Modul 3: Beyond Classical Search

Masayu Leylia Khodra

KK IF – Teknik Informatika – STEI ITB

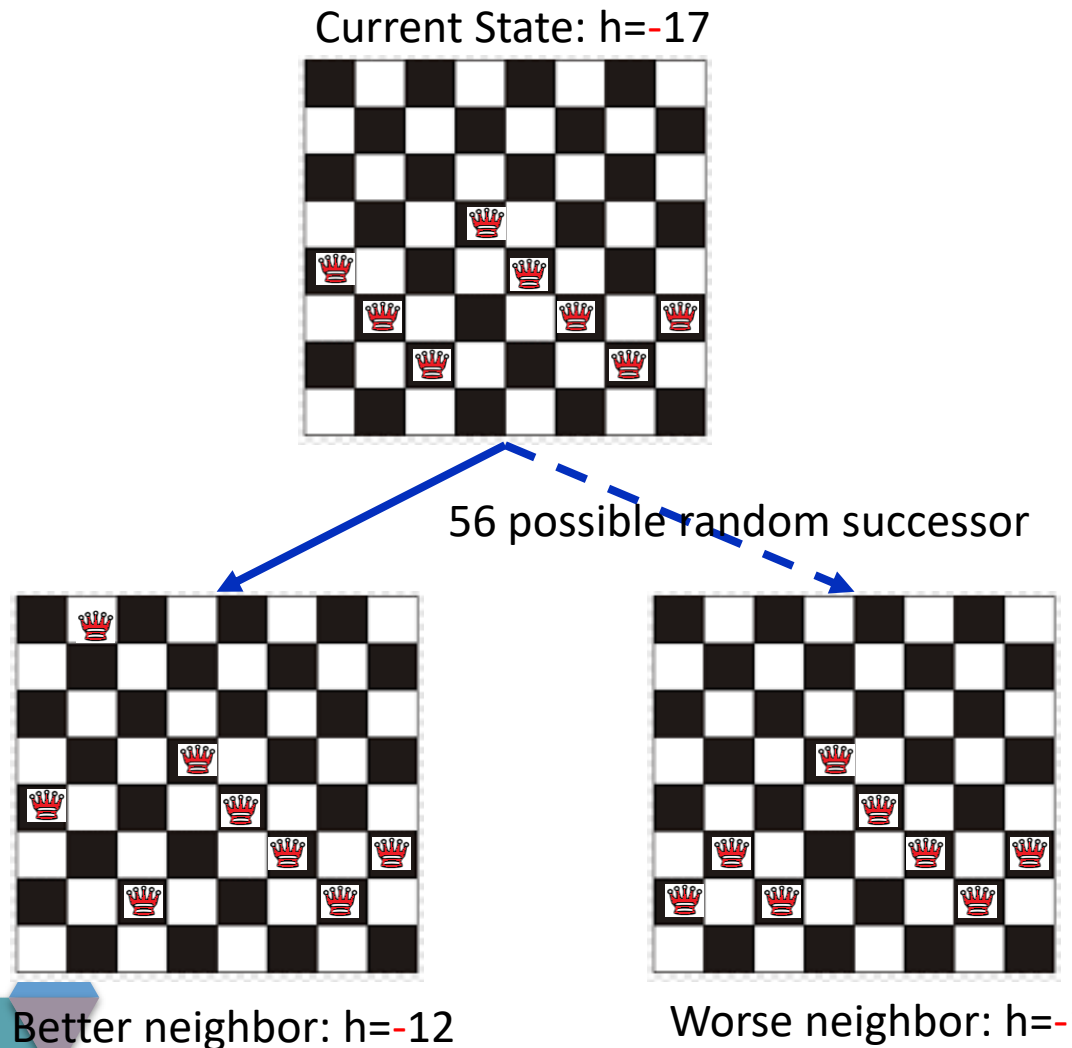
Simulated Annealing

Inteligensi Buatan
(*Artificial Intelligence*)

SA: Combining Completeness and Efficiency

- Simulated annealing combines hill climbing (efficient) with a random walk (complete)
- Idea: escape local maxima by allowing some "bad" moves but gradually decrease their frequency
- Simulated annealing is a version of stochastic hill climbing where some downhill moves are allowed.

Neighbor: One Random Successor of Current



- Stochastic hill-climbing:
  - Only move to better neighbor, and skip worse neighbor.
- Simulated annealing:
  - move to better neighbor, and allow move to worse neighbor that has probability $e^{\Delta E/T}$

Simulated Annealing

https://id.m.wikipedia.org/wiki/Berkas:Annealing_a_silver_strip.JPG

Annealing: heat (metal or glass) and allow it to cool slowly, in order to remove internal stresses and toughen it.

- T is the "temperature" of annealing that gradually decreases.
- ΔE=neighbor.value-current.value
- If better neighbor ($\Delta E>0$): move to neighbor (stochastic HC)
- If worse neighbor ($\Delta E<0$): probability move $e^{\Delta E/T}$
  - When T is high (e.g. ΔE=-5, T=100, prob=0.95), there is a lot of random motion ➔ random walk
  - When T approaches 0 (e.g. ΔE=-5, T=1, prob=0.007), randomness is decreased ➔ stochastic hill climbing.

Simulated Annealing (Russel & Norvig, 2010)

function SIMULATED-ANNEALING(*problem*, *schedule*) **returns** a solution state
  inputs: *problem*, a problem
    schedule, a mapping from time to "temperature"

  current ← MAKE-NODE(*problem*.INITIAL-STATE)
  for $t = 1$ **to** ∞ **do**
    $T \leftarrow schedule(t)$
    if $T = 0$ **then return** *current*
    next ← a randomly selected successor of *current*
    $\Delta E \leftarrow next.\text{VALUE} - current.\text{VALUE}$
    if $\Delta E > 0$ **then** *current* ← *next*
    else *current* ← *next* only with probability $e^{\Delta E/T}$

T: temperature as a function of time t

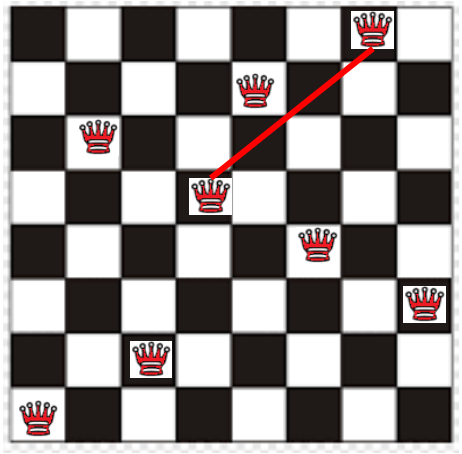
Terminates T=0

Move to better neighbor (stochastic HC)

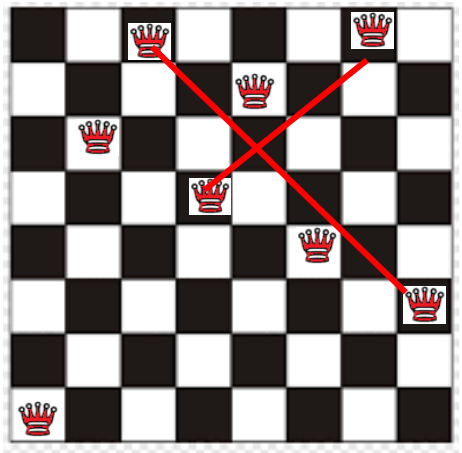
allowing some "bad" moves, depends on probability

Move to Worse State

Current State: h=-1



Column: C3
Row: 2➔8



$current \leftarrow next$ only with probability $e^{\Delta E/T}$

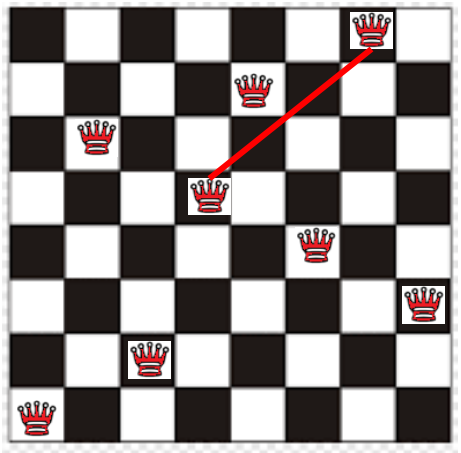
Compare to static value:
move probability > 0.5

Compare to random value:
move probability > random(0,1)

Next: h=-2, T=10
$\Delta E<0$: prob=$e^{(-2-(-1))/10}=e^{-0.1}$ =0.9

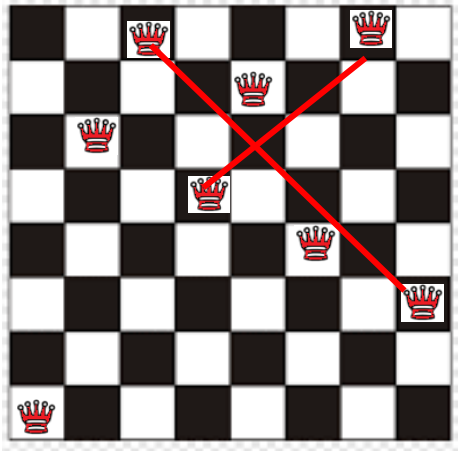
Simulated Annealing: Illustration

Current State: h=-1



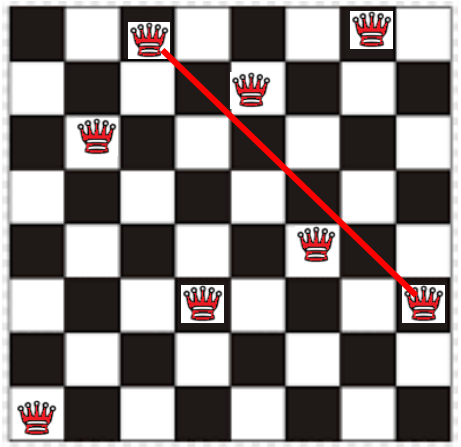
Column: C3
Row: 2➔8

Next: h=-2, T=10
Prob=$e^{(-2-(-1))/10}=e^{-0.1}$ =0.9
Current ← next



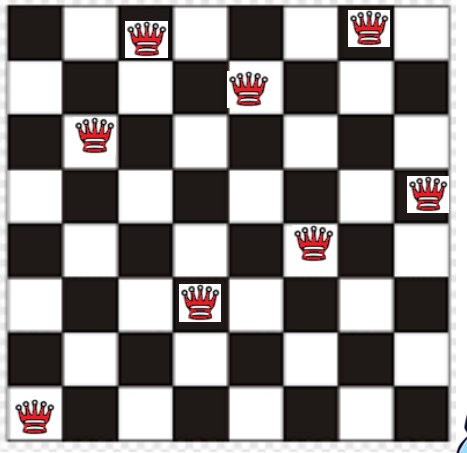
Column: C4
Row: 5➔3

Next: h=-1
Current ← next

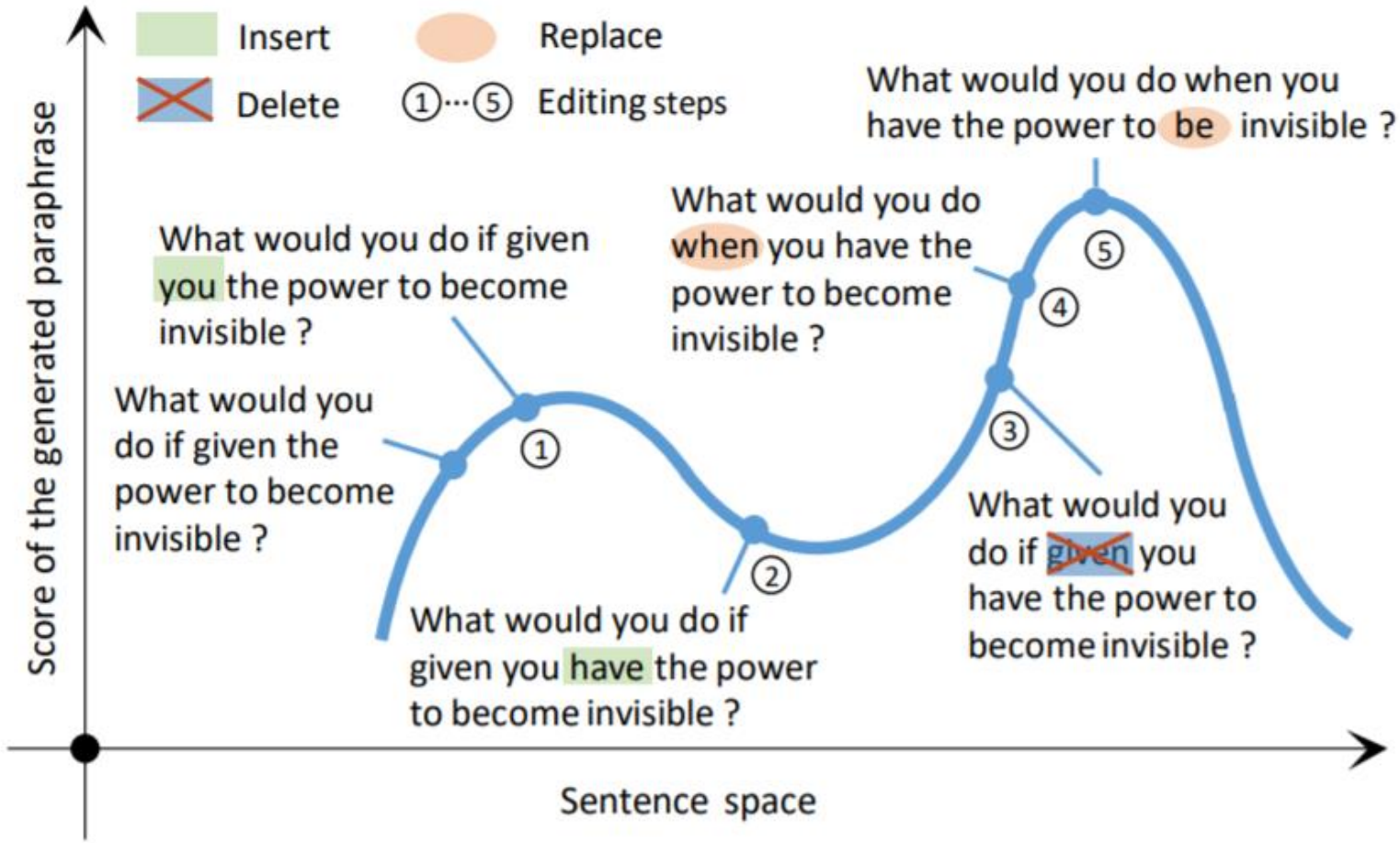


Column: C8
Row: 3➔5

Next: h=0
Current ← next



Application: Simulated Annealing for Paraphrase



Liu, X., Mou, L., Meng, F., Zhou, H., Zhou, J., & Song, S. (2020). Unsupervised paraphrasing by simulated annealing. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 302–312 July 5 - 10, 2020. c 2020 Association for Computational Linguistics. *arXiv preprint arXiv:1909.03588*.

Properties of simulated annealing search

One can prove: If *T* decreases slowly enough, then simulated annealing search will find a global optimum with probability approaching 1

Widely used in VLSI layout, airline scheduling, etc

Summary: Simulated Annealing

Simulated annealing is a version of stochastic hill climbing where some downhill moves are allowed.

If *T* decreases slowly enough, then simulated annealing search will find a global optimum with probability approaching 1

Next:

- Genetic Algorithm

THANK YOU