Introduction to Open-Source Software

Matthew Saltzman

Department of Mathematical Sciences
Clemson University
and
COIN-OR Foundation

INFORMS International, July 10, 2007

Outline

Disclaimer

I am not a lawyer!

Nothing I say in this presentation should be construed as constituting legal advice. If you have questions with legal ramifications, contact a real lawyer.

Laws discussed are US laws. Laws in other countries may differ.

The Constitution Says:

The Congress shall have Power. . . To *promote the Progress of Science and useful Arts*, by securing *for limited Times* to Authors and Inventors the exclusive Right to their respective Writings and Discoveries;

[Emphasis mine.]

Types of Intellectual Property

Copyright	Protects the expression of an idea.
Patent	Protects the design of an “invention”.
Trade Secret	Protects a design or expression of an idea.
Trademark	Protects a name, slogan, or symbol used to identify a company, product, or service.

Trademarks and Service Marks

- Of least interest to us.
- Someone infringes your trademark rights by misusing your registered name, slogan, or icon.
  - Pretending to be you or deliberately or accidentally confusing a consumer.
  - Using a trademarked name to refer to a generic product (e.g., Kleenex(tm) brand facial tissues, Xerox(tm) brand photocopying systems).

Trade Secrets

- Protection through obscurity.
- Unlike patents, no obligation to make public.
- No protection against accidental or independent discovery or reverse engineering.
- Some legal protection against outright theft, if you treat your secrets with care.

Patents

- Legal protection against others using your design without explicit license from you—even if discovered independently.
- In exchange for protection, you make your design public so others can learn and build on it.
- You must apply for protection. It's not cheap.
- Protection is (currently) 20 years from initial filing.
- Software and "business process" patents are a recent innovation (ca. 20 years).
- Patentable design must be novel and unobvious.

Copyrights

- Protect the expression, not the idea itself.
- Legal protection against others using or distributing your work without explicit license.
- Making the expression public is (generally) the point.
  - Writings, musical compositions, performances, etc.
  - Computer programs?
- Protection is (currently) automatic (registration optional).
- (Currently) the shorter of
  - 70 years after death of the author (120 years for corporate creations) or
  - 95 years after publication.
- Others can recreate the ideas independently.

Copyrights

Copyright holder has exclusive right:

- to produce copies or reproductions of the work and to sell those copies (including electronic copies);
- to import or export the work;
- to create derivative works (works that adapt the original work);
- to perform or display the work publicly;
- to sell or assign these rights to others;
- to transmit or display by means of digital audio/video transmission.

Copyright Exceptions

- *Fair use*: limited reproduction for scholarship and review.
- *Betamax decision*:
  - Time-shifting broadcast materials is fair use.
  - Manufacturers not liable for infringing use of their products.

  Consequences:
  - Backups are probably legal.
  - Media shifting is probably legal.

  But. . .
- *Digital Millennium Copyright Act* (DMCA):
  - Criminalizes circumvention of “digital rights management” (DRM) technology.
  - Restricts copying on the Internet.

Who Owns Your IP?

Hint: It may not be you.

- Your employer:
  - Probably owns any IP you create as part of your job;
  - May own any work-related IP you create:
    - even if not done as part of your job;
    - whether or not you use employer resources.
- You or your employer may transfer rights to a third party
  - e.g, books, journal publications, music, etc.
- Consulting:
  - depends on client-consultant relationship.
- etc.

Licenses

- Copyright holder can set terms and conditions for others to use or redistribute.
- Terms are set forth in a *license agreement*.
- User does not own software, he owns a right to use and (maybe) distribute under certain circumstances.

Proprietary Licenses

- “Shrink-wrap” and “click-through” licenses.
  - Implicit agreement by opening package or using software.
- Usually, no access to source code, no right to copy or redistribute.
- May have to install hardware or software keys.

Free and Open-Source Licenses

- Guarantee access to source code and the right to modify and use at will.
- Guarantee right to redistribute original and modified versions, provided license terms are preserved.

Brief History

- GNU (GNU's Not Unix) project
  - Started by Richard Stallman in 1980s.
  - Created GNU Public License to preserve software *freedom*.
    - "Free as in speech, not free as in beer."
  - Free versions of Unix utilities and development toolchain.
- Linux kernel in the 1990s.
  - Started by Linus Torvalds in 1990s.
  - Source released on the Internet.
  - Used GNU toolchain.
  - Licensed under GPL.
- Linux *distributions* package Linux kernel, GNU toolchain, other software.
  - e.g, Red Hat, Debian.
- *Open Source Definition*
  - Adapted from Debian Free Software Guidelines in 1988.
  - Open Source Initiative (http://www.opensource.org).

The Open Source Definition

- Free Redistribution
- Source Code
- Derived Works
- Integrity of the Author's Source Code
- No Discrimination Against Persons or Groups
- No Discrimination Against Fields of Endeavor
- Distribution of License
- License Must Not Be Specific to a Product
- License Must Not Restrict Other Software
- License Must Be Technology Neutral

Important Open-Source Licenses

Currently, 58 licenses are on the OSI-approved list. Nine are considered "widely used or with strong communities":

- Apache License, 2.0
- New BSD license
- GNU General Public License (GPL)
- GNU Library or "Lesser" General Public License (LGPL)
- MIT license
- Mozilla Public License 1.1 (MPL)
- Common Development and Distribution License
- Common Public License 1.0 (CPL)
- Eclipse Public License

If you are considering an open-source license, it's best to stick with these.
If you are considering publishing on COIN-OR, we strongly prefer the CPL.

All Open-Source Licenses Are Not the Same

- Open-source license terms apply only to redistribution.
  - Use for anything and mix and match code with different licenses.
  - May not be able to distribute the result.
- Licenses differ in the requirements made on derived works.
  - Some call the GPL “viral”:
    - if you use any GPL code (including linked libraries) in your code, your code must be distributed under the GPL.
  - Other licenses (LGPL, CPL) draw a clean line at linked libraries.
    - If your code links to library code, you are not required to use the same license for your code.
- Licenses differ in how they treat software patents.
- Other differences.

Open-Source Software as Adacemic Publication

Publishing source code is a highly efficient mechanism for knowledge dissemination/technology transfer.

Advantages Over Traditional Publications

Publishing papers about code is not enough. Publishing the code is better.

- All details are apparent.
- Reproducibility.
- Apply results immediately.
- Build effectively on previous results.
- Community development model can improve code quality, documentation, etc.

Existing Fora

- ACM TOMS, etc.
  - Oriented toward core algorithms—not complete, usable systems.
- Other journal code archives.
  - Reproducibility.
  - No active development.
  - Voluntary.
  - Usually not designed for usability.

Challenges for Code as Publication

What needs to happen to make publishing code academically rewarding?

- Overcome inertia on review committees.
- How to formalize peer review?
- Where are high-quality venues?
- Code quality and documentation standards are higher for reusable software.

Examples of Open Source for OR

- COIN-OR
  - 20+ interlocking and standalone projects
- GLPK/GMPL
- cdd
- Numerous others, mostly self-published, many just archival.
- Your project here.
  - Many advantages to hosting on COIN-OR.