Reaction Network Modeling

Molecular Programming tutorial
ASE 2015

Chemical Reaction Networks

A mathematical model.

Chemical Reaction Networks

A mathematical model.

A programming language for molecular programming in a well-mixed solution.

Chemical Reaction Networks

A mathematical model.

A programming language for molecular programming in a well-mixed solution.

"Computing in soup"

Chemical Reaction Networks

A mathematical model.

A programming language for molecular programming in a well-mixed solution.

"Computing in soup"

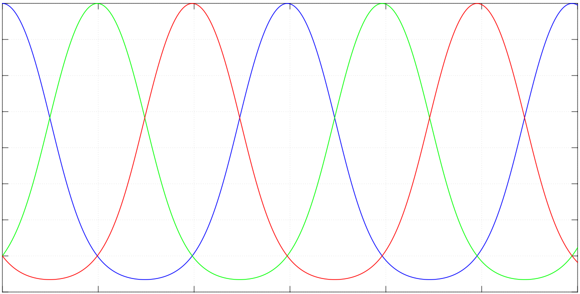
Essentially equivalent to

- Population protocols
- Petri nets
- Vector addition systems

Chemical Reaction Networks

Example: A three-phase oscillator ($\approx$ Lotka-Volterra)

$$X + Y \xrightarrow{k} 2Y$$
$$Y + Z \xrightarrow{k} 2Z$$
$$Z + X \xrightarrow{k} 2X$$



Four types of CRN research

1. Descriptive (20th century): Use CRNs to describe/explain chemical/biochemical processes.

Four types of CRN research

1. Descriptive (20th century): Use CRNs to describe/explain chemical/biochemical processes.
2. Mathematical (20th century): Use algebra, analysis, and probability to analyze structure and dynamics of CRNs.

Four types of CRN research

3. *Ad hoc* prescriptive (20th century):
    (a) Write down an (abstract) CRN that does something cool.

Four types of CRN research

3. *Ad hoc* prescriptive (20th century):
   (a) Write down an (abstract) CRN that does something cool.
   (b) Analyze it (math and/or simulation).

Four types of CRN research

3. *Ad hoc* prescriptive (20th century):
    (a) Write down an (abstract) CRN that does something cool.
    (b) Analyze it (math and/or simulation).
    (c) Find the chemicals that implement it.

Four types of CRN research

3. *Ad hoc* prescriptive (20th century):
   (a) Write down an (abstract) CRN that does something cool.
   (b) Analyze it (math and/or simulation).
   (c) Find the chemicals that implement it.
   Often takes years!

Four types of CRN research

3. *Ad hoc* prescriptive (20th century):
   (a) Write down an (abstract) CRN that does something cool.
   (b) Analyze it (math and/or simulation).
   (c) Find the chemicals that implement it.
   Often takes years!
4. Systematic prescriptive (21st century):

Four types of CRN research

3. *Ad hoc* prescriptive (20th century):
   (a) Write down an (abstract) CRN that does something cool.
   (b) Analyze it (math and/or simulation).
   (c) Find the chemicals that implement it.
   Often takes years!
4. Systematic prescriptive (21st century):

   Replace (c) with uniform translation into DNA strand displacement systems.

   Soloveichik, Seelig, and Winfree 2010

reaction network $\longrightarrow$ strand displacement system $\longrightarrow$ DNA strands

programming language	assembly language	machine language

Chen, Dalchau, Srinivas, Phillips, Cardelli, Soloveichik, and Seelig 2013.

"Real programmers code in chemistry!"

Soloveichik, 2011

The CRN Model

$\mathbf{S} = \{X_0, X_1, X_2, \ldots\}$ is the global set of species.

The CRN Model

$\mathbf{S} = \{X_0, X_1, X_2, \ldots\}$ is the global set of species.

A reaction over a finite set $S \subseteq \mathbf{S}$ is a triple

$$\rho = (\mathbf{r}, \mathbf{p}, k) \in \mathbb{N}^S \times \mathbb{N}^S \times (0, \infty)$$

with $\mathbf{r} \neq \mathbf{p}$.

The CRN Model

$\mathbf{S} = \{X_0, X_1, X_2, \ldots\}$ is the global set of species.

A reaction over a finite set $S \subseteq \mathbf{S}$ is a triple

$$\rho = (\mathbf{r}, \mathbf{p}, k) \in \mathbb{N}^S \times \mathbb{N}^S \times (0, \infty)$$

with $\mathbf{r} \neq \mathbf{p}$.

$$\begin{aligned} \mathbf{r} = \mathbf{r}(\rho) &= \text{ reactant vector of } \rho \\ \mathbf{p} = \mathbf{p}(\rho) &= \text{ product vector of } \rho \\ k = k(\rho) &= \text{ rate constant of } \rho \end{aligned}$$

The CRN Model

Example: Let $S = \{X, Y, Z\}$. The intuitive notation

$$X + Z \xrightarrow{k} 2Y + Z$$

refers to the formal reaction $(\mathbf{r}, \mathbf{p}, k)$, where

$$\mathbf{r}(X) = \mathbf{r}(Z) = 1, \;\; \mathbf{r}(Y) = 0,$$
$$\mathbf{p}(X) = 0, \;\; \mathbf{p}(Y) = 2, \;\; \mathbf{p}(Z) = 1.$$

The CRN Model

The two most widely used semantics (operational meanings) ascribed to a CRN $N = (S, R)$.

1. Deterministic mass action

   A state is a vector $\mathbf{x} \in [0, \infty)^S$.

   $\mathbf{x}(Y) =$ the concentration of Y in state $\mathbf{x}$.

   Behavior of N governed by autonomous polynomial ODEs.

The CRN Model

The two most widely used semantics (operational meanings) ascribed to a CRN $N = (S, R)$.

1. Deterministic mass action

    A state is a vector $\mathbf{x} \in [0, \infty)^S$.

    $\mathbf{x}(Y) =$ the concentration of Y in state $\mathbf{x}$.

    Behavior of N governed by autonomous polynomial ODEs.

Used to model systems with many molecules of each species.

Law of Mass Action (deterministic)

The rate of a reaction depends on its rate constant and the state!

Example. In state $\mathbf{x}$, the reactions

$$\rho_1 \equiv X + Z \xrightarrow{a} 2Y + Z$$

and

$$\rho_2 \equiv 2Y + Z \xrightarrow{b} X + Z$$

Law of Mass Action (deterministic)

The rate of a reaction depends on its rate constant and the state!

Example. In state $\mathbf{x}$, the reactions

$$\rho_1 \equiv X + Z \xrightarrow{a} 2Y + Z$$

and

$$\rho_2 \equiv 2Y + Z \xrightarrow{b} X + Z$$

have rates

$$rate_{\mathbf{x}}(\rho_1) = a\mathbf{x}(X)\mathbf{x}(Z)$$

and

Law of Mass Action (deterministic)

The rate of a reaction depends on its rate constant and the state!

Example. In state $\mathbf{x}$, the reactions

$$\rho_1 \equiv X + Z \xrightarrow{a} 2Y + Z$$

and

$$\rho_2 \equiv 2Y + Z \xrightarrow{b} X + Z$$

have rates

$$rate_{\mathbf{x}}(\rho_1) = a\mathbf{x}(X)\mathbf{x}(Z)$$

and

$$rate_{\mathbf{x}}(\rho_2) = b\mathbf{x}(Y)^2\mathbf{x}(Z).$$

Example (continued). The CRN

$$X + Z \xrightarrow{a} 2Y + Z$$
$$2Y + Z \xrightarrow{b} X + Z$$

is governed by the system

$$\frac{dx}{dt} = -axz + by^2z$$
$$\frac{dy}{dt} =$$
$$\frac{dz}{dt} =$$

Example (continued). The CRN

$$X + Z \xrightarrow{a} 2Y + Z$$
$$2Y + Z \xrightarrow{b} X + Z$$

is governed by the system

$$\frac{dx}{dt} = -axz + by^2 z$$
$$\frac{dy}{dt} = 2axz - 2by^2 z$$
$$\frac{dz}{dt} =$$

Example (continued). The CRN

$$X + Z \xrightarrow{a} 2Y + Z$$
$$2Y + Z \xrightarrow{b} X + Z$$

is governed by the system

$$\begin{aligned} \frac{dx}{dt} &= -axz + by^2 z \\ \frac{dy}{dt} &= 2axz - 2by^2 z \\ \frac{dz}{dt} &= 0 \quad (Z \text{ is a catalyst.}) \end{aligned}$$

The CRN Model

Open question: Are deterministic mass action CRNs Turing universal?

The CRN Model

Open question: Are deterministic mass action CRNs Turing universal?

Conjecture: Probably so. Stay tuned!

The CRN Model

Open question: Are deterministic mass action CRNs Turing universal?

Conjecture: Probably so. Stay tuned!

A harder question: If so, is this universality robust?

The CRN Model

Open question: Are deterministic mass action CRNs Turing universal?

Conjecture: Probably so. Stay tuned!

A harder question: If so, is this universality robust?

What is robustness here?

The CRN Model

Klinge, Lathrop, and Lutz 2015: Uniform translation of a nondeterministic finite automaton M into a deterministic mass action CRN N.

- N simulates M
- size(N) linear in size(M)
- N robust with respect to
  - input signal
  - initial concentrations
  - output
  - rate constants of reactions
- Proof, not simulation.

The CRN Model

2. Stochastic mass action

   A state is a vector $\mathbf{x} \in \mathbb{N}^S$.
   $\mathbf{x}(Y) =$ the count of Ys in state $\mathbf{x}$.
   Behavior is governed by continuous time Markov chains.

The CRN Model

2. Stochastic mass action

   A state is a vector $\mathbf{x} \in \mathbb{N}^S$.
   $\mathbf{x}(Y) =$ the count of Ys in state $\mathbf{x}$.
   Behavior is governed by continuous time Markov chains.

Used to model systems in which small counts of molecules are significant.

The CRN Model

2. Stochastic mass action

   A state is a vector $\mathbf{x} \in \mathbb{N}^S$.
   $\mathbf{x}(Y) =$ the count of Ys in state $\mathbf{x}$.
   Behavior is governed by continuous time Markov chains.

The CRN Model

2. Stochastic mass action

   A state is a vector $\mathbf{x} \in \mathbb{N}^S$.
   $\mathbf{x}(Y) =$ the count of Ys in state $\mathbf{x}$.
   Behavior is governed by continuous time Markov chains.

Theorem (Kurtz 1972). Roughly stated, deterministic mass action is the limiting case of stochastic mass action as populations and volumes go to infinity.

The CRN Model

The computational power of stochastic mass action CRNs.

- Angluin, Aspnes, and Eisenstat 2006;
  Chen, Doty, and Soloveichik 2012:

  Stable computation (correctness on all paths) can only compute semilinear predicates and functions.

The CRN Model

The computational power of stochastic mass action CRNs.

- Angluin, Aspnes, and Eisenstat 2006;
  Chen, Doty, and Soloveichik 2012:

  Stable computation (correctness on all paths) can only compute semilinear predicates and functions.

- Soloveichik, Cook, Winfree, and Bruck 2008:

  Allowing any error $\epsilon > 0$ makes the model Turing universal.

The CRN Model

The computational power of stochastic mass action CRNs.

- Angluin, Aspnes, and Eisenstat 2006;
  Chen, Doty, and Soloveichik 2012:

  Stable computation (correctness on all paths) can only compute semilinear predicates and functions.

- Soloveichik, Cook, Winfree, and Bruck 2008:

  Allowing any error $\epsilon > 0$ makes the model Turing universal.

- The known simulation is not robust.

The CRN Model

Recall For $\lambda \in (0, \infty)$, the exponential distribution with rate λ is the probability measure on $[0, \infty)$ given by the c.d.f.

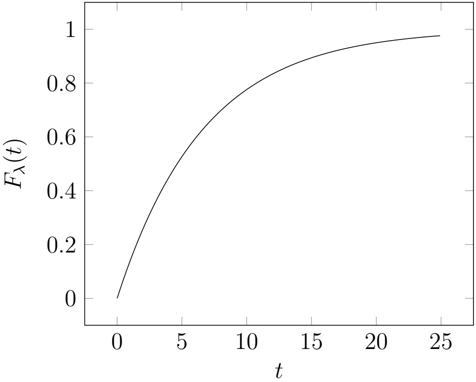
$$F_\lambda : [0, \infty) \to [0, 1]$$
$$F_\lambda(t) = 1 - e^{-\lambda t}$$

The CRN Model

Recall For $\lambda \in (0, \infty)$, the exponential distribution with rate λ is the probability measure on $[0, \infty)$ given by the c.d.f.

$$F_\lambda : [0, \infty) \to [0, 1]$$

$$F_\lambda(t) = 1 - e^{-\lambda t}$$



A continuous-time Markov chain (CTMC) is an ordered triple

$$C = (Q, \lambda, \pi),$$

where

- Q is a countable set of states;

A continuous-time Markov chain (CTMC) is an ordered triple

$$C = (Q, \lambda, \pi),$$

where

- Q is a countable set of states;
- $\lambda : Q \times Q$ is the rate matrix and satisfies $\lambda(q, q) = 0$ for all $q \in Q$;

A continuous-time Markov chain (CTMC) is an ordered triple

$$C = (Q, \lambda, \pi),$$

where

- Q is a countable set of states;
- $\lambda : Q \times Q$ is the rate matrix and satisfies $\lambda(q, q) = 0$ for all $q \in Q$;
- $\pi \in \Delta(Q)$ is the state initialization.

A continuous-time Markov chain (CTMC) is an ordered triple

$$C = (Q, \lambda, \pi),$$

where

- Q is a countable set of states;
- $\lambda : Q \times Q$ is the rate matrix and satisfies $\lambda(q, q) = 0$ for all $q \in Q$;
- $\pi \in \Delta(Q)$ is the state initialization.

$$\Delta(Q) = \{\text{probability measures on } Q\}.$$

Semantics of a CTMC $C = (Q, \lambda, \pi)$.

- At each time $t \in [0, \infty)$, C is probabilistically in some state state(t).

Semantics of a CTMC $C = (Q, \lambda, \pi)$.

- At each time $t \in [0, \infty)$, C is probabilistically in some state state(t).
- state(0) is chosen according to π.

Semantics of a CTMC $C = (Q, \lambda, \pi)$.

- At each time $t \in [0, \infty)$, C is probabilistically in some state state(t).
- state(0) is chosen according to π.
- For each state $q \in Q$,

$$\lambda_q = \sum_{r \in Q} \lambda(q, r)$$

is the rate out of q.

- If $\lambda_q = 0$, then state q is terminal.

- If $\lambda_q = 0$, then state q is terminal.
- if $\lambda_q > 0$, then state q is nonterminal If C enters state q at time t, then the sojourn time for which C remains in state q before moving to a new state has exponential distribution with rate λ_q.

- If $\lambda_q = 0$, then state q is terminal.
- if $\lambda_q > 0$, then state q is nonterminal If C enters state q at time t, then the sojourn time for which C remains in state q before moving to a new state has exponential distribution with rate λ_q.

  When C does move to a new state, this state is r with probability

$$p(q, r) = \frac{\lambda(q, r)}{\lambda_q}.$$

Stochastic mass action semantics of a CRN $N = (S, R)$
$\equiv$ semantics of the CTMC $C = (Q, \lambda, \pi)$ defined as follows

Stochastic mass action semantics of a CRN $N = (S, R)$
$\equiv$ semantics of the CTMC $C = (Q, \lambda, \pi)$ defined as follows

- $Q = \mathbb{N}^S$

Stochastic mass action semantics of a CRN $N = (S, R)$
$\equiv$ semantics of the CTMC $C = (Q, \lambda, \pi)$ defined as follows

- $Q = \mathbb{N}^S$

- $\pi \in \Delta(Q)$ is how you want to initialize the species counts in N.

Stochastic mass action semantics of a CRN $N = (S, R)$
$\equiv$ semantics of the CTMC $C = (Q, \lambda, \pi)$ defined as follows

- $Q = \mathbb{N}^S$

- $\pi \in \Delta(Q)$ is how you want to initialize the species counts in N.

- Still need $\lambda : \mathbb{N}^S \times \mathbb{N}^S \to [0, \infty)$.

Law of Mass Action (stochastic)

Intuition: The rate of a unary reaction

$$\rho \equiv X \xrightarrow{k} \text{<anything>}$$

in a state $\mathbf{x} \in \mathbb{N}^S$ is

$$rate_{\mathbf{x}}(\rho) = k\mathbf{x}(X).$$

Law of Mass Action (stochastic)

Intuition: The rate of a unary reaction

$$\rho \equiv X \xrightarrow{k} \text{<anything>}$$

in a state $\mathbf{x} \in \mathbb{N}^S$ is

$$rate_{\mathbf{x}}(\rho) = k\mathbf{x}(X).$$

The rate of a binary reaction

$$\rho \equiv X + Y \xrightarrow{k} \text{<anything>}$$

in a state $\mathbf{x} \in \mathbb{N}^S$ and volume V is

$$rate_{\mathbf{x}}(\rho) = \begin{cases} k\mathbf{x}(X)\mathbf{x}(Y)/V & \text{if } X \neq Y \\ k\mathbf{x}(X)(\mathbf{x}(X) - 1)/V & \text{if } X = Y \end{cases}$$

The rate of a reaction $\rho = (\mathbf{r}, \mathbf{p}, k)$ in a state $\mathbf{x} \in \mathbb{N}^S$ is

$$\mathsf{rate}_{\mathbf{x}}(\rho) = \frac{k}{V^{\|\mathbf{r}\|-1}} \prod_{Y \in S} \frac{\mathbf{x}(Y)!}{(\mathbf{x}(Y) - \mathbf{r}(Y))!},$$

where V is the volume of the solution and

$$\|\mathbf{r}\| = \sum_{Y \in S} \mathbf{r}(Y).$$

Gillespie, 1977

In the CTMC $C = (\mathbb{N}^S, \lambda, \pi)$ for a CRN $N = (S, R)$ we define the rate matrix $\lambda : Q \times Q \to [0, \infty)$ by

$$\lambda(\mathbf{x}, \mathbf{y}) = \sum \left\{ rate_{\mathbf{x}}(\rho) \mid \rho \in R \; takes \; \mathbf{x} \; to \; \mathbf{y} \right\}.$$

We have now completely specified the most often-used semantics (kinetics) for CRNs, namely,

deterministic mass action

and

stochastic mass action.

We have now completely specified the most often-used semantics (kinetics) for CRNs, namely,

deterministic mass action

and

stochastic mass action.

Caution: These are sometimes called

mass action

and

stochastic,

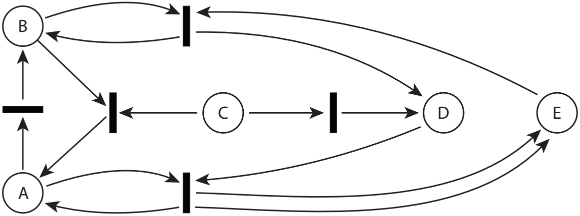
respectively.

Relating CRNs to Other Models

1. Petri nets (Petri 1962).

   A Petri net is a directed bipartite graph whose vertices are places (circles) and transitions (bars).

   Example (CSWB 2009)



Relating CRNs to Other Models

A state of a Petri net is a placement of zero or more tokens, all of which are identical, on its places.

Relating CRNs to Other Models

A state of a Petri net is a placement of zero or more tokens, all of which are identical, on its places.

A transition is enabled if there is a token available for each of its incoming edges.

Relating CRNs to Other Models

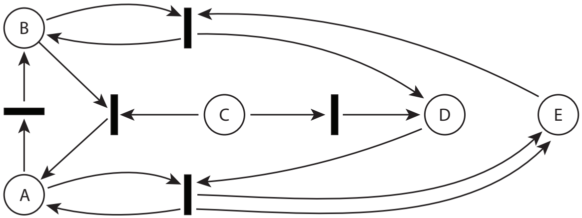
A state of a Petri net is a placement of zero or more tokens, all of which are identical, on its places.

A transition is enabled if there is a token available for each of its incoming edges.

In this case, the transition may fire, causing these tokens to be removed and tokens to be placed at the targets of its outgoing edges.

Relating CRNs to Other Models

It is routine to check that the Petri net



has the same behavior as the CRN

$$A \to B \qquad A + D \to A + 2E$$
$$C \to D \qquad B + E \to B + D$$
$$B + C \to A$$

Relating CRNs to Other Models

2. Vector addition systems (Karp and Miller 1969)

   Fix $n \in \mathbb{Z}^+$ and a finite set V of vectors $\mathbf{v} \in \mathbb{Z}^n$.
   Consider walks in $\mathbb{N}^n$ that start at a location $\mathbf{x}_0 \in \mathbb{N}^n$ and proceed in steps, each of which adds a vector in V to the current location, leading to a new location in $\mathbb{N}^n$.

Relating CRNs to Other Models

Example (CSWB 2009) The CRN

$$A \to B \qquad A + D \to A + 2E$$
$$C \to D \qquad B + E \to B + D$$
$$B + C \to A$$

is equivalent to the vector addition system

A	B	C	D	E	F	G
(-1	1	0	0	0	0	0)
(0	0	-1	1	0	0	0)
(1	-1	-1	0	0	0	0)
(-1	0	0	-1	0	1	0)
(1	0	0	0	2	-1	0)
(0	-1	0	0	-1	0	1)
(0	1	0	1	0	0	-1)

Relating CRNs to Other Models

3. Population protocols (Angluin, Aspnes, and Eisenstat 2008)

   These are essentially CRNs in which every reaction has exactly two reactants and two products.

   Note that this fixes the total population.

   Example (AAE 2008) The population protocol

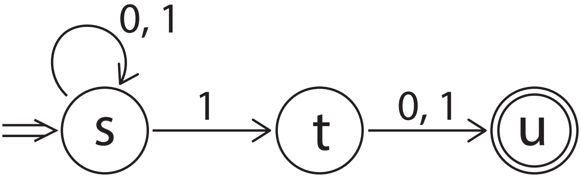
$$X + Y \to X + B \qquad\qquad X + B \to X + X$$
$$X + Y \to Y + B \qquad\qquad Y + B \to Y + Y$$

   computes approximate majority.

A CRN That Simulates an NFA

with Titus Klinge and Jim Lathrop

Let M be the NFA



Objective: Design a CRN N that simulates M in deterministic mass action.

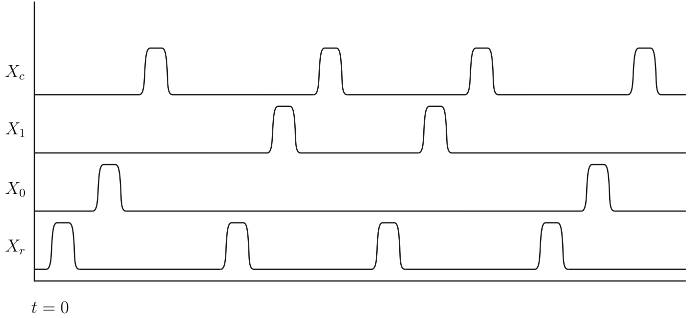
A CRN That Simulates an NFA

M gets its input $w \in \{0, 1\}^*$ sequentially, so N should get its input sequentially.

A CRN That Simulates an NFA

M gets its input $w \in \{0,1\}^*$ sequentially, so N should get its input sequentially.

The user will provide an input $w = b_0 b_1 \cdots b_{n-1} \in \{0,1\}^n$ approximately in the square-wave form



A CRN That Simulates an NFA

We thus have input species X_0, X_1, X_c, X_r.

A CRN That Simulates an NFA

We thus have input species X_0, X_1, X_c, X_r.
We also have

- state species $Y_s, Y_t, Y_u,$

A CRN That Simulates an NFA

We thus have input species X_0, X_1, X_c, X_r.
We also have

- state species Y_s, Y_t, Y_u,
- portal species Z_s, Z_t, Z_u, and

A CRN That Simulates an NFA

We thus have input species X_0, X_1, X_c, X_r.
We also have

- state species Y_s, Y_t, Y_u,
- portal species Z_s, Z_t, Z_u, and
- dual species $\overline{Y_s}, \overline{Y_t}, \overline{Y_u}, \overline{Z_s}, \overline{Z_t}, \overline{Z_u}$.

Reactions

- For each state q,

$$X_r + Z_q \to X_r + \overline{Z_q}.$$

reset

Reactions

- For each state q,

$$X_r + Z_q \to X_r + \overline{Z_q}. \qquad \text{reset}$$

- For each transition $q \xrightarrow{b} r$,

$$X_b + Y_q + \overline{Z_r} \to X_b + Y_q + Z_q. \qquad \text{compute}$$

Reactions

- For each state q,

$$X_r + Z_q \to X_r + \overline{Z_q}.$$

reset

- For each transition $q \xrightarrow{b} r$,

$$X_b + Y_q + \overline{Z_r} \to X_b + Y_q + Z_q.$$

compute

- For each state q

$$X_c + Z_q + \overline{Y_q} \to X_c + Z_q + Y_q$$
$$X_c + \overline{Z_q} + Y_q \to X_c + \overline{Z_q} + \overline{Y_q}$$

copy

Reactions

- For each state q,

$$2Y_q + \overline{Y_q} \to 3Y_q$$
$$2\overline{Y_q} + Y_q \to 3\overline{Y_q}$$

error correction

Summary

CRNs are programs.

Summary

CRNs are programs.

CRNs are fun.

Summary

CRNs are programs.

CRNs are fun.

CRNs are the future.

Summary

CRNs are programs.

CRNs are fun.

CRNs are the future.

CRNs are magnets for good students.

Summary

CRNs are programs.

CRNs are fun.

CRNs are the future.

CRNs are magnets for good students.

Thank you!