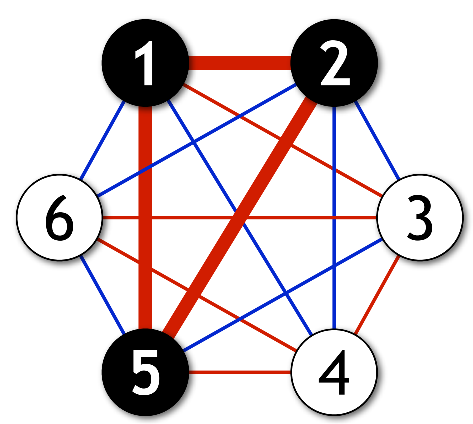
Ramsey's theorem and lower-bound results

Jukka Suomela

Adaptive Computing Group
Helsinki Institute for Information Technology HIIT
University of Helsinki

11 March 2010



=

$N = 6$, $k = 2$, $c = 2$				
{1,2}	{1,3}	{1,4}	**{1,5}**	{1,6}
	{2,3}	{2,4}	**{2,5}**	{2,6}
		{3,4}	{3,5}	{3,6}
			{4,5}	{4,6}
				{5,6}

Part I:
Ramsey's theorem

- A generalisation of the pigeonhole principle
- Frank P. Ramsey (1930):
  On a problem of formal logic
  - "... in the course of this investigation it is necessary to use certain theorems on combinations which have an independent interest..."

Basic definitions

- Assign a colour from $\{1, 2, \ldots, c\}$ to each k-subset of $\{1, 2, \ldots, N\}$

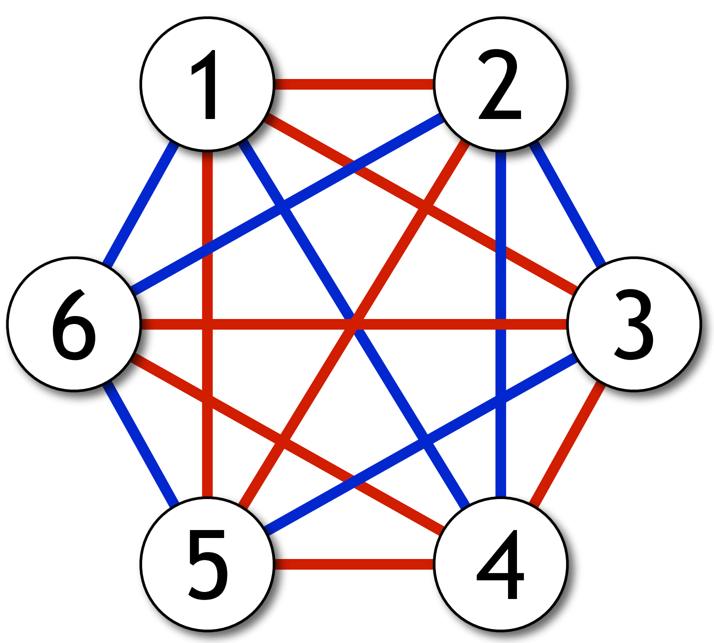
$N = 4, k = 3, c = 2$	
{1,2,3}	{1,2,4}
{1,3,4}	{2,3,4}

$N = 13, k = 1, c = 3$			
{1}	{2}	{3}	{4}
{5}	{6}	{7}	{8}
{9}	{10}	{11}	{12}
{13}			

$N = 6, k = 2, c = 2$				
{1,2}	{1,3}	{1,4}	{1,5}	{1,6}
	{2,3}	{2,4}	{2,5}	{2,6}
		{3,4}	{3,5}	{3,6}
			{4,5}	{4,6}
				{5,6}

Basic definitions

- Assign a colour from $\{1, 2, \ldots, c\}$ to each k-subset of $\{1, 2, \ldots, N\}$

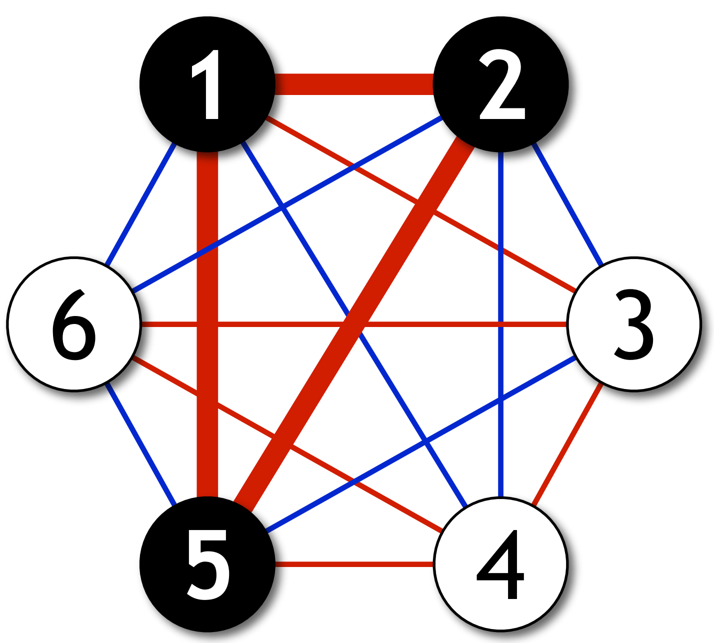


=

$N = 6$, $k = 2$, $c = 2$				
{1,2}	{1,3}	{1,4}	{1,5}	{1,6}
	{2,3}	{2,4}	{2,5}	{2,6}
		{3,4}	{3,5}	{3,6}
			{4,5}	{4,6}
				{5,6}

Basic definitions

- $X \subset \{1, 2, \ldots, N\}$ is a ***monochromatic subset*** if all k-subsets of X have the same colour



=

$N = 6$, $k = 2$, $c = 2$				
{1,2}	{1,3}	{1,4}	**{1,5}**	{1,6}
	{2,3}	{2,4}	**{2,5}**	{2,6}
		{3,4}	{3,5}	{3,6}
			{4,5}	{4,6}
				{5,6}

Ramsey's theorem

- Assign a colour from $\{1, 2, \ldots, c\}$ to each k-subset of $\{1, 2, \ldots, N\}$
- $X \subset \{1, 2, \ldots, N\}$ is a monochromatic subset if all k-subsets of X have the same colour
- **Ramsey's theorem**: For all c, k, and n there is a finite N such that ***any*** c-colouring of k-subsets of $\{1, 2, \ldots, N\}$ contains a monochromatic subset with n elements

Ramsey's theorem

- Assign a colour from $\{1, 2, \ldots, c\}$ to each k-subset of $\{1, 2, \ldots, N\}$
- $X \subset \{1, 2, \ldots, N\}$ is a monochromatic subset if all k-subsets of X have the same colour
- **Ramsey's theorem**: For all c, k, and n there is a finite N such that ***any*** c-colouring of k-subsets of $\{1, 2, \ldots, N\}$ contains a monochromatic subset with n elements
  - The smallest such N is denoted by $R_c(n; k)$

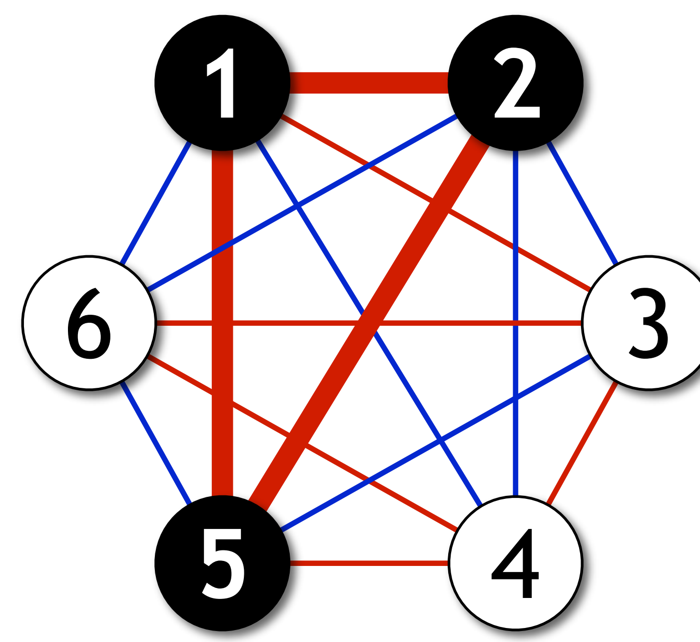
Ramsey numbers

Ramsey's theorem: $k = 1$

- $k = 1$: pigeonhole principle
- If we put N items into c slots, then at least one of the slots has to contain at least n items
  - Colour of the 1-subset $\{i\}$ = slot of the element i
  - Clearly holds if $N \geq c(n - 1) + 1$
  - Does not necessarily hold if $N \leq c(n - 1)$
  - $R_c(n; 1) = c(n - 1) + 1$

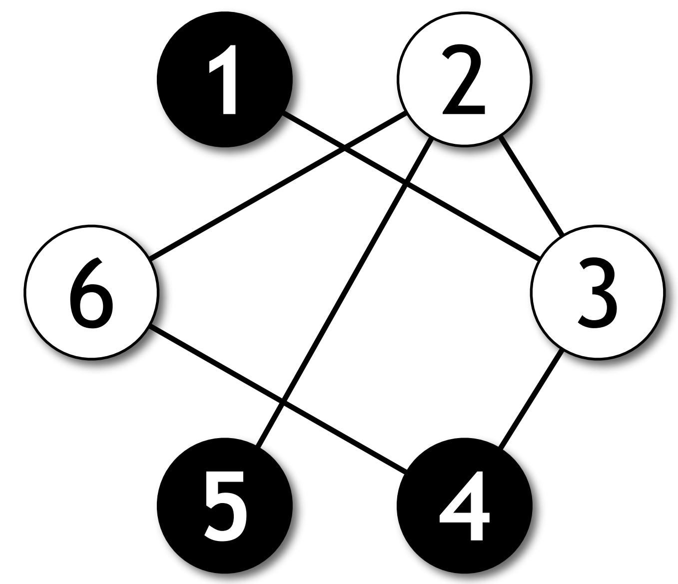
Ramsey's theorem: $k = 2$, $c = 2$

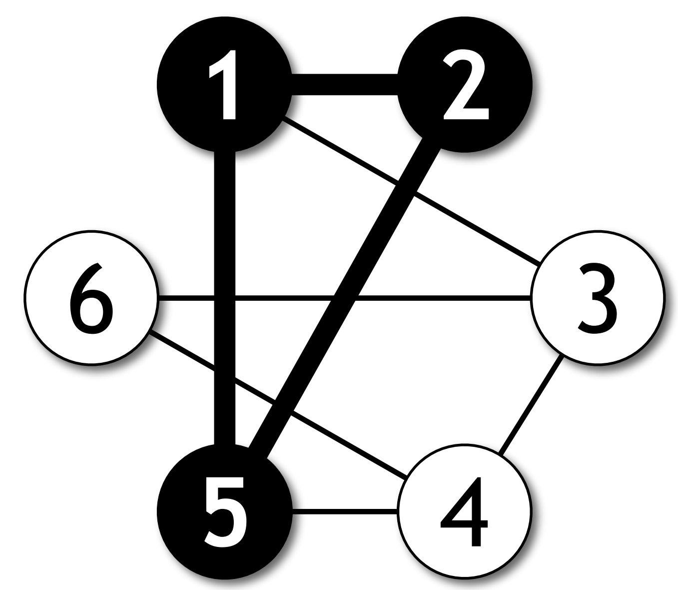
- Complete graphs, red and blue edges
- If the graph is large enough,
  there will be a monochromatic clique
  - For example, $R_2(2; 2) = 2$,
    $R_2(3; 2) = 6$, and $R_2(4; 2) = 18$
  - A graph with 2 nodes contains
    a monochromatic edge
  - A graph with 6 nodes contains
    a monochromatic triangle



Ramsey's theorem: $k = 2$, $c = 2$

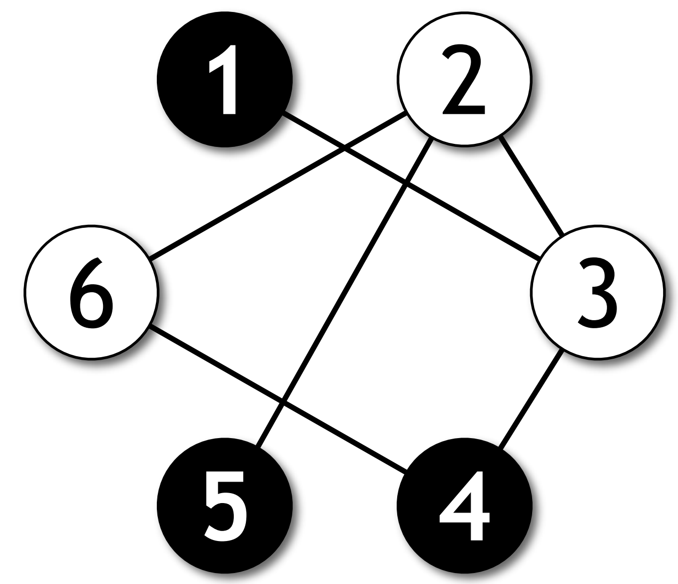
- Another interpretation: graphs
  - $\{u,v\}$ red: edge $\{u,v\}$ present
  - $\{u,v\}$ blue: edge $\{u,v\}$ missing
- Large monochromatic subset:
  - Large clique (red) or
    large independent set (blue)
  - Any graph with 6 nodes
    contains a clique with 3 nodes or
    an independent set with 3 nodes

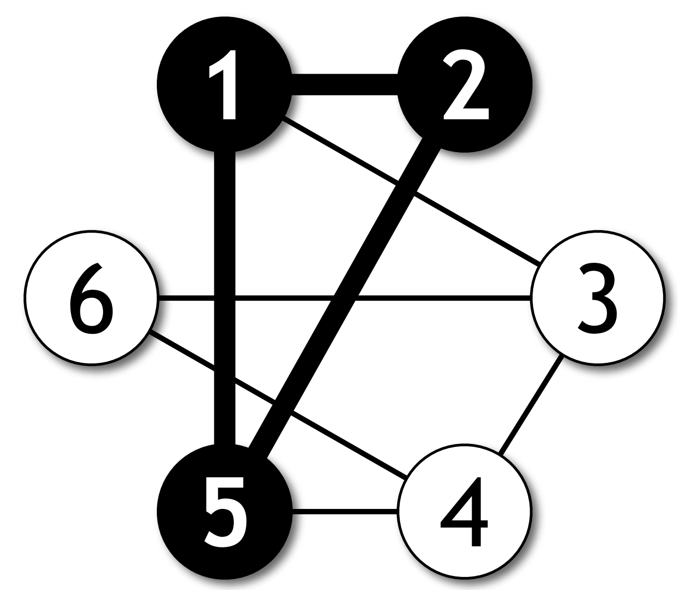




Ramsey's theorem: k = 2, c = 2

- Sufficiently large graphs (N nodes) contain large independents sets (n nodes) or large cliques (n nodes)
  - You can avoid one of these, but not both
  - However, Ramsey numbers are large: here N is exponential in n

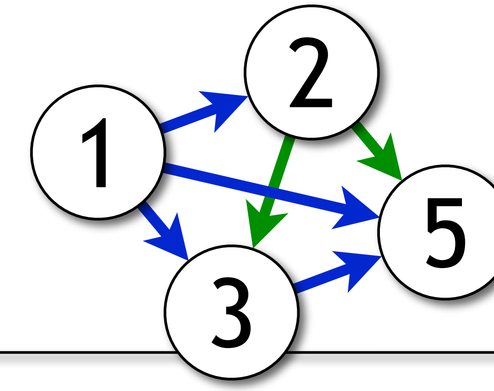




Part II:
Proof of Ramsey's theorem

- Following Nešetřil (1995)
- Notation from Radziszowski

Definitions



- X ⊂ {1, 2, …, *N*} is a ***monochromatic subset***:
  if *A* and *B* are *k*-subsets of *X*,
  then *A* and *B* have the same colour
- X ⊂ {1, 2, …, *N*} is a ***good subset***:
  if *A* and *B* are *k*-subsets of *X* and min(*A*) = min(*B*),
  then *A* and *B* have the same colour
  - An example with *c* = 2 and *k* = 2:
    {1,2,3,5} is good but not monochromatic in the colouring
    {1,2}, **{1,3}**, {1,4}, **{1,5}**, **{2,3}**, {2,4}, **{2,5}**, **{3,5}**, {4,5}

Definitions

- $X \subset \{1, 2, \ldots, N\}$ is a ***monochromatic subset***:
  if A and B are k-subsets of X,
  then A and B have the same colour
- $X \subset \{1, 2, \ldots, N\}$ is a ***good subset***:
  if A and B are k-subsets of X and $\min(A) = \min(B)$,
  then A and B have the same colour
  - $R_c(n; k)$ = smallest N s.t. ∃ monochromatic n-subset
  - $G_c(n; k)$ = smallest N s.t. ∃ good n-subset

Proof outline

- $R_c(n; k)$ = smallest N s.t. ∃ monochromatic n-subset
- $G_c(n; k)$ = smallest N s.t. ∃ good n-subset
- Theorem: $R_c(n; k)$ is finite for all c, n, k

  (i) $R_c(n; 1)$ is finite for all c, n

  (ii) If $R_c(n; k - 1)$ is finite for all c, n
  then $G_c(n; k)$ is finite for all c, n

  (iii) $R_c(n; k) \leq G_c(c(n - 1) + 1; k)$ for all c, n, k

Proof: step (i)

- Lemma: $R_c(n; 1)$ is finite for all c, n
- Proof:
  - Pigeonhole principle
  - $R_c(n; 1) = c(n - 1) + 1$

Proof: step (ii) — outline

- Lemma: if $R_c(n; k - 1)$ is finite for all c, n then $G_c(n; k)$ is finite for all c, n
- Proof (for each fixed c):
  - Induction on n
  - $G_c(k; k)$ is finite
  - Assume that $M = G_c(n - 1; k)$ is finite
  - Then we also have a finite $R_c(M; k - 1)$
  - Enough to show that $G_c(n; k) \leq 1 + R_c(M; k - 1)$

Proof: step (ii)

f:	{1,2,3}	{1,2,4}	{1,3,4}	{2,3,4}
f':	{2,3}	{2,4}	{3,4}	

- $G_c(n; k) \leq 1 + R_c(M; k-1)$ where $M = G_c(n-1; k)$
  - Let $N = 1 + R_c(M; k-1)$, consider any colouring f of k-subsets of $\{1, 2, \ldots, N\}$
  - Delete element 1: colouring f' of $(k-1)$-subsets of $\{2, 3, \ldots, N\}$
  - Find an f'-monochromatic M-subset $X \subset \{2, 3, \ldots, N\}$
  - Find an f-good $(n-1)$-subset $Y \subset X$
  - $\{1\} \cup Y$ is an f-good n-subset of $\{1, 2, \ldots, N\}$

Proof: step (ii)

In real life, these constants would be much larger…

- A fictional example: $N = 7$, $M = 5$, $n = 5$, $k = 3$
  - Original colouring f: {1,2,3}, {1,2,4}, {1,2,5}, {1,2,6}, {1,2,7}, …, {1,6,7}, {2,3,4}, …, {5,6,7}
  - Colouring f': {2,3}, {2,4}, {2,5}, {2,6}, {2,7}, …, {6,7}
  - f'-monochromatic M-subset {2,3,4,5,7} of {2,3,…,N}: {2,3}, {2,4}, {2,5}, {2,7}, …, {5,7}
  - f-good $(n-1)$-subset {2,4,5,7}: {2,4,5}, {2,4,7}, {4,5,7}
  - {***1***,2,4,5,7} is f-good: {***1***,2,4}, {***1***,2,5}, {***1***,2,7}, …, {***1***,5,7}, {2,4,5}, {2,4,7}, {4,5,7}

Proof: step (ii)

$N - 1 \geq R_c(M; k - 1)$

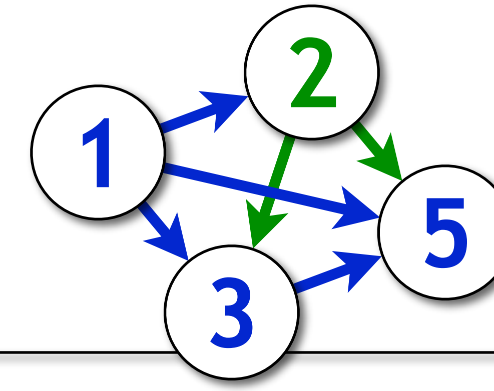
$M \geq G_c(n - 1; k)$

- A fictional example: $N = 7$, $M = 5$, $n = 5$, $k = 3$
  - Original colouring f: {1,2,3}, {1,2,4}, {1,2,5}, {1,2,6}, {1,2,7}, …, {1,6,7}, {2,3,4}, …, {5,6,7}
  - Colouring f': {2,3}, {2,4}, {2,5}, {2,6}, {2,7}, …, {6,7}
  - f'-monochromatic M-subset {2,3,4,5,7} of {2,3,…,N}: {2,3}, {2,4}, {2,5}, {2,7}, …, {5,7}
  - f-good $(n-1)$-subset {2,4,5,7}: {2,4,5}, {2,4,7}, {4,5,7}
  - {***1***,2,4,5,7} is f-good: {***1***,2,4}, {***1***,2,5}, {***1***,2,7}, …, {***1***,5,7}, {2,4,5}, {2,4,7}, {4,5,7}

Proof: step (ii) — summary

- Lemma: if $R_c(n; k - 1)$ is finite for all c, n then $G_c(n; k)$ is finite for all c, n
- Proof (for each fixed c):
  - Induction on n
  - $G_c(k; k)$ is finite
  - We have shown that if $G_c(n - 1; k)$ is finite then $G_c(n; k)$ is finite
    - Trick: show that $G_c(n; k) \leq 1 + R_c(G_c(n - 1; k); k - 1)$

Proof: step (iii)



- Lemma: $R_c(n; k) \le G_c(c(n-1)+1; k)$ for all c, n, k
- Proof:
  - If $N = G_c(c(n-1)+1; k)$, we can find a good subset X with $c(n-1)+1$ elements
  - If k-subset A of X has colour i, put $\min(A)$ into slot i
  - E.g.: **{1,2}**, **{1,3}**, **{1,5}**, **{2,3}**, **{2,5}**, **{3,5}**: put 1 and 3 to slot blue, 2 to slot green, 5 to any slot
  - Each slot is monochromatic and at least one slot contains n elements (pigeonhole)!

Ramsey's theorem: proof summary

- $R_c(n; k)$ = smallest N s.t. $\exists$ monochromatic n-subset
- $G_c(n; k)$ = smallest N s.t. $\exists$ good n-subset
- Theorem: $R_c(n; k)$ is finite for all c, n, k

  (i) $R_c(n; 1)$ is finite for all c, n

  (ii) If $R_c(n; k - 1)$ is finite for all c, n
  then $G_c(n; k)$ is finite for all c, n

  - Induction: $G_c(n; k) \leq 1 + R_c(G_c(n - 1; k); k - 1)$

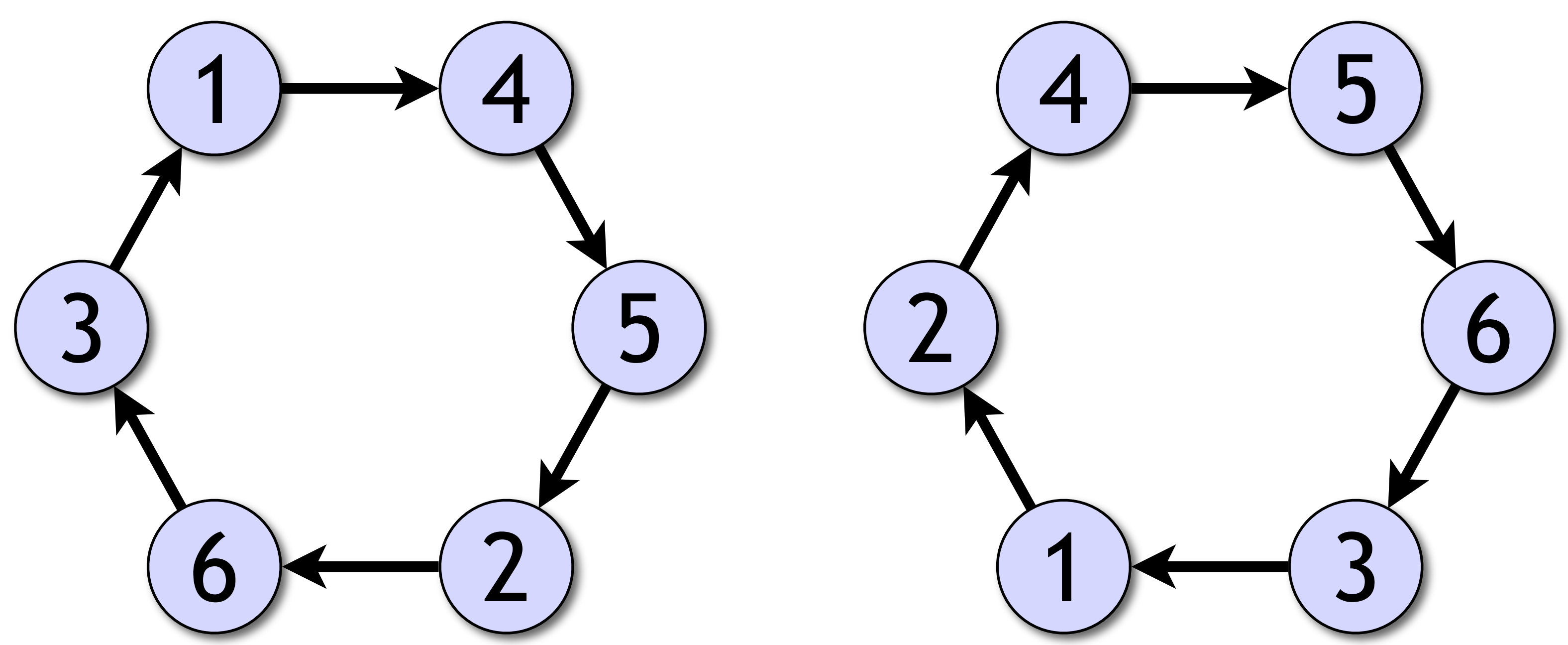
  (iii) $R_c(n; k) \leq G_c(c(n - 1) + 1; k)$ for all c, n, k

Part III: An application of Ramsey's theorem

- Czygrinow et al. (2008)
- A deterministic distributed algorithm can't find a $(2 - \varepsilon)$-approximation of vertex cover in constant time
- Holds even if we consider an n-cycle with unique identifiers from $1, 2, \ldots, n$

Lower-bound result for vertex cover approximation

- Numbered directed *n*-cycle:
  - directed *n*-cycle, each node has outdegree = indegree = 1
  - node identifiers are a permutation of {1, 2, ..., *n*}



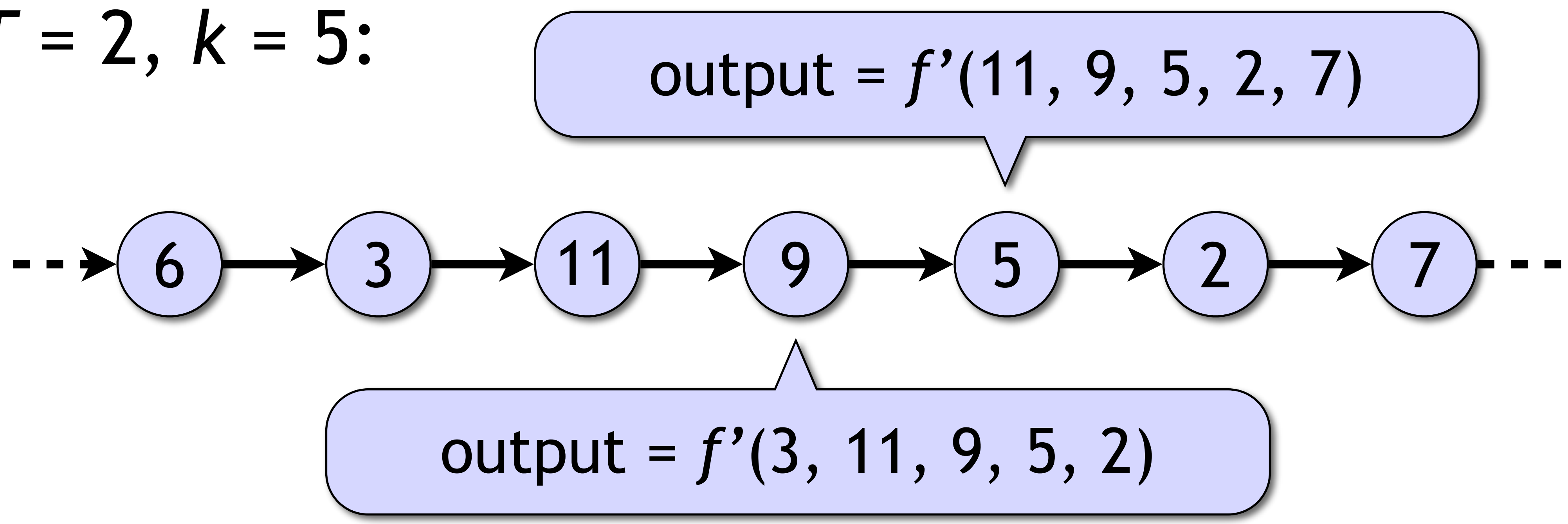
Lower-bound result for vertex cover approximation

- Fix any $\varepsilon > 0$ and a deterministic local algorithm A
  - Assumption: A finds a feasible vertex cover (at least in any numbered directed cycle)
- **Theorem:** For a sufficiently large n there is a numbered directed n-cycle C in which A outputs a vertex cover with $\geq (1 - \varepsilon)n$ nodes
- **Corollary:** Approximation ratio of A is at least $2 - 2\varepsilon$

Lower-bound result for vertex cover approximation

- Let T be the running time of A, let $k = 2T + 1$
- The output of a node is a function f' of a sequence of k integers (unique IDs)

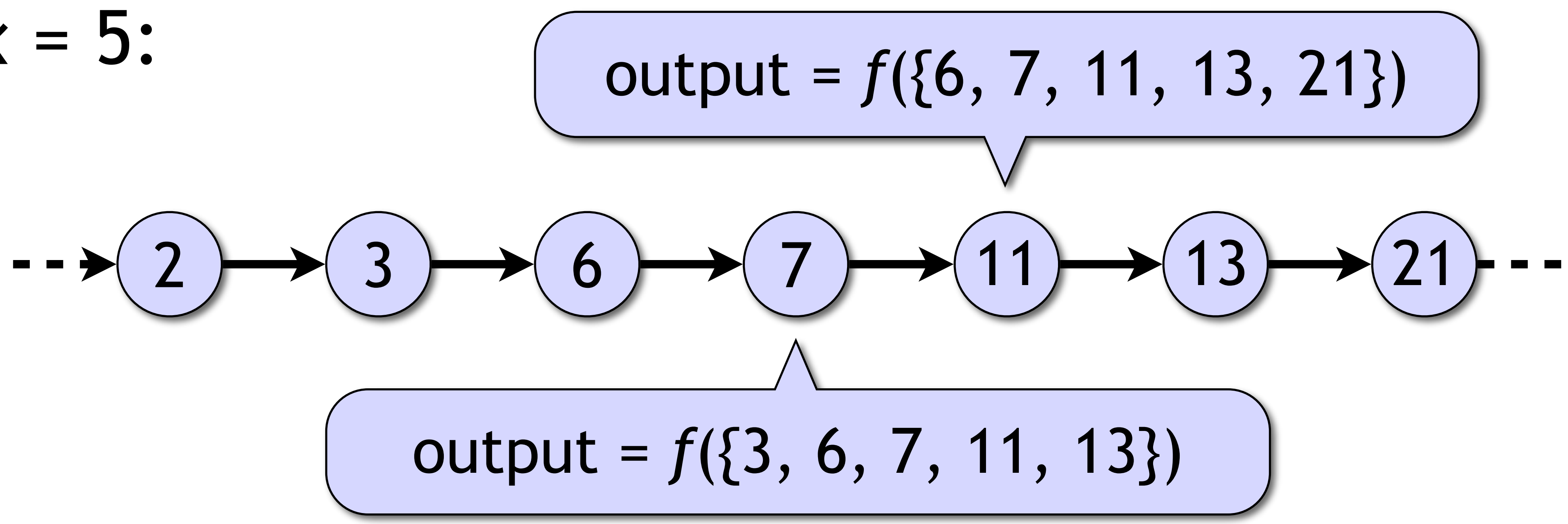
$T = 2$, $k = 5$:



Lower-bound result for vertex cover approximation

- Lets focus on **increasing** sequences of IDs
- Then the output of a node is a function *f* of a **set** of *k* integers

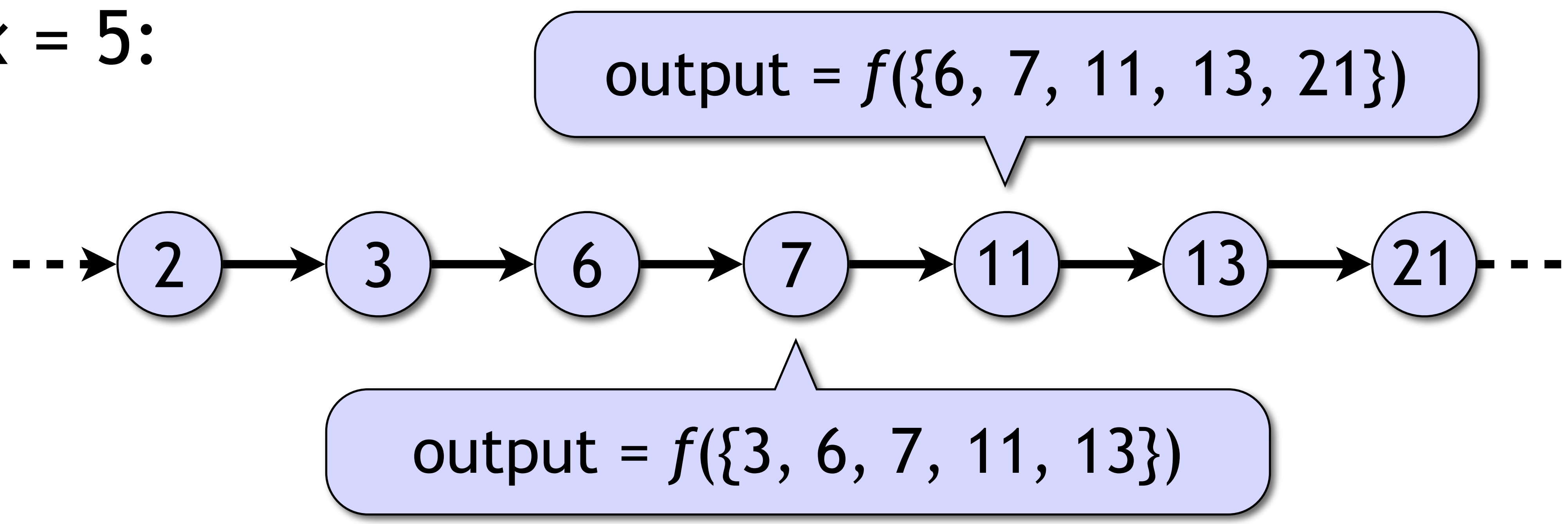
$k = 5$:



Lower-bound result for vertex cover approximation

- Hence we have assigned a colour $f(X) \in \{0, 1\}$ to each k-subset $X \subset \{1, 2, \ldots, n\}$

$k = 5$:

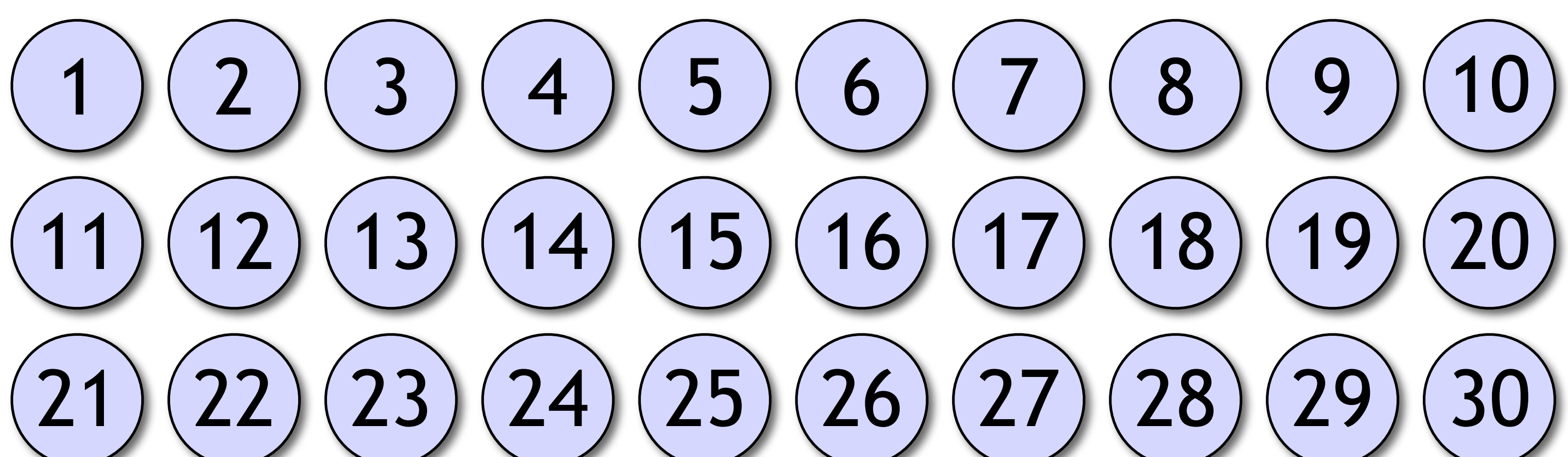


Lower-bound result for vertex cover approximation

- Hence we have assigned a colour $f(X) \in \{0, 1\}$ to each k-subset $X \subset \{1, 2, \ldots, n\}$
- Fix a large m (depends on k and ε)
- Ramsey: If n is sufficiently large, we can find an m-subset $A \subset \{1, 2, \ldots, n\}$ s.t. all k-subset $X \subset A$ have the same colour

Lower-bound result for vertex cover approximation

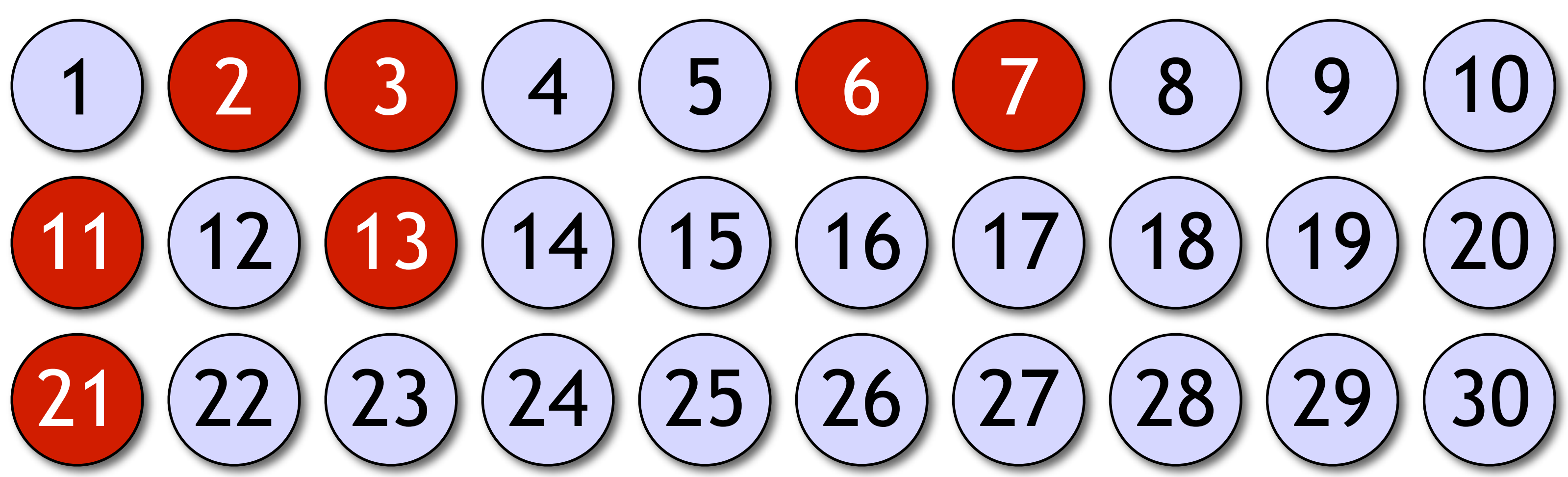
- That is, if the ID space is sufficiently large...



Lower-bound result for vertex cover approximation

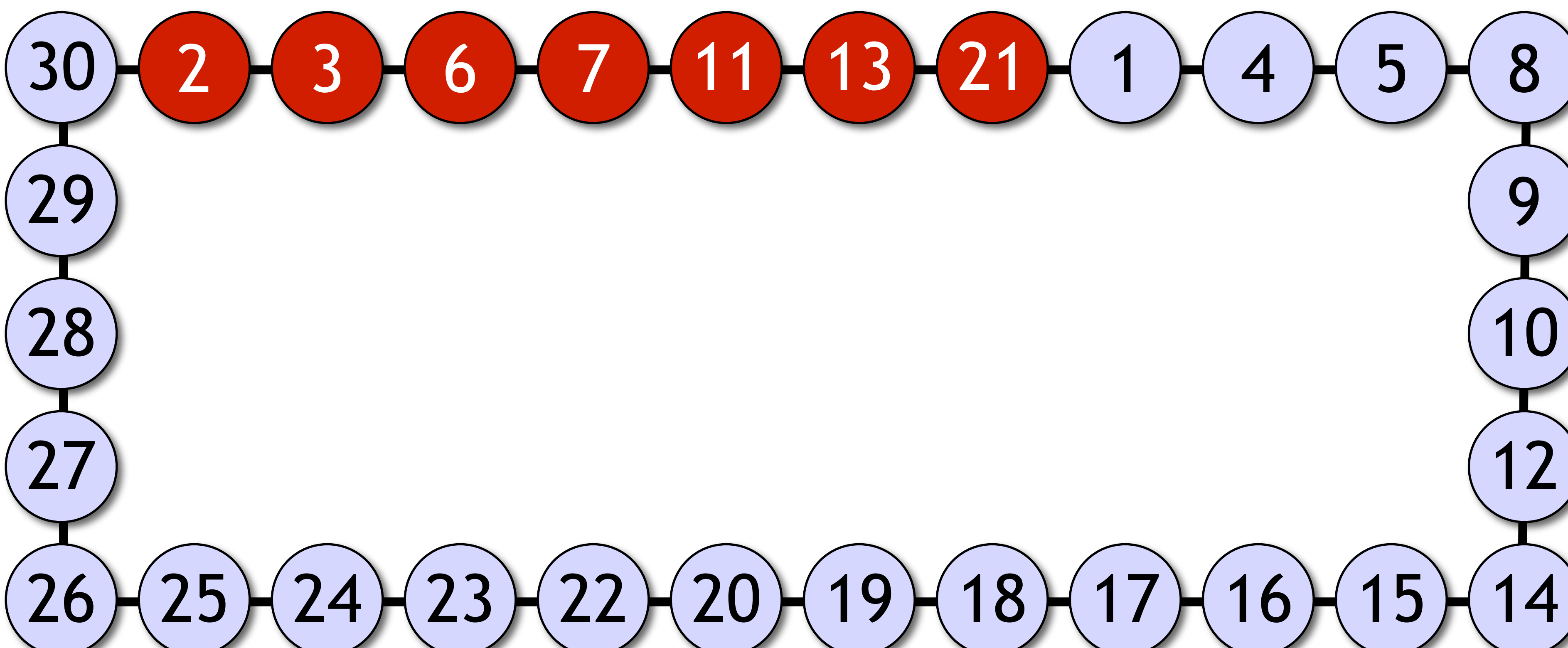
- That is, if the ID space is sufficiently large, we can find a monochromatic subset of *m* IDs...

$$f(\{2, 3, 6, 7, 11\}) = f(\{2, 3, 6, 7, 13\}) =$$
$$f(\{2, 3, 6, 7, 21\}) = f(\{2, 3, 6, 11, 13\}) =$$
$$\ldots = f(\{6, 7, 11, 13, 21\})$$



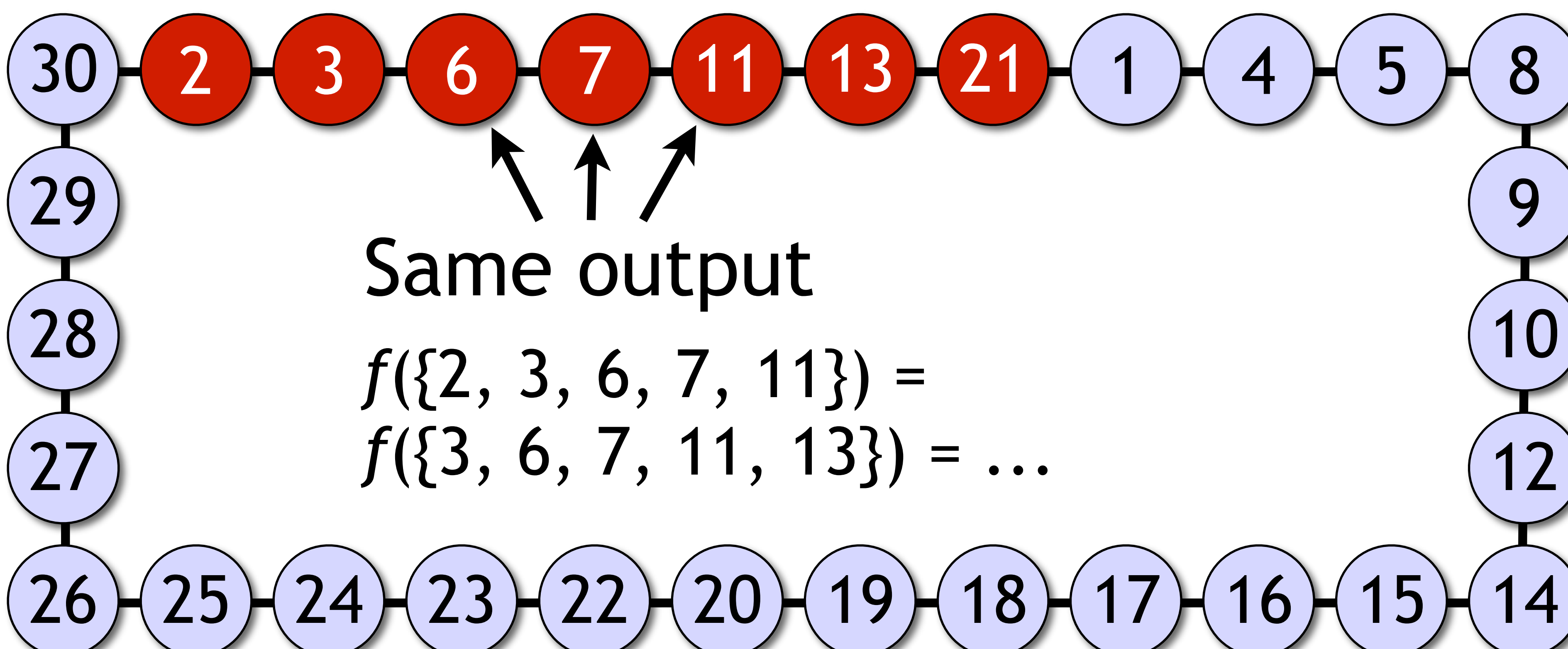
Lower-bound result for vertex cover approximation

- Construct a numbered directed cycle: monochromatic subset as consecutive nodes



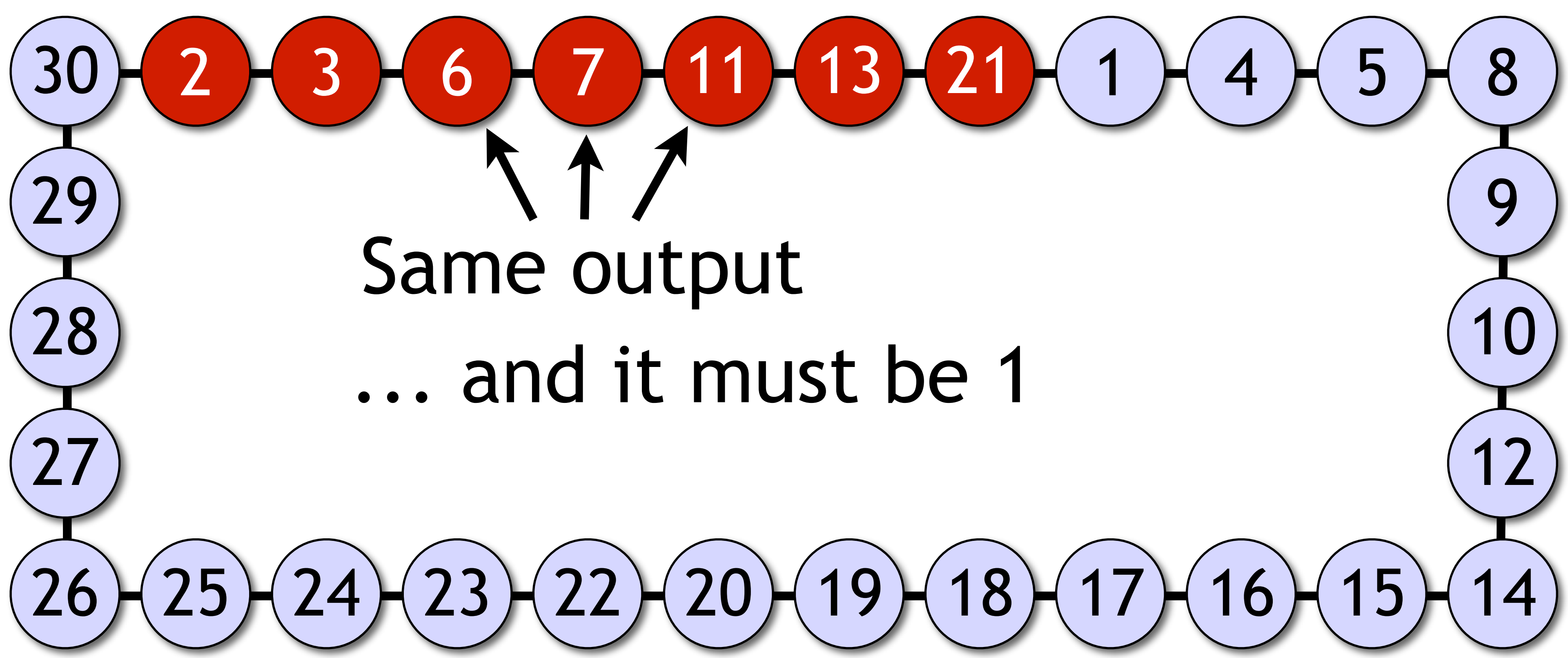
Lower-bound result for vertex cover approximation

- Construct a numbered directed cycle: monochromatic subset as consecutive nodes



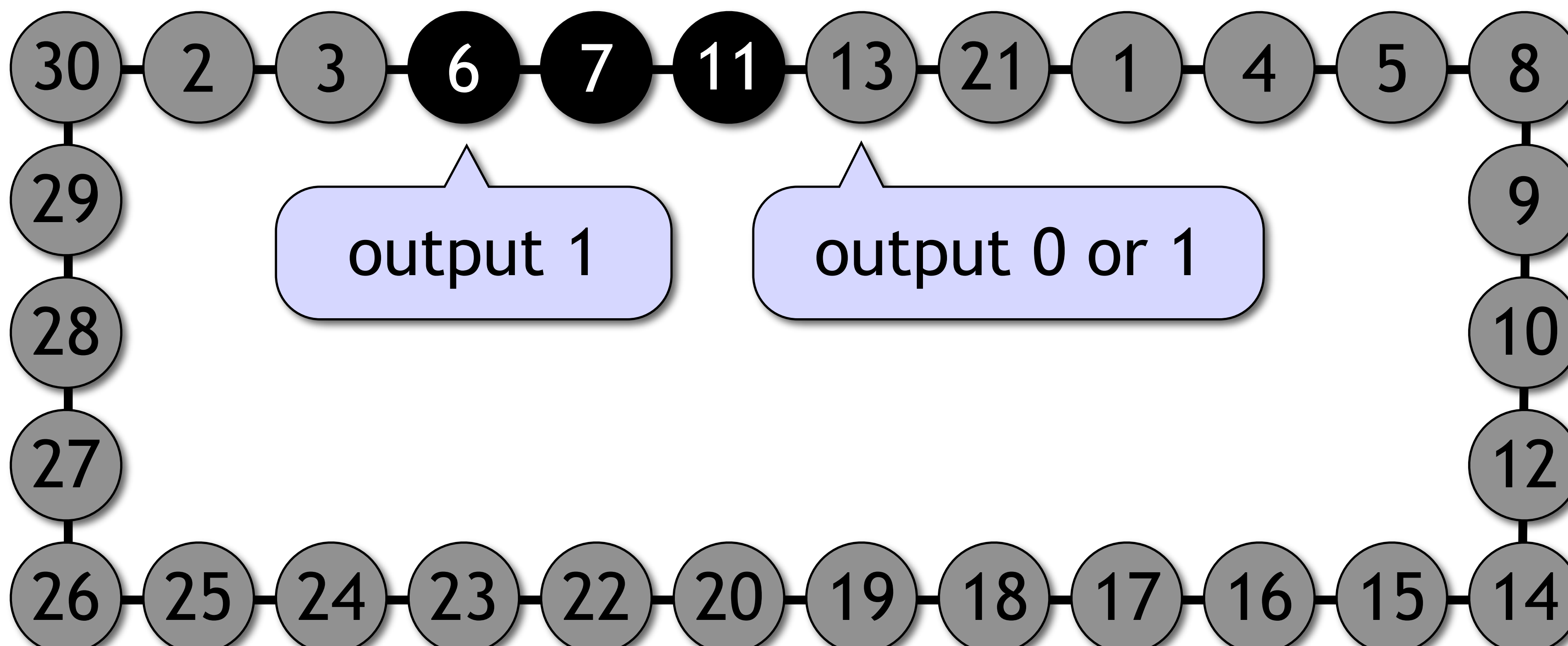
Lower-bound result
for vertex cover approximation

- Construct a numbered directed cycle:
  monochromatic subset as consecutive nodes



Lower-bound result for vertex cover approximation

- Hence there is an n-cycle with a chain of $m - 2T$ nodes that output 1

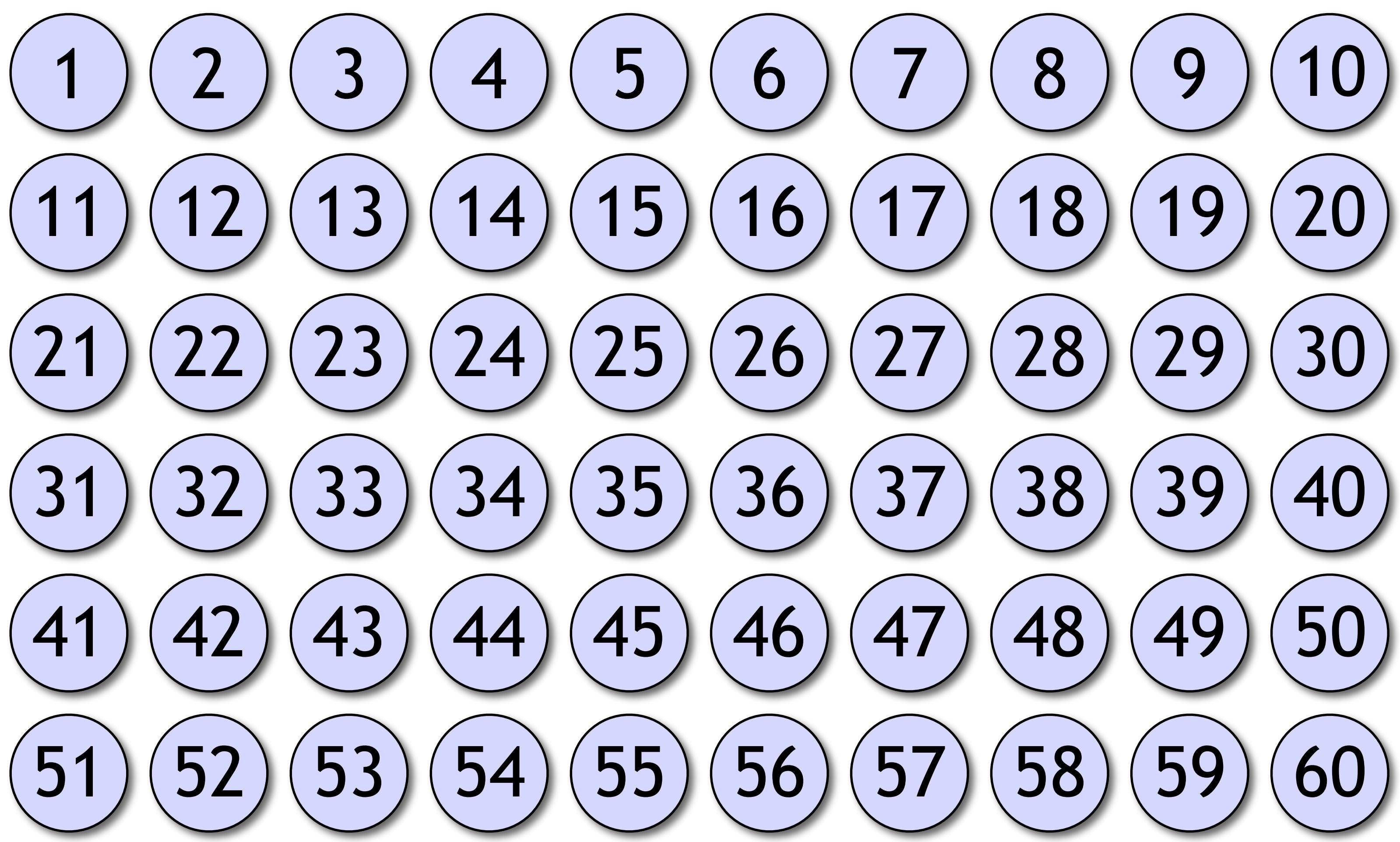


Lower-bound result for vertex cover approximation

- Hence there is an n-cycle with a chain of $m - 2T$ nodes that output 1
- We can choose as large m as we want
  - Good, more “black” nodes that output 1
- However, n increases rapidly if we increase m
  - Bad, more “grey” nodes that might output 0
- Trick: choose “unnecessarily large” n so that we can apply Ramsey’s theorem repeatedly

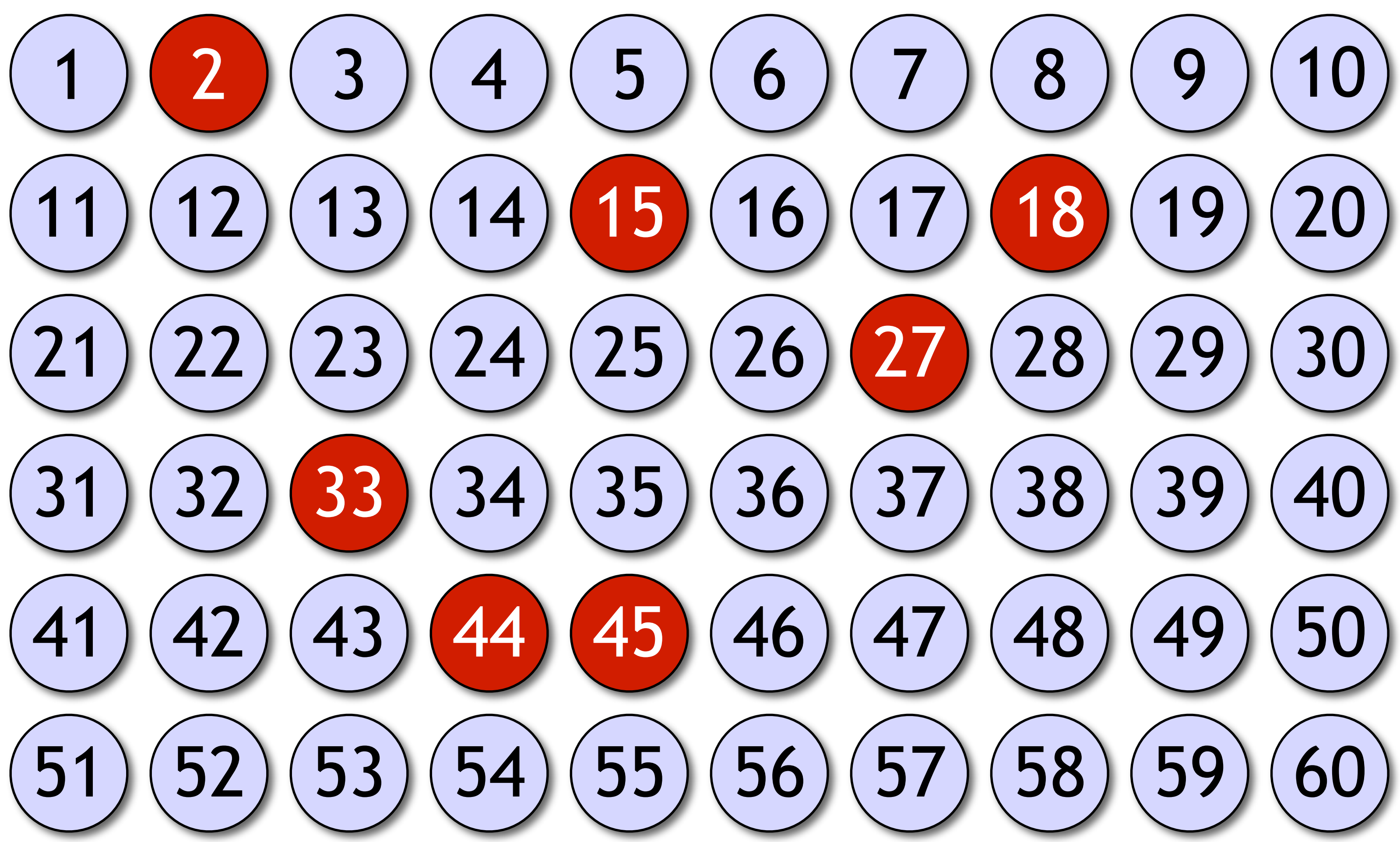
Lower-bound result for vertex cover approximation

- Huge ID space...



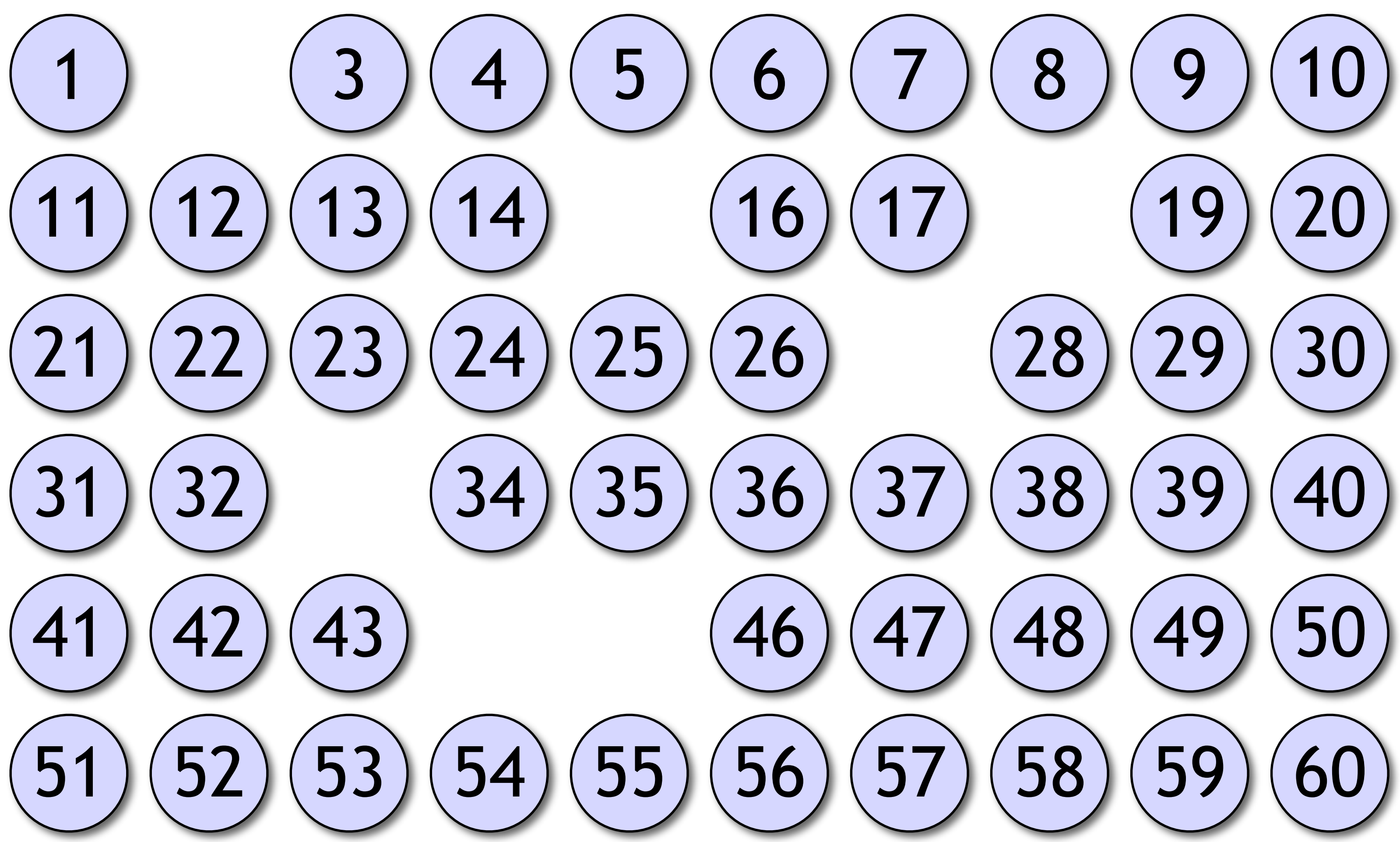
Lower-bound result for vertex cover approximation

- Find a monochromatic subset of size *m*...



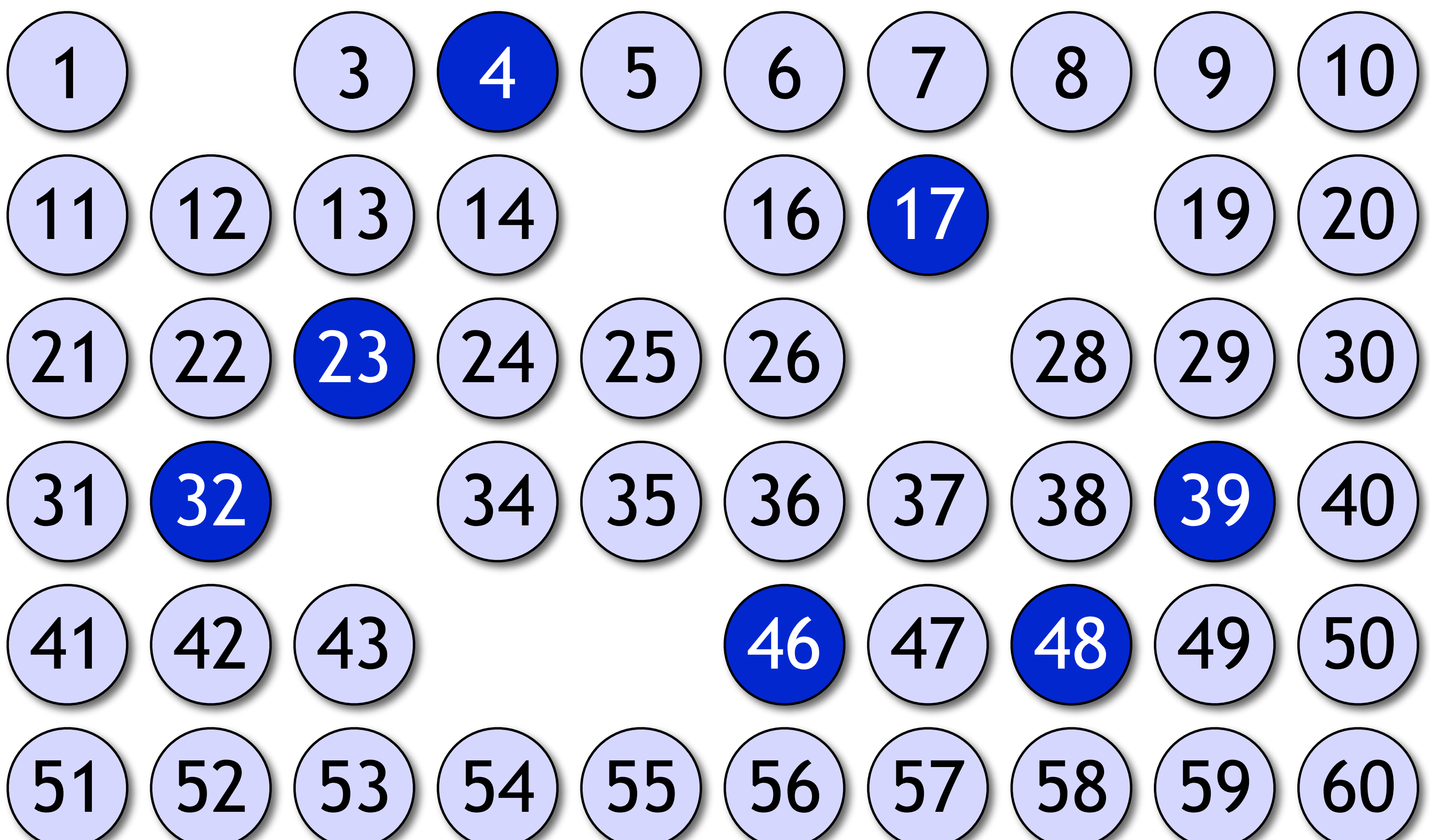
Lower-bound result for vertex cover approximation

- Delete these IDs...



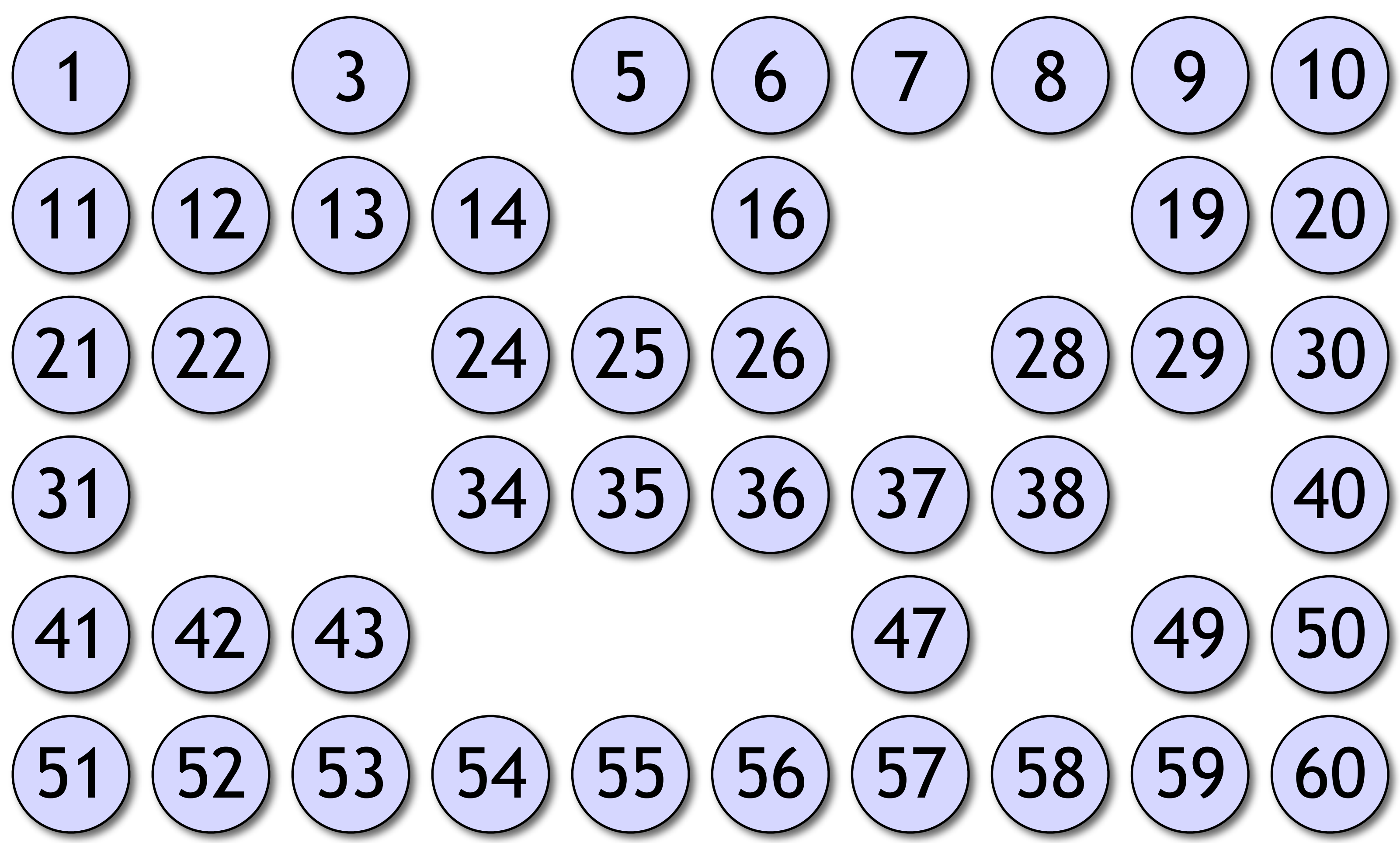
Lower-bound result for vertex cover approximation

- Still sufficiently many IDs to apply Ramsey...



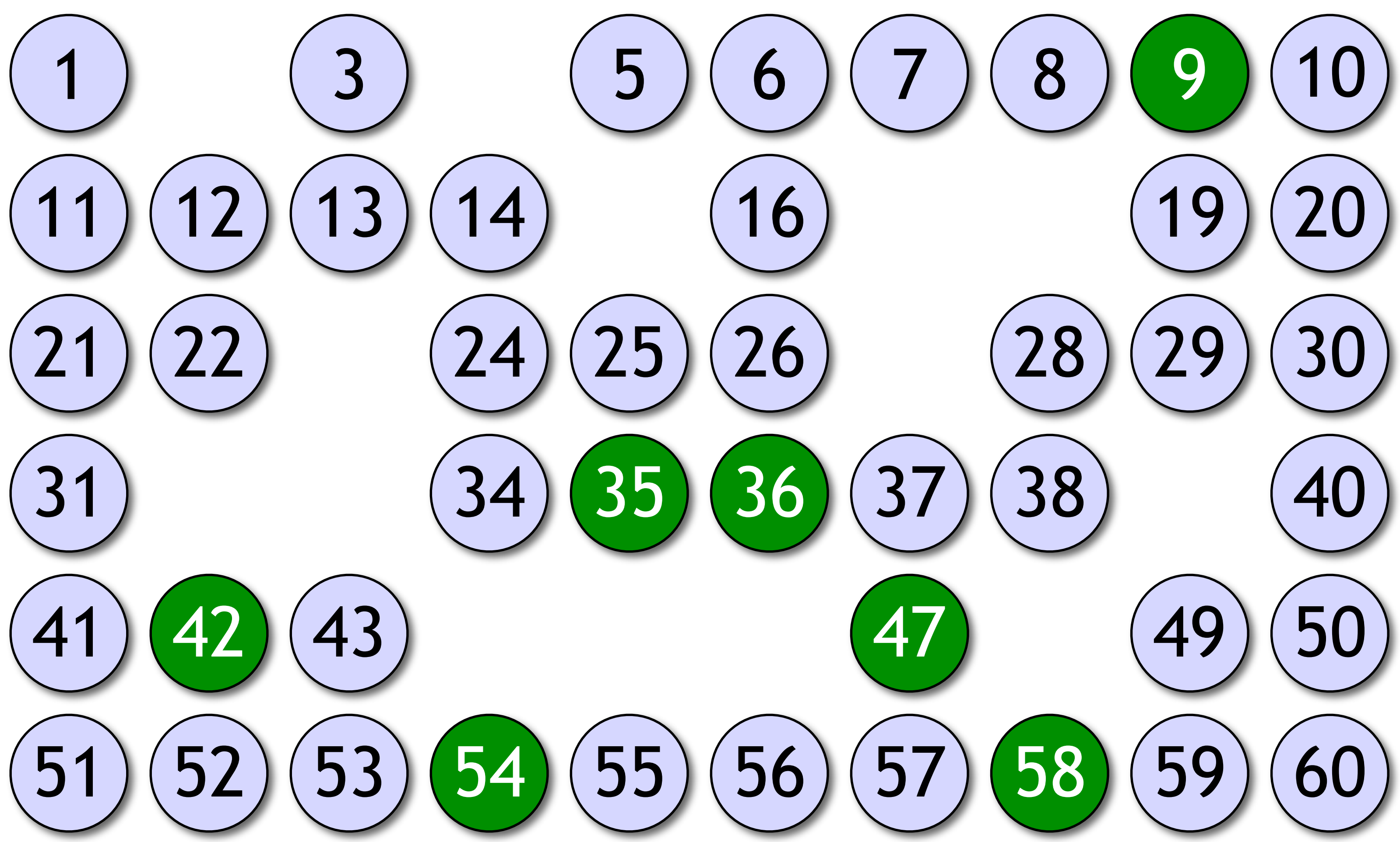
Lower-bound result for vertex cover approximation

- Repeat...



Lower-bound result for vertex cover approximation

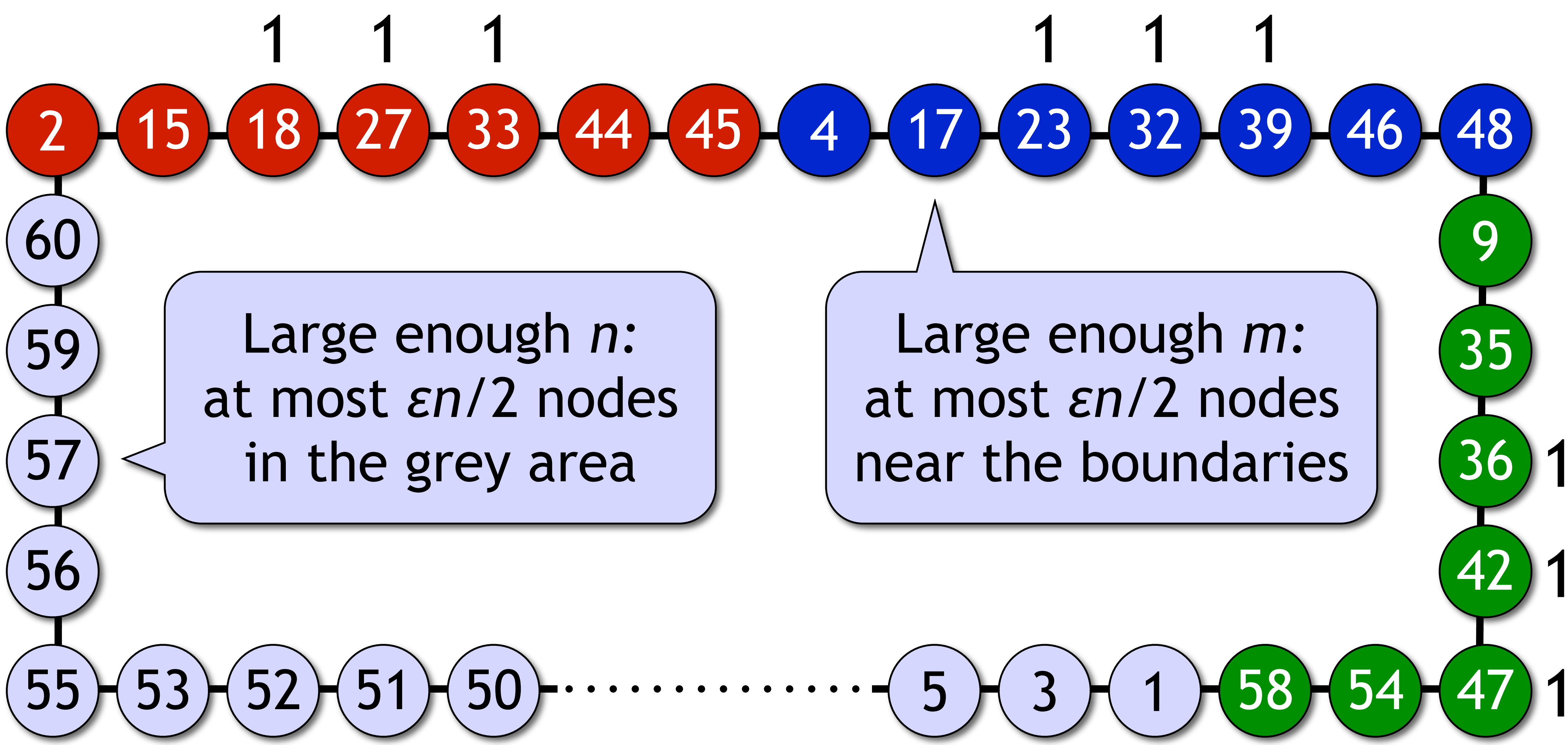
- Repeat until stuck



Lower-bound result for vertex cover approximation

- Several monochromatic subsets + some leftovers

Lower-bound result for vertex cover approximation



Lower-bound result for vertex cover approximation

- Thus A outputs a vertex cover with ≥ $(1 - \varepsilon)n$ nodes