My Favorite Proof on Boolean Functions:

Mykkeltveit's proof for Golomb's Conjecture

Tor Helleseth

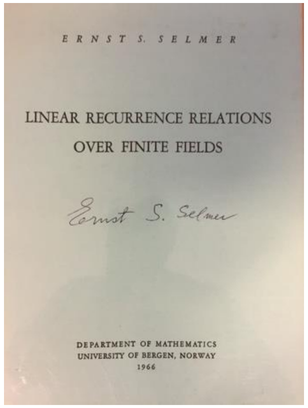
University of Bergen

NORWAY

My First Research Group

- 1969 – I started as master student at UiB
  - Ernst S. Selmer become my master supervisor
  - My first task reading his lecture notes on linear shift registers

    Linear Recurrence Relations over finite fields



  - In 1969/70 lectures on non-linear shift registers by visiting postdoc researchers Harold Fredricksen (former PhD student of Professor Solomon Golomb).
  - One PhD student (Johannes Mykkeltveit)
  - One PhD student (myself)

    (Some others like (Torleiv Kløve, Kjell Kjeldsen)

Some visitors now and them. Most notable Solomon Golomb

Solomon W. Golomb

- American researcher Solomon W. Golomb (1932-2016)
- Professor at University of Southern California (1962-2016)
- Fulbright scholar at University of Oslo (1955–1956)

”Selmer and I (in Oslo) had many interests in common, in prime number theory, sequence generation, combinatorics etc.”

- He was another pioneer with publications on shift registers in the 1960s.
  - Solomon Golomb, “Shift Register Sequences” (1967)

- Franklin Medal 2016
- National Medal of Honor 2014
- Hamming Medal 2000
- Shannon Award 1985

Outline

- In 1967 Solomon W. Golomb published a landmark book entitled: Shift Register Sequences
- S. Golomb studied linear and nonlinear shift registers
- Any Boolean function $f: F_2^n \to F_2$ of the form

  $$f(s_0, \ldots, s_{n-1}) = s_0 + \mathrm{g}(s_1, \ldots, s_{n-1})$$

  mapping

  $$(s_0, \ldots, s_{n-1}) \to (s_1, \ldots, s_{n-1}, s_0 + \mathrm{g}(s_1, \ldots, s_{n-1}))$$

  permutes the set B_n of all 2^n different binary n-tuples

  into distinct cycles.
- What is the maximum number of cycles that B_n can be decomposed into for all such Boolean functions f

Golomb's Conjecture

Among all $2^{2^{n-1}}$ nonsingular Boolean functions f the maximum number of cycles occurs for $f = s_0$ (i.e., for $g = 0$)

Golomb's Conjecture : The maximum number of cycles by any f occurs for $g = 0$ and equals

$$Z(n) = \frac{1}{n} \sum_{d|n} \varphi(d) 2^{\frac{n}{d}}$$

- Golomb's conjecture was based on computer search for $\mathrm{n} = 5$
- Improvements by Lempel for small cases like $n = 6,7,8$
- Further improvement Fredricsen and Mykkeltveit $n = 9,10,11,12$.
- Special cases solved in Fredricksen's thesis.

Finally solved by Mykkeltveit by a wondeful proof.

- One year of work.
- Published in Journal of Combinatorial Theory, Series B, 1972 (paper was 6-pages long and Mykkeltveit's 2nd paper as PhD)

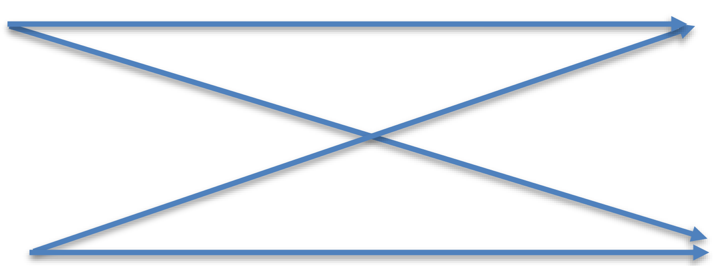
DeBruijn Graph B_n

- Nodes = Set of all 2^n binary n-tuples
- Directed edge iff

$$(s_0, s_1, \dots, s_{n-1}) \rightarrow (s_1, \dots, s_{n-1}, s_n)$$

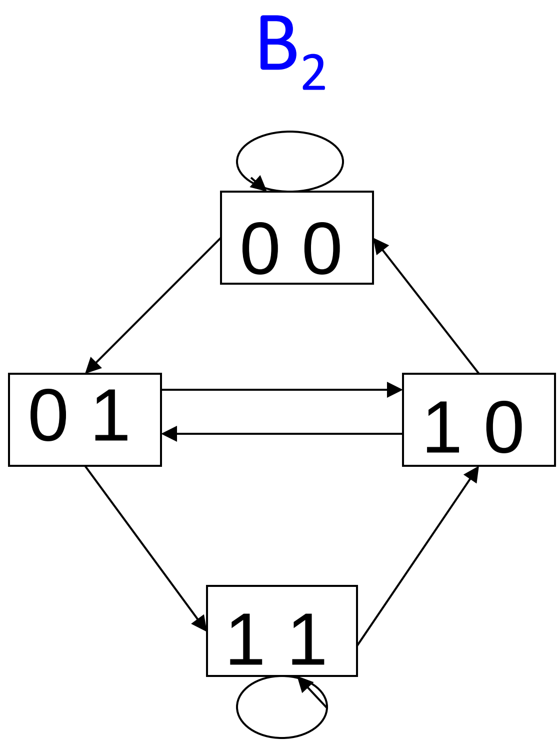
- Each node has two successors and two predecessors

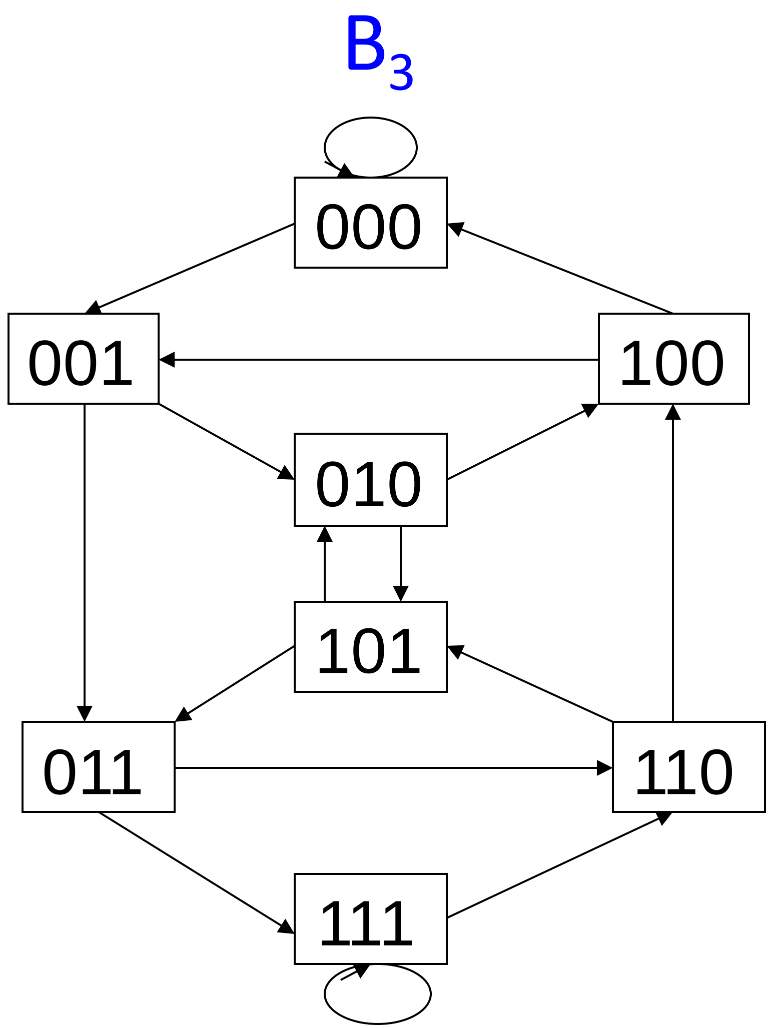
$(0, s_1, \dots, s_{n-1})$ $\qquad$ $(s_1, \dots, s_{n-1}, 0)$



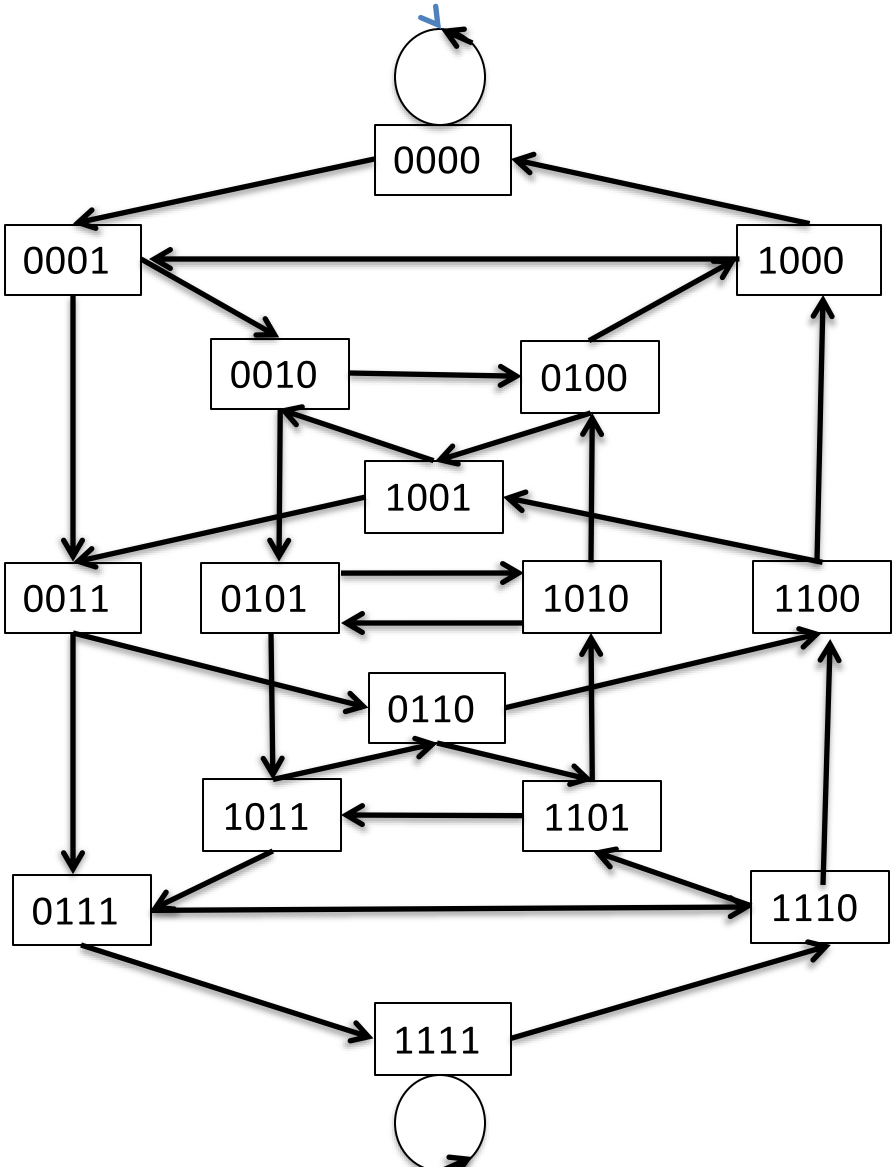
$(1, s_1, \dots, s_{n-1})$ $\qquad$ $(s_1, \dots, s_{n-1}, 1)$

DeBruijn Graphs (B_2 and B_3)





DeBruijn graph B_4



Pure Cycling Register (PCR_n)

- Let $f(s_0,s_1,\ldots,s_{n-1}) = s_0$ i.e., $g = 0$ (since $f=s_0 + g(s_1,\ldots,s_n)$)
  - Weight of truth table of g is 0
  - Cycle structure (PCR_n)

n=3 (0), (1), (001), (011)

n=4 (0), (1), (01), (0001), (0011), (0111)

- Number of cycles of B_n is well known to be

$$Z(n) = \frac{1}{n}\sum_{d|n} \varphi(d) 2^{\frac{n}{d}}$$

Pure Cycling Register (PCR_3) : ($f = s_0$)

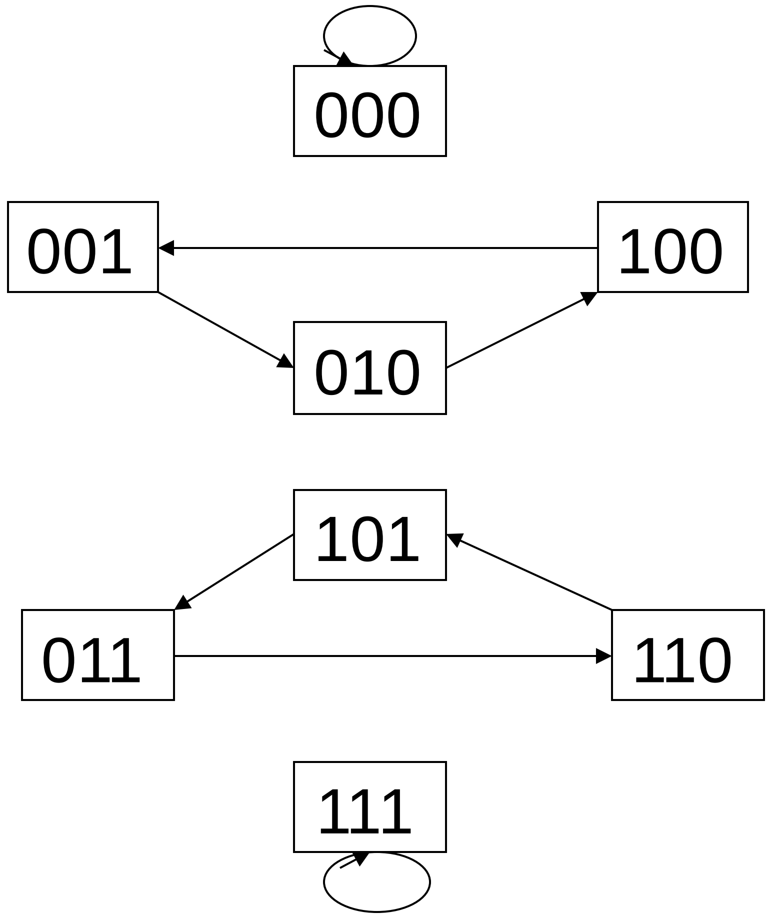
- Decomposition of B_3 for Boolean function $f=s_0$

(0)

(001)

(101)

(1)

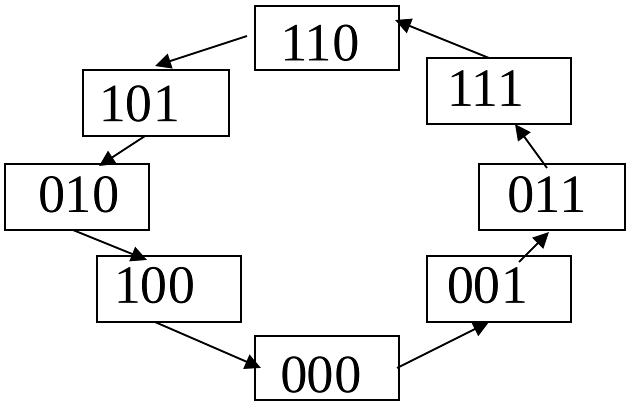


$f = s_0$

Number of cycles
Z(3) = 4

Example – de Bruijn Sequence

- Let $f(s_0,s_1,s_2) = 1+s_0+s_1+s_1s_2$



- This gives a maximal sequence of length 2^n

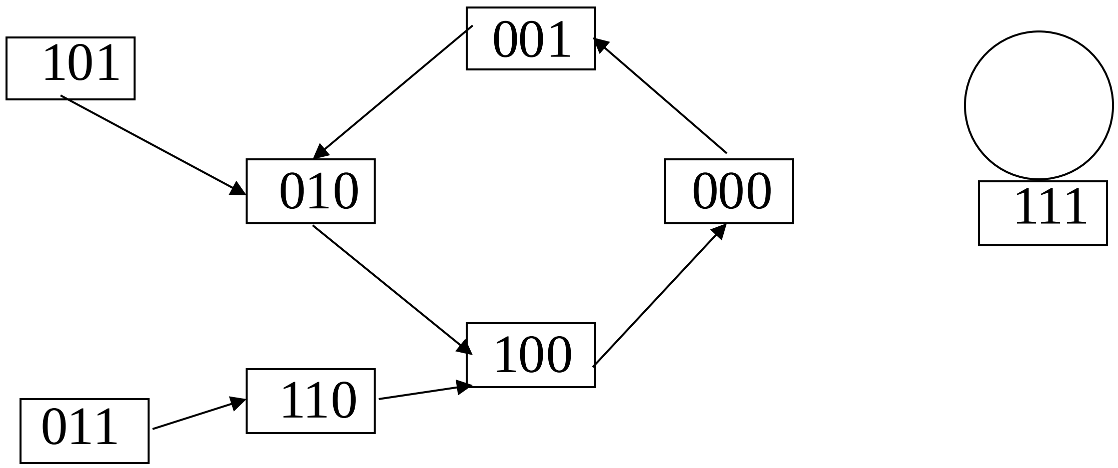
  … 11010001 …

  and is called a de Bruijn sequence

- Number of de Bruijn sequences of period 2^n are $2^{2^{n-1}-n}$

Example – Singular f

- Let $f(s_0,s_1,s_2) = 1+s_0+s_1+s_2+s_0s_1+s_0s_2+s_1s_2$

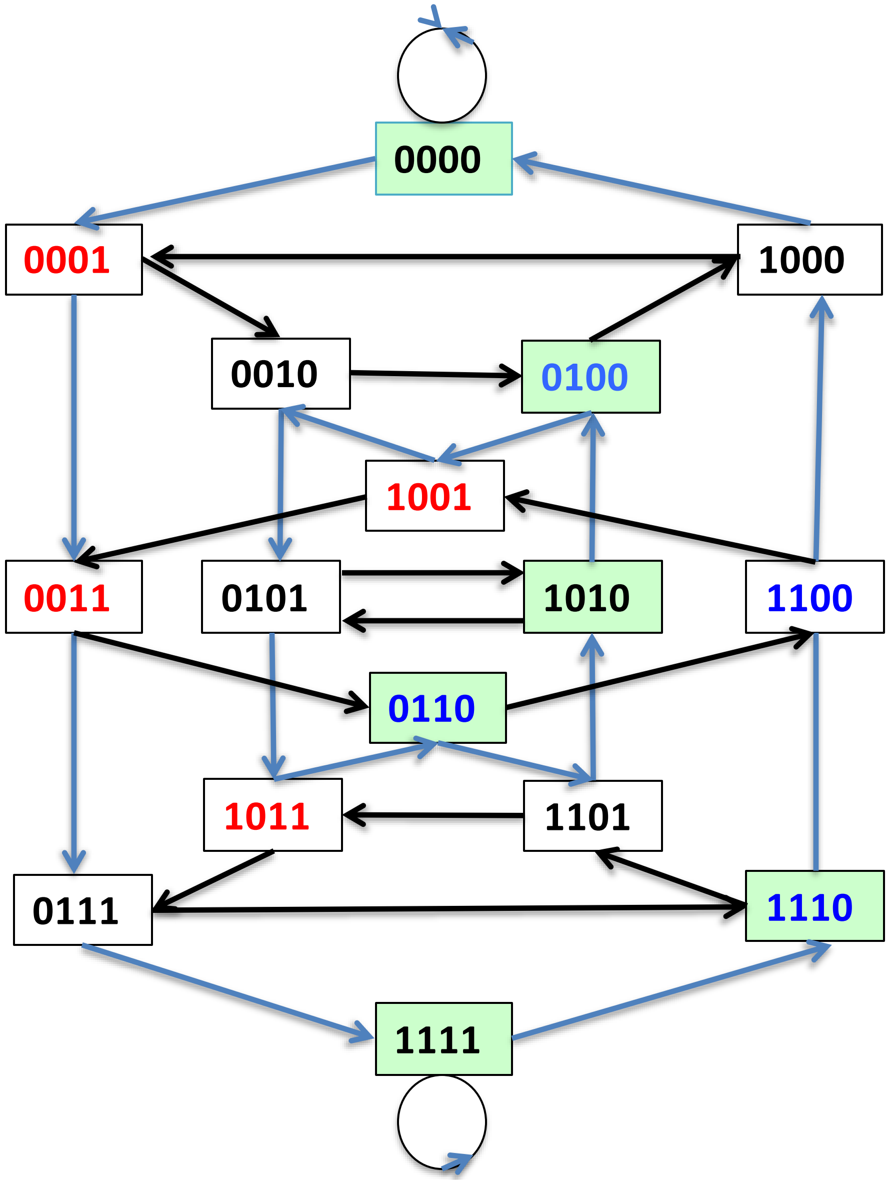


- Contains "branch point" and such an f is called singular
- f is nonsingular if and only if $f = s_0 + g(s_1,\ldots,s_{n-1})$
- Then $(s_0, s_1, \ldots, s_{n-1}) \rightarrow (s_1, s_2, \ldots, s_{n-1}, f(s_0,s_1,\ldots,s_{n-1}))$ is a permutation of B_n

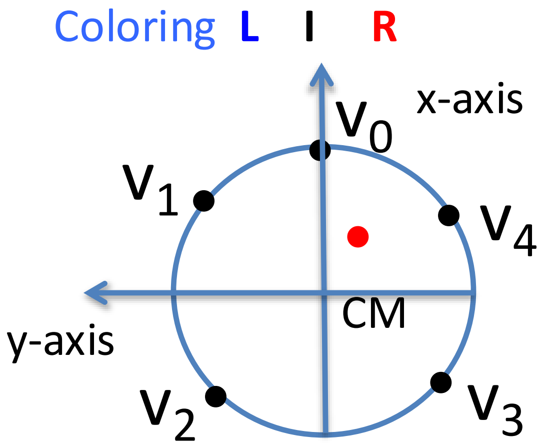
Mykkeltveit's Proof – Overview

1. Color all the nodes in B_n
2. Select one node on each of the Z(n) PCR_n cycle
3. Show that each cycle in B_n contains at least one selected node

Coloring deBruijn graph B_4



- Any cycle in B_4 contains at least one of the Z(4)=6 selected green colored nodes
- Coloring due to Mykkeltveit
- How to select these nodes with green color ?



How to for example color the node ($v_o\, v_1\, v_2\, v_3\, v_4$) ?

Compute center of mass CM for an n-tuple located around the unit circle

0 = 0 kg and **1** = 1 kg

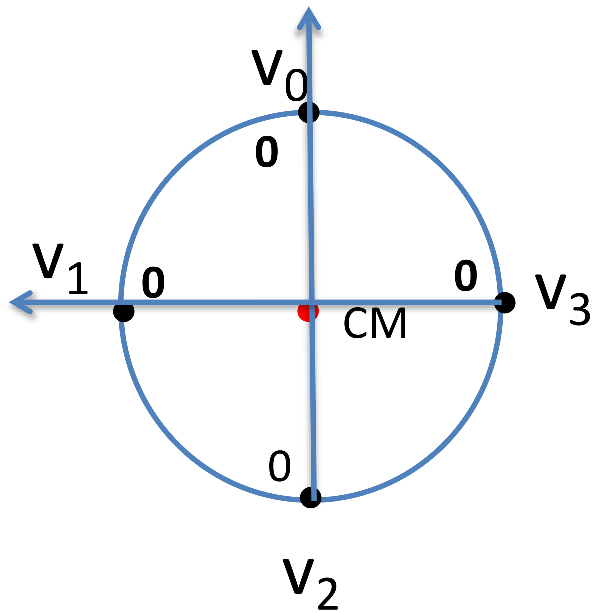
1. Color all the nodes in B_n

All nodes are colored **L, I,** or **R** according to whether the center of mass CM is **L**eft , **I**n or **R**ight of x-axis

Coloring B_4

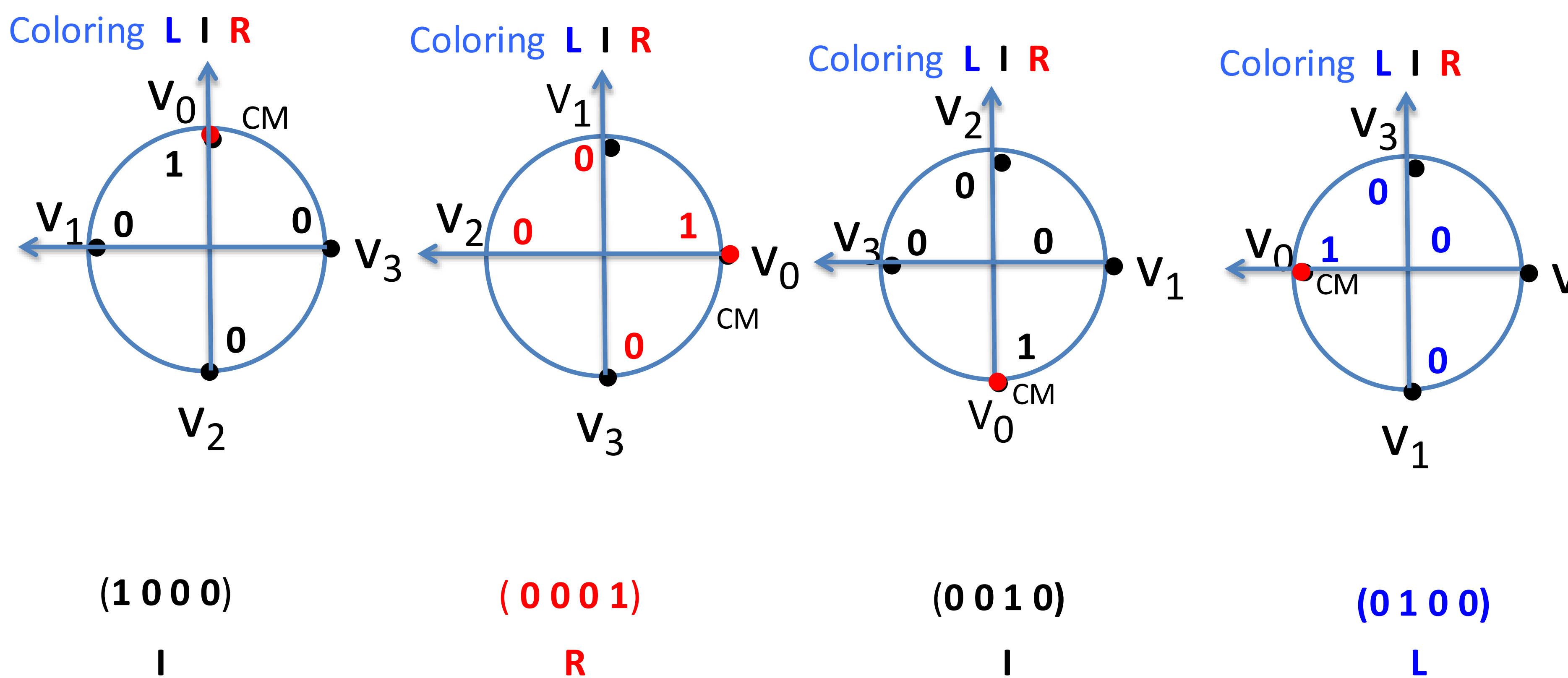
Coloring **L I R**

How to color node $(v_o\, v_1\, v_2\, v_3) = (\mathbf{0\ 0\ 0\ 0})$?

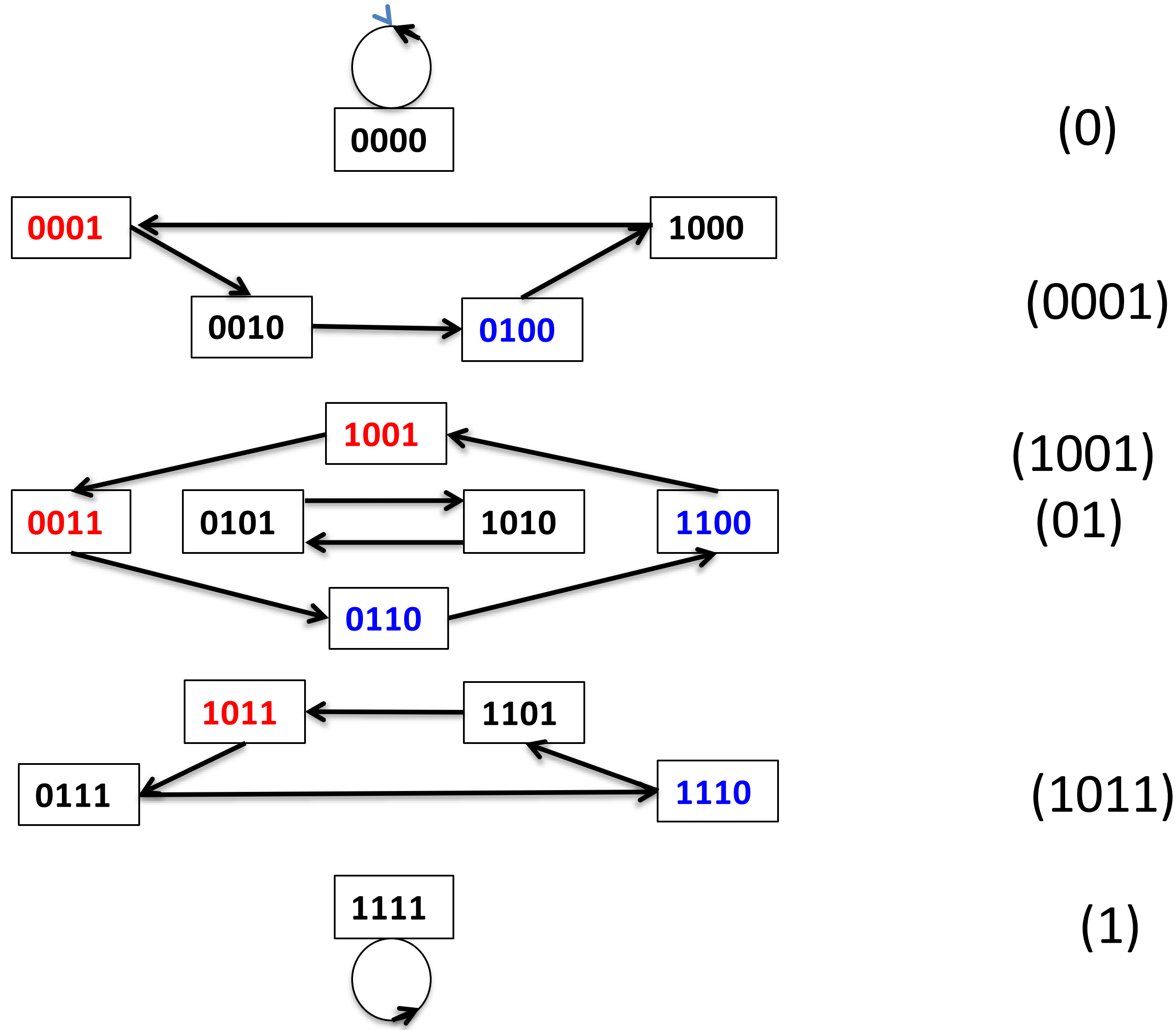


Coloring B_4

How to color nodes PCR_n cycles $(v_o\, v_1\, v_2\, v_3) = (\mathbf{1\;0\;0\;0})$?



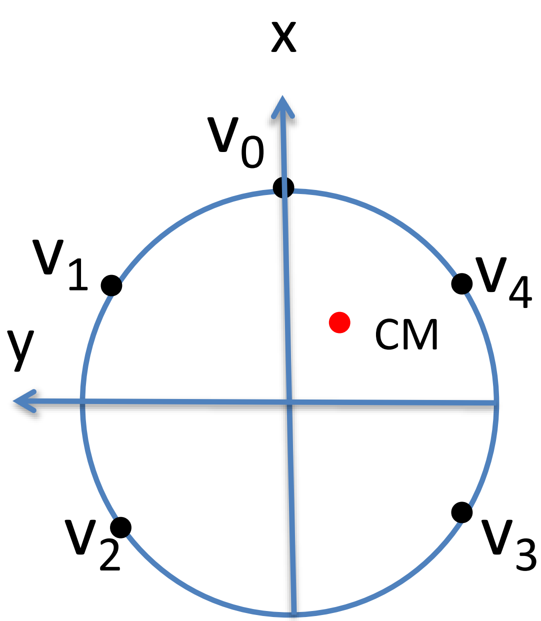
Pure Cycling Register (PCR_4) : $f = s_0$



Coloring of PCR_n cycles

Note that there are essentially only two possible ways of coloring all of the $Z(n)$ PCR_n cycles

CM of an n-tuple



Let $V_0=(v_0,v_1,v_2,v_3,v_4)$, (n=5)

Place v_t in coordinate position

$$(x,y)=\left(\cos\frac{2\pi t}{n},\sin\frac{2\pi t}{n}\right)$$

Compute CM=Center of mass

Moment y = $m_{V_0}=\sum_{t=0}^{n-1} v_t \sin\frac{2\pi t}{n}$

Color a vector $(v_0, v_1, \ldots, v_{n-1})$

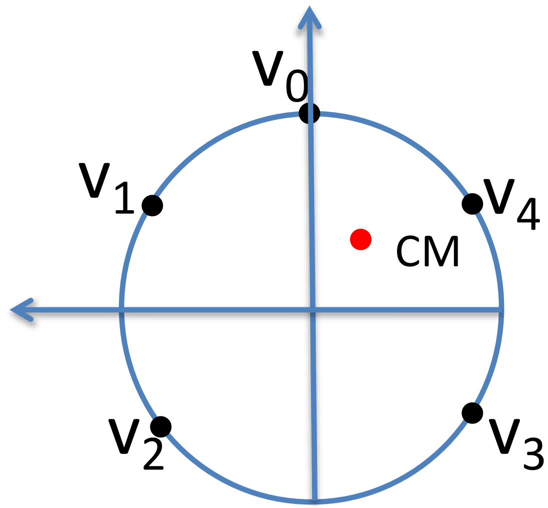
L = If CM on the **left** of the x-axis (y > 0)

I = If CM **on** the x-axis (y = 0)

R = If CM on the **right** of the x-axis (y < 0)

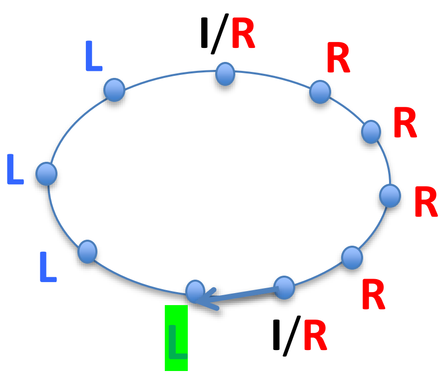
Coloring the PCR_n Cycles

Coloring **L** **I** **R**



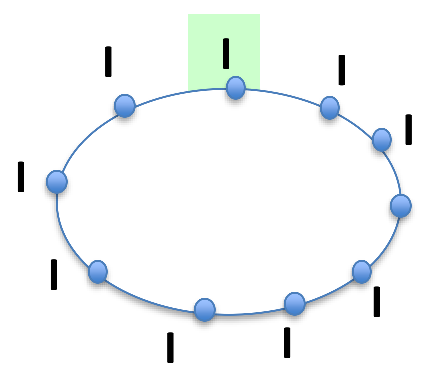
Type 1: (CM not in center of PCR cycle)

- **Select** unique node **L** with predecessor not **L**)



Type 2: (CM in the center of PCR cycle)

- Select **any** node colored **I**



2. Mark one node on each of the Z(n) PCR cycles

1. Color all the nodes in B_n
2. Select one node on each of the Z(n) PCR cycles
3. Show that each cycle in B_n contains at least one selected node

Mark one node on each PCR_n cycle

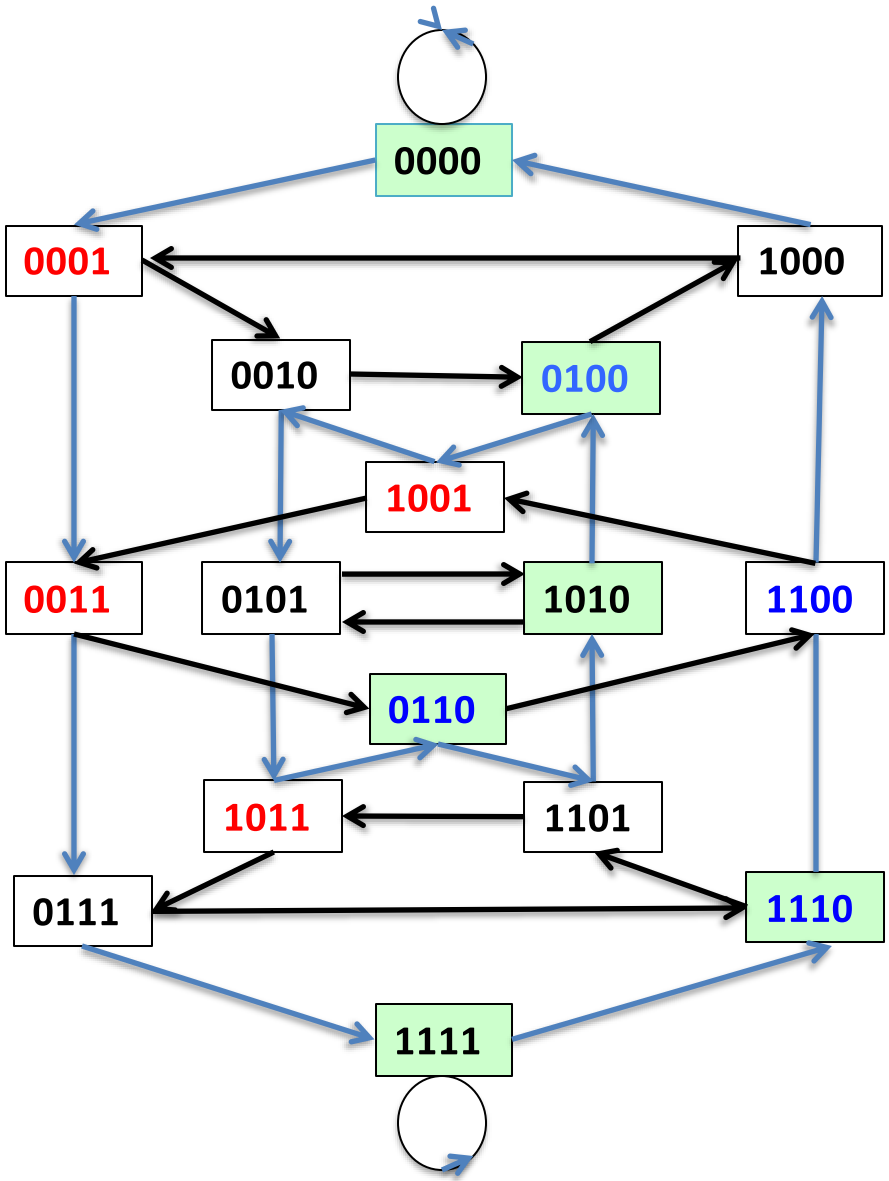
Case 1:

If all nodes on a PCR_n cycle have color **I** (i.e. CM in center) then select any node arbitrarily from cycle.

Case 2:

If a PCR_n has CM not in the center (i.e., has nodes of colors both **L** and **R**), then select the **UNIQUE** node **L** on the PCR_n cycle with a predecessor **not** colored **L.**

Coloring deBruijn graph B_4



- Any cycle in B_4 contains at least one of the Z(4)=6 green colored nodes
- Coloring due to Mykkeltveit
- How to select green color?

Properties of the coloring of the deBruijn graph B_n

This is important to prove that the coloring method works

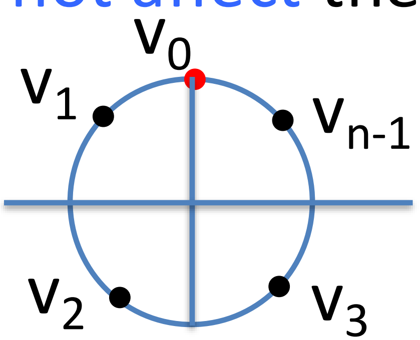
(surprising and very trivial properties)

The two predecessors of a node have the same color

Lemma

$(v_0, v_1, \ldots, v_{n-1})$ and $(v_0+1, v_1, \ldots, v_{n-1})$ have the same color.

Proof. The two n-vectors only differ in the red point on the x-axis that do not affect the y-coordinate of CM.



Corollary

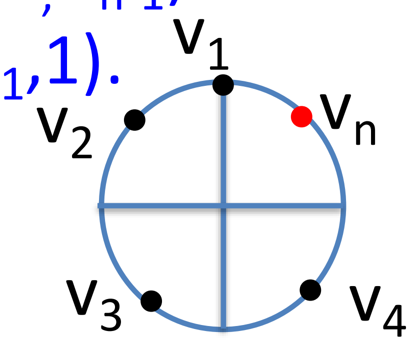
The two predecessors of any node $(v_1, v_2, \ldots, v_n)$ in the deBruijn graph have the same color.

The two successors of any node cannot both have color I

Lemma

The two successors of a node $(v_0, v_1, \ldots, v_{n-1})$ cannot have the same color **I**.

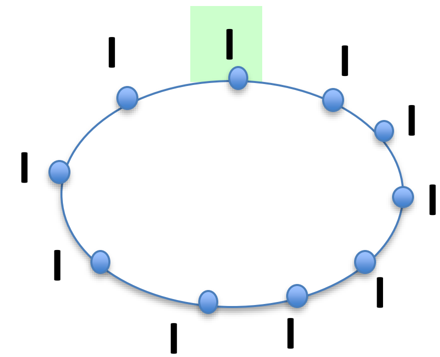
Proof. The two successors of $(v_0, v_1, \ldots, v_{n-1})$ are the two nodes $(v_1, v_2, \ldots, v_{n-1}, 0)$ and $(v_1, v_2, \ldots, v_{n-1}, 1)$.



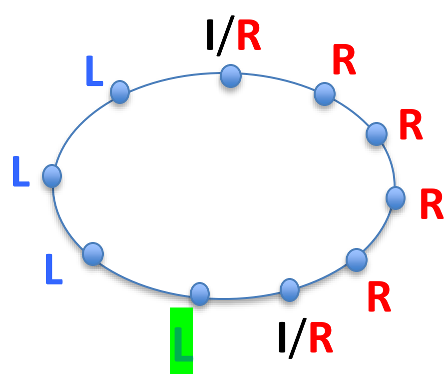
Since they only differ in the last coordinate (red point), they cannot both have CM on the x-axis and thus cannot have same color I.

There are two types of cycles on PCR_n

Type 1: All nodes on the PCR_n are **I**-nodes (i.e., CM is in the center)



Type 2: All nodes of the PCR_n cycle consist of one block of **L**-nodes and one block of R-nodes separated by at most one **I**-node



General cycles on B_n

Colors on a general cycle

Lemma 1

Let $(s_0, s_1, \ldots, s_{e-1})$ be a cycle of length e on B_n. The nodes (n-tuples) of the cycles are $\mathbf{S}_t = (s_t, s_{t+1}, \ldots, s_{t+n-1})$, t=0,1,…,e-1.

Then either

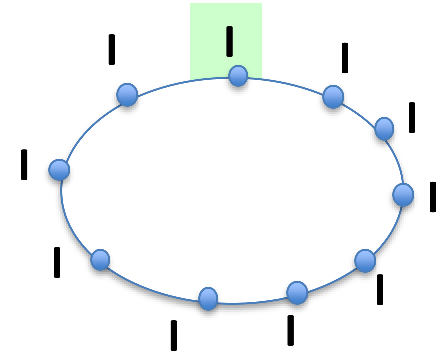
- All nodes on the cycle have the color **I**
- Cycle contains at least one **R** **and** one **L**

Proof. This follows since the sum of the y-coordinates on the nodes on a cycle is

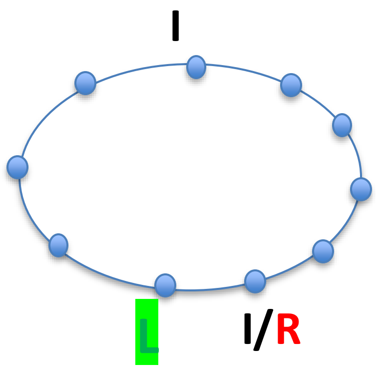
$$\sum_{t=0}^{e-1} m_{S_t} = \sum_{t=0}^{e-1} \sum_{t'=0}^{n-1} s_{t+t'} \sin\frac{2\pi i t'}{n} = \sum_{t=0}^{e-1} s_t \sum_{t'=0}^{n-1} \sin\frac{2\pi i t'}{n} = 0$$

There are two types of (general) cycles in B_n

Type G1: All nodes on the B_n are I-nodes (i.e., CM is in the center)

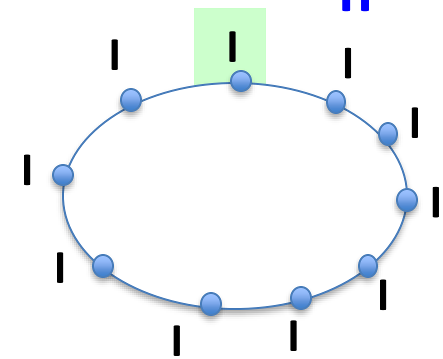


Type G2: The B_n cycle consist of at least one **L** and one **R** node



I

Cycles with only **I**-nodes in B_n



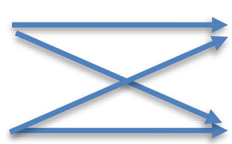
Lemma: A cycle in B_n with only **I**-nodes is a PCR_n cycle with CM in center

Proof: Any node in the cycle has an **I**-node as predecessor.

Therefore CM is in center since node is on a PCR_n cycle with at least two consecutive **I**-nodes.

Suppose cycle has **I**-nodes from two different PCR cycles C_1 and C_2.
Then an **I**-node on C_1 has successor on C_2

$(v_0 \quad v_1 \cdots v_{n-1})$ **I** $\qquad$ **I** $(v_1 \cdots \; v_{n-1}v_0)$ on C_1



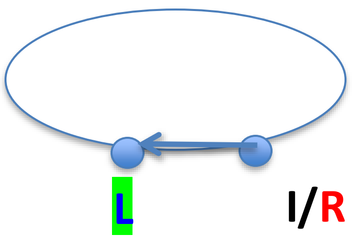
$(v_0{+}1\; v_1 \cdots v_{n-1})$ **I** $\qquad$ **I** $(v_1 \cdots v_{n-1}\, v_0{+}1)$ on C_2

Since **I**-node $(v_0v_1 \cdots v_{n-1})$ on C_1 has two possible **I**-node successors $(v_1v_2 \cdots v_{n-1}v_0)$ on C_1 and $(v_1 \cdots v_{n-1}v_0{+}1)$ on C_2 this is impossible.

Cycles with and L's (and R's)

Lemma

In a cycle with **L'**s and **R**'s let **V** be a node with color **L** with predecessor not in **L**. Then (in PCR_n) **V** is the first node on a block of **L**'s on the PCR_n.



Proof. Predecessor of **V** has color ≠ **L** on the cycle. Therefore, both predecessors of **V** (also the one on PCR_n) have color not being **L**. Hence, **V** is first node in a block of **L**'s on PCR_n.

Observation: Each cycle in B_n with the property above contains the first **L** node in some PCR_n cycle in a block of **L**'s

Final Remarks – Coloring Summary

- Shifting a node cyclically shifts **CM**
- The two predecessors for a node in B_n have the same color (since they only differ in 0-th coordinate on the x-axis).
- The two successors of a node can not both have color I (since they only differ in position n-1).
- A cycle in PCR_n has either:
  - All nodes colored I
  - One R block and one L block separated by at most one I.
- Any cycle $S = (s_0, s_1, \dots, s_{e-1})$ in B_n has (average moment = 0), i.e. has either:
  - All nodes colored I
  - At least one R and one L node