

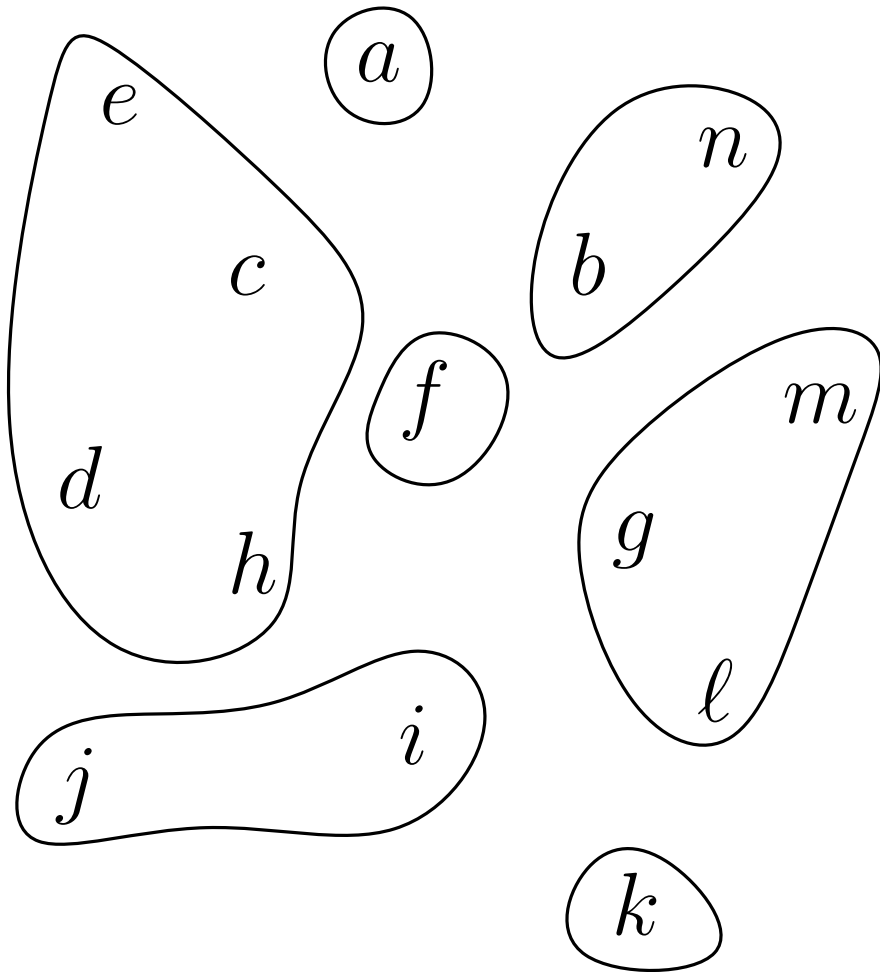
Union Find, Path Compression, and the Inverse Ackermann Function

The Data

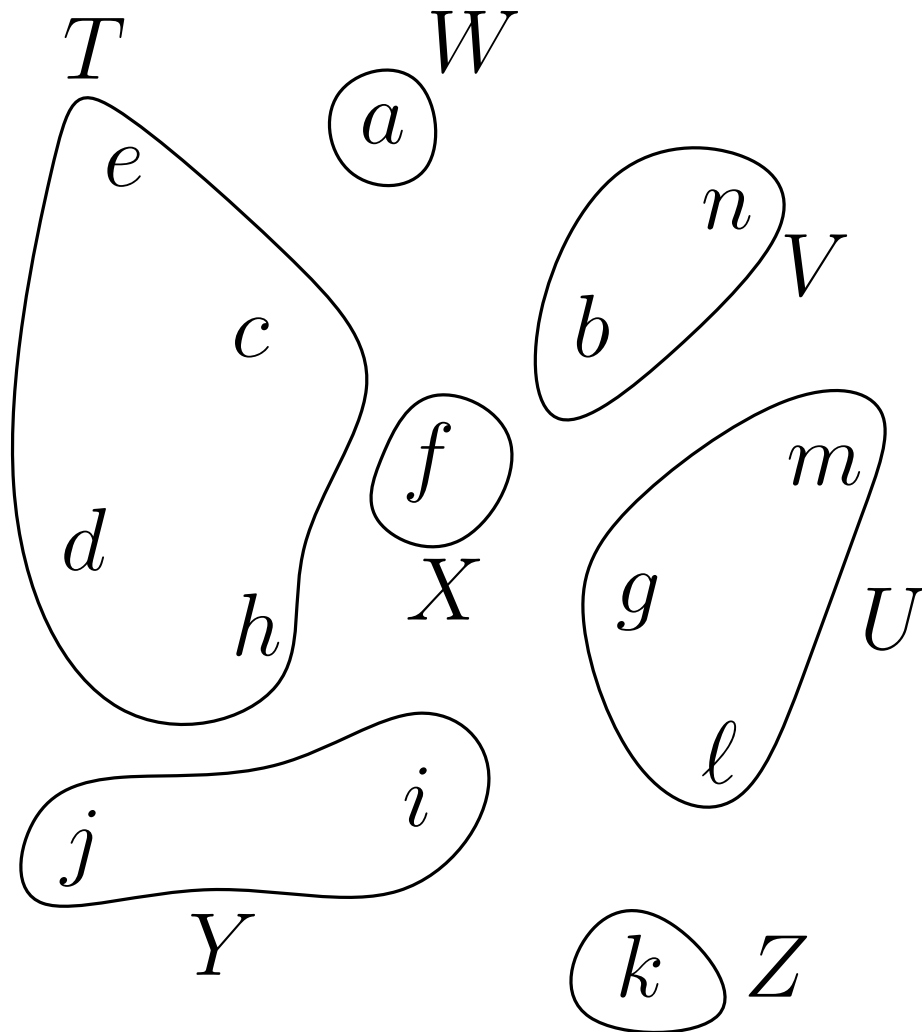
The Data

e *a*
c *b* *n*
d *f* *m*
h *g*
j *i* *l*
k

The Data

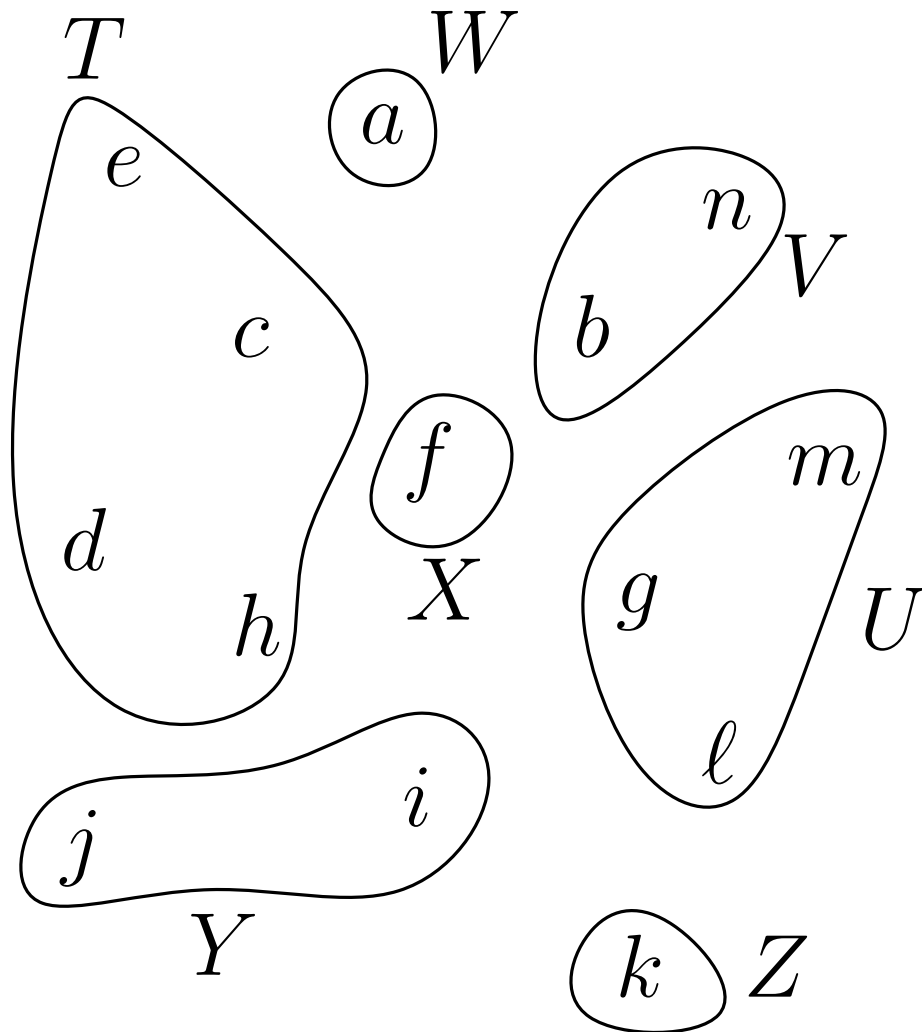


The Data



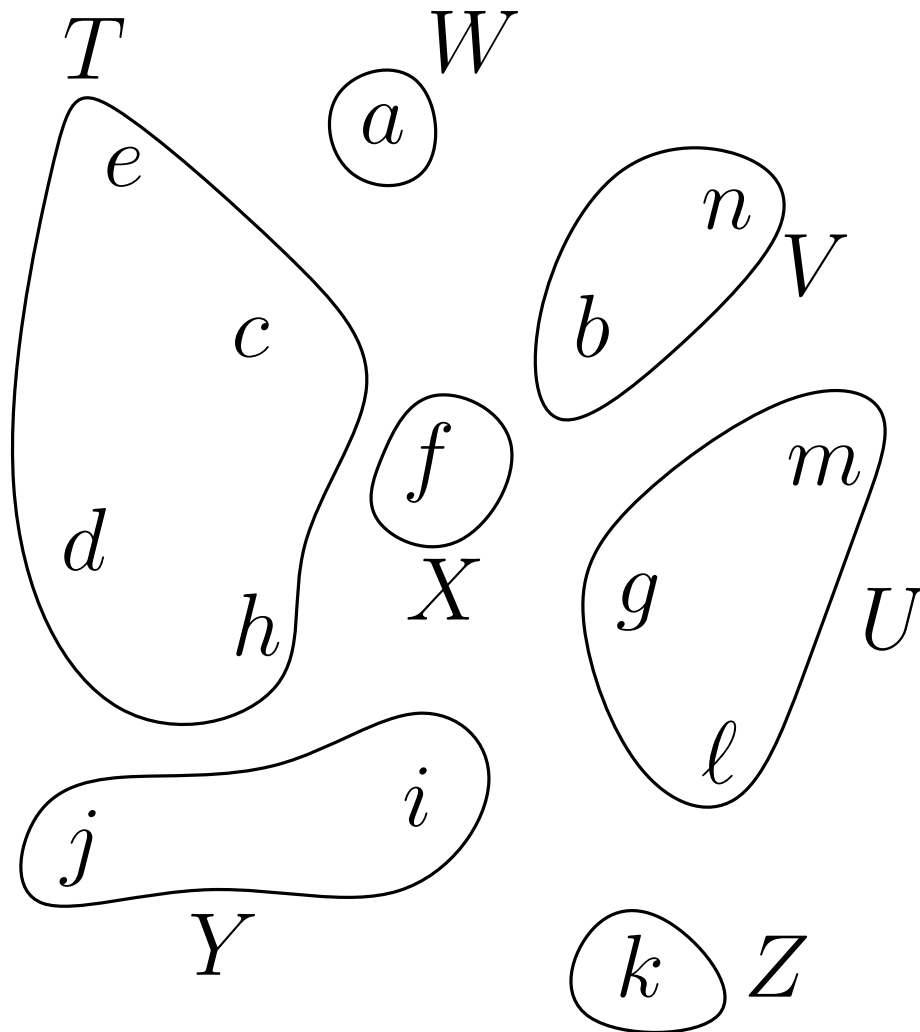
The Data

The Operations



The Data

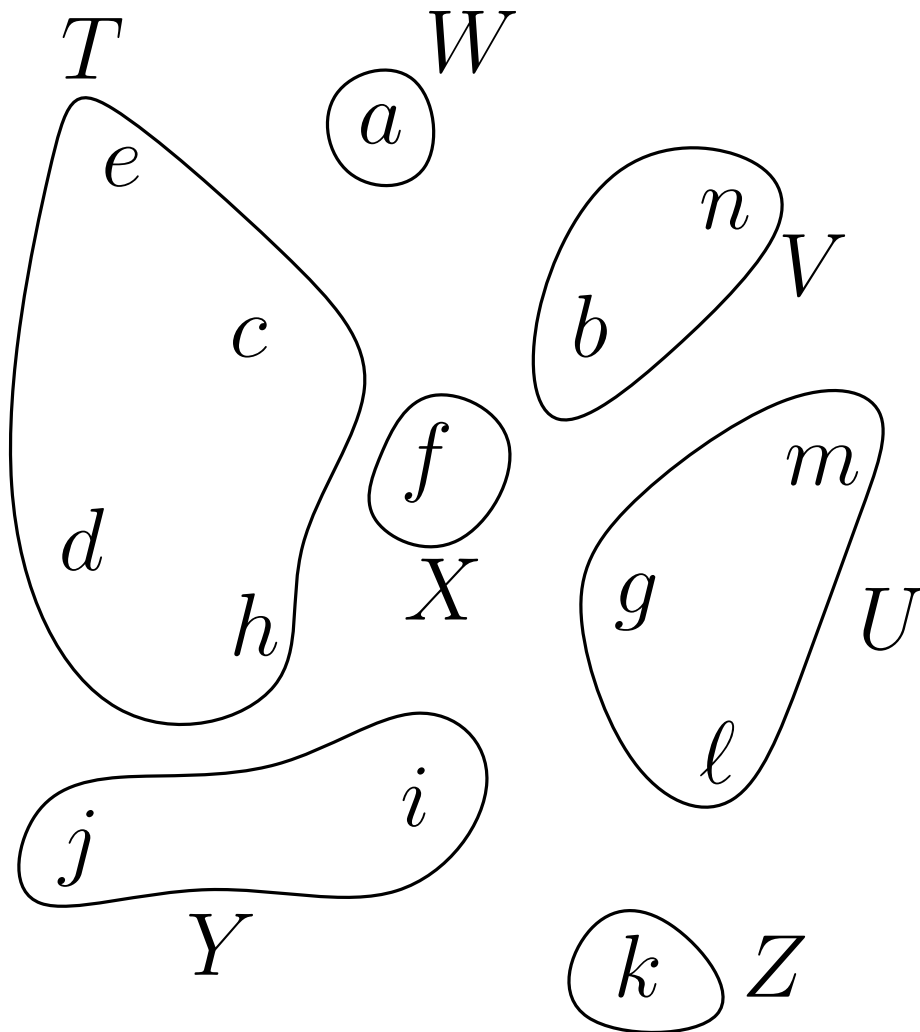
The Operations



`>>> find(h)`

The Data

The Operations



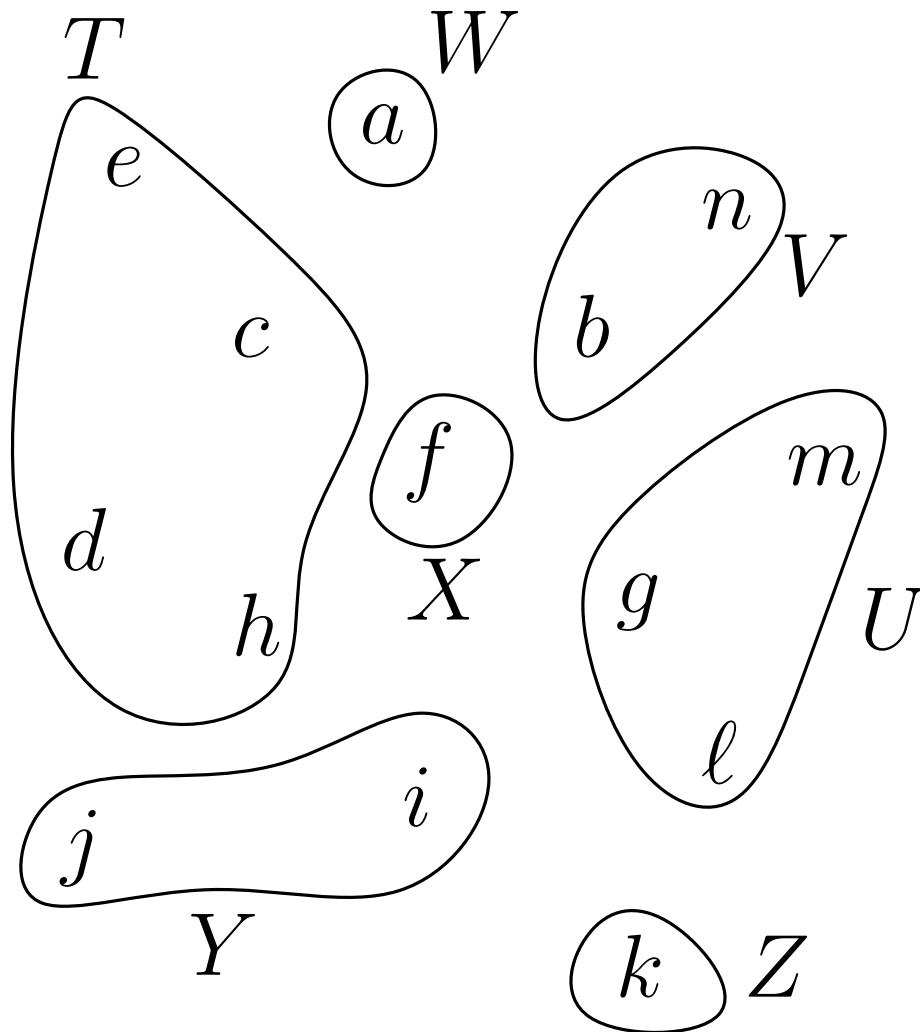
$\ggg \text{find}(h)$

T

\ggg

The Data

The Operations



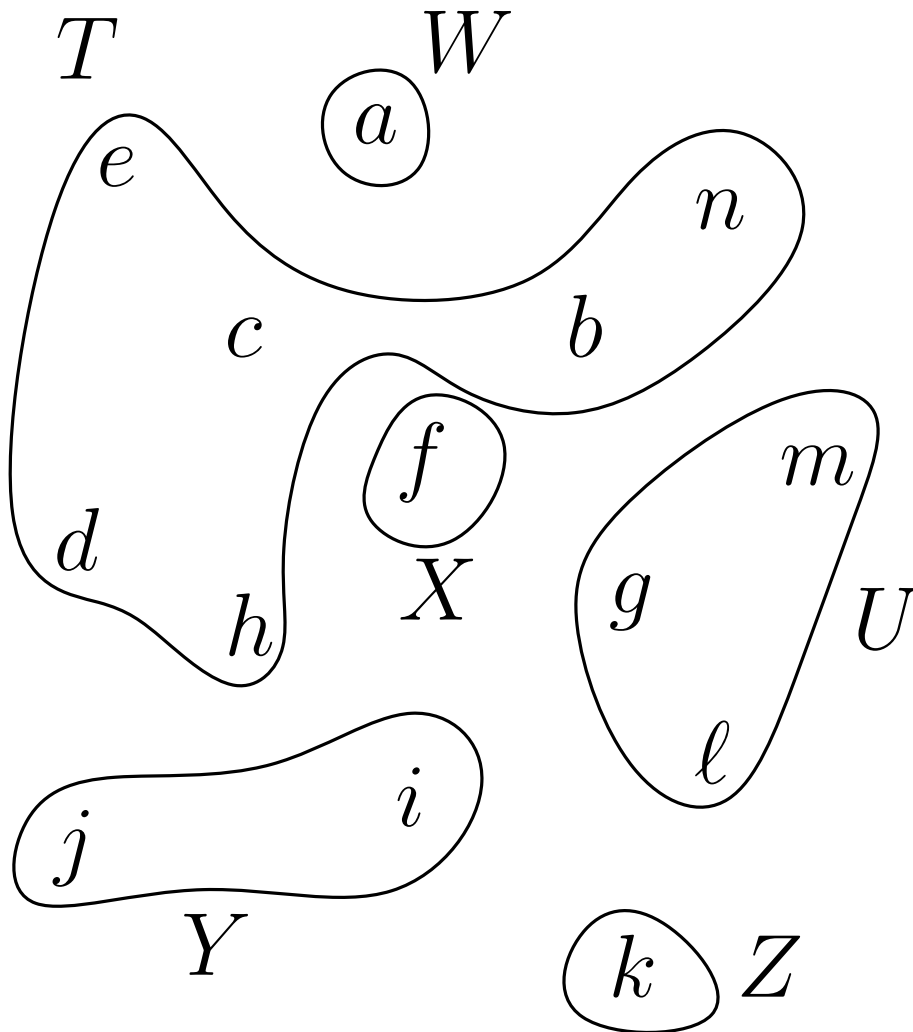
`>>> find(h)`

T

`>>> union(T, V)`

The Data

The Operations



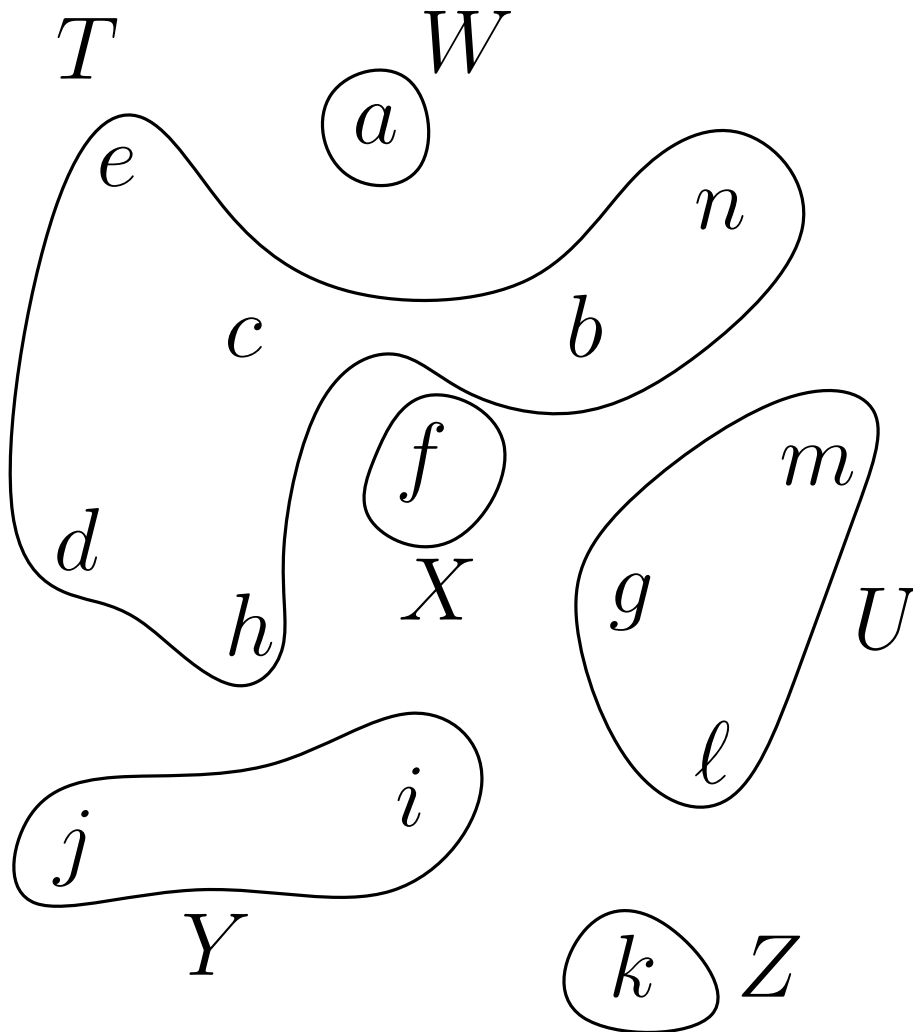
`>>> find(h)`

T

`>>> union(T, V)`

The Data

The Operations



`>>> find(h)`

T

`>>> union(T, V)`

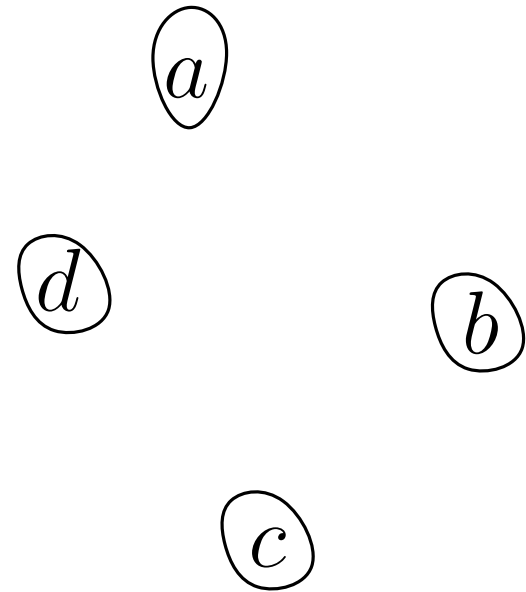
`>>>`

The Data

The Operations

```
>>> init({a, b, c, d})
```

The Data

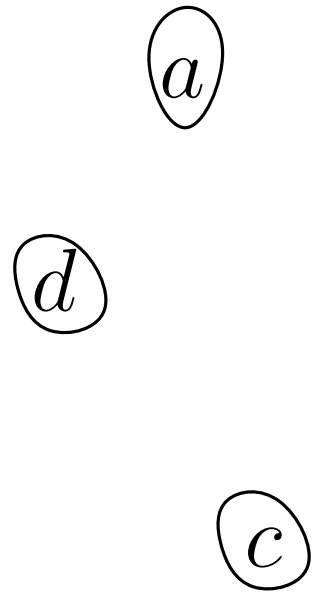


The Operations

```
>>> init({a, b, c, d})
```

```
>>>
```

The Data



The Operations

```
>>> init({a, b, c, d})
```

```
>>>
```

The names are arbitrary

The Data Structure: Union by Rank

The Data Structure: Union by Rank

e *a*
 n
c *b*
d *f* *m*
 h *g*
j *i* *l*
 k

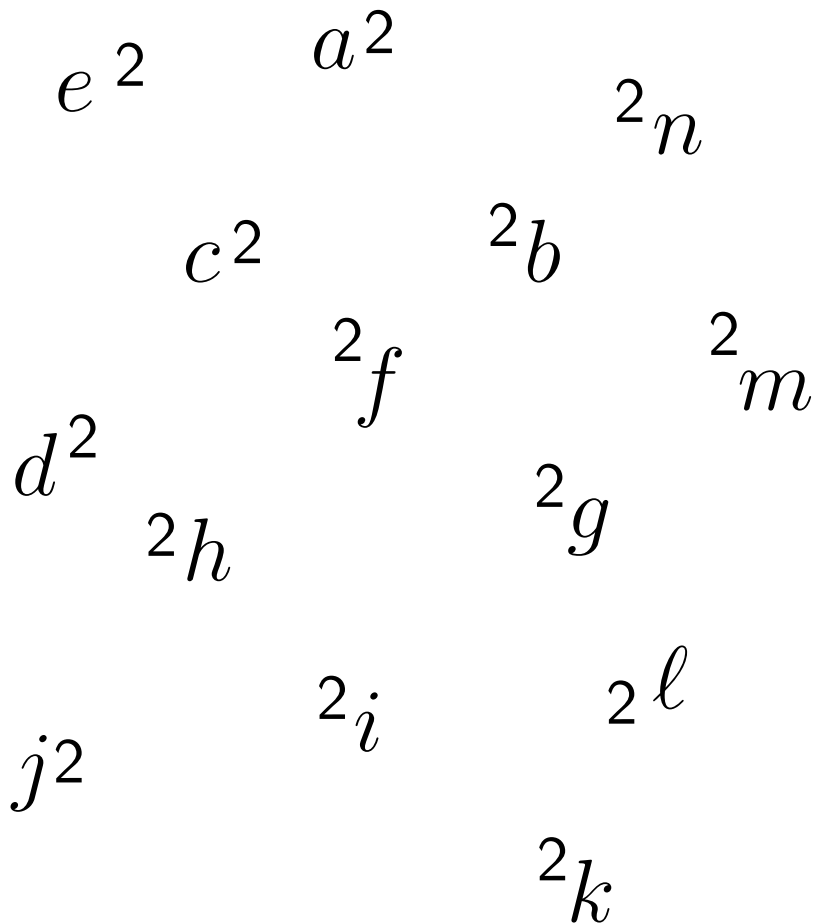
The Data Structure: Union by Rank

Every element has a *rank*.
It is initialized to 2 at the beginning.

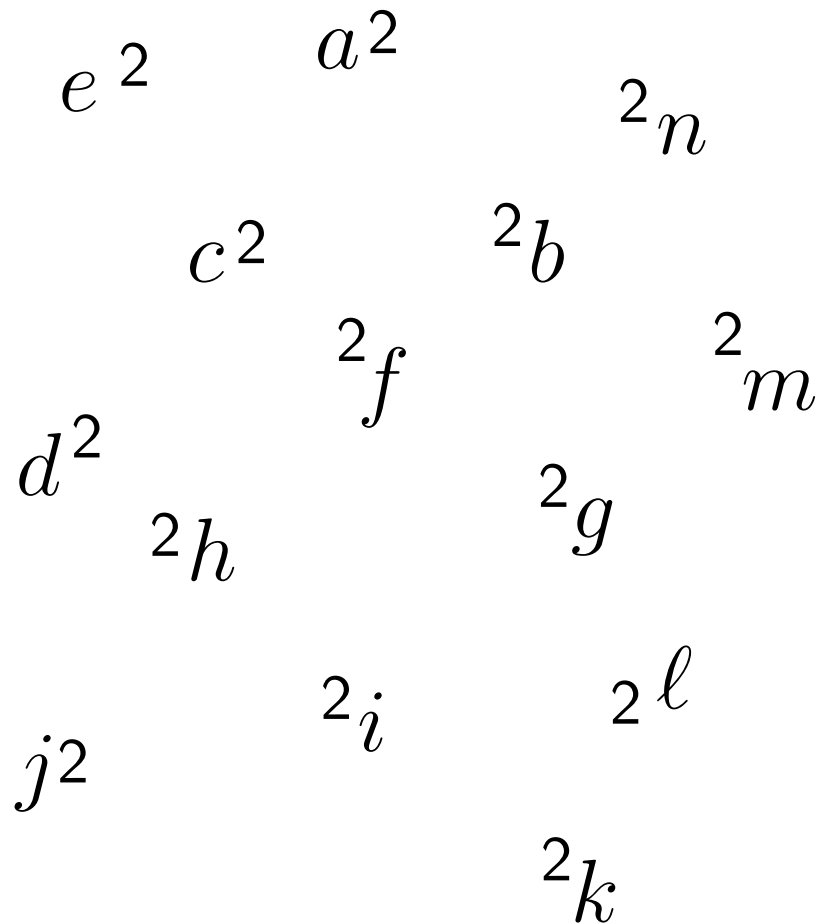
e *a*
 n
c *b*
d *f* *m*
 h *g*
j *i* *l*
 k

The Data Structure: Union by Rank

Every element has a *rank*.
It is initialized to 2 at the beginning.

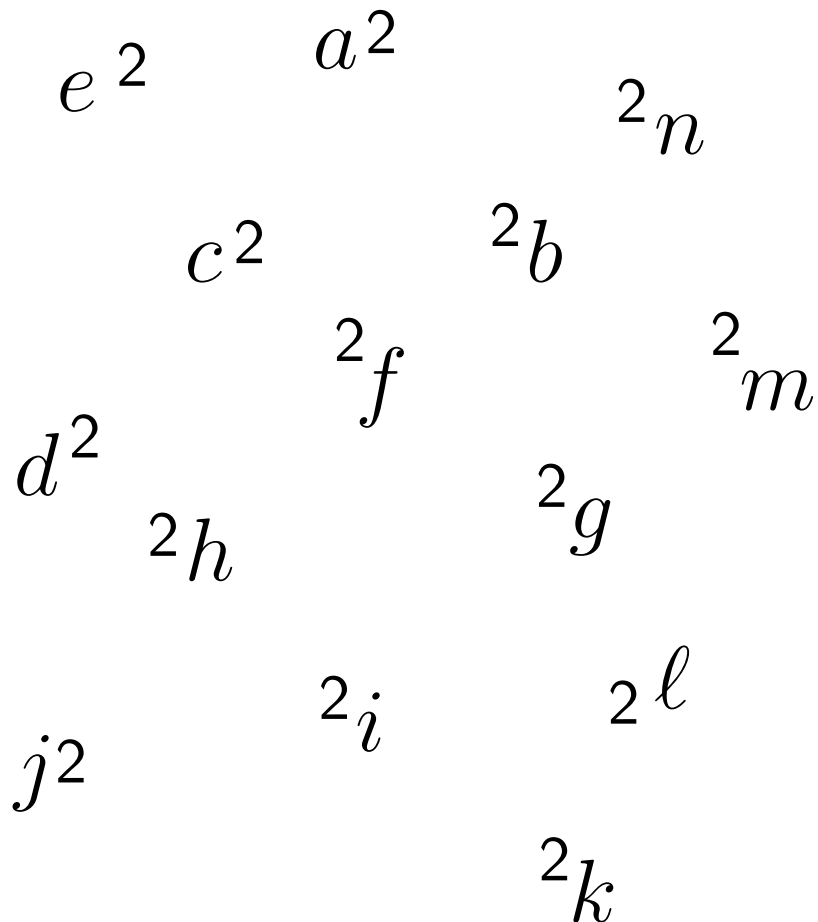


The Data Structure: Union by Rank



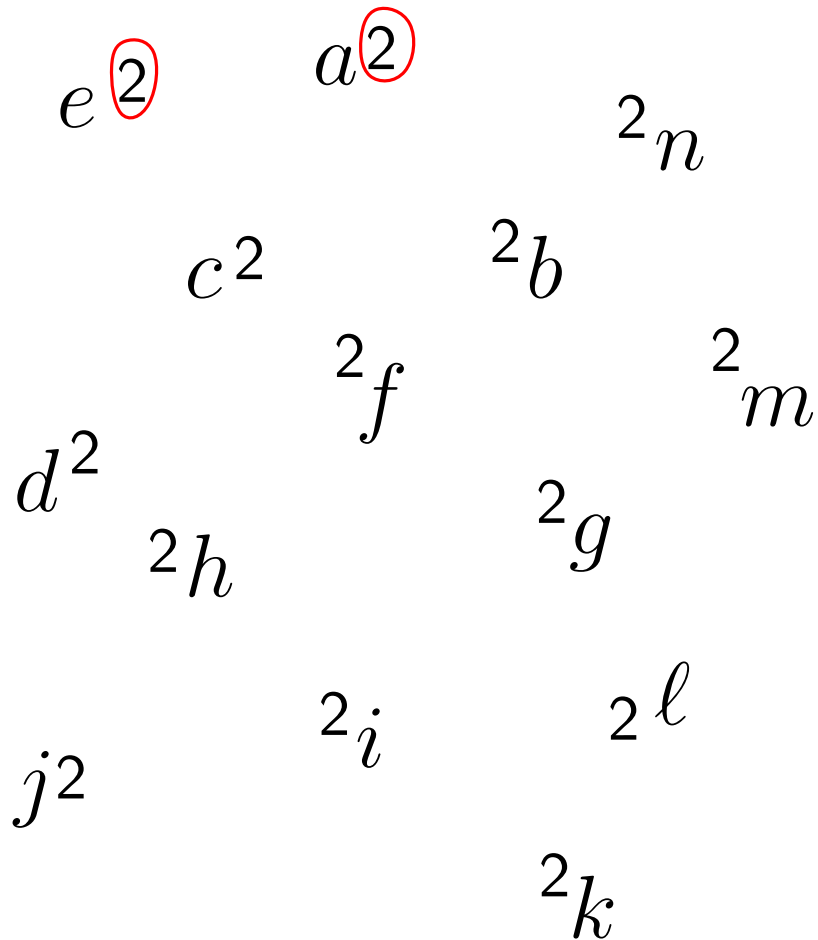
The Data Structure: Union by Rank

>>> union(*e*, *a*)

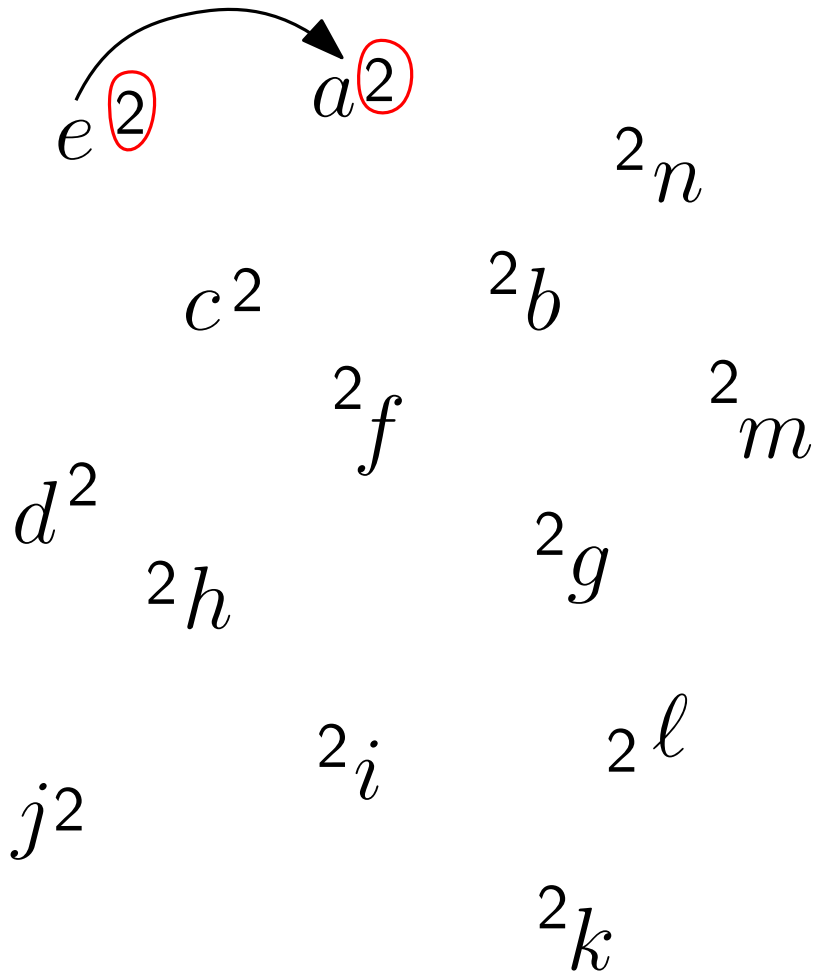


The Data Structure: Union by Rank

```
>>> union(e, a)  
if rank(e) = rank(a):
```



The Data Structure: Union by Rank

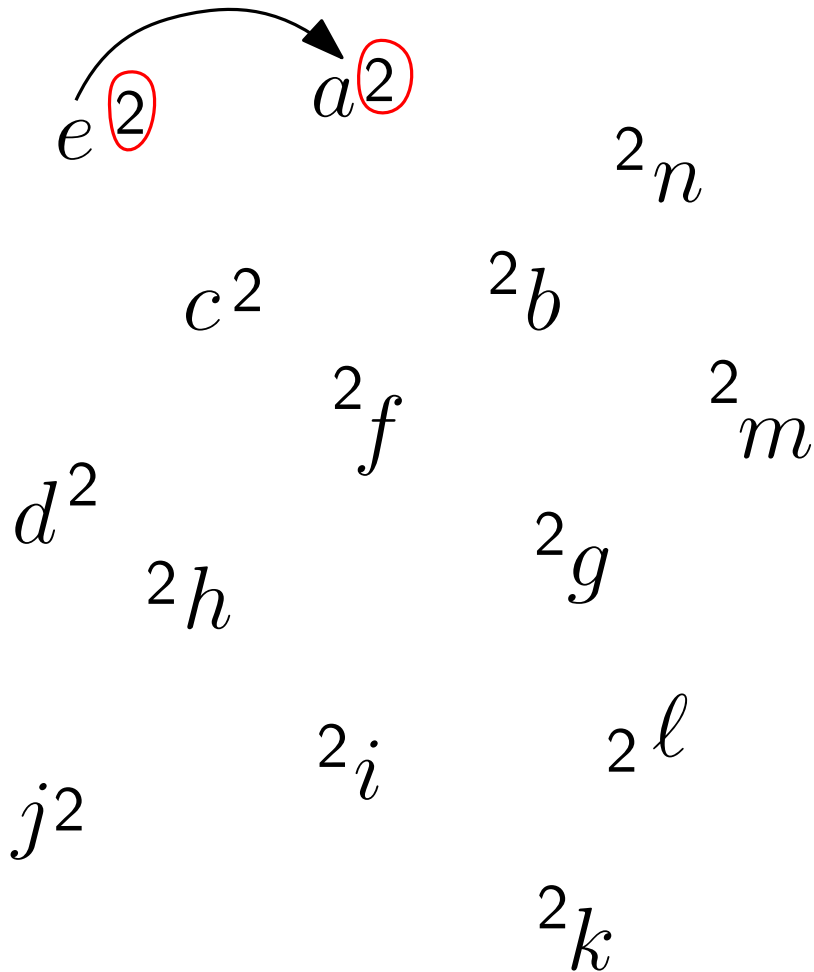


>>> union(e , a)

if rank(e) = rank(a):

edge from e to a

The Data Structure: Union by Rank



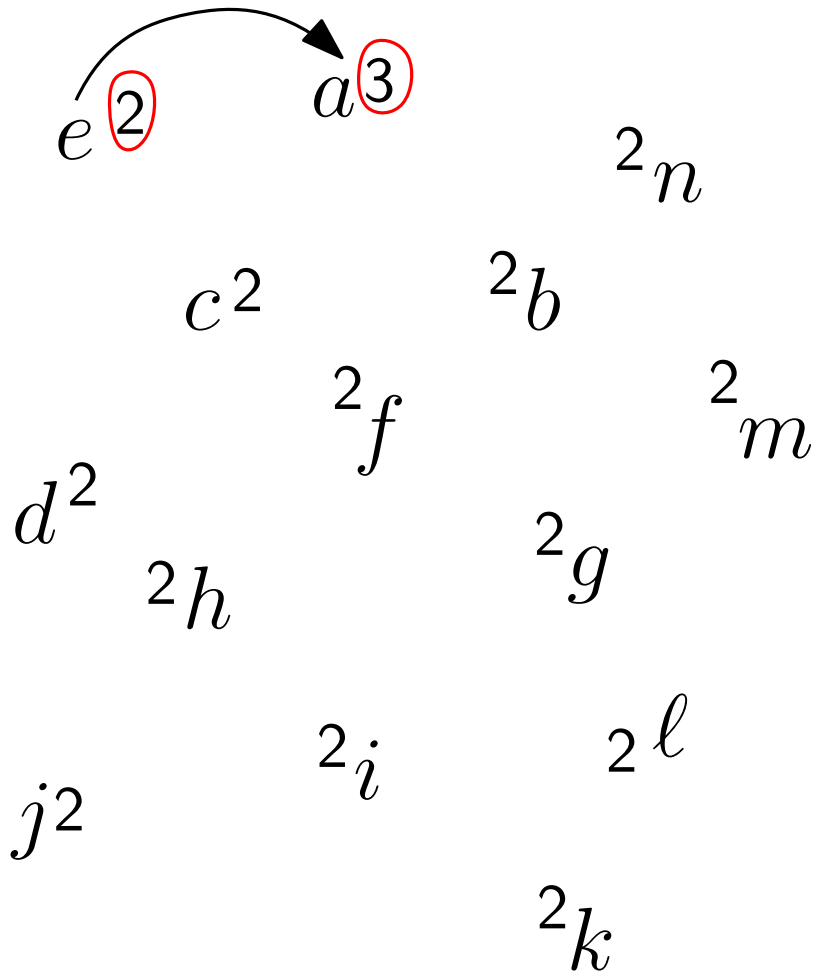
`>>> union(e, a)`

if $\text{rank}(e) = \text{rank}(a)$:

edge from e to a

increase rank of a

The Data Structure: Union by Rank



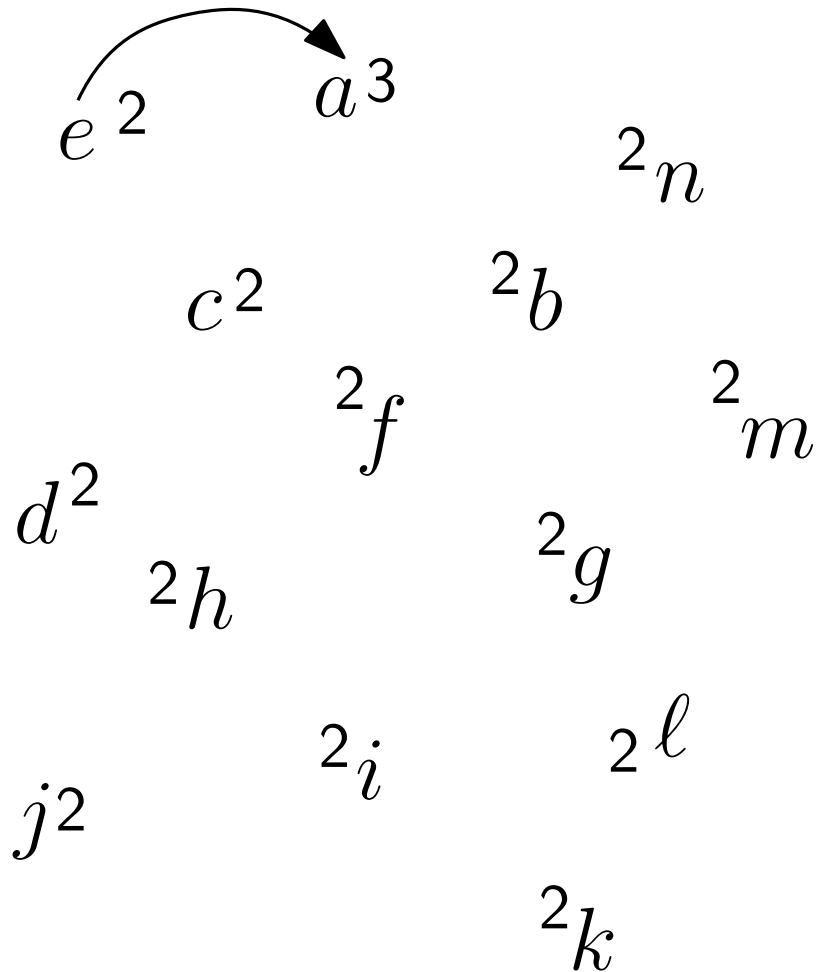
>>> union(e, a)

if rank(e) = rank(a):

edge from e to a

increase rank of a

The Data Structure: Union by Rank



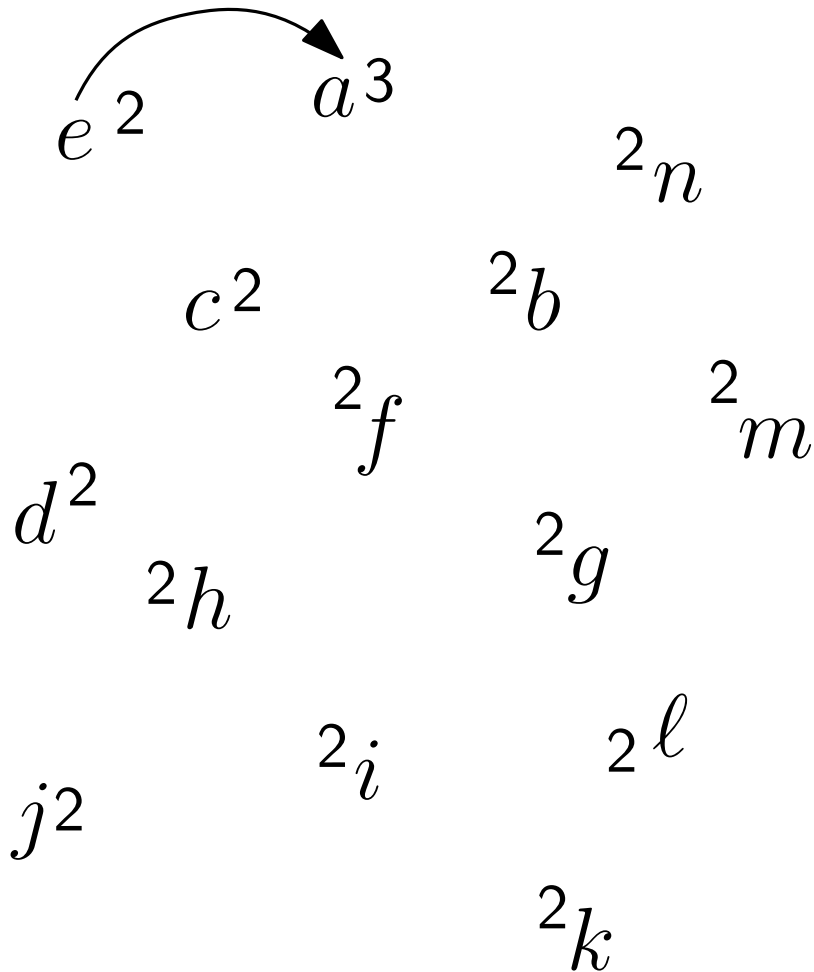
>>> union(e, a)

if rank(e) = rank(a):

edge from e to a

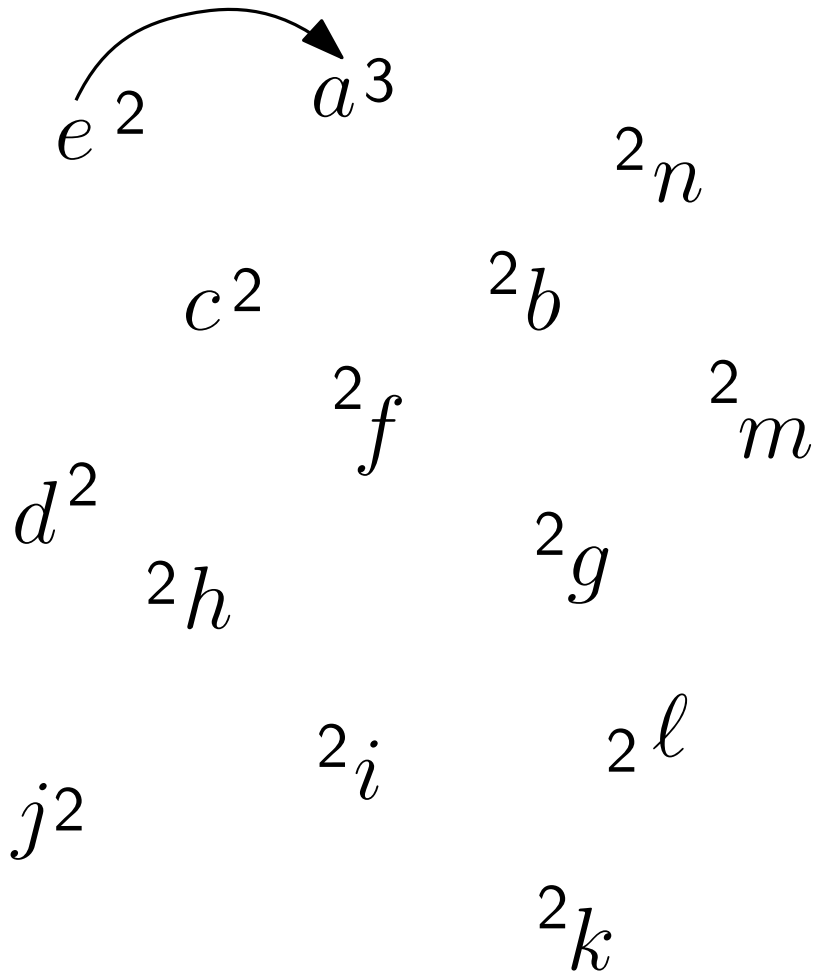
increase rank of a

The Data Structure: Union by Rank



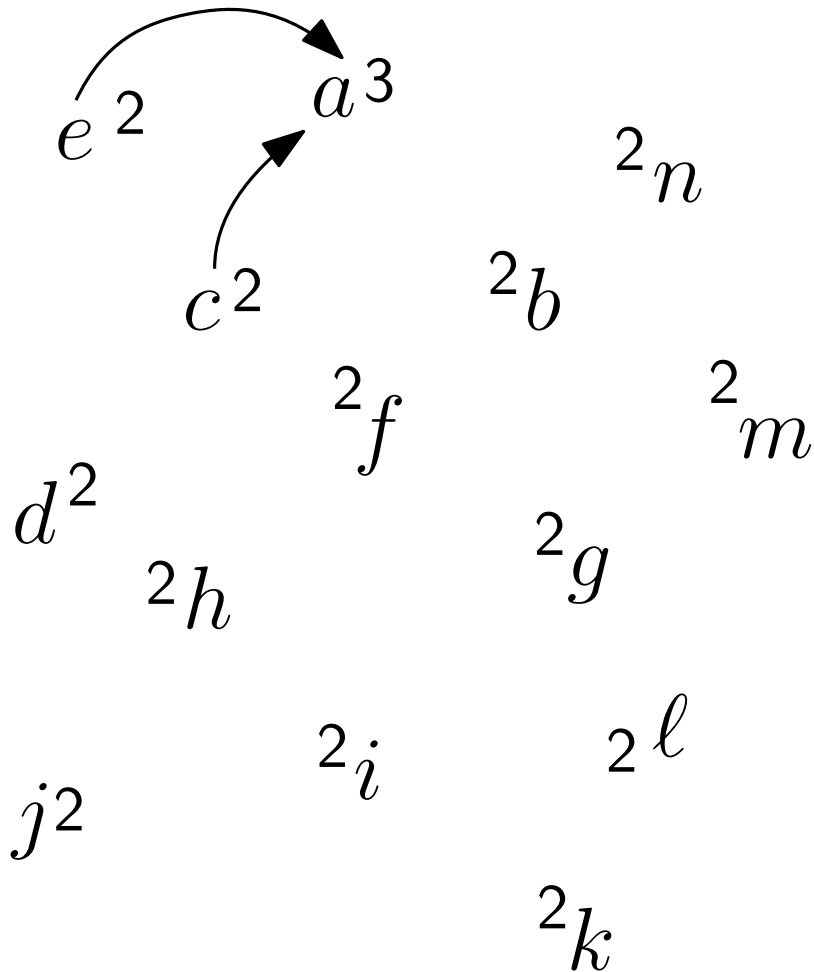
```
>>> union(e, a)
if rank(e) = rank(a):
    edge from e to a
    increase rank of a
>>> union(a, c)
```

The Data Structure: Union by Rank



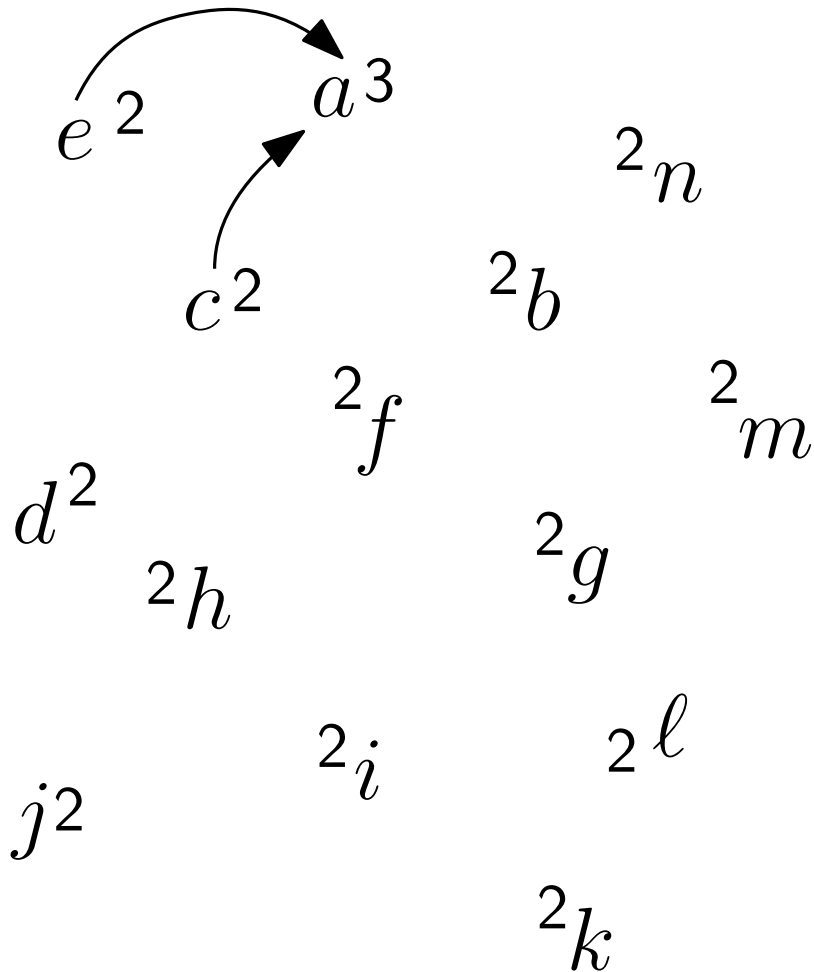
```
>>> union(e, a)
if rank(e) = rank(a):
    edge from e to a
    increase rank of a
>>> union(a, c)
if rank(e) ≠ rank(a):
```

The Data Structure: Union by Rank



```
>>> union(e, a)
if rank(e) = rank(a):
    edge from e to a
    increase rank of a
>>> union(a, c)
if rank(e) ≠ rank(a):
    smaller to larger
```

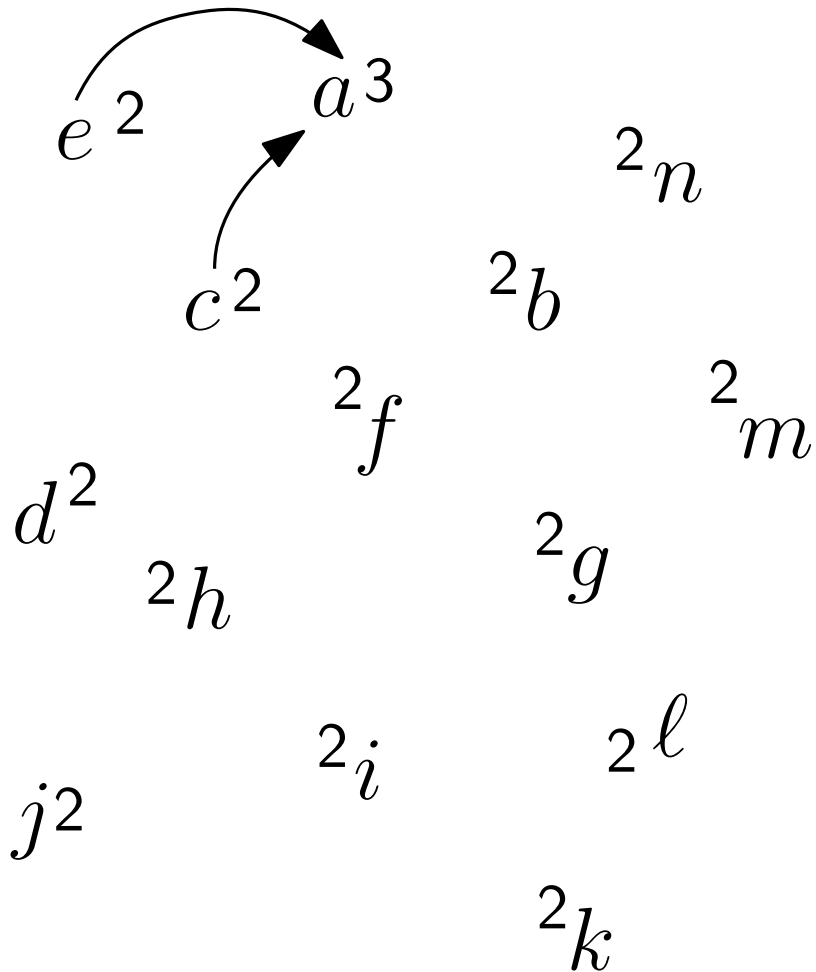
The Data Structure: Union by Rank



```
>>> union(e, a)
if rank(e) = rank(a):
    edge from e to a
    increase rank of a
>>> union(a, c)
if rank(e) ≠ rank(a):
    smaller to larger
    don't change rank
```

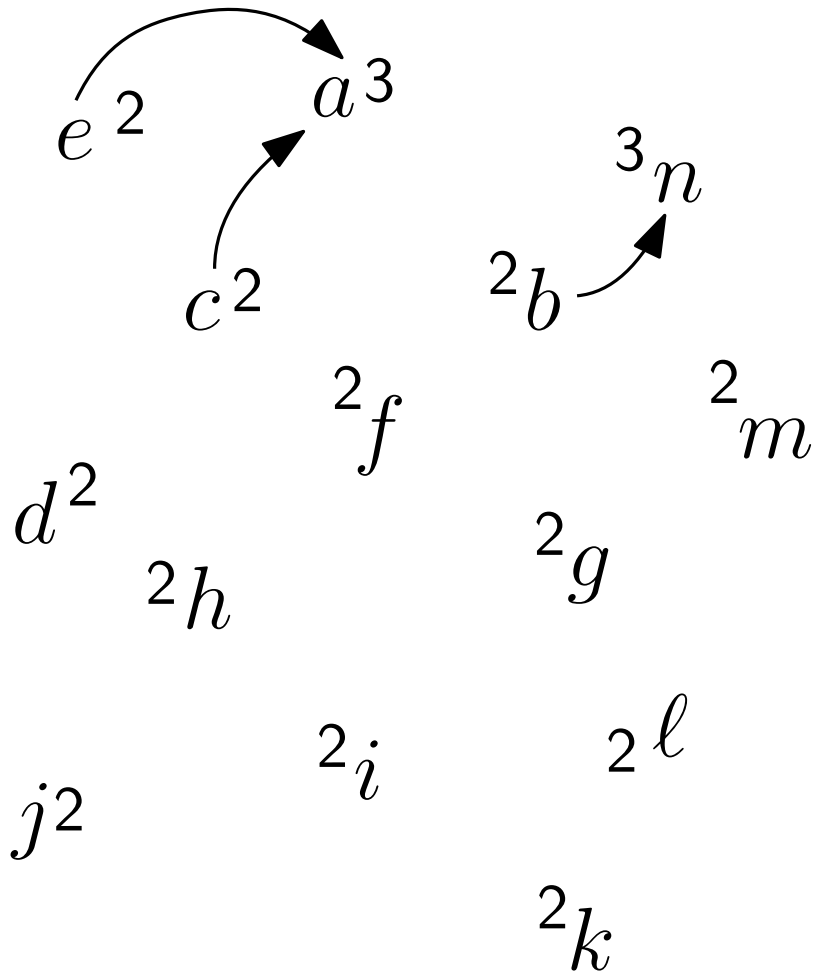
The Data Structure: Union by Rank

>>> union(*b*, *n*)

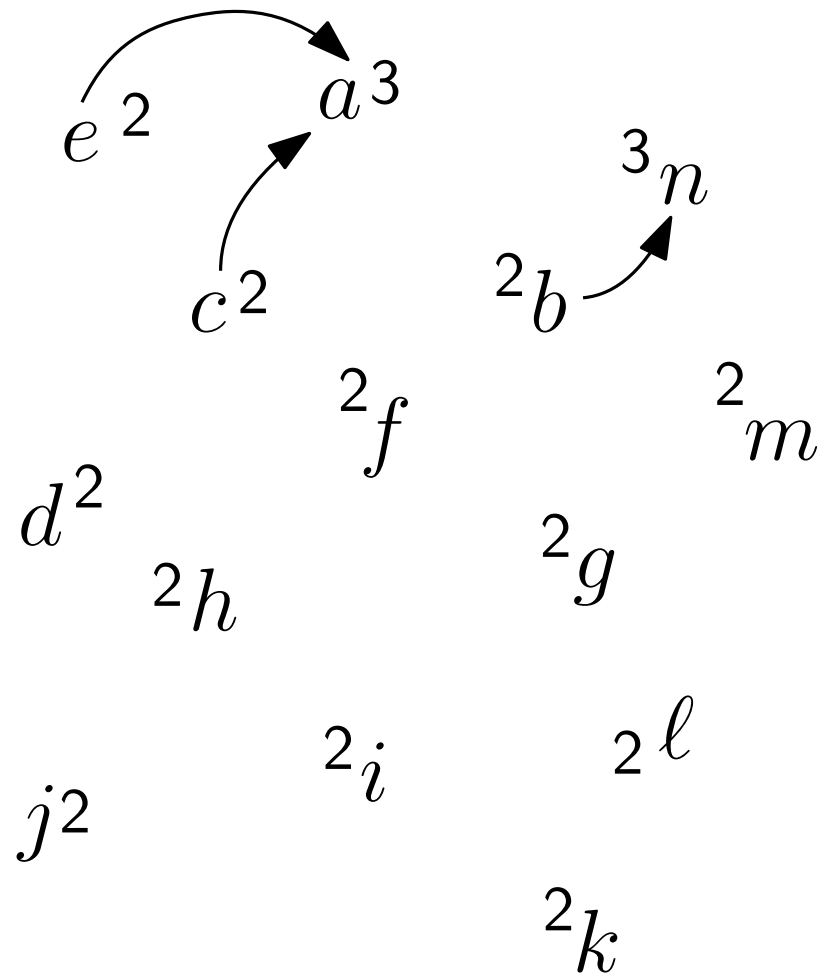


The Data Structure: Union by Rank

>>> union(*b*, *n*)



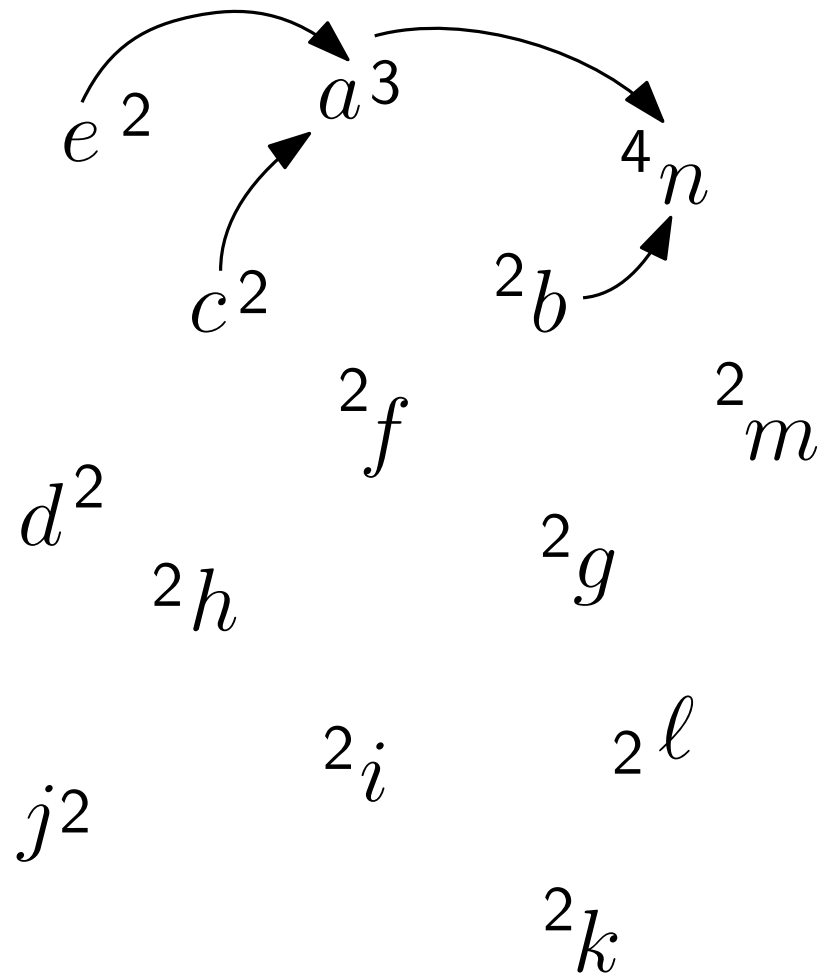
The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

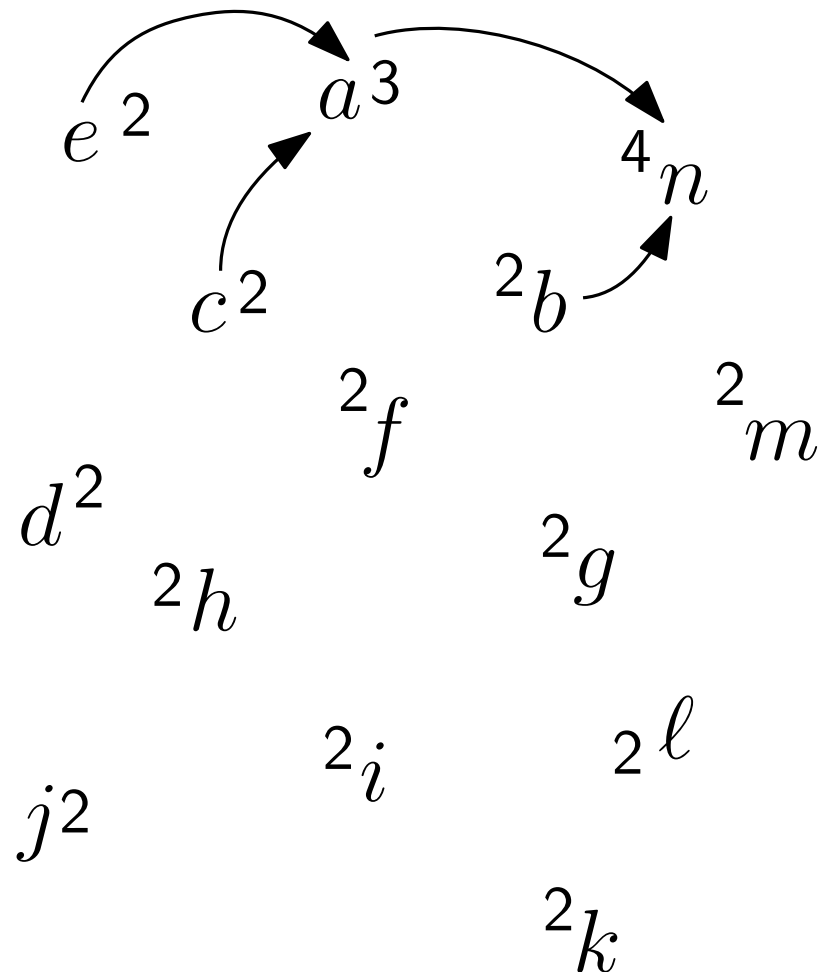

The Data Structure: Union by Rank



>>> union(b, n)

>>> union(a, n)

The Data Structure: Union by Rank

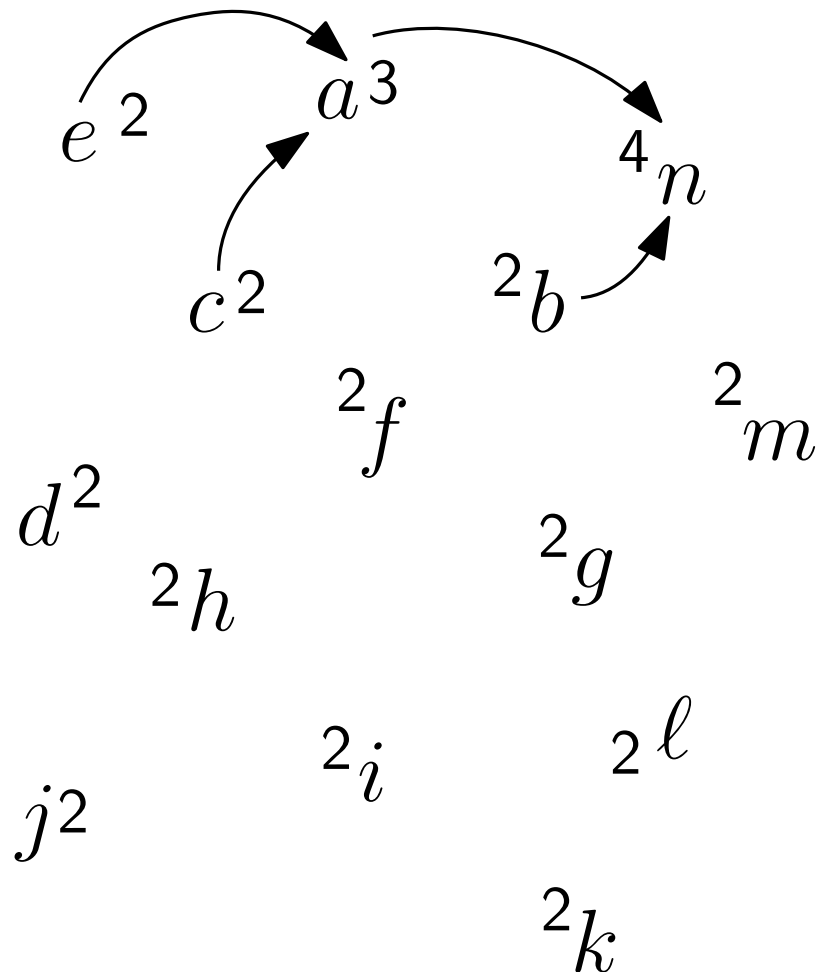


```
>>> union(b, n)
```

```
>>> union(a, n)
```

cost: $O(1)$

The Data Structure: Union by Rank



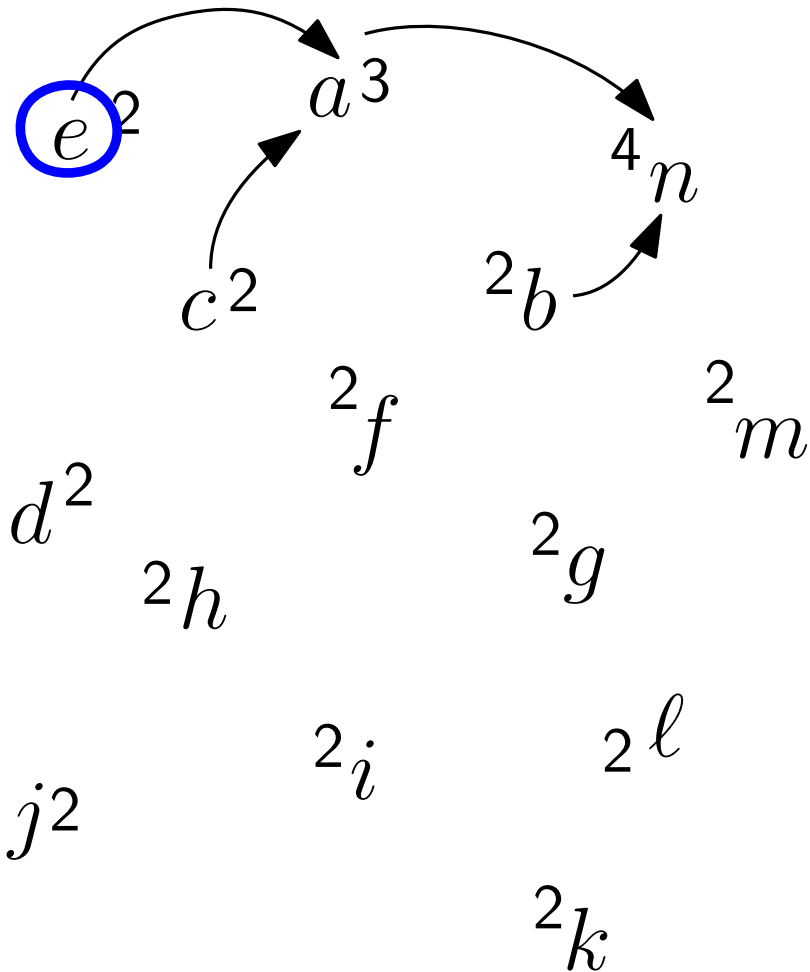
```
>>> union(b, n)
```

```
>>> union(a, n)
```

cost: $O(1)$

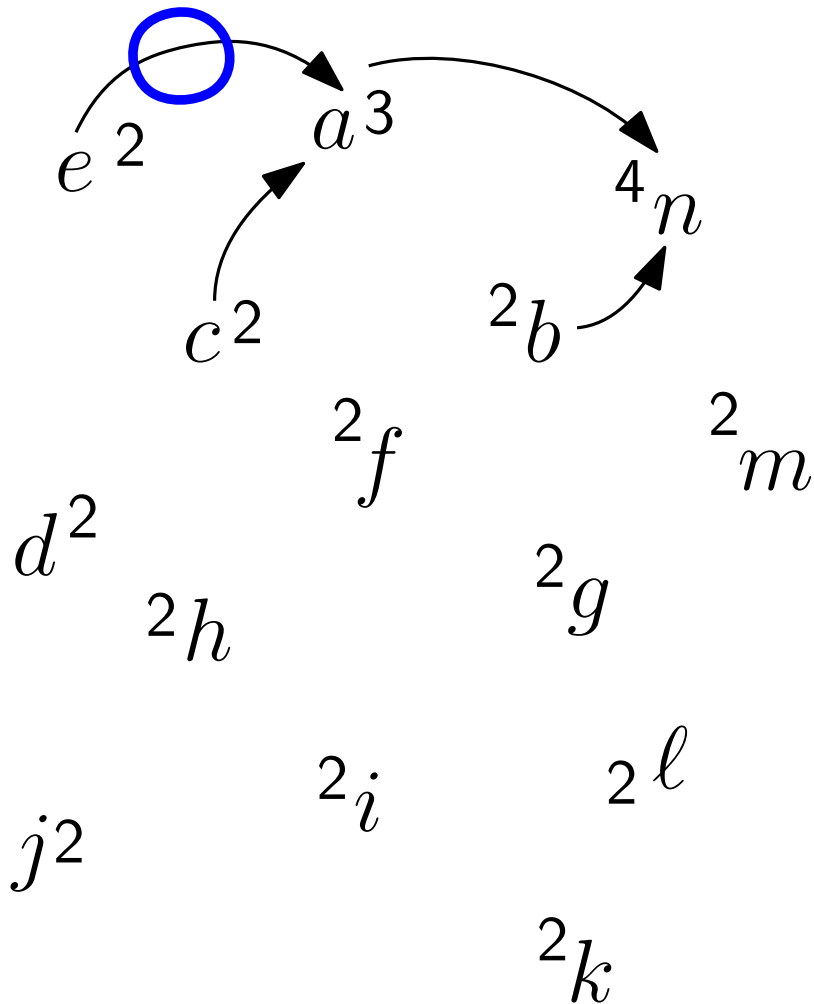
```
>>> find(e)
```

The Data Structure: Union by Rank



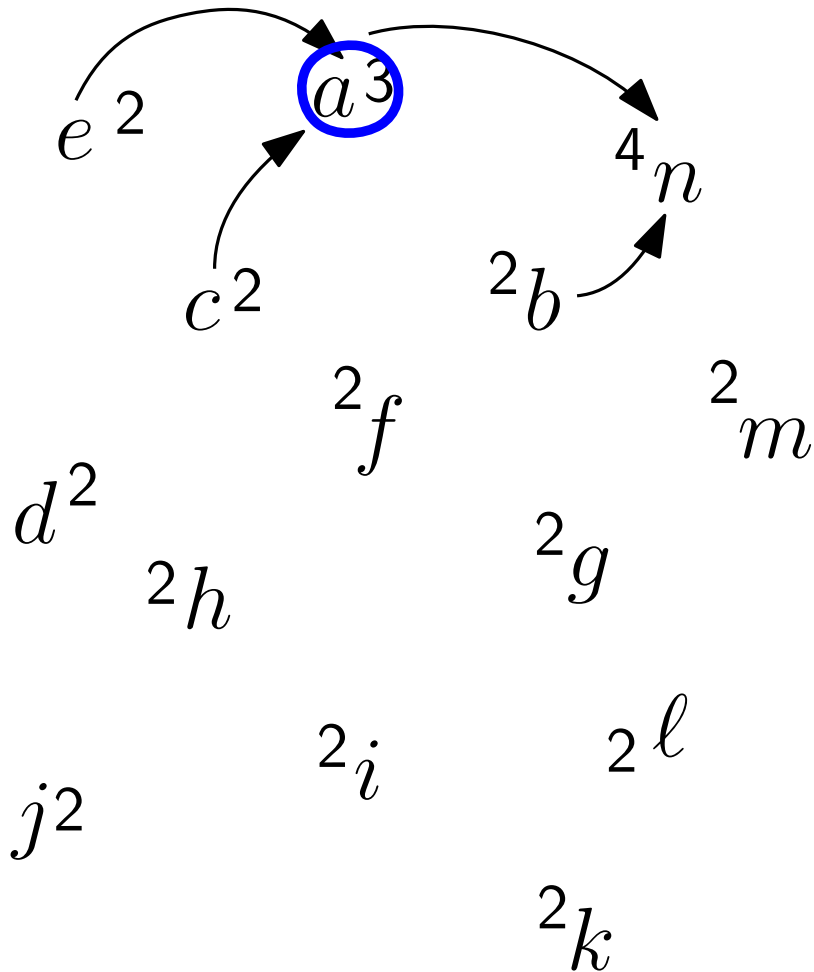
```
>>> union(b, n)
>>> union(a, n)
                        cost:  $O(1)$ 
>>> find(e)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
>>> union(a, n)
cost: O(1)
>>> find(e)
```

The Data Structure: Union by Rank



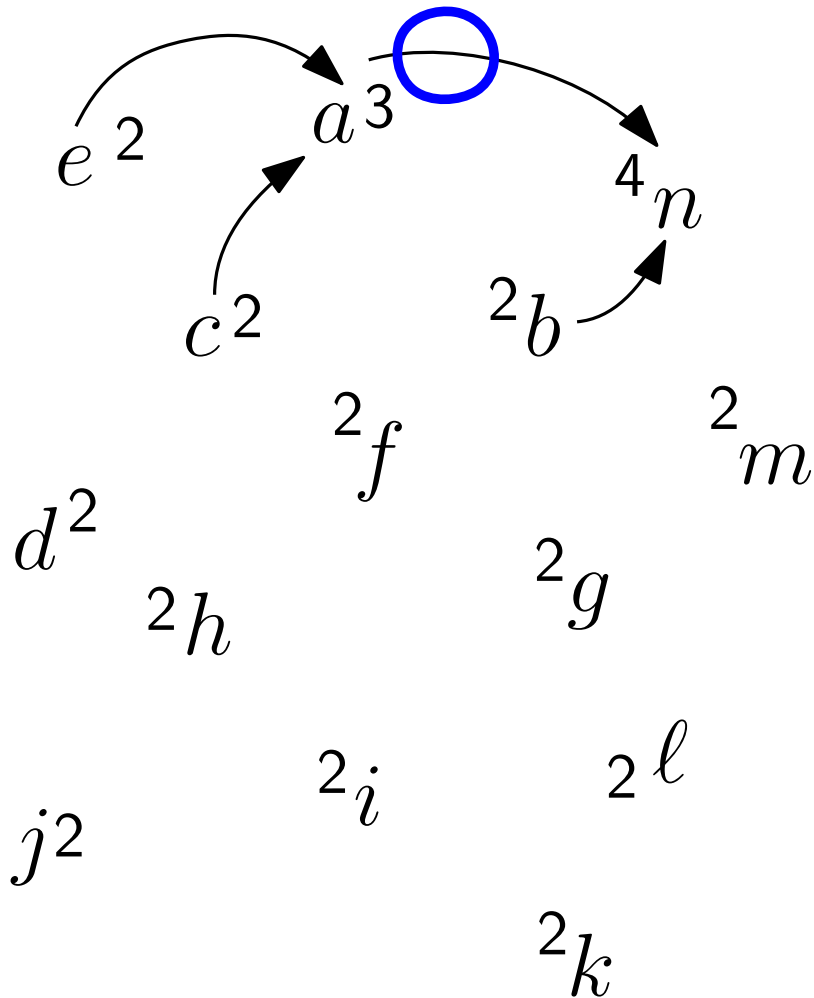
```
>>> union(b, n)
```

```
>>> union(a, n)
```

cost: $O(1)$

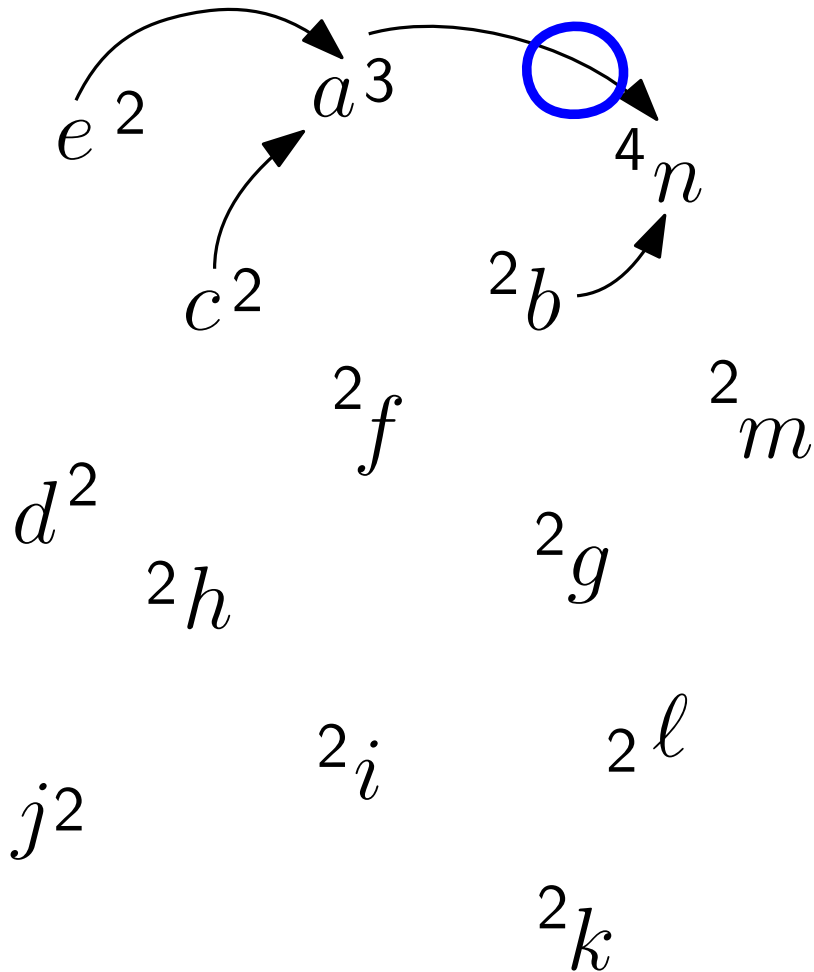
```
>>> find(e)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
>>> union(a, n)
                                cost:  $O(1)$ 
>>> find(e)
```

The Data Structure: Union by Rank



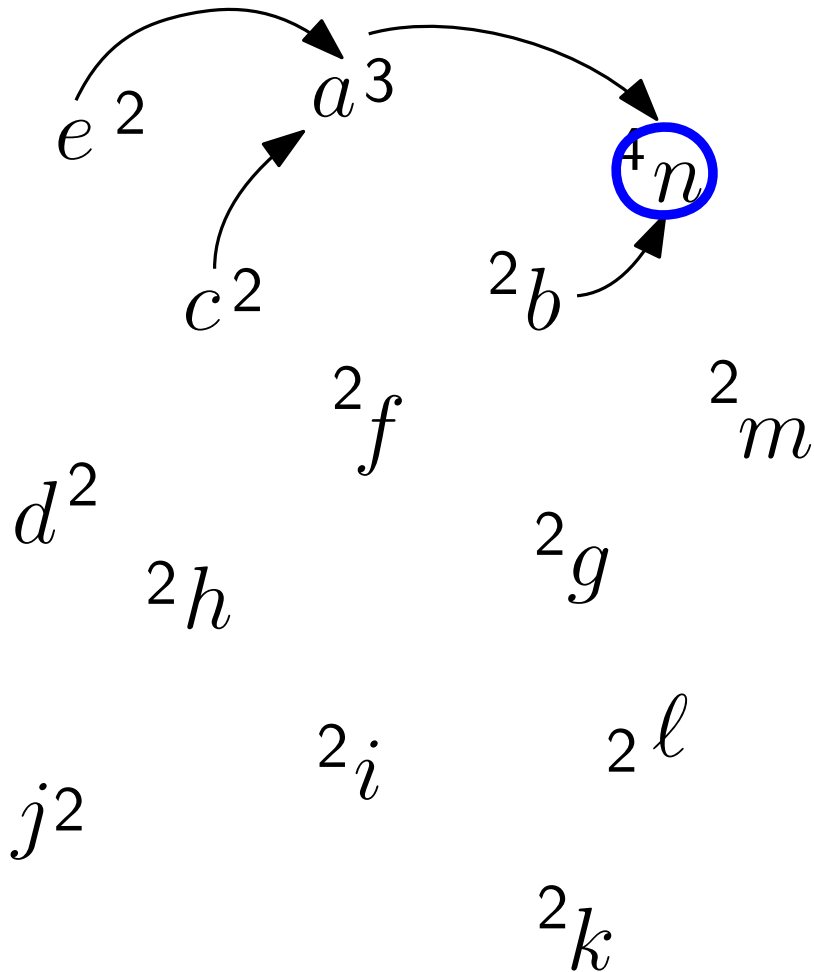
```
>>> union(b, n)
```

```
>>> union(a, n)
```

cost: $O(1)$

```
>>> find(e)
```


The Data Structure: Union by Rank



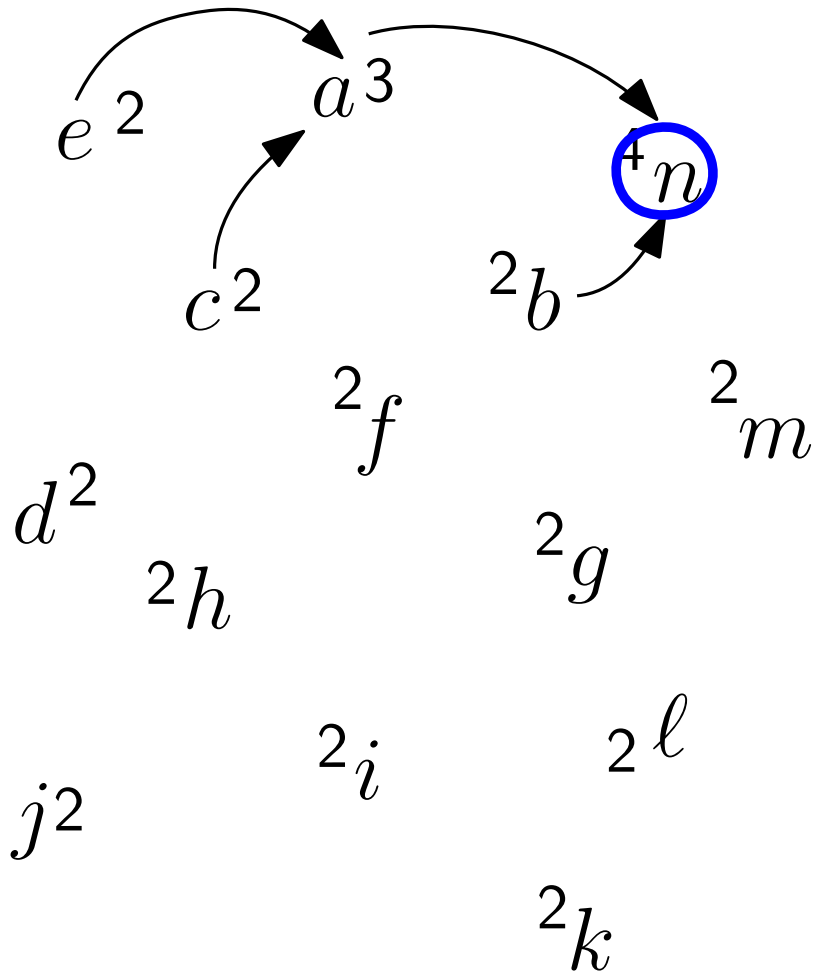
```
>>> union(b, n)
```

```
>>> union(a, n)
```

cost: $O(1)$

```
>>> find(e)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

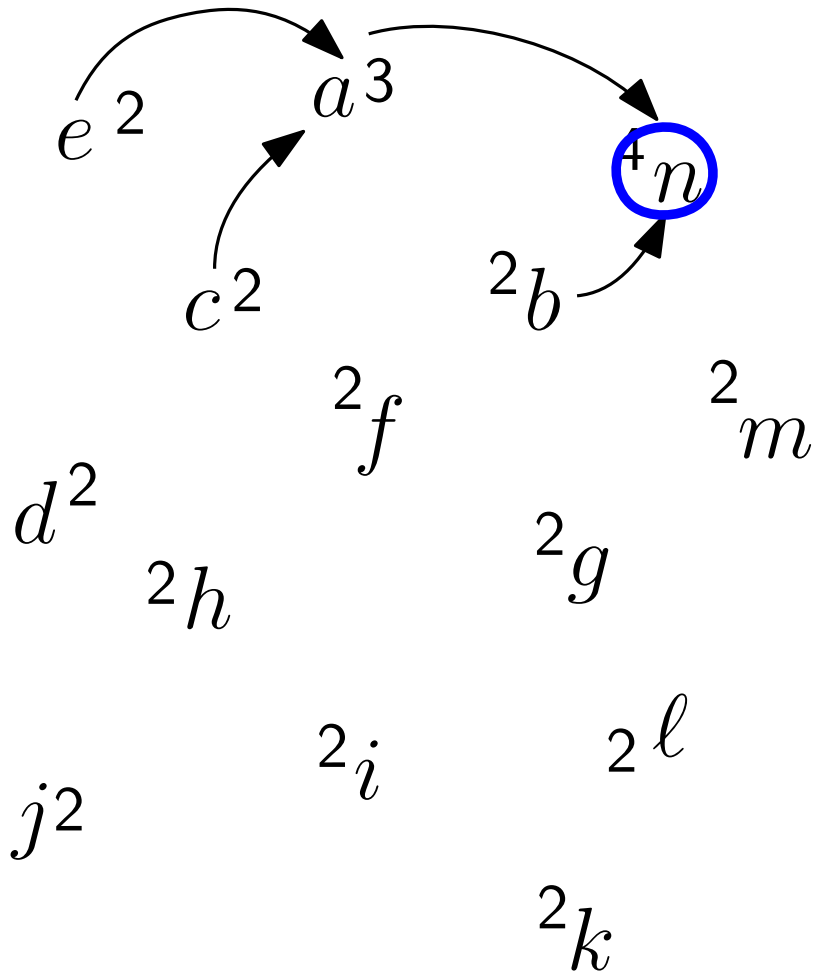
```
>>> union(a, n)
```

cost: $O(1)$

```
>>> find(e)
```

```
n
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

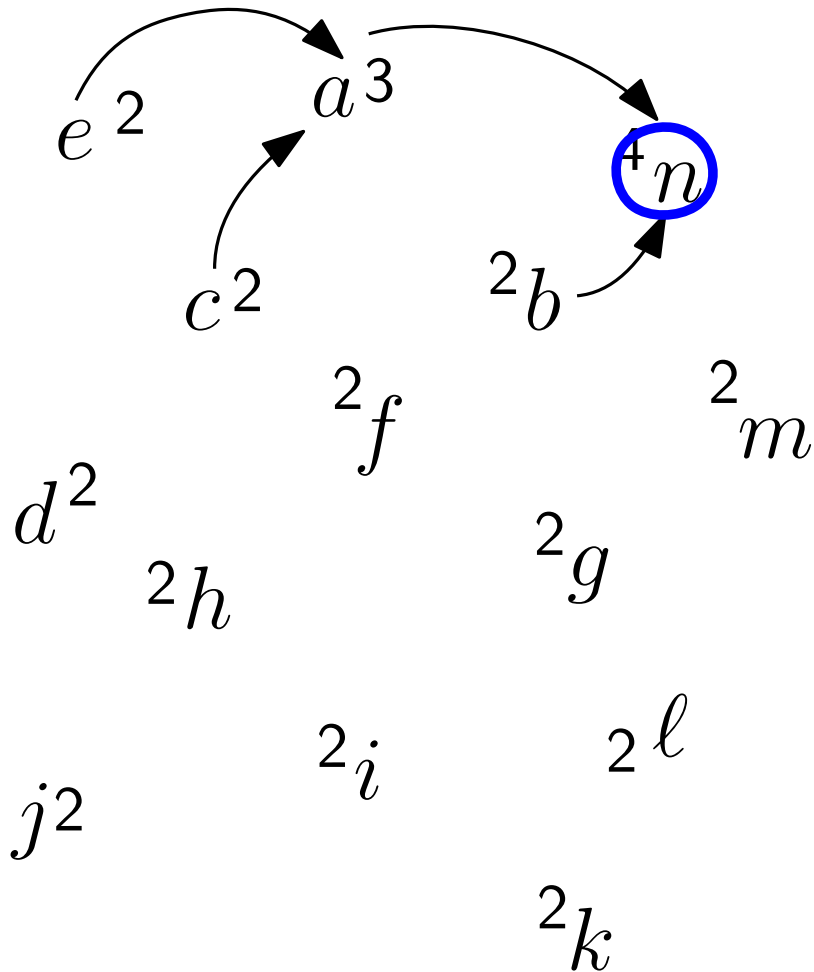
```
>>> union(a, n)
```

cost: $O(1)$

```
>>> find(e)
```

n
cost: $O(\text{length of path})$

The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

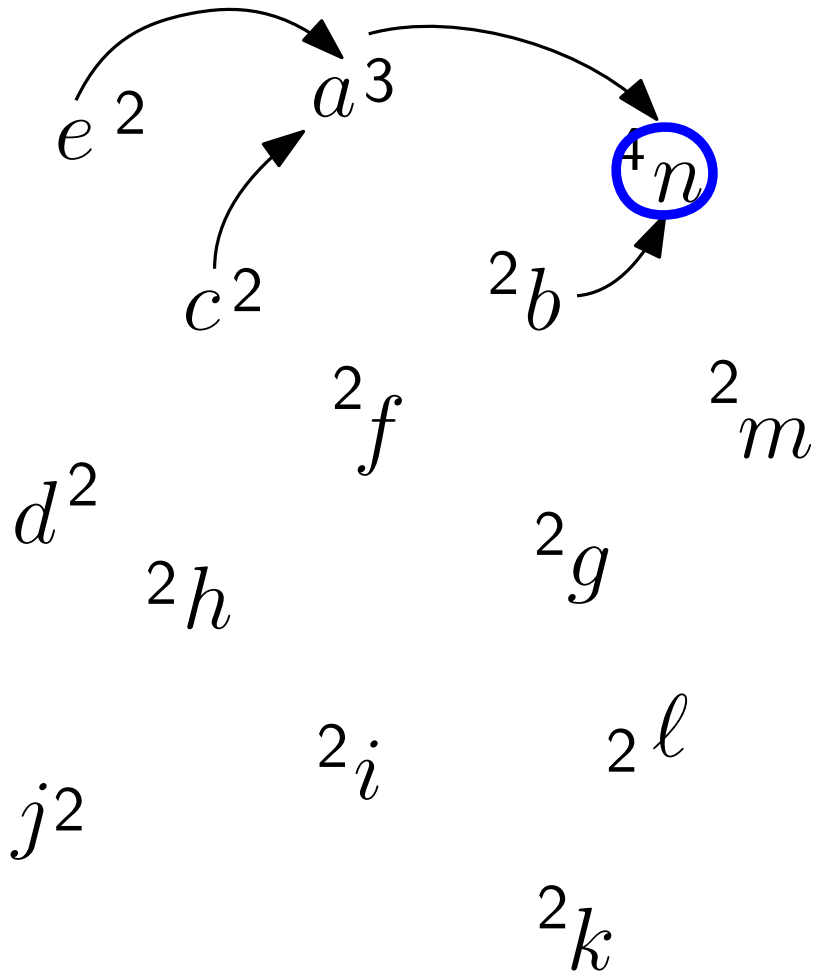
cost: $O(1)$

```
>>> find(e)
```

n

cost: $O(\text{length of path})$
 $\leq O(\text{height of tree})$

The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

cost: $O(1)$

```
>>> find(e)
```

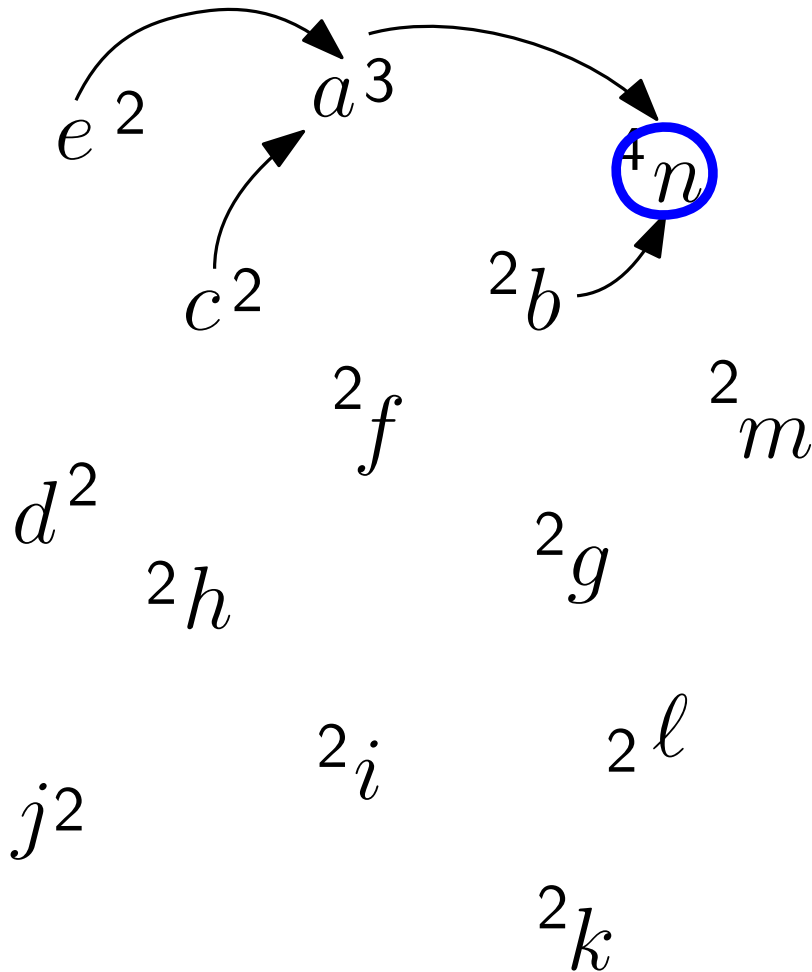
n

cost: $O(\text{length of path})$

$\leq O(\text{height of tree})$

$\leq O(\text{rank}(n))$

The Data Structure: Union by Rank



>>> union(b, n)

>>> union(a, n)

cost: $O(1)$

>>> find(e)

n

cost: $O(\text{length of path})$

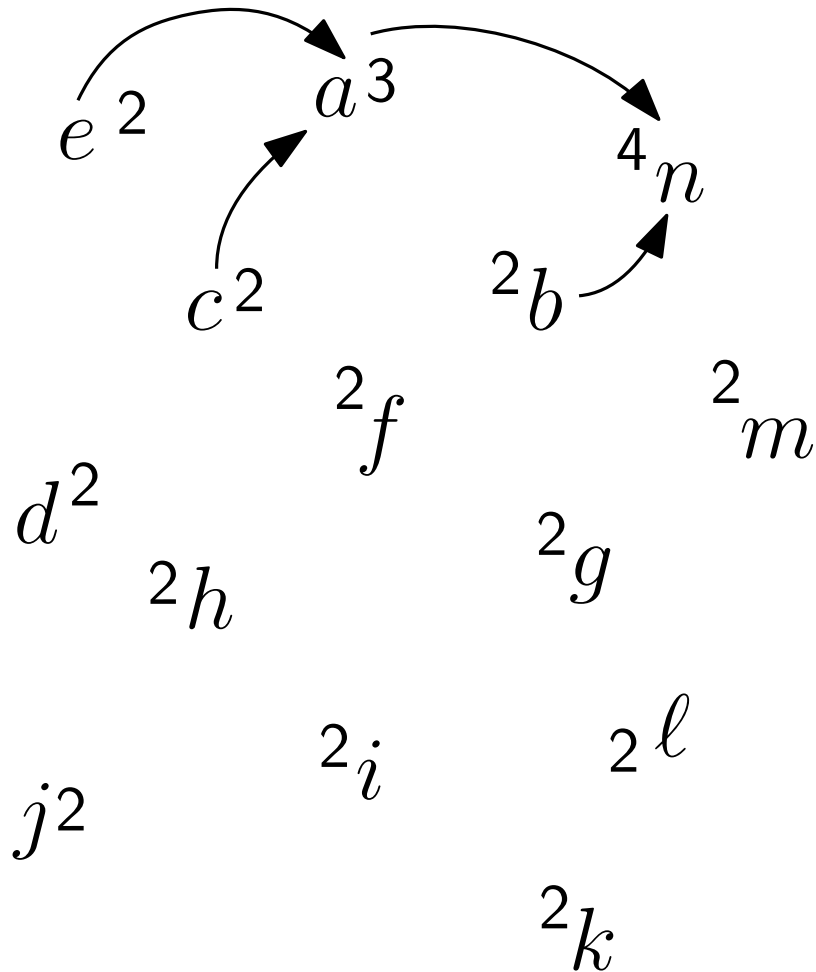
$\leq O(\text{height of tree})$

$\leq O(\text{rank}(n))$

because

Lemma: The height of the subtree rooted at x is $\text{rank}(x) - 2$.

The Data Structure: Union by Rank



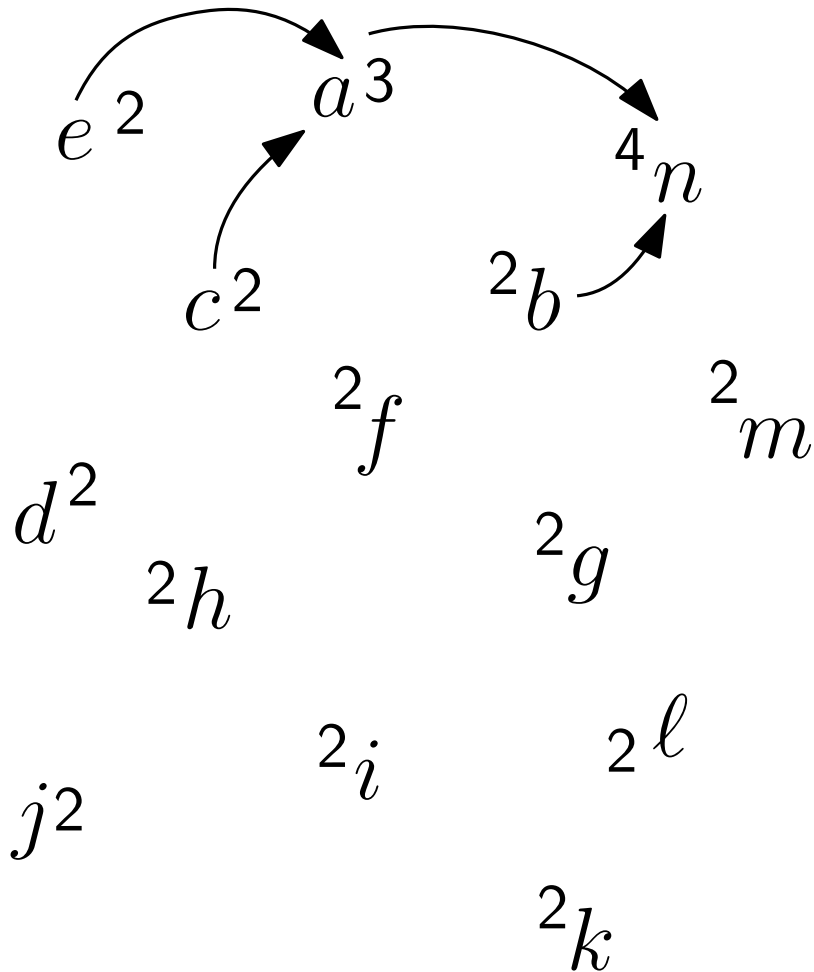
```
>>> union(b, n)
```

```
>>> union(a, n)
```

```
>>> find(e)
```

```
n
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

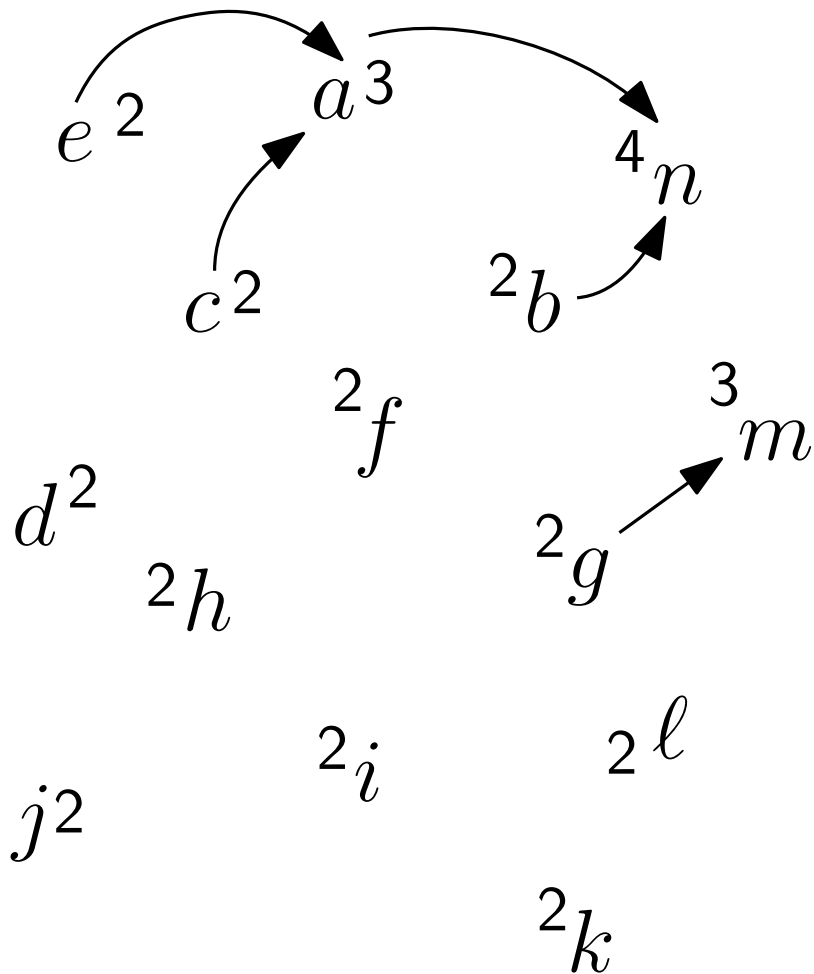
```
>>> union(a, n)
```

```
>>> find(e)
```

```
n
```

```
>>> union(g, m)
```


The Data Structure: Union by Rank



```
>>> union(b, n)
```

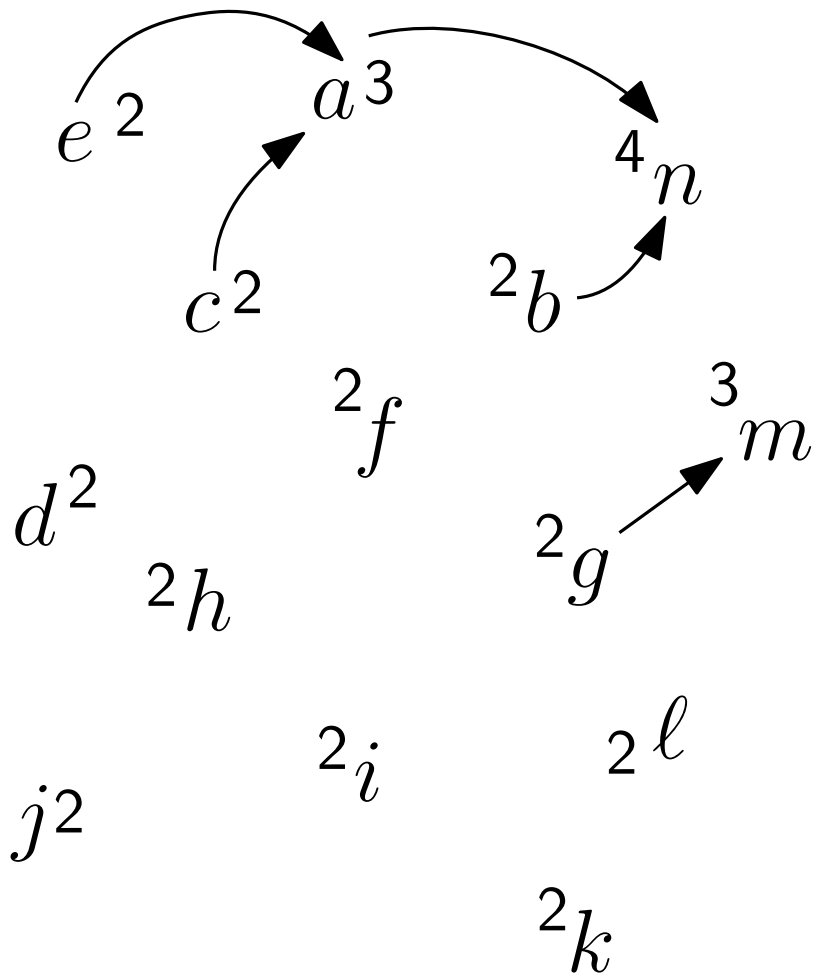
```
>>> union(a, n)
```

```
>>> find(e)
```

```
 $n$ 
```

```
>>> union(g, m)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

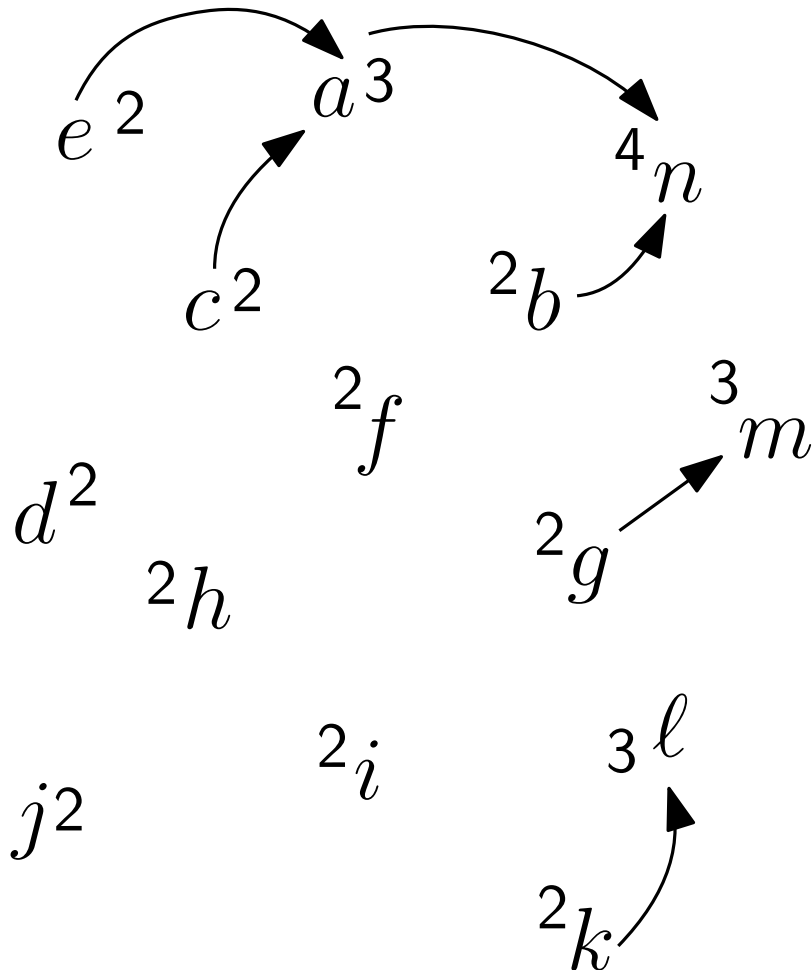
```
>>> find(e)
```

```
n
```

```
>>> union(g, m)
```

```
>>> union(k, l)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

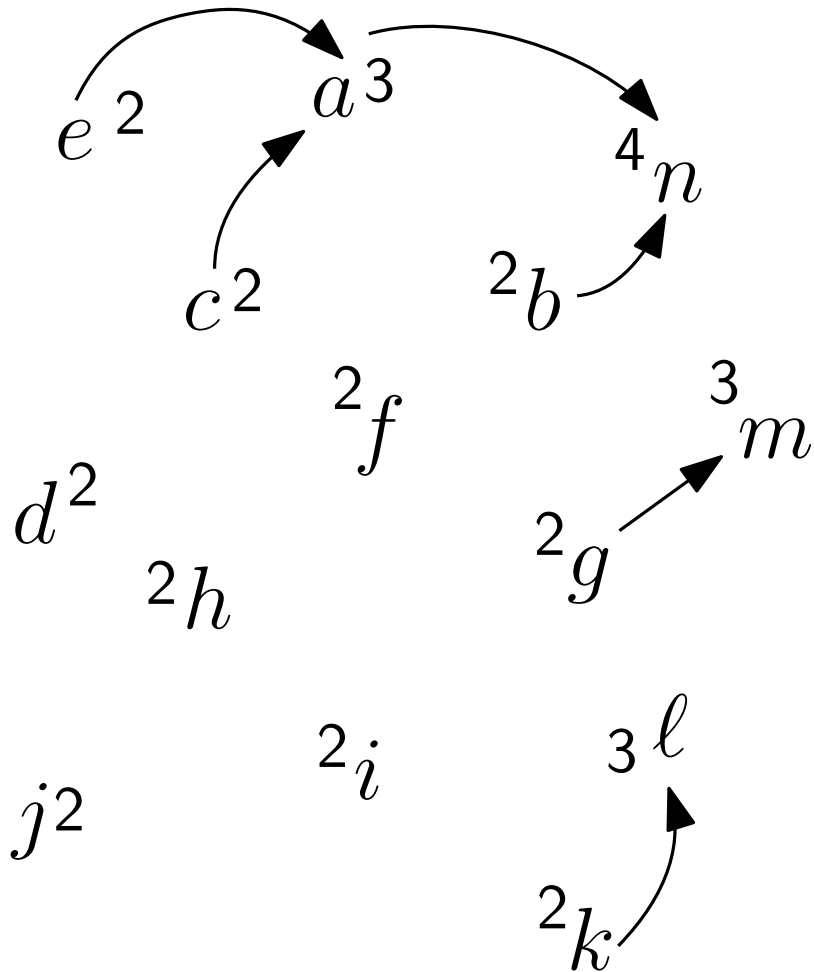
```
>>> find(e)
```

```
n
```

```
>>> union(g, m)
```

```
>>> union(k, l)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

```
>>> find(e)
```

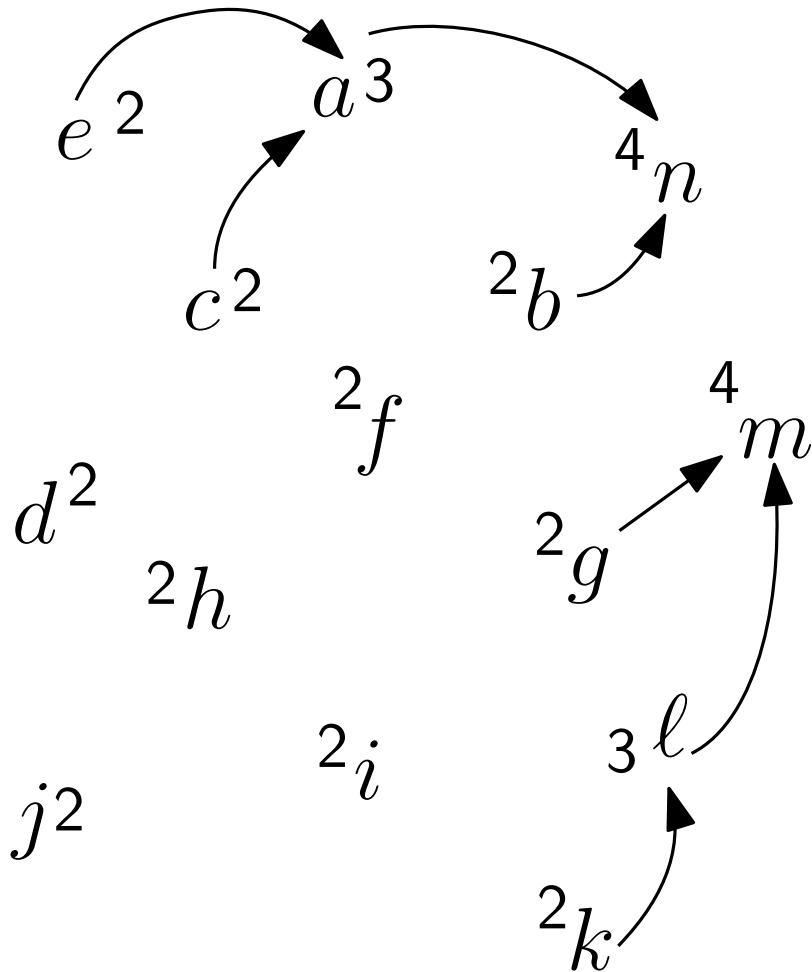
```
n
```

```
>>> union(g, m)
```

```
>>> union(k, l)
```

```
>>> union(l, m)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
```

```
>>> union(a, n)
```

```
>>> find(e)
```

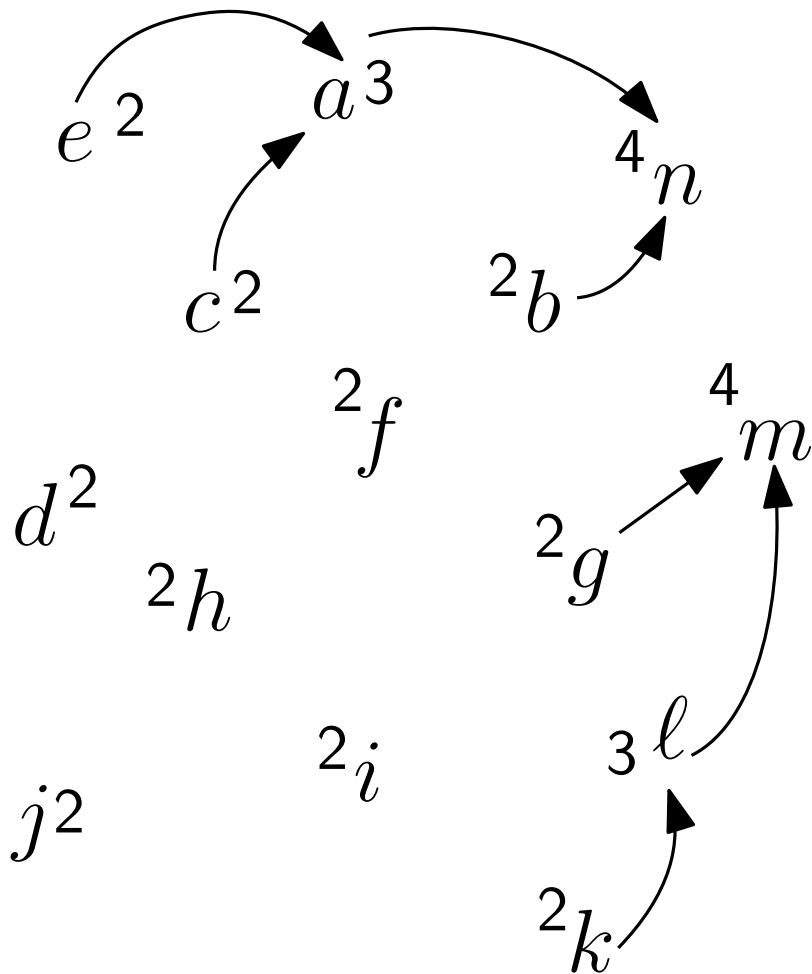
```
n
```

```
>>> union(g, m)
```

```
>>> union(k, l)
```

```
>>> union(l, m)
```

The Data Structure: Union by Rank

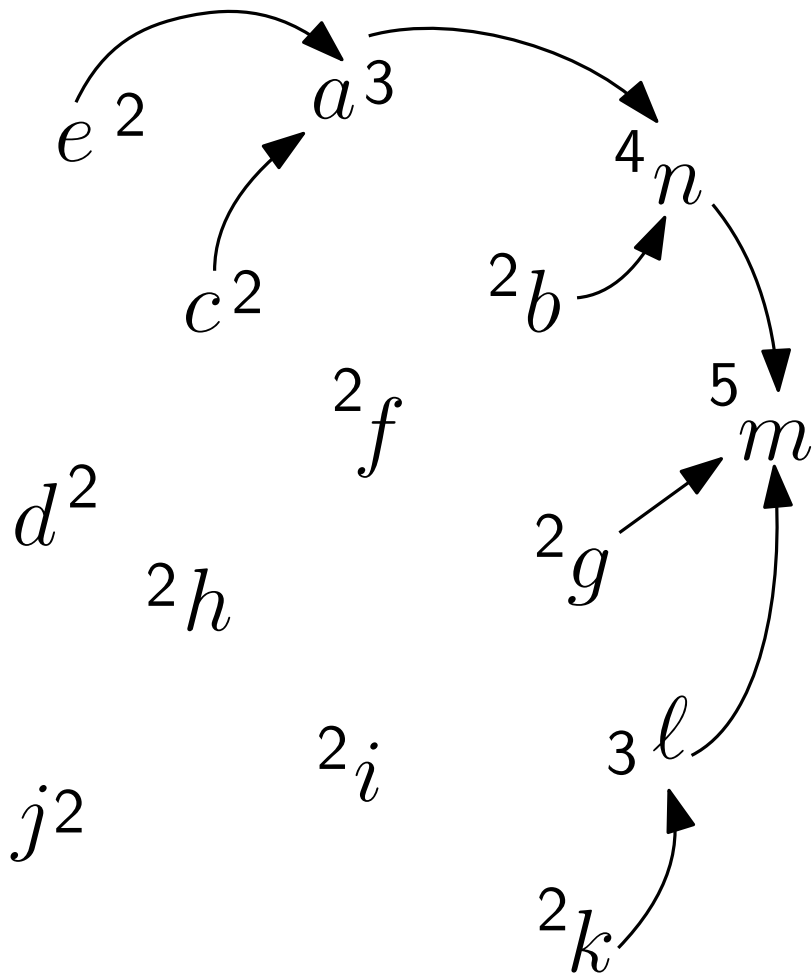


```
>>> union(b, n)
>>> union(a, n)

>>> find(e)
n

>>> union(g, m)
>>> union(k, l)
>>> union(l, m)
>>> union(n, m)
```

The Data Structure: Union by Rank



```
>>> union(b, n)
>>> union(a, n)

>>> find(e)
n

>>> union(g, m)
>>> union(k, l)
>>> union(l, m)
>>> union(n, m)
```

Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.

Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.

Lemma. The number of nodes in x 's subtree is at least $2^{\text{rank}(x)-2}$.

Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.

Lemma. The number of nodes in x 's subtree is at least $2^{\text{rank}(x)-2}$.

Lemma. The number of elements with rank r is at most $\frac{n}{2^{r-2}} = \frac{4n}{2^r}$.

Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.

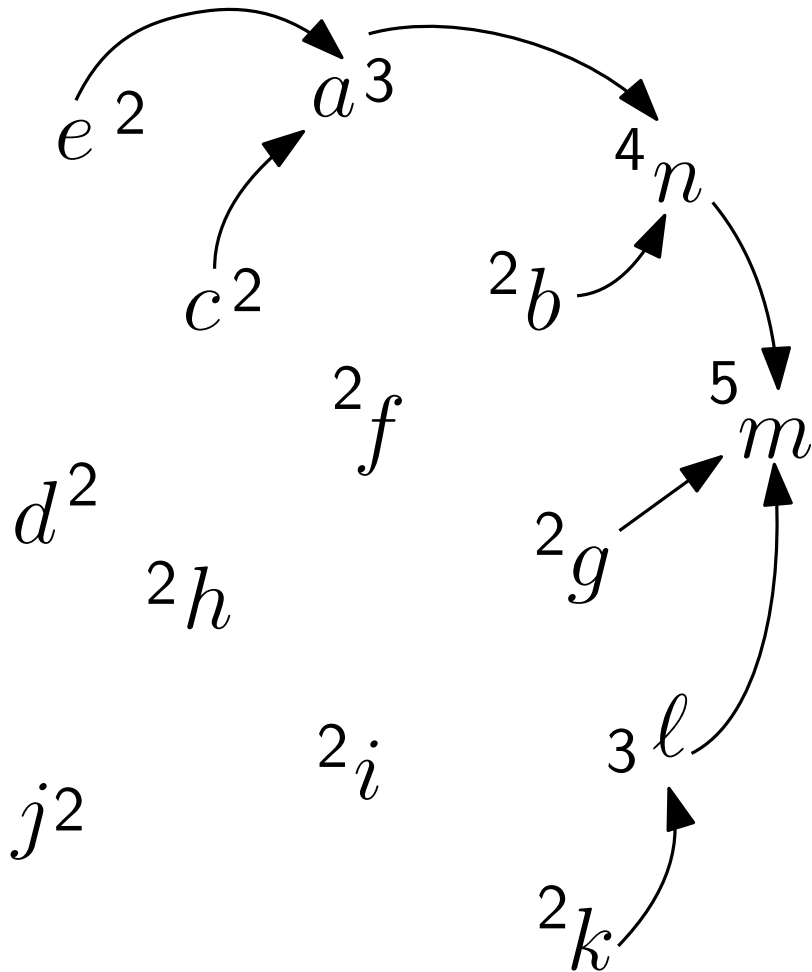
Lemma. The number of elements with rank r is at most $\frac{n}{2^{r-2}} = \frac{4n}{2^r}$.

Lemma. The number of nodes in x 's subtree is at least $2^{\text{rank}(x)-2}$.

Corollary. The maximum rank is at most $\log(n) + 2$. The maximum height is at most $\log(n)$. The operation $\text{find}(x)$ takes $O(\log n)$ steps.

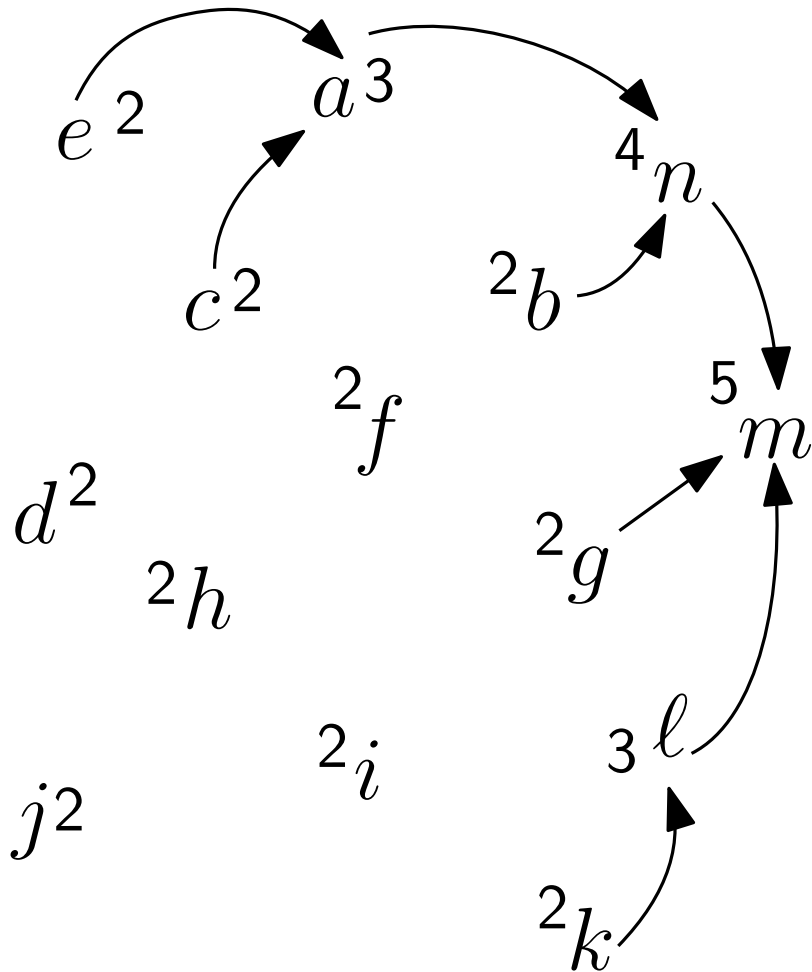
Path Compression

The Data Structure: Union by Rank



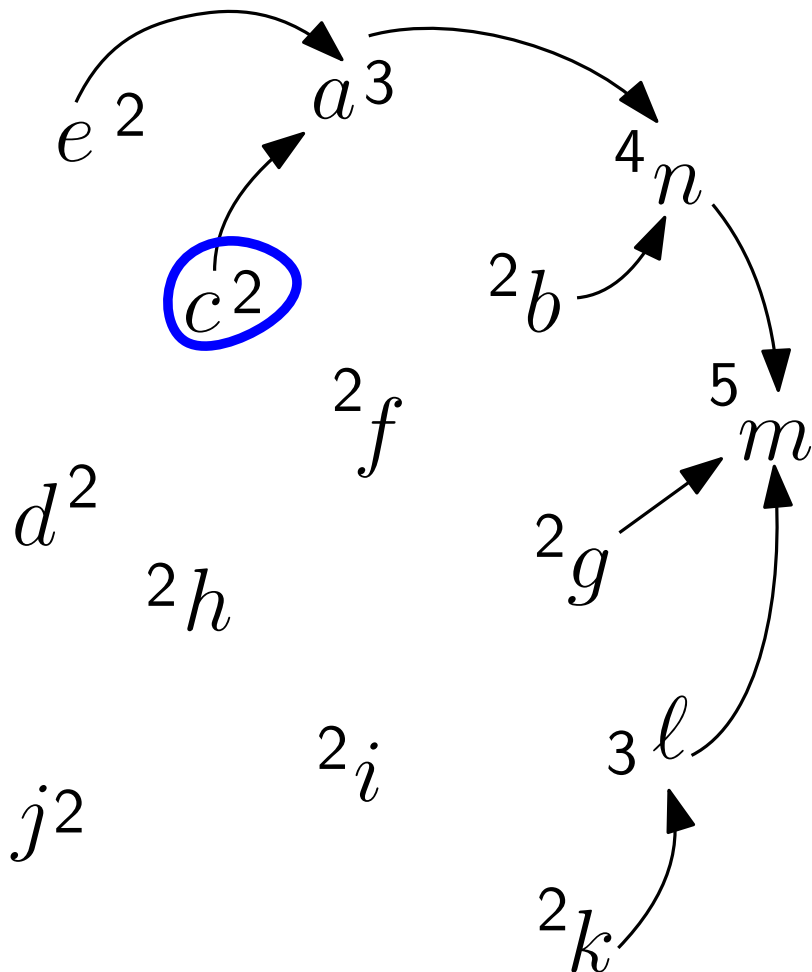
The Data Structure: Union by Rank

```
>>> find(c)
```



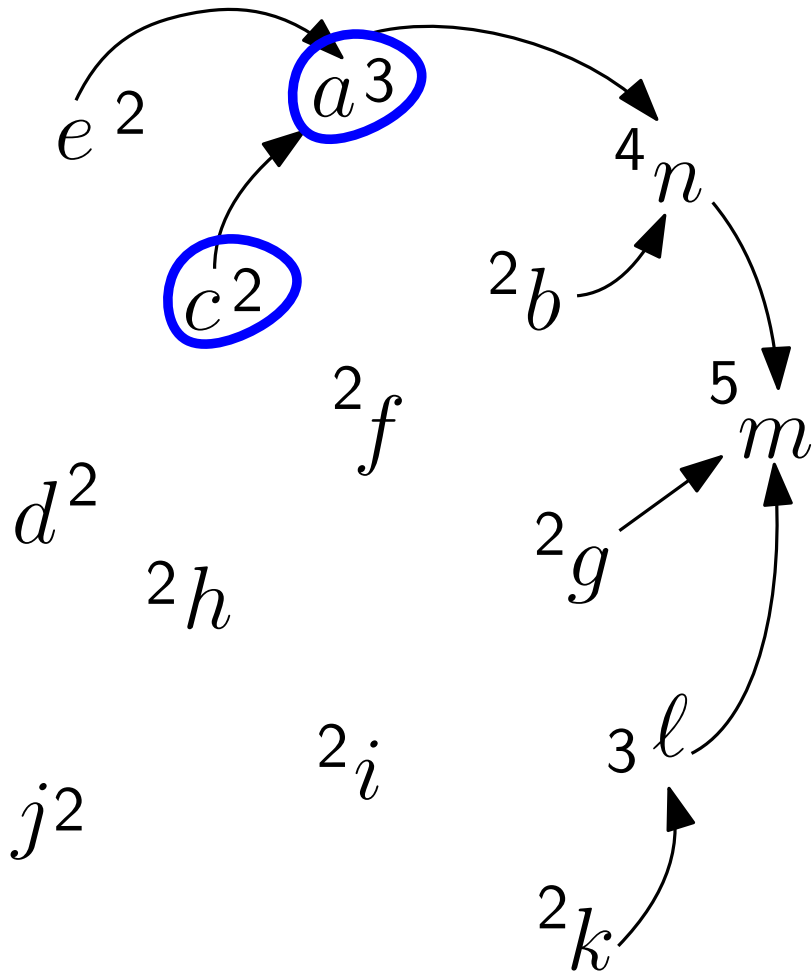
The Data Structure: Union by Rank

>>> find(*c*)



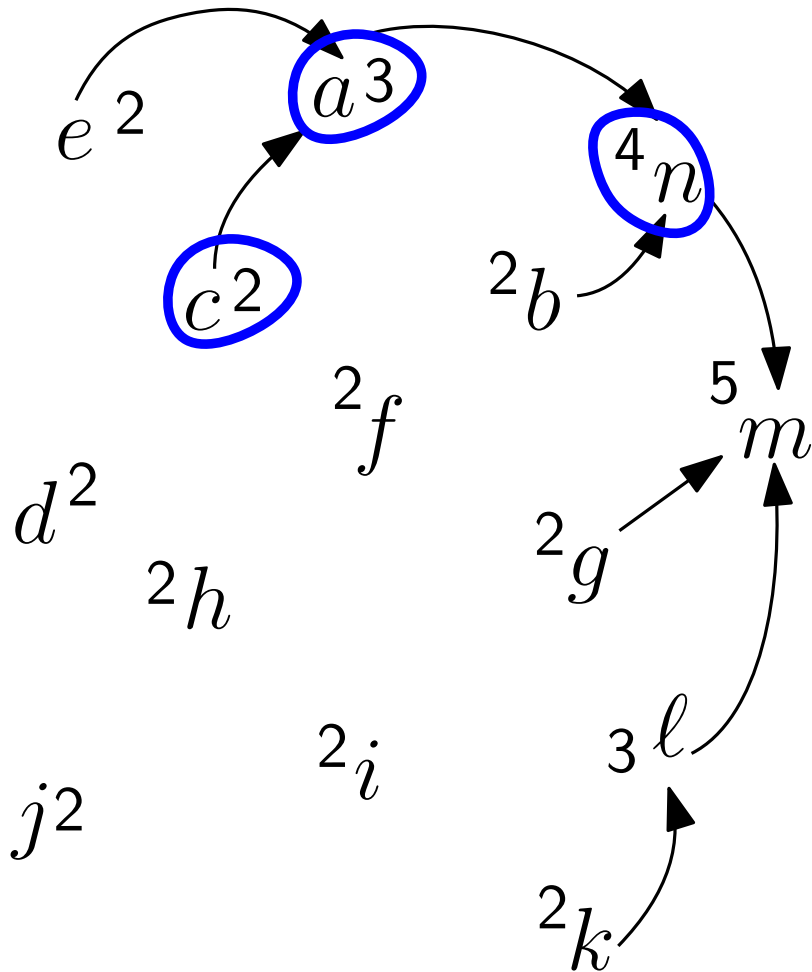
The Data Structure: Union by Rank

>>> find(*c*)



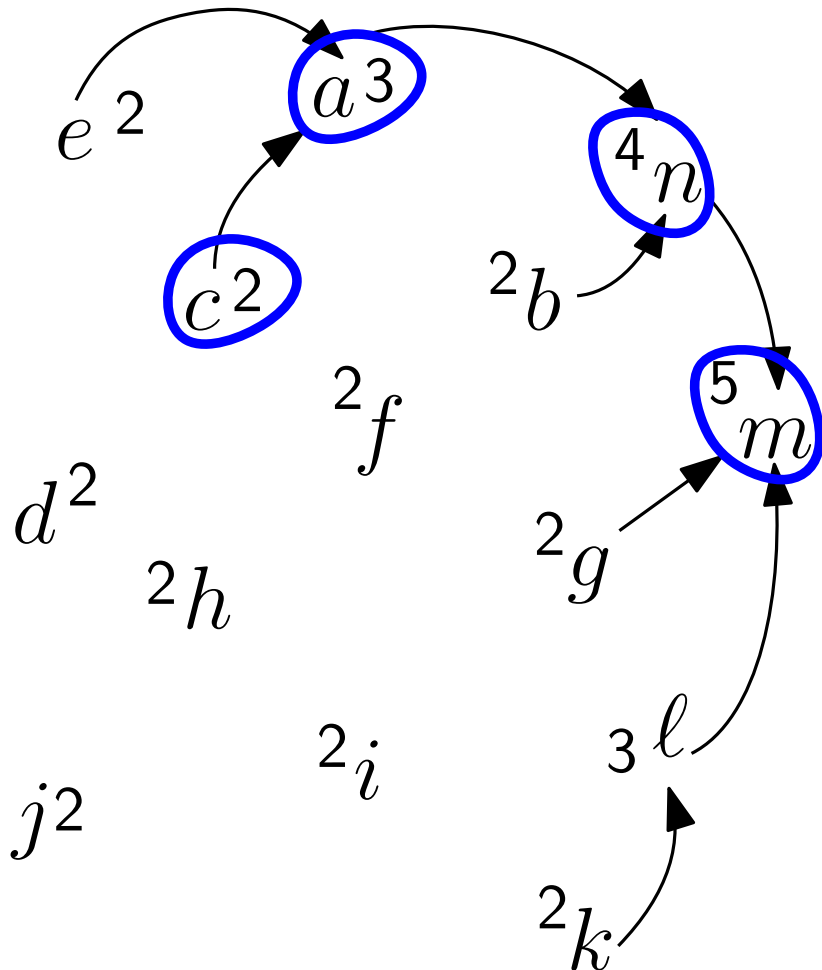
The Data Structure: Union by Rank

>>> find(*c*)

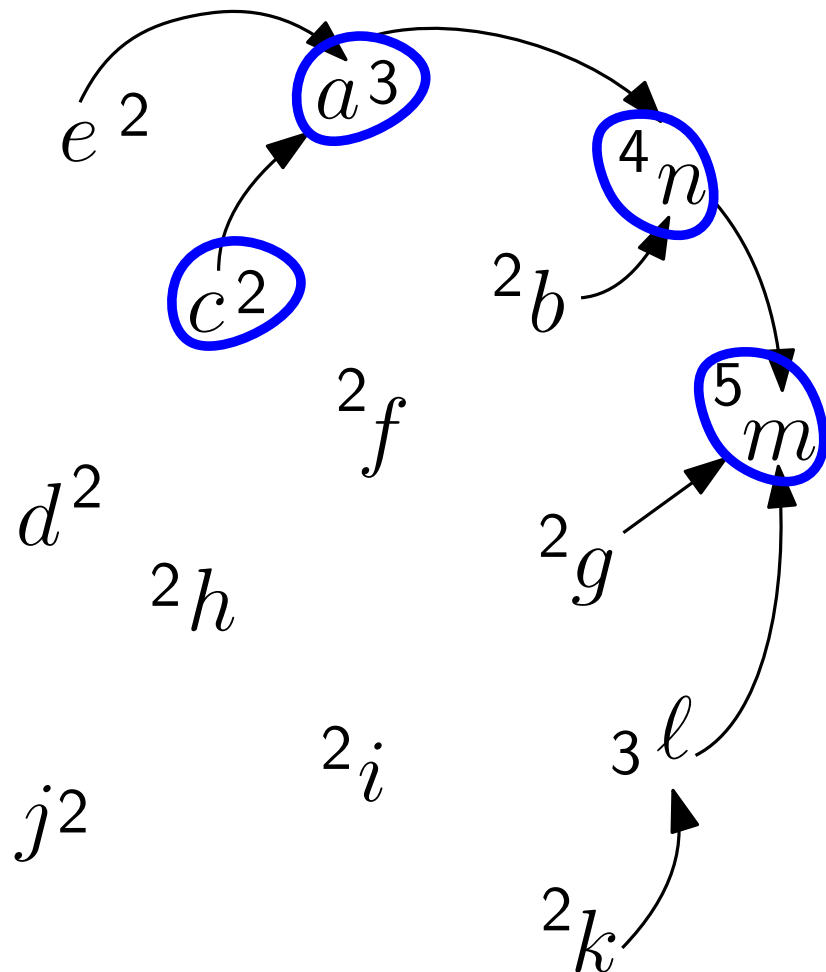


The Data Structure: Union by Rank

>>> find(*c*)



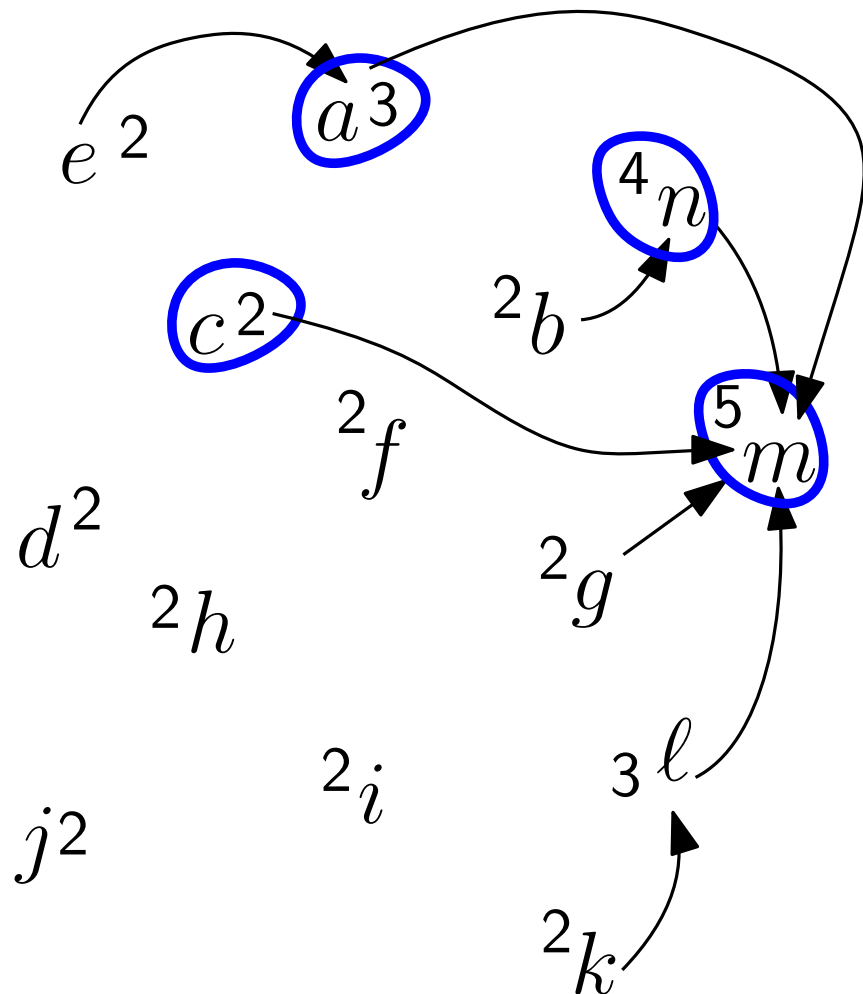
The Data Structure: Union by Rank



`>>> find(c)`

m

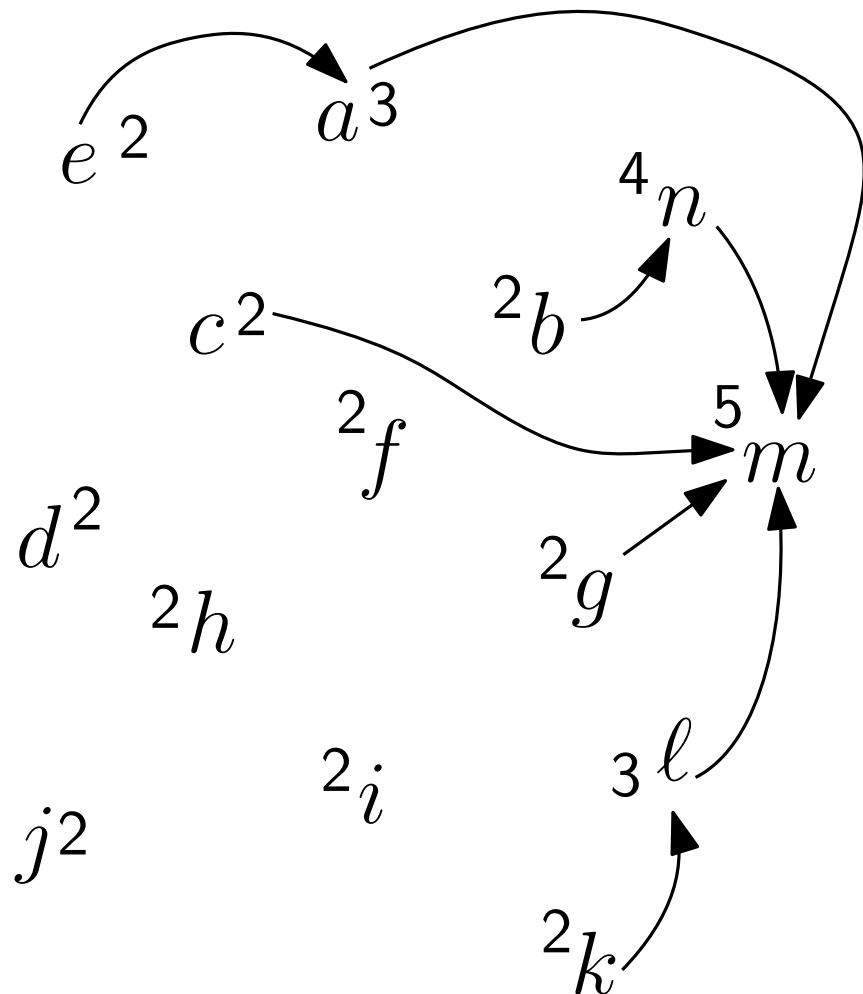
The Data Structure: Union by Rank



`>>> find(c)`

m

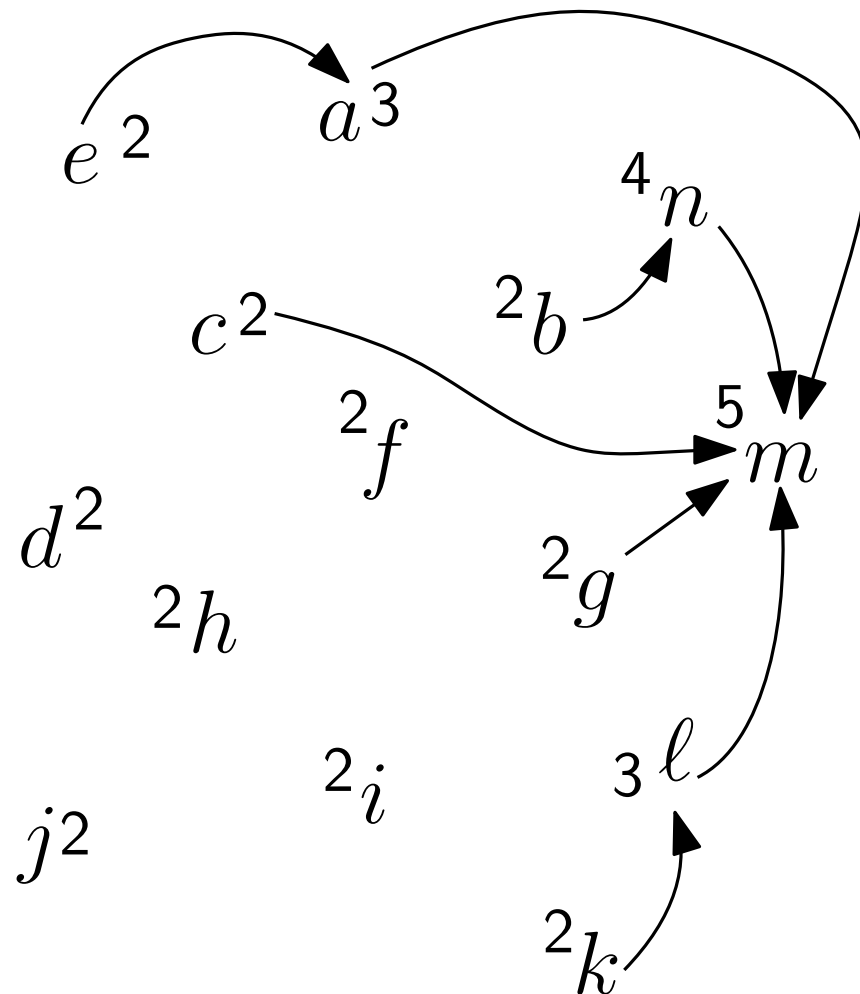
The Data Structure: Union by Rank



`>>> find(c)`

m

The Data Structure: Union by Rank



>>> find(*c*)

m

rank(*n*) is still 4. But it's subtree has height 1 and consists of 2 nodes only.

Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.

Lemma. The number of elements with rank r is at most $\frac{n}{2^{r-2}} = \frac{4n}{2^r}$.

Lemma. The number of nodes in x 's subtree is at least $2^{\text{rank}(x)-2}$.

Corollary. The maximum rank is at most $\log(n) + 2$. The maximum height is at most $\log(n)$. The operation $\text{find}(x)$ takes $O(\log n)$ steps.

~~Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.~~

Lemma. The number of elements with rank r is at most $\frac{n}{2^{r-2}} = \frac{4n}{2^r}$.

~~Lemma. The number of nodes in x 's subtree is at least $2^{\text{rank}(x)-2}$.~~

Corollary. The maximum rank is at most $\log(n) + 2$. The maximum height is at most $\log(n)$. The operation $\text{find}(x)$ takes $O(\log n)$ steps.

~~Lemma. The height of the subtree rooted at x is $\text{rank}(x) - 2$.~~

~~Lemma. The number of nodes in x 's subtree is at least $2^{\text{rank}(x)-2}$.~~

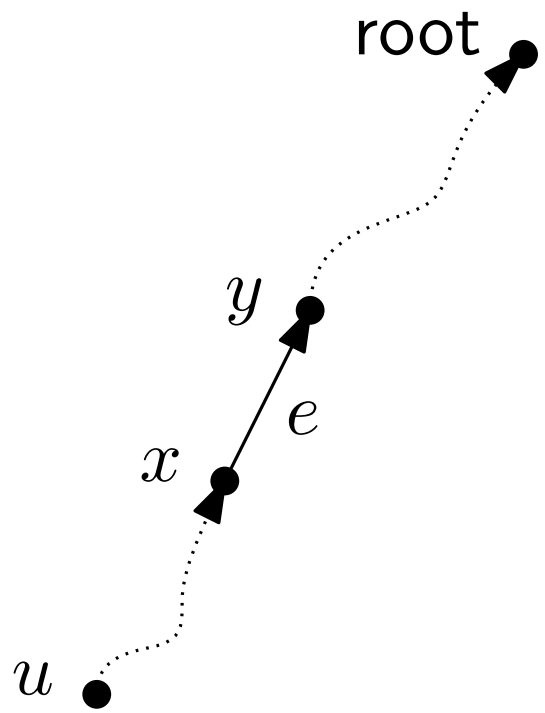
Lemma. The number of elements with rank r is at most $\frac{n}{2^{r-2}} = \frac{4n}{2^r}$.

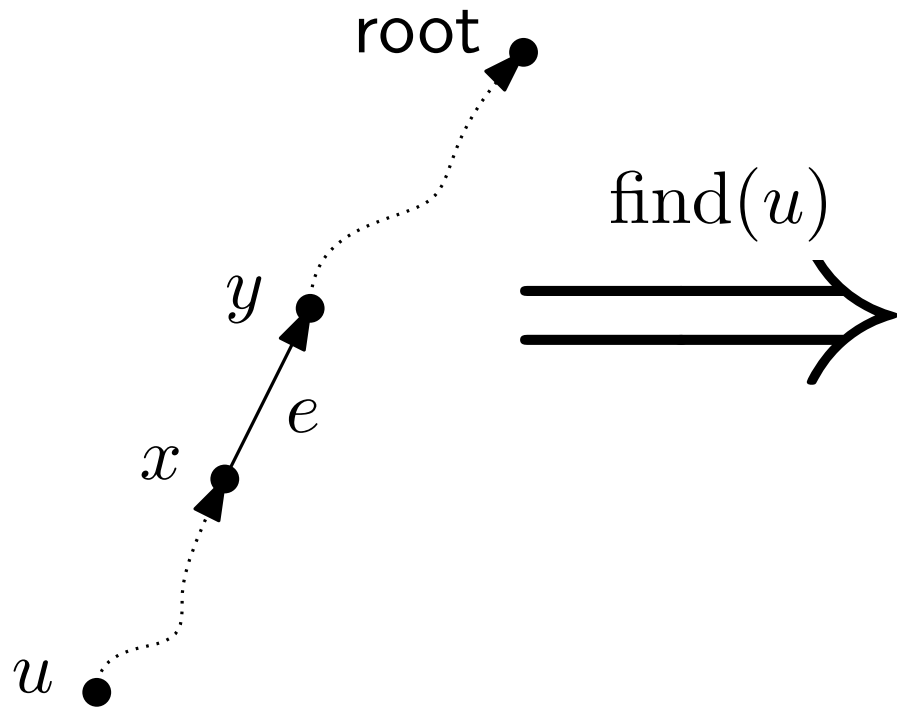
Corollary. The maximum rank is at most $\log(n) + 2$. The maximum height is at most $\log(n)$. The operation $\text{find}(x)$ takes $O(\log n)$ steps.

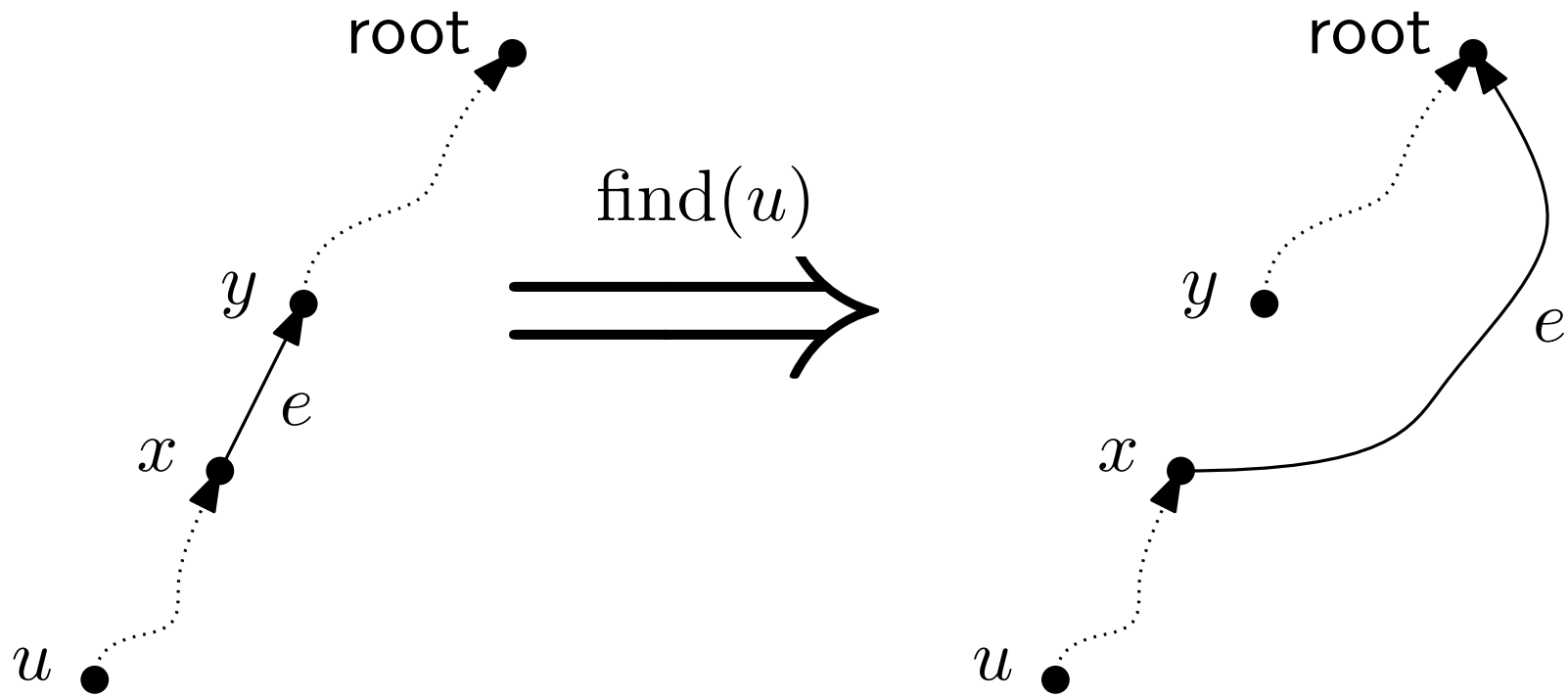
Running Time Analysis of the Path Compression Data Structure

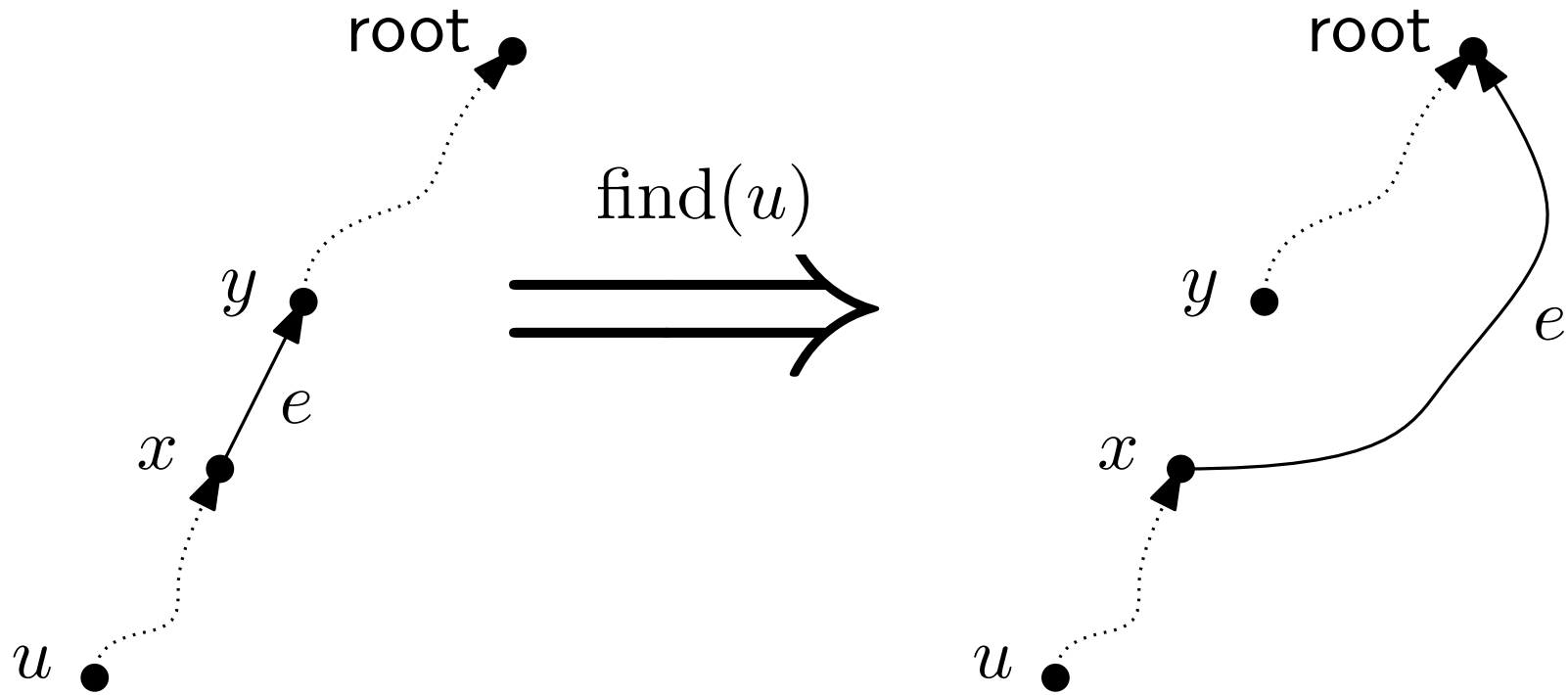
Running Time Analysis of the Path Compression Data Structure

First Attempt

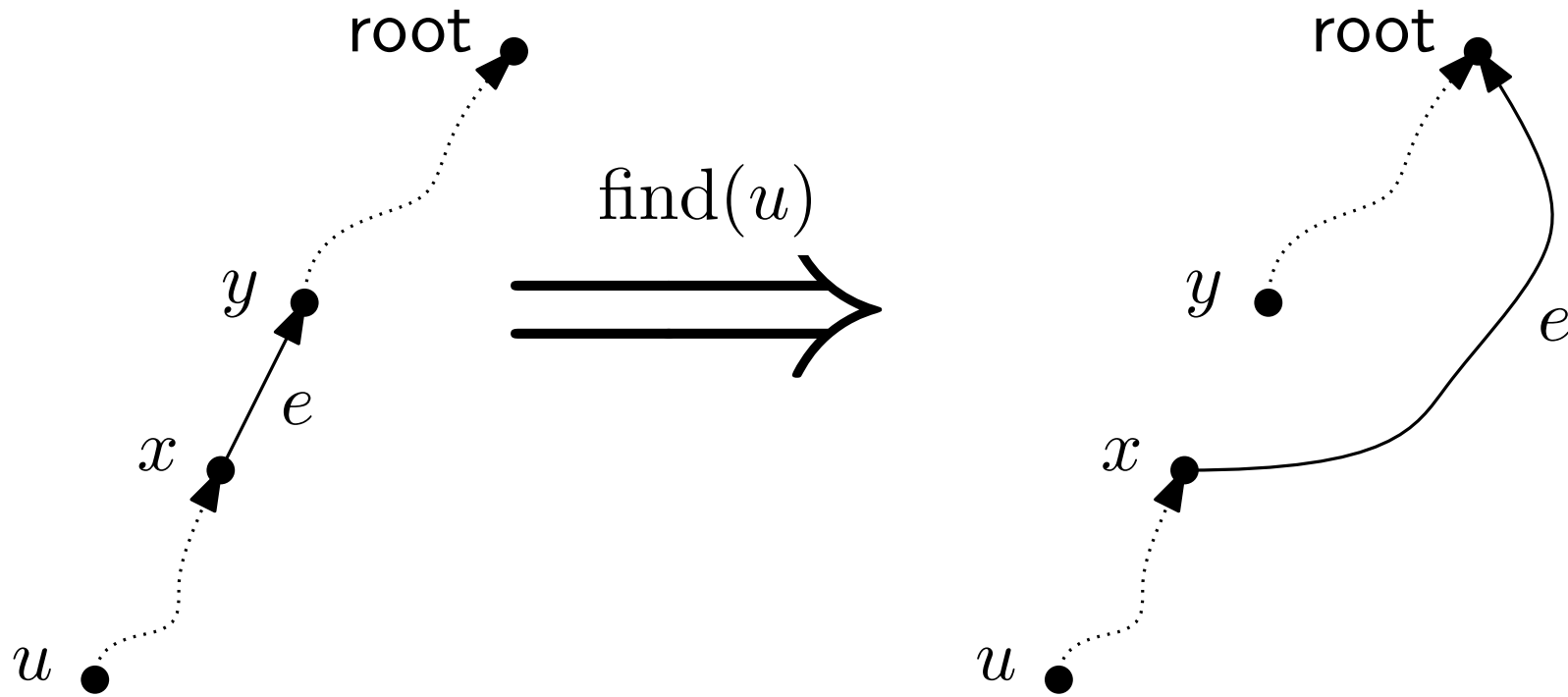






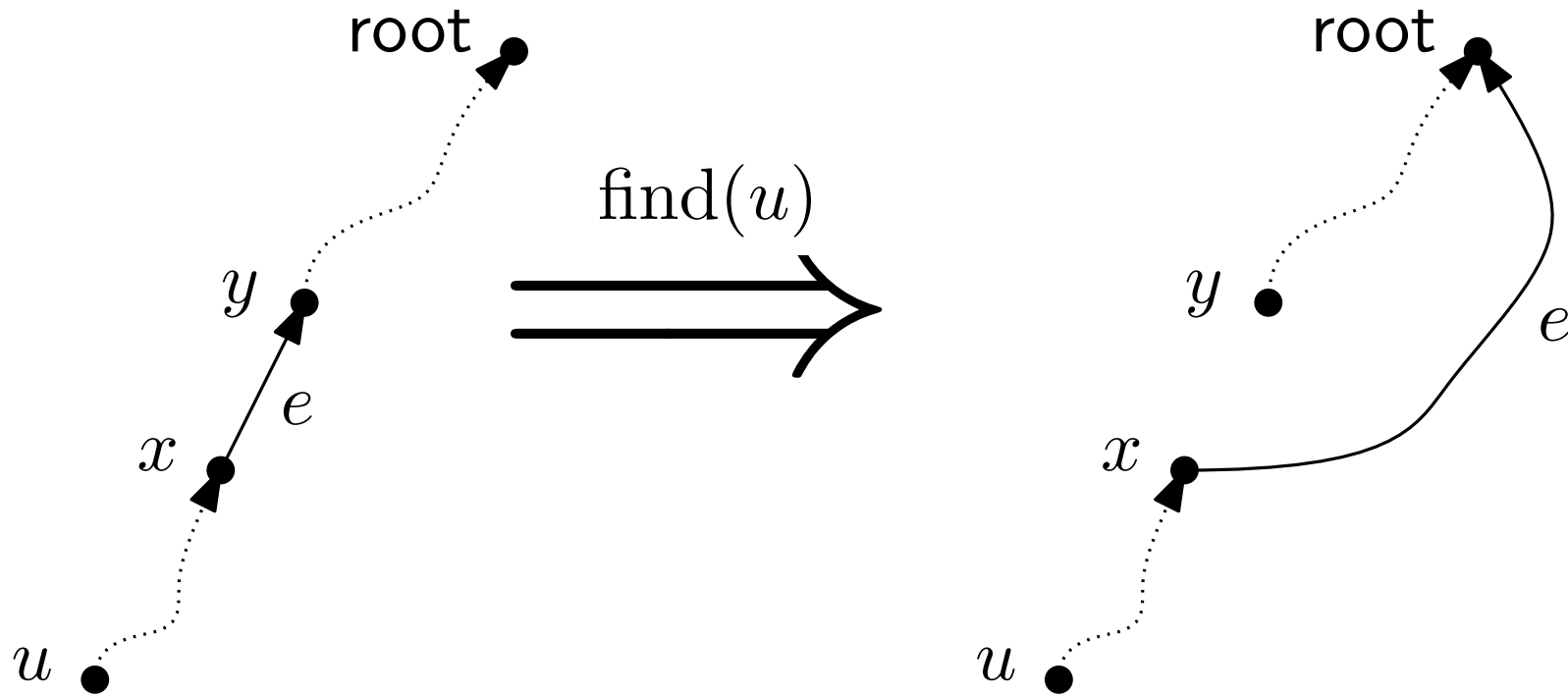


Edge e gets redirected.



Edge e gets redirected.

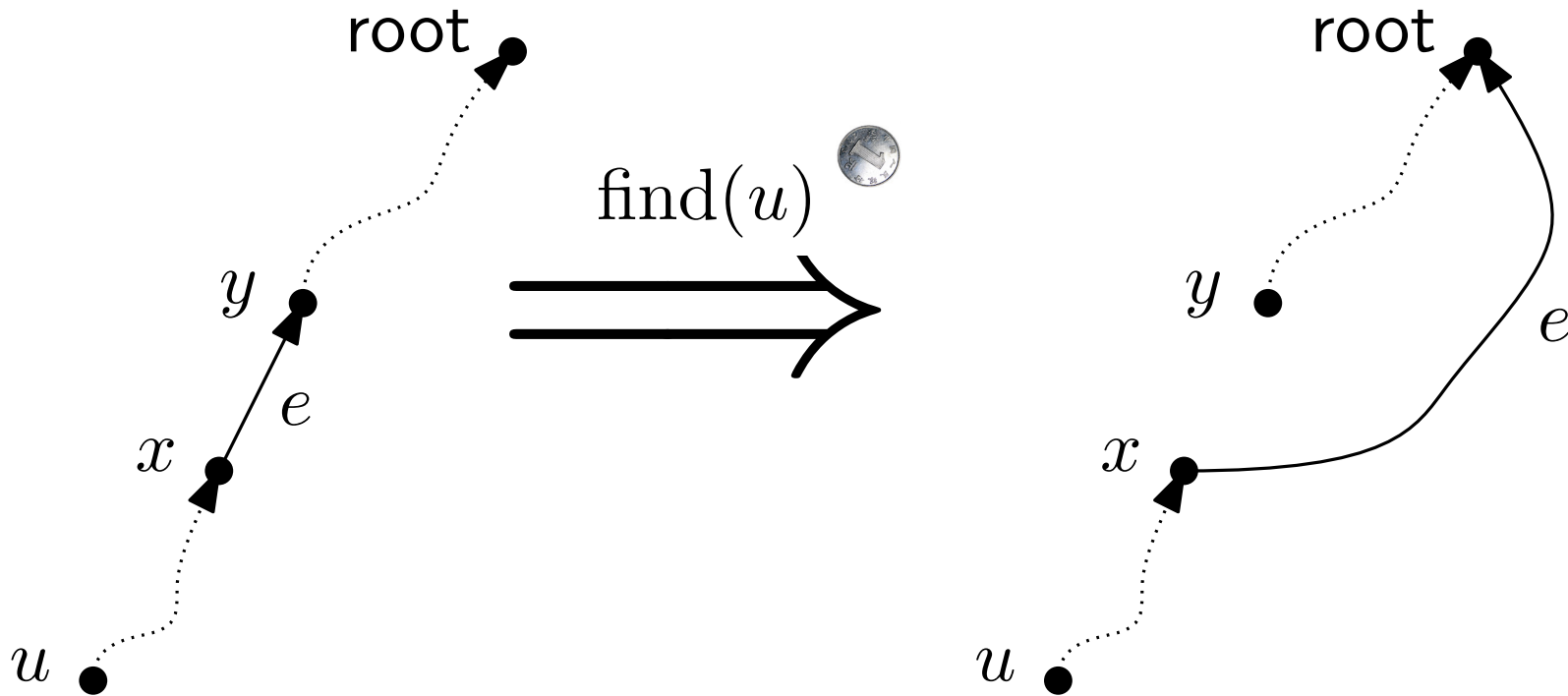
This operation costs 1 



Edge e gets redirected.

This operation costs 1 

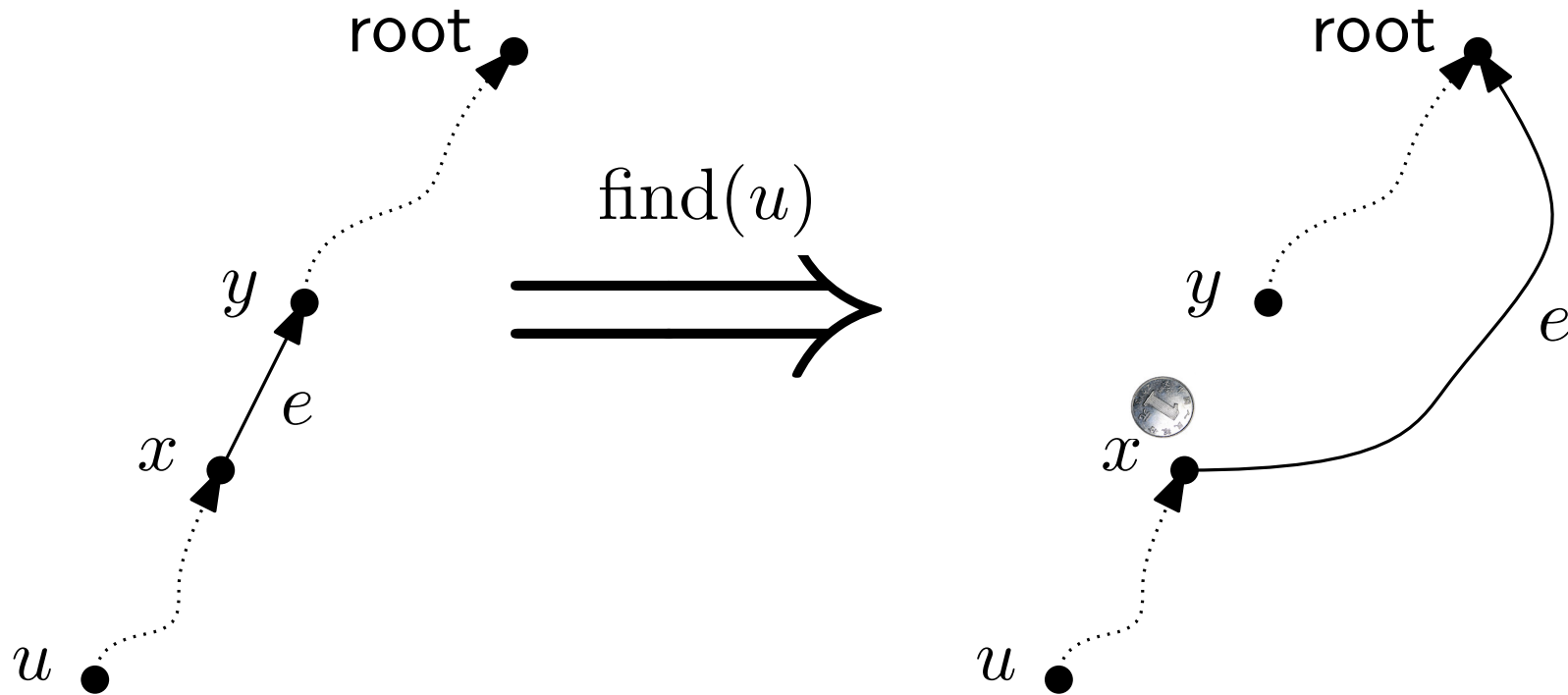
Who has to pay?



Edge e gets redirected.

This operation costs 1 

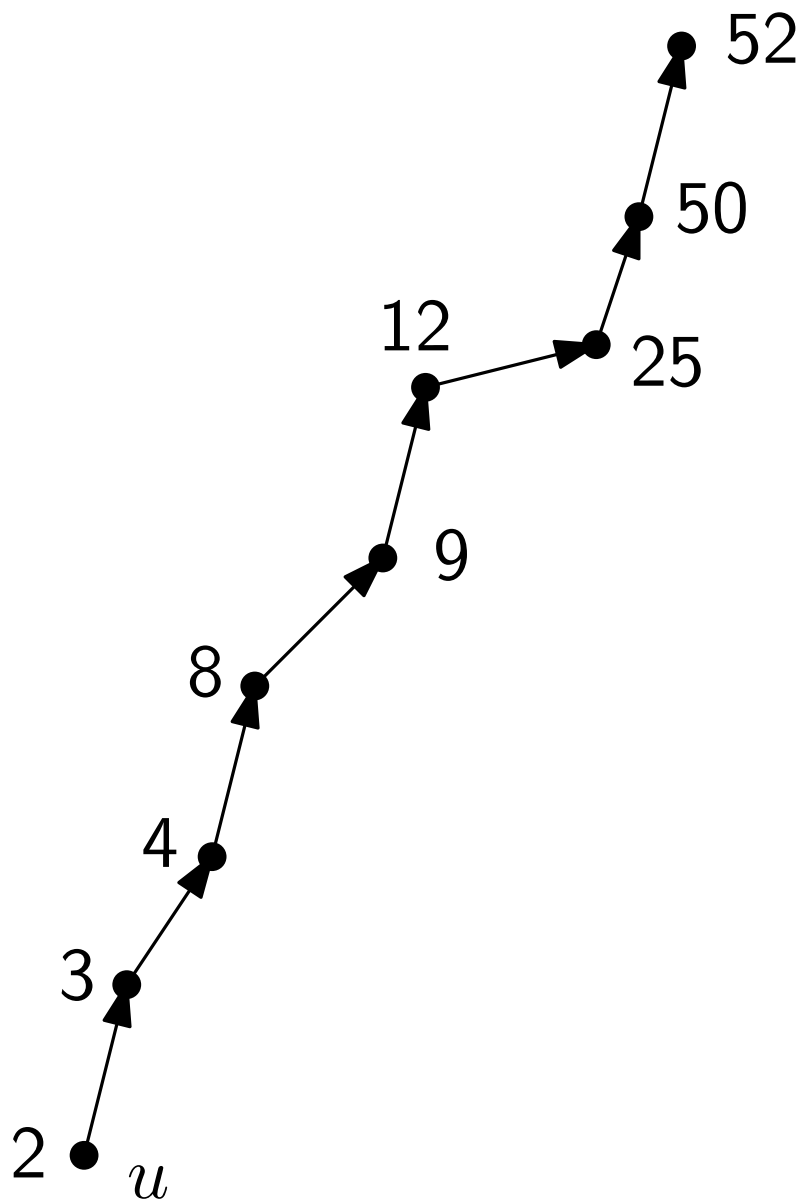
Who has to pay?

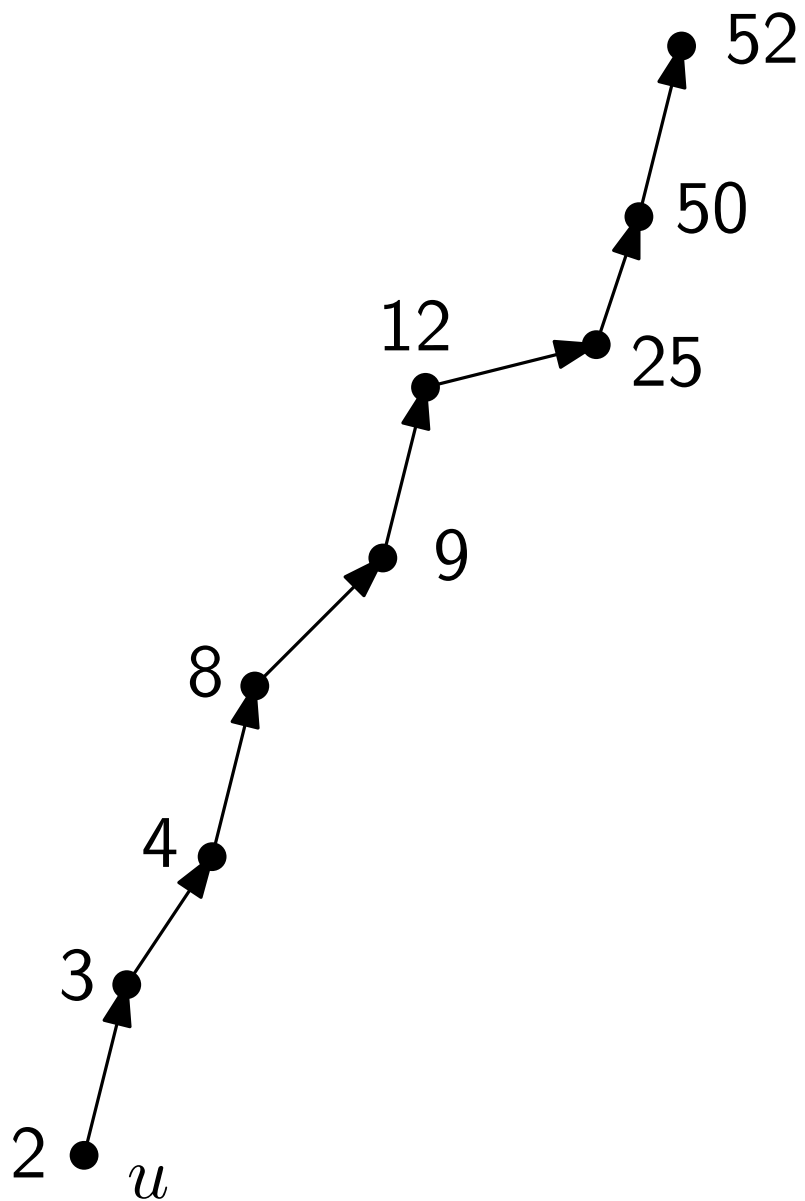


Edge e gets redirected.

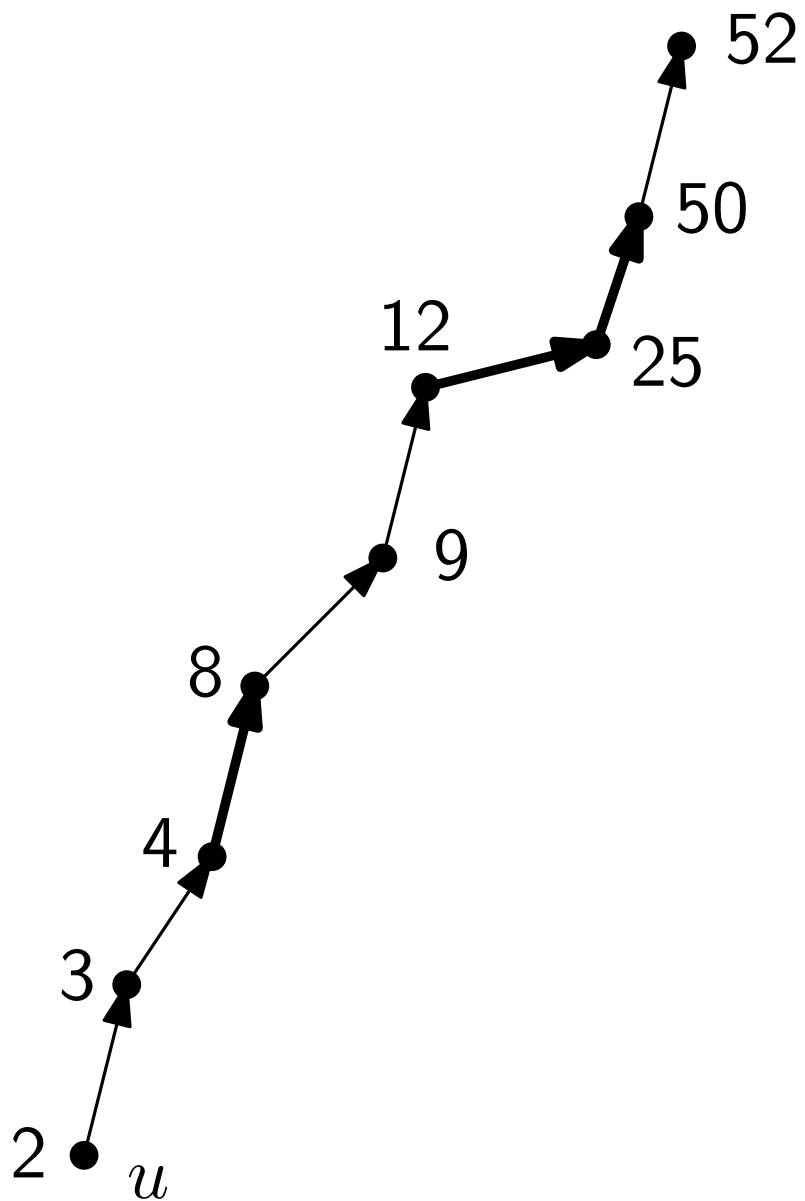
This operation costs 1 

Who has to pay?

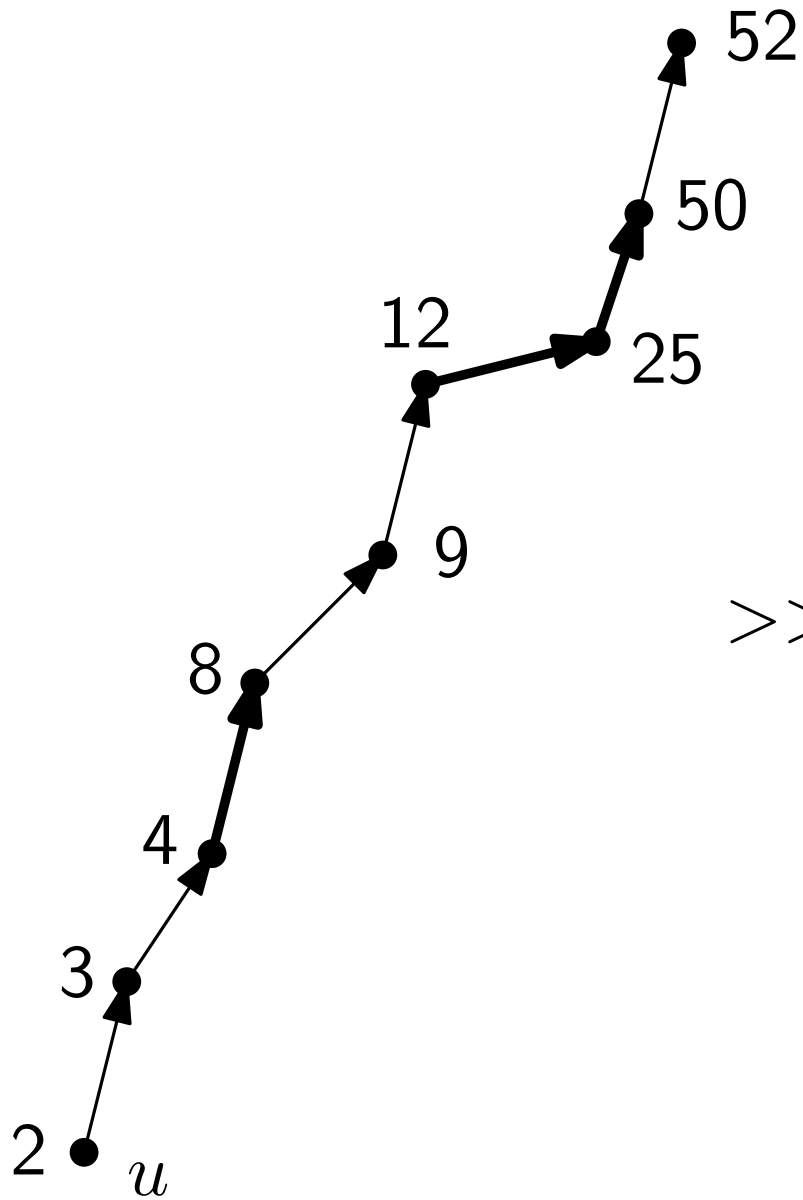




Definition. An edge $x \longrightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

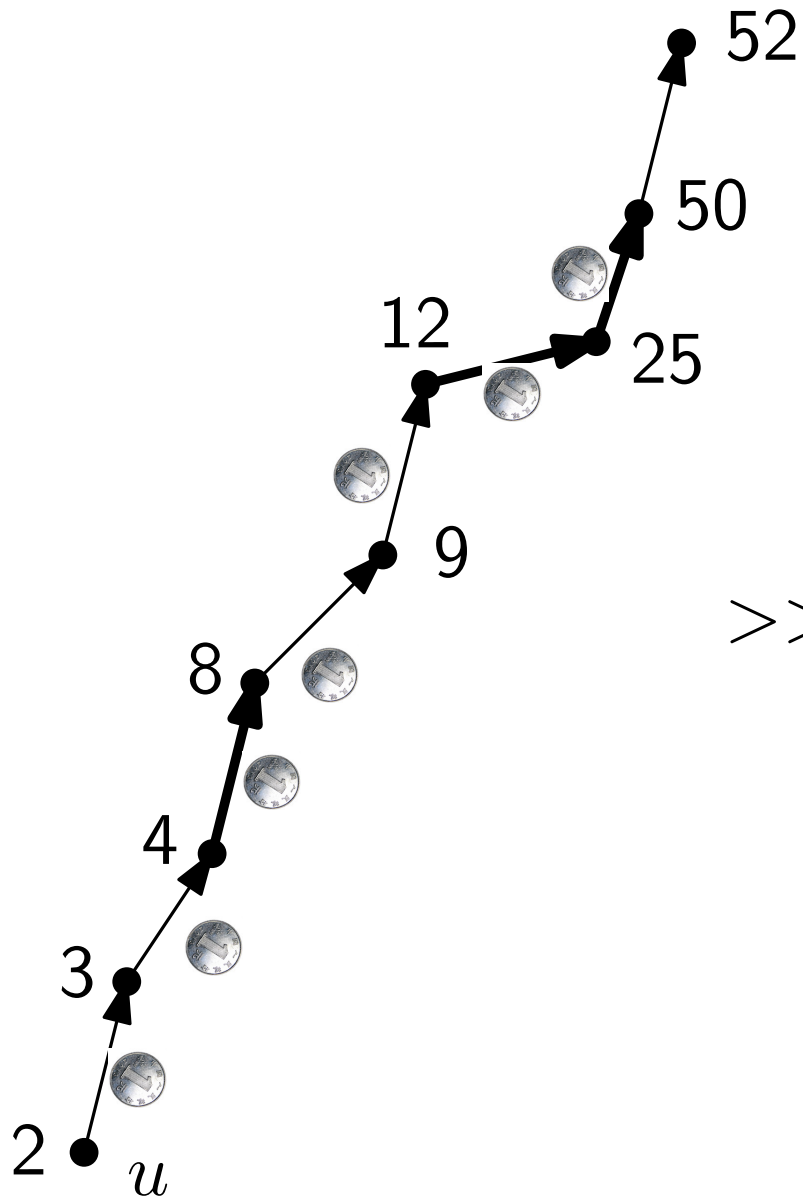


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



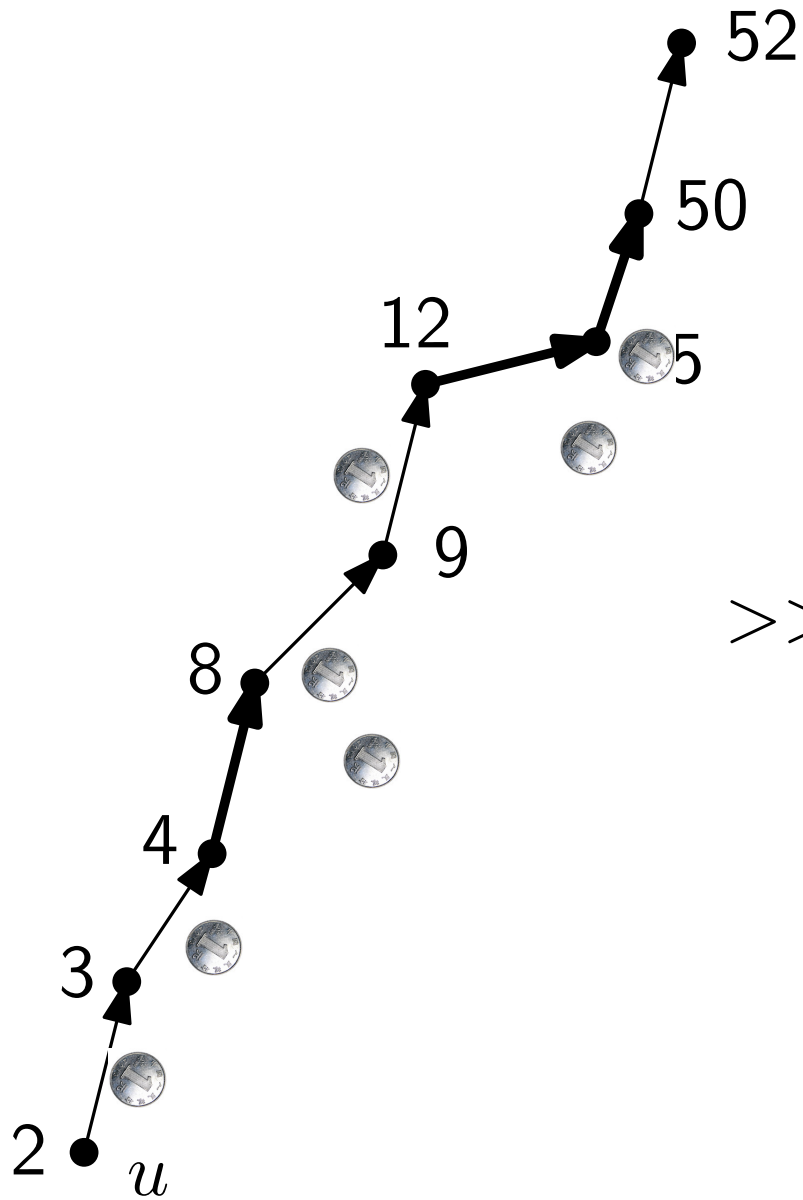
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$



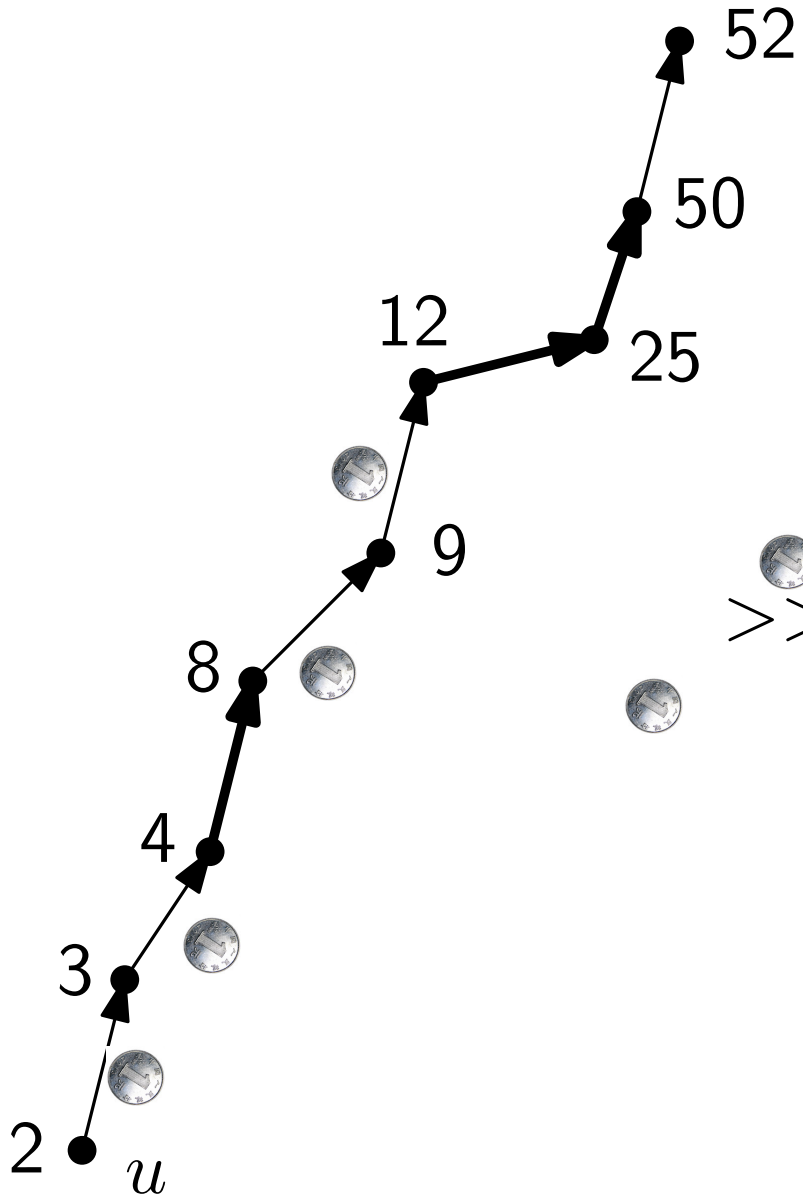
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

>>> find(u)



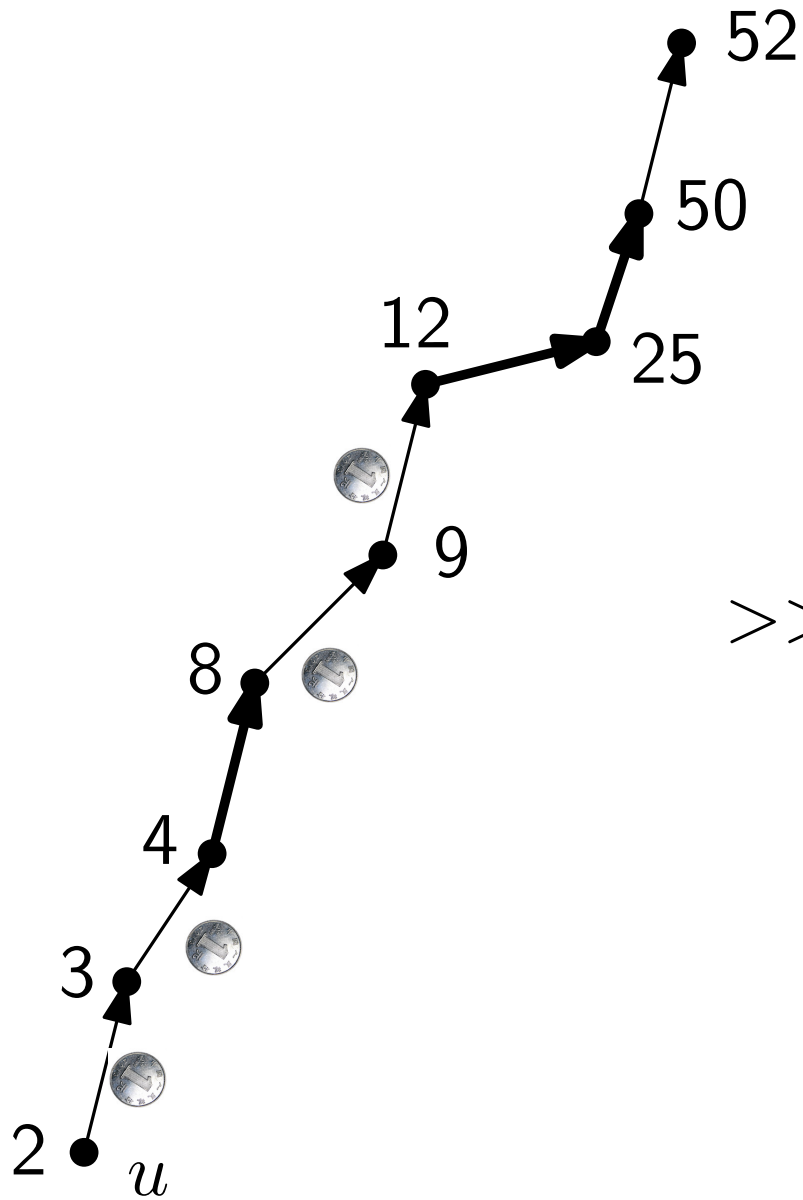
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$



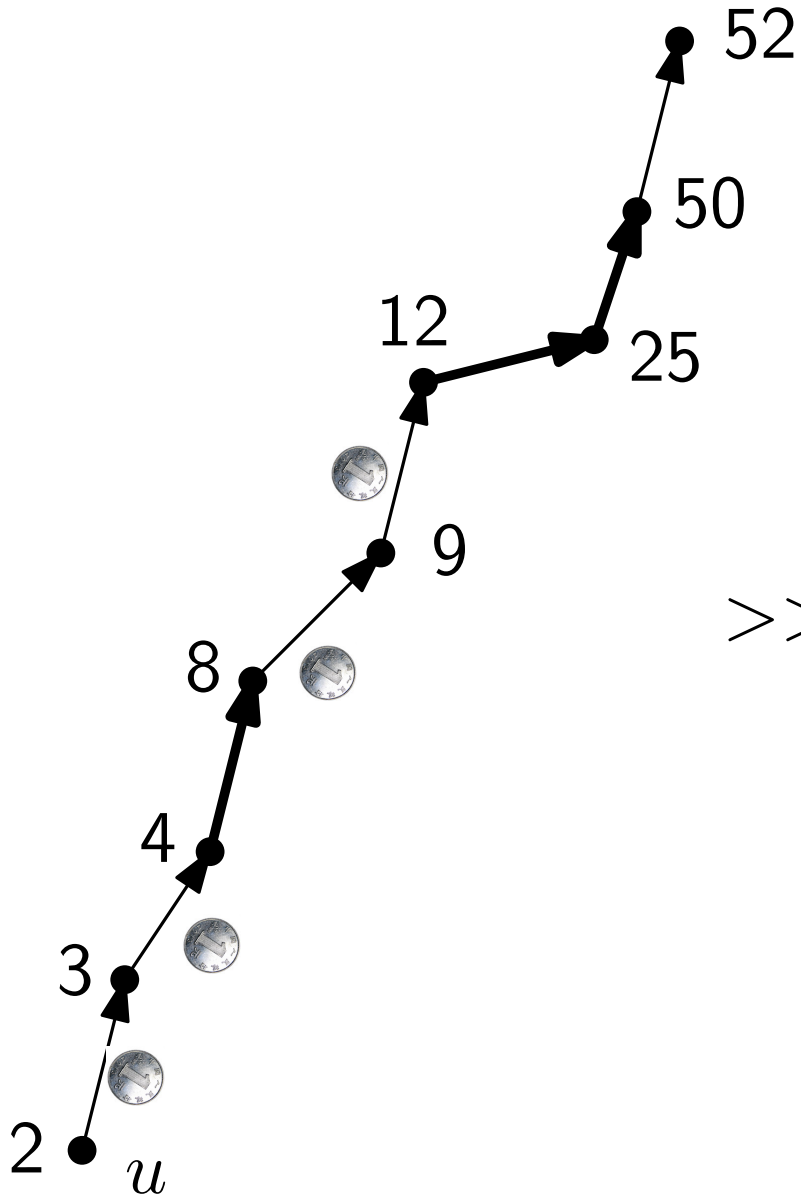
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$




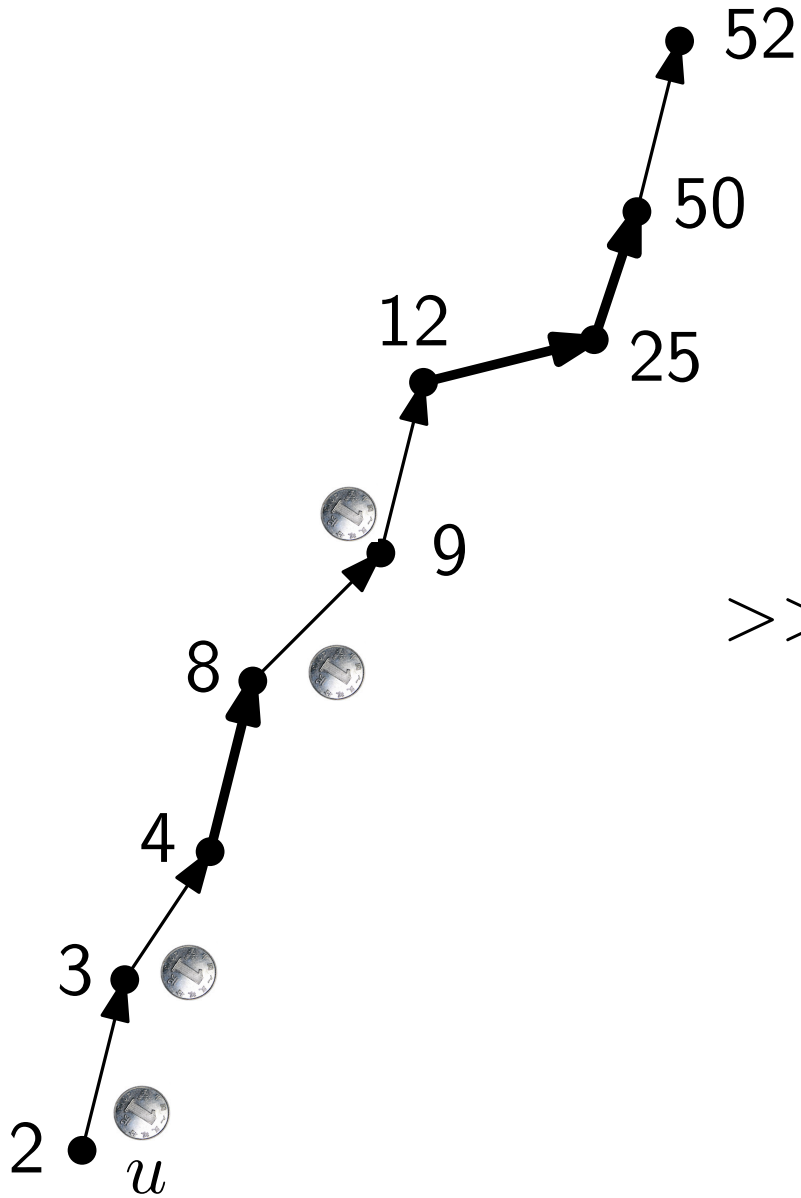
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$




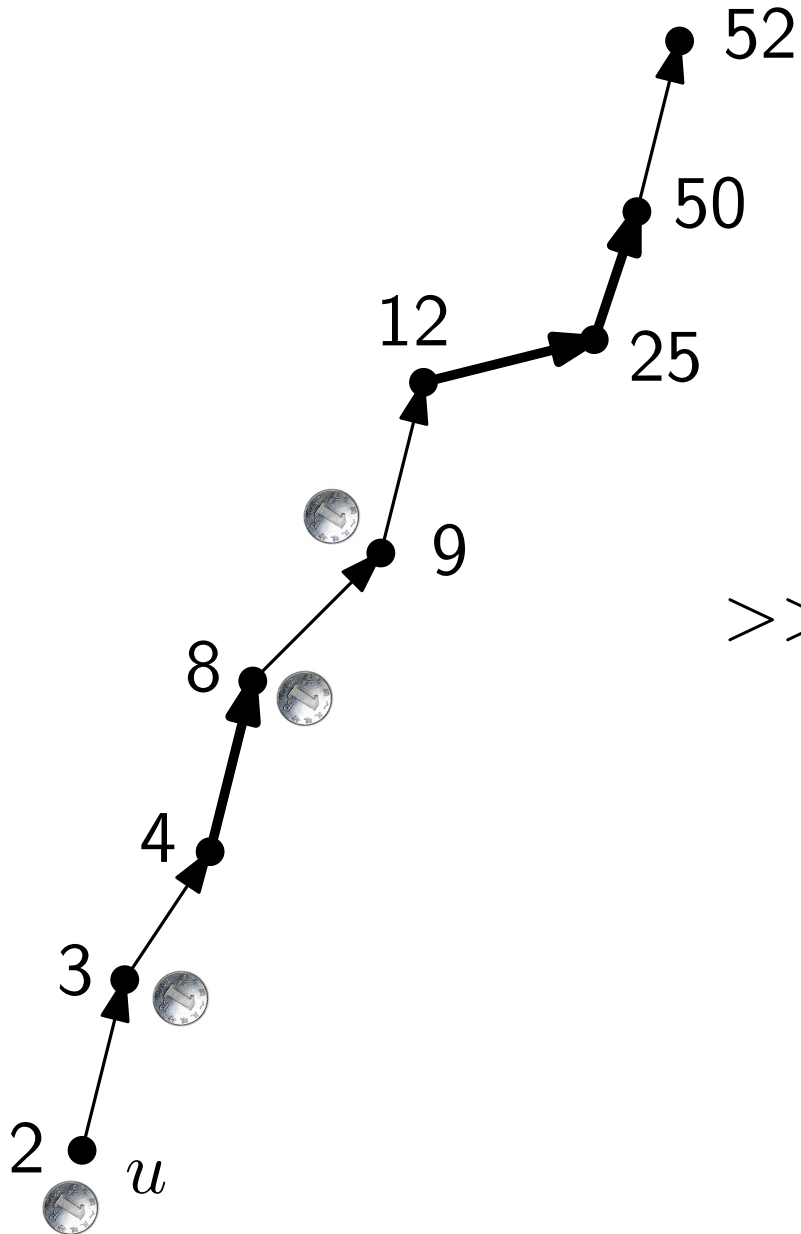
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

>>> find(u) 




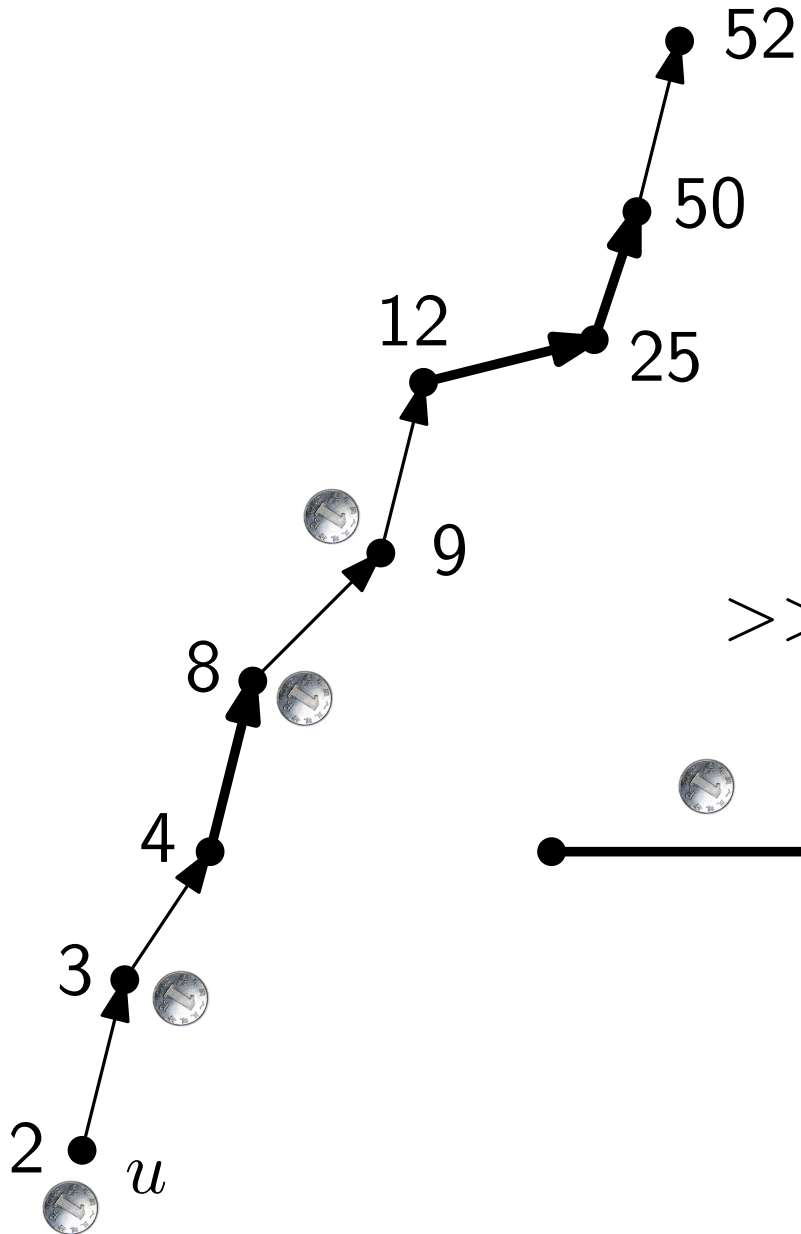
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

>>> find(u) 




Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

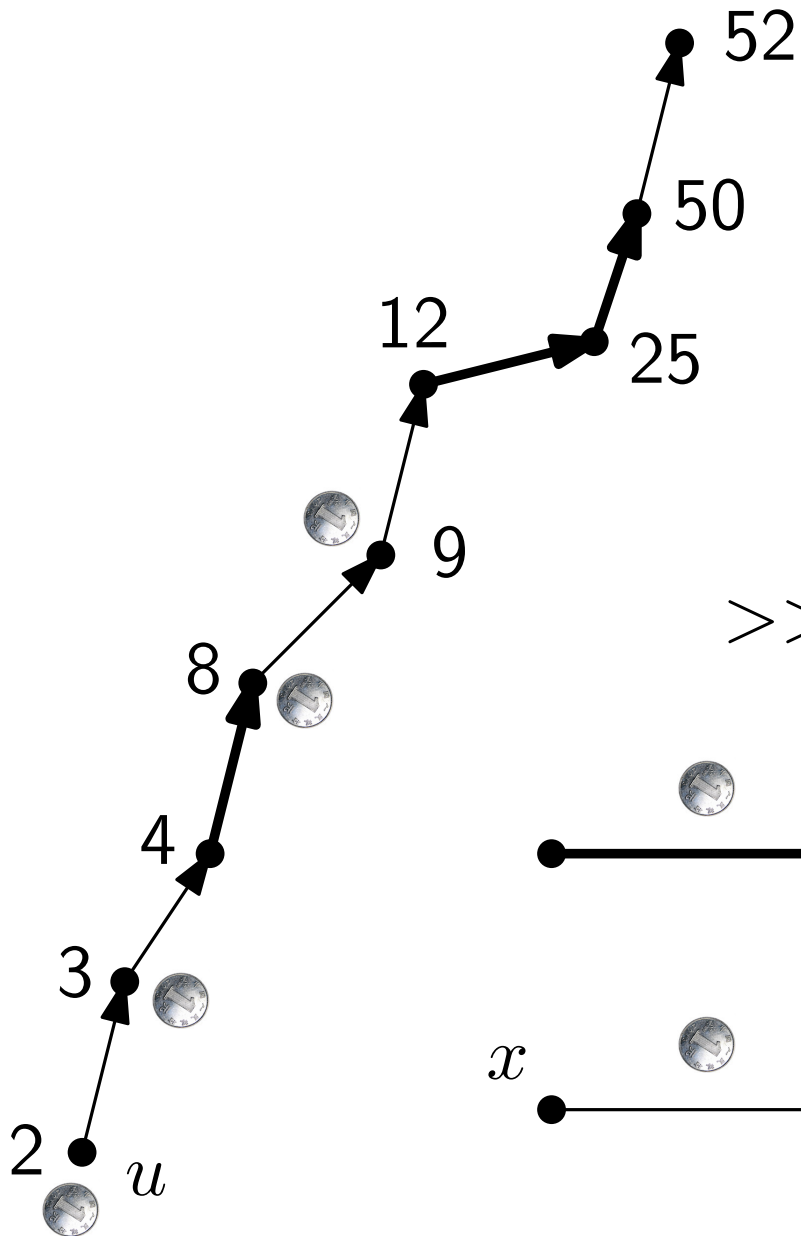
>>> find(u) 




Definition. An edge $x \longrightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$ 


 is paid by find

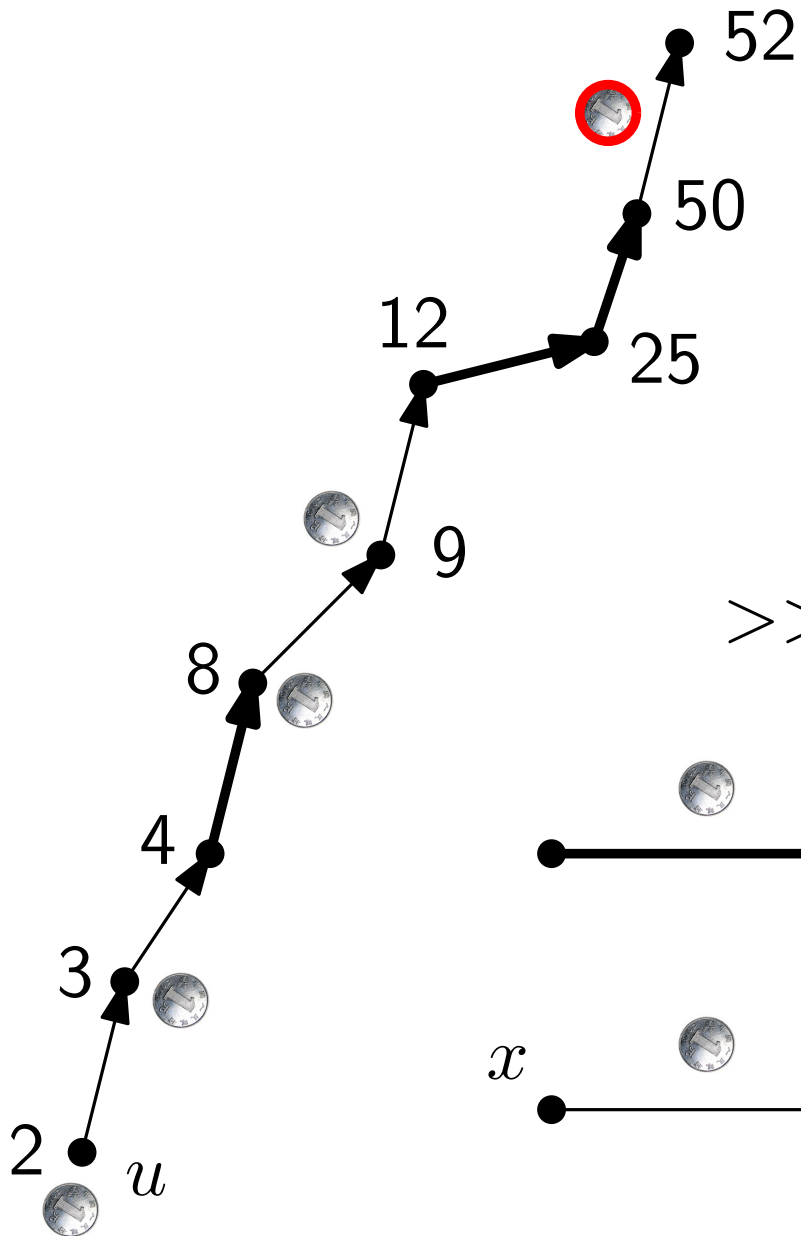


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$ 

 is paid by find

$x \rightarrow y$  is paid by x

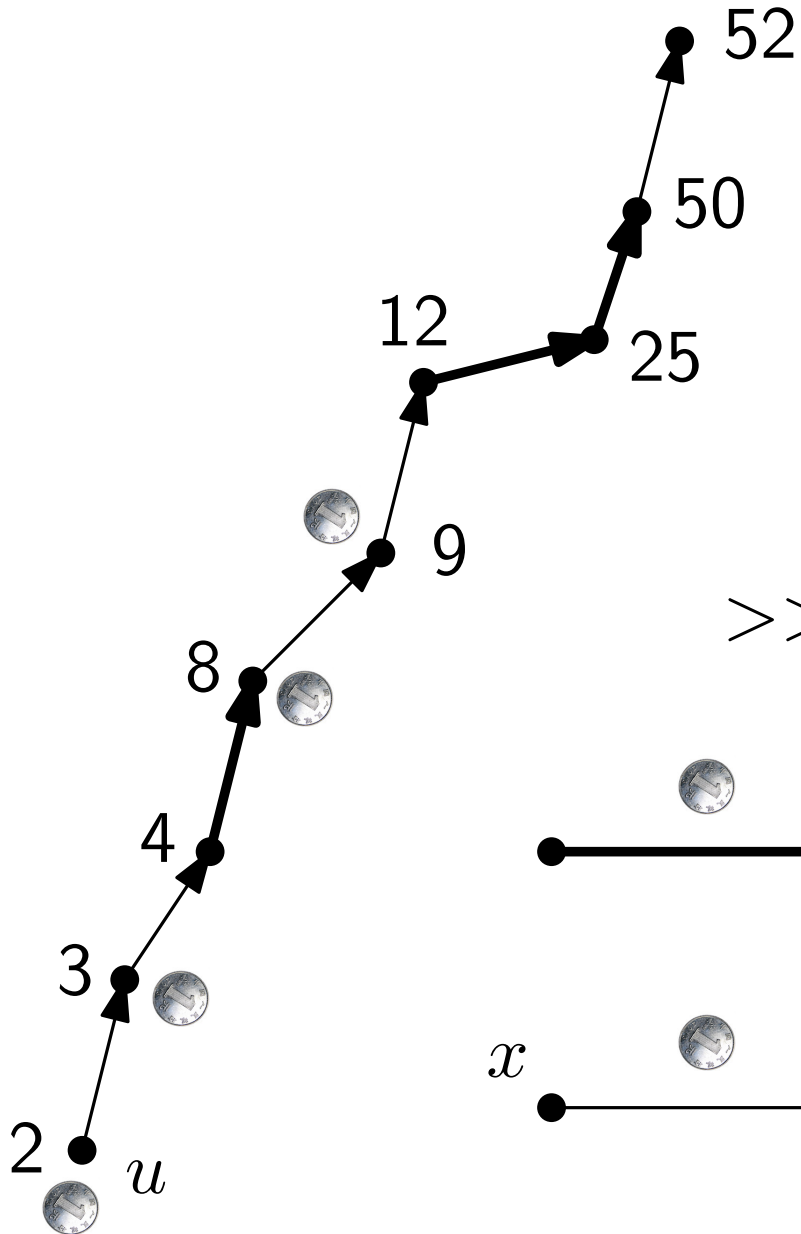


Definition. An edge $x \longrightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.


$\ggg \text{find}(u)$

is paid by find


x is paid by x

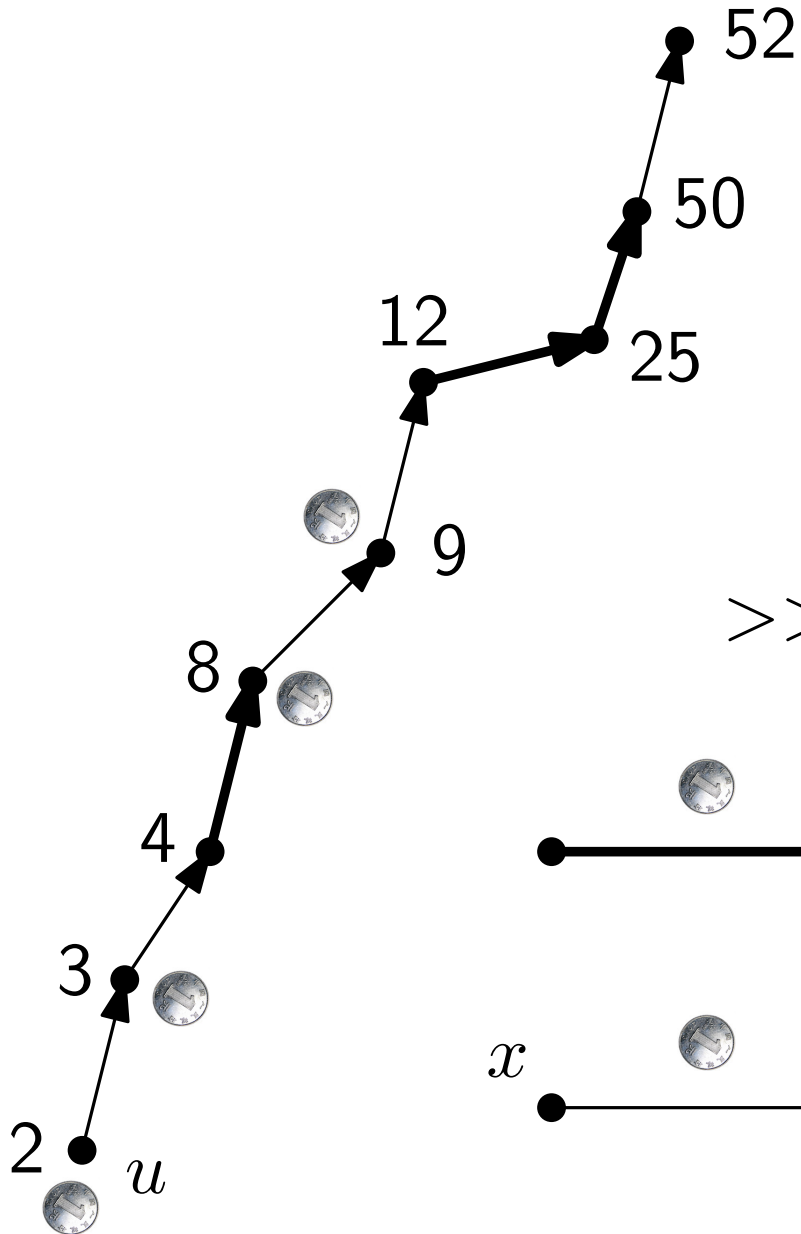


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

$\ggg \text{find}(u)$ 

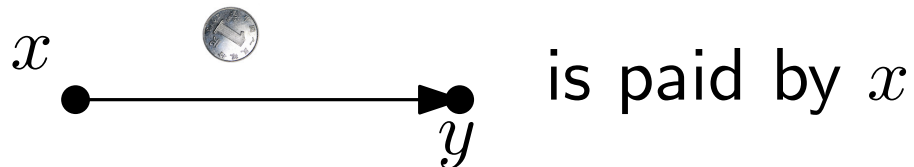
 is paid by find

$x \rightarrow y$  is paid by x

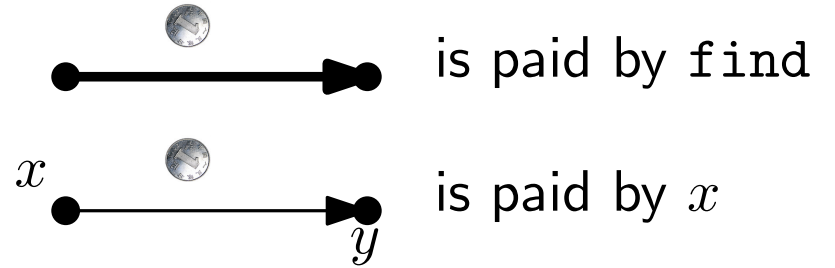


Definition. An edge $x \longrightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

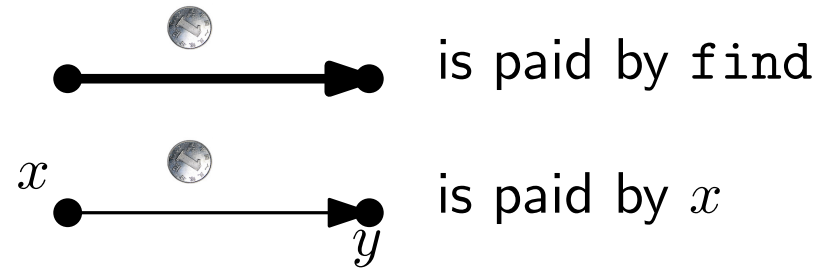
$\ggg \text{find}(u)$



Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

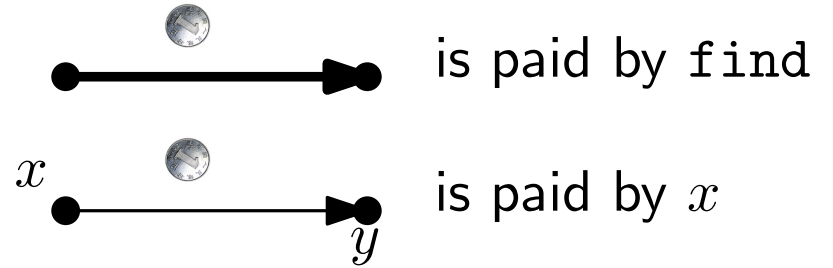


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

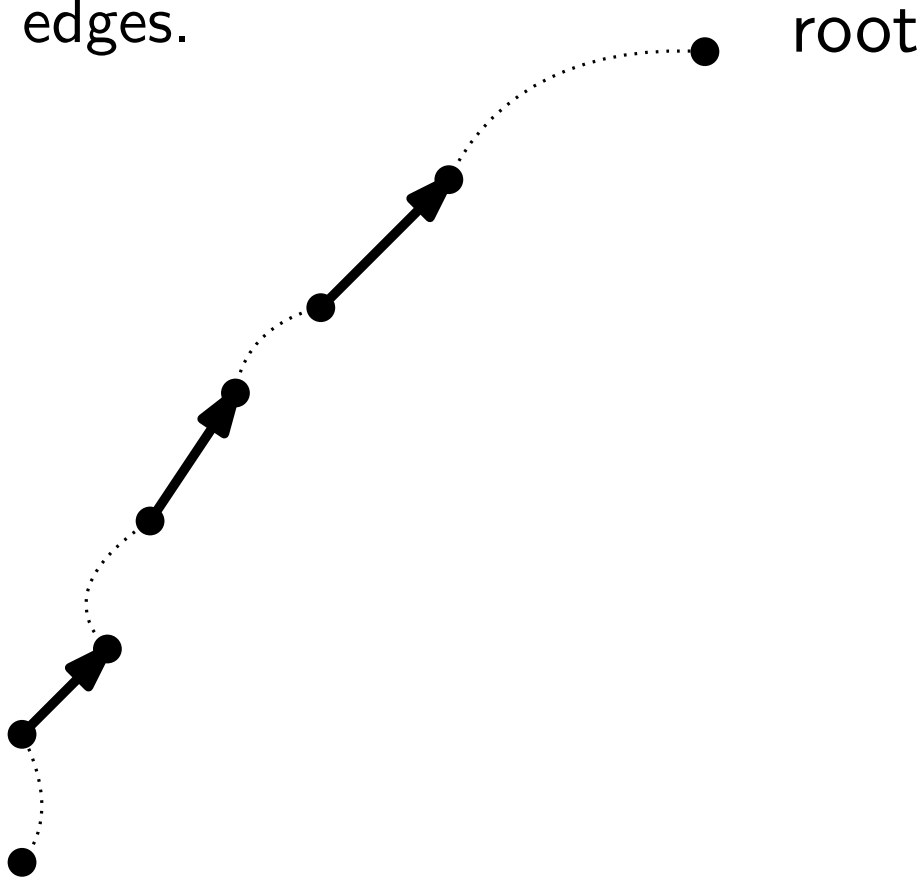


Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges.

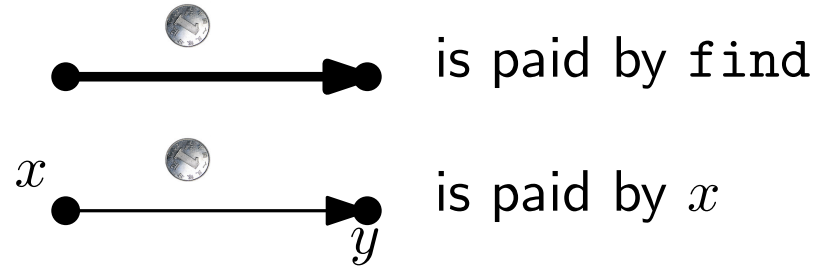
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



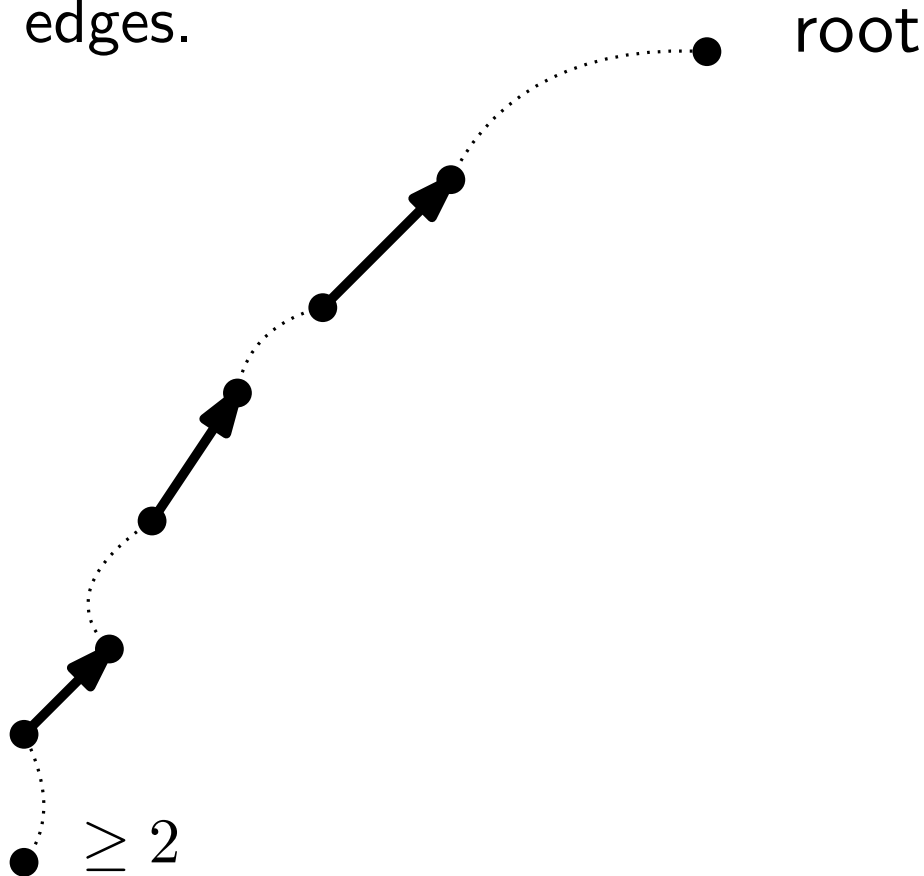
Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.



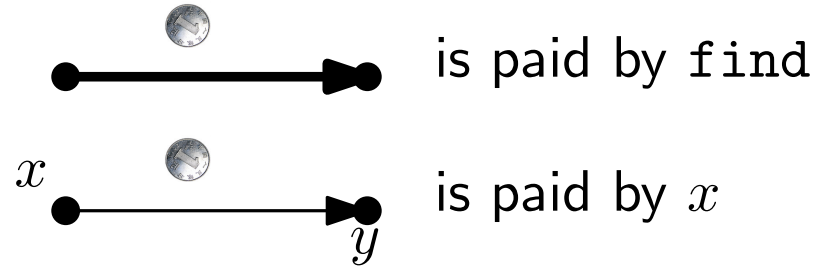
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



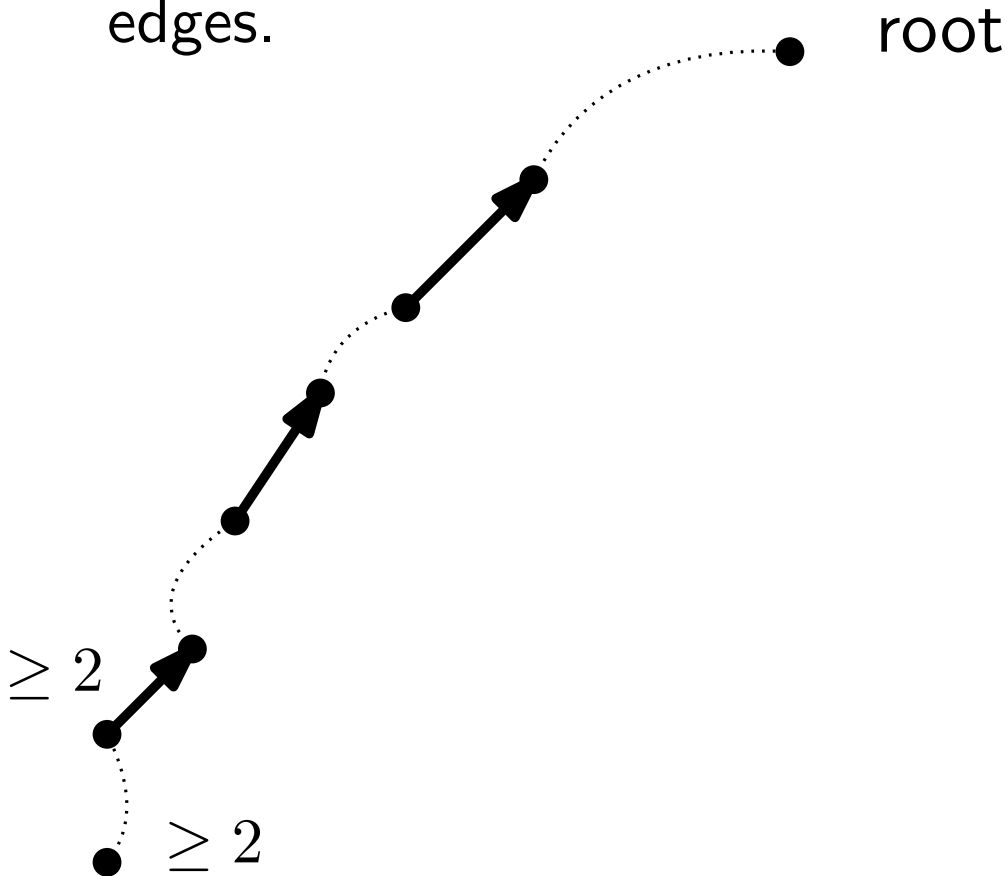
Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.



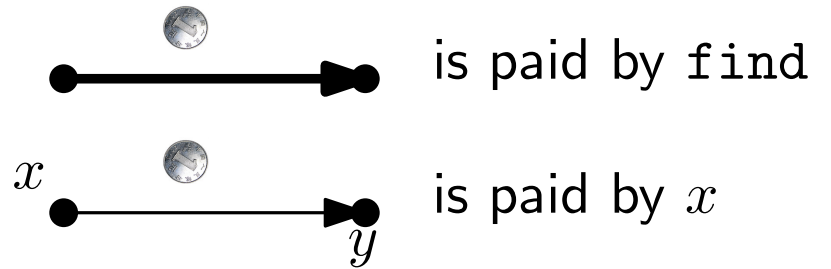
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



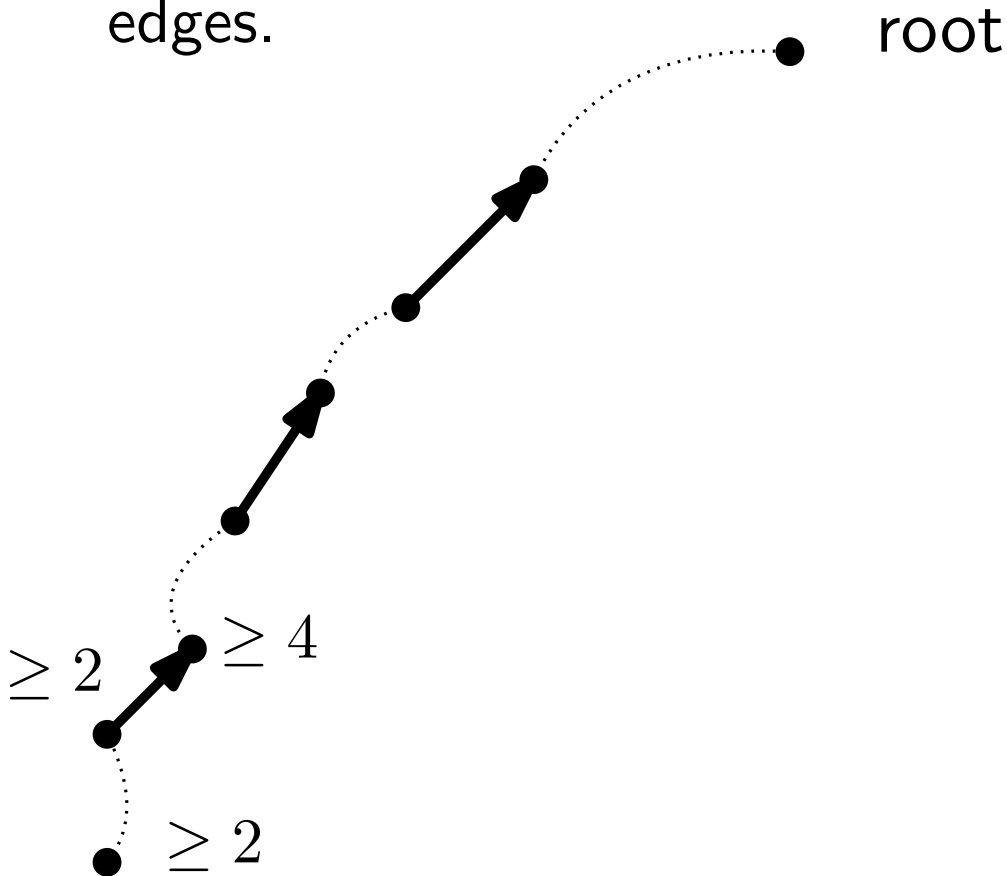
Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.



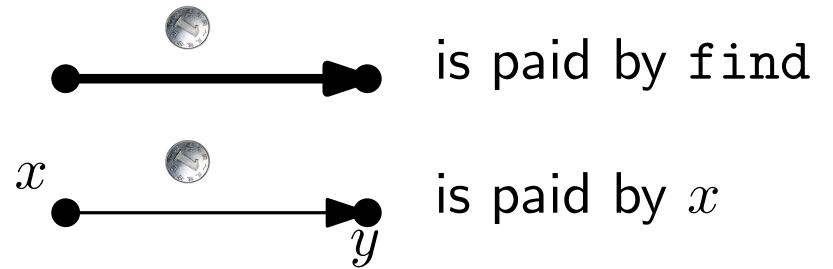
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



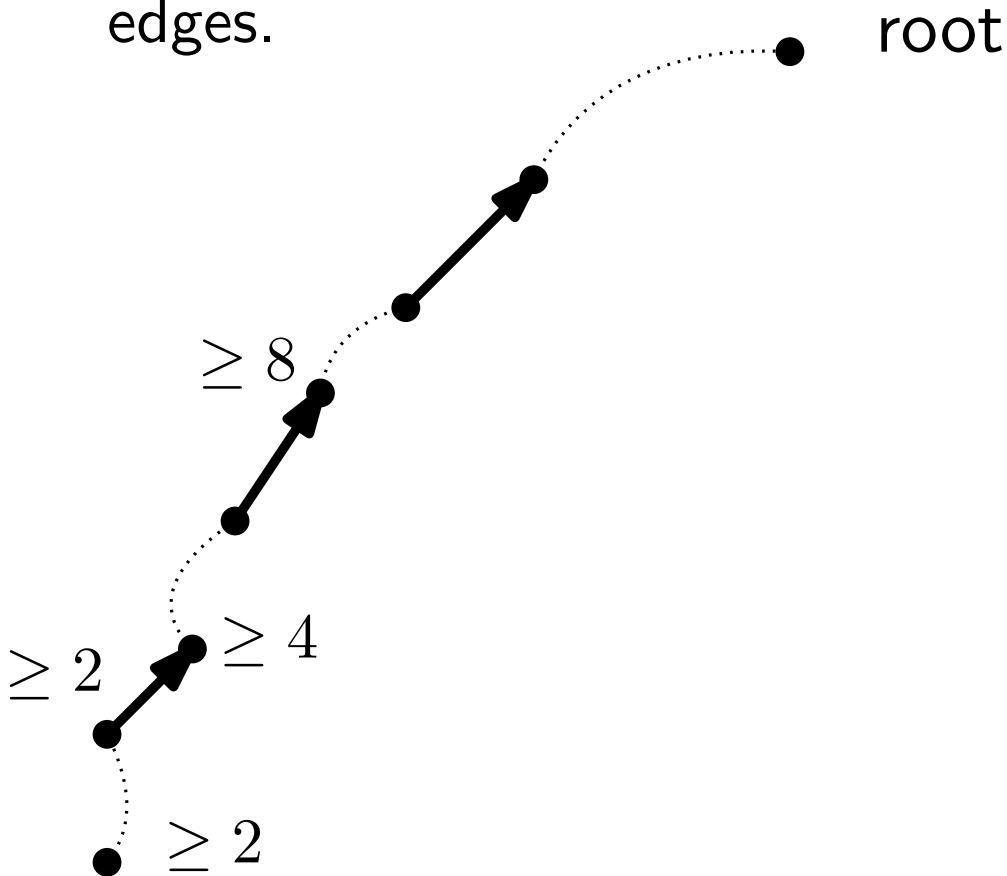
Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.



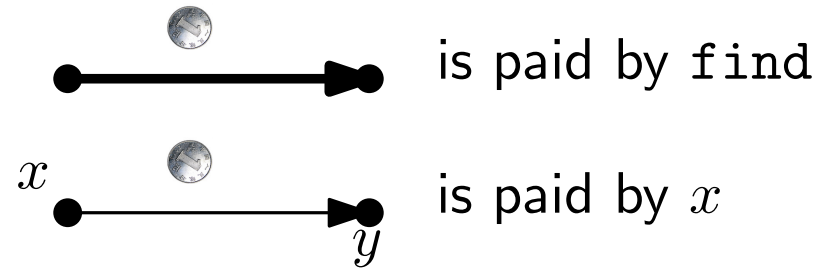
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



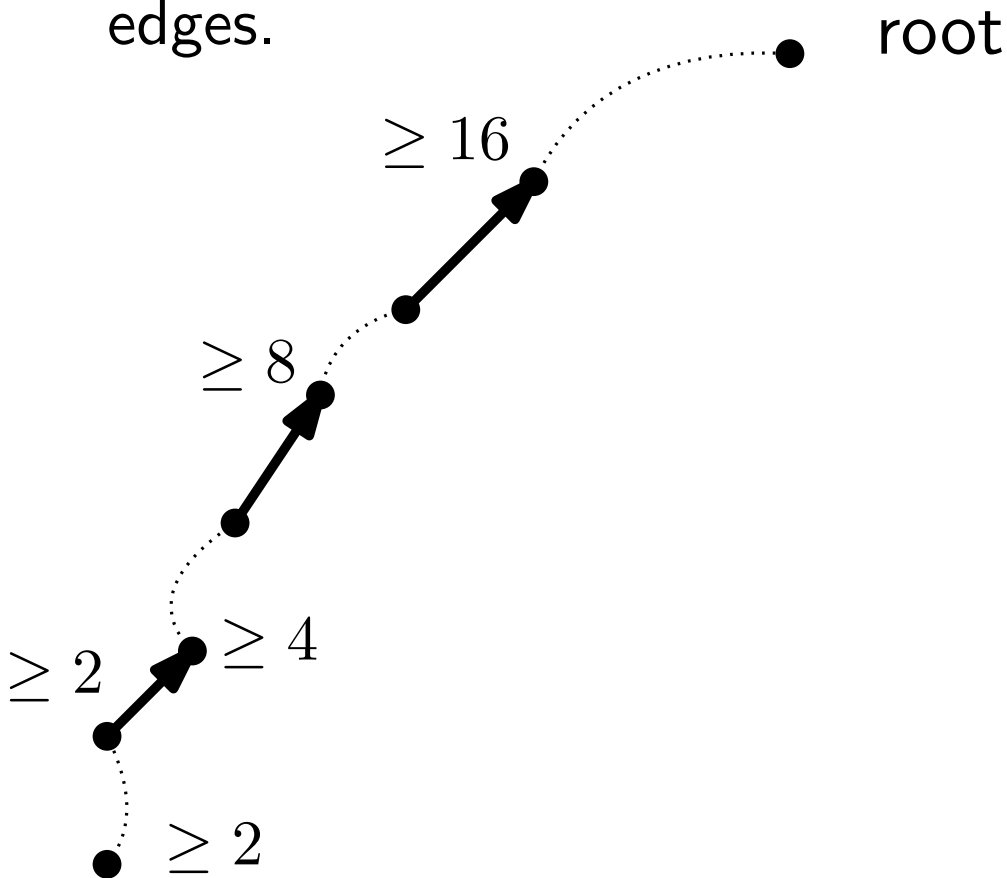
Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.



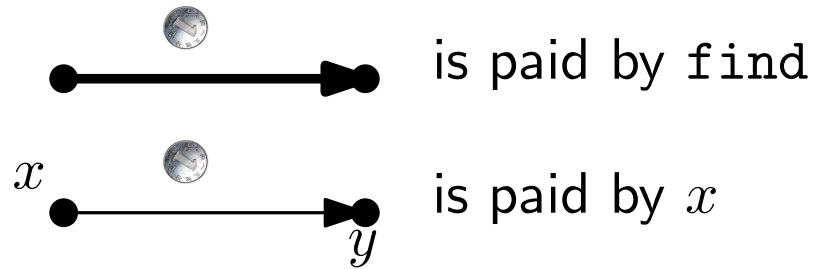
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



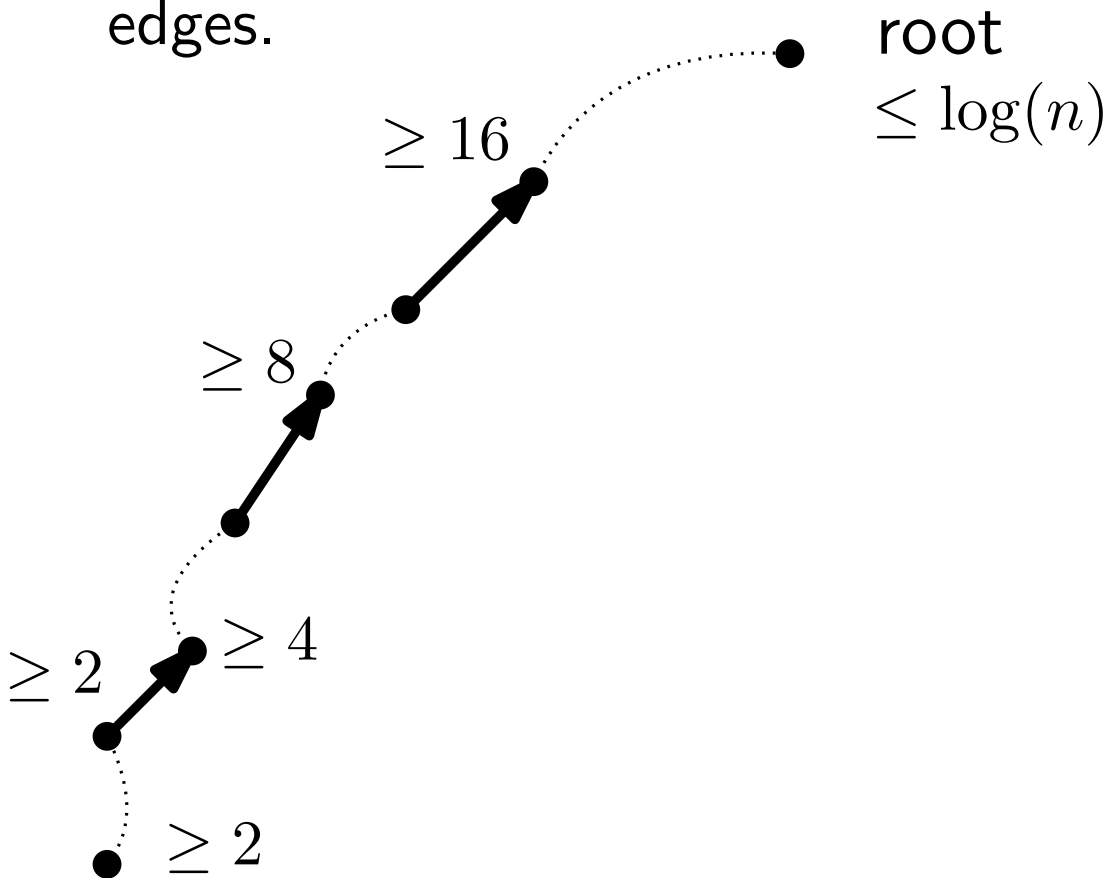
Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.



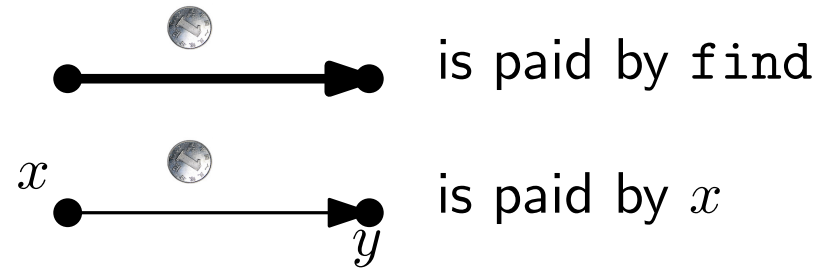
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



Lemma. Every find operation redirects at most $\log \log(n)$ thick edges.

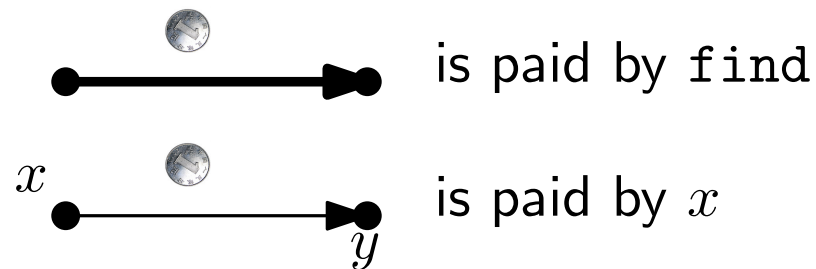



Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



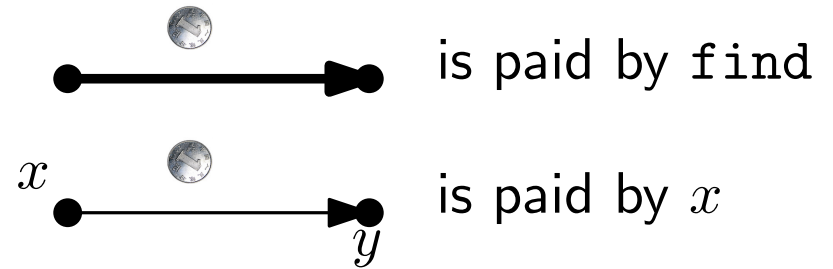
Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges.


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

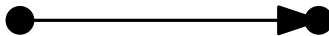


Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 

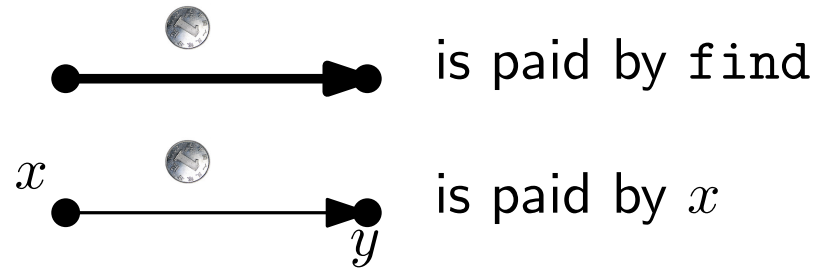
Definition. An edge $x \longrightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.




Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 

Lemma. Suppose edge $x \longrightarrow$  is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick.

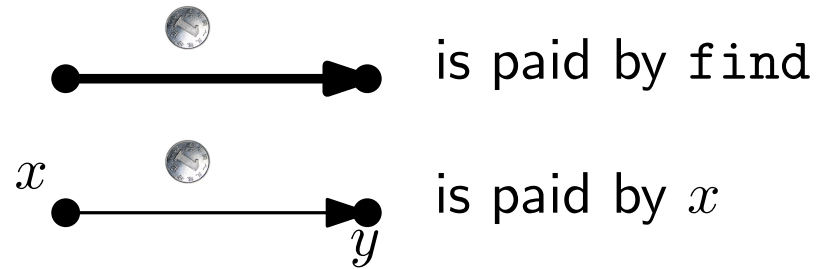
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.




Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 

Lemma. Suppose edge $x \xrightarrow{r}$ is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick.

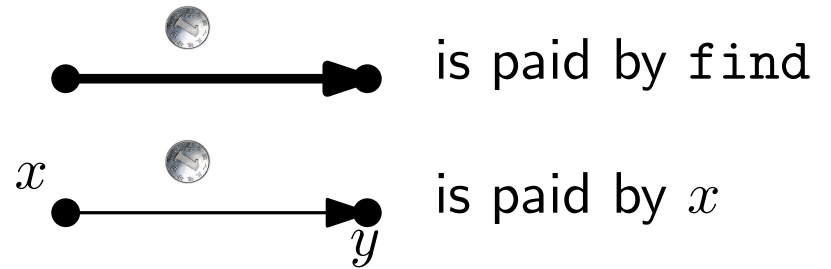
Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.




Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 

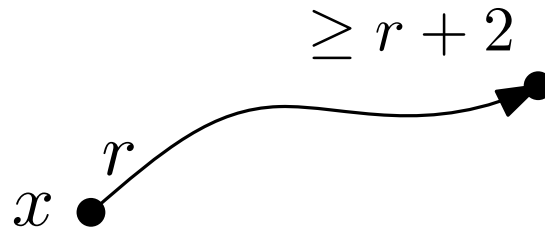
Lemma. Suppose edge $x \xrightarrow{r} \bullet$ is thin. It can be redirected at most $\geq r + 1$ $r := \text{rank}(x)$ times before it becomes thick.

Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

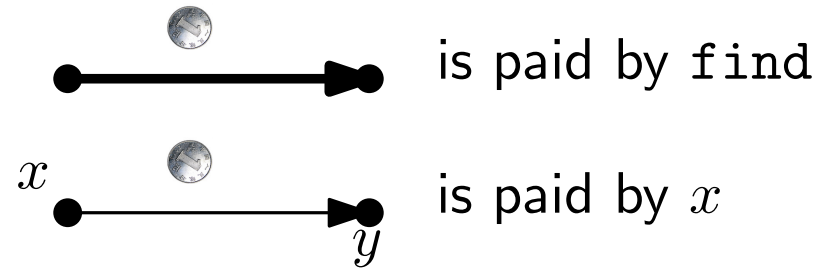



Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 

Lemma. Suppose edge $x \rightarrow y$ is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick.

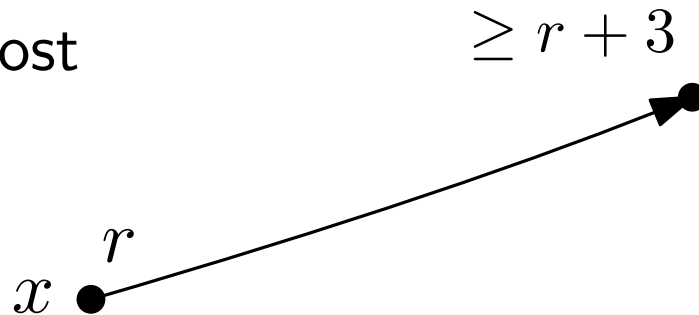


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

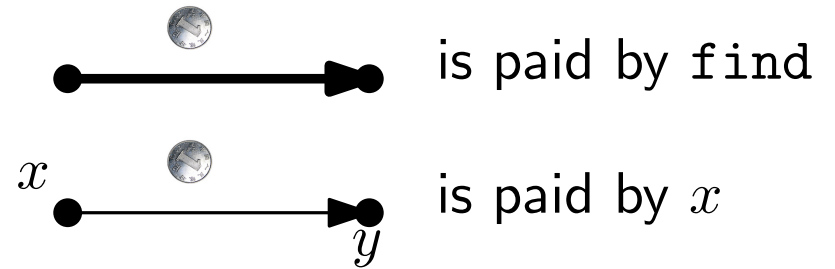



Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 

Lemma. Suppose edge $x \rightarrow y$ is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick.



Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

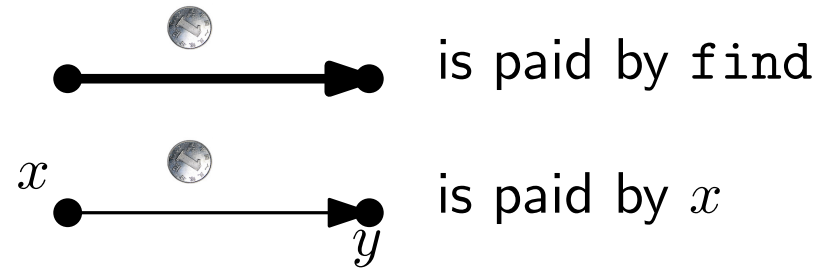



Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 

Lemma. Suppose edge $x \rightarrow y$ is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick.



Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.

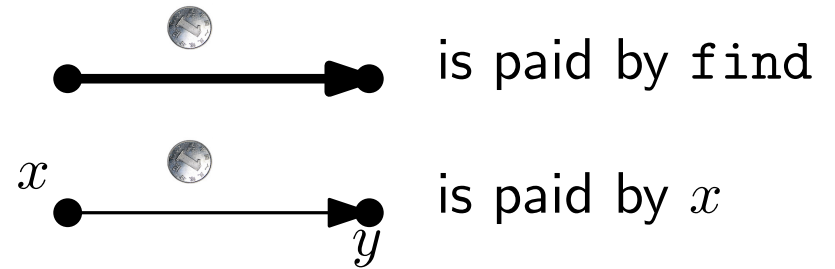



Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 


Lemma. Suppose edge $x \rightarrow y$ is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick.

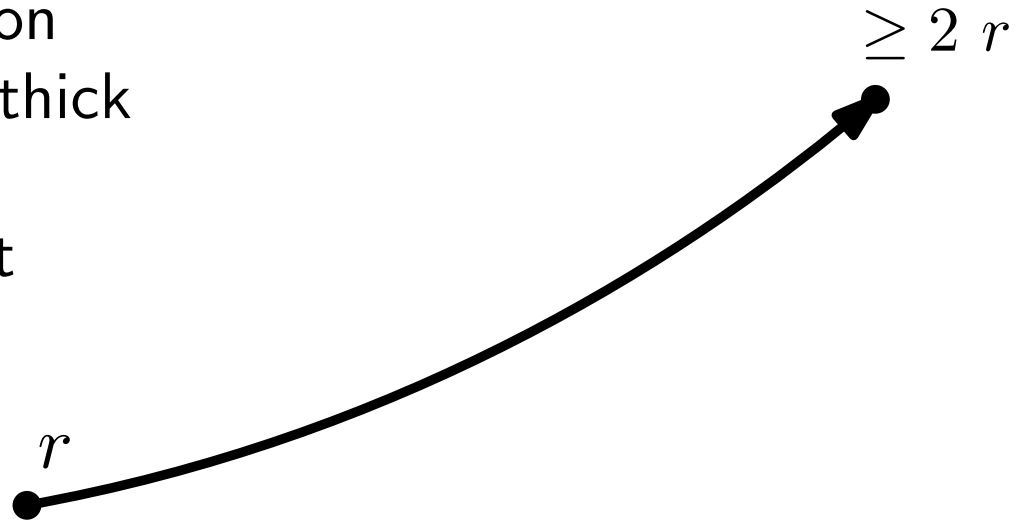


Definition. An edge $x \rightarrow y$ is *thick* if $\text{rank}(y) \geq 2 \text{rank}(x)$.



Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 

Lemma. Suppose edge $x \rightarrow y$ is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 

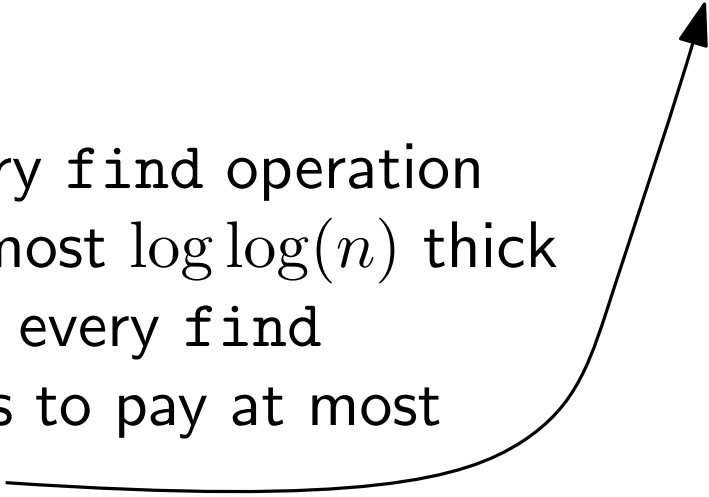


Lemma. Every `find` operation redirects at most $\log \log(n)$ thick edges. Thus, every `find` operation has to pay at most $\log \log(n)$ 🌐

Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

m find operations: $\leq m \log \log(n)$ 🌐

Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 🌐



Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

m find operations: $\leq m \log \log(n)$ 🌐

Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 🌐

🌐 paid by elements

Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

m find operations: $\leq m \log \log(n)$ 🌐

Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 🌐

Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

🌐 paid by elements
 $\leq \sum_x \text{rank}(x)$

m find operations: $\leq m \log \log(n)$ 🌐

Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 🌐

Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

🌐 paid by elements

$$\leq \sum_x \text{rank}(x)$$

$$= \sum_r r \cdot |\{\text{rank} - r - \text{elements}\}|$$

m find operations: $\leq m \log \log(n)$ 🌐

Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 🌐

Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

🌐 paid by elements

$$\leq \sum_x \text{rank}(x)$$

$$= \sum_r r \cdot |\{\text{rank} - r - \text{elements}\}|$$

$$\leq \sum_r r \cdot \frac{n}{2^{r-2}}$$

m find operations: $\leq m \log \log(n)$ 🌐

Lemma. Every find operation redirects at most $\log \log(n)$ thick edges. Thus, every find operation has to pay at most $\log \log(n)$ 🌐

Lemma. Suppose edge is thin. It can be redirected at most $r := \text{rank}(x)$ times before it becomes thick. Thus, x has to pay at most $\text{rank}(x)$ 🌐

🌐 paid by elements

$$\leq \sum_x \text{rank}(x)$$

$$= \sum_r r \cdot |\{\text{rank}-r\text{-elements}\}|$$

$$\leq \sum_r r \cdot \frac{n}{2^{r-2}} = 6n$$

Union-by-Rank: $O(n + m \log n)$

Union-by-Rank: $O(n + m \log n)$

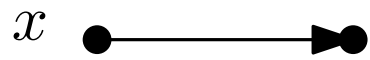
Union-by-Rank with path compression: $O(n + m \log \log n)$

Union-by-Rank: $O(n + m \log n)$

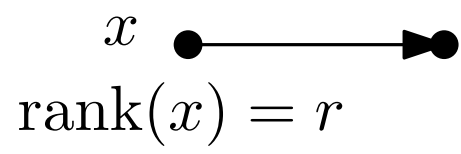
Union-by-Rank with path compression: $O(n + m \log \log n)$

Even Better Analysis

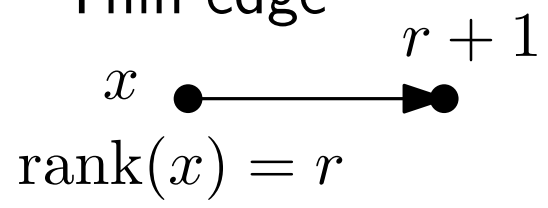
Thin edge

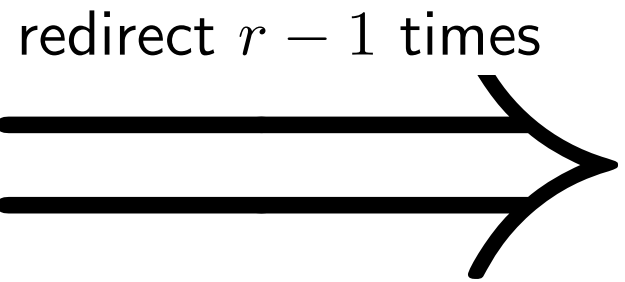
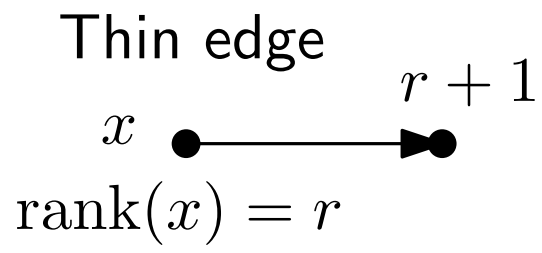


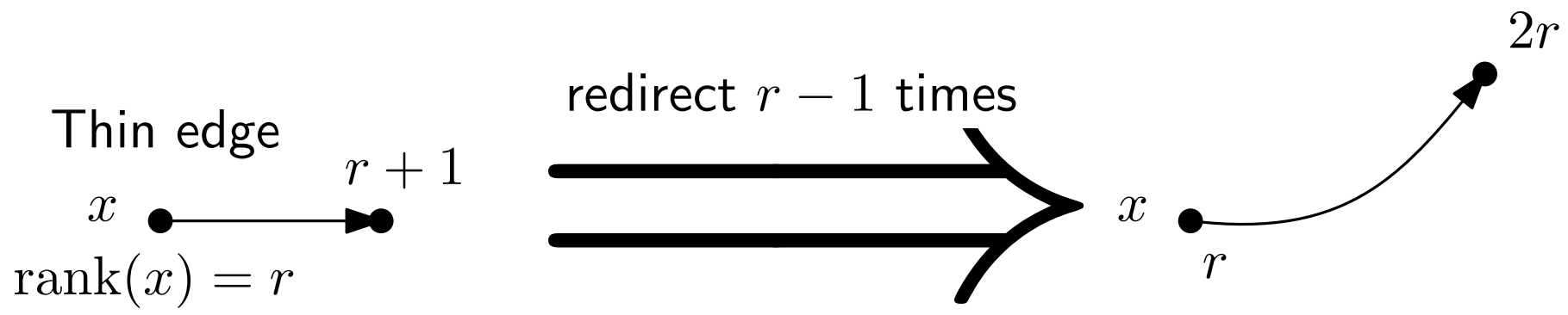
Thin edge

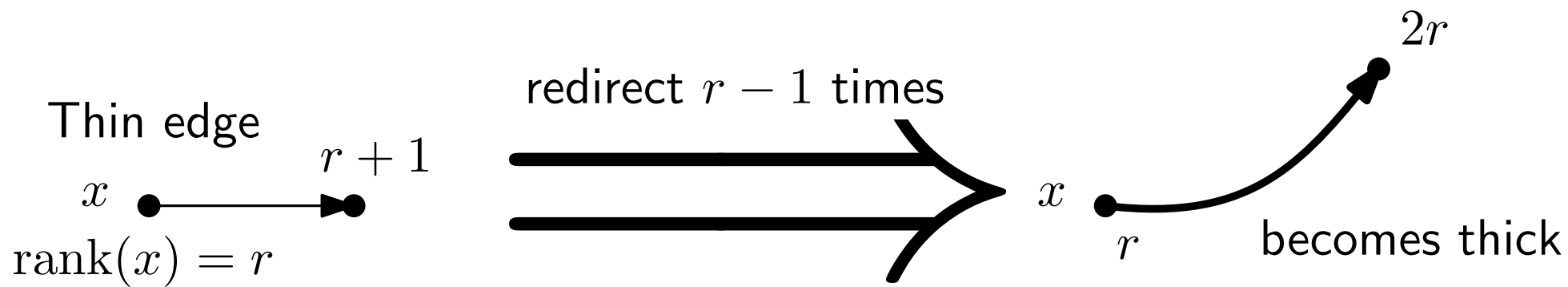


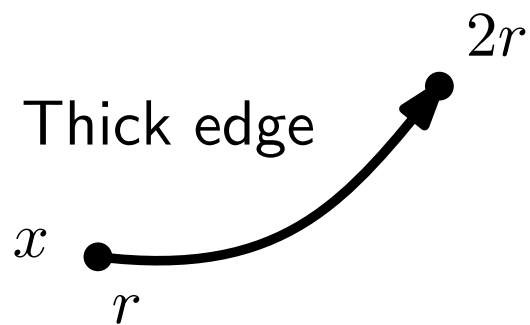
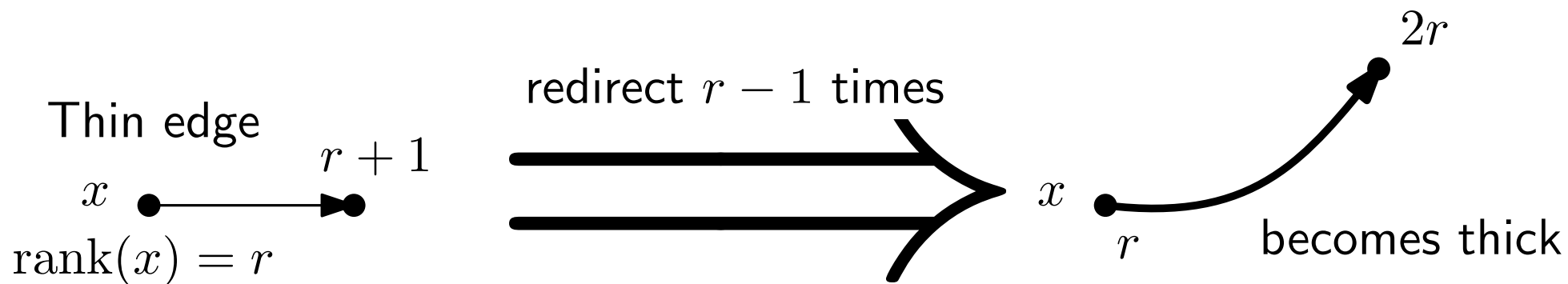
Thin edge

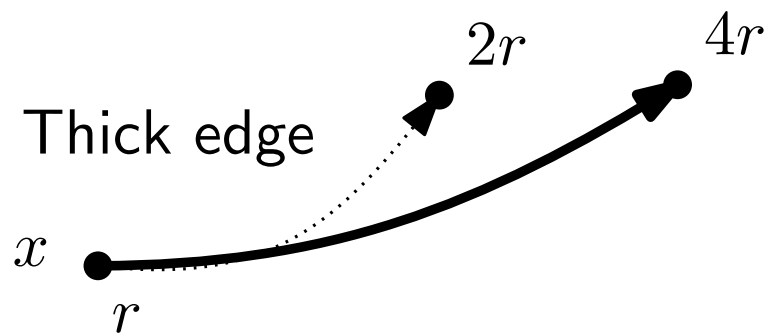
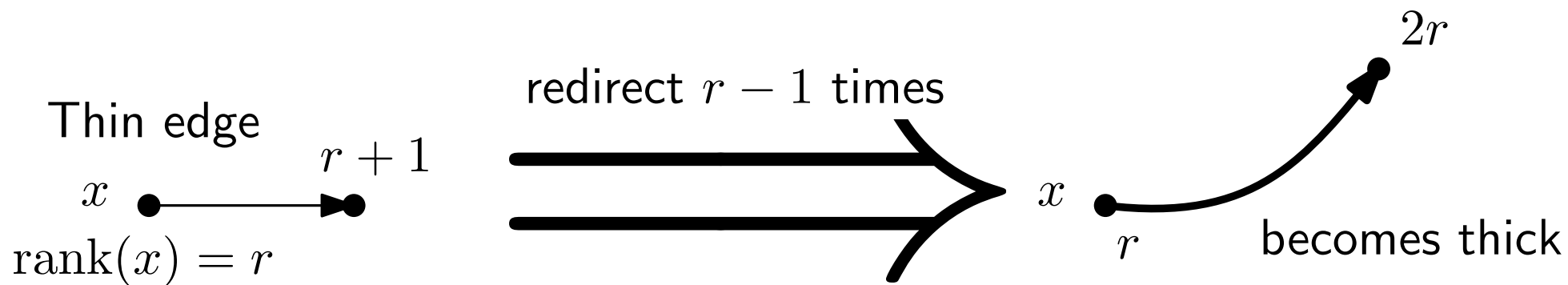


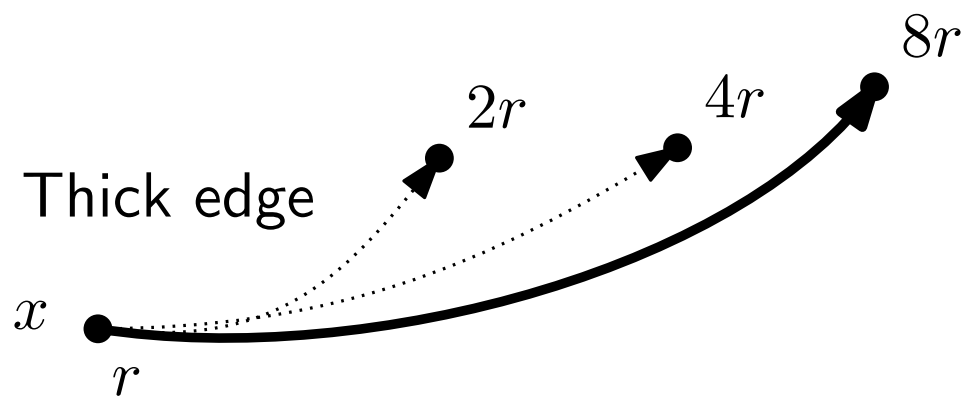
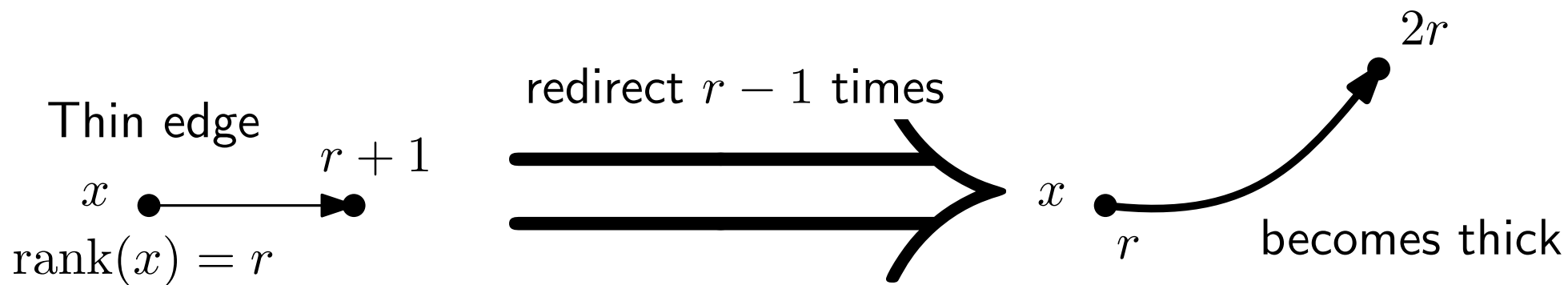


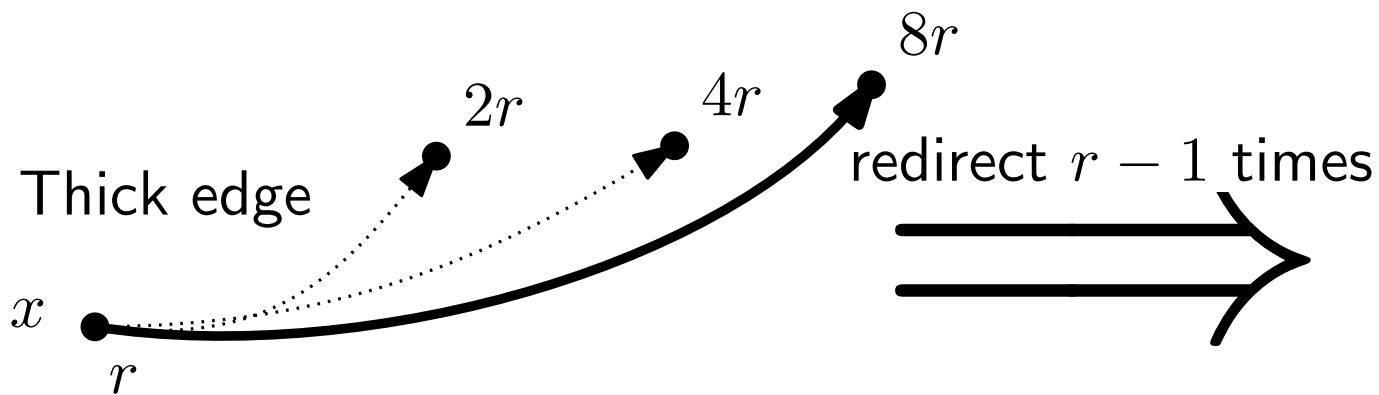
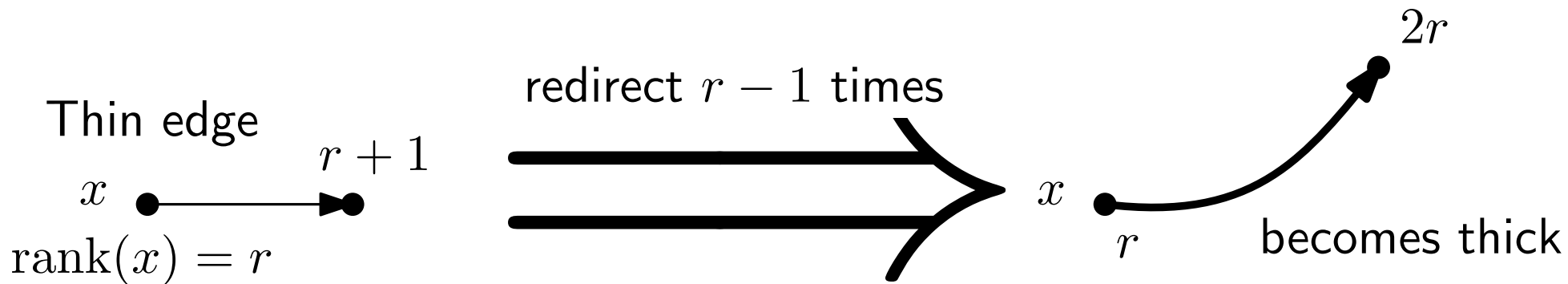


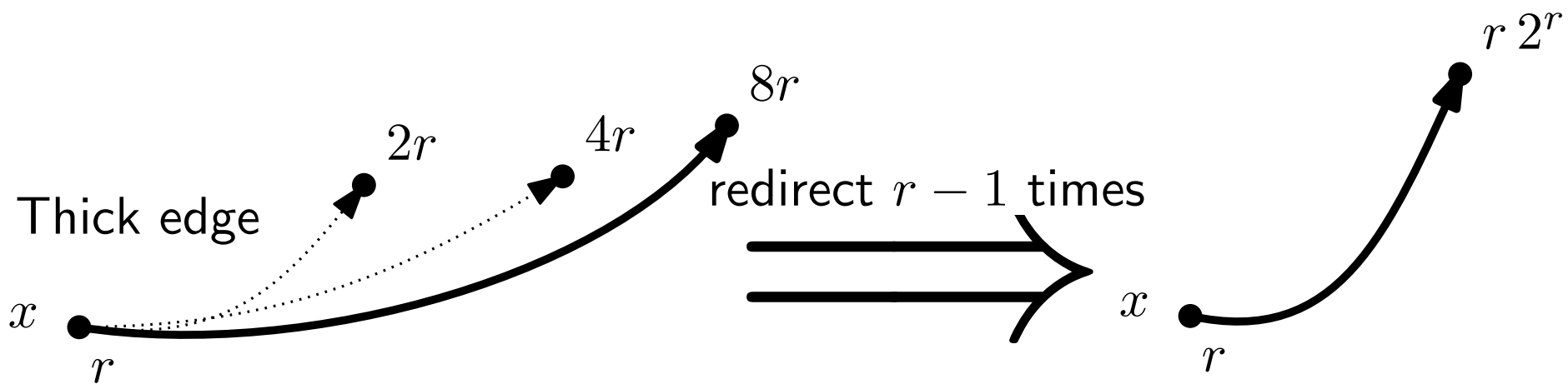
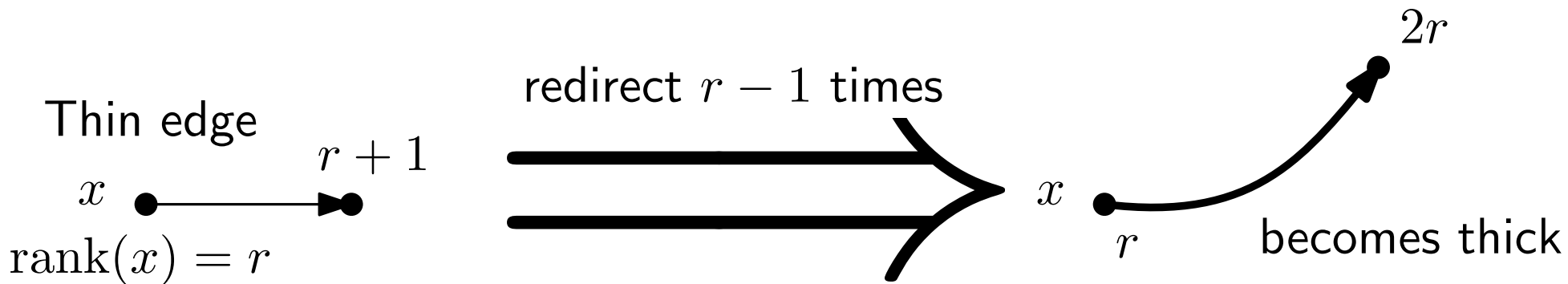


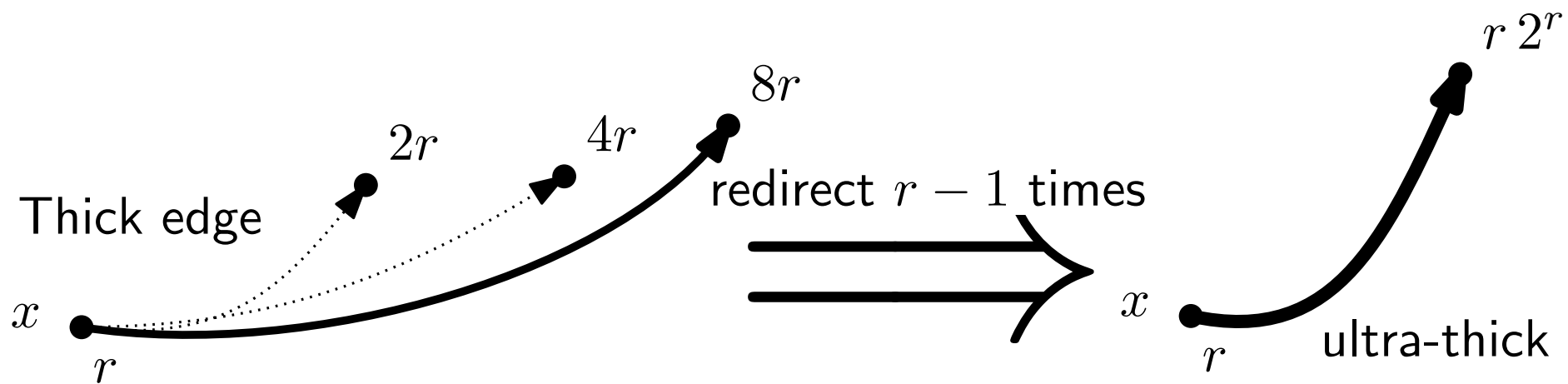
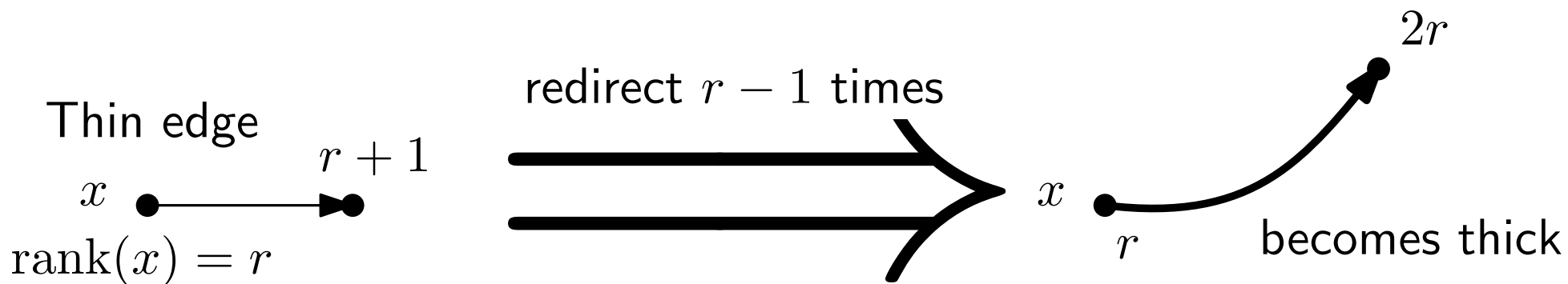


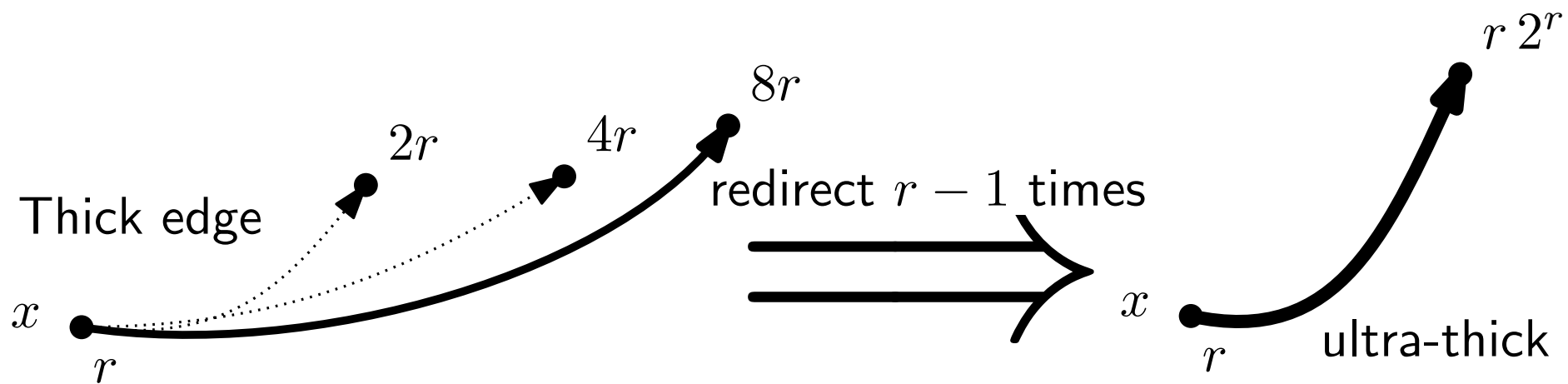
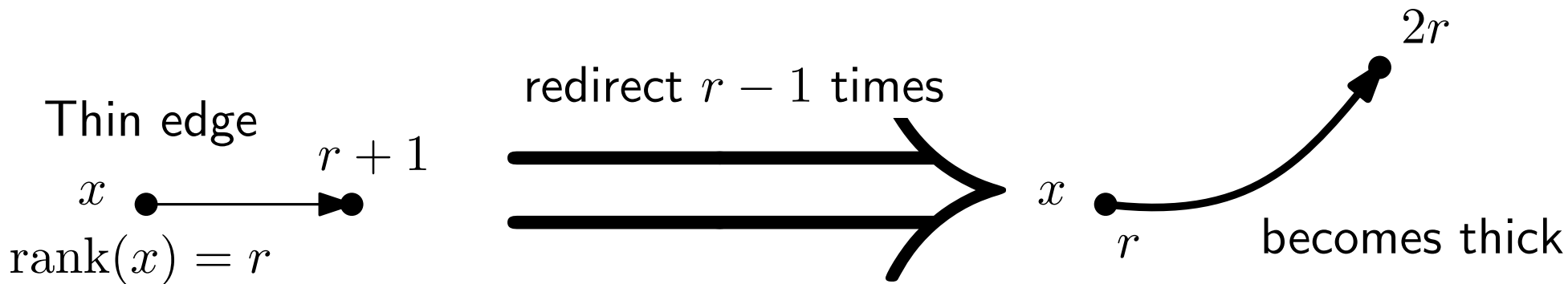




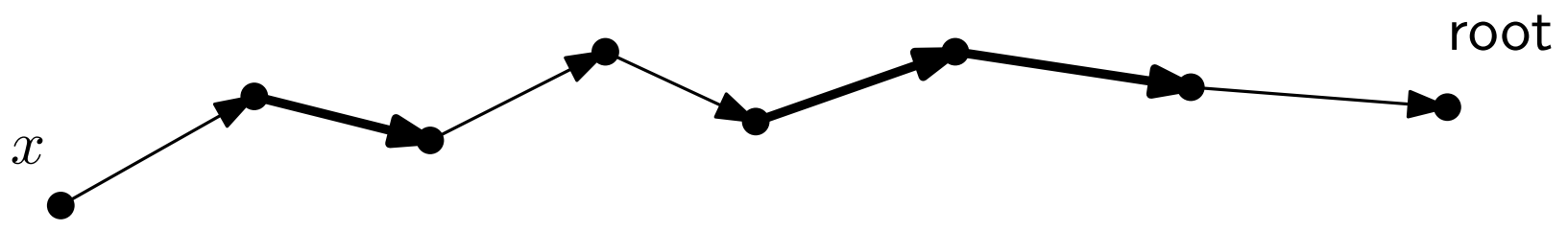


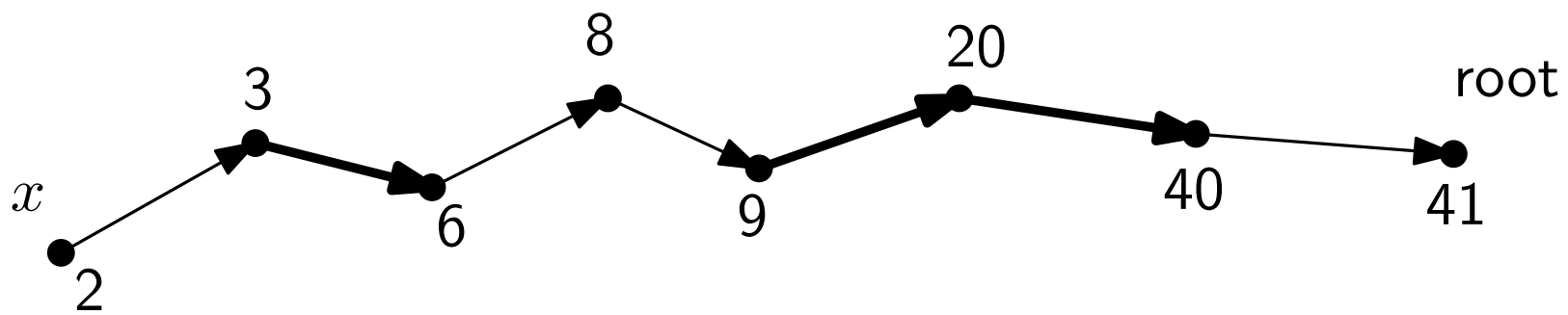


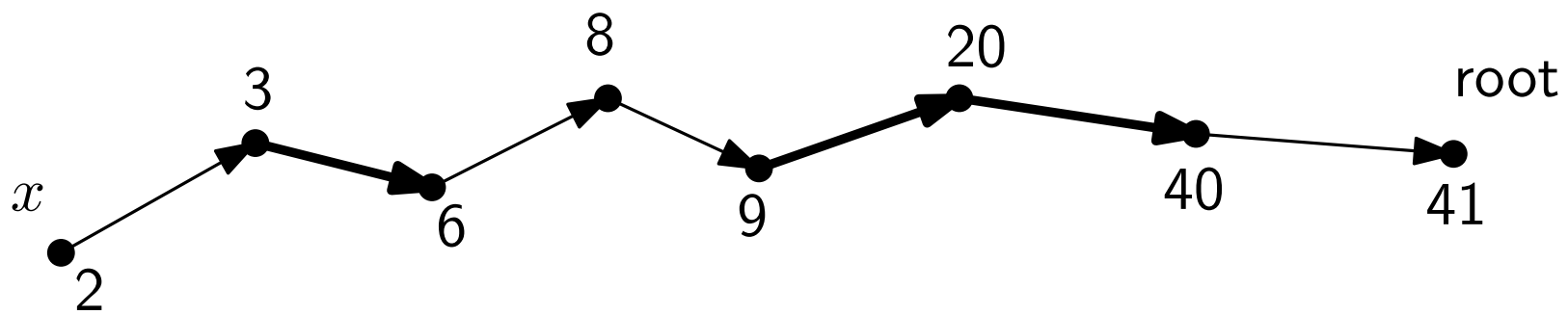




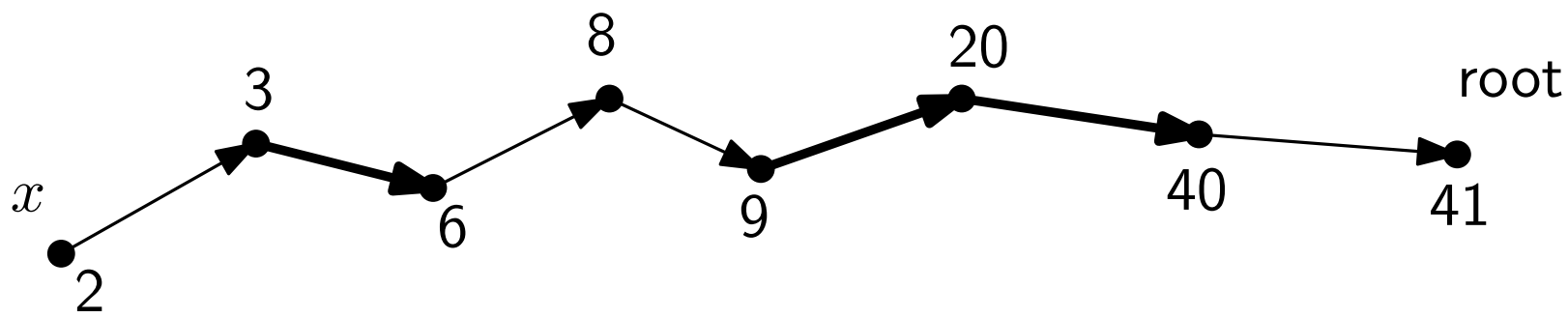
Spot the error!



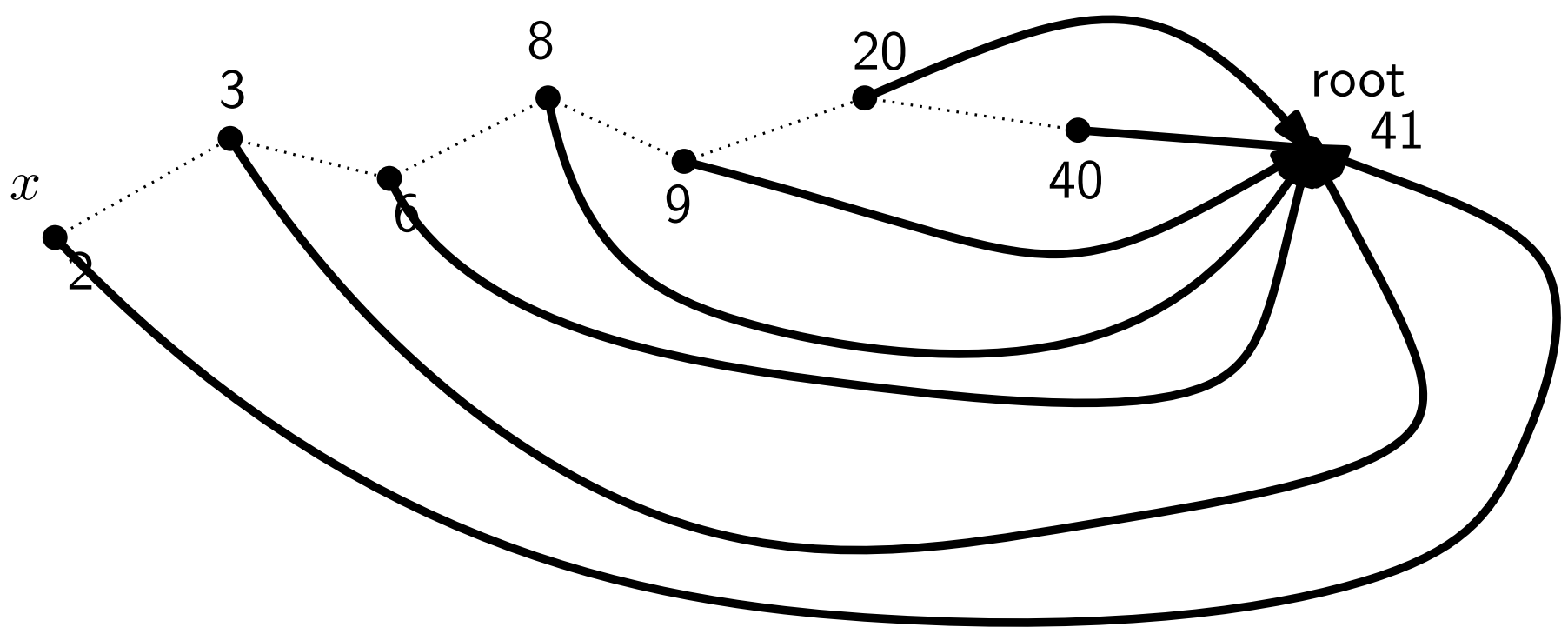


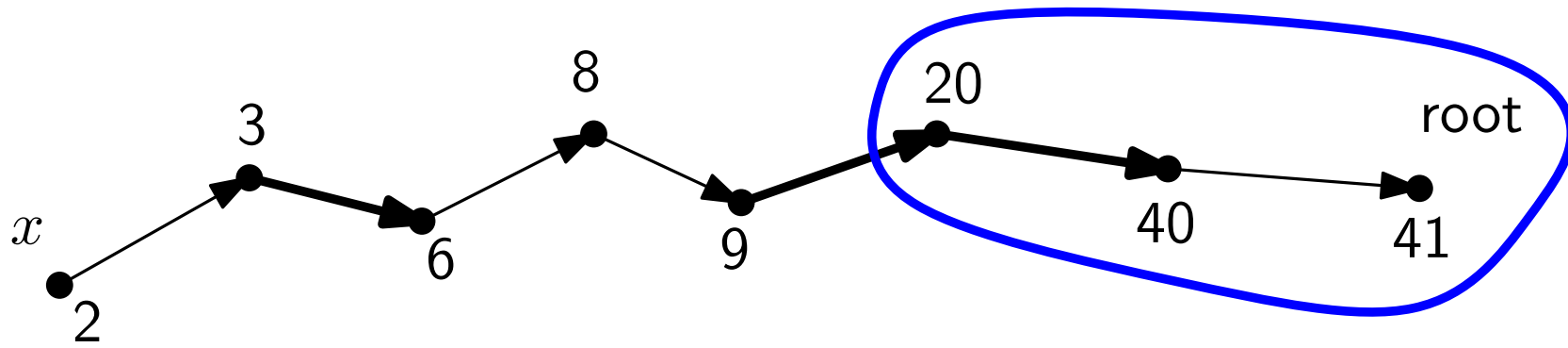


>>> find(x)

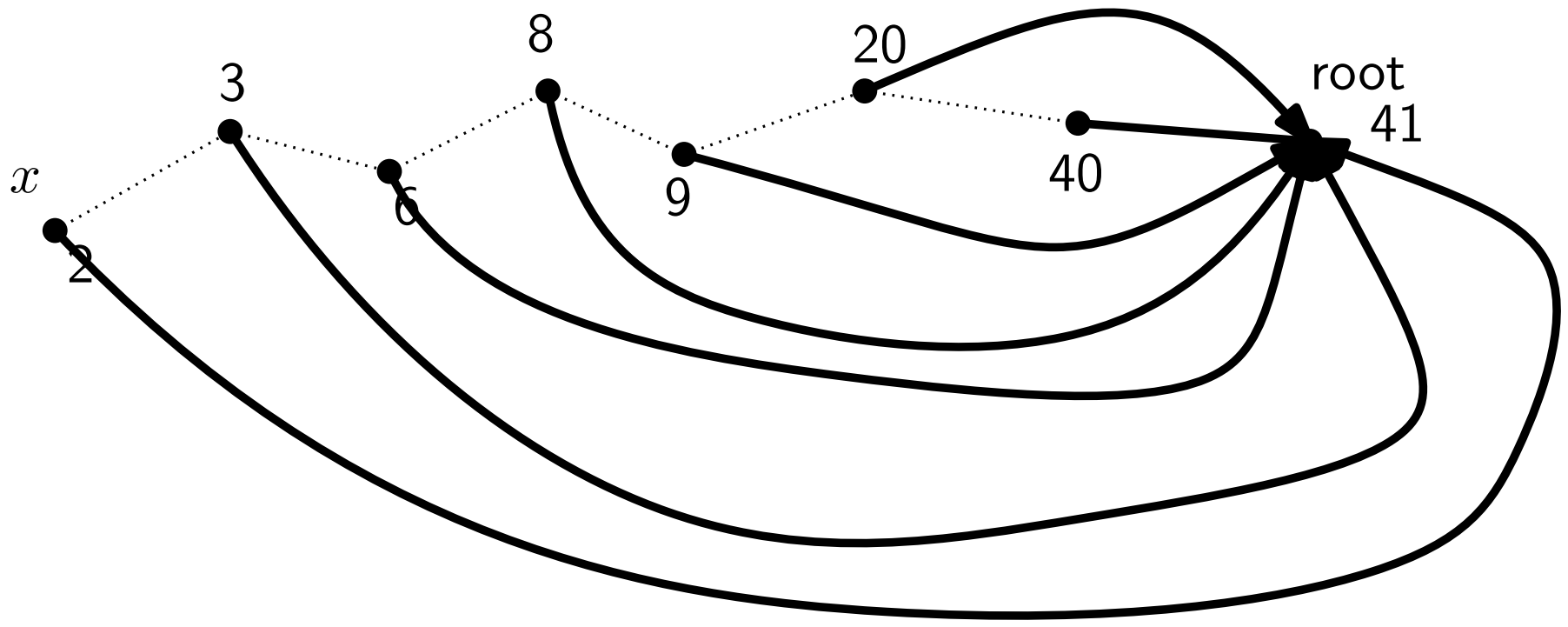


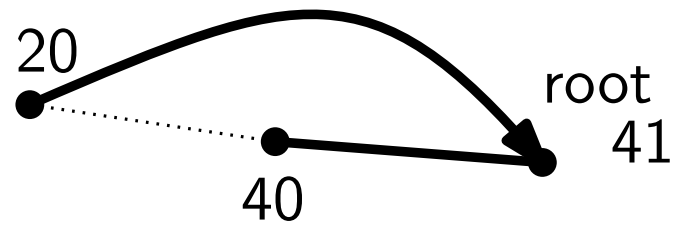
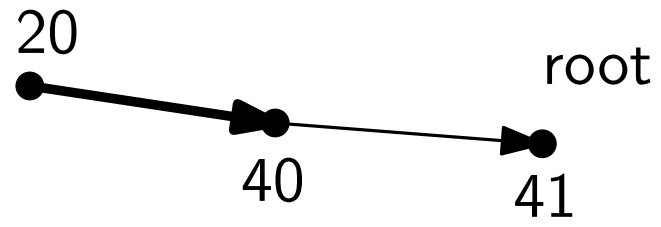
>>> find(x)

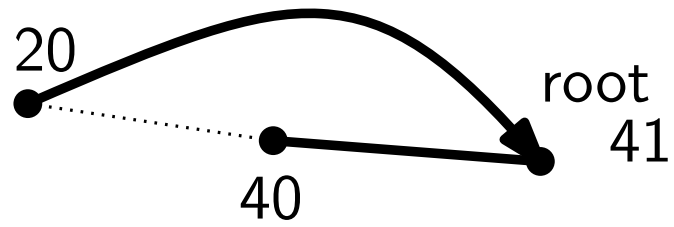
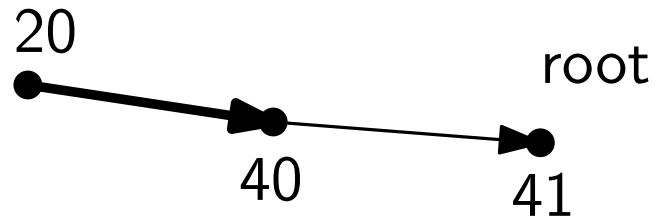




>>> find(x)

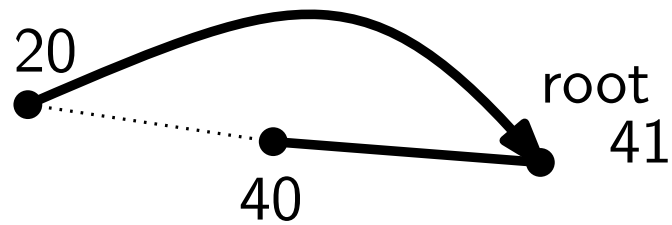
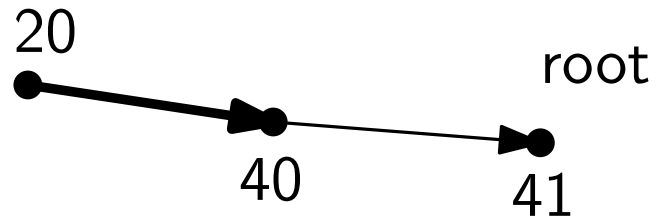




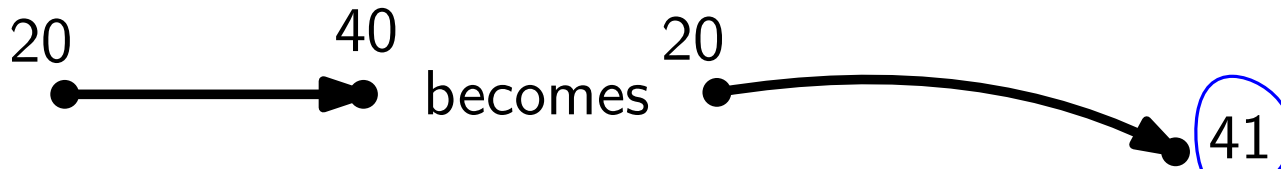


Problem:

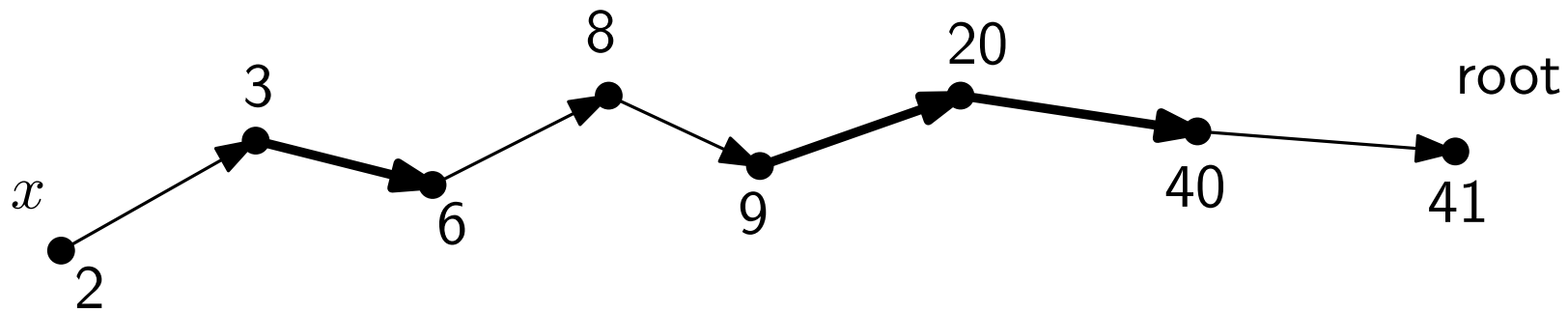




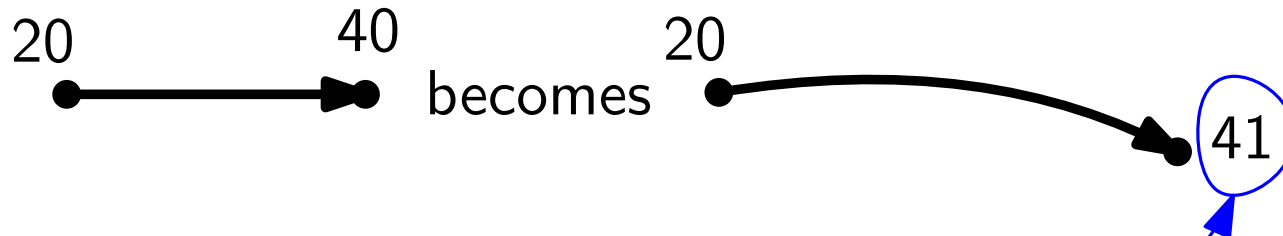
Problem:



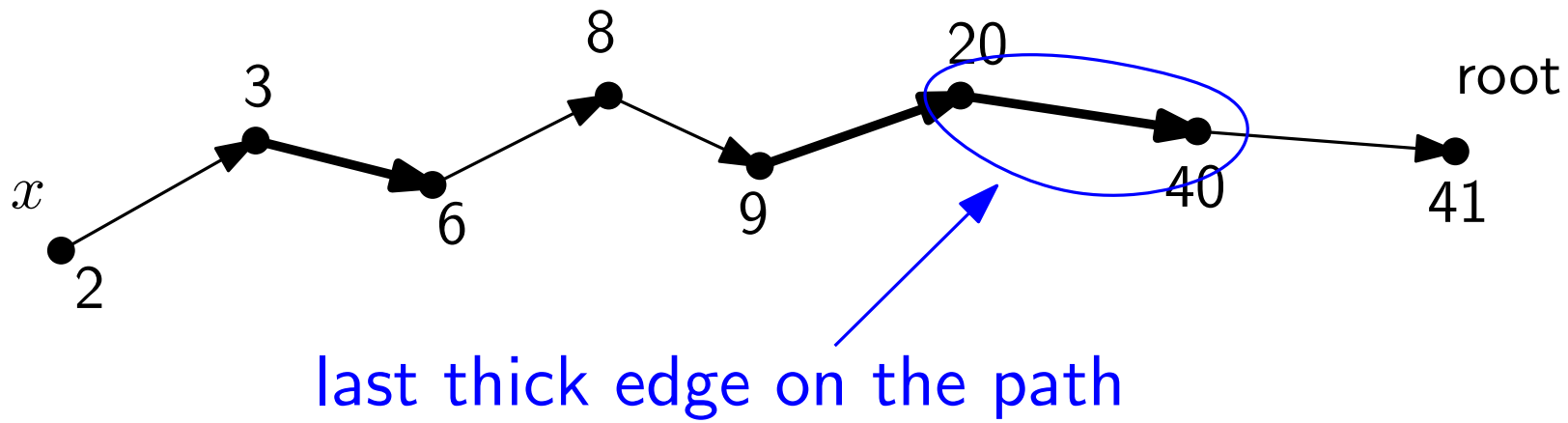
But I claimed this would be at least 80



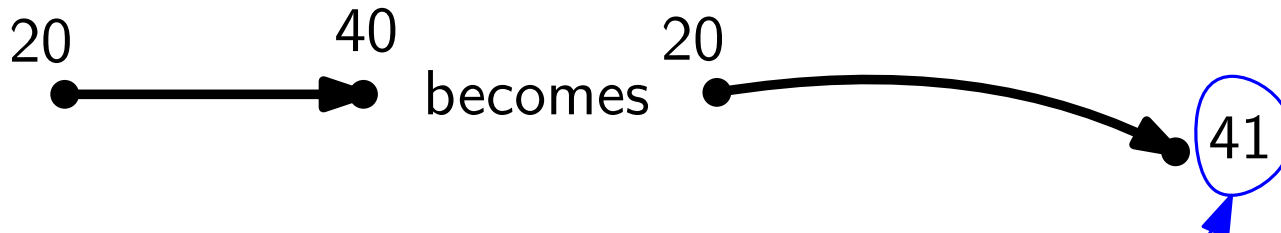
Problem:



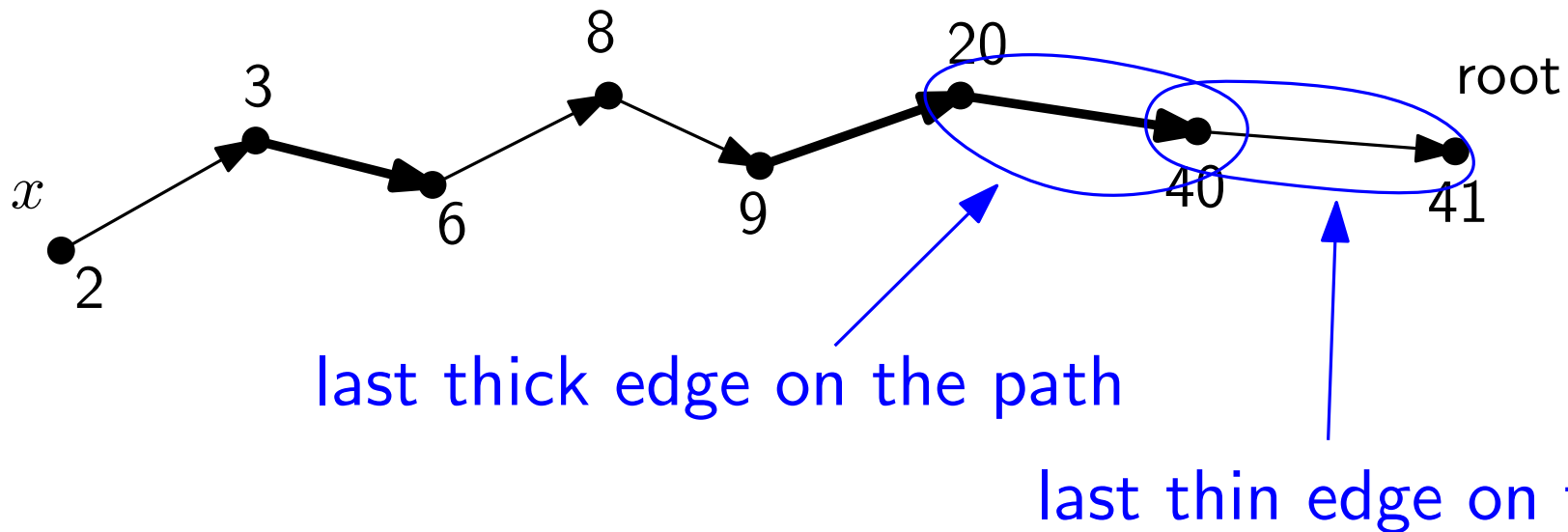
But I claimed this would be at least 80



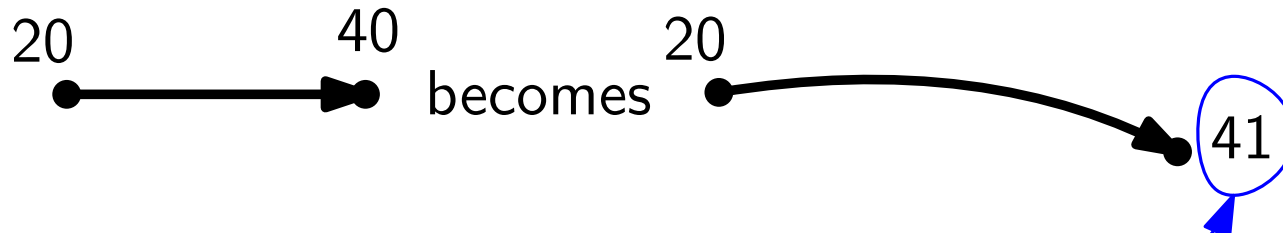
Problem:



But I claimed this would be at least 80



Problem:



But I claimed this would be at least 80

Let's summarize and generalize

Let's summarize and generalize

Different thickness types:

Let's summarize and generalize



Let's summarize and generalize

Different thickness types: 

$r + 1 \leq s < 2r$: thickness 1

Let's summarize and generalize

Different thickness types: 

$r + 1 \leq s < 2r$: thickness 1

$2r \leq s < r 2^r$: thickness 2

Let's summarize and generalize

Different thickness types: 

$r + 1 \leq s < 2r$: thickness 1

$2r \leq s < r 2^r$: thickness 2

$r 2^r \leq s$: thickness 3

Let's summarize and generalize

Different thickness types: 

$r + 1 \leq s < 2r$: thickness 1

$2r \leq s < r 2^r$: thickness 2

$r 2^r \leq s$: thickness 3

Why stop here?

Let's summarize and generalize

Different thickness types: 

$f_1(r)$

$r + 1 \leq s < 2r$: thickness 1

$2r \leq s < r 2^r$: thickness 2

$r 2^r \leq s$: thickness 3

Why stop here?

Let's summarize and generalize

Different thickness types: 

$f_1(r)$

$r + 1 \leq s < 2r$: thickness 1

$f_2(r)$ $2r \leq s < r 2^r$: thickness 2

$r 2^r \leq s$: thickness 3

Why stop here?

Let's summarize and generalize

Different thickness types: 

$f_1(r)$

$r + 1 \leq s < 2r$: thickness 1

$f_2(r)$ $2r \leq s < r 2^r$: thickness 2

$f_3(r)$ $r 2^r \leq s$: thickness 3

Why stop here?

Let's summarize and generalize

Different thickness types: 

$$f_1(r) \quad r + 1 \leq s < 2r: \text{ thickness 1}$$

$$f_2(r) \quad 2r \leq s < r 2^r: \text{ thickness 2}$$

$$f_3(r) \quad r 2^r \leq s < f_4(r): \text{ thickness 3}$$

Let's summarize and generalize

Different thickness types: 

$f_1(r)$

$r + 1 \leq s < 2r$: thickness 1

$f_2(r)$ $2r \leq s < r 2^r$: thickness 2

$f_3(r)$ $r 2^r \leq s < f_4(r)$: thickness 3

$f_4(r) \leq s < f_5(r)$: thickness 4

Let's summarize and generalize

Different thickness types: 

$f_1(r)$

$r + 1 \leq s < 2r$: thickness 1

$f_2(r)$ $2r \leq s < r 2^r$: thickness 2

$f_3(r)$ $r 2^r \leq s < f_4(r)$: thickness 3

$f_4(r) \leq s < f_5(r)$: thickness 4

⋮

Let's summarize and generalize



$f_1(r)$

$r + 1 \leq s < 2r$: thickness 1

$f_2(r) = f_1(f_1(\dots f_1(r)\dots))$

$f_2(r)$ $2r \leq s < r 2^r$: thickness 2

$f_3(r)$ $r 2^r \leq s < f_4(r)$: thickness 3

$f_4(r) \leq s < f_5(r)$: thickness 4

⋮

Let's summarize and generalize



$$f_1(r) \quad \textcircled{r+1} \leq s < 2r: \text{ thickness 1}$$

$$f_2(r) = \underbrace{f_1(f_1(\dots f_1(r) \dots))}_{r \text{ times}}$$

$$f_2(r) \quad \textcircled{2r} \leq s < r 2^r: \text{ thickness 2}$$

$$f_3(r) \quad \textcircled{r 2^r} \leq s < f_4(r): \text{ thickness 3}$$

$$f_4(r) \leq s < f_5(r): \text{ thickness 4}$$

⋮

Let's summarize and generalize



$$f_1(r) \quad \textcircled{r+1} \leq s < 2r: \text{ thickness 1}$$

$$f_2(r) \quad \textcircled{2r} \leq s < r 2^r: \text{ thickness 2}$$

$$f_3(r) \quad \textcircled{r 2^r} \leq s < f_4(r): \text{ thickness 3}$$

$$f_4(r) \leq s < f_5(r): \text{ thickness 4}$$

⋮

$$f_2(r) = \overbrace{f_1(f_1(\dots f_1(r)\dots))}^{r \text{ times}} \\ = f_1^{(r)}(r)$$

Let's summarize and generalize



$$f_1(r) \quad \textcircled{r+1} \leq s < 2r: \text{ thickness 1}$$

$$f_2(r) \quad \textcircled{2r} \leq s < r 2^r: \text{ thickness 2}$$

$$f_3(r) \quad \textcircled{r 2^r} \leq s < f_4(r): \text{ thickness 3}$$

$$f_4(r) \leq s < f_5(r): \text{ thickness 4}$$

⋮

$$f_2(r) = \overbrace{f_1(f_1(\dots f_1(r) \dots))}^{r \text{ times}}$$

$$= f_1^{(r)}(r)$$

$$f_3(r) = f_2^{(r)}(r)$$

Let's summarize and generalize



$$f_1(r) \\ (r + 1) \leq s < 2r: \text{ thickness 1}$$

$$f_2(r) (2r) \leq s < r 2^r: \text{ thickness 2}$$

$$f_3(r) (r 2^r) \leq s < f_4(r): \text{ thickness 3}$$

$$f_4(r) \leq s < f_5(r): \text{ thickness 4}$$

⋮

$$f_2(r) = \overbrace{f_1(f_1(\dots f_1(r) \dots))}^{r \text{ times}}$$

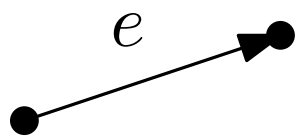
$$= f_1^{(r)}(r)$$

$$f_3(r) = f_2^{(r)}(r)$$

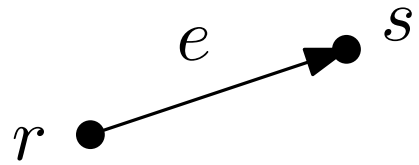
$$f_4(r) = f_3^{(r)}(r)$$

Weak and Strong Edge Redirections

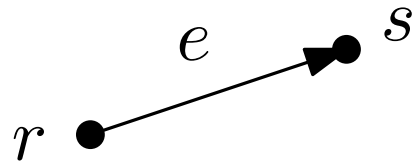
Weak and Strong Edge Redirections



Weak and Strong Edge Redirections

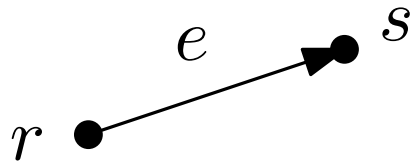


Weak and Strong Edge Redirections



$$\text{thickness}(e) = i$$

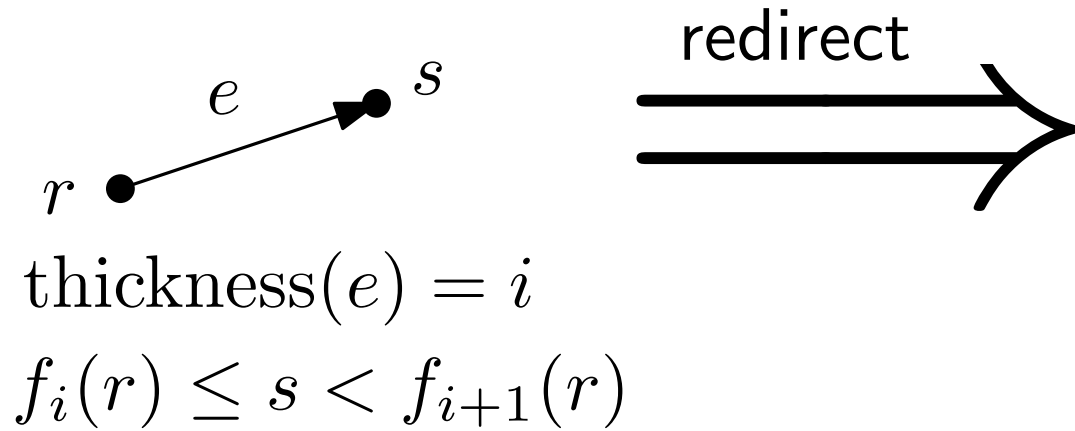
Weak and Strong Edge Redirections



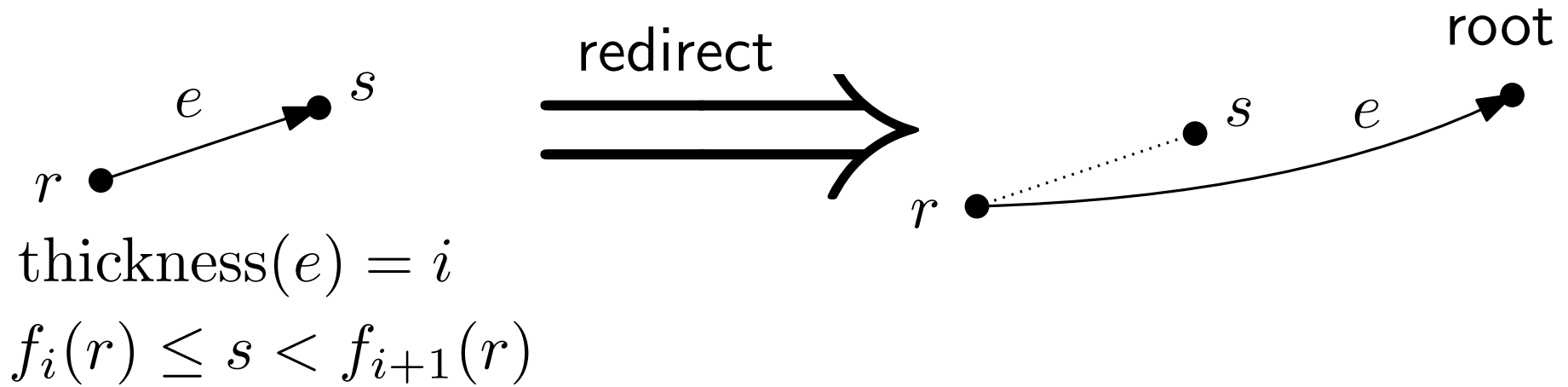
$$\text{thickness}(e) = i$$

$$f_i(r) \leq s < f_{i+1}(r)$$

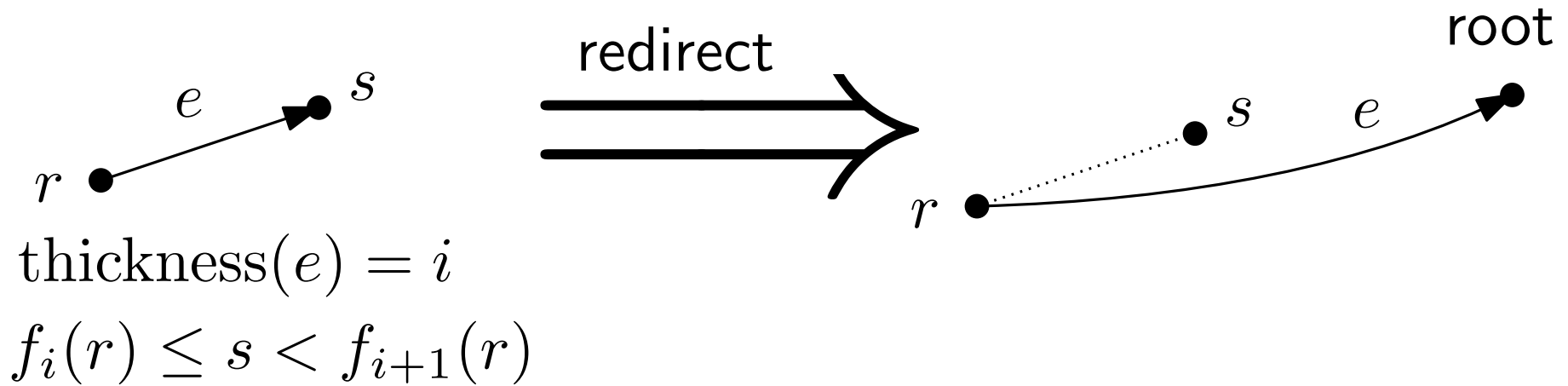
Weak and Strong Edge Redirections



Weak and Strong Edge Redirections

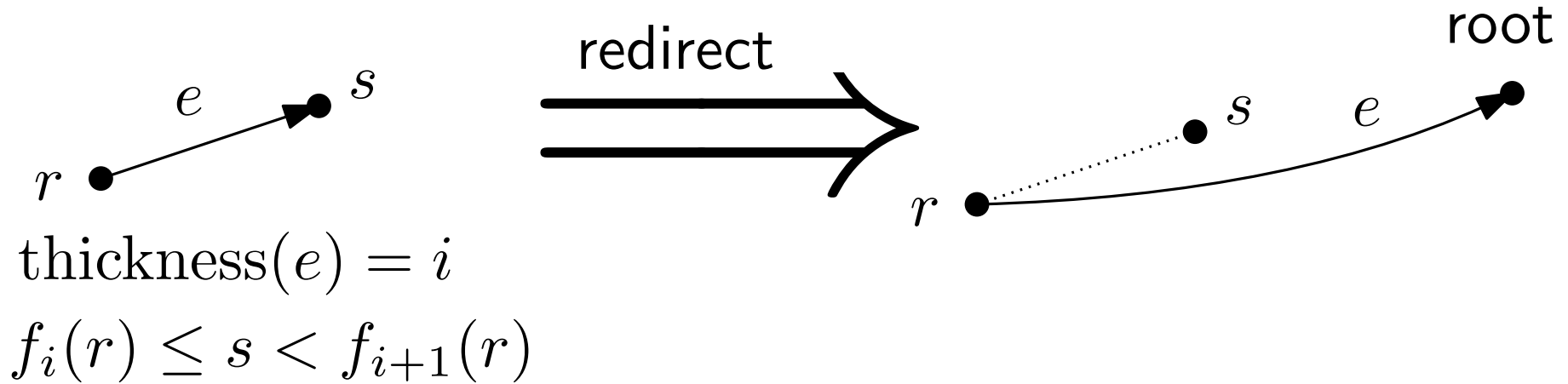


Weak and Strong Edge Redirections



Case 1: e is the last edge of thickness i on this path.

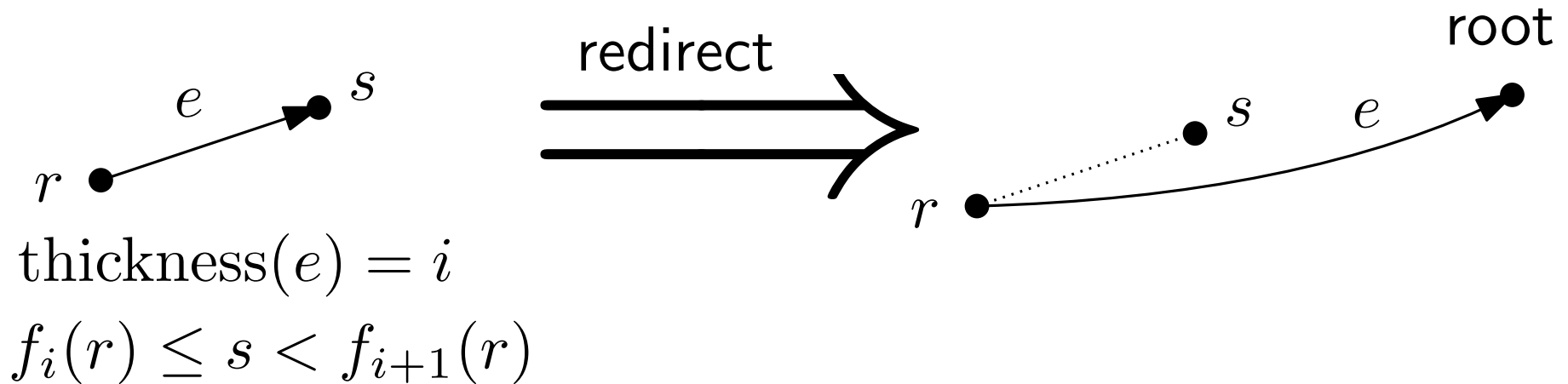
Weak and Strong Edge Redirections



Case 1: e is the last edge of thickness i on this path.

Then `find` operation pays one 🌐 for this.

Weak and Strong Edge Redirections

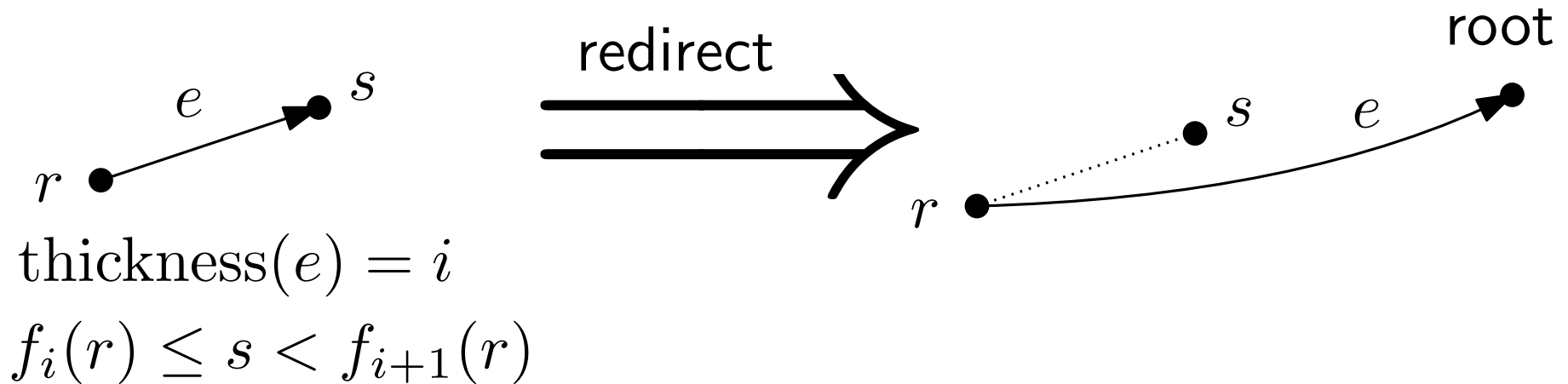


Case 1: e is the last edge of thickness i on this path.

Then `find` operation pays one 🌐 for this.

This is a *weak* edge redirection.

Weak and Strong Edge Redirections



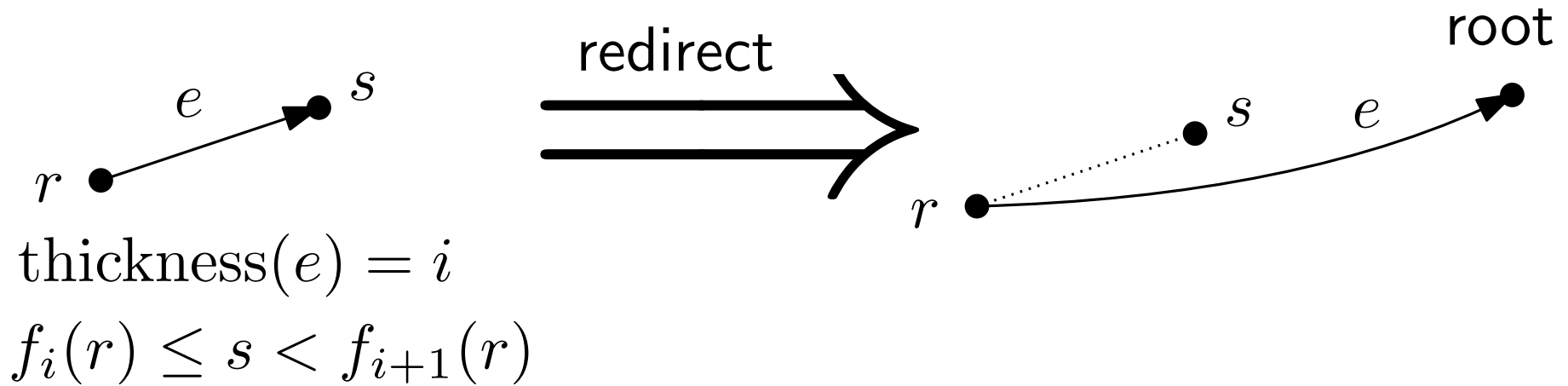
Case 1: e is the last edge of thickness i on this path.

Then `find` operation pays one 🌐 for this.

This is a *weak* edge redirection.

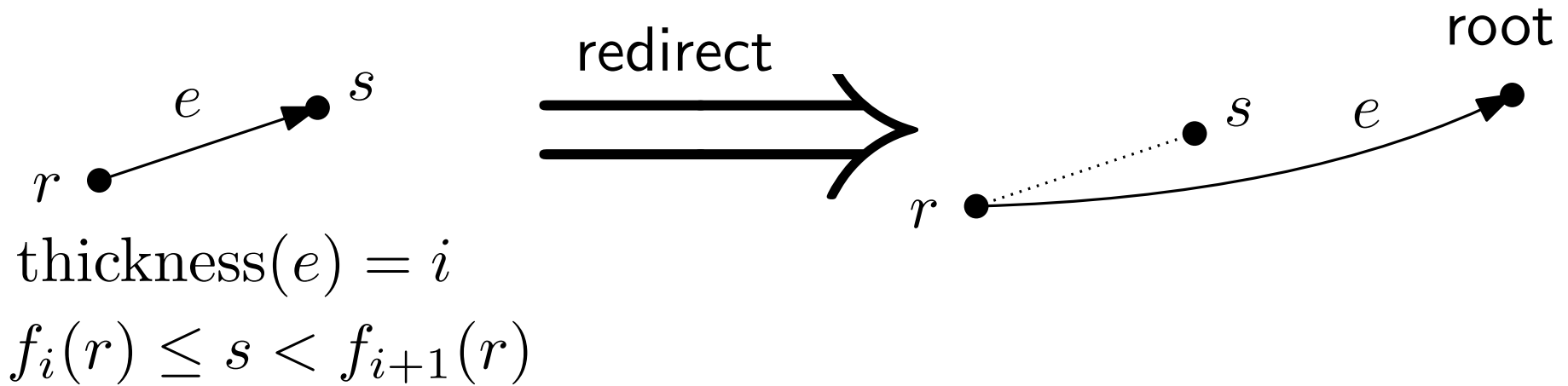
Total cost: number of thickness types.

Weak and Strong Edge Redirections

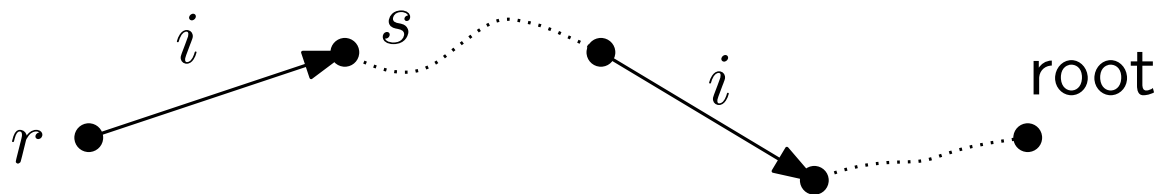


Case 2: e is not the last edge of thickness i on this path.

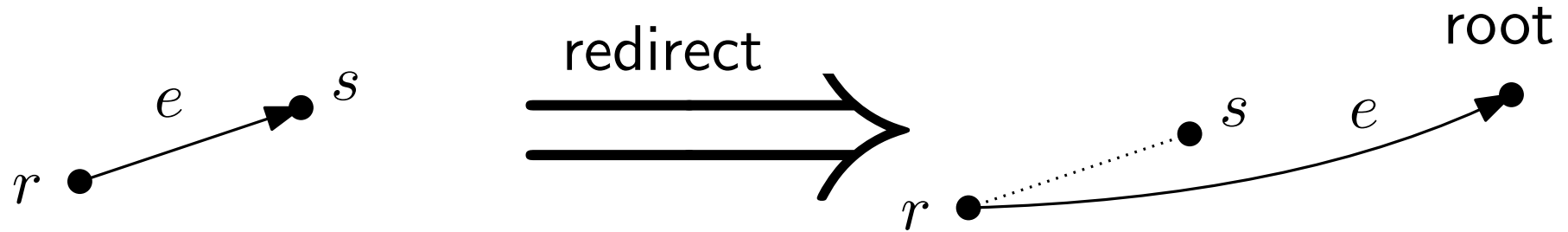
Weak and Strong Edge Redirections



Case 2: e is not the last edge of thickness i on this path.



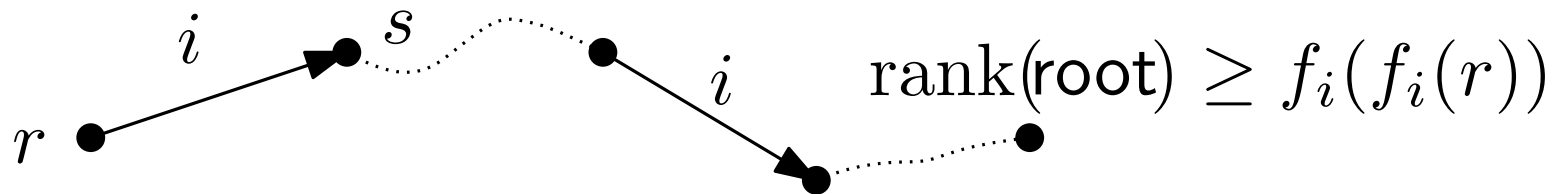
Weak and Strong Edge Redirections



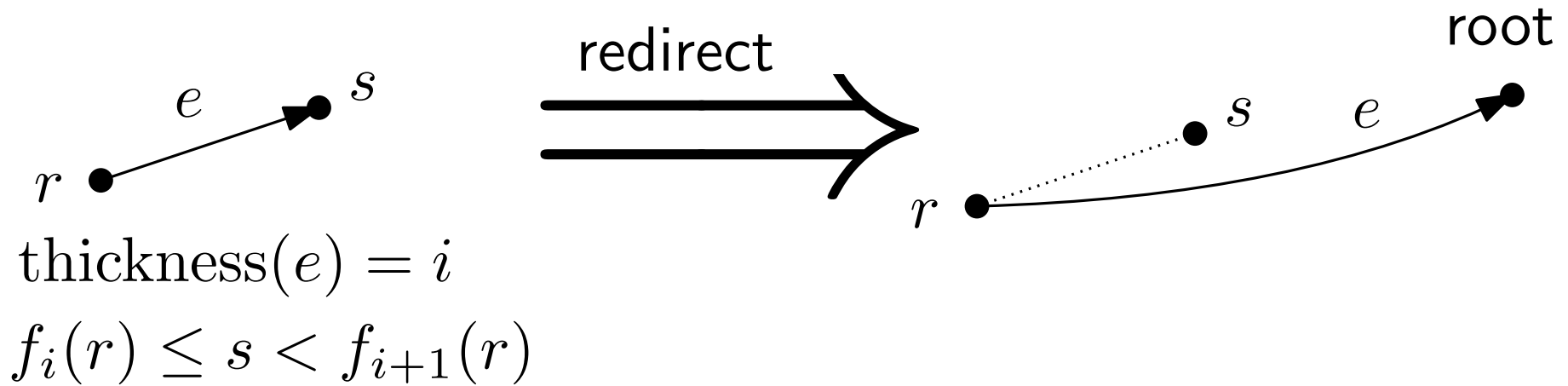
$$\text{thickness}(e) = i$$

$$f_i(r) \leq s < f_{i+1}(r)$$

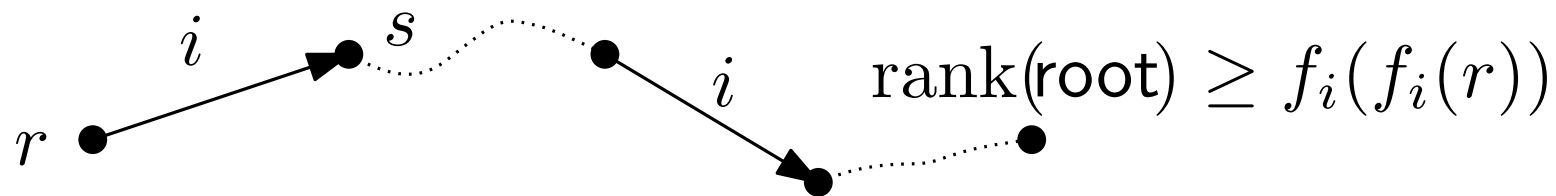
Case 2: e is not the last edge of thickness i on this path.



Weak and Strong Edge Redirections

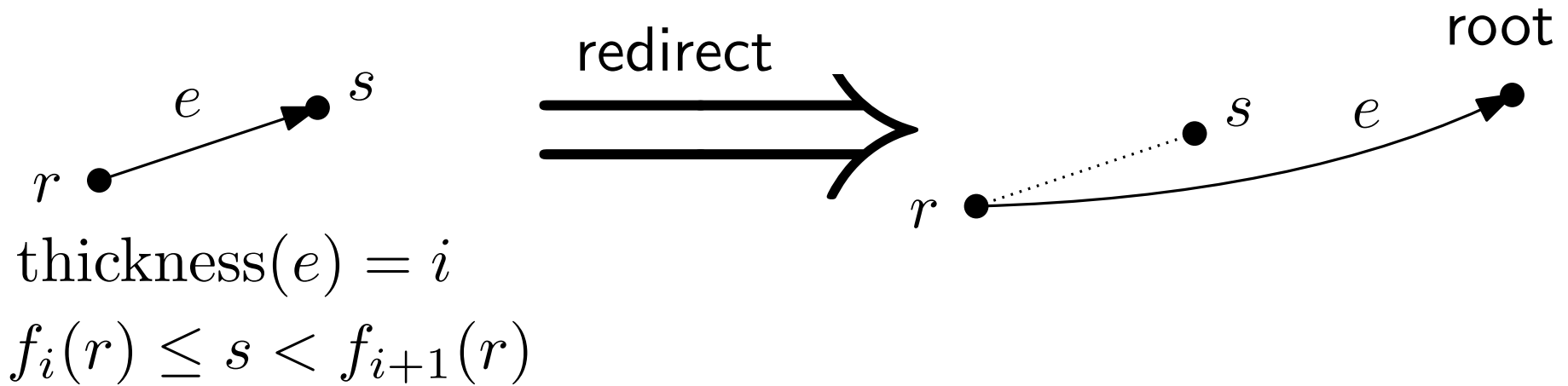


Case 2: e is not the last edge of thickness i on this path.

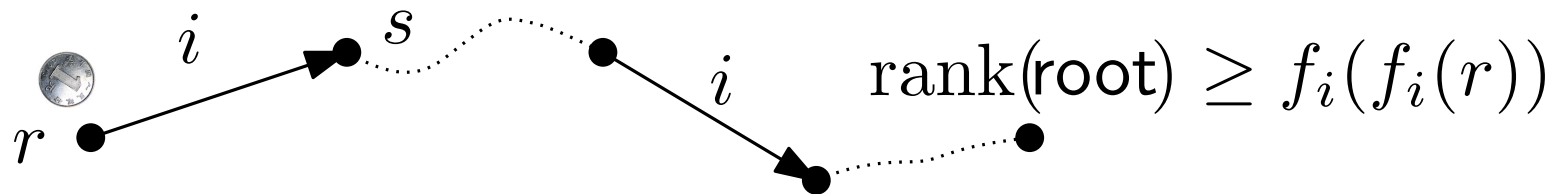


strong redirection

Weak and Strong Edge Redirections

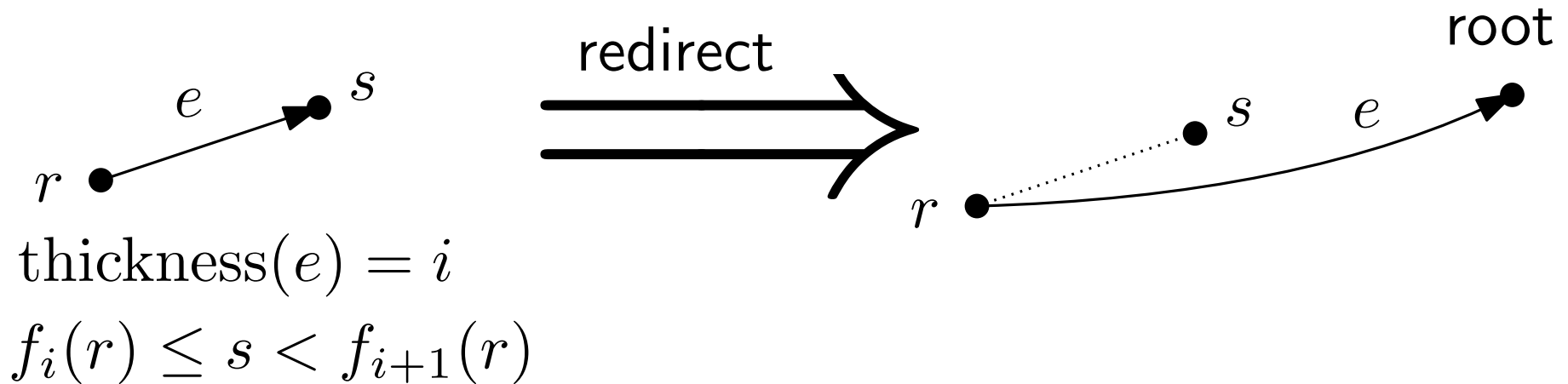


Case 2: e is not the last edge of thickness i on this path.

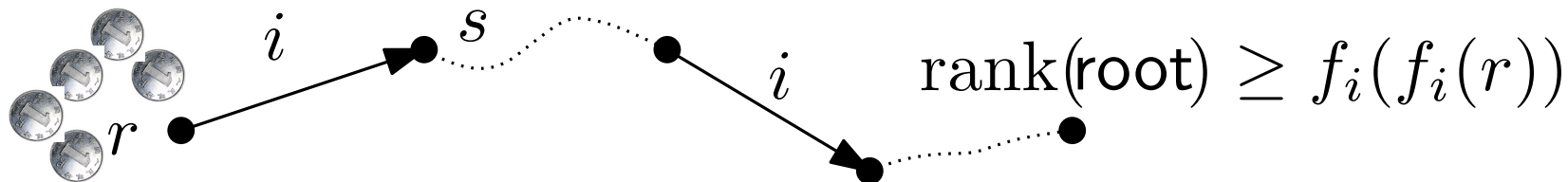


strong redirection

Weak and Strong Edge Redirections

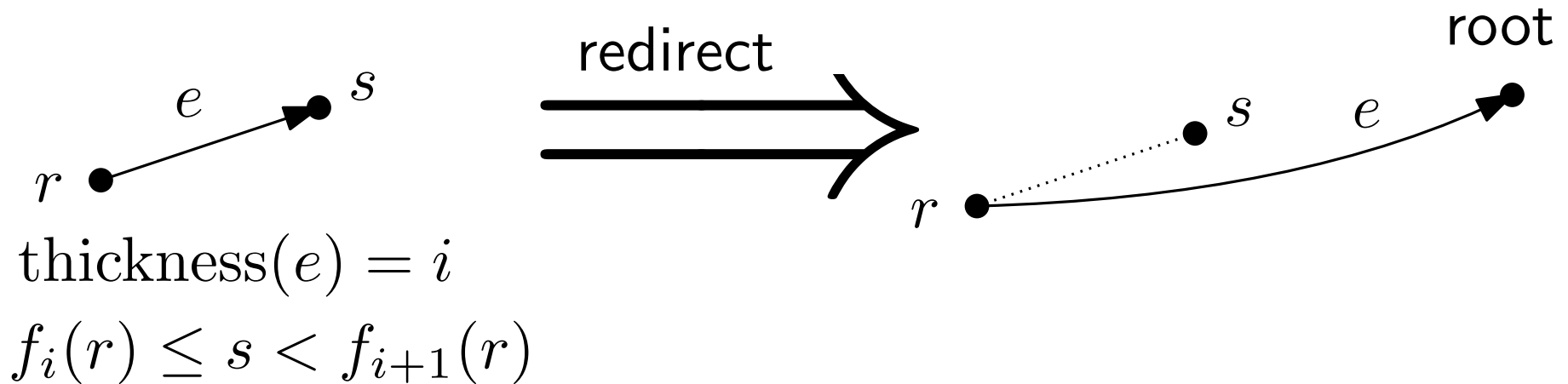


Case 2: e is not the last edge of thickness i on this path.

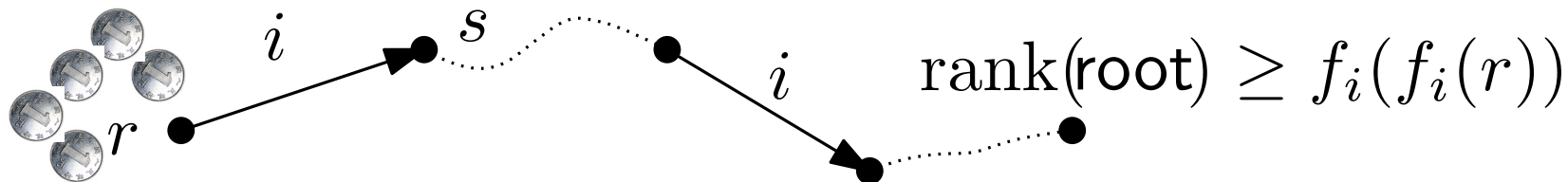


after $r - 1$ *strong* redirections: $\text{rank}(\text{root}) \geq f_i^{(r)}(r) = f_{i+1}(r)$

Weak and Strong Edge Redirections



Case 2: e is not the last edge of thickness i on this path.



after $r - 1$ *strong* redirections: $\text{rank}(\text{root}) \geq f_i^{(r)}(r) = f_{i+1}(r)$
 and the thickness of e increases to $i + 1$

Putting Everything Together

Putting Everything Together

let ℓ be the number of thickness types occurring.

Putting Everything Together

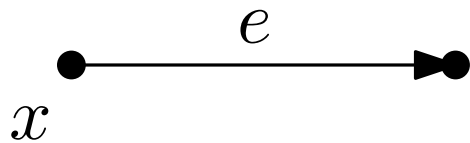
let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐



Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🌐
before the thickness of e increases.

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🌐
before the thickness of e increases.

Overall, x pays at most $\ell \cdot \text{rank}(x)$ 🌐

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🪙

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🪙
before the thickness of e increases.

Overall, x pays at most $\ell \cdot \text{rank}(x)$ 🪙

$$\sum_x \ell \cdot \text{rank}(x) = \ell \cdot \sum_r |\{\text{elements of rank } r\}|$$

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🌐
before the thickness of e increases.

Overall, x pays at most $\ell \cdot \text{rank}(x)$ 🌐

$$\begin{aligned} \sum_x \ell \cdot \text{rank}(x) &= \ell \cdot \sum_r |\{\text{elements of rank } r\}| \\ &\leq \ell \cdot \sum_r \frac{n}{2^{r-2}} \end{aligned}$$

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🌐
before the thickness of e increases.

Overall, x pays at most $\ell \cdot \text{rank}(x)$ 🌐

$$\begin{aligned}\sum_x \ell \cdot \text{rank}(x) &= \ell \cdot \sum_r |\{\text{elements of rank } r\}| \\ &\leq \ell \cdot \sum_r \frac{n}{2^{r-2}} \\ &= 6 \ell n.\end{aligned}$$

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🌐
before the thickness of e increases.

Overall, x pays at most $\ell \cdot \text{rank}(x)$ 🌐

$$\begin{aligned}\sum_x \ell \cdot \text{rank}(x) &= \ell \cdot \sum_r |\{\text{elements of rank } r\}| \\ &\leq \ell \cdot \sum_r \frac{n}{2^{r-2}} \\ &= 6 \ell n.\end{aligned}$$

Overall cost is $O(\ell \cdot (n + m))$.

Putting Everything Together

let ℓ be the number of thickness types occurring.

each find pays at most ℓ 🌐

x \xrightarrow{e} x pays at most $\text{rank}(x)$ 🌐
before the thickness of e increases.

Overall, x pays at most $\ell \cdot \text{rank}(x)$ 🌐

$$\begin{aligned}\sum_x \ell \cdot \text{rank}(x) &= \ell \cdot \sum_r |\{\text{elements of rank } r\}| \\ &\leq \ell \cdot \sum_r \frac{n}{2^{r-2}} \\ &= 6 \ell n.\end{aligned}$$

How large can ℓ become?

Overall cost is $O(\ell \cdot (n + m))$.

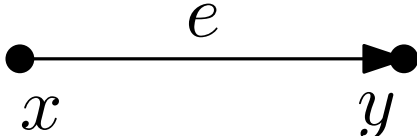
Overall running time is $O(\ell (n + m))$

Overall running time is $O(\ell (n + m))$

ℓ is the number of thickness levels

Overall running time is $O(\ell (n + m))$

ℓ is the number of thickness levels

Imagine  is an edge of thickness 5.

Overall running time is $O(\ell (n + m))$

ℓ is the number of thickness levels

Imagine  is an edge of thickness 5.

$$\text{rank}(y) \geq f_5(\text{rank}(x)) \geq f_5(2)$$

What is $f_5(2)$?

What is $f_5(2)$?

$$f_1(n) = n + 1$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_2(n) = 2n$$

$$f_1(2) = 3$$

$$f_2(2) = 4$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_2(n) = 2n$$

$$f_3(n) = n 2^n$$

$$f_1(2) = 3$$

$$f_2(2) = 4$$

$$f_3(2) = 8$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8)$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048)$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048) \geq f_3(f_3(f_3(\dots f_3(2048)\dots)))$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048) \geq \underbrace{f_3(f_3(f_3(\dots f_3(2048)\dots)))}_{2048 \text{ times}}$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048) \geq \underbrace{f_3(f_3(f_3(\dots f_3(2048)\dots)))}_{2048 \text{ times}}$$

$$\geq 2^{2^{2^{\dots 2^{2048}}}}$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048) \geq \underbrace{f_3(f_3(f_3(\dots f_3(2048)\dots)))}_{2048 \text{ times}}$$

$$f_5(2) \geq 2^{2^{2^{\dots 2^{2048}}}}$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048) \geq \underbrace{f_3(f_3(f_3(\dots f_3(2048)\dots)))}_{2048 \text{ times}}$$

$$\text{rank}(y) \geq f_5(2) \geq 2^{2^{2^{\dots 2^{2048}}}}$$

What is $f_5(2)$?

$$f_1(n) = n + 1$$

$$f_1(2) = 3$$

$$f_2(n) = 2n$$

$$f_2(2) = 4$$

$$f_3(n) = n 2^n$$

$$f_3(2) = 8$$

$$f_4(2) = f_3(f_3(2)) = f_3(8) = 8 \cdot 2^8 = 2048.$$

$$f_5(2) = f_4(f_4(2)) = f_4(2048) \geq \underbrace{f_3(f_3(f_3(\dots f_3(2048)\dots)))}_{2048 \text{ times}}$$

$$\log(n) \geq \text{rank}(y) \geq f_5(2) \geq 2^{2^{2^{\dots 2^{2048}}}}$$