True or false? (unless otherwise specified.)

Answer grid

Circle one answer for each question.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	B1	B2
True	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
False	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

1. The binary number $(101.101)_2$ is equal to the decimal number 5.625.
2. A real number has a finite representation in base 2 if and only if it is a rational number.
3. The binary representation of $\frac{1}{3}$ is $(0.\overline{01})_2$.
4. Let $(\bullet)_2$ denote binary representation. Then
$$(1000000000)_2 \times (0.001)_2 = (1000000)_2.$$
5. It holds that
$$(0.111111)_2 + (0.000001)_2 = (1.001)_2.$$
6. It holds that $(0.\overline{1100})_2 = \frac{4}{5}$.
7. Let ε_{32} denote the machine epsilon for the `Float32` format, i.e. `eps(Float32)`. Then the number 3 is representable exactly in this format, and the next representable number in the `Float32` format is $3+2\varepsilon_{32}$.
8. In Julia, `Float64(3.25) == Float32(3.25)` evaluates to `true`.
9. The spacing (in absolute value) between successive half-precision (`Float16`) floating point numbers is always greater than or equal to `eps(Float16)`.
10. In Julia, `Float64(.6) == Float32(.6)` evaluates to `true`.
11. There exists $n \in \mathbf{N}$ such that the machine epsilon for the `Float64` format equals 2^{-n}. In other words, the machine epsilon is an exact negative power of two.
12. In Julia, it holds that `eps()/3 + eps()/3 + 1 > 1` evaluates to `true`.
13. For any real number x such that $x \in \mathbf{F}_{64}$, it holds that $2x \in \mathbf{F}_{64}$.
14. Infinitely many distinct real numbers can be represented exactly in the `Float64` format, but only finitely many can be represented exactly in the `Float32` format.
15. Let x and y be two numbers in $\mathbf{F}_{64}$. The result of the machine multiplication $x \mathbin{\widehat{*}} y$ is sometimes exact and sometimes not, depending on the values of x and y.
16. **(Bonus)** Does the Julia expression `sqrt(1 + 2eps()) == 1` evaluate to `true` or `false`?

    Explain briefly:

17. **(Bonus)** Does the Julia expression `log2(2 + eps()) == 1` evaluate to `true` or `false`?

    Explain briefly:

Answer key

Questions 1–15: 1 point for a correct answer, 0 otherwise.

1. **True.** $(101.101)_2 = 4 + 1 + 1/2 + 1/8 = 5.625$.
2. **False.** A binary expansion is finite if and only if the number is a dyadic rational, i.e. of the form $m/2^k$ for integers m and k. Most rational numbers (e.g. 1/3) require an infinite expansion.
3. **True.** $1/3 = 0.010101\ldots_2$, i.e. $(0.\overline{01})_2$.
4. **True.** $(0.001)_2 = 2^{-3}$, so $2^9 \cdot 2^{-3} = 2^6 = (1000000)_2$.
5. **False.** The last carry propagates through all the bits: $(0.111111)_2 + (0.000001)_2 = (1.000000)_2 = 1$, not $(1.001)_2$.
6. **True.** The repeating block is 1100, whose value is 3/4, so $S = (3/4)/(1 - 1/16) = (3/4) \cdot (16/15) = 4/5$.
7. **True.** In `Float32` the spacing near 3 is $2 \cdot \varepsilon_{32}$ (the exponent is 1), so the next representable number is $3 + 2\varepsilon_{32}$.
8. **True.** $3.25 = 13/4$ is a dyadic rational, so it is represented exactly in both formats and converting one to the other is exact.
9. **False.** For numbers below 1, e.g. between 1/2 and 1, the spacing in `Float16` is 2^{-11}, which is smaller than `eps(Float16) = 2^{-10}`.
10. **False.** The decimal number 0.6 is not exactly representable in binary, and its nearest `Float64` and `Float32` values differ.
11. **True.** `eps(Float64)` is $2^{-52} = \left(\frac{1}{2}\right)^{52}$, an exact (negative) power of two.
12. **True.** Each `eps()/3` is represented almost exactly, and twice the value is about $2\varepsilon/3 \approx 1.48 \times 10^{-16}$, which is larger than the rounding threshold $\varepsilon/2 \approx 1.11 \times 10^{-16}$. Adding this quantity to 1 therefore rounds upward to $1 + \varepsilon$, so the left-hand side is strictly larger than 1. (Confirmed in Julia.)
13. **False.** Doubling can fall outside the representable range (overflow to ∞, underflow of subnormals) or introduce rounding: e.g. $2 \cdot x$ for the largest finite number is not finite.
14. **False.** Both formats represent only finitely many real numbers exactly.
15. **True.** The rounded result of the multiplication is exact when the product fits in the mantissa (e.g. powers of two) and is rounded otherwise.

Bonus.

1. **False.** The true value of $\sqrt{1 + 2\varepsilon}$ is approximately $1 + \varepsilon - \varepsilon^2/2$, which rounds to the nearest representable number `1 + eps()` (i.e. $1 + \varepsilon$), not to 1.
2. **True.** $2 + \varepsilon$ lies exactly at the midpoint between the consecutive representable numbers 2 and $2 + 2\varepsilon$ (the spacing near 2 is 2ε), so with round-to-nearest-even it rounds to 2. Consequently $\log_2(2 + \varepsilon)$ computes $\log_2(2) = 1$ exactly. (Confirmed in Julia.)