

Ιόνιο Πανεπιστήμιο – Τμήμα Πληροφορικής
Εισαγωγή στην Επιστήμη των Υπολογιστών

Αριθμοί Κινητής Υποδιαστολής

(παραδείγματα)

<http://mixstef.github.io/courses/csintro/>

Μ.Στεφανιδάκης



Αριθμοί κινητής υποδιαστολής

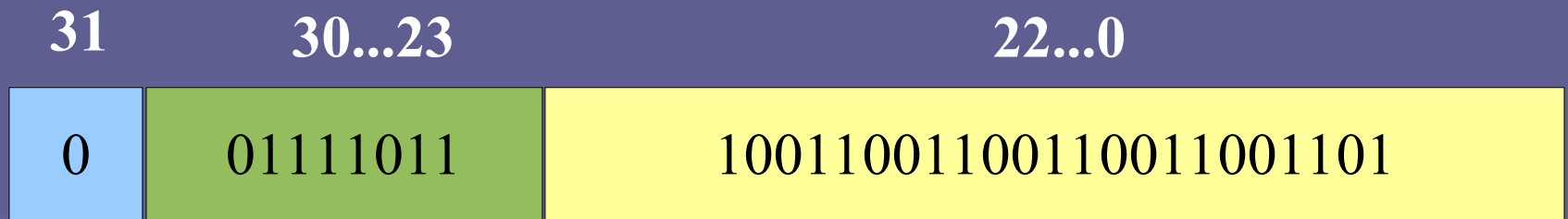
- Floating-point numbers
 - Οι πράξεις γίνονται στο υλικό (hardware)
 - Η αναπαράσταση είναι **προσεγγιστική** (δεν μπορεί να αναπαρασταθεί κάθε κλασματικός αριθμός!)
- Παγιωμένα πρότυπα (IEEE 754)
 - Αριθμοί απλής (C floats, 32 bits) και διπλής ακρίβειας (C doubles, 64 bits)
- Νέα πρότυπα
 - Με την άνθηση της μηχανικής μάθησης και των νευρωνικών δικτύων
 - Δεν απαιτείται μεγάλη ακρίβεια αλλά μεγάλη ταχύτητα επεξεργασίας/μικρότερο μέγεθος δεδομένων
 - Μειωμένος αριθμός bits (π.χ. 16 bit **bfloat16** format)

Αριθμοί μονής ακρίβειας (C floats)



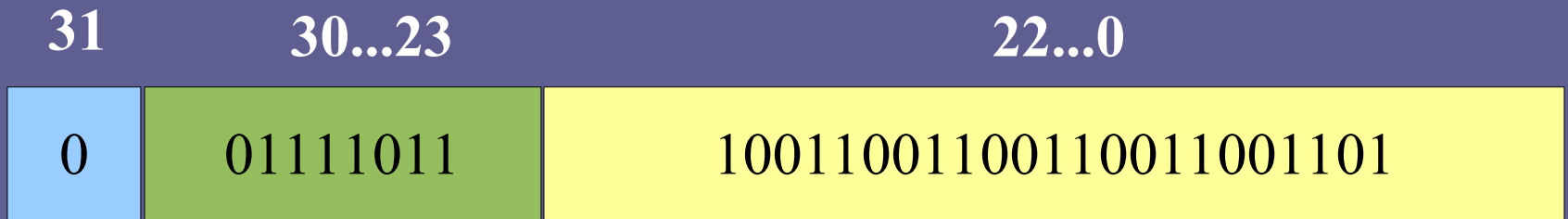
- 32 bits (4 bytes)
 - bit 31: Πρόσημο (Π)
 - bits 30...23: Εκθέτης (E)
 - bits 22...0: Σημαινόμενο Τμήμα (Σ)
- $-1^{\Pi} \times 2^{E-127} \times 1.\Sigma$

Αριθμοί μονής ακρίβειας (C floats)



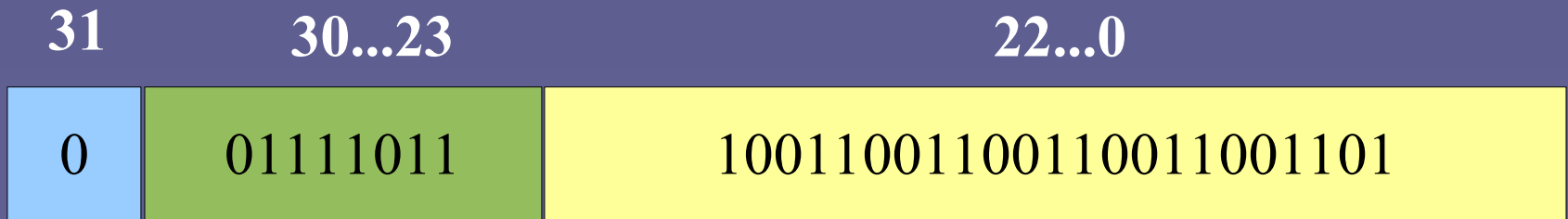
- Έστω αριθμός float με τα ακόλουθα bits: 3D CC CC CD
00111101 11001100 11001100 11001101
- Πρόσημο: 0 (+)

Αριθμοί μονής ακρίβειας (C floats)



- Έστω αριθμός float με τα ακόλουθα bits: 3D CC CC CD
00111101 11001100 11001100 11001101
- Πρόσημο: 0 (+)
- Εκθέτης: 01111011 (123), άρα $123 - 127 = -4$
($2^{-4} = 0,0625$)

Αριθμοί μονής ακρίβειας (C floats)



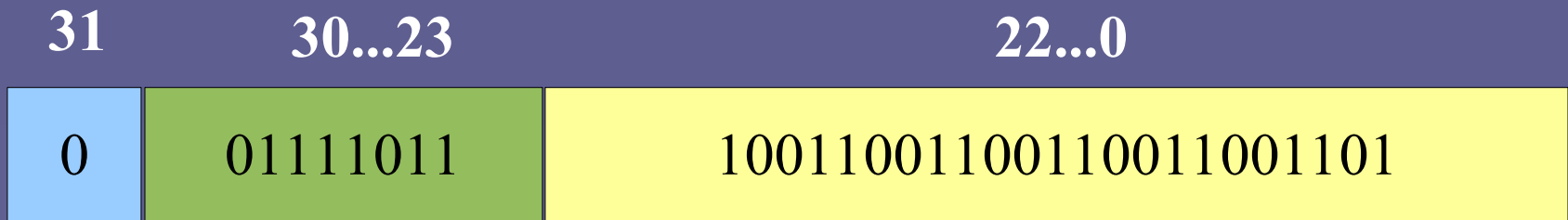
- Σημαινόμενο Τμήμα

1.10011001100110011001101

- Βάρη (από αριστερά προς τα δεξιά)

$1/2^1, 1/2^2, 1/2^3, 1/2^4, \dots 1/2^{23}$

Αριθμοί μονής ακρίβειας (C floats)



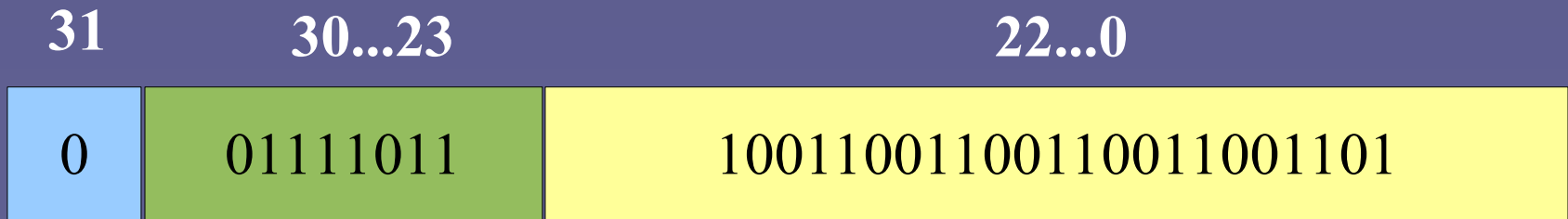
- Σημαινόμενο Τμήμα

$$1/2^1 + 1/2^4 + 1/2^5 + 1/2^8 + 1/2^9 + 1/2^{12} + 1/2^{13} + 1/2^{16} + 1/2^{17} + 1/2^{20} + 1/2^{21} + 1/2^{23}$$

$$= 0.5 + 0.0625 + 0.03125 + 0.00390625 + 0.001953125 + 0.000244140625 + 0.0001220703125 + 0.0000152587890625 + 0.00000762939453125 + 0.00000095367431640625 + 0.000000476837158203125 + 0.00000011920928955078125$$

$$= 0.60000002384185791015625 (+1.0)$$

Αριθμοί μονής ακρίβειας (C floats)



- Έστω αριθμός float με τα ακόλουθα bits: 3D CC CC CD
00111101 11001100 11001100 11001101
- Πρόσημο: +
Εκθέτης: 123, $2^{123-127} = 2^{-4} = 0,0625$
- Σημαινόμενο Τμήμα: 0.60000002384185791015625 (+1.0)
- Τελικός αριθμός: + 0.100000001490116119384765625