Question A.

$A=2$ is the base of the floating point implementation,
$p=24$ is the number of digits (of base β) of precision,
$e_{max} = +127$, $e_{min} = -126$ are the limits of the exponent.

The exponent takes $e_{max} - e_{min} + 3 = 256$ stored values, requiring 8 bits.
The significand is stored in 23 bits ($p = 23 + 1$ hidden bit).
The exponent e is stored as $e + e_{max}$. [6 marks]

Exponent	Fraction	Represents	
$e = e_{min} - 1$	$f = 0$	± 0	zero
$e = e_{min} - 1$	$f \neq 0$	$\pm 0.f \times 2^{e_{min}}$	denormal numbers
$e_{min} \leq e \leq e_{max}$	any f	$\pm 1.f \times 2^{e}$	normalised numbers
$e = e_{max} + 1$	$f = 0$	$\pm \infty$	infinities
$e = e_{max} + 1$	$f \neq 0$	NaN	Not a Number

[5 marks]

(a) $x \omega n$ evaluates to NaN in all cases
(b) $\pm\infty \omega n$ " " " " " "
(c) $x \omega \pm\infty$ evaluates to $\pm\infty$ for $\omega = +, -$ or $*$
" " " ± 0 for $\omega = /$
(d) $\sqrt{+\infty}$ evaluates to $+\infty$
$\sqrt{-\infty}$ " " NaN [6 marks]

(e) $z = 1 \times 2^{e_{min}} = 2^{-126}$

(f) Largest representable number smaller than z
= largest representable positive denormal number
$= 0.11\ldots1 \times 2^{e_{min}}$
$= (1 - 2^{-23}) \times 2^{-126}$

(g) Smallest representable positive number = smallest positive denormal number
$= 0.00\ldots01 \times 2^{e_{min}} = 2^{-23} \times 2^{e_{min}} = 2^{-149}$. [3 marks]