There exists a collection of risky assets in the portfolio $\mathcal{P}$. An oracle provides the current price $p_i \in \mathbb{R}_+$ for each asset $i \in \mathcal{P}$, and a binary action vector $a \in \{0, 1\}^{|\mathcal{P}|}$ indicating whether each asset is available for investment ($a_i = 1$) or not ($a_i = 0$). The goal of the investment agent is to allocate a fixed budget B across these assets to **maximize the utility** of the portfolio.

Utility Function: A utility function $U : \mathbb{R}_+^{|\mathcal{P}|} \to \mathbb{R}$ maps the vector of shares of each asset in the portfolio to a real-valued utility score that reflects the investor's satisfaction with that allocation. We use a Cobb-Douglas utility function to model investor preferences:

$$U(n_1, n_2, \ldots, n_P) = \kappa(\gamma) \prod_{i \in \mathcal{P}} n_i^{\gamma_i}$$

where $\gamma_i \in \mathbb{R}$ is the **preference coefficient** for asset i (we need to estimate these values), and $\kappa(\gamma)$ is a leading coefficient that sets the scale of the utility function. Let $n = (n_1, \ldots, n_P)^\top \in \mathbb{R}_+^{|\mathcal{P}|}$ be the vector of shares, where n_i is the **number of shares** of asset i in the portfolio (we need to estimate these values).