Section 2: Integer & Floating Point Numbers

- Representation of integers: unsigned and signed
- Unsigned and signed integers in C
- Arithmetic and shifting
- Sign extension

- Background: fractional binary numbers
- IEEE floating-point standard
- Floating-point operations and rounding
- Floating-point in C

Values for Different Word Sizes

	W			
	8	**16**	**32**	**64**
UMax	255	65,535	4,294,967,295	18,446,744,073,709,551,615
TMax	127	32,767	2,147,483,647	9,223,372,036,854,775,807
TMin	-128	-32,768	-2,147,483,648	-9,223,372,036,854,775,808

- **Observations**
  - | *TMin* | = *TMax* + 1
    - Asymmetric range
  - *UMax* = 2 * *TMax* + 1

- **C Programming**
  - #include <limits.h>
  - Declares constants, e.g.:
    - ULONG_MAX
    - LONG_MAX
    - LONG_MIN
  - Values are platform specific
  - See: /usr/include/limits.h on Linux

Signed vs. Unsigned in C

- **Constants**
  - By default are considered to be signed integers
  - Use “U” suffix to force unsigned:
    - `0U`, `4294967259U`

Signed vs. Unsigned in C

- **Casting**
  - `int tx, ty;`
  - `unsigned ux, uy;`
  - Explicit casting between signed & unsigned:
    - `tx = (int) ux;`
    - `uy = (unsigned) ty;`
  - Implicit casting also occurs via assignments and function calls:
    - `tx = ux;`
    - `uy = ty;`
    - The gcc flag *-Wsign-conversion* produces warnings for implicit casts, but *-Wall* does not!
  - How does casting between signed and unsigned work – what values are going to be produced?
    - *Bits are unchanged*, just interpreted differently!

Casting Surprises

- **Expression Evaluation**
  - If you mix unsigned and signed in a single expression, then ***signed values implicitly cast to unsigned***
  - Including comparison operations **<, >, ==, <=, >=**
  - Examples for *W* = 32: **TMIN = -2,147,483,648** **TMAX = 2,147,483,647**

Constant$_1$	**Constant$_2$**	**Relation**	**Evaluation**
0	0U	==	unsigned
-1	0	<	signed
-1	0U	>	unsigned
2147483647	-2147483648	>	signed
2147483647U	-2147483648	<	unsigned
-1	-2	>	signed
(unsigned)-1	-2	>	unsigned
2147483647	2147483648U	<	unsigned
2147483647	(int) 2147483648U	>	signed