Demostrar q' si $Q_t(a) = E[R_t \mid A_t = a] \Rightarrow \underbrace{A_t = \arg\max_a Q_t(a)}_{\text{Óptimo.}}$

Defino $G_T = \sum_{t=1}^{T} R_t$ Recompensa Total.

Retorno esperado $\boxed{E[G_T \mid A_1 = a_1, A_2 = a_2, \ldots, A_T = a_T]}$ (1)

Variable aleatoria X p.d.f $p(x)$ / $\int_{-\infty}^{+\infty} p(x)\,dx = 1$

$$\mu = E[X] = \int_{-\infty}^{+\infty} x\,p(x)\,dx$$

$$E[X \mid Y=y] = \int_{-\infty}^{+\infty} x\,p(x \mid Y=y)\,dx$$

$$\left.\begin{aligned} E[X+Y] &= E[X] + E[Y] \\ E[\alpha X] &= \alpha E[X] \end{aligned}\right\} \text{ linealidad } E[\cdot]$$

---

(1) $= \sum_{t=1}^{T} E[R_t \mid A_1 = a_1, \ldots] = \underbrace{E[R_1 \mid A_1 = a_1, \ldots]} + E[R_2 \mid A_2 = a_2, \ldots]$

$$= \underbrace{E[R_1 \mid A_1 = a_1] + E[R_2 \mid A_2 = a_2] + \cdots + E[R_T \mid A_T = a_T]}$$

$$\boxed{A_t = \arg\max_a Q_t(a) \quad \text{Greedy}}$$