Order Metrics for Semantic Knowledge Systems

Cliff Joslyn[1] and Emilie Hogan[2]

[1] National Security Directorate, Pacific Northwest National Laboratory, Seattle, Washington, 98109, USA
cjoslyn@pnl.gov
[2] Mathematics Department, Rutgers University

Abstract. Knowledge systems technologies, as derived from AI methods and used in the modern Semantic Web movement, are dominated by graphical knowledge structures such as ontologies and semantic graph databases. A critical but typically overlooked aspect of all of these structures is their admission to analyses in terms of formal hierarchical relations. The partial order representations of whatever hierarchy is present within a knowledge structure afford opportunities to exploit these hierarchical constraints to facilitate a variety of tasks, including ontology analysis and alignment, visual layout, and anomaly detection. We introduce the basic concepts of order metrics and address the impact of a hierarchical (order-theoretical) analysis on knowledge systems tasks.

1 Introduction

Knowledge systems technologies are dominated by graphical structures. *Semantic graph databases* [15] take the form of labeled directed graphs implemented in RDF [9]. Their OWL [10] ontological typing systems are also labeled directed graphs, frequently dominated by directed acyclic graph (DAG) and other hierarchical structures. Fig. 1 shows a toy example, where the ontology of classes on the left forms the typing system for the semantic graph of node and link instances on the right.

But where semantic taxonomies such as the Gene Ontology [2] include hierarchical class structures, other portions can be non-hierarchical. And more general knowledge structures like semantic graphs are not explicitly or necessarily hierarchical, but may contain large hierarchical components.

In practice, ontologies are dominated by their "hierarchical cores", specifically their class hierarchies connected by `is-a` subsumptive and `has-part` compositional links. And many of the most common links in RDF graphs are transitive, including `causes`, `implies`, and `precedes`. The partial order representation of whatever hierarchy is present within a knowledge structure affords opportunities to exploit these hierarchical constraints for a variety of tasks, including:

Clustering and Classification: Characterizing a portion of a hierarchy (e.g. groups of ontology nodes) to identify common characteristics [12,18],

Alignment: Casting ontology matching [7] as mappings between hierarchical structures [11,14].

E.S. Corchado Rodriguez et al. (Eds.): HAIS 2010, Part II, LNAI 6077, pp. 399–409, 2010.

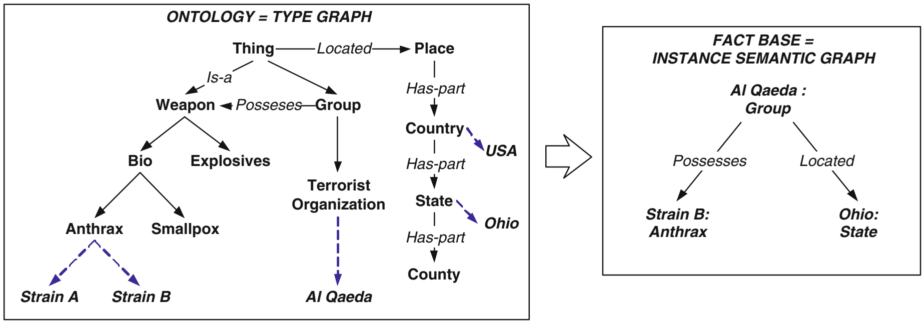

Fig. 1. Toy model of a semantic graph database. (Left) Ontological typing system as a labeled, directed graph of classes (sample instances shown below dashed links). (Right) Conforming instance sub-graph.

Visualization: Including exploiting the level structure of hierarchies to achieve a satisfactory layout [13].

In general, such a hierarchical analysis, when available, promises complexity reduction, improved user interaction with the knowledge base, and improved layout and visual analytics.

2 DAGs and Partial Orders

Hierarchies are represented as partially ordered sets (posets), which are reflexive, anti-symmetric, and transitive binary relations $\mathcal{P} = \langle P, \leq \rangle$ on an underlying finite set of nodes P [6]. While we typically think of hierarchies as tree structures, more general kinds of hierarchies have "multiple inheritance", where nodes can have more than one parent. These include lattice structures, where pairs of nodes have unique least common subsumers (and unique greatest lower bounds as well); partial orders where pairs of nodes can have an indefinite number of least common subsumers and greatest lower bounds; and finally general DAGs can also include "transitive links" which form shortcuts across paths.

Consider simple DAG in the top of Fig. 2. The two transitive links $1 \rightarrow H, 1 \rightarrow E$ connect the two paths $1 \rightarrow K \rightarrow H$ and $1 \rightarrow C \rightarrow I \rightarrow E$ respectively. Given a DAG $\mathcal{D}$, the DAG $\mathcal{P}(\mathcal{D})$ produced by including all possible transitive links consistent with its paths is its *transitive closure*, and determines an ordered set $\mathcal{P}(\mathcal{D}) = \langle P, \leq \rangle$ where $a \leq b \subseteq \mathcal{P}$ if there is a directed path from a to b in $\mathcal{D}$. The graph $\mathcal{V}(\mathcal{D})$ produced from a DAG $\mathcal{D}$ by removing all its transitive links (its transitive reduction [1]) determines a *cover relation* or *Hasse diagram*. Thus each cover relation $\mathcal{V}$ determines a unique poset $\mathcal{P}(\mathcal{V})$, and *vice versa* a poset $\mathcal{P}$ determines a unique cover $\mathcal{V}(\mathcal{P})$; each DAG $\mathcal{D}$ determines a unique poset $\mathcal{P}(\mathcal{D})$ and cover $\mathcal{V}(\mathcal{D})$; and each unique poset-cover pair determines a class of DAGs equivalent by transitive links.

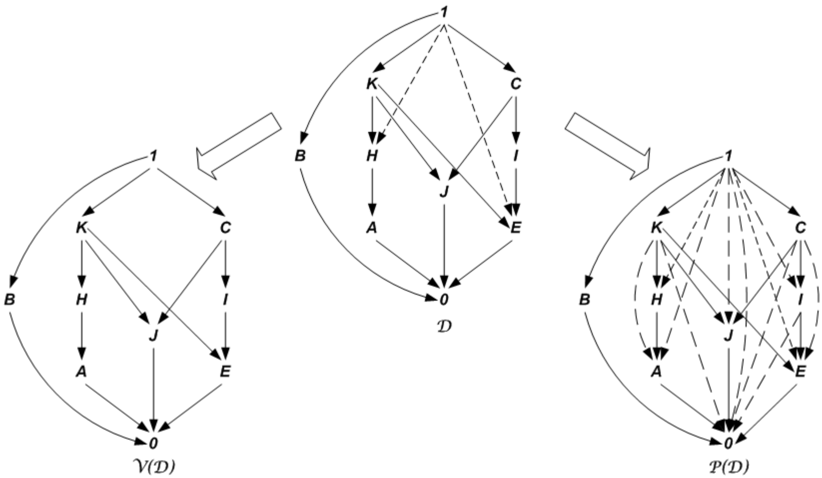

Fig. 2. (Top) A DAG $\mathcal{D}$. (Left) Transitive reduction $\mathcal{V}(\mathcal{D})$. (Right) Transitive closure $\mathcal{P}(\mathcal{D})$.

For a DAG $\mathcal{D}$ we can measure its *degree of transitivity* as

$$TR(\mathcal{D}) := \frac{|\mathcal{D} \setminus \mathcal{V}(\mathcal{D})|}{|\mathcal{P}(\mathcal{D}) \setminus \mathcal{V}(\mathcal{D})|}, \tag{1}$$

where $\setminus$ is set subtraction, we interpret each structure as the binary relation on P^2 of its incidence matrix, and $|\cdot|$ is cardinality, so that $|\cdot|$ is the number of links in $\cdot$, seen as a graph. $TR(\mathcal{D})$ measures the number $|\mathcal{D} \setminus \mathcal{V}(\mathcal{D})|$ of transitive links in $\mathcal{D}$ relative to the total possible number $|\mathcal{P}(\mathcal{D}) \setminus \mathcal{V}(\mathcal{D})|$ in its transitive closure $\mathcal{P}(\mathcal{D})$. In Fig. 2 we have $TR(\mathcal{D}) = \frac{2}{11}$, indicating a relatively low degree of transitivity.

In knowledge systems such as ontologies, our interpretation of the presence or absence of transitive links in DAGs is significant. If the link-type in question is anti-transitive, so that transitive links are disallowed, then clearly the presence of transitive links is in error. If, on the other hand, the link-type in question is atransitive, so that transitive links are allowed, but not required, then the $TR(\mathcal{D})$ measures this extent. But finally, if, as is the case with our subsumption and composition types, the link type represents a fully transitive property, then the presence of transitive links are irrelevant or erroneous. Effectively, such link types live in the trasitively equivalent class of DAGs, that is, in the partial order $\mathcal{P}(\mathcal{D})$, and $TR(\mathcal{D})$ can be used as an aid to the user or engineer to identify issues with the underlying ontology.

3 Measures on Hierarchical Graphs

Given a hierarchical DAG structure represented by its transitive closure poset $\mathcal{P}$, tools are available to measure this hierarchical structure. Here we discusses *interval-valued rank* measuring the vertical level of nodes, and *order metrics* measuring the distances between nodes. See [13] for more details.

Consider the hierarchy shown in Fig. 3. We are concerned with the proper representation of the vertical level of each node, as represented by its positioning in a layout. We note that all children of the root have the same "distance" from the root, but if these are *also* leaves then they should be positioned further down. In other words, we need to exploit the vertical distance from *both* the top *and* a global bottom, including a virtual node $0 \in P$ inserted below all the leaves.

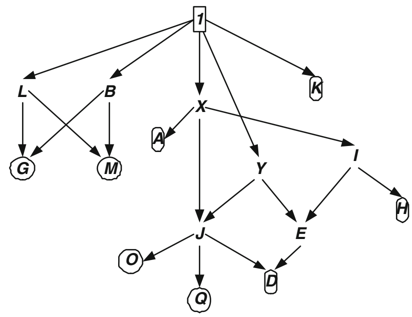

Fig. 3. A DAG displayed as a hierarchy

For $a, b \in P$, let $h^*(a, b)$ be the length of the maximum path from a to b. Then the distance of a node $a \in P$ from the root $1 \in P$ is the *top rank* $r^t(a) := h^*(a, 1)$. Dually we define the *bottom rank* $r^b(a) := h^*(0, 1) - h^*(0, b)$, where $h^*(0, 1)$ is the overall *height* of the structure. Then the *interval rank* $\bar{R}(a) := [r^t(a), r^b(a)]$ becomes available as an interval-valued measure of the vertical levels over which a can range, while the *rank width* $W(a) := r^b(a) - r^t(a)$ is a measure of that range [13].

We can exploit this vertical rank for hierarchical layout and visualization, as shown for our example in Fig. 4. Each node which sits on a complete chain (a path from 1 down to 0) of maximal size is placed horizontally at the center of the page. Nodes are laid out horizontally according to the size of their largest maximal chains. The result is to place maximal complete chains along a central axis, and short complete chains towards the outer edges. Nodes are placed vertically according to the mathematical quantity of the midpoint of their interval rank, but can be free to move between top rank $r^t(a)$ and bottom rank $r^b(a)$.

The result is that while nodes on maximal complete chains (all those intersecting the chain $0 \to D \to E \to I \to X \to 1$ in the example) exist at a single level, some (for example K) do not. While Fig. 4 shows a 2D layout, we have also deployed this concept in a 3D layout [13].

4 Order Metrics

Given the need to perform operations like clustering or alignment on ontologies represented as ordered sets $\mathcal{P} = \langle P, \leq \rangle$, it is essential to have a general sense of

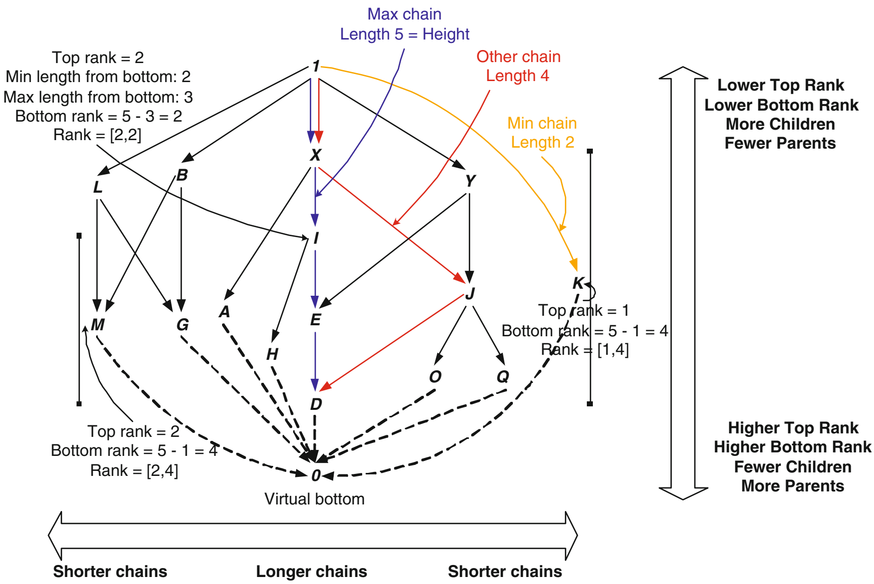

Fig. 4. Chain layout of the cyclic decomposition of the network in Fig. 3

distance $d(a, b)$ between two nodes $a, b \in P$. The knowledge systems literature has focused on *semantic similarities* to perform a similar function, which are available when $\mathcal{P}$ is equipped with a probability distribution, derived, for example, from the frequency with which terms appear in documents (for the Wordnet [8] thesaurus), or genes are annotated to GO nodes.

So assume a poset $\langle P, \leq \rangle$ with a base probability distribution $p\colon P \to [0, 1]$, $\sum_{a \in P} p(a) = 1$, and a "cumulative" function $\beta(a) := \sum_{b \leq a} p(a)$. We then generalize the join (least upper bound) and meet (greatest lower bound) operations in lattices as follows. Let $\uparrow a := \{b \geq a\}$ and $\downarrow a := \{b \leq a\}$ are the up-set (filter) and down-set (ideal) respectively of a node $a \in P$. Then for two nodes $a, b \in P$, let $a \nabla b := \uparrow a \cap \uparrow b$ and $a \Delta b := \downarrow a \cap \downarrow b$ be the set of nodes above or below respectively both of them. Then the generalized join $a \vee b$ is the set of minimal (lowest) nodes of $a \nabla b$, and the generalized meet $a \wedge b$ is the set of maximal (highest) nodes of $a \Delta b$. When $\mathcal{P}$ is a lattice, then $|a \vee b| = |a \wedge b| = 1$, recovering traditional join and meet.

Common choices for the semantic similarity $S(a, b)$ between two nodes include the measures of Resnik, Lin, and Jiang and Conrath [5]:

$$S(a, b) = \max_{c \in a \vee b} [-\log_2(\beta(c))] \tag{2}$$

$$S(a, b) = \frac{2 \max_{c \in a \vee b} [\log_2(\beta(c))]}{\log_2(\beta(a)) + \log_2(\beta(b))} \tag{3}$$

$$S(a,b) = 2 \max_{c \in a \vee b} [\log_2(\beta(c))] - \log_2(\beta(a)) - \log_2(\beta(b)) \tag{4}$$

respectively. But most of these are not metrics (not satisfying the triangle inequality), and all of these lack a general mathematical grounding and require a probabilistic weighting.

We use ordered set metrics [16,17], which are preferable to semantic similarities, because while they can use, they do not require, a quantitative weighting such as β; and because they always yield a metric. They are based on valuation functions $v \colon P \to \mathbb{R}^+$ which are, first, either isotone ($a \le b \to v(a) \le v(b)$) or antitone ($a \le b \to v(a) \ge v(b)$); and then semimodular, in that

$$v(a) + v(b) \bowtie v^{\nabla}(a,b) + v_{\Delta}(a,b), \tag{5}$$

where $\bowtie \in \{\le, \ge, =\}$, yielding super-modular, sub-modular, and modular valuations respectively; and

$$v^{\nabla}(a,b) := \min_{c \in a \nabla b} v(c), \quad v_{\Delta}(a,b) := \max_{c \in a \Delta b} v(c). \tag{6}$$

Whether a valuation v is antitone or isotone, and then sub- or super-modular, determines which of four distance functions is generated, e.g. the antitone, supermodular case yields $d(a,b) = v(a) + v(b) - 2v^{\nabla}(a,b)$. When $\mathcal{P}$ is a lattice, then this simplifies to $d(a,b) = v(a) + v(b) - 2v(a \vee b)$. See [17] for full details and proofs.

Typical valuations v include the cardinality of up-sets and down-sets: $v(a) = |\uparrow a|, v(a) = |\downarrow a|$, and the cumulative probabilities used in semantic similarities $v(a) = \beta(a)$. In this way, poset metrics generalize semantic similarities and provide a strong basis for various analytical tasks.

5 Order Metrics in Ontology Alignment

A good example of the utility of this order theoretical technology in knowledge systems tasks is in ontology alignment [7,11]. An ontology *alignment* is a mapping $f \colon \mathcal{P} \to \mathcal{P}'$ taking anchors $a \in P$ in one semantic hierarchy $\mathcal{P} = \langle P, \le \rangle$ into anchors $a' \in P'$ in another $\mathcal{P}' = \langle P', \le' \rangle$. In seeking a measure of the structural properties of the mapping f, our primary criterion is that f should not distort the metric relations of concepts, taking nodes that are close together and making them farther apart, or *vice versa*.

It should be noted that a "smooth" mapping f is neither necessary nor sufficient to be a good alignment: on the one hand, a good structural mapping may be available between structures from different domains; and on the other, differences in semantic intent between the two structures may be irreconcilable. Nonetheless, other things being equal, it is preferable to have a more smooth mapping than not.

So, for two ontology nodes $a, b \in \mathcal{P}$, consider the *lower cardinality distance* $d_l(a,b) := |\downarrow a| + |\downarrow b| - 2 \max_{c \in a \wedge b} |\downarrow c|$. We can measure the change in distance between $a, b \in P$ induced by f as the *distance discrepancy*

$$\delta(a,b) := |\bar{d}_l(a,b) - \bar{d}_l(f(a), f(b))|, \tag{7}$$

where $\bar{d}_l(a,b) := \frac{d_l(a,b)}{\mathrm{diam}_d(\mathcal{P})} \in [0,1]$ is the normalized lower distance between a and b in $\mathcal{P}$ given the diameter $\mathrm{diam}_d(\mathcal{P}) := \max_{a,b \in P} d(a,b)$. We can measure the entire amount of distance discrepancy at a node $a \in P$ compared to all the other anchors $b \in P$ by summing

$$\delta_f(a) := \sum_{b \in P} \delta(a,b) = \sum_{b \in P} |\bar{d}_l(a,b) - \bar{d}_l(f(a), f(b))|, \tag{8}$$

yielding the discrepancy $\delta(f) := \sum_{a \in P} \delta_f(a)$ of the alignment.

Consider the example in Fig. 5, with the partial alignment function f as shown, mapping only certain nodes $\{B, E, G\}$ from $\mathcal{P}$ to $\mathcal{P}'$. Then we have e.g. the lower normalized distance between nodes E and G as $\bar{d}_l(E,G) = 1/3$; the distance discrepancy between the two nodes E, G in virtue of f as $\delta(E,G) = |1/3 - 3/5| = .267$; the entire distance discrepancy at the node E as $\delta_f(E) = 2/5$; and finally the distance discrepancy for the entire alignment as $\delta(f) = .47$.

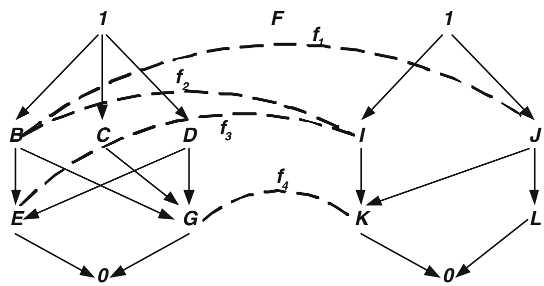

Fig. 5. An example alignment

6 Order Metrics for Ontology Clustering

Consider the following question. Assume a (portion of a) taxonomy is represented as a finite, non-empty poset $\mathcal{P} = \langle P, \leq \rangle$, and then we're given a collection of nodes $Q \subseteq P$. How "big an area" does Q "implicate" or "delineate" or "occupy" in the hierarchy? We are pursuing this question now in the context of determining the quality of ontology query returns: the "tighter" a set of ontology nodes returned from an ontology query, the stronger the quality of that set.

For two nodes $a, b \in P$, they are *comparable* $a \sim b$ if $a \leq b$ or $a \geq b$. If $a \leq b \in P$ then define the *order interval* $[a,b] := \{c \in P : a \leq c \leq b\}$. Note that $[a, b] = \uparrow a \cap \downarrow b$. Now consider two typical order metrics, the upper and lower cardinality metrics:

$$d_u(a,b) := |\uparrow a| + |\uparrow b| - 2|\uparrow a \cap \uparrow b|, \quad d_l(a,b) := |\downarrow a| + |\downarrow b| - 2|\downarrow a \cap \downarrow b|, \tag{9}$$

for $a, b \in P$. From the triangle inequality of d, we know that $\forall a, b, c \in P, d(a,b) \leq d(a,c) + d(c,b)$. So following [3,4], for two points $a, b \in P$ and metric d, we can define the *segment* $[[a,b]]_d$ as the set of all nodes which are "between" them in the metric sense:

$$[[a,b]]_d := \{c \in P : d(a,b) = d(a,c) + d(c,b)\}. \tag{10}$$

We know that $\forall a, b \in P, [[a,b]] \neq \emptyset$, since $a, b \in [[a,b]]$; and when nodes are comparable, segments collapse to order intervals: $a \sim b \rightarrow [[a,b]] = [a,b]$.

Consider the three-cube in Fig. 6, with d_u shown in Table 1, we have $d_u(B,G) = 4$, and $[[B,G]] = \{a \in P : d_u(a,B) + d_u(a,G) = 4\} = \{A, B, C, D, G\}$.

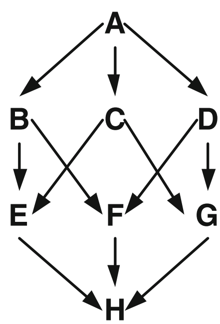

Fig. 6. The Boolean 3-cube

Table 1. Upper distance matrix d_u in the 3 cube

$d_u(a,b)$	A	B	C	D	E	F	G	H
A	0	1	1	1	3	3	3	7
B	1	0	2	2	2	2	4	6
C	1	2	0	2	2	4	2	6
D	1	2	2	0	4	2	2	6
E	3	2	2	4	0	4	4	4
F	3	2	4	2	4	0	4	4
G	3	4	2	2	4	4	0	4
H	7	6	6	6	4	4	4	0

Convexity is the idea that any nodes between other nodes in a collection $Q \subseteq P$ are also in that collection, so that a subset of nodes $Q \subseteq P$ is convex if $\forall a, b \in Q, [[a,b]] \subseteq Q$. We can then define $C(Q)$, the *convex hull* of Q, by the following iterative algorithm using the function $K(Q) = \bigcup_{a,b \in Q} [[a,b]]_d$.

```
Let Q̂ := Q
While Q̂ is not convex{
    Q̂ := K(Q̂)
}
RETURN Q̂
```

The convex hull $C(Q)$ is clearly convex, and includes the original set: $Q \subseteq C(Q)$.

Consider again a poset $\mathcal{P} = \langle P, \leq \rangle$ with metric d, and a subset of nodes $Q \subseteq P$. Then we can define the *exterior points* as those outside the convex hull: $E(Q) := P \setminus C(Q)$; and the *interior points* as those inside the convex hull, but not in the original collection: $I(Q) := C(Q) \setminus Q$.

For a subset of nodes $Q \subseteq P$ we have its *size*

$$S(Q) := |Q|, \quad \bar{S}(Q) := \frac{S(Q)}{S(P)} \tag{11}$$

in both un-normalized and normalized forms, and similarly the *dispersion*

$$D(Q) := \sum_{a,b \in C(Q)} d(a,b), \quad \bar{D}(Q) := \frac{D(Q)}{D(P)}. \tag{12}$$

Continuing our example from Fig. 6, still using d_u, consider the set $Q=\{B,E,G\}$. Then we have

$$\begin{aligned} C(Q) &= [[B,E]] \cup [[B,G]] \cup [[E,G]] \\ &= \{B,E\} \cup \{A,B,C,D,G\} \cup \{E,C,G\} = \{A,B,C,D,E,G\} \end{aligned} \tag{13}$$

This is shown in Fig. 7. We have exterior points $E(Q) = \{F,H\}$ and interior points $I(Q) = \{A,C,D\}$. We also have $D^-(Q) = 10, D(Q) = 35$, and $D(P) = D^-(P) = 91$. So, note that while the normalized dispersion is $\bar{D}(Q) = 35/91 = 0.385\%$, the relative size is $S(Q) = 3/8 = 0.375 \leq 0.385 = \bar{D}(Q)$.

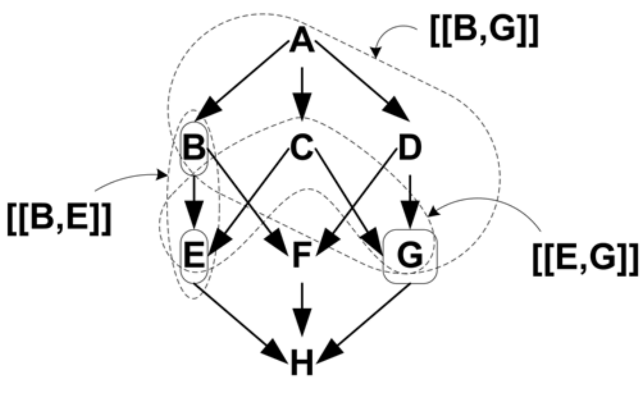

Fig. 7. The 3-cube identifying $C(\{B,E,G\})$

Now consider $Q' = \{E, G\} \subseteq Q$, then we have:

$$C(Q') = [[E, G]] = \{E, C, G\}, \quad E(Q') = \{A, B, D, F, H\}, \tag{14}$$

$$I(Q') = \{C\}, \quad V(Q') = C(Q) = \{E, C, G\}. \tag{15}$$

Note that this last result $V(Q) = C(Q)$ holds whenever $|Q| = 2$. Finally we have

$$S(Q') = 2, \qquad \bar{S}(Q') = 0.25, \tag{16}$$

$$D(Q') = D^-(Q') = 3, \qquad \bar{D}(Q') = \bar{D}^-(Q') = 3/91 = 0.033. \tag{17}$$

It is valuable to compare the above approach to a typical approach used in semantic analysis, which is to work not within the poset $\mathcal{P}$ as a directed graph, but rather the undirected, symmetrically-closed version of $\mathcal{P}$ wherein link directions and thus hierarchical structure are not recognized. Let $\mathcal{G}(\mathcal{P}) := \langle P, R \rangle$, where $R \subseteq P^2$ and $\forall a, b \in P, a, b \in R \leftrightarrow a \sim b$. The metric $d_p(a, b)$ is then the minimum path length in $\mathcal{G}(\mathcal{P})$ of a and b.

In our original example with $Q = \{B, E, G\}$, we have $E \sim B$, so that $[[B, E]]_{d_u} = [[B, E]]_{d_p} = \{B, E\}$. But $[[E, G]]_{d_p} = \{E, C, G, H\}$, and $[[B, G]]_{d_p} = P$, because B and G are inverses. Thus the convex hull is $C_{d_p}(Q) = P$, and Q can be said to be of maximal size. This is clearly inadequate.

References

1. Aho, A.V., Garey, M.R., Ullman, J.D.: The Transitive Reduction of a Directed Graph. SIAM Journal of Computing 1(2), 131–137 (1972)
2. Ashburner, M., Ball, C.A., Blake, J.A., et al.: Gene Ontology: Tool For the Unification of Biology. Nature Genetics 25(1), 25–29 (2000)
3. Bandelt, H.J.: Centroids and Medians of Finite Metric Spaces. J. Graph Theory 16(4), 305–317 (1992)
4. Bandelt, H.J., Chepoi, V.: Metric Graph Theory and Geometry: A Survey. In: Surveys on Discrete and Computational Geometry: Twenty Years Later, vol. 453, pp. 49–86. American Math. Soc., Providence (2008)
5. Butanitsky, A., Hirst, G.: Evaluating WordNet-based Measures of Lexical Semantic Relatedness. Computational Linguistics 32(1), 13–47 (2006)
6. Davey, B.A., Priestly, H.A.: Introduction to Lattices and Order, 2nd edn., Cambridge UP, Cambridge UK (1990)
7. Euzenat, J., Shvaiko, P.: Ontology Matching. Springer, Hiedelberg (2007)
8. Fellbaum, C. (ed.): Wordnet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
9. `http://www.w3.org/RDF`
10. `http://www.w3.org/TR/owl-features`
11. Joslyn, C., Baddeley, B., Blake, J., Bult, C., et al.: Automated Annotation-Based Bio-Ontology Alignment with Structural Validation. In: Smith, B. (ed.) Proc. Int. Conf. on Biomedical Ontology (ICBO 2009), pp. 75–78 (2009), doi:10.1038/npre.2009.3518.1

12. Joslyn, C., Mniszewski, S., Fulmer, A., Heaton, G.: The Gene Ontology Categorizer. Bioinformatics 20(s1), 169–177 (2004)
13. Joslyn, C., Mniszewski, S.M., Smith, S.A., Weber, P.M.: SpindleViz: A Three Dimensional, Order Theoretical Visualization Environment for the Gene Ontology. In: Joint BioLINK and 9th Bio-Ontologies Meeting, JBB 2006 (2006), http://www.bio-ontologies.org.uk/2006/download/Joslyn2EtAlSpindleviz.pdf
14. Joslyn, C., Paulson, P., White, A.: Measuring the Structural Preservation of Semantic Hierarchy Alignments. In: Proc. 4th Int. Wshop. on Ontology Matching (OM 2009), CEUR, vol. 551 (2009), http://ceur-ws.org/Vol-551/om2009_Tpaper6.pdf
15. McBride, B.: Jena: A Semantic Web Toolkit. IEEE Internet Computing 6(6), 55–59 (2002)
16. Monjardet, B.: Metrics on Partially Ordered Sets - A Survey. Discrete Mathematics 35, 173–184 (1981)
17. Orum, C., Joslyn, C.A.: Valuations and Metrics on Partially Ordered Sets (2009) (submitted), http://arxiv.org/abs/0903.2679v1
18. Verspoor, K.M., Cohn, J.D., Mniszewski, S.M., Joslyn, C.A.: A Categorization Approach to Automated Ontological Function Annotation. Protein Science 15, 1544–1549 (2006)