System modelling and simulation

Jakub Bujas
Dawid Dworak

https://github.com/ddworak/formin

formin

Project vision

Overview

Foraminifera are single-celled eukaryotes that occupy marine benthic and pelagic zones.They move to the areas rich of algaes simultaneously avoiding grouping. Their aggregations can be an indicator of oil wells underneath the seabed. Therefore accurate foraminifera habitat behaviour predictions are highly-sought and prove a challenge for the scientists. We intend to conduct a large-scale cluster simulation following realistic rules known from the nature to provide a scalable environment for such analysis.

Goals

1. **Functional requirements:**
    a. Based on predator-prey algorithm,
    b. Using a mechanism of smell spreading described in ‘Dynamic Assignment of Tasks to Mobile Robots in Presence of Obstacles’ article by Tomasz Sośnicki, Wojciech Turek, Krzysztof Cetnarowicz, Małgorzata Żabińska,
    c. The initial state is generated randomly at the beginning of program,

d. Each Akka Actor taking care of computation over multiple cells on the grid,

e. Some cells on the map disabled for algaes and foraminifera populations.

2. **Non-functional requirements:**

a. Using Akka Cluster,

b. Computation scaled out on a large-scaled cluster,

c. The most interesting cases - population above 10^9,

d. Using GitHub platform.

Specifications

The simulation of foraminiferal habitats presents much higher complexity than a Cellular Automata (e.g. Conway's Game of Life), since simple rules are not enough to model the population behaviour. The two-dimensional grid is only a basis for spatial distribution of the simulated entities. The proposed algorithm involves a signal spreading mechanism and obstacles imitating a real environment.

For full utilization of the target platform - a large-scale cluster - the application will be distributed and easily scalable, as well as asynchronous as much as it is possible. The concurrency will be built upon the Akka framework. It provides an Actor Model implementation including remoting and clustering modules out of the box. Each actor will be responsible for computations related to a part of the grid stored within its internal state. The communication and conflict resolution over such a large-scale mutable in-memory dataset is non-trivial, nevertheless the chosen model is well-suited to such task.

Milestones

1. Single node habitat simulation.

   In order to better understand the project domain and establish a basis for solving further goals, we will start by reducing the problem to a single node simulation on a single grid. This approach will ensure the focus on correct implementation of proposed algorithm, which is a core element of the end-product.

2. Foraminifera parallelized.

   The second step on the way to larger habitat simulations is the parallelization of the implemented algorithm using multiple local actors handling different parts of the grid. This involves creating a communication protocol for the workers, clever time-stepping, monitoring data collection and possibly conflict resolution, hence conflicts are an intrinsic part of an asynchronous approach to the problem.

3. Foraminifera clusterized.[1]

   Lastly, for full utilization of the available hardware, the computation will be scaled out on a large-scale cluster using the Akka Cluster. Large-scale cluster simulation of the habitat might provide unknown thus far insights into the behaviour of real-life sized foraminifera populations.

[1] Hopefully.