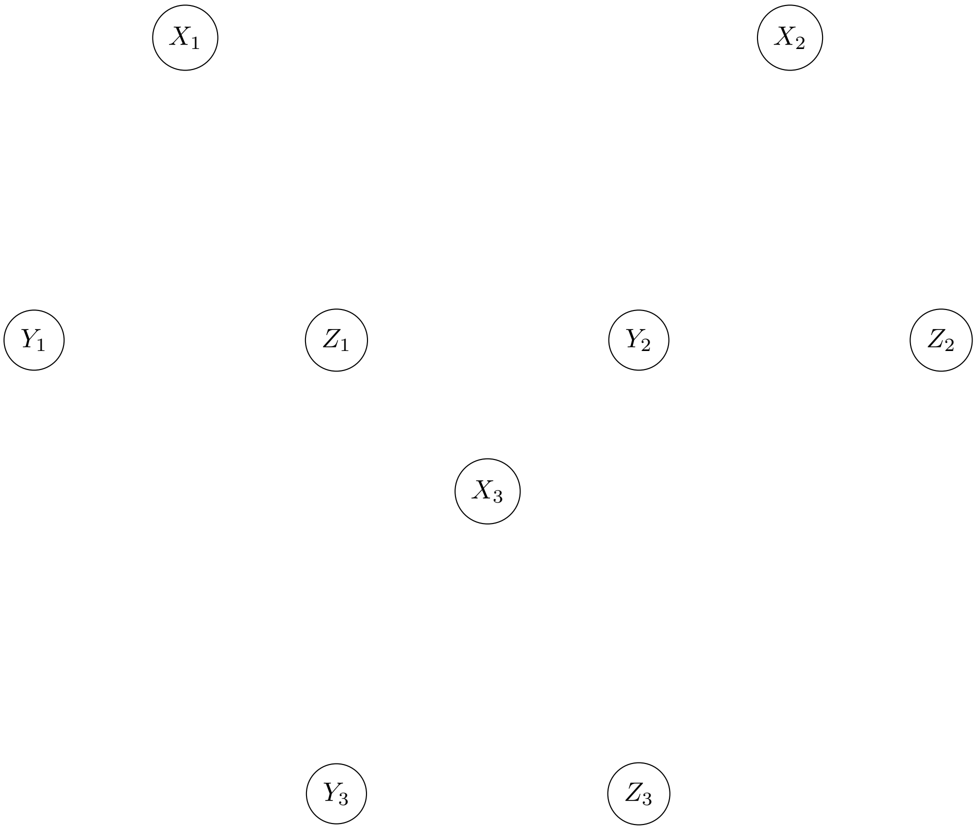
- Let D be the set of descriptors.
- Let R be the set of possible relations.
- Let $G = (V, E)$ be a directed graph where $V = D$ and E is a set of edges such that each $v \in V$ has $|R| \times |D|$ edges, one per relation per descriptor (exhaustively listing all possible relations between this descriptor 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:
  - Reasses 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, storing results in $Corr_{(l_1,l_2)}(c_1, c_2)$.

Event: Initialization



Correlation Table: (empty)

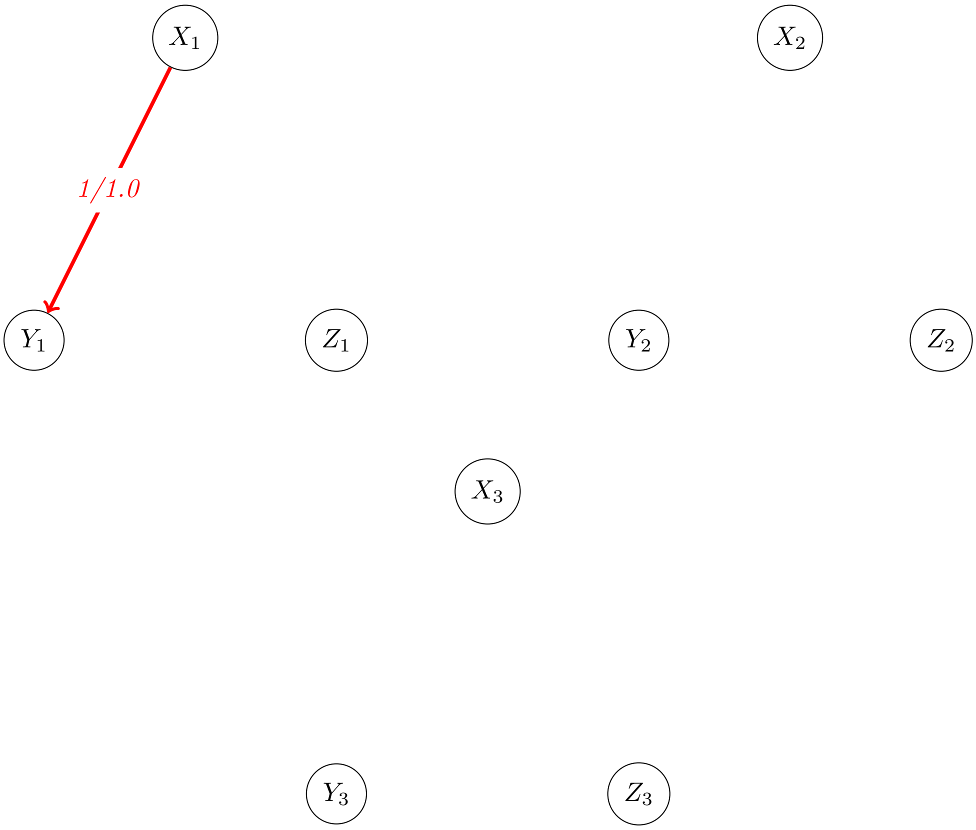
Event: Recieved from S: Confidence of 1.0 for relation 1 between X_1 and Y_1.

Extrapolation Stage:

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

Correlation Reassessment Stage:

- No correlations can be drawn.



Correlation Table: (empty)

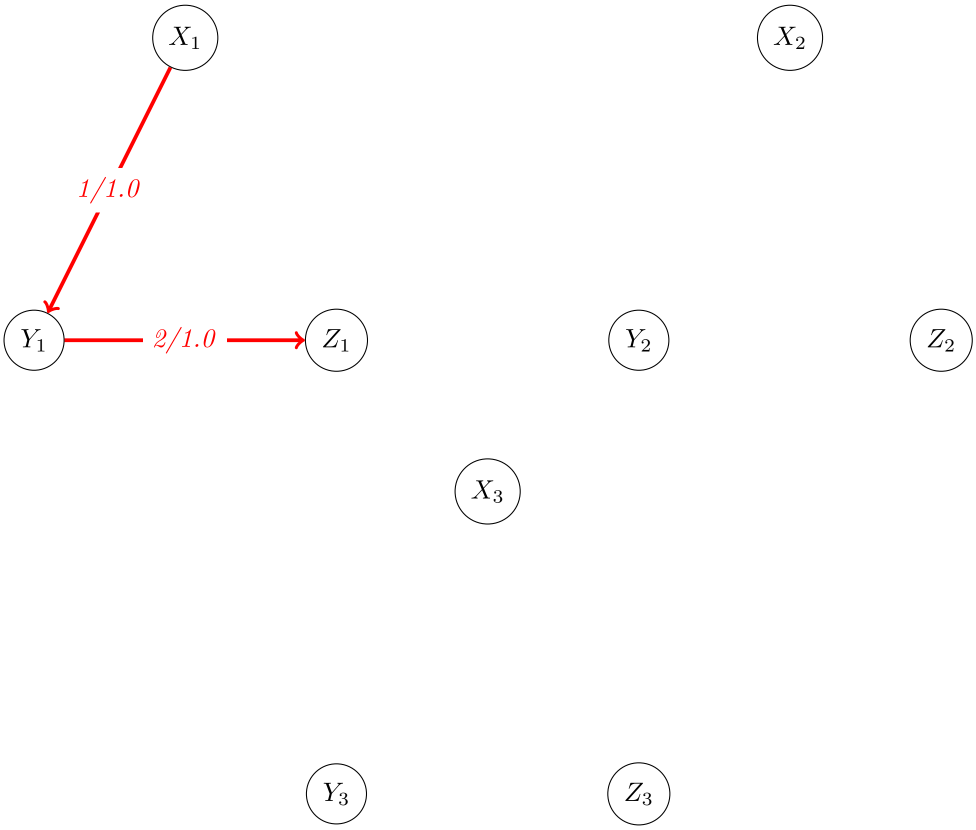
Event: Recieved from S: Confidence of 1.0 for relation 2 between Y_1 and Z_1.

Extrapolation Stage:

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

Correlation Reassessment Stage:

- No correlations can be drawn.



Correlation Table: (empty)

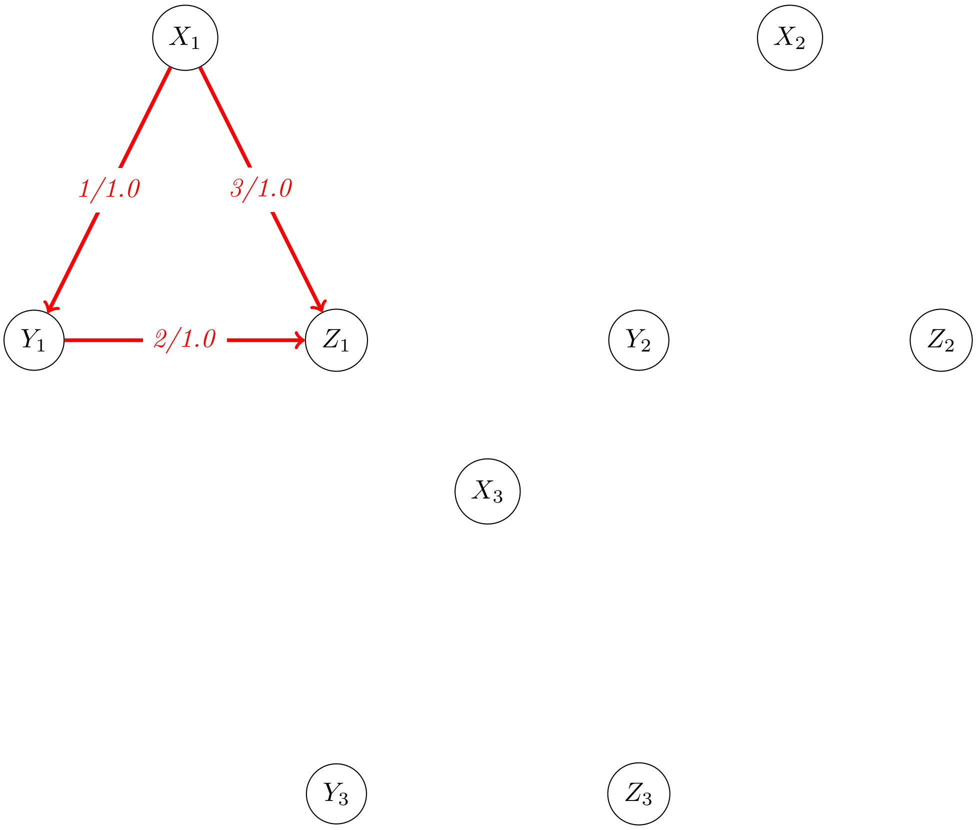
Event: Recieved from S: Confidence of 1.0 for relation 3 between X_1 and Z_1.

Extrapolation Stage:

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

Correlation Reassessment Stage:

- No correlations can be drawn.



Correlation Table: (empty)