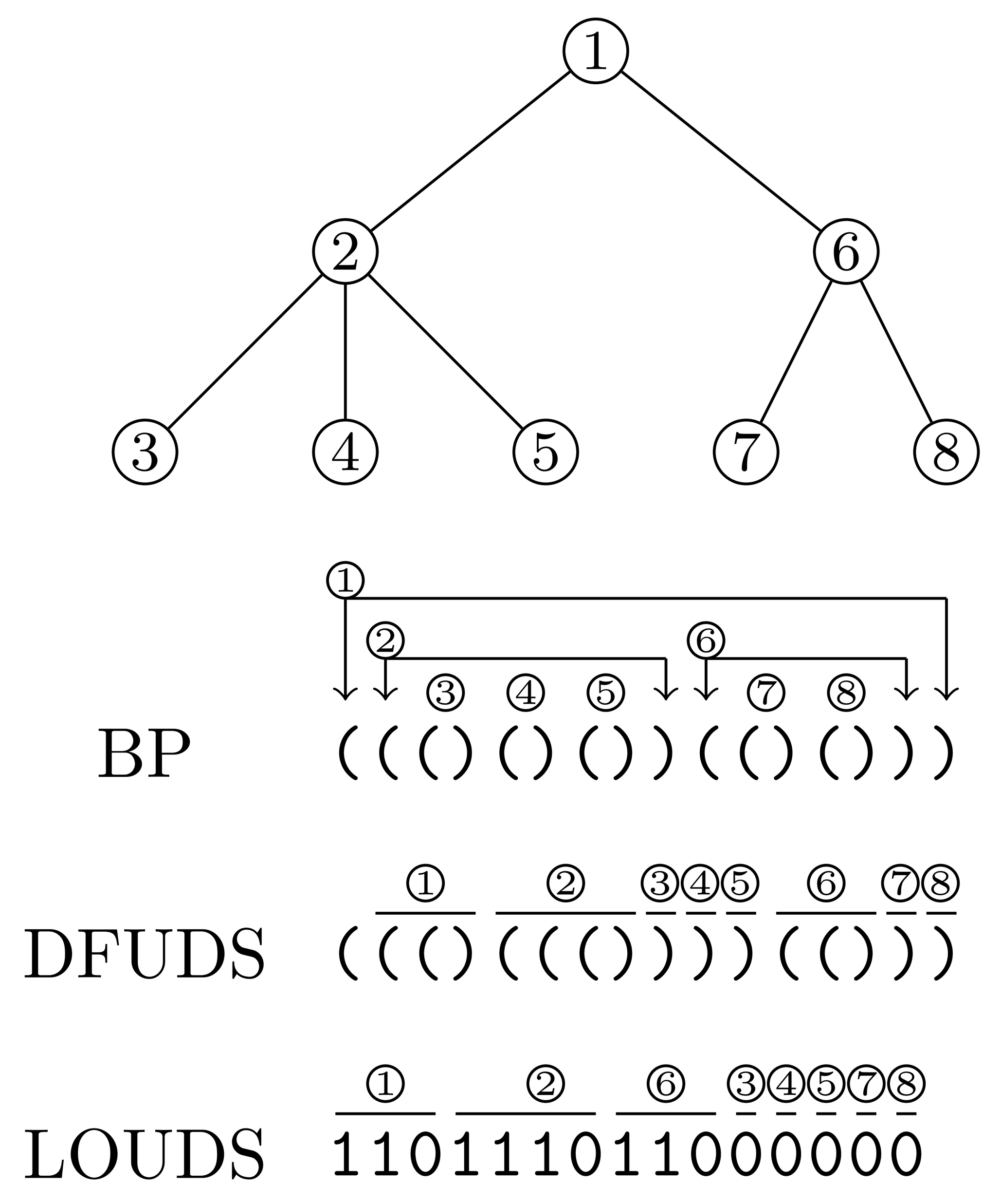
Succint Trees and Trees Compression

Gabriel Alfredo Carmona Tabja

Succint Trees



Space lower bound: $2n+\theta(\log n)$ **bits**
Classical solutions support operations 1-13 in constant time using $2n+o(n)$ bits [1, 2, 3]

Operations

1. `findclose(P, x)`
2. `findopen(P, x)`
3. `enclose(P, x)`
4. `parent(x)`
5. `firstchild(x)`
6. `sibling(x)`
7. `depth(x)`
8. `desc(x)`
9. `pre_rank(x)`
10. `pre_select(i)`
11. `level_ancestor(x, d)`
12. `lca(x, y)`
13. `degree(x)`
14. `child(x, i)`
15. `child_rank(x)`
16. `level_next(x)`
17. `level_prev(x)`
18. `level_lmost(x, d)`
19. `level_rmost(x, d)`
20. `deepest_node(x)`
21. `height(x)`
22. `in_rank(x)`
23. `post_rank(x)`
24. `in_select(i)`
25. `post_select(i)`
26. `leaf_rank(x)`
27. `leaf_select(i)`
28. `lmost_leaf(x)`
29. `rmost_leaf(x)`
30. `last_child(x)`

Labeled Trees

- XBWT by Ferragina et al. [4]
  - Consider k preceding labels along a path for any given label c.
- RL-XBWT by Prezza, 2001 [5]
  - Consider outgoing lables for a path of k preceding labels.

Repetitive Subtrees

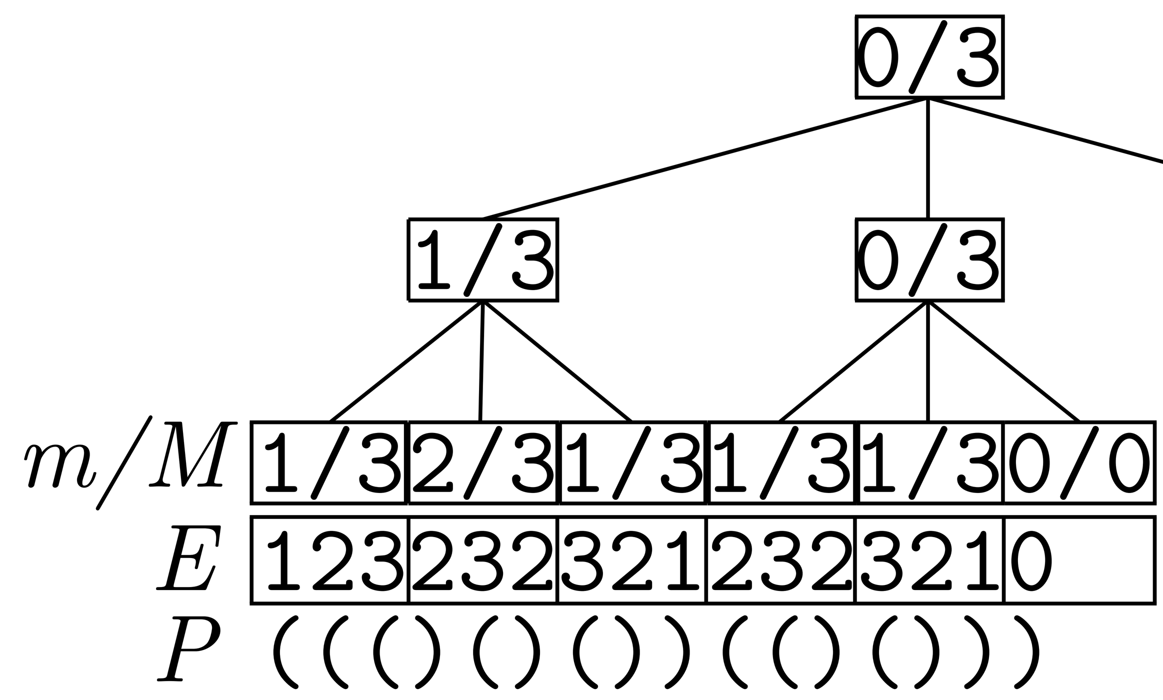
- DAG Compression
  - Can be done on $O(n)$ by Downey et al., 1980 [6] .
- Tree Grammar Compression by Lohrey and Maneth, 2006 [7]
  - Finding the minimal tree grammar is NP-Hard [8]

Degree Repetition

- Tree Degree Entropy by Jansson et al., 2007. [9]

$$H^*(T) = \sum_i \frac{n_i}{n} \log \frac{n}{n_i}$$

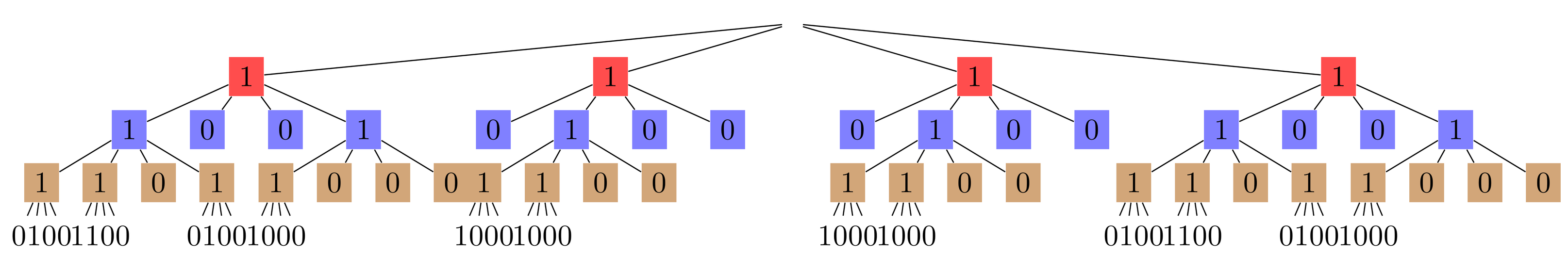
Range Min-Max Tree



- Fully Functional Succinct Tree by Sadakane and Navarro [10].
- All operations in constant time. Useful for BP and DFUDS.
- Space usage: $2n + O(n/log^c n)$ bits for $c > 0$.

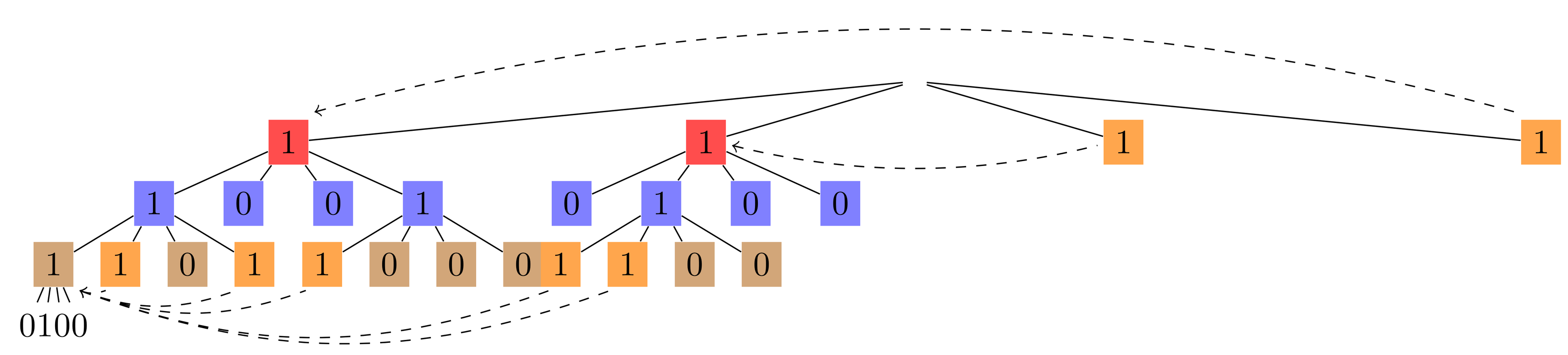
k^2-tree [11]

0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0
0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



- T: 1111 1001 0100 0100 1001 1101 1000 1100 1100 1101 1000
- L: 0100 1100 1000 1000 1000 1000 1000 0100 1100 0100 1000

Example Minimal Dag



Open Questions

- Can we combine succint representation with tree compression?
- Can we say how compressible a tree is, based on subtree repetitions in general trees?
- Is there a relation between repetitive data and repetitive subtrees?
- Can we do every operation in succint trees but in a compressed succint tree?

References

[1] Guy Jacobson. "Space-efficient static trees and graphs". In: *30th annual symposium on foundations of computer science*. IEEE Computer Society. 1989.

[2] J Ian Munro and Venkatesh Raman. "Succinct representation of balanced parentheses and static trees". In: *SIAM Journal on Computing* 31.3 (2001).

[3] David Benoit et al. "Representing trees of higher degree". In: *Algorithmica* 43 (2005).

[4] Paolo Ferragina et al. "Compressing and indexing labeled trees, with applications". In: *J. ACM* 57.1 (Nov. 2009). ISSN: 0004-5411.

[5] Nicola Prezza. "On locating paths in compressed tries". In: *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2021.

[6] Peter J Downey, Ravi Sethi, and Robert Endre Tarjan. "Variations on the common subexpression problem". In: *Journal of the ACM (JACM)* 27.4 (1980).

[7] Markus Lohrey and Sebastian Maneth. "The complexity of tree automata and XPath on grammar-compressed trees". In: *Theoretical Computer Science* 363.2 (2006).

[8] Moses Charikar et al. "The smallest grammar problem". In: *IEEE Transactions on Information Theory* 51.7 (2005).

[9] Jesper Jansson, Kunihiko Sadakane, and Wing-Kin Sung. "Ultra-succinct representation of ordered trees". In: *SODA*. Vol. 7. 2007.

[10] Gonzalo Navarro and Kunihiko Sadakane. "Fully functional static and dynamic succinct trees". In: *ACM Transactions on Algorithms (TALG)* 10.3 (2014).

[11] Nieves R Brisaboa, Susana Ladra, and Gonzalo Navarro. "k2-trees for compact web graph representation". In: *International symposium on string processing and information retrieval*. Springer. 2009.