

El problema del Bandido Multi Brazo



El problema del Bandido Multi Brazo



D1



D2



D3

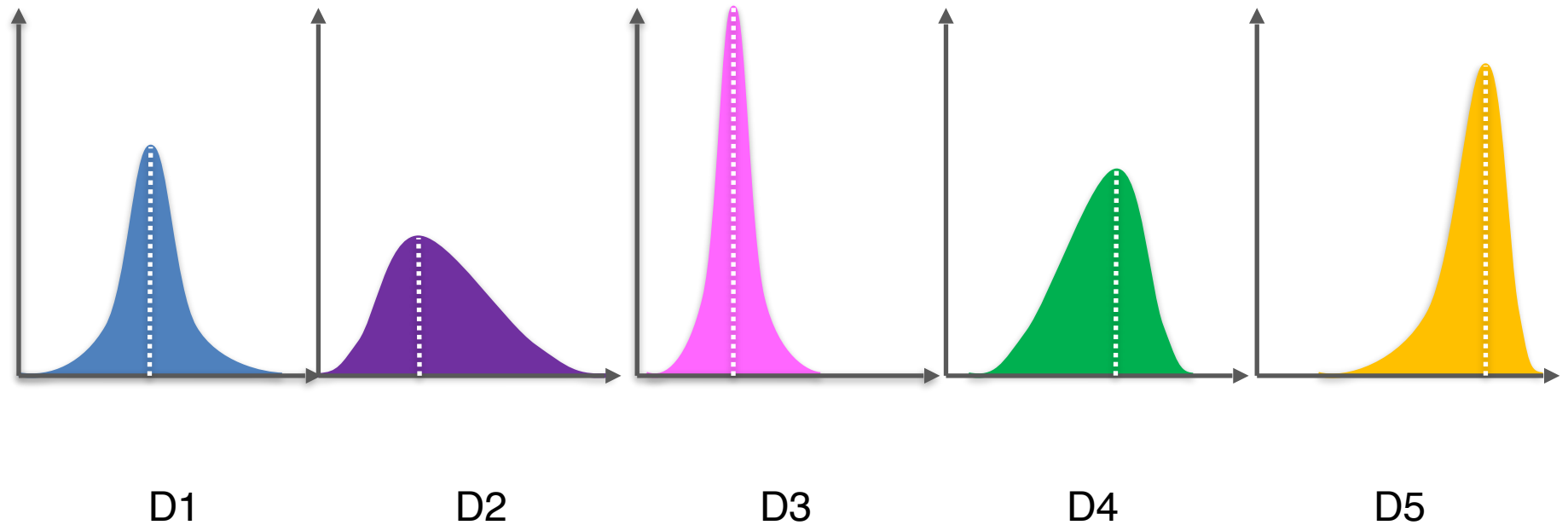


D4



D5

El problema del Bandido Multi Brazo



El problema del Bandido Multi Brazo



D1



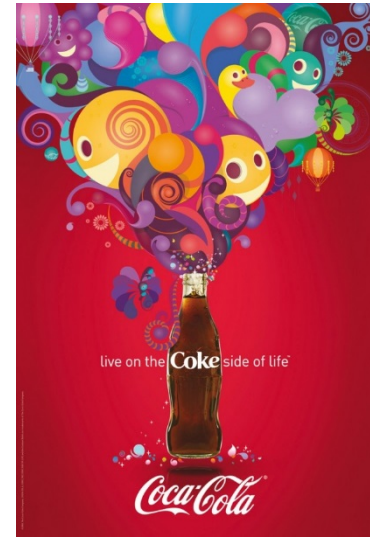
D2



D3



D4



D5

Examples used for educational purposes. No affiliation with Coca-Cola

Idea de Upper Confidence Bound (UCB)

El problema del Bandido Multi Brazo



D1



D2



D3



D4



D5

El problema del Bandido Multi Brazo

- Tenemos **d** brazos. Por ejemplo, los brazos son anuncios que mostramos a los usuarios cuando se conectan a una página web.
- Cada vez que un usuario se conecta a la página web, se desencadena una ronda.
- En cada ronda, **n**, se elige uno de los anuncios a ser mostrado al usuario.
- A cada ronda **n**, el anuncio **i** da una recompensa: $r_i(n) \in \{0, 1\}$

$r_i(n) = 1$ Si el usuario hace click en el anuncio **i** en la ronda **n**

$r_i(n) = 0$ Si el usuario no hace click en el anuncio **i** en la ronda **n**

- Nuestra meta es maximizar la recompensa a través de las rondas que se lleven a cabo.

Algoritmo Upper Confidence Bound

PASO 1: A cada ronda n , se consideran dos números para cada i

- $N_i(n)$ El número de veces que el anuncio i se selecciona hasta la ronda n ,
- $R_i(n)$ La suma de las recompensas del anuncio i hasta la ronda n

PASO 2: A partir de estos dos números calculamos

- La recompensa media del anuncio i hasta la ronda n

$$\bar{r}_i(n) = \frac{R_i(n)}{N_i(n)}$$

- El intervalo de confianza en la ronda n :

$$(\bar{r}_i(n) - \Delta_i(n), \bar{r}_i(n) + \Delta_i(n)) \quad \text{con} \quad \Delta_i(n) = \sqrt{\frac{3 \log(n)}{2N_i(n)}}$$

PASO 3: Se selecciona el anuncio i con mayor límite superior del intervalo de confianza (UCB):

Algoritmo Upper Confidence Bound



D1



D2



D3

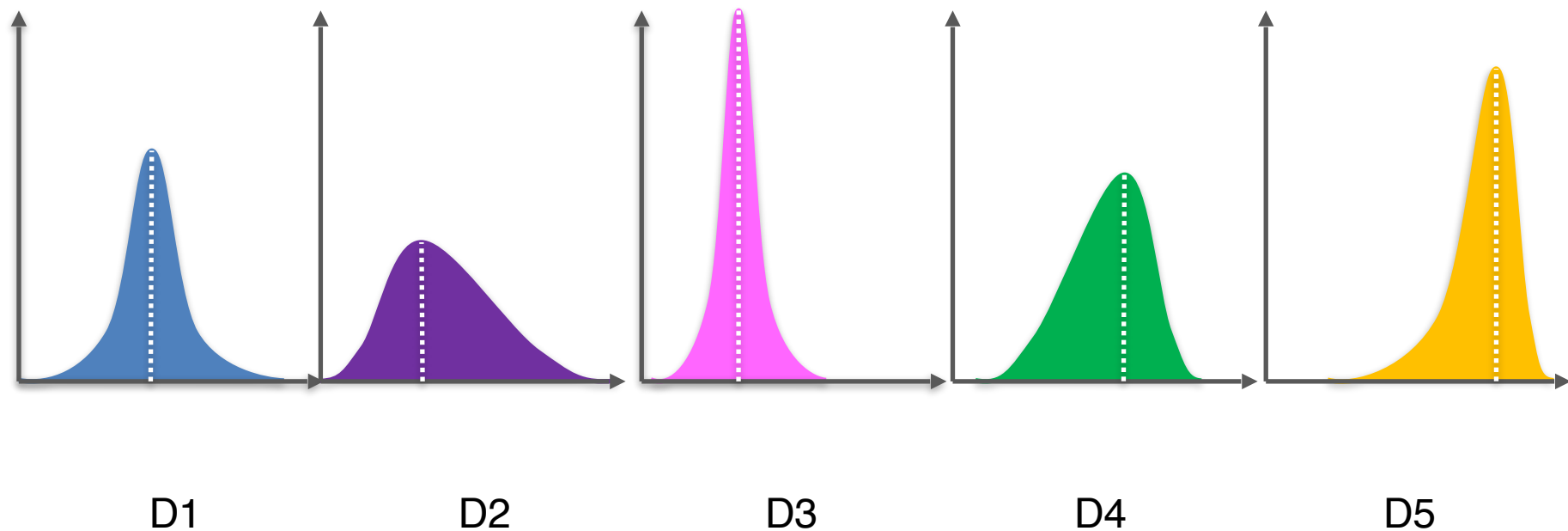


D4

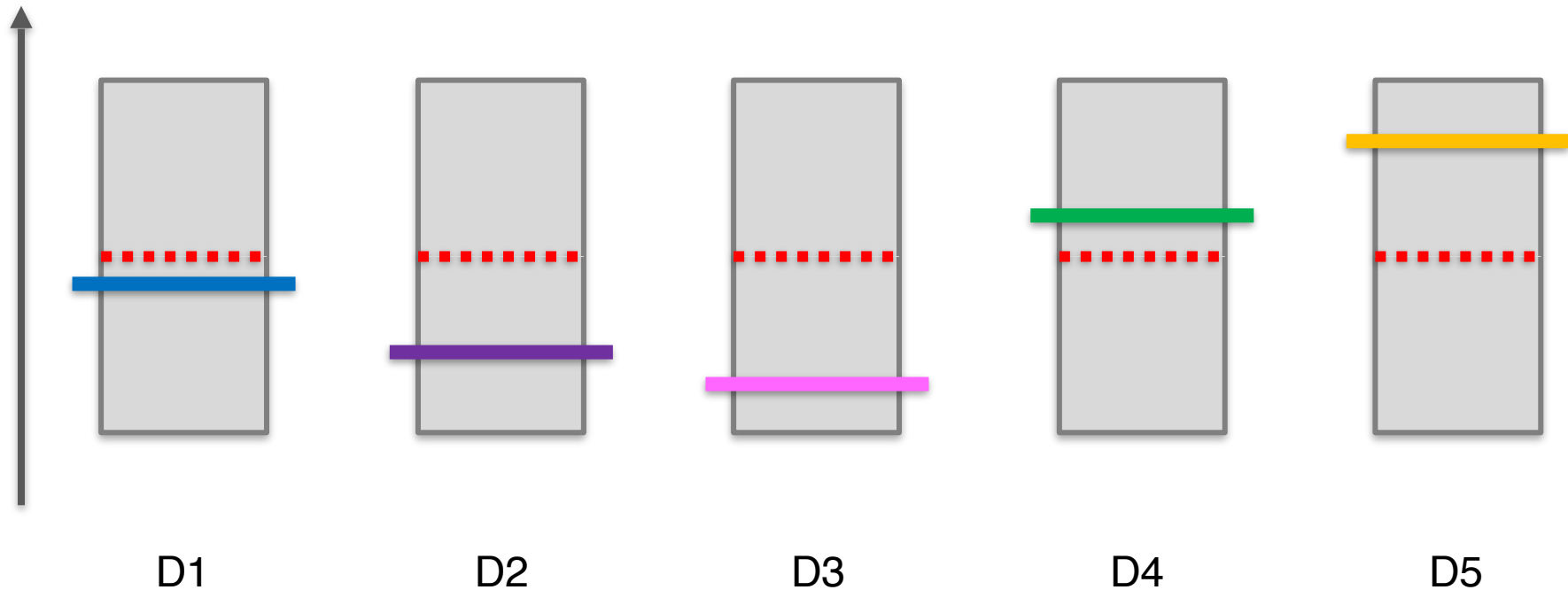


D5

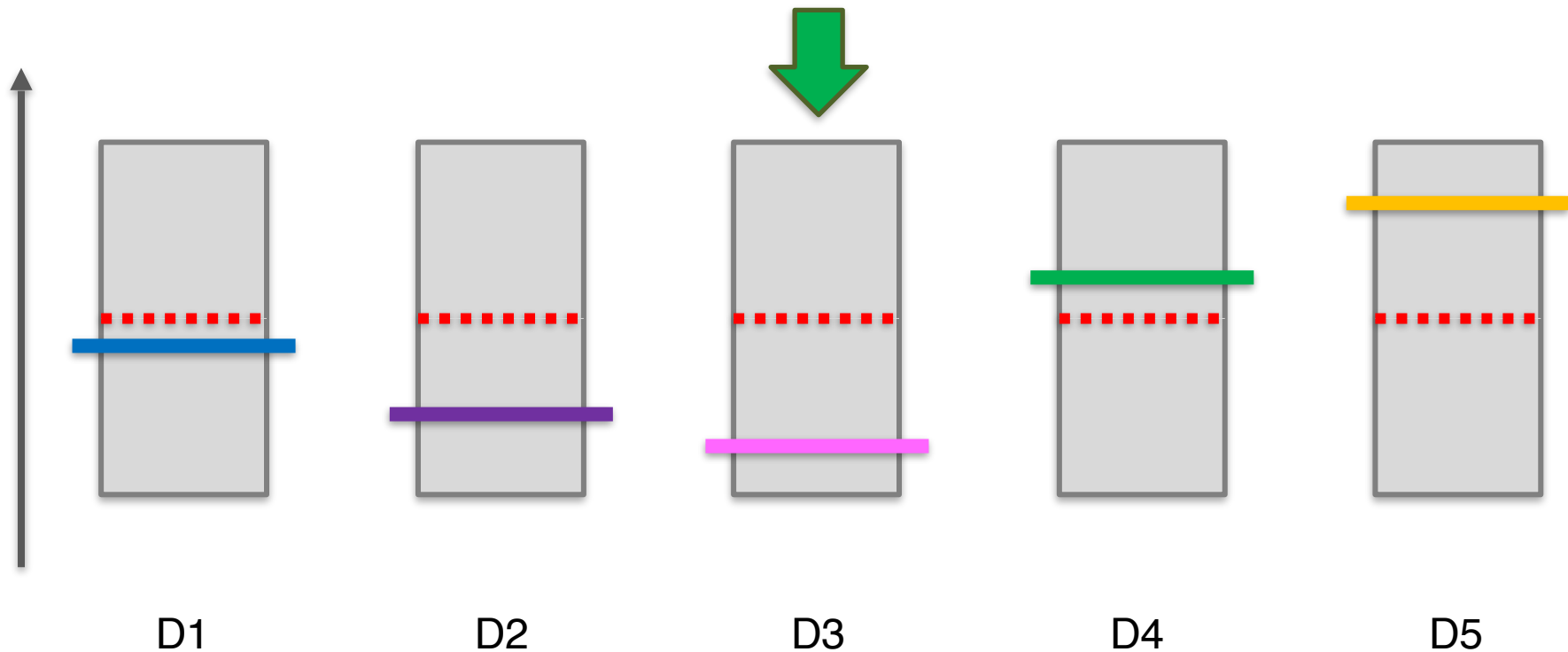
Algoritmo Upper Confidence Bound



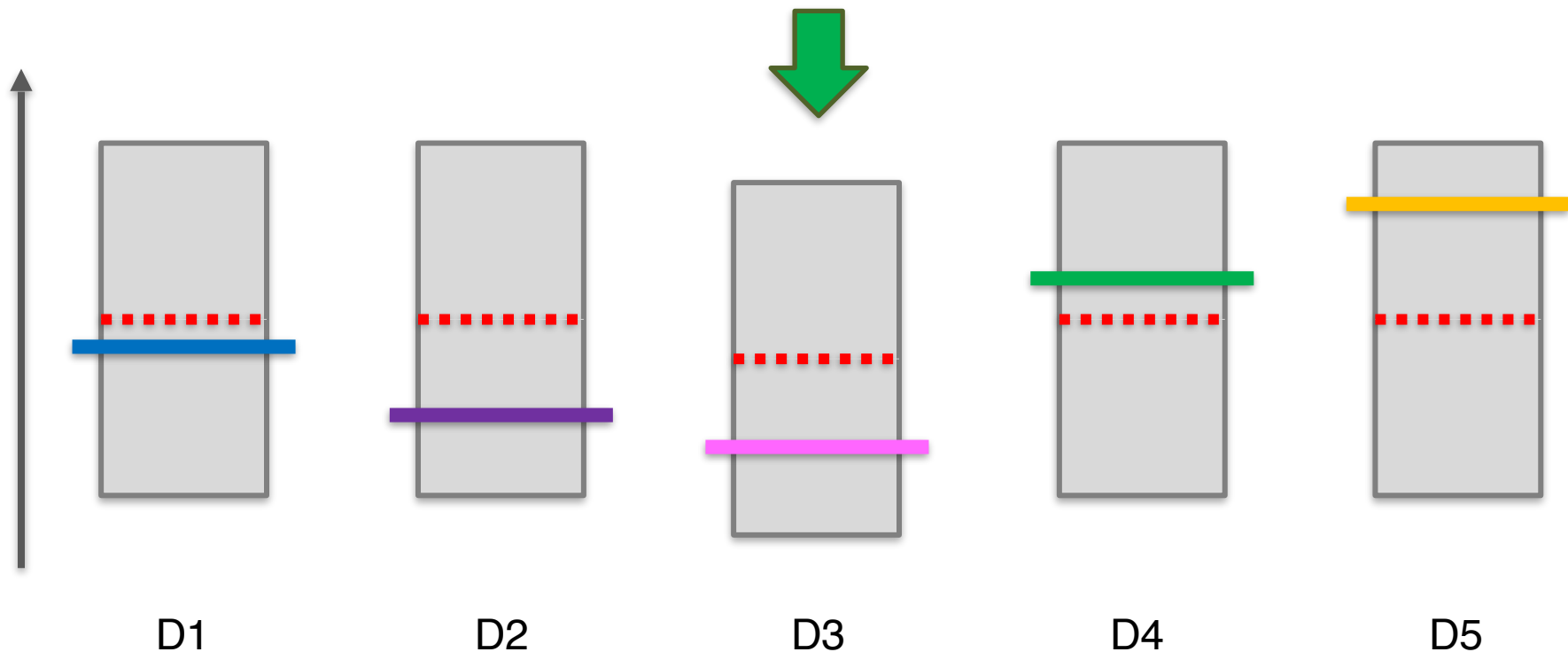
Algoritmo Upper Confidence Bound



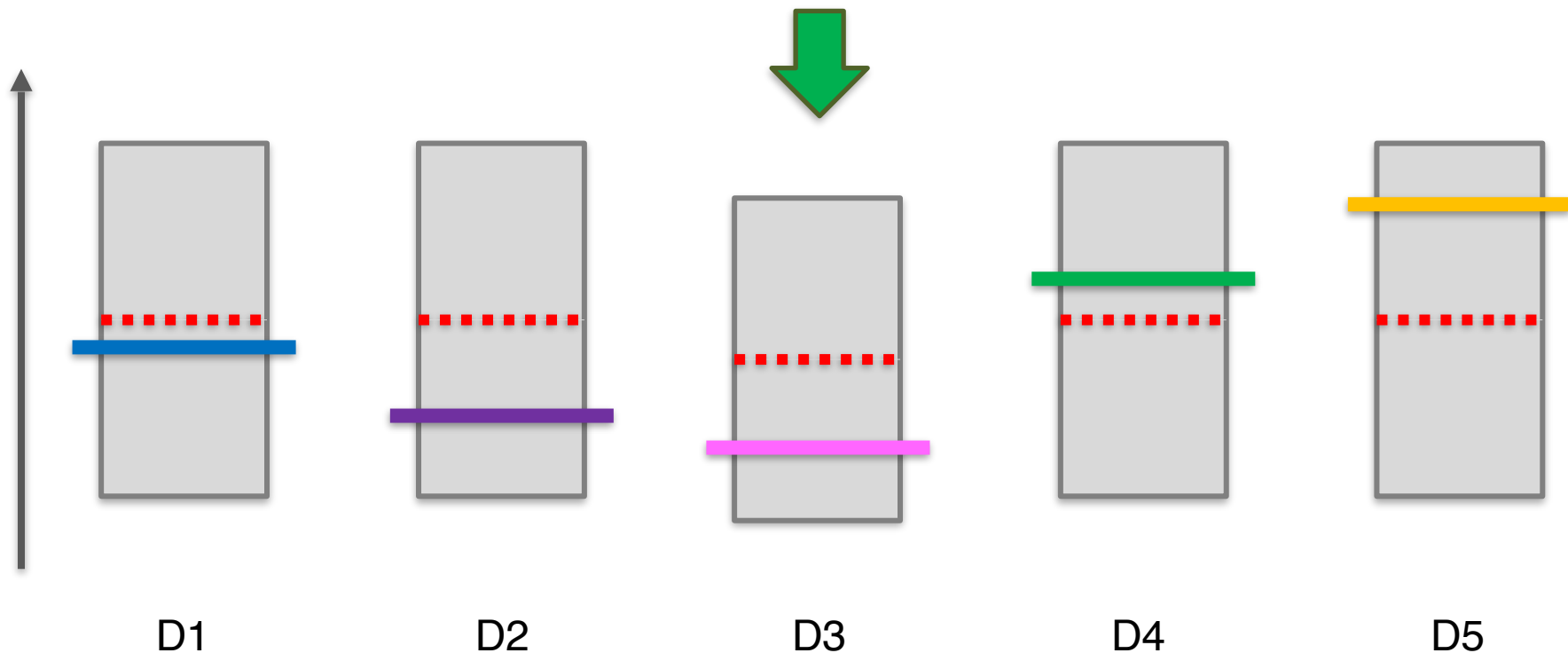
Algoritmo Upper Confidence Bound



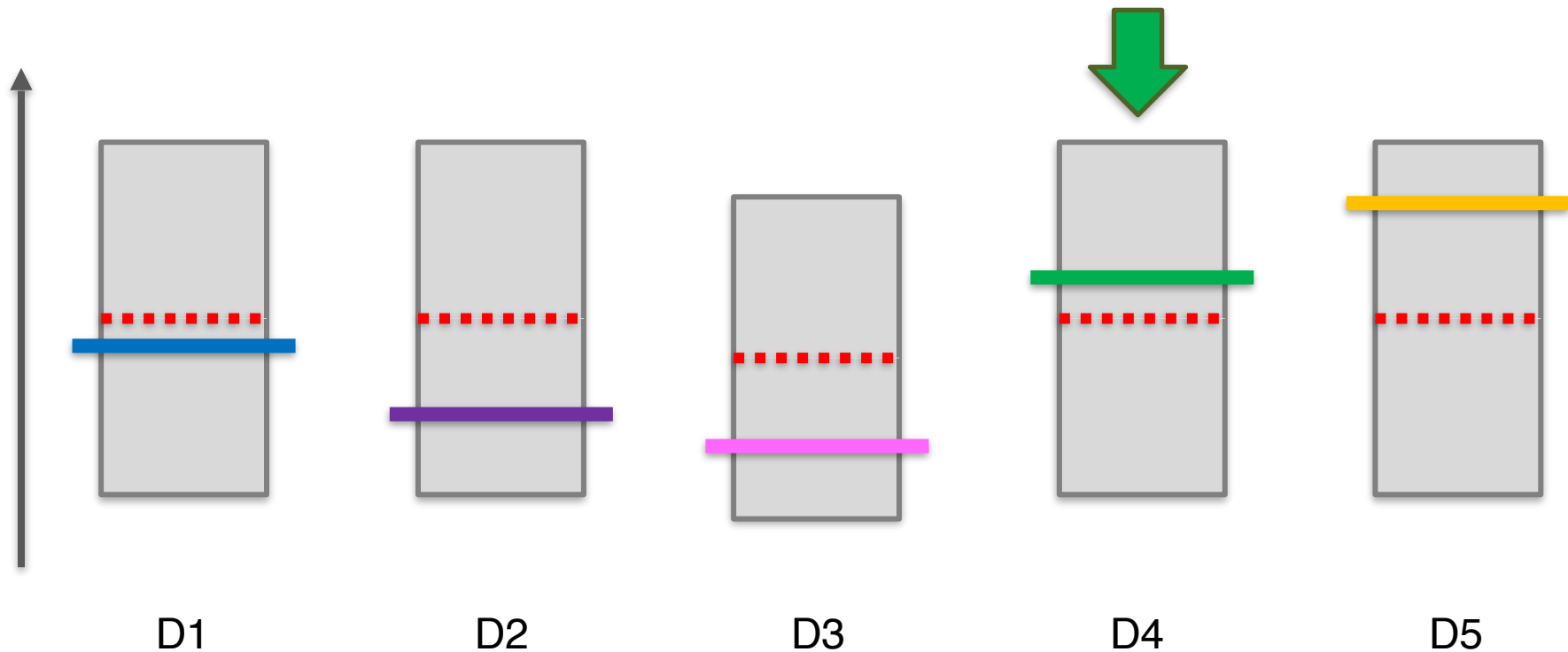
Algoritmo Upper Confidence Bound



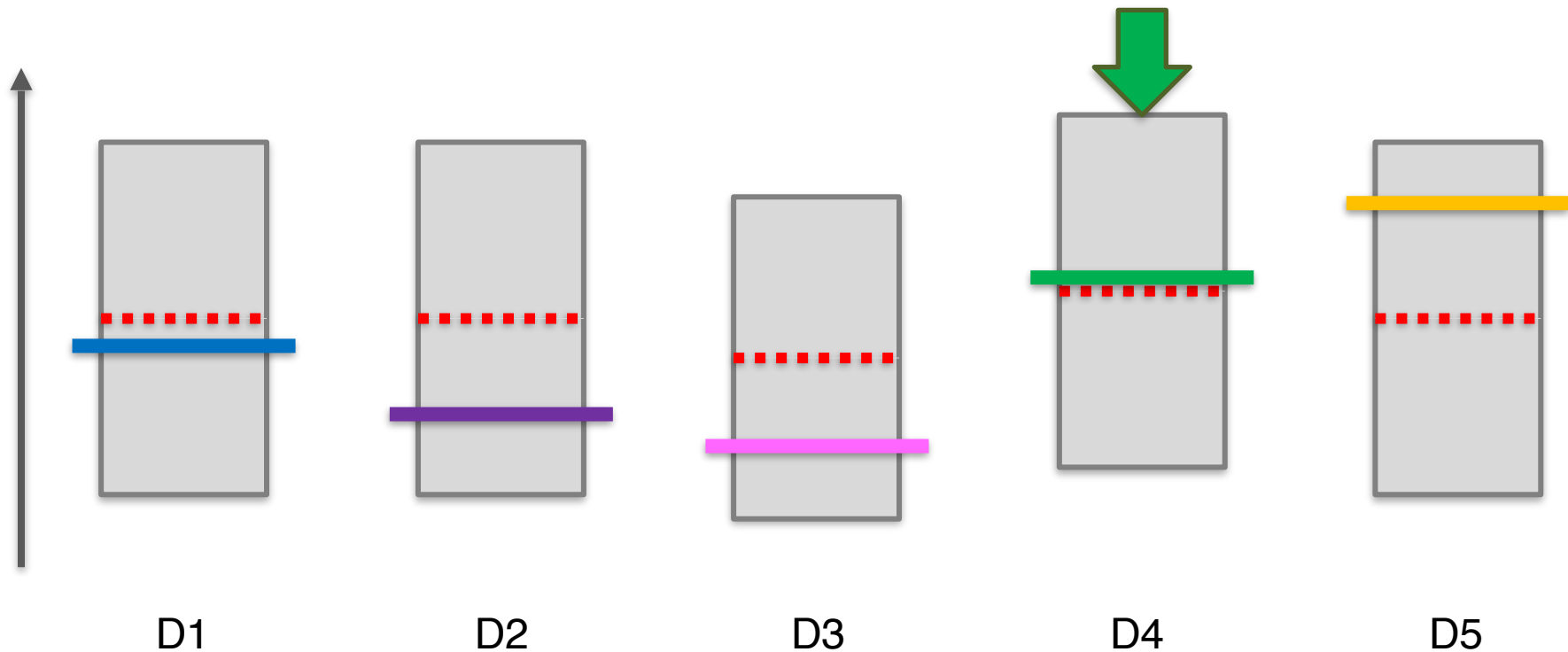
Algoritmo Upper Confidence Bound



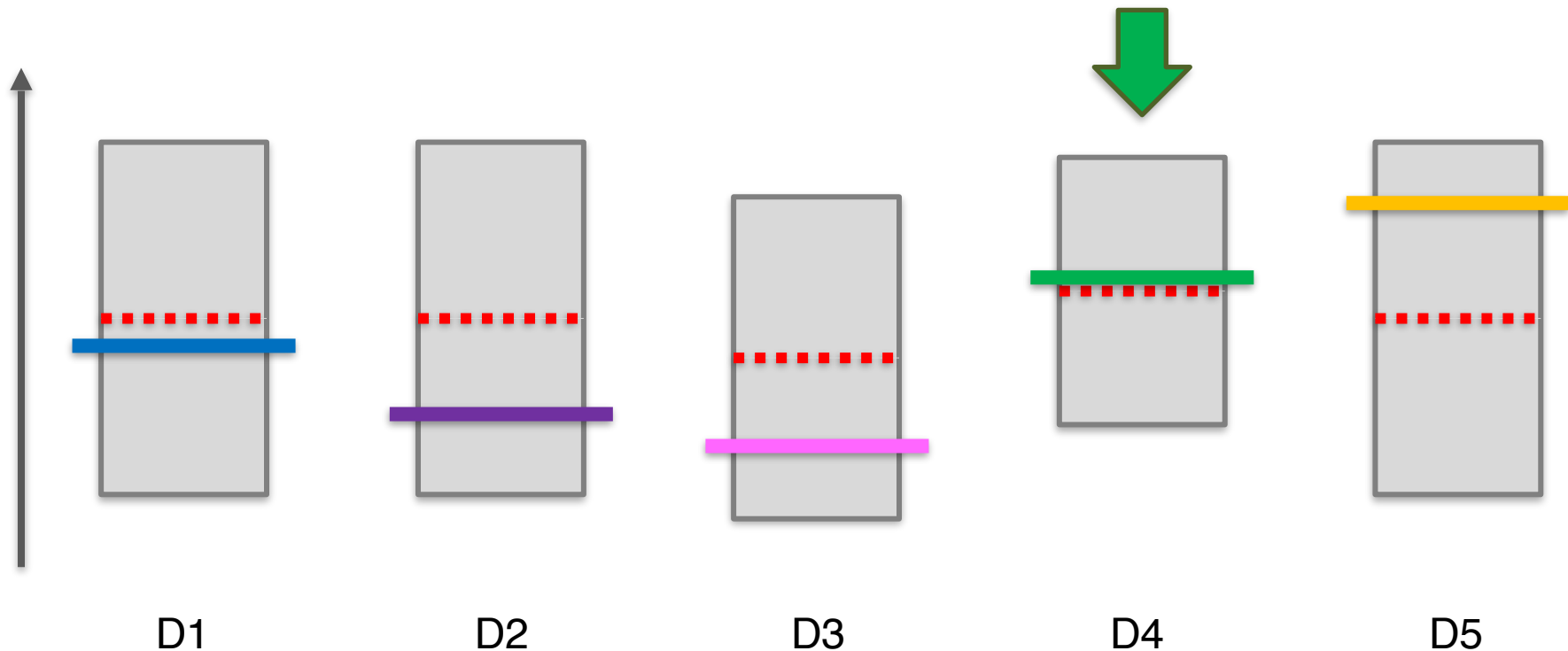
Algoritmo Upper Confidence Bound



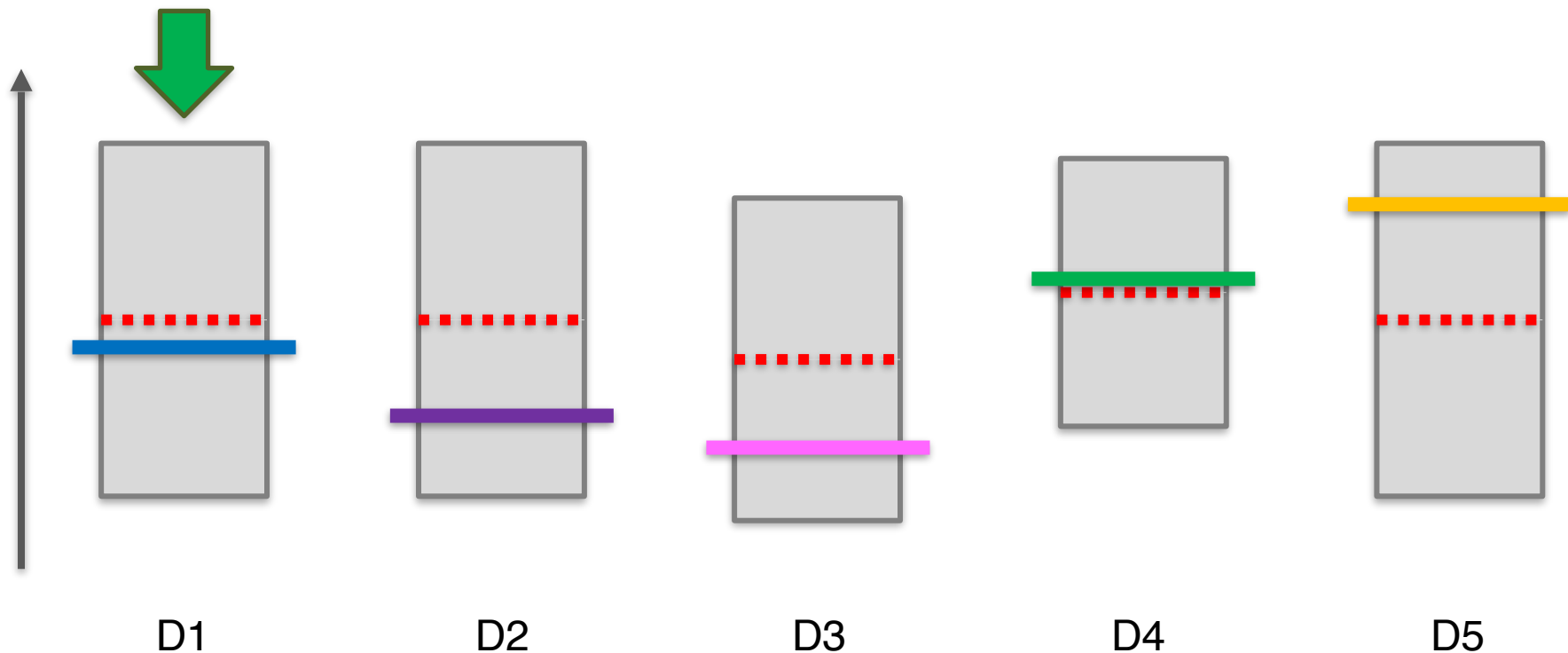
Algoritmo Upper Confidence Bound



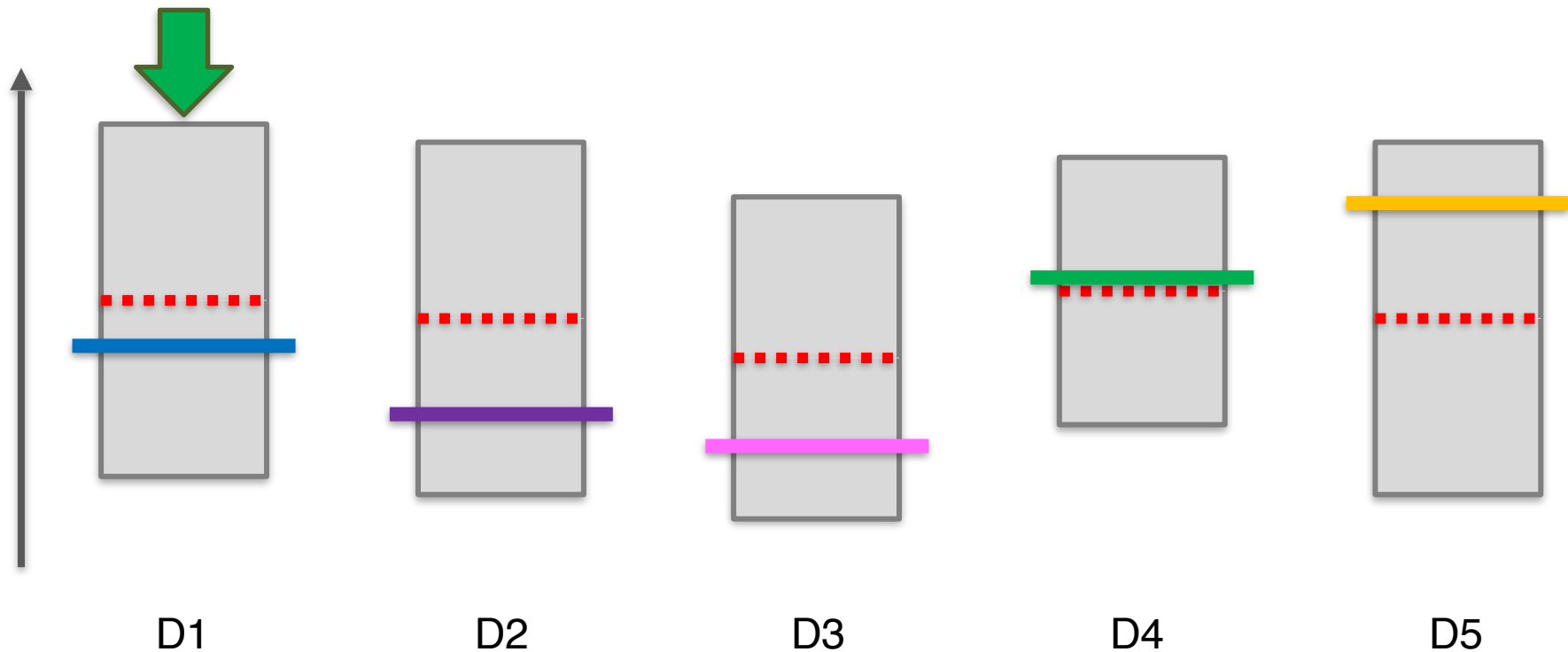
Algoritmo Upper Confidence Bound



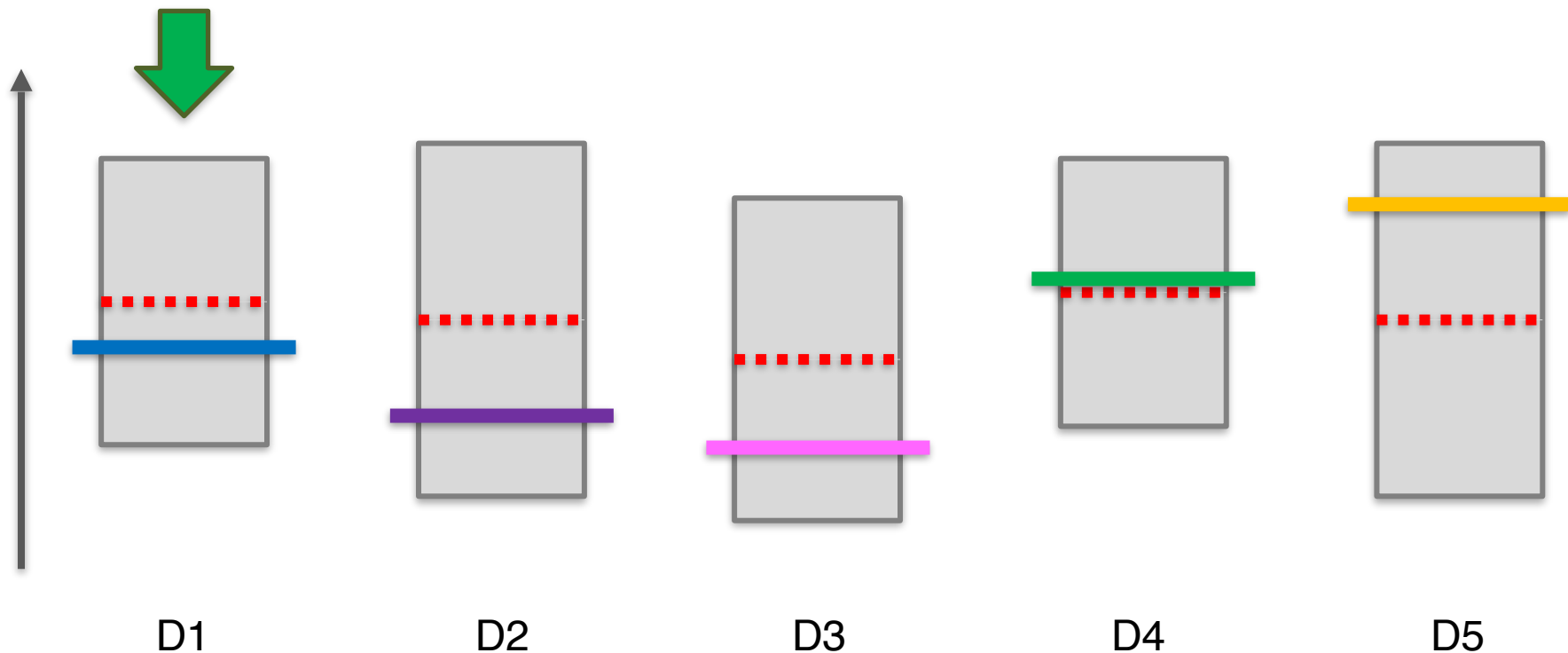
Algoritmo Upper Confidence Bound



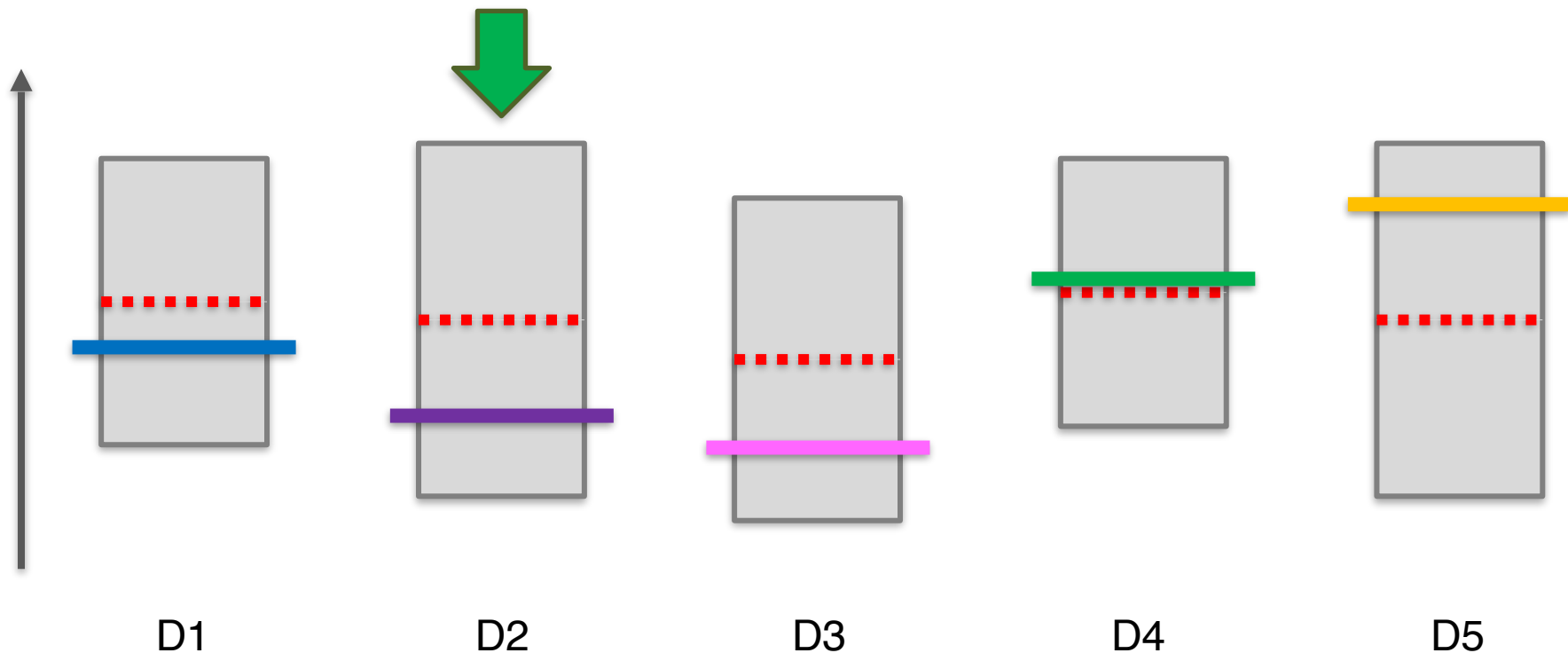
Algoritmo Upper Confidence Bound



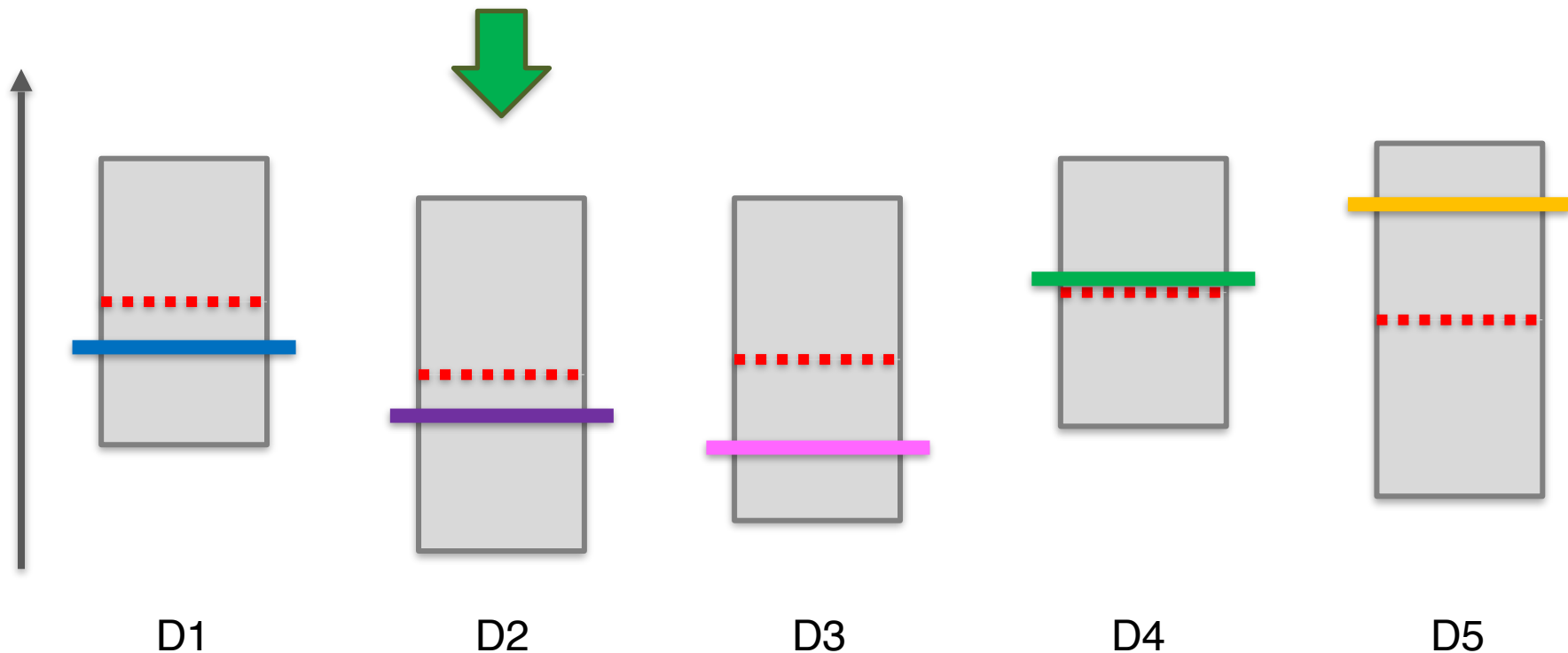
Algoritmo Upper Confidence Bound



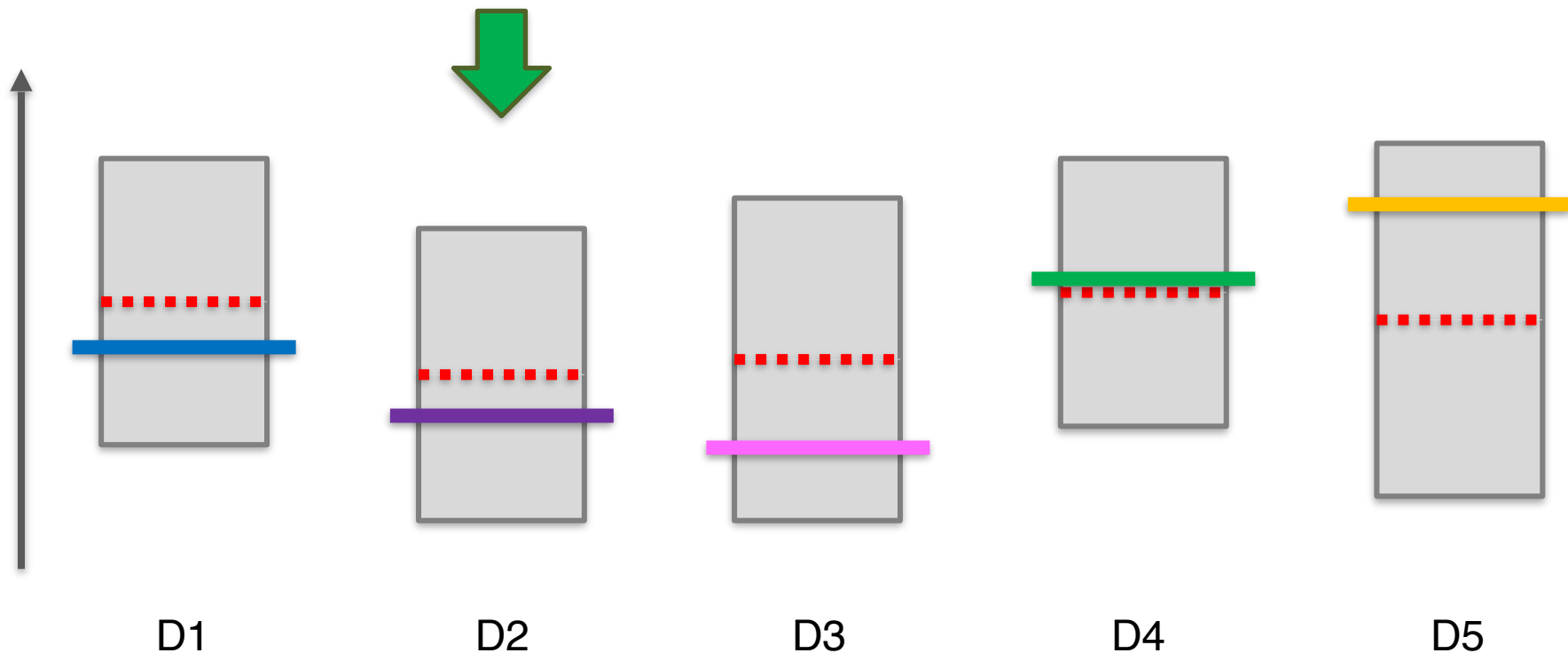
Algoritmo Upper Confidence Bound



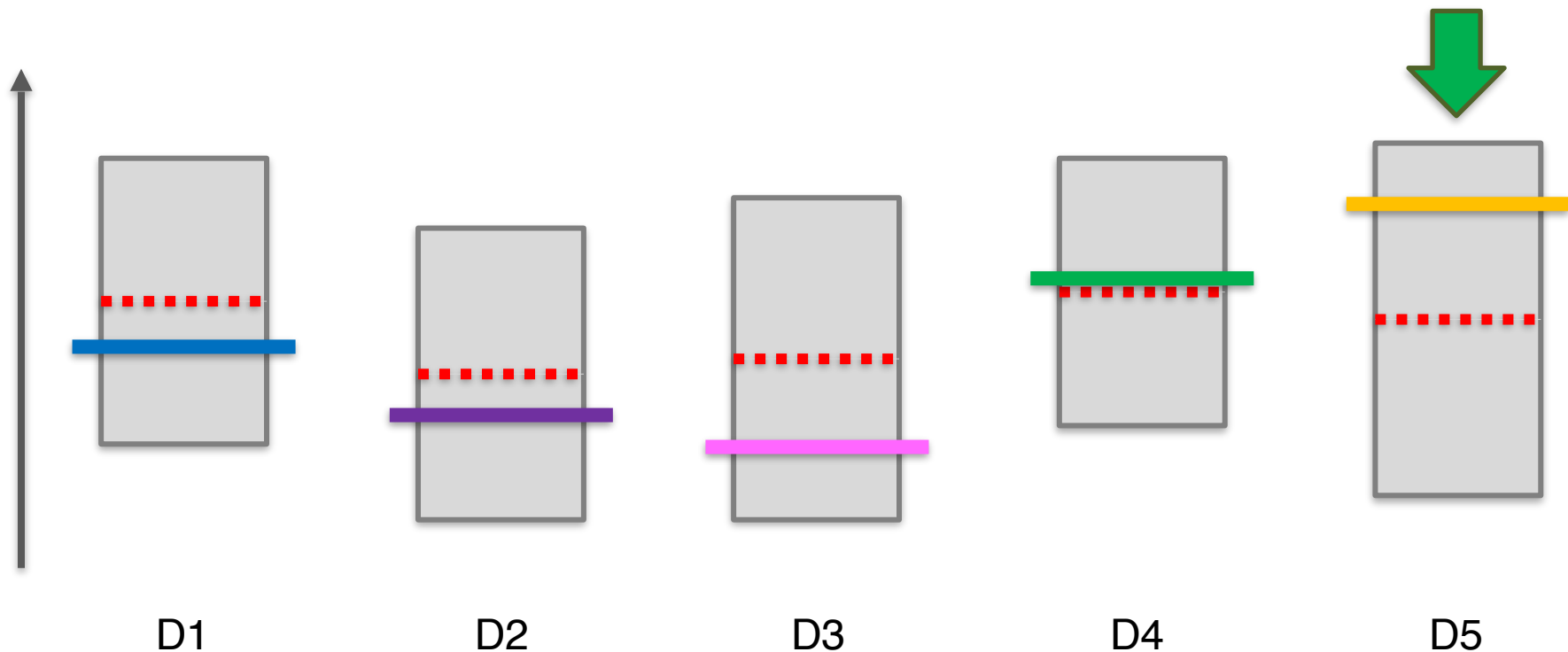
Algoritmo Upper Confidence Bound



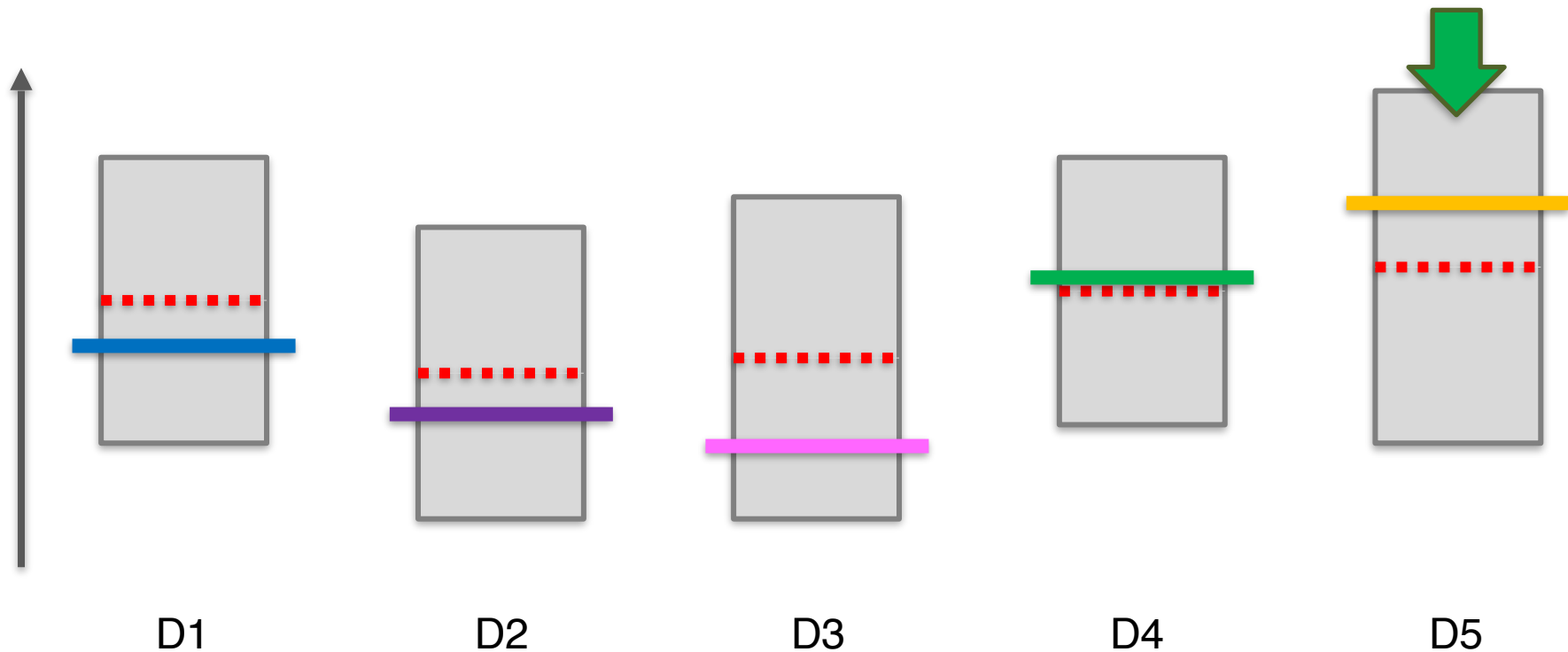
Algoritmo Upper Confidence Bound



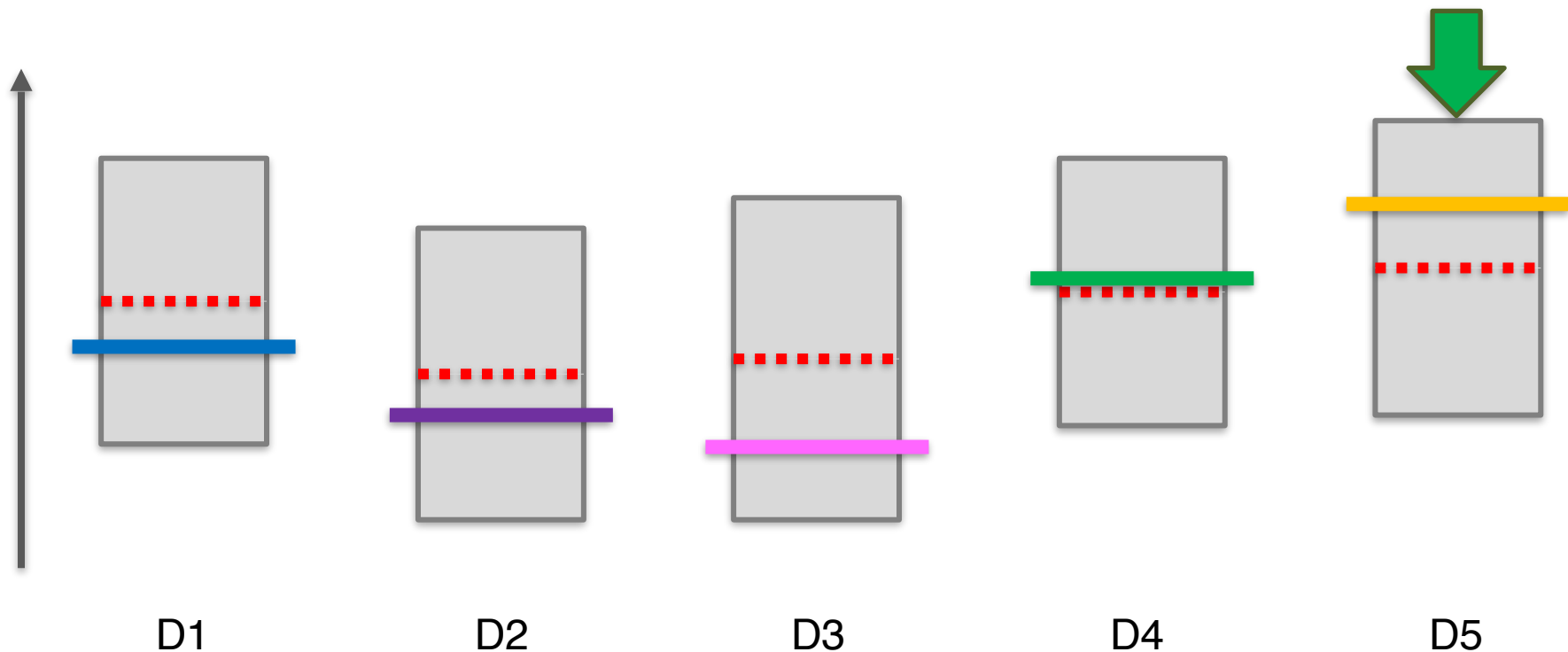
Algoritmo Upper Confidence Bound



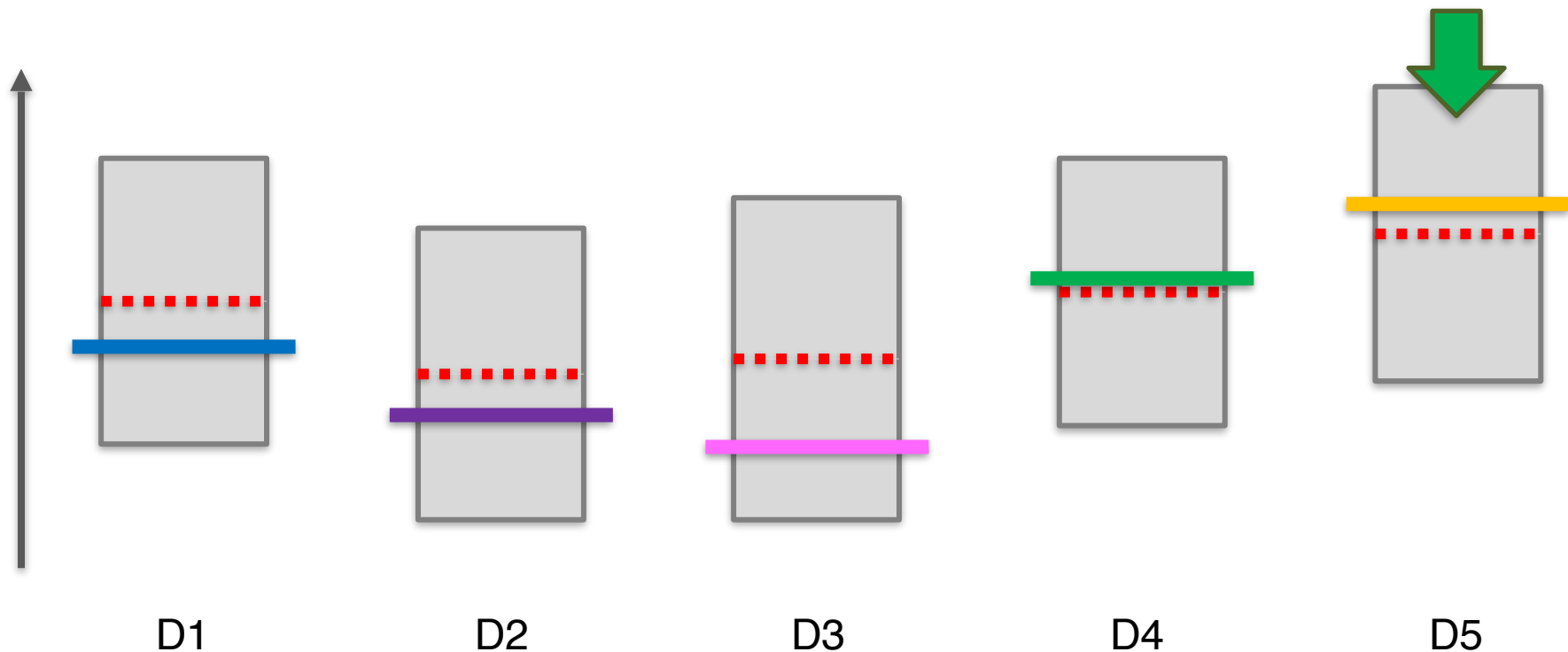
Algoritmo Upper Confidence Bound



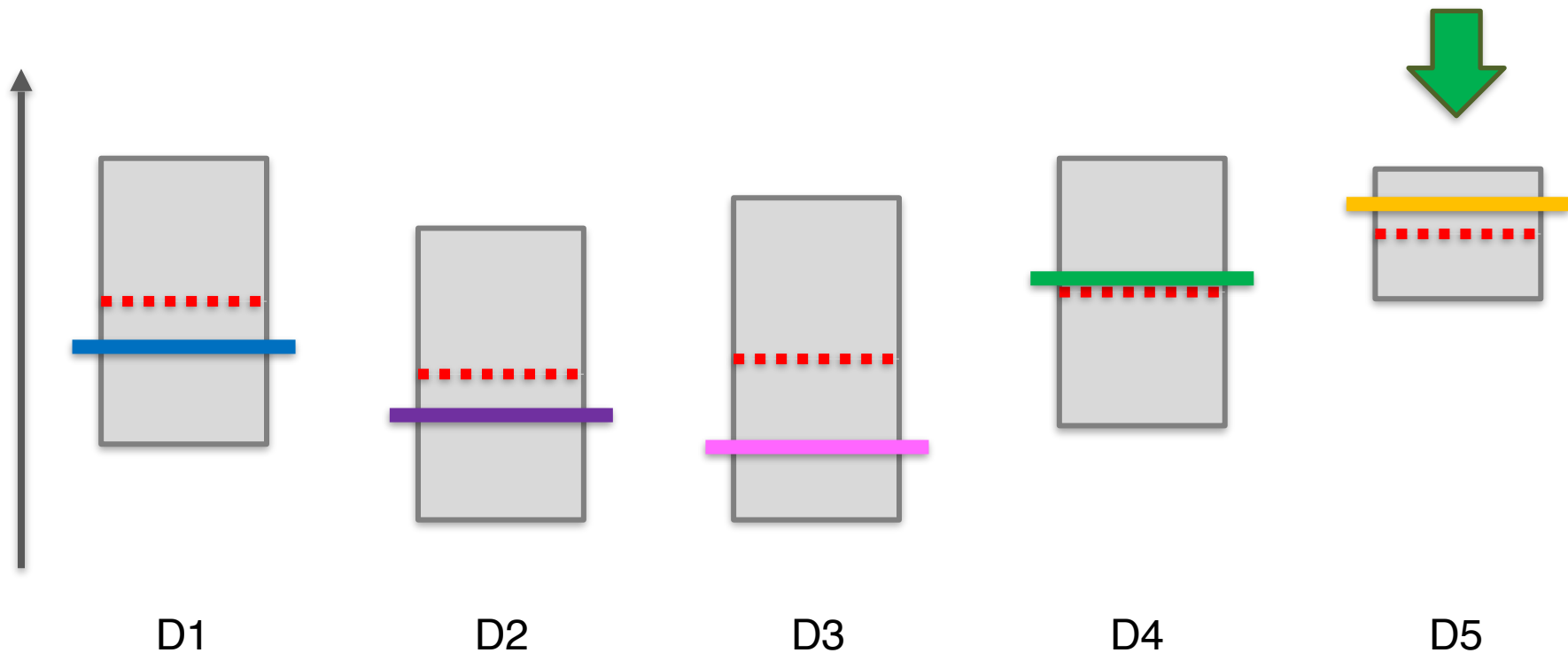
Algoritmo Upper Confidence Bound



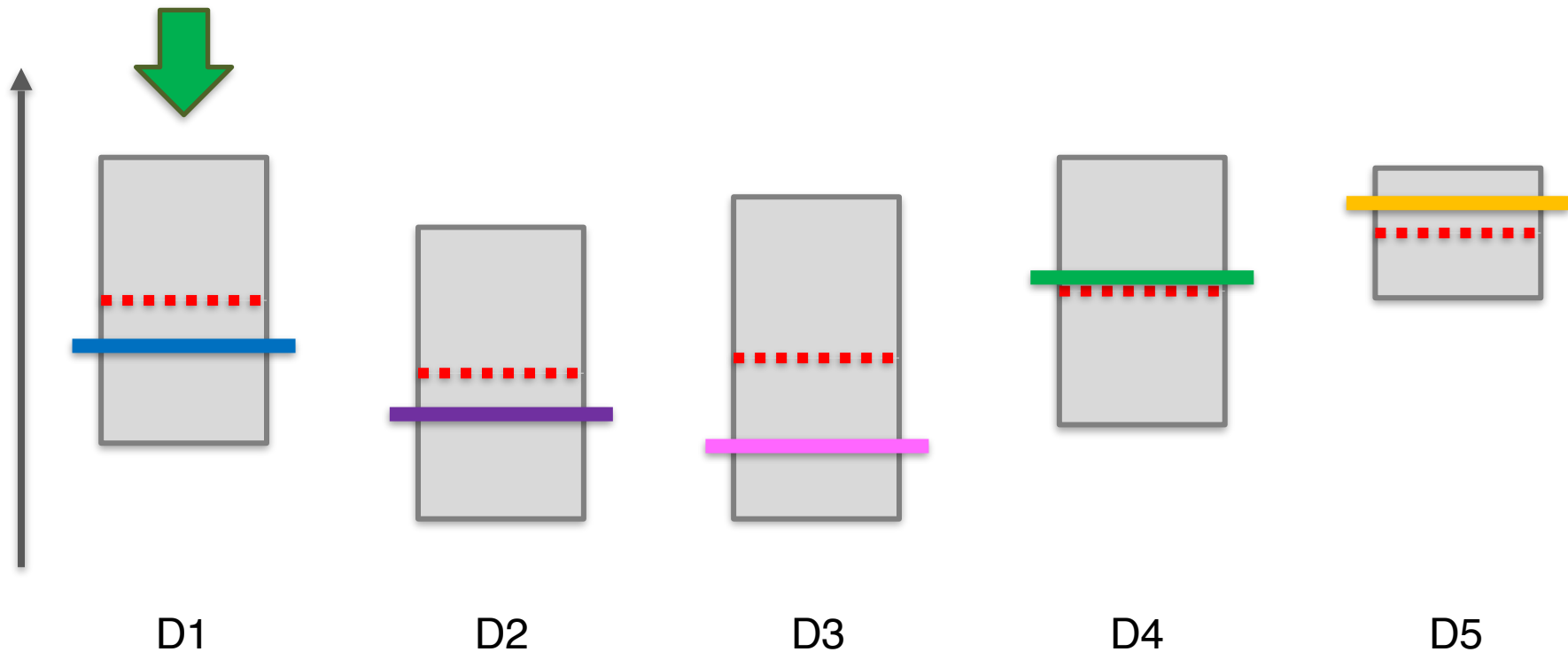
Algoritmo Upper Confidence Bound



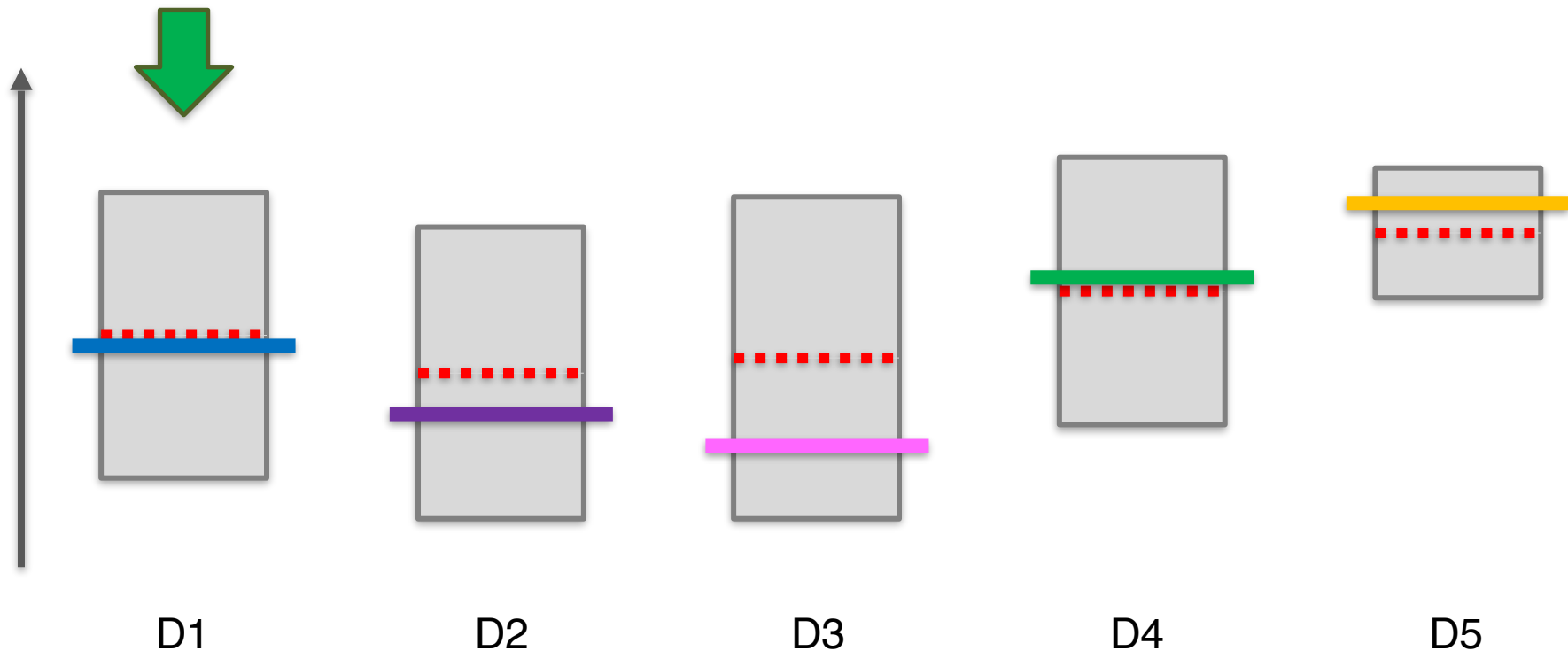
Algoritmo Upper Confidence Bound



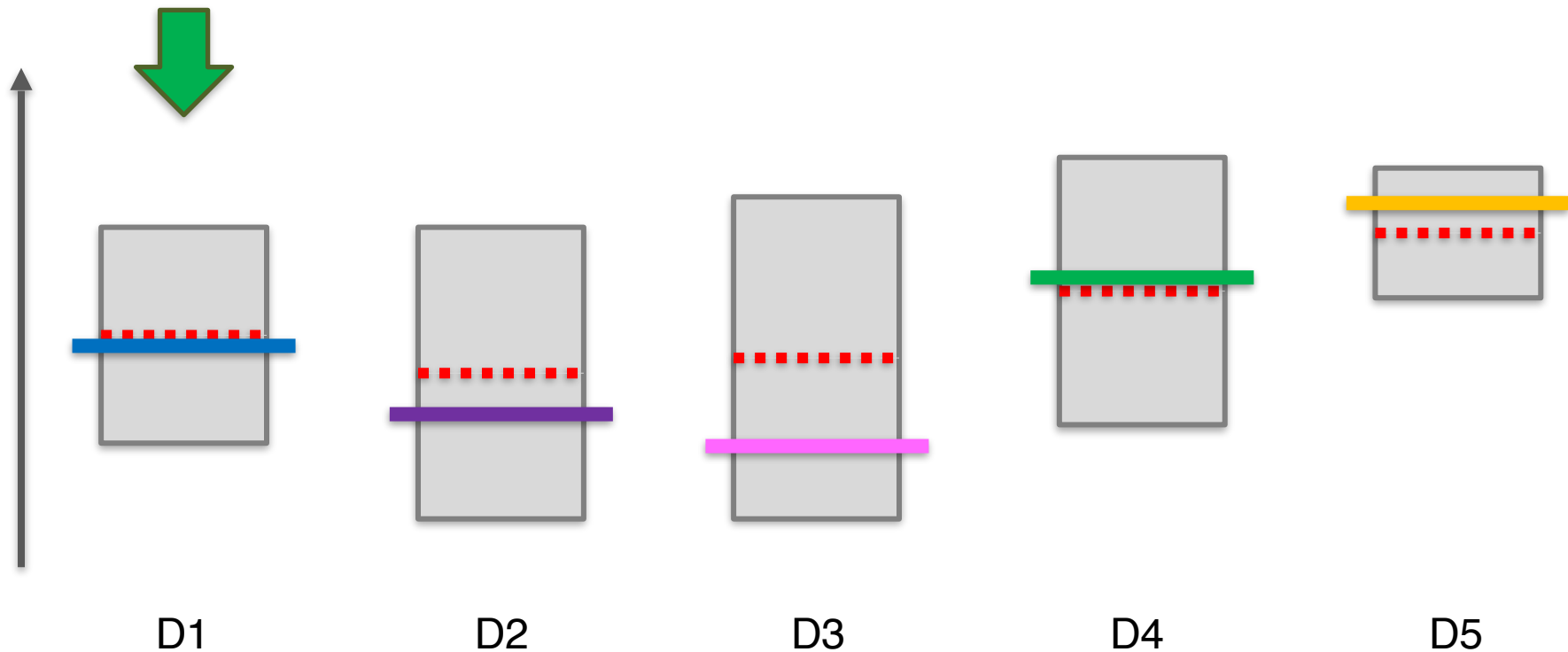
Algoritmo Upper Confidence Bound



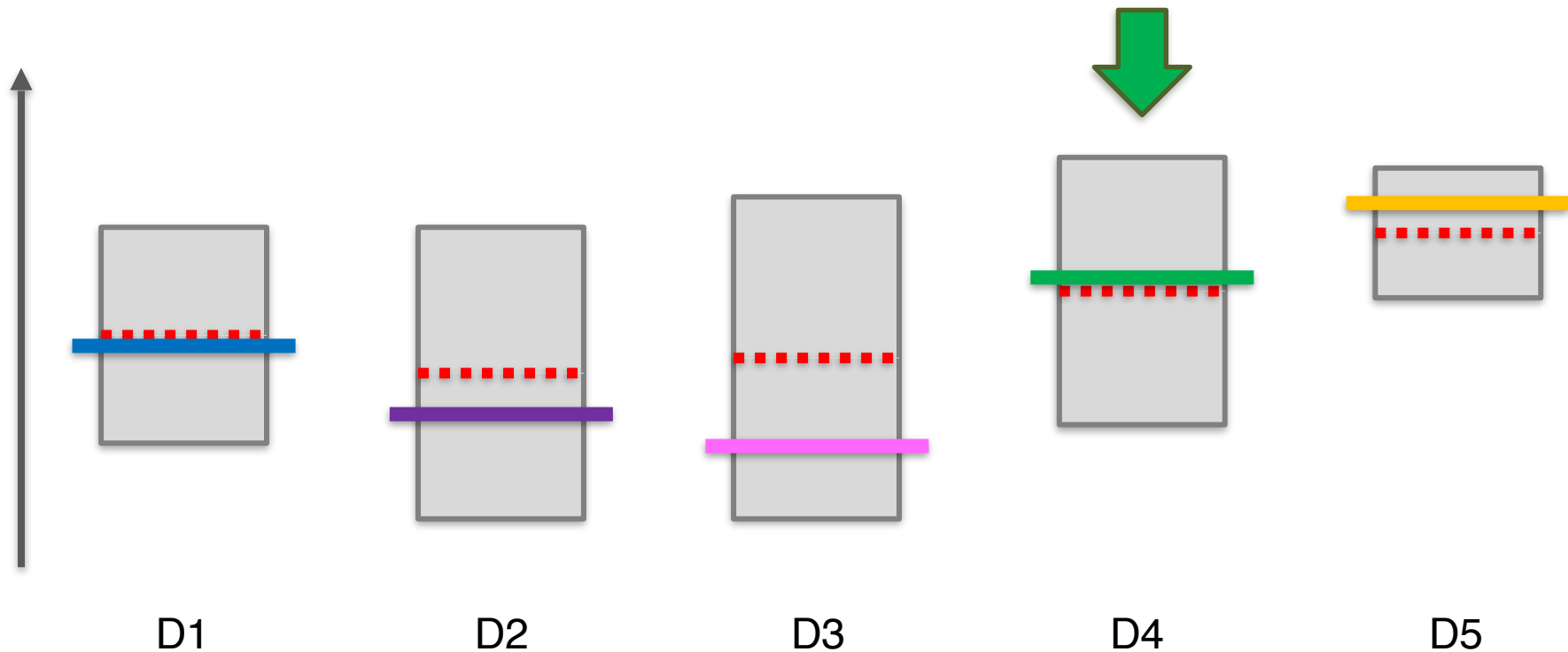
Algoritmo Upper Confidence Bound



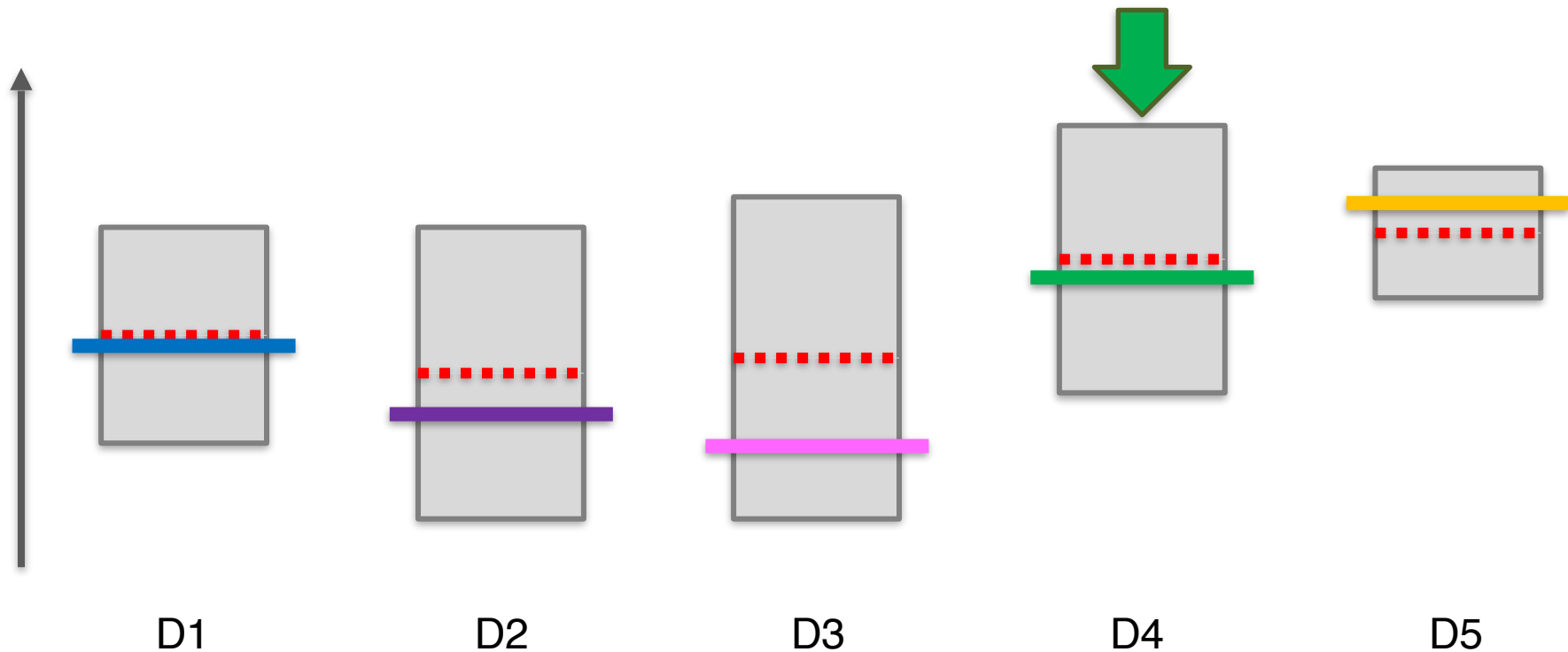
Algoritmo Upper Confidence Bound



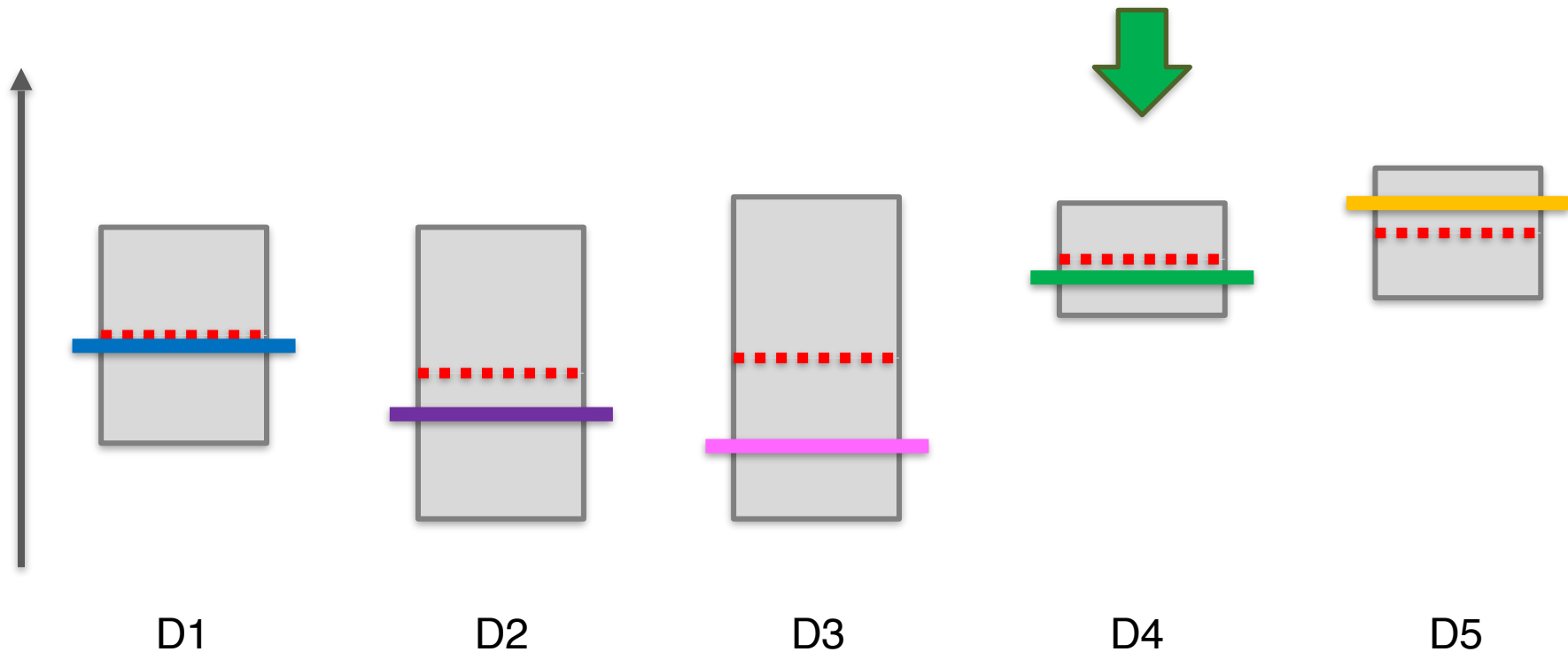
Algoritmo Upper Confidence Bound



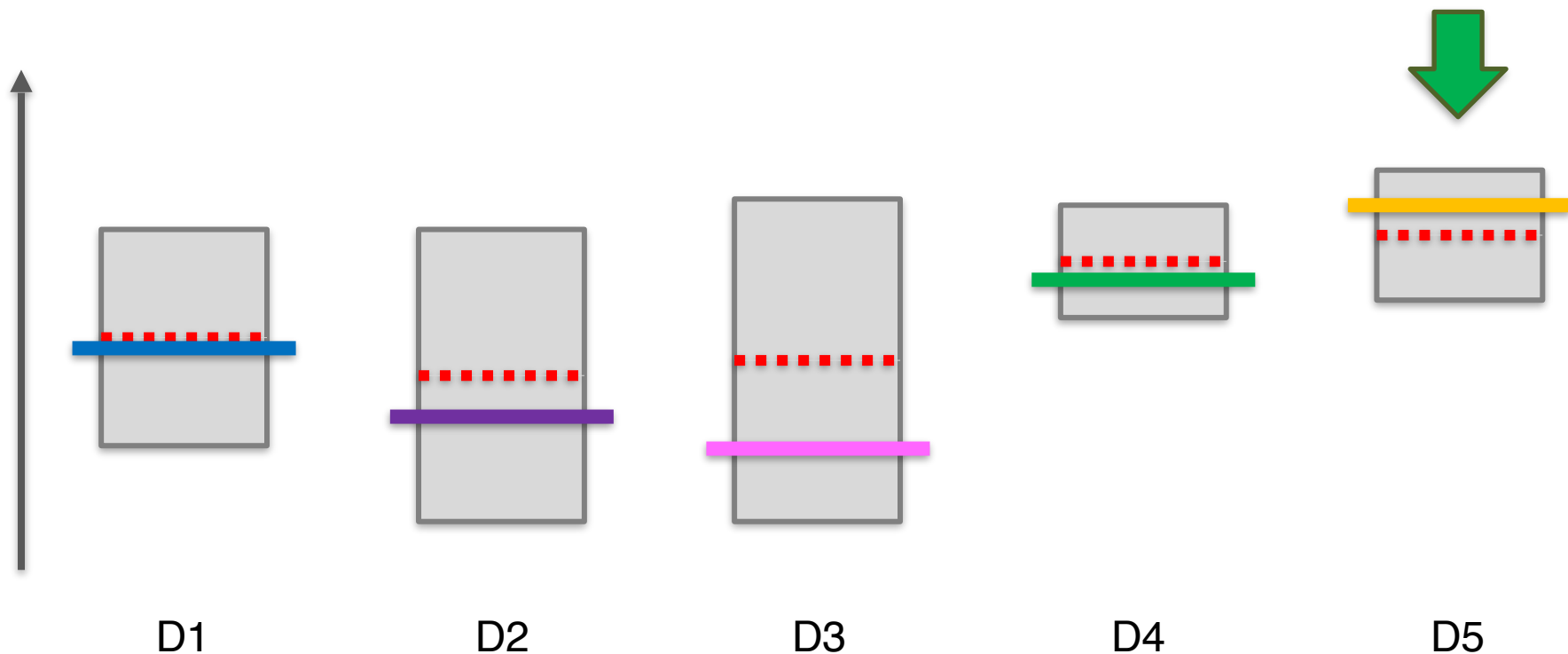
Algoritmo Upper Confidence Bound



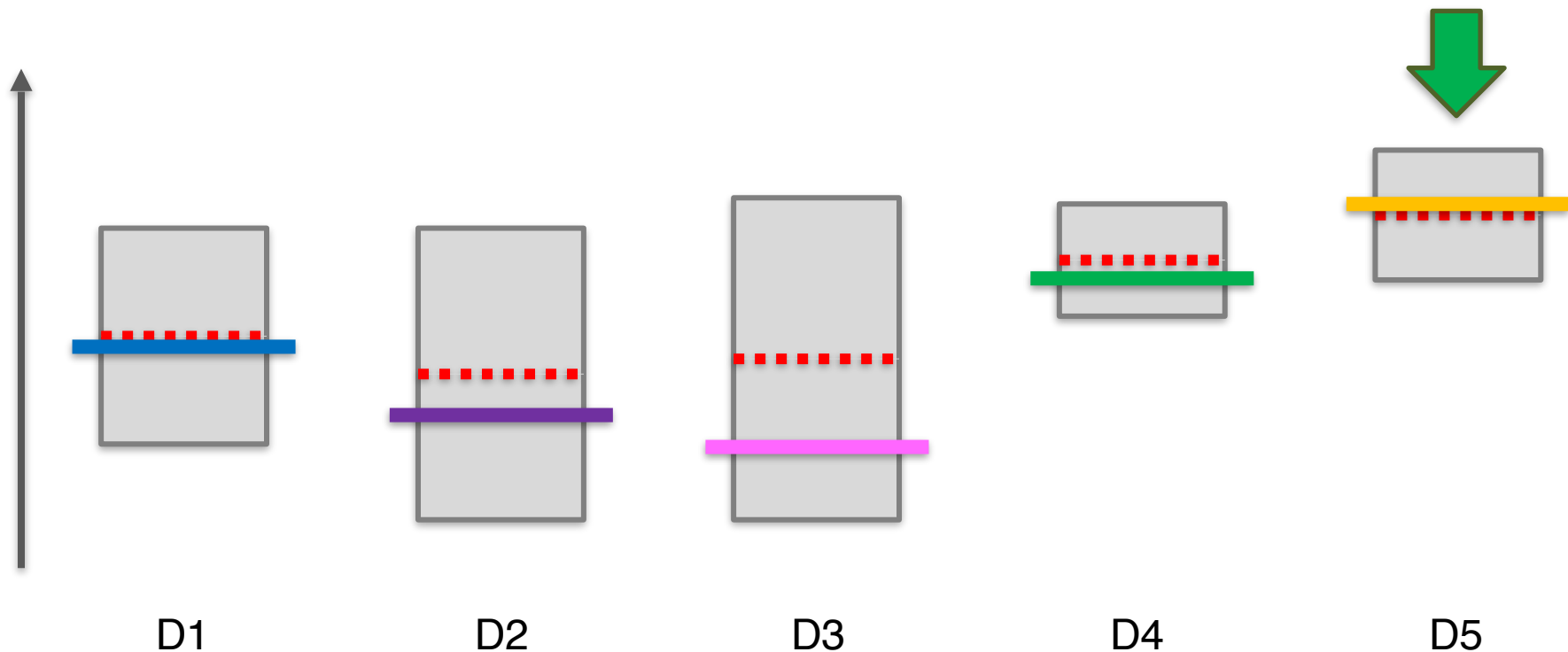
Algoritmo Upper Confidence Bound



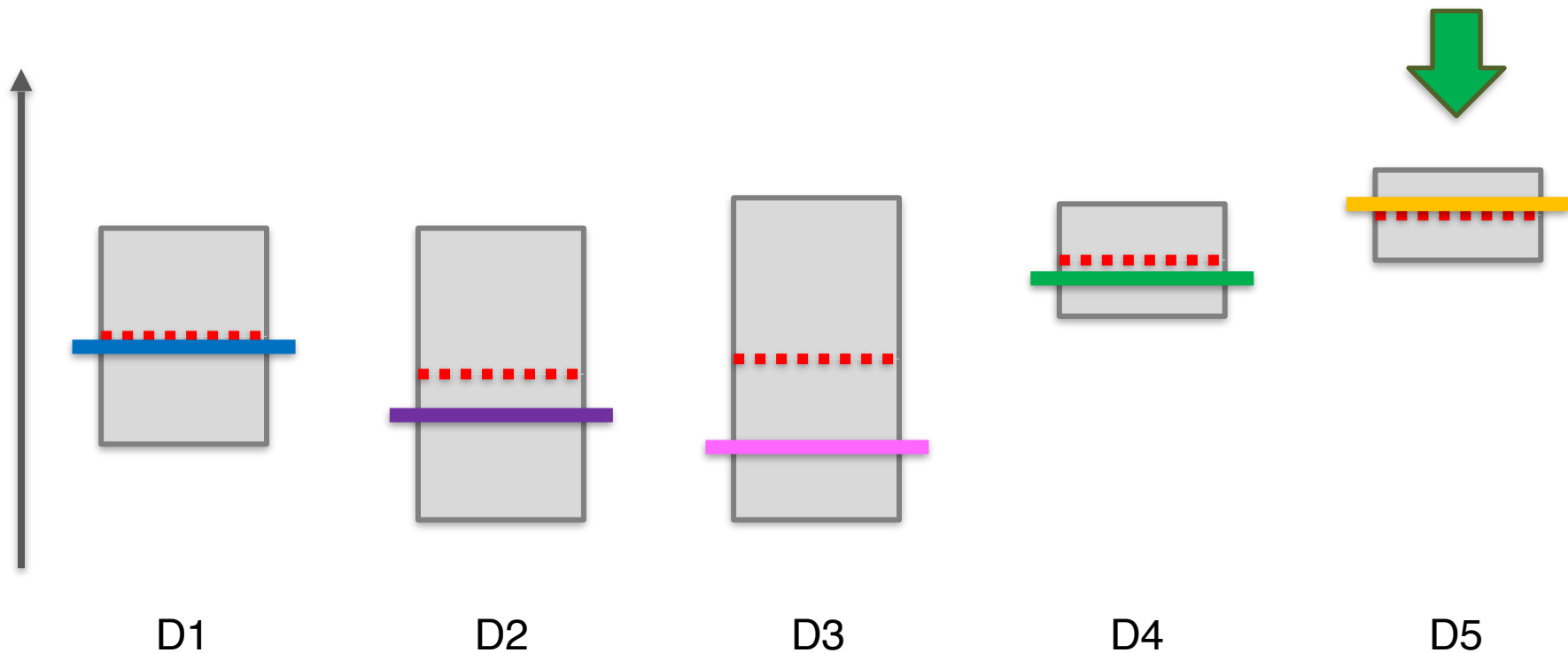
Algoritmo Upper Confidence Bound



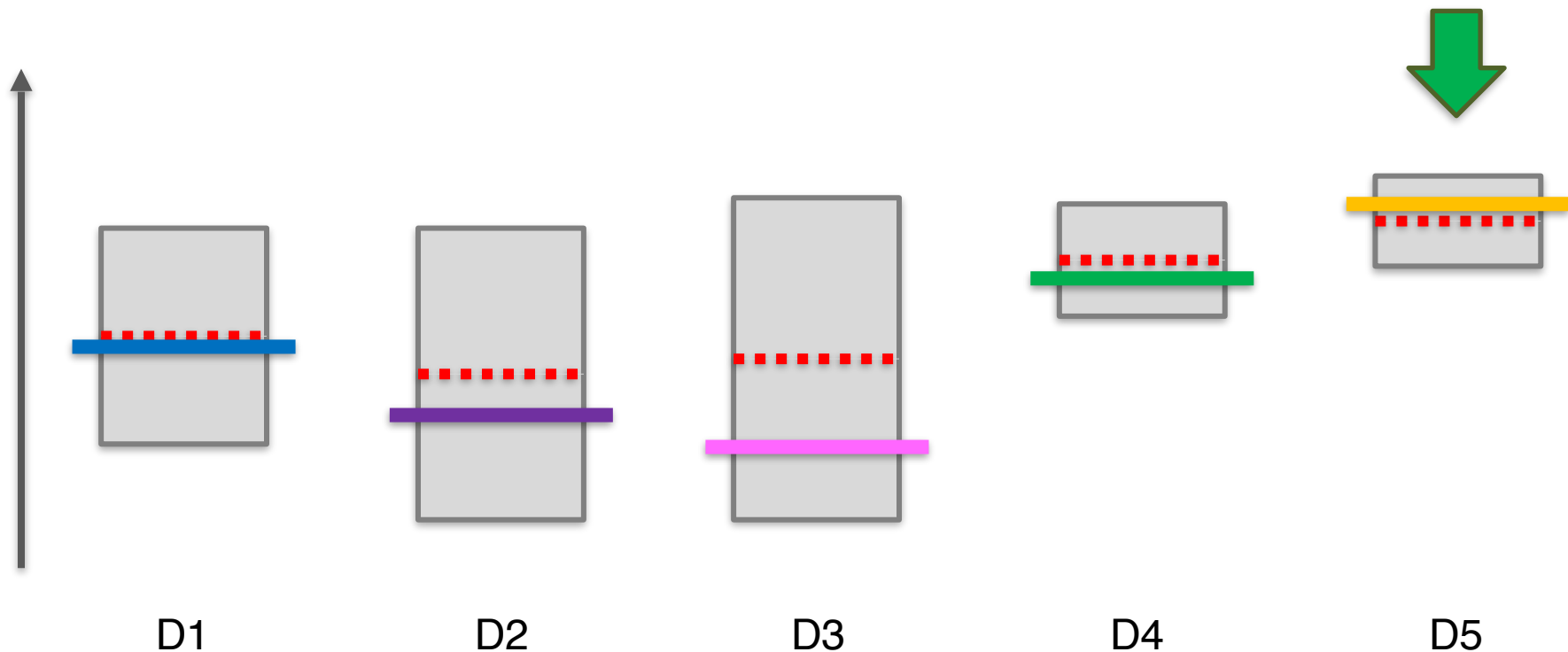
Algoritmo Upper Confidence Bound



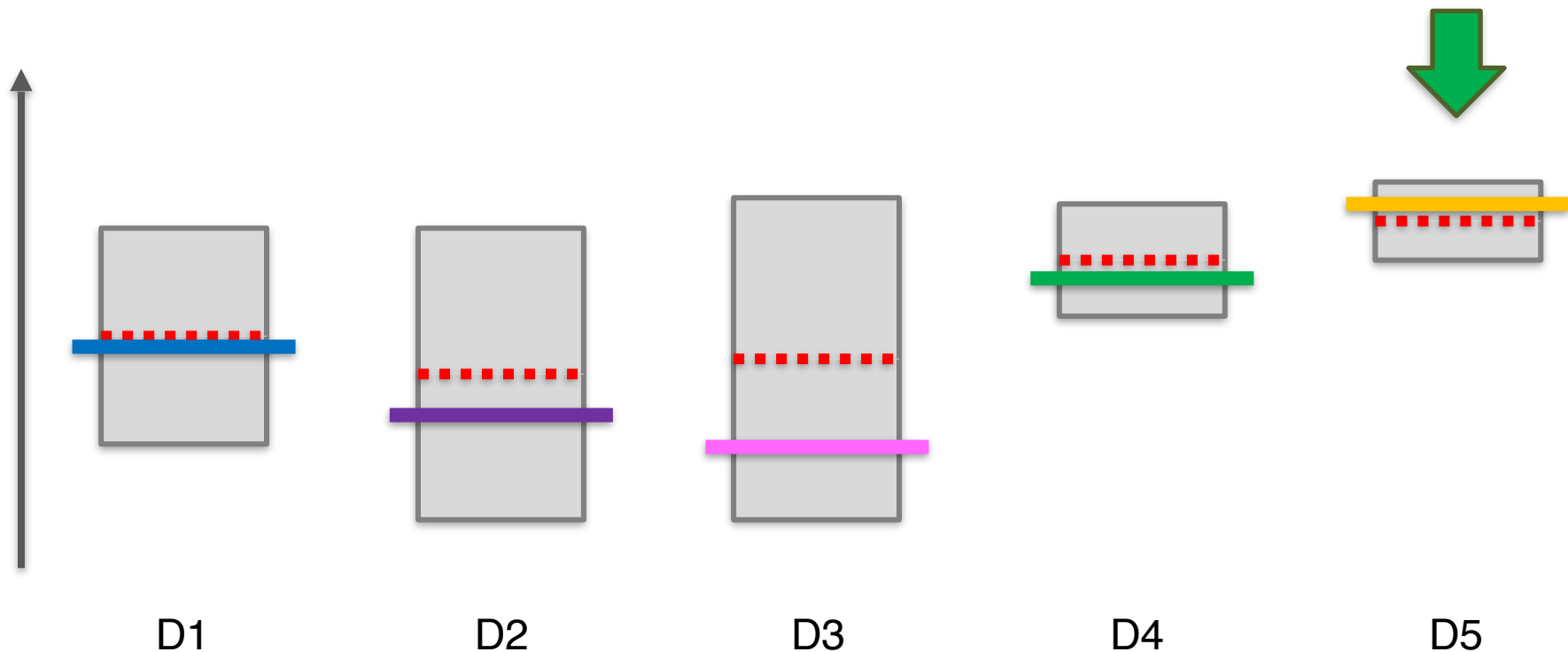
Algoritmo Upper Confidence Bound



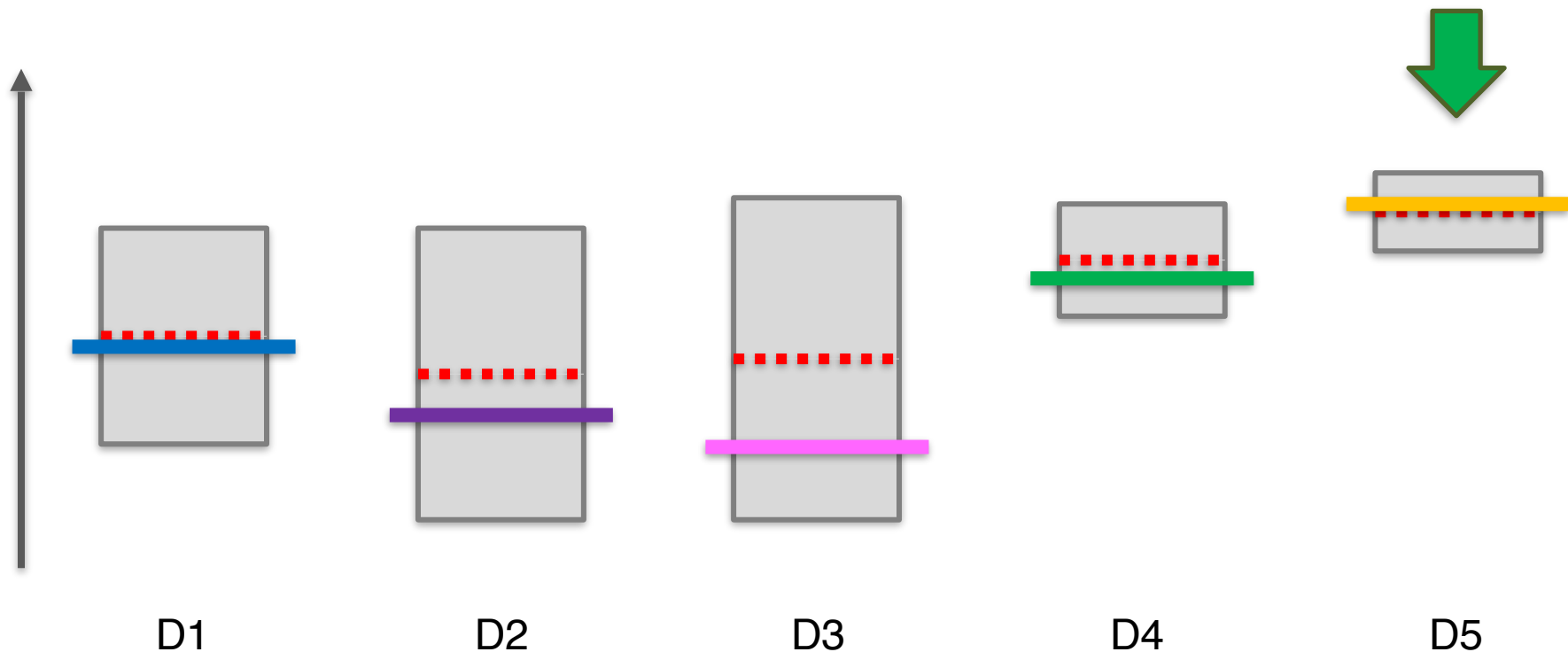
Algoritmo Upper Confidence Bound



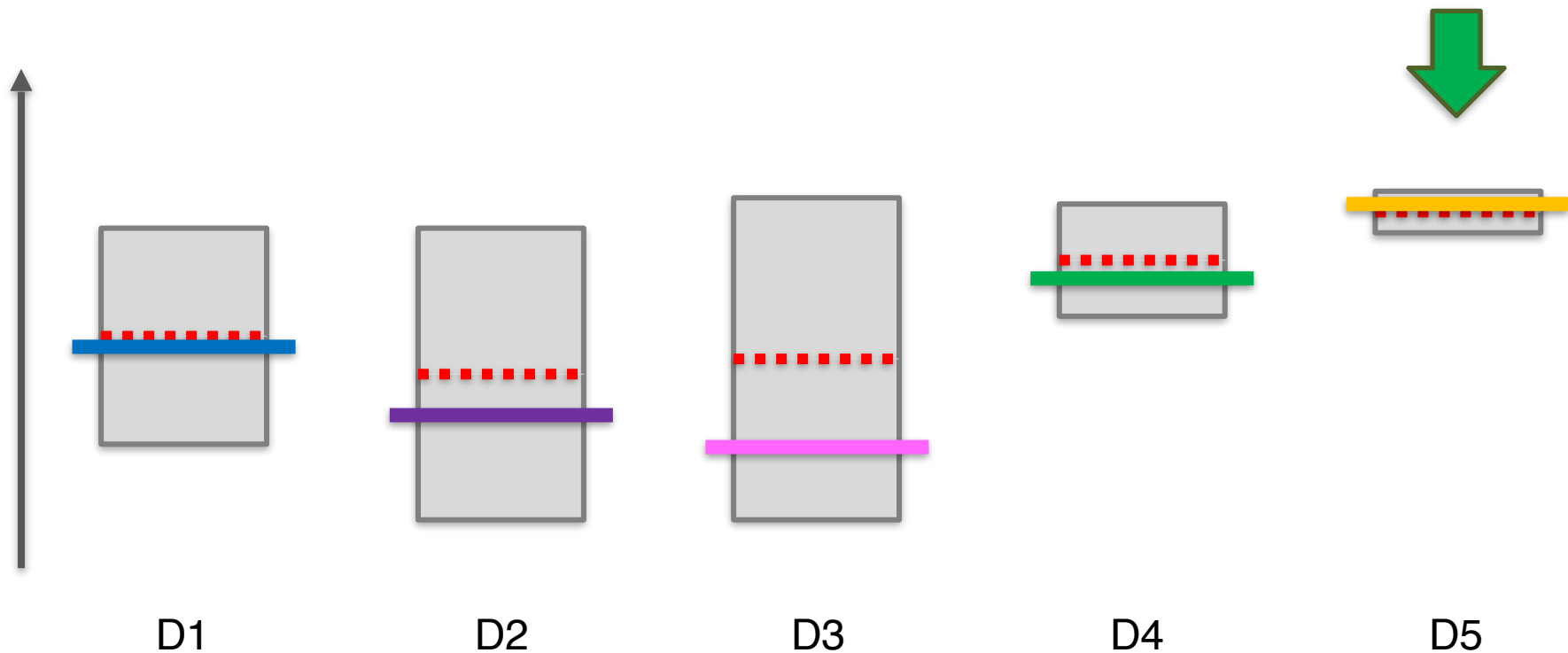
Algoritmo Upper Confidence Bound



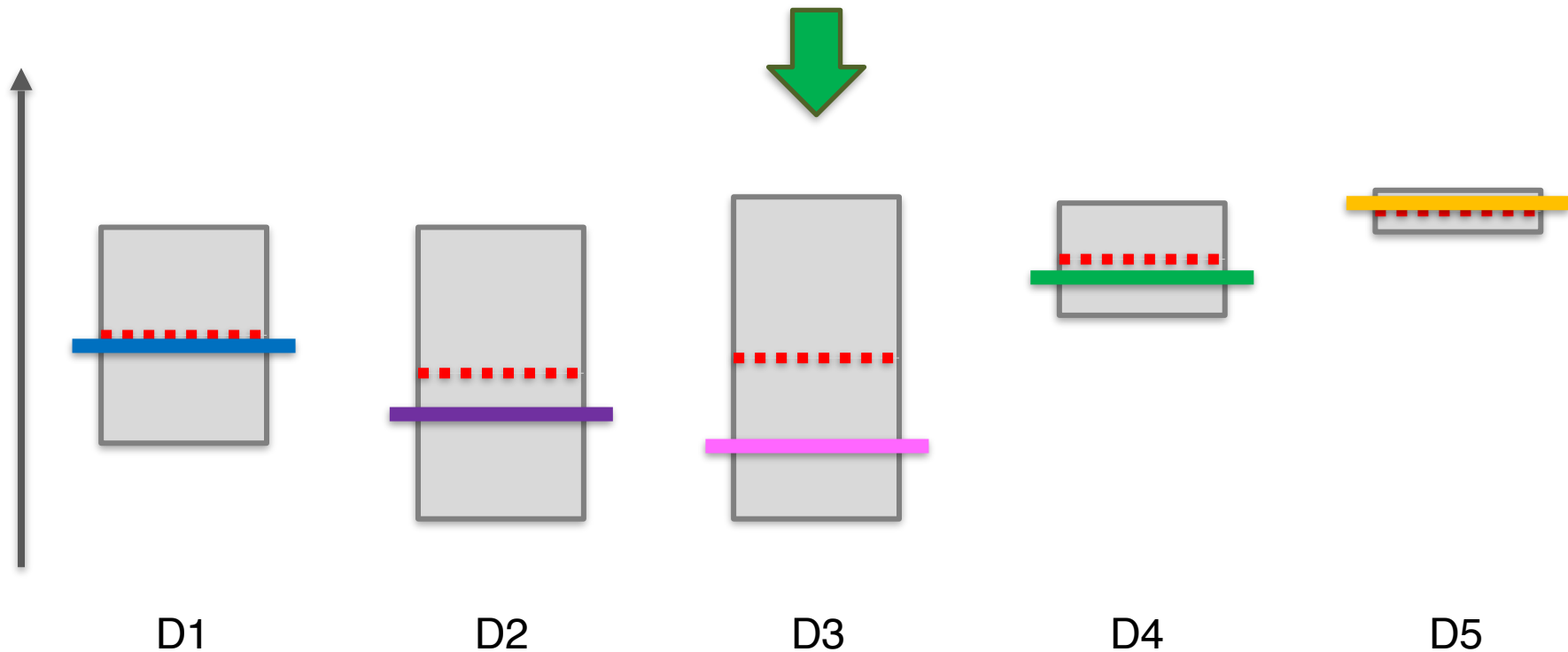
Algoritmo Upper Confidence Bound



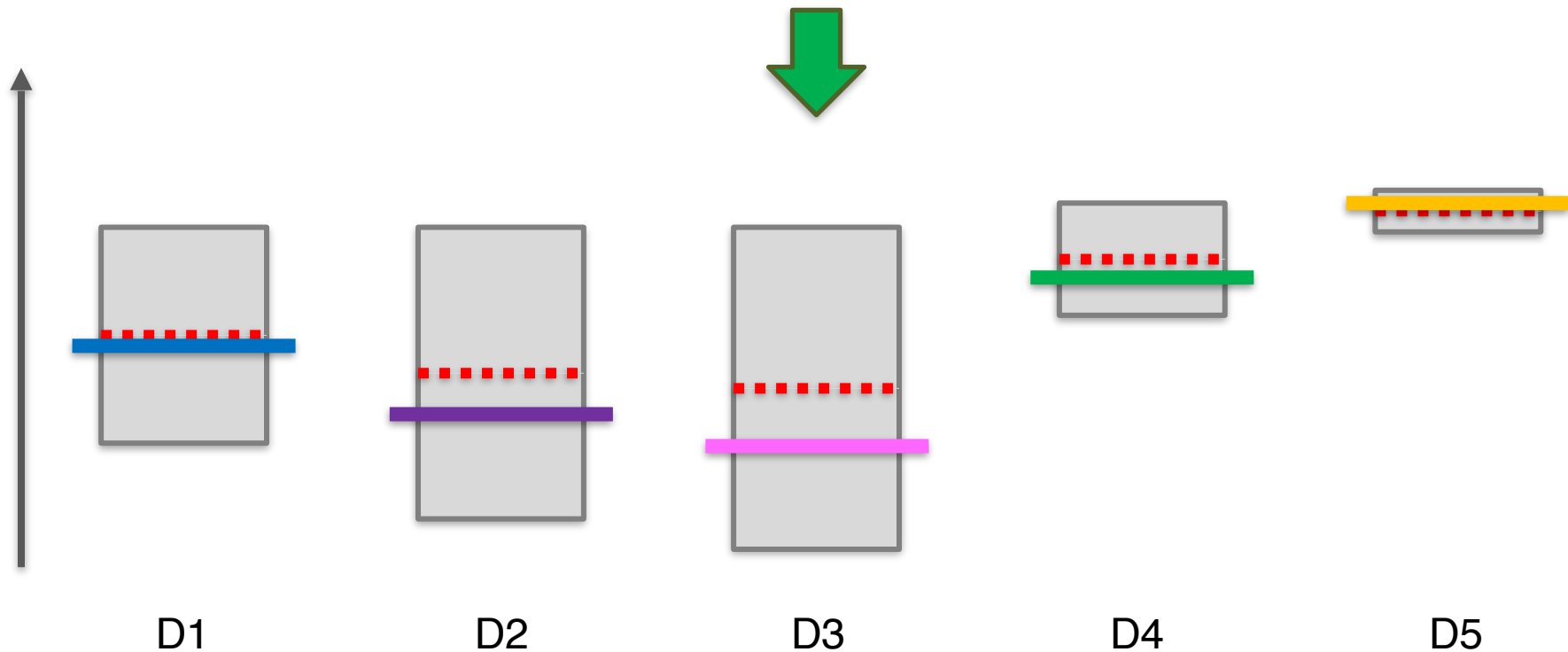
Algoritmo Upper Confidence Bound



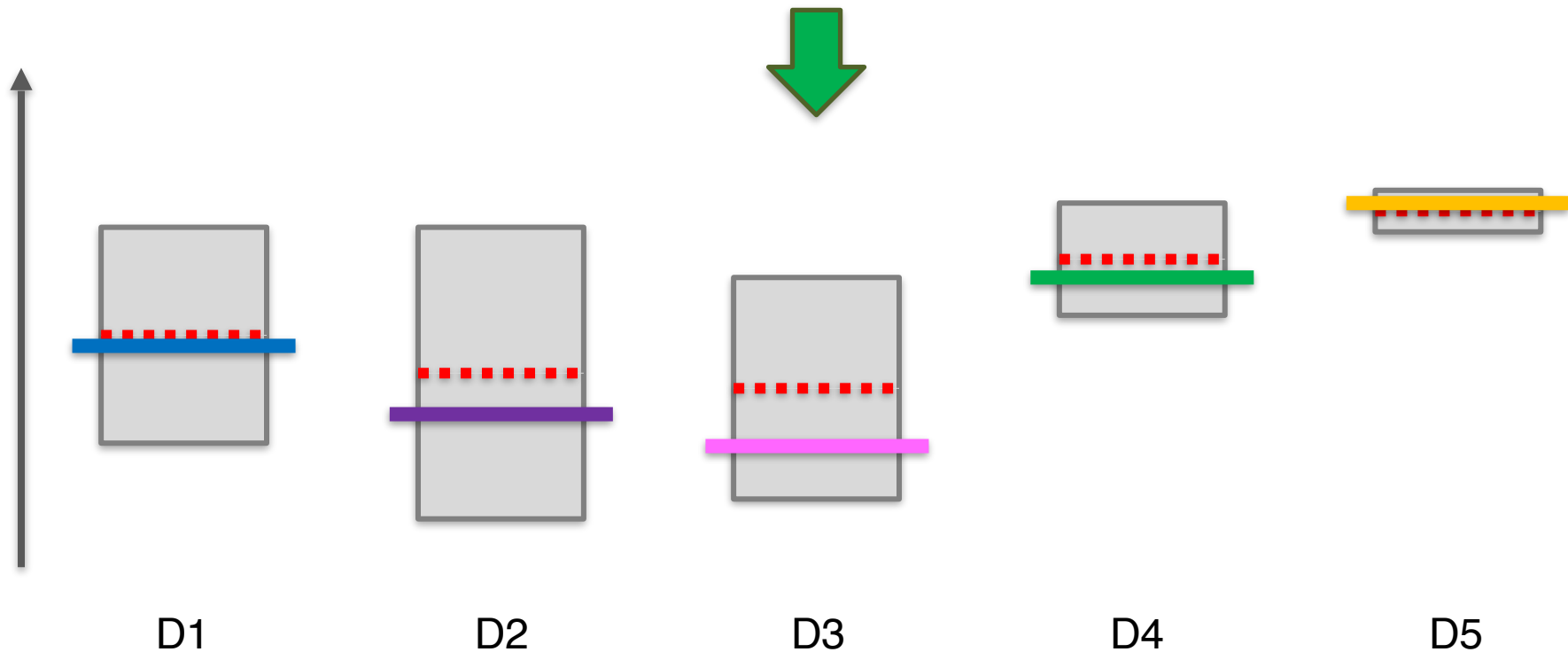
Algoritmo Upper Confidence Bound



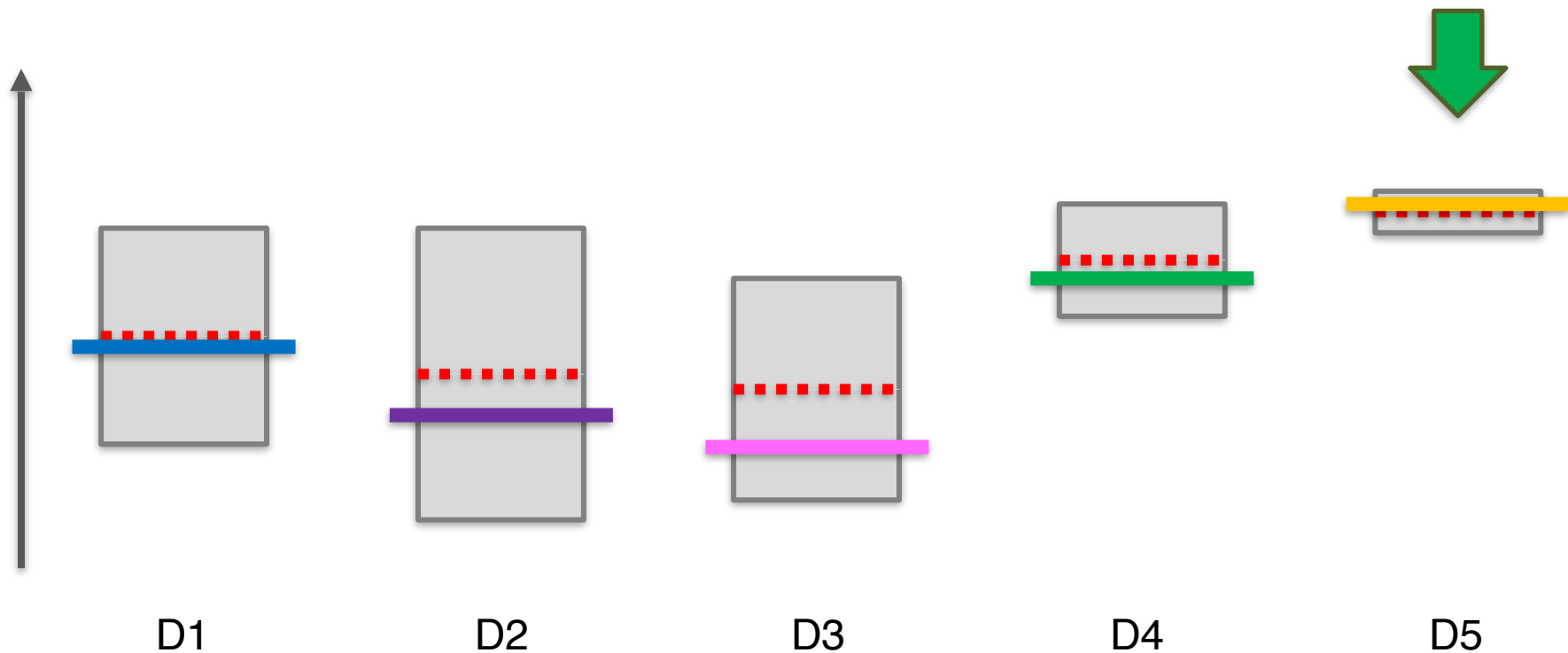
Algoritmo Upper Confidence Bound



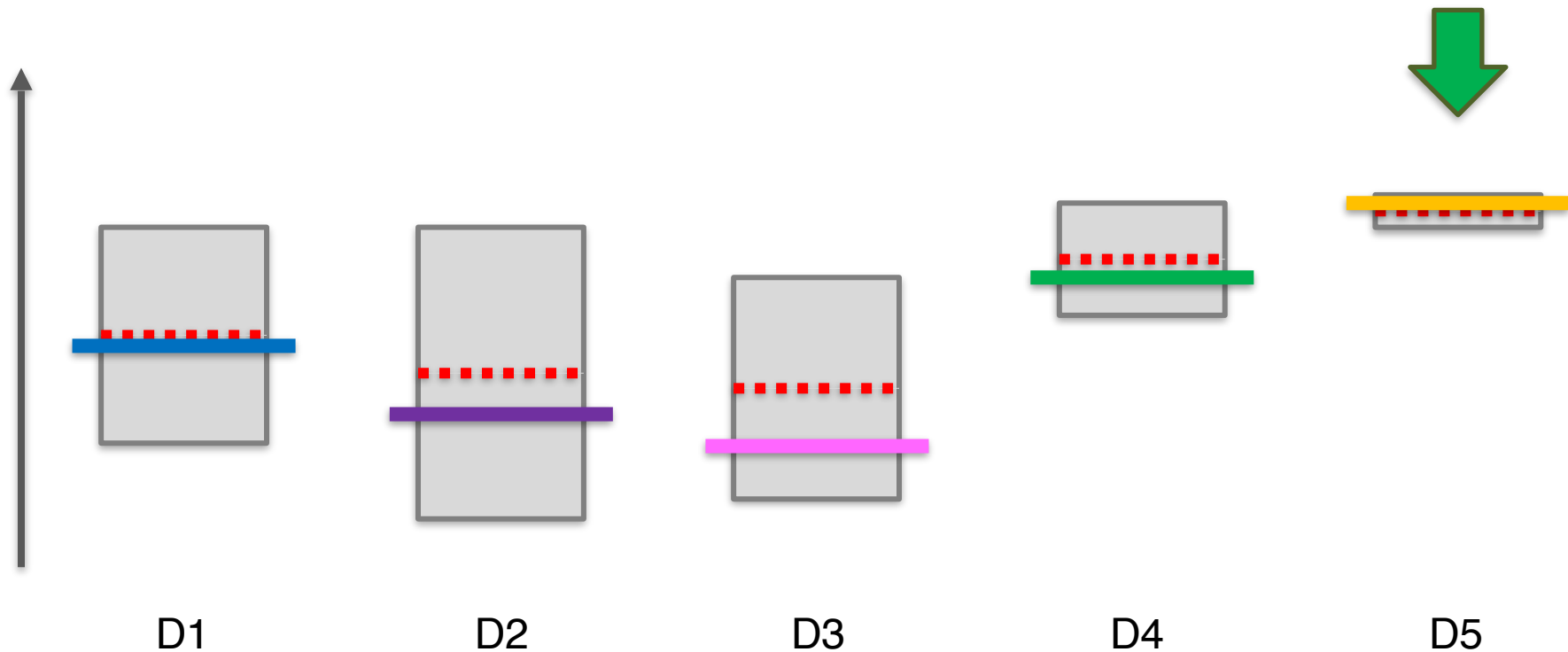
Algoritmo Upper Confidence Bound



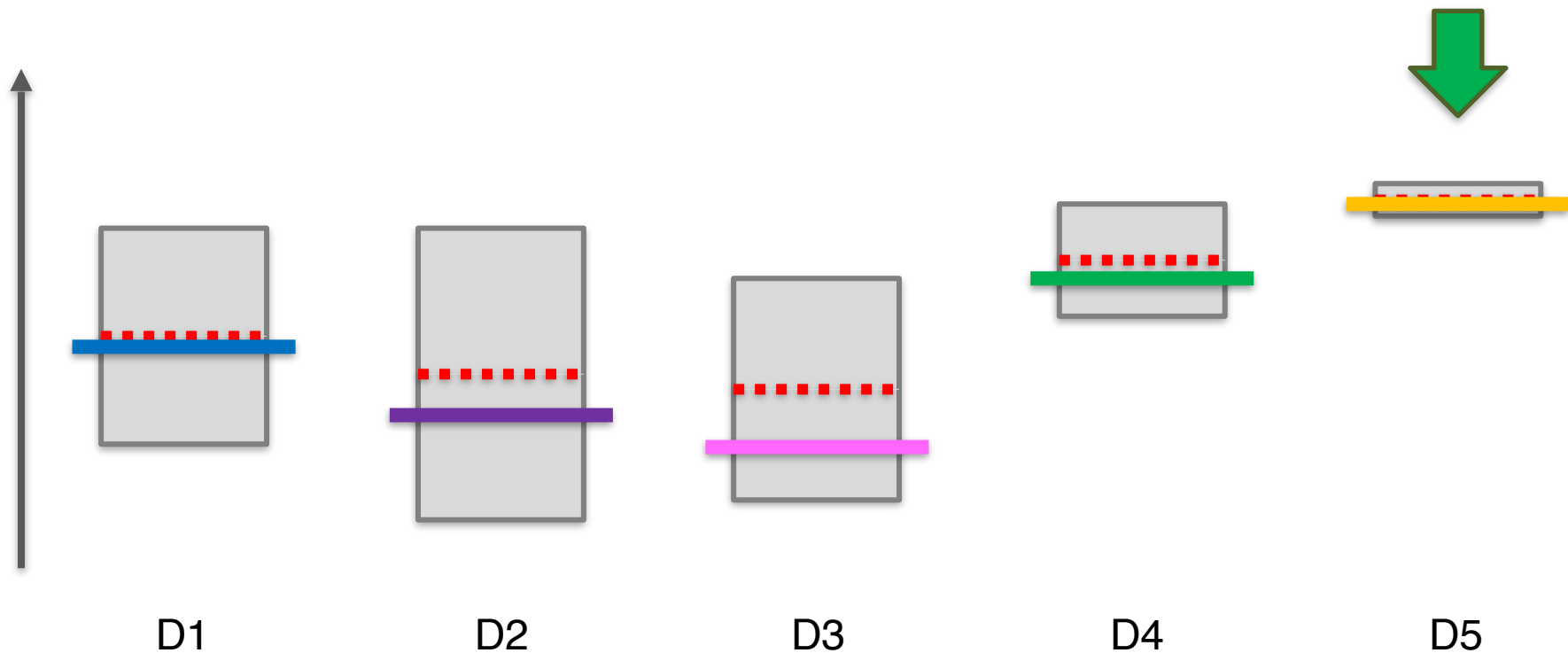
Algoritmo Upper Confidence Bound



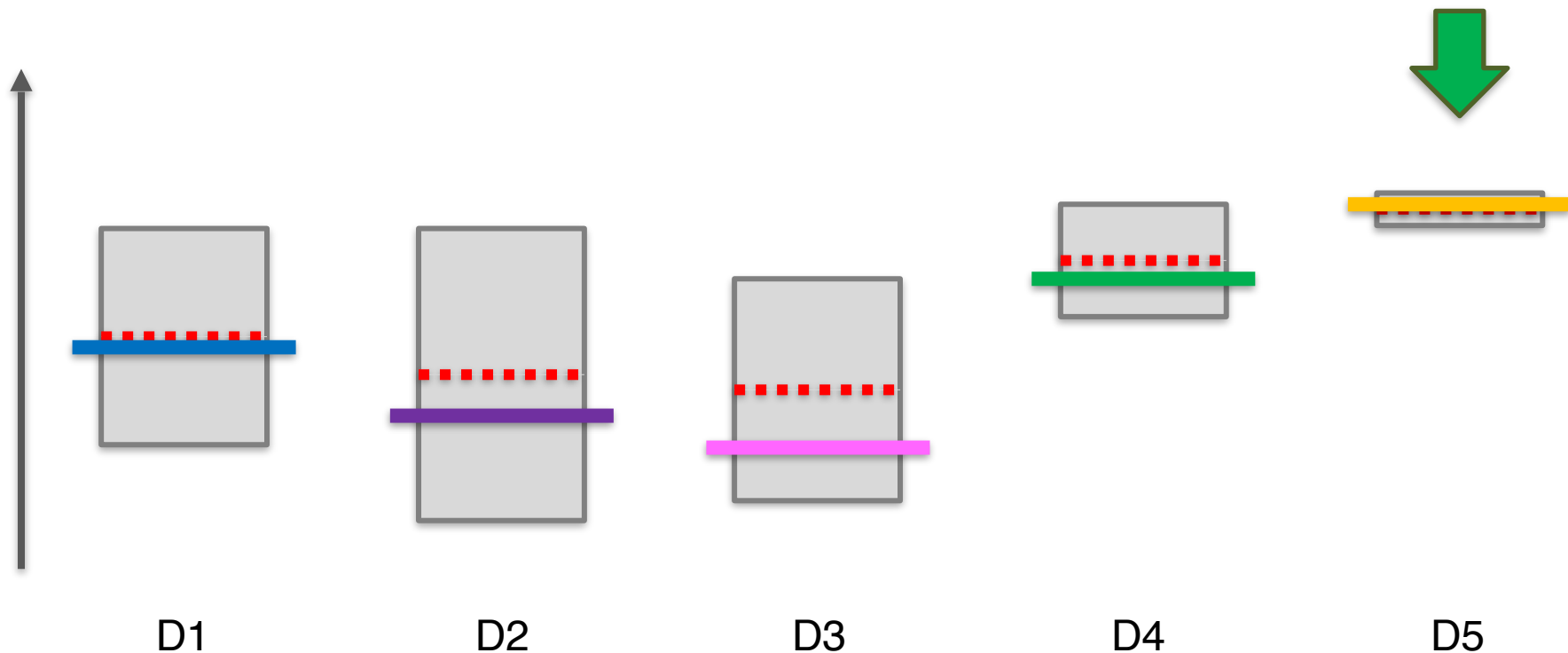
Algoritmo Upper Confidence Bound



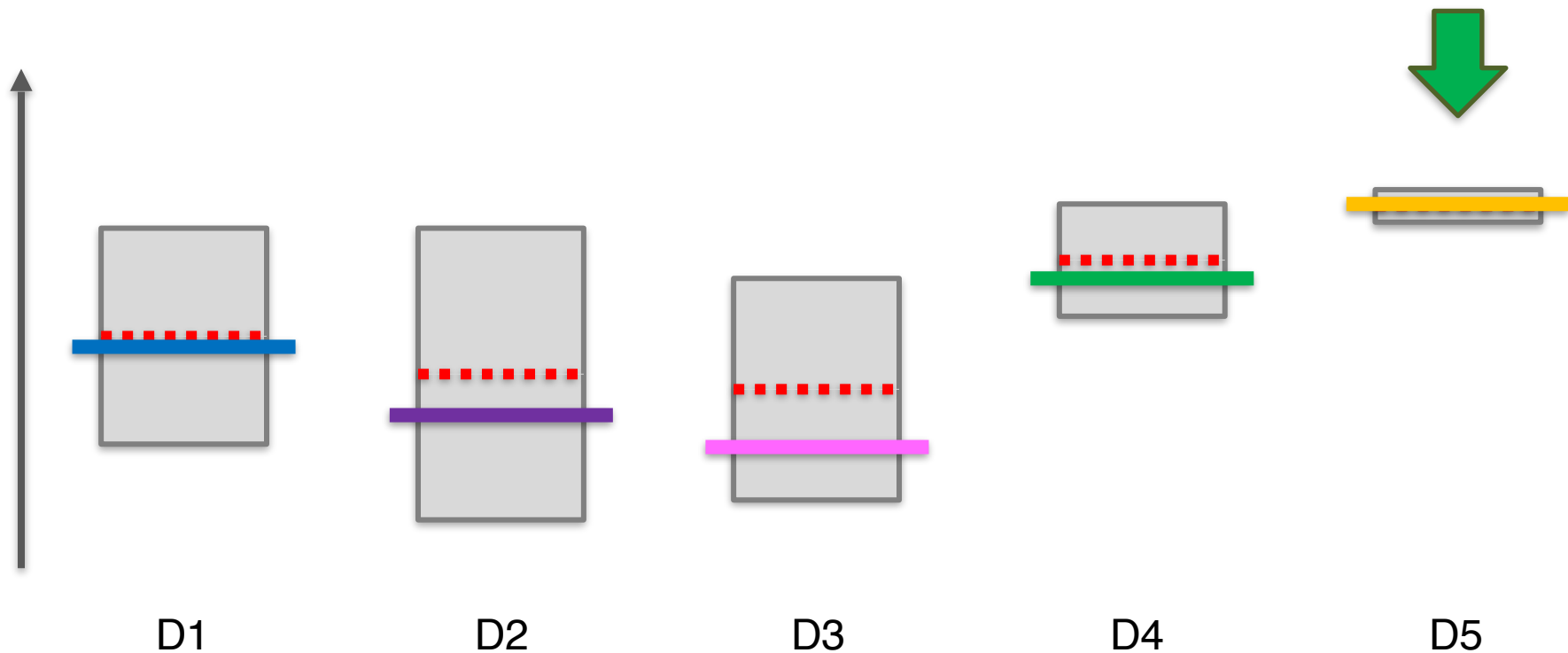
Algoritmo Upper Confidence Bound



Algoritmo Upper Confidence Bound



Algoritmo Upper Confidence Bound



Machine Learning A-Z

