Update Consistency for Wait-free Concurrent Objects

Matthieu Perrin   Achour Mostefaoui   Claude Jard

LINA, University of Nantes

May 26th, 2015
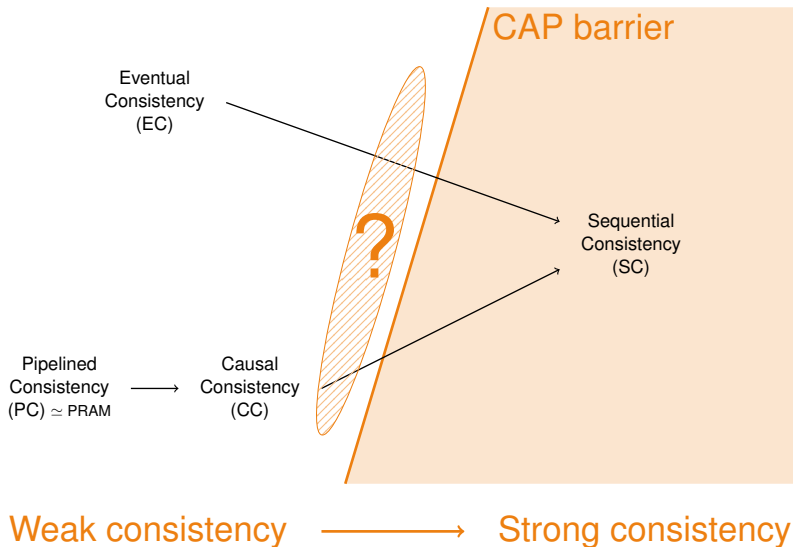
Architecture of a distributed application

Shared Objects in Distributed Systems

- Strong **C**onsistency
- **A**vailability
- **P**artition tolerance

Impossible [Gilbert, Lynch, 2002]

A map of weak consistency

Plan

1. Specifying shared objects

2. Towards Pipelined Convergence?

3. Update Consistency

Specifying shared objects

Sequential specification

▶ How the object behaves in a sequential environment
  - ▸ Memory
  - ▸ Set
  - ▸ Data base

Consistency criterion

▶ How distribution impacts the sequential behaviour
  - ▸ Causal consistency
  - ▸ Eventual consistency
  - ▸ Sequential consistency

Update-Query Abstract Data Type (UQ-ADT)

Transition system

- ▶ Update operations
  - ‣ change the state
  - ‣ have no output

- ▶ Query operation
  - ‣ loop over the states
  - ‣ return a value

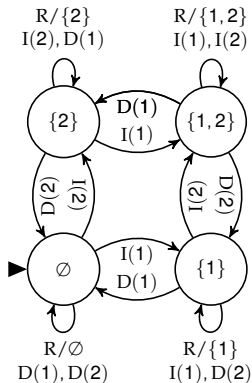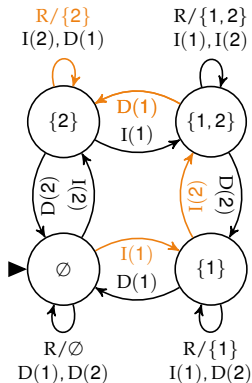# Update-Query Abstract Data Type (UQ-ADT)

Transition system

▶ **Update operations**
  ▸ change the state
  ▸ have no output

$\{I(1), I(2), D(1), D(2)\}$

▶ **Query operation**
  ▸ loop over the states
  ▸ return a value

$\{R/\emptyset, R/\{1\}, R/\{2\}, R/\{1, 2\}\}$

E.g. set $\mathscr{S}_{\{1,2\}}$

Update-Query Abstract Data Type (UQ-ADT)

Transition system

E.g. set $\mathscr{S}_{\{1,2\}}$

- ▶ Update operations
  - ▶ change the state
  - ▶ have no output

  $\{I(1), I(2), D(1), D(2)\}$

- ▶ Query operation
  - ▶ loop over the states
  - ▶ return a value

  $\{R/\emptyset, R/\{1\}, R/\{2\}, R/\{1,2\}\}$

Sequential specification

- ▶ Set of all the paths

$I(1)I(2)D(1)R/\{2\}^{\omega} \in L\left(\mathscr{S}_{\{1,2\}}\right)$

Distributed history

Distributed history $H \in \mathbb{H}$

- ▶ Partial order of events, labelled by updates and queries
- ▶ Each event has a finite past

Distributed history

Distributed history $H \in \mathbb{H}$

▶ Partial order of events, labelled by updates and queries

▶ Each event has a finite past

Wait-free system

▶ Asynchronous processes

▶ Asynchronous lossless channel

▶ Any number of crashes

Admitted history

Events: method calls

Labels: method name and returned value

Order: process order

Consistency criterion

Sequential specification

▶ $L :$ UQ-ADT $\rightarrow \mathscr{P}((U \cup Q)^{\omega} \cup (U \cup Q)^{\star})$

Consistency criterion

▶ $C :$ UQ-ADT $\rightarrow \mathscr{P}(\mathbb{H})$
▶ E.g. Sequential Consistency
   ▸ $SC : T \mapsto \{H \in \mathbb{H} : \text{lin}(H) \cap L(T) \neq \varnothing\}$
▶ Strength
   ▸ C_1 stronger than C_2:
      ▸ $\forall T, C_1(T) \subset C_2(T)$
   ▸ C strong:
      ▸ not all UQ-ADTs can be implemented in a wait-free system

Eventual Consistency

Definition

▶ If the processes stop updating, they eventually converge

$$EC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \lor \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$ and all the queries of $Q_H \setminus Q'$ can be done in the same state
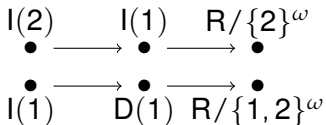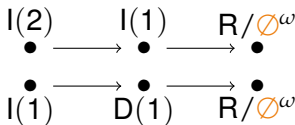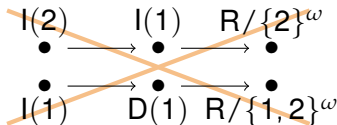
Eventual Consistency

Definition

▶ If the processes stop updating, they eventually converge

$$EC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$ and all the queries of $Q_H \setminus Q'$ can be done in the same state

Examples

$$
\begin{array}{ccc}
I(2) & I(1) & R/\{2\}^{\omega} \\
\bullet \longrightarrow & \bullet \longrightarrow & \bullet \\
\\
\bullet \longrightarrow & \bullet \longrightarrow & \bullet \\
I(1) & D(1) & R/\{1,2\}^{\omega}
\end{array}
$$

Eventual Consistency

Definition

▶ If the processes stop updating, they eventually converge

$$EC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$ and all the queries of $Q_H \setminus Q'$ can be done in the same state
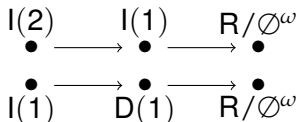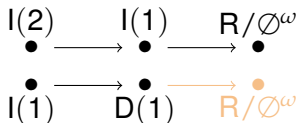
Examples

Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \mathrm{lin}(H_{U_H \cup p}) \cap L(T) \neq \varnothing\}$$
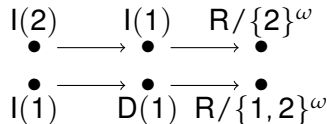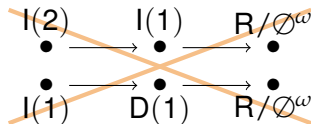
Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \text{lin}(H_{U_H \cup p}) \cap L(T) \neq \varnothing\}$$

Examples

$$
\begin{array}{ccc}
I(2) & I(1) & R/\varnothing^\omega \\
\bullet \longrightarrow & \bullet \longrightarrow & \bullet \\
\bullet & \bullet & \bullet \\
I(1) & D(1) & R/\varnothing^\omega
\end{array}
$$

Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \mathrm{lin}(H_{U_H \cup p}) \cap L(T) \neq \varnothing\}$$

Examples

$$
\begin{array}{ccc}
\mathsf{I}(2) & \mathsf{I}(1) & \mathsf{R}/\varnothing^\omega \\
\bullet \longrightarrow & \bullet \longrightarrow & \bullet \\
\bullet \longrightarrow & \bullet \longrightarrow & \bullet \\
\mathsf{I}(1) & \mathsf{D}(1) & \mathsf{R}/\varnothing^\omega
\end{array}
$$

Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \text{lin}(H_{U_H \cup p}) \cap L(T) \neq \emptyset\}$$
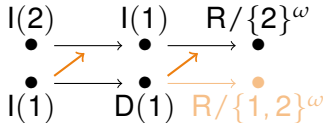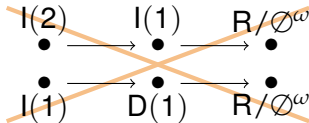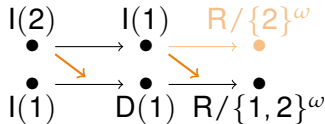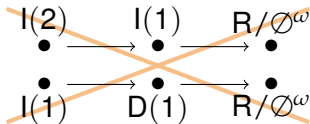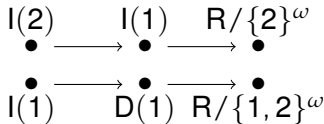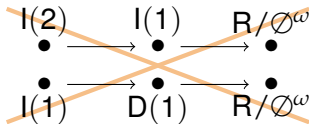
Examples

Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \mathrm{lin}(H_{U_H \cup p}) \cap L(T) \neq \varnothing\}$$
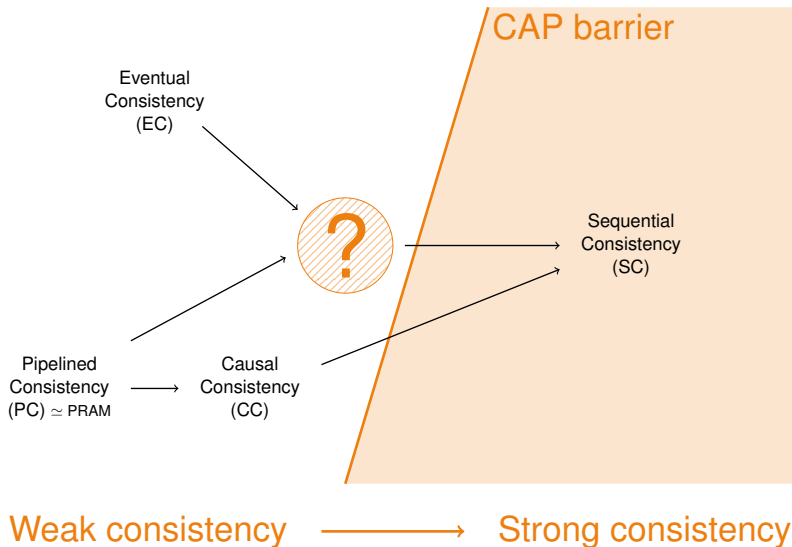
Examples

Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \text{lin}(H_{U_H \cup p}) \cap L(T) \neq \varnothing\}$$

Examples

Pipelined Consistency

Definition

▶ For each process, the history could be sequentially consistent if the other processes read correctly

$$PC : T \mapsto \{H \in \mathbb{H} : \forall p, \lin(H_{U_H \cup p}) \cap L(T) \neq \varnothing\}$$

Examples

A map of weak consistency

Weak consistency ⟶ Strong consistency

Towards Pipelined Convergence?

Definition

$$EC + PC : T \mapsto EC(T) \cap PC(T)$$

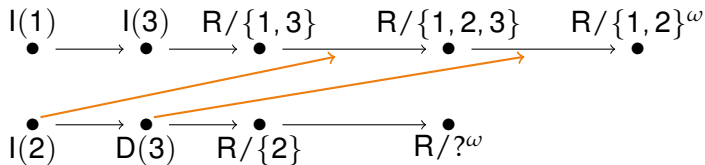# Towards Pipelined Convergence?

$$EC + PC : T \mapsto EC(T) \cap PC(T)$$

Strength

Towards Pipelined Convergence?

$$EC + PC : T \mapsto EC(T) \cap PC(T)$$

Strength

Towards Pipelined Convergence?

$$EC + PC : T \mapsto EC(T) \cap PC(T)$$

Strength

I(1) → I(3) → R/{1,3} → R/{1,2,3} → R/{1,2}$^\omega$

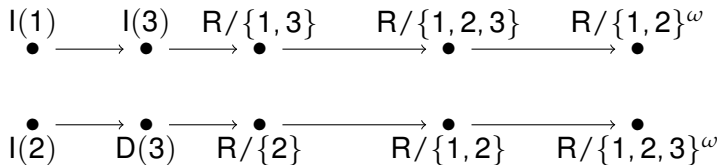I(2) → D(3) → R/{2} → R/?$^\omega$
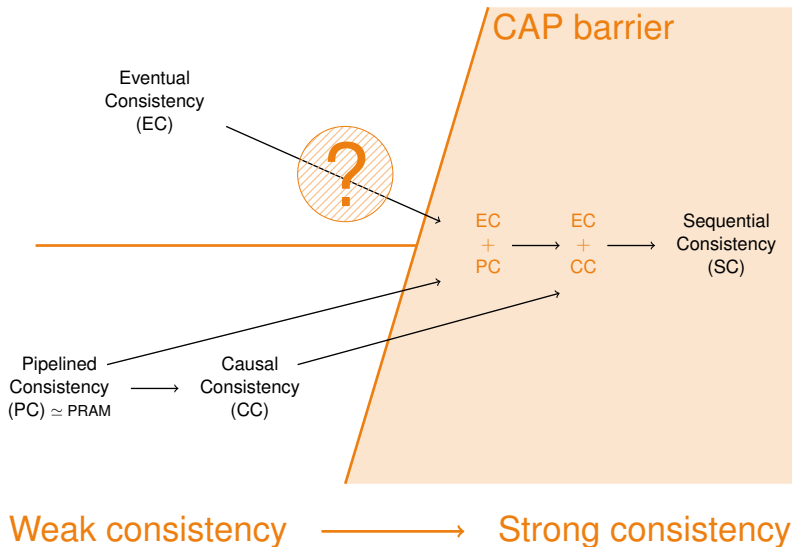
Towards Pipelined Convergence?

Definition

$$EC + PC : T \mapsto EC(T) \cap PC(T)$$

Strength

A map of weak consistency

Update Consistency

Definition of Eventual Consistency

$$EC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$:

all the queries of $Q_H \setminus Q'$ can be done in the same state

Update Consistency

Definition of Update Consistency

$$UC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$: $lin(H_{E \setminus Q'}) \cap L(T) \neq \varnothing$

all the queries of $Q_H \setminus Q'$ can be done in the
same state

Examples

$$\overset{I(1)}{\bullet} \xrightarrow{R/\{1\}} \overset{D(2)}{\bullet} \xrightarrow{} \overset{R/\{1\}^{\omega}}{\bullet}$$

$$\overset{\bullet}{I(2)} \xrightarrow{R/\{2\}} \overset{\bullet}{D(1)} \xrightarrow{} \overset{\bullet}{R/\{1\}^{\omega}}$$

Update Consistency
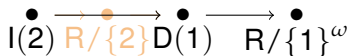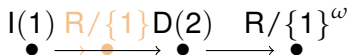
Definition of Update Consistency

$$UC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$: $lin(H_{E \setminus Q'}) \cap L(T) \neq \varnothing$

all the queries of $Q_H \setminus Q'$ can be done in the same state

Examples

$$\underset{\bullet}{I(1)} \xrightarrow{R/\{1\}} \underset{\bullet}{\bullet} \underset{\bullet}{D(2)} \xrightarrow{R/\{1\}^{\omega}} \underset{\bullet}{\bullet}$$

$$\underset{\bullet}{I(2)} \xrightarrow{R/\{2\}} \underset{\bullet}{D(1)} \xrightarrow{R/\{1\}^{\omega}} \underset{\bullet}{\bullet}$$

Update Consistency
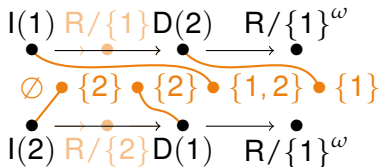
Definition of Update Consistency

$$UC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$: $lin(H_{E \setminus Q'}) \cap L(T) \neq \varnothing$

all the queries of $Q_H \setminus Q'$ can be done in the same state

Examples

I(1) $R/\{1\}$ D(2) $R/\{1\}^\omega$

$\varnothing$ {2} {2} {1,2} {1}

I(2) $R/\{2\}$ D(1) $R/\{1\}^\omega$

Update Consistency

Definition of Update Consistency

$$UC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$: $lin(H_{E \setminus Q'}) \cap L(T) \neq \emptyset$

all the queries of $Q_H \setminus Q'$ can be done in the
same state

Examples

$$\underset{\bullet}{I(1)} \xrightarrow{} \underset{\bullet}{R/\{1\}} \underset{\bullet}{D(2)} \xrightarrow{} \underset{\bullet}{R/\{1\}^\omega}$$

$$\underset{\bullet}{I(1)} \xrightarrow{} \underset{\bullet}{R/\{1\}} \underset{\bullet}{D(2)} \xrightarrow{} \underset{\bullet}{R/\{1,2\}^\omega}$$

$$\underset{\bullet}{I(2)} \xrightarrow{} \underset{\bullet}{R/\{2\}} \underset{\bullet}{D(1)} \xrightarrow{} \underset{\bullet}{R/\{1\}^\omega}$$

$$\underset{\bullet}{I(2)} \xrightarrow{} \underset{\bullet}{R/\{2\}} \underset{\bullet}{D(1)} \xrightarrow{} \underset{\bullet}{R/\{1,2\}^\omega}$$

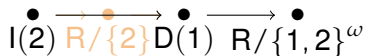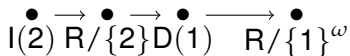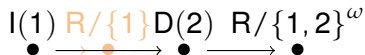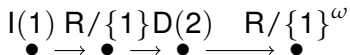# Update Consistency

Definition of Update Consistency

$$UC : T \mapsto \{H \in \mathbb{H} : \text{Write4Ever} \vee \text{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$: $lin(H_{E \setminus Q'}) \cap L(T) \neq \emptyset$

all the queries of $Q_H \setminus Q'$ can be done in the same state

Examples

Matthieu Perrin

Update Consistency

Definition of Update Consistency

$$UC : T \mapsto \{H \in \mathbb{H} : \textit{Write4Ever} \vee \textit{Converge}\}$$

Write4Ever $\equiv |U_H| = \infty$

Converge $\equiv \exists Q' \subset Q_H$ with $|Q'| < \infty$: $lin(H_{E \setminus Q'}) \cap L(T) \neq \emptyset$

all the queries of $Q_H \setminus Q'$ can be done in the
same state

Examples

Implementation

Idea

- ▶ Choose a total order *a priori*
- ▶ Update: notify everyone
- ▶ Store all the updates
- ▶ Query: execute the updates in the chosen order

Example

Implementation

Idea

▶ Choose a total order *a priori*
▶ Update: notify everyone
▶ Store all the updates
▶ Query: execute the updates in the chosen order

Example

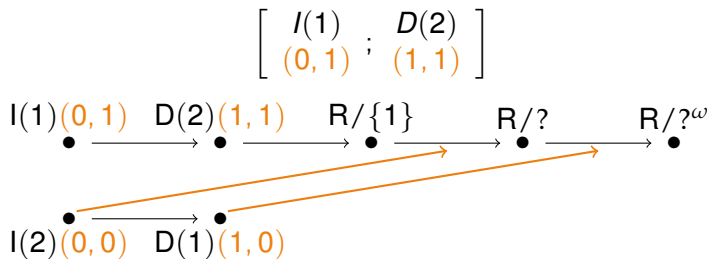$$\left[\begin{array}{cc} I(1) & D(2) \\ (0,1) & (1,1) \end{array}\right]$$

Implementation

Idea

▶ Choose a total order *a priori*
▶ Update: notify everyone
▶ Store all the updates
▶ Query: execute the updates in the chosen order

Example

$$\left[\begin{array}{ccc} I(2) & I(1) & D(2) \\ (0,0) & (0,1) & (1,1) \end{array} \right]$$

I(1)(0,1)   D(2)(1,1)   R/{1}   R/{1}   R/?$^\omega$

I(2)(0,0)   D(1)(1,0)

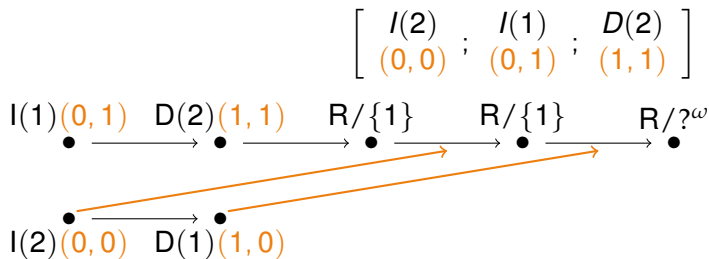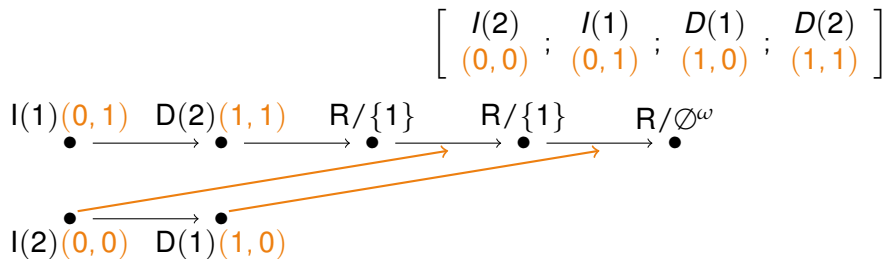# Implementation

Idea

▶ Choose a total order *a priori*
▶ Update: notify everyone
▶ Store all the updates
▶ Query: execute the updates in the chosen order

Example

$$\left[\begin{array}{cccc} I(2) & ; & I(1) & ; & D(1) & ; & D(2) \\ (0,0) & & (0,1) & & (1,0) & & (1,1) \end{array}\right]$$
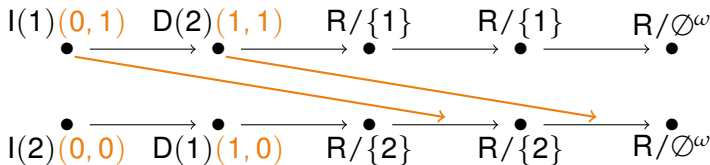
Implementation

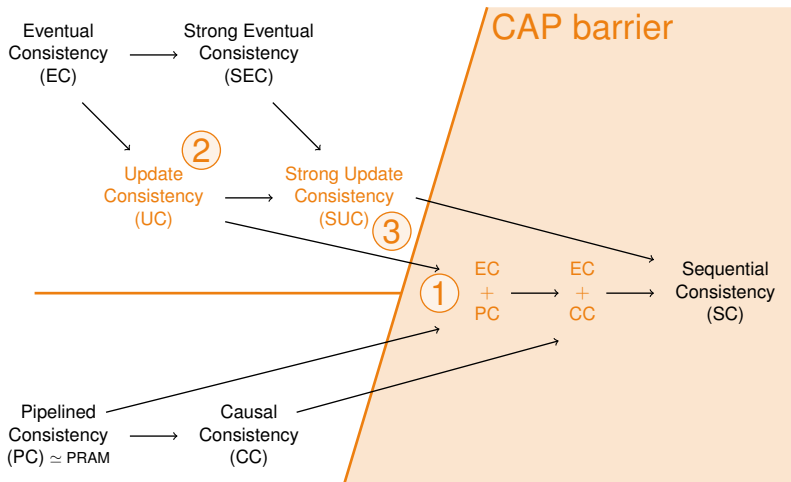## Idea

▶ Choose a total order *a priori*
▶ Update: notify everyone
▶ Store all the updates
▶ Query: execute the updates in the chosen order

Example

$I(1)(0,1)$ $D(2)(1,1)$ $R/\{1\}$ $R/\{1\}$ $R/\emptyset^\omega$

$I(2)(0,0)$ $D(1)(1,0)$ $R/\{2\}$ $R/\{2\}$ $R/\emptyset^\omega$

A map of weak consistency

Weak consistency ⟶ Strong consistency

Future work

Memory-aware implementation

- ▶ So far, only a theoretical proof
- ▶ Unbounded memory/computation

Topology of weak consistency space

- ▶ There is no maximum
- ▶ Are there maximal elements?

Add some causality

- ▶ Causal consistency only defined for memory
- ▶ Exploring the area between SUC and CC+EC