---

title: 'DE-Sim: an object-oriented, discrete-event simulation tool for complex, data-driven modeling in Python'

tags:

- dynamical modeling

- simulation

- discrete-event simulation

- object-oriented simulation

- parallel discrete-event simulation

- biochemical modeling

- whole-cell modeling

- Python

authors:

- name: Arthur P. Goldberg

orcid: 0000-0003-2772-1484

affiliation: "1"

- name: Jonathan R. Karr

orcid: 0000-0002-2605-5080

affiliation: "1"

affiliations:

- name: Icahn Institute for Data Science and Genomic Technology, and Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

index: 1

date: 3 August 2020

bibliography: paper.bib

---

# Summary

Recent advances in data collection and storage have created unprecedented opportunities to gain insights into complex systems such as the biochemical networks that generate cellular behavior.

Understanding the behavior of such systems will likely require larger and more comprehensive dynamical models that are based on a combination of first principles and large datasets.

Large models often contain multiple types of components, and multiple types of interactions between component types.

The first principles typically use a mechanistic approximation of the interactions in a system, for example, that reactants in a chemical reaction may bond when they collide, to derive mathematics that must be parameterized by data.

One of the most promising methods for building and simulating large, data-driven dynamical models is to simulate them with discrete-event simulation (DES) [@fishman2013discrete].

For example, discrete-event simulations are frequently used to study the dynamics of biochemical networks, characterize the performance of computer networks, evaluate potential military strategies, and forecast epidemics [@banks2005discrete].

However, it is difficult to design, encode and simulate large, comprehensive models of complex systems as discrete-event simulations because the existing DES tools lack adequate expressive power.

To address this problem, we present DE-Sim, an [open-source](https://github.com/KarrLab/de\_sim), object-oriented (OO), Python-based DES tool that makes it easier to create and use dynamical models of complex systems.

Because DE-Sim encodes discrete-event simulations in OO Python programs, numerous types of components and complex interactions can be modeled by leveraging Python's powerful OO features.

In addition, because DE-Sim is implemented in Python, DE-Sim models can easily use Python's powerful data science tools such as pandas and SciPy to help build complex models, store and analyze data during simulations, and analyze simulation results.

We anticipate that DE-Sim will enable a new generation of models which capture systems with unprecedented breadth and depth.

An initial example is a multi-algorithmic simulator that we are developing on DE-Sim to simulate whole-cell models that predict phenotype from genotype by capturing all of the biochemical activity in a biological cell.

# The need for tools that help researchers build and simulate complex models

Many scientific fields can now collect detailed data about the components of complex systems and their interactions. For example, the revolution in deep sequencing has dramatically increased the availability of molecular data.

When the measurements of a complex system generate a large, heterogeneous dataset, models of the system that use the measurements are invariably complex and large because the dataset contains measurements of many types of components, and many instances of each component.

For example, measurements of the biochemistry in a cell collect the properties of many types of molecules, many types of processes that transform or translocate molecules, and many instances of each type.

These data mirror the intrinsic complexity of cellular biology and lead to complex and large models of cells.

DE-Sim simplifies the construction and simulation of \*discrete-event models\* which represent the dynamics of a complex system as a sequence of instantaneous events.

Dynamical models constructed with DE-Sim can leverage Python's extensive suite of high-quality data science tools to easily manage and integrate large, heterogeneous, multidimensional data into models. For example, tools such as NumPy [@oliphant2006guide], pandas [@mckinney2010data], SciPy [@virtanen2020scipy], and SQLAlchemy [@bayer2020sqlalchemy] can be used by DE-Sim models to store and integrate model inputs, simplify analyses and data storage during simulation, and organize and save predictions for downstream analysis.

Discrete-event models of complex systems are challenging to construct. DE-Sim addresses this challenge by structuring discrete-event models as object-oriented programs. This approach, known as \*object-oriented discrete-event simulation\* (OO DES), implements the component types in a model as simulation object classes and implements their interaction events as \*event messages\* that schedule the simulation objects to execute events at specified times [@zeigler1987hierarchical]. With DE-Sim, users define classes of simulation object by subclassing DE-Sim’s simulation object class and specifying their behavior. Complex systems that contain multiple types of components can be easily modeled in DE-Sim by creating multiple classes of simulation objects. Users can model arbitrarily many instances of a type of component by creating multiple instances of the corresponding simulation object class. DE-Sim simulation object classes can exploit all the features of Python objects. For example, hierarchical relationships among the components in a system being modeled can be mirrored by subclass relationships among the simulation object classes that represent the components in a DE-Sim model. Thus, a model of a biochemical system that represents macromolecule components such as DNA and RNA could define a macromolecule class, and then define DNA and RNA subclasses of the macromolecule class.

In our experience, DE-Sim’s use of OO programming to construct discrete-event models simplifies and accelerates the design and development of models.

DE-Sim is designed for scientists and engineers who want to build and use quantitative, dynamical models to understand the properties of complex, discrete-time systems. DE-Sim's features address the needs of this audience: it uses Python, one of the most popular languages; it is open-source; it is easy to learn because it provides several tutorials, examples, and documentation; and it is thoroughly tested.

design,

Modelers who use DE-Sim write their models as object-oriented Python programs.simplifies the ion of

comprehensive Researchers who use DE-Sim can employ Python’s powerful, comprehensiveNumPy, pandas, SciPy, and SQLAlchemy easily and complex analyses into their

--

a

simulations can ily A model simply specifies

design,

DE-Sim automatically generates these diagrams from log files.

and output

, as documented in [a Jupyter notebook](<https://sandbox.karrlab.org/notebooks/de_sim/2.%20DE-Sim%20advanced%2C%20simultaneous%20events.ipynb>)

![\*\*DE-Sim can generate space-time visualizations of simulation trajectories.\*\*

This figure illustrates a space-time visualization of all of the events and messages in a simulation of the parallel hold (PHOLD) DES benchmark model [@fujimoto1990performance] with three simulation objects. The timeline (grey line) for each object shows its events (grey dots). Each event in PHOLD schedules one other event.

Each event \*e\* is illustrated by an event message (arrow) that the simulation sends from the event that schedules \*e\* to \*e\* itself.

The curved blue arrows indicate events scheduled by a simulation object for itself, while the straight purple arrows indicate event messages sent to other simulation objects. The programs for the PHOLD model and for visualizing the trajectory of any simulation are available in the DE-Sim Git repository.

\label{fig:phold\_space\_time\_plot}](phold\_space\_time\_plot.pdf)

The following sections describe how DE-Sim provides these features and how they help researchers build models of complex systems.

# Comparison of DE-Sim with existing discrete-event simulation tools

Multiple DES tools already exist.

\autoref{fig:comparison} lists some of the more popular tools for modeling nature and engineered systems.

All of these tools provide a programming environment for developing DES models, a simulator that simulates models, and methods for recording simulation predictions.

![\*\*Comparison of DE-Sim with important existing DES tools.\*\*

DE-Sim is the only open-source, object-oriented, discrete-event simulation tool based on Python.

This combination of features makes it uniquely suitable for creating discrete-event models that can be used to study complex systems.

\label{fig:comparison}](comparison.pdf)

SimEvents [@clune2006discrete] is a MATLAB-based tool that adds DES capabilities to the Simulink environment for modeling dynamical systems.

Its graphical interface enables visual definitions of queueing models of networks and processes.

MATLAB code can also specify a SimEvents model's behavior.

SimEvents obtains its OO modeling functionality from the OO features of MATLAB.

However, MATLAB lacks Python's extensive library of scientific data analysis packages and broad adoption by researchers, which limits SimEvents' suitability for analyzing complex systems.

SimPy [@matloff2008introduction] is an open-source, general-purpose DES tool which can leverage Python’s data-science packages and popularity because it’s written in Python.

However, SimPy models are defined with functions that do not have the flexibility and power of the objects that DE-Sim employs to define models of complex systems.

In addition, DE-Sim supports a uniform approach for scheduling events, whereas SimPy models that contain multiple processes that schedule events for themselves and each other must use two approaches.

A DE-Sim object always schedules an event by sending an event message to the object that will execute the event.

However, a SimPy process schedule events for itself by using a timeout call and Python's `yield` function, but schedules events for other processes by raising an interrupt exception.

SIMSCRIPT III [@rice2005simscript] is a commercial DES tool that describes models in SIMSCRIPT III, a proprietary, object-oriented language.

It’s well suited for modeling decision support systems in domains such as war-gaming, communications network analysis, and transportation and manufacturing.

But DE-Sim is more appropriate for modeling complex systems because it can access Python packages.

SIMUL8 [@concannon2003dynamic] specializes in modeling business processes.

Like other DES tools that use proprietary languages it is not well designed for building scientific models of complex systems.

Lastly, SystemC [@mueller2001simulation] is an OO DES tool based on C++ that is commonly used to model the design and performance of digital systems.

DE-Sim is more accessible to scientific researchers than SystemC, because DE-Sim builds upon Python whereas SystemC uses C++, a lower-level language.

An important benefit of OO DES tools like DE-Sim and the other OO DES tools in \autoref{fig:comparison} is that individual simulation runs may be sped up by parallel execution on multiple cores.

More precisely, the simulation of an OO DES model composed of simulation objects that only interact with each other via event messages and do not access shared memory might be sped up by distributing its objects across multiple cores and executing them in parallel.

This simulation would need to be synchronized by a parallel DES simulator, such as Time Warp [@Jefferson1985, @carothers2000ross].

SystemC has been parallelized [@schumacher2010parsc], for example.

Parallel DES simulations can achieve substantial speedup, as was demonstrated by running the PHOLD benchmark on nearly 2 million cores [@Barnes2013].

In summary, the primary advantages of DE-Sim are that i

# Tutorial: Building and simulating models with DE-Sim

A simple DE-Sim model can be defined in three steps: define an event message class; define a simulation object class; and build and run a simulation.

We illustrate this process with a model of a random walk on the integer number line (this example is available in a [Jupyter notebook](https://sandbox.karrlab.org/notebooks/de\_sim/1.%20DE-Sim%20tutorial.ipynb)).

1: Create an event message class by subclassing `EventMessage`.

Each DE-Sim event contains an event message that provides data to the simulation object which executes the event.

The random walk model sends event messages that contain the value of a random step.

```python

import de\_sim

class RandomStepMessage(de\_sim.EventMessage):

"An event message class that stores the value of a random walk step"

attributes = ['step']

```

The attribute `attributes` is a special attribute of a `EventMessage` that specifies the names of an event message class' attributes.

These names must be valid Python identifiers.

`attributes` is optional.

An event message class must be documented by a docstring, as illustrated.

2: Define a simulation object class by subclassing `SimulationObject`.

Simulation objects are like threads, in that a simulation's scheduler decides when to execute them, and their execution is suspended when they have no work to do.

But DES simulation objects and threads are scheduled by different algorithms.

Whereas a thread can be scheduled whenever it has work to do,

a DES scheduler schedules simulation objects to ensure that events occur in simulation time order, as summarized by the fundamental invariant of discrete-event simulation:

1. All events in a simulation are executed in non-decreasing time order.

By guaranteeing this behavior, the DE-Sim scheduler ensures that causality relationships between events are respected.

(The invariant says \*non-decreasing\* instead of \*increasing\* time order because events can occur simultaneously, as discussed above.)

This invariant has two consequences:

1. All synchronization between simulation objects is controlled by the simulation times of events.

2. Each simulation object executes its events in non-decreasing time order.

The Python classes that generate and handle simulation events are simulation object classes, which are defined as subclasses of `SimulationObject`.

DE-Sim provides a custom class creation method for `SimulationObject` that gives special meaning to certain methods and attributes.

Below, we define a simulation object class that models a random walk, and illustrates all key features of `SimulationObject`.

To add variety to its temporal behavior we modify the traditional random walk by randomly selecting the time delay between steps.

```python

import random

class RandomWalkSimulationObject(de\_sim.SimulationObject):

" A 1D random walk model, with random delays between steps "

def \_\_init\_\_(self, name):

super().\_\_init\_\_(name)

def init\_before\_run(self):

" Initialize before a simulation run; called by the simulator "

self.position = 0

self.history = {'times': [0],

'positions': [0]}

self.schedule\_next\_step()

def schedule\_next\_step(self):

" Schedule the next event, which is a step "

# A step moves -1 or +1 with equal probability

step\_value = random.choice([-1, +1])

# The time between steps is 1 or 2, with equal probability

delay = random.choice([1, 2])

# Schedule an event `delay` in the future for this object

# The event contains a `RandomStepMessage` with `step=step\_value`

self.send\_event(delay, self, RandomStepMessage(step\_value))

def handle\_step\_event(self, event):

" Handle a step event "

# Update the position and history

step = event.message.step

self.position += step

self.history['times'].append(self.time)

self.history['positions'].append(self.position)

self.schedule\_next\_step()

# `event\_handlers` contains pairs that map each event message class

# received by this simulation object to the method that handles

# the event message class

event\_handlers = [(RandomStepMessage, handle\_step\_event)]

# messages\_sent registers all message types sent by this object

messages\_sent = [RandomStepMessage]

```

DE-Sim simulation objects (subclasses of `SimulationObject`) use these special methods and attributes:

\* Special `SimulationObject` methods:

1. \*\*`init\_before\_run`\*\* (optional): immediately before a simulation run, after all simulation objects have been added to a simulator, the simulator calls each simulation object’s `init\_before\_run` method. In this method simulation objects can send initial events and perform other initializations. For example, in `RandomWalkSimulationObject`, `init\_before\_run` schedules the object's first event and initializes the object's position and history attributes.

2. \*\*`send\_event`\*\*: `send\_event(delay, receiving\_object, event\_message)` schedules an event to occur `delay` time units in the future at simulation object `receiving\_object`. `event\_message` must be an `EventMessage` instance. An event can be scheduled for any simulation object in a simulation, including the object scheduling the event, as `RandomWalkSimulationObject` does.

The event will be executed at its scheduled simulation time by an event handler in the simulation object `receiving\_object`.

The handler defines an `event` parameter.

Its value will be the scheduled event, which contains `event\_message` in its `message` attribute.

Object-oriented DES terminology often describes the event message as being sent by the sending object at the message's send time (the simulation time when the sending object schedules the event) and being received by the receiving object at the event's receive time (the simulation time when the event is executed). An event message can thus be viewed as a directed edge in simulation space-time from the pair (sending object, send time) to (receiving object, receive time), as illustrated by \autoref{fig:phold\_space\_time\_plot}.

3. \*\*event handlers\*\*: an event handler is a method that handles and executes a simulation event. Event handlers have the signature `event\_handler(self, event)`, where `self` is the simulation object that handles (receives) the event, and `event` is a simulation event. A subclass of `SimulationObject` must define at least one event handler, as illustrated by `handle\_step\_event` in the example above.

\* Special `SimulationObject` attributes:

1. \*\*`event\_handlers`\*\*: a simulation object can receive arbitrarily many types of event messages, and implement arbitrarily many event handlers. The attribute `event\_handlers` must contain an iterator over pairs that map each event message class received by a `SimulationObject` subclass to the event handler which handles the event message class. In the example above, `event\_handlers` associates `RandomStepMessage` event messages with the `handle\_step\_event` event handler.

2. \*\*`messages\_sent`\*\*: the types of messages sent by a subclass of `SimulationObject` must be listed in `messages\_sent`. It ensures that a simulation object doesn't send messages of the wrong `EventMessage` class.

3. \*\*`time`\*\*: `time` is a read-only attribute that always equals the current simulation time in every simulation object. For example, a `RandomWalkSimulationObject` saves the value of `time` when recording its history.

![\*\*Trajectory of a simulation of a model of a random walk on the integer number line.\*\*

The random walk model starts at position 0 and moves +1 or -1 with equal probability at each step.

Steps take place every 1 or 2 time units, also with equal probability.

This trajectory illustrates two key characteristics of discrete-event models. First, the state, in this case the position, changes at discrete times.

Second, since the state does not change between instantaneous events, the trajectory of any state variable is a step function.

The source code for this model is available in the DE-Sim Git repository.

\label{fig:random\_walk\_trajectory}](random\_walk\_trajectory.png)

3: Execute a simulation by creating and initializing a `Simulator`, and running the simulation.

The `Simulator` class simulates models.

Its `add\_object` method adds a simulation object to the simulator.

The `initialize` method, which calls each simulation object’s `init\_before\_run` method, must be called before a simulation starts.

At least one simulation object in a simulation must schedule an initial event--otherwise the simulation cannot start.

More generally, a simulation with no events to execute will terminate.

Finally, `run` simulates a model. It takes the maximum time of a simulation run. `run` also takes several optional configuration arguments, as described in the DE-Sim [API documentation](https://docs.karrlab.org/de\_sim/master/source/de\_sim.html#de\_sim.simulator.Simulator.simulate).

```python

# Create a simulator

simulator = de\_sim.Simulator()

# Create a random walk simulation object and add it to the simulation

random\_walk\_sim\_obj = RandomWalkSimulationObject('rand\_walk')

simulator.add\_object(random\_walk\_sim\_obj)

# Initialize the simulation

# This executes `init\_before\_run` in `random\_walk\_sim\_obj`

simulator.initialize()

# Run the simulation until time 10

max\_time = 10

simulator.run(max\_time)

# Plot the random walk

import matplotlib.pyplot as plt

plt.step(random\_walk\_sim\_obj.history['times'],

random\_walk\_sim\_obj.history['positions'])

plt.xlabel('time')

plt.ylabel('position')

plt.show()

```

This example runs a simulation for `max\_time` time units, and plots the random walk’s trajectory (\autoref{fig:random\_walk\_trajectory}).

This tutorial and additional examples are available in a [Jupyter notebook](https://sandbox.karrlab.org/notebooks/de\_sim/1.%20DE-Sim%20tutorial.ipynb).

# Performance of DE-Sim

\autoref{fig:performance} shows the performance of DE-Sim simulating a model of a cyclic messaging network over range of network sizes.

The code for this performance test is available in the DE-Sim Git repository and a [Jupyter notebook that runs the test](https://sandbox.karrlab.org/notebooks/de\_sim/4.%20DE-Sim%20performance%20test.ipynb).

![\*\*Performance of DE-Sim simulating a model of a cyclic messaging network over a range of network sizes.\*\* Each statistic represents the average of three simulation runs in a Docker container on a 2.9 GHz Intel Core i5 processor.

The cyclic messaging network model consists of a ring of simulation objects. Each simulation object executes an event at every time unit and schedules an event for the next object in the ring 1 time unit in the future.

The number of simulation objects in the ring is given by \*\*Nodes\*\*.

Each simulation run executes for 100 time units.

\label{fig:performance}](performance.pdf)

# Case study: a multi-algorithmic simulation tool for whole-cell modeling implemented using DE-Sim

We have used DE-Sim to develop WC-Sim [@goldberg2020wc\_sim], a multi-algorithmic simulator for comprehensive whole-cell models of the biochemical dynamics inside biological cells [@karr2015principles; @goldberg2018emerging; @karr2012whole].

Whole-cell models which predict phenotype from genotype by representing all of the biochemical activity in a cell have great potential to help scientists elucidate the basis of cellular behavior, help bioengineers rationally design biosensors and biomachines, and help physicians personalize medicine.

Due to the diverse timescales of the reactions inside cells, one promising way to simulate whole-cell models is to simulate each reaction with an appropriate algorithm for its timescale. For example, slow biochemical reactions, such as transcription, can be simulated with the Stochastic Simulation Algorithm (SSA, @gillespie1977exact). Faster processes, such as signal transduction, can be simulated with ordinary differential equations (ODEs). Metabolism, another fast process, can be simulated with flux-balance analysis (FBA, @orth2010flux). Simulating entire cells requires co-simulating SSA, ODE and FBA. However, tools for co-simulating these algorithms did not exist before we created WC-Sim.

To accelerate whole-cell modeling, we have created WC-Sim, a tool for simulating multi-algorithmic whole-cell models described in the WC-Lang language [@karr2020wc\_lang].

We implemented WC-Sim by using DE-Sim to construct separate simulation object classes for SSA, ODE, and FBA.

A cell’s state is given by the populations of its molecular species, which are stored in an object that is shared by all simulation objects.

DE-Sim event messages schedule the activities of each simulation object, while the exact simulation time of events is used to coordinate the objects’ shared access to the cell’s state.

DE-Sim’s object-oriented modeling functionality made it easy to separately develop SSA, ODE, and FBA simulation object classes and compose them into a multi-algorithmic simulator.

DE-Sim’s discrete-event framework provided the control needed to synchronize the interactions between these classes.

In addition, accessing Python's data-science tools reduced the effort required to build WC-Sim.

It uses NumPy's random number generator for stochastic simulation, and its arrays to store and compare molecule counts; pandas DataFrames store simulation predictions and transfer them to and from files; directed graphs and the DFS algorithm in networkx analyze reaction network dependencies [@hagberg2008exploring]; the ODE solver in scikits.ODES determines the rate of change of species populations modeled by ODEs [@malengier2018odes]; and matplotlib visualizes simulation predictions [@Hunter:2007].

We anticipate that WC-Sim will enable researchers to conduct unprecedented simulation studies of cellular biochemistry.

# Conclusions

[DE-Sim](https://pypi.org/project/de-sim/) is an open-source, object-oriented, discrete-event simulation tool implemented in Python.

We encourage researchers who use discrete-event models to understand the dynamics of complex systems to try DE-Sim because it combines two features that are not available together in other tools.

First, discrete-event models defined in DE-Sim are constructed from multiple, interacting simulation objects.

This gives modelers the ability to intuitively model a complex system that contains multiple types of interacting components by encoding each component type as a different DE-Sim simulation object class.

Second, to conveniently store, manage and analyze the large, heterogeneous data needed by models of complex systems, modelers using DE-Sim can access Python’s extensive library of data science packages.

We anticipate that DE-Sim will enable the construction and use of ambitiously detailed models of complex systems, and that simulation studies conducted with these models contribute to important engineering innovations and scientific discoveries.

# Availability of DE-Sim

DE-Sim is freely and openly available under the MIT license at the locations below.

\* Python package: [PyPI: de-sim](https://pypi.org/project/de-sim/)

\* Docker image: [DockerHub: karrlab/de\_sim](https://hub.docker.com/r/karrlab/de\_sim)

\* Tutorials: Jupyter notebooks at [https://sandbox.karrlab.org](https://sandbox.karrlab.org/tree/de\_sim)

\* Installation instructions and documentation of DE-Sim’s API: [docs.karrlab.org](https://docs.karrlab.org/de\_sim/)

\* Issue tracker: [GitHub: KarrLab/de\_sim](https://github.com/KarrLab/de\_sim/issues/)

\* Source code: [GitHub: KarrLab/de\_sim](https://github.com/KarrLab/de\_sim/)

\* Guide to contributing to DE-Sim and code of conduct for developers: [GitHub: KarrLab/de\_sim](https://github.com/KarrLab/de\_sim/)

\* Continuous integration: [CircleCI: gh/KarrLab/de\_sim](http://circleci.com/gh/KarrLab/de\_sim/)

DE-Sim requires [Python](https://www.python.org/) 3.7 or higher and [pip](https://pip.pypa.io/).

This article discusses version 0.0.8 of DE-Sim.

# Acknowledgements

We thank Yin Hoon Chew for her helpful feedback on this paper. This worked was supported by the National Science Foundation [award 1649014 to J.R.K.], the National

Institutes of Health [award R35GM119771 to J.R.K], and the Icahn Institute for Data Science and Genomic Technology.

# References