

This is the *PlusCal* specification of the deconstructed bakery algorithm in the paper

Deconstructing the Bakery to Build a Distributed State Machine

In this version of the specification, the choice of a ticket number is performed non-atomically, using an explicit loop over processes. There is one simplification that has been made in the *PlusCal* version: the registers $localCh[i][j]$ have been made atomic, a read or write being a single atomic action. This doesn't affect the derivation of the distributed bakery algorithm from the deconstructed algorithm, which also makes the simplifying assumption those registers are atomic because they disappear from the final algorithm.

Here are some of the changes made to the paper's notation to conform to *PlusCal*/TLA+. Tuples are enclosed in $\langle \rangle$, so we write $\langle i, j \rangle$ instead of (i, j) . There's no upside down "?" symbol in TLA+, so that's replaced by the identifier *qm*.

The pseudo-code for main process *i* has two places in which subprocesses $\langle i, j \rangle$ are forked and process *i* resumes execution when they complete. *PlusCal* doesn't have subprocesses. This is represented in *PlusCal* by having a single process $\langle i, j \rangle$ executing concurrently with process *i*, synchronizing appropriately using the variable *pc*.

Here is the basic idea:

This pseudo-code for process *i*:

```
main code ;
process j # i \in S
  s1: subprocess code
end process
p2: more main code
```

is expressed in *PlusCal* as follows:

```
In process i
  main code ;
  p2: await \A j # i : pc[<<i,j>>] = "s2"
  more main code

In process <i, j>
  s1: await pc[i] = "p2"
  subprocess code ;
  s2: ...
```

Also, processes have identifiers and, for reasons that are not important here, we can't use *i* as the identifier for process *i*, so we use $\langle i \rangle$. So, $pc[i]$ in the example above should be $pc[\langle i \rangle]$. In the pseudo-code, process *i* also launches asynchronous processes $\langle i, j \rangle$ to set $localNum[j][i]$ to 0. In the code, these are another set of processes with ids $\langle i, j, "wr" \rangle$.

We could simplify this algorithm by not waiting for $localNum[j][i]$ to equal 0 in subprocess $\langle i, j \rangle$ and having the asynchronous write of 0 not do anything if process *i* has begun the write to $localCh[i][j]$ that sets its value to $number[i]$. However, I think I like the algorithm in the paper the way it is because it makes the pseudo-code more self-contained.

EXTENDS *Data, Integers*

```
--algorithm Decon{
  variables number = [p ∈ Procs ↦ 0],
           localNum = [p ∈ Procs ↦ [q ∈ OtherProcs(p) ↦ 0]],
```

```

    localCh    = [p ∈ Procs ↦ [q ∈ OtherProcs(p) ↦ 0]];

fair process ( main ∈ ProcIds )
  variable unRead = {}, v = 0;
  {
    ncs:- while ( TRUE ) {
      skip; noncritical section
      M: await ∀ p ∈ SubProcsOf(self[1]) : pc[p] = "test" ;
        unRead := OtherProcs(self[1]);
      M0: while ( unRead ≠ {} ) {
        with ( j ∈ unRead ) {
          if ( localNum[self[1]][j] ≠ qm ) {
            v := Max(v, localNum[self[1]][j]) } ;
            unRead := unRead \ {j}
          }
        } ;
        with ( n ∈ {m ∈ Nat : m > v} ) {
          number[self[1]] := n ;
          localNum := [j ∈ Procs ↦
            [i ∈ OtherProcs(j) ↦
              IF i = self[1] THEN qm
              ELSE localNum[j][i]];
          } ;
          v := 0 ;
          L: await ∀ p ∈ SubProcsOf(self[1]) : pc[p] = "ch" ;
          cs: skip; critical section
          P: number[self[1]] := 0 ;
          localNum := [j ∈ Procs ↦
            [i ∈ OtherProcs(j) ↦
              IF i = self[1] THEN qm
              ELSE localNum[j][i]];
          }
        }
      }
    }

fair process ( sub ∈ SubProcs ) {
  ch: while ( TRUE ) {
    await pc[⟨self[1]⟩] = "M" ;
    localCh[self[2]][self[1]] := 1 ;
  test: await pc[⟨self[1]⟩] = "L" ;
    localNum[self[2]][self[1]] := number[self[1]] ;
  Lb: localCh[self[2]][self[1]] := 0 ;
  L2: await localCh[self[1]][self[2]] = 0 ;
  L3:- See below for an explanation of why there is no fairness here.
    await (localNum[self[1]][self[2]] ∉ {0, qm}) ⇒
      (⟨number[self[1]], self[1]⟩ ≪

```

$\langle localNum[self[1]][self[2]], self[2] \rangle$

The await condition is written in the form $A \Rightarrow B$ rather than $A \vee B$ because when *TLC* is finding new states, when evaluating $A \vee B$ it evaluates B even when A is true, and in this case that would produce an error if $localNum[self[1]][self[2]]$ equals qm .

}
}

We allow process $\langle i, j, "wr" \rangle$ to set $localNum[j][i]$ to 0 only if it has not already been set to qm by process $\langle i \rangle$ in action $M0$. We could also allow it to write 0 after that write of qm but before process $\langle i, j \rangle$ executes statement test. Such a write just decreases the possible executions, so eliminating this possibility doesn't forbid any possible executions.

```

fair process ( wrp  $\in$  WrProcs ) {
  wr: while ( TRUE ) {
    await  $\wedge$  localNum[self[2]][self[1]] = qm
           $\wedge$  pc[ $\langle$ self[1] $\rangle$ ]  $\in$  { "ncs", "M", "M0" };
    localNum[self[2]][self[1]] := 0;
  }
}

```

BEGIN TRANSLATION ($chksum(pcal) = "ffdaa638" \wedge chksum(tla) = "814037c2"$)

VARIABLES *number, localNum, localCh, pc, unread, v*

vars \triangleq $\langle number, localNum, localCh, pc, unread, v \rangle$

ProcSet \triangleq (*ProcIds*) \cup (*SubProcs*) \cup (*WrProcs*)

Init \triangleq Global variables

\wedge *number* = [$p \in$ *Procs* \mapsto 0]
 \wedge *localNum* = [$p \in$ *Procs* \mapsto [$q \in$ *OtherProcs*(p) \mapsto 0]]
 \wedge *localCh* = [$p \in$ *Procs* \mapsto [$q \in$ *OtherProcs*(p) \mapsto 0]]

Process main

\wedge *unread* = [$self \in$ *ProcIds* \mapsto {}]
 \wedge *v* = [$self \in$ *ProcIds* \mapsto 0]
 \wedge *pc* = [$self \in$ *ProcSet* \mapsto CASE $self \in$ *ProcIds* \rightarrow "ncs"
 \square $self \in$ *SubProcs* \rightarrow "ch"
 \square $self \in$ *WrProcs* \rightarrow "wr"]

ncs(*self*) \triangleq \wedge *pc*[*self*] = "ncs"
 \wedge TRUE
 \wedge *pc'* = [*pc* EXCEPT ![*self*] = "M"]
 \wedge UNCHANGED $\langle number, localNum, localCh, unread, v \rangle$

M(*self*) \triangleq \wedge *pc*[*self*] = "M"
 \wedge $\forall p \in$ *SubProcsOf*(*self*[1]) : *pc*[p] = "test"
 \wedge *unread'* = [*unread* EXCEPT ![*self*] = *OtherProcs*(*self*[1])]

$$\begin{aligned}
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"M0"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, localCh, v \rangle \\
M0(self) \triangleq & \wedge pc[self] = \text{"M0"} \\
& \wedge \text{IF } unRead[self] \neq \{\} \\
& \quad \text{THEN } \wedge \exists j \in unRead[self] : \\
& \quad \quad \wedge \text{IF } localNum[self[1]][j] \neq qm \\
& \quad \quad \quad \text{THEN } \wedge v' = [v \text{ EXCEPT } ![self] = Max(v[self], localNum[self[1]][j])] \\
& \quad \quad \quad \text{ELSE } \wedge \text{TRUE} \\
& \quad \quad \quad \wedge v' = v \\
& \quad \quad \wedge unRead' = [unRead \text{ EXCEPT } ![self] = unRead[self] \setminus \{j\}] \\
& \quad \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"M0"}] \\
& \quad \wedge \text{UNCHANGED } \langle number, localNum \rangle \\
& \text{ELSE } \wedge \exists n \in \{m \in Nat : m > v[self]\} : \\
& \quad \wedge number' = [number \text{ EXCEPT } ![self[1]] = n] \\
& \quad \wedge localNum' = [j \in Procs \mapsto \\
& \quad \quad [i \in OtherProcs(j) \mapsto \\
& \quad \quad \quad \text{IF } i = self[1] \text{ THEN } qm \\
& \quad \quad \quad \text{ELSE } localNum[j][i]]] \\
& \quad \wedge v' = [v \text{ EXCEPT } ![self] = 0] \\
& \quad \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"L"}] \\
& \quad \wedge \text{UNCHANGED } unRead \\
& \wedge \text{UNCHANGED } localCh \\
L(self) \triangleq & \wedge pc[self] = \text{"L"} \\
& \wedge \forall p \in SubProcsOf(self[1]) : pc[p] = \text{"ch"} \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"cs"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, localCh, unRead, v \rangle \\
cs(self) \triangleq & \wedge pc[self] = \text{"cs"} \\
& \wedge \text{TRUE} \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"P"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, localCh, unRead, v \rangle \\
P(self) \triangleq & \wedge pc[self] = \text{"P"} \\
& \wedge number' = [number \text{ EXCEPT } ![self[1]] = 0] \\
& \wedge localNum' = [j \in Procs \mapsto \\
& \quad [i \in OtherProcs(j) \mapsto \\
& \quad \quad \text{IF } i = self[1] \text{ THEN } qm \\
& \quad \quad \quad \text{ELSE } localNum[j][i]]] \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"ncs"}] \\
& \wedge \text{UNCHANGED } \langle localCh, unRead, v \rangle \\
main(self) \triangleq & ncs(self) \vee M(self) \vee M0(self) \vee L(self) \vee cs(self) \\
& \vee P(self) \\
ch(self) \triangleq & \wedge pc[self] = \text{"ch"}
\end{aligned}$$

$$\begin{aligned}
& \wedge pc[\langle self[1] \rangle] = \text{"M"} \\
& \wedge localCh' = [localCh \text{ EXCEPT } ![self[2]][self[1]] = 1] \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"test"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, unRead, v \rangle \\
test(self) & \triangleq \wedge pc[self] = \text{"test"} \\
& \wedge pc[\langle self[1] \rangle] = \text{"L"} \\
& \wedge localNum' = [localNum \text{ EXCEPT } ![self[2]][self[1]] = number[self[1]]] \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"Lb"}] \\
& \wedge \text{UNCHANGED } \langle number, localCh, unRead, v \rangle \\
Lb(self) & \triangleq \wedge pc[self] = \text{"Lb"} \\
& \wedge localCh' = [localCh \text{ EXCEPT } ![self[2]][self[1]] = 0] \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"L2"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, unRead, v \rangle \\
L2(self) & \triangleq \wedge pc[self] = \text{"L2"} \\
& \wedge localCh[self[1]][self[2]] = 0 \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"L3"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, localCh, unRead, v \rangle \\
L3(self) & \triangleq \wedge pc[self] = \text{"L3"} \\
& \wedge (localNum[self[1]][self[2]] \notin \{0, qm\}) \Rightarrow \\
& \quad (\langle number[self[1]], self[1] \rangle \ll \\
& \quad \quad \langle localNum[self[1]][self[2]], self[2] \rangle) \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"ch"}] \\
& \wedge \text{UNCHANGED } \langle number, localNum, localCh, unRead, v \rangle \\
sub(self) & \triangleq ch(self) \vee test(self) \vee Lb(self) \vee L2(self) \vee L3(self) \\
wr(self) & \triangleq \wedge pc[self] = \text{"wr"} \\
& \wedge \wedge localNum[self[2]][self[1]] = qm \\
& \quad \wedge pc[\langle self[1] \rangle] \in \{\text{"ncs"}, \text{"M"}, \text{"M0"}\} \\
& \wedge localNum' = [localNum \text{ EXCEPT } ![self[2]][self[1]] = 0] \\
& \wedge pc' = [pc \text{ EXCEPT } ![self] = \text{"wr"}] \\
& \wedge \text{UNCHANGED } \langle number, localCh, unRead, v \rangle \\
wrp(self) & \triangleq wr(self) \\
Next & \triangleq (\exists self \in ProcIds : main(self)) \\
& \quad \vee (\exists self \in SubProcs : sub(self)) \\
& \quad \vee (\exists self \in WrProcs : wrp(self)) \\
Spec & \triangleq \wedge Init \wedge \square [Next]_{vars} \\
& \wedge \forall self \in ProcIds : WF_{vars}((pc[self] \neq \text{"ncs"}) \wedge main(self)) \\
& \wedge \forall self \in SubProcs : WF_{vars}((pc[self] \neq \text{"L3"}) \wedge sub(self)) \\
& \wedge \forall self \in WrProcs : WF_{vars}(wrp(self))
\end{aligned}$$

END TRANSLATION

In statement *L3*, the await condition is satisfied if process $\langle i, j \rangle$ reads $localNum[self[1]][self[2]]$ equal to qm . This is because that's a possible execution, since the process could "interpret" the qm as 0. For checking safety (namely, mutual exclusion), we want to allow that because it's a possibility that must be taken into account. However, for checking liveness, we don't want to require that the statement must be executed when $localNum[self[1]][self[2]]$ equals qm , since that value could also be interpreted as $localNum[self[1]][self[2]]$ equal to 1, which could prevent the wait condition from being true. So we omit that fairness condition from the formula *Spec* produced by translating the algorithm, and we add weak fairness of the action when $localNum[self[1]][self[2]]$ does not equal qm . This produces the TLA+ specification *FSpec* defined here.

$$FSpec \triangleq \wedge Spec \\ \wedge \forall q \in SubProcs : WF_{vars}(L3(q) \wedge (localNum[q[1]][q[2]] \neq qm))$$

$$TypeOK \triangleq \wedge number \in [Procs \rightarrow Nat] \\ \wedge \wedge DOMAIN localNum = Procs \\ \wedge \forall i \in Procs : localNum[i] \in [OtherProcs(i) \rightarrow Nat \cup \{qm\}] \\ \wedge \wedge DOMAIN localCh = Procs \\ \wedge \forall i \in Procs : localCh[i] \in [OtherProcs(i) \rightarrow \{0, 1\}]$$

$$MutualExclusion \triangleq \forall p, q \in ProcIds : (p \neq q) \Rightarrow (\{pc[p], pc[q]\} \neq \{\text{"cs"}\})$$

$$StarvationFree \triangleq \forall p \in ProcIds : (pc[p] = \text{"M"}) \rightsquigarrow (pc[p] = \text{"cs"})$$

Checking the invariant in the appendix of the paper.

$$inBakery(i, j) \triangleq \vee pc[\langle i, j \rangle] \in \{\text{"Lb"}, \text{"L2"}, \text{"L3"}\} \\ \vee \wedge pc[\langle i, j \rangle] = \text{"ch"} \\ \wedge pc[\langle i \rangle] \in \{\text{"L"}, \text{"cs"}\}$$

$$inCS(i) \triangleq pc[\langle i \rangle] = \text{"cs"}$$

In TLA+, we can't write both $inDoorway(i, j, w)$ and $inDoorway(i, j)$, so we change the first to $inDoorwayVal$. Its definition differs from the definition of $inDoorway(i, j, w)$ in the paper to avoid having to add a history variable to remember the value of $localNum[self[1]][j]$ read in statement *M0*. It's a nicer definition, but it would have required more explanation than the definition in the paper.

The definition of $inDoorway(i, j)$ is equivalent to the one in the paper. It is obviously implied by $\exists w \in Nat : inDoorwayVal(i, j, w)$, and type correctness implies the opposite implication.

$$inDoorwayVal(i, j, w) \triangleq \vee \wedge pc[\langle i \rangle] = \text{"M0"} \\ \wedge j \notin unRead[\langle i \rangle] \\ \wedge v[\langle i \rangle] \geq w \\ \vee \wedge pc[\langle i \rangle] = \text{"L"} \\ \wedge pc[\langle i, j \rangle] = \text{"test"} \\ \wedge number[i] > w \quad \text{sm: replaced } \geq \text{ by } > \text{ (Aug 24)}$$

$$inDoorway(i, j) \triangleq \vee \wedge pc[\langle i \rangle] = \text{"M0"} \\ \wedge j \notin unRead[\langle i \rangle]$$

$$\begin{aligned} & \vee \wedge pc[\langle i \rangle] = \text{"L"} \\ & \wedge pc[\langle i, j \rangle] = \text{"test"} \quad \text{sm: added Aug 23, 2021} \end{aligned}$$

$$Outside(i, j) \triangleq \neg(inDoorway(i, j) \vee inBakery(i, j))$$

$$\begin{aligned} passed(i, j, LL) \triangleq & \text{IF } LL = \text{"L2"} \text{ THEN } \vee pc[\langle i, j \rangle] = \text{"L3"} \\ & \vee \wedge pc[\langle i, j \rangle] = \text{"ch"} \\ & \wedge pc[\langle i \rangle] \in \{\text{"L"}, \text{"cs"}\} \\ & \text{ELSE } \wedge pc[\langle i, j \rangle] = \text{"ch"} \\ & \wedge pc[\langle i \rangle] \in \{\text{"L"}, \text{"cs"}\} \end{aligned}$$

$$\begin{aligned} Before(i, j) \triangleq & \wedge inBakery(i, j) \\ & \wedge \vee Outside(j, i) \\ & \vee inDoorwayVal(j, i, number[i]) \\ & \vee \wedge inBakery(j, i) \\ & \wedge \langle number[i], i \rangle \ll \langle number[j], j \rangle \\ & \wedge \neg passed(j, i, \text{"L3"}) \end{aligned}$$

$$\begin{aligned} Inv(i, j) \triangleq & \wedge inBakery(i, j) \Rightarrow Before(i, j) \vee Before(j, i) \\ & \vee inDoorway(j, i) \\ & \wedge passed(i, j, \text{"L2"}) \Rightarrow Before(i, j) \vee Before(j, i) \\ & \wedge passed(i, j, \text{"L3"}) \Rightarrow Before(i, j) \end{aligned}$$

$$I \triangleq \forall i \in Procs : \forall j \in OtherProcs(i) : Inv(i, j)$$

The following is for testing. Since the spec allows the values of $number[n]$ to get arbitrarily large, there are infinitely many states. The obvious solution to that is to use models with a state constraint that $number[n]$ is at most some value $TestMaxNum$. However, TLC would still not be able to execute the spec because the with statement in action M allows an infinite number of possible values for $number[n]$. To solve that problem, we have the model redefine Nat to a finite set of numbers. The obvious set is $0 \dots TestMaxNum$. However, trying that reveals a subtle problem. Running the model produces a bogus counterexample to the $StarvationFree$ property.

This is surprising, since constraints on the state space generally fail to find real counterexamples to a liveness property because the counterexamples require large (possibly infinite) traces that are ruled out by the state constraint. The remaining traces may not satisfy the liveness property, but they are ruled out because they fail to satisfy the algorithm's fairness requirements. In this case, a behavior that didn't satisfy the liveness property $StarvationFree$ but shouldn't have satisfied the fairness requirements of the algorithm did satisfy the fairness requirement because of the substitution of a finite set of numbers for Nat .

Here's what happened: In the behavior, two nodes kept alternately entering the critical section in a way that kept increasing their values of `num` until one of those values reached `TestMaxNum`. That one entered its critical section while the other was in its noncritical section, re-entered its noncritical section, and then the two processes kept repeating this dance forever. Meanwhile, a third process's subprocess was trying to execute action `M`. Every time it tried to execute that action, it saw that another process's number equaled `TestMaxNum`. In a normal execution, it would just set its value of `num` larger than `TestMaxNum` and eventually enter its critical section. However, it couldn't do that because the substitution of $0 \dots TestMaxNum$ for `Nat` meant that it couldn't set `num` to such a value, so the enter step was disabled. The fairness requirement on the enter action is weak fairness, which requires an action eventually to be taken only if it's continually enabled. Requiring strong fairness of the action would have solved this problem, because the enabled action kept being enabled and strong fairness would rule out a behavior in which that process's enter step never occurred. However, it's important that the algorithm satisfy starvation freedom without assuming strong fairness of any of its steps.

The solution to this problem is to substitute $0 \dots (TestMax + 1)$ for `Nat`. The state constraint will allow the enter step to be taken, but will allow no further steps from that state. The process still never enters its critical section, but now the behavior that keeps it from doing so will violate the weak fairness requirements on that process's steps.

$TestMaxNum \triangleq 6$
 $TestNat \triangleq 0 \dots (TestMaxNum + 1)$

```
*****
Old Version, with statement M atomic Test Results Default fairness (without the correction to
L3 fairness):
N = 2, TestMaxNum = 6, 2,388 states 0:05 on Azure [Default fairness]
N = 3, TestMaxNum = 4, 5,119,808 states in 27:05 + 7:20 on Azure

Correct Fairness
N = 3, TestMaxNum = 5, 9,382,640 states in 40:34 + 5:57 on Azure
N = 3, TestMaxNum = 6, 15,530,720 states in 1:06:31 + 9:26 on Azure
N = 4, TestMaxNum = 2, on Azure [safety only] killed, it would have taken days

Version of 27 April 2021 with M deconstructed
N = 2, TestMaxNum = 6, 3,844
N = 3, TestMaxNum = 3, 12,127,440 states 1:07:06 + 12:06 on Azure (testing  $\square \diamond inCS$ )
N = 3, TestMaxNum = 4, 38,818,800 states 2:44:00 + 0:26:01 on Azure
N = 3, TestMaxNum = 3, 12,127,440 states on Azure (testing invariance of I

Version of 28 April 2021 with handling of asynchronous writing fixed all checking I, Mutex &
StarvationFree
N = 2, TestMaxNum = 6, 2500 states
N = 3, TestMaxNum = 3, 1,794,168 states in 08:07 + 1:52 on Azure
N = 3, TestMaxNum = 4, 3,211.104 states in 14:06 + 3:07 on Azure
N = 3, TestMaxNum = 5, 12,071,392 states in 17:05 + 6:58 on Azure
N = 4, TestMaxNum = 2 killed because it would have taken days.
*****
```

```
\ * Modification History
\ * Last modified Wed Nov 17 18:42:50 CET 2021 by merz
\ * Last modified Thu Jul 01 12:24:37 CEST 2021 by merz
\ * Last modified Wed Apr 28 18:06:24 PDT 2021 by lamport
\ * Created Sat Apr 24 09:45:26 PDT 2021 by lamport6
```