Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

B. Awerbuch, L.M. Kirousis, E. Kranakis, P.M.B. Vitanyi

On proving register atomicity

Computer Science/Department of Algorithmics & Architecture Report CS-R8707 May

---

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

On Proving Register Atomicity

Baruch Awerbuch

Massachusetts Institute of Technology
Department of Mathematics and Laboratory for Computer Science
Cambridge, Massachusetts (baruch@theory.lcs.mit.edu)

Lefteris M. Kirousis

University of Patras, Department of Mathematics, Patras, Greece
and
Centre for Mathematics and Computer Science
Amsterdam, The Netherlands (lefteris@cwi.nl)

Evangelos Kranakis

Centre for Mathematics and Computer Science
Amsterdam, The Netherlands (eva@cwi.nl)

Paul M. B. Vitanyi

Centre for Mathematics and Computer Science
Amsterdam, The Netherlands (paulv@cwi.nl)

Concurrent access of shared variables by asynchronous processes does not require mutual exclusion, but can be solved with no waiting. A fruitful paradigm in this context is the notion of a shared register satisfying a niceness condition called atomicity. The model is rigorously presented, and then a method is given for proving register atomicity. It is then used to give simple proofs of the atomicity of two register constructions, the matrix register and the Bloom register, without assuming the existence of a global clock. (The matrix construction shows how to implement an atomic, n-writer, n-reader register with value domain V from n^2 atomic - or even regular - 1-writer, 1-reader registers with value domain $\mathbf{N}\times V$, with $\mathbf{N}$ the set of nonnegative integers. Bloom's construction shows how to implement an atomic, 2-writer, n-reader register with value domain V, from two atomic, 1-writer, $n-1$-reader registers with value domain $\{0,1\}\times V$. These constructions are so simple that they may be even practical.)

1980 Mathematics Subject Classification: 68C05, 68C25, 68A05, 68B20.
CR Categories: B.3.2, B.4.3, D.4.1., D.4.4.
Keywords and Phrases: Register, run, atomic, regular, reader, writer, proof method
Note: This paper is submitted for publication elsewhere.

The work of the first author was supported in part by the Air Force Office of Scientific Research under Contract TNDGAFOSR-86-0078. The work of the fourth author was supported in part by the Office of Naval Research under Contract N00014-85-K-0168, by the Office of Army Research under Contract DAAG29-84-K-0058, by the National Science Foundation under Grant DCR-83-02391, and by the Defence Advanced Research Projects Agency (DARPA) under Contract N00014-83-K-0125.

Report CS-R8707
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

1. Introduction

We are interested in true concurrency in the context of shared register access by asynchronous processors. Concurrency control of asynchronous processes is often realized by actively serializing concurrent actions, using synchronization primitives like mutual exclusion, semaphores, and locking. Thus, although it *seems* that the actions are executed concurrently, in the system they are *actually* executed serially in some order. It has been pointed out in [3] that to implement such primitives we first need interprocess communication through a shared memory unit, which we shall call a *register*, even if the processors communicate by message passing. This suggests that the problem of simultaneous memory access needs to be solved without recourse to synchronisation primitives. It is desired that such a solution involves *no waiting* by one operator for another one. Thus we kill two birds with one stone, since it is the waiting involved in synchronization methods to control the communication between asynchronous participants, which may make such solutions unacceptable. Note, that asynchrony need not be due solely to hardware, but can also be caused by multiple users on the various machines. The problem of providing general wait-free asynchronous communication interfaces becomes more acute, as more and more hardware from different technologies, scale and speed continue to be connected in computer networks and other complexes. The purpose of the present investigation is to examine the feasibility of such general interfaces. In particular, we analyse the problem of how to implement a shared register which can be read by different asynchronous processors (the readers) and be written by different asynchronous processors (the writers) in a truly concurrent fashion. That is, without any restrictions to prevent simultaneous access and making no assumptions, either about the relative durations of the reads and writes, or about the actual timing of the lower level constituent operation executions.

More precisely, we are given some registers with certain restrictions on their mode of operation, e.g. that only a certain number of operators are allowed to access each one of them. We are asked to construct a more powerful (*compound*) register without some of the original restrictions, while retaining some of the positive characteristics of the subregisters, e.g. their serializable mode of operation (otherwise called *atomicity*). These compound registers will comprise a set of registers (i.e. *subregisters*) and an operation execution on the compound register will consist of a sequence of operation executions on the subregisters that follow a given *protocol*. A well-known instance of the above problem is when the subregisters can be read or written by only one operator, and the compound register should be accessible to many readers or writers. In this case, each read (or write) on the compound register will comprise a sequence of reads and writes on the subregisters. Suppose now that the subregisters do have a serial (or atomic) mode of operation. Here we allow that the absence of concurrency in the subregisters is not actual, i.e. the results are *as if* the operations were executed serially. The atomicity of the subregisters does not imply *a priori* that the operation executions on the compound register are serializable. This is a consequence of the fact that each operation on the compound register generally will comprise more than one operation on the subregisters. One way to make the compound register operate atomically is, essentially, to *coerce* it to operate consistently (i.e. in a serializable fashion) via the mentioned synchronization primitives. Serializability in concurrent databases is usually enforced in this way [6]. These methods, however, entail some sort of synchronization among the operators, which makes necessary the slowing down of the fast operators to the pace of the slow ones (see [9] for a detailed exegesis).

Here, we assume complete asynchronicity among the operators, and no waiting. All we require from the constructed protocol is that it guarantees the *existence* of some total (i.e., linear)

order in which the operation executions on the compound register could have taken place (external consistency). This order, in some sense, represents the succession these operations *seemingly* follow. This idea has been successfully applied by Bloom, Lamport, Peterson, and Vitanyi and Awerbuch (see [1], [3], [7], [9], [10]). Of course, for such a total order to be meaningful, it must satisfy certain additional requirements. For example, there should be no second write placed by this order between a read and the write it reads (internal consistency). If we assume that there is a global (i.e. referring to all registers) time-reference system (otherwise, *a global clock*), and if all subactions of an operation execution on the compound register precede in time all the subactions of a second operation execution on the compound register, then this order must place the second operation execution after the first one. In general, we have a relation on the operation executions which is naturally imposed by the problem (e.g., an acyclic relation that tells if an operation execution can have an influence on another), and we desire the existence of a total order that extends this relation, but without violating the above restrictions. If this is possible for each scenario of operation executions of a proposed register, then the register is atomic. Atomic access of compound shared registers is related to version management in distributed systems. E.g., [3], [7] deal with techniques for distributing a shared variable, while [1], [3], [8], [9], [10] and the current paper deal with some aspects of concurrency control for replicated shared variables.

In the next section we present the model rigourously, and give two general atomicity criteria that are suitable for proving register atomicity (the second criterion is a simple variation of the first, but is more suitable for 1-writer registers). We use the general 'causality' model proposed by Lamport [3], specialized towards shared registers and atomicity, and we have no need to assume the existence of a global clock. We also investigate how the assumption of global time affects these criteria, by proving a rather general shrinking function theorem. In Section 3 we prove the atomicity of a multiwriter, multireader compound register constructed from atomic 1-writer, 1-reader subregisters. We also show that the construction is optimal in the number of subregisters used, and the fact that it suffices if the subregisters are 'regular' (regularity is a weaker requirement than atomicity). This 'matrix' register appeared first in [9]. In Section 4 we prove the atomicity of a 2-writer, n-reader compound register constructed from atomic 1-writer, n-reader registers. This 'Bloom' register appeared first in [3]. The atomicity proofs of the matrix and Bloom registers given here, are the first such proofs that are based solely on causality considerations. They are influenced by the proofs in [1], [3], [9], [10].

Finding the 'right' formalism and level of rigour is of major importance in this complicated area. Ideally, as stated in [7], one designs protocols "sufficiently simple that there is no need to provide complicated, and possibly not understandable, formal proofs." Despite their apparent simplicity, the protocols treated here are subtle, and their correctness is by no means obvious. We therefore need rigorous proof. However, excess of formality may prove counterproductive, make proofs incomprehensible and introduce errors. Our goal is to develop *simple* formal criteria and proof methods, and demonstrate their usefulness by proving these nontrivial protocols correct in a *convincing* manner.

2. The Model

A **proto-register** is an abstract data type, capable of holding **values** out of a given **domain** of values. Initially the proto-register is empty. The operations that can be performed on the proto-register are writes and reads. A *write* of a value puts that value in the proto-register. A *read* reports a value from the domain.

We assume that such values have an identity apart from a value. That is, values written by

different write operation executions may have the same value, e.g. 0, but they are not identical. The **identity** $id(v)$ of a value v written by a write w is defined by $id(v)=w$. (Thus, $v=v'$ and $id(v)\neq id(v')$ may be both true.) If a read operation execution reports a value v, then either $id(v)=w$ for a particular write w or else $id(v)$ is undefined. If a read r reports v with $id(v)=w$ for some write w, then we say that r reports the value written by w. The proto-register can be implemented by a **multiset** over a given domain. That is, an unordered list of elements, where the same element can occur more than once. A proto-register has associated with it a finite set of processors called the **writers** and a finite set of processors called the **readers**. A processor can be both a reader and a writer. A write can only be performed by a writer, while a read can only be performed by a reader.

A **sequential** register is a proto-register where all operations are executed in sequence, and an execution of a read operation reports the value written by the execution of the last write operation that precedes it. A sequential register can be implemented by a **linear list**. Originally, the list is empty. A write adds an element to the end of the list, and a read reports the element at the end of the list.

We address the problems arising from true concurrency, where we allow simultaneous operation executions by different processors. However, simultaneous operation executions by the same processor are excluded. Informally, we aim at a specification of a general **concurrent** register, register for short, which corresponds as closely as possible to that of the sequential register. In the atomic register defined below, the operations may be **actually** executed concurrently, yet it will **seem** as if they were executed in sequence. For read operations to report a value which was written by a write operation to the register, there must be causal relations between the operations. We define an 'apparent' precedence relation ($\rightarrow$) on the set of operation executions, which captures the crucial aspect of the causal relations between operation executions to the same register.* To define various degrees of niceness conditions on a register with simultaneous operation executions, we need some formal definitions first. We use 'action' as synonym for 'operation execution'.

A **run** $\rho = (A, \rightarrow, \pi)$ consists of the following:

(R1) A finite or countably infinite set A of **read** and **write** actions. If R is the set of read actions and W is the set of write actions, which were actually performed during the course of the run, then $A=W\cup R$ and $W\cap R=\emptyset$.

(R2) A **reading mapping** which is a partial function $\pi : R \rightarrow W$.

(R3) An irreflexive partial order $\rightarrow$ on the set A of actions.† We call $\rightarrow$ a **precedence relation**.

---

* Lamport [3] defines two precedence relations $\rightarrow$ and $---\rightarrow$, the semantics of which are intended to be problem independent. If a "precedes" relation on the subactions of a, b is defined, then $a ---\rightarrow b$ means "some subaction of a precedes some subaction of b", and $a \rightarrow b$ means "each subaction of a precedes each subaction of b." In the context of shared register access there is always an intended way for the actions to interact, which ensures correctness of the algorithm. Therefore, it is advantageous to reflect this essential causal relation between the actions of a particular algorithm by a single made-to-measure precedence relation. This relation is our $\rightarrow$, not to be confused with Lamport's $\rightarrow$, and will have an algorithm dependent semantics.

† It is not *a priori* obvious that the precedence relation needs to be cycle free. Suppose we have defined $\rightarrow$ such that it is transitive and contains cycles. In our framework this is supposed to have a physical interpretation. Accordingly, in the following discussion, we use some well-known notions from modern physics, in particular the geometry of space-time as in special relativity, as in [4]. Occurrence of events a, b such that $a\rightarrow b$ & $b\rightarrow a$, is intended to have the physical meaning that subevents of a precede subevents of b, and subevents of b precede subevents of a, in space-time. E.g, a and b involve the same point in space at alternating times, first a, then b, and then a again. In contrast, occurrence of events a, b which are not related in $\rightarrow$ at all, i.e., $\neg(a\rightarrow b$ or $b\rightarrow a)$, can mean that no subevent of a pre-

If $a \rightarrow b$ then we say a **precedes** b. To initialize the run, there is an **initial** write that precedes all other actions. We moreover require that, for each $a \in A$, there are only finitely many $b \in A$ such that $\neg(a \rightarrow b)$. Informally, this means that a run begins at some point in time, rather than extending in the infinite past [3], and that an action cannot be infinitely long or infinitely small in duration.

Intuitively, $a \rightarrow b$ will imply that, in the aspect we deem important, a may influence b, but b cannot influence a. Two actions a, b are called **concurrent** if $\neg(a \rightarrow b$ or $b \rightarrow a)$. I.e., if they are incomparable in the relation $\rightarrow$. If w is a write and r is a read , then w **directly precedes** r, if $w \rightarrow r$ and there is no write w', such that $w \rightarrow w' \rightarrow r$.

Irreflexive orders. All **orders** in this paper are irreflexive. For convenience, "total order" and "partial order" will henceforth mean "irreflexive total order" and "irreflexive partial order," respectively.

A run $\rho = (A, \rightarrow, \pi)$ can now be classified into the following categories according to how well it behaves under concurrent operations. The definitions below closely follow the presentation of Lamport [3]. The 'normal' run is a new category we found advantageous to introduce.

1. (**safe**) For each read r, that has no concurrent writes, $\pi(r)$ is defined, and directly precedes r.*
2. (**normal**) For each read r, $\pi(r)$ is defined, and $\pi(r)$ either precedes r or is concurrent with r.
2. (**regular**) For each read r, $\pi(r)$ is defined, and $\pi(r)$ directly precedes r or is concurrent with r. (Hence a regular run is both safe and normal.)
4. (**atomic**) A run is atomic if it is normal and there is a total order $\Rightarrow$, which we call an **atomic precedence** relation, on the set A of actions, as follows.

   (i) if $a \rightarrow b$ then $a \Rightarrow b$ (**external consistency**), and

   (ii) for each read r, $\pi(r)$ is the write directly $\Rightarrow$-preceding r (**internal consistency**).

   We say that $\Rightarrow$ **atomically extends** the precedence relation $\rightarrow$.

Without proof we state the hierarchy involved. Every atomic run is regular, but not every regular run is atomic. By definition, regular runs are exactly the ones which are both safe and normal. There are runs which are safe but not normal, and there are runs which are normal but not safe.

A run is a possible set of operation executions by a register, a possible 'history'. We now tie up the notion of a register and the notion of a run. Intuitively, a register is a deterministic 'black

---

cedes any subevent of b, and no subevent of b precedes any subevent of a, in space-time. E.g., a and b take place at different points in space, and a light ray sent by a to b's point in space arrives after b has finished, while a light ray sent by b to a's point in space arrives after a has finished. Both cases imply concurrency albeit of a different kind. For us, it is not necessary to distinguish between them, due to the later restrictions (R7)-(R9) on the choice of $\rightarrow$. We therefore replace a transitive $\rightarrow$ with cycles by a relation $\rightarrow'$ defined as: $a \rightarrow' b$ iff $(a \rightarrow b)$ & $\neg(b \rightarrow a)$. Then $\rightarrow'$ is an irreflexive partial order. W.l.o.g., therefore, we define $\rightarrow$ from the outset as an irreflexive partial order.

* In the definition of a run we assumed that the reading function π is partial, since it is perfectly possible in general to have a read that returns a value that has not been written by a write. For an example, consider a safe register with a domain of values $\{0,1\}$ where all writes write a 0. If a read overlappes a write, then it is legitimate by our definitions to have this read return the value 1, a value not written by any write. In most cases though, the protocols one designs instruct a read to return the value written by a particular write. Similarly, a read can return a value, actually written by a different write than the one it ostensibly ought to choose, as long as this value is the right one. This requires a more elaborate definition of π than we need here.

box' that reports a value in response to a read query. We can view the function of this black box as associating a reading mapping with a given pair $(A,\rightarrow)$. Since $(A,\rightarrow)$ is a high-level description, it is actually an equivalence class of different finer grained descriptions. These differences may give rise to different responses to the same read query. Therefore, the register associates a *set* of reading mappings with each $(A,\rightarrow)$. Let Π be the set of all its possible reading mappings. Formally, a **register mapping** $REG: \{(A,\rightarrow)\}\rightarrow 2^{\Pi}$ is a total mapping, that associates a nonempty set of reading mappings π with each pair $(A,\rightarrow)$ satisfying (R1), (R2) and (R3). With each register we associate a register mapping. We assume that each processor actually executes operations to the register serially. This assumption is embodied in requirement (R4) below. If K is a register and REG_K is its associated register mapping, then a run $\rho=(A,\rightarrow,\pi)$ of K satisfies

(R4) if a and b are different actions by the same processor then either $a\rightarrow b$ or $b\rightarrow a$, and this total $\rightarrow$-order on the actions by the same processor is identical with the serial order in which a processor executes its actions in A;

(R5) $\pi \in REG_K(A,\rightarrow)$; and

(R6) if a read r returns a value v and $id(v)=w$, then $\pi(r)=w$.

A register is **atomic** (respectively **regular, normal, safe**) if each of its runs is atomic (respectively regular, normal, safe). Obviously, the atomic register is the ideal register; the operations may be concurrent, yet they seem to be executed in a serial fashion, extending the given precedence relation (external consistency) and consistent with the reading function (internal consistency). The notion of register atomicity is closely related to what is called '(strict) serializability' in conventional concurrency control, in particular in the context of databases with concurrent 'transactions'.* See e.g. [6]. Given a particular implementation of a data structure, and having selected the particular precedence relation $\rightarrow$ we wish to employ,† it is often simple to check whether it is a safe, a normal, a regular register or none of these. However, proving atomicity using the given definition, or its 'shrinking' variant which we will meet below, turns out to be a difficult matter. Therefore, in the next section we propose simple criteria which are necessary and sufficient for atomicity. In later sections we show how to use these atomicity

---

* Concurrent transactions are usually called atomic if they are both serializable and **recoverable**. Recoverability means that each transaction appears all-or-nothing: either it executes to completion (in which case we say that it **commits**) or it cannot influence other transactions (in which case we say that it **aborts**). Recoverability is a problem only in the presence of failures. We assume that registers are failure-free, so we do not consider recoverability.

† In our definition, (R1) through (R6) do not preclude that $\rightarrow$ is empty, apart from the total $\rightarrow$-order on the set of actions by each processor (R4). The sets of actions by different processors are necessarily disjoint. Suppose we construct a register and choose $\rightarrow$ such that, for each associated run $\rho=(A,\rightarrow,\pi)$, we have:

(i) if a,b are actions by different processors, then $\neg(a\rightarrow b$ or $b\rightarrow a)$, and

(ii) if r is a read by processor p, then $\pi(r)$ is a write by processor q, $q\neq p$.

Then this register is vacuously atomic. This example shows clearly, that the significance of the properties of a register is derived from the meaningfulness of our precedence relation ($\rightarrow$). In section 2.2 we show that this abstract approach suffices to cover the conventional global time semantics by a 'natural' choice for $\rightarrow$. In [3], Lamport proceeds differently. As already noted in a previous footnote, the two precedence relations $\rightarrow$ and $---\rightarrow$ distinguished by him are not chosen freely. Ultimately, his precedence relations have a physical foundation with its roots in the notion of causality. The intuitive concept of register implies the possibility of communication by means of read and write operations to the register. But this requires some causality ($---\rightarrow$) relations between reads and writes of the register. For 1-writer registers, Lamport restricts registerhood to constructions satisfying his axiom B1: "for any read r and write w to the same register, either $r---\rightarrow w$ or $w---\rightarrow r$ (or both)". The causality relation $---\rightarrow$ refers to subactions of the two actions involved. Therefore, we defer the introduction of our version of B1 to the discussion of the compound register, where it appears as (R9).

criteria for verifying that proposed constructions implement atomic registers.

2.1. Atomicity Criteria

For a proposed data structure K to be an atomic register, it suffices to prove that each of its runs, as defined in (R1) through (R6), is atomic. Let $\rho = (A, \rightarrow, \pi)$ be a normal run of K, so π is total. We divide the set of actions A into equivalence classes induced by π. Each such equivalence class, called a clan, is associated with a write. The **clan** associated with a write w is the set $[w] = \{w\} \cup \{r \in R : \pi(r) = w\}$. For any two writes w, w' define $[w] \rightarrow^{\pi} [w']$ if and only if $w \neq w'$ and there exist actions $a \in [w]$ and $a' \in [w']$ such that $a \rightarrow a'$. Note that $\rightarrow^{\pi}$ is not necessarily acyclic. The following theorem is basic for proving the atomicity of runs.

Theorem 2.1. (Atomicity Criterion) *Let $\rho = (A, \rightarrow, \pi)$ be a run. The following statements are equivalent:*
(1) ρ is atomic.
(2) ρ is normal and $\rightarrow^{\pi}$ is acyclic.

Proof. *(1) implies (2).* Let ρ be atomic. By definition, atomicity implies normality, which shows the first part of (2). To show the second part of (2), let $\Rightarrow$ be a total order that atomically extends $\rightarrow$. I.e., for each read r, we have

(i) $\pi(r) \Rightarrow r$, and

(ii) there is no write w with $\pi(r) \Rightarrow w \Rightarrow r$.

We prove that $\rightarrow^{\pi}$ is extendible to a total order, which implies acyclicity of $\rightarrow^{\pi}$. It is enough to show that for any two writes w, w', if $[w] \rightarrow^{\pi} [w']$ then $w \Rightarrow w'$.

Since $\Rightarrow$ is a total order, the negation of $w \Rightarrow w'$ is equivalent to $w' \Rightarrow w$. Therefore, we only need to show that for any two writes w, w', if $w' \Rightarrow w$ then $\neg([w] \rightarrow^{\pi} [w'])$. Thus, suppose $w' \Rightarrow w$. Exhaustive analysis of all cases shows that then the combination of (i) and (ii) implies $w' \Rightarrow r' \Rightarrow w \Rightarrow r$, for all reads $r \in [w]$ and $r' \in [w']$. Hence, there are no $a \in [w]$ and $a' \in [w']$ such that $a \rightarrow a'$. Therefore, $\neg([w] \rightarrow^{\pi} [w'])$.

(2) implies (1). Assume (2) holds. It is clear that the transitive closure of $\rightarrow^{\pi}$ is a partial order, which in turn can be extended to a total order $\Rightarrow^{\pi}$. Since ρ is normal, for each read $r \in [w]$, we have $\neg(r \rightarrow w)$. Hence, there is a total order $\Rightarrow_{[w]}$ on each $[w]$ atomically extending $\rightarrow$ and such that $w \Rightarrow_{[w]} r$, for each read $r \in [w]$. Define a unique relation $\Rightarrow$ on the set A as follows. For all $a, a' \in A$, $a \Rightarrow a'$ if and only if either

(i) $a, a' \in [w]$ and $a \Rightarrow_{[w]} a'$, or

(ii) $a \in [w]$, $a' \in [w']$, and $[w] \Rightarrow^{\pi} [w']$.

Clearly, $\Rightarrow$ is a total order atomically extending $\rightarrow$. It follows that ρ is atomic. •

The second atomicity theorem refers to registers with only one writer. In this case, for each pair of different writes $w, w' \in A$, either $w \rightarrow w'$ or $w' \rightarrow w$, by (R4). The theorem is similar to a corresponding theorem in Lamport [3].

Theorem 2.2. (1-writer Atomicity Criterion) *Assume that K is a register with only one writer. Then for each run $\rho = (A, \rightarrow, \pi)$ of K the following statements are equivalent:*
(1) ρ is atomic.
(2) ρ is regular and π is weakly monotonic (i.e., if $r \rightarrow r'$, then either $\pi(r) \rightarrow \pi(r')$ or $\pi(r) = \pi(r')$).

Proof. *(1) implies (2).* Let ρ be atomic. Atomicity implies regularity. Therefore we only

need to prove weak monotonicity of π.

Let $r, r' \in A$ be different reads with $r \rightarrow r'$. Since there is only one writer, we have by (R4) that either $\pi(r) \rightarrow \pi(r')$ or $\pi(r) = \pi(r')$ or $\pi(r') \rightarrow \pi(r)$. Atomically extend $\rightarrow$ to a total order $\Rightarrow$, as in the definition of an atomic run. Exhaustive case analysis shows that, by the properties of $\Rightarrow$, either $\pi(r) \Rightarrow r \Rightarrow \pi(r') \Rightarrow r'$, or $\pi(r) = \pi(r')$.

(2) implies (1). Let ρ be regular and π be weakly monotonic. By Theorem 2.1, if we prove that $\rightarrow^{\pi}$ is acyclic, then we are done. Assume to the contrary, there is a cycle

$$[w] \rightarrow^{\pi} [w'] \rightarrow^{\pi} \cdots \rightarrow^{\pi} [w] .$$

Since $[w] \rightarrow^{\pi} [w']$, there are $a \in [w]$ and $a' \in [w']$, $w \neq w'$, such that $a \rightarrow a'$. If both a and a' are writes then $w \rightarrow w'$. If a, a' are both reads, then by weak monotonicity of π, we have $w \rightarrow w'$. If a is a read and $a' = w'$, then $a \rightarrow w' \rightarrow w$ $(= \pi(a))$ contradicts normality of ρ. Therefore $w \rightarrow w'$ by (R4). If $a = w$ and a' is a read, then $w' \rightarrow w \rightarrow a'$ $(\pi(a') = w')$ contradicts safety of ρ, and therefore $w \rightarrow w'$ by (R4) again. Hence, $[w] \rightarrow^{\pi} [w']$ implies $w \rightarrow w'$. Since this argument holds for all pairs of adjacent clans in the cycle, we obtain a cycle $w \rightarrow w' \rightarrow \cdots \rightarrow w$. This contradicts that $\rightarrow$ is a partial order. •

2.2. Global Time, Intervals and Shrinking

In [1], [7], [8], [9], atomicity is related to the assumption of a global time reference frame (also called global clock). We show that the theory as developed here, along the lines of [3], is more general. In particular, a register is atomic in global time if and only if it is atomic for a particular choice of the $\rightarrow$ precedence relation. This precedence relation turns out to be the interval order induced by the time intervals representing the actions. The notions developed and results proved in the rest of the present section are not needed for the results of the remainder of the paper.

An important aspect of atomic runs is the following property. Although their actions have a duration on a global time scale, and such durations may overlap, each action may be considered to take place instantaneously, i.e., as if it happened completely at a particular time instant. If all of these time instants are distinct, then the apparent time instants of the actions orders the actions totally. This relates the order approach to atomicity with the global time approach.

Time is represented by the set of real numbers $\mathbf{R}$, ordered as usual. Assume that every action a is **represented** by an open time interval $(s(a), f(a))$, $s(a) < f(a)$, within the bounds of which the action is supposed to have taken place. $s(a)$ (respectively, $f(a)$) is a real number called the **starting** (respectively, **finishing**) time of the action a.*

Define the precedence relation $\rightarrow$, as the natural relation $a \rightarrow b$ iff $f(a) \leq s(b)$. A relation which is so induced by a set of intervals of the real line $\mathbf{R}$, satisfies the axioms of a special type of partial order called interval order. Formally, an **interval order** on a set A is an irreflexive relation $\rightarrow$ that satisfies

---

* To exclude some technical difficulties, there is usually an assumption that $s(a) \neq s(b)$, $s(a) \neq f(b)$ and $f(a) \neq f(b)$, for any two distinct actions $a, b \in A$, and $s(a) \neq f(a)$ for each action $a \in A$. The fact that we should be allowed to assume that no two starting or finishing times are equal, is justified by appeal to the sensibility of natural law [3]. "No physical meaningful result could depend on upon completely accurate knowledge of these times. (It makes no physical sense to specify starting and finishing times of an operation execution down to the fraction of a micropicosecond.)" By excluding the starting and finishing times from the duration associated with action a, to obtain the desired effect in the mathematical framework, we may come closer to the spirit of physics. Thus, we choose to represent durations of actions as open intervals.

$$a \rightarrow b \ \& \ c \rightarrow d \text{ implies } a \rightarrow d \text{ or } c \rightarrow b, \text{ for all } a,b,c,d \in A \text{ (see [2])}. \qquad (2.1)$$

Every interval order is a partial order, and hence the previously developed theory applies. Since not every (irreflexive) partial order is an interval order, the global time approach requires more from the precedence relation ($\rightarrow$) than the general approach in (R3). We now proceed with the definitions of this more conventional global time approach to atomicity. The relation between the two approaches is analysed in Theorem 2.3.

A **shrinking function** on the set of actions of a run $\rho=(A,\rightarrow,\pi)$, with $\rightarrow$ the interval order induced by the set of intervals $\{(s(a),f(a))\subseteq\mathbf{R}: a\in A\}$, is a one-to-one function σ that associates with each action a of the run a time instant (i.e., a real number) $\sigma(a)$ such that:

(S1) $\sigma(a)$ belongs to the interval $(s(a),f(a))$ of a.

A shrinking function gives a possible serialization of the actions. Condition (S1) enforces external consistency of the serialization. In the order approach, external consistency follows from the fact that the serialisation extends $\rightarrow$. Define the precedence relation $\rightarrow_\sigma$, induced by σ, as $a\rightarrow_\sigma b$ iff $\sigma(a)<\sigma(b)$. Obviously, $\rightarrow_\sigma$ is a total order on A. Then (S1) implies that $\rightarrow_\sigma$ extends $\rightarrow$. A shrinking function σ is consistent with the reading mapping π if

(S2) $(A,\rightarrow_\sigma,\pi)$ is atomic.

A run ρ is **shrinking** atomic if there is a shrinking function σ such that (S1) and (S2) are satisfied.

Theorem 2.3. (Shrinking Function Theorem) *Let $\rho=(A,\rightarrow,\pi)$ be a run, and let $\rightarrow$ be the interval order induced by a representation of open (time) intervals of the actions in A. The following statements are equivalent:*
(1) ρ is atomic, and
(2) ρ is shrinking atomic.

Proof. *(1) implies (2).* Suppose (1) holds. Let $\Rightarrow$ be a total order which atomically extends $\rightarrow$. Define $Q(a)=\{b:\neg(a\rightarrow b)\ \&\ \neg(b\Rightarrow a)\}$. Note that, by (R3), $Q(a)$ is finite, and that $Q(a)$ is nonempty since $a\in Q(a)$. Define, by induction on a, $\sigma(a)$ to be a real number such that:

(i) if $b\Rightarrow a$ then $\sigma(a)>\sigma(b)$,

(ii) $\sigma(a)>s(a)$, and

(iii) $\sigma(a)<\mu$, with $\mu=\min\{f(b): b\in Q(a)\}$.

Note that (ii) and (iii) imply (S1), and (i) implies (S2). Induction is possible if:

(a) if $b\rightarrow a$ then $\sigma(b)<\mu$, and

(b) $s(a)<\mu$.

Let $b_{\min}\in Q(a)$ be an action such that $\mu=f(b_{\min})$.

Ad (a). Assume $b\rightarrow a$. If $b\rightarrow b_{\min}$ then $\sigma(b)<f(b)\leq s(b_{\min})<f(b_{\min})=\mu$.
If $\neg(b\rightarrow b_{\min})\ \&\ \neg(b_{\min}\rightarrow b)$ then, since $b\rightarrow a$, we have $\neg(b_{\min}\Rightarrow b)$. Therefore, $b_{\min}\in Q(b)$. Then, $\sigma(b)<\min\{f(c): c\in Q(b)\}\leq f(b_{\min})=\mu$.
If $b_{\min}\rightarrow b$ then $b_{\min}\rightarrow b\rightarrow a$, contradicting $b_{\min}\in Q(a)$.

Ad (b). Since $\neg(a\rightarrow b_{\min})\ \&\ \neg(b_{\min}\Rightarrow a)$ we have $s(a)<f(b_{\min})=\mu$.

(2) implies (1). Assume (2). Since $\rightarrow_\sigma$ is a total order extending $\rightarrow$ and satisfying (S2), atomicity of ρ is immediate. ∎

2.3. Compound Register

The most obvious approach to constructing a register is to build it from simpler ones. The existence of such a simpler register is either postulated, or it is constructed from still simpler registers. More precisely, a **compound register** consists of a finite number of registers, called **subregisters**. The set of readers and writers of the compound register is the union of the set of readers and writers of the subregisters. The subregisters are allowed to hold a value out of a given domain. We can distinguish essentially two cases. In one case the subregisters are simpler than the compound register in that their domain of values is smaller than the value domain of the compound register. Then the construction for the compound register **distributes** the value to be stored piecemeal over the subregisters. E.g., a positive integer n can be distributed in $\log n$ bits over $\log n$ boolean subregisters. In the other case the set of readers and writers associated with the compound register is larger than the set of readers and writers associated with each subregister. Then the construction for the compound register **replicates** the value to be stored as versions in several subregisters. A reader has to determine the 'latest' version among the versions it obtains from different subregisters. To make this possible, extra information such as a 'timestamp' is attached to each version. As a result, the value domain of each subregister has to be larger than the value domain of the compound register. The constructions of compound registers in this paper are of the latter type. To express their complexity we use the following cost measures. Let V be the value domain of the compound register, $v=|V|$, and let there be n readers and m writers associated with the compound register. Let $S,T: V\times\mathbb{N}\times\mathbb{N}\to\mathbb{N}$ be total cost functions, with $\mathbb{N}$ the set of nonnegative integers. Let the value domain of each subregister of the compound register be (isomorphically) contained in $TAG\times V$, with $|TAG| = S(v,n,m)$, the number of elements in TAG. Then the **space** complexity of the compound register is $\log S(v,n,m)$. The processors execute read or write actions on the compound register, independently of each other but following a protocol. Let each read or write action by a given processor on the compound register consist of at most $T(v,n,m)$ read and/or write actions on the subregisters. Then the **time** complexity of the compound register is $T(v,n,m)$. An action on the compound register is considered to be a higher-level operation execution of the same nature as its subactions. Thus, with each run of the compound register is associated a run of each subregister which constitutes the compound register. This means that we associate with each subregister a set of subactions related by a precedence relation. Sets of subactions associated with different subregisters are disjoint. We assume that a processor actually executes all its subactions in serial order. The disjoint precedence relations of the subactions on respective subregisters are related by the order in which each processor executes its subactions. For the compound register we define a transitive precedence relation ($-\!\!\gg$) on the set of all subactions involved, as follows.

Let K be a compound register comprising subregisters $K_1, \ldots, K_n$. Let $\rho=(A,\to,\pi)$ be a run of K and let $\rho_i=(A_i,\to_i,\pi_i)$ be the associated run of subregister K_i, $1\le i\le n$. The **precedence** relation $-\!\!\gg$ on the set $\bigcup_{i=1}^{n} A_i$, is defined as the minimal transitive relation that extends all precedence relations $\to_i$, $1\le i\le n$, such that

(R7) if α and β are different subactions by the same processor, then either $\alpha-\!\!\gg\beta$ or $\beta-\!\!\gg\alpha$, but not both, and this total $-\!\!\gg$-order on the subactions by the same processor is identical with the actual serial order in which the processor executes these subactions; and

(R8) if $a,b\in A$ and for each subaction α of a and each subaction β of b holds $\alpha-\!\!\gg\beta$, then $a\overset{\pi}{\to}b$.

Lemma 2.4. *$-\!\!\gg$ is a partial order.*

Proof. Clearly, (R7) precludes $-\gg$-cycles containing two subactions by the same processor. Therefore, since the sets of subactions on the same subregisters are disjoint, any $-\gg$-cycle contains only subactions on the same subregister. But these subactions are partially ordered, which contradicts such a $-\gg$-cycle. •

Finally, we need to express the 'registerhood' of the compound by suitably restricting the choice of $\rightarrow$. That this is necessary can be seen from the following example. Let K be a compound register consisting of subregisters K_1,K_2. Let p be a writer associated with subregister K_1, and let q be a reader associated with subregister K_2, $p \neq q$. Then there is no way that q can read what p has written. Yet runs of K can satisfy (R1) through (R8) and even be atomic. For example, atomicity of K_1,K_2 implies atomicity of K. Such anomalies are due to the fact that we have not yet required the existence of causal relations between actions by different processors. There must be some causal relation between a write and a read, since otherwise a reader cannot report what a writer wrote. There must be some causal relation between two writes, because otherwise a writer cannot replace the value in the register by the value it wants to write. However, it is not necessary to have a causal relation between two reads; this is because neither do reads have to change the value contained by the register, nor do they need to report what the another read wrote. The following condition expresses these requirements on the compound register in terms of subregisters. Assuming the general setting above:

(R9) if $a,b \in A$ are not both reads, then there are subactions α of a and β of b, α,β are not both subreads, and some i $(1 \leq i \leq n)$, such that $\alpha,\beta \in A_i$. (α and β act on the same subregister K_i.)

It follows that a choice of $\rightarrow$ satisfying (R9) is 'proper' if the choices of the $\rightarrow_i$'s are 'proper.' This can be argued as follows. Assume that the ultimate subsub..subregister is atomic. If a,b are not both reads, then there are subactions α of a and β of b, not both reads, which act on the same subregister, and so on. At the atomic subsub..subregister level the subsub..subactions involved have an apparent total order. Choose this as the precedence relation. For convenience, let the K_i's be the basic atomic subregisters, so the $\rightarrow_i$'s are total orders. Then either $\alpha \rightarrow_i \beta$ or $\beta \rightarrow_i \alpha$, but not both, by (R7). Suppose $\alpha \rightarrow_i \beta$. If $c,d \in A$, $c \rightarrow a$ and $b \rightarrow d$, then by (R8) we have $c \rightarrow d$. Suppose $\beta \rightarrow_i \alpha$. If $c,d \in A$, $a \rightarrow c$ and $d \rightarrow b$, then by (R8) we have $d \rightarrow c$. Using the precedence relations at the previous level, we induce in this fashion a 'coarse' precedence relation at each next higher level compound register. Our choice of $\rightarrow$ is constrained to be an extension of this coarse precedence relation. That is, (R7) through (R9) restrict the freedom of our choice of $\rightarrow$ appropriately, by ultimately reducing the constraints on our choice of precedence $\rightarrow$ to precedence at the elemental level.

2.4. Naming of Registers

Unfortunately, the naming conventions for types of registers are inconsistent. For a 1-writer register, the operator who writes can simply remember the value it wrote last. Therefore, the name '1-writer, 1-reader' register is used for a register that can be read by both writer and reader [3]. By analogy, we use '1-writer, $(n-1)$-reader' register for a register that can be written by one writer and read by $n-1$ readers that cannot write. The writer can always read as above. However, in an m-writer register, with $m>1$, while a writer can remember what it wrote last, this value can have been overwritten by a later write of another writer. Hence, here we might as well have writers that cannot read (in addition to readers that cannot write). We will, however, only consider registers where the writers can also read. For us, an 'm-writer, n-reader' register, $m>1$, designates a register that can be written by m processors, and read by n processors including the m writers $(n \geq m)$.

3. The Matrix Register

The matrix register is a compound n-writer, n-reader register constructed as a matrix of atomic 1-writer, 1-reader subregisters. The domain of values of the subregisters is the cartesian product of the domain of values of the compound register with the nonnegative integers. The matrix register was the first general atomic multiwriter register, and was initially defined in [9]. It may be a register of practical importance, because of its simplicity, elegance and low complexity. Its atomicity was derived in [9] by proving that four special conditions were satisfied, using global time. The present proof is an application of the atomicity Theorem (Theorem 2.1). Shrinking atomicity is an immediate corollary, by Theorem 2.3.

Architecture. Let $p_1,...,p_n$ be n processors and let K be an $n \times n$ matrix register consisting of n^2 atomic, 1-reader, 1-writer registers $K_{i,j}$, $i,j = 1,...,n$. Each p_i is a writer of (i.e., is connected to the write terminal of) each $K_{i,j}$. Each p_i is also a reader (i.e., is connected to the read terminal) of each $K_{j,i}$. Let V be the domain of values of the compound register. Then $\mathbf{N} \times \{1, \ldots, n\} \times V$, with $\mathbf{N}$ the nonnegative integers, is the domain of values of each subregister. A **tag** is a pair (k,i), where k is a nonnegative integer and $i \in \{1,...,n\}$. We say that each subregister can hold a tag, next to a value from the domain V of the compound register. All subregisters are initialized with tag (0,1) and value 0. Moreover, each run of the compound register starts with a write action, which precedes all other actions, as required by (R3). The architecture is depicted in Figure 1.

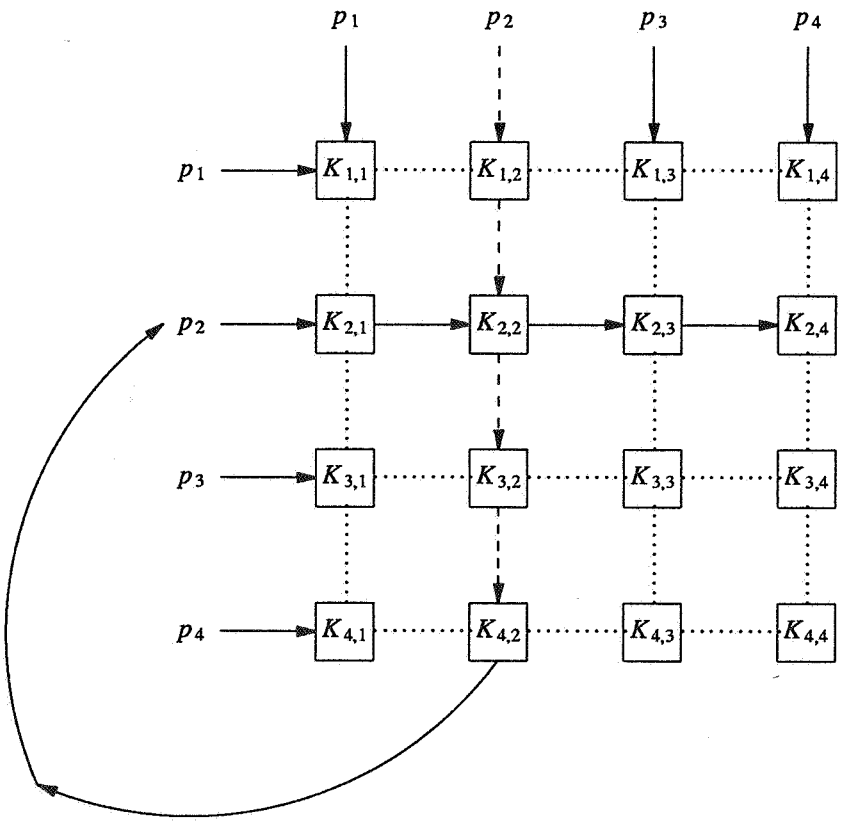

Figure 1: An action by processor p_2 in the 4-reader, 4-writer, matrix register.

Protocol. The register K obeys the following protocol.

p_i writes the value v:

1. for all $j = 1,...,n$ read $K_{j,i}$ (i.e., read the i th column);
2. determine the lexicographically largest tag (k_{max},m);
3. set own tag to $(k_{max}+1,i)$;
4. for all $j = 1,...,n$ write on $K_{i,j}$ (i.e., write to the i th row) the new tag, as well as the value v.

p_i reads:

1. for all $j = 1,...,n$ read $K_{j,i}$ (i.e., read the i th column);
2. determine the lexicographically largest tag (k_{max},m) and let v_m be the value contained in a register with such a tag;
3. set own tag to (k_{max},m);
4. for all $j = 1,...,n$ write to $K_{i,j}$ (i.e., write to the i th row) the new tag, as well as the value v_m, which was determined in 2. (Also, report v_m.)

Each action a by processor p_i consists of a set of subreads $R(a,1,i),\ldots,R(a,n,i)$ followed by a set of subwrites $W(a,i,1),\ldots,W(a,i,n)$, where the last two indices i,j indicate the subregister $K_{i,j}$ on which the subaction took place. The order in which these subreads and subwrites take place is arbitrary, but for the fact that each subread precedes each subwrite. Each subregister $K_{i,j}$ $(1\le i,j\le n)$ of K is atomic. Let $\rho=(A,\rightarrow,\pi)$ be a run of K. Let A^i be the subset of actions in A that are executed by p_i $(1\le i\le n)$. Define, for all $1\le i,j\le n$, $\rho_{i,j}=(A_{i,j},\rightarrow_{i,j},\pi_{i,j})$, the run of $K_{i,j}$ associated with ρ, where

$$A_{i,j} = R_{i,j}\cup W_{i,j},$$

$$R_{i,j}=\{R(a,i,j): a\in A^j\},$$

$$W_{i,j}=\{W(a,i,j): a\in A^i\};$$

$R_{i,j}$ is the set of subreads, and $W_{i,j}$ is the set of subwrites on subregister $K_{i,j}$. Since $K_{i,j}$ is atomic, there is an atomic extension $\Rightarrow_{i,j}$ of $\rightarrow_{i,j}$, for all $1\le i,j\le n$. This atomic extension is a total order on the subactions executed on the subregister concerned. Moreover, if a subread reads a subwrite (on a subregister), then there is no other subwrite placed between them by this order. The orders on the disjoint sets of subactions associated with each subregister are related by the orders on the disjoint sets of subactions by each processor. Let $-\!\gg$ be the minimal transitive relation on the subactions in $\bigcup_{i,j=1}^{n} A_{i,j}$ extending the $\Rightarrow_{i,j}$'s and satisfying (R7). By (R7), the subactions by the same processor p are totally ordered by $-\!\gg$. This order is the serial execution order of the subactions by p. In particular, $-\!\gg$ must satisfy:

$$R(a,i,j) -\!\gg W(a,j,k), \tag{3.1}$$

for all $a\in A^j$ and all $1\le i,j,k\le n$. We now define a precedence relation $\rightarrow$ on A. For any two actions a and b on the compound register K, by p_i and p_j, respectively, let $\rightarrow$ be the transitive closure of $\rightarrow'$:

$$a \rightarrow' b \text{ iff } W(a,i,j) -\!\gg R(b,i,j). \tag{3.2}$$

Clearly, this satisfies (R8) and (R9).

Lemma 3.1. *$\rightarrow$ is a partial order on A.*

Proof. Existence of a $\rightarrow$-cycle containing $a \in A^j$, implies $W(a,j,k) -\!\!\gg R(a,i,j)$, for some k,i $(1 \leq k,i \leq n)$. This contradicts (3.1), since $-\!\!\gg$ is a partial order by Lemma 2.4. •

Remark. If $-\!\!\gg$ extends the original partial orders $\rightarrow_{i,j}$, instead of the apparent total orders $\Rightarrow_{i,j}$, then Lemma 3.1 still holds for the $\rightarrow$ resulting from (3.2). This will be useful in the proof of Theorem 3.5.

The following theorem is the main result of this section.

Theorem 3.2. *The matrix register K is an atomic, n-writer, n-reader compound register, which is implemented with n^2 atomic, 1-writer, 1-reader registers.*

Proof.

Let $\rho = (A, \rightarrow, \pi)$ be a run of K. Examine the write protocol. For a write $w \in A$, let $v(w)$ be the value written by w to the compound register, and, if $t(w)$ denotes the tag determined in step 3 of w, let $(t(w), v(w))$ be the value written to the subregisters in step 4 of w's execution.* For a read $r \in A$, let $v(r)$ be the value reported by r from the compound register, and, if $t(r)$ is the tag associated with $v(r)$, let $(t(r), v(r))$ be the value written to the subregisters in step 4 of its execution. The pair $(t(r), v(r))$ is selected in step 3 of the read protocol.

Claim. For each read r, there is a write w, such that

$$t(r) = t(w) \;\&\; id(v(r)) = id(v(w)) \;(=w). \tag{3.3}$$

If $id(v(r)) = w$ then $\pi(r) = w$. Hence, π is total.

Proof of Claim. The subregisters are initialized with tag (0,1), and there is a write preceding all other actions, by (R3). Hence, there is an $a \in A$, such that $(t(r), v(r)) = (t(a), v(a))$. If a is a write then we are done, else a is a read and we repeat the argument. If $r \rightarrow a$, then, by (3.2) and atomicity of the subregisters, r can not read the value written by a to the subregister involved. There are only finitely many a, such that $\neg(r \rightarrow a)$, $\rightarrow$ is a partial order, and there is an initial write preceding all other actions, by (R3). Hence, we need only finitely many repetititions of the argument before we find a write w such that (3.3) holds. If $id(v(r)) = w$ then $\pi(r) = w$ by (R6). Since this holds for each read r, π is total. This proves the Claim.

Let $<_{lx}$ be the irreflexive **lexicographic** order on pairs of integers. If $a, b \in A$ such that $a \rightarrow b$ (therefore $a \neq b$), then it follows by (3.2) and the choosing of the new tag in step 3 of the write and read protocols, that:

$$t(a) \leq_{lx} t(b) \quad (t(a) <_{lx} t(b) \text{ if } b \in W). \tag{3.4}$$

The prove atomicity, by Theorem 2.1, we only need to prove that ρ is normal and that there is a total order extending the $\rightarrow^{\pi}$ relation among the clans. Intuitively, we proceed by *first* choosing a plausible total order on the set of writes, and *next* showing that the corresponding total order on the set of clans extends $\rightarrow^{\pi}$. In the matrix register, the obvious total order on the set of writes is the lexicographical order of the associated tags. Proceeding this way, the conclusion of the theorem follows from Theorem 2.1 and by the following lemma.

Lemma 3.3.

---

* In the next section on Bloom's algorithm, it is useful to distinguish between $t(a)$, the *identity* of a tag, and $val(t(a))$, the *value* of a tag. In this algorithm, however, $val(t(a)) = val(t(b))$ iff $t(a) = t(b)$, so we do not want to load the notation unnecessarily.

(1) ρ *is normal, and*
(2) if $[w] \rightarrow^{\pi} [w']$, *then* $t(w) <_{lx} t(w')$. *In particular,* $\rightarrow^{\pi}$ *is acyclic.*

Proof. *(1).* By the Claim above, π is a total function. Let $\pi(r)=w$ (i.e., $r \in [w]$). Then, by (3.3), $t(r)=t(w)$. However, if $r \rightarrow w$, then by (3.4) we have $t(w) >_{lx} t(r)$, which is a contradiction. Hence, $\neg(r \rightarrow w)$, i.e., ρ is normal.

(2). Let $[w] \rightarrow^{\pi} [w']$. By definition of $\rightarrow^{\pi}$, there exist actions $a \in [w]$ and $b \in [w']$ such that $a \rightarrow b$.

Suppose $b = w'$. Then by (3.3) and (3.4) it follows that $t(w') >_{lx} t(w)$, which is as claimed.

Suppose that $a = w$ and b is a read. By (3.3) and (3.4), $t(w) \leq_{lx} t(b) = t(w')$. If w, w' are writes by different processors, then their tags have different processor numbers; if they are writes by the same processor then, since $w \neq w'$, one of them $\rightarrow$-precedes the other. Therefore, by (3.4), they must have different tags. In both cases, $t(w) \neq t(w')$, which is as claimed.

Suppose both a, b are reads. Then $t(a) = t(w) \leq_{lx} t(b) = t(w')$, by (3.3) and (3.4). The proof of $t(w) \neq t(w')$ is now exactly as before. This proves the lemma.

The proof of the theorem is finished. •

Corollary. The matrix register is shrinking atomic.

Proof sketch. The argument goes as follows. Assume global time. The interval representations of the subactions induce the $\rightarrow_{i,j}$ precedence relations on the subregisters. Each such relation is therefore an interval order. The intervals associated with the subactions of each processor are linearly ordered (do not overlap) by definition. Since each subregister $K_{i,j}$ is atomic, each run $\rho_{i,j}=(A_{i,j}, \rightarrow_{i,j}, \pi_{i,j})$ has a shrinking function $\sigma_{i,j}$ such that $(A_{i,j}, \rightarrow_{\sigma_{i,j}}, \pi_{i,j})$ is shrinking atomic, by Theorem 2.3. Since the associated intervals are open, we can always choose the $\sigma_{i,j}$'s such that $\sigma' = \bigcup_{i,j=1}^{n} \sigma_{i,j}$ is one-to-one. Define $-\gg$ as the total order of the real images of the subactions under σ', i.e., $-\gg$ agrees with the usual total order $<$ on the reals. Then $-\gg$ satisfies (R7) and (3.1). Define $a \rightarrow'' b$ iff $\sigma'(\alpha) < \sigma'(\beta)$ for all subactions α of a and β of b. Then $\rightarrow''$ is an interval order. This satisfies (R8) and (R9), and $\rightarrow$ is a refinement of $\rightarrow''$. The proof of Theorem 3.2 goes through exactly as before, with interval order $\rightarrow''$ instead of $\rightarrow$, which implies that register K is shrinking atomic by the Shrinking Function theorem (Theorem 2.3). •

3.1. Complexity and Optimality

The time complexity of the matrix register is $2n$ (or rather $2n-2$, as follows from Theorem 3.4 below) which seems to be as low as it can possibly be.* The space complexity of the matrix register is unbounded. In theory this is pretty bad. In practice, however, this solution uses far less space than many solutions which theoretically do better. For instance, in [9] a solution has been proposed where the space complexity of the compound register is $5n^2 \log n$. However, we can assume that a system executes only a limited number of actions on the compound register in its total lifetime. If we set a generous bound of at most 2^{50} such actions, the matrix solution is superior in terms of space complexity, with respect to the mentioned bounded space solution, for any number $n \geq 3$ of associated processors. Thus the matrix register has effectively a lower space complexity than comparable solutions with bounded space complexity, even for solutions which solve only subproblems of the one addressed by the matrix solution. An exception is the Bloom

* Cf. related lower bounds on distributed match-making in [5].

register in the next section (with only two writers), which both effectively and theoretically cannot be improved in space and time complexity.

Another complexity criterion is the number of subregisters of a certain type used in the compound register. Leaving out the subregisters on the main diagonal, which are redundant, the matrix solution is optimal in the number of 1-writer, 1-reader subregisters used.

Theorem 3.4. (Optimality) *The implementation of a compound safe n-writer, n-reader register from 1-writer, 1-reader subregisters, requires at least $n(n-1)$ such subregisters (atomic or not). Register K, minus the subregisters on the main diagonal, is such an optimal implementation.*

Proof. Suppose we have implemented a safe compound n-writer, n-reader register R, with associated processors $p_1, \ldots, p_n$, from 1-writer, 1-reader subregisters. For each ordered pair of processors (p_i, p_j), $1 \le i, j \le n$ and $i \ne j$, we can consider a run $(\{w, r\}, \rightarrow, \pi)$ of R, consisting solely of two nonoverlapping operation executions: a write w by p_i, followed by a read r by p_j. Since R is safe, $\pi(r) = w$. Since $i \ne j$, there must be a subregister $R_{i,j}$, such that p_i is the associated writer and p_j is the associated reader. There are $n(n-1)$ different ordered pairs (p_i, p_j), $i \ne j$. In each such ordered pair the first element is a writer and the second element is a reader. No subregister $R_{i,j}$ can be associated with more than one such (writer, reader) pair, since the subregisters have only one associated writer and one associated reader other than the writer. Hence, there must also be $n(n-1)$ different subregisters $R_{i,j}$ in the compound register R. This is exactly achieved by the presented matrix register K, noting that the subregisters on the main diagonal are superfluous. I.e., p_i can remember what it wrote last in $K_{i,i}$. •

The assumption of atomicity of the subregisters of the compound matrix register is convenient in the proof, but is it necessary? It turns out that regularity of subregisters suffices to obtain atomicity of the compound matrix register. First we note that the relation between the operation of the subregisters and the operation of the compound register can be expressed as follows. Let $\max_{lx}$ mean the lexicographic maximum. The value written by a write subaction satisfies:

$$v(W(a,i,j)) = (t(a), v(a)),\ 1 \le j \le n, \text{ with} \tag{3.5}$$

$$t(a) = \max_{lx}\{t(b) : \pi_{k,i}(R(a,k,i)) = W(b,k,i),\ 1 \le k \le n\} \text{ if } a \text{ is a read, and}$$

$$t(a) >_{lx} \max_{lx}\{t(b) : \pi_{k,i}(R(a,k,i)) = W(b,k,i),\ 1 \le k \le n\} \text{ if } a \text{ is a write.}$$

Theorem 3.5. (Regularity of Subregisters) *The registers $K_{i,j}$ do not need to be atomic; regularity is sufficient to guarantee atomicity of K.*

Proof. Let ρ and $\rho_{i,j}$ be defined as before. Define $-\!\gg$ as the minimal transitive relation on the subactions $\bigcup_{i,j=1}^{n} A_{i,j}$ extending the $\rightarrow_{i,j}$'s and satisfying (R7). Then (3.1) holds for this new $-\!\gg$, and we define a new $\rightarrow$ in terms of the new $-\!\gg$ as in (3.2). This satisfies (R8) and (R9) again. The new $\rightarrow$ is a partial order on A, since the proof of Lemma 3.1 goes through unchanged for the new relations $-\!\gg$ and $\rightarrow$.

Since the subregisters are regular, by Theorem 2.2 there is only one way in which associated runs can fail to be atomic. Namely, if the reading mapping associated with such a run is not weakly monotonic. We need to show that this does not result in a reading mapping that is inconsistent with atomicity of the run of the compound register. The proof is by first assuming that all subregisters are atomic, and then replacing the atomic subregisters by regular subregisters, one by one, retaining an atomic compound register after each step. In the step involving subregister $K_{i,j}$,

we use induction on the ordered set of read subactions in $A_{i,j}$. So, assume first that all subregisters $K_{i,j}$ $(1 \le i,j \le n)$ of K are atomic.

Base Case: $i+j \le 1$. Run ρ using atomic subregisters is atomic by Theorem 3.2.

Induction: $n \ge i,j \ge 1$. Assume the register mapping induced by the compound register K stays atomic (i.e., all runs are atomic) under replacement of the atomic subregisters $K_{p,q}$ by regular ones, for all p,q such that $(p,q) <_{lx} (i,j)$. By way of contradiction, assume that the register mapping becomes nonatomic if we replace also $K_{i,j}$ by a regular subregister. I.e., there is a nonatomic run $\rho=(A,\rightarrow,\pi)$. By the inductive assumption, the nonatomicity of ρ must be due to a first nonatomic event in $\rho_{i,j}$. By Theorem 2.2 this must be as follows.

By (R7), $-\!\!\gg$ is a total order on the subactions by the same processor. Let $R(a,i,j)$ and $R(b,i,j)$ be the $-\!\!\gg$-least pair of consecutive read subactions by p_j on $K_{i,j}$, such that

$$R(a,i,j) \rightarrow_{i,j} R(b,i,j) \;\&\; W(c,i,j) \rightarrow_{i,j} W(d,i,j), \tag{3.6}$$

$$\pi_{i,j}(R(a,i,j)) = W(d,i,j) \;\&\; \pi_{i,j}(R(b,i,j)) = W(c,i,j).$$

For ρ to be nonatomic as a result of (3.6), $\pi(b)$ must be different from a write it could have been in case $\rho_{i,j}$ were atomic. By choice of $R(a,i,j)$ and $R(b,i,j)$, it follows by (R7) and (R8) that a and b must be $\rightarrow$-consecutive actions by p_j.

Since $a \rightarrow b$, if $\rho_{i,j}$ were atomic, then, by Theorem 2.2,

$$\text{if } \pi_{i,j}(R(a,i,j)) = W(d,i,j) \text{ then } \pi_{i,j}(R(b,i,j)) = W(e,i,j), \tag{3.7}$$

for some action e such that $d \rightarrow e$ or $d=e$. Now consider the subregisters on the main diagonal. Processor p_j is both the only writer and the only reader associated with $K_{j,j}$. In particular, regardless of whether the subregisters are regular or atomic, we have for ρ:

$$\pi_{j,j}(R(b,j,j)) = W(a,j,j). \tag{3.8}$$

Intuitively, even though $R(b,i,j)$ scans too old a value, the way $\pi(b)$ is determined as in (3.5) will show no difference whether we assume (3.6) or (3.7), because of (3.8). More formally, consider

$$S = \{t(x): W(x,k,j) = \pi_{k,j}(R(b,k,j)),\ 1 \le k \le n\}$$

$$S' = (S - \{t(c)\}) \cup \{t(e)\}.$$

Then S' is the set we would have instead of S, if $K_{i,j}$ were not replaced by a regular subregister. By the induction hypothesis, S' gives rise to an atomic compound register. But, S contains also $t(a)$ by (3.8), and by (3.6) and (3.5) we have $t(a) \ge_{lx} t(d)$ Hence, the $\max_{lx}(S)$ is invariant under whether $\pi_{i,j}(R(b,i,j)) = W(c,i,j)$ or $\pi_{i,j}(R(b,i,j)) = W(d,i,j)$, i.e., whether (3.6) or (3.7) hold. Consequently, atomicity of the compound register is invariant under (3.6) or (3.7) as well. This ends the induction, and therefore the proof of the theorem. ●

Using regular subregisters, one may get the impression that the main diagonal subregisters in the matrix register cannot be left out. However, again a processor can remember what it wrote last, and so we can dispense with the main diagonal.

In the matrix construction, the reads on the compound register, comprise writes on the subregisters. A natural question to ask is, whether we can implement a compound atomic register such that the reads on the compound register do not contain writes on a subregister. In [3] it is proved that to implement an atomic register from **regular** ones, the reads of the compound

register necessarily include writes of the subregisters. Hence, writes of subregisters are unavoidable in reads of the atomic matrix register, assuming it is constructed from regular subregisters (as in Theorem 3.5). In contrast, in Bloom's register below, where the subregisters are **atomic**, reads of the compound register do not involve writes of a subregister.

4. The Bloom Register

The Bloom register is a compound 2-writer, n-reader register, constructed as a pair of atomic 1-writer, n-reader subregisters. The domain of values of the subregisters is the cartesian product $\{0,1\}\times V$, with V the domain of values of the compound register. This register, the first atomic 2-writer register, was designed by Bard Bloom, and is e.g. defined in [1]. It is remarkable in the simplicity of its protocol, and that the domain of values of the subregister is only twice the domain of values of the compound register. The argument for its atomicity, as given in [1], uses global time and is not easy to follow. It assigns a value $\sigma(a)$ to the time interval $I(a)$ associated with each action a (i.e., a shrinking value). The present proof is based only on causality considerations. The atomicity of the register follows by Theorem 2.1. If we assume global time, then the precedence relation $\rightarrow$ defined below is an interval order. Hence, shrinking atomicity of Bloom's register follows by Theorems 2.1 and 2.3.

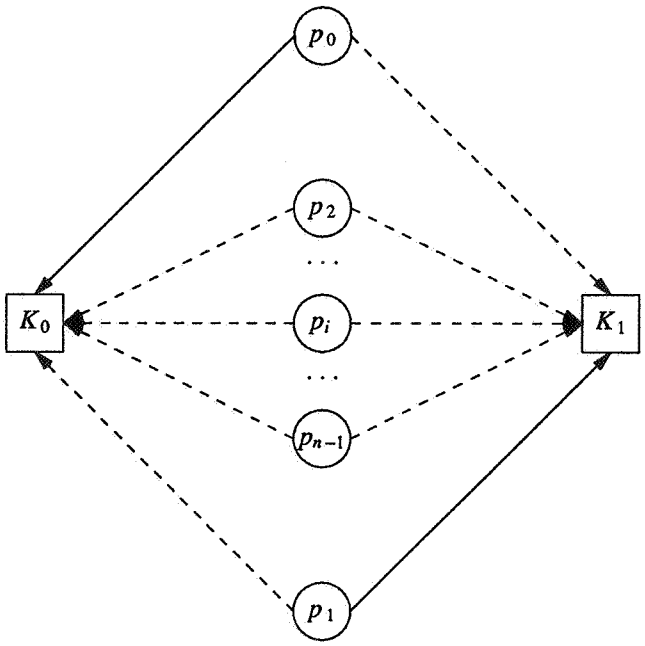

Figure 2: The 2-writer, n-reader, Bloom register.

Throughout the present section $\oplus$ denotes modulo 2 addition. Let $p_0,p_1,\ldots,p_{n-1}$ be n processors and let K be a compound register consisting of two atomic, 1-writer, n-reader subregisters, say K_0, K_1. Processor p_0 is a writer of K_0, and processor p_1 is a writer of K_1. All processors p_i are readers of both K_0 and K_1, $i=0,\ldots,n-1$. Let V be the value domain of the compound register K. Then $\{0,1\}\times V$ is the value domain of each subregister. A **tag** is an element of $\{0,1\}$. We say that each subregister can hold a tag, next to a value from the domain V of the compound register. Denote the tag variable in K_i by, say t_i, $i=0,1$. Both registers are initialized with tag 0 and value 0. Moreover, each run of the compound register starts with a write action, as

required by (R3). The architecture is depicted in figure 2. It implements an n-reader, 2-writer register K. The register obeys the following protocol.

p_i writes the value a $(i = 0,1)$:

1. read $t_{i\oplus 1}$;
2. $t_i := i \oplus t_{i\oplus 1}$;
3. write t_i and value a in register K_i

p_i reads value:

1. read t_0 from K_0;
2. read t_1 from K_1;
3. read the value from register K_j, for $j = t_0 \oplus t_1$;

Ignoring step 2 of the write protocol, because it is no register access operation, every read or write action a by processor p_i consists of two or three atomic read or write subactions.*

Let $\rho = (A, \rightarrow, \pi)$ be a run associated with K. Let $a \in A$. Denote by $p(a)$ the number of the processor that executes a. Denote by $R(w, i\oplus 1)$ the subaction of a write w that reads from subregister $K_{i\oplus 1}$, and denote by $W(w)$ the subaction of w that writes to subregister K_i, for $i = p(w)$. Denote by $R_s(r, i)$ the first subaction of read r that reads subregister K_i, i=0,1, and denote by $R_f(r, j)$ the final subaction of r that reads K_j, for j is either 0 or 1. Each subregister K_i (i=0,1) of K is atomic. Let ρ=$(A, \rightarrow, \pi)$ be a run of K. Define, for i=0,1, ρ_i=$(A_i, \rightarrow_i, \pi_i)$, the run of K_i associated with ρ. Since K_i is atomic, we can extend $\rightarrow_i$ to a total atomic precedence relation $\Rightarrow_i$, i=0,1. This means that there is an apparent total order on the subactions executed on the same subregister. Moreover, if a subread reads a subwrite (on a subregister) then there is no other subwrite placed between them by this order.

Lemma 4.1. *Let $-\!\!\gg$ be the minimal transitive relation on the subactions in $A_1 \cup A_2$, extending $\Rightarrow_0$ and $\Rightarrow_1$, and satisfying (R7). Then $-\!\!\gg$ is a partial order on A.*

Proof. By Lemma 2.4. •

Define the precedence relation $\rightarrow$ on A by

$$a \rightarrow b \text{ iff } \alpha -\!\!\gg \beta, \tag{4.1}$$

for all subactions α of a and β of b. Obviously, relation $\rightarrow$ is a partial order and satisfies (R8) and (R9). It is convenient to use the following property.

Lemma 4.2. *$\rightarrow$ is an interval order on the set of writes W.*

Proof. Since $\rightarrow$ is a partial order, it is irreflexive, and it remains to show that its restriction to W satisfies (2.1). Note that there are only two processors involved, the writers p_0 and p_1. Suppose $a \rightarrow b$ and $c \rightarrow d$. Assume w.l.o.g. that all four actions are distinct. Using (R7),

if a, c are by the same processor, then $a \rightarrow c \rightarrow d$ or $c \rightarrow a \rightarrow b$;

---

* In the read protocol we can delete step 1 if i=0 and step 2 if i=1, because the executing processor p_i can remember the t_i it has written last. Because p_i can also remember the value it has written last, we can delete step 3 of the read protocol if j=i. Thus, such a fine tuned read protocol may consist of only one atomic read of $t_{i\oplus 1}$ from $K_{i\oplus 1}$ by p_i, in case $t_0 \oplus t_1 \neq i$. Correctness follows because this version is obviously equivalent to the original one.

if b,d are by the same processor, then $a \to b \to d$ or $c \to d \to b$;

if a,d are by the same processor, then $a \to d$ or $c \to d \to a \to b$;

if b,c are by the same processor, then $a \to b \to c \to d$ or $c \to b$.

W.l.o.g., the only remaining case to check is a,b are by p_0 and c,d are by p_1. If $W(a) \Rightarrow_0 R_s(d,0)$, then each subaction of a $-\gg$-precedes each subaction of d, i.e., $a \to d$. If $R_s(d,0) \Rightarrow_0 W(a)$ then each subaction of c $-\gg$-precedes each subaction of b, i.e., $c \to b$. •

Define the clans $[w]$ and the $\to^\pi$ relation between them as before.

Theorem 4.3. *The Bloom register K is an atomic, 2-writer, n-reader register, which is implemented by two atomic, 1-writer, $n-1$-reader registers. The value domain of the subregisters is the cartesian product of $\{0,1\}$ and the value domain of the compound register.*

Proof. Let $\rho = (A, \to, \pi)$ be an arbitrary run according to the protocol. The conclusion of the theorem follows immediately from Lemma 4.4, Lemma 4.5 and Theorem 2.1. Namely, by Lemma 4.5 (1) ρ is normal, and Lemma 4.5 (2) implies acyclicity of $\to^\pi$ by Lemma 4.4.

Like in the previous section, the key idea is to totally order the write actions first. Denote the *identity* of the tag written in step 3 of a write w by $t(w)$, and its *value* (0 or 1) by $val(t(w))$. Consider only the set W of writes. Define a relation $\ll_{ia}$ on the set of writes by $w \ll_{ia} w'$ iff $w \neq w'$ and either

(a) $w \to w'$, or

(b) $\neg(w \to w')$ & $\neg(w' \to w)$ and $val(t(w)) \oplus val(t(w')) = p(w')$.

Let $\ll$ be the transitive closure of $\ll_{ia}$.

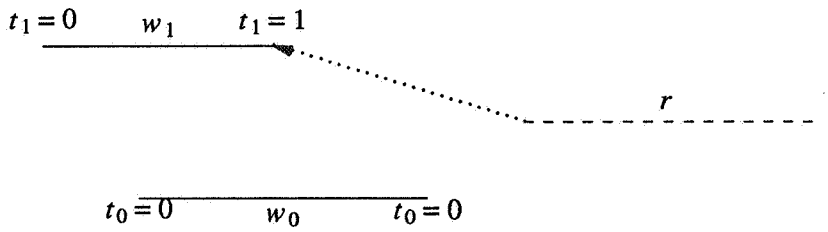

Figure 3: r reads w_1 since $t_0 \oplus t_1 = 1$.

It may come as a surprise, that the writes can be ordered by $\ll$ in a way which contradicts the timing of the final subwrites. See figure 3 for an example where write w_0 ends later than write w_1, and yet the value written by the former cannot be read by any read.

Lemma 4.4. *$\ll$ is a total order on W.*

Proof. It suffices to prove that $\ll_{ia}$ is acyclic. Assume to the contrary that there is a minimal length cycle:

$$w_0 \ll_{ia} w_1 \ll_{ia} \cdots \ll_{ia} w_m \ll_{ia} w_0 .$$

Since writes by the same processor are totally ordered by $\to$, the cycle must contain writes by both p_0 and p_1. Suppose the cycle contains writes $a \to b$ by p_0 and writes $c \to d$ by p_1. Since $\to$ is an interval order on W, it follows either $a \to d$ or $c \to b$, which contradicts that the cycle is of minimal length. Therefore, the cycle contains precisely one write by one of the two processors, say w_0 by p_0. Consequently, since the writes by p_1 are totally ordered by $\to$, we have

$w_1 \rightarrow \cdots \rightarrow w_m$. Since the cycle has minimal length, and $\rightarrow$ is transitive, it follows $m=1$ or $m=2$.

Case 1. Suppose $m=1$. Then, $w_0 \ll_{ia} w_1$ and $w_1 \ll_{ia} w_0$, which contradicts the antisymmetry of $\ll_{ia}$.

Case 2. Suppose $m=2$. Then,

$$w_0 \ll_{ia} w_1 \ll_{ia} w_2 \ll_{ia} w_0 , \tag{4.2}$$

in which w_1 and w_2 are writes by p_1. Since $w_1 \rightarrow w_2$, the only way for this cycle to occur is if

$$\neg(w_0 \rightarrow w_1) \;\&\; \neg(w_1 \rightarrow w_0) \;\&\; val(t(w_0)) \oplus val(t(w_1)) = 1 \;\&\; \tag{4.3}$$
$$\neg(w_0 \rightarrow w_2) \;\&\; \neg(w_2 \rightarrow w_0) \;\&\; val(t(w_0)) \oplus val(t(w_2)) = 0 \; .$$

We now derive a contradiction by induction on $\rightarrow$. We only give one half of the argument. The other half, with the roles of p_0 and p_1 reversed, is symmetric.

Base case. Assume w_0 is the first write by p_0. Then, since both w_1, w_2 are $\rightarrow$-incomparable to w_0, they obtain the same initial tag from K_0. Therefore, the tags written by w_1, w_2 are equal, i.e., $val(t(w_1)) = val(t(w_2))$, which contradicts (4.3).

Induction. Assume there is no 3-cycle like (4.2) containing a write w by p_0, with $w \rightarrow w_0$. That is, (4.2) is $\rightarrow$-minimal in this sense.

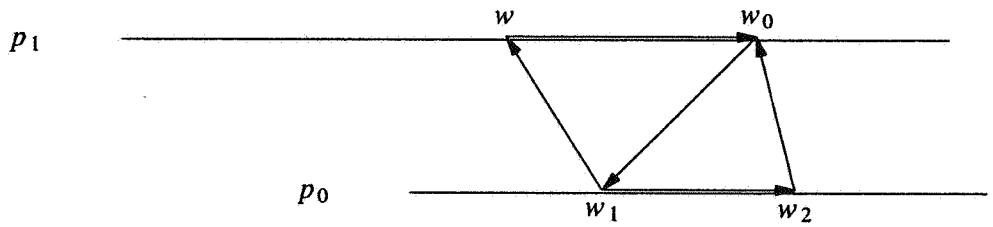

Figure 4: Reducing to an earlier 3-cycle.

Let w be a write by p_0, with $w \rightarrow w_0$. Since $w_1 \rightarrow w_2$, and $\rightarrow$ is an interval order on W, either $w \rightarrow w_2$ or $w_1 \rightarrow w_0$. Since w_1, w_0 are $\rightarrow$-incomparable, we must have $w \rightarrow w_2$. Assume that w is the write by p_0 directly preceding w_0. Then w is also the write by p_0 which directly precedes w_2. Therefore, w_2 reads $t(w)$ from K_0. Hence, in write w_2, processor p_1 chooses $t(w_2)$ such that $val(t(w)) \oplus val(t(w_2)) = 1$. Then, by (4.3),

$$val(t(w)) \oplus val(t(w_1)) = 0 . \tag{4.4}$$

We can now finish the argument by showing that each way w and w_1 can be $\rightarrow$-related leads to a contradiction.

Suppose $w \rightarrow w_1$. Since both pairs w_0, w_1 and w_0, w_2 were $\rightarrow$-incomparable, both w_1 and w_2 must read tag $t(w)$. Then, $val(t(w_1)) = val(t(w_2))$, which contradicts (4.3).

Suppose $w_1 \rightarrow w$. Then $w_1 \rightarrow w_0$, which contradicts (4.3) again.

Suppose w and w_1 are $\rightarrow$-incomparable. Then, by (4.4) and the definition of $\ll_{ia}$, we have $w_1 \ll_{ia} w$. This, together with $w \rightarrow w_0$ and part of cycle (4.2), creates a new cycle

$$w \ll_{ia} w_0 \ll_{ia} w_1 \ll_{ia} w \; .$$

However, this is a 3-cycle contradicting the assumption that (4.2) is the 3-cycle containing the $\rightarrow$-least write by p_0. This concludes the induction and the proof of the lemma. •

Lemma 4.5.

(1) ρ *is normal.*

(2) If $[w] \rightarrow^{\pi} [w']$ *then* $w \ll w'$.

Proof. *(1).* By construction of the protocol, and because there is an initial write preceding all other actions, each read reports a value written by a write. I.e., π is a total function. If $w=\pi(r)$, then r returns v with $id(v)=w$, by (R6). Without loss of generality, let $p(w)=0$. Thus, considering the run $\rho_0=(A_0,\Rightarrow_0,\pi_0)$ of subregister K_0, it must be the case that $\pi_0(R_f(r,0))=W(w)$, and therefore $W(w)\Rightarrow_0 R_f(r,0)$ by the definition of atomicity. Hence, by definition (4.1) of $\rightarrow$, we have $\neg(r \rightarrow \pi(r))$.

(2). It is convenient to prove the following Claim first.

Claim. If $[w] \rightarrow^{\pi} [w']$ and $\neg(w \rightarrow w')$, then $w \rightarrow r$ for some read $r \in [w']$.

Proof of the claim. If $[w] \rightarrow^{\pi} [w']$ then $a \rightarrow b$ with $a \in [w]$ and $b \in [w']$.

Case a. If $a=w$ and b is a read, then there is nothing to prove.

Case b. Let $a=w$ and $b=w'$. This contradicts assumption $\neg(w \rightarrow w')$.

Case c. Let a be a read. I.e., $\pi(a)=w$. Without loss of generality, let $p(w)=0$. Then, $\pi_0(R_f(a,0))=W(w)$ and therefore $W(w)\Rightarrow_0 R_f(a,0)$. Therefore, by (4.1), we have $w \rightarrow b$. If $b=w'$ then this case reduces to Case b. If b is a read then this case reduces to $w \rightarrow b$ with $\pi(b)=w'$, which is what we had to prove. This completes the proof of the Claim.

To prove part 2 of the lemma, assume to the contrary that $[w] \rightarrow^{\pi} [w']$ and $w' \ll w$. Suppose first that w and w' are executed by the same processor. Then, by definition of $\ll$, $w' \rightarrow w$. By the claim, $w \rightarrow r$ for some read $r \in [w']$. Since the value written by w overwrites the value written by w' in K_0, a read r with $R_f(r,0)$ such that $W(w)\Rightarrow_0 R_f(r,0)$, cannot read the value written in K_0 by w', which contradicts $r \in [w']$. Therefore, we must assume that w and w' are executed by different processors, say p_0 and p_1, respectively.

Case 1. Suppose $w \rightarrow w'$. Then it follows that $w \ll w'$, which is a contradiction.

Case 2. Suppose $w' \rightarrow w$. By the claim, for some r, we have $w \rightarrow r$ and $\pi(r)=w'$. Therefore, $w' \rightarrow w \rightarrow r$. Since $\pi(r)=w'$, p_1 does not write in K_1 in between $W(w')$ and $R_f(r,1)$. Hence when read r scans the tag in register K_1, in subread $R_s(r,1)$, it must obtain $t(w')$. Each write $\overline{w}$ by p_0 (in particular $\overline{w}=w$, but also all other writes $\overline{w}$ satisfying the following conditions), such that $w' \rightarrow \overline{w}$ and $W(\overline{w})\Rightarrow_0 R_s(r,0)$, scanning the tag in K_1, also must obtain $t(w')$. Therefore, p_0 chooses $val(t(\overline{w}))$ equal $val(t(w'))$ in write $\overline{w}$. Hence, subread $R_s(r,0)$ scans a tag $t(\overline{w})$ such that $val(t(\overline{w})) = val(t(w'))$. The final subread in r, therefore, must read from K_0, i.e., must be $R_f(r,0)$, since $val(t(\overline{w}))\oplus val(t(w')) = 0$. This contradicts the assumption that r returns the value written by w' in K_1.

Case 3. Suppose neither $w' \rightarrow w$ nor $w \rightarrow w'$. Since $w' \ll w$, it must be the case that

$$val(t(w))\oplus val(t(w')) = 0 . \tag{4.5a}$$

By the Claim we have $w \rightarrow r$ with $\pi(r)=w'$. Thus, r returns the value from K_1. This implies that r scans a tag $t(w_0)$, written by some write w_0 by p_0, and a tag $t(w_1)$, written by some write w_1 by p_1, such that

$$val(t(w_0))\oplus val(t(w_1)) = 1. \tag{4.5b}$$

It turns out, that any possible $\rightarrow$-relation between w_0 and w (both by p_0), in combination with any possible $\rightarrow$-relation between w_1 and w' (both by p_1), leads to a contradiction.

We first observe that neither $w_0 \rightarrow w$ nor $w' \rightarrow w_1$. Namely, if $w_0 \rightarrow w$, then $w_0 \rightarrow w \rightarrow r$ and r cannot scan $t(w_0)$ in K_0, since it is already overwritten by w when r scans. Therefore, $w = w_0$ or $w \rightarrow w_0$. Secondly, because r returns the value written by w' in K_1, there can be no $W(w_1)$ between $W(w')$ and $R_f(r,1)$. Therefore, $w' = w_1$ or $w_1 \rightarrow w'$. Now we check off the remaining possibilities.

Subcase 3.1. Suppose $val(t(w_0)) = val(t(w))$ and $val(t(w_1)) = val(t(w'))$. This contradicts (4.5) straightaway.

Subcase 3.2. Suppose $val(t(w_1)) \neq val(t(w'))$. Then $w_1 \rightarrow w'$ and therefore $w_1 \ll w'$. Since we have assumed $w' \ll w$, we have $w_1 \ll w$. If w_1, w were $\rightarrow$-incomparable, then $val(t(w_1)) \oplus val(t(w)) = 0$ by definition of $\ll$. Therefore $val(t(w')) \oplus val(t(w)) = 1$, which contradicts (4.5). Hence $w_1 \rightarrow w$. However, we have assumed that $w \rightarrow r$ and r scans $t(w_1)$. Then, also w must obtain $t(w_1)$, and again $val(t(w_1)) \oplus val(t(w)) = 0$, contradicting (4.5).

Subcase 3.3. Suppose $val(t(w)) \neq val(t(w_0))$. Then $w \rightarrow w_0$ and therefore $w \ll w_0$. Since $w' \ll w$ we have $w' \ll w_0$. If w_0, w' were $\rightarrow$-incomparable, then $val(t(w_0)) \oplus val(t(w')) = 0$ by definition of $\ll$. Then, $val(t(w)) \oplus val(t(w')) = 1$, which contradicts (4.5). The only other way to satisfy $w' \ll w_0$ is $w' \rightarrow w_0$. Since r scans $t(w_0)$, we also have $W(w_0) \rightarrow\!\!\gg R_s(r,0)$. But r reports the value written by w', and therefore there is no write on K_1 in between $W(w')$ and $R_f(r,1)$. So w_0 must scan $t(w')$. Hence, $val(t(w')) \oplus val(t(w_0)) = 0$, contradicting (4.5).

This finishes the proof of the lemma and hence of the theorem. •

Corollary. The Bloom register is shrinking atomic.

Proof: sketch. This follows similarly as in the matrix construction by assuming that $\rightarrow$ is an interval order, e.g., by assuming global time. Then we have 'shrinking' atomicity by Theorem 2.3. •

5. Conclusion

In this paper we propose a method of proving atomicity of shared registers in an order setting. It seems to us that it can be applied in many cases where we have to prove atomicity. In outline:

1. Find an appropriate partial order $\rightarrow$ between the high level reads and writes defined in terms of the assumed partial order between the lower level reads and writes. Induce the $\rightarrow^{\pi}$ relation on the set of clans defined by the reading mapping.
2. Find a way to totally order the writes, using the intuition which makes you believe the protocol works correctly. Use this total order to prove that $\rightarrow^{\pi}$ is acyclic.

Leslie Lamport has suggested expressing atomicity in terms of orders before. He treats the single writer case in [3]. On a metalevel, our method of first totally ordering the writes as an auxiliary construction for proving acyclicity of $\rightarrow^{\pi}$, may be viewed as a reduction to the single writer case. The method is present in embryonic form in [9], and was used in its present form in the presentation of that paper (and distributed) at that conference as [10]. Nancy Lynch has devised a related proof for the matrix algorithm using time and Input/Output automata. The 'knowledge graph' in [8] is the set of clans with the $\rightarrow^{\pi}$ relation induced by the natural interval order.

Acknowledgement

Conversations with Bard Bloom, Leslie Lamport, Arjen Lenstra and Nancy Lynch are gratefully acknowledged. Lambert Meertens' comments had a profound influence on this paper.

References

[1] Bloom, B., *Constructing Two-writer Atomic Registers,* Manuscript, Massachusetts Institute of Technology, June 1986.

[2] Fishburn, P.C., *Interval Orders and Interval Graphs*, Wiley, 1985.

[3] Lamport, L., *On Interprocess Communication, Part I: Basic Formalism, Part II: Algorithms,* Distributed Computing, vol. 1, pp. 77-101, 1986.

[4] Lamport, L., *The mutual exclusion problem, part I - A theory of interprocess communication*, Journal ACM, vol. 33, pp.313-326, 1986.

[5] Mullender, S.J., and P.M.B. Vitanyi, *Distributed match-making for processes in computer networks.* In: Proceedings 4th ACM Symposium on Principles of Distributed Computing, 1985, 261-271.

[6] Papadimitriou, C., *The serializability of concurrent database updates*, Journal ACM, vol. 26, pp. 631-653, 1979.

[7] Peterson, G. L., *Concurrent Reading While Writing,* ACM Transactions on Programming Languages and Systems, Vol. 5, No. 1, Jan. 1983, pp. 46-55.

[8] Peterson, G.L. and J.E. Burns, *Concurrent Reading While Writing II, the Multiwriter Case*, Tech. Rept. CIT-ICS-86/26, Georgia Institute of Technology, December 1986.

[9] Vitanyi, P. M. B., and Awerbuch, B., *Atomic Shared Register Access by Asynchronous Hardware,* Proceedings 27th IEEE Symposium on Foundations of Computer Science, 1986, 233-243.

[10] Vitanyi, P. M. B., and Awerbuch, B., *Atomic Shared Register Access by Asynchronous Hardware using Unbounded Tags,* Manuscript, Massachusetts Institute of Technology, October 1986.