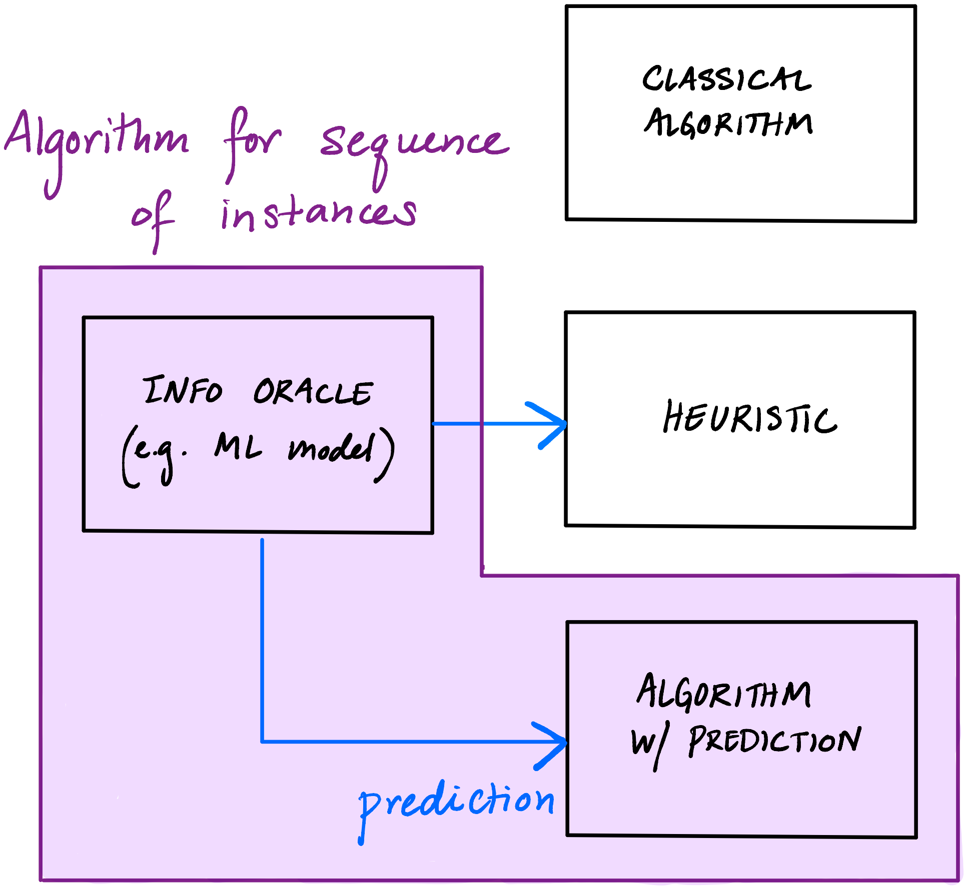
Competitive strategies to use "warm start" algorithms with predictions

Vaidehi Srinivas
Northwestern University,
Computer Science

joint work with
Avrim Blum
Toyota Technological Institute, Chicago

ALGORITHMS WITH PREDICTIONS



Pessimistic:

Worst-case guarantee on every input

Optimistic:

Performs well when information is good

Best of both worlds:

Good performance when info is good <u>and</u> worst-case guarantee

$\boxed{?}$ How can we learn good information for algorithms with predictions?

↳ "End to end" guarantees for solving sequences of related instances

WARM STARTS

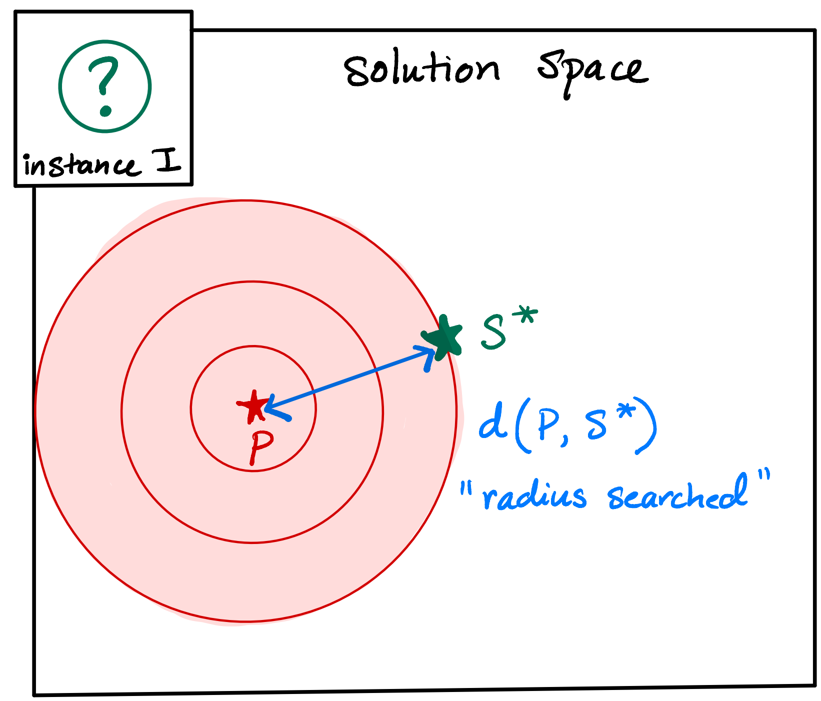
Assume an instance space $\mathcal{I}$, metric solution space $\mathcal{S}$

Definition. A warm start algorithm $A(I, P)$ for $I \in \mathcal{I}$, $P \in \mathcal{S}$ finds S^* (true solution for instance I) in time at most $d(P, S^*)$.

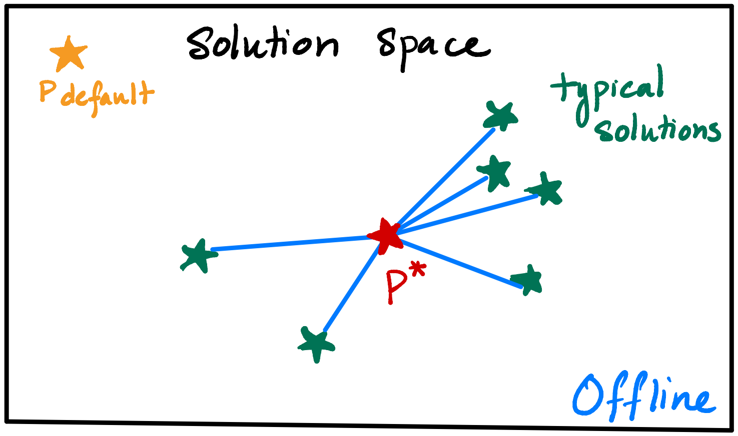
Examples.

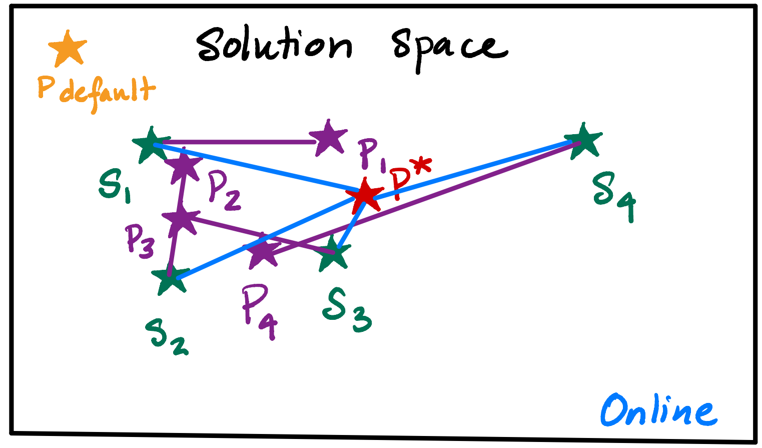
- Bipartite matching via learned duals
  [Dinitz Im Lavastida Moseley Vassilvitskii 21],
  [Chen Silwal Vakilian Zhang 22]

- Max Flow via Ford Fulkerson
  [Polak Zub 22],
  [Davies Moseley Vassilvitskii Wang 23]



TYPICAL USE CASES





Offline. Assume distribution $\mathcal{D}$ over instance solution pairs $\mathbb{I} \times \mathbb{S}$.
Learn a fixed prediction P^* to minimize

$$\mathbb{E}_{(I,S)\sim \mathcal{D}} \left[d\,(P^*, S) \right].$$

Online. For a sequence of instances $I_1, \ldots, I_t$ arriving online,
select predictions to compete with best fixed prediction in hindsight.
[Khodak Balcan Talwalkar Vassilvitskii 22] solve via online convex optimization

⚠ Require solutions to be clustered to provide strong guarantees

? Are there other forms of structure we can take advantage of?

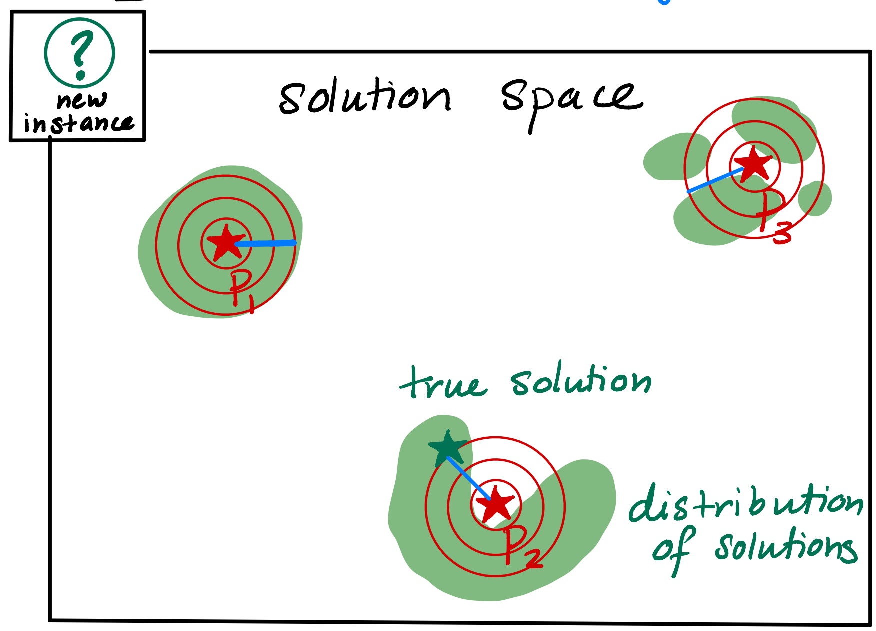
A SIMPLE OBSERVATION

? What if there are multiple types of instances?

Offline. Assume distribution $\mathcal{D}$ over instance solution pairs $\mathbb{I} \times \mathbb{S}$.
Learn fixed predictions $P_1, \dots, P_k$ to minimize

$$\mathop{\mathbb{E}}_{(I,S)\sim\mathcal{D}} \left[\min_i \, d(P_i, S) \right]$$

(k-medians clustering cost)



Cost: $3 \times$ radius

Algorithm

For a fresh $(I, S) \sim \mathcal{D}$,

run $\mathcal{A}(I, P_1), \dots, \mathcal{A}(I, P_k)$ "in parallel"

output computation that completes first.

Expected work is $O(k)$ times k-median clustering cost

? Can we get around the $O(k)$ factor by using other structure?

? Online setting?

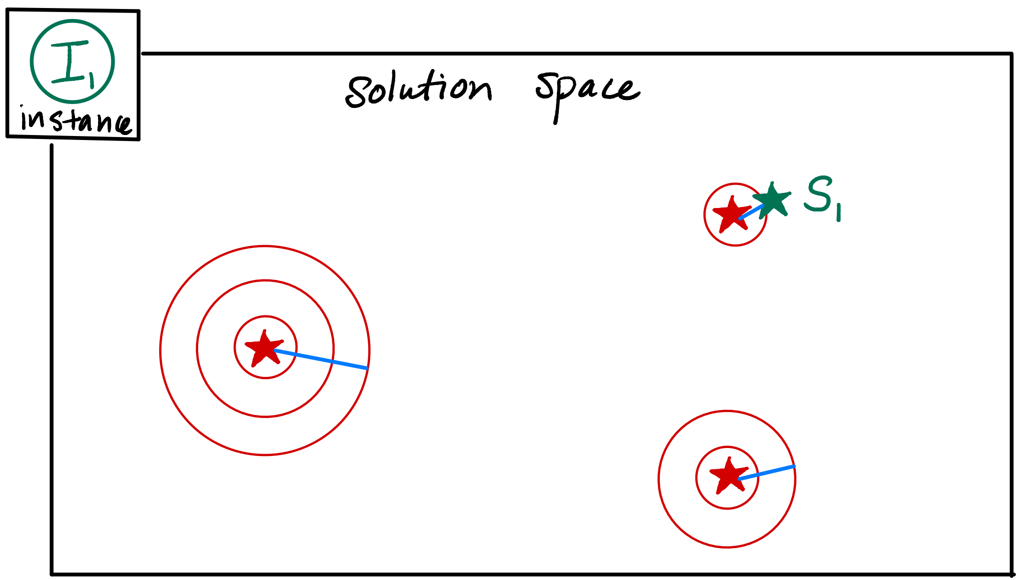
ONLINE FORMULATION

Problem.

On each day $t \in [T]$, instance $I_t \in \mathcal{I}$ arrives online

Algorithm with access to warm start $\mathcal{A}(\cdot, \cdot)$ must solve I_t.

Goal: minimize total runtime over all days



total radius searched:

6

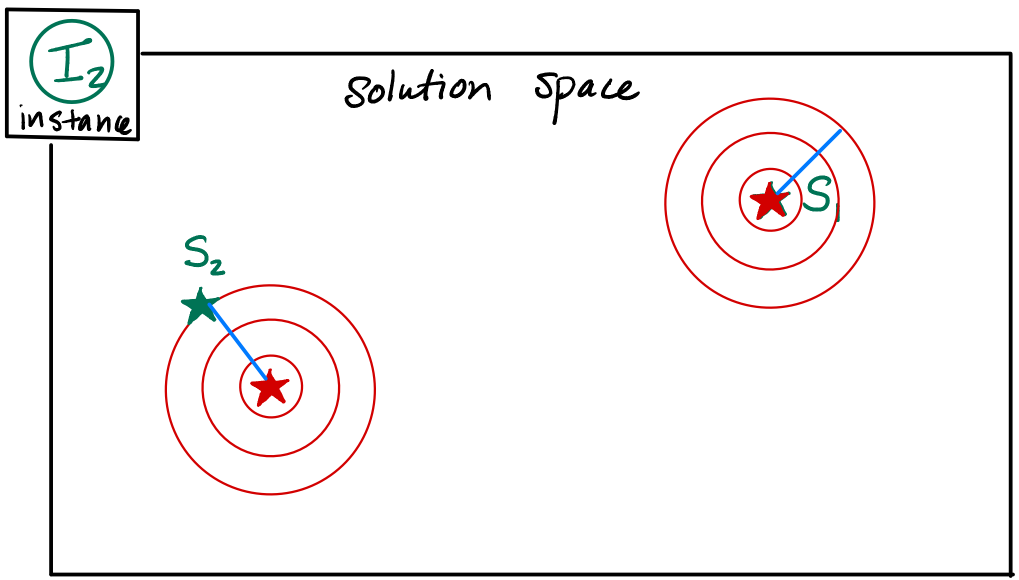
ONLINE FORMULATION

Problem.

On each day $t \in [T]$, instance $I_t \in \mathcal{I}$ arrives online

Algorithm with access to warm start $\mathcal{A}(\cdot, \cdot)$ must solve I_t.

Goal: minimize total runtime over all days



total radius searched:

$6 + 6 = 12$

Want to do competitive analysis

Problem: not enough structure assumed $\Longrightarrow$ offline strategy too powerful

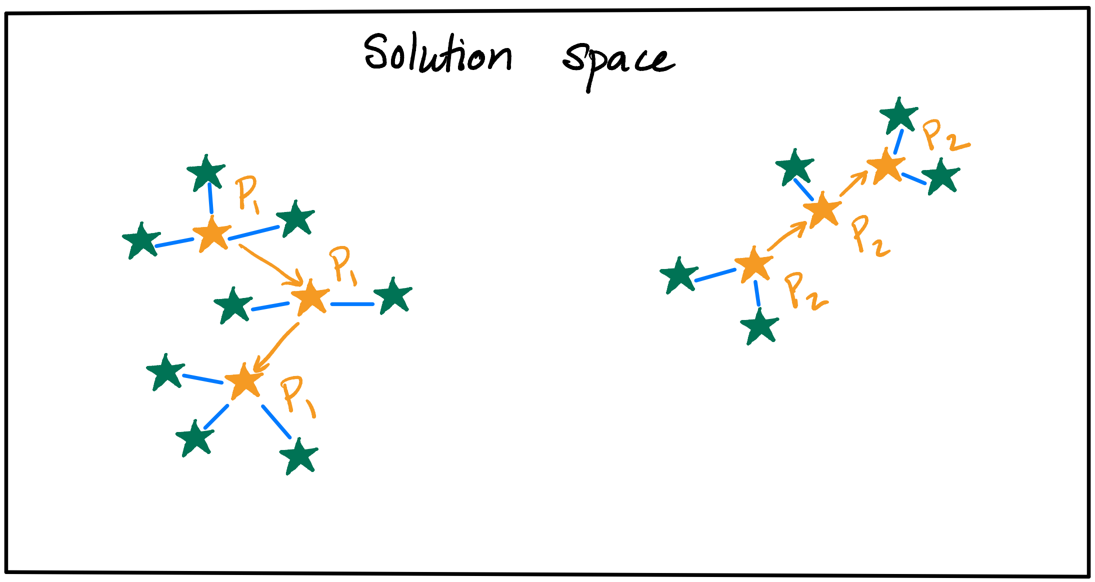
BASELINES

Definition [k offline trajectories]

Offline strategy that maintains a set of k predictions, and on each day pays

hit cost : distance from today's solution to closest prediction

movement cost : distance moved by all predictions



- Sequences that are "clusterable" have good baselines
- **WANT:** Algorithm is competitive with baselines

RESULT

Theorem [S. Blum 24].

We give an algorithm that is $O(k^4 \ln^2 k)$-competitive with any set of k offline trajectories in radius searched.

Furthermore,

- total runtime of algorithm is bounded by $O(1)$ times the radius searched
- algorithm is deterministic (robust to adaptive adversary)
- guarantee holds for all k simultaneously

- End-to-end guarantee for sequences of instances that display "clustering" structure

THANKS!

vaidehi@u.northwestern.edu