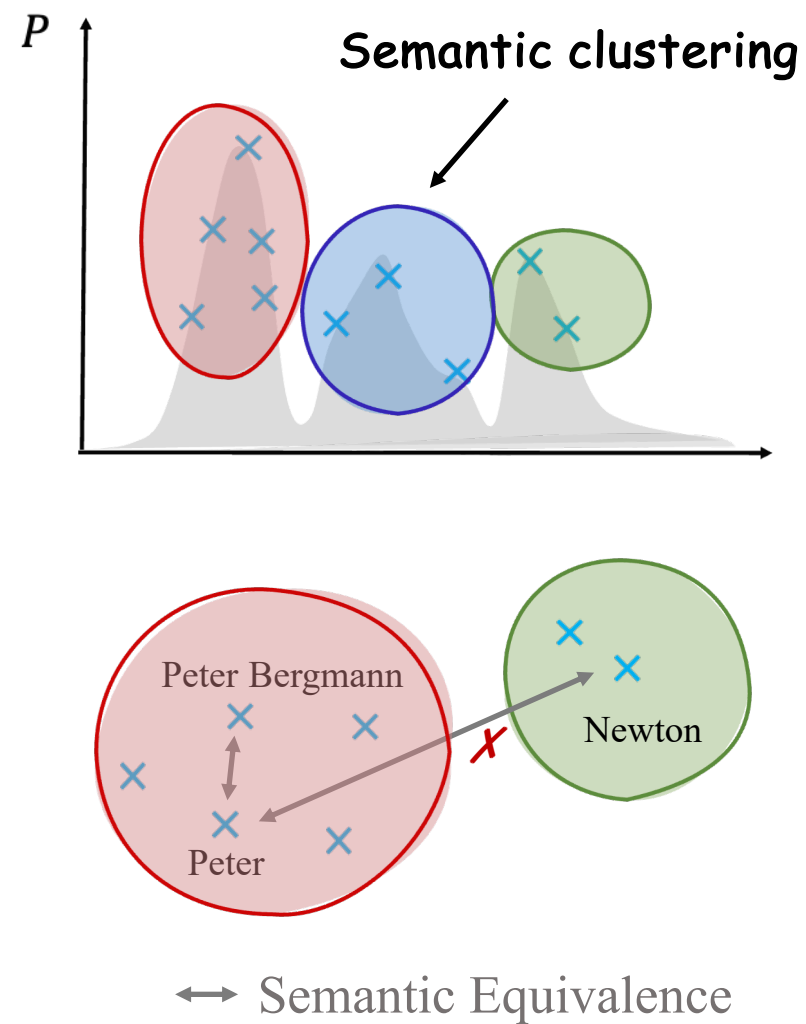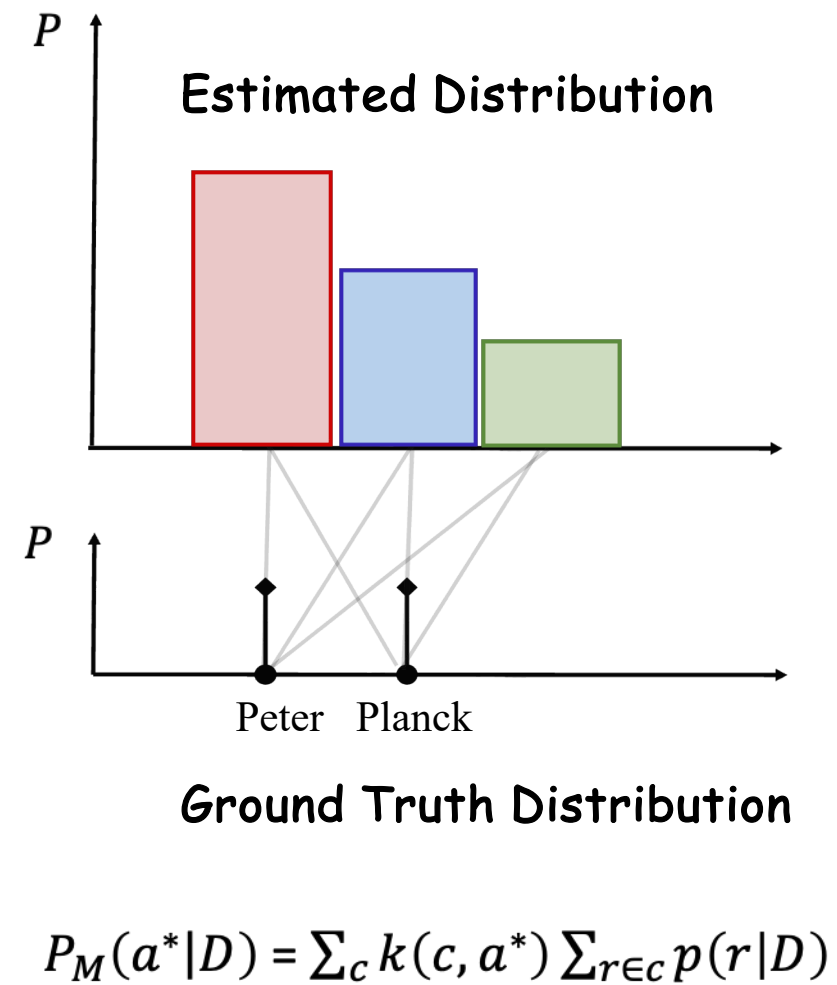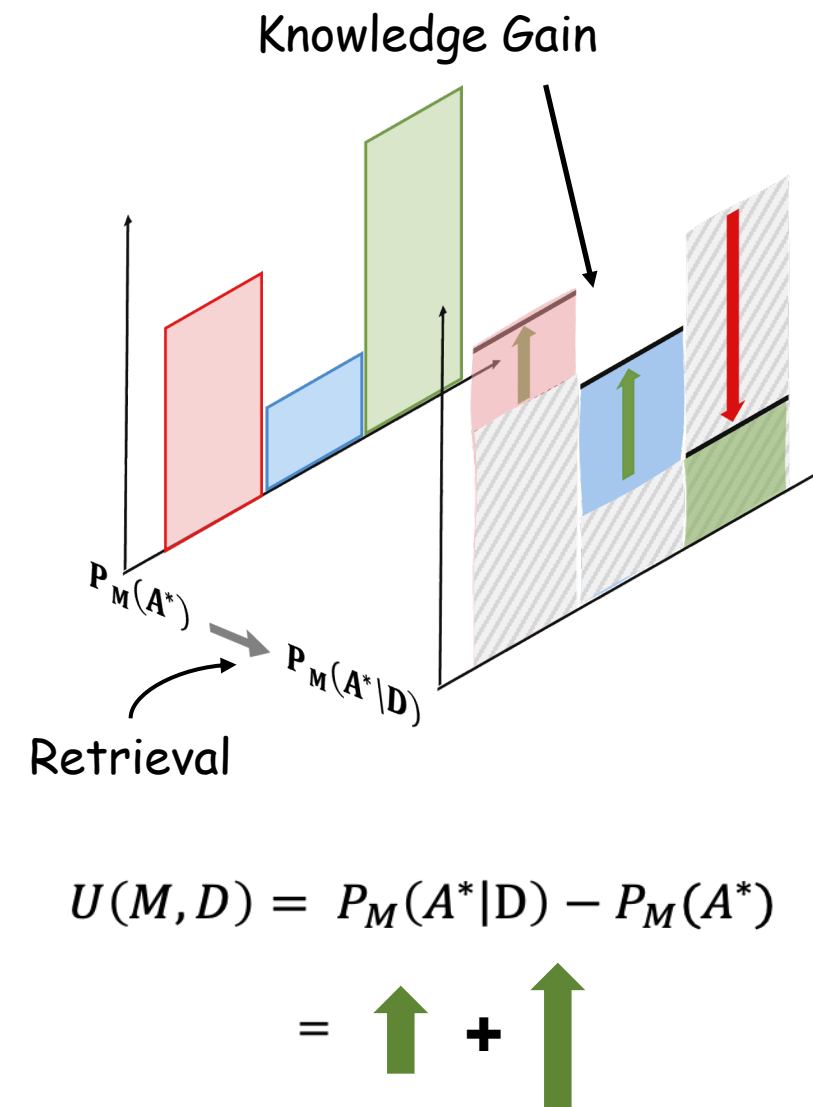① **Monte-Carlo Sampling**   ② **Semantic Clustering**   ③ **Semantic Perplexity (SePer)**   ④ **Utility Score Calculation**

Knowledge Distribution

Peter

Peter Bergmann

Newton

Ludwig Planck

LLM

$\times N$

Q #1:
Who did *Albert Einstein work with in 1933*?

Doc #1:
… Albert Einstein joined Princeton's Institute in 1933. …

Semantic clustering

Peter Bergmann

Newton

Peter

↔ Semantic Equivalence

Estimated Distribution

Peter   Planck

Ground Truth Distribution

$$P_M(a^*|D) = \sum_c k(c, a^*) \sum_{r \in c} p(r|D)$$

Knowledge Gain

$P_{M(A^*)}$

$P_{M(A^*|D)}$

Retrieval

$$U(M, D) = P_M(A^*|D) - P_M(A^*)$$