Herman's Algorithm: A Brief Survey and Some Recent Results

LSV, ENS de Cachan, February 2012

Stefan Kiefer[1] Andrzej S. Murawski[2] Joël Ouaknine[1]
James Worrell[1] Lijun Zhang[3]

[1]Department of Computer Science, University of Oxford, UK

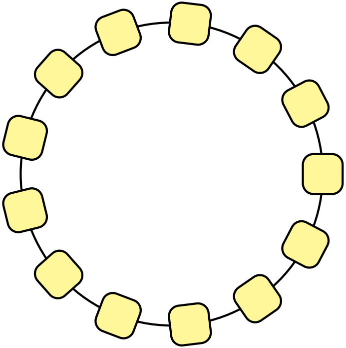
[2]Department of Computer Science, University of Leicester, UK

[3]DTU Informatics, Technical University of Denmark, Denmark

Particle Collision: CERN

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

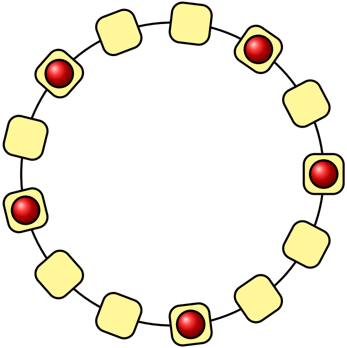


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

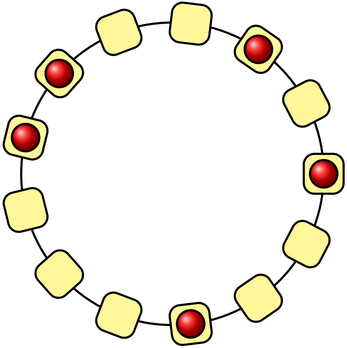


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and move it clockwise.
- When a processor has two tokens: both disappear.

Once there is only one token: this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

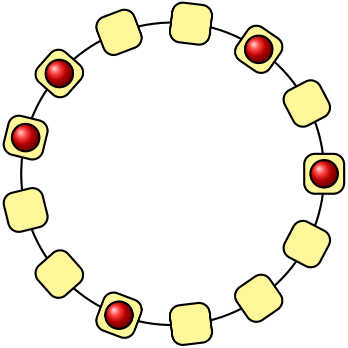


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and move it clockwise.
- When a processor has two tokens: both disappear.

Once there is only one token: this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

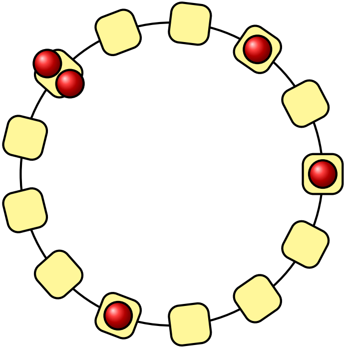


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

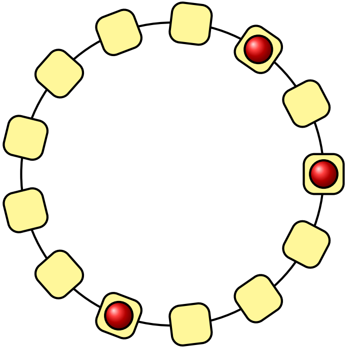


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

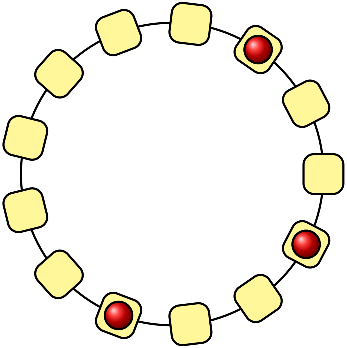


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.

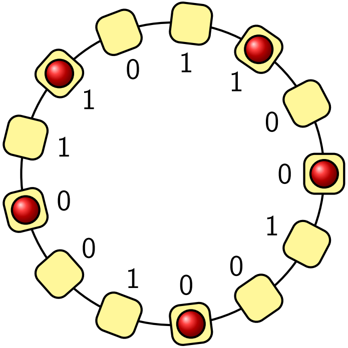


- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



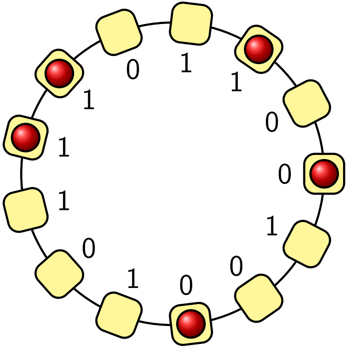
- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token = same bit as counter-clockwise neighbour

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



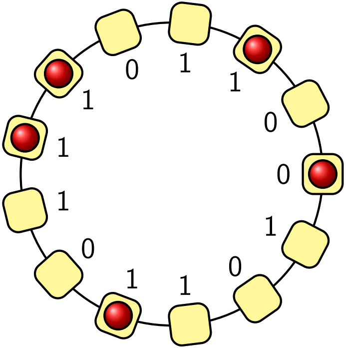
- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token = same bit as counter-clockwise neighbour

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



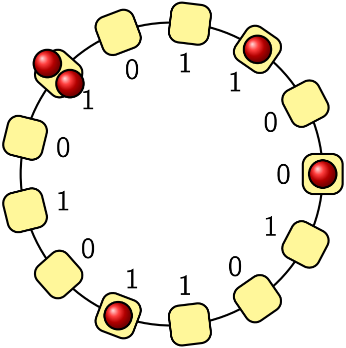
- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token = same bit as counter-clockwise neighbour

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



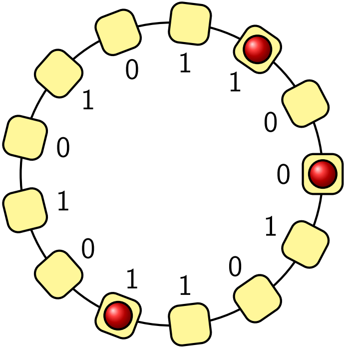
- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token = same bit as counter-clockwise neighbour

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



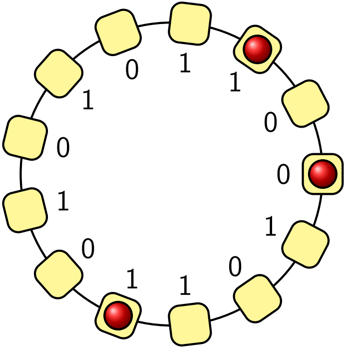
- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token = same bit as counter-clockwise neighbour

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



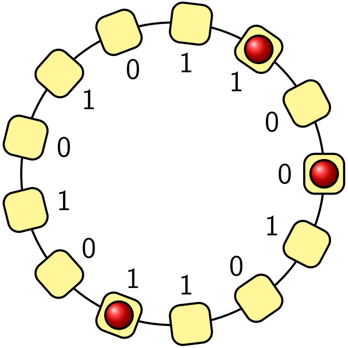
- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token = same bit as counter-clockwise neighbour
$\Longrightarrow$ Number of tokens is odd. At any time. Guaranteed!

Herman's Algorithm

Herman's algorithm is a randomized leader election protocol.



- Ring of N processors (N is odd)
- Some processors have a token.
- Pick a token randomly and
  move it clockwise.
- When a processor has two tokens:
  both disappear.

Once there is only one token:
this processor is the leader.

Implementation: token $=$ same bit as counter-clockwise neighbour
$\Longrightarrow$ Number of tokens is odd. At any time. Guaranteed!

Main Question:

How long does it take – on average – until a leader is elected?

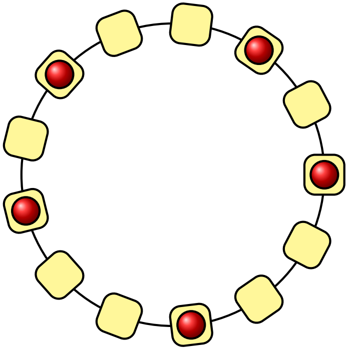
Herman's Algorithm

Two versions of Herman's algorithm:

- Asynchronous: Each processor with a token
  passes the token to its clockwise neighbour with rate λ
  $\Longrightarrow$ continuous-time Markov chain
  $\Longrightarrow$ (as seen on the previous slides)
- Synchronous: Each processor with a token
  with probability $1/2$: passes it (i.e., flips its bit)
  with probability $1/2$: keeps it (i.e., keeps its bit)
  $\Longrightarrow$ discrete-time Markov chain

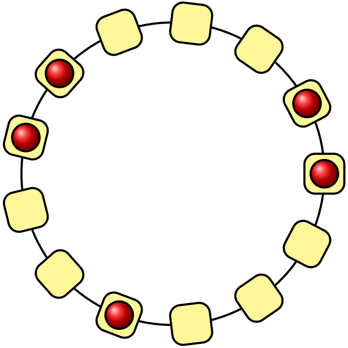
Herman's Algorithm

Illustration of the synchronous version:



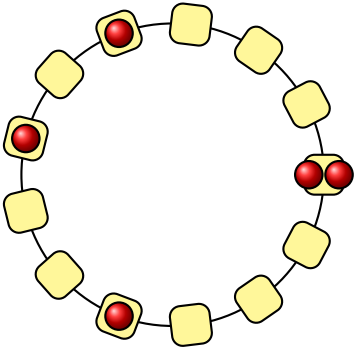
Herman's Algorithm

Illustration of the synchronous version:



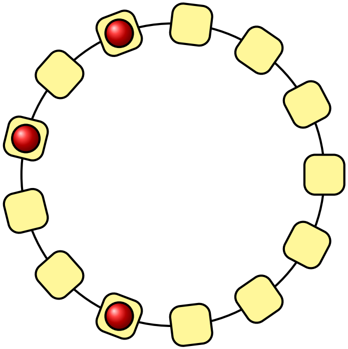
Herman's Algorithm

Illustration of the synchronous version:



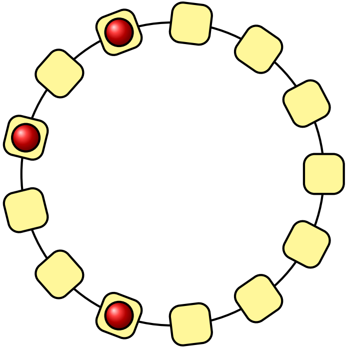
Herman's Algorithm

Illustration of the synchronous version:



Herman's Algorithm

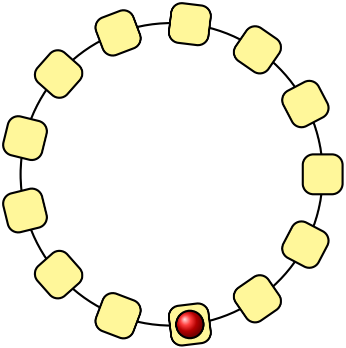
Illustration of the synchronous version:



The synchronous version is standard.
The probability of passing is usually chosen as $1/2$,
but is a parameter in our analysis.
We study both versions (asynchronous and synchronous).

Self-Stabilization

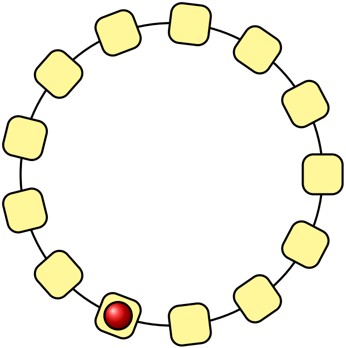
Self-stabilization: a concept of fault-tolerance:
Any configuration leads to a "legitimate" configuration
(in Herman: a single-token configuration).



Dijkstra (1974): In a token ring
there is no symmetric deterministic self-stabilizing algorithm.

Self-Stabilization

Self-stabilization: a concept of fault-tolerance:
Any configuration leads to a “legitimate” configuration
(in Herman: a single-token configuration).



Dijkstra (1974): In a token ring
there is no symmetric deterministic self-stabilizing algorithm.

Time to Stabilization

Let $\mathbf{T} :=$ time until a 1-token configuration is reached.

What is $\mathbb{E}\mathbf{T}$?

Markov Chain argument $\Longrightarrow$ $\mathbb{E}\mathbf{T}$ is finite for any initial conf.

Time to Stabilization

Let $\mathbf{T}$:= time until a 1-token configuration is reached.

What is $\mathbb{E}\mathbf{T}$?

Markov Chain argument $\Rightarrow$ $\mathbb{E}\mathbf{T}$ is finite for any initial conf.

- 1990 Herman: invented the protocol
  showed $\mathbb{E}\mathbf{T} \leq 0.5N^2 \log N$
  mentions improvement to $\mathbb{E}\mathbf{T} \leq c\,N^2$
- 2004 Fribourg et al.: $\mathbb{E}\mathbf{T} \leq 2N^2$
- 2005 McIver/Morgan: $\mathbb{E}\mathbf{T} \leq 2N^2$ (independently)
- 2005 Nakata: $\mathbb{E}\mathbf{T} \leq 0.93N^2$

Time to Stabilization

Let $\mathbf{T}$:= time until a 1-token configuration is reached.

What is $\mathbb{E}\mathbf{T}$?

Markov Chain argument $\Rightarrow$ $\mathbb{E}\mathbf{T}$ is finite for any initial conf.

- 1990 Herman: invented the protocol
  showed $\mathbb{E}\mathbf{T} \leq 0.5 N^2 \log N$
  mentions improvement to $\mathbb{E}\mathbf{T} \leq c\, N^2$
- 2004 Fribourg et al.: $\mathbb{E}\mathbf{T} \leq 2N^2$
- 2005 McIver/Morgan: $\mathbb{E}\mathbf{T} \leq 2N^2$ (independently)
- 2005 Nakata: $\mathbb{E}\mathbf{T} \leq 0.93 N^2$

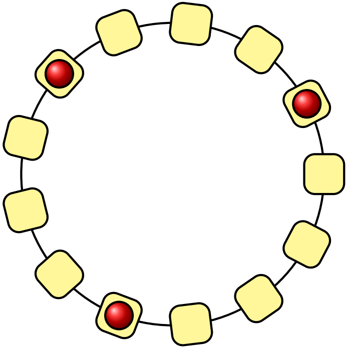
Theorem

Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous).

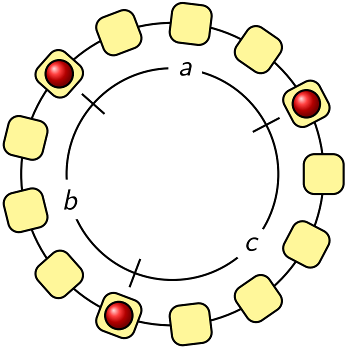
Then
$$\mathbb{E}\mathbf{T} \leq \left(\frac{\pi^2}{8} - \frac{29}{27}\right) \cdot \frac{N^2}{D}.$$

In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.64 N^2$.

The 3-Token Case



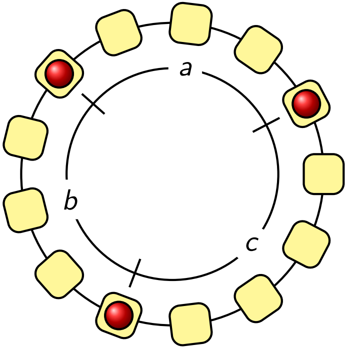
The 3-Token Case



Precise formula for 3 tokens [MM'05]:

$$\mathbb{E}\mathbf{T} = 4abc/N$$

The 3-Token Case

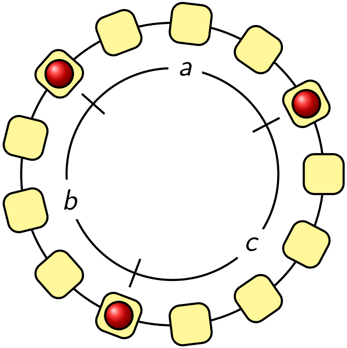


Precise formula for 3 tokens [MM'05]:

$$\mathbb{E}\mathbf{T} = 4abc/N$$

Maximized for equilateral configuration with $a = b = c = N/3$
$\Longrightarrow \mathbb{E}\mathbf{T} = \frac{4}{27}N^2 = 0.148N^2$

The 3-Token Case



Precise formula for 3 tokens [MM'05]:

$$\mathbb{E}\mathbf{T} = 4abc/N$$

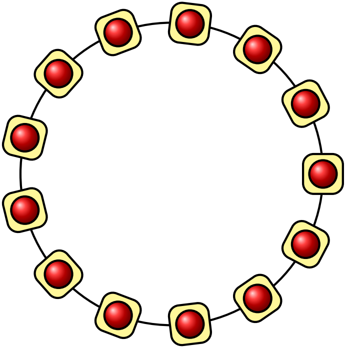
Maximized for equilateral configuration with $a = b = c = N/3$
$\Rightarrow \mathbb{E}\mathbf{T} = \frac{4}{27}N^2 = 0.148N^2$

Open conjecture [MM'05]:

3-token equilateral conf. is worse than any other conf.

If true, then $\mathbb{E}\mathbf{T} \leq 0.148N^2$ (cf. $\mathbb{E}\mathbf{T} \leq 0.64N^2$).

The Full Configuration



Theorem

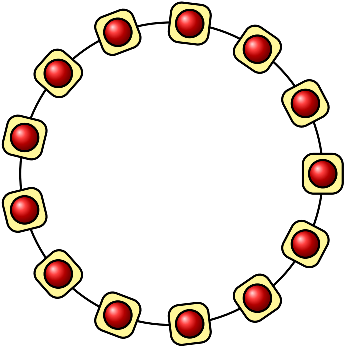
Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous). Starting in the full configuration, for almost all odd N:

$$\mathbb{E}\mathbf{T} \leq 0.0285 N^2 / D\,.$$

In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.114 N^2$.

(cf. for general conf.: $\mathbb{E}\mathbf{T} \leq 0.64 N^2$
for equilateral conf.: $\mathbb{E}\mathbf{T} = 0.148 N^2$)

The Full Configuration



Theorem

Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous). Starting in the full configuration, for almost all odd N:

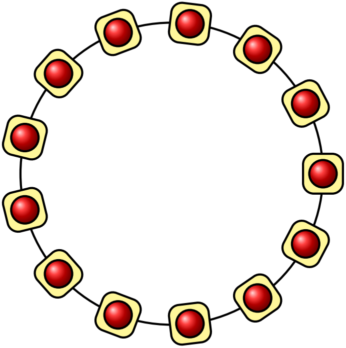
$$\mathbb{E}\mathbf{T} \leq 0.0285 N^2 / D\,.$$

In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.114 N^2$.

(cf. for general conf.: $\mathbb{E}\mathbf{T} \leq 0.64 N^2$
for equilateral conf.: $\mathbb{E}\mathbf{T} = 0.148 N^2$)

What about a random initial configuration?

The Full Configuration



Theorem

Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous). Starting in the full configuration, for almost all odd N:

$$\mathbb{E}\mathbf{T} \leq 0.0285 N^2 / D \, .$$

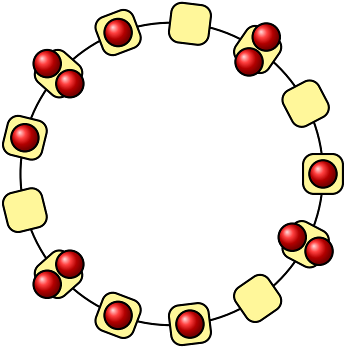
In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.114 N^2$.

(cf. for general conf.: $\mathbb{E}\mathbf{T} \leq 0.64 N^2$
for equilateral conf.: $\mathbb{E}\mathbf{T} = 0.148 N^2$)

What about a random initial configuration?

A random initial configuration is reached after the first step.

The Full Configuration



Theorem

Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous). Starting in the full configuration, for almost all odd N:

$$\mathbb{E}\mathbf{T} \leq 0.0285 N^2 / D\,.$$

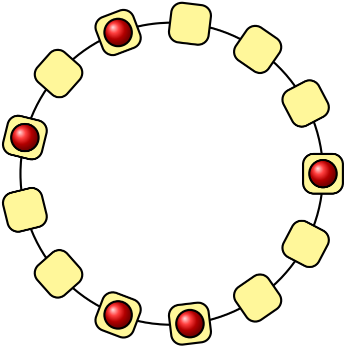
In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.114 N^2$.

(cf. for general conf.: $\mathbb{E}\mathbf{T} \leq 0.64 N^2$
for equilateral conf.: $\mathbb{E}\mathbf{T} = 0.148 N^2$)

What about a random initial configuration?

A random initial configuration is reached after the first step.

The Full Configuration



Theorem

Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous). Starting in the full configuration, for almost all odd N:

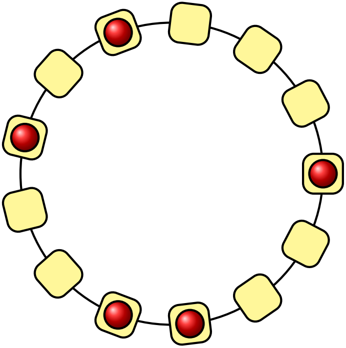
$$\mathbb{E}\mathbf{T} \leq 0.0285 N^2 / D \,.$$

In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.114 N^2$.

(cf. for general conf.: $\mathbb{E}\mathbf{T} \leq 0.64 N^2$
for equilateral conf.: $\mathbb{E}\mathbf{T} = 0.148 N^2$)

What about a random initial configuration?
A random initial configuration is reached after the first step.

The Full Configuration



Theorem

Let $D = r(1-r)$ (for synchronous) or $D = \lambda$ (for asynchronous). Starting in the full configuration, for almost all odd N:

$$\mathbb{E}\mathbf{T} \leq 0.0285 N^2/D\,.$$

In particular, for synchronous and $r = 1/2$, we have $\mathbb{E}\mathbf{T} \leq 0.114 N^2$.

(cf. for general conf.: $\mathbb{E}\mathbf{T} \leq 0.64 N^2$
for equilateral conf.: $\mathbb{E}\mathbf{T} = 0.148 N^2$)

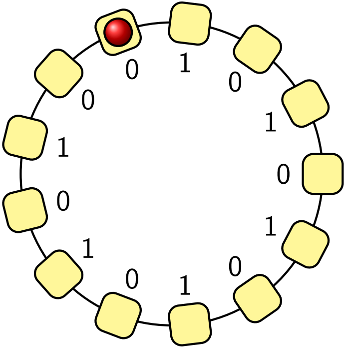
What about a random initial configuration?

A random initial configuration is reached after the first step.

$\Longrightarrow$ no need for a different analysis

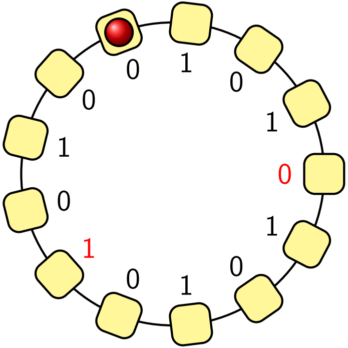
Restabilization

Imagine: starting from 1 token . . .



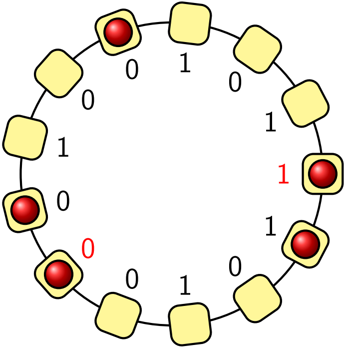
Restabilization

Imagine: starting from 1 token . . . 2 bit-errors occur



Restabilization

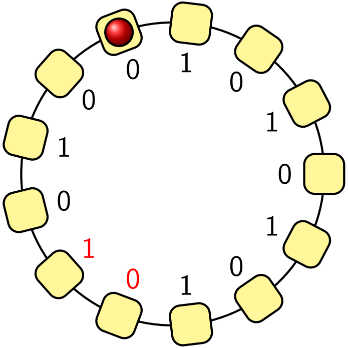
Imagine: starting from 1 token . . . 2 bit-errors occur



⇒ “flip-2 configuration”

Restabilization

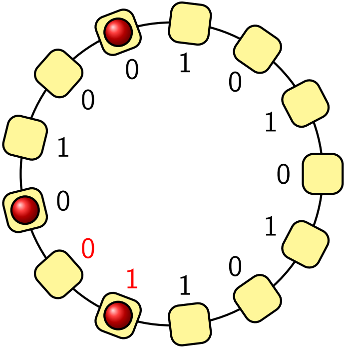
Imagine: starting from 1 token . . . 2 bit-errors occur



⇒ “flip-2 configuration”

Restabilization

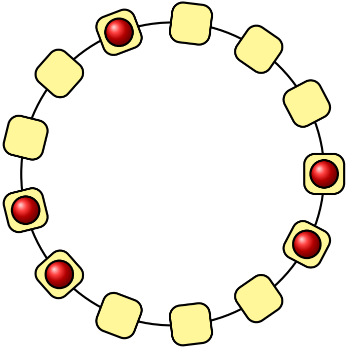
Imagine: starting from 1 token . . . 2 bit-errors occur



$\Rightarrow$ “flip-2 configuration”

Restabilization

Imagine: starting from 1 token . . . 2 bit-errors occur



$\Rightarrow$ "flip-2 configuration"

Theorem

Consider the synchronous protocol with $r \in (0.12, 0.88)$. *Fix* m.
Then for any flip-m configuration: $\mathbb{E}\mathbf{T} = O(N)$

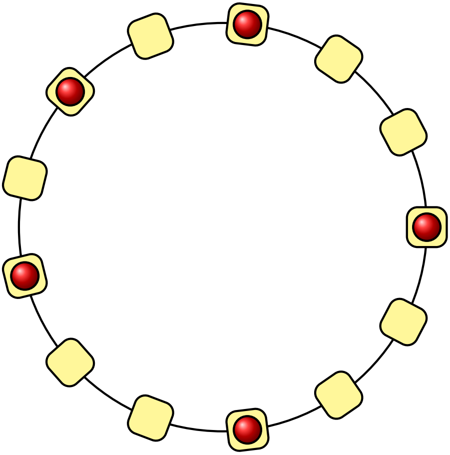
(cf. $O(N^2)$ for all other bounds)

Interacting Particle Systems

Interacting particles under random motion occur in:

- statistical mechanics
- neural networks
- spread of infections
- tumor growth
- . . . (other physical and medical sciences)
- physical chemistry [Balding'88]

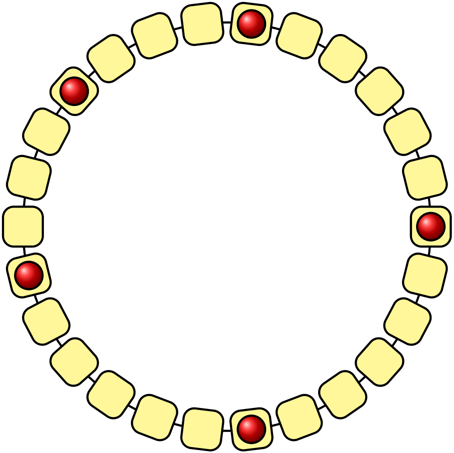
Analysis Techniques



Key ingredients:

- Taking the limit $N \to \infty$ leads to
  Brownian motion.

Analysis Techniques

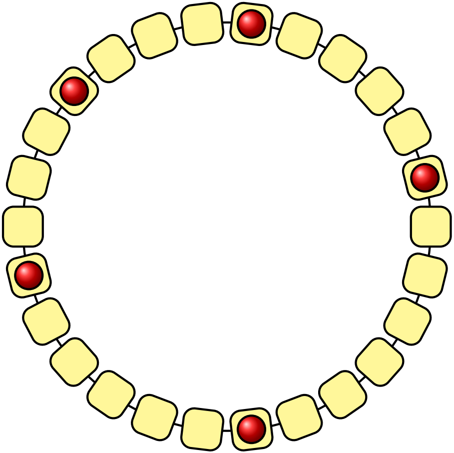


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

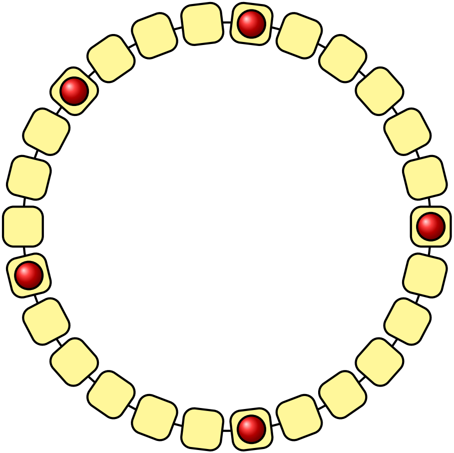
Analysis Techniques



Key ingredients:

- Taking the limit $N \to \infty$ leads to
  Brownian motion.

Analysis Techniques

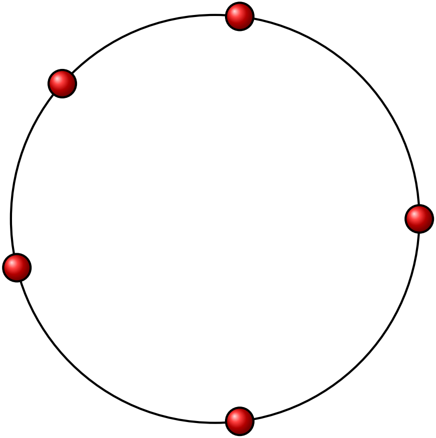


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

Analysis Techniques

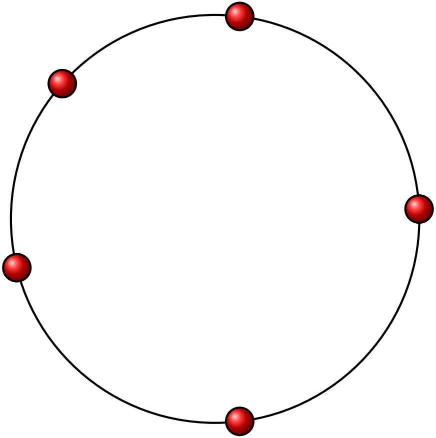


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

Analysis Techniques

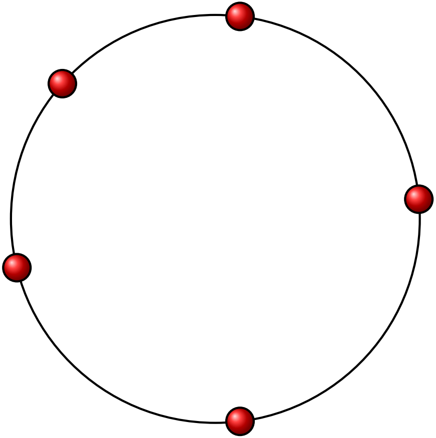


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

Analysis Techniques

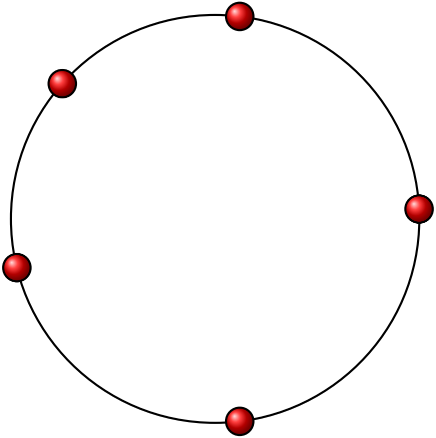


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

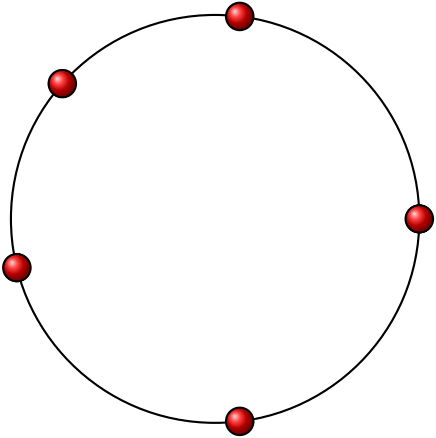
Analysis Techniques



Key ingredients:

- Taking the limit $N \to \infty$ leads to
  Brownian motion.

Analysis Techniques

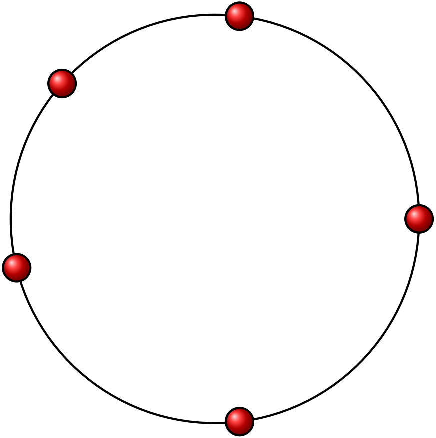


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

Analysis Techniques



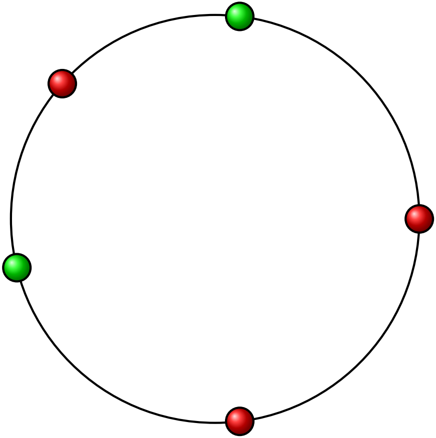
Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

- Tokens with two colours simplify the analysis of the possible interactions.

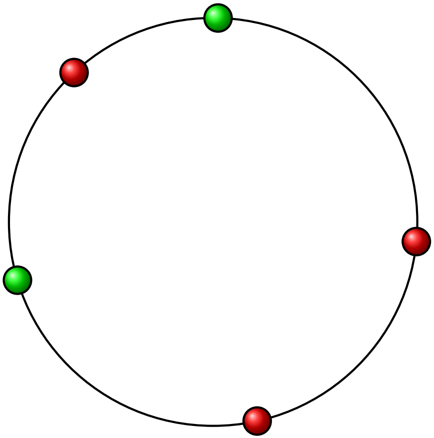
Analysis Techniques



Key ingredients:

- Taking the limit $N \to \infty$ leads to
  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

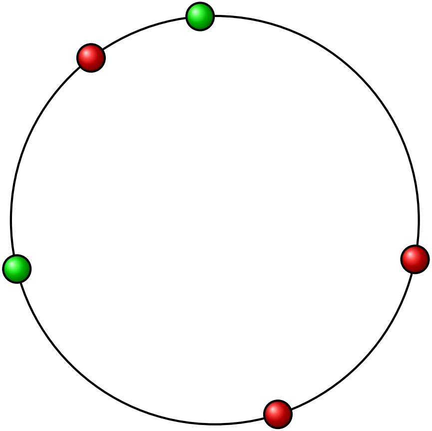


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

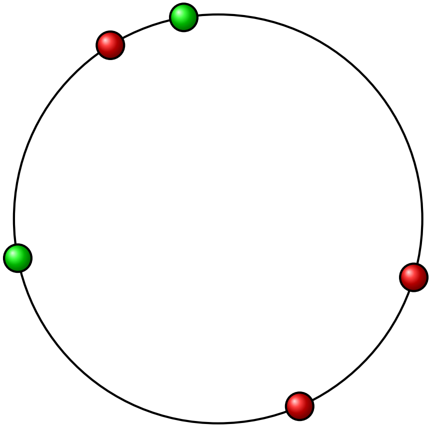


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

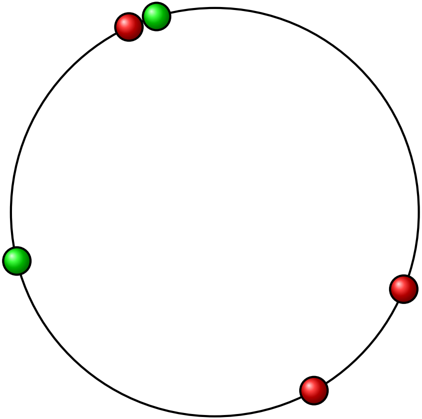


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

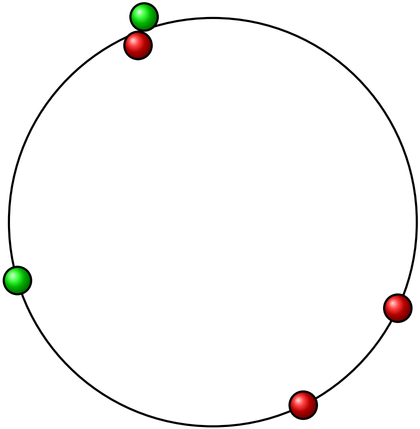


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques



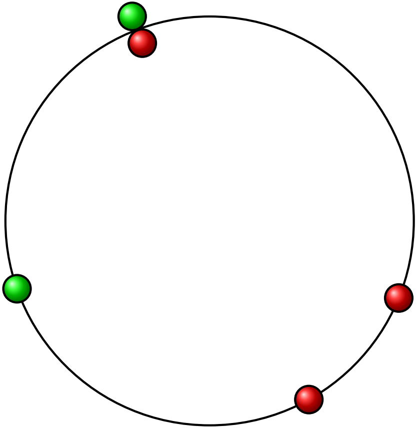
Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

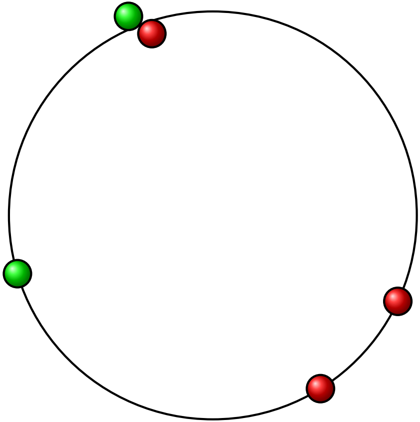


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

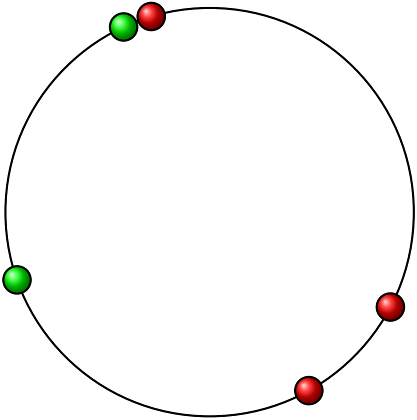


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

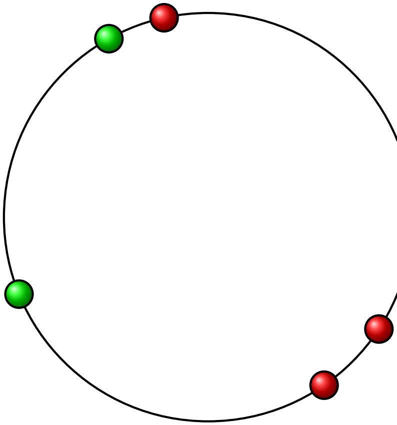


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques



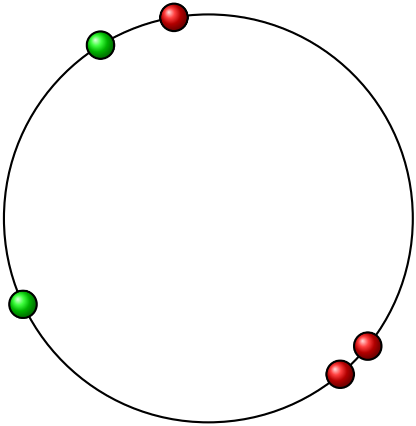
Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques



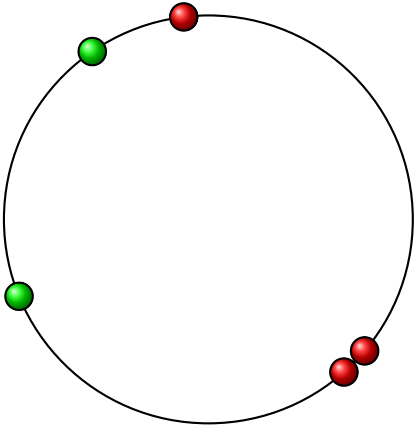
Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

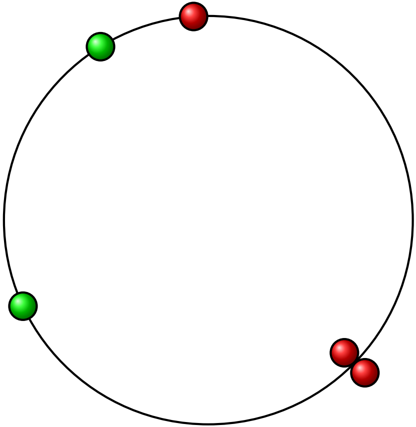


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques



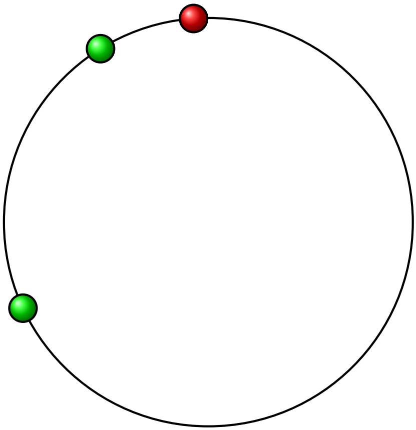
Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.

- Tokens with two colours simplify the analysis of the possible interactions.

Analysis Techniques

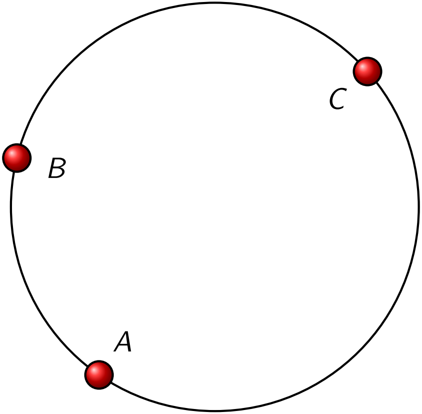


Key ingredients:

- Taking the limit $N \to \infty$ leads to

  Brownian motion.
- Tokens with two colours simplify the analysis of the possible interactions.

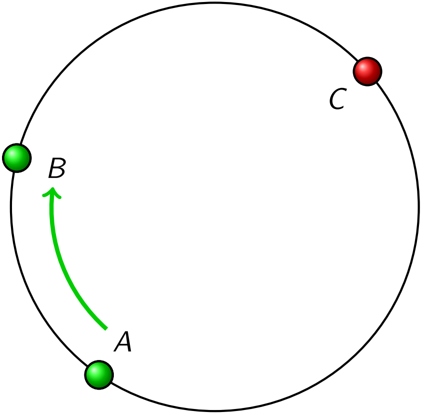
Two Colours: 3 Tokens



For any time $t \geq 0$:

$$\begin{aligned}\Pr(\mathbf{T} \leq t) &= \Pr(A \to B) - \Pr(B \to A) \\ &\quad + \Pr(B \to C) - \Pr(C \to B) \\ &\quad + \Pr(C \to A) - \Pr(A \to C)\end{aligned}$$

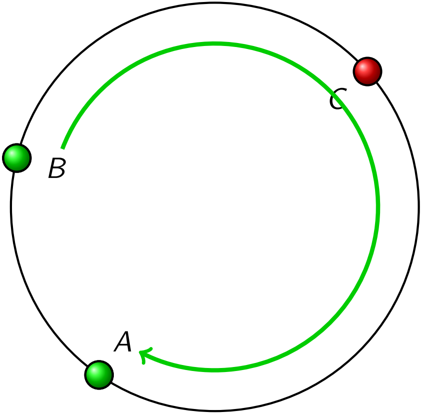
Two Colours: 3 Tokens



For any time $t \geq 0$:

$$\begin{aligned}
&\Pr(\mathbf{T} \leq t) \\
&= \quad \Pr(A \to B) - \Pr(B \to A) \\
&\quad + \Pr(B \to C) - \Pr(C \to B) \\
&\quad + \Pr(C \to A) - \Pr(A \to C)
\end{aligned}$$

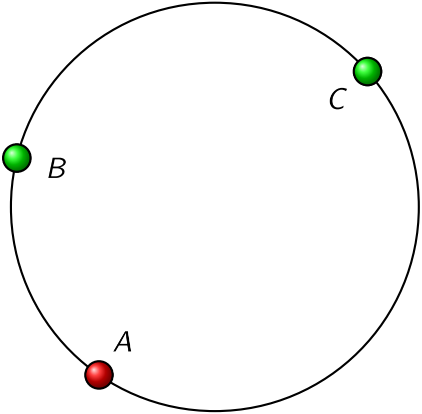
Two Colours: 3 Tokens



For any time $t \geq 0$:

$$
\begin{aligned}
\Pr(\mathbf{T} \leq t) \\
= \quad & \Pr(A \to B) - \Pr(B \to A) \\
& + \Pr(B \to C) - \Pr(C \to B) \\
& + \Pr(C \to A) - \Pr(A \to C)
\end{aligned}
$$

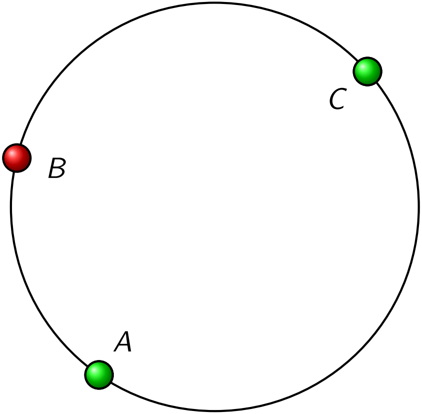
Two Colours: 3 Tokens



For any time $t \geq 0$:

$$
\begin{aligned}
\Pr(\mathbf{T} \leq t) \\
= \quad & \Pr(A \to B) - \Pr(B \to A) \\
& + \Pr(B \to C) - \Pr(C \to B) \\
& + \Pr(C \to A) - \Pr(A \to C)
\end{aligned}
$$

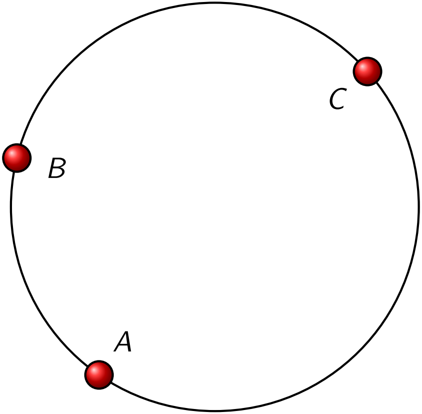
Two Colours: 3 Tokens



For any time $t \geq 0$:

$$\begin{aligned}\Pr(\mathbf{T} \leq t) \\ = \quad & \Pr(A \to B) - \Pr(B \to A) \\ & + \Pr(B \to C) - \Pr(C \to B) \\ & + \Pr(C \to A) - \Pr(A \to C)\end{aligned}$$

Two Colours: 3 Tokens



For any time $t \geq 0$:

$$
\begin{aligned}
\Pr(\mathbf{T} \leq t) \\
= \quad & \Pr(A \rightarrow B) - \Pr(B \rightarrow A) \\
& + \Pr(B \rightarrow C) - \Pr(C \rightarrow B) \\
& + \Pr(C \rightarrow A) - \Pr(A \rightarrow C)
\end{aligned}
$$

$\Pr(\mathbf{T} \leq t)$ in terms of 1-D random walk with two absorbing barriers

- distributions well-known

Principle generalizes to more than 3 tokens.

We use Mathematics . . .

with

$$e_0 = \overbrace{\frac{r\cdot(1-r)\cdot\pi^{k+2}}{2^k}}^{c_1^k\cdot(-d_2)}\cdot j_1\cdot\ldots\cdot j_k\cdot(j_1^2+\cdots+j_k^2) \qquad \text{and}$$

$$e_i = j_1\cdots j_k\cdot E_k((\boldsymbol{j}^2)^{(i+2)/2}) \qquad \text{for } i=0,2,\ldots$$

Since $e_0>0$, the power series $e_0+e_2u^2+e_4u^4+\cdots$ can be inverted. The inversion formula yields

$$\overline{F}_k(\boldsymbol{j};u) = a_0 + a_2u^2 + a_4u^4 + \cdots$$

with

$$a_0 = \frac{1}{e_0} = \frac{2^k}{r\cdot(1-r)\cdot\pi^{k+2}\cdot j_1\cdots j_k\cdot(j_1^2+\cdots+j_k^2)} \qquad \text{and}$$

$$a_i = -a_0\cdot\sum_{\ell=0,2,\ldots,i-2} a_\ell e_{i-\ell} \qquad \text{for } i=2,4,\ldots$$

It follows by an easy induction that

$$a_i = \frac{E_k((\boldsymbol{j}^2)^i)}{j_1\cdots j_k\cdot(j_1^2+\cdots+j_k^2)^{(i+2)/2}} \qquad \text{for } i=0,2,\ldots$$

Using further values of the Taylor coefficients c_i, d_i from above, a straightforward but tedious computation shows that

$$a_4 = \frac{2^{k-4}P(\boldsymbol{j})}{45\cdot\pi^{k-2}\cdot r\cdot(1-r)\cdot j_1\cdots j_k\cdot(j_1^2+\cdots+j_k^2)^3},$$

where

$$\begin{aligned} P(\boldsymbol{j}) = &\sum_{1\le i_1\le k} -3j_{i_1}^8 + \sum_{1\le i_1<i_2\le k} -9(j_{i_1}^6 j_{i_2}^2 + j_{i_1}^2 j_{i_2}^6) + \\ &\sum_{1\le i_1<i_2\le k} (720(r(1-r))^2 - 240r(1-r)+8)\, j_{i_1}^4 j_{i_2}^4 + \\ &\sum_{1\le i_1<i_2<i_3\le k} (720(r(1-r))^2 - 300r(1-r)+13)(j_{i_1}^4 j_{i_2}^2 j_{i_3}^2 + j_{i_1}^2 j_{i_2}^4 j_{i_3}^2 + j_{i_1}^2 j_{i_2}^2 j_{i_3}^4) + \\ &\sum_{1\le i_1<i_2<i_3<i_4\le k} (1440(r(1-r))^2 - 720r(1-r)+60)\, j_{i_1}^2 j_{i_2}^2 j_{i_3}^2 j_{i_4}^2 . \end{aligned}$$

We have determined the above coefficients of $P(\boldsymbol{j})$ using the computer algebra system Maple. Now it is straightforward to verify that all coefficients of $P(\boldsymbol{j})$ are negative, if $\frac{3-\sqrt{3}}{12} < r\cdot(1-r) \le \frac{1}{4}$. Those inequalities hold, if $r\in\left(\frac{1}{2}-\frac{\sqrt[4]{27}}{6}, \frac{1}{2}+\frac{\sqrt[4]{27}}{6}\right)$. □

The following lemma is used as an induction step in the proof of Lemma 10 below.

Lemma 13. *If $k\in\{2,3,\ldots\}$ and $u>0$, then*

$$\lim_{j_k\to 0}\left(j_k\cdot\overline{F}_k(j_1,\ldots,j_k;u)\right) = \frac{2}{\pi}\overline{F}_{k-1}(j_1,\ldots,j_{k-1};u),$$

where the j_i vary over the nonnegative reals. Consequently, with the Taylor expansion $\overline{F}_k(\boldsymbol{j};u) = a_{k;0}(\boldsymbol{j}) + a_{k;2}(\boldsymbol{j})u^2 + a_{k;4}(\boldsymbol{j})u^4+\cdots$ from Lemma 12 we also have

$$\lim_{j_k\to 0}\left(j_k\cdot a_{k;i}(j_1,\ldots,j_k)\right) = \frac{2}{\pi}a_{k-1;i}(j_1,\ldots,j_{k-1}).$$

Proof. As $h(0,u)=1$, it suffices to show that $\lim_{j_k\to 0}\frac{j_k\sin(j_k\pi u)}{1-\cos(j_k\pi u)} = \frac{2}{\pi u}$. This follows easily from l'Hopital's rule:

$$\begin{aligned} \lim_{j_k\to 0}\frac{j_k\sin(j_k\pi u)}{1-\cos(j_k\pi u)} &= \lim_{j_k\to 0}\frac{\sin(j_k\pi u)+j_k\cos(j_k\pi u)\cdot\pi u}{\sin(j_k\pi u)\cdot\pi u} \\ &= \frac{1}{\pi u} + \lim_{j_k\to 0}\frac{j_k\cos(j_k\pi u)}{\sin(j_k\pi u)} \\ &= \frac{1}{\pi u} + \lim_{j_k\to 0}\frac{\cos(j_k\pi u)-j_k\sin(j_k\pi u)}{\cos(j_k\pi u)\cdot\pi u} \\ &= \frac{1}{\pi u}+\frac{1}{\pi u} = \frac{2}{\pi u} \end{aligned}$$

□

Now we can prove Lemma 10 which is restated here.

Lemma 10. *For any fixed $k\in\mathbb{N}_+$ and $r\in\left(\frac{1}{2}-\frac{\sqrt[4]{27}}{6}, \frac{1}{2}+\frac{\sqrt[4]{27}}{6}\right)$, we have*

$$\sum_{\boldsymbol{j}\in\{1,\ldots,N-1\}^k}\left|\overline{F}_k(\boldsymbol{j},\boldsymbol{y};1/N) - a_0(\boldsymbol{j})\right| = O\left(\frac{(\log N)^k}{N^2}\right).$$

Proof. By Lemma 12 it is equivalent to prove

$$\sum_{\boldsymbol{j}\in\{1,\ldots,N-1\}^k}\left|\frac{a_2(\boldsymbol{j})}{N^2}+\frac{a_4(\boldsymbol{j})}{N^4}+\cdots\right| = O\left(\frac{(\log N)^k}{N^2}\right).$$

Notice that an easy induction shows that

$$\sum_{\boldsymbol{j}\in\{1,\ldots,N-1\}^k}\frac{1}{j_1\cdots j_k} = \sum_{j_k=1}^{N-1}\frac{1}{j_k}\sum_{\boldsymbol{j}\in\{1,\ldots,N-1\}^{k-1}}\frac{1}{j_1\cdots j_{k-1}} = O((\log N)^k).$$

Hence, with Lemma 12 we have

$$\sum_{\boldsymbol{j}\in\{1,\ldots,N-1\}^k}\left|\frac{a_2(\boldsymbol{j})}{N^2}\right| = \frac{1}{N^2}\cdot\sum_{\boldsymbol{j}\in\{1,\ldots,N-1\}^k}\frac{O(1)}{j_1\cdots j_k} = O\left(\frac{(\log N)^k}{N^2}\right).$$

Summary

Upper bounds for $\mathbb{E}\mathbf{T}$ for various initial configurations:

- arbitrary configuration (improved bound)
- full configuration (new bound)
- random configuration (new bound)
- flip-m configurations (new bound, $O(N)$)

Also considered:

- asynchronous version
- synchronous version with passing prob. $\neq 1/2$

Main techniques used:

- limit $\rightarrow$ Brownian motion
- two colours

Main open question is still:

- Is the 3-token equilateral conf. the worst case?

Particle Collision: CERN