

Euphemia Reproducibility

MESC
Brussels, 9 December 2016

Euphemia Reproducibility

Reproducibility vs Optimality

Based on the MPs' previous feedback, PCR has focused on reducing PRBs and increasing welfare. However, some of the algorithmic methods now used or planned to be used inhibit the ability to reproduce results.

- Euphemia 9.3 (go-live 29 October 2015) first introduced the PRB reinsertion method aimed at reducing PRBs while not harming economic welfare. Compared to E9.2 (no PRB reinsertion) on 2014 MRC data:
 - Average number of PRBs (across MRC) decreased by 9.6 (~30% reduction)
 - Welfare improved on average by 1700€ per day
- PRB reinsertion introduces new internal processing time deadlines that make it very hard to reproduce results based simply on the same inputs
- PCR has communicated the information on PRB reinsertion transparently to stakeholders, and will continue to do so in the subsequent MESC meetings

Euphemia Reproducibility

Reproducibility: old definition

The capability of the algorithm to reproduce the same results upon request. On the **same machine**, two subsequent runs with the **same input data** should find the **same solutions**, meaning that the intermediate/final solutions found at iteration 'X' are the same. In other words, when the stopping criterion is the number of investigated solutions, a reproducible algorithm can guarantee to obtain the same final result when run on the same machine. However, when the stopping criterion is a time limit, a faster computer will allow the algorithm to investigate more solutions than a slower one. In this case, the reproducibility consists in investigating on the faster computer at least the same set of solutions as the ones investigated on the slower computer.

*Definition from Euphemia public documentation

Euphemia Reproducibility

Sources of non-reproducibility

| Source | Simulation Facility impacted | Production impacted |
|---|--|--|
| Different inputs (even if only ordering is different) | Mitigated by securing identical ordering | |
| Different machines | No impact (since the same machine used) | Reproduction only possible on coordinator machine |
| Different Euphemia version | Not applicable provided same versions is used for reproduction | |
| Time limit and speed of the machine | Impact in case the machine is slowed down due to other process running in parallel | Mitigated by having dedicated systems for Euphemia and the implementation of a machine certification procedure |
| Local search (or PRB reinsertion) | Full reproducibility cannot be guaranteed | Full reproducibility cannot be guaranteed |
| E10 and parallel processing (not yet in production) | Parallelization will bring further reproducibility challenges | |

Euphemia Reproducibility

Conclusions and next steps

- With PRB reinsertion Euphemia is no longer fully compliant with old reproducibility definition;
 - More apparent on non-production environments (such as the simulation facility)
 - On production environments reproducibility is not guaranteed, but the impact seems more limited (based on a limited sample size)
- With Euphemia 10 and parallelization we depart further from reproducibility
- PCR parties believe the appropriate balance has been made when pursuing quality (welfare and PRBs) at the expense of full reproducibility
- Next steps:
 - Public Euphemia description to be updated to include PRB reinsertion method and impact on reproducibility
 - Proposal to meet CACM requirement that price coupling results are *repeatable* (Article 38(1)(e)) as follows:
 - a) During the course of the calculation process, information relevant to repeat the resulting solution will be logged;
 - b) The algorithm can repeat the calculation using the same version of the algorithm on the same machine with the same input data and the information logged under a)