Madgraph4gpu cudacpp WIP: copyright and licensing

Andrea Valassi (CERN)

Madgraph on GPU development meeting, 25th April 2023
https://indico.cern.ch/event/1275610

DISCLAIMER: I am not a lawyer!
My proposal is based on my <u>limited</u> understanding of these issues...

(I will not repeat this disclaimer explicitly on each slide, but you should assume that I have done that implicitly!)

MG5aMC CUDACPP copyright and licensing

- Why now?
  - The CERN IT Innovation Review Board in March recommended that we mark/protect "the intellectual property of CERN" in the developments involving CERN personnel
  - We will soon prepare public (pre-)release(s) of the software, we should solve this before
  - We may want to repackage the software repository (and/or create new ones)

- **I prepared a draft PR: https://github.com/madgraph5/madgraph4gpu/pull/643**
  - The proposal covers two related pieces of software for CUDACPP developments
    - the CUDACPP code-generating plugin (Python plugin and CUDA/C++ templates)
    - the CUDACPP-generated code
  - The proposal covers three aspects of the problem for CUDACPP developments
    - Copyright – I propose "CERN and UCLouvain" (rather than the individual authors)
    - License – I propose LGPL v3 (the GNU Lesser General Public License version 3)
    - Authors – I propose 5 authors and a list of collaborators
  - The proposal covers the issues above at two levels:
    - With "central" files COPYRIGHT, COPYING, COPYING.LESSER, AUTHORS
    - With copyright/license/author headers in each source file

Where do we start? the MG5aMC license

- See https://github.com/mg5amcnlo/mg5amcnlo/blob/main/madgraph/LICENSE :

```
The MadGraph5_aMC@NLO software comes with the following Open Source License,
adapted from the University of Illinois/NCSA license.

Copyright (c) 2009, 2013, the MadTeam.  All rights reserved.

Developed by the "MadTeam", i.e.

Johan ALWALL
Rikkert FREDERIX
Stefano FRIXIONE
Michel HERQUET
Valentin HIRSCHI
Fabio MALTONI
Olivier MATTELAER
Tim STELZER
Paolo TORRIELLI
Marco ZARO

Permission is hereby granted, free of charge, to any person obtaining a copy
of this software and associated documentation files (the "Software"), to
deal with the Software without restriction, including without limitation the
rights to use, copy, modify, merge, publish, distribute, sublicense, and/or
sell copies of the Software, and to permit persons to whom the Software is
furnished to do so, subject to the following conditions:
  1. Redistributions of source code must retain the above copyright notice,
     this list of conditions and the following disclaimers.
  2. Redistributions in binary form must reproduce the above copyright
     notice, this list of conditions and the following disclaimers in the
     documentation and/or other materials provided with the distribution.
  3. Neither the name "MadTeam", the names of the authors, their institution,
     nor the names of its contributors may be used to endorse or promote
     products derived from this Software without specific prior written
     permission.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR
IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY,
FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT.  IN NO EVENT SHALL THE
CONTRIBUTORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER
LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS
WITH THE SOFTWARE.
```

This is distributed with the code-generating framework, but not in generated code

Where do we start? the MG5aMC authors

- See https://github.com/mg5amcnlo/mg5amcnlo/blob/main/madgraph/AUTHORS :

```
***********************************************************************
**                                                                   **
**                        A U T H O R S                              **
**                                                                   **
**                              &                                    **
**                                                                   **
**                    C O N T R I B U T O R S                        **
**                                                                   **
***********************************************************************

CORE TEAM
---------

Rikkert FREDERIX
Stefano FRIXIONE
Valentin HIRSCHI
Fabio MALTONI
Olivier MATTELAER
Tim STELZER
Paolo TORRIELLI
Marco ZARO


CONTRIBUTORS
------------

Johan ALWALL
Pierre ARTOISENET
Diogo B FRANZOSI
Alexis KALOGEROPOULOS
Michel HERQUET
Robert RIETKERK
Chia-Hsien SHEN
```

This is distributed with the code-generating framework, but not in generated code

Where do we start? the MG5aMC code files

- e.g. https://github.com/mg5amcnlo/mg5amcnlo/blob/main/Template/LO/bin/madevent :
  - also copied as-is to generated code, e.g.
    https://github.com/madgraph5/madgraph4gpu/blob/master/epochX/cudacpp/gg_tt.mad/bin/madevent

```
#! /usr/bin/env python3

################################################################################
#
# Copyright (c) 2009 The MadGraph5_aMC@NLO Development team and Contributors
#
# This file is a part of the MadGraph5_aMC@NLO project, an application which
# automatically generates Feynman diagrams and matrix elements for arbitrary
# high-energy processes in the Standard Model and beyond.
#
# It is subject to the MadGraph5_aMC@NLO license which should accompany this
# distribution.
#
# For more information, visit madgraph.phys.ucl.ac.be and amcatnlo.web.cern.ch
#
################################################################################
```

This kind of headers is distributed with the code-generating framework, and in some (not all) generated code

Constraints on madgraph4gpu developments

- Our work is a _derivative_ of MG5aMC (and is used in _combination_ with it)
  - I think that we should acknowledge the authors and their copyright
  - But we are not forced to use the same NCSA license

- Our work will be used with other software in _combined_ projects
  - We should avoid licenses like GPL which forbid linking our libraries to non-GPL projects

Proposed license: LGPL

- Our libraries can be linked to any software, free or proprietary (unlike with GPL)

- Our source code can be modified, but derivative code must remain free and LGPL

- In practice
  - mention LGPL in the central COPYRIGHT file and in each file header
  - add central COPYING and COPYING.LESSER files
  - the above, both in the code-generating plugin and in generated code

Proposed COPYRIGHT file

This would be distributed with the code-generating plugin, and also in generated code

Proposed AUTHORS file

This would be distributed with the code-generating plugin.

But not in generated code?
(because otherwise I think we should add the MG5aMC authors file too)

Proposed central files

Code-generating plugin,
only in madgraph4gpu

```
bash > pwd
<...>/madgraph4gpuX/epochX/cudacpp

bash > ls -1F
AUTHORS@
CODEGEN/
COPYING@
COPYING.LESSER@
COPYRIGHT@
ee_mumu.mad/
ee_mumu.sa/
gg_ttggg.mad/
gg_ttggg.sa/
gg_ttgg.mad/
gg_ttgg.sa/
gg_ttg.mad/
gg_ttg.sa/
gg_tt.mad/
gg_tt.sa/
gq_ttq.mad/
gq_ttq.sa/
heft_gg_h.sa/
smeft_gg_tttt.sa/
tmad/
tput/
```

Code-generating plugin,
now in madgraph4gpu,
later in mg5amcnlo

```
bash > pwd
<...>/madgraph4gpuX/epochX/cudacpp/CO
DEGEN/PLUGIN/CUDACPP_SA_OUTPUT

bash > ls -1F
aloha/
AUTHORS@
COPYING@
COPYING.LESSER@
COPYRIGHT@
__init__.py
madgraph/
model_handling.py
output.py
```

```
bash > pwd
<...>/madgraph4gpuX/epochX/cudacpp/CO
DEGEN/PLUGIN/CUDACPP_SA_OUTPUT/
madgraph/iolibs/template_files

bash > ls -1F
AUTHORS
CMake/
COPYING
COPYING.LESSER
COPYRIGHT
cpp_model_parameters_cc.inc
cpp_model_parameters_h.inc
gpu/
read_slha.cc
read_slha.h
bash >
```

Generated code

```
bash > pwd
<...>/madgraph4gpuX/epochX/cudacpp/gg_tt.mad

bash > ls -1F
bin/
Cards/
CMake/
CMakeLists.txt
CODEGEN_mad_gg_tt_log.txt
COPYING
COPYING.LESSER
COPYRIGHT
Events/
lib/
MGMEVersion.txt
README
README.systematics
Source/
src/
SubProcesses/
TemplateVersion.txt
```

Should we also
- add the MG5aMC LICENSE?
(referenced in COPYRIGHT)
- add the MG5aMC AUTHORS?
(or reference it in COPYRIGHT?)
- or copy them verbatim in our files?

Should we also
- add the CUDACPP AUTHORS?

(Aside – about the structure of generated code)

- In the context of copyright but not only, I considered some directory changes

- Currently generated code (e.g. gg_tt.mad) has these directories
  – Source, with Fortran/MadEvent code
  – src, with CUDACPP code
  – SubProcesses/P1_etc_etc, mixing both Fortran/MadEvent code and CUDACPP code

- I considered a different structure:
  – Source and SubProcesses/P1_etc_etc with Fortran/MadEvent code
  – *me_cudacpp, encapsulating all of the CUDACPP generated code*
    - in practice, add me_cudacpp/src and me_cudacpp/ SubProcesses/P1_etc_etc
    - some advantages: cleanly separate source code and builds, cleaner COPYRIGHT/AUTHORS...

- I am not at all happy of the result, and I dropped the idea for the moment
  – It is also not yet clear if src will eventually be a library (now HelAmps.h is header-only!)
  – Maybe we will come back to something similar in the future – but not now!

Proposed file headers

- e.g. CPPProcess.cc (modified MG5aMC code – derivative work)

See https://github.com/madgraph5/madgraph4gpu/blob/c2d22aebd45348e56a0477f0c85bab3d8eab18ab/epochX/cudacpp/gg_tt.mad/SubProcesses/P1_gg_ttx/CPPProcess.cc

```
// Copyright (C) 2010 The MadGraph5_aMC@NLO development team and contributors.
// Created by: J. Alwall (Oct 2010) for the MG5aMC CPP backend.
//========================================================================
// Copyright (C) 2020-2023 CERN and UCLouvain.
// Licensed under the GNU Lesser General Public License (version 3 or later).
// Modified by: S. Roiser (Feb 2020) for the MG5aMC CUDACPP plugin.
// Further modified by: S. Hageboeck, O. Mattelaer, S. Roiser, A. Valassi, Z. Wettersten
(2020-2023) for the MG5aMC CUDACPP plugin.
//========================================================================
// This file has been automatically generated for CUDA/C++ standalone by
// MadGraph5_aMC@NLO v. 3.5.0_lo_vect, 2023-01-26
// By the MadGraph5_aMC@NLO Development Team
// Visit launchpad.net/madgraph5 and amcatnlo.web.cern.ch
//========================================================================
...
```

This kind of headers would be distributed with the code-generating plugin, and in all generated code

- e.g. MadgraphTest.h (brand new code – not a derivative from MG5aMC)

See https://github.com/madgraph5/madgraph4gpu/blob/c2d22aebd45348e56a0477f0c85bab3d8eab18ab/epochX/cudacpp/gg_tt.mad/SubProcesses/MadgraphTest.h

```
// Copyright (C) 2020-2023 CERN and UCLouvain.
// Licensed under the GNU Lesser General Public License (version 3 or later).
// Created by: S. Hageboeck (Dec 2020) for the MG5aMC CUDACPP plugin.
// Further modified by: S. Hageboeck, A. Valassi (2020-2023) for the MG5aMC CUDACPP plugin.
```

Proposed next steps

- Discuss <u>PR #643</u>, amend any obvious issues if needed, *merge it asap*
  - I am leaving on holiday on Thursday, I would like to get done with this "stuff" before...

- *NB: THIS IS A FIRST STEP, WE CAN (PROBABLY WILL?) AMEND IT...*
  - We had a first exchange with the CERN IT OSS expert, but we may need to iterate
  - We need to discuss with the full MG5aMC team (Gargnano workshop in Sep 2023?)

- For the SYCL plugin (and others if we want to release them)
  - Identify authors of the plugin itself (Nathan, Taylor... ?)
  - Identify copyright of the plugin itself (Argonne... ?)
  - *License: LGPL if CUDACPP is LGPL, this is a derivative*
  - Create similar central COPYRIGHT and AUTHORS files
  - Add "Created by ... for the MG5AMC SYCL plugin" for new code
  - Add "Modified by ... for the MG5AMC SYCL plugin" for derived code
    - Below the CUDACPP (and MG5aMC where present) headers
  - Proposed timescale: before any release of the SYCL plugin
    - And/or before branching off to a separate repository if this is foreseen