Event similarity

We use an ontology-based semantic similarity measure (except the temporal information)

For given two values x and y of the same object type a is calculated as

$$sim_a(x, y) = \frac{2l_a(x, y)}{l_a(x) + l_a(y)} - \lambda\sigma_a, \tag{1}$$

where

- $l_a(v)$ is the depth (or level) of a node v in the taxonomy tree H_a
- $l_a(v_1, v_2)$ is the depth of the closest common ancestor of v_1 and v_2 in H_a
- λ takes a binary value, where a score of 0 means x and y are identical or hierarchically linked and 1 otherwise
- σ_a is a penalization factor, used when $\lambda = 1$

Event similarity

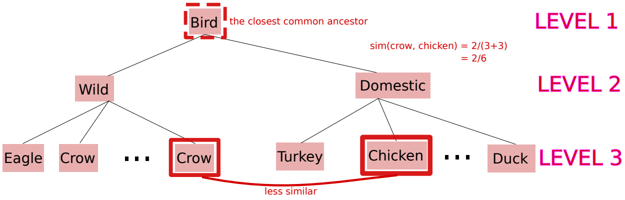
We use an ontology-based semantic similarity measure (except the temporal information)

For given two values x and y of the same object type a is calculated as

$$sim_a(x,y) = \frac{2l_a(x,y)}{l_a(x) + l_a(y)} - \lambda\sigma_a, \tag{1}$$

where

- $l_a(v)$ is the depth (or level) of a node v in the taxonomy tree H_a
- $l_a(v_1, v_2)$ is the depth of the closest common ancestor of v_1 and v_2 in H_a
- λ takes a binary value, where a score of 0 means x and y are identical or hierarchically linked and 1 otherwise
- σ_a is a penalization factor, used when $\lambda = 1$



Event similarity

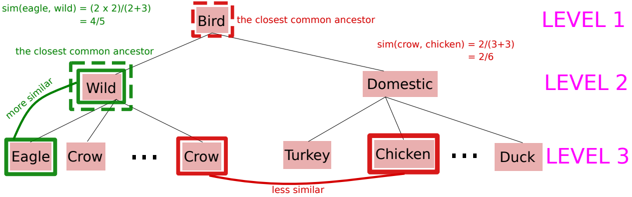
We use an ontology-based semantic similarity measure (except the temporal information)

For given two values x and y of the same object type a is calculated as

$$sim_a(x, y) = \frac{2l_a(x, y)}{l_a(x) + l_a(y)} - \lambda\sigma_a, \quad (1)$$

where

- $l_a(v)$ is the depth (or level) of a node v in the taxonomy tree H_a
- $l_a(v_1, v_2)$ is the depth of the closest common ancestor of v_1 and v_2 in H_a
- λ takes a binary value, where a score of 0 means x and y are identical or hierarchically linked and 1 otherwise
- σ_a is a penalization factor, used when $\lambda = 1$



Event similarity

We use an ontology-based semantic similarity measure (except the temporal information)

For given two values x and y of the same object type a is calculated as

$$sim_a(x,y) = \frac{2l_a(x,y)}{l_a(x) + l_a(y)} - \lambda\sigma_a, \quad (1)$$

where

- $l_a(v)$ is the depth (or level) of a node v in the taxonomy tree H_a
- $l_a(v_1, v_2)$ is the depth of the closest common ancestor of v_1 and v_2 in H_a
- λ takes a binary value, where a score of 0 means x and y are identical or hierarchically linked and 1 otherwise
- σ_a is a penalization factor, used when $\lambda = 1$

$$sim_t(t_1, t_2) = 1 - \frac{|t_2 - t_1|}{L = 21} \quad (2)$$

$$sim(e_1, e_2) = \sum_{a \in \{\ell, t, d, h\}} sim_a(a_i, a_j), \quad (3)$$

Event matching

- **Approach:** Modeling as an assignment problem
- Let D_1 (resp. D_2) be an event dataset associated with an EBS platform, containing n_1 (resp. n_2) events. Also, we assume $n_1 \leq n_2$ without loss of generality.
- Moreover, let S be the $n_1 \times n_2$ similarity matrix of D_1 and D_2. The term $S_{i,j}$, with $1 \leq i \leq n_1$ and $1 \leq j \leq n_2$, represents the similarity score between events e_i and e_j
- We look for a bijection $f : \{1, 2, ..., n_1\} \rightarrow \{1, 2, ..., n_2\}$ such that the objective is to maximize the similarity between D_1 and D_2, i.e.

$$Max \sum_{i=1}^{n_1} S_{i,f(i)}. \tag{4}$$

- **Post-processing:** Keep only the assignments, whose score is greater than some threshold value