

# Checking Causal Consistency of Distributed Databases<sup>\*</sup>

Rachid Zennou<sup>1,2</sup>, Ranadeep Biswas<sup>1</sup>, Ahmed Bouajjani<sup>1</sup>,  
Constantin Enea<sup>1</sup>, and Mohammed Erradi<sup>2</sup>

<sup>1</sup> Université de Paris, IRIF, CNRS, F-75013 Paris, France  
{ranadeep,abou,cenea}@irif.fr

<sup>2</sup> ENSIAS, Mohammed V University, Rabat, Morocco,  
{rachid.zennou,mohamed.erradi}@gmail.com

**Abstract.** The **CAP Theorem** shows that (strong) Consistency, Availability, and Partition tolerance are impossible to be ensured together. Causal consistency is one of the weak consistency models that can be implemented to ensure availability and partition tolerance in distributed systems. In this work, we propose a tool to check automatically the conformance of distributed/concurrent systems executions to causal consistency models. Our approach consists in reducing the problem of checking if an execution is causally consistent to solving Datalog queries. The reduction is based on complete characterizations of the executions violating causal consistency in terms of the existence of cycles in suitably defined relations between the operations occurring in these executions. We have implemented the reduction in a testing tool for distributed databases, and carried out several experiments on real case studies, showing the efficiency of the suggested approach.

**Keywords:** Causal Consistency · Causal Memory · Causal Convergence · Distributed Databases · Formal verification · Testing

## 1 Introduction

Causal consistency [21] is one of the most implemented models for distributed systems. Contrary to strong consistency [18] (Linearizability [19] and Sequential Consistency [20]), causal consistency can be implemented in the presence of faults while ensuring availability. Several implementations of different variants of causal consistency (such as causal convergence[23] and causal memory [7, 24]) have been developed i.e., [8, 12, 13, 22, 25, 26]. However, the development of such implementations that meet both consistency requirements and availability and performance requirements is an extremely hard and error prone task. Hence,

---

<sup>\*</sup> This is an extended version of a paper published in NETYS’2019 proceeding [32]. This work is supported in part by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 678177). This work is also supported by Centre National pour la Recherche Scientifique et Technique (CNRST) Morocco.

developing efficient approaches to check the correctness of executions w.r.t consistency models such as causal consistency is crucial. This paper presents an approach and a tool for checking automatically the conformance of the computations of a system to causal consistency. More precisely, we address the problem of, given a computation, checking its conformance to causal consistency. We consider this problem for three variants of causal consistency that are used in practice. Solving this problem constitutes the cornerstone for developing dynamic verification and testing algorithms for causal consistency.

Bouajjani et al. [9] studied the complexity of checking causal consistency for a given computation and showed that it is polynomial time. In addition, they formalized the different variations of causal consistency and proposed a reduction of this problem to the occurrence of a finite number of small "bad-patterns" in the computations (i.e., some small sets of events occurring in the computations in some particular order). In this paper, we build on that work in order to define a practical approach and a tool for checking causal consistency, and to apply this tool to real-life case studies. Our approach consists basically in reducing the problem of detecting the existence of bad patterns defined in [9] in computations to the problem of solving a Datalog queries. The fact that solving Datalog queries is polynomial time and that our reduction is polynomial in the size of the computation, allows to solve the conformance checking for causal consistency in polynomial time ( $O(n^3)$ ). We have implemented our approach in an efficient testing tool for distributed systems, and carried out several experiments on real distributed databases, showing the efficiency and performance of this approach.

The rest of this paper is as follows, Section 2 presents preliminaries that include the used notations and the system model. Section 3 is dedicated to defining the causal consistency models. Section 3.2 recalls the characterization of causal consistency violations introduced in [9]. Section 5 presents our reduction of the problem of conformance checking for causal consistency to the problem of solving Datalog queries. Section 6 describes our testing tool, the case studies we have considered, and the experimental results we obtained. Section 7 presents related work, and finally conclusions are drawn in Section 8.

## 2 Preliminaries

**Notations.** Given a set  $O$  and a relation  $\mathcal{R} \subseteq O \times O$ , we use the notation  $(o_1, o_2) \in \mathcal{R}$  to denote the fact that  $o_1$  and  $o_2$  are related by  $\mathcal{R}$ . If  $\mathcal{R}$  is an order, it denotes the fact that  $o_1$  precedes  $o_2$  in this order. The transitive closure of  $\mathcal{R}$  is denoted by  $\mathcal{R}^+$ , which is the composition of one or multiple copies of  $\mathcal{R}$ .

Let  $O'$  be a subset of  $O$ . Then  $\mathcal{R}|_{O'}$  is the relation  $\mathcal{R}$  projected on the set  $O'$ , that is  $\{(o_1, o_2) \in \mathcal{R} \mid o_1, o_2 \in O'\}$ . The set  $O' \subseteq O$  is said to be *downward-closed* w.r.t a relation  $\mathcal{R}$  if  $\forall o_1, o_2$ , if  $o_2 \in O'$  and  $(o_1, o_2) \in \mathcal{R}$  then  $o_1 \in O'$  as well. A relation  $\mathcal{R} \subseteq O \times O$  is a *strict partial order* if it is transitive and irreflexive. Given a strict partial order  $\mathcal{R}$  over  $O$ , a *poset* is a pair  $(O, \mathcal{R})$ . Notice here that we consider the strict version of posets (not the ones where the underlying partial order is *weak*, i.e. reflexive, transitive and antisymmetric. Given a set  $\Sigma$ ,

a poset  $(O, \mathcal{R})$ , and a labeling function  $\ell : O \rightarrow \Sigma$ , the  $\Sigma$  *labeled poset*  $\rho$  is a tuple  $(O, \mathcal{R}, \ell)$ .

We say that  $\rho'$  is a *prefix* of  $\rho$  if there exists a downward closed set  $A \subseteq O$  w.r.t. relation  $\mathcal{R}$  such that  $\rho' = (A, \mathcal{R}, \ell)$ . If the relation  $\mathcal{R}$  is a strict total order, we say that a (resp., labeled) *sequential poset* (sequence for short) is a (resp., labeled) poset. The concatenation of two sequential posets  $e$  and  $e'$  is denote by  $e.e'$ .

Consider a set of methods  $\mathbb{M}$  from a domain  $\mathbb{D}$ . For  $m \in \mathbb{M}$  and  $arg, rv \in \mathbb{D}$ , and  $o \in O$ ,  $\ell(o) = (m, arg, rv)$  means that operation  $o$  is an invocation of  $m$  with input  $arg$  which returns  $rv$ . The label  $\ell(o)$  is sometimes denoted  $m(arg, rv)$ . Let  $\rho = (O, \mathcal{R}, \ell)$  be a  $\mathbb{M} \times \mathbb{D} \times \mathbb{D}$  labeled poset and  $O' \subseteq O$ . We denote by  $\rho\{O'\}$  the labeled poset where we only keep the return values of the operations in  $O'$ . Formally,  $\rho\{O'\}$  is the  $(\mathbb{M} \times \mathbb{D}) \cup (\mathbb{M} \times \mathbb{D} \times \mathbb{D})$  labeled poset  $(O, \mathcal{R}, \ell')$  where for all  $o \in O'$ ,  $\ell'(o) = \ell(o)$ , and for all  $o \in O \setminus O'$ , if  $\ell(o) = (m, arg, rv)$ , then  $\ell'(o) = (m, arg)$ . Now, we introduce a relation on labeled posets, denoted  $\preceq$ . Let  $\rho = (O, \mathcal{R}, \ell)$  and  $\rho' = (O, \mathcal{R}', \ell')$  be two posets labeled by  $(\mathbb{M} \times \mathbb{D}) \cup (\mathbb{M} \times \mathbb{D} \times \mathbb{D})$  (the return values of some operations in  $O$  might not be specified). The notation  $\rho' \preceq \rho$  means that  $\rho'$  has less order and label constraints on the set  $O$ . Formally,  $\rho' \preceq \rho$  if  $\mathcal{R}' \subseteq \mathcal{R}$  and for all operation  $o \in O$ , and for all  $m \in \mathbb{M}$ ,  $arg, rv \in \mathbb{D}$ ,  $\ell(o) = \ell'(o)$ , or  $\ell(o) = (m, arg, rv)$  implies  $\ell'(o) = (m, arg)$ .

**System model.** We consider a distributed system model in which a system is composed of several processes (sites) connected over a network. Each process performs operations on objects (variables)  $\text{Var} = \{x, y, \dots\}$ . These objects are called replicated objects and their state is replicated at all processes. Clients interact with the system by performing operations. Assuming an unspecified set of values  $\text{Val}$  and a set of operation identifiers  $\text{IdO}$ . We define the set of operations as  $\text{Op} = \{\text{read}_i(x, v), \text{write}_i(x, v) : i \in \text{IdO}, x \in \text{Var}, v \in \text{Val}\}$ . Where  $\text{read}_i(x, v)$  is a read operation reading a value  $v$  from a variable  $x$  and  $\text{write}_i(x, v)$  is a write operation writing a value  $v$  on a variable  $x$ . The set of read, resp., write, operations in a set of operations  $O$  is  $\mathbb{R}(O)$ , resp.,  $\mathbb{W}(O)$ . The variable accessed by an operation  $o$  is denoted by  $\text{var}(o)$ .

**Histories.** We consider an abstract notion of an execution called *history* which includes write and read operations. The operations performed by the same process are ordered by a *program order*  $\text{po}$ . We assume that histories include a *write-read* relation that matches each read operation to the write operation written its return value.

Formally, a *history*  $\langle O, \text{po}, \text{wr} \rangle$  is a set of read or write operations  $O$  along with a partial *program order*  $\text{po}$  and a *write-read* relation  $\text{wr} \subseteq \mathbb{W}(O) \times \mathbb{R}(O)$ , such that if  $(\text{write}(x, v), \text{read}(x', v')) \in \text{wr}$ , then  $x = x'$  and  $v = v'$ . For  $o_1, o_2 \in O$ ,  $(o_1, o_2) \in \text{po}$  means that  $o_1, o_2$  were issued by the same process and  $o_1$  was submitted before  $o_2$ . We mention that the *write-read* relation can only be defined for differentiated histories.

**Differentiated histories.** A history  $\langle O, \text{po}, \text{wr} \rangle$  is differentiated if each value is written at most once, i.e., for all write operations  $\text{write}(x, v)$  and  $\text{write}(x, v')$ ,  $v \neq v'$ .

**Data Independence.** An implementation is data-independent if its behavior does not depend on the handled values. We consider in this paper implementations that are data-independent which is a natural assumption that corresponds to a wide range of existing implementations. Under this assumption, it is good enough to consider differentiated histories [9]. Thus, all histories in this paper are differentiated.

In addition, we assume that each history contains a write operation writing the initial value (the value 0) of variable  $x$ , for each variable  $x$ . These write operations precede all other operations in  $\text{po}$ .

**Specification.** The consistency of a replicated object is defined w.r.t. some specification, determining the correct behaviors of that object in a sequential setting. In this work, we consider the read/write memory for which the specification  $S_{RW}$  is inductively defined as the smallest set of sequences closed under the following rules ( $x \in \text{Var}$  and  $v \in \text{Val}$ ):

1.  $\varepsilon \in S_{RW}$ ,
2. if  $\rho \in S_{RW}$ , then  $\rho.\text{write}(x, v) \in S_{RW}$ ,
3. if  $\rho \in S_{RW}$  contains no write on  $x$ , then  $\rho.\text{read}(x, 0) \in S_{RW}$ ,
4. if  $\rho \in S_{RW}$  and the last write in  $\rho$  on variable  $x$  is  $\text{write}(x, v)$ , then  $\rho.\text{read}(x, v) \in S_{RW}$ .

### 3 Causal Consistency

#### 3.1 Causal Consistency definitions

Causal consistency [21] is one of the most used models for replicated objects. It guarantees that, if two operations  $o_1$  and  $o_2$  are *causally related* (some process is aware of  $o_1$  when executing  $o_2$ ), then  $o_1$  should be executed before  $o_2$  in all processes. Operations that are not causally related may be seen in different orders by different processes. We present in the following three variations of causal consistency, weak causal consistency, causal convergence and causal memory. We use the same definitions as in [9].

**Weak causal consistency** The weakest variation of causal consistency is called *weak causal consistency* (CC, for short). A history is CC if there *exists* a causal order that explains the return value of all operations in the history [9]. Formally,

**Definition 1.** *A history  $h$  satisfies CC w.r.t a specification  $S$  if there exists a strict partial order, called causal order,  $\text{co} \subseteq O \times O$ , such that, for all operations  $o \in O$  in  $h$ , there exists a specification sequence  $\rho_o \in S$  such that axioms  $\text{AxCausal}$  and  $\text{AxCausalValue}$  hold (see 1).*

Axiom  $\text{AxCausal}$  states that the causal order should at least include the program order. Axiom  $\text{AxCausalValue}$  states that, for each operation  $o \in O$ , a valid sequence of the specification  $S$  can be obtained by sequentializing the causal history of  $o$  i.e., all operations that precede  $o$  in the causal order. In addition,

AxCausal	$po \subseteq co$
AxArb	$co \subseteq arb$
AxCausalValue	$CausalHist(o)\{o\} \preceq \rho_o$
AxCausalSeq	$CausalHist(o)\{POPast(o)\} \preceq \rho_o$
AxCausalArb	$CausalArb(o)\{o\} \preceq \rho_o$

where:

$$CausalHist(o) = (CausalPast(o), co, \ell)$$

$$CausalArb(o) = (CausalPast(o), arb, \ell)$$

$$CausalPast(o) = \{o' \in O \mid (o', o) \in co^*\}$$

$$POPast(o) = \{o' \in O \mid (o', o) \in po^*\}$$

Table 1: Axioms used in the causal consistency definitions.

this sequentialization must also preserve the constraints provided by the causal order.

Formally, the *causal past* of  $o$ ,  $CausalPast(o)$ , is the *set* of operations that precede  $o$  in the causal order. The *causal history* of  $o$ ,  $CausalHist(o)$ , is the restriction of the causal order to the operations in its causal past  $CausalPast(o)$ . The notation  $CausalHist(o)\{o\}$  means that only the return value of operation  $o$  is kept. The axiom  $AxCausalValue$  uses  $CausalHist(o)\{o\}$  because a process is not required to be consistent with the values it has returned in the past or the values returned by the other processes.

The notations  $CausalHist(o)\{o\} \preceq \rho_o$  means that  $CausalHist(o)\{o\}$  can be sequentialized to a sequence  $\rho_o$  in the specification. We will formally define these last two notations in the next sections.

For a better understanding of this model, consider the following examples.

*Example 1.* The history 1d is **CC**, we can consider that  $write(x, 1)$  is not causally-related to  $write(x, 2)$ . Therefore,  $p_2$  can execute them in any order.

*Example 2.* The history 1e is not **CC**. The reason is that, a causal order that explains the return values of all operations in the history cannot be found. Intuitively, since  $read(y, 1)$  reads the value from  $write(y, 1)$ , in any causal order,  $write(y, 1)$  should precede  $read(y, 1)$ . By transitivity of the causal order and because any causal order should include the program order,  $write(x, 1)$  precedes  $write(x, 2)$  in the causal order ( $write(x, 1)$  and  $write(x, 2)$  are causally related). However, process  $p_3$  inverse this order. This is a contradiction with the informal definition of **CC** which requires that every process should see causally related operations in the same order.

**Causal convergence** *Causal convergence* (**CCv**, for short) is stronger than **CC**. It ensures that, as long as no new updates are submitted, all processes eventually converge towards the same state. In addition of seeing causally related operations in the same order (**CC**), causal convergence uses a total order over all the operations in a history to agree on how to order operations which are *not*

$p_1:$ $\text{write}(z, 1)$ $\text{write}(x, 1)$ $\text{write}(y, 1)$	$p_2:$ $\text{write}(x, 2)$ $\text{read}(z, 0)$ $\text{read}(y, 1)$ $\text{read}(x, 2)$	$p_1:$ $\text{write}(x, 1)$ $\text{read}(x, 2)$	$p_2:$ $\text{write}(x, 2)$ $\text{read}(x, 1)$
(a) CCv but not CM			
$p_1:$ $\text{write}(x, 1)$ $\text{read}(y, 0)$ $\text{write}(y, 1)$ $\text{read}(x, 1)$	$p_2:$ $\text{write}(x, 2)$ $\text{read}(y, 0)$ $\text{write}(y, 2)$ $\text{read}(x, 2)$	$p_1:$ $\text{write}(x, 1)$	$p_2:$ $\text{write}(x, 2)$ $\text{read}(x, 1)$ $\text{read}(x, 2)$
(b) CM but not CCv			
(c) CC , CCv and CM			
$p_1:$ $\text{write}(x, 1)$ $\text{write}(y, 1)$	$p_2:$ $\text{read}(y, 1)$ $\text{write}(x, 2)$	$p_3:$ $\text{read}(x, 2)$ $\text{read}(x, 1)$	
(d) CC but not CCv nor CM			
(e) not CC (nor CCv, nor CM)			

Fig. 1: Histories illustrating the differences between the causal consistency models CC, CCv, and CM.

causally related[9]. This order is called the *arbitration order* and denoted by  $arb$ . Similarly to the causal order, the arbitration order is existentially quantified in the CCv definition. Formally,

**Definition 2.** A history is *CCv* w.r.t a specification  $S$  if there exist a strict partial order  $co \subseteq O \times O$  and a strict total order  $arb \subseteq O \times O$  such that, for each operation  $o \in O$  in  $h$ , there exists a specification sequence  $\rho_o \in S$  such that the axioms  $AxCausal$ ,  $AxArb$ , and  $AxCausalArb$  hold.

Axiom  $AxArb$  states that the arbitration order  $arb$  should at least include the causal order  $co$ . Axiom  $AxCausalArb$  states that, sequentializing the operations that are in the causal past of  $o$  to explain the return value of an operation  $o$ , should respect the arbitration order  $arb$ .

We now present two examples, one which satisfies CCv and another one which violates it.

*Example 3.* The history 1a is CCv, we can set an arbitration order in which  $\text{write}(x, 1)$  is ordered before  $\text{write}(x, 2)$ .

*Example 4.* The history 1b is not CCv. In order to read  $\text{read}(x, 2)$ ,  $\text{write}(x, 1)$  must be ordered before  $\text{write}(x, 2)$  in the arbitration order. On the other hand, to read  $\text{read}(x, 1)$ ,  $\text{write}(x, 2)$  must be ordered before  $\text{write}(x, 1)$  in the arbitration order, that is not possible.

**Causal memory** The third model we consider is *causal memory* (CM, for short) that is also stronger than CC. It guarantees that each process should observe concurrent operations in the same order. In addition, this order should be maintained throughout its whole execution, but it can differ from one process to another [9]. Formally,

**Definition 3.** *A history  $h$  is CM w.r.t. a specification  $S$  if there exists a strict partial order  $\text{co} \subseteq O \times O$  such that, for each operation  $o \in O$  in  $h$ , there exists a specification sequence  $\rho_o \in S$  such that axioms  $\text{AxCausal}$  and  $\text{AxCausalSeq}$  hold.*

Compared to CC, CM requires that each process should be consistent with the return values it has returned in the past. However, a process is not required to be consistent with respect to the return values provided by other processes. Therefore,  $\text{AxCausalSeq}$  states:

$$\text{CausalHist}(o)\{\text{POPast}(o)\} \preceq \rho_o$$

where  $\text{CausalHist}(o)\{\text{POPast}(o)\}$  is the causal history where we only keep the return values of the operations that precede  $o$  in the program order (in  $\text{POPast}(o)$ ).

As we noticed above, CC is weaker than CCv and CM. For instance, the history in Figure 1d is CC but not CCv nor CM. It is CC, we can consider that  $\text{write}(x, 1)$  is not causally-related to  $\text{write}(x, 2)$ . On the other hand, for reading the value 1 the process  $p_2$  decides to order  $\text{write}(x, 2)$  before  $\text{write}(x, 1)$ , then it changes this order to read the value 2. This is not allowed under CM nor under CCv. Both CCv and CM require that each process should observe concurrent operations in the same order. In CM this order can differ from one process to another. It seems that this intuitive description implies that CCv is stronger than CM but these two models are actually incomparable. The following examples illustrate the differences between these models.

*Example 5.* For instance, the history in Figure 1b is CM, but not CCv. It is not CCv because reading the value 1 from  $x$  in the  $p_1$  implies that  $\text{write}(x, 1)$  is ordered after  $\text{write}(x, 2)$  while reading the value 2 from  $x$  in  $p_2$  implies that it  $\text{write}(x, 2)$  is ordered after  $\text{write}(x, 1)$ . This is allowed by CM as different processes can observe concurrent write operations in different orders.

*Example 6.* The history in Figure 1a is CCv but not CM. CCv requires that concurrent operations should be observed in the same order by all processes. Thus, a possible order for concurrent write operations  $\text{write}(x, 1)$  and  $\text{write}(x, 2)$  is to order  $\text{write}(x, 2)$  after  $\text{write}(x, 1)$ . Under CM, in order to read  $\text{read}(z, 0)$ ,  $\text{write}(x, 1)$  should be ordered after  $\text{write}(x, 2)$  while to read 2 from  $x$ ,  $\text{write}(x, 2)$  must be ordered after  $\text{write}(x, 1)$  ( $\text{write}(x, 1)$  must have been observed because  $p_2$  reads 1 from  $y$  and the writes on  $x$  and  $y$  are causally-related).

The Figure 2 summarizes the relationships between the causal consistency models presented in this section.

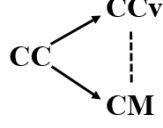


Fig. 2: Relationships between causal consistency models. Directed arrows denote the “weaker-than” relation while dashed lines connect incomparable models.

### 3.2 Causal consistency violations

Now, we will see, for each definition of causal consistency, how to characterize histories that are not conform to causal consistency through the presence of some specific sets of operations. In [9], computations that are violations of CC, CCv or CM are characterized by the occurrence of a finite number of particular (small) sets of ordered events, called bad-patterns. In this section, we recall the bad-patterns corresponding to each model and their definitions (Table 2 and 3).

**CC Bad-patterns.** We now give the CC bad-patterns as defined in [9]. These bad-patterns are defined using the relation of causality  $\text{co}$  which is given by the *program order*  $\text{po}$  or the *write-read* relation  $\text{wr}$  or any transitive composition of these relations i.e.,  $\text{co} = (\text{po} \cup \text{wr})^+$ .

**Lemma 1.** ([9]) *A history is CC if it does not contain any of the bad-patterns  $\text{CyclicCO}$ ,  $\text{WriteCOInitRead}$ ,  $\text{ThinAirRead}$  and  $\text{WriteCORead}$ .*

*Example 7.* The history in Figure 1e contains the bad-pattern  $\text{WriteCORead}$ . The  $\text{write}(x, 1)$  is causally ordered before  $\text{write}(x, 2)$  by the transitivity. On the other hand, the process  $p_3$ ,  $\text{read}(x, 1)$  from  $\text{write}(x, 1)$  ( $\text{wr}(\text{write}(x, 1), \text{read}(x, 1))$ ). The read  $\text{read}(x, 1)$  is also causally-related to  $\text{write}(x, 2)$  by transitivity. The history in Figure 1c does not contain any of the bad-patterns, so it is CC, CCv and CM.

**CCv bad-patterns.** As we have seen before, CCv is stronger than CC. Therefore, CCv excludes all the CC bad-patterns we have seen above ( $\text{CyclicCO}$ ,  $\text{WriteCOInitRead}$ ,  $\text{ThinAirRead}$  and  $\text{WriteCORead}$ ). In addition, CCv excludes another bad pattern called  $\text{CyclicCF}$ , defined in terms of a conflict relation  $\text{cf}$ . Intuitively, two writes  $w_1$  and  $w_2$  on the same variable are in conflict, if  $w_1$  is causally-related to a read taking its value from  $w_2$ . Formally,  $\text{cf}$  is defined as

$$(\text{write}(x, v), \text{write}(x, v')) \in \text{cf} \text{ iff } (\text{write}(x, v), \text{read}(x, v')) \in \text{co} \text{ and } (\text{write}(x, v'), \text{read}(x, v')) \in \text{wr}, \text{ for some } \text{read}(x, v')$$

Then,

**Lemma 2.** ([9]) *A history is CCv if it is CC and does not contain the bad-pattern  $\text{CyclicCF}$ .*

*Example 8.* The History in Figure 1b is not CCv as it contains the bad-pattern  $\text{CyclicCF}$ . In order to read  $\text{read}(x, 2)$ ,  $\text{write}(x, 2)$  must precedes  $\text{write}(x, 2)$  in the conflict order. On the other hand, to read  $\text{read}(x, 1)$ ,  $\text{write}(x, 2)$  must be ordered before  $\text{write}(x, 1)$  in the conflict order. Thus, we get a cycle in  $\text{cf}$ .



CC	CCv	CM
CyclicCO	CyclicCO	CyclicCO
WriteCOInitRead	WriteCOInitRead	WriteCOInitRead
ThinAirRead	ThinAirRead	ThinAirRead
WriteCOWRead	WriteCOWRead	WriteCOWRead
	CyclicCF	WriteHBInitRead
		CyclicHB

Table 2: Bad-patterns for each causal consistency model

CyclicCO	the causality relation $\text{co}$ is cyclic
WriteCOInitRead	a $\text{read}(x, 0)$ is causally preceded by a $\text{write}(x, v)$ (i.e., $(\text{write}(x, v), \text{read}(x, 0)) \in \text{co}$ ) such that $v \neq 0$
ThinAirRead	there is a $\text{read}(x, v)$ operation that reads a value $v$ such that $v \neq 0$ that it is never written before i.e., it can not be related to any write by a $\text{wr}$ relation.
WriteCOWRead	there exist write operations $w_1, w_2$ such that $\text{var}(w_1) = \text{var}(w_2)$ and a read operation $r_1$ such that $\text{wr}(w_1, r_1)$ . In addition, $(w_1, w_2) \in \text{co}$ and $(w_2, r_1) \in \text{co}$ .
WriteHBInitRead	there exist a $\text{read}(x, 0)$ and a $\text{write}(x, v)$ ( $v \neq 0$ ) such that $(\text{write}(x, v), \text{read}(x, 0)) \in \text{hb}_o$ for some operation $o$ , with $(r, o) \in \text{po}^*$ .
CyclicHB	the $\text{hb}_o$ relation is cyclic for some operation $o$
CyclicCF	the union of $\text{cf}$ and $\text{co}$ ( $\text{cf} \cup \text{co}$ ) is cyclic

Table 3: Bad-patterns definitions

**CM bad-Patterns.** As we have seen above, CM is also stronger than CC. Therefore, CM excludes all the CC bad-patterns (CyclicCO, WriteCOInitRead, ThinAirRead and WriteCOWRead). In addition, CM excludes two additional bad-patterns (WriteHBInitRead and CyclicHB), defined using a happened-before relation per operation  $\text{hb}_o$ . Formally,  $\text{hb}_o$  is defined as follows.

**Definition 4.** Let  $h = \langle O, \text{po}, \text{wr} \rangle$  be a history. For every operation  $o$  in  $h$ , let  $\text{hb}_o$  be the smallest transitive relation such that:

1.  $\text{co}_{|\text{CausalPast}(o)} \subseteq \text{hb}_o$ , which means that if two operations are causally related and each one is causally related to  $o$ , then they are related by  $\text{hb}_o$  i.e.,  $(o_1, o_2) \in \text{hb}_o$  if  $(o_1, o_2) \in \text{co}$ ,  $(o_1, o) \in \text{co}$  and  $(o_2, o) \in \text{co}^*$  (where  $\text{co}^*$  is the reflexive closure of  $\text{co}$ ), and
2. two writes  $w_1$  and  $w_2$  are related by  $\text{hb}_o$  if  $w_1$  is  $\text{hb}_o$ -related to a read taking its value from  $w_2$  and that read is done by the same thread executing  $o$  and before  $o$ , i.e.,  $(\text{write}(x, v), \text{write}(x, v')) \in \text{hb}_o$  if  $(\text{write}(x, v), \text{read}(x, v')) \in \text{hb}_o$ ,  $(\text{write}(x, v'), \text{read}(x, v')) \in \text{wr}$  and  $(\text{read}(x, v'), o) \in \text{po}^*$  for some  $\text{read}(x, v')$ .

Then,

**Lemma 3.** ([9]) *A history is CM if it is CC and does not contain any of the bad-patterns WriteHBInitRead and CyclicHB.*

*Example 9.* 1. The history 1a contains the bad-pattern WriteHBInitRead so it is not CM. Let's consider  $hb = hb_{\text{read}(x,2)}$ . We have  $(\text{write}(z,1), \text{write}(x,1)) \in \text{po}$  and  $(\text{write}(x,1), \text{write}(x,2)) \in hb$  ( $(\text{write}(x,1), \text{read}(x,2)) \in \text{co}$  implies  $(\text{write}(x,1), \text{write}(x,2)) \in \text{co}$ ) and  $(\text{write}(x,2), \text{read}(z,0)) \in \text{po}$ , thus by transitivity we have  $(\text{write}(z,1), \text{read}(z,0)) \in hb$ .

2. The history in Figure 1c does not contain any of the bad-patterns, so it is CC, CCv and CM.

## 4 An improved characterization of CM

The proposed characterization of CM in [9] requires computing the  $hb_o$  relation for all operations and then check for CM bad-patterns. Let's call this approach CM\_1. Now, we will show that it is enough to check the CM-bad patterns for only a small set of operations (for only last operation in each process). We propose in the following a succinct but equivalent approach for checking CM. We show that it improves the verification runtime (Section 6). Let's call this new approach CM\_2.

To prove the equivalence between the two approaches, we have to prove some intermediate results. First, we define  $hb_o^i$  to denote a controlled saturated version of  $hb_o$ .

**Definition 5.** *Let  $h = \langle O, \text{po}, \text{wr} \rangle$  be a history. For every operation  $o$  in  $h$ ,*

1. *let  $hb_o^0$  be the relation such that if two operations are causally related and each one is causally related to  $o$ , then they are related by  $hb_o^0$  i.e.,  $(o_1, o_2) \in hb_o^0$  if and only if  $(o_1, o_2) \in \text{co}$ ,  $(o_1, o) \in \text{co}$  and  $(o_2, o) \in \text{co}^*$  (where  $\text{co}^*$  is the reflexive closure of  $\text{co}$ ),*
2. *let  $hb_o^i$  for  $i > 0$  be the transitive relation if two writes  $w_1$  and  $w_2$  are related by  $hb_o^i$  if  $w_1$  is  $(\cup_{j < i} hb_o^j)^+$  (transitive closure of all the previous  $hb_o^j$ ) related to a read taking its value from  $w_2$  and that read is done by the same thread executing  $o$  and before  $o$ , i.e.,  $(\text{write}(x, v), \text{write}(x, v')) \in hb_o^i$  if and only if  $(\text{write}(x, v), \text{read}(x, v')) \in (\cup_{j < i} hb_o^j)^+$ ,  $(\text{write}(x, v'), \text{read}(x, v')) \in \text{wr}$  and  $(\text{read}(x, v'), o) \in \text{po}^*$  for some  $\text{read}(x, v')$ .*

**Theorem 1.** *For all  $o$ ,  $hb_o = (\cup_{i \geq 0} hb_o^i)^+$*

*Proof.* By construction,  $(\cup_{i \geq 0} hb_o^i)^+$  satisfies definition 4. Because  $hb_o$  is the smallest one,  $hb_o \subseteq (\cup_{i \geq 0} hb_o^i)^+$ . Also, by construction, all the relations in  $(\cup_{i \geq 0} hb_o^i)^+$  must be present in  $hb_o$  because they are constructed statically from  $\text{co}$  and  $\text{wr}$ . So  $hb_o \supseteq (\cup_{i \geq 0} hb_o^i)^+$ .

Now, we prove that  $hb_o$  is included in  $hb_{o'}$  if  $o$  is executed before  $o'$  in a same thread. Then, checking  $hb_o$  acyclicity of  $\text{po}$ -maximal operations is enough to decide for all operations. To prove this, we use the  $hb_o^i$  definition.

**Lemma 4.** *If  $(o, o') \in \text{po}$  then  $\text{hb}_o^i \subseteq \text{hb}_{o'}^i$  for  $i \geq 0$*

*Proof.* We will prove by induction on  $i$ .

- Base case.  $i = 0$ . Since  $(o, o') \in \text{po} \subseteq \text{co}$ ,  $(o_1, o) \in \text{co}$  and  $(o_2, o) \in \text{co}^*$  implies,  $(o_1, o') \in \text{co}$  and  $(o_2, o') \in \text{co}^*$ . So all the  $\text{hb}_o^0 \subseteq \text{hb}_{o'}^0$ .
- Inductive step. If there exists two writes  $\text{write}(x, v), \text{write}(x, v')$  and a read  $\text{read}(x, v')$  with  $(\text{write}(x, v), \text{read}(x, v')) \in (\cup_{j < i} \text{hb}_o^j)^+$ ,  $(\text{write}(x, v'), \text{read}(x, v')) \in \text{wr}$  and  $(\text{read}(x, v'), o) \in \text{po}^*$  to force  $(\text{write}(x, v), \text{write}(x, v')) \in \text{hb}_o^i$  relation, then it is also true that,  $(\text{write}(x, v), \text{read}(x, v')) \in (\cup_{j < i} \text{hb}_{o'}^j)^+$  (induction hypothesis) and  $(\text{read}(x, v'), o') \in \text{po}^*$ . So  $(\text{write}(x, v), \text{write}(x, v')) \in \text{hb}_{o'}^i$  is forced as well. Thus,  $\text{hb}_o^i \subseteq \text{hb}_{o'}^i$ .

**Corollary 1.** *If  $(o, o') \in \text{po}$  then  $\text{hb}_o \subseteq \text{hb}_{o'}$ .*

*Proof.* Direct consequence of theorem 1 and lemma 4.

Finally, we can prove the equivalence between the two CM verification approaches. Both of the approaches requires the history to be CC. So, we just need to do it for the acyclicity of  $\text{hb}_o$  and for the WriteHBInitRead bad-pattern.

CM\_1 requires  $\text{hb}_o$  for all  $o$  to be acyclic, whereas CM\_2 requires  $\text{hb}_o$  for a subset of operations  $o$  (po-maximal operations) to be acyclic. So trivially CM\_1 implies CM\_2.

For the other direction, we use corollary 1. If  $(o, o') \in \text{po}$  then  $\text{hb}_o \subseteq \text{hb}_{o'}$ . Hence, a cycle in  $\text{hb}_o$  for some  $o$  (if  $o$  is po-maximal operation, then we are done) will be also present in  $\text{hb}_{o'}$  for the po-maximal  $o'$  after  $o$  because  $(o, o') \in \text{po}$ .

Now, we will prove that it is enough to check the WriteHBInitRead bad-pattern for only po-maximal operations as well.

Consider two operations  $o_1$  and  $o_2$  in a history  $h$  such that  $(o_1, o_2) \in \text{po}$ . Suppose that there exists a bad-pattern WriteHBInitRead for  $o_1$  i.e., there exist a  $\text{read}(x, 0)$  and a  $\text{write}(x, v)$  ( $v \neq 0$ ) such that  $(\text{write}(x, v), \text{read}(x, 0)) \in \text{hb}_{o_1}$ , with  $(r, o_1) \in \text{po}^*$ . Since  $(o_1, o_2) \in \text{po}$  and using the corollary 1, we have  $\text{hb}_{o_1} \subseteq \text{hb}_{o_2}$ , then  $(\text{write}(x, v), \text{read}(x, 0)) \in \text{hb}_{o_2}$  and  $(r, o_2) \in \text{po}^*$  (because of  $(r, o_1) \in \text{po}^*$ ). Therefore, the bad-pattern WriteHBInitRead exists for  $o_2$  as well.

Then,

**Theorem 2.** *CM\_1 and CM\_2 are equivalent.*

Notice that CM\_2 is characterized by the same CM bad-patterns described above, it is just that the  $\text{hb}_o$  is only computed for each po-maximal operation  $o$  in the history not for all operations (see next section).

## 5 Reduction to Datalog queries solving

In this section, we show our reduction of the problem of checking whether a given computation is a CC, CCv or CM violation to the problem of Datalog queries

solving. Datalog is a logic programming language that does not allow functions as predicate arguments. The advantage of using Datalog is that it provides a high level language for naturally defining constraints on relations and that solving Datalog queries is polynomial time [28].

### 5.1 Datalog

A rule in datalog is a statement of the following form:

$$r_1(v_1) \text{ :- } r_2(v_2), \dots, r_i(v_i)$$

Where  $i \geq 1$ ,  $r_i$  are the names of predicates (relations) and  $v_i$  are arguments. A Datalog program is a finite set of Datalog rules over the same schema [6]. The LHS is called the rule head and represents the outcome of the query, while the RHS is called the rule body.

*Example 10.* For instance, this Datalog program computes the transitivity closure of a given graph.

```
trans(X,Y) :- edge(X,Y).
trans(X,Y) :- trans(X,Z), trans(Z,Y).
```

Where the fact  $\text{edge}(a,b)$  means that there exists a direct edge from  $a$  to  $b$ .

In the literature, there are three definitions for the semantics of Datalog programs, *model theoretic*, *proof-theoretic* and *fixpoint semantics* [6]. In this paper, we have considered the *fix-point semantics*.

**Fix-point semantics.** This approach is based on the fix-point theory. Given a function  $f()$ , its fix-point is an element  $e$  from its domain which is mapped by the function  $f$  to itself i.e.,  $f(e) = e$ . Each Datalog program has an associated operator called immediate consequence operator. Applying repeatedly this operator on existing facts generates new facts until getting a fixed point.

### 5.2 Histories Encoding

In our approach, extracted relations from a history ( $\text{po}$ ,  $\text{wr} \dots$ ) are represented as predicates called facts, while the algorithm for fixed point computation is formulated as Datalog recursive relations called inference rules.

We first introduce all the facts. For instance, consider the fact  $\text{po}(a,b)$  which represents the program order from the operation  $a$  to the operation  $b$  (similarly  $\text{po}(b,c)$ ),

```
po(a,b).
po(b,c).
```

Now, we define the necessary relations (axioms) for our approach.

- $rd(X)$ ,  $X$  is a read operation

- $wrt(X)$ ,  $X$  is a write operation
- $po(X, Y)$ ,  $X$  precedes  $Y$  in the **po** order.
- $wr(X, Y)$ ,  $Y$  reads the value from a write operation  $X$  (**wr** relation)
- $sv(X, Y)$ , the operations  $X$  and  $Y$  accessed to the same variable.

Then, we define the inference rules used to generate derived relations. For instance, the following rule states that the causal relation **co** is transitive.

```
co(X,Z) :- co(X,Y), co(Y,Z).
```

### 5.3 Bad-patterns Encoding

We have expressed all the bad-patterns as Datalog inference rules, except **ThinAirRead** that we verify externally as it contains a universal quantification over all operations. There exist two kinds of bad-patterns. The first type is related to the existence of a cycle in a relation. For instance, the bad-pattern **CyclicCO** that can be expressed as

```
:- co(X,Y), co(Y,X).
```

Intuitively, this means that there exist no operations  $X$  and  $Y$  such that  $X$  precedes  $Y$  in the causal order and  $Y$  also precedes  $X$  in the causal order. Since **co** is transitive, we can simply write it as

```
:- co(X,X).
```

The second type is related to the occurrence of some operations in some particular order. For instance, **WriteC0Read** can be expressed as follows

```
:- co(X,Y), co(Y,Z), wr(X,Z), wrt(X), wrt(Y), rd(Z), sv(X,Y), sv(Y,Z).
```

Intuitively, this means that there exist no write operations  $X$  and  $Y$  on the same variable and a read operation  $Z$  which takes the value from  $X$  such that  $X$  precedes  $Y$  in the causal order and  $Y$  precedes  $Z$  in the causal order.

**CC bad-patterns encoding.** In addition of the **CyclicCO** bad-pattern we have seen above, we will see how the other CC bad-patterns are encoded. Consider the following example which presents the Datalog program corresponding to an execution.

*Example 11.* This example represents the history 1b Datalog program for checking CC:

```
% Facts
wrt("w(x,1,id0)").
po("w(x,1,id0)","r(x,2,id1)").
sv("r(x,2,id1)","w(x,1,id0)").
sv("w(x,2,id2)","w(x,1,id0)").
```

```

sv("r(x,1,id3)","w(x,1,id0)").
rd("r(x,2,id1)").
sv("w(x,1,id0)","r(x,2,id1)").
wr("w(x,2,id2)","r(x,2,id1)").
sv("w(x,2,id2)","r(x,2,id1)").
sv("r(x,1,id3)","r(x,2,id1)").
wrt("w(x,2,id2)").
sv("w(x,1,id0)","w(x,2,id2)").
sv("r(x,2,id1)","w(x,2,id2)").
po("w(x,2,id2)","r(x,1,id3)").
sv("r(x,1,id3)","w(x,2,id2)").
rd("r(x,1,id3)").
wr("w(x,1,id0)","r(x,1,id3)").
sv("w(x,1,id0)","r(x,1,id3)").
sv("r(x,2,id1)","r(x,1,id3)").
sv("w(x,2,id2)","r(x,1,id3)").
initread("r(a,0,ida)").
% Inference rules
co(X,Y) :- po(X,Y).
co(X,Y) :- wr(X,Y).
co(X,Z) :- co(X,Y), co(Y,Z). % Transitivity
% CC bad-patterns
:- co(X,X). % CyclicCO
:- co(X,Y), wrt(X), initread(Y), sv(X,Y). % WriteCOInitRead
:- co(X,Y), co(Y,Z), wr(X,Z), wrt(X), wrt(Y), rd(Z), sv(X,Y), sv(Y,Z).
% WriteCORead

```

Notice that, since the bad pattern WriteCOInitRead includes a predicate `initread(Y)`, we add the `initread("r(a,0,ida)")` to the programs that do not contain a read which reads the initial value.

The result of running this Datalog program using the online clingo version (<https://potassco.org/clingo/run/>) is shown in the following:

```

clingo version 5.5.0
Reading from stdin
Solving...
Answer: 1
po("w(x,1,id0)","r(x,2,id1)") po("w(x,2,id2)","r(x,1,id3)") co("w(x,1,id0)",
"r(x,2,id1)") co("w(x,2,id2)","r(x,1,id3)") co("w(x,2,id2)","r(x,2,id1)")
co("w(x,1,id0)","r(x,1,id3)") wr("w(x,2,id2)","r(x,2,id1)")
wr("w(x,1,id0)","r(x,1,id3)") sv("r(x,2,id1)","w(x,1,id0)") sv("w(x,2,id2)",
"w(x,1,id0)") sv("r(x,1,id3)","w(x,1,id0)") sv("w(x,1,id0)",
"r(x,2,id1)") sv("w(x,2,id2)","r(x,2,id1)") sv("r(x,1,id3)","r(x,2,
id1)") sv("w(x,1,id0)","w(x,2,id2)") sv("r(x,2,id1)","w(x,2,id2)") sv
("r(x,1,id3)","w(x,2,id2)") sv("w(x,1,id0)","r(x,1,id3)") sv("r(x,2,
id1)","r(x,1,id3)") sv("w(x,2,id2)","r(x,1,id3)") initread("r(a,0,ida
)") wrt("w(x,1,id0)") wrt("w(x,2,id2)") rd("r(x,2,id1)") rd("r(x,1,
id3)")
SATISFIABLE

```

```

Models      : 1
Calls       : 1
Time        : 0.008s (Solving: 0.00s 1st Model: 0.00s Unsat: 0.00s)
CPU Time    : 0.000s

```

Now, let's see how CCv and CM bad-patterns are encoded. Since the CCv/CM bad-patterns include CC bad-patterns, each CCv/CM Datalog program should contain CC bad-patterns which we have already seen above in addition of some other rules we will see in the next sections.

**CCv bad-patterns encoding.** The CCv bad-patterns are encoded as follows:

```

% CCv bad-patterns
cf(X,Y) :- co(X,Z), wr(Y,Z), wrt(X), sv(X,Y), sv(X,Z). %Conflict order CF
cf(X,Y) :- cf(X,Z), cf(Z,Y). %Transitivity
cfco(X,Y) :- co(X,Y). %cfco= CF U CO, cfco is the union of cf and co.
cfco(X,Y) :- cf(X,Y). %cfco= CF U CO
%CCv bad-pattern
:- cfco(X,Y), cfco(Y,X). %CyclicCF CF U CO

```

Let's consider an example of CCv Datalog programs. As we have seen, the example 1b is not CCv so the following Datalog program for checking CCv is not satisfiable.

```

% Facts
wrt("w(x,1,id0)").
po("w(x,1,id0)","r(x,2,id1)").
sv("r(x,2,id1)","w(x,1,id0)").
sv("w(x,2,id2)","w(x,1,id0)").
sv("r(x,1,id3)","w(x,1,id0)").
rd("r(x,2,id1)").
sv("w(x,1,id0)","r(x,2,id1)").
wr("w(x,2,id2)","r(x,2,id1)").
sv("w(x,2,id2)","r(x,2,id1)").
sv("r(x,1,id3)","r(x,2,id1)").
wrt("w(x,2,id2)").
sv("w(x,1,id0)","w(x,2,id2)").
sv("r(x,2,id1)","w(x,2,id2)").
po("w(x,2,id2)","r(x,1,id3)").
sv("r(x,1,id3)","w(x,2,id2)").
rd("r(x,1,id3)").
wr("w(x,1,id0)","r(x,1,id3)").
sv("w(x,1,id0)","r(x,1,id3)").
sv("r(x,2,id1)","r(x,1,id3)").
sv("w(x,2,id2)","r(x,1,id3)").
initread("r(a,0,ida)").
% CC inference rules
co(X,Y) :- po(X,Y).

```

```

co(X,Y) :- wr(X,Y).
co(X,Z) :- co(X,Y), co(Y,Z). % Transitivity
% CC bad-patterns
:- co(X,X). % CyclicCO
:- co(X,Y), wrt(X), initread(Y), sv(X,Y). % WriteCOInitRead
:- co(X,Y), co(Y,Z), wr(X,Z), wrt(X), wrt(Y), rd(Z), sv(X,Y), sv(Y,Z).
% WriteCOWRead

% CCv inference rules
cf(X,Y) :- co(X,Z), wr(Y,Z), wrt(X), sv(X,Y), sv(X,Z). %Conflict order CF
cf(X,Y) :- cf(X,Z), cf(Z,Y). %Transitivity
cfco(X,Y) :- co(X,Y). %cfco= CF U CO, cfco is the union of cf and co.
cfco(X,Y) :- cf(X,Y). %cfco= CF U CO
%CCv bad-pattern
:- cfco(X,Y), cfco(Y,X). %CyclicCF CF U CO

clingo version 5.5.0
Reading from stdin
Solving...
UNSATISFIABLE

Models      : 0
Calls       : 1
Time        : 0.009s
CPU Time    : 0.000s

```

**CM bad-patterns encoding.** The CM bad-patterns are encoded as follows:

```

%CM inference rules
hb(X,0,0) :- co(X,0). %hbo initialized to causal order
hb(X,Y,0) :- hb(Y,0,0), co(X,Y).
hb(X,Y,0) :- hb(X,Z,0), po(Z,0), wr(Y,Z), wrt(X), sv(X,Y).% hbo
    definition
hb(X,Y,0) :- hb(X,Z,0), wr(Y,Z), wrt(X), sv(X,Y). % hbo definition
hb(X,Z,0) :- hb(X,Y,0), hb(Y,Z,0). %Transitivity
%CM bad-patterns
:- hb(X,Y,0), wrt(X), sv(X,Y), po(Y,0), initread(Y). %WriteHBInitRead
:- hb(X,Y,0), hb(Y,X,0). %CyclicHB

```

As we have mentioned in the section 5, CM<sub>1</sub> and CM<sub>2</sub> are characterized by the same CM bad-patterns described above. The only difference is that for the CM<sub>2</sub>, we have added a function which identifies the po-maximal operation in each thread. We replace then the operation "O" in the CM bad-patterns above by these identified operations (last operation in each thread) instead of replacing it by all read/write operations in the history (CM<sub>1</sub>).

For a better understanding, consider the instantiation of the CM bad-patterns for CM<sub>1</sub> and CM<sub>2</sub>.



– For CM<sub>1</sub>: We replace "O" by all operations in the history.

```
%CM inference rules
hb(X,w(x,1,id0),w(x,1,id0)) :- co(X,w(x,1,id0)).
hb(X,Y,w(x,1,id0)) :- hb(Y,w(x,1,id0),w(x,1,id0)), co(X,Y).
hb(X,Y,w(x,1,id0)) :- hb(X,Z,w(x,1,id0)), po(Z,w(x,1,id0)), wr(Y,Z), wrt(
    X), sv(X,Y).
hb(X,Y,w(x,1,id0)) :- hb(X,Z,w(x,1,id0)), wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,w(x,1,id0)) :- hb(X,Y,w(x,1,id0)), hb(Y,Z,w(x,1,id0)).
%CM bad-patterns
:- hb(X,Y,w(x,1,id0)), wrt(X), sv(X,Y), po(Y,w(x,1,id0)), initread(Y).
:- hb(X,Y,w(x,1,id0)), hb(Y,X,w(x,1,id0)).
%CM inference rules
hb(X,w(x,2,id2),w(x,2,id2)) :- co(X,w(x,2,id2)).
hb(X,Y,w(x,2,id2)) :- hb(Y,w(x,2,id2),w(x,2,id2)), co(X,Y).
hb(X,Y,w(x,2,id2)) :- hb(X,Z,w(x,2,id2)), po(Z,w(x,2,id2)), wr(Y,Z), wrt(
    X), sv(X,Y).
hb(X,Y,w(x,2,id2)) :- hb(X,Z,w(x,2,id2)), wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,w(x,2,id2)) :- hb(X,Y,w(x,2,id2)), hb(Y,Z,w(x,2,id2)).
%CM bad-patterns
:- hb(X,Y,w(x,2,id2)), wrt(X), sv(X,Y), po(Y,w(x,2,id2)), initread(Y).
:- hb(X,Y,w(x,2,id2)), hb(Y,X,w(x,2,id2)).
%CM inference rules
hb(X,"r(x,2,id1)","r(x,2,id1)") :- co(X,"r(x,2,id1)").
hb(X,Y,"r(x,2,id1)") :- hb(Y,"r(x,2,id1)","r(x,2,id1)", co(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)", po(Z,"r(x,2,id1)", wr(Y,Z)
    , wrt(X), sv(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)", wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,2,id1)") :- hb(X,Y,"r(x,2,id1)", hb(Y,Z,"r(x,2,id1)").
%CM bad-patterns
:- hb(X,Y,"r(x,2,id1)", wrt(X), sv(X,Y), po(Y,"r(x,2,id1)", initread(Y)
    .
:- hb(X,Y,"r(x,2,id1)", hb(Y,X,"r(x,2,id1)").
%CM inference rules
hb(X,"r(x,1,id3)","r(x,1,id3)") :- co(X,"r(x,1,id3)").
hb(X,Y,"r(x,1,id3)") :- hb(Y,"r(x,1,id3)","r(x,1,id3)", co(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)", po(Z,"r(x,1,id3)", wr(Y,Z)
    , wrt(X), sv(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)", wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,1,id3)") :- hb(X,Y,"r(x,1,id3)", hb(Y,Z,"r(x,1,id3)").
%CM bad-patterns
:- hb(X,Y,"r(x,1,id3)", wrt(X), sv(X,Y), po(Y,"r(x,1,id3)", initread(Y)
    .
:- hb(X,Y,"r(x,1,id3)", hb(Y,X,"r(x,1,id3)").
```

- For CM\_2: We replace "*O*" by the last operation in each process in the history.

```
%CM inference rules
hb(X,"r(x,2,id1)","r(x,2,id1)") :- co(X,"r(x,2,id1)").
hb(X,Y,"r(x,2,id1)") :- hb(Y,"r(x,2,id1)","r(x,2,id1)"), co(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)"), po(Z,"r(x,2,id1)"), wr(Y,Z)
    , wrt(X), sv(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)"), wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,2,id1)") :- hb(X,Y,"r(x,2,id1)"), hb(Y,Z,"r(x,2,id1)").

%CM bad-patterns
:- hb(X,Y,"r(x,2,id1)"), wrt(X), sv(X,Y), po(Y,"r(x,2,id1)"), initread(Y)
    .
:- hb(X,Y,"r(x,2,id1)"), hb(Y,X,"r(x,2,id1)").

%CM inference rules
hb(X,"r(x,1,id3)","r(x,1,id3)") :- co(X,"r(x,1,id3)").
hb(X,Y,"r(x,1,id3)") :- hb(Y,"r(x,1,id3)","r(x,1,id3)"), co(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)"), po(Z,"r(x,1,id3)"), wr(Y,Z)
    , wrt(X), sv(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)"), wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,1,id3)") :- hb(X,Y,"r(x,1,id3)"), hb(Y,Z,"r(x,1,id3)").

%CM bad-patterns
:- hb(X,Y,"r(x,1,id3)"), wrt(X), sv(X,Y), po(Y,"r(x,1,id3)"), initread(Y)
    .
:- hb(X,Y,"r(x,1,id3)"), hb(Y,X,"r(x,1,id3)").
```

Now let's consider the whole Datalog programs and their running results.

- For CM\_1:

```
% Facts
wrt("w(x,1,id0)").
po("w(x,1,id0)","r(x,2,id1)").
sv("r(x,2,id1)","w(x,1,id0)").
sv("w(x,2,id2)","w(x,1,id0)").
sv("r(x,1,id3)","w(x,1,id0)").
rd("r(x,2,id1)").
sv("w(x,1,id0)","r(x,2,id1)").
wr("w(x,2,id2)","r(x,2,id1)").
sv("w(x,2,id2)","r(x,2,id1)").
sv("r(x,1,id3)","r(x,2,id1)").
wrt("w(x,2,id2)").
sv("w(x,1,id0)","w(x,2,id2)").
sv("r(x,2,id1)","w(x,2,id2)").
po("w(x,2,id2)","r(x,1,id3)").
sv("r(x,1,id3)","w(x,2,id2)").
rd("r(x,1,id3)").
wr("w(x,1,id0)","r(x,1,id3)").
sv("w(x,1,id0)","r(x,1,id3)").
```

```

sv("r(x,2,id1)","r(x,1,id3)").
sv("w(x,2,id2)","r(x,1,id3)").
initread("r(a,0,ida)").
% Inference rules
co(X,Y) :- po(X,Y).
co(X,Y) :- wr(X,Y).
co(X,Z) :- co(X,Y), co(Y,Z). % Transitivity
% CC bad-patterns
:- co(X,X). % CyclicCO
:- co(X,Y), wrt(X), initread(Y), sv(X,Y). % WriteCOInitRead
:- co(X,Y), co(Y,Z), wr(X,Z), wrt(X), wrt(Y), rd(Z), sv(X,Y), sv(Y,Z).
% WriteCORRead

%CM inference rules
hb(X,w(x,1,id0),w(x,1,id0)) :- co(X,w(x,1,id0)).
hb(X,Y,w(x,1,id0)) :- hb(Y,w(x,1,id0),w(x,1,id0)), co(X,Y).
hb(X,Y,w(x,1,id0)) :- hb(X,Z,w(x,1,id0)), po(Z,w(x,1,id0)), wr(Y,Z), wrt(
    X), sv(X,Y).
hb(X,Y,w(x,1,id0)) :- hb(X,Z,w(x,1,id0)), wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,w(x,1,id0)) :- hb(X,Y,w(x,1,id0)), hb(Y,Z,w(x,1,id0)).
%CM bad-patterns
:- hb(X,Y,w(x,1,id0)), wrt(X), sv(X,Y), po(Y,w(x,1,id0)), initread(Y).
:- hb(X,Y,w(x,1,id0)), hb(Y,X,w(x,1,id0)).
% CM inference rules
hb(X,w(x,2,id2),w(x,2,id2)) :- co(X,w(x,2,id2)).
hb(X,Y,w(x,2,id2)) :- hb(Y,w(x,2,id2),w(x,2,id2)), co(X,Y).
hb(X,Y,w(x,2,id2)) :- hb(X,Z,w(x,2,id2)), po(Z,w(x,2,id2)), wr(Y,Z), wrt(
    X), sv(X,Y).
hb(X,Y,w(x,2,id2)) :- hb(X,Z,w(x,2,id2)), wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,w(x,2,id2)) :- hb(X,Y,w(x,2,id2)), hb(Y,Z,w(x,2,id2)).
%CM bad-patterns
:- hb(X,Y,w(x,2,id2)), wrt(X), sv(X,Y), po(Y,w(x,2,id2)), initread(Y).
:- hb(X,Y,w(x,2,id2)), hb(Y,X,w(x,2,id2)).
% CM inference rules
hb(X,"r(x,2,id1)","r(x,2,id1)") :- co(X,"r(x,2,id1)").
hb(X,Y,"r(x,2,id1)") :- hb(Y,"r(x,2,id1)","r(x,2,id1)", co(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)", po(Z,"r(x,2,id1)", wr(Y,Z)
    , wrt(X), sv(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)", wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,2,id1)") :- hb(X,Y,"r(x,2,id1)", hb(Y,Z,"r(x,2,id1)").
%CM bad-patterns
:- hb(X,Y,"r(x,2,id1)", wrt(X), sv(X,Y), po(Y,"r(x,2,id1)", initread(Y)
    .
:- hb(X,Y,"r(x,2,id1)", hb(Y,X,"r(x,2,id1)").
% CM inference rules
hb(X,"r(x,1,id3)","r(x,1,id3)") :- co(X,"r(x,1,id3)").
hb(X,Y,"r(x,1,id3)") :- hb(Y,"r(x,1,id3)","r(x,1,id3)", co(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)", po(Z,"r(x,1,id3)", wr(Y,Z)
    , wrt(X), sv(X,Y).

```

```

hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)", wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,1,id3)") :- hb(X,Y,"r(x,1,id3)", hb(Y,Z,"r(x,1,id3)").
%CM bad-patterns
:- hb(X,Y,"r(x,1,id3)", wrt(X), sv(X,Y), po(Y,"r(x,1,id3)", initread(Y)
.
:- hb(X,Y,"r(x,1,id3)", hb(Y,X,"r(x,1,id3)").

clingo version 5.5.0
Reading from stdin
Solving...
Answer: 1
po("w(x,1,id0)","r(x,2,id1)") po("w(x,2,id2)","r(x,1,id3)") co("w(x,1,id0)
","r(x,2,id1)") co("w(x,2,id2)","r(x,1,id3)") co("w(x,2,id2)","r(x
,2,id1)") co("w(x,1,id0)","r(x,1,id3)") wr("w(x,2,id2)","r(x,2,id1)")
wr("w(x,1,id0)","r(x,1,id3)") sv("r(x,2,id1)","w(x,1,id0)") sv("w(x
,2,id2)","w(x,1,id0)") sv("r(x,1,id3)","w(x,1,id0)") sv("w(x,1,id0)",
"r(x,2,id1)") sv("w(x,2,id2)","r(x,2,id1)") sv("r(x,1,id3)","r(x,2,
id1)") sv("w(x,1,id0)","w(x,2,id2)") sv("r(x,2,id1)","w(x,2,id2)") sv
("r(x,1,id3)","w(x,2,id2)") sv("w(x,1,id0)","r(x,1,id3)") sv("r(x,2,
id1)","r(x,1,id3)") sv("w(x,2,id2)","r(x,1,id3)") initread("r(a,0,ida
)") wrt("w(x,1,id0)") wrt("w(x,2,id2)") rd("r(x,2,id1)") rd("r(x,1,
id3)") hb("w(x,2,id2)","r(x,1,id3)","r(x,1,id3)") hb("w(x,1,id0)","r(
x,1,id3)","r(x,1,id3)") hb("w(x,2,id2)","w(x,1,id0)","r(x,1,id3)") hb
("w(x,1,id0)","r(x,2,id1)","r(x,2,id1)") hb("w(x,2,id2)","r(x,2,id1)"
,"r(x,2,id1)") hb("w(x,1,id0)","w(x,2,id2)","r(x,2,id1)")
SATISFIABLE

Models      : 1
Calls       : 1
Time        : 0.029s
CPU Time    : 0.000s

```

– For CM\_2:

```

% Facts
wrt("w(x,1,id0)").
po("w(x,1,id0)","r(x,2,id1)").
sv("r(x,2,id1)","w(x,1,id0)").
sv("w(x,2,id2)","w(x,1,id0)").
sv("r(x,1,id3)","w(x,1,id0)").
rd("r(x,2,id1)").
sv("w(x,1,id0)","r(x,2,id1)").
wr("w(x,2,id2)","r(x,2,id1)").
sv("w(x,2,id2)","r(x,2,id1)").
sv("r(x,1,id3)","r(x,2,id1)").
wrt("w(x,2,id2)").
sv("w(x,1,id0)","w(x,2,id2)").
sv("r(x,2,id1)","w(x,2,id2)").
po("w(x,2,id2)","r(x,1,id3)").

```

```

sv("r(x,1,id3)","w(x,2,id2)").
rd("r(x,1,id3)").
wr("w(x,1,id0)","r(x,1,id3)").
sv("w(x,1,id0)","r(x,1,id3)").
sv("r(x,2,id1)","r(x,1,id3)").
sv("w(x,2,id2)","r(x,1,id3)").
initread("r(a,0,ida)").
% CC inference rules
co(X,Y) :- po(X,Y).
co(X,Y) :- wr(X,Y).
co(X,Z) :- co(X,Y), co(Y,Z). % Transitivity
% CC bad-patterns
:- co(X,X). % CyclicCO
:- co(X,Y), wrt(X), initread(Y), sv(X,Y). % WriteCOInitRead
:- co(X,Y), co(Y,Z), wr(X,Z), wrt(X), wrt(Y), rd(Z), sv(X,Y), sv(Y,Z).
% WriteCORead

% CM inference rules

hb(X,"r(x,2,id1)","r(x,2,id1)") :- co(X,"r(x,2,id1)").
hb(X,Y,"r(x,2,id1)") :- hb(Y,"r(x,2,id1)","r(x,2,id1)", co(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)", po(Z,"r(x,2,id1)", wr(Y,Z)
, wrt(X), sv(X,Y).
hb(X,Y,"r(x,2,id1)") :- hb(X,Z,"r(x,2,id1)", wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,2,id1)") :- hb(X,Y,"r(x,2,id1)", hb(Y,Z,"r(x,2,id1)").
%CM bad-patterns
:- hb(X,Y,"r(x,2,id1)", wrt(X), sv(X,Y), po(Y,"r(x,2,id1)", initread(Y)
.
:- hb(X,Y,"r(x,2,id1)", hb(Y,X,"r(x,2,id1)").

% CM inference rules
hb(X,"r(x,1,id3)","r(x,1,id3)") :- co(X,"r(x,1,id3)").
hb(X,Y,"r(x,1,id3)") :- hb(Y,"r(x,1,id3)","r(x,1,id3)", co(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)", po(Z,"r(x,1,id3)", wr(Y,Z)
, wrt(X), sv(X,Y).
hb(X,Y,"r(x,1,id3)") :- hb(X,Z,"r(x,1,id3)", wr(Y,Z), wrt(X), sv(X,Y).
hb(X,Z,"r(x,1,id3)") :- hb(X,Y,"r(x,1,id3)", hb(Y,Z,"r(x,1,id3)").
%CM bad-patterns
:- hb(X,Y,"r(x,1,id3)", wrt(X), sv(X,Y), po(Y,"r(x,1,id3)", initread(Y)
.
:- hb(X,Y,"r(x,1,id3)", hb(Y,X,"r(x,1,id3)").

clingo version 5.5.0
Reading from stdin
Solving...
Answer: 1
po("w(x,1,id0)","r(x,2,id1)") po("w(x,2,id2)","r(x,1,id3)") co("w(x,1,id0)
","r(x,2,id1)") co("w(x,2,id2)","r(x,1,id3)") co("w(x,2,id2)","r(x
,2,id1)") co("w(x,1,id0)","r(x,1,id3)") wr("w(x,2,id2)","r(x,2,id1)")
wr("w(x,1,id0)","r(x,1,id3)") sv("r(x,2,id1)","w(x,1,id0)") sv("w(x

```

```

,2,id2)","w(x,1,id0)") sv("r(x,1,id3)","w(x,1,id0)") sv("w(x,1,id0)",
"r(x,2,id1)") sv("w(x,2,id2)","r(x,2,id1)") sv("r(x,1,id3)","r(x,2,
id1)") sv("w(x,1,id0)","w(x,2,id2)") sv("r(x,2,id1)","w(x,2,id2)") sv
("r(x,1,id3)","w(x,2,id2)") sv("w(x,1,id0)","r(x,1,id3)") sv("r(x,2,
id1)","r(x,1,id3)") sv("w(x,2,id2)","r(x,1,id3)") initread("r(a,0,ida
)") wrt("w(x,1,id0)") wrt("w(x,2,id2)") rd("r(x,2,id1)") rd("r(x,1,
id3)") hb("w(x,2,id2)","r(x,1,id3)","r(x,1,id3)") hb("w(x,1,id0)","r(x,
x,1,id3)","r(x,1,id3)") hb("w(x,2,id2)","w(x,1,id0)","r(x,1,id3)") hb
("w(x,1,id0)","r(x,2,id1)","r(x,2,id1)") hb("w(x,2,id2)","r(x,2,id1)"
,"r(x,2,id1)") hb("w(x,1,id0)","w(x,2,id2)","r(x,2,id1)")
SATISFIABLE
Models      : 1
Calls       : 1
Time        : 0.011s
CPU Time    : 0.000s

```

As we have mentioned before, our new approach (CM\_2) computes the  $hb_o$  relation for a small set of operations (po-maximal operations) compared to CM\_1. As can be seen above, the size of the Datalog program was considerably reduced when we use CM\_2 for a small history. Let alone long histories that contains hundreds of operations. This will be seen in the experimental results (Section 6).

#### 5.4 A procedure for checking Causal Consistency

Let's name the procedure which implements the reduction we have seen in the previous section REDUC-to-DATALOG. This procedure takes as input the history  $h$  and the causal consistency model  $M$  to check, and returns the corresponding Datalog program  $\mathcal{D}$ . Then, we call another procedure named DATALOG-SOLVER which verifies whether  $\mathcal{D}$  is SATISFIABLE or not.

```

Input: A history  $h = \langle O, po, wr \rangle$  and a causal consistency model  $M$ 
Output: SAT iff  $h$  satisfies  $M$ 
1 REDUC-to-DATALOG( $h, M$ )
2 if DATALOG-SOLVER( $REDUC\text{-}to\text{-}DATALOG(h, M)$ ) then
3   | return true;
4 else
5   | return false;
6 end

```

**Algorithm 1:** Checking Causal Consistency.

**Theorem 3.** *Algorithm 1 returns true iff the input history  $h$  satisfies the causal consistency model  $M$ .*

The correctness of this theorem is ensured by the fact that our reduction is a simple and direct encoding of bad patterns to Datalog and these bad-patterns were proven in [9] to capture exactly the causal consistency violations.

### 5.5 Complexity

The complexity of a Datalog program is  $\mathcal{O}(n^k)$  [29], where  $n$  is the number of constants in the input data, and  $k$  is the maximum number of variables in a clause. As we have seen in the previous section, given a history  $\langle O, po, wr \rangle$ , the maximum number of variables in a rule in our Datalog programs is 3, thus the complexity of our approach is  $\mathcal{O}(n^3)$ , where  $n$  is the size of the computation (the number of operations). Our approach's complexity is better than the one defined in [9] in which the complexity of checking CC, CCv and CM was shown to be  $\mathcal{O}(n^5)$ .

## 6 Experimental Evaluation

We have investigated the efficiency and scalability of our tool (named *CausalC-Checker*) by applying it to two real-life distributed transactional databases, CockroachDB [1] and Galera [2].

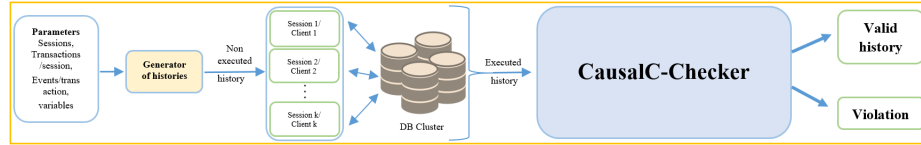


Fig. 3: The General architecture of the histories checking procedure

**Histories generation:** The Figure 3 presents the general architecture of the used testing procedure in the next experiments. Histories are generated using random clients with the parameters, the number of sessions, the number of transactions per session, the number of events per transaction (in this paper, we consider one event per transaction), and the number of variables. A client is generated by the generator of histories (Algorithm 2) by choosing randomly the type of operation (read or write) in each transaction, the variable and a value for write operations. That constitutes non executed histories that are the histories which do not contain the return values of read operations. Each client performs a session, communicates with the database cluster by executing operations (read-/write) and gets the return values for read operations. The recorded histories are called executed histories in the Figure 3. We ensure that all histories are differentiated. These histories are the input of our *CausalC-Checker*.

### 6.1 Case study 1: CockroachDB.

We have used the highly available and strongly consistent distributed database CockroachDB [1] (v2.1.0) that is built on a transactional strongly-consistent key-value store, so it is expected to be causally consistent. Considering one operation

```

Input: nClient, nTransaction, nEvent, nVariable
Output: A non executed history

1  lastWrite  $\leftarrow \emptyset$ ;
2  foreach  $v \in 1..nVariable$  do
3    | lastWrite( $v$ )  $\leftarrow 0$ ;
4  end
5  history  $\leftarrow \emptyset$ ;
6  foreach  $1..nClient$  do
7    | Client  $\leftarrow \emptyset$ ;
8    | foreach  $1..nTransaction$  do
9      | Transaction  $\leftarrow \emptyset$ ;
10     | foreach  $1..nEvent$  do
11       | Event  $\leftarrow new(Event)$ ;
12       | Event.operation  $\leftarrow uniformly\_choose(\{Read, Write\})$ ;
13       | Event.variable  $\leftarrow uniformly\_choose(\{1..nVariable\})$ ;
14       | if Event.operation = Write then
15         | Event.value  $\leftarrow lastWrite(Event.variable) + 1$ ;
16         | lastWrite(Event.variable)  $\leftarrow lastWrite(Event.variable) + 1$ ;
17       | end
18       | Transaction.push(Event);
19     | end
20     | Client.push(Transaction);
21   | end
22   | history.push(Client);
23 end
24 return history;

```

**Algorithm 2:** The histories generator algorithm

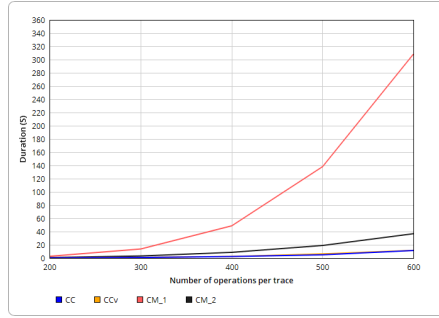
per transaction lead to our model.

We have examined the effect of the number of operations on runtime for a fixed number of processes (4 processes) and the effect of the number of processes. We have tested 200 histories for each configuration and calculated the average runtime.

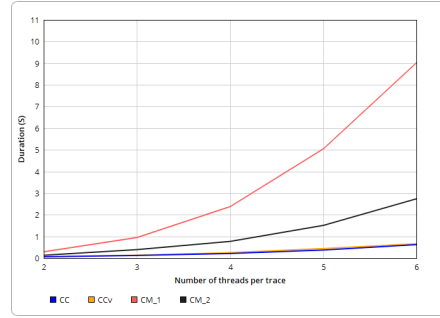
We have checked CC, CCv and CM, using its two definitions CM\_1 and CM\_2, for all generated histories. Figure 4 shows the results. The graphs 4a, 4c, 4e and 4g show the runtime while increasing the number of operations from 100 to 600, in augmentations of 100 (with a fixed number of processes, 4 processes). The graphs 4b, 4d, 4f and 4h report the runtime when increasing the number of processes from 2 to 6, in augmentations of 1. For each number of processes  $x$  we have considered  $50x$  operations, so increasing the number of processes increases the number of operations in the history as well.

The graph 4a resp., 4b shows a comparison between CC, CCv, CM\_1 and CM\_2 verification runtimes while varying the number of operations resp., the number of processes. The graph 4c resp., 4d, presents the running time of CM\_2 verification compared to CC and CCv verification running time. The graph 4e resp., graph 4g, shows the evolution of CC and CCv verification resp., CM\_1 and CM\_2 verification, runtime while increasing the number of operations. The

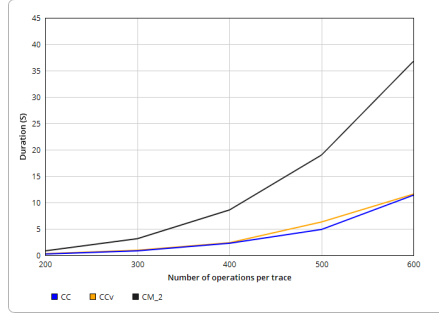




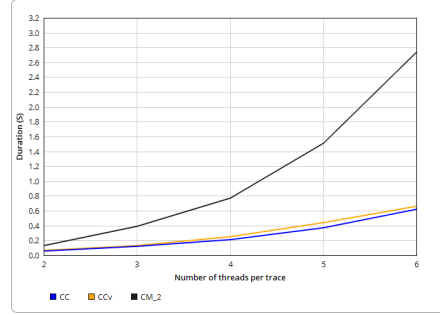
(a) Checking Causal Consistency while varying the number of operations.



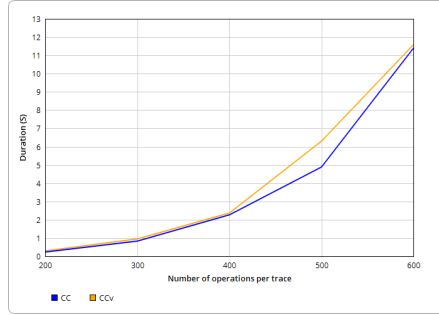
(b) Checking Causal Consistency while varying the number of processes.



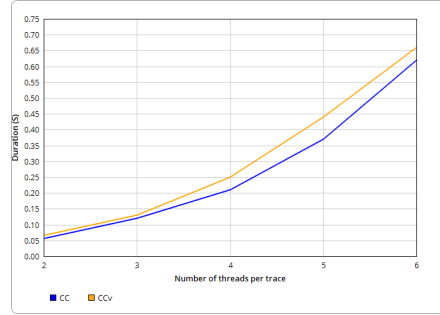
(c) Checking CC, CCv and CM\_2 while varying the number of operations.



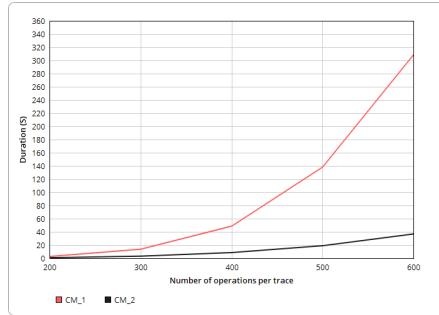
(d) Checking CC, CCv and CM\_2 while varying the number of processes.



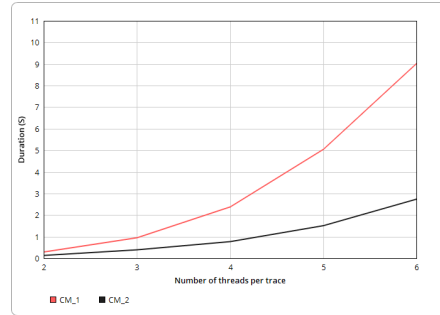
(e) Checking CC and CCv while varying the number of operations.



(f) Checking CC and CCv while varying the number of processes.



(g) Comparing CM\_1 and CM\_2 run-times while varying the number of operations.



(h) Comparing CM\_1 and CM\_2 run-times while varying the number of processes.

Fig. 4: Checking Causal Consistency for CockroachDB histories.

graph 4f resp., graph 4h, shows the evolution of CC and CCv verification resp., CM\_1 and CM\_2 verification, runtime while increasing the number of processes.

Our approach is more efficient in the case of CC and CCv verification compared to the CM\_1 case (graphs 4a and 4b). The figure 4c resp., 4d, is a zoom on CC, CCv and CM\_2 of figure 4a resp., 4b. It shows that the CM\_2 improves the running time but costs more compared to CC and CCv as well. The figure 4e resp., 4f, is a zoom on CC and CCv of figure 4a resp., 4b. It shows that CC and CCv verification are very efficient and terminates in less than 11.6 seconds for all histories we have tested. As we have noticed above, the results shown in 4g and 4h show that CM\_2 has better performance, by factors of 8 times in the case of 600 operations. As expected, all the tested histories were valid w.r.t. all the considered causal consistency models.

## 6.2 Case study 2: Galera.

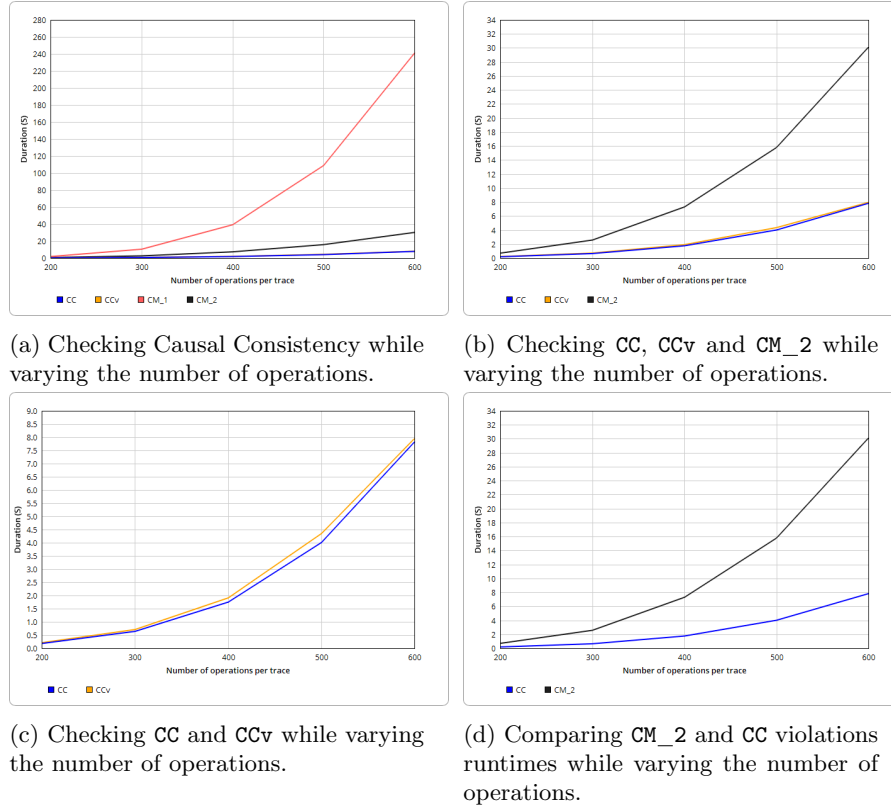


Fig. 5: Checking Causal Consistency for Galera histories.

We have also used the cluster called Galera [2] (v3.20). Galera Cluster is a database cluster based on synchronous replication and Oracle's InnoDB/MySQL.

It is expected to implement *Snapshot isolation* when transactions are processed in separated nodes.

Similarly to the first case study, we have studied the evolution of runtime while increasing the number of operations from 100 to 600, in augmentations of 100. We have verified 200 histories for each number of operations and compute the runtime average.

The graphs in Figure 5 show the impact of increasing the number of operations on runtime while fixing the number of processes (4 processes). The graph 5a shows the comparison of CC, CCv, CM\_1 and CM\_2 verification runtimes. The graph 5b presents a zoom on graph 5a in order to compare CM\_2 to CC and CCv. The graph 5c reports the evolution of CC and CCv verification runtime. Finally, the graph 5d presents a comparison between CC and CM\_2 running times.

Similarly to the CockroachDB case study, our approach is more efficient in the case of CC and CCv either while increasing the number of operations or processes. The graph 5a shows that our new definition CM\_2 outperforms CM\_1, but still less efficient compared to CC and CCv (graph 5b).

Our approach allows capturing violations on the Galera database. We have found that 1.25% of the tested Galera histories violate causal consistency, that confirms the bugs submitted on Github[3]. We mention that 73.3% of the detected CM violations are also CC violations. The suggested approach scales well and detects violations on the used version of Galera DB.

The experiments show that our approach is efficient for both verification of valid computations and detection of violations, especially in the case of CC and CCv. The gap between CC (CCv) and CM\_1 runtimes reported in the graphs 4a, 4b and 5a is due to the fact that in CM\_1 we compute the  $hb_o$  relation and check the bad-patterns for each operation. This gap is reduced using the new definition CM\_2 (graphs 4a, 4b and 5a) in which we compute the  $hb_o$  relation and check the bad-patterns for only the last operation of each thread.

## 7 Related Work

Several works have considered the problem of checking strong consistency models such as *Linearizability* and *Sequential consistency* (SC) [4, 5, 11, 14, 16, 17, 27, 30]. Our recent works [33, 31] address the problem of verifying SC and TSO (*Total store ordering*) gradually by using several variants of causal consistency (and other weak consistency models) including the ones we have considered in this work. However, few have addressed the problem of checking weak consistency models. Emmi and Enea [15] propose an algorithm to optimize the consistency checking based on the notion of minimal-visibility. However, their work relies on some specific relaxations in those criteria, leading to the naive enumeration in the context of strong consistency models such as SC and TSO. Bouajjani et al. [10] presents a formalization of eventual consistency for replicated objects and reduces the problem of checking eventual consistency to reachability and model checking problems.

Bouajjani et al. [9] considers the problem of checking causal consistency. They

present the formalization of the different variations of causal consistency (CC, CCv and CM) we use in this work and a complete characterization of the violations of those models. In addition, they show that checking if an execution satisfies one of those models is polynomial time ( $\mathcal{O}(n^5)$ ). However, this work does not propose any implementation. Our work presents an implementation based on a reduction to Datalog queries solving which improves the complexity from  $\mathcal{O}(n^5)$  to  $\mathcal{O}(n^3)$ .

## 8 Conclusion

We have presented a tool for checking automatically that given computations of a system are causally consistent. Our procedure for solving this conformance problem is based on implementing the theoretical approach introduced in [9] where causal consistency violations are characterized in terms of the occurrence of some particular bad-patterns. We build on this work by reducing the problem of detecting the existence of these patterns in computations to the problem of solving Datalog queries. We have applied our algorithm to two real-life case studies. The experimental results show that in the case of CC and CCv our approach is efficient and scalable. In the CM case, the cost grows polynomially but much faster than in the case of CC and CCv. In order to improve the CM checking performance, an optimized definition (CM\_2) of the original definition [9] has been proposed. Our experimental results show that this new definition reduce considerably the cost of CM verification. It reduces the CM verification runtime by more than 7 times (for histories with 600 operations). However, this optimized CM definition still less efficient compared to CC and CCv. Nevertheless, it turned out that interestingly, most of the CM violations (73.3%) that we found are in fact CC violations, and therefore can be caught using a more efficient procedure in which one can start by verifying CC first.

## References

1. <https://www.cockroachlabs.com>, [Retrieved 15/11/2018]
2. <http://galeracluster.com>, [Retrieved 15/11/2018]
3. <https://github.com/codership/galera/issues/336>, [Retrieved 15/11/2018]
4. Abdulla, P.A., Atig, M.F., Jonsson, B., Lång, M., Ngo, T.P., Sagonas, K.: Optimal stateless model checking for reads-from equivalence under sequential consistency. *Proc. ACM Program. Lang.* **3**(OOPSLA) (Oct 2019). <https://doi.org/10.1145/3360576>, <https://doi.org/10.1145/3360576>
5. Abdulla, P.A., Haziza, F., Holík, L.: Parameterized verification through view abstraction. *STTT* **18**(5), 495–516 (2016). <https://doi.org/10.1007/s10009-015-0406-x>, <https://doi.org/10.1007/s10009-015-0406-x>
6. Abiteboul, S., Hull, R., Vianu, V. (eds.): *Foundations of Databases: The Logical Level*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edn. (1995)

7. Ahamad, M., Neiger, G., Burns, J.E., Kohli, P., Hutto, P.W.: Causal memory: Definitions, implementation, and programming. *Distrib. Comput.* **9**(1), 37–49 (Mar 1995). <https://doi.org/10.1007/BF01784241>, <https://doi.org/10.1007/BF01784241>
8. Bailis, P., Ghodsi, A., Hellerstein, J.M., Stoica, I.: Bolt-on causal consistency. In: *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. pp. 761–772. SIGMOD '13, ACM, New York, NY, USA (2013). <https://doi.org/10.1145/2463676.2465279>, <http://doi.acm.org/10.1145/2463676.2465279>
9. Bouajjani, A., Enea, C., Guerraoui, R., Hamza, J.: On verifying causal consistency. In: Castagna, G., Gordon, A.D. (eds.) *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages, POPL 2017, Paris, France, January 18–20, 2017*. pp. 626–638. ACM (2017), <http://dl.acm.org/citation.cfm?id=3009888>
10. Bouajjani, A., Enea, C., Hamza, J.: Verifying eventual consistency of optimistic replication systems. In: *Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*. pp. 285–296. POPL '14, ACM, New York, NY, USA (2014). <https://doi.org/10.1145/2535838.2535877>, <http://doi.acm.org/10.1145/2535838.2535877>
11. Burckhardt, S., Dern, C., Musuvathi, M., Tan, R.: Line-up: a complete and automatic linearizability checker. In: Zorn, B.G., Aiken, A. (eds.) *Proceedings of the 2010 ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2010, Toronto, Ontario, Canada, June 5–10, 2010*. pp. 330–340. ACM (2010). <https://doi.org/10.1145/1806596.1806634>, <https://doi.org/10.1145/1806596.1806634>
12. Du, J., Elnikety, S., Roy, A., Zwaenepoel, W.: Orbe: Scalable causal consistency using dependency matrices and physical clocks. In: *Proceedings of the 4th Annual Symposium on Cloud Computing*. pp. 11:1–11:14. SOCC '13, ACM, New York, NY, USA (2013). <https://doi.org/10.1145/2523616.2523628>, <http://doi.acm.org/10.1145/2523616.2523628>
13. Du, J., Iorgulescu, C., Roy, A., Zwaenepoel, W.: Gentlerain: Cheap and scalable causal consistency with physical clocks. *Proceedings of the 5th ACM Symposium on Cloud Computing, SOCC 2014 (11 2014)*. <https://doi.org/10.1145/2670979.2670983>
14. Eiriksson, Á.T., McMillan, K.L.: Using formal verification/analysis methods on the critical path in system design: A case study. In: Wolper, P. (ed.) *Computer Aided Verification, 7th International Conference, Liège, Belgium, July, 3–5, 1995, Proceedings. Lecture Notes in Computer Science*, vol. 939, pp. 367–380. Springer (1995). [https://doi.org/10.1007/3-540-60045-0\\_63](https://doi.org/10.1007/3-540-60045-0_63), [https://doi.org/10.1007/3-540-60045-0\\_63](https://doi.org/10.1007/3-540-60045-0_63)
15. Emmi, M., Enea, C.: Monitoring weak consistency. In: Chockler, H., Weissenbacher, G. (eds.) *Computer Aided Verification - 30th International Conference, CAV 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 14–17, 2018, Proceedings, Part I. Lecture Notes in Computer Science*, vol. 10981, pp. 487–506. Springer (2018). [https://doi.org/10.1007/978-3-319-96145-3\\_26](https://doi.org/10.1007/978-3-319-96145-3_26), [https://doi.org/10.1007/978-3-319-96145-3\\_26](https://doi.org/10.1007/978-3-319-96145-3_26)
16. Emmi, M., Enea, C.: Sound, complete, and tractable linearizability monitoring for concurrent collections. *PACMPL* **2**(POPL), 25:1–25:27 (2018). <https://doi.org/10.1145/3158113>, <https://doi.org/10.1145/3158113>

17. Emmi, M., Enea, C., Hamza, J.: Monitoring refinement via symbolic reasoning. In: Grove, D., Blackburn, S. (eds.) *Proceedings of the 36th ACM SIGPLAN Conference on Programming Language Design and Implementation*, Portland, OR, USA, June 15-17, 2015. pp. 260–269. ACM (2015). <https://doi.org/10.1145/2737924.2737983>, <https://doi.org/10.1145/2737924.2737983>
18. Gilbert, S., Lynch, N.: Brewer’s conjecture and the feasibility of consistent, available, partition-tolerant web services. *SIGACT News* **33**(2), 51–59 (Jun 2002). <https://doi.org/10.1145/564585.564601>, <http://doi.acm.org/10.1145/564585.564601>
19. Herlihy, M.P., Wing, J.M.: Linearizability: A correctness condition for concurrent objects. *ACM Trans. Program. Lang. Syst.* **12**(3), 463–492 (Jul 1990). <https://doi.org/10.1145/78969.78972>, <http://doi.acm.org/10.1145/78969.78972>
20. Lamport, L.: How to make a multiprocessor computer that correctly executes multiprocess programs. *IEEE Trans. Comput.* **28**(9), 690–691 (Sep 1979). <https://doi.org/10.1109/TC.1979.1675439>, <https://doi.org/10.1109/TC.1979.1675439>
21. Lamport, L.: Time, clocks, and the ordering of events in a distributed system. *Commun. ACM* **21**(7), 558–565 (Jul 1978). <https://doi.org/10.1145/359545.359563>, <http://doi.acm.org/10.1145/359545.359563>
22. Lloyd, W., Freedman, M.J., Kaminsky, M., Andersen, D.G.: Don’t settle for eventual: Scalable causal consistency for wide-area storage with cops. In: *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles*. pp. 401–416. SOSP ’11, ACM, New York, NY, USA (2011). <https://doi.org/10.1145/2043556.2043593>, <http://doi.acm.org/10.1145/2043556.2043593>
23. Mahajan, P., Alvisi, L., Dahlin, M.: Consistency, availability, convergence. Tech. rep. (2011)
24. Perrin, M., Mostefaoui, A., Jard, C.: Causal consistency: Beyond memory. In: *Proceedings of the 21st ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. pp. 26:1–26:12. PPoPP ’16, ACM, New York, NY, USA (2016)
25. Petersen, K., Spreitzer, M.J., Terry, D.B., Theimer, M.M., Demers, A.J.: Flexible update propagation for weakly consistent replication. *SIGOPS Oper. Syst. Rev.* **31**(5), 288–301 (Oct 1997). <https://doi.org/10.1145/269005.266711>, <http://doi.acm.org/10.1145/269005.266711>
26. Preguiça, N., Zawirski, M., Bieniussa, A., Duarte, S., Balesgas, V., Baquero, C., Shapiro, M.: Swiftcloud: Fault-tolerant geo-replication integrated all the way to the client machine. In: *Proceedings of the 2014 IEEE 33rd International Symposium on Reliable Distributed Systems Workshops*. pp. 30–33. SRDSW ’14, IEEE Computer Society, Washington, DC, USA (2014). <https://doi.org/10.1109/SRDSW.2014.33>, <https://doi.org/10.1109/SRDSW.2014.33>
27. Qadeer, S.: Verifying sequential consistency on shared-memory multiprocessors by model checking. *IEEE Trans. Parallel Distrib. Syst.* **14**(8), 730–741 (2003). <https://doi.org/10.1109/TPDS.2003.1225053>, <https://doi.org/10.1109/TPDS.2003.1225053>
28. Vardi, M.Y.: The complexity of relational query languages (extended abstract). In: *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*. pp. 137–146. STOC ’82, ACM, New York, NY, USA

- (1982). <https://doi.org/10.1145/800070.802186>, <http://doi.acm.org/10.1145/800070.802186>
29. Warren, D.S.: Programming in tabled prolog. In: Symposium Program. p. 62 (1999)
30. Wing, J.M., Gong, C.: Testing and verifying concurrent objects. *J. Parallel Distrib. Comput.* **17**(1-2), 164–182 (1993). <https://doi.org/10.1006/jpdc.1993.1015>, <https://doi.org/10.1006/jpdc.1993.1015>
31. Zennou, R., Atig, M.F., Biswas, R., Bouajjani, A., Enea, C., Erradi, M.: Boosting sequential consistency checking using saturation. In: Hung, D.V., Sokolsky, O. (eds.) *Automated Technology for Verification and Analysis*. pp. 360–376. Springer International Publishing, Cham (2020), [https://doi.org/10.1007/978-3-030-59152-6\\_20](https://doi.org/10.1007/978-3-030-59152-6_20)
32. Zennou, R., Biswas, R., Bouajjani, A., Enea, C., Erradi, M.: Checking causal consistency of distributed databases. In: Atig, M.F., Schwarzmann, A.A. (eds.) *Networked Systems*. pp. 35–51. Springer International Publishing, Cham (2019)
33. Zennou, R., Bouajjani, A., Enea, C., Erradi, M.: Gradual consistency checking. In: Dillig, I., Tasiran, S. (eds.) *Computer Aided Verification*. pp. 267–285. Springer International Publishing, Cham (2019)