

04. Tipi di dato

Corso di Algoritmi e Linguaggi di Programmazione Python/C

Outline

- I dati numerici interi
 - Segno nei dati di tipo intero
- I dati numerici reali
 - Fixed point
 - Floating point
- Stringhe e caratteri

I dati numerici interi

- Esistono diversi tipi di dati numerici interi, differenziati in base al numero di bit (e quindi al range di valori) rappresentabili.
- Anche se possono esserci delle piccole differenze tra diversi linguaggi, nella maggior parte dei casi i valori rispettano quelli elencati nella tabella a destra.

Tipo	Lunghezza (bit)	Range di valori
bit	1	$[0, 1]$
ubyte	8	$[0, 2^8 - 1]$
ushort	16	$[0, 2^{16} - 1]$
uint	32	$[0, 2^{32} - 1]$
ulong	64	$[0, 2^{64} - 1]$

Segno nei dati di tipo intero

- Nella precedente tabella, abbiamo anteposto a ciascun tipo (tranne il *bit*) la lettera *u*, abbreviazione di *unsigned*.
- Significa che li interi considerati *non hanno segno*, e possono rappresentare solo valori positivi.
- Per rappresentare valori negativi, dovremo usare dati *signed*.

Tipo	Lunghezza (bit)	Range di valori
byte	8	$[-2^7, 2^7 - 1]$
short	16	$[-2^{15}, 2^{15} - 1]$
int	32	$[-2^{31}, 2^{31} - 1]$
long	64	$[-2^{63}, 2^{63} - 1]$

I dati numerici reali – Fixed point

- I numeri reali possono essere rappresentati in due modi.
- Nella rappresentazione a virgola fissa (**fixed point**), dati W bit a disposizione, i primi W_r rappresentano la parte intera, mentre gli ultimi W_f la parte decimale.
- Ad esempio:

$$W = 8, N = 1.11, W_r = 4, W_f = 4 \Rightarrow$$

$$\Rightarrow N_2 = 0001.1011$$

- Ovviamente, se $N > 15$, avremo bisogno di 5 bit, per cui il numero non potrà più essere rappresentato con l'attuale configurazione.

I dati numerici reali – Floating point

- La virgola fissa ha svantaggi notevoli legati alla scarsa flessibilità.
- Per sopperire a ciò, si usa la rappresentazione a virgola mobile (**floating point**).
- Questa si basa sul concetto di **mantissa** ed **esponente**.
- Ad esempio:

$$N = 1.11 = 0.111 \cdot 10^2$$

- Stiamo esprimendo il numero in funzione di una parte reale (**mantissa**) e di un **esponente** moltiplicativo.

I dati numerici reali – Floating point

- Possiamo quindi scrivere che:

$$N = M \cdot 10^e \Rightarrow$$

$$\Rightarrow N_b = M \cdot b^e$$

- dove b è la base considerata.
- Il vantaggio di questa notazione non è il richiedere un numero inferiore di bit, ma l'avere una maggiore flessibilità, in quanto il numero di bit da destinare a mantissa ed esponente può variare.

Stringhe e caratteri

- I dati non numerici (ad eccezione dei valori booleani) sono considerati come **caratteri**.
- Per i caratteri, abbiamo diverse possibili rappresentazioni, tra le cui più utilizzate sono le codifiche ASCII, UTF-8 ed UNICODE.
- Queste rappresentazioni si rifanno comunque a dei particolari codici numerici, quindi in realtà sono interpretate dal calcolatore come numeri.
- **Curiosità:** l'UNICODE, che è una rappresentazione estremamente completa dei caratteri esistenti (inclusendo, ad esempio, il cirillico, il kanji ed il mandarino), richiede soltanto 16 bit.

Domande?

42