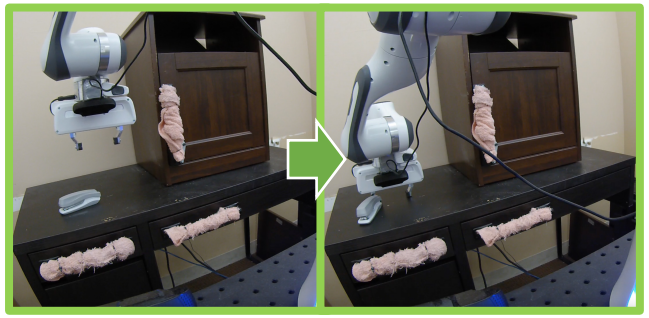
Learning Language-Conditioned Robot Behavior from Offline Data and Crowd-Sourced Annotation

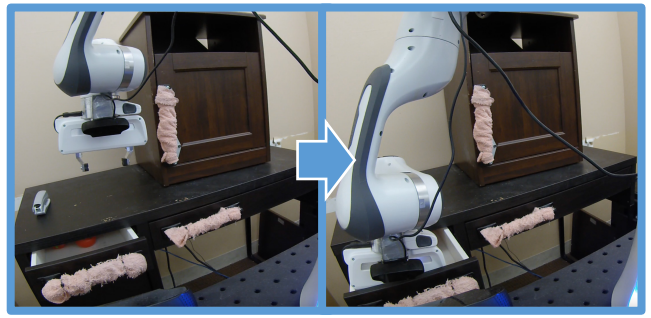
Suraj Nair[1], Eric Mitchell[1], Kevin Chen[1], Brian Ichter[2], Silvio Savarese[1], Chelsea Finn[1,2]

[1]Stanford University, [2]Robotics at Google

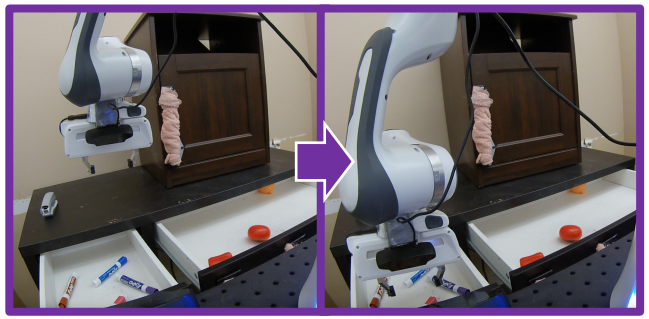
"Move the stapler"



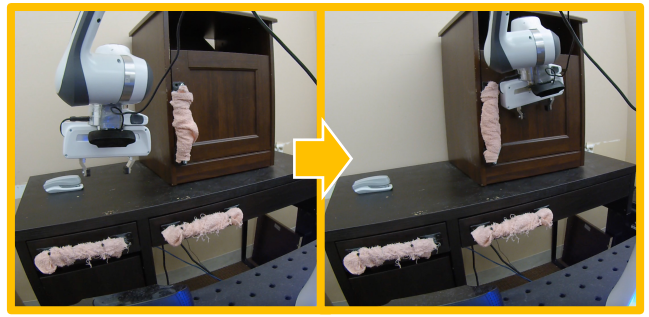
"Open the left drawer"



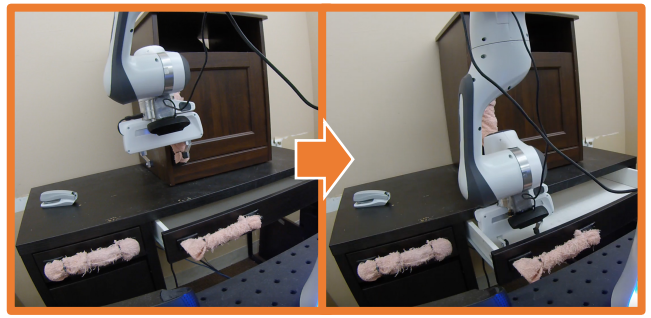
"Reach the marker"



"Reach the cabinet"

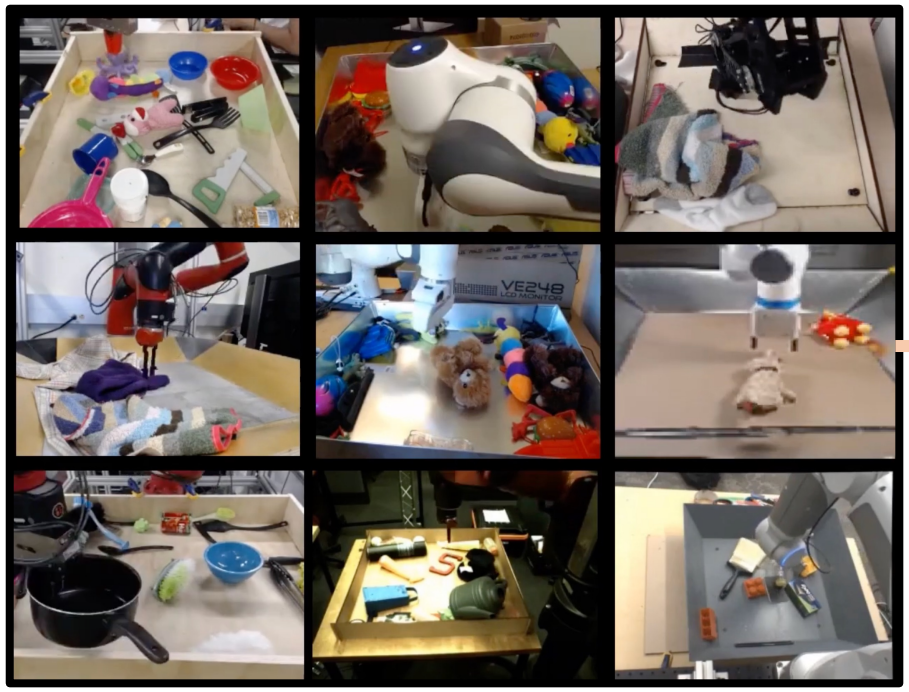


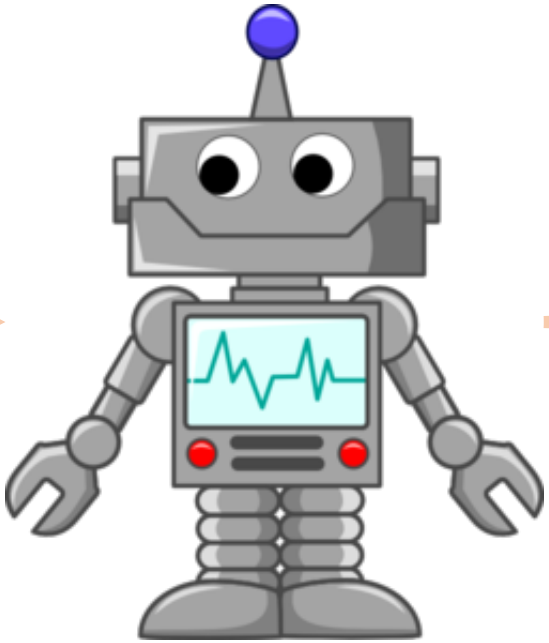
"Open the right drawer"



Multi-Task Learning from Offline Data

Offline Data

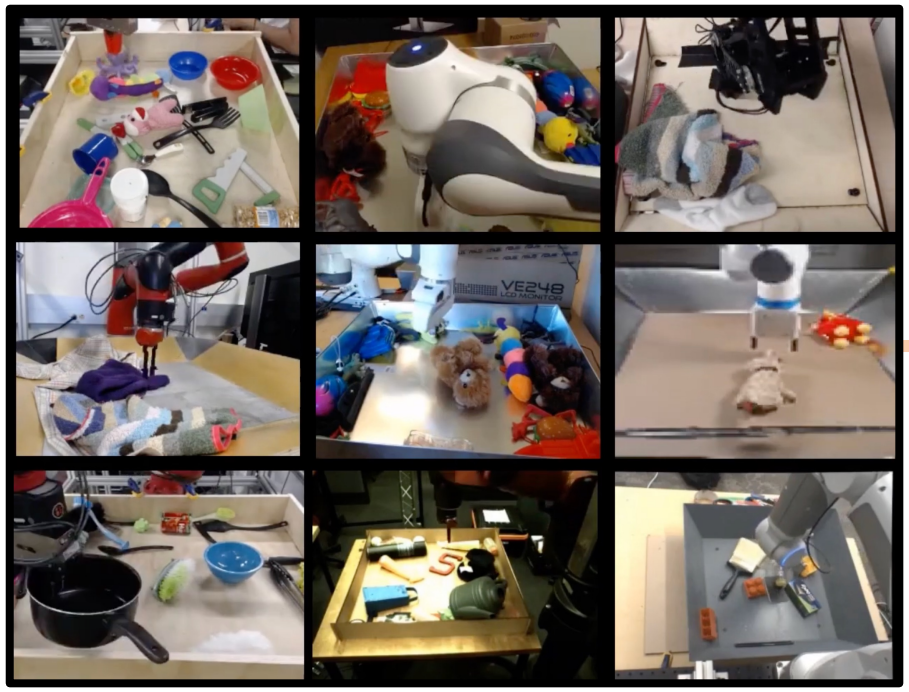


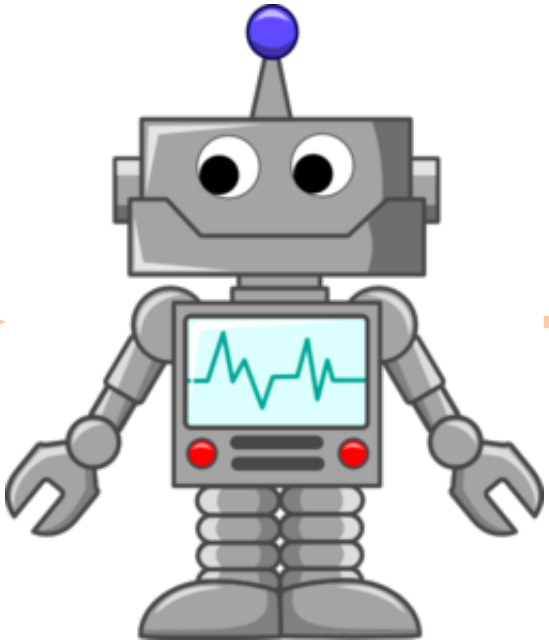


Dasari et al. RoboNet: Large-Scale Multi-Robot Learning. CoRL 2019

Multi-Task Learning from Offline Data

Offline Data





Humans need easy and effective ways of specifying tasks

Dasari et al. RoboNet: Large-Scale Multi-Robot Learning. CoRL 2019

Goal Image Task Specification

Initial Image

Goal Image

Goal Image Task Specification

Initial Image

Goal Image

Requires human effort

Goal Image Task Specification

Initial Image

Goal Image

Requires human effort
Task over-specification

Goal Image Task Specification

Initial Image

Goal Image

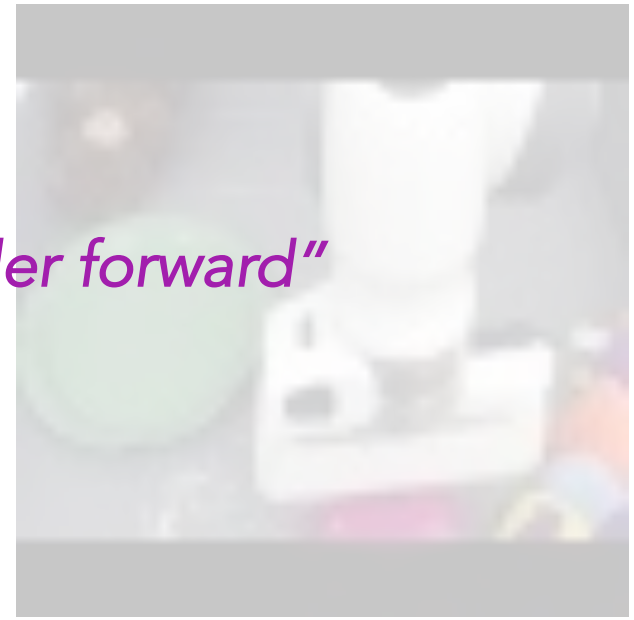
Requires human effort
Task over-specification
Cannot specify persistent behavior

Natural Language Task Specification

Initial Image

"Push the pink stapler forward"

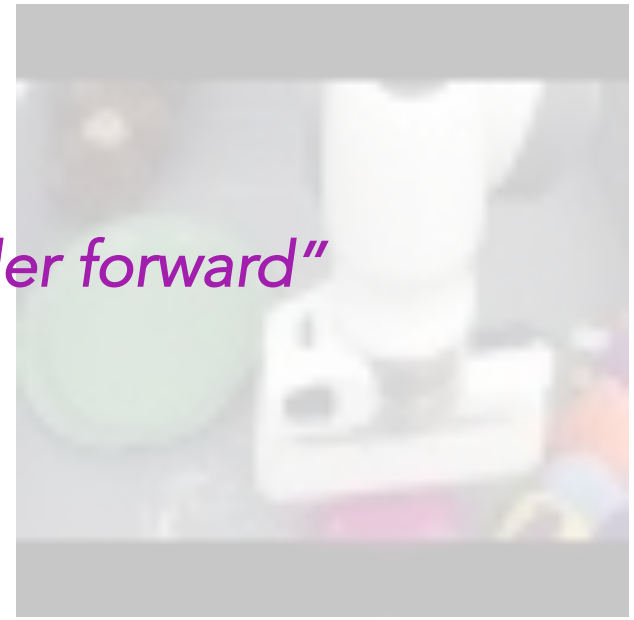
Goal Image



Natural Language Task Specification

Initial Image

Goal Image



"Push the pink stapler forward"

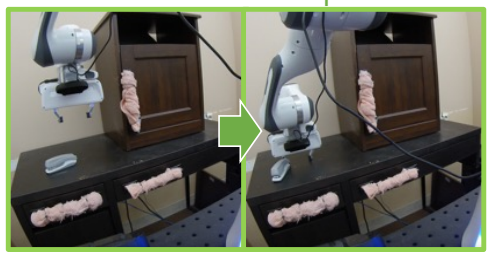
Easy for humans to provide
Can flexibly represent tasks

Our Objective

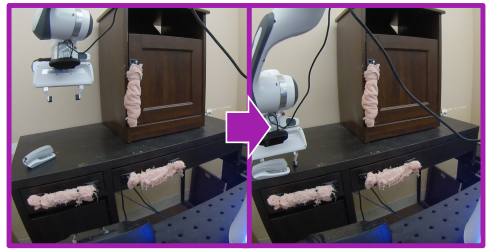
Learn ***language-conditioned visuomotor manipulation skills*** from:
offline datasets +
crowd-sourced annotation

Language-Conditioned Task Completion

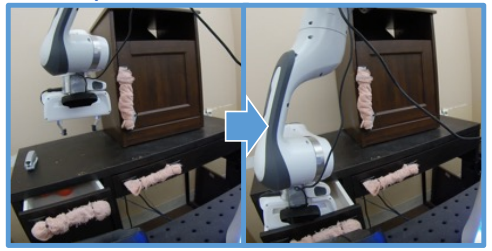
"Move the stapler"



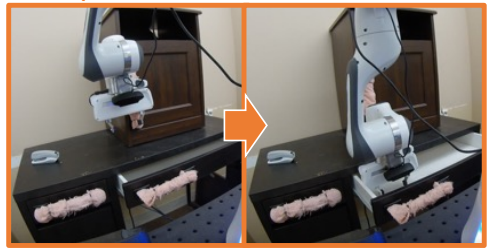
"Push the small gray stapler around on top of the black desk"



"Open the left drawer"



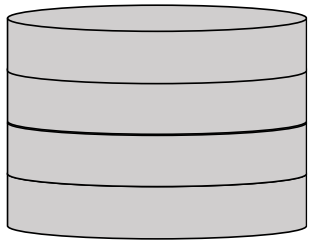
"Open the right drawer"



Our Objective

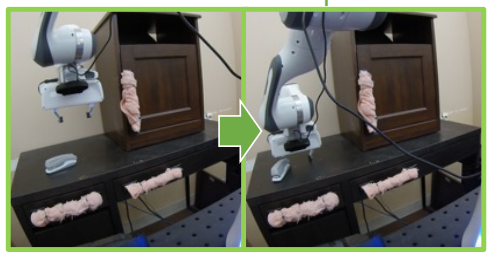
Learn ***language-conditioned visuomotor manipulation skills*** from:
offline datasets +
crowd-sourced annotation

Offline Robot Data

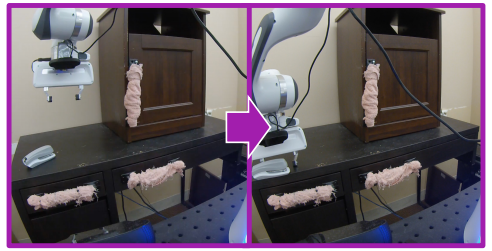


Language-Conditioned Task Completion

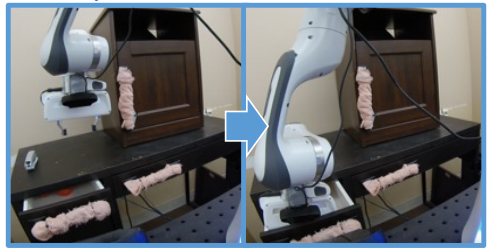
"Move the stapler"



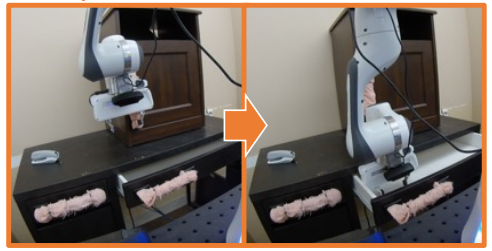
"Push the small gray stapler around on top of the black desk"



"Open the left drawer"



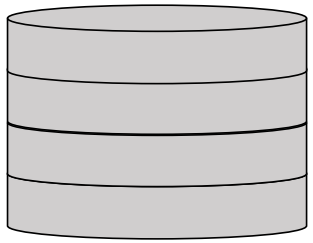
"Open the right drawer"



Our Objective

Learn ***language-conditioned visuomotor manipulation skills*** from:
offline datasets +
crowd-sourced annotation

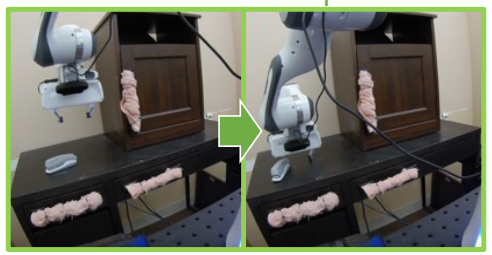
Offline Robot Data



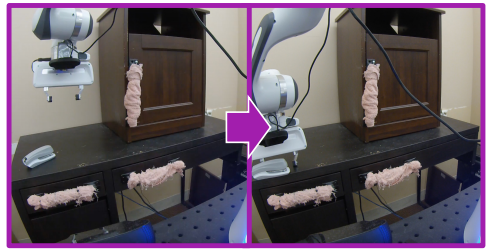
Crowd-Sourced Annotation

Language-Conditioned Task Completion

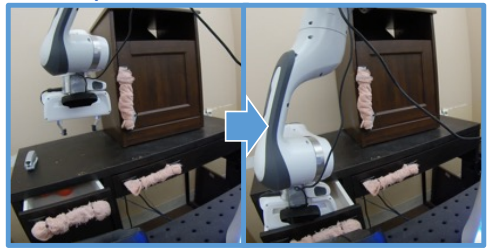
"Move the stapler"



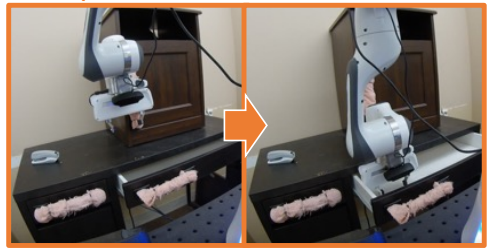
"Push the small gray stapler around on top of the black desk"



"Open the left drawer"



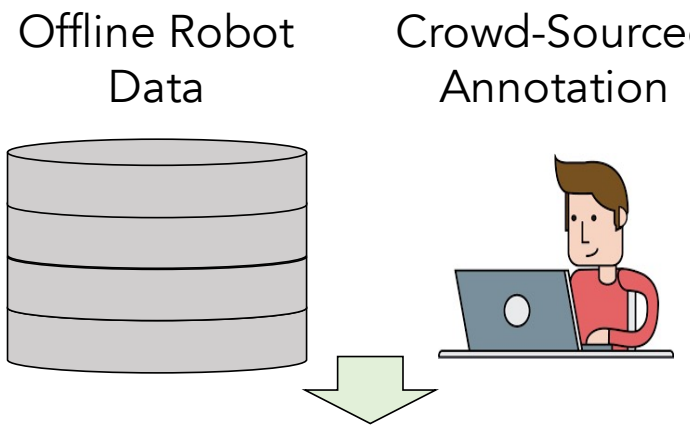
"Open the right drawer"



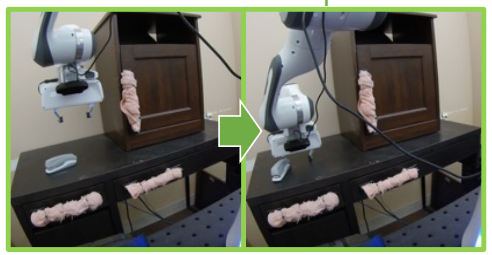
Our Objective

Learn ***language-conditioned visuomotor manipulation skills*** from:
offline datasets +
crowd-sourced annotation

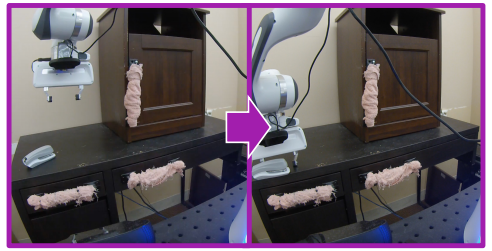
Offline data:
- Highly suboptimal
- From autonomous exploration /existing replay buffers



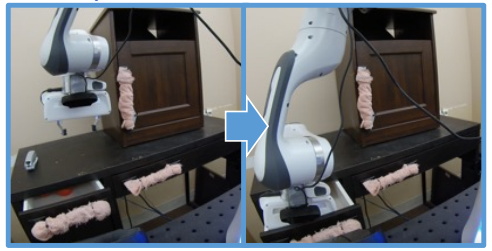
"Move the stapler"



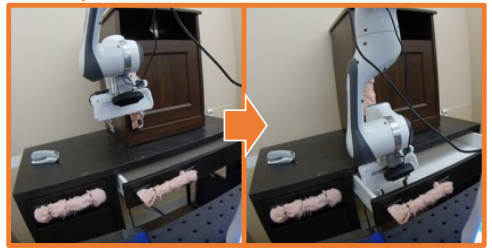
"Push the small gray stapler around on top of the black desk"



"Open the left drawer"



"Open the right drawer"

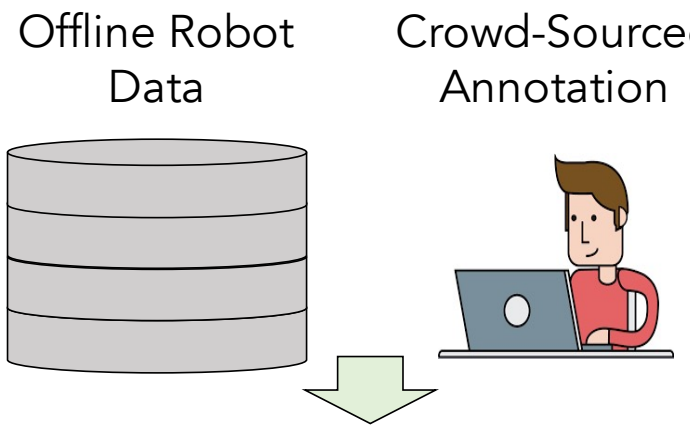


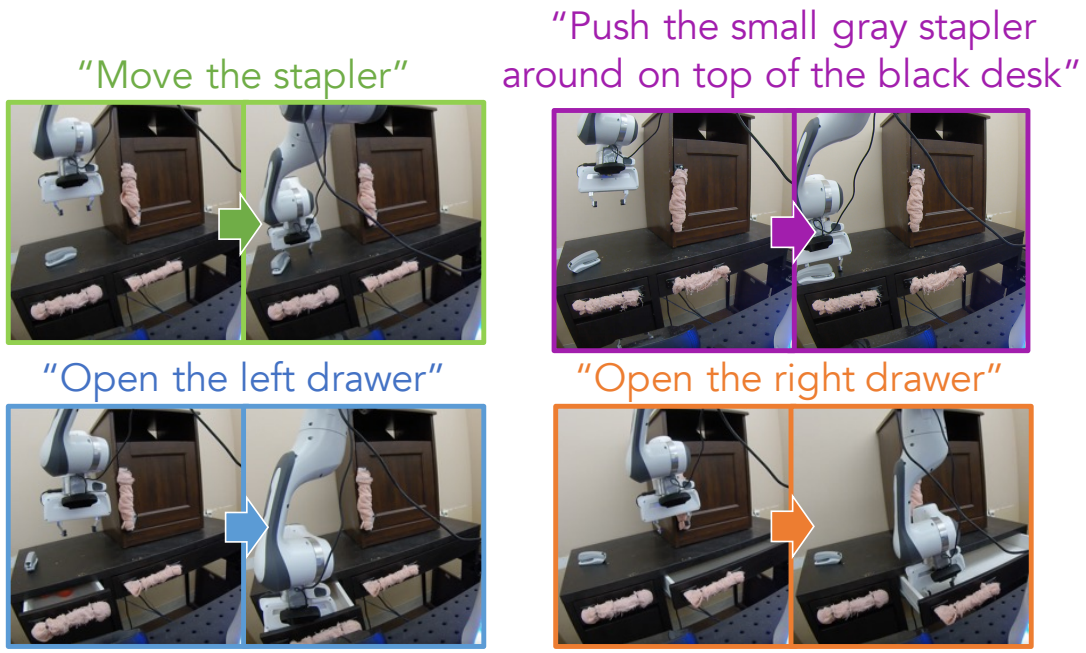
Our Objective

Learn ***language-conditioned visuomotor manipulation skills*** from:
offline datasets +
crowd-sourced annotation

Offline data:
- Highly suboptimal
- From autonomous exploration /existing replay buffers

Crowd-sourced annotation:
- Describes what task (if any) is completed





Key Idea

Actions in data cannot be treated as optimal

Key Idea

Actions in data cannot be treated as optimal

start → ***finish*** of episode
completes the annotated instruction

Key Idea

Actions in data cannot be treated as optimal

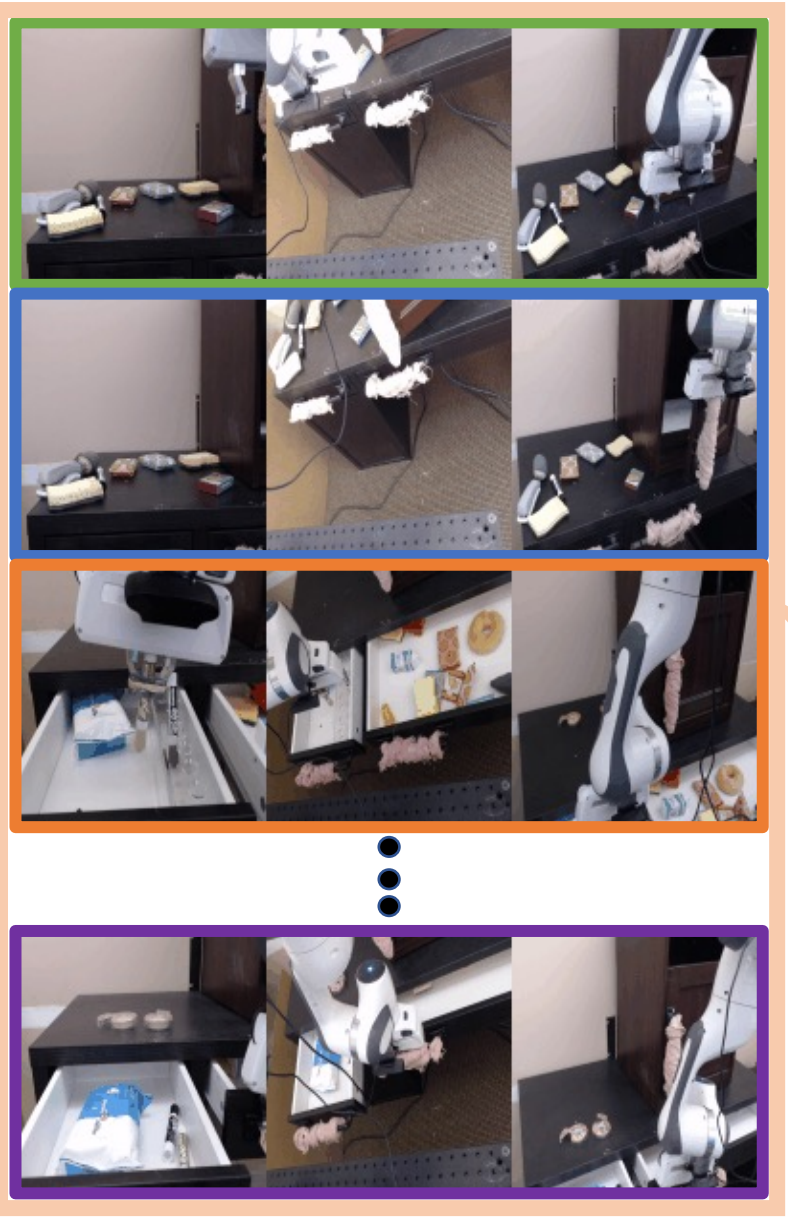
start → ***finish*** of episode
completes the annotated instruction

Learn ***language-conditioned reward function*** for offline RL

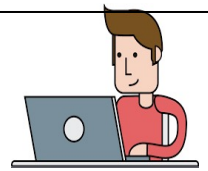
Offline Data

View 1 View 2 View 3

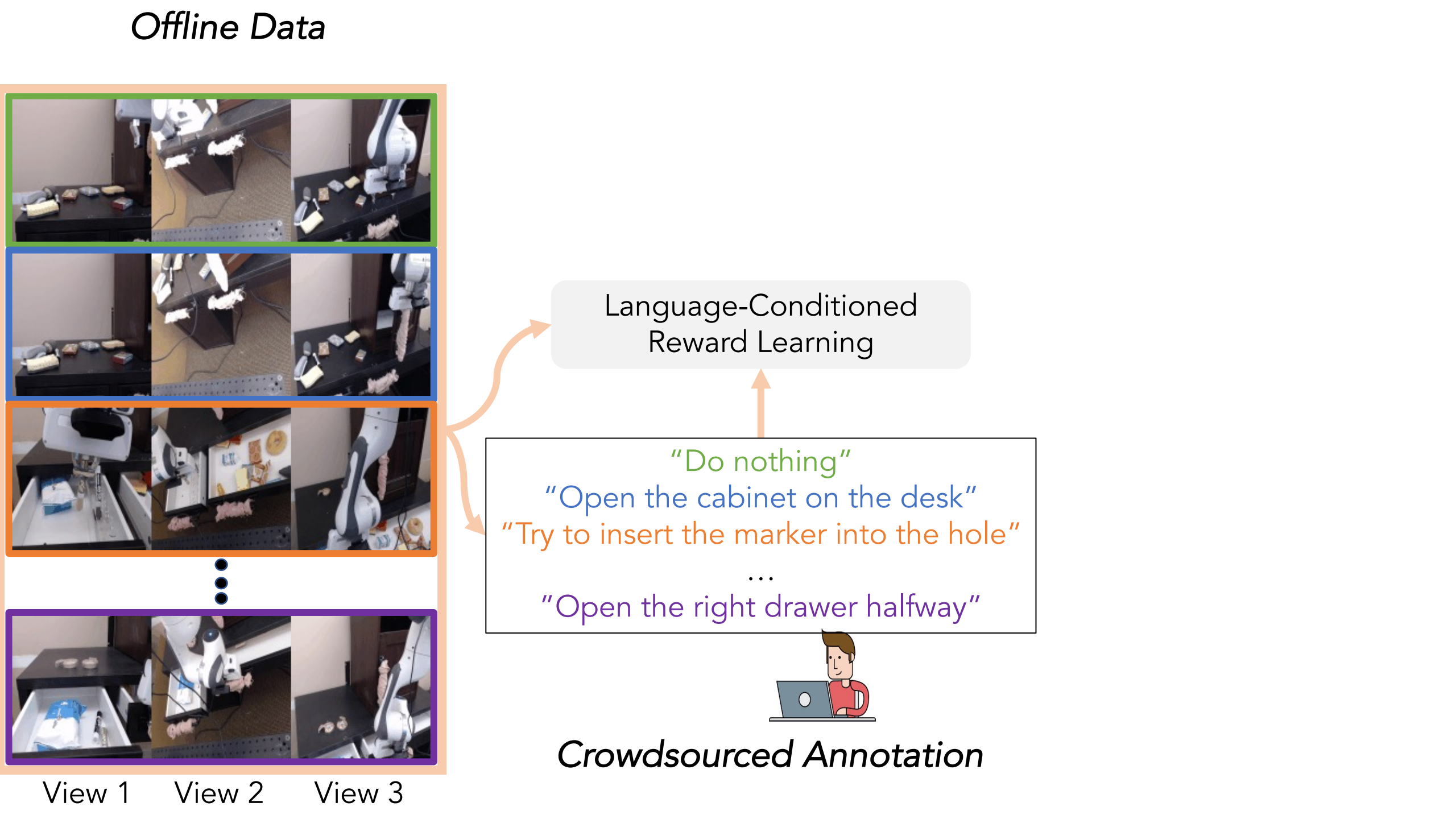
Offline Data

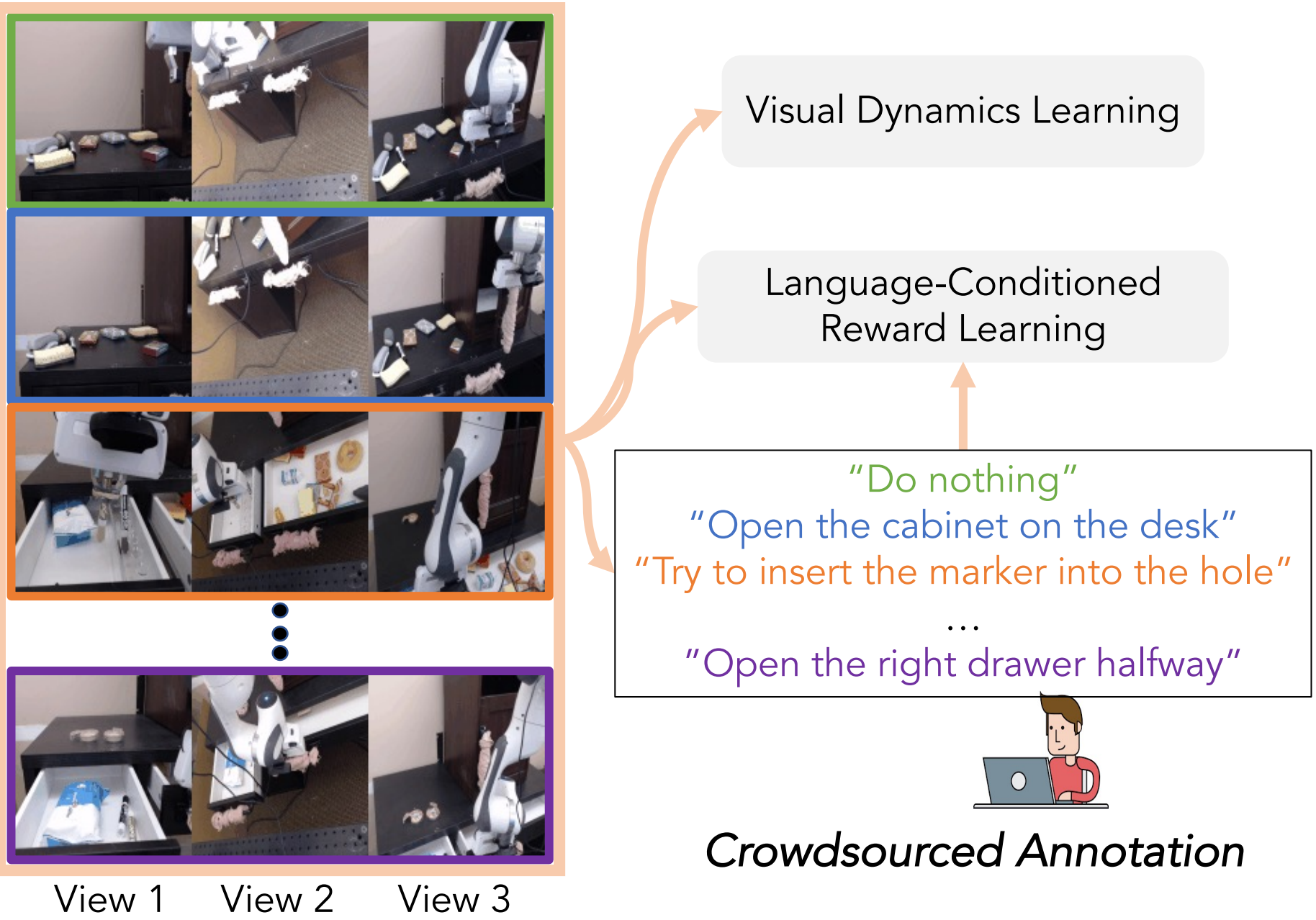


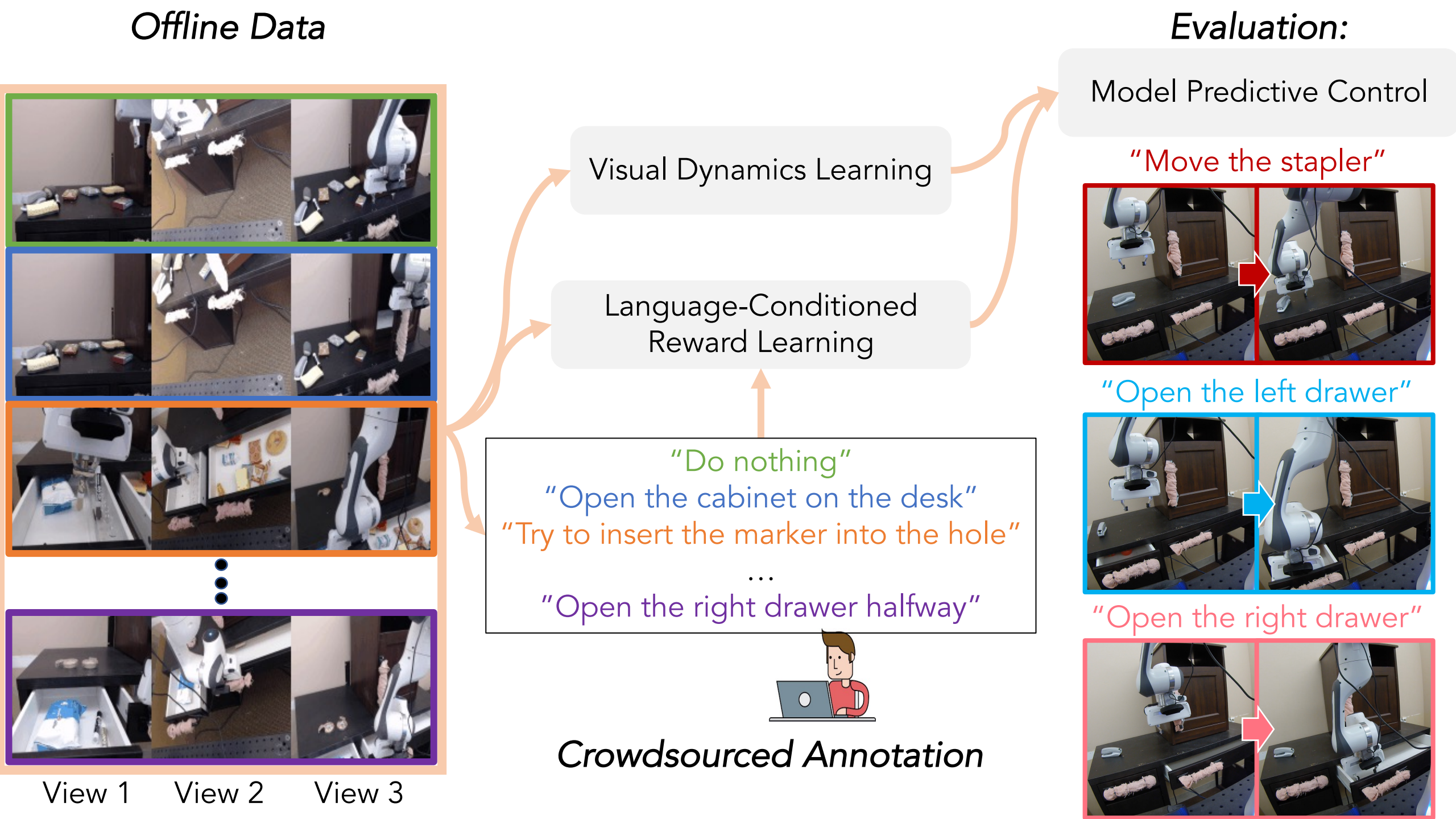
"Do nothing"
"Open the cabinet on the desk"
"Try to insert the marker into the hole"
...
"Open the right drawer halfway"



Crowdsourced Annotation


Offline Data
View 1
View 2
View 3
Language-Conditioned
Reward Learning
"Do nothing"
"Open the cabinet on the desk"
"Try to insert the marker into the hole"
...
"Open the right drawer halfway"
Crowdsourced Annotation


Offline Data
Visual Dynamics Learning
Language-Conditioned Reward Learning
"Do nothing"
"Open the cabinet on the desk"
"Try to insert the marker into the hole"
...
"Open the right drawer halfway"
Crowdsourced Annotation
View 1
View 2
View 3


Offline Data
View 1
View 2
View 3
Visual Dynamics Learning
Language-Conditioned Reward Learning
"Do nothing"
"Open the cabinet on the desk"
"Try to insert the marker into the hole"
...
"Open the right drawer halfway"
Crowdsourced Annotation
Evaluation:
Model Predictive Control
"Move the stapler"
"Open the left drawer"
"Open the right drawer"

Language-Conditioned Offline Reward Learning (LOReL)

Learn binary classifier of instruction completion

Instruction: "Open the cabinet on the desk"

s_0

s_T

LOReL

Label = 1

Language-Conditioned Offline Reward Learning (LOReL)

Learn binary classifier of instruction completion

Positives:
- Episode and corresponding annotations

Instruction: "Open the cabinet on the desk"

s_0

s_T

LOReL

Label = 1

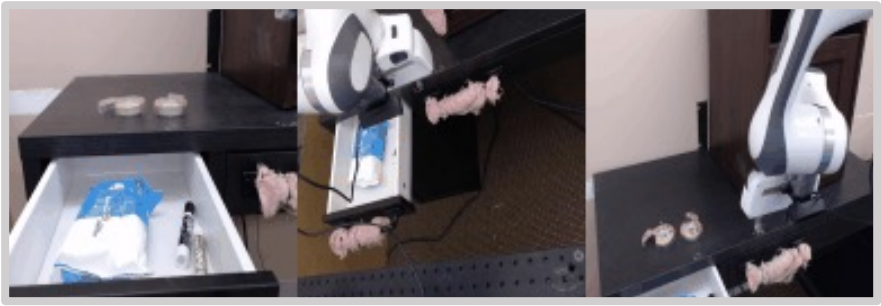
Language-Conditioned Offline Reward Learning (LOReL)

Learn binary classifier of instruction completion

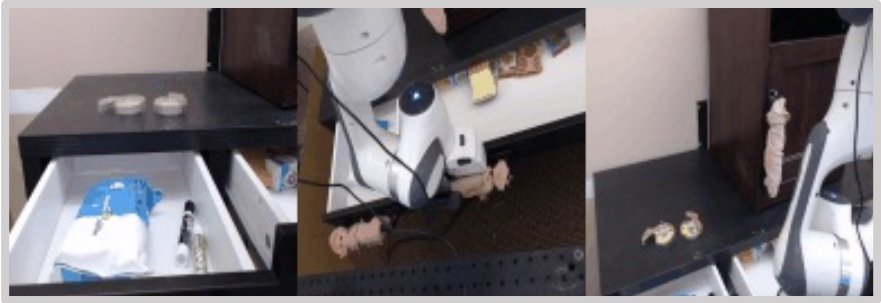
Negatives:
- Different episodes

Instruction: "Open the cabinet on the desk"

s_0



s_T



LOReL

Label = 0

Language-Conditioned Offline Reward Learning (LOReL)

Learn binary classifier of instruction completion

Negatives:
- Different episodes
- Episode reversed

Instruction: "Open the cabinet on the desk"

s_0

s_T

LOReL

Label = 0

Language-Conditioned Offline Reward Learning (LOReL)

Learn binary classifier of instruction completion

Pretrained language models to enable generalization from limited data

Instruction: "Open the cabinet on the desk"

s_0

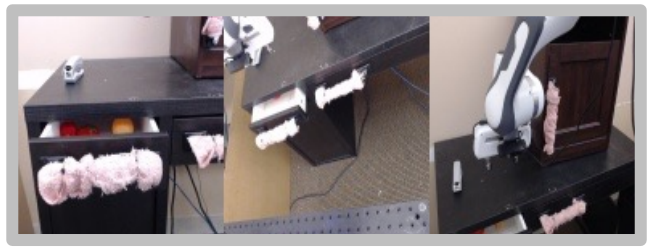
s_T

LOReL

Label = 0

Language-Conditioned Offline Reward Learning (LOReL)

Initial State

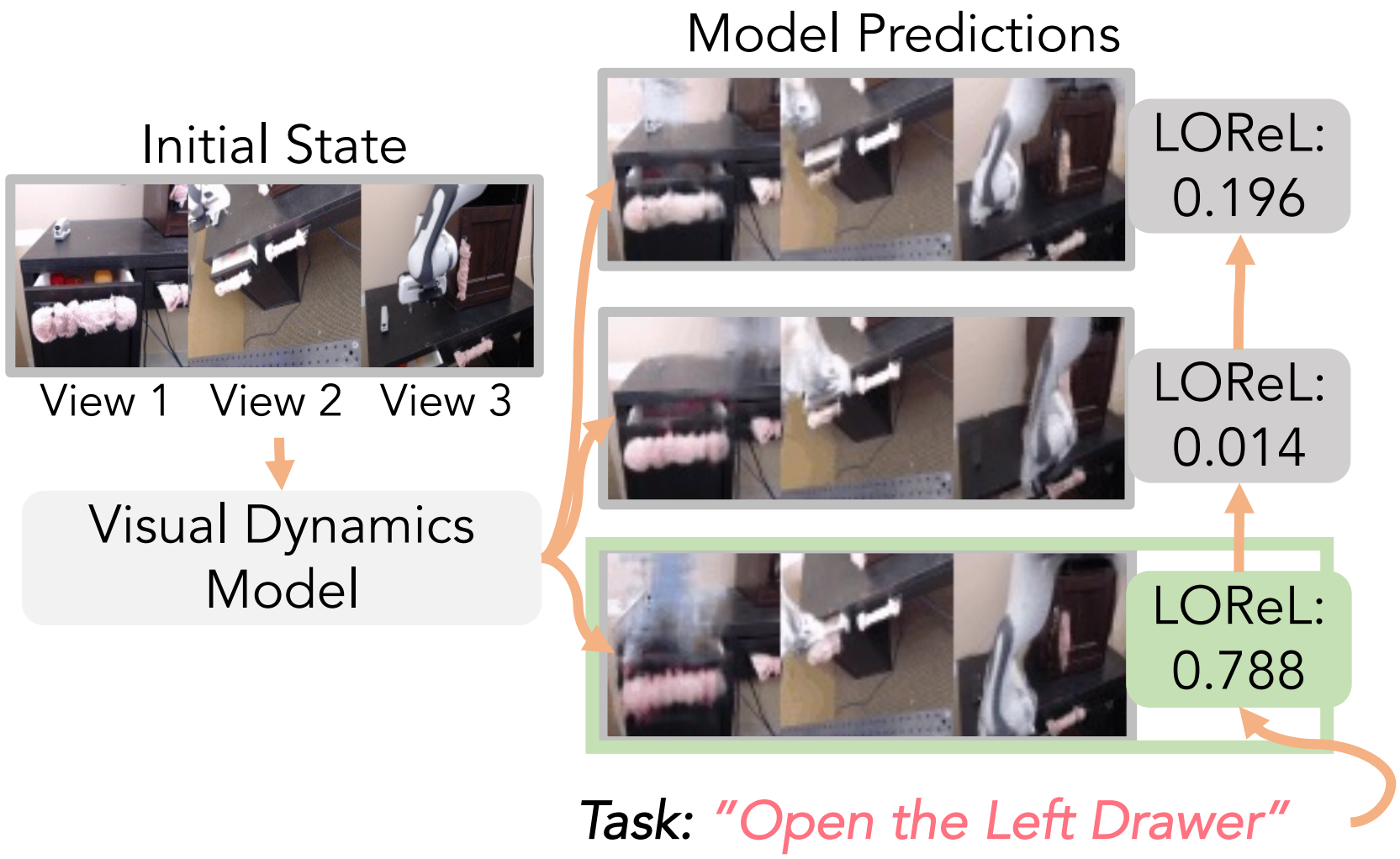


View 1 View 2 View 3

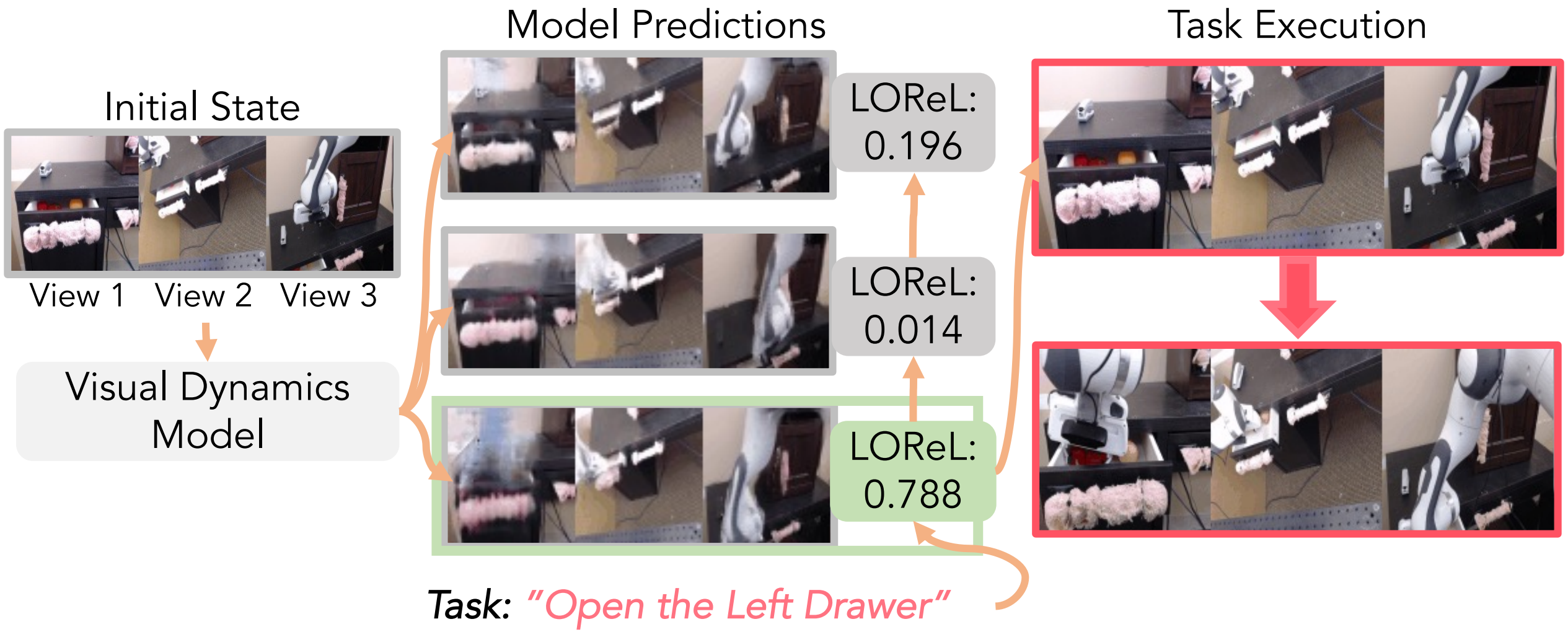
Visual Dynamics Model

Model Predictions

Language-Conditioned Offline Reward Learning (LOReL)



Language-Conditioned Offline Reward Learning (LOReL)



Experiment 1:
How does LOReL enable effective language-conditioned behavior compared to prior methods?

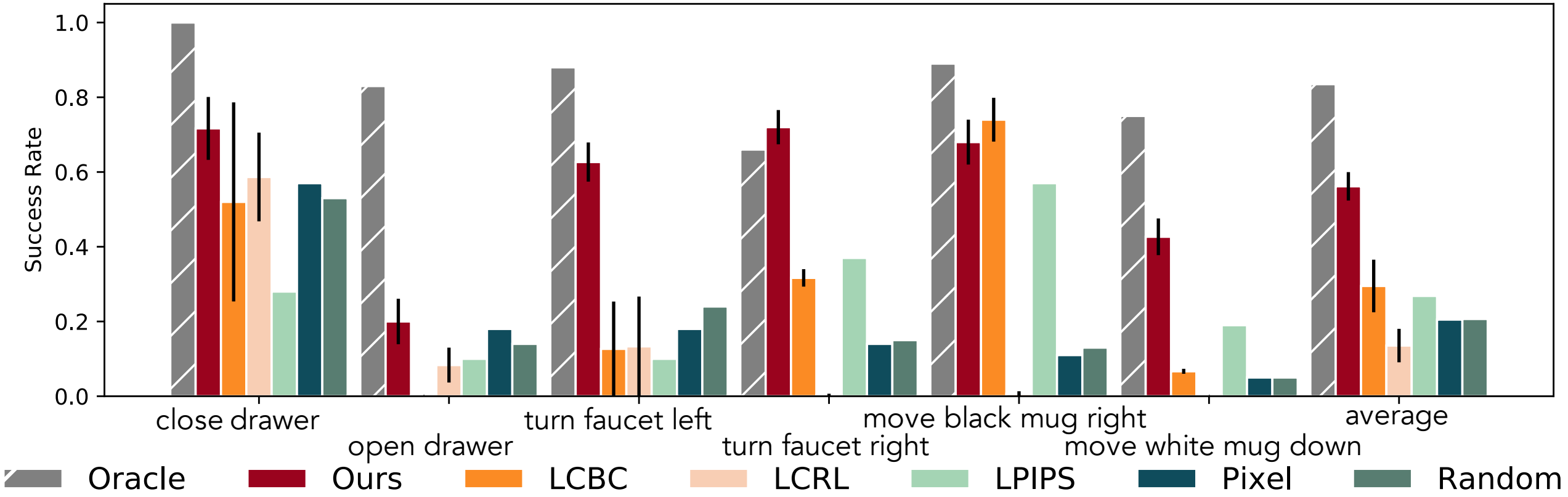
Simulated Environment

Procedurally Generated Annotations

- Close the drawer
- Move the white mug up and turn the faucet left
- Open the drawer and turn the faucet right
- Do nothing
- Move the white mug right
- Move the black mug down and left

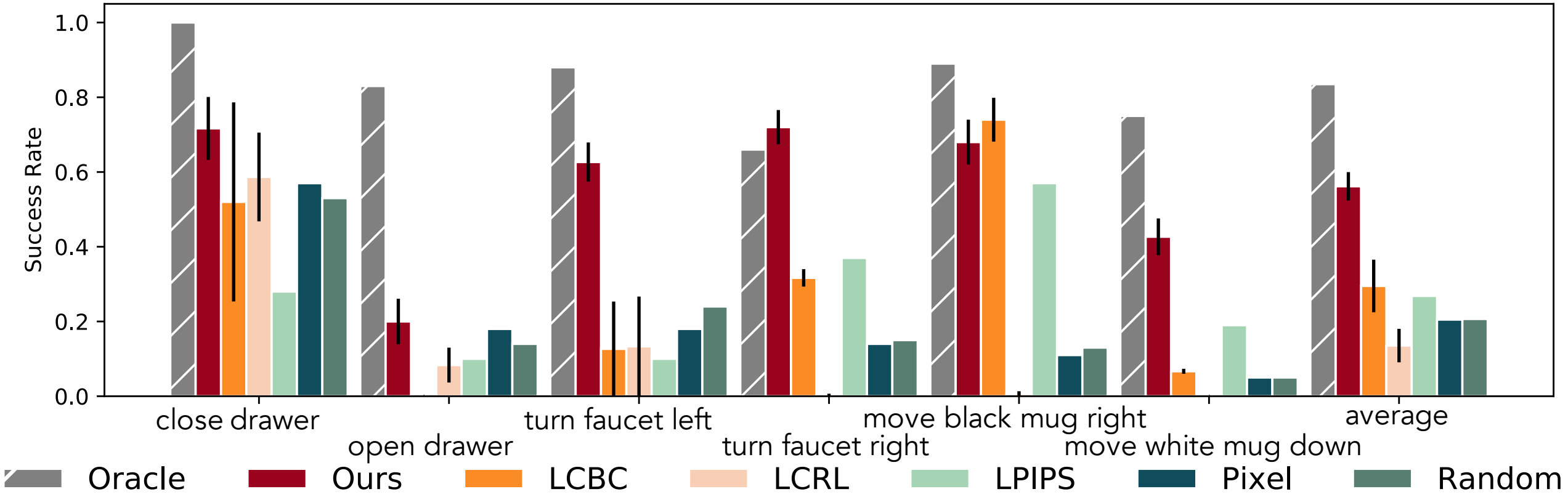
⋮

Experiment 1:
How does LOReL enable effective language-conditioned behavior compared to prior methods?



- *LOReL* outperforms *language conditioned imitation learning/RL* by *25+%*

Experiment 1:
How does LOReL enable effective language-conditioned behavior compared to prior methods?



- *LOReL* outperforms *language conditioned imitation learning/RL* by *25+%*
- *LOReL* outperforms goal image *Pixel/LPIPS* specification by *29+%*

Experiment 2:
Can LOReL generalize to unseen natural language commands?

Success Rate	Ours	Ours (-PM)
Original	**56 ± 1%**	40 ± 1%
Unseen Verb	**51 ± 3%**	33 ± 2%
Unseen Noun	**51 ± 1%**	39 ± 4%
Unseen Verb + Noun	**47 ± 2%**	17 ± 3%
Unseen Natural Language	**46 ± 2%**	21 ± 1%

- LOReL can generalize to unseen instructions effectively (<10% drop in performance)
- Leveraging a pretrained model is essential for this generalization

Instruction	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
Seen	Close drawer	Open drawer	Turn faucet left	Turn faucet right	Move black mug right	Move white mug down
Unseen Verb	shut drawer	pull drawer	rotate faucet left	rotate faucet right	push black mug right	push white mug down
Unseen Noun	close container	open container	turn tap left	turn tap right	move dark cup right	move light cup down
Unseen Verb+ Noun	shut container	pull container	rotate tap left	rotate tap right	push dark cup right	push light cup down
Human Provided	Push the drawer shut Push the drawer Shut the drawer shut drawer Slide the drawer closed Shut the drawer shut the dresser Shut the drawer. Shut the cupboard	Pull the drawer open Pull the handle Pull the drawer handle Pull the drawer open Pull open the drawer open the dresser Pull the drawer. Unclose the cabinet	Rotate the tap counterclock wise Turn faucet away from camera Rotate nozzle left faucet counterclock wise Rotate the faucet left Turn the faucet to the left rotate handle to the left Turn faucet counter clockwise. Spin nozzle left	Rotate tap clockwise Turn faucet towards camera Rotate nozzle right faucet clockwise Rotate the faucet right Turn the faucet to the right rotate handle rightward Turn faucet clockwise. Twirl valve right	Translate the black cup to the right Move black mug away from drawer Push black cup right black mug right Slide the black mug right Move the dark mug to the right push black cup right Move black mug right. Shift dark cup right	Translate the white cup down Move with mug closer to the faucet Bring white cup down white mug down Push the white mug down and left Move the lighter mug down shift white mug down Pull white mug to the front. Reposition white glass down

Experiment 3:
Is LOReL effective on a real robot?

Robot Domain

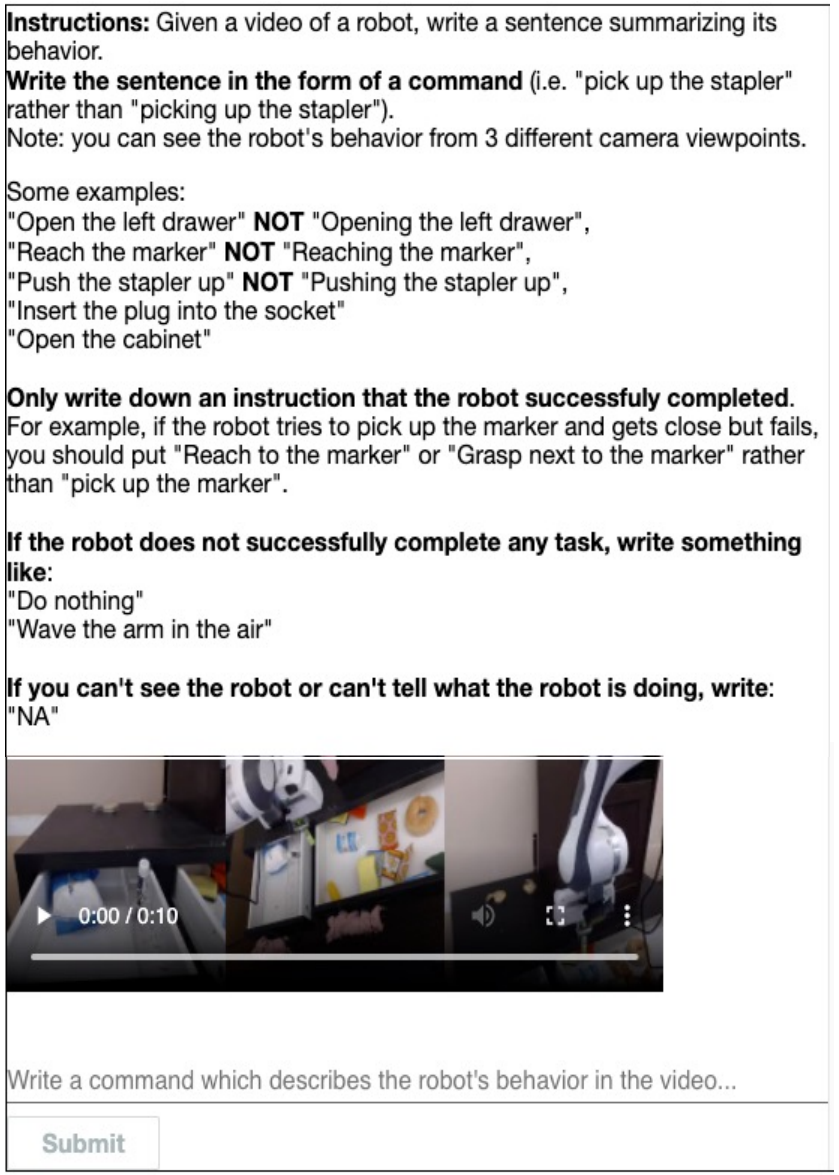
3000 episodes taken from concurrent work performing online RL

Replay buffer consists of successful and unsuccessful attempts of many skills

2 annotations per episodes collected using AMT

Experiment 3:
Is LOReL effective on a real robot?

Robot Domain



Experiment 3:
Is LOReL effective on a real robot?

Robot Domain

Crowd-sourced Annotations

Insert pin in object
Reach item in left drawer.
close the drawer
Rub drawer handle
Do nothing
open the top left drawer
Reach to the dark green marker.
pull the drawer out
touch the tabletop
touch surface of table
Hit the rack with the marker

⋮

Experiment 3:
Is LOReL effective on a real robot?

Task (10 Trials Each)	LOReL
"Open the left drawer"	**90%**

"Open the left drawer"

Experiment 3:
Is LOReL effective on a real robot?

Task (10 Trials Each)	LOReL
"Open the left drawer"	**90%**
"Open the right drawer"	**40%**

"Open the right drawer"

Experiment 3:
Is LOReL effective on a real robot?

Task (10 Trials Each)	LOReL
"Open the left drawer"	**90%**
"Open the right drawer"	**40%**
"Move the stapler"	**50%**

"Move the stapler"

Experiment 3:
Is LOReL effective on a real robot?

Task (10 Trials Each)	LOReL
"Open the left drawer"	**90%**
"Open the right drawer"	**40%**
"Move the stapler"	**50%**
"Reach the marker"	**70%**

"Reach the marker"

Experiment 3:
Is LOReL effective on a real robot?

Task (10 Trials Each)	LOReL
"Open the left drawer"	**90%**
"Open the right drawer"	**40%**
"Move the stapler"	**50%**
"Reach the marker"	**70%**
"Reach the cabinet"	**80%**
Average over tasks	**66%**

"Reach the cabinet"

Robustness to Instruction Complexity

"Open the small black and white drawer on the left fully"

70% Success Rate

"Push the small gray stapler around on top of the black desk"

50% Success Rate

Key Takeaways

To learn useful multi-task robot policies humans need effective task-specification like natural language

Key Takeaways

To learn useful multi-task robot policies humans need effective task-specification like natural language

By annotating sub-optimal offline data with crowd-sourced instructions, LOReL can learn language-conditioned visuomotor skills on a real robot

Key Takeaways

To learn useful multi-task robot policies humans need effective task-specification like natural language

By annotating sub-optimal offline data with crowd-sourced instructions, LOReL can learn language-conditioned visuomotor skills on a real robot

By using pre-trained language models LOReL can generalize to unseen natural language instructions + handle crowdsourced annotations