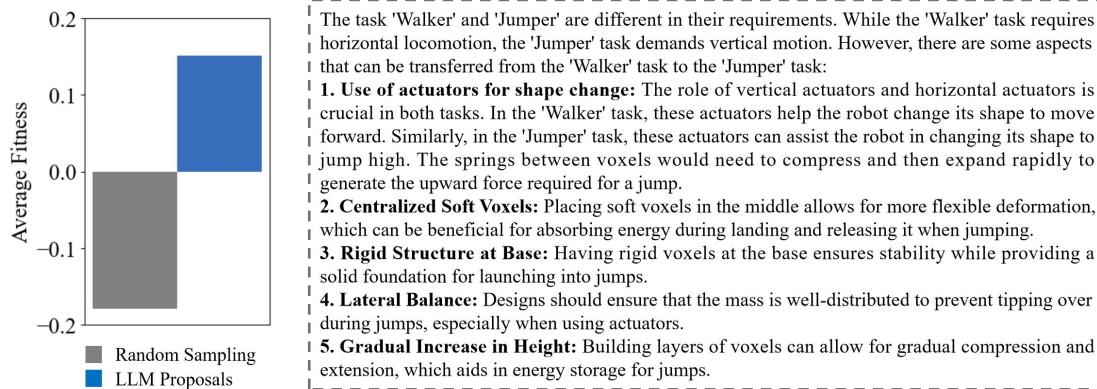


Supplementary material 11: Generalization across different tasks

While in our paper, we focused on transferring design experience between task instances that are intuitively similar (i.e. Walker-v0, BridgeWalker-v0 and UpStepper-v0), here we demonstrate that this prior knowledge of task relationships is not strictly necessary for successful inter-task generalization. Specifically, we performed an additional zero-shot design experiment (**Supplementary figure 11**) where the LLM was given elite samples of Walker-v0, but instructed to propose designs for a significantly different task, Jumper-v0, in which the robot is required to jump as high as possible. Despite the discrepancy between task objectives, we find that the LLM is able to dig deeper into the underlying inter-task associations and identify rather general, low-level design principles that are still relevant to the new task. Moreover, the zero-shot designs generated by LLM still outperform randomly sampled ones. These findings suggest that LLMs possess substantial potential for generalizing experience across seemingly different optimization problems, as long as they share some common ground and are not completely irrelevant to one other (such as robot design and TSP).



(a) Zero-shot performance (b) Design principles that LLM generalizes from Walker-v0 to Jumper-v0
Supplementary figure 11. Additional zero-shot design experiment on Jumper-v0.

The zero-shot performance is averaged over 10 robot designs for both random sampling and LLM proposals. Despite the task discrepancy, the LLM is still able to identify general principles, such as the importance of actuators for shape change and rigid voxels for stability, to transfer to the new task.