Overhauling SC atomics in C11 and OpenCL

Abstract

Despite the conceptual simplicity of sequential consistency (SC), the semantics of SC atomic operations and fences in the C11 and OpenCL memory models is subtle, leading to convoluted prose descriptions that translate to complex axiomatic formalisations. We conduct an overhaul of SC atomics in C11, reducing the associated axioms in both number and complexity. A consequence of our simplification is that the SC operations in an execution no longer need to be totally ordered. This relaxation enables, for the first time, efficient and exhaustive simulation of litmus tests that use SC atomics. We use our improved C11 model to present the first rigorous memory model formalisation for OpenCL (which extends C11 with support for heterogeneous many-core programming). In the OpenCL setting, we refine the SC axioms further to give a sensible semantics to SC operations that employ a ‘memory scope’ to restrict their visibility to specific threads. Our overhaul requires slight strengthenings of both the C11 and the OpenCL memory models, including some behaviours to become disallowed. We argue that these strengthenings are natural, and prove that all with the original model, making

1. Introduction

**Atoms and memory models.** C11 and OpenCL both define a collection of atomic operations, or ‘atomics’, which can be used by experts to program high-performance, lock-free algorithms in a portable manner. Atomics accept a memory order parameter, which controls the exposure of certain relaxed memory behaviours that modern CPUs and GPUs natively exhibit.

The C11 and OpenCL specifications [15, 17] define the semantics of atomics via axiomatic memory models; that is, sets of rules that govern the reading and writing of shared memory locations. These memory models are complex, stretching to about 19 and 30 pages, respectively, of convoluted prose. This complexity makes it extremely challenging to reason about the correctness of programs that are written in, and compilers that implement, these languages. Correctness in any relaxed memory setting is notoriously evasive; indeed, the subtleties of relaxed memory have previously led to confirmed bugs in language specifications [5, 8], deployed processors [1], compilers [20, 32] and vendor-endorsed programming guides [3]. Yet the importance of correctness in the context of C11 is well-known. In OpenCL, which is an open standard for heterogeneous programming developed and supported by major hardware vendors such as Altera, AMD, ARM, Intel, Nvidia and Qualcomm, correctness is just as crucial, OpenCL being a key player in the recent drive to exploit GPUs in general-purpose computing – including in safety-critical domains such as medical imaging [26] and autonomous navigation [18].

We seek in our work to tame the complexity of these memory models through formalisation.

**The C11 memory model** has been formalised by several researchers, in varying degrees of completeness, and with varying degrees of fidelity to the standard [2, 5, 30]. These formalisation efforts have proved fruitful; they have, for instance, enabled the construction of simulators that automatically explore the allowed behaviours of small C11 programs (called litmus tests) [2, 5, 9, 22], underpinned the design of program logics for specifying and verifying C11 programs [29, 30], and they provide a firm foundation for ongoing debate about the design of the C11 memory model itself [8, 31].

**The OpenCL memory model** (introduced in version 2.0 of the standard) has received comparatively little academic attention, with the notable exception of the work of Gaster et al. [12], which we discuss further in §7. OpenCL provides a framework for CPU programs to delegate the execution of massively-parallel kernel functions, written in a variant of C, to one or more accelerator devices, such as GPUs. Threads that execute these kernels are organised into a hierarchy: threads are grouped into work-groups, and work-groups are grouped by device. The OpenCL memory model is broadly similar to that of C11, but is extended with features such as memory regions (which contain locations that are accessible only to a certain subtree of the thread hierarchy), and memory scopes (which, when applied to an atomic operation, confine its visibility to a certain subtree of threads).

**SC atomics.** Our work is distinguished by its focus on the sequentially consistent (SC) fragment of these memory models; that is, the semantics of atomics whose memory order is memory_order_seq_cst. The chief guarantee provided by this memory order is that all SC atomics in a given execution will execute in some order (say, S) on which all threads mutually agree. Note that these memory models do not construct S; they merely postulate the existence of a suitable S.

Sequential consistency is known for its simplicity [19], and indeed, any C11 or OpenCL program using exclusively SC atomics would enjoy a simple interleaving semantics. However, when combined with the more relaxed memory orders that C11 and OpenCL also provide, the semantics of SC atomics becomes highly complex, and it is this complexity that we tackle in this paper.

SC atomics are in widespread use, partly because the SC memory order is used when no other is specified, and partly because programmers are routinely advised to use SC atomics prior to optimising their code with the more relaxed memory orders [34 (p. 221)]. Algorithms that make use of SC atomics include Dekker’s mutual exclusion algorithm [10], and more generally, multiple-producer-multiple-consumer algorithms that require every consumer to observe the actions of every producer in the same order. As such, it is important that the semantics of SC atomics is clear to programmers, to allow smooth transitioning between the exclusive use of SC (for ease of reasoning) to a mixture of SC and weaker-than-SC atomics (for performance optimisation).
In theory, SC atomics can be avoided by replacing them with mutex-protected non-atomic operations (and simple spinlock mutexes can be implemented using just release and acquire atomics [34 (p. 111)]). In practice, support for SC atomics is non-negotiable if software is to make use of concurrency libraries. This is because the aforementioned replacement of SC atomics must be performed throughout the entire program – in both library code and client code alike – and with the same mutex variable for every operation. Moreover, programs that make extensive use of spinlocks could prove less efficient than those that rely on native SC atomics, and accidental misuse of locks may lead to deadlock.

1.1 Main contributions

Our work aims to provide clearer, simpler foundations for reasoning about C11, enabling a clean extension to OpenCL for heterogeneous programming, and facilitating efficient simulation.

1. Overhauling SC atomics in C11 (§3). The C11 specification devotes around 276 words to explaining the semantics of SC atomics. In our work, we have translated these words into mathematical axioms, carefully strengthened these axioms (without imposing unreasonable demands on the compiler), and then restructured them so that they are expressed as simply as possible. Our revised model

- is shorter (requiring just 93 words in the same prose style),
- is simpler (because it does away with the total order $S$), and
- is amenable to more efficient simulation (because it is no longer necessary to search for a suitable $S$).

2. Overhauling SC atomics in OpenCL (§5). Our simplifications to the rules governing SC atomics in C11 can be carried over directly to OpenCL, where the same three benefits listed above can be reapplied. In the OpenCL setting, however, there is an additional complexity in the semantics of SC atomics. Specifically, the total order $S$ in which all of a program’s SC atomics execute is only guaranteed to exist when one of two conditions holds: either all SC atomics in the program’s execution use the widest-possible memory scope and only access memory shared between devices, or all SC atomics have their memory scope limited to the current device and never access memory shared between devices. We find that this semantics is unhelpful to programmer-writers: a loop-hole that we discovered in the second condition above means that even device-scoped SC atomics must be implemented using expensive inter-device synchronisation.

We have amended the rules that govern SC atomics in OpenCL, so that the SC guarantees do not vanish immediately in the presence of a differently-scoped SC atomic somewhere in the program, but instead degrade gracefully. Our revised rules

- are shorter and simpler (we can replace 391 words in the specification with 106 words in the same prose style),
- enable new programming patterns in OpenCL (such as programs that use SC atomics in a natural manner, yet a manner that violates the overly restrictive conditions above),
- let device-scoped SC atomics be efficiently implemented, and
- improve the compositionality of OpenCL semantics, and hence the ability to write concurrency libraries (because the behaviour of SC atomics no longer depends on unstable, global conditions).

3. Proving the implementability of our revised models (§3.3, §5.1). Our improvements to the SC axioms in the C11 and OpenCL memory models hinge on slight strengthenings of the models; that is, tweaking some of the axioms so that fewer executions are allowed. This increases the demands on compilers that implement these memory models, so it is important to check that our changes do not invalidate existing compilation schemes. To this end, we provide proofs that all of the formalised compilation schemes of which we are aware (namely: the C11 schemes for Power [6] and x86 [5] machines, and the OpenCL scheme for AMD GPUs [33]) remain sound after our change.

1.2 Supporting contributions

In order to justify the claims we make in our main contributions, we have established several supporting artefacts, which we believe are also valuable in their own right.

4. Formalising the OpenCL memory model (§5). The OpenCL specification contains numerous ambiguities, omissions and inconsistencies, which makes it a shaky structure upon which to build an argument about the correctness of an OpenCL program or compiler. The lack of clarity may lead programmers and compiler-writers to cautiously opt for low-efficiency implementations that are easier to guarantee correct; moreover, there are instances where the OpenCL specification authors have made unnecessarily conservative, programmer-unfriendly decisions in the design of rules for memory consistency. We provide the first mechanised formalisation of the OpenCL memory model. Our formalisation serves to clarify the specification, and can be henceforth be used to underpin future program logics for verifying OpenCL kernels, and to inform further refinements to the memory model. In particular, we use our rigorous memory model to show that the design decisions of the specification can be made less conservative, offering programmers more flexibility, without placing any additional burden on efficient implementation of the language.

5. Formalising the memory models in .cat (§2, §4). We have encoded the C11 and OpenCL memory models in the .cat framework [2]. Previous formalisations of the C11 memory model exist, in Isabelle [5], Lem [6] and Coq [30]; here we contribute the first version in .cat. We conduct our development work in the .cat language because it is the native input format to the HERD memory model simulator, which has a proven record of efficiently simulating a range of CPU machine-level memory models [2 (§8.3)].

6. Extending the HERD memory model simulator (§6). During memory modelling work, tool support for simulating alternative memory models against litmus tests is invaluable. HERD is able to simulate any memory model expressed as a .cat file, but in its original incarnation, it supported only machine-level models of CPUs [2] and GPUs [3]. To explore our proposed changes to the C11 and OpenCL memory models, we extended HERD with a module for generating executions of C11 and OpenCL programs, and support for language-level memory models that incorporate ‘undefined behaviour’ (a notion that is absent from machine-level models). This involved adding around 8000 lines to the original HERD codebase. All of the examples in this paper have been automatically checked with HERD. Using HERD, we have evaluated the impact of our changes to the SC axioms, and found an exponential improvement in simulation performance.

Online material. At http://tinyurl.com/popl16 we provide instructions for downloading HERD and our .cat formalisations.

2. The C11 memory model in .cat

This section describes formally the current C11 memory model.

The semantics of multi-threaded C11 programs is formalised in two stages; the first concerning the thread-local semantics, and

---

2 Indeed, our memory model has already been built upon by researchers investigating a proposed extension to OpenCL [33].

3 As estimated by git log.
the second capturing the memory model. Roughly speaking, the first stage takes as input a C11 program and calculates its set of executions (§2.1, §2.2); the second stage then compares each execution to the memory model to determine which executions are actually allowed (§2.3).

There exist several prior formalisations of the C11 memory model [2, 5, 30]. The novelty of this section is the first comprehensive formalisation of the model in the . cat framework [2], which enables the use of the efficient HERD simulator [2]. For reasons of space, and because they are orthogonal to the thrust of our contributions, we omit our treatment of the ‘consume’ memory order, unsequenced races and C11 locks from the paper. However, our .cat-based formalisation fully accounts for these features, and is provided as an appendix (§D).

2.1 C11 programs

C11 programs manipulate a set of shared memory locations.

Definition 1 (Memory locations). Each memory location is declared with either a non-atomic or an atomic type. That is, type(l) ∈ {atomic, non-atomic} for every memory location l.

Definition 2 (Structure of C11 programs). We consider C11 programs of the form P = ℓT pT, where T is a set of thread identifiers, pT is a piece of sequential code, and T is parallel composition.

Atomic locations can be accessed via atomic operations; these include reads, writes, and read-modify-writes (RMWs). C11 also defines fence operations. Atomic operations and fences expose the programmer to relaxed memory behaviours; which behaviours are exposed is controlled by the operation’s memory order parameter.

Definition 3 (Memory orders). The available memory orders in C11 are:

o := RLX (relaxed)
| ACQ (acquire, only for reads/RMWs)
| REL (release, only for writes/RMWs)
| AR (acquire+release, only for RMWs)
| SC (sequentially consistent, the default).

Example 1 (A C11 program). We give below a contrived C11 program that operates on two atomic locations, x and y, using atomic store and load operations with a variety of memory orders.

```c
atomic_int *x; atomic_int *y;
store(x,1,RLX); store(y,2,RLX);
store(x,2,SC); store(y,1,SC);
store(x,1,SC); store(y,2,SC);
```

2.2 C11 executions

The C11 memory model is defined in terms of program executions. An execution X takes the form of a mathematical graph, where each node e ∈ E is labelled with a run-time memory event (see Def. 4), and the edges connect events performed by the same thread in program order. In other words, an execution is a partial order over the second capturing the memory model. Roughly speaking, the first stage takes as input a C11 program and calculates its set of executions (§2.1, §2.2); the second stage then compares each execution to the memory model to determine which executions are actually allowed (§2.3).

The novelty of this section is the first comprehensive formalisation of the model in the . cat framework [2], which enables the use of the efficient HERD simulator [2]. For reasons of space, and because they are orthogonal to the thrust of our contributions, we omit our treatment of the ‘consume’ memory order, unsequenced races and C11 locks from the paper. However, our .cat-based formalisation fully accounts for these features, and is provided as an appendix (§D).

The labels represent (reading down): non-atomic writes, atomic writes, non-atomic reads, atomic reads, RMWs (which are always atomic), and memory fences. Where relevant, labels contain (reading across): the location being accessed, the value being read, the value being written, and the memory order specified by the programmer. A -mark on the right-hand side of the table indicates that an event with this label belongs to the set R (resp. W, F, A) of events that read (resp. write, are a fence, are atomic). Let L denote the set of labels.

Definition 5 (Executions). An execution is a tuple X = (E, I, lbl, thd, sb) with the following components.

- E is a set of event identifiers.
- lbl ∈ E → L associates each event with a label. For each event e, loc(e) projects the loc attribute of lbl(e) (if applicable), and so on.
- I ⊆ E is a set of initial events. Every initial event e ∈ I is a non-atomic write of zero; that is, kind(e) = Wna and rval(e) = 0. Moreover, there is exactly one initial event per location.
- thd ⊆ (E \ I)2 is an equivalence relation on non-initial events that relates events from the same thread.
- sb ⊆ thd is the sequenced-before relation: a strict partial order (i.e., irreflexive and transitive) between events from the same thread, that captures the program order.

Let X be the set of all executions. Next, we define a number of derived sets and relations over the events of an execution that will prove useful in describing the memory model.

Definition 6 (Derived sets and relations). In the context of an execution (E, I, lbl, thd, sb), we define the relation =loc as {l | {e, e′} ∈ E \ F2 | loc(e) = loc(e′)}; it holds between non-fence events that access the same location. The relation =val, defined as {l | {e, e′} ∈ X × R | rval(e) = rval(e′)}, holds when the first event writes the value that the second reads. For each memory order o, we abbreviate the set {e ∈ A | ord(e) = o} as just o. We also define the set nal = {e ∈ E \ F | type(loc(e)) = non-atomic} of events that access a non-atomic location.

Example 2 (A C11 execution). The diagram below depicts one execution of the program given in Example 1. The initial events, a and b, are placed above the events of the four parallel threads, and reflexive and transitive edges are elided.

```
<table>
<thead>
<tr>
<th>a: Wna(x, 0)  b: Wna(y, 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c: W(x, 1, RLX)  d: R(x, 1, RLX)</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>f: W(x, 2, SC)  h: W(y, 1, SC)</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>thd↓↓sb  thd↓↓sb</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>e: R(x, 2, RLX)  g: R(y, 0, SC)</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>thd↓↓sb  thd↓↓sb</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>i: R(x, 1, SC)</td>
</tr>
</tbody>
</table>
```

Basic executions. The first stage of the C11 semantics translates a program into its basic set of executions. Each execution in this set is compatible with the instructions of the individual threads, but the set is constructed without considering the behaviour of shared memory, so it provides only an approximation of the executions that will ultimately be allowed to happen once the whole program and the memory model are taken into account. For instance, the execution in Example 2 is a basic execution of the program in Example 1: the values of the write events correspond to the program text, but the values of the read events are arbitrary and the basic set of all executions ranges over all choices. We do not define formally how the basic executions are constructed, and simply assume their existence for any program we wish to consider. Practical tools such as HERD and Cppmem [5] implement this construction as part of

---

3 This set is sometimes called the ‘pre-executions’ [5] or the ‘opsem’ [31].
Candidate executions. The second stage of the C11 semantics, which is the focus of this paper, takes as input a program’s basic execution set and returns the set of allowed executions. In order to build the allowed executions, we employ an intermediate structure called a candidate execution, which extends an execution with a witness that comprises three additional relations, called rf (read-from), mo (modification order) and S (sequential consistency order).

Definition 7 (Candidate executions). A candidate execution is a pair \((X, w)\) where \(X = (E, I, lbl, thd, sb)\) is an execution, and \(w = (rf, mo, S)\) is a witness comprising three relations \(rf, mo, S \subseteq E^2\). A candidate execution is well-formed, written \(wf(X, w)\), if:

- every read event observes exactly one write event, and the locations and values match; that is,
  \[
  \forall e \in R. \exists! e' \in W. (e', e) \in rf
  \text{ and } \quad rf \subseteq \mathcal{I}_{loc} \cap \mathcal{I}_{val}
  \]

- where \(\mathcal{I}\) means ‘exists unique’;

- the modification order relates, in a strict total order, all and only those events that write atomically to the same location; that is,
  \[
  (mo \cap mo^{-1}) = (\mathcal{I}_{loc} \cap W^2 \cap A^2 \setminus \text{id})
  \text{ and } \quad acy(mo)
  \]

- the \(S\) relation relates, in a strict total order, all and only the SC events in an execution; that is,
  \[\text{acy}(S) \text{ and } (S \cup S^{-1}) = (SG^2 \setminus \text{id})\] (WFs)

Example 3 (A C11 candidate execution). The diagram below the extends the execution in Example 2 with a witness. We elide the \(thd\) edges (each column corresponds to one thread). The candidate execution is well-formed, and consistent with the axioms of the memory model (presented next).

\[
\begin{align*}
a: W_{na}(x, 0) & \quad b: W_{na}(y, 0) \\
c: W(x, 1, RLX) & \quad d: R(x, 1, RLX) & f: W(x, 2, SC) & \quad h: W(y, 1, SC) \\
e: R(x, 2, RLX) & \quad g: R(y, 0, SC) & & \quad i: R(x, 1, SC)
\end{align*}
\]

2.3 C11 axioms

A candidate execution is deemed consistent with the memory model if it satisfies the 12 consistency axioms of Def. 11. We express these axioms using the .cat language [2], which means that the axioms are expressed in a concise language based on the propositional fragment of Tarski’s relation calculus [28].

Definition 8 (The .cat language). The cat language supports the definition of relations using the operators union, intersection, difference, complement (¬), inverse \((r^{-1})\), reflexive closure \((r^r)\), transitive closure \((r^+)\), and relational composition \((.\)) which is defined such that \(x, z \in r_1 \cap r_2\) if \((x, y) \in r_1\) and \((y, z) \in r_2\) for some \(y\). It also provides the syntax \([s] = \{(e, e) \mid e \in s\}\) for the identity relation \((id)\) restricted to the set \(s\). Each axiom of the memory model must be expressed in the form of an acyclicity \((acy\ r)\), irreflexivity \((irr\ r)\), or emptiness \((empty\ r)\) constraint on some relation \(r\) constructed using these operators.

In order to define these axioms, we need first to introduce several derived relations.

Remark 9. In the following, we justify our formal definitions by referring to the C11 standard [15], using the notation \(\text{§}N:n\) for section \(N\), paragraph \(n\). We refer to the C++11 standard [14], whenever a clause was erroneously omitted from C11. (C11 inherits its memory model from C++11. Similarly, we refer to the C++14 standard [16] in the case of an erroneous omission from C++11. We include these omitted parts because doing so leads to a cleaner model that we believe to be closer to the designers’ intent.

Definition 10 (Further derived sets and relations). In the context of a candidate execution \((E, I, lbl, thd, sb, rf, mo, S)\), we define the following subsets of \(E\) and relations over \(E\):

\[
\begin{align*}
acq & \equiv (ACU AR \cup SC) \cap (R \cup F) \\
rel & \equiv (RELU AR \cup SC) \cap (W \cup F) \\
rh & \equiv (rf^{-1} ; mo) \setminus \text{id} \\
Fsb & \equiv [F] ; sb \\
sb & \equiv sb ; [F] \\
rs' & \equiv thd \cup (E^2 \cap [R \cap W]) \\
rS & \equiv mo \cap rs' \setminus ((mo \cap rs') ; mo) \\
sW & \equiv ([rel] ; Fsb; \ [W \cap A]) ; rs' ; rf ; \\
rsW & \equiv [W \cap A] ; sbF^2 ; [acq] \setminus thd \\
hb & \equiv (sb \cup (I \times -I) \cup sw) \\
hb & \equiv hb \cap \mathcal{I}_{loc} \\
vis & \equiv (W \times R) \cap hb \setminus (hb \setminus [W] ; hb) \\
cnf & \equiv ((W \times X) \cup (W \times R) \cup (R \times X)) \cap \mathcal{I}_{loc} \\
Dr & \equiv cnf \setminus hb \setminus hb^{-1} \setminus A^2 \cap thd
\end{align*}
\]

Commentary. The set \(acq\) (resp. \(rel\)) contains all events that behave as an acquire (resp. a release).

The relation \(rs\) captures the release sequence, using \(rs'\) as a helper. The release sequence of \(e\) comprises those events that form a maximal \(mo\)-chain, starting from \(e\), of events that either are in \(e\)'s thread or are RMWs.

Release/acquire synchronisation is captured by the \(sw\) relation. This relates an atomic write-release event to an atomic read-acquire event in a different thread if the read obtains its value from the write or its release sequence. If the acquire (resp. release) is a fence, the synchronisation happens via an atomic read (resp. write) sequenced before (resp. after) the fence.

Happens-before \((hb)\) is a transitive relation that includes sequenced-before and synchronisation edges, and puts initial events before all other events. A write is visible \((vis)\) to a read if it is the most recent write to that location in happens-before.

Two events are in conflict \((cnf)\) if they access the same location and at least one is a write; these events go on to form a data race \((dr)\) if they are unrelated by happens-before, they are not both atomic, and they are in different threads.

We now use the derived relations of Def. 10 to formalise what it means for an execution to be consistent.

Definition 11 (Consistency). A candidate execution \((X, w) = (E, I, lbl, thd, sb, rf, mo, S)\) is consistent, written \(\text{consistent}(X,\ w)\),

\[\text{[15 (§7.17.3.3–4)], [15 (§7.17.4.1.2)] \quad [6 (§5.3)] \quad [7 (§5.1.2.4), [15 (§7.17.4.1.2)] \quad [8 (§5.1.2.4.11)] \quad [9 (§7.17.4.2–4)] \quad [10 (§5.1.2.4.18)], simplified in the absence of memory_order_consume \quad [11 (§5.1.2.4.19)] \quad [12 (§5.1.2.4.4)] \quad [13 (§5.1.2.4.25)]\]
w), if it is well-formed and it satisfies all of the following axioms:

\[ \text{irr}(hh) \] (Hb)
\[ \text{irr}(rf^{-1}) \cap: mo; rf^* ; hh \] (Coh)
\[ \text{irr}(rf ; hh) \] (Rf)
\[ \text{empty}(rf \cap: [nll] \backslash vis) \] (NaRf)
\[ \text{irr}(rf \cup (mo ; mo ; rf^{-1}) \cup (mo ; rf)) \] (Rmw)
\[ \text{irr}(S ; r_1) \] where \( r_1 = hh \) (S1)
\[ \text{irr}(S ; r_2) \] where \( r_2 = Fsb^* ; mo; sbF^* \) (S2)
\[ \text{irr}(S ; r_3) \] where \( r_3 = rf^{-1} \cap [SC] ; mo \) (S3)
\[ \text{irr}(S \backslash (mo ; S) ; r_4) \] where \( r_4 = rf^{-1} \cap hbl ; [W] \) (S4)
\[ \text{irr}(S ; r_5) \] where \( r_5 = Fsb ; rb \) (S5)
\[ \text{irr}(S ; r_6) \] where \( r_6 = rb ; sbF \) (S6)
\[ \text{irr}(S ; r_7) \] where \( r_7 = Fsb ; rb ; sbF \) (S7)

Commentary: These axioms are equivalent to the corresponding definitions in Batty et al.'s Lem formalisation [6], the fidelity of which has been endorsed by the C11 standards committee. Being expressed in the .cat language, however, they are markedly more concise. We now explain each in turn.

Happens-before must contain no cycles.\(^{14}\) Requiring irreflexivity here is sufficient (Hb), since \( hh \) is transitive. Coherence (Coh) governs the relationship between \( hh \) and \( mo \): if the write \( e_1 \) is \( mo \)-before the write \( e_2 \), then \( e_2 \) (and any events that read from \( e_2 \)) must not happen before \( e_1 \) (nor any events that read from \( e_1 \)).\(^{15}\) A read must not observe a write that happens after it (Rf).\(^{16}\) and a read of a non-atomic location must observe a visible write (NaRf).\(^{11}\) An RMW must observe the immediately-preceding write in \( mo \) (Rmw).\(^{17}\) that is, not itself (first disjunct), nor a too-early write (second disjunct), nor a too-late write (third disjunct).

This leaves the SC axioms, which we present using where-clauses for ease of reference later. Axiom S1 states that \( S \) must be consistent with happens-before.\(^{18}\) Axiom S2 governs the relationship between \( S \) and \( mo \): if the write \( e_1 \) is \( mo \)-before the write \( e_2 \), then \( e_2 \) (and any fences sequenced after \( e_2 \)) must not come before \( e_1 \) (nor any fences sequenced before \( e_1 \)) in \( S \).\(^{19}\)

Axioms S3 and S4 constrain the values that an SC read \( e_1 \) of a location \( l \) may observe. If there are any SC writes to \( l \) preceding \( e_1 \) in \( S \), then \( e_1 \) must read either from the most recent of these in \( S \) — call this \( e_2 \) — or from a non-SC write that does not happen before \( e_2 \).\(^{18}\) We encode this requirement as two irreflexivity constraints. First, we wish to rule out reading from an SC write that is not the most recent in \( S \); that is, we wish to forbid cycles of the shape depicted below left, where \( S_{loc} \equiv S \cap =_{loc} \). Axiom S3 does this, using the simplified form shown below right.

Second, we require \( e_1 \) not to read from a write that happens before \( e_2 \); that is, we wish to forbid cycles of the shape depicted below left. Axiom S4 does this, using the simplified form shown below right.

\[ \text{irr}(S ; r_1) \] (S1)
\[ \text{irr}(S ; r_2) \] (S2)
\[ \text{irr}(S ; r_3) \] (S3)
\[ \text{irr}(S \backslash (mo ; S) ; r_4) \] (S4)
\[ \text{irr}(S ; r_5) \] (S5)
\[ \text{irr}(S ; r_6) \] (S6)
\[ \text{irr}(S ; r_7) \] (S7)

from a write to \( l \) that is \( mo \)-earlier than the last write to \( l \) that precedes the fence in \( S \).\(^{20}\) In fact, the ‘last write’ here can be safely generalised to ‘some write’, because being \( mo \)-earlier than some write to \( l \) that precedes the fence in \( S \) implies being \( mo \)-earlier than the last write, since \( mo \) is total (S5). If a write \( e_2 \) to location \( l \) is sequenced before an SC fence, then any SC read of \( l \) that follows the fence in \( S \) must not read from a write to \( l \) that is \( mo \)-earlier than \( e_2 \) (S6).\(^{21}\) Finally, if a read \( e_1 \) of location \( l \) is sequenced after an SC fence, and a write \( e_2 \) to \( l \) is sequenced before another SC fence that precedes the first fence in \( S \), then \( e_1 \) must not read from a write \( mo \)-earlier than \( e_2 \) (S7).\(^{22}\)

A final axiom formalises what it means for an execution to exhibit a fault.

Definition 12 (Faultiness). A candidate execution \((X, w)\) is faulty, written faulty\((X, w)\), if it is consistent and does not satisfy the following axiom:

\[ \text{empty}(dr) \] (Dr)

If any execution can be extended to a faulty candidate execution, then the entire program’s behaviour is ‘undefined’ and any execution is allowed. Otherwise, the allowed executions are those executions that can be extended to a consistent candidate execution.

Definition 13 (Allowed executions). Given a set \( X \) of a program’s executions, we obtain the program’s allowed executions as:

\[ \text{allowed}(X) \equiv \begin{cases} \text{if } \exists X \in X, \exists w. \text{ faulty}(X, w) \text{ then } X \\ \text{else } \{ \exists X \in X \mid \exists w. \text{ consistent}(X, w) \} \end{cases} \]

3. Overhauling the SC axioms in C11

The rules for SC axioms in C11, as demonstrated in the previous section, are highly convoluted. In this section, we describe how these rules can be improved in two fairly orthogonal ways. In §3.1, we describe how the total order over SC operations can be replaced with a partial order; this simplification will be demonstrated in §6.2 to dramatically improve the efficiency with which the model can be simulated. In §3.2, we describe a slight strengthening of the model that enables significant simplifications to be made. These simplifications lead to a model that is easier to understand, and should prove easier to work with in a formal setting.

3.1 Reducing \( S \) from a total to a partial order

We observe that all but one of the seven SC axioms (Def. 11) can be written in the form \( \text{irr}(S ; r) \) for some relational expression \( r \). These \( r \)'s can be seen as the constraints on the total order \( S \). Axiom S4 is not quite of this form. However, replacing its ‘\( S \backslash (mo ; S) \)’ with just ‘\( S \)’, to obtain the axiom S4a given below, happens to coincide exactly with an amendment to the model already proposed by Vafeiadis et al. to lend the model more desirable mathematical properties [31 (§4.2)].

\[ \text{irr}(S ; r_1) \] (S4a)

Where axiom S4 forbids an SC read to observe any write that happens before the most recent SC write in \( S \), axiom S4a forbids it to observe any write that happens before any SC write in \( S \). Let us assume here that their uncontroversial amendment will be accommodated by the C standards committee. We are now in a position to replace the seven irreflexivity axioms with a single acyclicity axiom.

\[ \text{Theorem 1. There exists a strict total order on SC events that satisfies axioms S1, S2, S3, S4a, S5, S6, and S7, if and only if} \]

\[ [15 (§7.17.3:9)] \]

\[ [15 (§7.17.3:10)] \]

\[ [15 (§7.17.3:11)] \]
the following $S_{\text{partial}}$ axiom (which states that the union of all the constraints on $S$, when restricted to SC operations, is acyclic) holds:

\[ \text{acy}(SC^2 \cap (r_1 \cup r_2 \cup r_3 \cup r_4 \cup r_5 \cup r_6 \cup r_7) \setminus \text{id}) \]  \hspace{1cm} (S_{\text{partial}}) \]

Proof. We are to prove

\[ (\exists S. \text{WfS} \land S1 \land S2 \land S3 \land S4a \land S5 \land S6 \land S7) = S_{\text{partial}}. \]

In doing so, we make use of the order extension principle: that any (strict) partial order can be extended to a (strict) total order.

\[ \exists S. \text{WfS} \land S1 \land S2 \land S3 \land S4a \land S5 \land S6 \land S7 = [\text{basic properties of relations}] \]

\[ \exists S. \text{WfS} \land \text{irr}(S; (r_1 \cup \ldots \cup r_7)) = [\text{since } S \text{ is a strict total order on SC}] \]

\[ \exists S. (\text{WfS} \land (SC^2 \cap (r_1 \cup \ldots \cup r_7) \setminus \text{id}) \subseteq S) = [\text{by the order extension principle}] \]

\[ S_{\text{partial}} \]

Having replaced axioms S1–S7 with the new $S_{\text{partial}}$ axiom, we no longer require the $S$ relation in execution witnesses. Memory model simulators, such as HERD, typically work by enumerating all executions of a program and then filtering out the consistent subset. Removing the need to iterate through all possible total orders of SC events – a computation that is exponential in the number of SC events – allows simulation performance to be greatly improved, as demonstrated in §6.2.

### 3.2 A stronger and simpler SC axiom

We now show that it is possible to strengthen the SC semantics without requiring changes to the compilation schemes of any of the C11 target architectures that have an established formal memory model, that is: x86 and Power. The strengthening we propose simplifies the $S_{\text{partial}}$ axiom significantly and provides stronger guarantees to the programmer.

The proposal for this simplification arises from the observation that the relations considered in the $S_{\text{partial}}$ axiom are nearly symmetric in $hb$, $mo$ and $rb$. In particular, both $hb$ and $mo$ constrain the $S$ order between any combination of SC fences and atomics. The treatment of $rb$ is different: for $rb$ edges that begin or end at a fence, the axioms $S5$, $S6$ and $S7$ ensure that the SC order is constrained to match. When two SC atomics are related by an $rb$ edge ($S3$ and $S4$), ordering is only provided when the intermediate access that forms the $rb$ is itself an SC atomic (rule 3S), or when the $mo$ edge from the intermediate access of the $rb$ to its target is also covered by a $hb$ edge (rule 4a).

Our proposal is to strengthen the $S_{\text{partial}}$ axiom, to add these missing constraints so that every $rb$ edge between SC atomics contributes to the $S$ order. We achieve this in our model by removing the $SC$ restriction from $S3$, which results in the following axiom:

\[ \text{irr}(S; rb). \]  \hspace{1cm} (S3a)

This change permits a significant simplification to the SC rules that we establish in the following theorem.

**Theorem 2.** If rule S3 is replaced by S3a (that is, if $r_3$ is replaced with $rb$ in the $S_{\text{partial}}$ axiom) then $S_{\text{partial}}$ becomes equivalent to:

\[ \text{acy}(SC^2 \cap (Fsb^7 \cup (hb \cup rb \cup mo) \cup sbF^7)). \]  \hspace{1cm} (S_{\text{simp}})

**Proof.** First note that

\[ r_1 \cup r_2 \cup rb \cup r_4 \cup r_5 \cup r_6 \cup r_7 = [\text{unfolding definitions and combining } rb, r_3, r_6 \text{ and } r_7] \]

\[ hb \cup (Fsb^7 \cup mo \cup sbF^7) \cup (Fsb^7 \cup rb \cup sbF^7) \cup r_4 = [\text{since } r_4 \subseteq rb, \text{ by WfMo}] \]

\[ hb \cup (Fsb^7 \cup mo \cup sbF^7) \cup (Fsb^7 \cup rb \cup sbF^7) = [\text{since } hb = (Fsb^7 \cup rb \cup sbF^7)] \]

\[ Fsb^7 \cup (hb \cup rb \cup mo) \cup sbF^7. \]

This gives us the axiom $\text{acy}(SC^2 \cap (Fsb^7 \cup (hb \cup rb \cup mo) \cup sbF^7) \setminus \text{id})$, which can be simplified to $S_{\text{simp}}$, because $Fsb^7 \cup (hb \cup rb \cup mo) \cup sbF^7$ is already irreflexive.

**Programming impact.** The change presented here does strengthen the memory model; there are executions that were previously allowed that are now forbidden. The simplest we found, which is similar to one used by Vafeiadis et al. [31 (Fig. 6)], is presented in Example 3. We believe Example 3 to be a counterintuitive execution, because the read event $i$ does not observe the most recent write to $x$ in $S$ (namely, $f$), but $c$, which is $mo$-earlier than $f$. The execution is forbidden by axiom S3a because of its $f \xrightarrow{rb} i \xrightarrow{rb} c$ cycle. Although the current C11 model allows this execution, mapping this example to the formalised targets of C11 (Power and x86) never yields programs that exhibit it.

### 3.3 Soundness of existing C11 compilation schemes

There are two C11 targets with formal architectural memory models: x86 and Power. In this subsection, we establish that for both of these architectures, the strengthening does not require a stronger compilation mapping. In both cases, we rely on an existing proof of soundness from the literature. We need only establish that our strengthened $S_{\text{simp}}$ axiom holds.

To establish the soundness of our strengthening for x86, we build on the soundness proof of Batty et al. [5], which uses the axiomatic model of x86 of Owens et al. [24]. To obtain soundness for Power, we build on the soundness proof of Batty et al. [6], which uses the operational Power model of Sarkar et al. [25].

**Theorem 3.** Let $P$ be a C11 program that has no faulty executions. If we compile $P$ to x86 according to the mapping given by Batty et al. [5], then every valid x86 execution corresponds to a C11 execution where $S_{\text{simp}}$ holds. If we compile $P$ to Power according to the mapping given by Batty et al. [6], then every valid Power trace is observationally equivalent to a C11 execution where $S_{\text{simp}}$ holds.

[Proof in §C]

In summary, this section has described how, having strengthened the original set of axioms (S1 through S7) to use Vafeiadis et al.’s S4a in place of S4, the behaviour of SC operations can be captured by a single axiom ($S_{\text{partial}}$) that allows the total order $S$ to be eliminated from the model. Moreover, if the axioms are further strengthened to use our S3a in place of S3, then that axiom can be greatly simplified ($S_{\text{simp}}$), while still respecting current compilation schemes. We include in §A.2 a suggestion for how the wording of the standard could be changed to accommodate our proposal; our wording is considerably shorter without being stylistically different from the original text.

### 4. Formalising the OpenCL memory model

A principal aim of the OpenCL initiative is to provide functional portability across a plethora of heterogeneous many-core devices. The standard is implemented by CPU, GPU and FPGA vendors,
and aims to allow applications to be device-agnostic. The OpenCL memory model, introduced in the 2.0 revision of the standard, is inherited from that of C11, but is specialised and extended for heterogeneous programming. The memory model is the sole mechanism for correctly implementing fine-grained concurrent algorithms in a device-agnostic manner. Rigorous foundations for this model are thus vital.

We now describe how our formalisation of the C11 memory model (§2, §3) can be extended to yield the first mechanised formalisation of the full OpenCL memory model. We describe the form of OpenCL programs (§4.1), their executions (§4.2), and the axioms against which these executions are judged (§4.3). We then discuss some interesting features of the memory model: some innocuous quirks (§4.4) and some serious shortcomings (§4.5). The most serious shortcoming relates to the axioms that govern SC atomics, and we propose how to fix this in §5.

For reasons of space, and because they are orthogonal to the thrust of our contributions, we omit our treatment of barrier synchronisation operations and the associated issue of barrier divergence. As with the omitted C11 features mentioned in §2, our cat-based formalisation of the OpenCL memory model, provided as an appendix (§E), fully accounts for these features.

### 4.1 OpenCL programs

**Definition 14 (Structure of OpenCL programs).** Building on Def. 2, we consider OpenCL programs of the form

\[ P = \prod_{d \in D} \prod_{w \in W} \prod_{t \in T} p_{d,w,t} \]

where \( D, W, \) and \( T \) are sets of devices, work-group, and thread identifiers,\(^{23}\) and each \( p_{d,w,t} \) is a piece of sequential code.

Using the notation above, we can write \( p \parallel p' \) to denote a litmus test comprising two threads to be executed on different devices, \( p \parallel p' \) for two threads in different work-groups in the same device, and \( p \parallel p' \) for two threads in the same work-group. We can also write, for example, \( p_1 \parallel p_2 \parallel p_3 \parallel p_4 \parallel p_5 \parallel p_6 \parallel p_7 \parallel p_8 \), to denote a litmus test comprising two devices, each executing two work-groups, each containing two threads.

**Remark 15 (Limitations).** This program structure does not account for sub-groups, an optional extension in OpenCL 2.0 that allows threads to synchronise with one another at a level of granularity finer than that of a work-group.\(^ {24}\) Nor for further non-OpenCL threads (e.g., POSIX threads) running on the host platform. Moreover, \( w \) and \( t \) can actually be 1-, 2-, or 3-dimensional vectors, but we make the simplifying assumption that all identifiers are natural numbers.

Recall that locations in C11 are either non-atomic or atomic (Def. 1). OpenCL locations are further declared to reside in a memory region.

**Definition 16 (Memory regions).** We have \( \text{region}(l) \in \{\text{local}, \text{global}, \text{global}_\text{fgb}\} \) for every location \( l \), where \( \text{fgb} \) stands for fine-grained shared virtual memory (SVM) buffer.\(^ {25}\) There is one \text{local} region per work-group, containing locations accessible only to that work-group. Locations in the \text{global} or \text{global}_\text{fgb} region are accessible to all devices. Fences can be performed on either the global or the local memory, or both simultaneously.

The distinction between \text{global} and \text{global}_\text{fgb} locations is that the former must not be shared between different devices, whereas the latter enable inter-device communication. Unlike C11, in which any memory location can be shared between threads, the OpenCL memory model prevents certain sharing patterns. For instance, threads from different devices are forbidden from conflicting on global memory, but are able to do so as a result of a programmer fault; in contrast, threads from different work-groups are unable to conflict on local memory: the language provides no mechanism through which such a conflict can arise.

**Definition 17 (Memory scopes).** Atomics in OpenCL are parameterised by a memory scope. The three options are

\[ s : = \text{WG} \quad \text{(work-group scope)} \]
\[ \text{DV} \quad \text{(device scope)} \]
\[ \text{ALL} \quad \text{(system scope)} \]

A memory scope specifies how widely visible the effects of the operation should be.

**Example 4.** The use of memory scopes is illustrated by the following code, which implements the message-passing idiom between two threads in the same work-group.

\[
\begin{align*}
global \text{ int } *x; \; & \text{global atomic int } *y; \\
*x &= 42; & \text{if}(y, \text{ACQ}, \text{WG}) == 1) \\
\text{store}(y, 1, \text{REL}, \text{WG}); & \quad r = *x;
\end{align*}
\]

Since all accesses to the global location \( y \) come from the same work-group, those accesses can be performed at \( \text{WG} \) scope (which means that on implementations where each work-group caches global memory, it suffices to read/write those cached values). This scope would be insufficient, and the program deemed faulty, if the threads were in different work-groups – both scopes would have to be upgraded to \( \text{DV} \).

### 4.2 OpenCL executions

OpenCL executions extend C11 executions as follows.

**Definition 18 (OpenCL event labels).** We extend C11 event labels (Def. 4) with an additional scope attribute, which assigns a memory scope \( s \) to all atomic events. We also subdivide the F label in order to represent fences on global (\( F_G \)), local (\( F_L \)) and both-global-and-local memory (\( F_{GL} \)). The updated table is as follows:

<table>
<thead>
<tr>
<th>kind</th>
<th>loc</th>
<th>rval</th>
<th>wval</th>
<th>ord</th>
<th>scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_{na} )</td>
<td>(l, v,</td>
<td>o, s)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>( W )</td>
<td>(l, v,</td>
<td>a, s)</td>
<td></td>
<td></td>
<td>✓ ✓</td>
</tr>
<tr>
<td>( R_{na} )</td>
<td>(l, v,</td>
<td>o, s)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>( R )</td>
<td>(l, v,</td>
<td>a, s)</td>
<td></td>
<td></td>
<td>✓ ✓</td>
</tr>
<tr>
<td>( RMW )</td>
<td>(l, v,</td>
<td>a, s)</td>
<td></td>
<td></td>
<td>✓ ✓</td>
</tr>
<tr>
<td>( F_G )</td>
<td>(o, s)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>( F_L )</td>
<td>(o, s)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>( F_{GL} )</td>
<td>(o, s)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

**Definition 19 (OpenCL executions).** An OpenCL execution is a tuple \((E, I, \text{bb}, \text{thd}, \text{wg}, \text{dv}, \text{sb})\) where \((E, I, \text{bb}, \text{thd}, \text{sb})\) is a C11 execution as in Def. 5, and \( \text{wg}, \text{dv} \subset (E \setminus I)^2 \) are equivalence relations on non-initial events that relate events from the same work-group and device, respectively. In order to enforce the privacy of local locations to a single work-group, we require that if \( \text{loc}(e) = \text{loc}(e') = l \) and \( \text{region}(l) = \text{local} \), then \( (e, e') \in \text{wg} \).

**Definition 20 (Derived sets and relations).** In the context of an OpenCL execution \((E, I, \text{bb}, \text{thd}, \text{wg}, \text{dv}, \text{sb})\), we define

\[
\begin{align*}
f_{\text{fgb}} & \seteq \{ e \in E \setminus F \mid \text{region}(\text{loc}(e)) = \text{global}_\text{fgb} \} \\
G & \seteq \{ e \in F \mid \text{kind}(e) \in \{F_G, F_{GL}\} \} \cup \\
& \quad \{ e \in E \setminus F \mid \text{region}(\text{loc}(e)) = \text{global} \} \cup \text{fgb} \\
L & \seteq \{ e \in F \mid \text{kind}(e) \in \{F_L, F_{GL}\} \} \cup \\
& \quad \{ e \in E \setminus F \mid \text{region}(\text{loc}(e)) = \text{local} \}
\end{align*}
\]
as the sets of events that access, respectively: fine-grained atomic SVM buffers, global memory, and local memory. Also, for each scope \(s\), we abbreviate the set \(\{e \in A \mid \text{scope}(e) = s\}\) as just \(s\).

**Definition 21** (OpenCL candidate executions). Candidate executions in OpenCL, and their well-formedness, are defined in the same way as in C11 (Def. 7).

### 4.3 OpenCL axioms

We now define and discuss the consistent and faulty predicates for the OpenCL memory model, paying particular attention to each of the departures from C11. We justify our formal definitions by reference to the OpenCL specification [17], writing \(n/m\) to denote line \(n\) on page \(m\).

**Definition 22** (Further derived sets and relations). In the context of a candidate execution \((E, l, lbd, thd, sb, rf, mo, S)\), we define the following subsets of \(E\) and relations over \(E\):

- \(\text{incl}\) \(\equiv (W^2 \cap ug) \cup (D^2 \cap df) \cup \text{ALL}^2\)
- \(\text{rsuw}(r)\) \(\equiv ([r \cap rel] ; \text{Fsh}^2 ; [W \cap A] \cup [r] ; [r] ; rf \cup [R \cap A] ; \text{sbF}^2 ; [r \cap acq]) \cap \text{incl} \cap \text{thd}\)
- \(\text{scf}\) \(\equiv \text{SCf}^2 \cup (G \cap L \cap F)^2\)
- \(\text{gsw}\) \(\equiv \text{rsuw}(G) \cup \text{scf} \cap \text{rsuw}(L)\)
- \(\text{lsw}\) \(\equiv \text{rsuw}(L) \cap \text{scf} \cap \text{rsuw}(G)\)
- \(\text{ghb}\) \(\equiv (G^2 \cap (sb \cup (I \times I))) \cup \text{gsw}^+\)
- \(\text{lhb}\) \(\equiv (L^2 \cap (sb \cup (I \times I))) \cup \text{lsw}^+\)
- \(\text{ghbl}\) \(\equiv \text{ghb} = \text{incl}\)
- \(\text{lhbl}\) \(\equiv \text{lhb} = \text{incl}\)
- \(\text{gvis}\) \(\equiv (W \times R) \cap \text{ghbl} \setminus (\text{ghbl} ; [W] ; \text{ghb})\)
- \(\text{lvis}\) \(\equiv (W \times R) \cap \text{lhbl} \setminus (\text{lhbl} ; [W] ; \text{lhb})\)
- \(\text{hr}\) \(\equiv \text{cnf} \setminus (\text{ghb} \cup \text{lhb}) \setminus (\text{ghbl} \cup \text{lhbl}) \cup \text{incl} \cap \text{thd}\)
- \(\text{iddr}\) \(\equiv \text{cnf} \setminus (G^2 \cap \text{fgb}^2)\)
- \(\text{sc-all}\) \(\equiv -((E^2 ; [SC \setminus (\text{ALL} \cap \text{fgb})]) ; E^2)\)
- \(\text{sc-dv}\) \(\equiv -((E^2 ; [SC \setminus (DV \cap \text{fgb})]) ; E^2)\)

**Commentary.** In OpenCL, only events that have inclusive scopes (incl) can synchronise: either the events have \(W^2\) scope and are in the same work-group, or they have \(D^2\) scope and are in the same device, or they have ALL scope.\(^{26}\) We shall explain in §4.5 how this notion of scope inclusion is unnecessarily conservative.

The synchronisation relation (rsuw) is parameterised by a region \(r\) (global or local). The global synchronises-with relation (gsw) includes events that synchronise on global memory;\(^{27}\) but also includes events that synchronise on local memory, providing both events have memory order \(SC\),\(^{28}\) or both are global-and-local fences.\(^{29}\) Local synchronises-with (lsw) is analogous. Example 6 shows how synchronisation works in the presence of global-and-local fences.

Happens-before is partitioned into global and local versions: global happens-before (ghb) contains global synchronises-with and sequenced-before edges between events on global memory,\(^{30}\) and local happens-before (lhb) is analogous.\(^{31}\) See Example 5 for a discussion of the repercussions of this definition of happens-before. Visibility is also split into global (gvis) and local (lvis) versions.\(^{32}\)

The heterogeneous race (hr)\(^{33}\) generalises C11’s data race (dr, Def. 10), to reflect the fact that in OpenCL, even atomic operations can race when memory scopes are used incorrectly.\(^{34}\) If two events from different devices conflict on a location that is not in a fine-grained atomic SVM buffer, then they form an *inter-device data race* (iddr); such races cannot be ruled out by happens-before edges.\(^{35}\)

This leaves the *sc-all* and *sc-dv* relations. In OpenCL, the total order \(S\) is only required to exist when \(SC \subseteq \text{ALL} \cap \text{fgb} \) or \(SC \subseteq \text{DV} \setminus \text{fgb}\).

The first condition holds when every SC event has ALL scope and accesses a global,\(^{36}\) the second holds when every SC event has DV scope and does not access a global.\(^{37}\) The relation *sc-all* (resp. *sc-dv*) is the universal relation if the first (resp. *sc-dv*) condition holds and is the empty relation otherwise. In §4.5, we shall criticise these conditions as being simultaneously too strong for programmers and too weak for compiler-writers.

**Definition 23** (Consistency axioms in OpenCL). There are nine consistency axioms. Departures from the C11 consistency axioms (Def. 11) are highlighted.

- \(\text{irr}(ghb)\) (O-HbG)
- \(\text{irr}(lhb)\) (O-HbL)
- \(\text{irr}((rf^{-1})^3 ; mo ; rf^{-1}) ; ghb)\) (O-CohG)
- \(\text{irr}((rf^{-1})^3 ; mo ; rf^{-1}) ; lhb)\) (O-CohL)
- \(\text{irr}(rf ; (ghb \cup lhb))\) (O-Rf)
- \(\text{empty}(rf \cup (G \cap \text{nal})) \setminus \text{gvis}\) (O-NaRG)
- \(\text{empty}(rf \cup (\text{L} \cap \text{nal})) \setminus \text{lvis}\) (O-NaRL)
- \(\text{irr}(rf \cup (mo ; rf^{-1}) \cup (mo ; rf))\) (O-Rmw)
- \(\text{acy}(SC^2 \cap (\text{sc-all} \cup \text{sc-dv}) \cap (\text{Fsh}^2 ; (rb \cup mo \cap ghb \cup lhb) ; \text{sbF}^2))\) (O-Simp)

**Commentary.** Both happens-before relations are required to be acyclic (O-HbG, O-HbL).\(^{38}\) OpenCL requires coherence for both global and local happens before separately (O-CohG, O-CohL).\(^{39}\) The axioms governing the reads-from relation are carried over from C11 (O-Rf, O-NaRfG, O-NaRFL, O-Rmw), but appropriately divided into global and local versions.\(^{40}\)

OpenCL defines the same SC axioms as we saw in Def. 11 (§1–7), but uses ghb \(\cup lhb\) in place of hb. We have incorporated into axiom O-Simp the simplifications that we already discussed in the context of C11 (§3). Intersecting with the *sc-all* and *sc-dv* conditions means that the acyclicity constraint is only enforced when one of those conditions holds.\(^{41}\)

**Definition 24** (Faultiness in OpenCL). A candidate OpenCL execution is faulty if it is consistent and does *not* satisfy both of the following axioms:

- \(\text{empty}(hr)\) (O-Hr)
- \(\text{empty}(iddr)\) (O-Iddr)

### 4.4 Quirks in the memory model

We present three worked examples that illustrate features of the memory model that may not be obvious from a cursory glance at its axioms. These ‘quirks’ in the model are distinguished from the technical shortcomings that we save for §4.5.

Our first example illustrates an interesting consequence of OpenCL’s separation of happens-before into two distinct relations.

**Example 5.** Suppose the code of Example 4 was changed so that \(y\) were declared local rather than global. Executions such as the one below would then become consistent, which means that a stale value

\(^{26}\) [17 (47/16–26)] \(^{27}\) [17 (51/1–9)] \(^{28}\) [17 (51/32–33)]
\(^{29}\) [17 (54/13–16)] \(^{30}\) [17 (49/3–7)] \(^{31}\) [17 (49/8–11)]
\(^{32}\) [17 (49/21–26)] \(^{33}\) This terminology is due to Hower et al. [13].
\(^{34}\) [17 (49/29–33)]
\(^{35}\) [17 (58/24–27)] \(^{36}\) [17 (51/5–17)] \(^{37}\) [17 (51/18–20)]
\(^{38}\) [17 (49/12–13)] \(^{39}\) [17 (50/11–24)] \(^{40}\) [17 (49/26–27)]
\(^{41}\) [17 (50/8–9)], [17 (52/22–23)]
of x can be read (event f), even when successful release/acquire
synchronisation (between d and e) has occurred.

\[ a: W_{na}(x, 0) \quad b: W_{na}(y, 0) \]
\[ c: W_{na}(x, 1) \quad d: W(y, 1, REL, WG) \]
\[ e: R(y, 1, ACQ, WG) \]
\[ f: R_{na}(x, 0) \]

This execution is consistent because the sb edges no longer induce
either variety of happens-before, since they link events that act
on different memory regions. Worse still, there is now a data race
between c and f, which renders the entire program undefined.

We learn from Example 5 that a flag in one memory region
cannot be used to protect data in another region. To address this
issue, OpenCL provides fences that act on both global and local
memory simultaneously. These are illustrated in Example 6.

**Example 6.** The following program uses relaxed (RLX) accesses
on the local flag y, relying instead on the fences to synchronise
the threads and enable the global data x to be passed.

```
global int *x; local atomic int *y;
*x = 42; if(load(y,RLX, WG))=1
fence(GL, REL, WG); \{ fence(GL, ACQ, WG); 
store(y,1,RLX, WG); \} 
```

The fence instructions successfully prevent the stale value of x
being read, because the following execution is inconsistent.

\[ a: W_{na}(x, 0) \quad b: W_{na}(y, 0) \]
\[ c: W_{na}(x, 1) \quad d: F_{GL}(REL, WG) \]
\[ e: W(y, 1, REL, WG) \quad f: R_{na}(x, 0) \]

The execution is inconsistent because it has a cycle \( h \xrightarrow{rf} a \xrightarrow{mo} c \xrightarrow{ghb} h \), in violation of O-CohG. Note that \( c \xrightarrow{ghb} h \) holds here
because, firstly, \((d, g)\) is in \( rs\) of \( L \) and hence in \( gsw \) and \( ghb \), and
secondly, \((e, d)\) and \((g, h)\) are both in \( sb \cap G^2 \) and hence in \( ghb \).

In Example 7, we illuminate the relationship between memory
scopes and non-atomic operations. Since scopes can be used to
limit atomic operations to certain groups of threads, it is tempting
to introduce an additional ‘work-item’ scope, \( WI \), and encode non-
atomic events as atomic events whose scope is limited to the current
thread. This would make the \( W_{na} \) and \( R_{na} \) labels redundant. An
ordinary data race can then be cast as a failure of scope inclusion.
However, the differences between non-atomic and atomic operations
go beyond racy behaviours, as we shall see in the following example.

**Example 7.** Consider the following load-buffering litmus test:

```
int *x, *y;
if (*x==1) *y=1; if (*y==1) *x=1;
```

The execution of this program that exhibits the relaxed behaviour,
in which both comparisons succeed (shown below left), is not
consistent: both of its reads observes that are not visible,
in violation of the NaRf axiom. If the non-atomic locations become
atomic and the non-atomic operations become work-item-scoped
atomics, then this relaxed behaviour (shown below right) becomes
consistent, since the NaRf restriction no longer applies.

4.5 Problems with the memory model

We present three shortcomings in the OpenCL memory model,
which we discovered as a direct result of our formalisation efforts.

**Scope inclusion is too strong.** The specification provides an
overly conservative notion of scope inclusion: two events only have
inclusive scopes if their scopes match exactly. This leads to such
surprises as the following example.

**Example 8.** Suppose the code of Example 4 were changed so that
the store to y now occurs at DW scope, but the load of y remains at
WG scope. Although the release scope is clearly ‘wide enough’, it
does not match the acquiring scope, so no synchronisation edge is
induced. This leads to two data races: both between the non-atomic
accesses of x, and between the ill-scoped atomic accesses of y.

A resolution proposed by Gaster et al. is to allow the annotated
scopes to differ, as long as both are sufficiently wide [12 (§3.1)].
This enables, for instance, a DW-scoped write to synchronise with a
WG-scoped read in the same work-group. The proposal can be
formalised in our framework by changing the definition of the incl
relation (Def. 22) as follows:

\[ incl \overset{\text{def}}{=} (WG) \cup ([DV] ; [dv]) \cup ([ALL] ; E^2) \]
\[ new-incl \overset{\text{def}}{=} incl \cap incl^{-1} \]

The idea here is to define a one-sided version of scope inclusion
first, so that \((e_1, e_2)\) is in incl if \( e_1 \) has a wide enough scope to
‘reach’ \( e_2 \). Requiring this to hold in both directions ensures that both
events have sufficient scopes, if not necessarily the same.

**The SC axioms are too weak.** As encoded in our O-Samp axiom
(Def. 23), SC operations in OpenCL are only guaranteed to provide
SC behaviour when one of the sc-all and sc-dev conditions holds.
Since these are conditions on the whole program, we have a “clear
composability problem” [12].

We find several reasons why these conditions are problematic.
First, they mean that the default memory scope (which is DV)
is not sufficient to ensure SC semantics in all situations. Second,
any program that includes a WG-scoped SC atomic, such as
store(x,1,SC,WG), immediately violates the conditions. Third,
the conditions are mutually exclusive, so a program that satisfies
sc-all can be combined with another that satisfies sc-dv, with the
result satisfying neither. Finally, consider the following example.

**Example 9.** The following program, comprising two threads in
different work-groups on the same device, has SC semantics, which
means that it cannot exhibit the relaxed behaviour \( r_0 = r_1 = 0 \):

```
global atomic int *x, *y;
```
```
a: store(x,1); \quad b: store(y,1);
```
```
if (!(*x==1)) *y=1; if (!(*y==1)) *x=1;
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
be invisible to the programmer, if they can see only the kernel code: the assignment of locations to SVM buffers occurs only on the host side, and such locations are only marked in a kernel as global.

The SC axioms are too strong. Following discussion with members of the Khronos OpenCL working group, we understand that the purpose of condition sc-dv is to enable efficient implementations of DV-scoped SC atomics. The intention of the condition is that if no SC atomic accesses memory shared between devices, they can be implemented without expensive inter-device synchronisation. It was thought not to matter that the specification requires implementations to establish a total order between SC events on different devices, because it is not possible to observe this order without creating an inter-device data race.

In fact, this is not the case. We present in Example 10 a program that satisfies condition sc-dv, and yet is still able to observe the order between SC events in different devices – even though these events are DV-scoped and access no memory shared between devices.

Example 10. Consider the following program, which comprises two devices, both executing two threads (stacked vertically). It can be thought of as a “twisted” version of the store-buffering test.

```
global atomic_int *x, *y;
global_flag atomic_int *z1, *z2;
store(x,1,SC,DV);
store(z1,1,REL,ALL);
r1 = load(z2,ACQ,ALL)?
    load(x,SC,DV); : 1;
```

Two threads in different devices write, using DV scope, to distinct global locations x and y, and then write to global_flag flags, using ALL scope, to signal that they are done. Meanwhile, two partner threads try to acquire these signals from the opposite device, and if they are successful, they read the location their partner (in the same device as they) wrote to. We are interested in whether these reads can both obtain 0; that is, whether the final state \( r1 = r2 = 0 \) is allowed. This final state could only be obtained via the following execution:

\[ a: W(x,1,SC,DV) \]
\[ b: W(z1,1,REL,ALL) \]
\[ c: R(z2,1,ACQ,ALL) \]
\[ d: R(x,0,SC,DV) \]
\[ e: W(y,1,SC,DV) \]
\[ f: W(z2,1,REL,ALL) \]
\[ g: R(z1,1,ACQ,ALL) \]
\[ h: R(y,0,SC,DV) \]

where the outer dotted rectangles delimit \( dv \) equivalence classes and the inner ones delimit \( lhd \) equivalence classes.

The execution is inconsistent, and therefore must be forbidden by a compiler. To see this, observe that each \( rf \) edge induces a synchronisation \( gsw \) edge, and hence global happens-before. Since \( sb \) edges also contribute to global happens-before, we obtain the cycle \( gsw \). This makes the execution fall foul of O-S\textsubscript{unp}, which is non-vacuous here because the condition sc-dv is satisfied.

That the execution in the example above is not allowed implies that OpenCL implementations must make the order of SC write operations visible to all devices, even when those writes are only performed with DV scope. In other words, the current phrasing of the OpenCL memory model demands too much from the compiler-writer to permit an efficient implementation of DV-scoped SC atomics, while in other respects offering too little to the programmer.

The SC axioms are too strong. Building on an informal suggestion by Gaster et al. [12 (§7.2)], we propose to eradicate the stringent conditions on the existence of the SC order by simply intersecting the constraints on the SC order with the scope-inclusion relation. This essentially means that the orderings imposed between events by the SC axioms only take effect if those events have inclusive scopes. Under this proposal, which recalls the way C11’s synchronisation relation (\( sw, df \)) is intersected with scope-inclusion when producing OpenCL’s version (\( rsw, df \)), we do not need to restrict the programmer’s usage of SC atomics to certain scopes; instead, the guarantees provided by those SC atomics degrade gracefully as their scopes narrow.

Definition 25 (Proposed SC axiom for OpenCL). The following axiom for SC atomics in OpenCL is obtained from O-S\textsubscript{unp} by removing the sc-all and sc-dv conditions and instead intersecting with \textit{incl}:

\[ acy(SC^2 \cap (Fab^\circ) ; (g,hb \sqcup lb \sqcup rb \sqcup mo) ; sfF^\circ) \cap \textit{incl}(O-S_{\text{scoped}}) \]

We include in §B.2 a suggestion for changing the wording of the OpenCL specification to accommodate our proposal.

5. Overhauling the SC axioms in OpenCL

We describe how the handling of SC atomics in OpenCL can be changed to address the shortcomings identified in §4.5. By changing the handling of SC atomics to allow inter-device use, we can avoid the need for expensive inter-device synchronisation.

To summarise: the intent of the Khronos working group was to enable efficient implementation of DV-scoped SC atomics by compilers, at the expense of programmer inconvenience. Instead, our formalisation shows that we have the worst of both worlds: the programmer is inconvenienced, and yet a correct compiler is obliged to enforce inter-device orderings on DV-scoped SC atomics.

5.1 Implementability of the new SC axiom

The new O-S\textsubscript{scoped} axiom is stronger than the original O-S\textsubscript{unp} axiom, so we must confirm that our proposal does not place undue demands on compilers that implement the memory model.

The only published compilation scheme of the OpenCL 2.0 memory model of which we are aware is that published by AMD [23] and later formalised by Wickerson et al. [33]. The scheme compiles the release/acquire fragment of OpenCL atomics, and its soundness has been verified against an operational model of an AMD GPU [33]. In this subsection we describe how the scheme can be extended to support SC atomics, and we demonstrate via a series of examples that the extended scheme meets the requirements of our revised SC axiom. The original compilation scheme does not cater for multiple devices, and does not include fences, and we do not attempt here to extend the scheme to cover these features. As such, this scheme does not engage directly with the problems of inter-device SC atomics that we noted in the previous section; however, it does illustrate how WG- and DV-scoped SC atomics can co-exist.

The AMD compilation scheme. The operational model is quite simple. Each work-group has its own L1 cache, and each device has
its own L2 cache. Since the compilation scheme considers only the single-device case, the L2 cache can be safely thought of as the main memory. No instruction reordering is permitted. At any time, the environment can flush a dirty L1 cache entry to the L2 (and thereby make it clean), can fetch an L2 entry to replace a clean L1 entry, and can evict a clean L1 entry.

The semantics of the various assembly instructions can be summarised as follows. LD r x loads into register r from the nearest cache that contains a valid entry for x; ST r x stores from r into the local L1 cache, first flushing x’s entry therein if it is invalid; INC_L1 r x increments x in the local L1 cache; INC_L2 r x increments x in the L2 cache, first flushing any dirty entry for x in the local L1 cache; FLU_L1 flushes all dirty entries in the local L1 cache; and INV_L1 marks all entries in the local L1 cache as invalid.

The extensions to the compilation scheme are given in Tab. 1. Here, fetch_inc stands for ‘atomic fetch and increment’, and provides a representative of RMW operations in OpenCL.

Correctness of the compilation scheme. Most of the flush and invalidate instructions in the compilation scheme are necessary to ensure correct release/acquire semantics. For SC atomics, we need add only two further instructions: the INV_L1 before the load in row θ, and the FLU_L1 after the store in row θ. The need for these instructions can be motivated by considering the following two examples, which correspond to the classic store-buffering and IRIW litmus tests.

The memory model requires the program in Example 9 not to produce the final state r0 = r1 = 0. With only release/acquire semantics, the compilation scheme inserts no flush or invalidate instructions between the store and the load in each thread, and the relaxed behaviour can be observed: both threads might pre-fetch x = y = 0 into their respective L1 caches (the threads are in different work-groups, so they have different L1 caches), then perform their stores, and finally load the L1-cached values of x and y. However, placing a FLU_L1 after the store and an INV_L1 before the load ensures that no sequence of fetching and flushing can lead to the relaxed behaviour. We do not need FLU_L1 or INV_L1 instructions before or after the SC increment instruction, because INC_L2 writes directly to the L2, invalidating the L1 as it does so.

The memory model also requires the IRIW litmus test
global atomic_int *x; global atomic_int *y;
store(x,1); store(y,1); r0=load(x); r2=load(y);
not to produce the final state \(\{x_0 = r_2 = 1, r_1 = r_3 = 0\}\). (Recall that these store and load operations use memory order SC and scope DV by default.) Here, an INV_L1 instruction between each pair of loads is sufficient to rule out such executions.\(^2\)

### 6. Simulating the memory models with HERD

Our overhaul of SC atomics avoids the requirement for the S relation to be explicitly constructed in execution witnesses. Our hypothesis was that this would lead to improved efficiency in the process of exhaustively enumerating the allowed behaviour of litmus tests that use SC atomics. We now explain how we extended the HERD memory model simulator in order to enable investigation of C11 and OpenCL litmus tests (§6.1), and present experimental results using HERD to compare the efficiency of simulation before and after our overhaul, and also in comparison to the CDSChecker memory model simulator [22] (§6.2). For a family of litmus tests derived from Dekker’s algorithm, our results show that our revised axioms lead to an exponential speedup in simulation time using HERD, bringing performance using HERD, which is general-purpose and exhaustive, much closer to that of CDSChecker, which is tuned for C++11 and is not guaranteed to be exhaustive.

#### 6.1 Extensions to HERD

The version of HERD described by Alglave et al. [2, 3] supports only assembly code: sequences of labelled instructions and gotos. In order to simulate our formalisations of the C11 and OpenCL memory models, we have extended the .cat format to support the definition of faulty axioms, and the HERD tool with both a routine for alerting the user when a faulty execution is detected and a module for translating C11 and OpenCL programs into their executions.

We model only a small fragment of the C11 language: enough to encode the litmus tests we found useful for testing our formalisation. We exclude, for example, the address-of operator, compound types, and function calls. We include if and while blocks, pointer dereferencing, simple expressions, and built-in atomic functions such as atomic_thread_fence (C11) and atomic_work_item_fence (OpenCL).

#### 6.2 Simulating the C11 model: performance evaluation

We now compare the performance of HERD in enumerating the behaviours of litmus tests (a) when equipped with the original SC axioms in C11 vs. (b) when equipped with our revised SC axioms. We also provide performance results gathered using CDSChecker, a custom-built simulator for the C++11 memory model [22]. The HERD tool guarantees exhaustive enumeration of allowed behaviours for a litmus test; CDSChecker aims for high coverage of behaviours, but is known to be non-exhaustive in general [22].

Recall that Dekker’s mutual exclusion algorithm [10] is a key use case for SC atomics. The essential idiom underlying an N-threaded version of Dekker’s algorithm is captured by the following N-threaded store-buffering litmus tests:

\[
P_N \overset{\text{def}}{=} \left\{ \text{store}(x_1,1); \text{store}(x_1,1); \cdots \text{store}(x_1,1) \right\}
\]

that operate on a collection \(\{x_1, \ldots, x_N\}\) of atomic locations initialised to zero. Recall that atomic store and load operations use memory order SC by default. Dekker’s algorithm requires that it is not possible to observe the final state where \(r_1 = \cdots = r_N = 0\); only SC is strong enough to rule out this relaxed behaviour.

We use the family \(P_N\)\(^{x\in\mathbb{N}}\) to assess the scalability of the two versions of HERD and of CDSChecker. Figure 1 shows the time each tool takes to simulate \(P_N\) as \(N\) increases. Experiments were conducted on a 3.1 GHz MacBook Pro, and each data point represents the mean of ten runs. We do not include error bars because the standard deviation is negligible. The original memory model,
naively implemented in HERD, times out on just 4 threads. This is because it iterates over all $(2N)!$ orders of the $2N$ SC events that are in every execution of $P_N$. When HERD is provided with our revised memory model, simulation times greatly improve. Bearing in mind the logarithmic $\gamma$-axis, the performance of both HERD on the revised memory model and CDSChecker appears to scale exponentially with $N$, which meets expectations since $P_N$ has $2^N – 1$ unique final states. Still, CDSChecker significantly outperforms HERD when simulating $P_N$, and on several other programs that we tried. This is because CDSChecker, unlike HERD, is optimised specifically for the C11 memory model, through the use of such techniques as the early elimination of infeasible executions, and a variant of dynamic partial order reduction (DPOR) [11] on the $S$ order. In fact, we conjecture that the use of DPOR here has an effect similar to our proposal to rephrase the memory model with $S$ as a partial order.

Figure 1 demonstrates that simply by tweaking the axioms that define the memory model, simulation time can be dramatically decreased, without the need to implement complex optimisations, such as DPOR, that make it difficult to assess the soundness and completeness of the tool.

7. Related work
The C11 memory model has been formalised several times. Batty et al. [5] present a comprehensive formalisation using Lem [21]. Vafeiadis et al. [30, 31] and Batty et al. [7] have also formalised slightly simplified variations. Alglave et al. have formalised a release/acquire fragment of the C11 model (without release sequences, fences, non-atomics, or data races) in the .cat language, and have shown it to be an instance of their generic axiomatic memory model [2]. We use the .cat language in our work too, but our comprehensive model, which incorporates undefined behaviours and a richer language of events, no longer fits within their generic framework.

Criticisms of the C11 model. Batty et al. describe a fundamental problem in the structure of the C11, C++11, C++14 and OpenCL memory models: the so-called “thin-air” executions [8]. This is a difficult open problem requiring a radically different approach: we do not address it here. Vafeiadis et al. note that the current rules governing SC atomics break desirable properties of the memory model, harming the prospect of reasoning above it, and they propose a strengthening of the model to fix this [31]. Our proposal builds on theirs (§3.1), but goes further (§3.2), arriving at a substantially simpler model. A similar proposal was in fact considered by Vafeiadis et al. in the context of the original total-order SC axioms [31], but abandoned over concerns that it would invalidate the existing Power compilation scheme. In our work, we have demonstrated that such a proposal is in fact valid on Power (and x86).

The OpenCL 2.0 memory model has recently been described by Gaster et al. [12], as an instance of a heterogeneous race-free (HRF) model [13]. Our work improves on theirs in several ways. A key shortcoming of their work is its relative informality: it lacks the mathematical precision that is required to resolve all the details of the OpenCL memory model. Our formalisation, in contrast, is precise enough to be executed by a machine (cf. §6). Moreover, their characterisation of the OpenCL memory model has several technical issues. It replaces the specification’s modification order (which orders atomic write events) with a coherence order (which orders both read and write events) without proving that the intent of the specification is preserved by this change. Another infidelity to the specification is the omission of release sequences, which prohibits the correct treatment of release-fences. Indeed, Gaster et al. include no formal treatment of fences at all, describing their behaviour only in prose. Our .cat presentation of the OpenCL memory model (§E) treats release sequences and fences in full. Its informality aside, Gaster et al.’s work contains numerous insights into the design and workings of the OpenCL memory model, and provides a valuable basis for our formalisation efforts.

Wickerson et al. have already built upon our formalisation of the OpenCL memory model, as part of their investigations into the semantics of a proposed extension to OpenCL [33].

Implementations of the OpenCL memory model We are aware of only one formalised implementation of the OpenCL memory model: for AMD GPUs [33]. Alglave et al. present an experimentally-validated axiomatic model of an Nvidia GPU [3], which could provide another compilation target for our OpenCL memory model. However, we find that their model is too weak to admit an efficient mapping from OpenCL. Specifically, it does not provide the property of cumulativity: synchronisation at one scope can not be chained with further synchronisation at a wider scope to induce overall synchronisation between the two end-points. Since cumulativity is a property required by the OpenCL memory model, we deduce that the OpenCL compiler must, very expensively, treat all operations as having the widest scope.

Memory model simulators other than HERD that are capable of handling the C11 model include Cppmem [5], Nitpick [9] and CDSChecker [22]. Owing to difficulties obtaining the tools, we were unable to include Cppmem and Nitpick in our tool comparison (§6.2), but informal experiments with Cppmem’s web interface suggest that its performance is comparable to HERD on the original C11 model. For instance, Cppmem takes 23 seconds to calculate the executions of the classic IRIW litmus test when all accesses are SC, while HERD takes 5 seconds on the original memory model, and is almost instantaneous on our revised model (§3).

CDSChecker continues to outperform HERD even on the revised C11 model, because it is highly optimised for the C11 memory model. HERD on the other hand is deliberately designed not to be optimised for a particular model, but to be instead a generic memory model simulator. A key advantage of using a generic memory model simulator like HERD is that it is easy to tinker with the model during the development process: one must only modify a text file and restart HERD in order to explore the impact of a proposed change. Indeed, this ease of modification, together with the challenge of expressing the C11 model in the very concise .cat language, inspired our discovery of the simpler SC axioms described in this paper. Moreover, where CDSChecker is designed for efficiency, sometimes at the cost of fidelity to the memory model (the lack of self-satisfying conditionals, for instance, is a source of incompleteness in CDSChecker), our formalisation and simulator are designed primarily to represent the memory model as closely as possible.

8. Conclusion
Our overhaul of the semantics of SC atomics and fences provides four main benefits in relation to the C11 and OpenCL memory models: more efficient exploration of the behaviours of litmus tests (cf. §3.1); refined specification text that we argue is easier for programmers and compiler-writers to understand (cf. §3.2); improved usability of the languages by programmers (cf. §5); and opportunities for compiler-writers to produce more efficient implementations (cf. §5). Our proposed changes to the memory models have been validated against all of the formalised C11 and OpenCL compilation schemes of which we are aware.

A topic for future research is the consideration of memory consistency between OpenCL devices and the host application that launches kernels on these devices; our treatment in this paper focuses solely on interactions between kernel threads.
References


A. Rules for SC atomics in C11

A.1 Original
The following text is reproduced verbatim from the C11 standard [15 (§7.17.3, paragraphs 6 and 9–11)].

6. There shall be a single total order $S$ on all `memory_order_seq_cst` operations, consistent with the “happens before” order and modification orders for all affected locations, such that each `memory_order_seq_cst` operation $B$ that loads a value from an atomic object $M$ observes one of the following values:
   - the result of the last modification $A$ of $M$ that precedes $B$ in $S$, if it exists, or
   - if $A$ exists, the result of some modification of $M$ in the visible sequence of side effects with respect to $B$ that is not `memory_order_seq_cst` and that does not happen before $A$, or
   - if $A$ does not exist, the result of some modification of $M$ in the visible sequence of side effects with respect to $B$ that is not `memory_order_seq_cst`.

9. For an atomic operation $B$ that reads the value of an atomic object $M$, if there is a `memory_order_seq_cst` fence $X$ sequenced before $B$, then $B$ observes either the last `memory_order_seq_cst` modification of $M$ preceding $X$ in the total order $S$ or a later modification of $M$ in its modification order.

10. For atomic operations $A$ and $B$ on an atomic object $M$, where $A$ modifies $M$ and $B$ takes its value, if there is a `memory_order_seq_cst` fence $X$ such that $A$ is sequenced before $X$ and $B$ follows $X$ in $S$, then $B$ observes either the effects of $A$ or a later modification of $M$ in its modification order.

11. For atomic operations $A$ and $B$ on an atomic object $M$, where $A$ modifies $M$ and $B$ takes its value, if there are `memory_order_seq_cst` fences $X$ and $Y$ such that $A$ is sequenced before $X$, $Y$ is sequenced before $B$, and $X$ precedes $Y$ in $S$, then $B$ observes either the effects of $A$ or a later modification of $M$ in its modification order.

For reference, we include with this passage its Flesch–Kincaid (FK) reading ease score. A higher score indicates easier readability. Scores usually range between 0 and 100.

A.2 Our proposal
Section 3.2 presented our proposal for simplifying the sequential consistency axioms in the C11 model. We give here our suggestion for how the specification document can be rephrased to accommodate our proposal, while maintaining the prose style used throughout the rest of the document.

Specifically, the paragraphs quoted above can be removed and replaced with the following three:

1. A value computation $A$ of an object $M$ reads before a side effect $B$ on $M$ if $A$ and $B$ are different operations and $B$ follows, in the modification order of $M$, the side effect that $A$ observes.

2. If $X$ reads before $Y$, or happens before $Y$, or precedes $Y$ in modification order, then $X$ (as well as any fences sequenced before $X$) is SC-before $Y$ (as well as any fences sequenced after $Y$).

3. If $A$ is SC-before $B$, and $A$ and $B$ are both `memory_order_seq_cst`, then $A$ is restricted-SC-before $B$.

4. There must be no cycles in restricted-SC-before.

[93 words; FK reading ease 73.1]
B. Rules for SC atoms in OpenCL

B.1 Original

The following text is reproduced verbatim from the OpenCL 2.1 standard. We exclude the citations to the C11 standard.

If one of the following two conditions holds:

- All `memory_order_seq_cst` operations have the scope `memory_scope_all_svm_devices` and all affected memory locations are contained in system allocations or fine grain SVM buffers with atomics support.
- All `memory_order_seq_cst` operations have the scope `memory_scope_device` and all affected memory locations are not located in system allocated regions or fine-grain SVM buffers with atomics support.

then there shall exist a single total order $S$ for all `memory_order_seq_cst` operations that is consistent with the modification orders for all affected locations, as well as the appropriate global-happens-before and local-happens-before orders for those locations, such that each `memory_order_seq_cst` operation $B$ that loads a value from an atomic object $M$ in global or local memory observes one of the following values:

- the result of the last modification $A$ of $M$ that precedes $B$ in $S$, if it exists, or
- if $A$ exists, the result of some modification of $M$ in the visible sequence of side effects with respect to $B$ that is not `memory_order_seq_cst` and that does not happen before $A$, or
- if $A$ does not exist, the result of some modification of $M$ in the visible sequence of side effects with respect to $B$ that is not `memory_order_seq_cst`.

[...]

If the total order $S$ exists, the following rules hold:

- For an atomic operation $B$ that reads the value of an atomic object $M$, if there is a `memory_order_seq_cst` fence $X$ sequenced-before $B$, then $B$ observes either the last `memory_order_seq_cst` modification of $M$ preceding $X$ in the total order $S$ or a later modification of $M$ in its modification order.
- For atomic operations $A$ and $B$ on an atomic object $M$, where $A$ modifies $M$ and $B$ takes its value, if there is a `memory_order_seq_cst` fence $X$ such that $A$ is sequenced-before $X$ and $B$ follows $X$ in $S$, then $B$ observes either the effects of $A$ or a later modification of $M$ in its modification order.
- For atomic operations $A$ and $B$ on an atomic object $M$, where $A$ modifies $M$ and $B$ takes its value, if there are `memory_order_seq_cst` fences $X$ and $Y$ such that $A$ is sequenced-before $X$, $Y$ is sequenced-before $B$, and $X$ precedes $Y$ in $S$, then $B$ observes either the effects of $A$ or a later modification of $M$ in its modification order.
- For atomic operations $A$ and $B$ on an atomic object $M$, if there are `memory_order_seq_cst` fences $X$ and $Y$ such that $A$ is sequenced-before $X$, $Y$ is sequenced-before $B$, and $X$ precedes $Y$ in $S$, then $B$ occurs later than $A$ in the modification order of $M$.

B.2 Our proposal

Section 5 presented our proposal for simplifying the sequential consistency axioms in the OpenCL model. We give here our suggestion for how the specification document can be rephrased to accommodate our proposal, while maintaining the prose style used throughout the rest of the document.

Specifically, the paragraphs quoted above can be removed and replaced with the following three.

1. A value computation $A$ of an object $M$ reads before a side effect $B$ on $M$ if $A$ and $B$ are different operations and $B$ follows, in the modification order of $M$, the side effect that $A$ observes.
2. If $X$ reads before $Y$, or global happens before $Y$, or local happens before $Y$, or precedes $Y$ in modification order, then $X$ (as well as any fences sequenced before $X$) is SC-before $Y$ (as well as any fences sequenced after $Y$).
3. If $A$ is SC-before $B$, and $A$ and $B$ have inclusive scopes, then $A$ is restricted-SC-before $B$.
4. There must be no cycles in restricted-SC-before.

The only departures from our proposal for the C11 memory model (Sec. A.2) are the requirement of inclusive scopes, and the splitting of happens-before into its global and local versions.
C. Proofs

The following theorem states that the x86 and Power compilation schemes for C11, as given in Tab. 2, remain sound in the presence of our revised SC axiom, $S_{\text{simp}}$.

<table>
<thead>
<tr>
<th>C11 operation</th>
<th>x86</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = \text{load}(x, SC)$</td>
<td>lock xadd(0) sync; id; cmp; bc; isync</td>
<td></td>
</tr>
<tr>
<td>$\text{store}(x, r, SC)$</td>
<td>lock xchg sync; st</td>
<td></td>
</tr>
<tr>
<td>$r = \text{fence}(x, SC)$</td>
<td>mfence sync</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Compiling the C11 SC atoms

Theorem 3. Let $P$ be a C11 program that has no faulty executions. If we compile $P$ to x86 according to the mapping given by Batty et al. [5], then every valid x86 execution corresponds to a C11 execution where $S_{\text{simp}}$ holds. If we compile $P$ to Power according to the mapping given by Batty et al. [6], then every valid Power trace is observationally equivalent to a C11 execution where $S_{\text{simp}}$ holds.

Proof (x86 case). The axiomatic model of Owens et al. restricts the behaviour of memory using a partial order over x86 memory events called $\text{memory-order}$. The proof of Batty et al. [5] constructs the relations of the C11 execution using $\text{memory-order}$ – modification order ($mo$) and reads-from ($rf$), in particular, are projected from it. Here we rely on several properties of $\text{memory-order}$ as set out by Owens et al.; when restricted to writes, it is a linear order; program order between two events is included in memory order if there is an intervening fence or if either instruction is locked; program-order edges from reads to later events are included; reads read from the most recent preceding write in $\text{memory-order}$.

We proceed by contradiction, showing that given the construction of $rf$ and $mo$ used in the proof of Batty et al., any cycle in $SC^2 \cap (F_{sb} \cap \{hb \cup rb \cup mo\} \cup F_{sfb}) \setminus \text{id}$ implies either the existence of a cycle in $\text{memory-order}$, or an inconsistent $rf$ edge.

Any cycle in the relation is made up of $mo$, $hb$, and $rb$ edges, possibly linked with $sb$ edges. The $mo$, $hb$ and $sb$ edges all imply corresponding $\text{memory-order}$ edges. To see this note: $\text{memory-order}$ is a linear order over writes; any $hb$ edge in the $S_{\text{simp}}$ relation begins with either a fence or a locked instruction, so $sb$ edges correspond to $\text{memory-order}$ edges, then there may be a chain of edges in $mo$; $rf$; $sb$, where the final $sb$ edge is headed by a read – transitivity implies that $hb$ corresponds to $\text{memory-order}$; finally, any $sb$ edges are either between locked instructions or have a fence between accesses, and correspond to $\text{memory-order}$ edges.

Finally, if the cycle contains an $rb$ edge, then $\text{memory-order}$ cannot contradict this; the read would become inconsistent in the x86 execution. Then for a given $S_{\text{simp}}$ cycle, we have a sequence of $\text{memory-order}$ edges that would form a cycle if not for holes corresponding to $rb$ edges.

We now show that there is indeed a cycle in $\text{memory-order}$. Consider an $rb$ edge: its head is a read, so the preceding edge must either be a $sb$ edge or a $hb$ edge. If it is a $sb$ edge then either the head of that edge is a write, or the edge that precedes that is an $hb$ edge. In all cases, the read at the head of the $rb$ edge is preceded in $hb$ by a write. As $\text{memory-order}$ is total over writes, it must order this preceding write’s x86 counterpart before the write in the tail of the $rb$ edge. We use this fact to construct a cycle in $\text{memory-order}$, a contradiction.

Proof (Power case). In their proof, Batty et al. construct a C11 execution from a Power trace such that Power coherence and $rf$ edges match their constructed C11 counterparts. In proving the SC axioms hold over the execution, they prove a property, $\text{good-sc}$, that establishes a total order over the SC atoms of the execution that contains $po$, $co$, $fr$ and an extended variant of reads from, $erf$, each restricted to the SC atoms. The SC fences are added to this relation in a way that is consistent with $co$ and $rf$ in the rest of the execution, preserving the invariant that it is a strict partial order. The following edges become part of the SC order: $[SC]$; $po^1$; $(rf^{-1})^1$; $co$; $rf^1$; $po^2$; $[F \cap SC]$ and $[F \cap SC]$; $po^3$; $(rf^{-1})^3$; $co$; $rf^3$; $po^3$; $[SC]$. The proof goes on to show that $hb$ restricted to the SC actions is a subset of the total order. The construction of $mo$ and $rf$ in the C11 execution follow the Power trace directly, so $\text{good-sc}$ together with the addition of the fence edges, which contain $F_{sb}^3$; $fr$; $F_{sfb}^3$ and $F_{sb}^3$; $co$; $F_{sfb}^3$, show the acyclicity of $SC^2 \cap (F_{sb}^3 \cap \{hb \cup rb \cup mo\} \cup F_{sfb}^3) \setminus \text{id}$ directly.

\[ \square \]
D. Raw .cat files for the C11 axioms

We give here the .cat files that encode the C11 axioms presented in this paper. The models presented here depart slightly from the one presented in §2, partly for syntactic restrictions in HErD, and partly because it includes locks, ‘consume’ atomics, and ‘unsequenced races’, which we do not discuss in the paper.

D.1 c11_base.cat

This file contains definitions and axioms that are common to the various versions of the C11 memory model that we present. Throughout the file, definitions carry comments that name the corresponding definitions of Batty et al.’s Lem model.

C++11 withinit

let sb = po
let mo = co
let rb = (rf^-1 ; mo) \ id
let cacq = acq | (sc & (R | F)) | acq_rel | (F & con)
let crel = rel | (sc & (W | F)) | acq_rel
let ccon = R & con
let dd = (data | addr)+
let Fsb = [F] ; sb
let sbF = sb ; [F]

(* release_acquire_fenced_synchronizes_with, hypothetical_release_sequence_set, release_sequence_set *)
let rs = [crel] ; Fsb? ; [A & W] ;
((mo ; [rmw]) | coi) & ~(coe ; ![rmw] ; mo))?
let swra = ext (rs ; rf ; [A] ; sbF? ; [cacq])
let swul = ext ([ul] ; lo ; [lk])
let pp_asw = asw \ (asw ; sb)
let sw = pp_asw | swul | swra

(* with_consume_cad_set, dependency_ordered_before *)
let cad = ((rf & sb) \ dd)+
let dob = (ext ([W] ; rs ; rf ; [ccon])); cad?

(* happens_before, inter_thread_happens_before, consistent hb *)
let ithbr = sv | dob | (sw ; sb)
let ithb = (ithbr | (sb ; ithbr))+
let hb = sb | ithb
acyclic hb as Hb

(* coherent_memory_use *)
let hbl = hb & loc
irreflexive ((rf^-1)? ; mo ; rf? ; hb) as Coh

irreflexive (rf | (mo ; mo ; rf^-1)
| (mo ; rf)) as Rmw

(* visible_side_effect_set *)
let vis = ([W] ; hbl ; [R])
& ~(hbl; [W]; hbl)

D.2 c11_orig.cat

This file describes the original C11 memory model.

C++11 withinit withsc

include c11_base.cat

let Simm = S & ~(mo ; S)

irreflexive (S ; hb)
irreflexive (S ; Fsb? ; mo ; sbF?)
irreflexive (S ; rf^-1; [sc] ; mo)
irreflexive (Simm ; rf^-1 ; hbl ; [W])
irreflexive (S ; Fsb ; rb)
irreflexive (S ; rb ; sbF)
irreflexive (S ; Fsb ; rb ; sbF)

undefined_unless empty dr as Dr
undefined_unless empty ur as unsequencedRace
undefined_unless empty bl as badLock
undefined_unless empty lu as badUnlock

D.3 c11_partialSC.cat

This file describes the C11 memory model after it has been revised to use our Spartial axiom (cf. §3.1).

C++11 withinit

include c11_base.cat

let r1 = hb
let r2 = Fsb? ; mo ; sbF?
let $r_3 = r_f^{-1}$; $[sc]$; $mo$
let $r_4 = r_f^{-1}$; $hbl$; $[W]$
let $r_5 = F_{sb}$; $rb$
let $r_6 = rb$; $sbF$
let $r_7 = F_{sb}$; $rb$; $sbF$

let $scp = r_1|r_2|r_3|r_4|r_5|r_6|r_7$

acyclic ("[\_] & ([sc] ; scp ; [sc])") as $S_{partial}$
show scp

undefined_unless empty dr as Dr
undefined_unless empty ur as unsequencedRace
undefined_unless empty bl as badLock
undefined_unless empty lu as badUnlock

D.4 c11_simp.cat
This file describes the C11 memory model after it has been further revised to use our $S_{simp}$ axiom (cf. §3.2).

C++11 within it

include c11_base.cat

let $scp = [sc]$; $F_{sb}$?; $(mo | rb | hb)$; $sbF$?; $[sc]$

acyclic scp as $S_{simp}$
show scp

undefined_unless empty dr as Dr
undefined_unless empty ur as unsequencedRace
undefined_unless empty bl as badLock
undefined_unless empty lu as badUnlock
E. Raw .cat files for the OpenCL axioms

We give here the .cat files that encode the OpenCL axioms presented in this paper. The models presented here depart slightly from the ones presented in §4, partly for syntactic restrictions in HERD, and partly because they include barrier divergence, which we do not discuss in the paper.

E.1 opencl_base.cat

This file contains definitions and axioms that are common to the various versions of the OpenCL memory model that we present.

OpenCL
withoutsc

let mo = co & (nonatomicloc^2)
let sb = po
let rb = (rf^-1; mo) \ id

(* Access modes *)
let mo_acq = memory_order_acquire
let mo_rel = memory_order_release
let mo_acq_rel = memory_order_acq_rel
let mo_rlx = memory_order_relaxed
let mo_sc = memory_order_seq_cst

(* Scope annotations *)
let s_wi = memory_scope_work_item
let s_wg = memory_scope_work_group
let s_dev = memory_scope_device
let s_all = memory_scope_all_svm_devices

(*******************)
(* Synchronisation *)
(*******************)

let acq = (mo_acq | mo_acq_rel | mo_sc) & (R | F)
let rel = (mo_rel | mo_acq_rel | mo_sc) & (W | F)

(* Fences sequenced before or after *)
let Fsb = [F]; sb
let sbF = sb; [F]

(* Release sequence *)
let rs' = wi | (unv; [R & W])
let rs = mo & rs' & ~(mo & `rs'); mo

(* Inclusive scopes *)
let incl_wg = wg & s_wg^-2 | dev & s_dev^-2 | s_all^-2

(* Inclusive scopes, less conservative (unused) version *)
let incl1 = ([s_wg];wg) | ([s_dev];dev)
| ([s_all];unv)
let incl1' = incl1 & incl1^-1

(* Release-acquire synchronisation *)
let ra_sw(r) =
  ([r & rel]; Fsb?; [W \ s_wi]; rs?; [r]; rf;
   [R \ s_wi]; sbF?; [acq & r]) & incl & ~wi

(* Barrier synchronisation *)
let bar_sw(r) = (entry_fence * exit_fence) &
  same_B & `wi & wg & r^2

(* Allowed to synchronise on the other region *)

let scf = mo_sc^-2 | (G & L & F)^2

(* Global and local synchronises-with *)
let gsw = ra_sw(G) | bar_sw(G) | (scf & ra_sw(L))
let lsw = ra_sw(L) | bar_sw(L) | (scf & ra_sw(G))

(************************)
(* Consistency of reads *)
(************************)

(* A load can only read from a store that already happened. *)
irreflexive rf; (ghb | lhb) as O-Rf

(* Visible side effects *)
let vis(hb) = (W * R) & hb & loc &
  ~(hb & loc); [W]; hb

(* An non-atomic load can only read from a store that is visible. *)
empty (rf;[G & nonatomicloc]|vis(hb)) as O-NaRfG
empty (rf;[L & nonatomicloc]|vis(hb)) as O-NaRfL

(* Consistency of RMWs *)
irreflexive rf | (mo;mo;rf^-1) | (mo;rf) as O-Rmw

E.2 opencl.cat

This file describes the OpenCL memory model after it has inherited our revisions to the C11 model.

OpenCL
withoutsc

include opencl_base.cat

(****************************)
(* Sequential consistency, simplified *)
(*******************************)

let scp = Fsb?; (rb | mo | (ghb | lhb)) \ sbF?

(* The following evaluates to unv if every SC event has scope s_all and accesses fine-grained buffers with atomics support,
and to empty otherwise *)
let S_cond1 = ~(unv; \[mo_sc \ (s_all & fga)\]; unv)

(* The following evaluates to unv if every SC event has scope s_dev and does not access fine-grained buffers with atomics support, and to empty otherwise *)
let S_cond2 = ~(unv; \[mo_sc \ (s_dev & !fga)\]; unv)

acyclic (mo_sc-2) & scp & (S_cond1|S_cond2) as O-Ssimp

(*********)
(* Races *)
(*********)

let cnf = ((W * W) | (W * R) | (R * W)) & loc

(* Data races *)
let dr = cnf & ~(ghb | lhb) & ~(ghb | lhb)-1 & ~incl
undefined_unless empty dr as O-Dr
show dr

(* Barrier divergence *)
let bsw = bar_sw(G) | bar_sw(L)
let bd = \[entry_fence\] & ((~wi & wg);unv)

undefined_unless empty bd as O-Bd
show bd

E.3 opencl_scopedsc.cat
This file describes the OpenCL memory model after it has been revised to use our O-S Scoped axiom (cf. §5).

OpenCL withoutsc

include opencl_base.cat

(***********************)
(* Sequential consistency, simplified, *)
(* with scoped SC axioms *)
(***********************)

let scp = Fsb?; (rb | mo | (ghb | lhb)); sbF?

acyclic (mo_sc-2) & scp & incl as O-Sscoped

(*********)
(* Races *)
(*********)

let cnf = ((W * W) | (W * R) | (R * W)) & loc

(* Data races *)
let dr = cnf & ~(ghb | lhb) & ~(ghb | lhb)-1 & ~incl
undefined_unless empty dr as O-Dr
show dr

(* Barrier divergence *)
let bsw = bar_sw(G) | bar_sw(L)