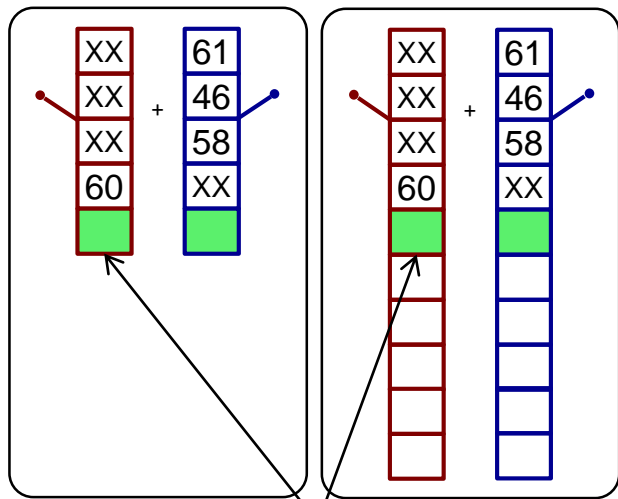


A

unequal information [1 3]

horizon 1

horizon 6



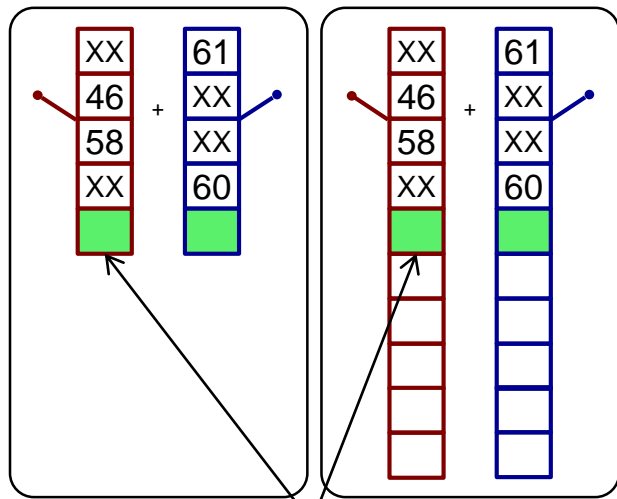
high information choice
 $p(\text{high info})$
directed exploration

B

equal information [2 2]

horizon 1

horizon 6



low mean reward choice
 $p(\text{low mean})$
random exploration