#### simmer

#### Discrete-Event Simulation for R



**Iñaki Úcar**, postdoctoral fellow @ UC3M-Santander Big Data Institute (IBiDat) Bart Smeets, data scientist @ dataroots.io

November 15-16, 2019



#### Simulation

From R. Shannon (1975), simulation is

the process of designing a **model of a real system** and conducting experiments with this model for the purpose either of **understanding the behavior of the system** or of **evaluating various strategies** [...] for the operation of the system.



#### Simulation

From R. Shannon (1975), simulation is

the process of designing a **model of a real system** and conducting experiments with this model for the purpose either of **understanding the behavior of the system** or of **evaluating various strategies** [...] for the operation of the system.

Taxonomy, from Law and Kelton (2000):

- 1. deterministic vs. stochastic
- 2. (time component?) static vs. dynamic
- 3. (if dynamic) continuous vs. discrete



#### Simulation

From R. Shannon (1975), simulation is

the process of designing a **model of a real system** and conducting experiments with this model for the purpose either of **understanding the behavior of the system** or of **evaluating various strategies** [...] for the operation of the system.

Taxonomy, from Law and Kelton (2000):

- 1. deterministic vs. stochastic
- 2. (time component?) static vs. dynamic
- 3. (if dynamic) continuous vs. discrete

#### Examples:

- deterministic + dynamic + continuous = *Dynamical Systems*
- stochastic + static = Monte Carlo Simulation
- stochastic + dynamic + discrete = Discrete-Event Simulation (DES)



#### Simulation

**DES** 

What can be modelled as a Discrete-Event Simulation (DES)?

- An **event** is an instantaneous occurrence that may change the **state of the system**
- The number of events is countable
- Between events, all the state variables remain constant
- Output: snapshots of the state of the system over time



#### Simulation

#### DES

What can be modelled as a Discrete-Event Simulation (DES)?

- An **event** is an instantaneous occurrence that may change the **state of the system**
- The number of events is countable
- Between events, all the state variables remain constant
- Output: snapshots of the state of the system over time

#### Common examples:

- customers arriving at a bank,
- products being manipulated in a supply chain,
- packets traversing a network,
- ..

and many more applications from manufacturing systems, construction engineering, project management, logistics, transportation systems, business processes, healthcare, telecommunications networks...



#### Simulation

DES

Programming styles (Banks 2005):

- Activity-oriented: fixed time increments; scan and verify conditions
- **Event-oriented**: event routines; event list ordered by time of ocurrence
- **Process-oriented**: life cycle of entities of the real system; triggered by events



#### Simulation

#### DES

Programming styles (Banks 2005):

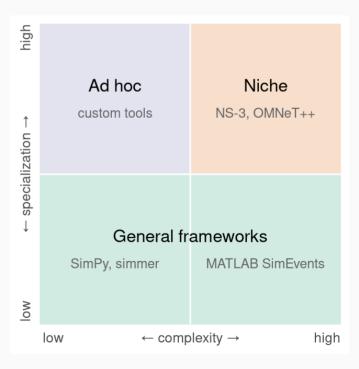
- Activity-oriented: fixed time increments; scan and verify conditions
- **Event-oriented**: event routines; event list ordered by time of ocurrence
- Process-oriented: life cycle of entities of the real system; triggered by events

#### Spectrum of tools:

• High **complexity** and **specialization** generally means more accuracy

#### but

- More specialization requires more effort
- More complexity requires more effort





#### Simulation

DES

simmer



#### Main characteristics:

- General versatile framework for fast prototyping
- Rich and user-friendly R API over a fast C++ simulation core
- Process-oriented trajectory-based modelling
- Automatic monitoring capabilities
- Integration with R: repeatability, analysis, visualization



#### Simulation

DES

simmer



#### Main characteristics:

- General versatile framework for fast prototyping
- Rich and user-friendly R API over a fast C++ simulation core
- Process-oriented trajectory-based modelling
- Automatic monitoring capabilities
- Integration with R: repeatability, analysis, visualization

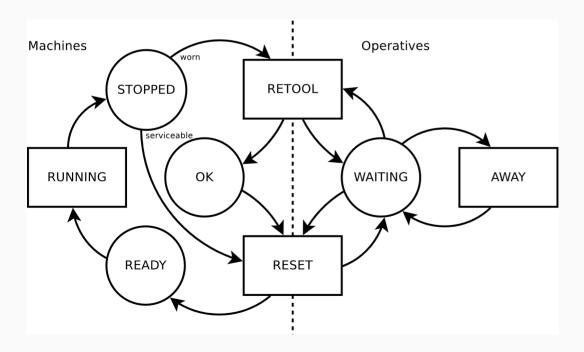
#### Resources:

- Online documentation (manual + 10 vignettes): https://r-simmer.org
- **Ucar I**, Smeets B, Azcorra A (2019). "simmer: Discrete-Event Simulation for R." *Journal of Statistical Software*, 90(2), 1-30. doi: 10.18637/jss.v090.i02.
- **Ucar I**, Hernández JA, Serrano P, Azcorra A (2018). "Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping." *IEEE Communications Magazine*, 56(11), 145-151. doi: 10.1109/MCOM.2018.1700960.



#### From M. Pidd (1988), Section 5.3.1:

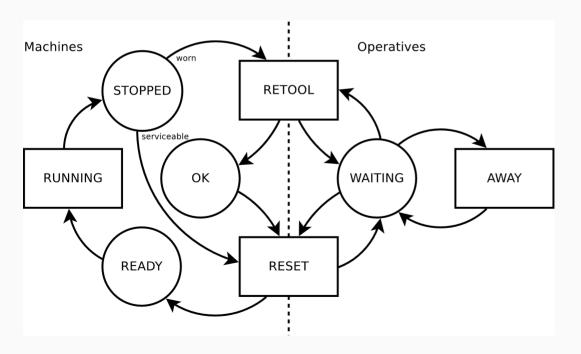
- Jobs are allocated to the first available machine.
- Machines need to be retooled (sometimes) and reset by an operative.
- Operatives may be away attending other tasks.





From M. Pidd (1988), Section 5.3.1:

- Jobs are allocated to the first available machine.
- Machines need to be retooled (sometimes) and reset by an operative.
- Operatives may be away attending other tasks.

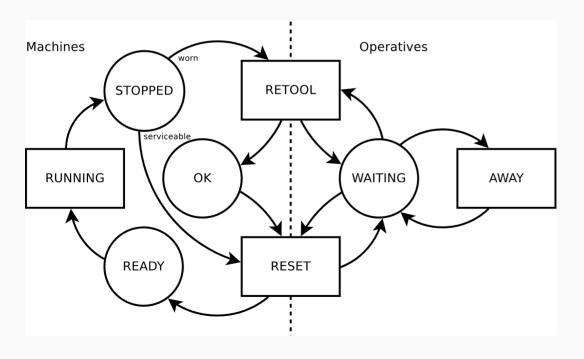


```
job ← trajectory() %>%
  seize("machine") %>%
  timeout(RUNNING) %>%
  branch(
    CHECK_WORN, continue = TRUE,
    trajectory() %>%
      seize("operative") %>%
      timeout(RETOOL) %>%
     release("operative")
  ) %>%
  seize("operative") %>%
  timeout(RESET) %>%
  release("operative") %>%
  release("machine")
```



From M. Pidd (1988), Section 5.3.1:

- Jobs are allocated to the first available machine.
- Machines need to be retooled (sometimes) and reset by an operative.
- Operatives may be away attending other tasks.



```
job ← trajectory() %>%
  seize("machine") %>%
  timeout(RUNNING) %>%
  branch(
    CHECK WORN, continue = TRUE.
    trajectory() %>%
      seize("operative") %>%
     timeout(RETOOL) %>%
     release("operative")
  ) %>%
  seize("operative") %>%
  timeout(RESET) %>%
  release("operative") %>%
  release("machine")
```

```
task ← trajectory() %>%
  seize("operative") %>%
  timeout(AWAY) %>%
  release("operative")
```



```
library(simmer); set.seed(1234)

RUNNING ← function() rexp(1, 1)
CHECK_WORN ← function() runif(1) < 0.2

RETOOL ← function() rexp(1, 2)

RESET ← function() rexp(1, 3)

AWAY ← function() rexp(1, 1)</pre>
```

```
job ← trajectory() %>%
  seize("machine") %>%
  timeout(RUNNING) %>%
  branch(
   CHECK_WORN, continue = TRUE,
    trajectory() %>%
     seize("operative") %>%
     timeout(RETOOL) %>%
     release("operative")
  ) %>%
  seize("operative") %>%
  timeout(RESET) %>%
  release("operative") %>%
  release("machine")
```

```
task ← trajectory() %>%
  seize("operative") %>%
  timeout(AWAY) %>%
  release("operative")
```



```
library(simmer); set.seed(1234)
RUNNING ← function() rexp(1, 1)
CHECK_WORN ← function() runif(1) < 0.2</pre>
RETOOL \leftarrow function() rexp(1, 2)
RESET \leftarrow function() rexp(1, 3)
AWAY \leftarrow function() rexp(1, 1)
job ← trajectory() %>%
task ← trajectory() %>%
   • • •
NEW_JOB \leftarrow function() rexp(1, 5)
NEW_TASK ← function() rexp(1, 1)
```



```
library(simmer): set.seed(1234)
RUNNING \leftarrow function() rexp(1, 1)
CHECK WORN ← function() runif(1) < 0.2
RETOOL \leftarrow function() rexp(1, 2)
RESET \leftarrow function() rexp(1, 3)
AWAY \leftarrow function() rexp(1, 1)
job ← trajectory() %>%
task ← trajectory() %>%
   ...
NEW JOB \leftarrow function() rexp(1, 5)
NEW TASK \leftarrow function() rexp(1, 1)
```

```
env 		 simmer("Job Shop") %>%
   add_resource("machine", 10) %>%
   add_resource("operative", 5) %>%
   add_generator("job", job, NEW_JOB) %>%
   add_generator("task", task, NEW_TASK) %>%
   run(until=1000)
   env

### simmer environment: Job Shop | now: 1000 | next: 1000.118
### { Monitor: in memory }
### { Resource: machine | monitored: TRUE | server status: 94
### { Resource: operative | monitored: TRUE | server status: 94
### { Source: job | monitored: 1 | n_generated: 4954 }
```

## { Source: task | monitored: 1 | n generated: 1041 }

## 2 operative 3.389082 0.3436009



```
library(simmer): set.seed(1234)
RUNNING \leftarrow function() rexp(1, 1)
CHECK WORN \leftarrow function() runif(1) < 0.2
RETOOL \leftarrow function() rexp(1, 2)
RESET \leftarrow function() rexp(1, 3)
AWAY \leftarrow function() rexp(1, 1)
job ← trajectory() %>%
task ← trajectory() %>%
   ...
NEW JOB \leftarrow function() rexp(1, 5)
NEW TASK \leftarrow function() rexp(1, 1)
```

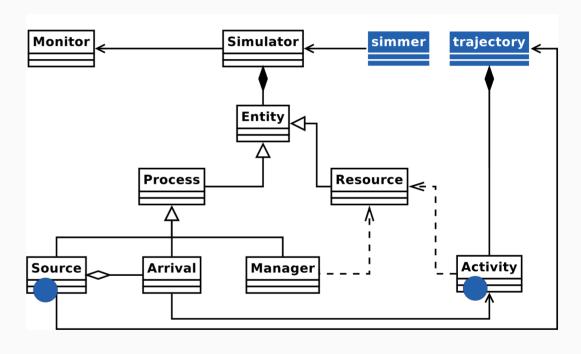
```
env ← simmer("Job Shop") %>%
  add resource("machine", 10) %>%
  add resource("operative", 5) %>%
  add generator("job", job, NEW JOB) %>%
  add generator("task", task, NEW TASK) %>%
  run(until=1000)
env
## simmer environment: Job Shop | now: 1000 | next: 1000.11
## { Monitor: in memory }
## { Resource: machine | monitored: TRUE | server status: 9
## { Resource: operative | monitored: TRUE | server status:
## { Source: job | monitored: 1 | n generated: 4954 }
## { Source: task | monitored: 1 | n generated: 1041 }
```

```
aggregate(cbind(server, queue) ~ resource, get_mon_resources(env), mean)
### resource server queue
### 1 machine 7.632616 0.5598666
```



# Terminology

#### Architecture

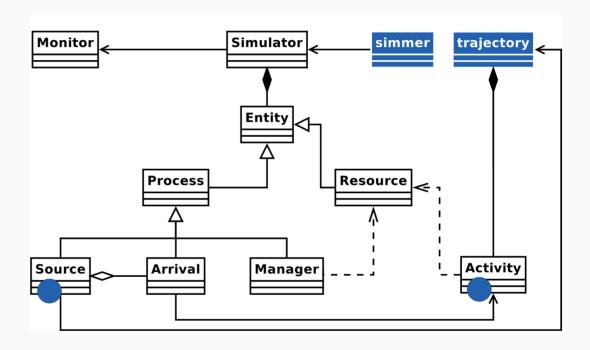




## Terminology

- **Resource**: server (configurable capacity) + priority queue (configurable size), supports preemption
- Manager: modifies resources at run time (schedule)

#### Architecture

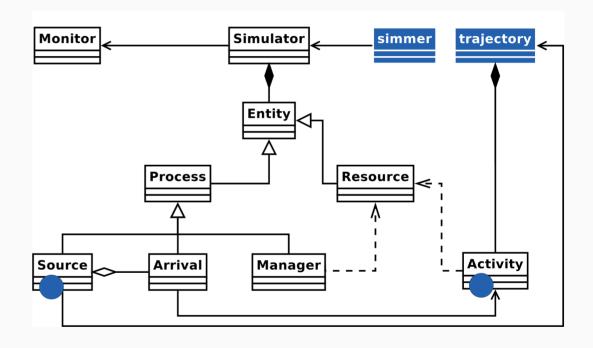




#### Terminology

- **Resource**: server (configurable capacity) + priority queue (configurable size), supports preemption
- Manager: modifies resources at run time (schedule)
- **Source**: creates new arrivals following some distribution of inter-arrival times
- Arrival: interacting processes, with attributes and prioritization values

#### Architecture

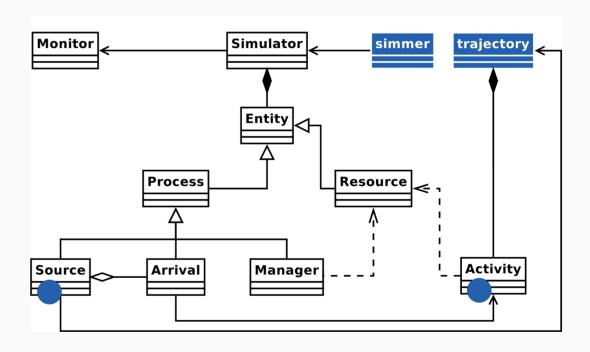




#### Terminology

- **Resource**: server (configurable capacity) + priority queue (configurable size), supports preemption
- Manager: modifies resources at run time (schedule)
- **Source**: creates new arrivals following some distribution of inter-arrival times
- **Arrival**: interacting processes, with attributes and prioritization values
- **Trajectory**: interlinkage of activities, a common path for arrivals of the same type
- Activity: unit of action, main building block

#### Architecture





#### **Trajectory**

Similar to dplyr for data manipulation. In the words of H. Wickham,

- by constraining your options, it simplifies how you can think about [something]
- Trajectories are recipes, lists of activities defining the life time of arrivals
- Activities are common functional DES blocks



#### **Trajectory**

Similar to dplyr for data manipulation. In the words of H. Wickham,

- by constraining your options, it simplifies how you can think about [something]
- Trajectories are recipes, lists of activities defining the life time of arrivals
- Activities are common functional DES blocks

Fixed vs. dynamic parameters:

```
traj0 ← trajectory() %>%
  log_("Entering the trajectory") %>%
  timeout(10) %>%
  log_("Leaving the trajectory")
```

```
traj1 ← trajectory() %>%
  log_(function() "Entering the trajectory") %>%
  timeout(function() 10) %>%
  log_(function() "Leaving the trajectory")
```



## **Trajectory**

#### **Activities**

- Spend time in the system
  - o timeout, timeout\_from\_attribute, timeout\_from\_global
- Modify arrival properties
  - o set\_attribute, set\_global
  - o set\_prioritization
- Interaction with resources
  - seize, release
  - o set\_capacity, set\_queue\_size
  - ∘ select, seize\_selected...
- Interaction with sources
  - activate, deactivate
  - o set\_trajectory, set\_source
- Loops
  - rollback

- Branching
  - branch
  - clone, synchronize
- Batching
  - batch, separate
- Inter-arrival communication
  - send, trap, untrap, wait
- Reneging
  - o leave
  - renege\_in, renege\_if,
    renege\_abort
  - handle\_unfinished
- Debugging
  - o log\_
  - o stop\_if



## **Trajectory**

#### **Activities**

- Spend time in the system
  - o timeout, timeout\_from\_attribute, timeout\_from\_global
- Modify arrival properties
  - o set\_attribute, set\_global
  - o set\_prioritization
- Interaction with resources
  - seize, release
  - o set\_capacity, set\_queue\_size
  - select, seize\_selected...
- Interaction with sources
  - activate, deactivate
  - o set\_trajectory, set\_source
- Loops
  - rollback

- Branching
  - branch
  - clone, synchronize
- Batching
  - batch, separate
- Inter-arrival communication
  - send, trap, untrap, wait
- Reneging
  - o leave
  - renege\_in, renege\_if,
    renege\_abort
  - handle\_unfinished
- Debugging
  - o log\_
  - o stop\_if

(Plus many getters to retrieve parameters at run time)



## Trajectory

Activities

- Create a simulator and attach a monitor
  - o simmer
  - monitor\_mem (default), monitor\_csv ... (extensible, see simmer.mon on GitHub)



## Trajectory

#### **Activities**

- Create a simulator and attach a monitor
  - o simmer
  - monitor\_mem (default), monitor\_csv ... (extensible, see simmer.mon on GitHub)
- Add sources of arrivals
  - add\_generator: based on a distribution function
  - add\_dataframe: based on a data frame (additional columns as attributes)



## Trajectory

#### **Activities**

- Create a simulator and attach a monitor
  - simmer
  - monitor\_mem (default), monitor\_csv ... (extensible, see simmer.mon on GitHub)
- Add sources of arrivals
  - add\_generator: based on a distribution function
  - add\_dataframe: based on a data frame (additional columns as attributes)
- Add resources
  - add\_resource: priority resource, with capacity and queue size; optional preemption



## Trajectory

#### **Activities**

- Create a simulator and attach a monitor
  - simmer
  - monitor\_mem (default), monitor\_csv ... (extensible, see simmer.mon on GitHub)
- Add sources of arrivals
  - add generator: based on a distribution function
  - add\_dataframe: based on a data frame (additional columns as attributes)
- Add resources
  - add\_resource: priority resource, with capacity and queue size; optional preemption
- Add global variables
  - add\_global



# Trajectory

#### **Activities**

- Create a simulator and attach a monitor
  - o simmer
  - monitor\_mem (default), monitor\_csv ... (extensible, see simmer.mon on GitHub)
- Add sources of arrivals
  - add generator: based on a distribution function
  - add\_dataframe: based on a data frame (additional columns as attributes)
- Add resources
  - add\_resource: priority resource, with capacity and queue size; optional preemption
- Add global variables
  - add\_global
- Run the simulation
  - o run, stepn



#### Trajectory

**Activities** 

Simulation environment

#### Monitoring

simmer automatically records every change in the state of the system. All these statistics can be retrieved after the simulation:

```
names( get_mon_arrivals(simmer(), per resource=FALSE) )
## [1] "name"
                       "start time"
                                       "end time"
                                                       "activity time"
## [5] "finished"
names( get mon arrivals(simmer(), per resource=TRUE) )
## [1] "name"
                       "start time"
                                       "end time"
                                                       "activity time"
## [5] "resource"
names( get mon attributes(simmer()) )
## [1] "time" "name" "key" "value"
names( get mon resources(simmer()) )
## [1] "resource" "time"
                                 "server"
                                              "queue"
                                                           "capacity"
## [6] "queue size" "system"
                                 "limit"
```



#### Queueing systems

Natural way to simulate CTMC and birth-death processes:

```
set.seed(1234)
lambda \leftarrow 2
mu \leftarrow 4
rho \leftarrow lambda/mu
```

```
mm1.traj ← trajectory() %>%
  seize("mm1.resource", 1) %>%
  timeout(function() rexp(1, mu)) %>%
  release("mm1.resource", 1)
```

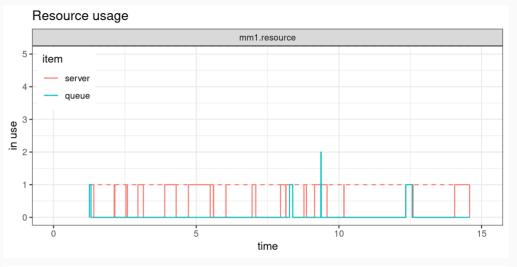


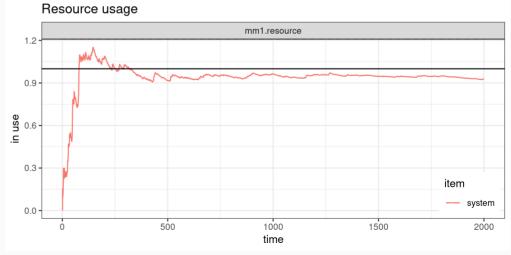
#### Queueing systems

Natural way to simulate CTMC and birth-death processes:

```
set.seed(1234)
lambda \leftarrow 2
mu \leftarrow 4
rho \leftarrow lambda/mu
```

```
mm1.traj ← trajectory() %>%
  seize("mm1.resource", 1) %>%
  timeout(function() rexp(1, mu)) %>%
  release("mm1.resource", 1)
```







#### Replication

Easy replication with standard R functions:



#### Replication

Easy replication with standard R functions:

#### **Parallelization**

Even easier parallelization of replicas:



## Replication

Easy replication with standard R functions:

#### Parallelization

Even easier parallelization of replicas:

```
head(get_mon_arrivals(mm1.envs), 3)

## name start_time end_time activity_time finished replication

## 1 arrival0 0.1455818 0.4810091 0.3354273 TRUE 1

## 2 arrival1 0.2380988 0.6834863 0.2024772 TRUE 1

## 3 arrival2 0.4526185 0.8724942 0.1890079 TRUE 1
```

# Modelling



#### Best practices

There are usually multiple valid ways of mapping the identified resources and processes into the simmer API

## Modelling



#### Best practices

There are usually multiple valid ways of mapping the identified resources and processes into the simmer API

#### Design pattern 1

```
beep ← trajectory() %>%
  log_("beeeep!")

env ← simmer() %>%
  add_generator("beep", beep, function() 1) %>%
  run(2.5)

## 1: beep0: beeeep!
## 2: beep1: beeeep!
```

## Modelling



#### Best practices

There are usually multiple valid ways of mapping the identified resources and processes into the simmer API

#### Design pattern 1

```
beep ← trajectory() %>%
  log_("beeeep!")

env ← simmer() %>%
  add_generator("beep", beep, function() 1) %>%
  run(2.5)

## 1: beep0: beeeep!
## 2: beep1: beeeep!
```

#### Design pattern 2

```
alarm 		 trajectory() %>%
   timeout(1) %>%
   log_("beeeep!") %>%
   rollback(2)

env 		 simmer() %>%
   add_generator("alarm", alarm, at(0)) %>%
   run(2.5)

## 1: alarm0: beeeep!
## 2: alarm0: beeeep!
```



Comparison with similar frameworks (out-of-date!):

- SimPy 3.0.9, Python 2.7
- SimJulia 0.3.14, Julia 0.5.1

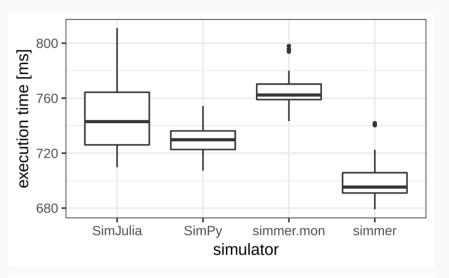
#### Heavy M/M/1, $ho \approx 0.9$ :

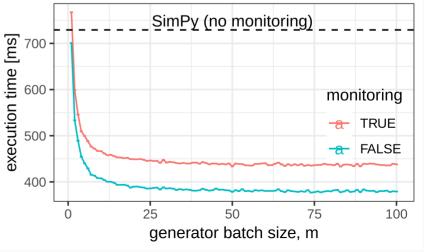


Comparison with similar frameworks (out-of-date!):

- SimPy 3.0.9, Python 2.7
- SimJulia 0.3.14, Julia 0.5.1

Heavy M/M/1,  $\rho \approx 0.9$ :







#### The cost of calling R from C++ revisited

Very simple deterministic test to study the impact:

```
test_simmer \leftarrow function(n, delay) {
  test \leftarrow trajectory() %>%
    timeout(delay)
  simmer() %>%
    add_generator("test", test, at(1:n)) %>%
    run() %>%
    get_mon_arrivals()
}

test_simmer(5, 1)[,1:5]
```



### The cost of calling R from C++ revisited

Very simple deterministic test to study the impact:

```
test_simmer \( \) function(n, delay) {
  test \( \) trajectory() %>%
    timeout(delay)
  simmer() %>%
    add_generator("test", test, at(1:n)) %>%
    run() %>%
    get_mon_arrivals()
}

test_simmer(5, 1)[,1:5]
```

##		name	start_time	end_time	activity_time	finished
##	1	test0	1	2	1	TRUE
##	2	test1	2	3	1	TRUE
##	3	test2	3	4	1	TRUE
##	4	test3	4	5	1	TRUE
##	5	test4	5	6	1	TRUE

Original benchmark in the JSS paper:

Expr	Min	Mean	Median	Max
test_simmer(n, 1)	429.8663	492.365	480.5408	599.3547
test_simmer(n, function() 1)	3067.9957	3176.963	3165.6859	3434.7979
test_R_for(n)	2053.0840	2176.164	2102.5848	2438.6836



### The cost of calling R from C++ revisited

Very simple deterministic test to study the impact:

```
test_simmer \( \) function(n, delay) {
  test \( \) trajectory() %>%
    timeout(delay)
  simmer() %>%
    add_generator("test", test, at(1:n)) %>%
    run() %>%
    get_mon_arrivals()
}

test_simmer(5, 1)[,1:5]
```

##	name	start_time	end_time	activity_time	finished
## 1	test0	1	2	1	TRUE
## 2	test1	2	3	1	TRUE
## 3	test2	3	4	1	TRUE
## 4	test3	4	5	1	TRUE
## 5	test4	5	6	1	TRUE

Original benchmark in the JSS paper:

Expr	Min	Mean	Median	Max
test_simmer(n, 1)	429.8663	492.365	480.5408	599.3547
test_simmer(n, function() 1)	3067.9957	3176.963	3165.6859	3434.7979
test_R_for(n)	2053.0840	2176.164	2102.5848	2438.6836

Update with -DRCPP\_USE\_UNWIND\_PROTECT:

Expr	Min	Mean	Median	Мах
test_simmer(n, 1)	467.8971	481.213	476.1667	521.4916
test_simmer(n, function() 1)	498.2631	583.777	561.6798	816.1343
test_R_for(n)	1158.9348	1201.460	1196.7223	1244.4041



- Generic yet powerful process-oriented Discrete-Event Simulation framework for R [1, 2]
- Combines a robust and fast C++ simulation core with a rich and flexible R API

<sup>[1]</sup> **Ucar I**, Smeets B, Azcorra A (2019). "simmer: Discrete-Event Simulation for R." *Journal of Statistical Software*, 90(2), 1-30. doi: 10.18637/jss.v090.i02.

<sup>[2]</sup> **Ucar I**, Hernández JA, Serrano P, Azcorra A (2018). "Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping." *IEEE Communications Magazine*, 56(11), 145-151. doi: 10.1109/MCOM.2018.1700960.



- Generic yet powerful process-oriented Discrete-Event Simulation framework for R [1, 2]
- Combines a robust and fast C++ simulation core with a rich and flexible R API
- Broad set of **activities**, the basic building block; extensible via custom routines
- Activities are chained into a **trajectory**, a common path for processes of the same type

<sup>[1]</sup> **Ucar I**, Smeets B, Azcorra A (2019). "simmer: Discrete-Event Simulation for R." *Journal of Statistical Software*, 90(2), 1-30. doi: 10.18637/jss.v090.i02.

<sup>[2]</sup> **Ucar I**, Hernández JA, Serrano P, Azcorra A (2018). "Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping." *IEEE Communications Magazine*, 56(11), 145-151. doi: 10.1109/MCOM.2018.1700960.



- Generic yet powerful process-oriented Discrete-Event Simulation framework for R [1, 2]
- Combines a robust and fast C++ simulation core with a rich and flexible R API
- Broad set of **activities**, the basic building block; extensible via custom routines
- Activities are chained into a trajectory, a common path for processes of the same type
- Automatic monitoring: focus on modelling

[1] **Ucar I**, Smeets B, Azcorra A (2019). "simmer: Discrete-Event Simulation for R." *Journal of Statistical Software*, 90(2), 1-30. doi: 10.18637/jss.v090.i02.

[2] **Ucar I**, Hernández JA, Serrano P, Azcorra A (2018). "Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping." *IEEE Communications Magazine*, 56(11), 145-151. doi: 10.1109/MCOM.2018.1700960.



- Generic yet powerful process-oriented Discrete-Event Simulation framework for R [1, 2]
- Combines a robust and fast C++ simulation core with a rich and flexible R API
- Broad set of **activities**, the basic building block; extensible via custom routines
- Activities are chained into a trajectory, a common path for processes of the same type
- Automatic monitoring: focus on modelling
- Integration: easy replication, parallelization, analysis...
- [1] **Ucar I**, Smeets B, Azcorra A (2019). "simmer: Discrete-Event Simulation for R." *Journal of Statistical Software*, 90(2), 1-30. doi: 10.18637/jss.v090.i02.
- [2] **Ucar I**, Hernández JA, Serrano P, Azcorra A (2018). "Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping." *IEEE Communications Magazine*, 56(11), 145-151. doi: 10.1109/MCOM.2018.1700960.

# Thanks, and happy simmer ing!

