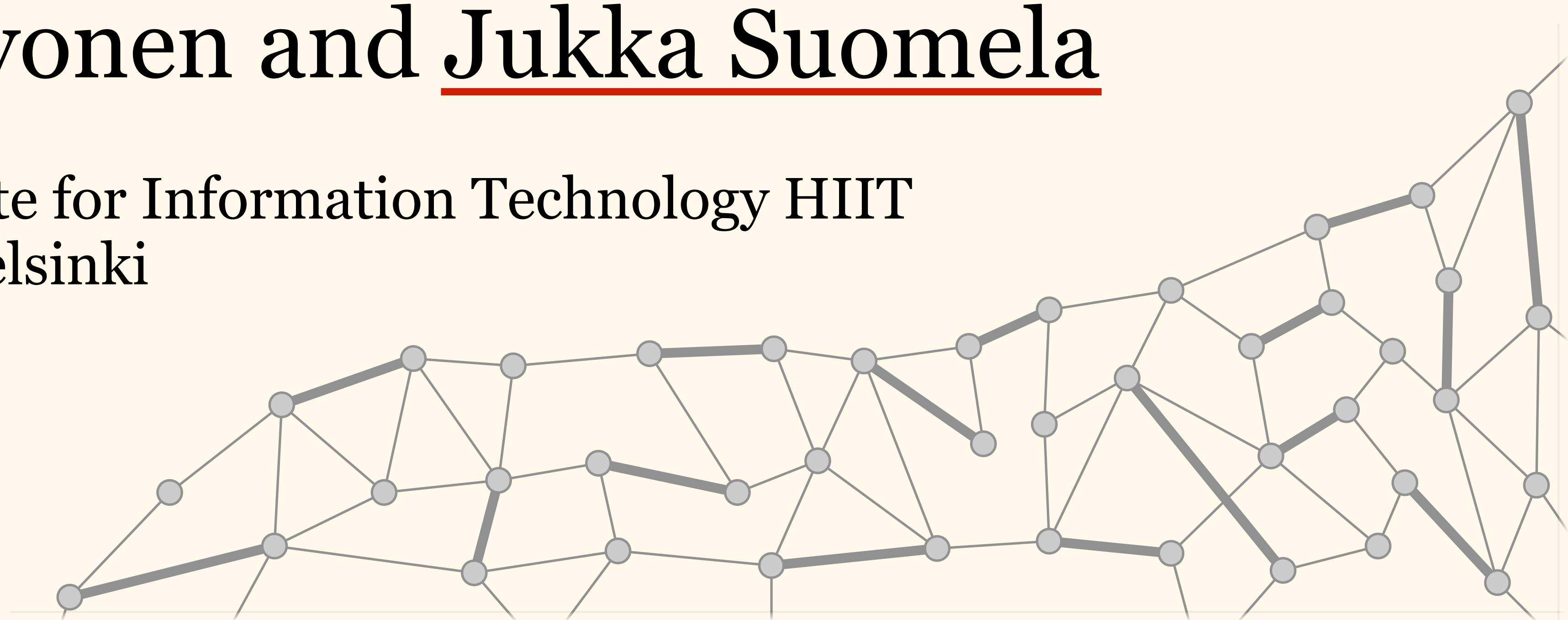
Distributed Maximal Matching: Greedy is Optimal

Juho Hirvonen and Jukka Suomela

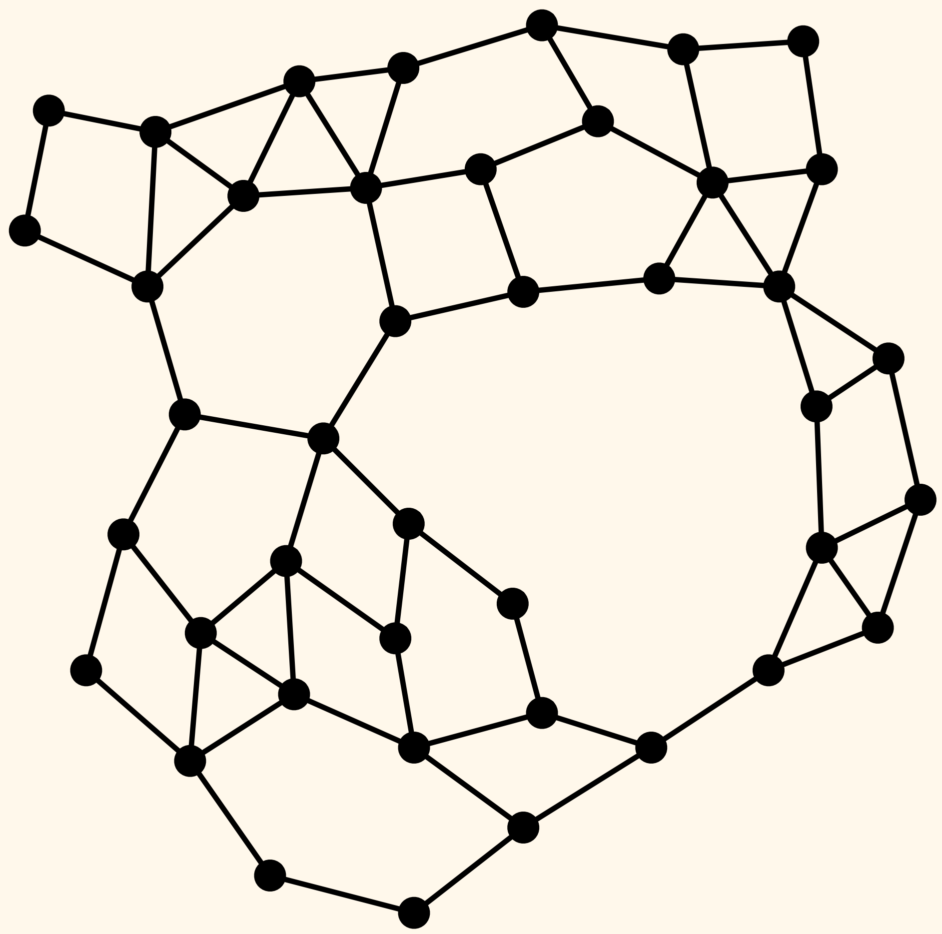
Helsinki Institute for Information Technology HIIT
University of Helsinki
Finland



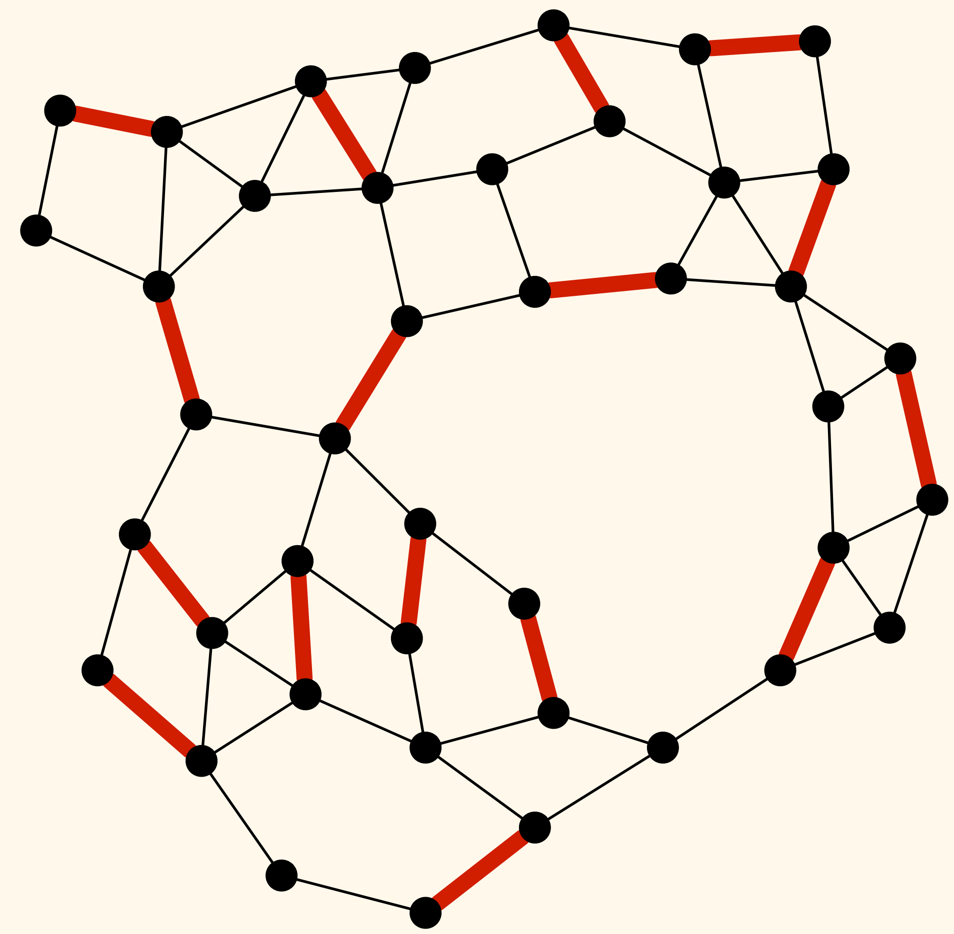
17 July 2012
PODC

Maximal Matchings

Input

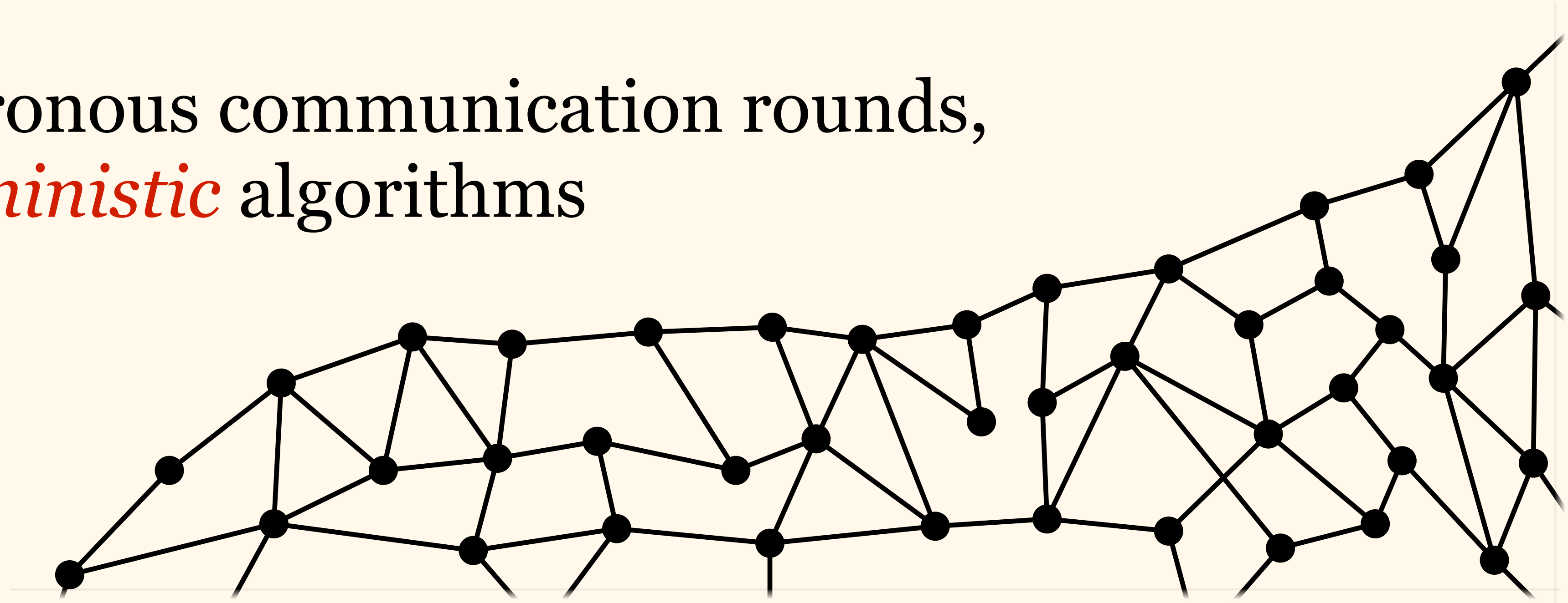


Output



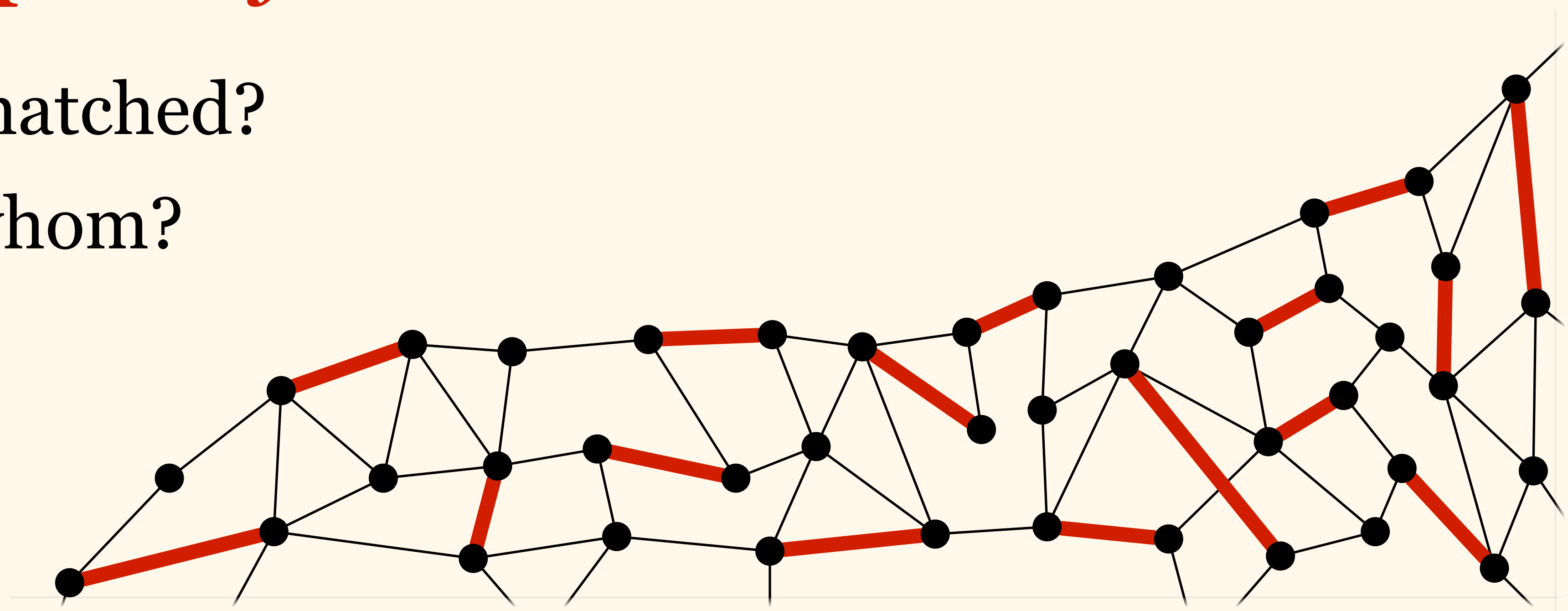
Distributed Algorithms

- Graph G = input = communication network
  - node = computer
  - edge = communication link
  - synchronous communication rounds, *deterministic* algorithms



Distributed Algorithms

- Graph G = input = communication network
- Each node has to stop and output its *own part of the solution*
  - am I matched?
  - with whom?



Distributed Algorithms

- Time = number of communication rounds
- Equivalent:
  - *running time is T*
  - all nodes stop after *T communication rounds*
  - output of node v = f(*radius-T neighbourhood* of v)
- **How fast** can we find a maximal matching?

Time = $f(n, \Delta)$

- $\boldsymbol{n}$ = number of nodes
- $\boldsymbol{\Delta}$ = maximum degree
- Focus: $\boldsymbol{\Delta \ll n}$

problem	*upper bound*	*lower bound*
maximal matching	$\Delta + \log^* n$	$\text{polylog}(\Delta) + \log^* n$
(Δ+1)-vertex colouring	$\Delta + \log^* n$	$\log^* n$
(2Δ–1)-edge colouring	$\Delta + \log^* n$	$\log^* n$
maximal edge packing	Δ	$\text{polylog}(\Delta)$
vertex cover 2-approx.	Δ	$\text{polylog}(\Delta)$

positive: *Panconesi–Rizzi* (2001), *Barenboim–Elkin* (2009), *Kuhn* (2009), *Åstrand–Suomela* (2010), ...

negative: *Linial* (1992), *Kuhn et al.* (2004, 2006)

Time = $f(n, \Delta)$

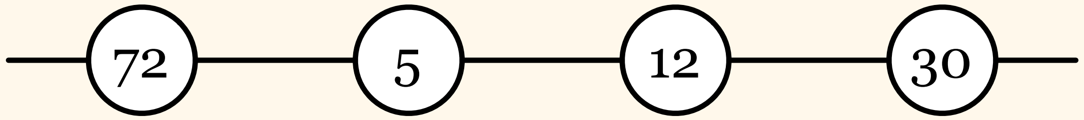
- Fairly well-understood as a function of n
  - tight upper and lower bounds if $\Delta = O(1)$
- Wide open as a function of Δ
  - exponential gap
- *Linear-in-Δ lower bounds missing*

Plan

1. Study a simpler model
    - unique node identifiers make things complicated
    - but *what is the right model?*
2. Generalise
    - future work…
    - but see the next talk for a promising technique!

Simpler Model?

- Unique identifiers



  - standard model
  - complicated to analyse directly…
- Node colouring

  - weaker than unique identifiers

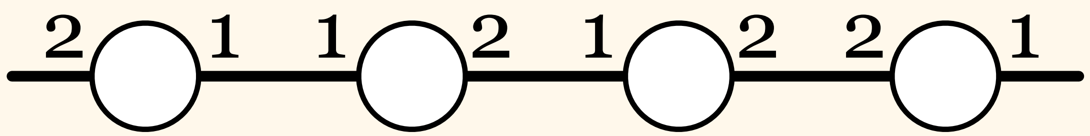
Simpler Model?

- Unique identifiers
  - standard model
  - complicated to analyse directly...
- Node colouring

  - weaker than unique identifiers
  - *too weak — cannot find a maximal matching!*

Simpler Model?

- Port numbering



  - another popular model
  - too weak...

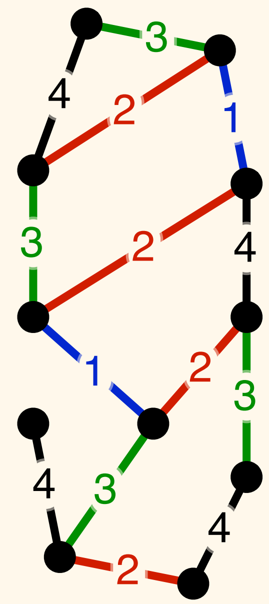
- **Edge colouring**

  - stronger than port numbering
  - *just right for our purposes!*

Model

- Given: *k*-edge-coloured graph

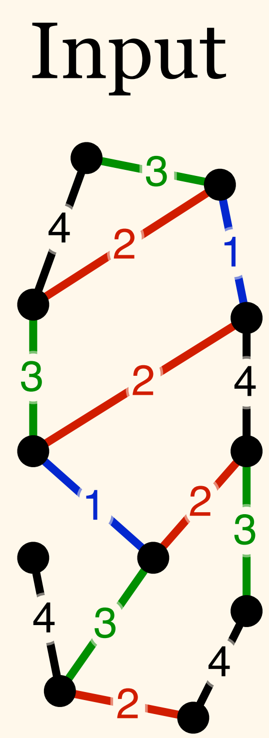
Input

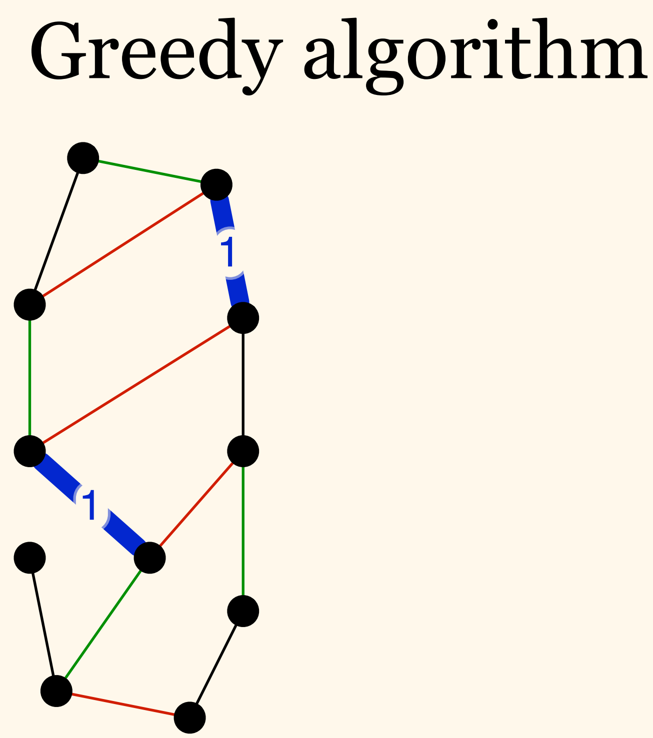


- proper edge colouring:
  adjacent edges have different colours
- colour palette: $\{1, 2, ..., k\}$
- anonymous nodes
- nodes can use edge colours to
  refer to their neighbours

Greedy Algorithm

- Greedily add edges of colour **1**, …

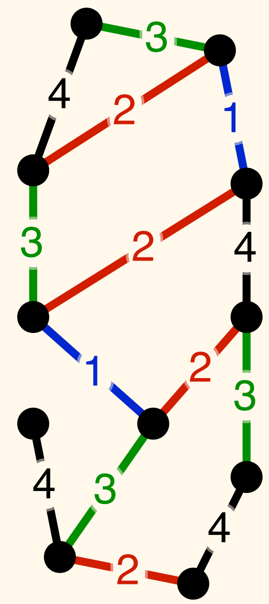




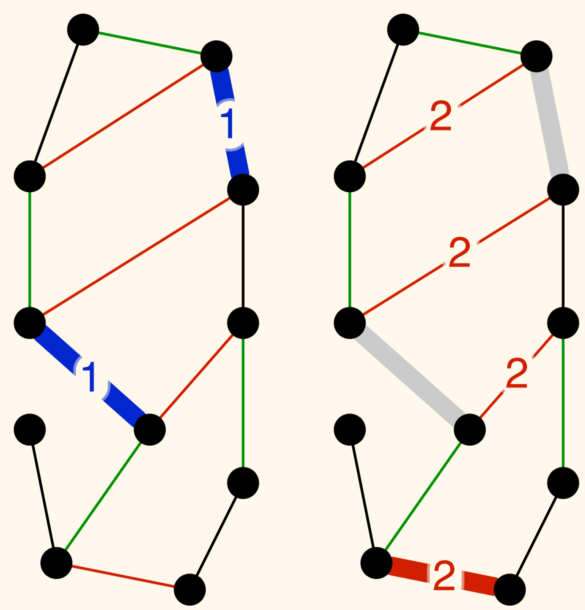
Greedy Algorithm

- Greedily add edges of colour 1, **2**, …

Input

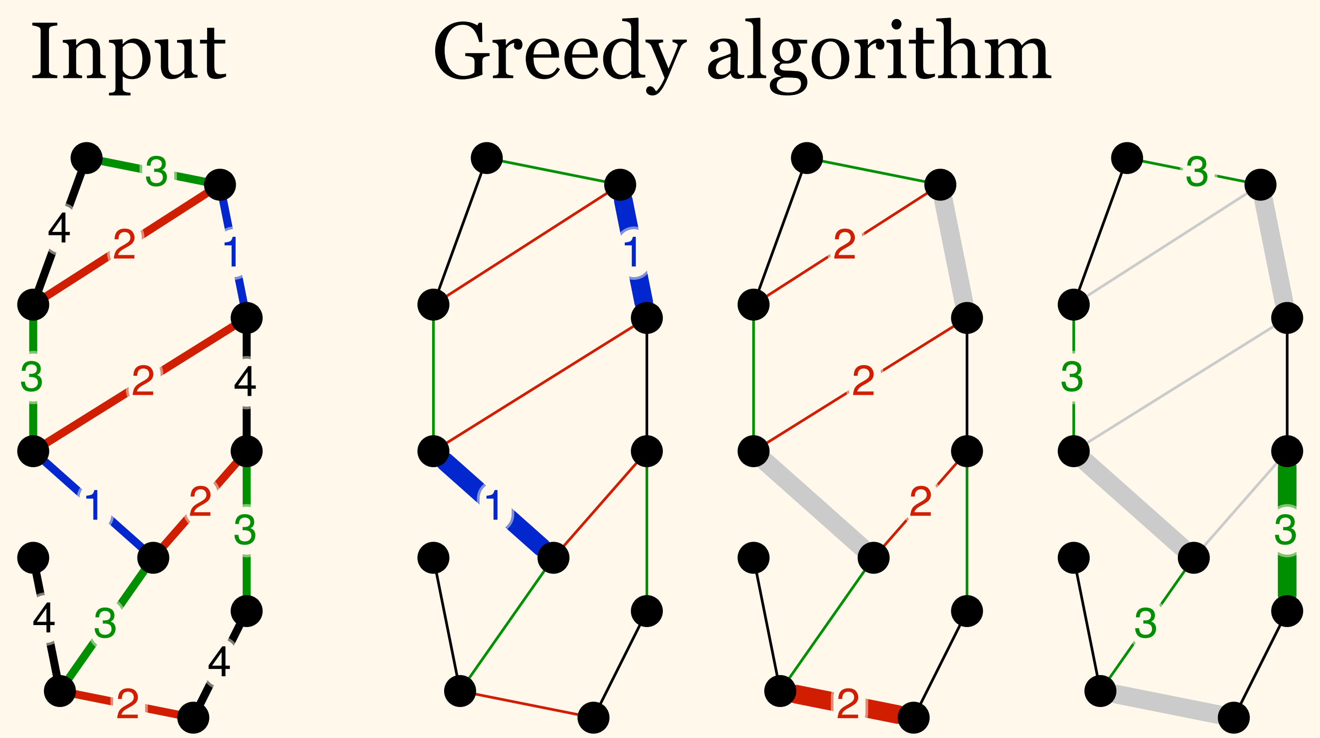


Greedy algorithm



Greedy Algorithm

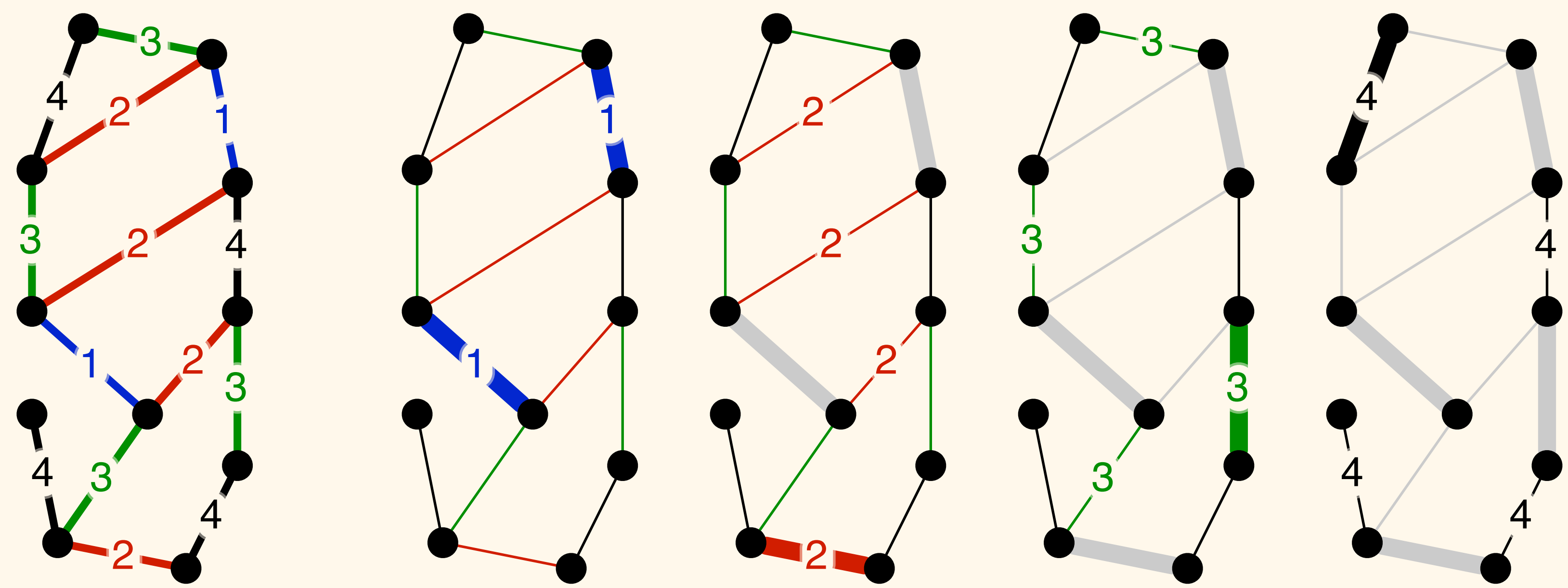
- Greedily add edges of colour 1, 2, **3**, …



Greedy Algorithm

- Greedily add edges of colour 1, 2, …, k

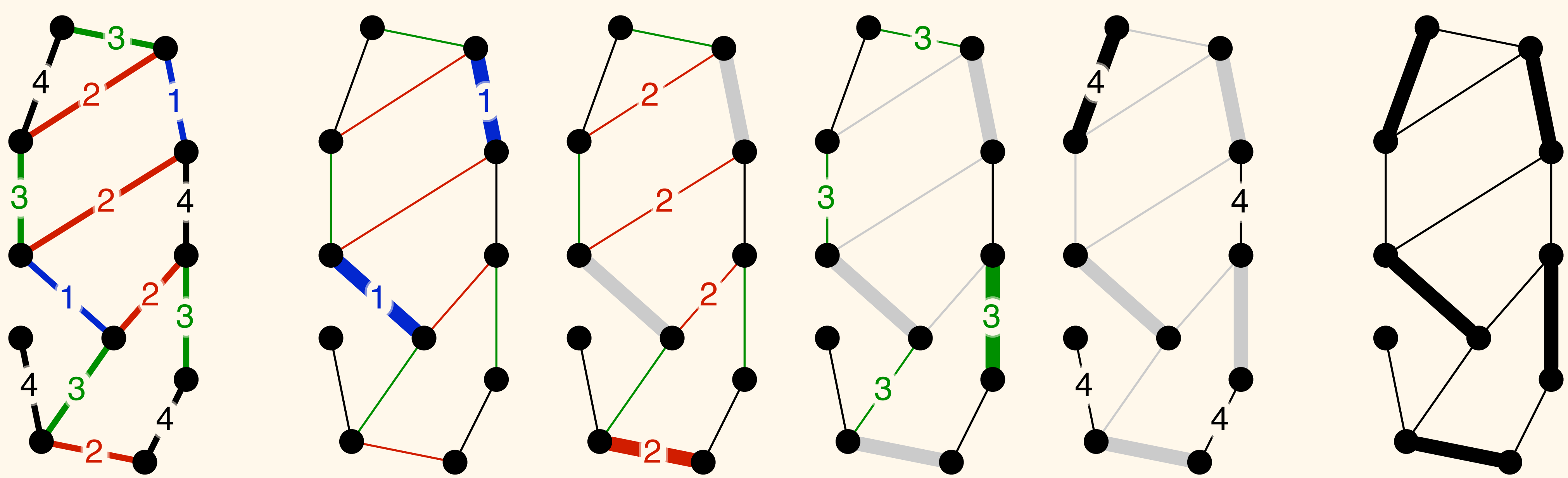
Input Greedy algorithm



Greedy Algorithm

- That’s it – we have a maximal matching

Input Greedy algorithm Output



Greedy Algorithm

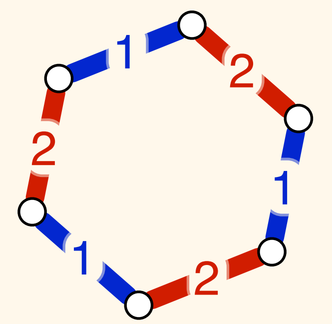
- Running time is exactly $k - 1$ rounds
  - initially each node knows the colours of incident edges
  - analysis is tight
- But *is there a faster algorithm*?

Contributions

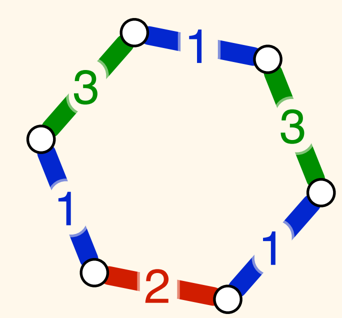
- Maximal matchings in k-edge-coloured graphs
- General graphs: $\geq k - 1$ rounds
  - matching upper bound: greedy
- Bounded-degree graphs: $\Omega(\Delta + \log^* k)$
  - matching upper bound: adaptation of *Panconesi–Rizzi* (2001)

Lower Bound

- *d-regular, k-edge-coloured* graphs
- $d = k$:
  - trivial to find a maximal matching in constant time (pick a colour class)



- $d = k - 1$:
  - as difficult as the general case!
  - we show that we need at least d rounds

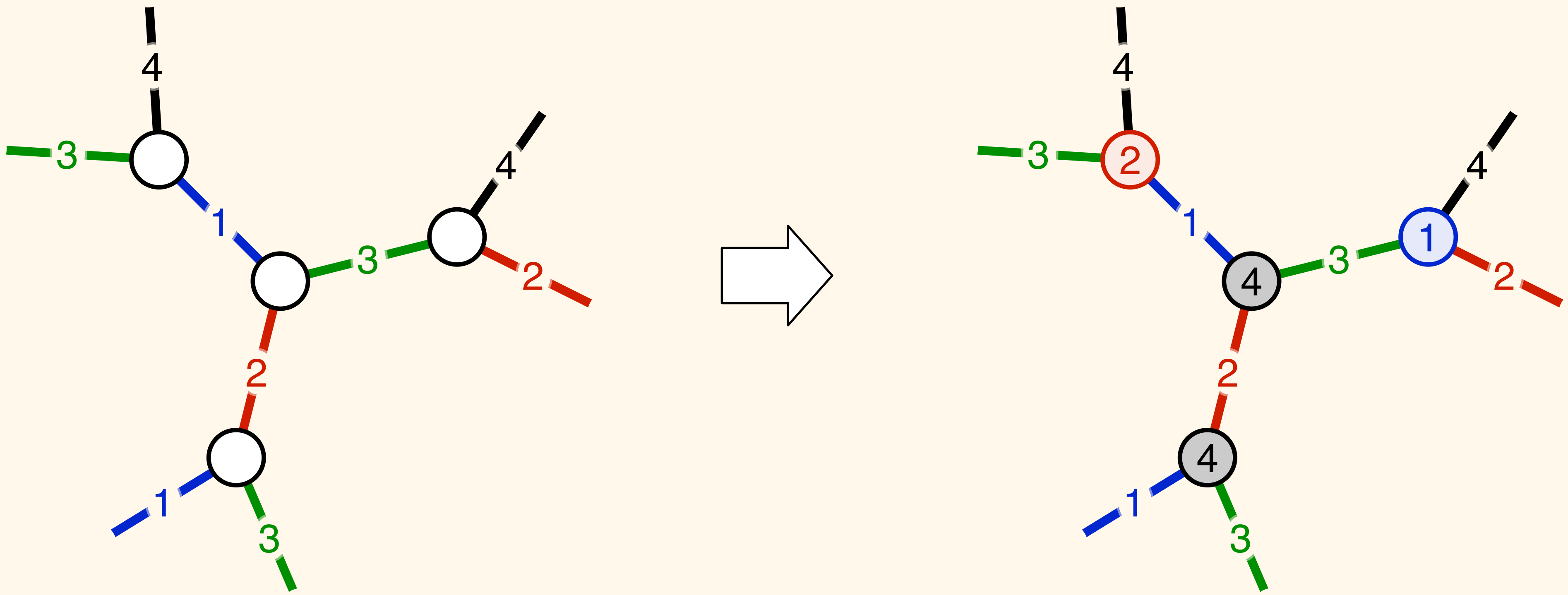


Lower Bound

- Given an algorithm A
- Construct two d-regular trees T_1 and T_2:
  - root nodes have *identical $(d - 1)$-neighbourhoods*
  - root nodes produce *different outputs*
- Running time of A is at least $d = k - 1$

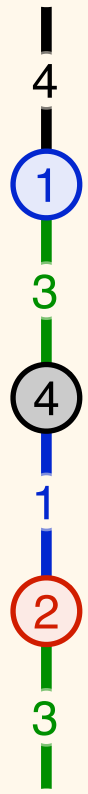
Node Colours

- Node colour = the unique "missing colour"



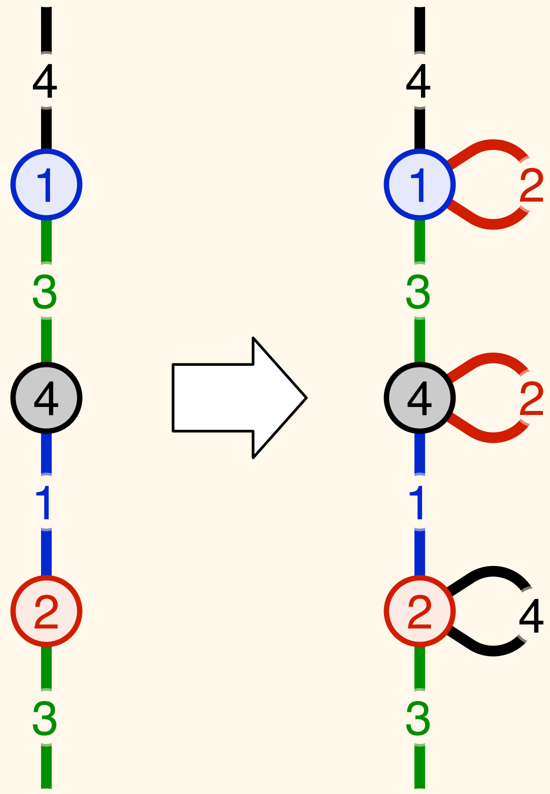
Templates

- Degree < d



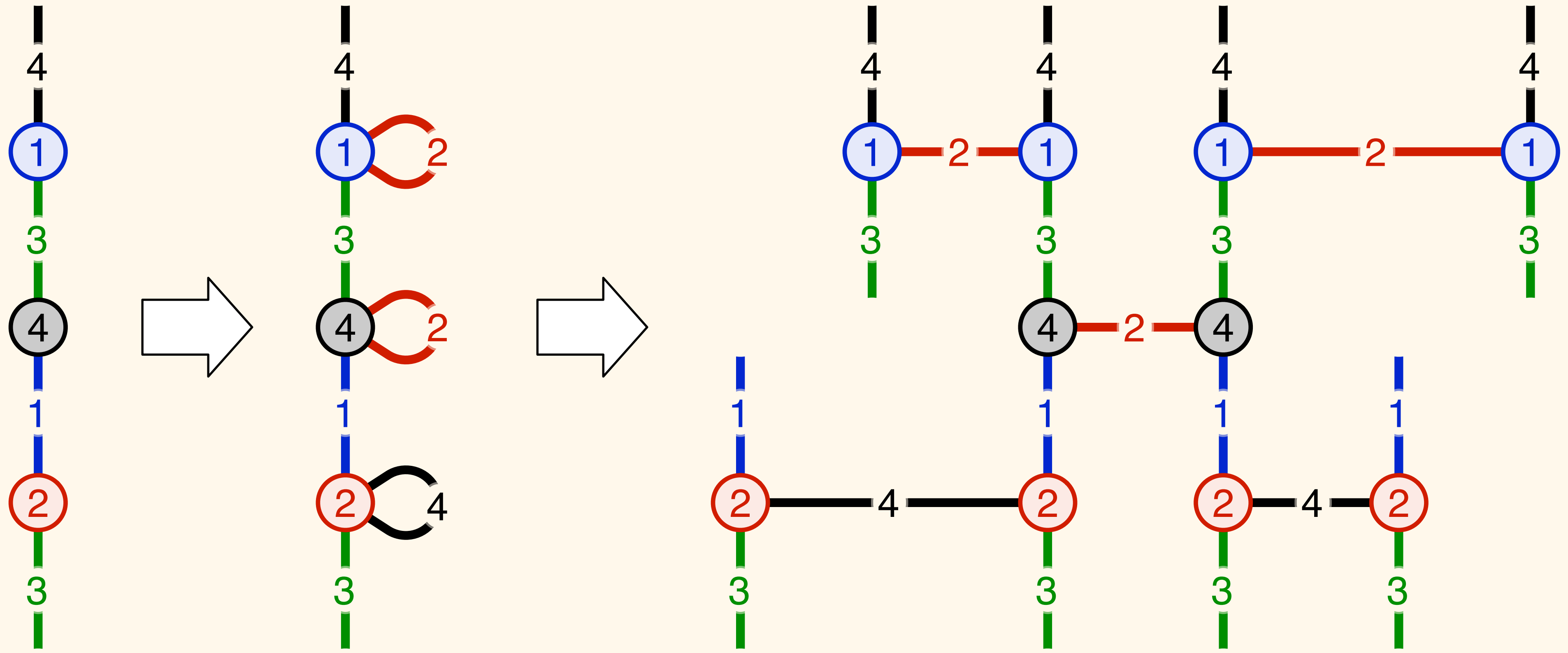
Templates

- Degree < d: add loops



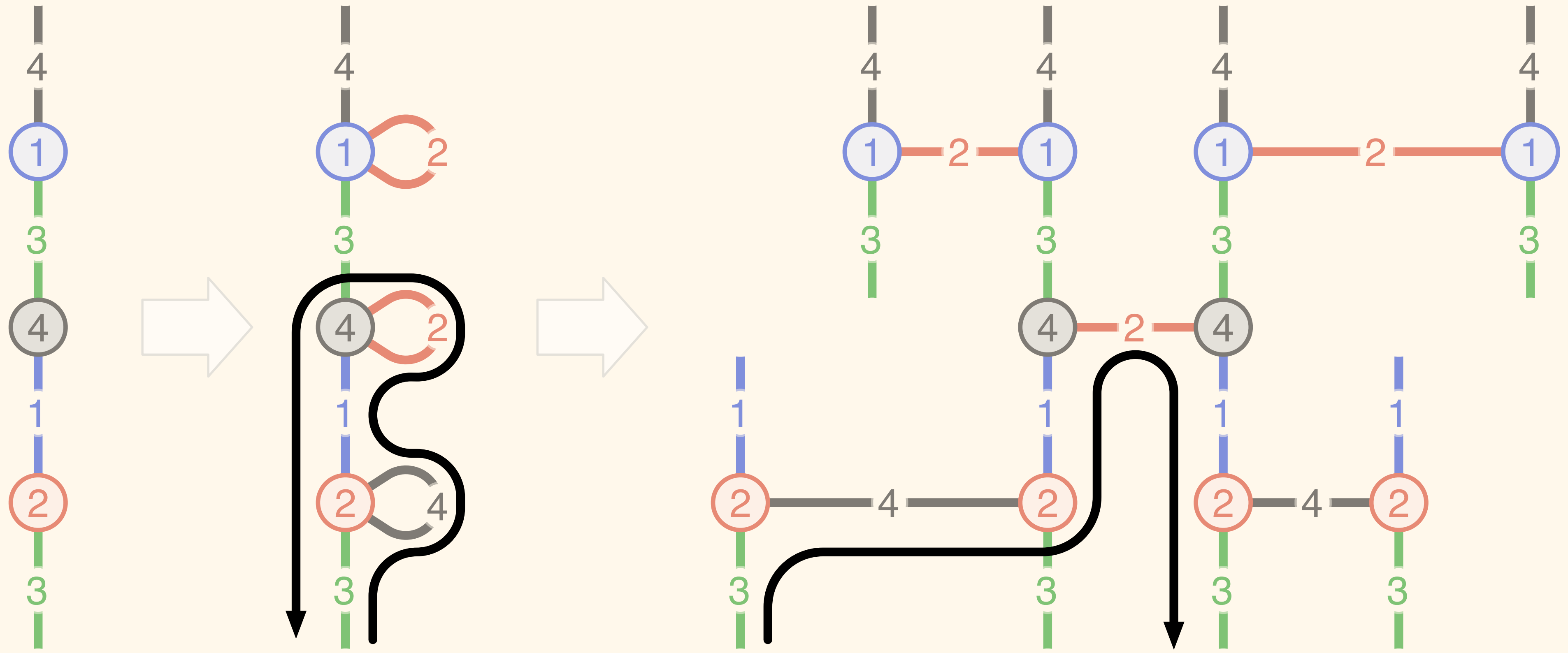
Templates

- Degree < d: add loops, unfold loops



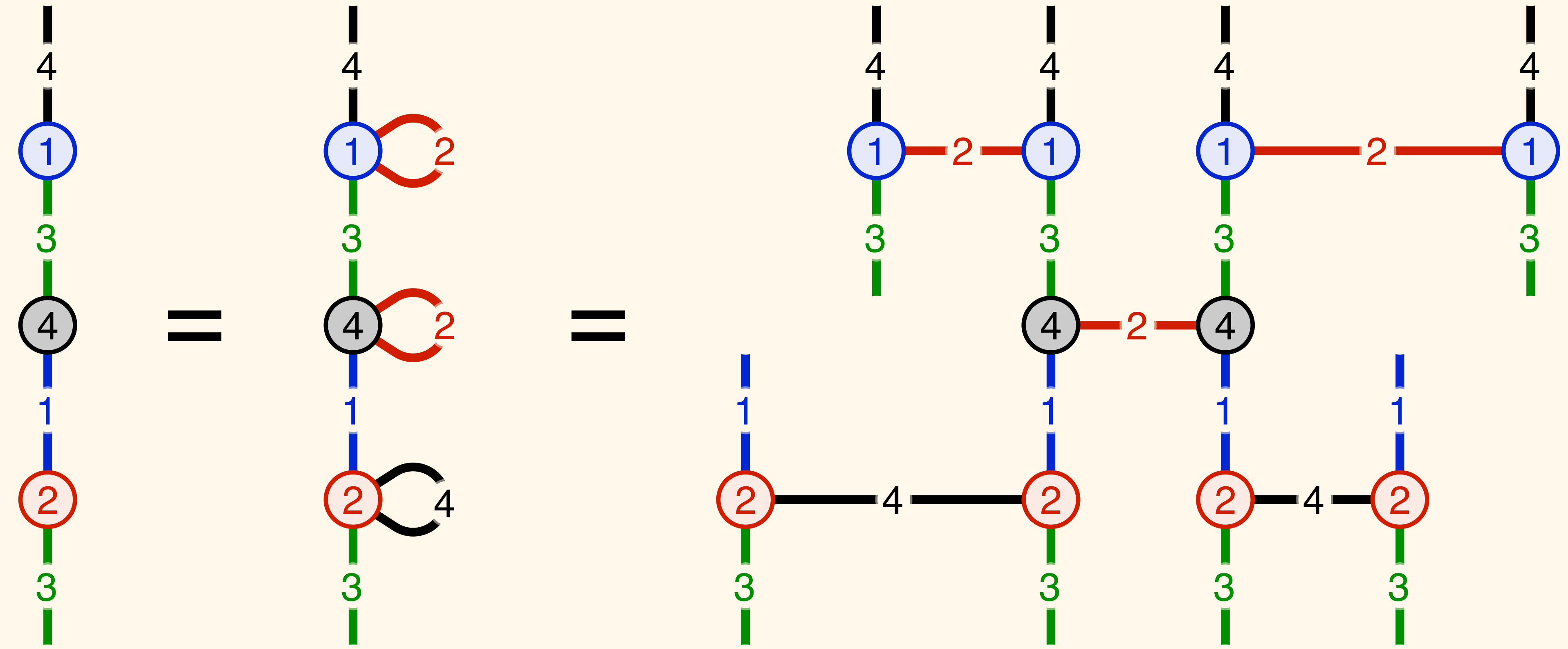
Templates

- Unfolding preserves traversals

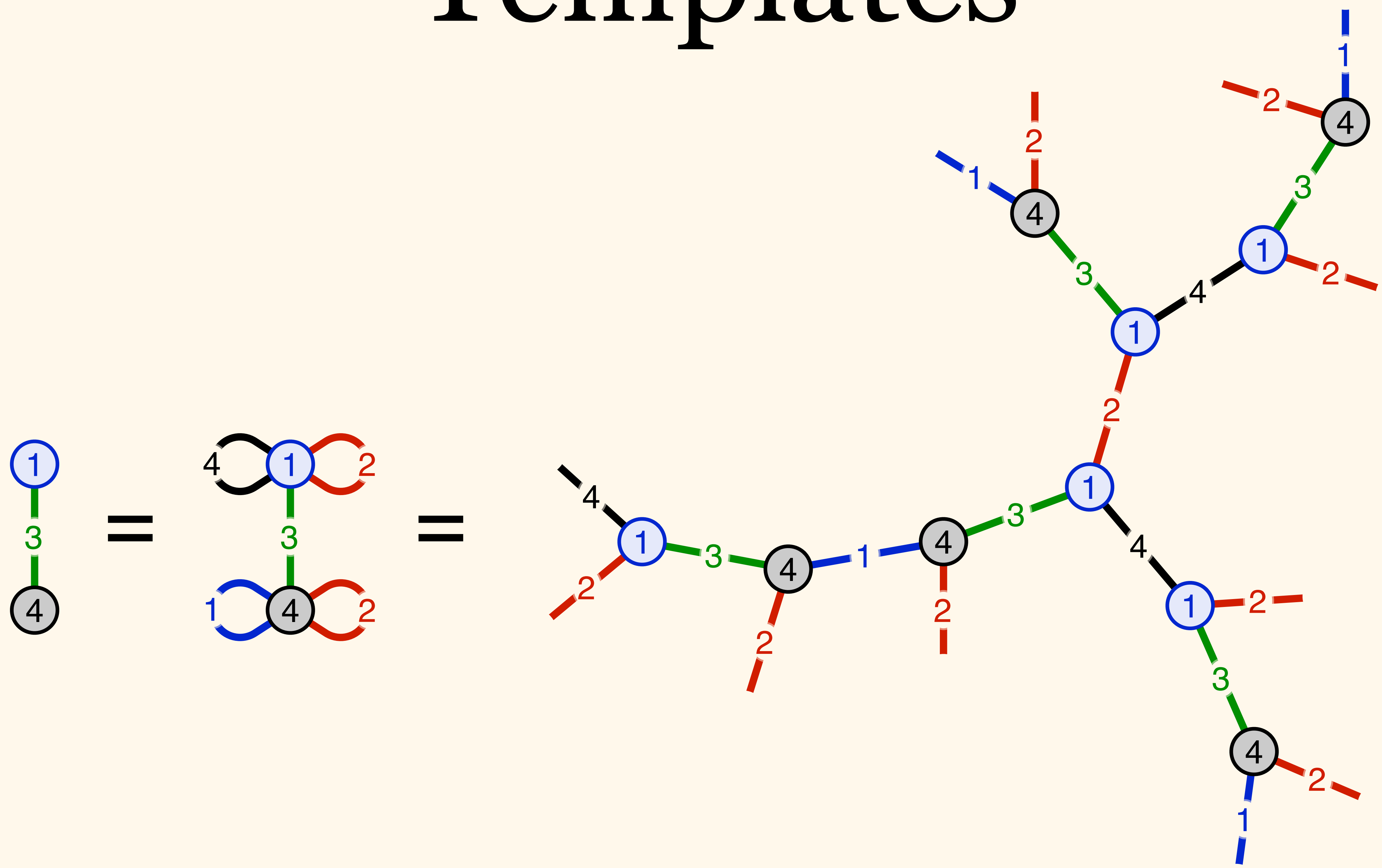


Templates

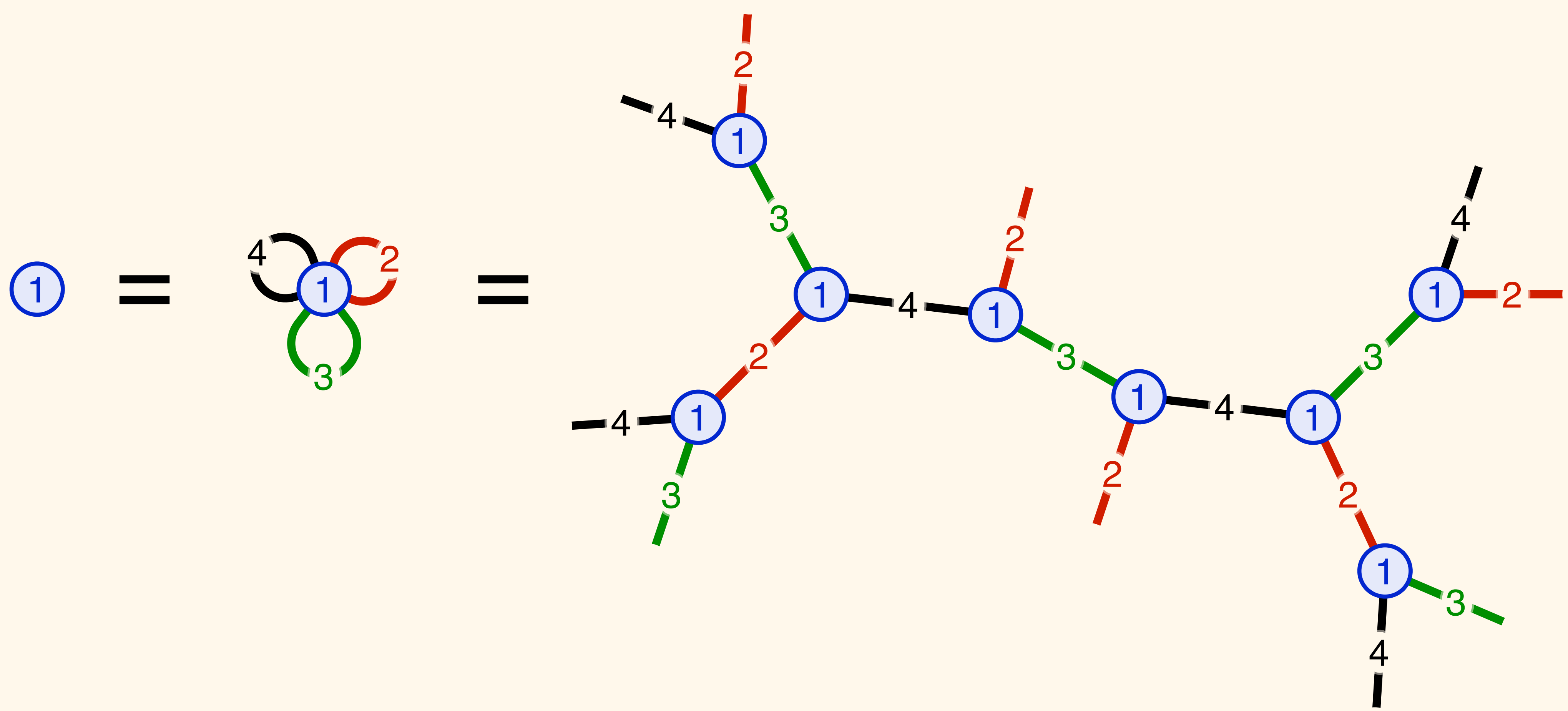
- Compact representations of trees



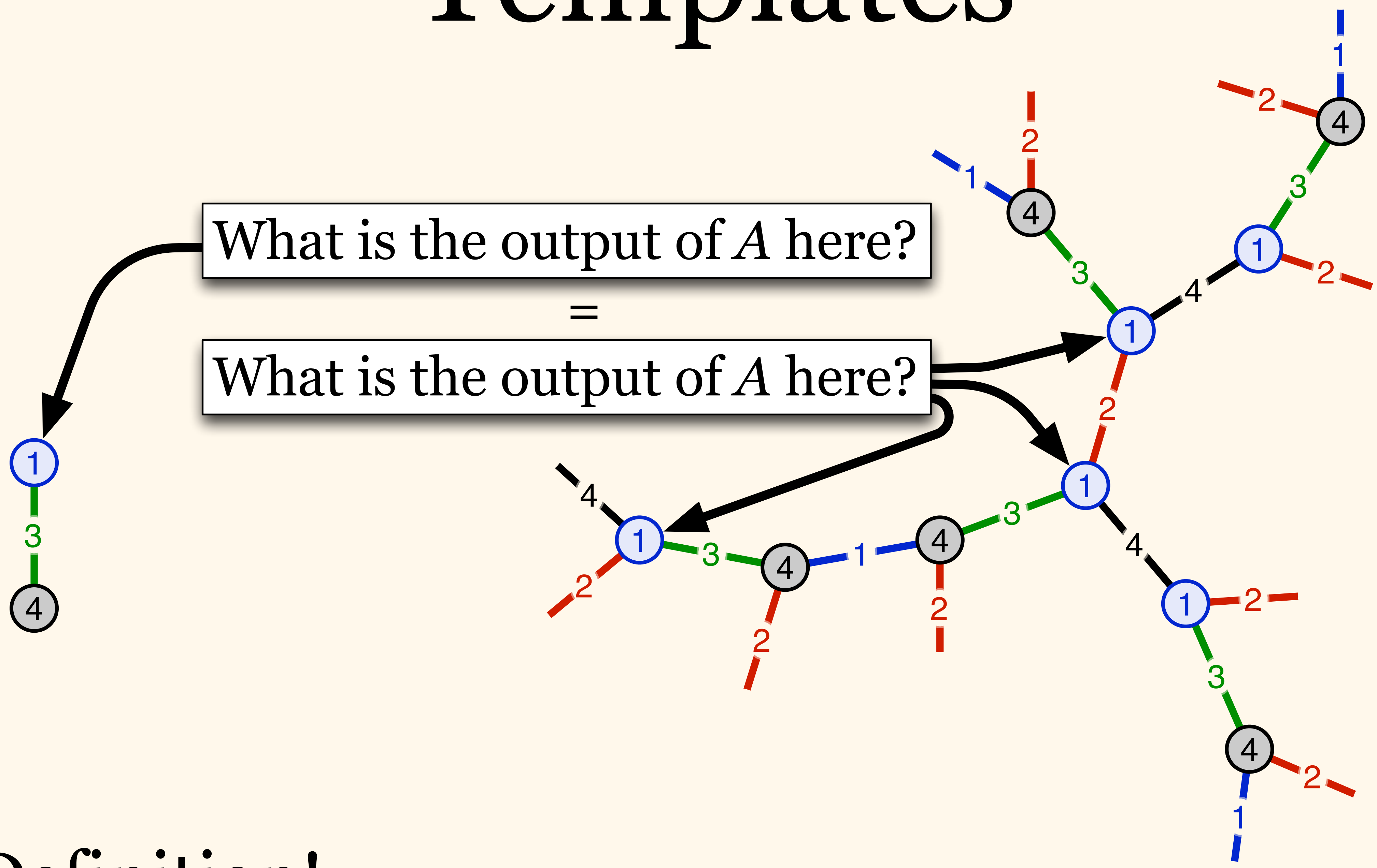
Templates



Templates



Templates

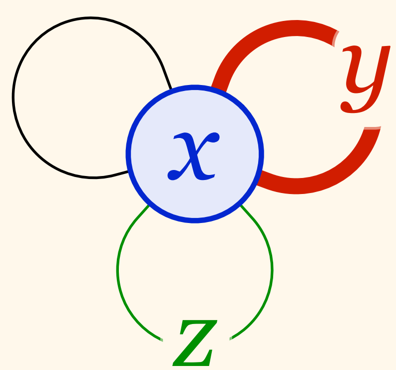


Definition!

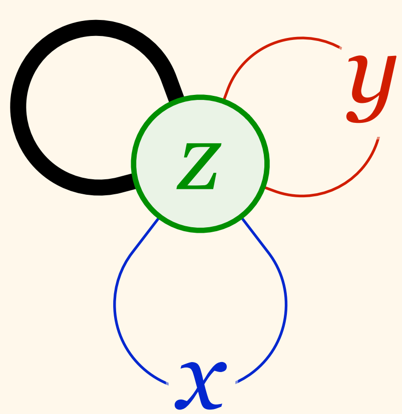
Induction

- Degree i templates:
  - root nodes produce different outputs
  - identical neighbourhoods up to distance $i - 1$
- $i = 1$: base case
- $i > 1$: by induction
- $i = d$: main result

Base Case

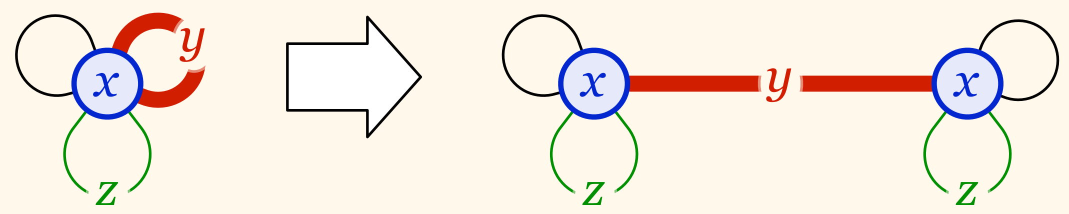


edge of colour y exists,
in matching

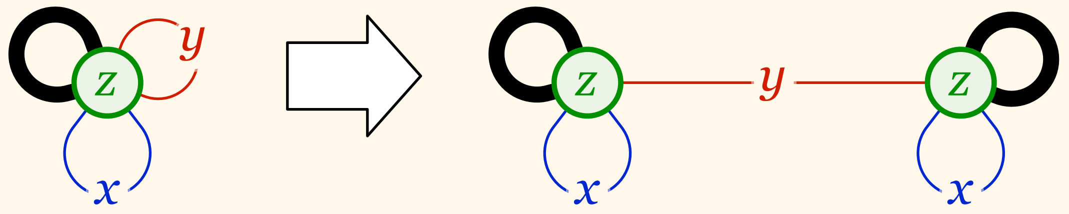


edge of colour y exists,
but not in matching

Base Case

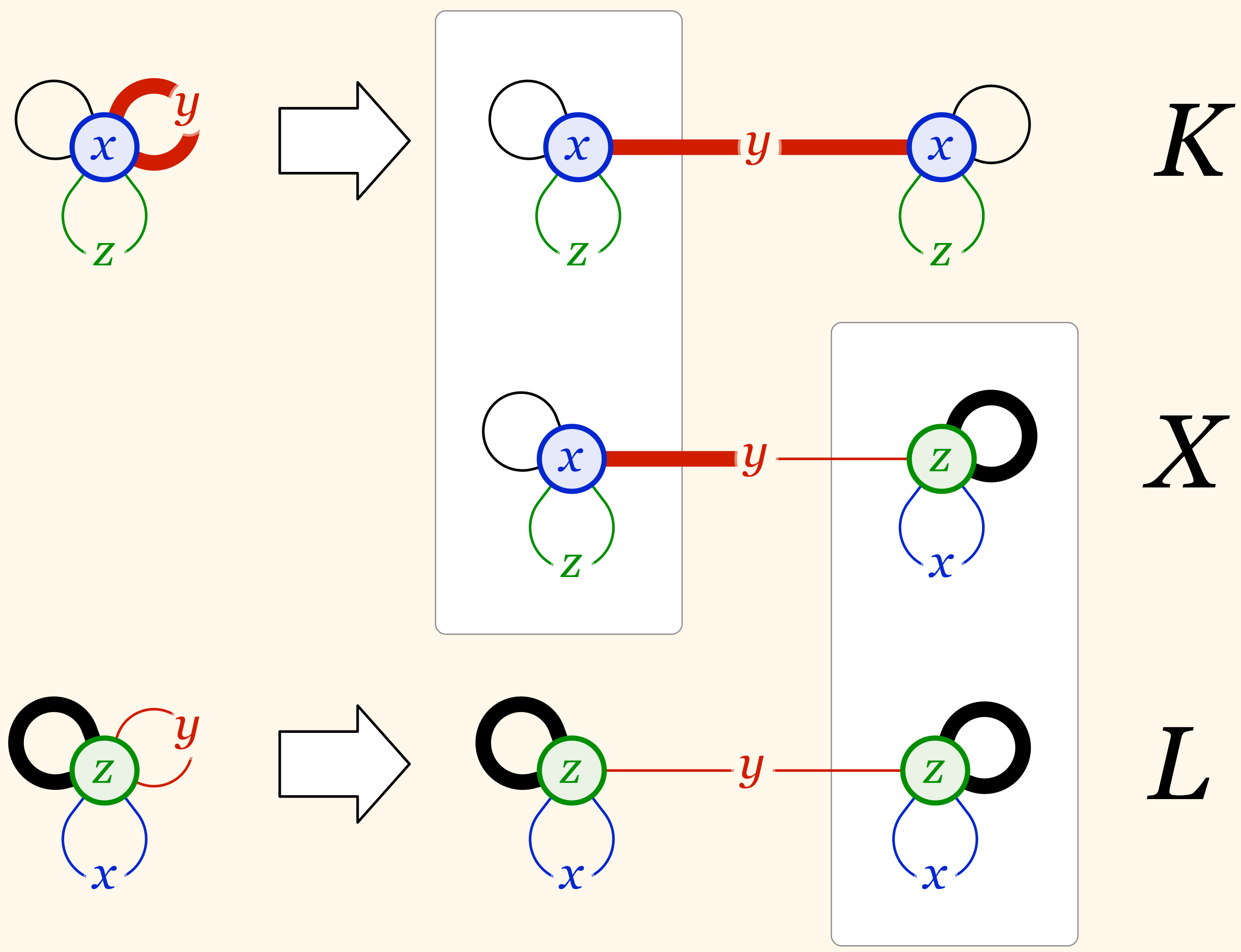


K



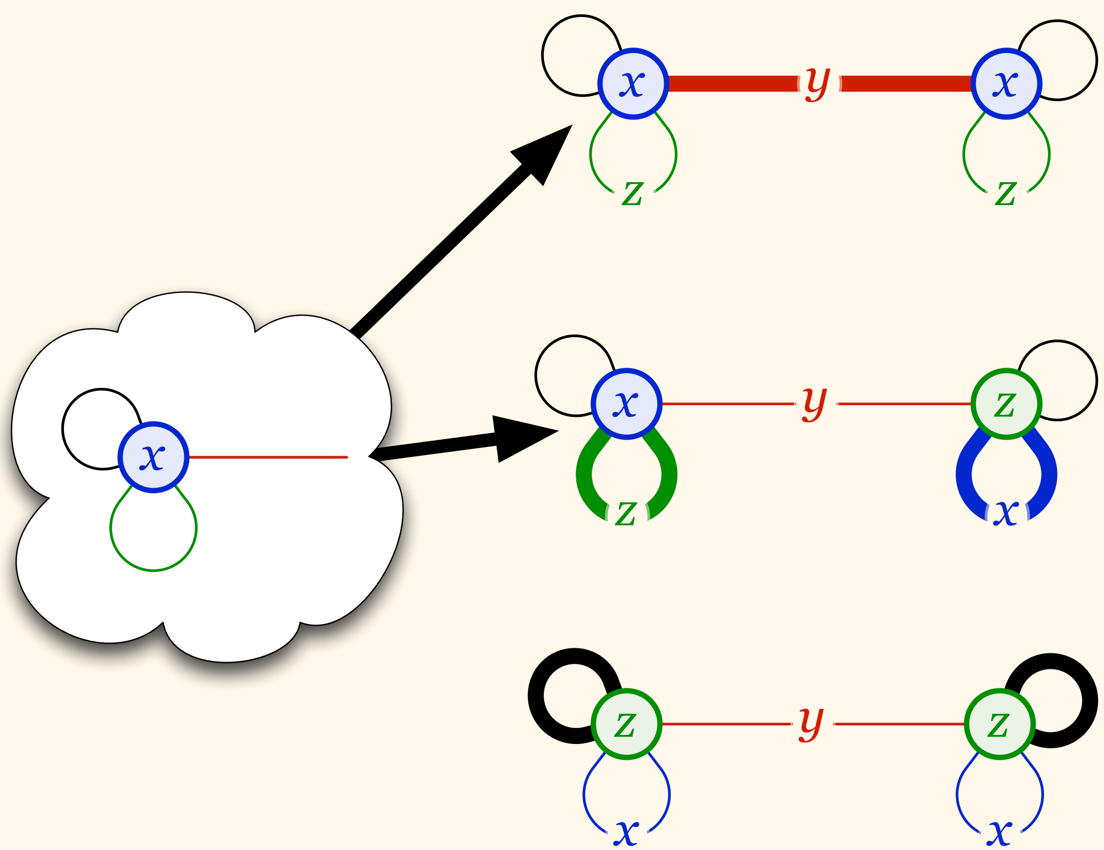
L

Base Case



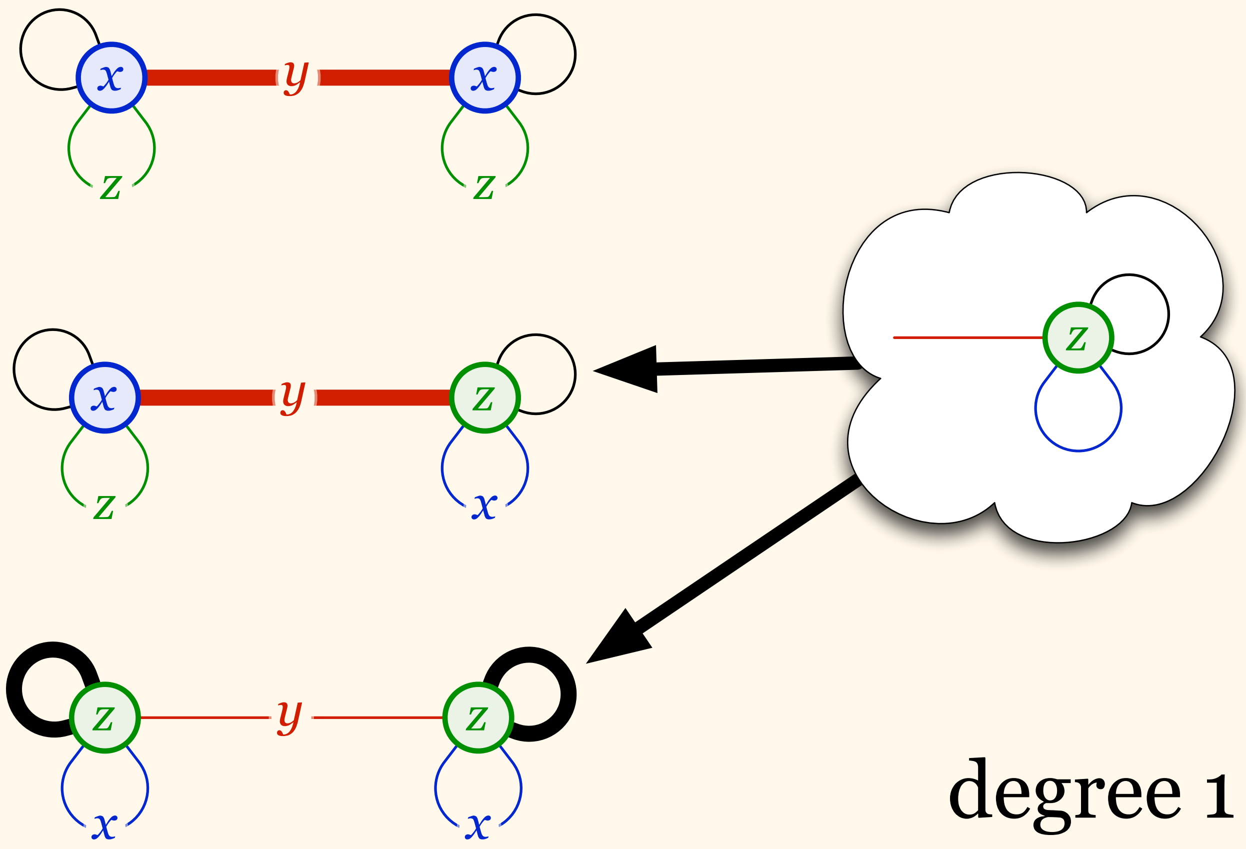
output in X cannot
be copied from K & L
– something must change!

Base Case



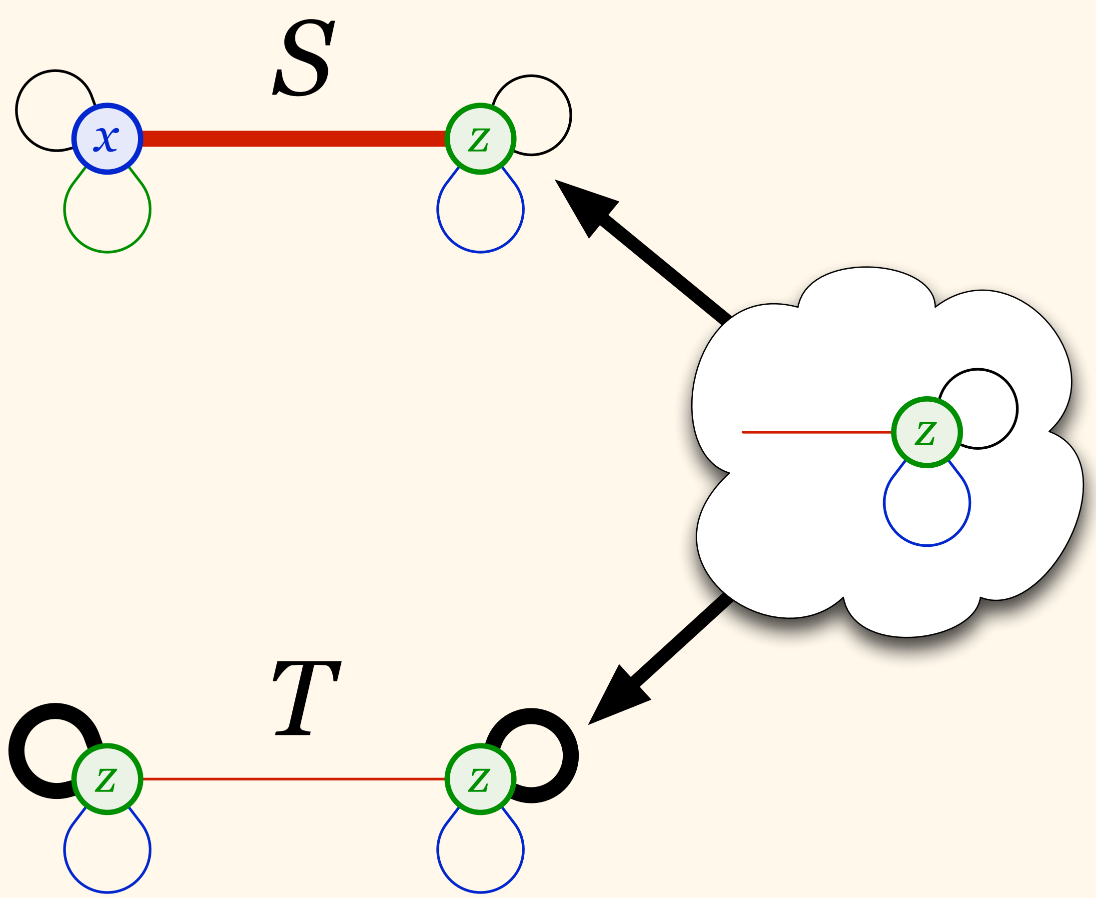
degree 1 templates,
same radius-0 view,
different output

Base Case



degree 1 templates,
same radius-0 view,
different output

Inductive Step

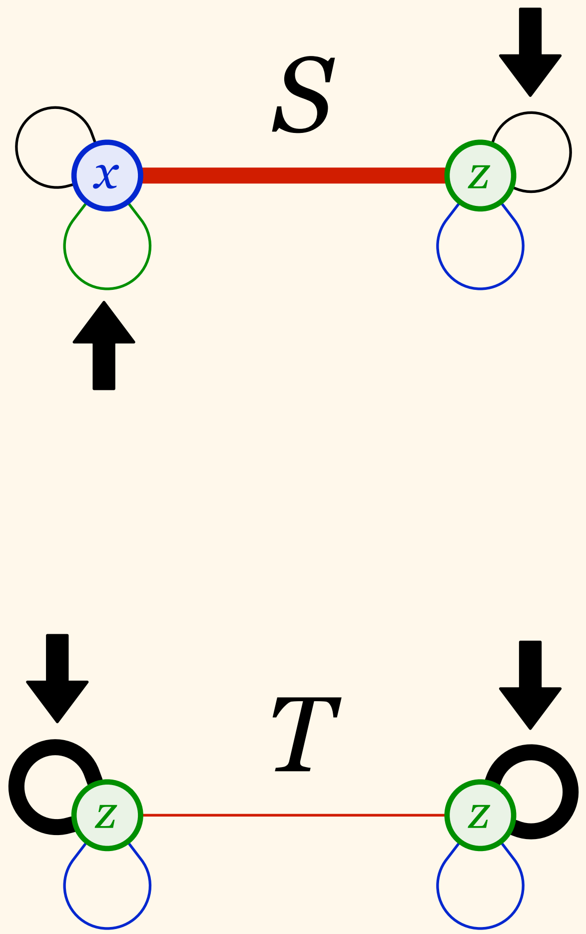


Given:
degree i templates,
same radius-(i–1) view,
different output

Construct:
degree i+1 templates,
same radius-i view,
different output

(here i = 1)

Inductive Step



Choose one loop per node

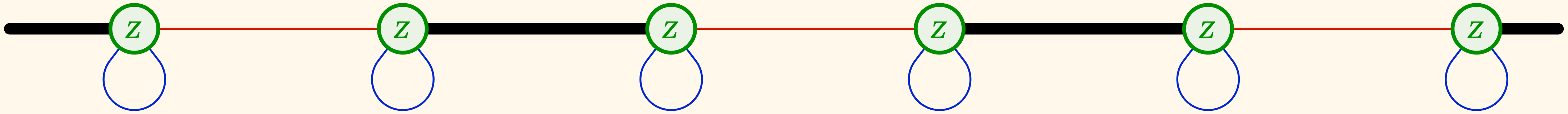
Prefer loops that are matched in T

Then unfold these loops…

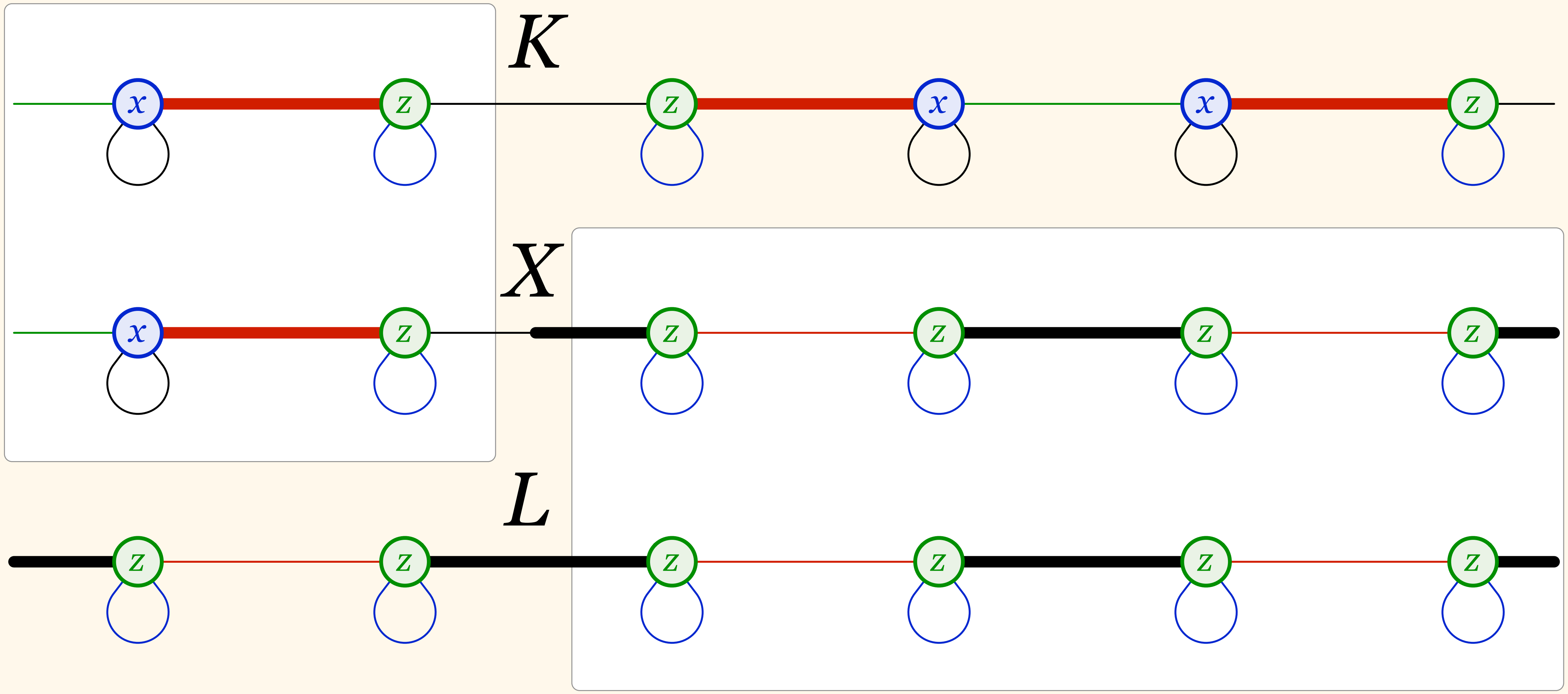
Inductive Step

K

L

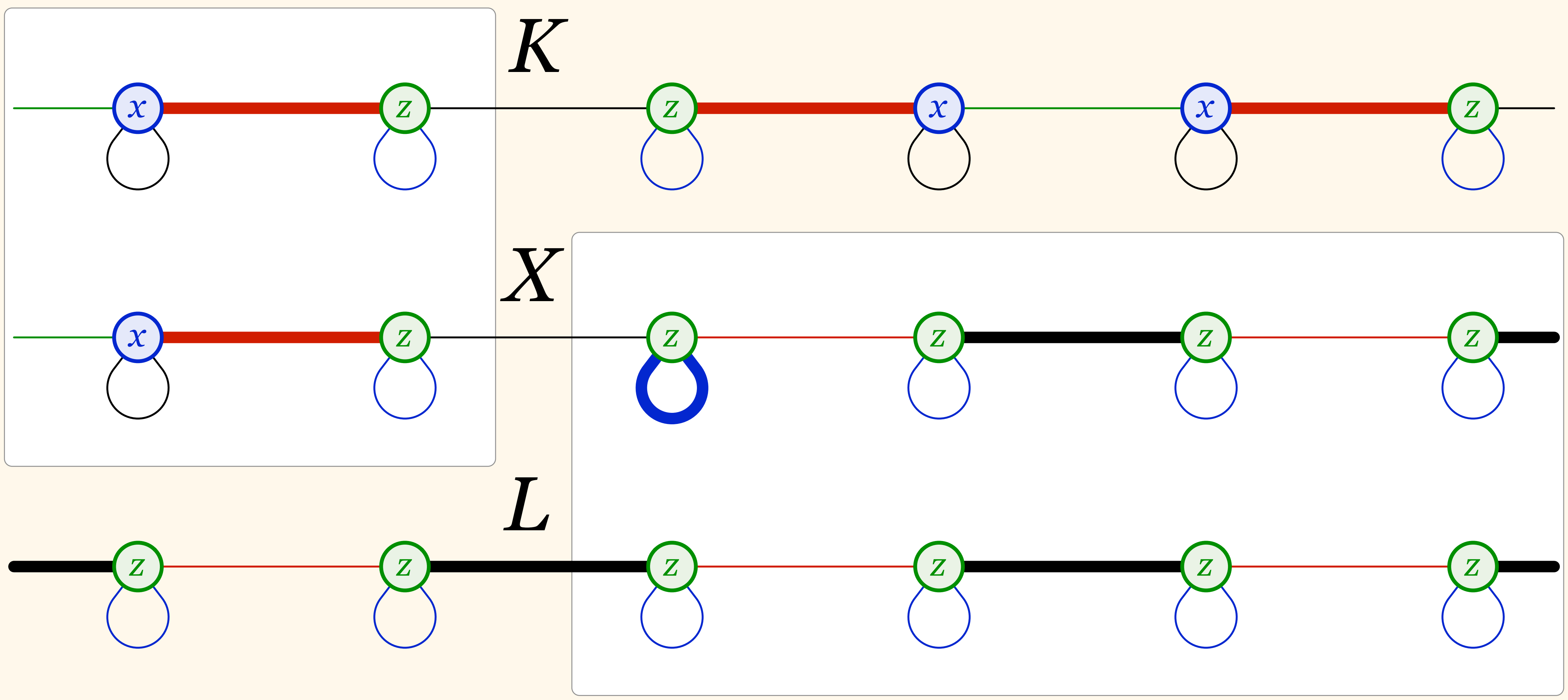


Inductive Step

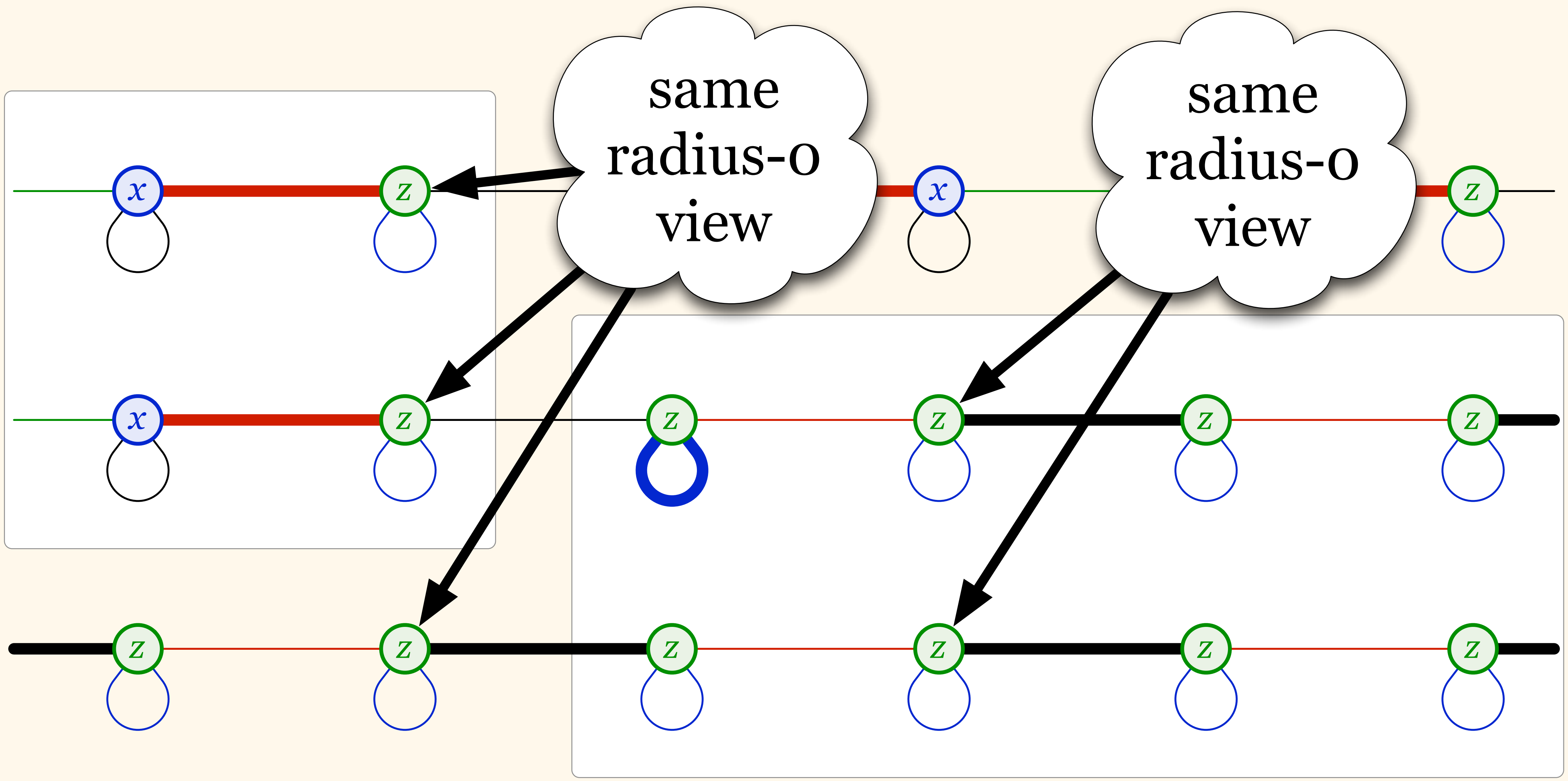


… again, something must change in the output!

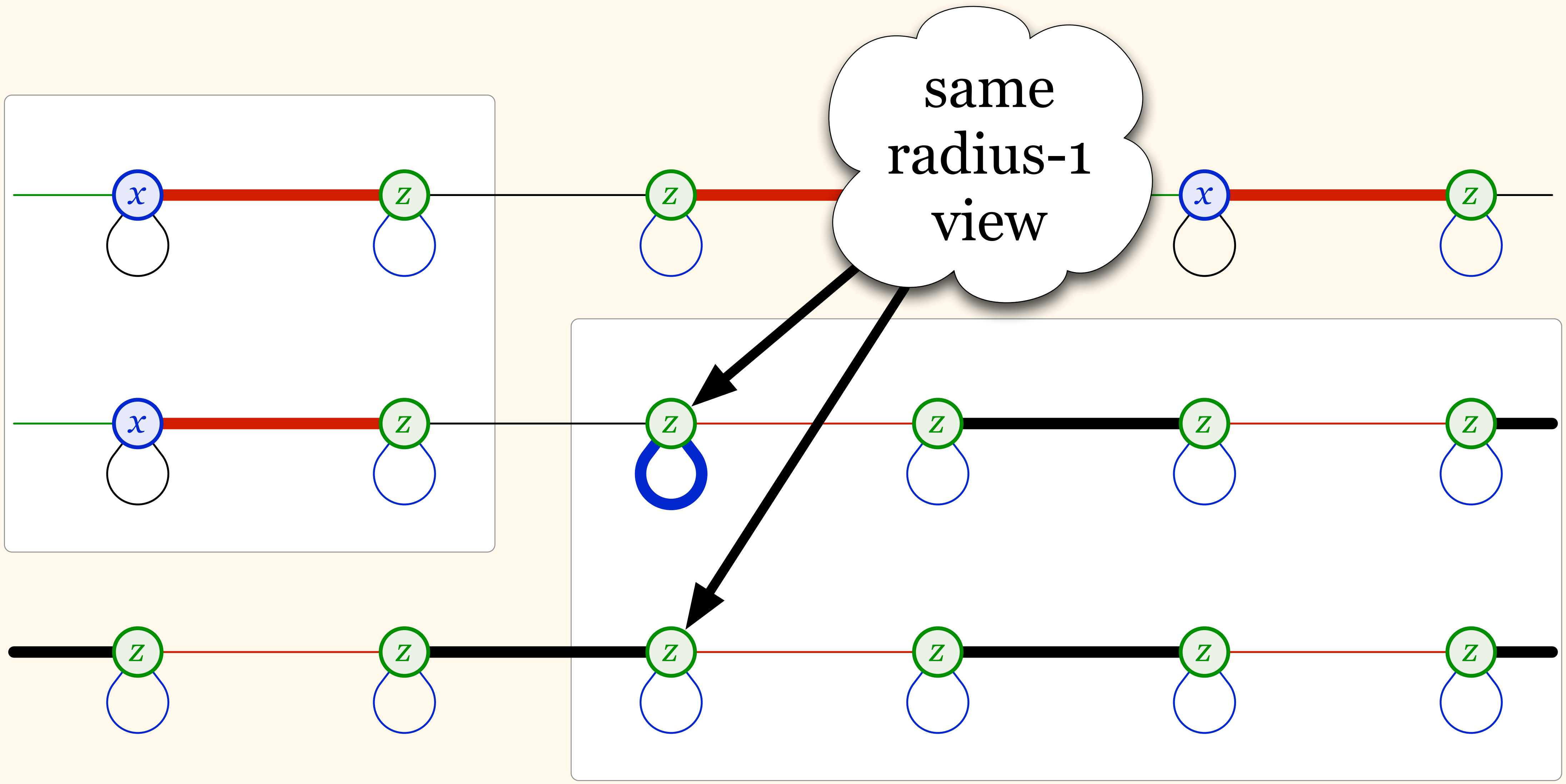
Inductive Step



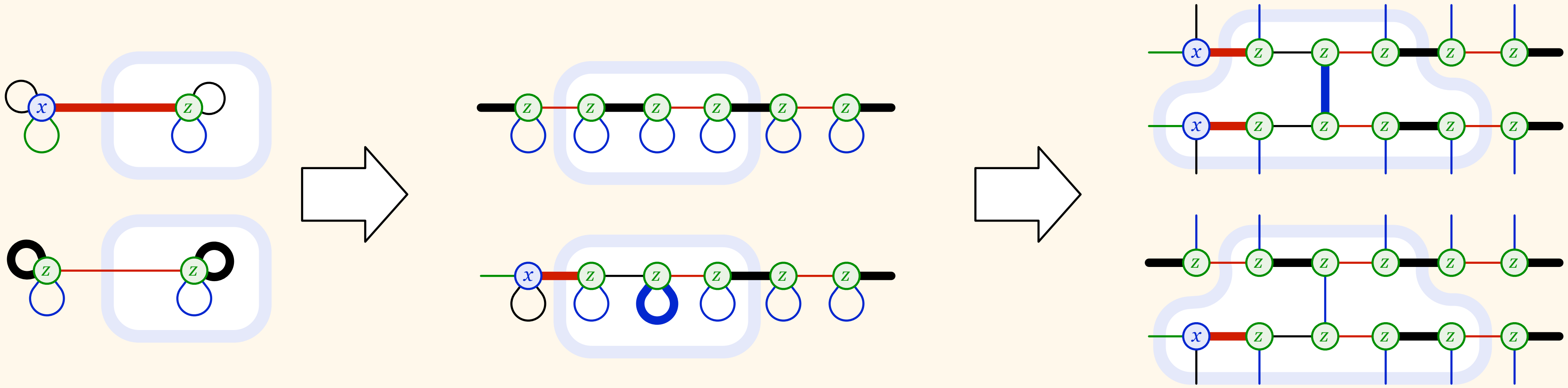
Inductive Step



Inductive Step



Conclusions



- By induction, we can construct:
  - two degree-d trees
  - same radius-(d–1) view
  - different output

Conclusions

- Maximal matching in ***k*-edge-coloured** graphs requires:
  - $k - 1$ communication rounds in general
  - $\Theta(\Delta + \log^* k)$ rounds in graphs of degree $\leq \Delta$
- What if we have **unique identifiers**?
  - *in progress*: tight bounds for *maximal edge packings*...
  - *still open*: tight bounds for *maximal matchings*?