Build tools

Compilers

Historical compilers

- Proprietary
  - Intel C++ Compiler (ICC, 1970's?)
  - Microsoft Visual C++ (MSVC, 1993)
  - ARM Compiler (ARMCC, 2005)
  - AMD Optimizing C/C++ Compiler (AOCC, 2017)

- Open source
  - GNU Compiler Collection (GCC, 1987)
  - LLVM (2003–)

Evolution of compilers

- 2014: ARM Compiler rebased on LLVM
- 2017: AMD Compiler was always based on LLVM
- 2021: Intel C++ Compiler rebased on LLVM

Current major compilers

- Microsoft Visual C++
  - default on MS Windows (in MS Visual Studio)
- GCC
  - default on most open source OSs
- LLVM (for C/C++: Clang)
  - base for hardware vendor compilers (Intel, ARM, AMD, nVidia)
  - default on MacOS, iOS (in Apple X Code)
  - default for native applications on Android

Components of a compiler

- Front-end (parses and analyses code – language-specific)
- Intermediate representation (IR) (most code optimization happens here)
- Back-end (writes assembly or machine code – ISA-specific)

- LLVM frontends:
  - C and C++ (Clang), Fortran (Flang), Rust, Zig, Swift
- LLVM backends:
  - Intel/AMD/ARM compilers, nVidia CUDA compiler, AMD ROCm

LLVM IR

```
define dso_local noundef i32 @square(int)(i32 noundef %num) #0 !dbg !10 {
entry:
  %num.addr = alloca i32, align 4
  store i32 %num, ptr %num.addr, align 4
  call void @llvm.dbg.declare(metadata ptr %num.addr, metadata !16, metadata !DIExpression()), !dbg !17
  %0 = load i32, ptr %num.addr, align 4, !dbg !18
  %1 = load i32, ptr %num.addr, align 4, !dbg !19
  %mul = mul nsw i32 %0, %1, !dbg !20
  ret i32 %mul, !dbg !21
}

declare void @llvm.dbg.declare(metadata, metadata, metadata) #1

attributes #0 = { mustprogress noinline nounwind optnone uwtable "frame-pointer"="all" "min-legal-vector-width"="0" "no-trapping-math"="true"
  "stack-protector-buffer-size"="8" "target-cpu"="x86-64" "target-features"="+cx8,+fxsr,+mmx,+sse,+sse2,+x87" "tune-cpu"="generic" }
attributes #1 = { nocallback nofree nosync nounwind speculatable willreturn memory(none) }
```

Compiling

Compiler invocation (1)

- Use `man gcc` / `man clang` for help.
- Compile and link:

```
gcc -o executable source_code.c
```

- Compile only:

```
gcc -c -o file.o file.c
```

- Link only

```
gcc -o executable file0.o file1.o file2.o file3.o
```

- Write assembly

```
gcc -S assembly.S source_code.c
```

- Internally, `gcc` runs other tools (assembler: `as`, linker: `ld`)

Compiler invocation (2)

- Enable warnings:

```
gcc -Wall -c -o file.o file.c
```

- Enable optimization:

```
gcc -Wall -O3 -c -o file.o file.c
```

Note for MacOS

Install `binutils`:

- from Homebrew https://brew.sh/

```
brew install binutils
```

- or from MacPorts https://www.macports.org

```
port install binutils
```

Utilities may be prefixed by a g:

`objdump` $\longrightarrow$ `gobjdump`

Tools

- `hexdump` dump hexadecimal representation of any file
  - `hexdump -C` also print ASCII for valid ASCII bytes
  - `hexdump -C | less` “pipe” outout to pager
  - `hexdump -C > file.hex` write outout to a file

- `objdump` dump contents of object file
  - `objdump -p` print object information (dynamically-linked libraries, symbols)
  - `objdump -M intel -d` disassembles object file, prints assembly code

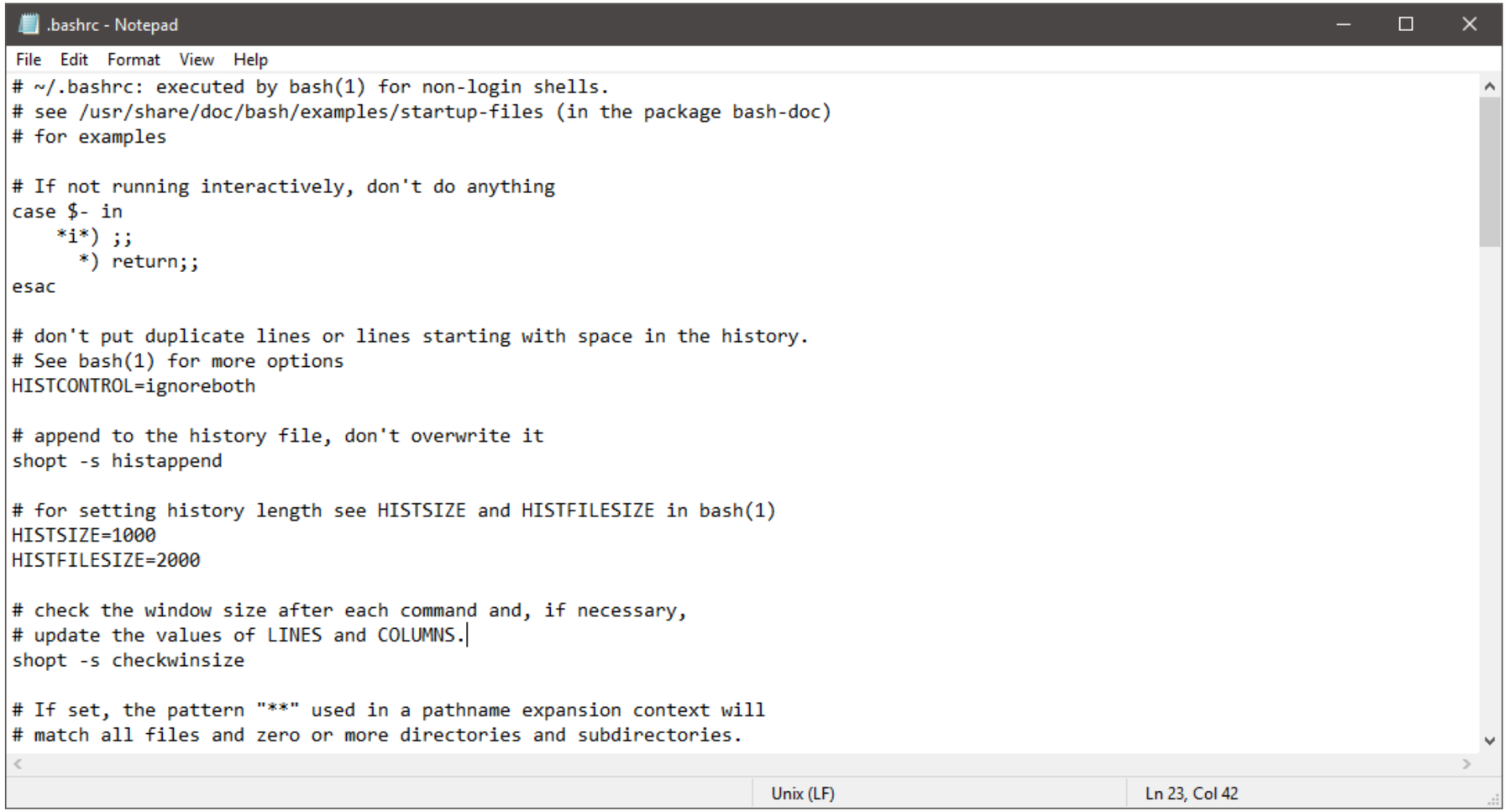
- or online: http://godbolt.org

Editing code

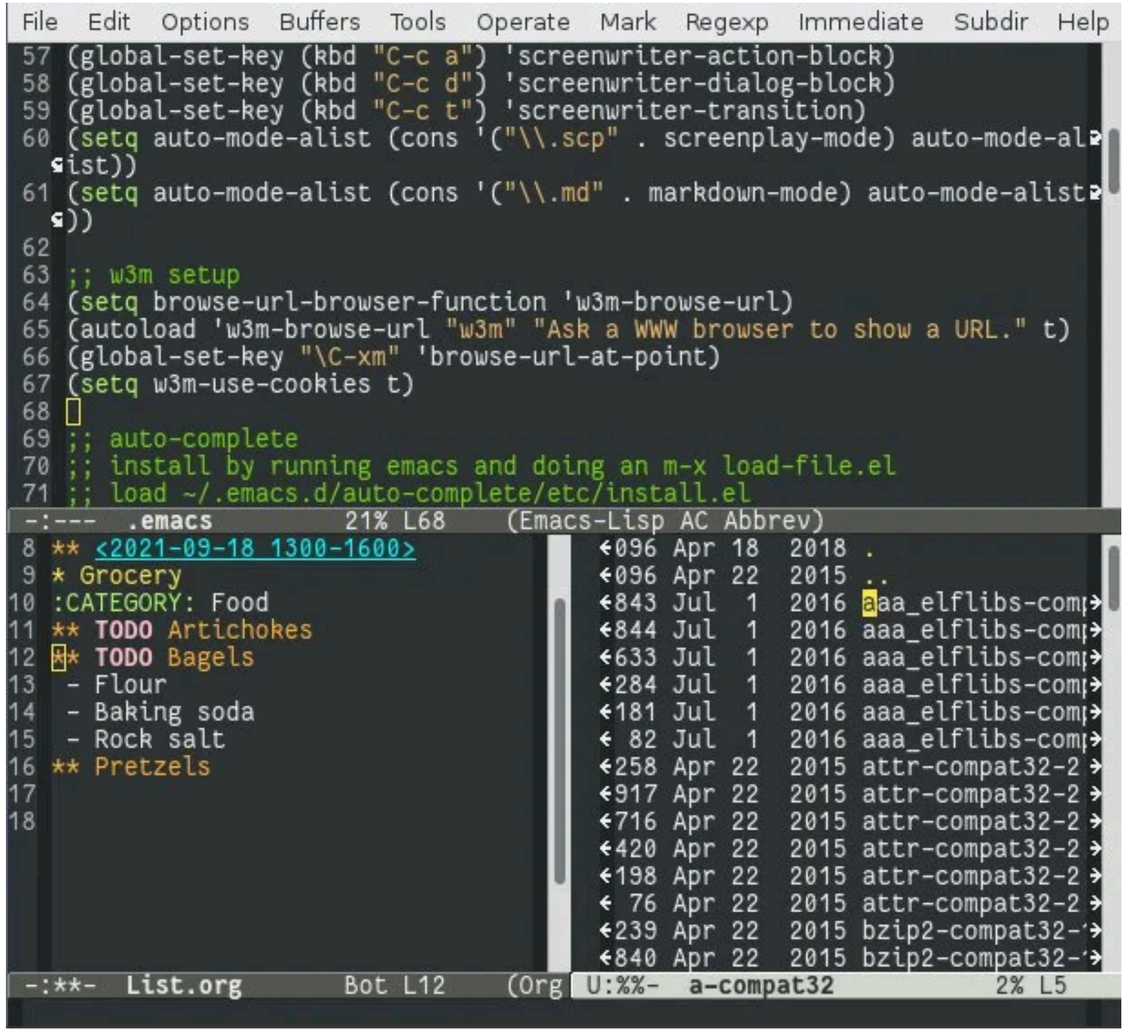
Applications for writing code

- Text editors
- Code editors
- Integrated development environment (IDE)

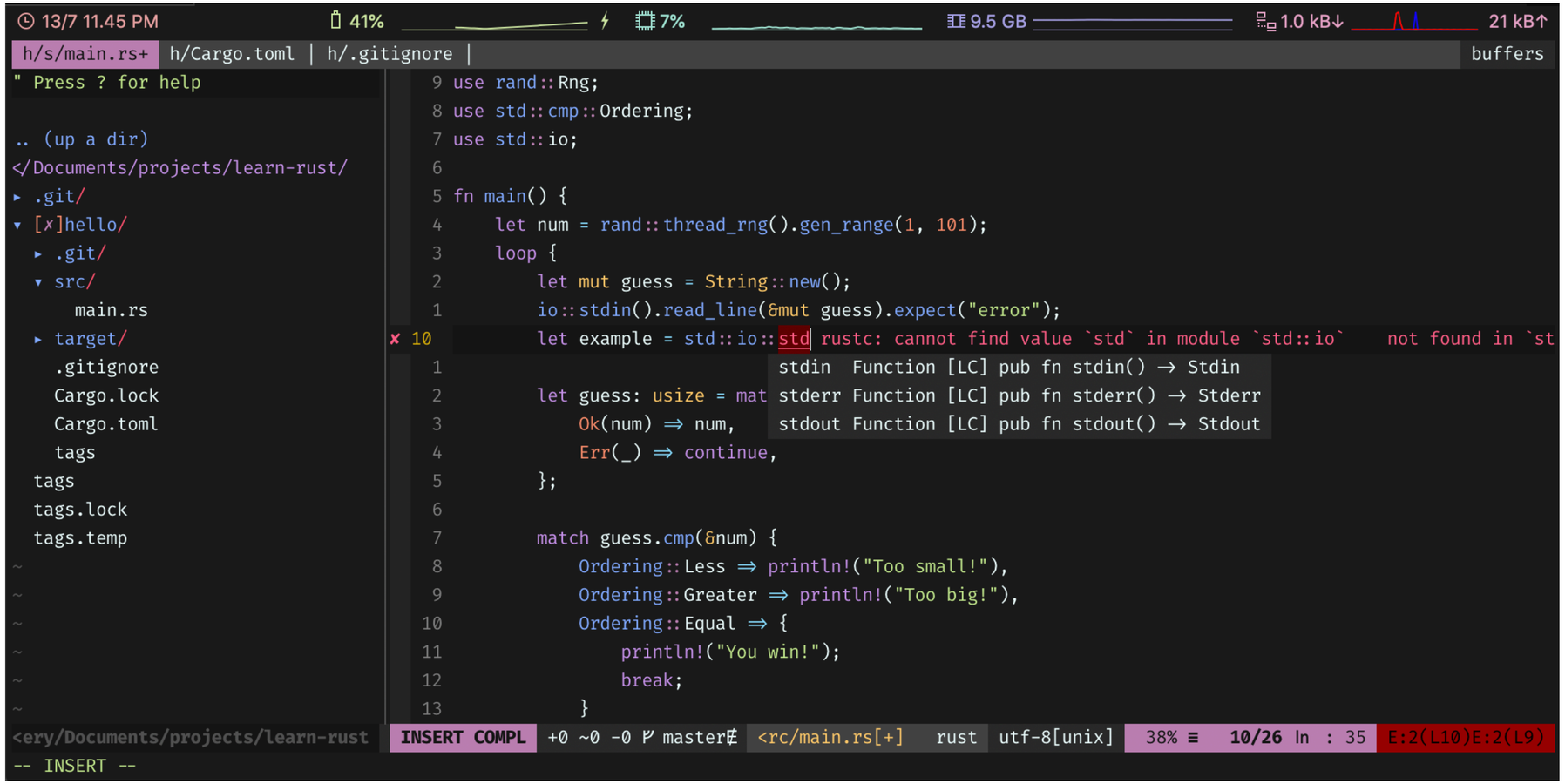
Text editor: Notepad



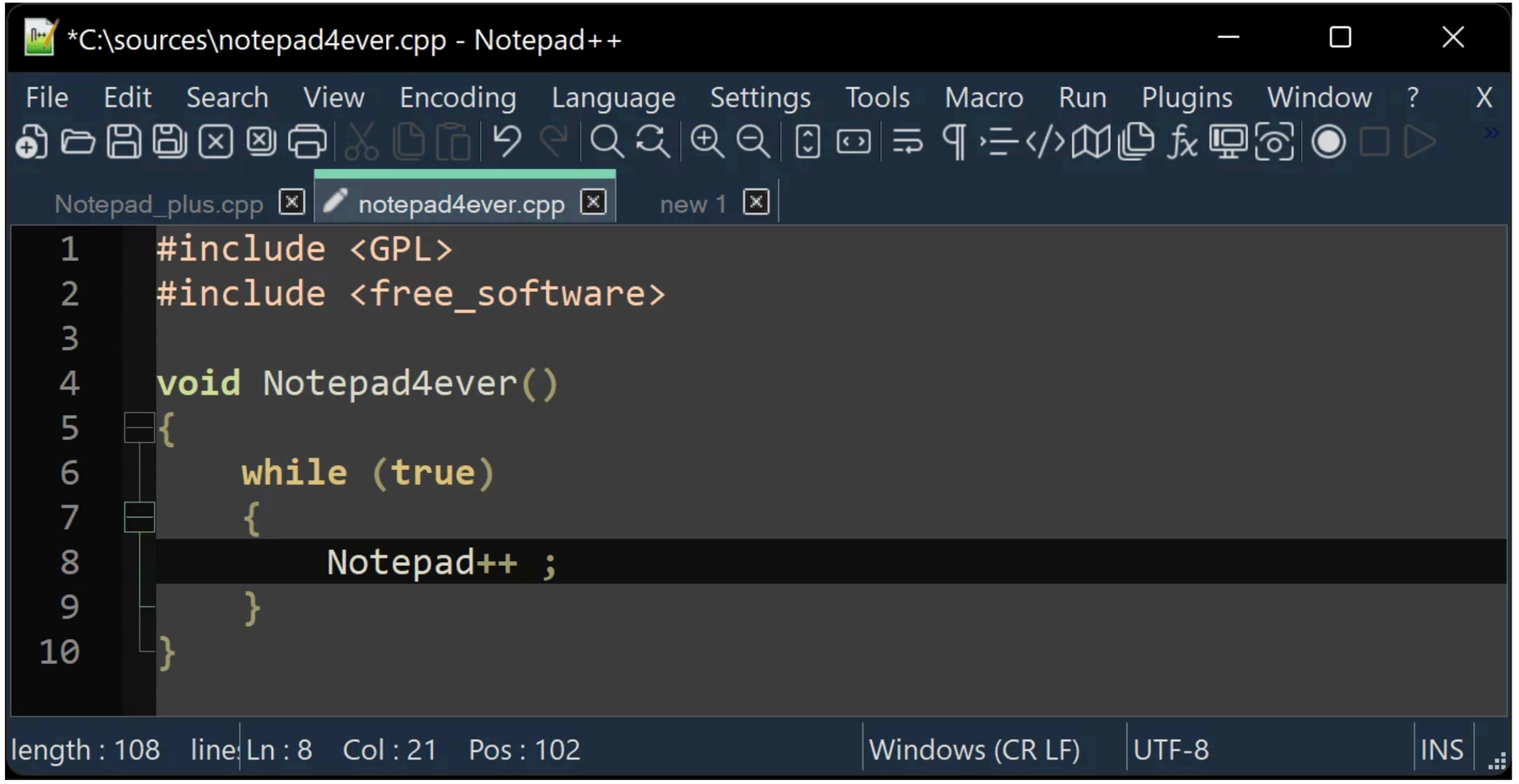
Code editor: emacs



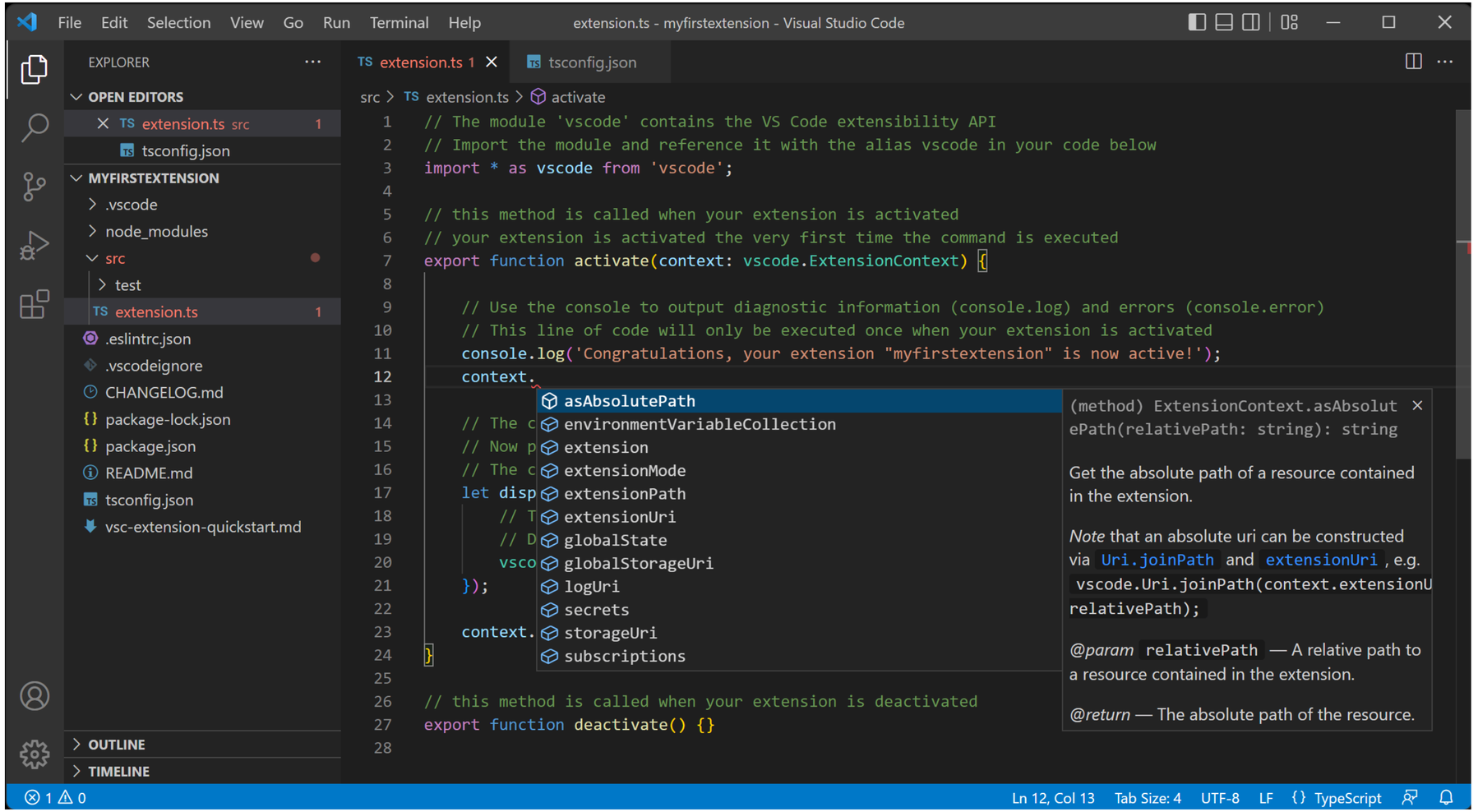
Code editor: vi / vim / neovim



Code editor: Notepad++



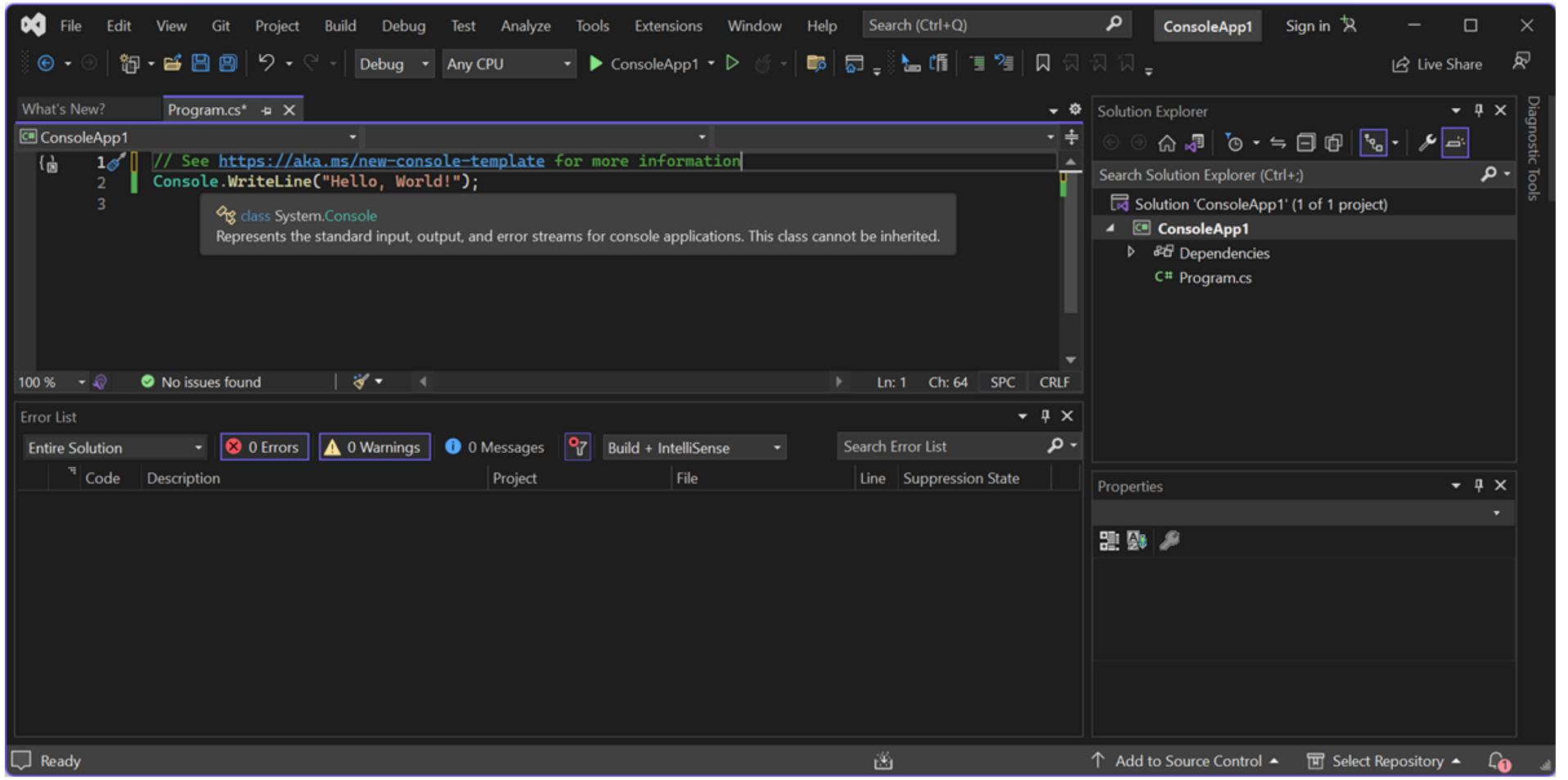
Code editor: Visual Studio Code



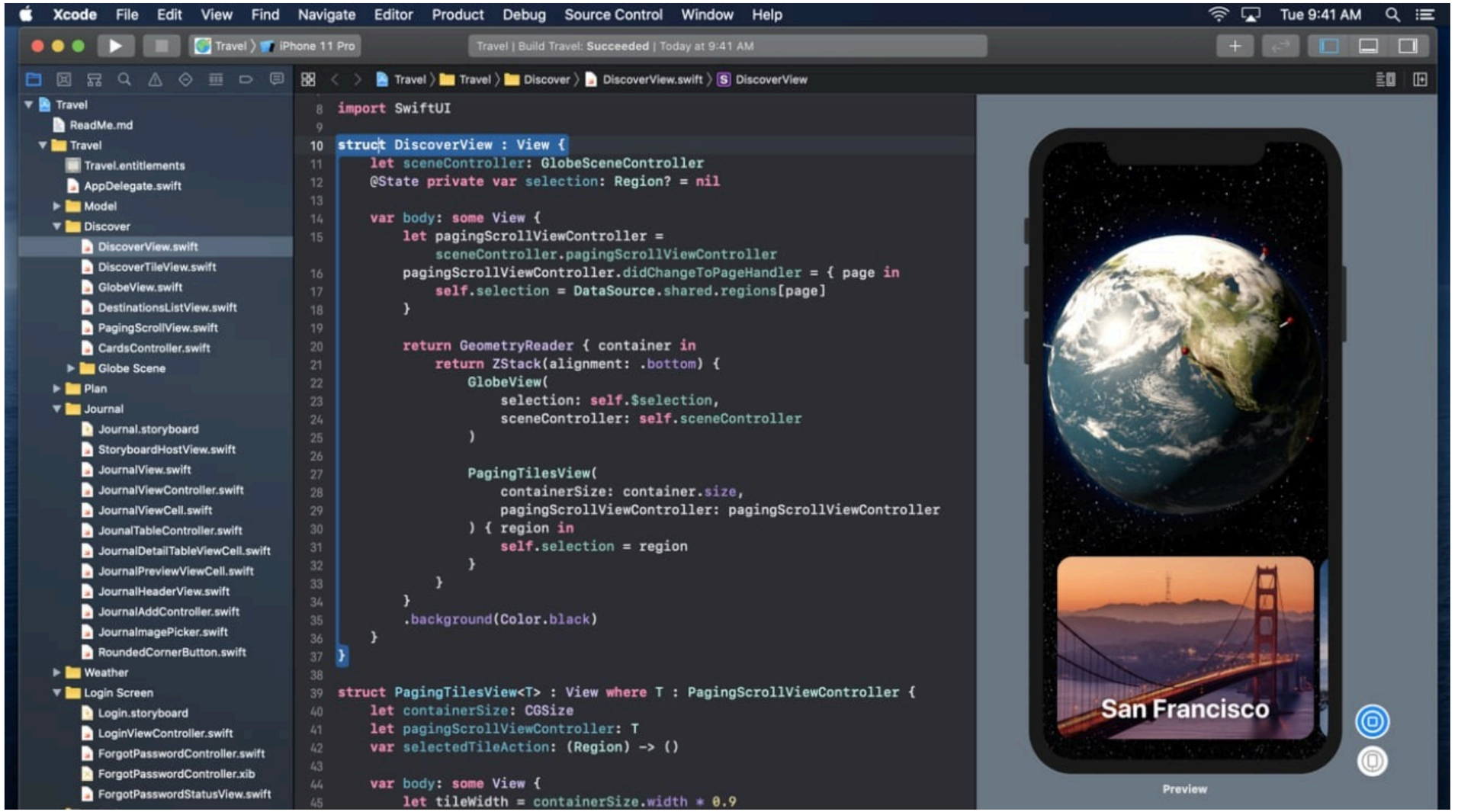
More code editors

- gedit
- Kate
- Sublime Text (paid)
- many more…

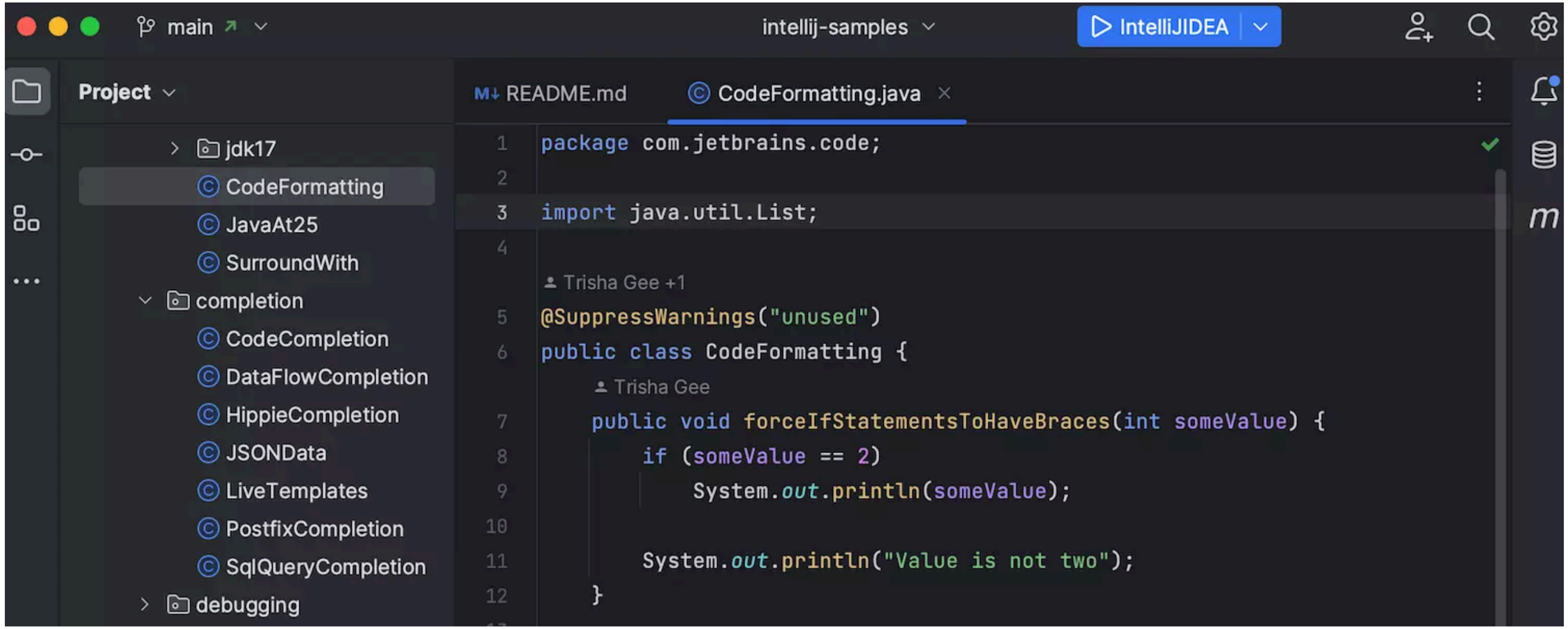
IDE: Microsoft Visual Studio (paid)



IDE: Apple Xcode



IDE: IntelliJ IDEA (paid)



More IDEs

- PyCharm (Python, paid)
- Android Studio (paid)
- KDevelop
- QtCreator
- Dev-C++
- Spyder (Python)
- …

Code editor vs. IDE

IDE pros:

- one-click compile
- IDE aware of whole project
  - can suggest code completions from different files
- integrated tools (e.g. debugger)

IDE cons:

- Project setup takes time and effort
- “Walled garden” problem
  - By default, anyone who wants to compile your project needs the same IDE.

Build systems

How do we compile a complex project?

- Option 1:

```
gcc -Wall -O3 -c -o ggml.o ggml.c
gcc -Wall -O3 -c -o ggml-alloc.o ggml-alloc.c
g++ -Wall -O3 -c -o llama.o llama.cpp
g++ -Wall -O3 -c -o common.o common/common.c
g++ -Wall -O3 -c -o console.o common/console.c
g++ -Wall -O3 -c -o grammar-parser.o common/grammar-parser.c
g++ -Wall -O3 -shared -fPIC -o libllama.so ggml.o ggml-alloc.o llama.o \
                              common.o console.o grammar-parser.o
```

- Option 2
  - Put above commands in a “shell script” file, e.g. `compile.sh`
  - Run:

```
./compile.sh
```

  - Problems:
    - Difficult to modify (e.g. change compiler options)
    - We recompile everything everytime

Build automation

- IDE integrated:
  - Visual Studio
  - Xcode

- Stand-alone:
  - `make`
  - Bazel (based on Google's internal tool Blaze) / Buck (Facebook)
  - Ninja (Google, for Chrome)
  - CMake (uses `make`, Ninja,…), `qmake` (uses `make`), Meson (uses Ninja, …)

Make

- Create a file named `Makefile`:

```
ggml.o: ggml.c ggml.h ggml-cuda.h
	gcc -Wall -O3 -c -o ggml.o ggml.c

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	gcc -Wall -O3 -c -o ggml-alloc.o ggml-alloc.c

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	g++ -Wall -O3 -c -o llama.o llama.cpp

common.o: common/common.cpp common/common.h build-info.h common/log.h
	g++ -Wall -O3 -c -o common.o common/common.cpp

console.o: common/console.cpp common/console.h
	g++ -Wall -O3 -c -o console.o common/console.cpp

grammar-parser.o: common/grammar-parser.cpp common/grammar-parser.h
	g++ -Wall -O3 -c -o grammar-parser.o common/grammar-parser.cpp

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	g++ -Wall -O3 -shared -fPIC -o libllama.so ggml.o ggml-alloc.o llama.o \
	                               common.o console.o grammar-parser.o
```

- Run

```
make libllama.so
```

Make rule syntax

```
target: source0 source1 source2 ...
        recipe
```

Whenever one of the `sources` was modified after the `target`,
run the recipe (to rebuild the `target`).

Otherwise, consider `target` up-to-date and do nothing.

Beware: `recipe` must be offset to the right using a TAB character!

Make variables

We can create variables in a Makefile with

```
VARIABLE_NAME := string content of the variable
```

Then, `$(VARIABLE_NAME)` can be used,
and will be expanded into `string content of the variable`.

Example:

```
contents_of_my_variable.txt:
	echo $(VARIABLE_NAME) > contents_of_my_variable.txt
```

```
ggml.o: ggml.c ggml.h ggml-cuda.h
	gcc -Wall -O3 -c -o ggml.o ggml.c

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	gcc -Wall -O3 -c -o ggml-alloc.o ggml-alloc.c

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	g++ -Wall -O3 -c -o llama.o llama.cpp

common.o: common/common.cpp common/common.h build-info.h common/log.h
	g++ -Wall -O3 -c -o common.o common/common.cpp

console.o: common/console.cpp common/console.h
	g++ -Wall -O3 -c -o console.o common/console.cpp

grammar-parser.o: common/grammar-parser.cpp common/grammar-parser.h
	g++ -Wall -O3 -c -o grammar-parser.o common/grammar-parser.cpp

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	g++ -Wall -O3 -shared -fPIC -o libllama.so ggml.o ggml-alloc.o llama.o \
	                              common.o console.o grammar-parser.o
```

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

ggml.o: ggml.c ggml.h ggml-cuda.h
	$(CC) $(CFLAGS) -c -o ggml.o ggml.c

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	$(CC) $(CFLAGS) -c -o ggml-alloc.o ggml-alloc.c

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	$(CXX) $(CXXFLAGS) -c -o llama.o llama.cpp

common.o: common/common.cpp common/common.h build-info.h common/log.h
	$(CXX) $(CXXFLAGS) -c -o common.o common/common.cpp

console.o: common/console.cpp common/console.h
	$(CXX) $(CXXFLAGS) -c -o console.o common/console.cpp

grammar-parser.o: common/grammar-parser.cpp common/grammar-parser.h
	$(CXX) $(CXXFLAGS) -c -o grammar-parser.o common/grammar-parser.cpp

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	$(CXX) $(CXXFLAGS) -shared -fPIC -o libllama.so ggml.o ggml-alloc.o llama.o \
	                                    common.o console.o grammar-parser.o
```

Special make variables

- `$(@)` the target of the current rule
- `$(<)` the first source of the current rule
- `$(^)` all the sources of the current rule

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

ggml.o: ggml.c ggml.h ggml-cuda.h
	$(CC) $(CFLAGS) -c -o ggml.o ggml.c

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	$(CC) $(CFLAGS) -c -o ggml-alloc.o ggml-alloc.c

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	$(CXX) $(CXXFLAGS) -c -o llama.o llama.cpp

common.o: common/common.cpp common/common.h build-info.h common/log.h
	$(CXX) $(CXXFLAGS) -c -o common.o common/common.cpp

console.o: common/console.cpp common/console.h
	$(CXX) $(CXXFLAGS) -c -o console.o common/console.cpp

grammar-parser.o: common/grammar-parser.cpp common/grammar-parser.h
	$(CXX) $(CXXFLAGS) -c -o grammar-parser.o common/grammar-parser.cpp

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	$(CXX) $(CXXFLAGS) -shared -fPIC -o libllama.so ggml.o ggml-alloc.o llama.o \
	                                    common.o console.o grammar-parser.o
```

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

ggml.o: ggml.c ggml.h ggml-cuda.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

common.o: common/common.cpp common/common.h build-info.h common/log.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

console.o: common/console.cpp common/console.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

grammar-parser.o: common/grammar-parser.cpp common/grammar-parser.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	$(CXX) $(CXXFLAGS) -shared -fPIC -o $(@) $(^)
```

Static pattern rules

- Static pattern syntax:

```
target0 target1 target2 ... : target-pattern : source-pattern
        recipe
```

- Target pattern contains %, which will match anything
- Source pattern also contains %, which is replaced by the match in target
- Example:

```
some_file.o other_file.o third_file.o : %.o : %.c
        recipe
```

is equivalent to:

```
some_file.o: some_file.c
        recipe

other_file.o: other_file.c
        recipe

third_file.o: third_file.c
        recipe
```

```
ggml.o: ggml.c ggml.h ggml-cuda.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)
```

becomes

```
ggml.o ggml-alloc.o: %.o: %.c %.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

ggml.o: ggml-cuda.h # Additional sources
ggml-alloc.o: ggml.h # Additional sources
```

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

ggml.o: ggml.c ggml.h ggml-cuda.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

ggml-alloc.o: ggml-alloc.c ggml.h ggml-alloc.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

common.o: common/common.cpp common/common.h build-info.h common/log.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

console.o: common/console.cpp common/console.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

grammar-parser.o: common/grammar-parser.cpp common/grammar-parser.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	$(CXX) $(CXXFLAGS) -shared -fPIC -o $(@) $(^)
```

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

ggml.o ggml-alloc.o: %.o: %.c %.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

ggml.o: ggml-cuda.h # Additional sources
ggml-alloc.o: ggml.h # Additional sources

llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h llama.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

common.o console.o grammar-parser.o: %.o: common/%.cpp common/%.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

common.o: build-info.h common/log.h # Additional sources

libllama.so: ggml.o ggml-alloc.o llama.o common.o console.o grammar-parser.o
	$(CXX) $(CXXFLAGS) -shared -fPIC -o $(@) $(^)
```

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

COBJS := ggml.o ggml-alloc.o
CXXOBJS_LLAMA := llama.o
CXXOBJS_COMMON := common.o console.o grammar-parser.o
CXXOBJS := $(CXXOBJS_LLAMA) $(CXXOBJS_COMMON)

# Build rules
$(COBJS): %.o: %.c %.h
        $(CC) $(CFLAGS) -c -o $(@) $(<)

$(CXXOBJS_LLAMA): %.o: %.cpp %.h
        $(CXX) $(CXXFLAGS) -c -o $(@) $(<)

$(CXXOBJS_COMMON): %.o: common/%.cpp common/%.h
        $(CXX) $(CXXFLAGS) -c -o $(@) $(<)

libllama.so: $(COBJS) $(CXXOBJS)
        $(CXX) $(CXXFLAGS) -shared -fPIC -o $(@) $(^)

# Additional sources
ggml.o: ggml-cuda.h
ggml-alloc.o: ggml.h
llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h
common.o: build-info.h common/log.h
```

Phony and default targets

- A “phony” target does not necessarily correspond to a file name:

```
.PHONY: clean

clean:
	rm libllama.so
```

- If no target is provided to the `make` command, the default target is the first one. A common pattern is:

```
.PHONY: default

default: libllama.so
```

```
CC := gcc
CXX := g++
CFLAGSS := -Wall -O3
CXXFLAGS := -Wall -O3

COBJS := ggml.o ggml-alloc.o
CXXOBJS_LLAMA := llama.o
CXXOBJS_COMMON := common.o console.o grammar-parser.o
CXXOBJS := $(CXXOBJS_LLAMA) $(CXXOBJS_COMMON)
LIBTARGET := libllama.so

.PHONY: default clean

# Build rules
default: $(LIBTARGET)

clean:
	rm -f $(COBJS) $(CXXOBJS) $(LIBTARGET)

$(COBJS): %.o: %.c %.h
	$(CC) $(CFLAGS) -c -o $(@) $(<)

$(CXXOBJS_LLAMA): %.o: %.cpp %.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

$(CXXOBJS_COMMON): %.o: common/%.cpp common/%.h
	$(CXX) $(CXXFLAGS) -c -o $(@) $(<)

$(LIBTARGET): $(COBJS) $(CXXOBJS)
	$(CXX) $(CXXFLAGS) -shared -fPIC -o $(@) $(^)

# Additional sources
ggml.o: ggml-cuda.h
ggml-alloc.o: ggml.h
llama.o: llama.cpp ggml.h ggml-alloc.h ggml-cuda.h ggml-metal.h
common.o: build-info.h common/log.h
```

Using shell commands

- The syntax is:

```
$(shell any-shell-command)
```

- For example:

```
TODAY := $(shell date)
C_FILES := $(shell ls *.c)
```

String replacement in variables

- The syntax is:

```
$(variable:pattern=replacement)
```

- The pattern contains %, which will match any substring
- The replacement may contain %, which will be replaced by the matched substring
- For example:

```
C_FILES := $(shell ls *.c)
O_FILES := $(C_FILES:%.c=%.o)
```

For more about make

```
# Using make
man make

# Writing Makefiles
info make
```