Synchronization: Await

CS 536: Science of Programming, Fall 2020

A. Why?

- Avoiding interference isn't the same as coordinating desirable activities.
- It's common for one thread to wait for another thread to reach a desired state.

B. Objectives

At the end of this lecture you should know

- The syntax and semantics of the `await` statement.
- How to draw an evaluation diagram for a parallel program that uses `await`.
- How to recognize deadlocked configurations in an evaluation diagram.
- How to list the potential deadlock predicates for a parallel program that uses `await`.

C. Synchronization

- We've looked at parallel programs whose threads avoid bad interactions
  - They don't interfere because they don't interact (disjoint programs/conditions).
  - They interact but don't interfere (interference-freedom).
- To supporting good interaction between threads, we often have to have one thread wait for another one. Some examples:
  - Thread 1 should wait until thread 2 is finished executing a certain block of code.
  - Thread 1 has to wait until some buffer is not empty
  - Thread 2 has to wait until some buffer is not full.
- The general problem is that we often want threads to synchronize: We want one thread to wait until some other thread makes a condition come true.

Example 1: For a more specific example, in the following program, the calculation of u doesn't start until we finish calculating z, even though u doesn't depend on z.

```plaintext
[x := ... | | y := ... | | z := ...]; u = f(x, y); v := g(u, z)
```

On the other hand, we can't nest parallel programs, so we can't write

```plaintext
[[[x := ... | | y := ...]; u = f(x, y) | | z := ...]; v := g(u, z)
```

which would be a natural way to do the calculations of u and z in parallel. In some sense, what we'd like is to run something like

```plaintext
[x := ... | | y := ... | | wait for x and y; u = f(x, y) | | z := ...]; v := g(u, z)//
```
D. The Await Statement

- It's time to introduce a new statement, the `await` statement, whose semantics implements the notion of waiting until some condition is true.
  - Could use busy wait loops like `while ¬B do skip od {B}` but this is wasteful.

**Syntax:** `await B then S end` where `B` is a boolean expression and `S` is a statement.

- `S` isn't allowed to have loops, `await` statements, or atomic regions.
- `await` statements can only appear in parallel programs (in some thread `S₁ || S₂ || ...`).

An `await` statement is a **conditional atomic region**. If we choose to run the thread that begins with `await B then S end`, then

- If `B` is true, immediately and atomically execute `S`.
- If `σ(B) = T` and `<S, σ> →* ⟨E, τ⟩`, then `<await B then S end, σ> → ⟨E, τ⟩`.
- No other thread can run between the `await` noticing `σ(B) = T` and evaluating `S`, nor can another thread run while `S` executes.
- We can see this from the → notation because each → step is a indivisible unit of execution.
- The nondeterministic choice of which thread to run is made after one arrow ends and the next arrow begins.
- On the other hand, when we make the nondeterministic choice of which thread to run, `σ(B)` being true doesn't force the choice of executing the `await`. If some other thread besides this one can also run, then that choice is allowed.
- If `B` is not true, wait until it is, then continue as above.
  - If `σ(B) = F`, there's no transition rule and no out arrow from the current configuration.
  - Nothing occurs and the thread is **blocked** (= the thread waits) until `B` becomes true.
  - If `B` never becomes true, we wait forever and the program **deadlocks**.

**Example 2:** Here are some examples of configurations and transitions.

- This `await` is blocked, so the configuration has no out arrows.
  ```
  ⟨await z ≠ 0 then x := 0; y := 1 end, {z = 0}⟩.
  ```
- This `await` can execute; note the body is executed atomically with the test
  ```
  ⟨await z ≠ 0 then x := 0; y := 1 end, {z = 5}⟩ → ⟨E, {z = 5, x = 0, y = 1}⟩
  ```
- This is a conditional statement; note that one step of execution has us jumping to the start of the true branch, so another thread could execute before we continue with `x := 0`.
  ```
  ⟨if z ≠ 0 then x := 0; y := 1 fi, {z = 5}⟩ → ⟨x := 0; y := 1, {z = 5}⟩
  ```

**Abbreviations:**

- `< S >` (S as an atomic statement) is an abbreviation for `await T then S end`. (It doesn't wait, but it does execute `S` atomically.)
- `wait B` is an abbreviation for `await B then skip end`. (Once it finishes waiting, it's done.)
There's an important difference between `wait B; S` and `await B then S end`.

`wait B; S` means `await B then skip end`; `S`, so it allows another thread to be executed after the `wait` but before running `S`. Because of this, in general, you can't rely on `B` being true when you execute `S`.

**Await Statement Proof rule**

- Here's the proof rule for the `await` statement:

**Await Statement (a.k.a. Synchronization Rule)**

1. `{p ∧ B} S {q}
2. `{p} await B then S end {q}

- Proof Outlines with `await`
  - Minimal outline: `{p} await B then S end {q}
  - Full outline: `{p} await B then {p ∧ B} S* {q} end {q} where S* is a proof outline for S.

- Weakest Precondition rule: \(\text{wp}(\text{await } B \text{ then } S \text{ end}, q) \equiv B \rightarrow \text{wp}(S, q)\).
  - This guarantees \({B \rightarrow \text{wp}(S, q)}\) `await B then {wp(S, q)} S* {q} end {q}`

- Note: Don't let `{p} \rightarrow \neg B` because it guarantees blockage:
  
  \[\{p ∧ \neg B\} \text{ await } B \text{ then } \{p ∧ \neg B ∧ B\} S* \{q\} \text{ end } \{q\}\]

**E. Example: Producer/Consumer Problem**

- The Producer/Consumer Problem (a.k.a. Bounded Buffer Problem) is a standard problem in parallel programming.

- We have two threads running in parallel: The producer creates things and puts them into a buffer; the consumer removes things from the buffer and does something with them.

- The problem is that if the buffer is full, the producer shouldn't add anything to the buffer; if the buffer is empty, the consumer shouldn't remove anything from the buffer.

- **Example 3**: The rough code to solve this is

  ```
  Initialize(buffer);
  [while ¬done do // Producer
    thing_p := Create();
    await NotFull(buffer) then
      BufferAdd(buffer, thing_p)
    end
  od
  || while ¬done do // Consumer
    await NotEmpty(buffer) then
      thing_c := BufferRemove(buffer);
    end;
    Consume(thing_c)
  od
  ]
  ```
• Buffer operations need to be synchronized because the threads share the buffer. The threads don’t share the individual thing objects, so the Create and Consume calls can go outside the await and interleave execution.

F. Deadlock

• Threads can block themselves (trivial example: await false then S end).
• More often, threads block because they're waiting for conditions they expect other threads to establish. E.g., if we’re running in a state where y = 0 and x = 0, then these two threads deadlock:
  • Thread 1: \( \{p_1\} \text{ await } y \neq 0 \text{ then } x := 1; \ldots \)
  • Thread 2: \( \{p_2\} \text{ await } x \neq 0 \text{ then } y := 1; \ldots \)

• Definition: A parallel program is deadlocked if it has not finished execution and all its threads that have not completed are waiting at await statements. I.e., all of the threads are either complete or blocked, and at least one thread is blocked.

• A program might deadlock under all execution paths or only certain execution paths.
• Example 4: The program
  \[
  \text{[await } y \neq 0 \text{ then } x := 1 \text{ end } | | \text{ await } x \neq 0 \text{ then } y := 1 \text{ end]}
  \]
deadlocks iff you execute in a state where x and y are both zero.
• Example 5: If thread 1 sets \( x := 0 \) before thread 2 evaluates its wait \( x \), then thread 2 will block. (Recall \( \text{ wait } x = \text{ await } x \text{ then skip end.} \))
  \[
  \{T\} x := 1; y := 1;
  \text{[wait } y = 1; x := 0 | | \text{ wait } x = 1; y := 0]
  \{x = 0 \land y = 0\}
  \]
Here’s an execution graph for this program in state \( \{x = 1, y = 1\} \) (somewhat abbreviated). There are two deadlocking paths (and four paths that terminate correctly).

• Note: There's a big difference between \( \text{wait } y = 1; x := 0 \) (i.e., \( \text{await } y = 1 \text{ then skip end; } x := 0 \)) and \( \text{await } y = 1 \text{ then } x := 0 \text{ end.} \) (See the activity questions for what happens with the example above if we change each \( \text{wait } B; \) assignment to \( \text{await } B \text{ then } \text{assignment end.} \))
- Obviously, we'd like to know if a program is going to deadlock. The following test identifies a set of predicates that indicate potential problems with a program; if none of these predicates is satisfiable, then deadlock is guaranteed not to occur.
- If one or more of these predicates is satisfiable, then we can't guarantee that deadlock will not occur, but we aren't guaranteeing that deadlock must occur. (So the deadlock conditions are sufficient to show deadlock is impossible but they are not necessary conditions.)
- Let \( \langle p \rangle \ \langle \{ p_1 \} S_1^* \{ q_1 \} \ | \ | \ (p_2) S_2^* \{ q_2 \} \ | \ | \ ... \ | \ | \ (p_n) S_n^* \{ q_n \} \rangle \{ q \} \) be a full outline for a parallel program, where \( p = p_1 \land \ldots \land p_n \) and \( q = q_1 \land \ldots \land q_n \).
- **Definition:** A (potential) deadlock condition for the program outline above is a predicate of the form \( r'_1 \land r'_2 \land \ldots \land r'_n \) where each \( r'_k \) is either
  - \( q_k \), the postcondition for thread \( S_k \) or
  - \( p \land \neg B \) where \( \langle p \rangle \) await \( B \ldots \) appears in the proof outline for thread \( S_k \).
  - In addition, at least one of the \( r'_k \) must involve waiting. I.e., \( q = q_1 \land \ldots \land q_n \) is not a potential deadlock condition.
- A program outline is **deadlock-free** if every one of its potential deadlock conditions is unsatisfiable (i.e., a contradiction):
  - I.e., for each deadlock condition \( r' \), we have \( \vdash \neg r' \) (or the equivalent \( \vdash r' \rightarrow F \)).

**Parallelism with Deadlock Freedom**

1. \( \{ p_1 \} S_1^* \{ q_1 \} \)
2. \( \{ p_2 \} S_2^* \{ q_2 \} \)
\[ n \cdot \{ p_n \} \cdot S_n^* \cdot \{ q_n \} \]
\[ n+1 \cdot \{ p_1 \land p_2 \land \ldots \land p_n \} \]
\[ \{ q_1 \land q_2 \land \ldots \land q_n \} \]

where the \{ p_k \} \cdot S_k^* \cdot \{ q_k \} are pairwise interference-free standard proof outlines and the parallel program outline is deadlock-free.

**G. Example 6: An Example of Deadlock Conditions**

- Let's take the program from Example 4:
  \[
  [\text{await } y \neq 0 \text{ then } x := 1 \text{ end} \mid | \text{await } x \neq 0 \text{ then } y := 1 \text{ end}]
  \]

  and develop an annotation for it:
  \[
  \{ T \}
  \]
  \[
  [\{ T \} \text{ await } y \neq 0 \text{ then } \{ y \neq 0 \} \cdot x := 1 \cdot \{ x \neq 0 \land y \neq 0 \} \cdot \text{end} \cdot \{ x \neq 0 \land y \neq 0 \}]
  \mid | \{ T \} \text{ await } x \neq 0 \text{ then } \{ x \neq 0 \} \cdot y := 1 \cdot \{ x \neq 0 \land y \neq 0 \} \cdot \text{end} \cdot \{ x \neq 0 \land y \neq 0 \}
  \]
  \[
  \{ x \neq 0 \land y \neq 0 \}
  \]

- Let set \( D_1 = \{ x \neq 0 \land y \neq 0, y = 0 \} \) be the choices for \( p_1 \).
  - \( x \neq 0 \land y \neq 0 \) is the thread postcondition
  - \( y = 0 \) indicates thread 1 is blocked at the \textit{await} statement.

- Similarly, let set \( D_2 = \{ x \neq 0 \land y \neq 0, x = 0 \} \) be the choices for \( p_2 \) (the postcondition of thread 2 and the blocking condition for its \textit{await}).

- There are three choices for the potential deadlock predicate \( r_1 \land r_2 \):
  - \( (x \neq 0 \land y \neq 0) \land (x = 0) \), which is a contradiction.
  - \( (y = 0) \land (x \neq 0 \land y \neq 0) \), which is a contradiction.
  - \( (y = 0) \land (x = 0) \), which is not a contradiction, therefore, it's a potential deadlock condition, and our program does not pass the deadlock-freedom test.

- Recall \( (x \neq 0 \land y \neq 0) \land (x \neq 0 \land y \neq 0) \) is not a potential deadlock predicate because it says that the two threads have both completed.

- One way out of this predicament is to make the initial precondition the negation of \( y = 0 \land x = 0 \). Let \( \{ p \} \) be \( (x \neq 0 \lor y \neq 0) \) in
  \[
  \{ p \}
  \]
  \[
  [\{ p \} \text{ await } y \neq 0 \text{ then } \{ p \land y \neq 0 \} \cdot x := 1 \cdot \{ x \neq 0 \land y \neq 0 \} \cdot \text{end} \cdot \{ x \neq 0 \land y \neq 0 \}]
  \mid | \{ p \} \text{ await } x \neq 0 \text{ then } \{ p \land x \neq 0 \} \cdot y := 1 \cdot \{ x \neq 0 \land y \neq 0 \} \cdot \text{end} \cdot \{ x \neq 0 \land y \neq 0 \}
  \]
  \[
  \{ x \neq 0 \land y \neq 0 \}
  \]

- Let \( D_1 = \{ x \neq 0 \land y \neq 0, p \land y = 0 \} \) and let \( D_2 = \{ x \neq 0 \land y \neq 0, p \land x = 0 \} \).

- The three potential deadlock predicates are now contradictory
  - \( (x \neq 0 \land y \neq 0) \land (p \land x = 0) \) (is false because of \( x \neq 0 \land x = 0 \))
• \((p \land y = 0) \land (x \neq 0 \land y \neq 0)\) (is false because of \(y = 0 \land y \neq 0\))

• \((p \land y = 0) \land (p \land x = 0)\)  
  \(= ((x \neq 0 \lor y \neq 0) \land y = 0) \land ((x \neq 0 \lor y \neq 0) \land x = 0)\)  
  \(\Rightarrow (x \neq 0 \land y = 0) \land (y \neq 0 \land x = 0)\)  
  \(\Rightarrow F\)

**H. Example 7: Another Example of Deadlock Conditions**

• Here's an example with three threads: Thread 1 has one `await` statement, thread 2 has two `await` statements, and thread 3 has no `await` statements.

```plaintext
[ ... {p₁₁} await B₁₁ ... {q₁} ]  
| | ... {p₂₁} await B₂₁ ... {p₂₂} await B₂₂ ... {q₂}  
| | ... {q₃} ]
```

• Let \(D₁ = \{p₁₁ \land \neg B₁₁, q₁\}\), let \(D₂ = \{p₂₁ \land \neg B₂₁, p₂₂ \land \neg B₂₂, q₂\}\), and let \(D₃ = \{q₃\}\).

• Let \(D = \{r₁ \land r₂ \land r₃ | r₁ \in D₁, r₂ \in D₂, r₃ \in D₃\} \setminus \{q₁ \land q₂ \land q₃\}\) be the set of deadlock conditions:

- \((p₁₁ \land \neg B₁₁) \land (p₂₁ \land \neg B₂₁) \land q₃\),  -- Thread 1 blocked, thread 2 blocked at 1st await
- \((p₁₁ \land \neg B₁₁) \land (p₂₂ \land \neg B₂₂) \land q₃\),  -- Thread 1 blocked, thread 2 blocked at 2nd await
- \((p₁₁ \land \neg B₁₁) \land q₂ \land q₃\),  -- Thread 1 blocked
- \(q₁ \land (p₂ \land \neg B₂₁) \land q₃\),  -- Thread 2 blocked at 1st await
- \(q₁ \land (p₂₂ \land \neg B₂₂) \land q₃\)  -- Thread 2 blocked at 2nd await

• The program is deadlock-free if every predicate in \(D\) is a contradiction (i.e., unsatisfiable).

**I. Strengthening Deadlock Conditions**

• Having all deadlock conditions be contradictory is sufficient for guaranteeing that no program execution will deadlock.

• It's not a necessary condition, however. Just because some \(r \in D\) is satisfiable, that doesn't mean that there exists a program execution that can get to the corresponding deadlocked configuration.

**J. Example 8: Proving Deadlock Freedom by Strengthening Conditions**

• This program doesn't deadlock:

```plaintext
\{T\} n := 0; \{\text{await } n = 0 \text{ then } n := 1 \text{ end } | | \text{ wait } n = 1\} \{n > 0\}
```

• If we annotate the program as below, we have sequential correctness for each thread, plus the threads are interference-free:

```plaintext
\{T\} n := 0; {T}  
[ {T} \text{ await } n = 0 \text{ then } n := 1 \text{ end } \{n > 0\}  
| | {T} \text{ wait } n = 1 \{n > 0\}  
] \{n > 0\}
```
• On the other hand, we can't prove deadlock freedom. In fact, all \(2 \times 2 - 1 = 3\) deadlock conditions are satisfiable:
  - \(n \neq 0 \land n \neq 1\) — Both threads blocked
  - \(n \neq 0 \land n > 0\) — Thread 1 blocked
  - \(n > 0 \land n \neq 1\) — Thread 2 blocked
• The problem here is that the proof outline's conditions are too weak. We want each deadlock condition to be logically equivalent to false, the strongest predicate.
• To make a conjunctive formula stronger, we need to strengthen its conjuncts. For a deadlock-freedom test, we have two kinds of conjuncts:
  - (postcondition of thread)
  - (precondition of \textit{await} statement) \land \neg (test of \textit{await} statement)
• By strengthening the postcondition of the initial assignment of \(n := 0\) from true to \(n = 0\), we can strengthen the precondition of the first \textit{await}:
  
  \[
  \{T\} \ n := 0; \ \{n = 0\} \\
  [ \ {n = 0} \ await \ n = 0 \ then \ {n = 0 \land n = 0} \ n := 1 \ end \ {n > 0} \\
  | | \ {T} \ await \ n = 1 \ then \ {n = 1} \ skip \ {n = 1} \ end \ {n > 0} ] \\
  \{n > 0\}
  \]
• The potential deadlock conditions for the proof outline above are now
  - \((n = 0 \land n \neq 0) \land n \neq 1\) — Both threads blocked (contradiction)
  - \((n = 0 \land n \neq 0) \land n > 0\) — Thread 1 blocked (contradiction)
  - \(n > 0 \land n \neq 1\) — Thread 2 blocked (satisfiable)
• So two of the conditions are contradictory, but one condition is still satisfiable. To prove deadlock-freedom, we need to strengthen the conditions even more to include the state we get to when the first thread has executed and the second thread hasn't.
• Unfortunately, if we annotate the two threads as
  - \{n = 0\} \textit{await} \ n = 0 \ then \ n := 1 \ end \ {n = 1} \\
  - \{n = 1\} \textit{wait} \ n = 1 \ {n = 1}
• Then the precondition of the parallel program has to be \((n = 0) \land (n = 1)\), which doesn't follow from the strongest postcondition of \(n := 0\) and worse yet, isn't possible anyway.
  
  \[
  \{T\} \ n := 0; \\
  \{n = 0 \land n = 1\} // \leftarrow \text{error} \\
  [ \ {n = 0} \ await \ n = 0 \ then \ n := 1 \ end \ {n = 1} \\
  | | \ {n = 1} \textit{wait} \ n = 1 \ {n = 1} \\
  ] \ {n = 1 \land n = 1} \ {n = 1}
  \]
• Before thread 2 runs, it sees \(n = 0\) or \(n = 1\) depending on whether thread 1 has run yet. If we use that as the precondition for thread 2, then we get \(n = 0 \land (n = 0 \lor n = 1)\) as the precondition for the parallel program, which works:
\{T\} \ n := 0;
\{n = 0 \land (n = 0 \lor n = 1)\}
\[\{n = 0\} \text{ await } n = 0 \text{ then } n := 1 \{n = 1\}\]
\| \{n = 0 \lor n = 1\} \text{ wait } n = 1 \{n = 1\}\]
\{n = 1 \land n = 1\} \{n = 1\}

- Better still, the deadlock conditions are now all contradictions, so we have deadlock-freedom
  - \( (n = 0 \land n \neq 0) \land ((n = 0 \lor n = 1) \land n \neq 1) \) — Both blocked (contradiction)
  - \( (n = 0 \land n \neq 0) \land n = 1 \) — Thread 1 blocked (contradiction)
  - \( n = 1 \land ((n = 0 \lor n = 1) \land n \neq 1) \) — Thread 2 blocked (contradiction)

- Unfortunately, one of the interference freedom tests now fails:
  - Pass: \( \{n = 0 \land (n = 0 \lor n = 1)\} \text{ await } n = 0 \text{ then } n := 1 \{n = 0 \lor n = 1\}\]
  - Pass: \( \{n = 0 \land n = 1\} \text{ await } n = 0 \text{ then } n := 1 \{n = 1\}\]
  - Fail: \( \{(n = 0 \lor n = 1) \land n = 0\} \text{ wait } n = 1 \{n = 0\}\]

- We can solve this problem by adding an auxiliary variable to say whether or not the first thread has run and set \( n = 1 \).