Developing Domain-Specific Languages for Ocean Modeling

EMLS’21 – Project OceanDSL

Reiner Jung, Sven Gundlach, Serafim Simonov, Wilhelm Hasselbring

Kiel-University

23rd February 2021
Introduction

- Project Goal: DSLs to support Ocean Modeling
- Domain analysis: Thematic Analysis [Braun and Clarke 2006]
- Example: Configuration and Parameterization DSL
What is Ocean Modeling?

- Sediment Management
- External Forcings
- Biogeochemical Model
- Transport Model
- Ocean Model
- Coupler
- Atmosphere Model

Jung, Gundlach, Simonov, Hasselbring
Different Scales in Modeling

~100 m

~1 km

~10 km

~100 km

~1,000 km

~10,000 km

MITgcm Project 2020
Modeling Example: Surface Water and Ocean Topography

Fluxes of Heat, Carbon and Oxygen at SWOT Scales

[Smith and Abernathey 2017]
Simplified Ocean Modeling Process

Modeling Trigger

- Mathematical Modeling
- Deployment & Execution
- Analysis

Does the model work as expected?

- Model Coding
- Test Process

- yes
- no, recheck model setup and code

Should the resulting model become part of the upstream model project?

- yes
- no

Integrate into Upstream Model

Gatekeeper

Scientific Modeler

Model Developer

Jung, Gundlach, Simonov, Hasselbring
Domain Characteristics

Models

- Long-living systems
- Implemented in Fortran 77, 90, C++ and Python
- Feature management by #ifdef

Editors

- Vi, Vim, Emacs and Xcode
- In general no IDEs
  (except Emacs, and PyCharm in rare cases)

Build system

- make, cmake, shell scripts, perl
Typical Aspects & Views in Ocean Modeling

DSLs in Ocean Modeling
- External DSLs, e.g., Dusk/Dawn MeteoSwiss [MeteoSwiss 2020]
- Embedded DSLs, e.g., Psyclone [Adams et al. 2019]

Views & Aspects
- Transport Model Specification
- Bio-geo-chemical Modeling
- Configuration and Parameterization
- Deployment
include size

barotropic_gyre : mitgcm
Global Parameters and Features

features ALLOW_FRICTIONHEATING

parameters

PARM01:
  viscAh = 4.E2
  f0 = 1.E-4
  beta = 1.E-11
  rhoConst = 1000.0
  gBaro = 9.81
module cost:
  features ALLOW_EGM96_ERROR_COV

cost_nml:
  mult_atl = 0.
  mult_test = 0.

diagnostics:
  diagMdsDir = "some-dir"
  format = netcdf
  diagSt_regMaskFile = "regMask_lat24.bin"
  set_regMask(1:3) = [ 1, 1, 1 ]
  val_regMask(1:3) = [ 1., 2., 3. ]

"first-out.log":
  logmode = snap
  frequency = 10
  missing_value = 5.0
  fields(1:2) = [ SDIAG1, SDIAG2 ]
  levels(1:2) = [ 1, 2 ]
Summary

- Introduced the domain of **ocean modeling**
  - Main process
  - Domain properties
- Presented the **configuration and parameterization DSL**
Questions

Language related aspects
- Which syntactical style should we use?
  - YAML, CPP, C, Python
  - Familiarity might be relevant
  - Structures must be as clear and simple
- How could we address modularization of the configuration?
  - Include, override, interfaces, immutuals

Technical and social aspects
- How to introduce DSLs into the domain?
  - Are there methods from other domains which we could use here?
  - What are usual methods and arguments to hinder the introduction of DSL?
  - How can we address them?
- How to organize maintenance after the project ends?
  - How to motivate institutions to commit themselves?
  - How to minimize maintenance?


MeteoSwiss (2020). Dawn – Compiler toolchain to enable generation of high-level DSLs for geophysical fluid dynamics models. URL: https://github.com/MeteoSwiss-APN/dawn.
