



while mám čas
 $\text{leaf} \leftarrow \text{select}(\text{root})$
 $\text{child} \leftarrow \text{expand}(\text{leaf})$
 $\text{result} \leftarrow \text{simulate}(\text{child})$
 $\text{backpropagate}(\text{result})$

podle max UCB

n
 (L/P)

$$UCB(n) = \frac{L(n)}{P(n)} + c \sqrt{\frac{\log P(\text{root}(n))}{P(n)}}$$

rollout

export

explore

explore

const

$\sqrt{2}$

Oproti minimaxu si vybíráme včelol s nejlepšími poměry,
 takže jeden chybní simulace mi jen malinko změní hodnocení
 včelol, zatímco u minimaxu to ovlivní úplně.