Making Graphs Really Hard to Say Out Loud

Graph Labelings with Non-Disjoint Vertex and Edge Sets

and how they can be used for encryption, poetry, and breaking mathematics

Gabriel Chuang (gtchuang@andrew.cmu.edu)
Carnegie Mellon University

***Abstract*— Modern communication systems have greatly increased the ease with which we can describe our ideas to others. In an effort to slow this progress and hinder scientific advancement, we examine labelings of graphs that *reduce* clarity by reusing edge labels for vertices. We bound how much overlap can be achieved, show how such labelings are cryptographically secure, and prove some theorems that suggest that all of mathematics is broken. Finally, we present some applications to poetry and propose a new naming convention for graphs based on these labelings, which we hope will be widely adopted.**

I. Introduction

In Chapter 1.1 of Diestel's widely-used textbook *Graph Theory*, a graph is defined to be a pair of sets (E, V) with $E \subseteq [V]^2$. Diestel follows this up with the following note:

> To avoid notational ambiguities, we shall always assume tacitly that $V \cap E = \varnothing$.

Needless to say, this is very disappointing to fans of notational ambiguity. There are many settings where being notationally ambiguous is beneficial, such as:

- When brevity is of the essence (e.g. when lecturing at a blackboard),
- When trying to sneak faulty proofs past reviewers,
- When trying to confuse students, and
- When trying to confuse yourself.

In this work, we will explore labelings of graphs in which we seek to maximize the overlap between the labels of the vertex set V and edge set E. (Note that we will continue to require that edges are two-element subsets of the vertices). See Fig. 2 for an example of such a graph. Note that that graph has an edge $\{a, b\}$, but also a *vertex* $\{a, b\}$!

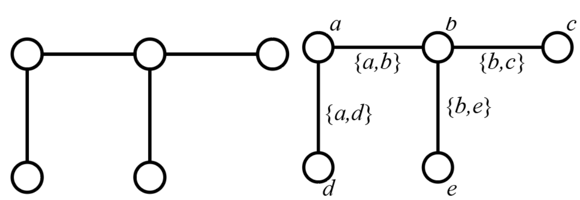
Such labelings have many desirable attributes; most notably, graphs represented in this way are extremely notationally ambiguous, which can allow us to prove interesting theorems. They are also very hard to say out loud, due to repetition in their vertex and edge sets.

These labelings may also be of interest to those seeking to keep their graphs concealed from adversaries who can only communicate verbally. We present a cryptographic protocol that is in fact secure against *all* adversaries.

Finally (and of purely academic interest)[1], such labelings of graphs often resemble rhyme schemes; we present several graphs whose edge sets correspond to common rhyme schemes or famous poems, and propose an alternative naming system based on poetic forms.

II. Motivation

Consider the following graph, and a natural labeling of its vertices:



(a) A simple undirected graph G.

(b) A "natural" labeling of G, with vertex set $V = \{a, b, c, d, e\}$.

Fig. 1: A simple undirected graph, and a natural labeling.

Unfortunately, this labeling is extremely boring: there is no overlap between V and E! Instead, consider the following labeling:

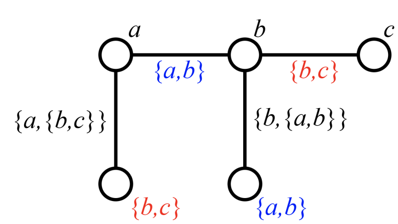


Fig. 2: A much more interesting labeling of the graph. The vertex set is $V = \{a, b, c, \{a, b\}, \{b, c\}\}$.

[1] just like the rest of this paper, to be clear

We've managed to achieve an overlap of size 2: $V \cap E = \{\{a,b\},\{b,c\}\}$. That is, there are two vertices that share a label with an edge. Can we do better on this graph? As it turns out, we can:

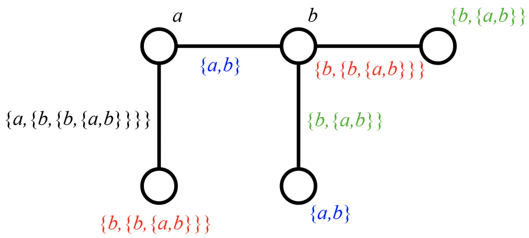


Fig. 3: An overlap-maximal labeling of the graph. The vertex set is $V = \{a, b, \{a,b\}, \{b,\{a,b\}\}, \{b,\{b,\{a,b\}\}\}\}$. In this case, three of four edges share a label with a vertex.

This labeling of the graph is has many desirable properties. For example, the verbalization is extremely ambiguous:

> "Vertices: A, B, A, B, B, A, B, and B, B, A, B. Edges: A to B, B to A, B, B to B, A, B, and A to B, B, A, B." [2]

In general, it is not obvious how much overlap can be achieved for arbitrary graphs. Fig. 3 achieves overlap on 3 of 4 edges; can one do better? What classes of graphs enable maximum overlap?

A central point of this paper will be to carefully define such labelings and to show upper and lower bounds on how much overlap can be achieved.

III. Definitions and Central Claim

Definition 1 (Overlap number)**.** *For a graph $G = (V,E)$, let the overlap number $\psi(G)$ be the maximum size of $E \cap V$ over all possible labelings L of G, i.e., the maximum number of edges that can share a name with a vertex (or vice versa). That is,*

$$\psi(G) = \max_{L:V\to\mathbb{U}} |L(E) \cap L(V)|.$$

Definition 2 (Overlap ratio)**.** *The overlap ratio $\rho(G)$ of a graph $G = (V,E)$ is defined to be*

$$\rho(G) = \frac{\psi(G)}{|E|}$$

In the example graph of Fig. 3, $\psi(G) = 3$ and $\rho(G) = \frac{3}{4}$.[3]

[2]We could've done even better by naming the vertices "and" and "to" instead of A and B.

[3]In the spirit of being notationally ambiguous, we will abuse notation and use $\rho(G)$ for non-maximal labelings as well.

Theorem 1. *For any graph $G = (V,E)$,*

$$\psi(G) = \min(|E|, |V| - 2).$$

In particular, since connected graphs have $|E| \geq |V| - 1$,

Corollary 1. *For a connected graph $G = (V,E)$,*

$$\psi(G) = |V| - 2$$

The proof of Theorem 1 shows that $\min(|E|, |V| - 2)$ is both an upper and lower bound for $\psi(G)$, using a partial order argument for the upper bound and an explicit construction for the lower bound. To spare you the gory details, the proof has been relegated to Appendix A.

IV. Overlap Ratio and Pretty Pictures

In this section, we will show some examples of maximum-overlap labelings of a few common classes of graphs, and discuss the overlap ratio induced by those labelings.[4]

A. Complete graphs K_n

K_n has $\binom{n}{2}$ edges. Theorem 1 implies that

$$\rho(K_n) = \frac{\psi(K_n)}{\binom{n}{2}} = \frac{2(n-2)}{n(n-1)} = \Theta\left(\frac{1}{n}\right).$$

Maximum-overlap labelings of K_5 and K_7 are shown in Fig. 4.

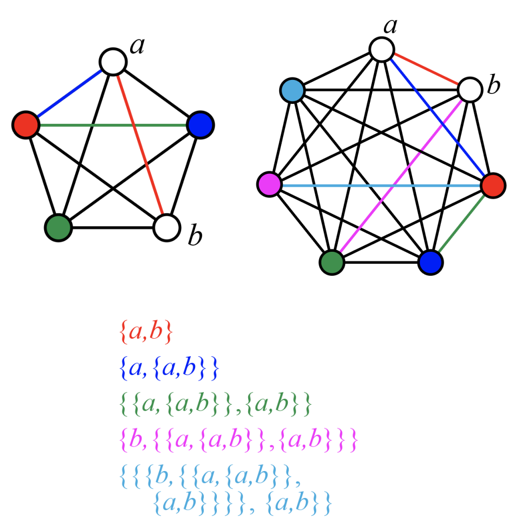


Fig. 4: Overlap-maximal labelings for K_5 and K_7. Note that very few edges are reused for vertex labels.

The edge set for a max-overlap labeling of K_7 can be found in Appendix B.

[4]More importantly, we will have many colorful figures, to maximize appeal to children, advertisers, and color theory gremlins.

B. Cycles C_n

C_n has n edges. Theorem 1 implies that

$$\rho(K_n) = \frac{\psi(C_n)}{n} = \frac{n-2}{n} = 1 - \frac{2}{n} \sim 1.$$

An overlap-maximal labeling of C_n is shown in Fig. 5

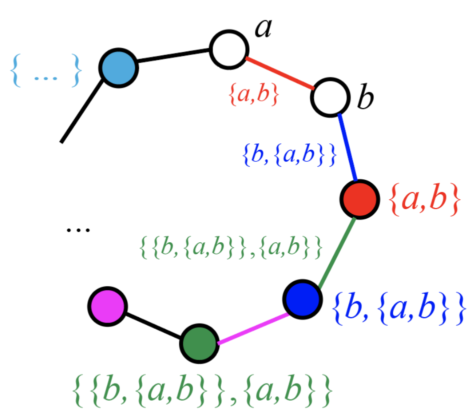


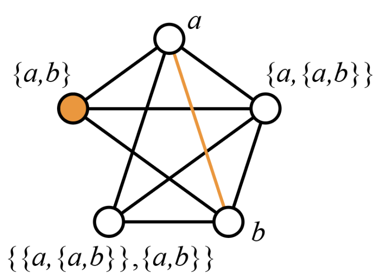
Fig. 5: Overlap-maximal labeling for C_n. Nearly every edge label is reused.

V. Notational Ambiguity Enables Chaos

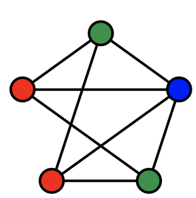
Now that we have our framework for labeling graphs, we will, naturally, attempt to prove some theorems.

A. Graph coloring

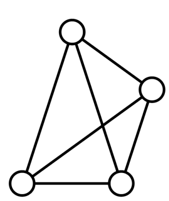
Let G be the graph shown below in Fig. 6a.



(a) A nice, well behaved graph G on 5 vertices. $\{a,b\}$ is highlighted in orange.



(b) $G \setminus \{\{a,b\}\}$



(c) $G \setminus \{\{a,b\}\}$

Fig. 6: Our graph G, and the subgraph $G \setminus \{\{a,b\}\}$.

Here are two theorems, both of which are true:

Theorem 2. *$G \setminus \{\{a,b\}\}$ is 3-colorable.*

Proof: A 3-coloring is shown in Fig. 6b. □

Theorem 3. *$G \setminus \{\{a,b\}\}$ is NOT 3-colorable.*

Proof: $G \setminus \{\{a,b\}\}$ is the complete graph K_4, as shown in Fig. 6c, and K_4 is not 3-colorable. □

Obviously, this is somewhat problematic. We might be tempted to claim that k-colorability is in fact either poorly defined or undecidable. However, complexity theorists have put a lot of work into proving that 3COL is NP-complete and thus in NP, which, by necessity, means that it is decidable.

In fact, simple extensions of Theorem 2 and Theorem 3 show that

$$\chi(G \setminus \{\{a,b\}\}) = 3 \quad \text{and} \quad \chi(G \setminus \{\{a,b\}\}) = 4,$$

where $\chi(G)$ is the chromatic number of G (i.e., the minimum number of colors needed to color G).

We are forced to conclude that $3 = 4$, and that in fact $0 = 1$.

B. Adjacency

One might be tempted to cut out the notion of coloring altogether, in the hope that the rest of graph theory is salvageable. Unfortunately, even the notion of adjacency yields contradiction. Let H be this graph:

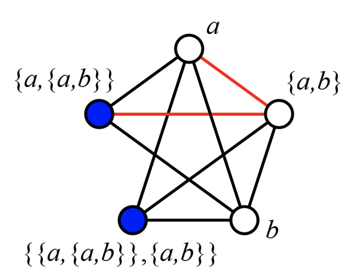


Fig. 7: Another perfectly well-behaved graph H.

Theorem 4. *In H, $\{\{a,\{a,b\}\},\{a,b\}\}$ is adjacent to $\{a,\{a,b\}\}$.*

Proof: $\{\{a,\{a,b\}\},\{a,b\}\} \cap \{a,\{a,b\}\} = \{\{a,b\}\}$, which is nonempty, so $\{\{a,\{a,b\}\},\{a,b\}\} \sim \{a,\{a,b\}\}$. □

Theorem 5. *In H, $\{\{a,\{a,b\}\},\{a,b\}\}$ is NOT adjacent to $\{a,\{a,b\}\}$.*

Proof: By the definition of H (as in Fig. 7), $\{\{a,\{a,b\}\},\{a,b\}\}$ is not adjacent to $\{a,\{a,b\}\}$. □

Thus, even adjacency, the most basic foundation of graph construction, leads to contradiction.

VI. Max-Overlap Labelings for Encryption

Maximum-overlap labelings also have great potential for use in encryption protocols.

Theorem 6. *A maximum-overlap labeling of a graph $G = (V, E)$ is cryptographically secure against all adversaries.*

Proof: Suppose that Alice wants to transmit graph $G = (V, E)$ to Bob. Alice computes the max-overlap labeling L of the graph and sends $(L(V), L(E))$, which Eve intercepts. One of two cases must occur:

- Eve attempts to read the input and falls asleep, or
- Eve concludes that Alice and Bob are insane, and that their transmitted graph must be worthless.

In both cases, Eve gains no information about G. □

VII. Poetry

A. The Limerick Graph

Let's take a look the max-overlap labeling of P_2, the length-2 path, as shown in 8:

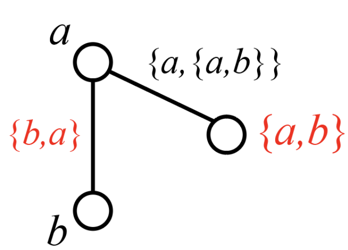


Fig. 8: Max-overlap labeling for P_2. As it turns out, this is the Limerick Graph $AABBA$.

The vertex and edge sets are

$$V = \{a, b, \{a, b\}\}$$

$$E = \{\{a, \{a, b\}\}, \{b, a\}\}$$

Verbalized, the edge set is "A, A, B, B, A", or "AABBA." This looks an awful lot like a rhyme scheme, doesn't it? Specifically, AABBA is the rhyme scheme for a limerick[5]. In general, it seems like we ought to be able to get graphs that correspond to other rhyme schemes.

[5]For example:
To label a 3-vertex path
For usage in improper math
Use $\{\{a, b\}, a, b\}$
On v_1 through $_3$;
Voila! It's a limerick graph.

B. The Sicilian Octave Graph

A Sicilian Octave is an eight-line poem with alternating rhymes; i.e., its rhyme scheme is ABABABAB. Consider this max-overlap labeling of the 3-clique K_3:

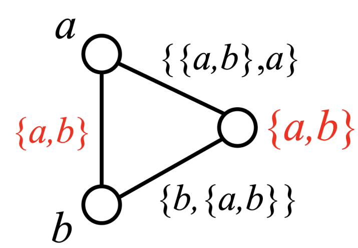


Fig. 9: Max-overlap labeling for K_3, the Sicilian Octave graph, with $\rho(G) = \frac{1}{3}$.

The edge set is

$$E = \{\{a, b\}, \{\{a, b\}, a\}, \{b, \{a, b\}\}\},$$

that is, ABABABAB.

C. The Shakespearean Sonnet Graphs (order 1 and 2)

We now explore several rhyme-scheme encoding graphs that may not necessarily be overlap-maximal, and which allow self-loops. This allows us to further bridge the (currently, very large) gap between mathematics and poetry.[6]

Shakespearean sonnets have a rhyme scheme of ABAB CDCD EFEF GG. Fig. 10 displays a graph with vertex and edge sets

$$V = \{a, b, c, d, e, f, g, \{a, b\}, \{c, d\}, \{e, f\}, \{g, g\}\}$$

$$E = \{\{a, b\}, \{\{a, b\}, \{c, d\}\}, \{c, d\}, \{e, f\}, \{\{e, f\}, \{g, g\}\}\}.$$ [7]

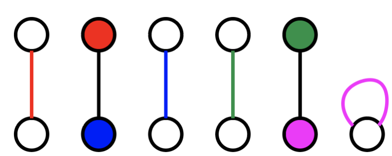


Fig. 10: The order-1 Shakespearean sonnet graph is a forest of K_2s, plus one self loop. It achieves $\rho(G) = \frac{2}{3}$.

This is a rather disappointing graph. Luckily for us, Shakespeare wrote more than one sonnet. Fig. 11 shows the graph corresponding to *two* Shakespearean sonnets.

(Shakespeare wrote 153 sonnets; the order-153 Shakespearean sonnet graph is left as an exercise to the reader.)

[6]The two fields have been slowly growing apart since Omar Khayyam, although Lewis Carroll made a valiant effort to bring them back together.

[7]I know this overlaps the margin of the page. shhhhh.

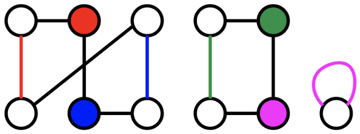


Fig. 11: The order-2 Shakespearean sonnet graph, with edges $E = \{\{a, b\}, \{\{a, b\}, \{c, d\}\}, \{c, d\}, \{e, f\}, \{\{e, f\}, \{g, g\}\}, \{\{a, b\}, a\}, \{b, c\}, \{d, \{c, d\}\}, \{\{e, f\}, e\}, \{f, \{g, g\}\}\}$, i.e. ABAB CDCD EFEF GG ABAB CDCD EFEF GG. $\rho(G) = \frac{4}{11}$.

D. *Other Poems (Selected)*

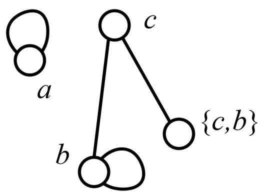


Fig. 12: Rhyme scheme graph of *The Raven* by Edgar Allen Poe, encoding rhyme scheme AABCCCBBB. $E = \{\{a, a\}, \{b, c\}, \{c, \{c, b\}\}, \{b, b\}\}$ with $\rho(G) = \frac{1}{4}$.

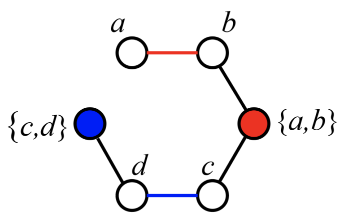


Fig. 13: Rhyme scheme graph of *The New Colossus* by Emma Lazarus (*"Give me your tired, your poor, your huddled masses yearning to breathe free"*), encoding rhyme scheme ABBAABBACDCDCD with $\rho(G) = \frac{2}{5}$.

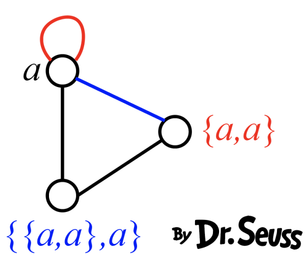


Fig. 14: Rhyme scheme graph of *Fox in Socks* by Dr. Seuss, encoding rhyme scheme AAAAAAAAAAAA. [8]

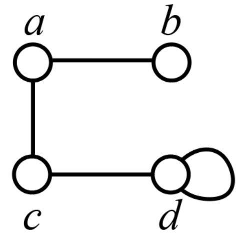


Fig. 15: Rhyme scheme graph of *Toxic* by Britney Spears, encoding rhyme scheme ABACDCDD with $\rho(G) = 0$.

VIII. Conclusion

In this work, we presented an overlap-maximization technique for labeling graphs. In order to counter the increasing ease of communication facilitated by the internet, we propose an alternate naming convention for graphs, in order to minimize clarity when spoken aloud.

In particular, scholars should strive to use either the name for a poem whose rhyme scheme encodes the graph, or directly list out the edge set of the maximal-overlap labeling of the graph.

For example:

1) the triangle graph could be validly called "AB, ABA, ABB" or "The Fox in Socks graph without self-loops";
2) the 4-cycle could be called "the second-smallest connected component of the Order 2 Shakespeare graph" or "AB, BAB, BABAB, ABAB."

Creativity is encouraged.

[8] not to be confused with the sound emitted by distressed humans

IX. References

[i] Google.

[ii] Wikipedia.

[iii] The Poetry Foundation.

All figures created in totally legal, non-pirated Adobe Illustrator.

X. Appendix A: Proof of Theorem 1

A. Upper bounding $\psi(G)$

First, we show that $\psi(G) \leq |V| - 2$.

Suppose for the sake of contradiction that $\psi(G) \geq |V| - 1$, i.e. that there are $|V| - 1$ vertices which share a label with an edge. Let these vertices be $V' = \{v_1, v_2, \cdots v_{|V|-1}\}$.

Note that these vertices are all of the form $\{x, y\}$ for some $x, y \in V$. So, $\subset$ is a strict partial order on the finite set V', so there exists at least one minimum element under $\subset$.

Let $m = \{m_1, m_2\}$ be such a minimum element. m_1 and m_2 must be vertices, since the vertex m overlaps with an edge. Since m is the minimum, $m_1, m_2 \notin V'$. So we have two vertices that are not in V', a contradiction. $\implies \impliedby$

Second, we show that $\psi(G) \leq |E|$. Note that

$$\begin{aligned} \psi(G) &= \max_{L:V\to\mathbb{U}} |L(E) \cap L(V)| && \text{(Definition 1)} \\ &\leq \max_{L:V\to\mathbb{U}} |L(E)| \\ &= \max_{L:V\to\mathbb{U}} |E| && (L \text{ bijective}) \\ &= |E| \end{aligned}$$

as desired.

B. Constructively lower bounding $\psi(G)$

We show a recursive procedure to assign labels to a graph to ensure $\min(|E|, |V| - 2)$ overlap between edge and vertex labels.

In particular, it suffices to iteratively build up the graph by adding vertices and keeping the following invariant:

Invariant: A graph $G = (V, E)$ with $|V| \geq 2$ can be $\min(|E|, |V| - 2)$-overlap-labeled, with the labeling determined entirely by the labels of two "base" vertices (in particular, the two "base" vertices can be arbitrarily renamed, and the rest of the graph relabeled accordingly, while preserving the overlap).

Base cases: The theorem trivially holds for $|V| \in \{0, 1, 2\}$ [9].

[9] Whether $|V| = 0$ is a valid graph is up for debate; see "Is the null-graph a pointless concept?" by Harary and Read, 2006

For $|V| = 3$: The labelings displayed in Fig. 16 cover all cases. Note that a, b could be labeled anything, and the rest of the graph relabeled accordingly, while keeping the invariant true.

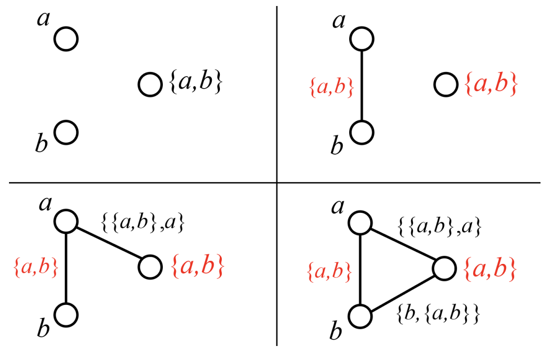


Fig. 16: Overlap-maximal labelings for $|V| = 3$, the base case for our recursive labeling scheme. Note that in each case the amount of overlap $\psi(G)$ is less than $\min(|E|, |V| - 2)$.

Inductive step. Consider some graph $G = (V, E)$ with $|V| \geq 4$. We split into cases:

- $|V| - 2 < |E|$. Then, there must be some vertex v with minimum degree. Inductively label $G \setminus v$, then choose some edge $e \in G \setminus V$ as the label for v.
- $|E| \leq |V| - 2$. Then, the graph is disconnected.
  - If one connected component is a singleton vertex, then label it arbitrarily, and inductively label the rest.
  - Otherwise, partition the graph into some connected component C and $G \setminus C$. Inductively label C and $G \setminus C$.
    Choose two edges $e_1, e_2 \in E(C) \setminus V(C)$ (i.e., they do not yet share a name with a vertex), and rename $G \setminus C$, using these e_1, e_2's labels as the "base" labels for $G \setminus C$.

It's easy to verify that this process preserves the invariant; the details are left to the reader.

XI. Appendix B: The Edge Set of K_7 Under Max-Overlap Relabeling

$$
\begin{aligned}
E(K_7) = \{ & \\
& \{a, b\}, \\
& \{a, \{a, b\}\}, \\
& \{a, \{a, \{a, b\}\}\}, \\
& \{a, \{\{a, \{a, b\}\}, \{a, b\}\}\}, \\
& \{a, \{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \\
& \{a, \{\{\{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \{a, b\}\}\}, \\
& \{b, \{a, b\}\}, \\
& \{b, \{a, \{a, b\}\}\}, \\
& \{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}, \\
& \{b, \{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \\
& \{b, \{\{\{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \{a, b\}\}\}, \\
& \{\{a, b\}, \{a, \{a, b\}\}\}, \\
& \{\{a, b\}, \{\{a, \{a, b\}\}, \{a, b\}\}\}, \\
& \{\{a, b\}, \{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \\
& \{\{a, b\}, \{\{\{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \{a, b\}\}\}, \\
& \{\{a, \{a, b\}\}, \{\{a, \{a, b\}\}, \{a, b\}\}\}, \\
& \{\{a, \{a, b\}\}, \{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \\
& \{\{a, \{a, b\}\}, \{\{\{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \{a, b\}\}\}, \\
& \{\{\{a, \{a, b\}\}, \{a, b\}\}, \{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \\
& \{\{\{a, \{a, b\}\}, \{a, b\}\}, \{\{\{b, \{\{a, \{a, b\}\}, \\
& \quad \{a, b\}\}\}\}, \{a, b\}\}\}, \\
& \{\{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}, \\
& \quad \{\{\{b, \{\{a, \{a, b\}\}, \{a, b\}\}\}\}, \{a, b\}\}\} \\
& \}
\end{aligned}
$$