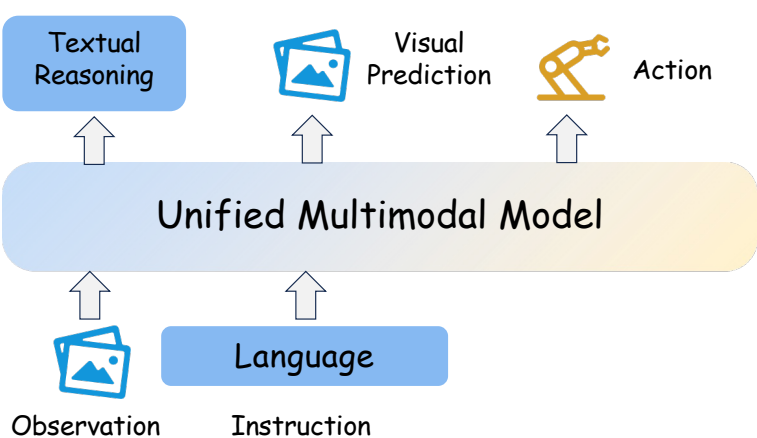
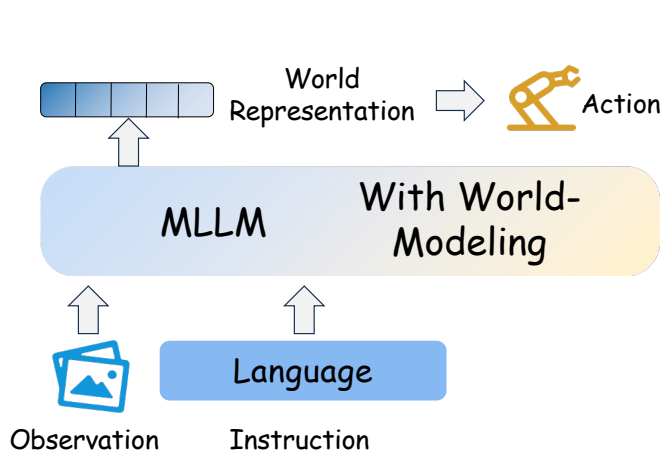




(a) Unified Vision-Language-Action Model



(b) Latent World Modeling Policy