Space-Complexity of Hypertries

Lemma 1. *Consider a hypertrie $h \in H(d, A, E)$, that stores equal slices only once. Its number of edges is bound by $\mathcal{O}(z(h) \cdot 2^{d-1} \cdot d)$ and its number of nodes is bound by $\mathcal{O}(z(h) \cdot (2^d - 1))$.*

Proof. The idea for the proof is (1) to identify a worst-case entry set and to affiliate (2) the number of nodes per entry and level from definition 3 as well as (3) the number of edges per node by level for such a entry set. (4) Finally, (2) and (3) can be combined to an upper bound.

(1) Consider a set $K \subset A^d$ of keys such that for every key part of every key, no other key has the same key part in the same position, i.e., $\forall k, k' \in K, k \neq k' : k_p \neq k'_p$. A hypertrie mapping values by these keys requires a maximum number of nodes as by construction two different keys share no common node but the root node.

(2) As subhypertries for equal slices are stored only once, the slice keys for nodes with depth $(d - i)$ are given by fixing the key parts at any i key positions. This results in $\binom{d}{i}$ nodes with depth $(d - i)$ being stored, one for every such slice key.

(3) By definition a node with depth $(d - i), i > 0$ has at least one outgoing edge for every position $1 \ldots (d - i)$. As every node of a hypertrie storing values by the keys K is reachable only on one path from the root node, every hypertrie with depth $d - i, i > 0$ represents exactly one subkey of one key. Thus, for every position it has exactly one edge that points to a slice by one key part. In total, this results in $(d - i)$ edges per node with depth $(d - i), i > 0$. For ever key, the root node has d outgoing edges as the key parts for every positions are unique by precondition.

(4) Summing up the nodes per level per key by level results in the total amount of nodes per key:

$$\sum_{1}^{d} \binom{d}{i} = 2^d - 1$$

Summing up the outgoing edges from nodes per level per key by level results in the total amount of nodes per key:

$$\sum_{0}^{d} \binom{d}{i} \cdot (d - i) = 2^{d-1} \cdot d$$

This results in upper bounds of $\mathcal{O}(z(h) \cdot 2^{d-1} \cdot d)$ edges and $\mathcal{O}(z(h) \cdot (2^d - 1))$ nodes.