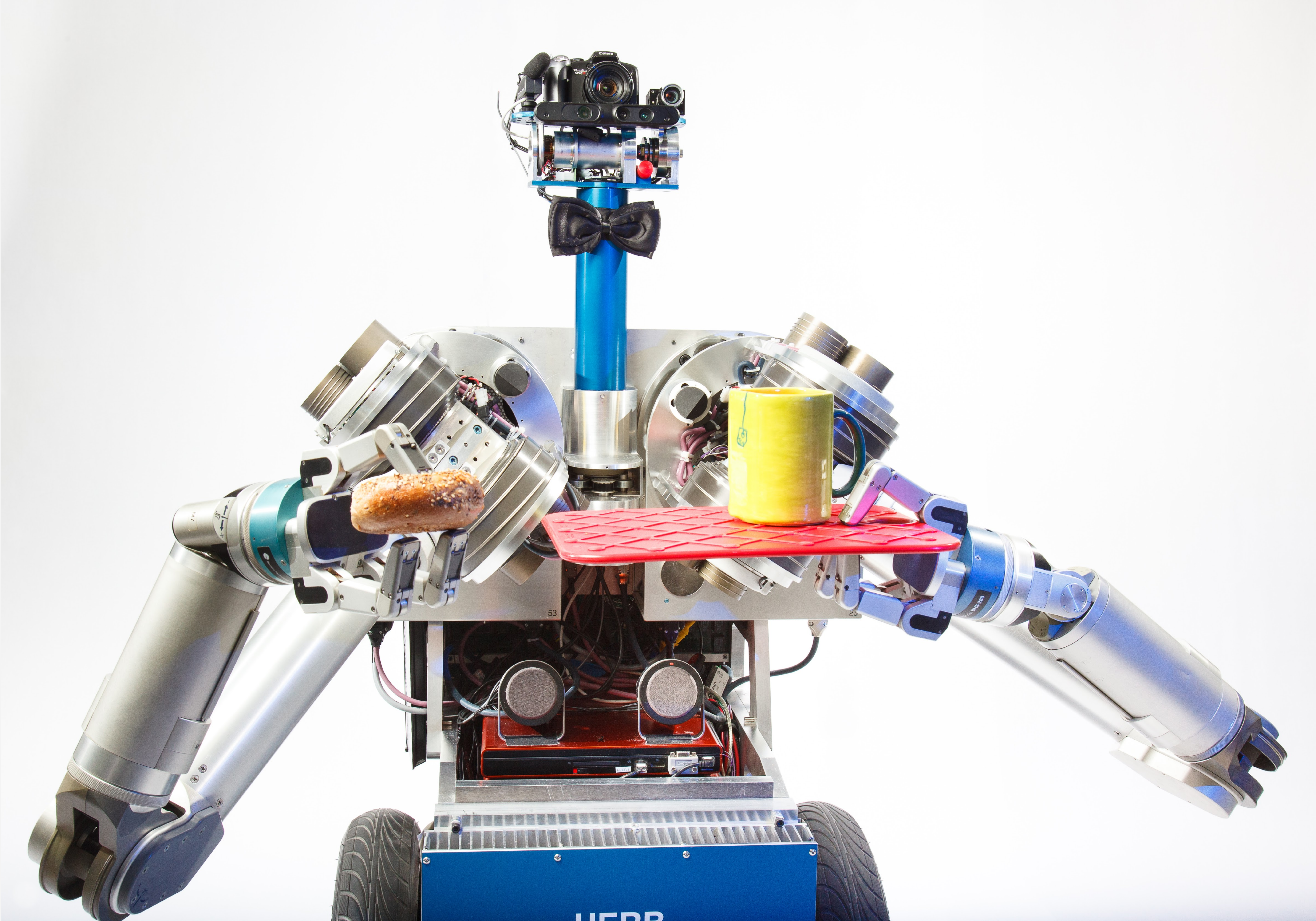
Software Engineering Institute
Carnegie Mellon

Automatically Evaluating and Generating Clear Robot Explanations

Shen Li

Thesis committee:
Dr. Siddhartha Srinivasa (co-chair)
Dr. Stephanie Rosenthal (co-chair)
Dr. Reid Simmons
Stefanos Nikolaidis



Making it easier for humans to **understand** robots.

Important to understand robots

Important to understand robots

Seamless
Efficient
Collaboration

Pellegrinelli, S., Admoni, H., Javdani, S., & Srinivasa, S. Human-Robot SharedWorkspace Collaboration via Hindsight Optimization. IROS. 2016.

It is critical to **Understand** robots.

Important to understand robots

Physical Conflict

Pellegrinelli, S., Admoni, H., Javdani, S., & Srinivasa, S. Human-Robot SharedWorkspace Collaboration via Hindsight Optimization. IROS. 2016.

Important to understand robots

Mental Conflict

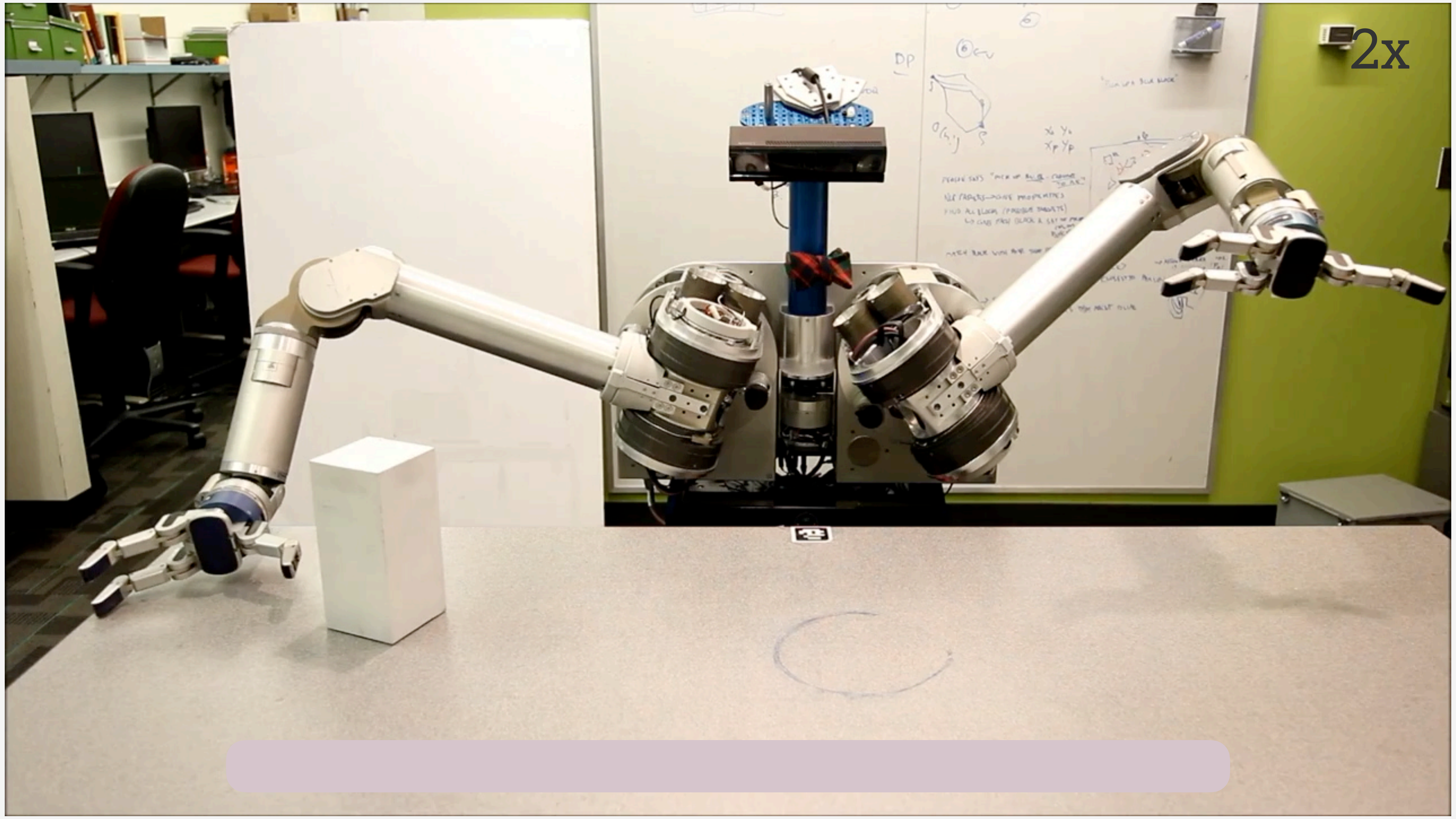
Pellegrinelli, S., Admoni, H., Javdani, S., & Srinivasa, S. Human-Robot SharedWorkspace Collaboration via Hindsight Optimization. IROS. 2016.
Adrian Bussone, Simone Stumpf, and Dympna O'Sullivan. The role of explanations on trust and reliance in clinical decision support sys- tems. ICHI. 2015.

It is **Critical** to understand robots.

It is **Not** easy to
understand robots.

It is **Not** easy to understand robots.

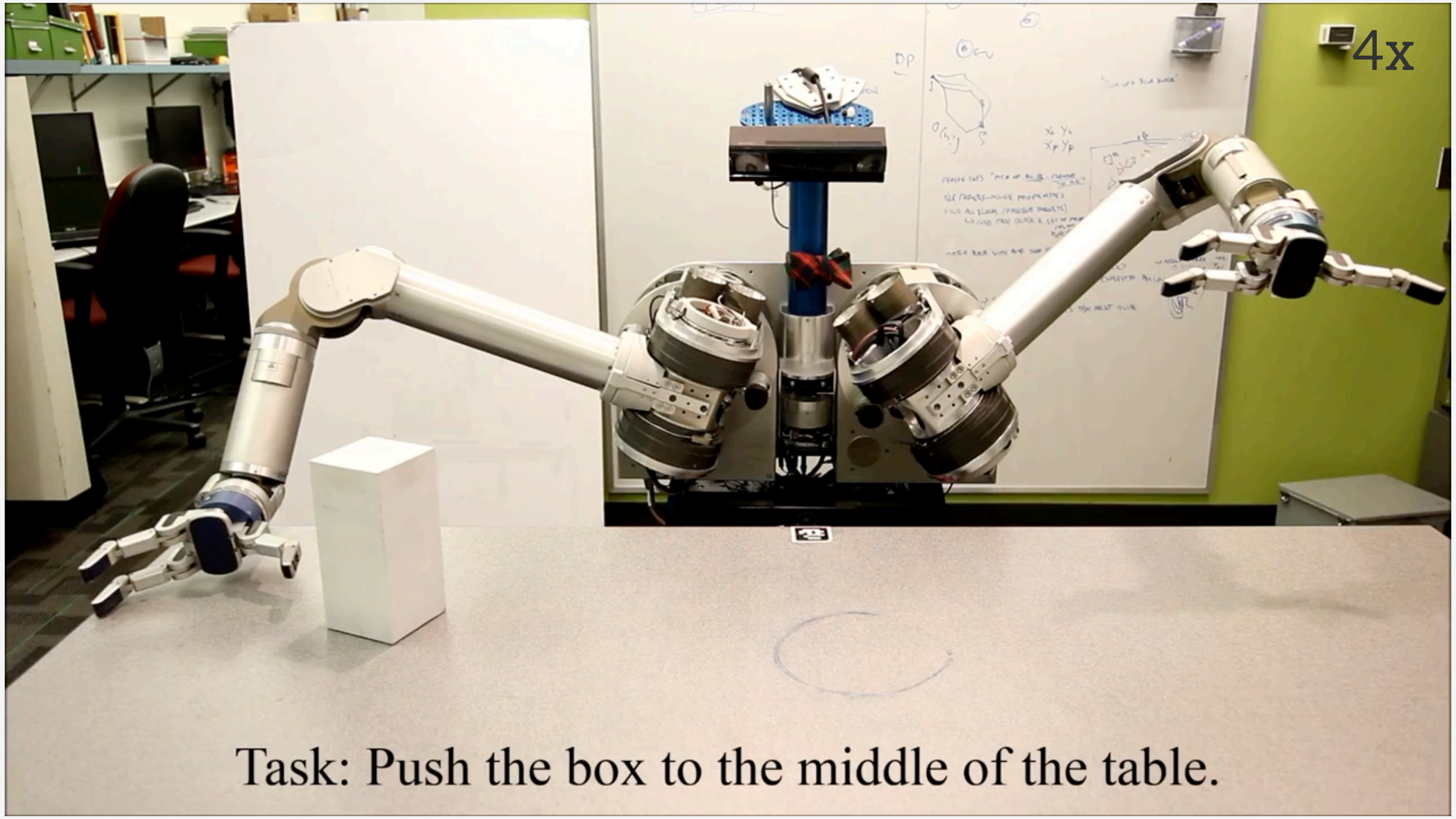
Not easy to understand robots



???

Johnson, A. M., King, J. E., & Srinivasa, S. Convergent planning. IEEE Robotics and Automation Letters. 2016.

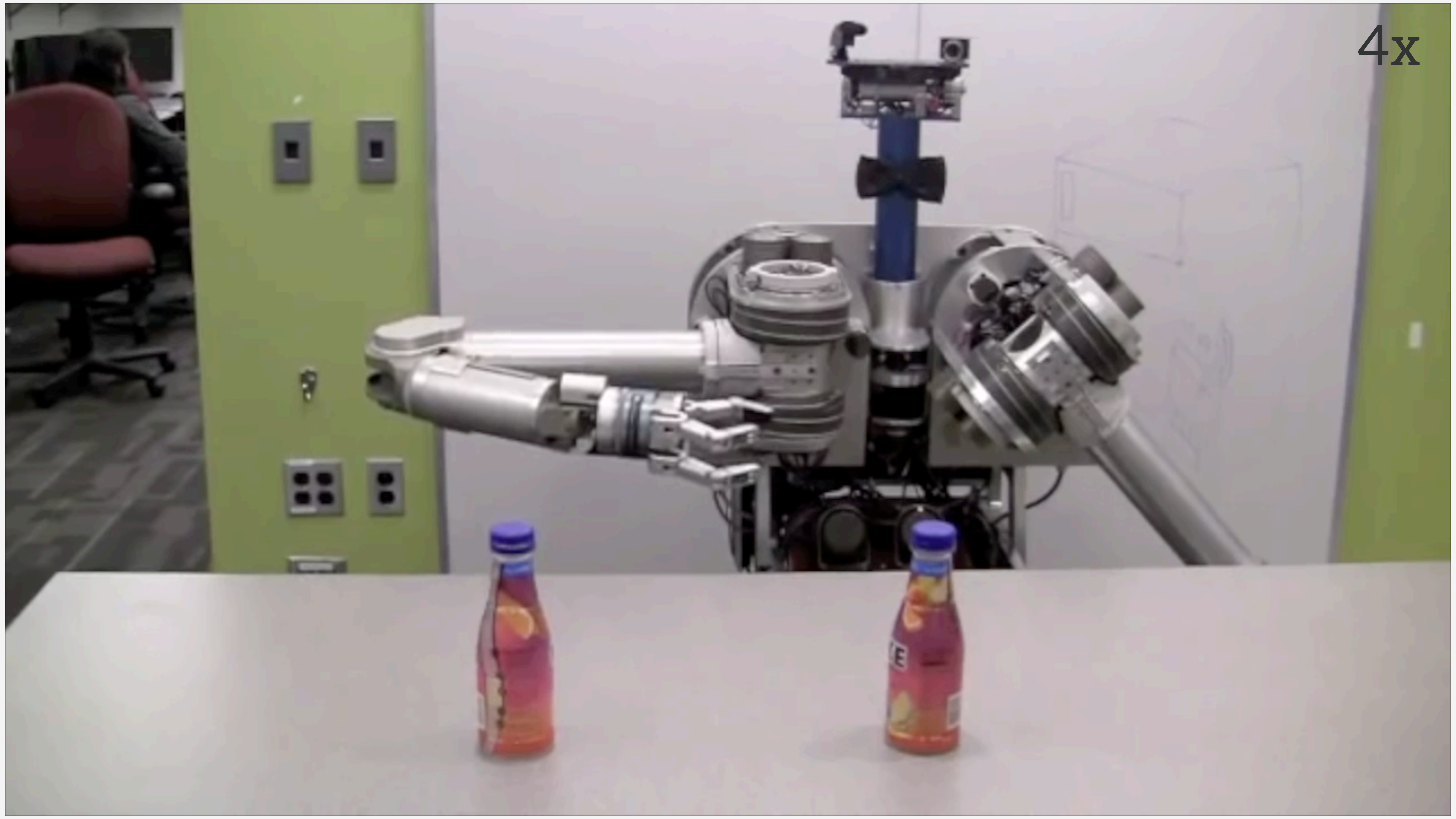
Not easy to understand robots



Robot Intention

Johnson, A. M., King, J. E., & Srinivasa, S. Convergent planning. IEEE Robotics and Automation Letters. 2016.

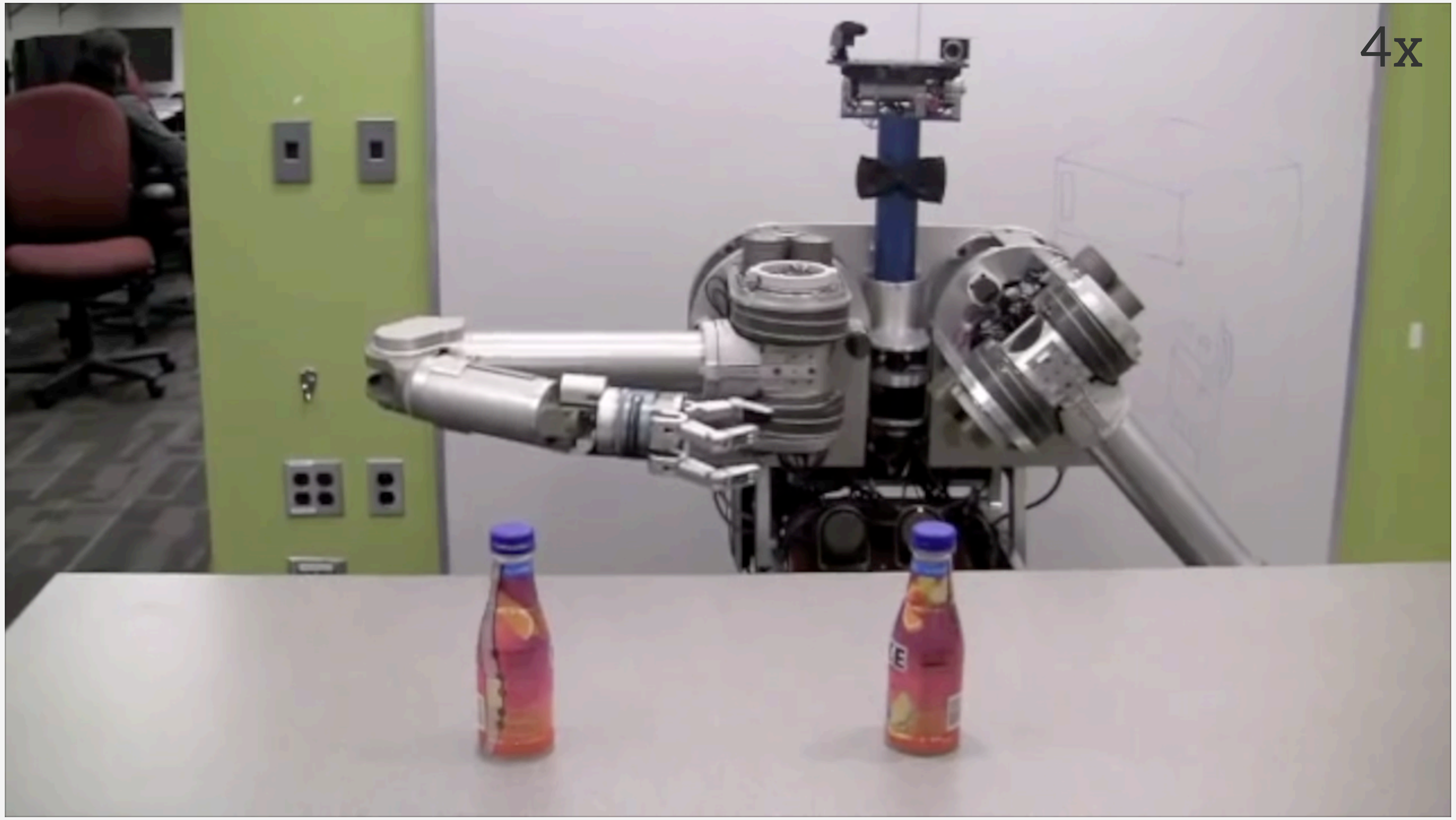
Not easy to understand robots



???

Dragan, A. D., Holladay, R. M., & Srinivasa, S. S. An Analysis of Deceptive Robot Motion. RSS, 2014

Not easy to understand robots



Robot Intention

Dragan, A. D., Holladay, R. M., & Srinivasa, S. S. An Analysis of Deceptive Robot Motion. RSS, 2014

Not easy to understand robots

Dogar, M., & Srinivasa, S. A framework for push-grasping in clutter. RSS, 2011

Not easy to understand robots

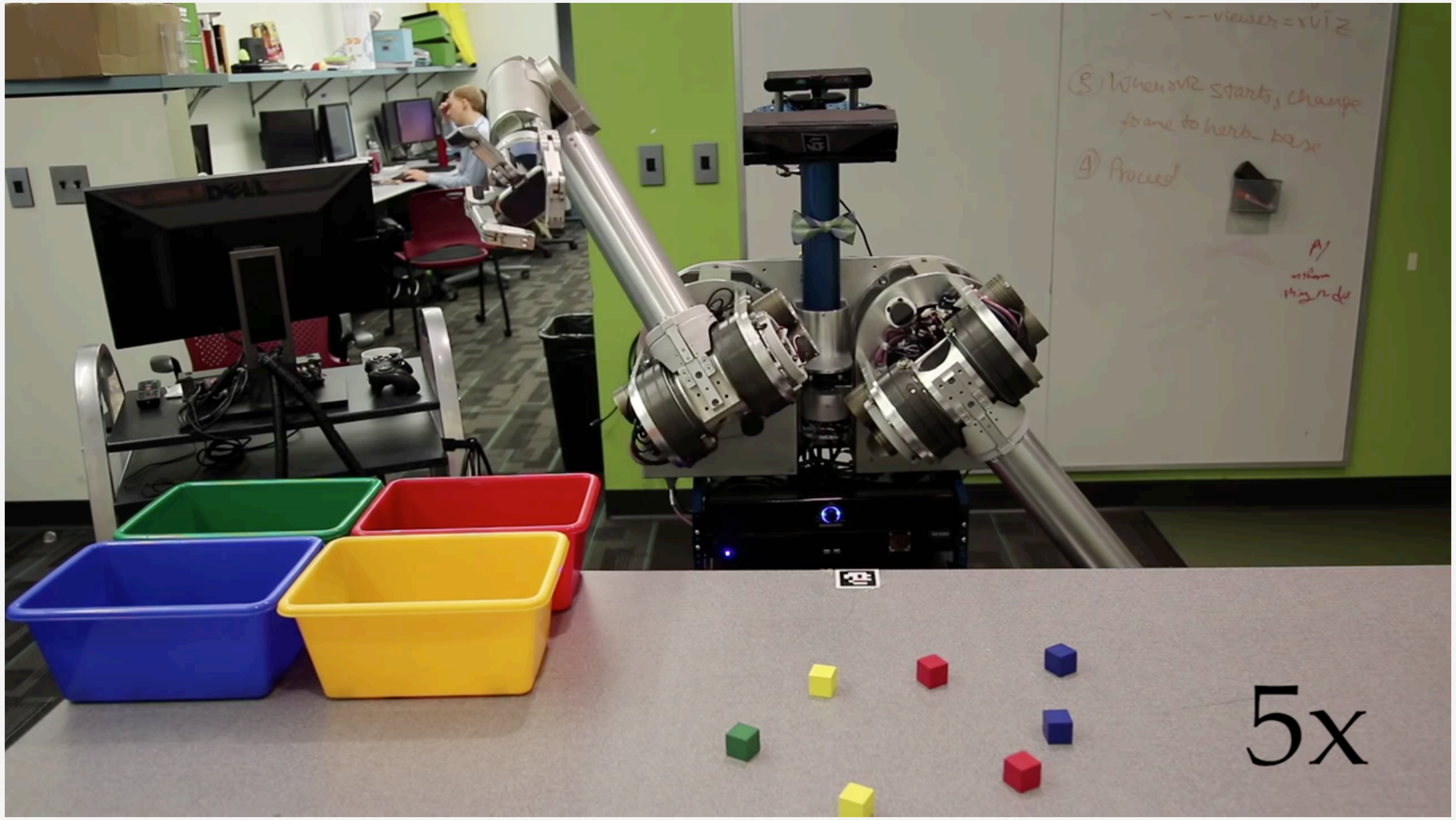
Dogar, M., & Srinivasa, S. A framework for push-grasping in clutter. RSS, 2011

Not easy to understand robots

Robot Intention

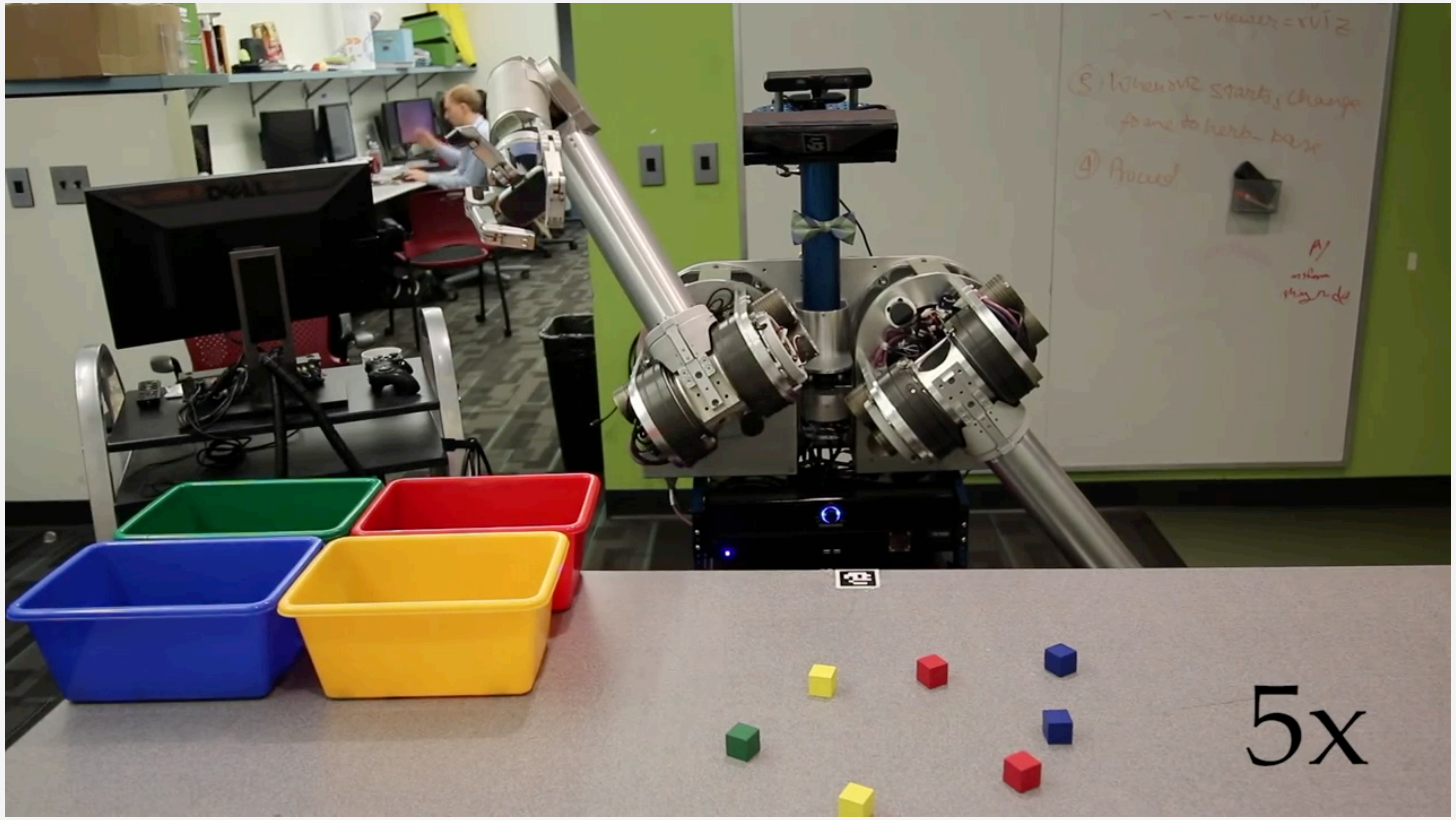
Dogar, M., & Srinivasa, S. A framework for push-grasping in clutter. RSS, 2011

Not easy to understand robots



HERB Sorts Colored Blocks

Not easy to understand robots



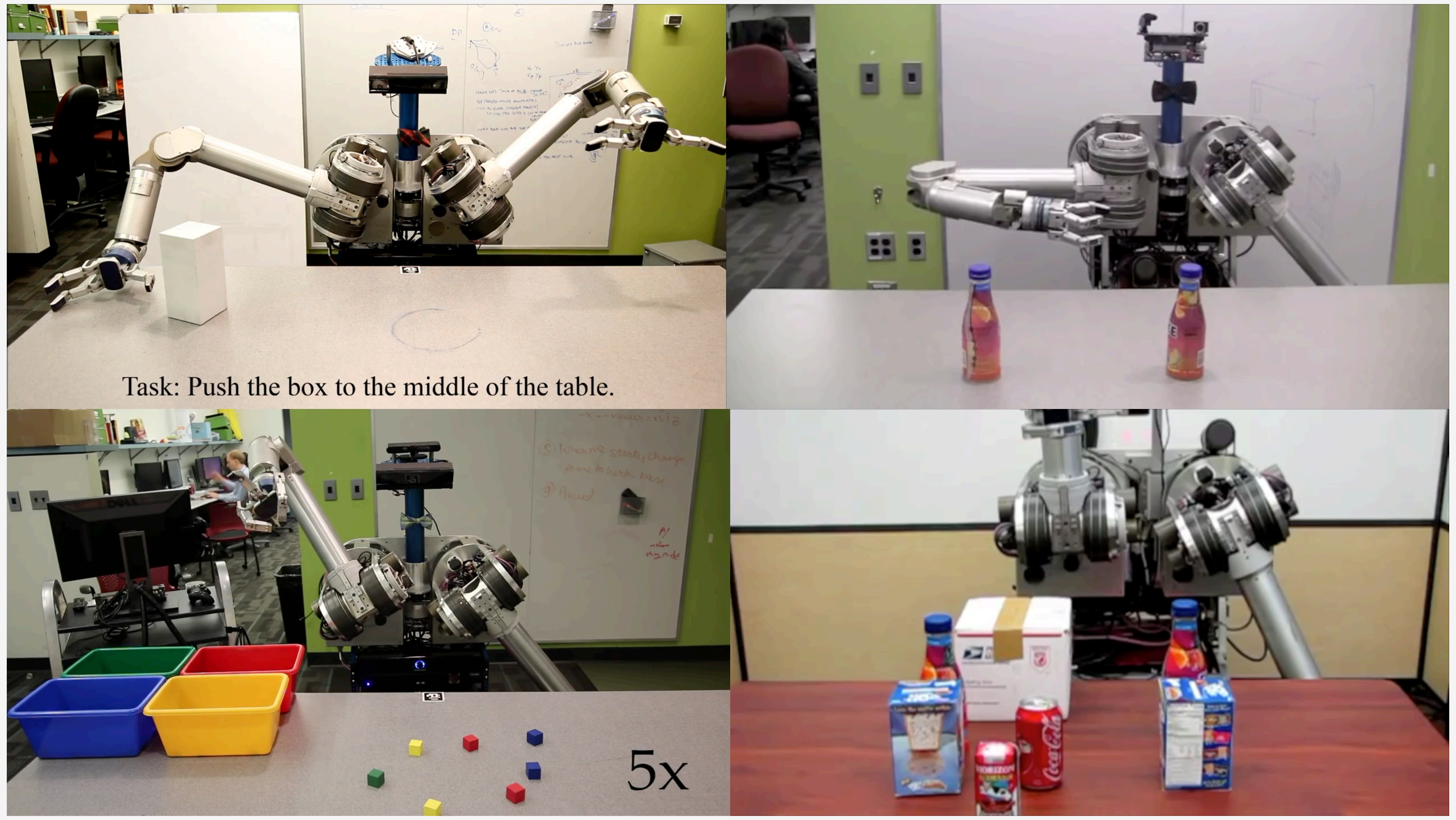
HERB Sorts Colored Blocks

Robot Intention

Not easy to understand robots

Robot intention

Understanding robot intentions helps people understand and anticipate robot behavior based on the **rationality principle**.

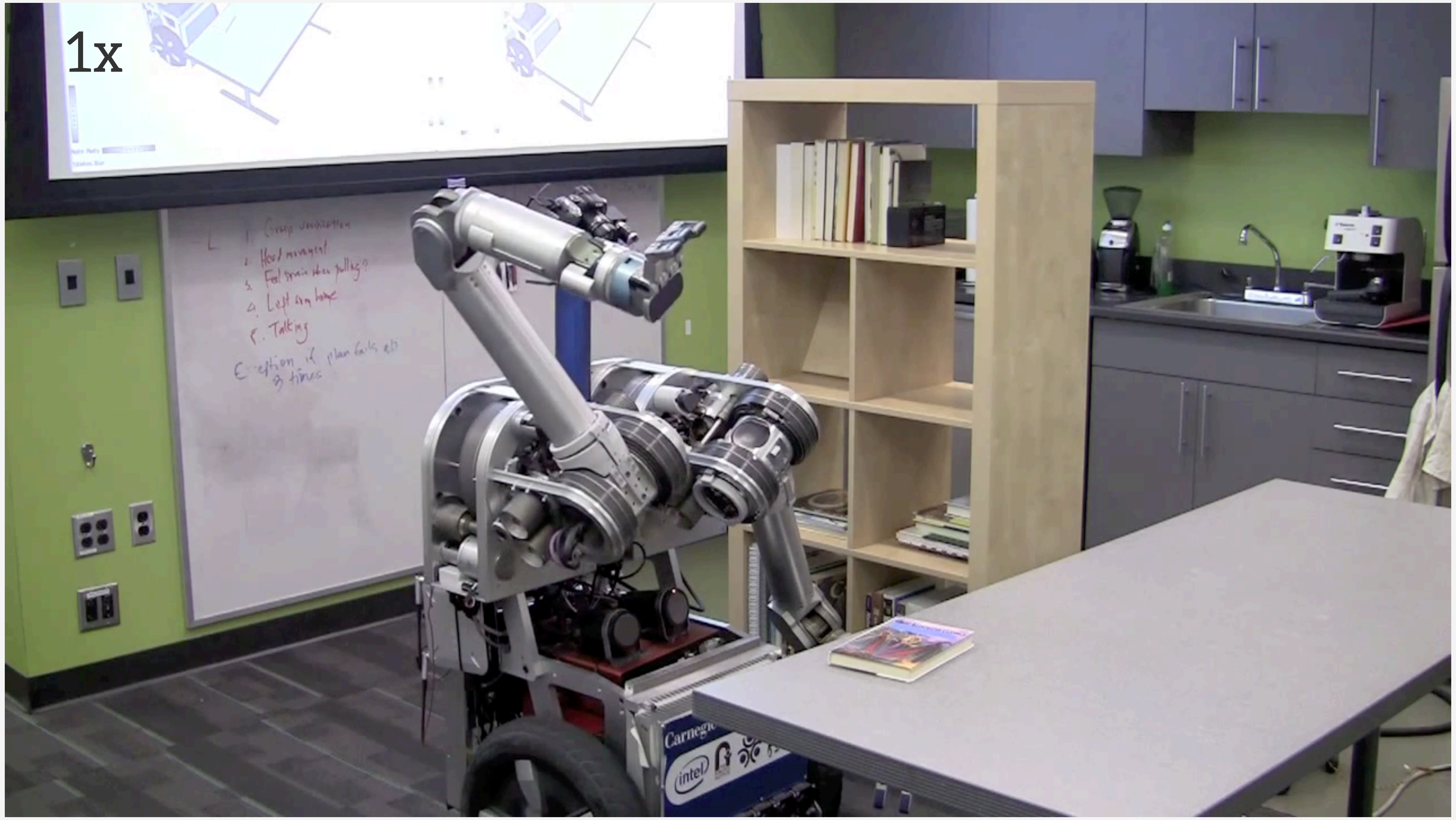


Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. Taking the intentional stance at 12 months of age. Cognition. 1995.
Dennett, Daniel Clement. The intentional stance. MIT press, 1989.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. Cognitive Development. 2005.

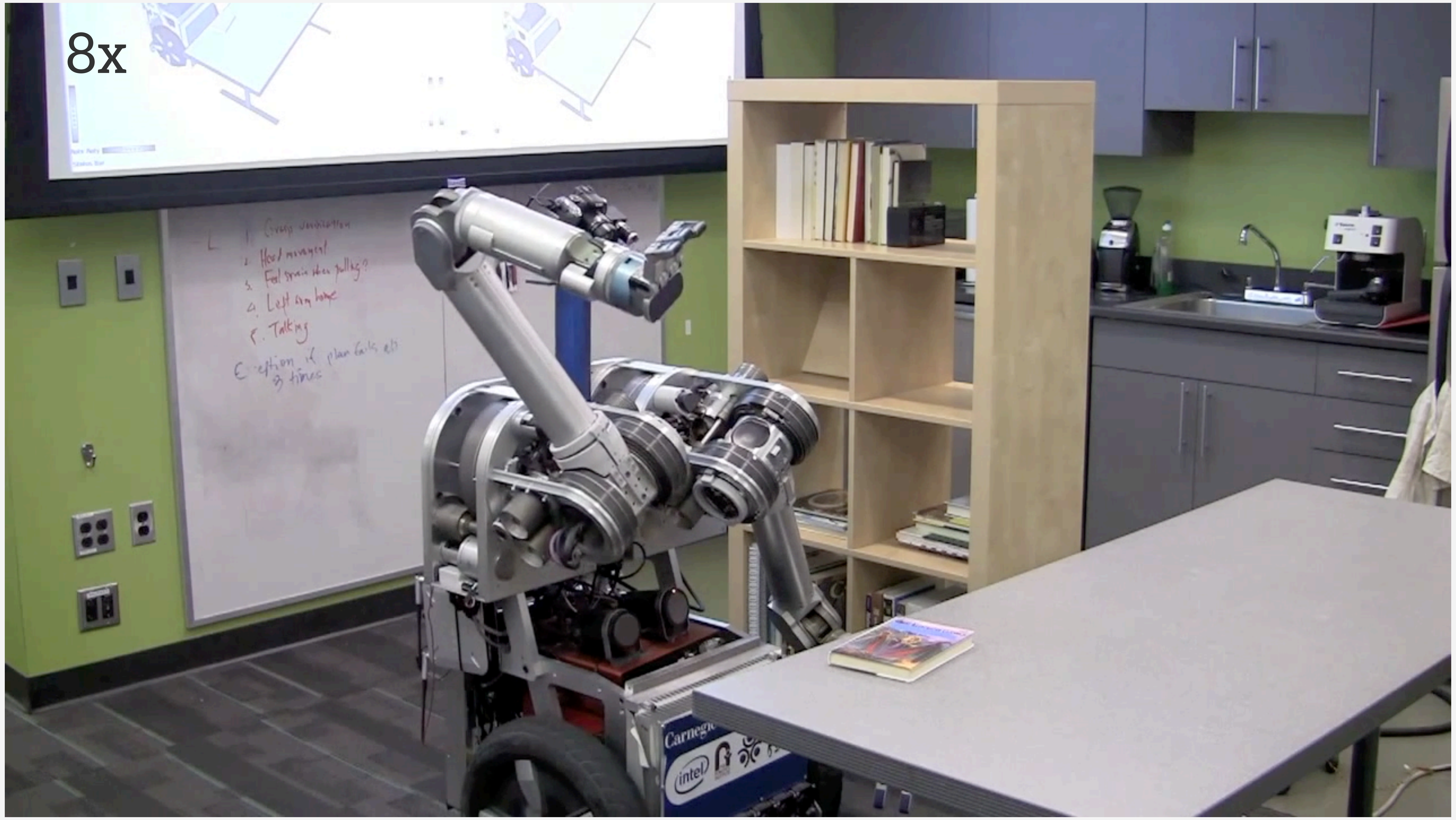
Intentions, **enough**?

Not easy to understand robots



HERB manages a library

Not easy to understand robots



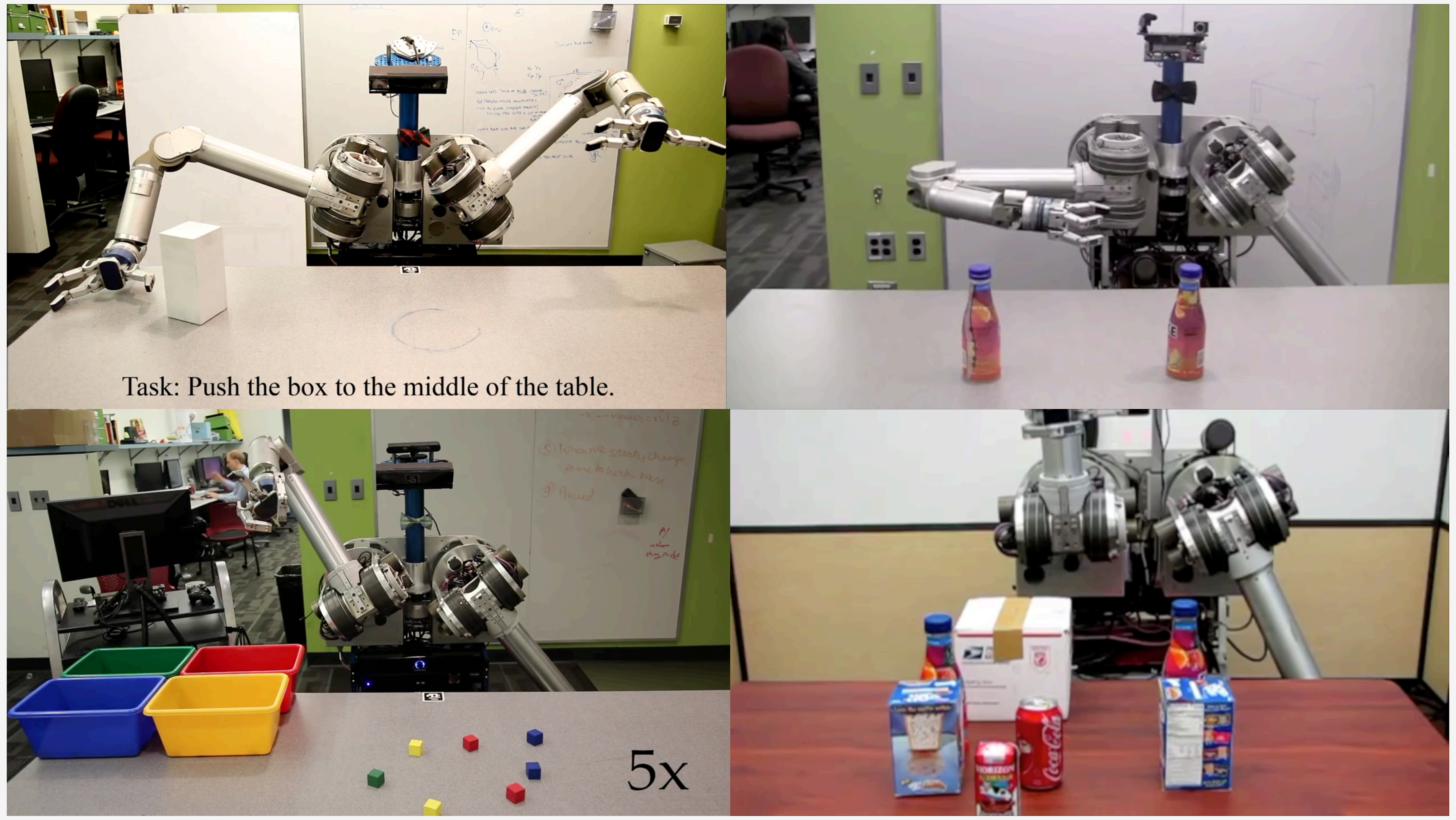
HERB manages a library

Robot Reasoning

Not easy to understand robots

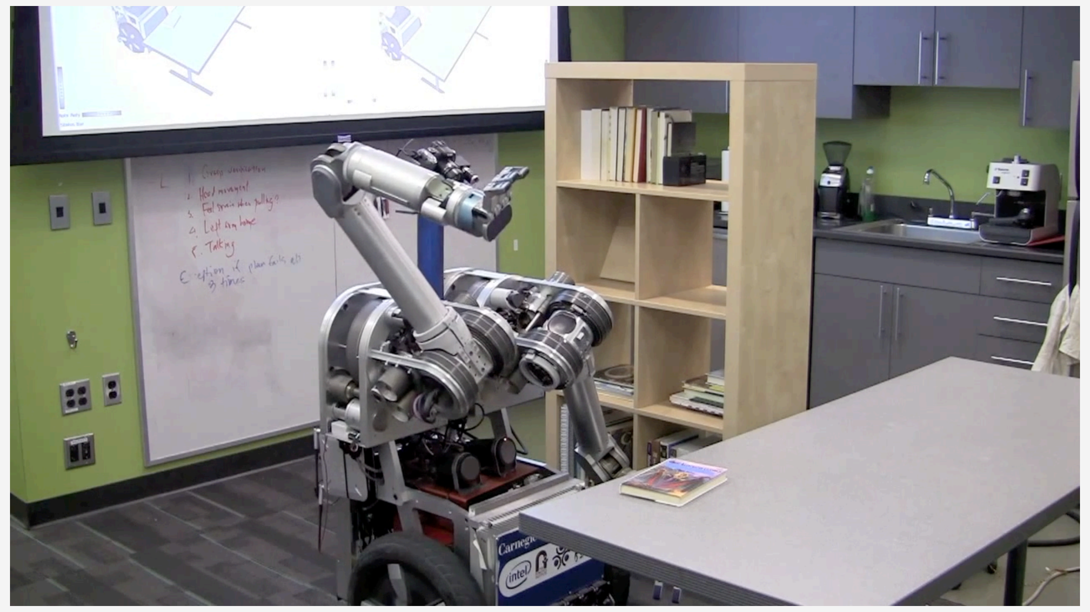
Robot intention

Understanding robot intentions helps people understand and anticipate robot behavior based on the **rationality principle**.



Robot reasoning

The robot preferences, constraints, cost, objective functions which affect robot plans.

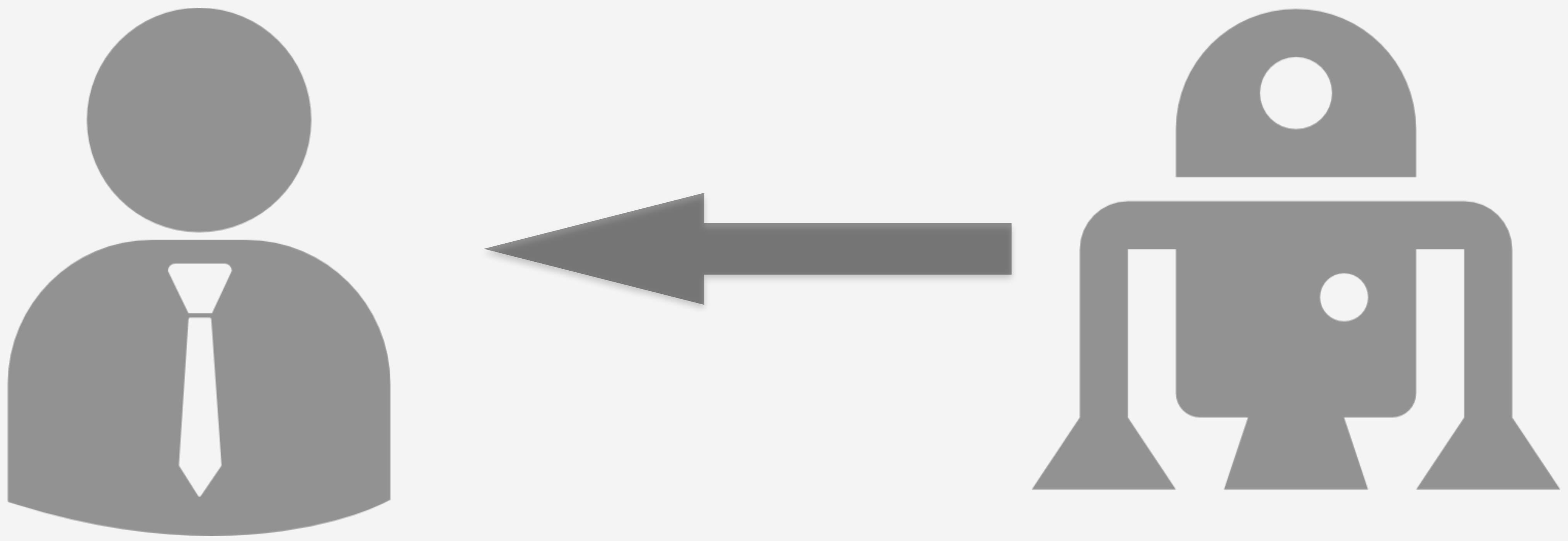


Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. Taking the intentional stance at 12 months of age. Cognition. 1995.
Dennett, Daniel Clement. The intentional stance. MIT press, 1989.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. Cognitive Development. 2005.

Make it easier for humans to
understand
robot intentions and reasoning

Explanations



Effects of explanations on the user mental models
Todd Kulesza, Simone Stumpf, Margaret Burne , and Irwin Kwan. Tell me more?: the effects of mental model soundness on personalizing an intelligent agent. In Proc. SIGCHI Conference on Human Factors in Computing Systems, pages 1–10. ACM, 2012.

Explaining its decision-making
Pat Langley. Explainable agency in human-robot interaction. 2016.

Intelligibility in context-aware systems
Anind K Dey. Explanations in context-aware systems. In ExaCt, pages 84–93, 2009.

Effects of explanations on the user mental models
Todd Kulesza, Simone Stumpf, Margaret Burne , Sherry Yang, Irwin Kwan, and Weng-Keen Wong. Too much, too little, or just right? ways explanations impact end users' mental models. In Proc. Visual Languages and Human-Centric Computing (VL/HCC), pages 3–10. IEEE, 2013

Convert visualization to verbalization
Stephanie Rosenthal, Sai P Selvaraj, and Manuela Veloso. Verbalization: Narration of autonomous robot experience. In Proc. IJCAI, pages 862– 868. AAAI Press, 2016.

Language-based explanation

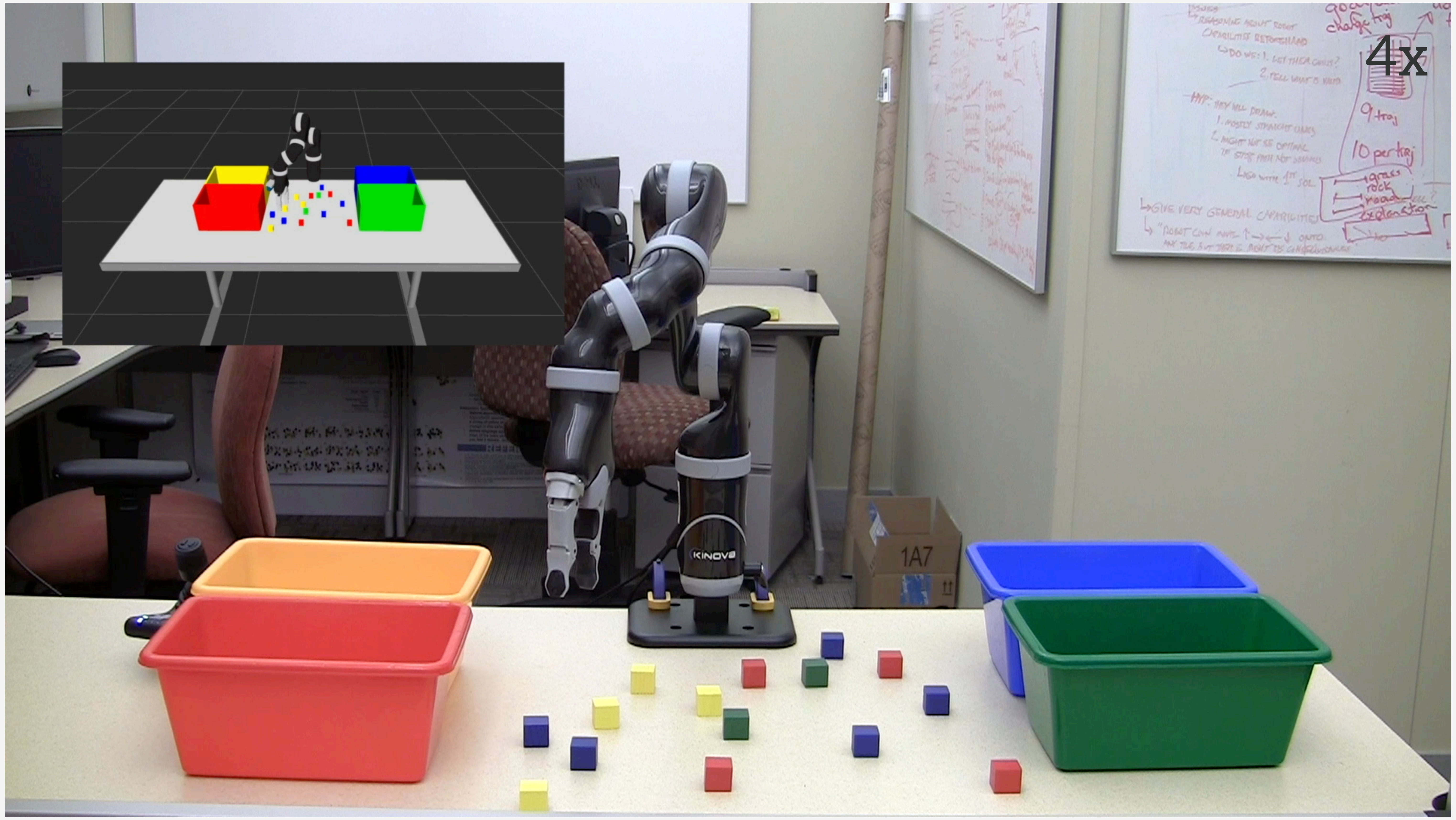
Effects of explanations on the user mental models
Adrian Bussone, Simone Stumpf, and Dympna O'Sullivan. The role of explanations on trust and reliance in clinical decision support systems. In Proc. International Conference on Healthcare Informatics (ICHI), pages 160–169. IEEE, 2015.

Predict the variability of utterances for different humans.
Vi orio Perera, Sai P Selveraj, Stephanie Rosenthal, and Manuela Veloso. Dynamic generation and refinement of robot verbalization. In Proc. RO-MAN, pages 212–218. IEEE, 2016.

Generating explanations in context-aware systems
Brian Y Lim, Anind K Dey, and Daniel Avrahami. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In Proc. SIGCHI Conference on Human Factors in Computing Systems, pages 2119–2128. ACM, 2009.

Multi-model explanation generation as a prolonged interaction.
Tathagata Chakraborti, Sarath Sreedharan, Yu Zhang, and Subbarao Kambhampati. Explanation generation as model reconciliation in multi-model planning. arXiv preprint arXiv:1701.08317, 2017.

Language-based explanation



Kinova Mico Arm

Language-based explanation

- No extra tools needed
- Compensates for **visual** communication
- Easy to be monitored and recorded
- Reach multiple agents in various positions
- Signal emotional states

J.C. Simon. Spoken Language Generation and Understanding. Springer. 1980.
Ma hias Scheu , Paul Schermerhorn, and James Kramer. The utility of affect expression in natural language interactions in joint human-robot tasks. HRI. 2006.
Ayelet N Landau, Lisa Aziz-Zadeh, and Richard B Ivry. The influence of language on perception: listening to sentences about faces affects the perception of faces. Journal of Neuroscience. 2010.

Expressively convey robot capabilities
Stefanos Nikolaidis, Swaprava Nath, Ariel D Procaccia, and Siddhartha Srinivasa. Game-theoretic modeling of human adaptation in human- robot collaboration. HRI. 2017.

Expressively convey object physical property
Alessandra Sciu i, Laura Patane, Francesco Nori, and Giulio Sandini. Understanding object weight from human and humanoid lifting actions. IEEE Transactions on Autonomous Mental Development, 6(2):80– 92, 2014.

Expressively convey robot goal
Michael J Gielniak, C Karen Liu, and Andrea L Thomaz. Generating human-like motion for robots. IJRR, 32(11):1275–1301, 2013.

Expressively convey robot goal
Anca Dragan and Siddhartha Srinivasa. Integrating human observer inferences into robot motion planning. Autonomous Robots, 37(4): 351– 368, 2014.

Expressively convey robot goal
Daniel Szafir, Bilge Mutlu, and Terrence Fong. Communication of intent in assistive free flyers. In Proc. HRI, pages 358–365. ACM/IEEE, 2014.
Leila Takayama, Doug Dooley, and Wendy Ju. Expressing thought: im- proving robot readability with animation principles. In Proc. HRI, pages 69–76. ACM, 2011.

Demonstration-based explanation

Model explicability as the distances between robot plans and the human approximation of robot plan.
Anagha Kulkarni, Tathagata Chakraborti, Yantian Zha, Satya Gautam Vadlamudi, Yu Zhang, and Subbarao Kambhampati. Explicable robot planning as minimizing distance from expected behavior. arXiv preprint arXiv:1611.05497, 2016.

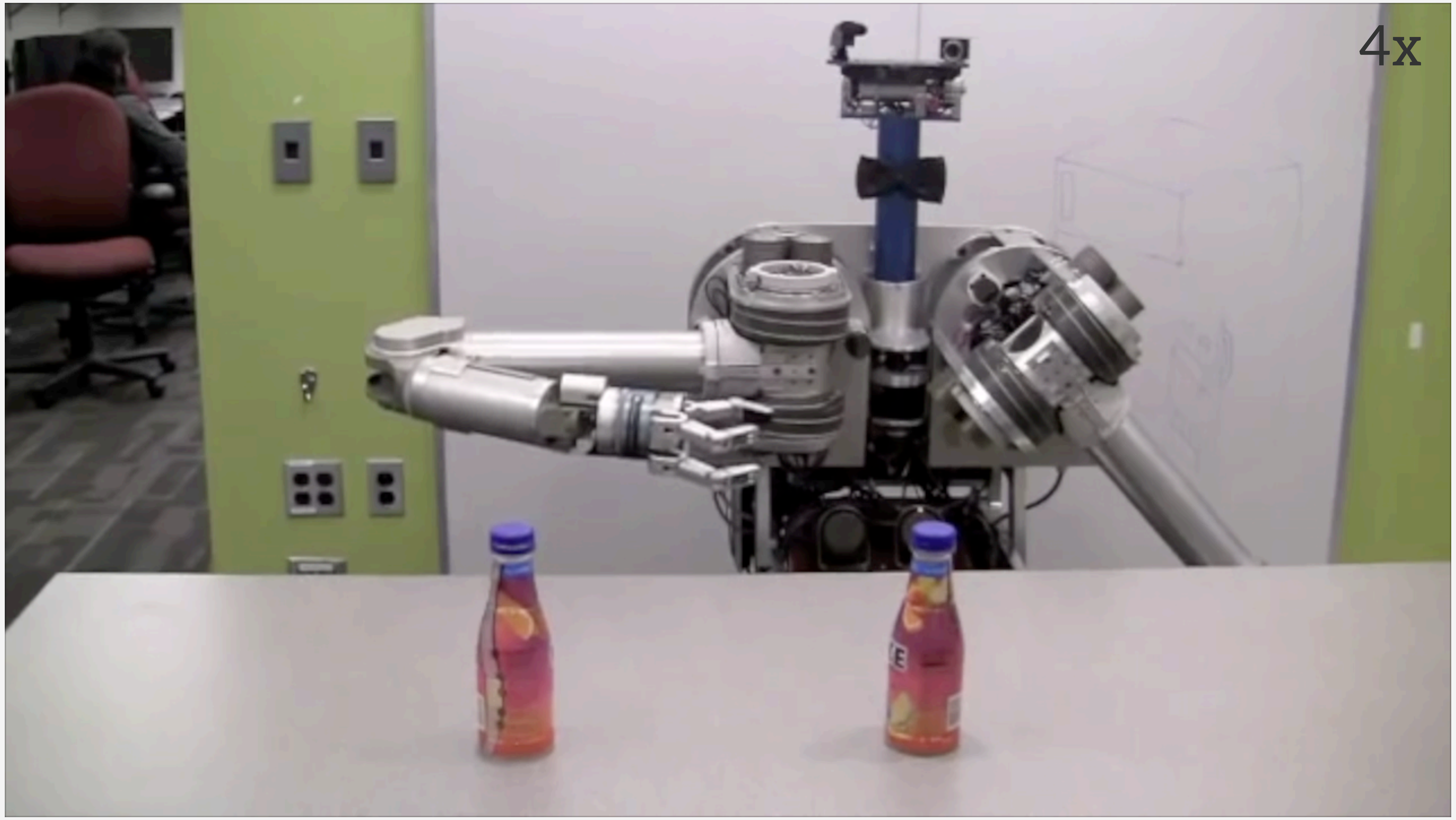
Expressively convey object physical property
Allan Zhou, Dylan Hadfield-Menell, Anusha Nagabandi, and Anca D Dragan. Expressive robot motion timing. In Proc. HRI, 2017.

Plan explicability and predictability
Yu Zhang, Hankz Hankui Zhuo, and Subbarao Kambhampati. Plan explainability and predictability for cobots. CoRR, abs/1511.08158, 2015.

Expressively convey robot goal
Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz Hankui Zhuo, and Subbarao Kambham- pati. Plan explicability for robot task planning. In Proc. RSS Workshop on Planning for Human-Robot Interaction: Shared Autonomy and Collaborative Robotics, 2016.

Expressively convey robot learning progress
Monica N Nicolescu and Maja J Mataric. Natural methods for robot task learning: Instructive demonstrations, generalization and practice. AAMAS. 2003.

Demonstration-based explanation



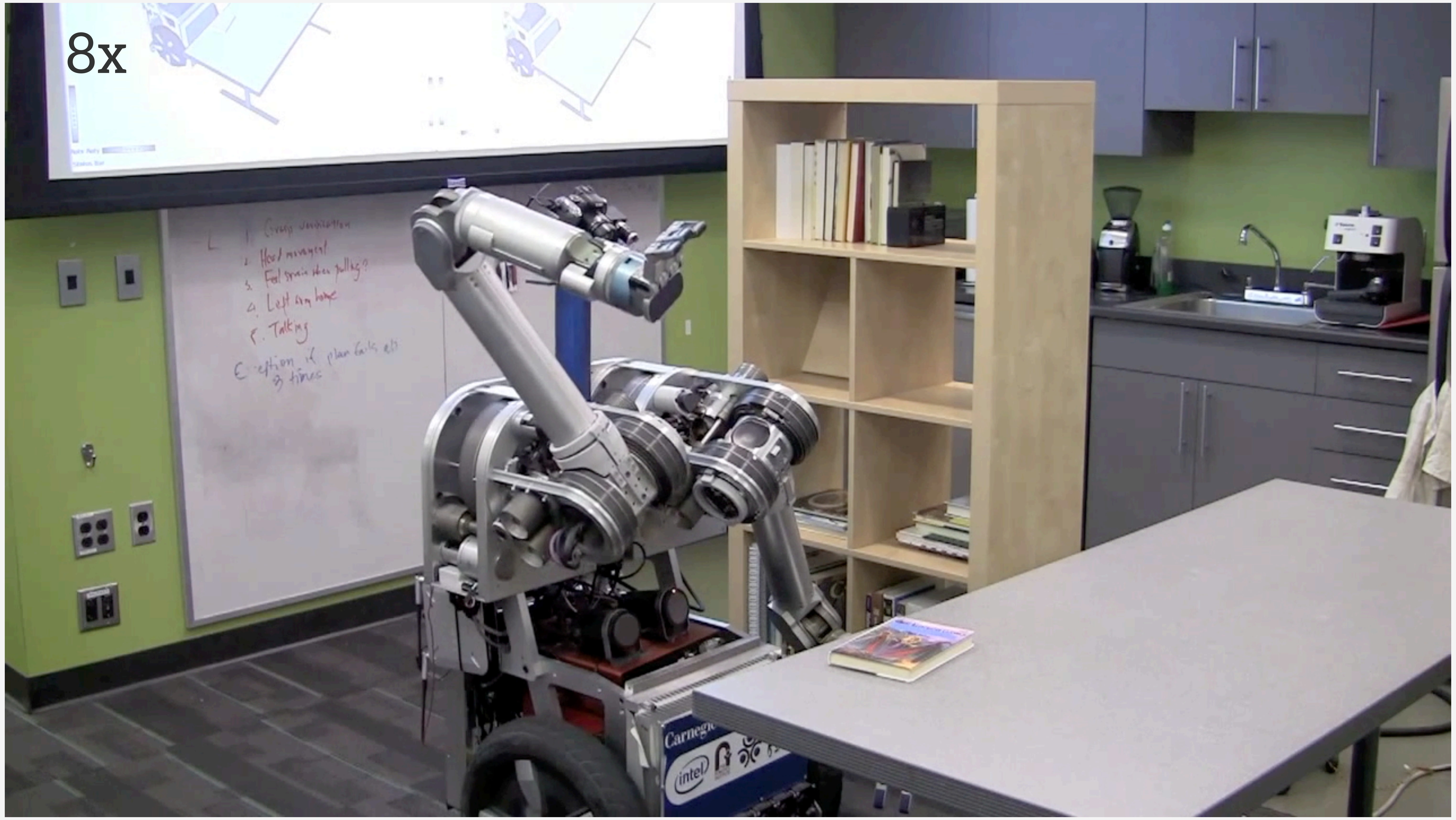
Explaining

Dragan, A. D., Holladay, R. M., & Srinivasa, S. S. An Analysis of Deceptive Robot Motion. RSS, 2014

Demonstration-based explanation

- Compensates for **verbal** communication
    - Environment
    - Task
- Reach multiple agents in various positions
- Signal emotional states

Demonstration is easier than language

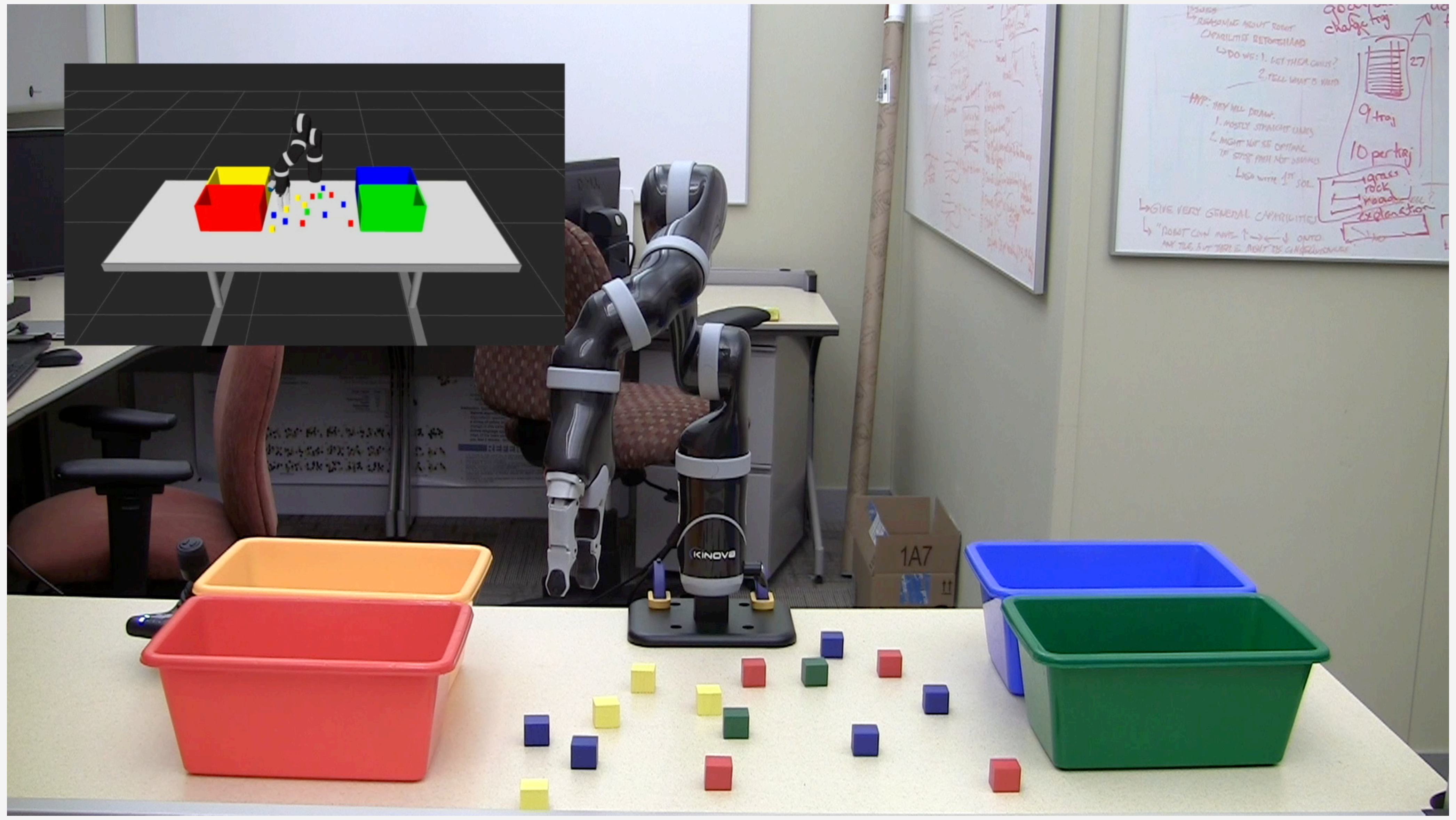


HERB manages a library

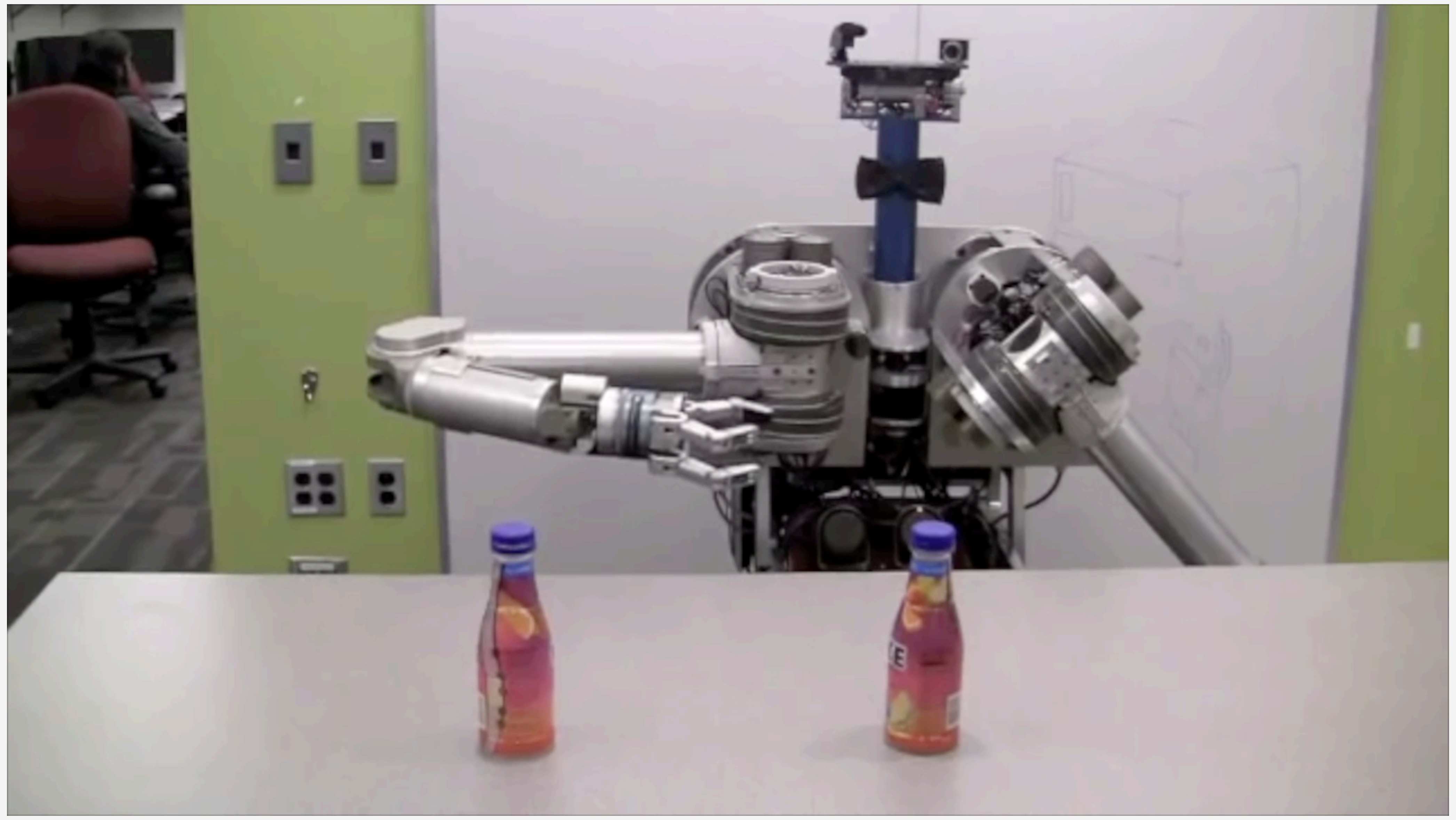
Robot Reasoning

Language Based Explanation

Demonstration Based Explanation

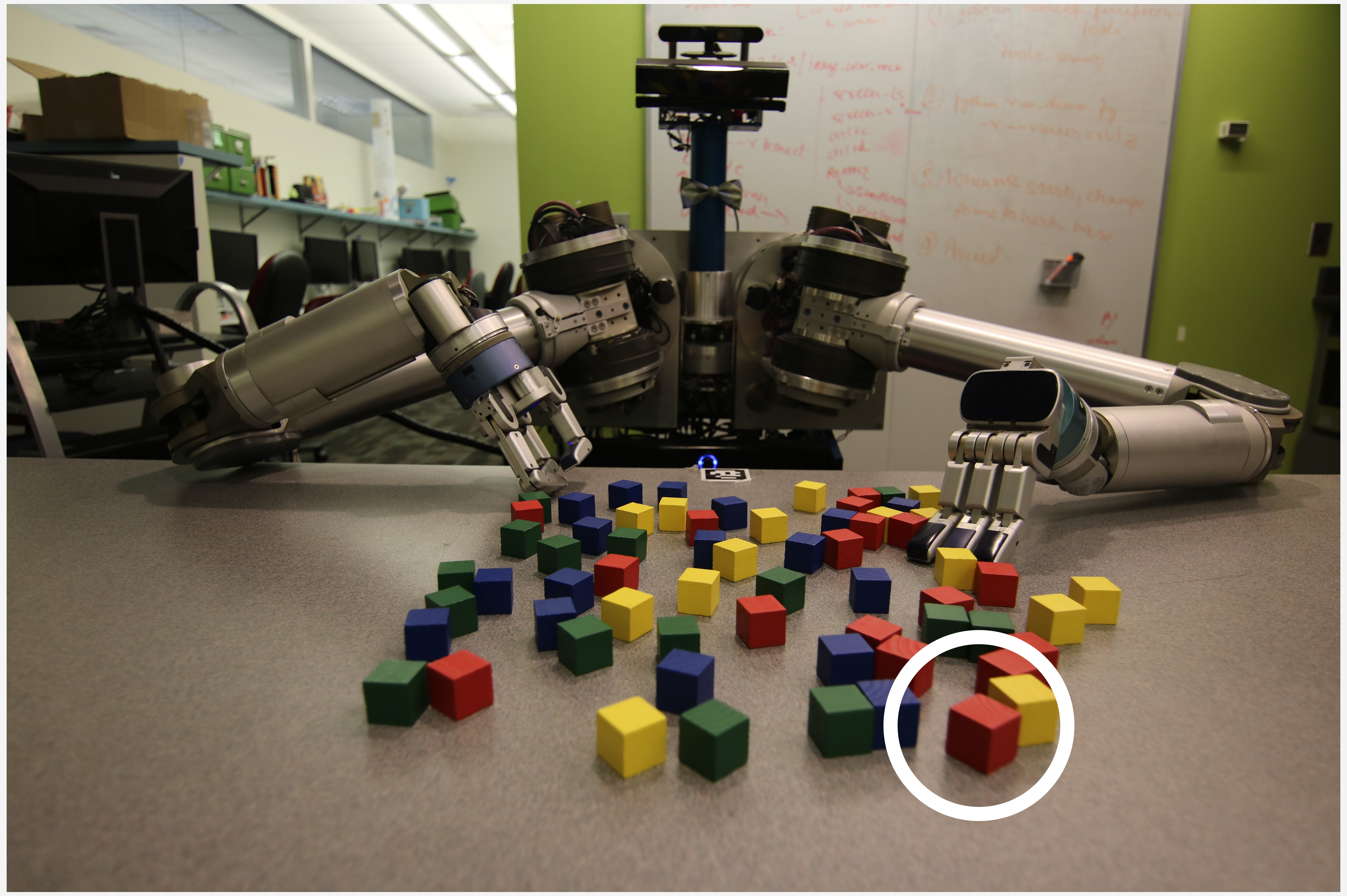


Kinova Mico Arm



Dragan, A. D., Holladay, R. M., & Srinivasa, S. S. An Analysis of Deceptive Robot Motion. RSS, 2014

Language-based explanation for robot intentions



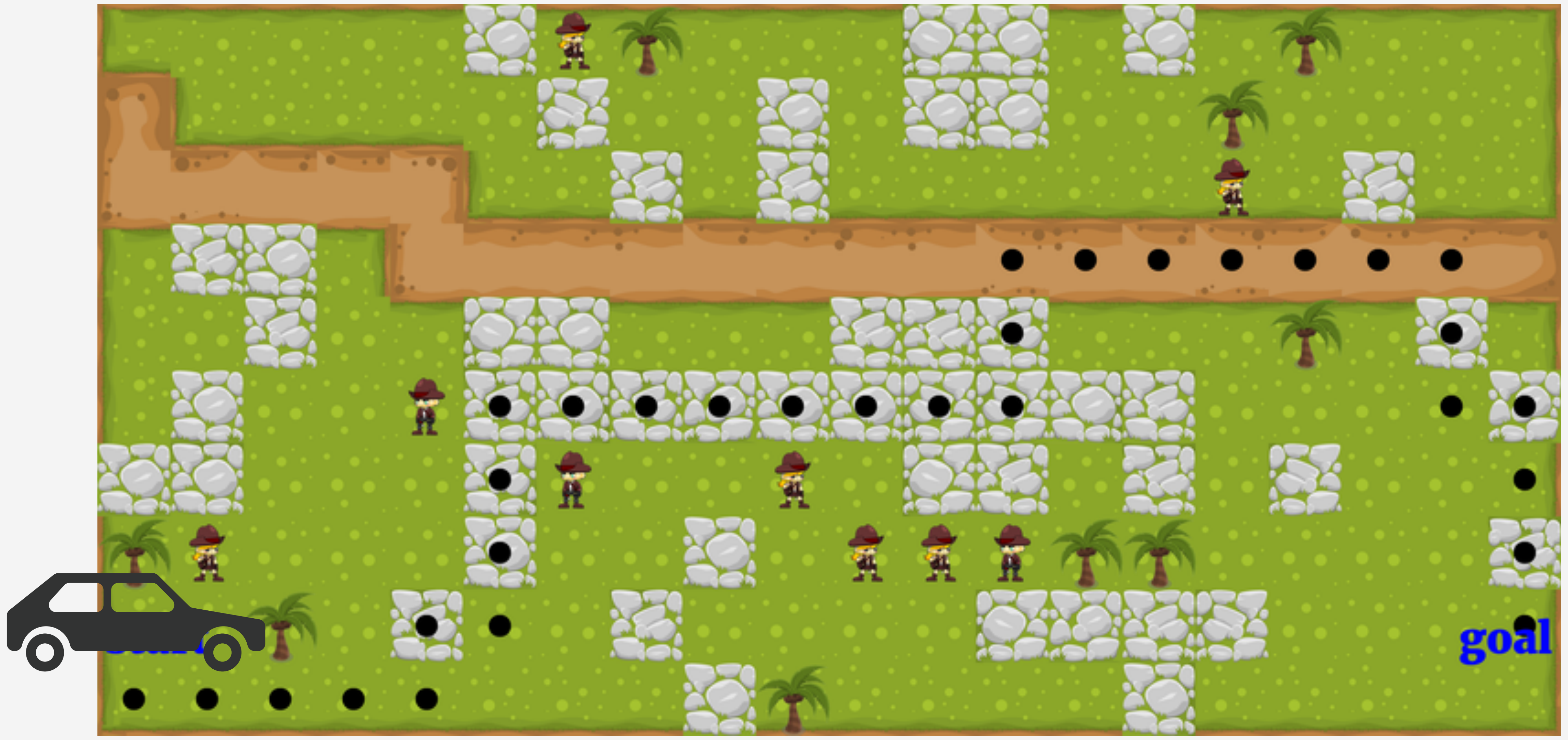
"I am picking up the red block closest to you."

Language-based explanation for intentions

"I am picking up the red block closest to you."

RO-MAN'16

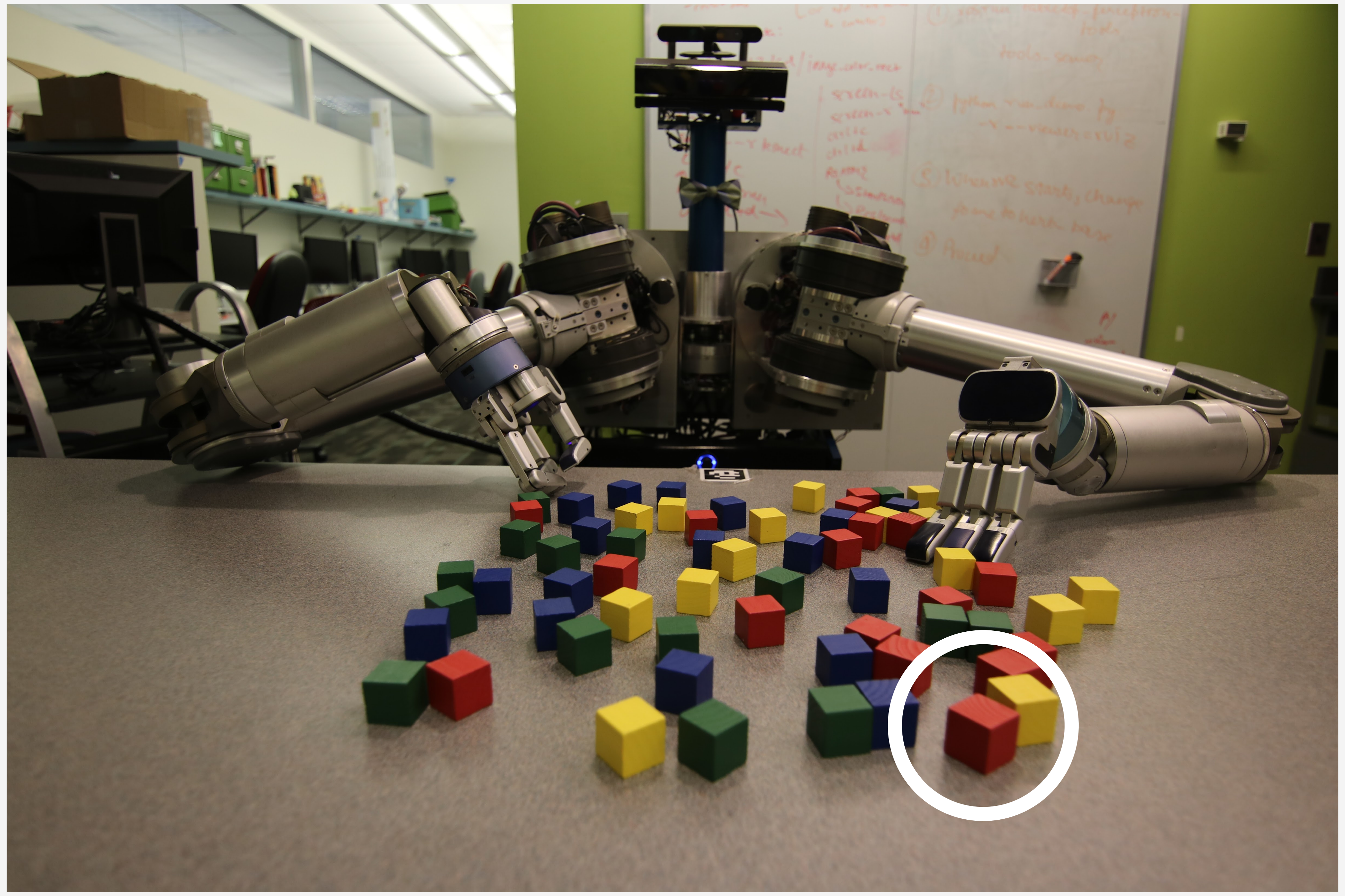
Demonstration-based explanation for reasoning



The robot trajectory is indicated as the black dots, which indicates that **it prefers rocks.**

Submitted to RO-MAN'17

Language-based explanation for intentions



"I am picking up the red block closest to you."

RO-MAN'16

Demonstration-based explanation for reasoning

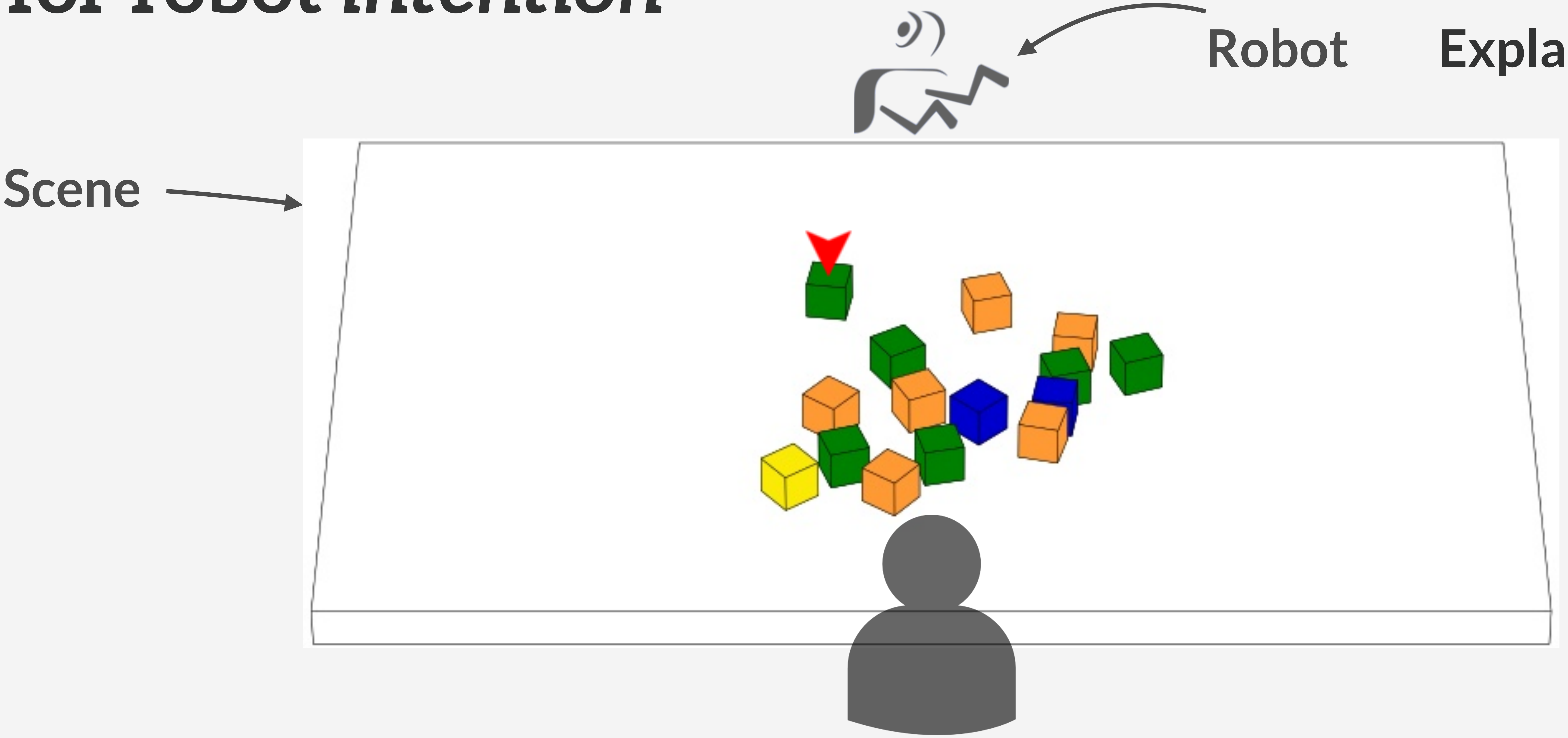
The robot trajectory is indicated as the black dots, which indicates that **it prefers rocks.**

Submitted to RO-MAN'17

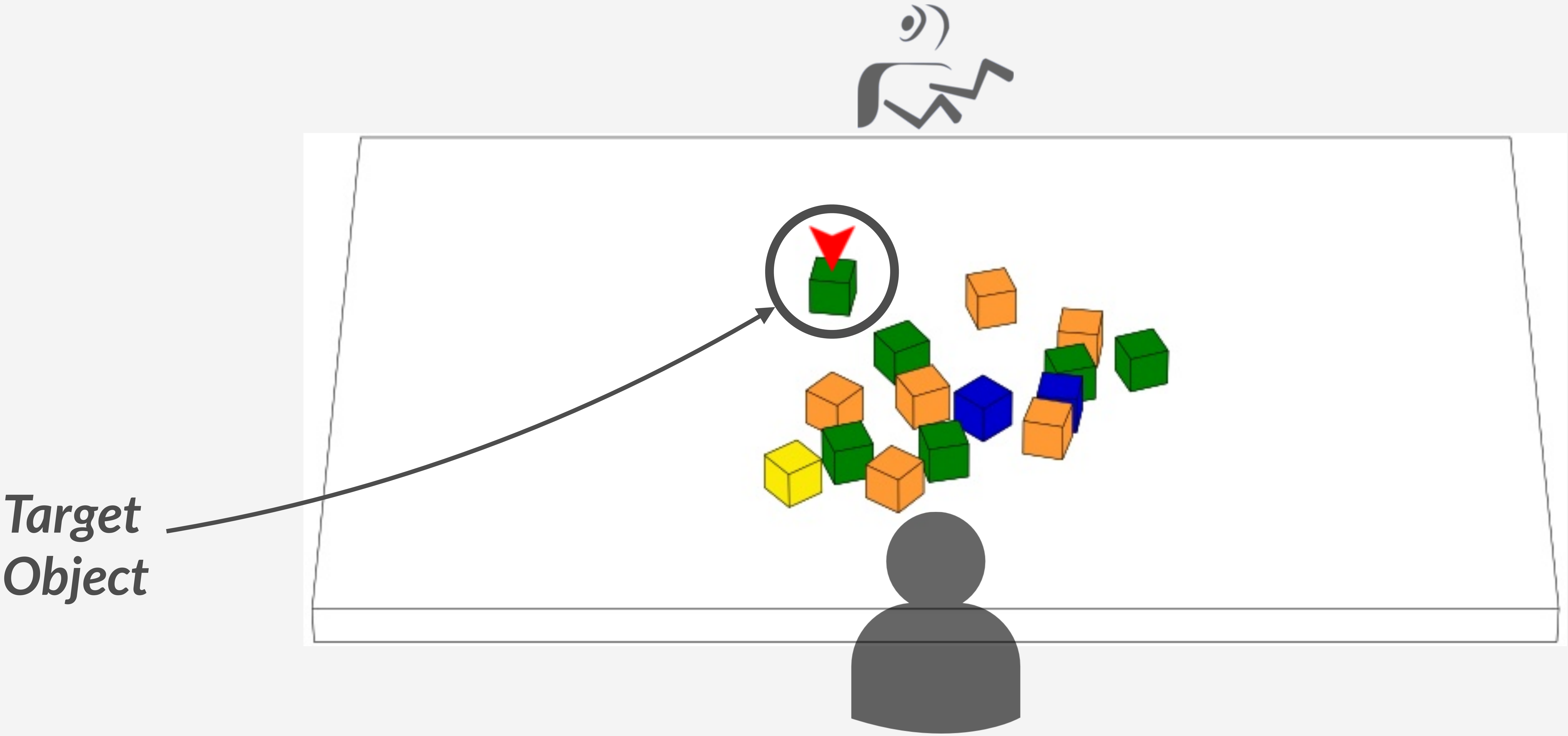
Basic concepts

Language-based explanation
for robot *intention*

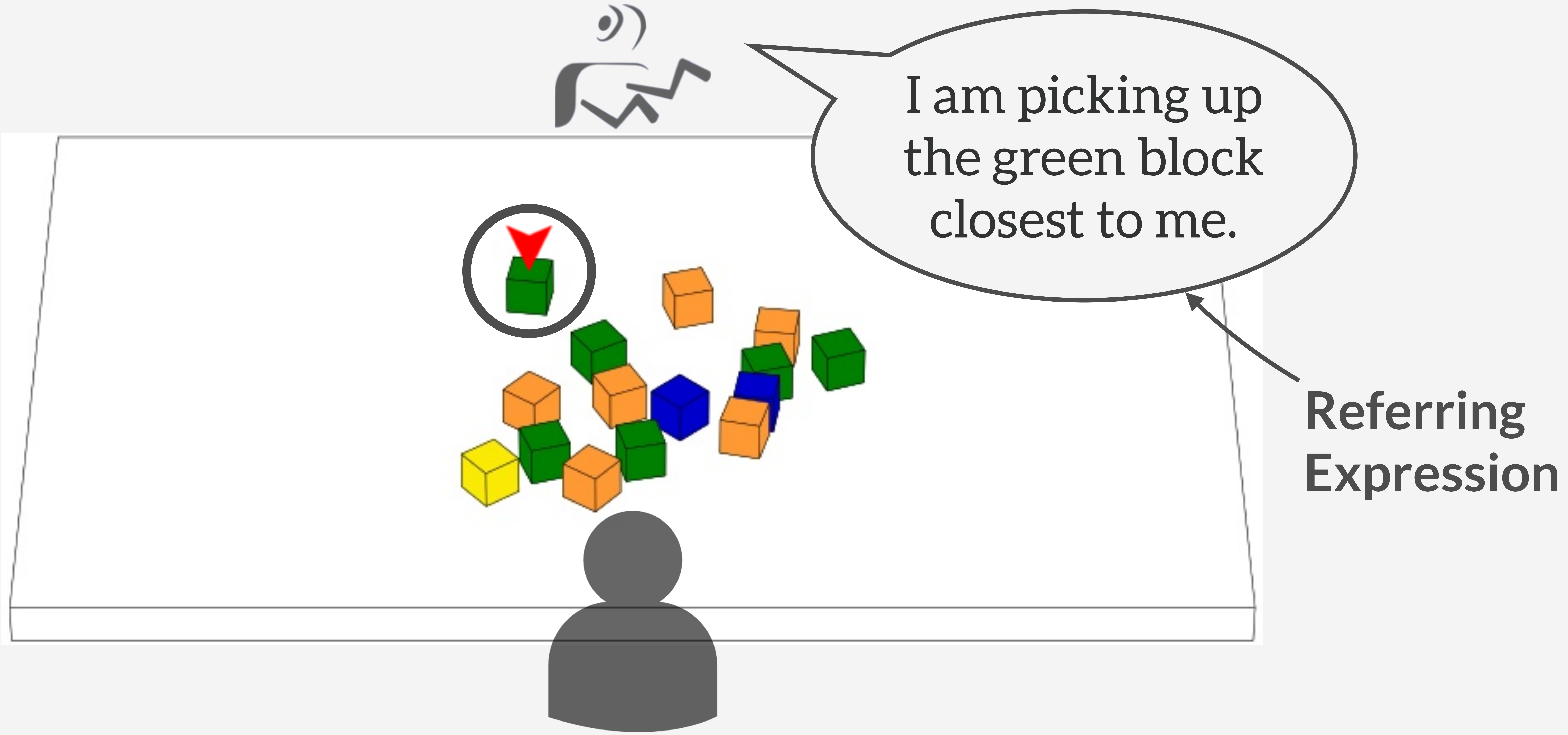
Explain intention



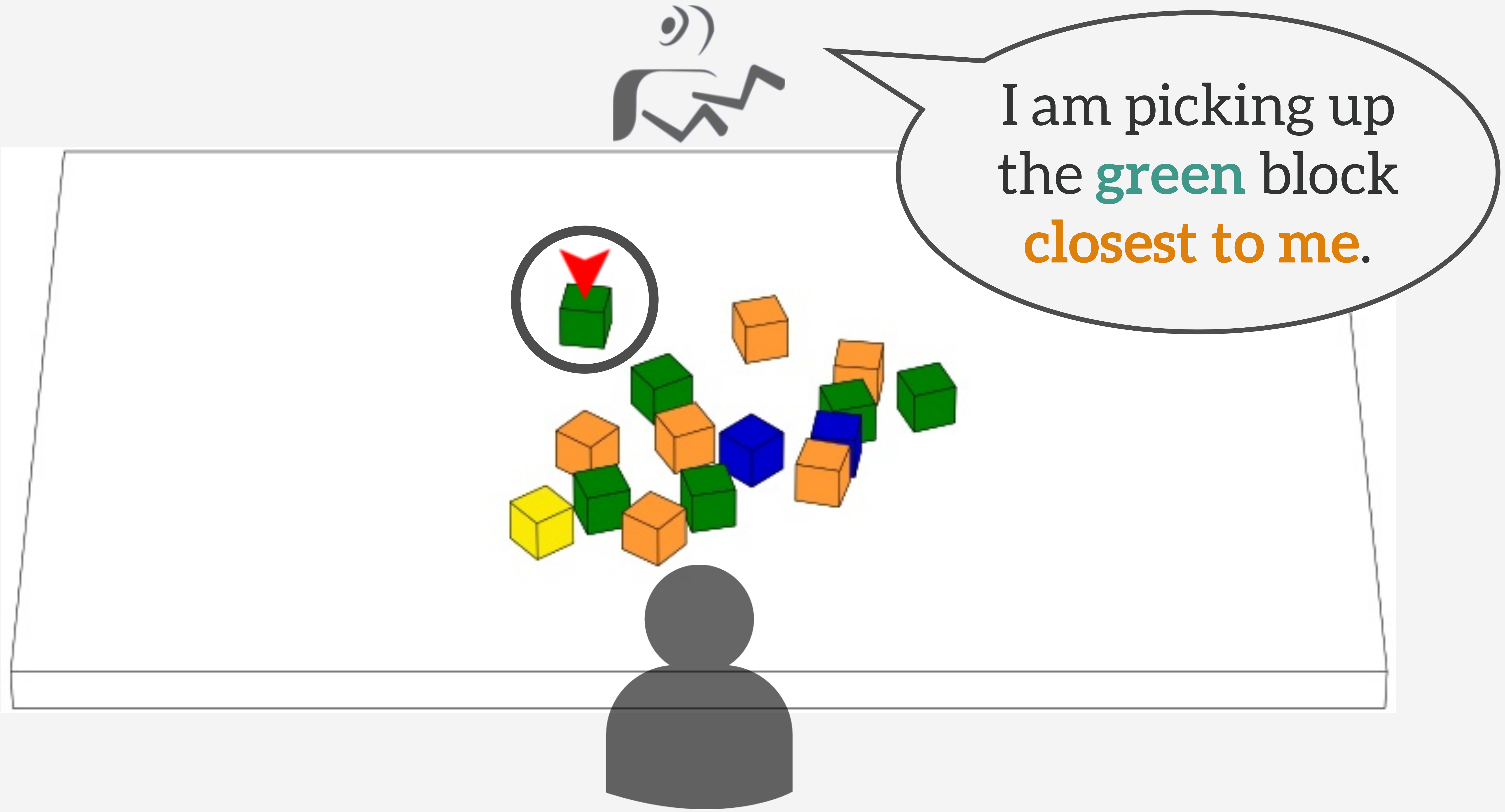
Intention = target object



Referring Expression (RE)

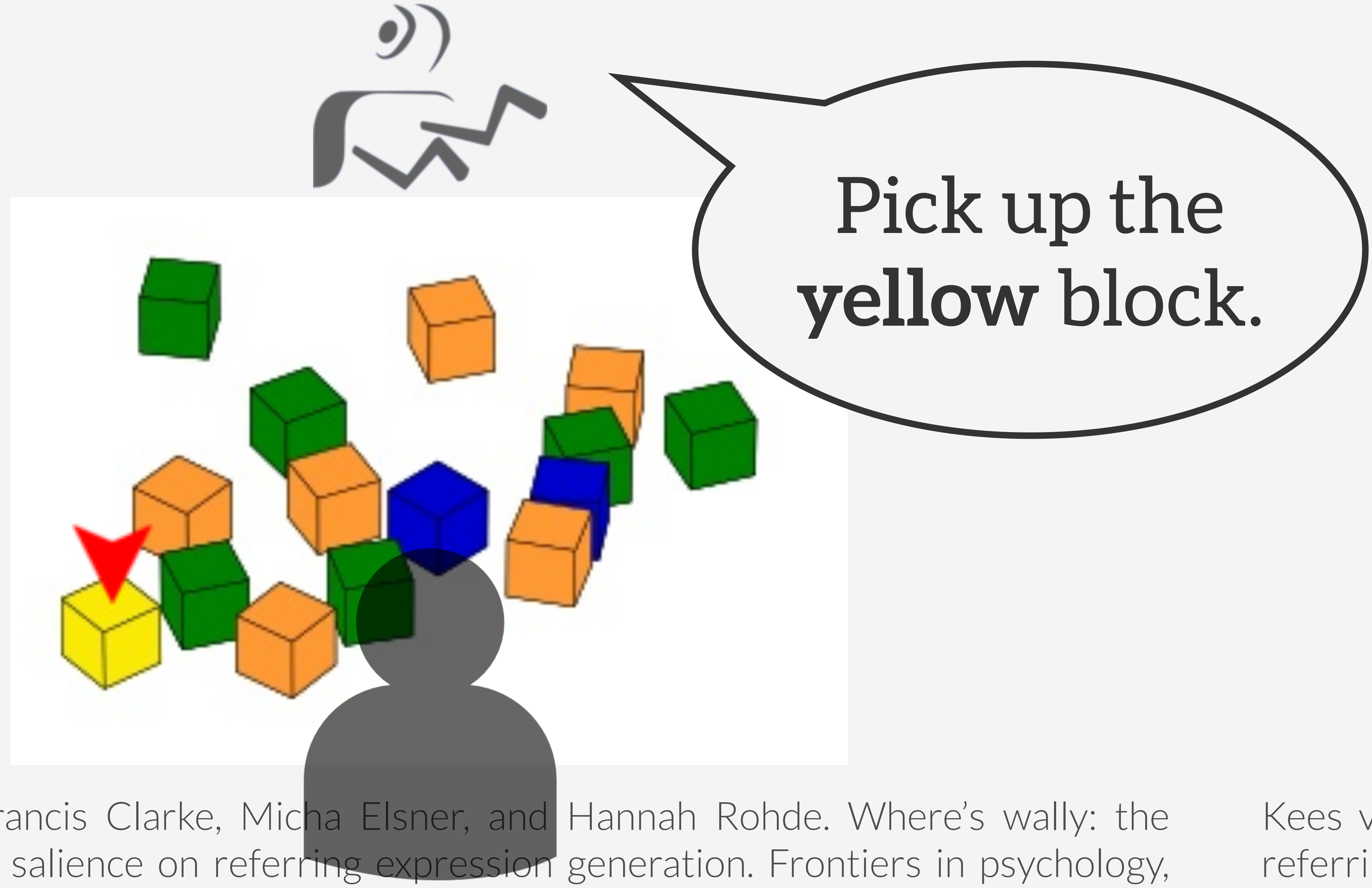


RE = a set of features



Features

- Visual features
  - *Color*. e.g. green, yellow, red
  - *Type*. e.g. block, box, spoons



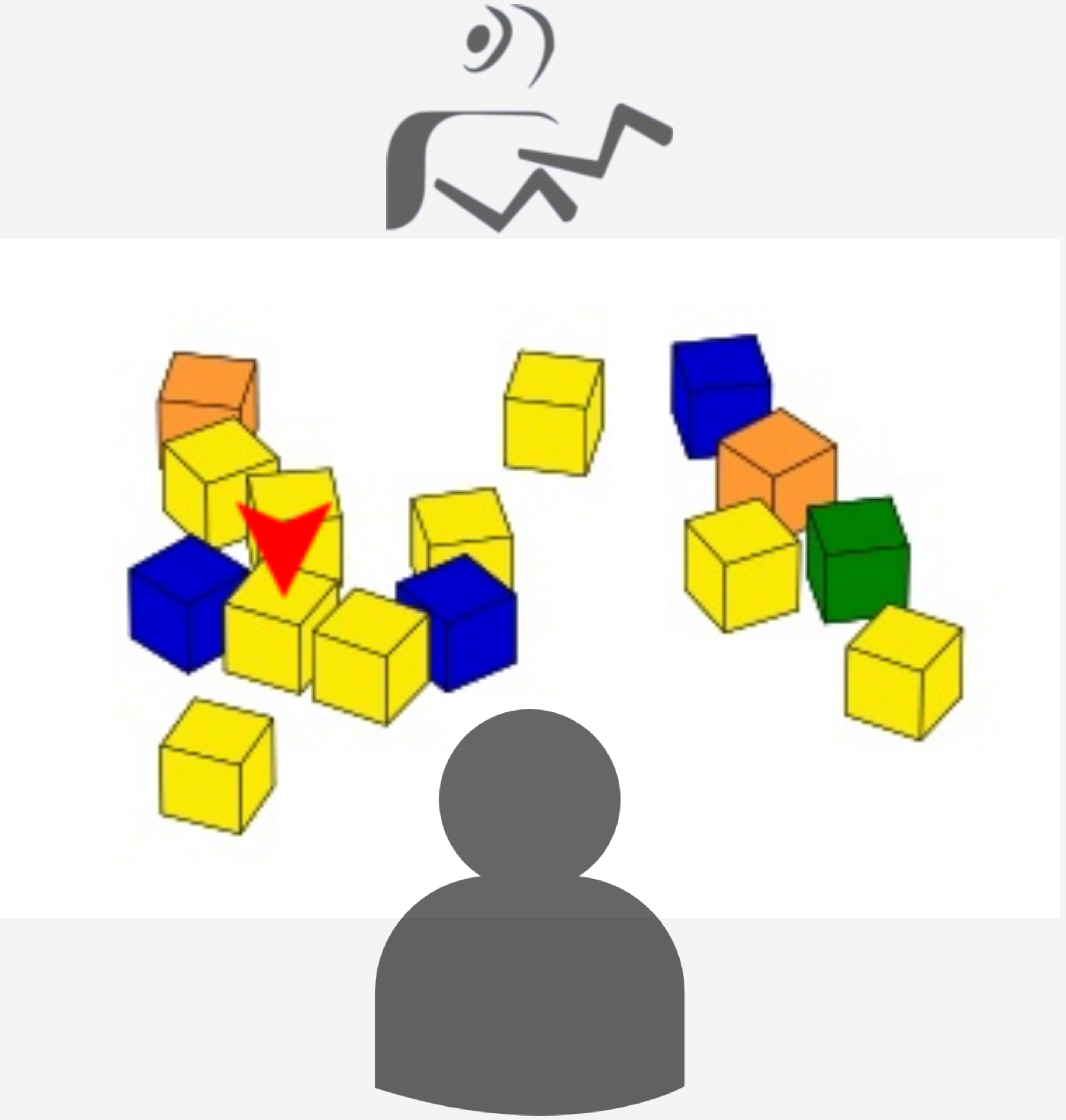
Alasdair Daniel Francis Clarke, Micha Elsner, and Hannah Rohde. Where's wally: the influence of visual salience on referring expression generation. Frontiers in psychology, 4:329, 2013.

Kees van Deemter, Albert Ga , Ielka van der Sluis, and Richard Power. Generation of referring expressions: Assessing the incremental algorithm. Cognitive science, 36(5):799–836, 2012.

Features

- Visual features
  - *Color*. e.g. green, yellow, red
  - *Type*. e.g. block, box, spoons
- Spatial relations
  - *Distance*. e.g. close, far
  - *Orientation*. e.g. left to, behind

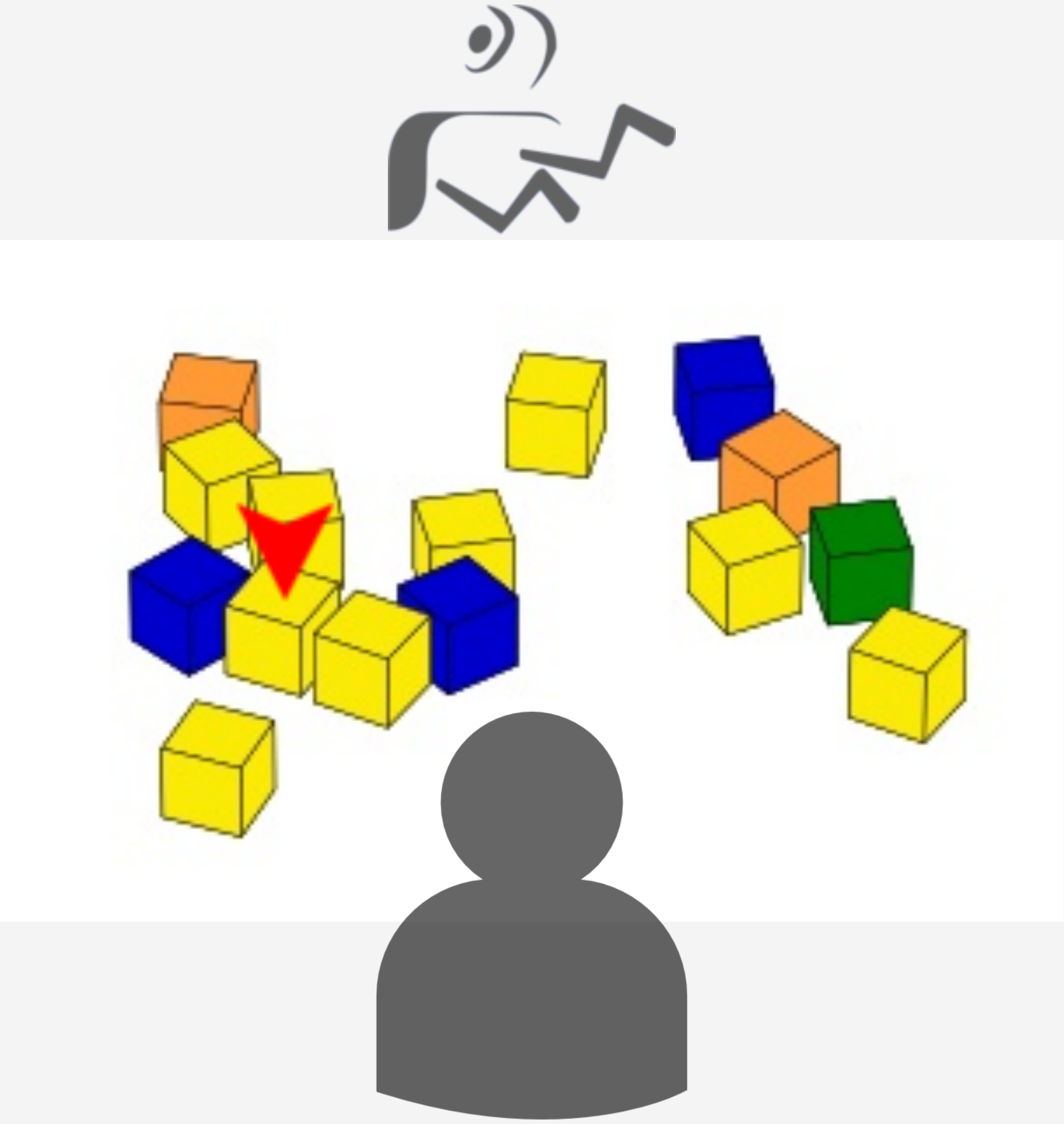
Find the two yellow blocks that are in **between two blue blocks**. Of those two yellow blocks, pick up the one on your right.



Features

- Visual features
  - *Color*. e.g. green, yellow, red
  - *Type*. e.g. block, box, spoons
- Spatial features
  - *Distance*. e.g. close, far
  - *Orientation*. e.g. left to, behind

Find the two yellow blocks that are in between two blue blocks. Of those two yellow blocks, pick up the one on your right.

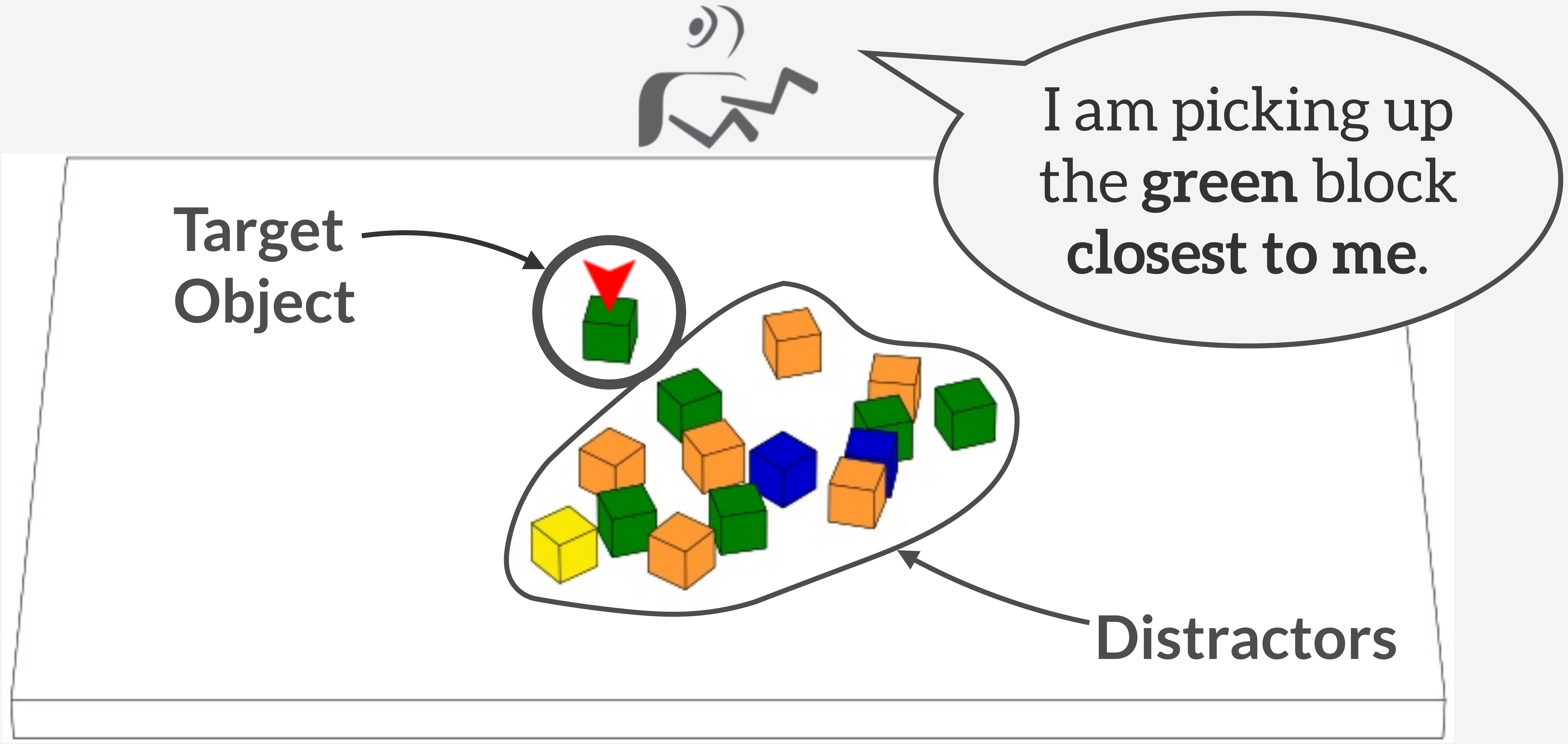


Target Object → **Landmarks**

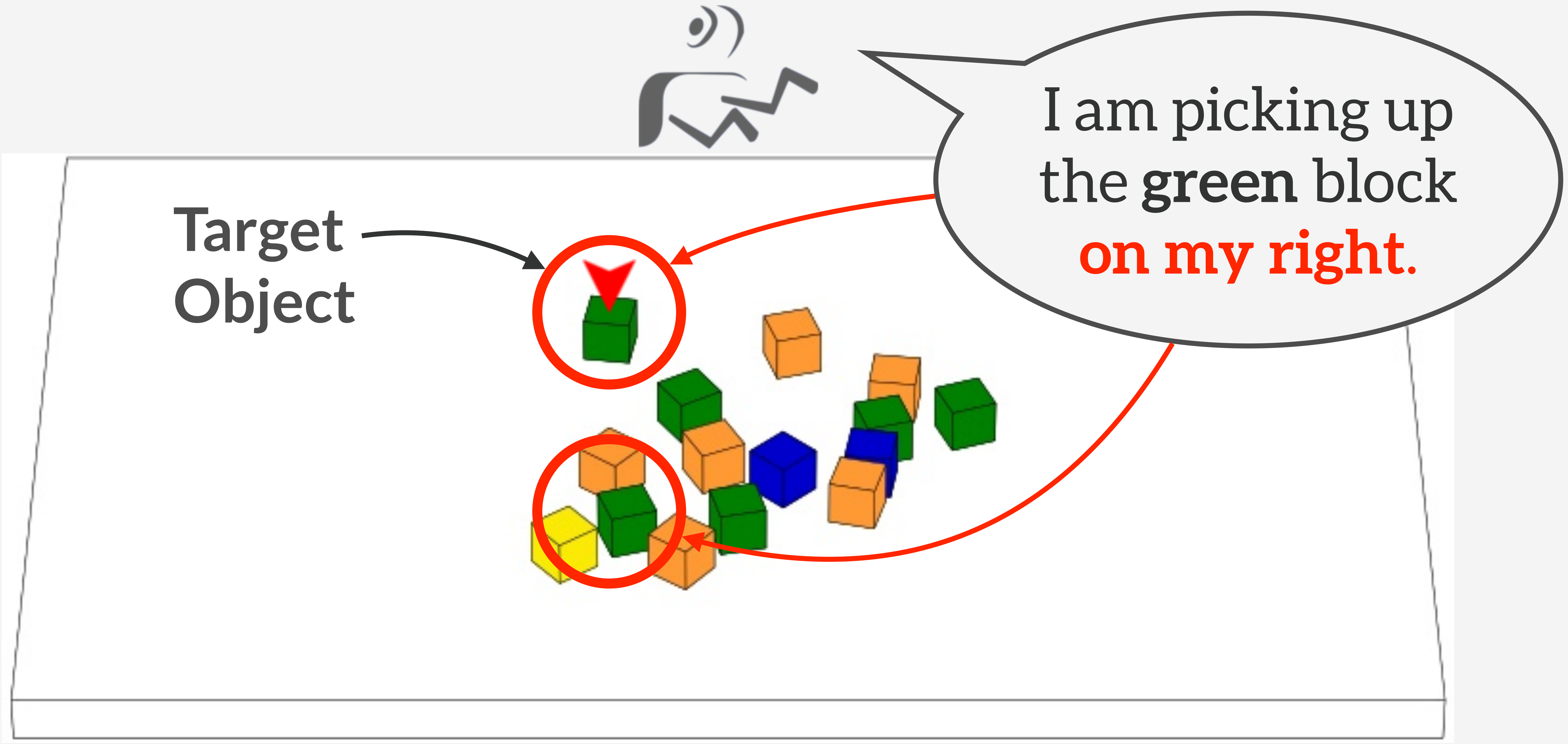
Spatial Relation

two blue blocks

Clear referring expressions



Ambiguous referring expressions



Referring expression generation
(REG)

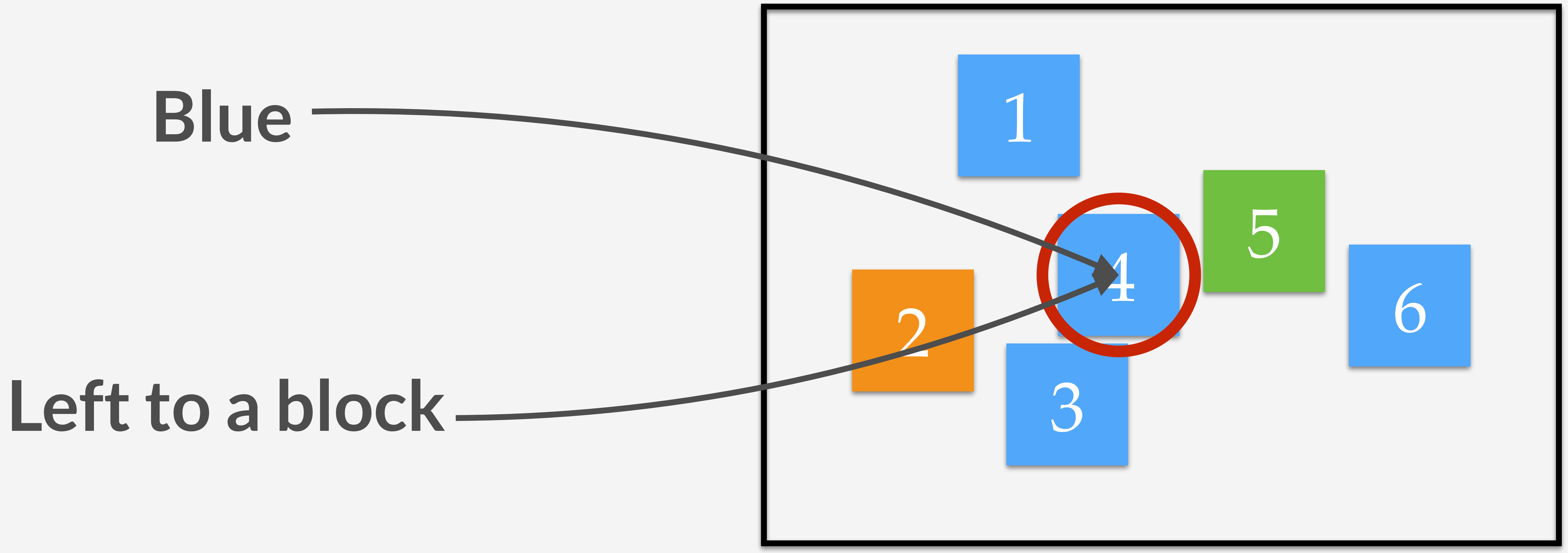
Referring expression generation (REG)

- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
  - Graph structure

Previous work on REG

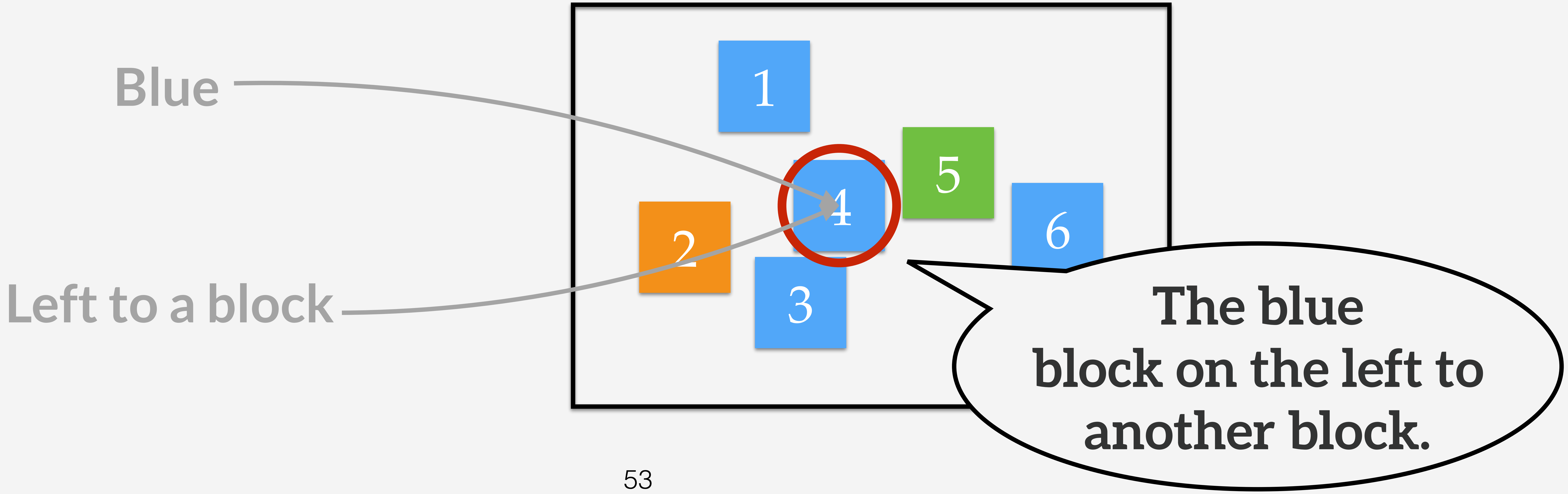
Previous work on REG

- Referring expression generation
  - Content selection/determination
    - What features you select to distinguish the target object from distractors?



Previous work on REG

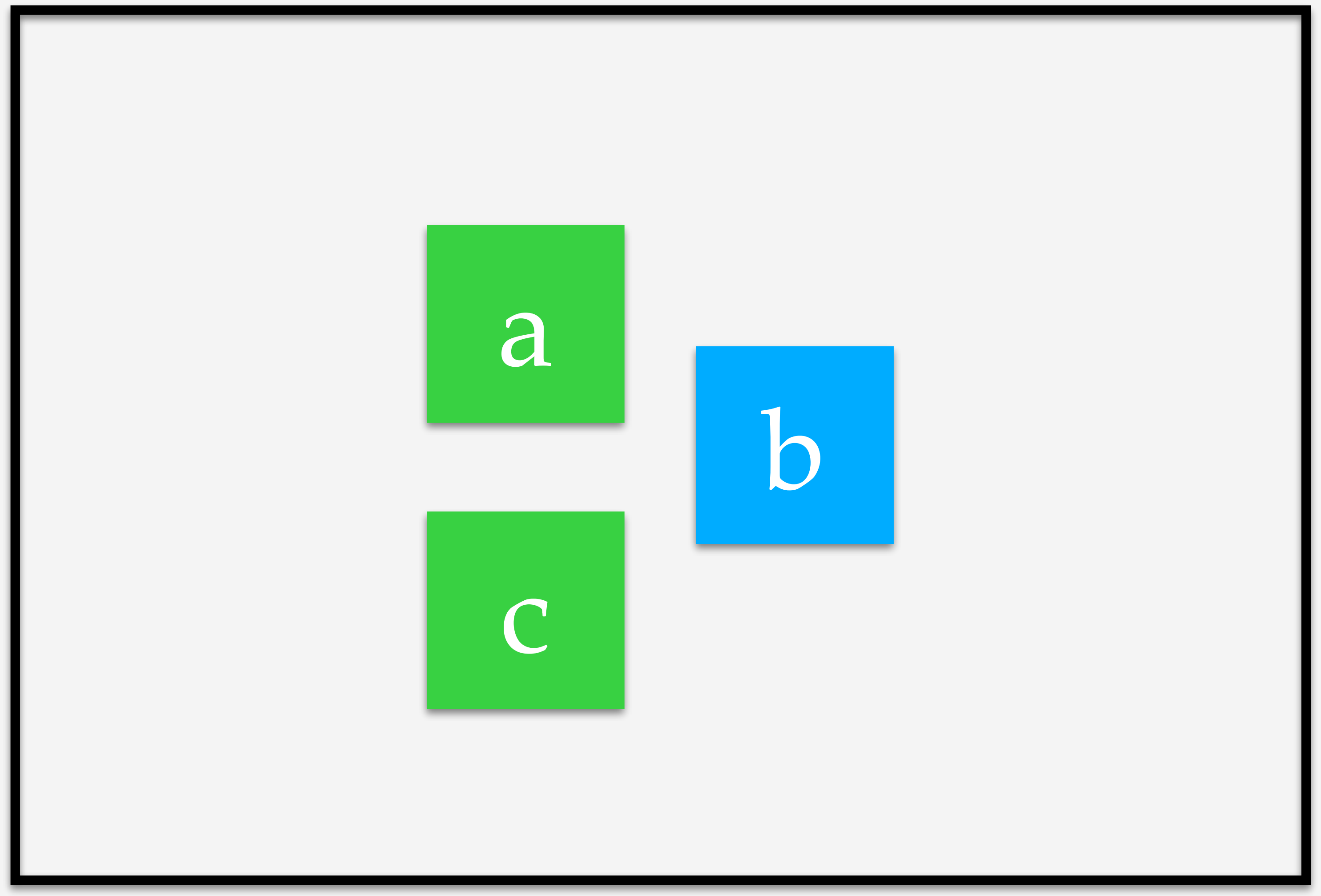
- Referring expression generation
  - Content selection/determination
  - Surface realization
    - Realize the set of features into natural language

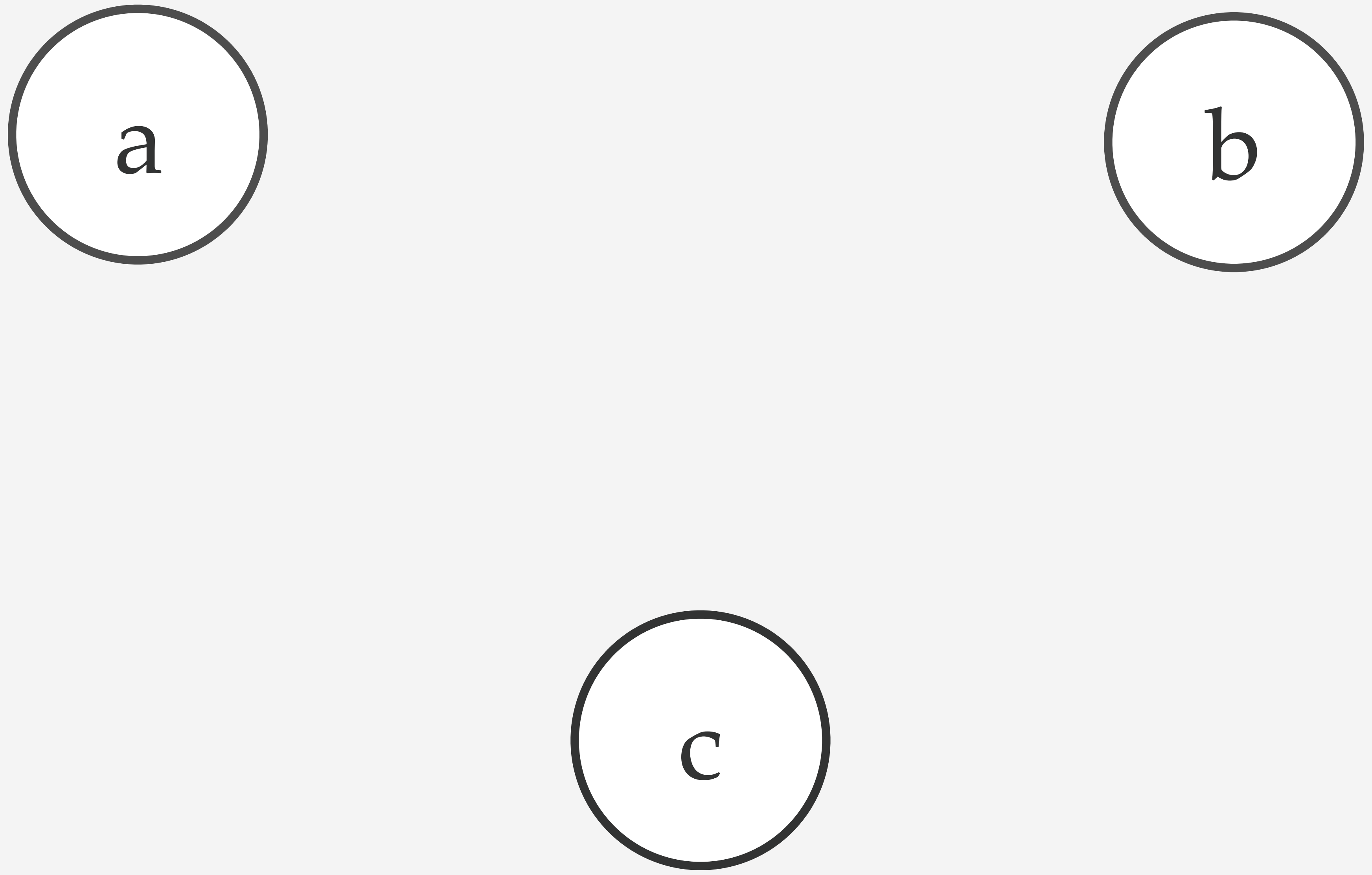


Previous work on REG

- Referring expression generation
  - Content selection/determination
    - Graph-based REG
  - Surface realization

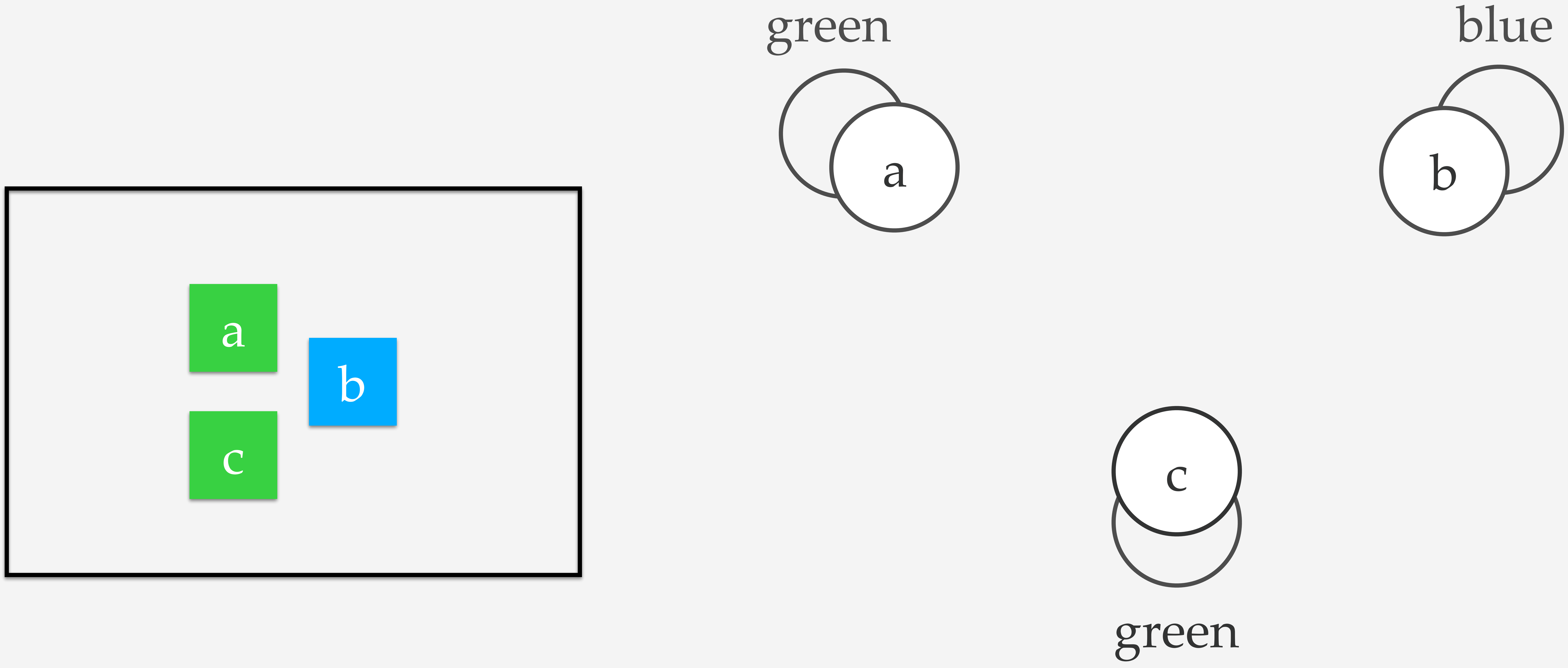
REG graph





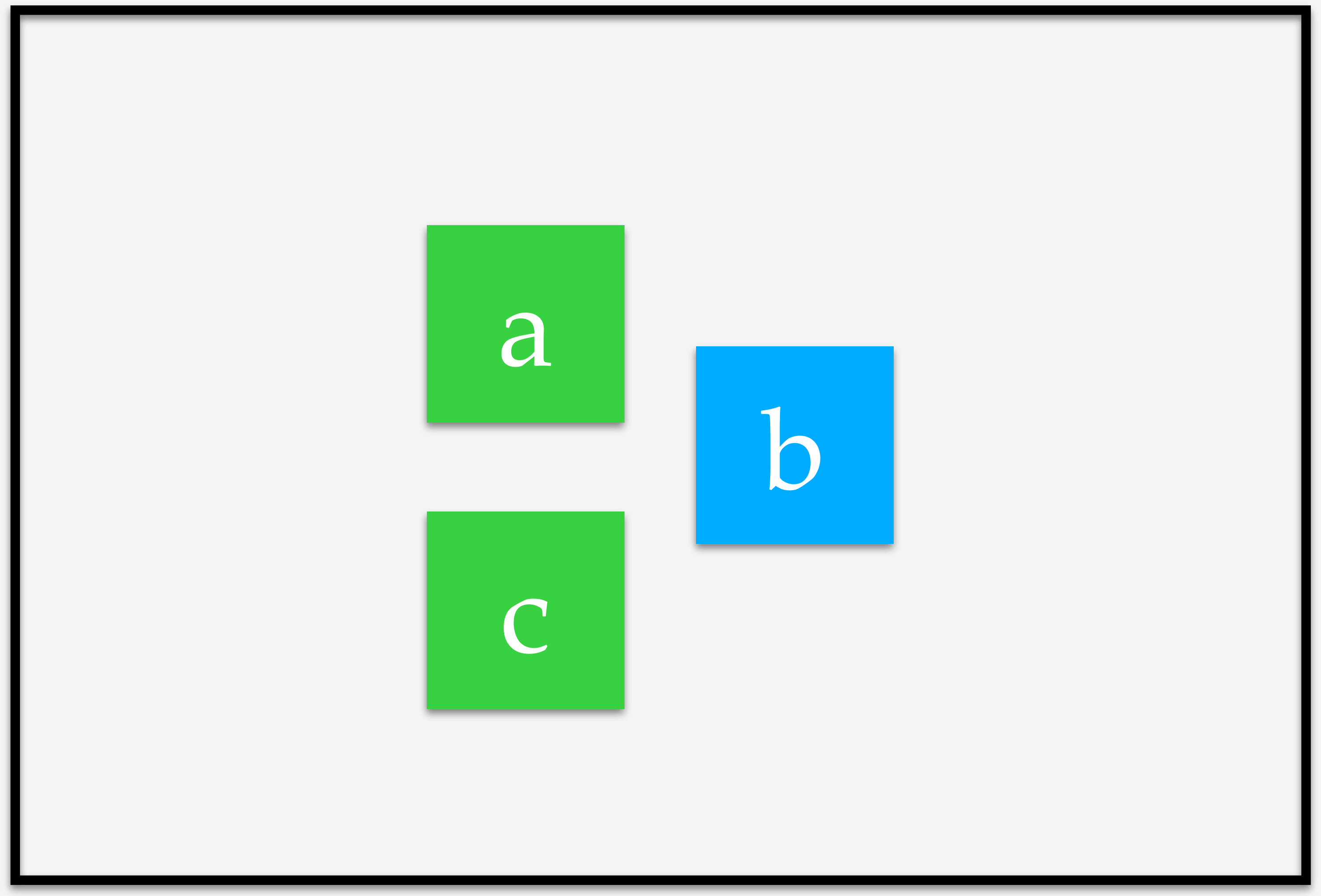
Object ⟺ Node

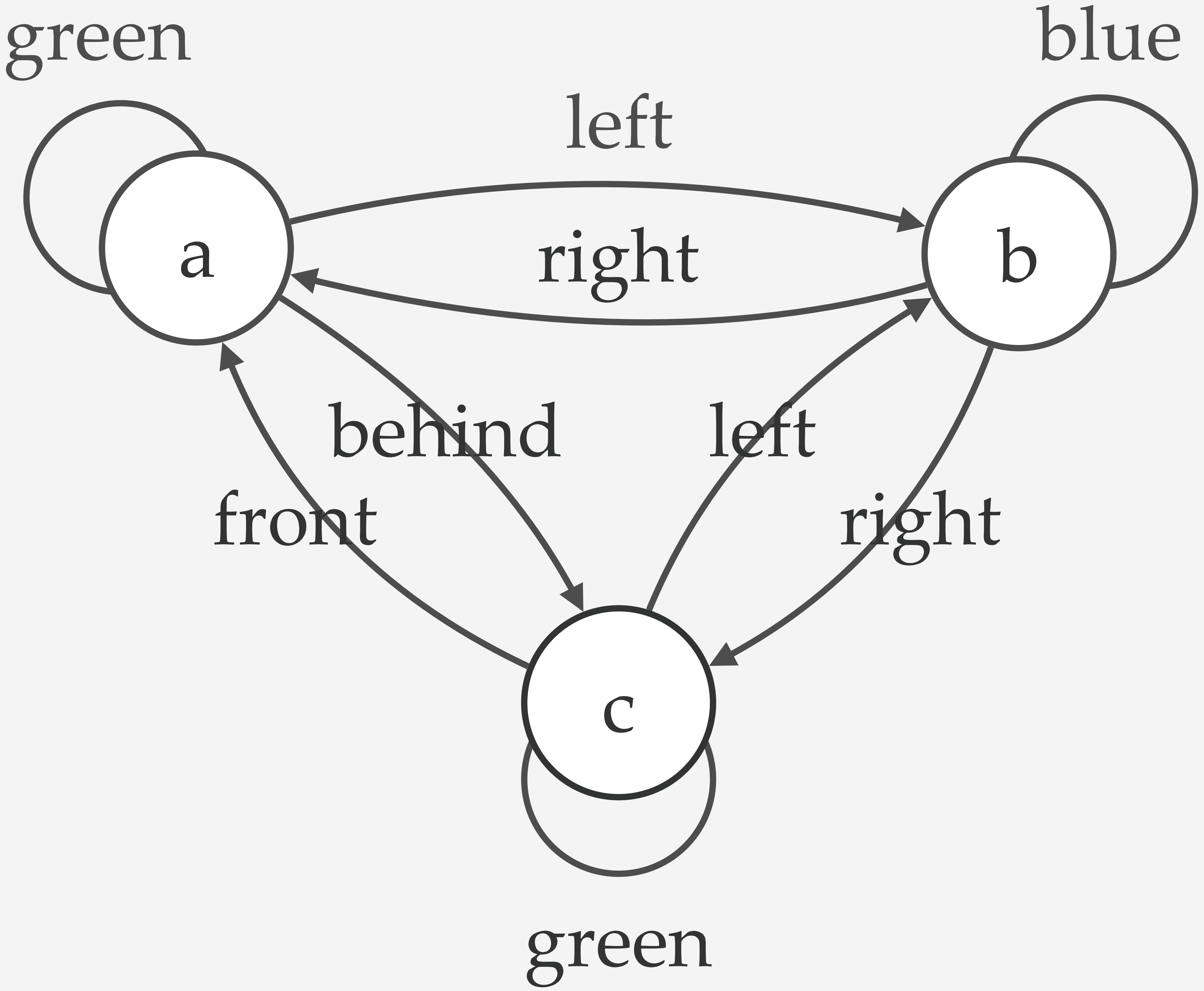
REG graph



Visual features ⟺ Self-loops

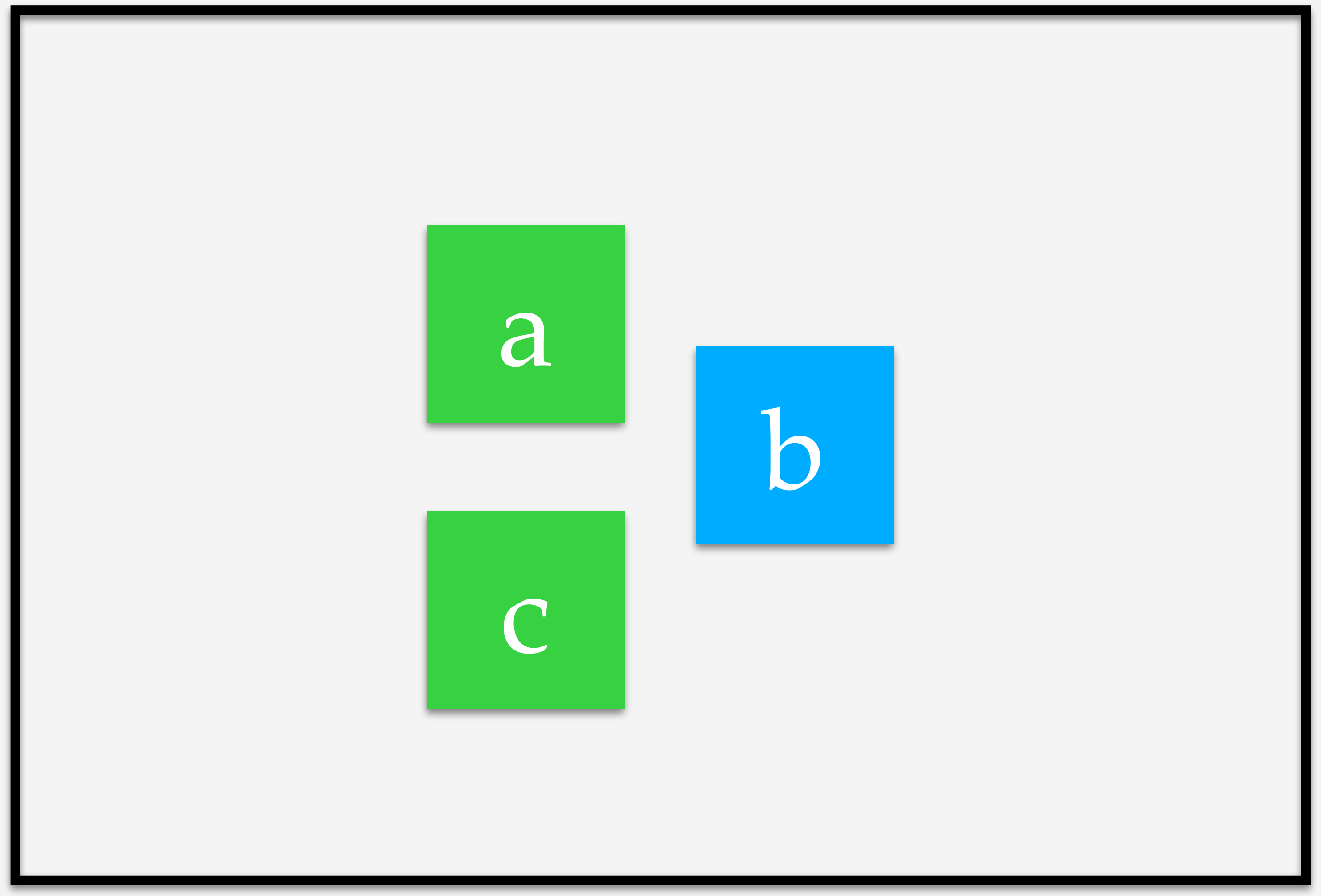
REG graph

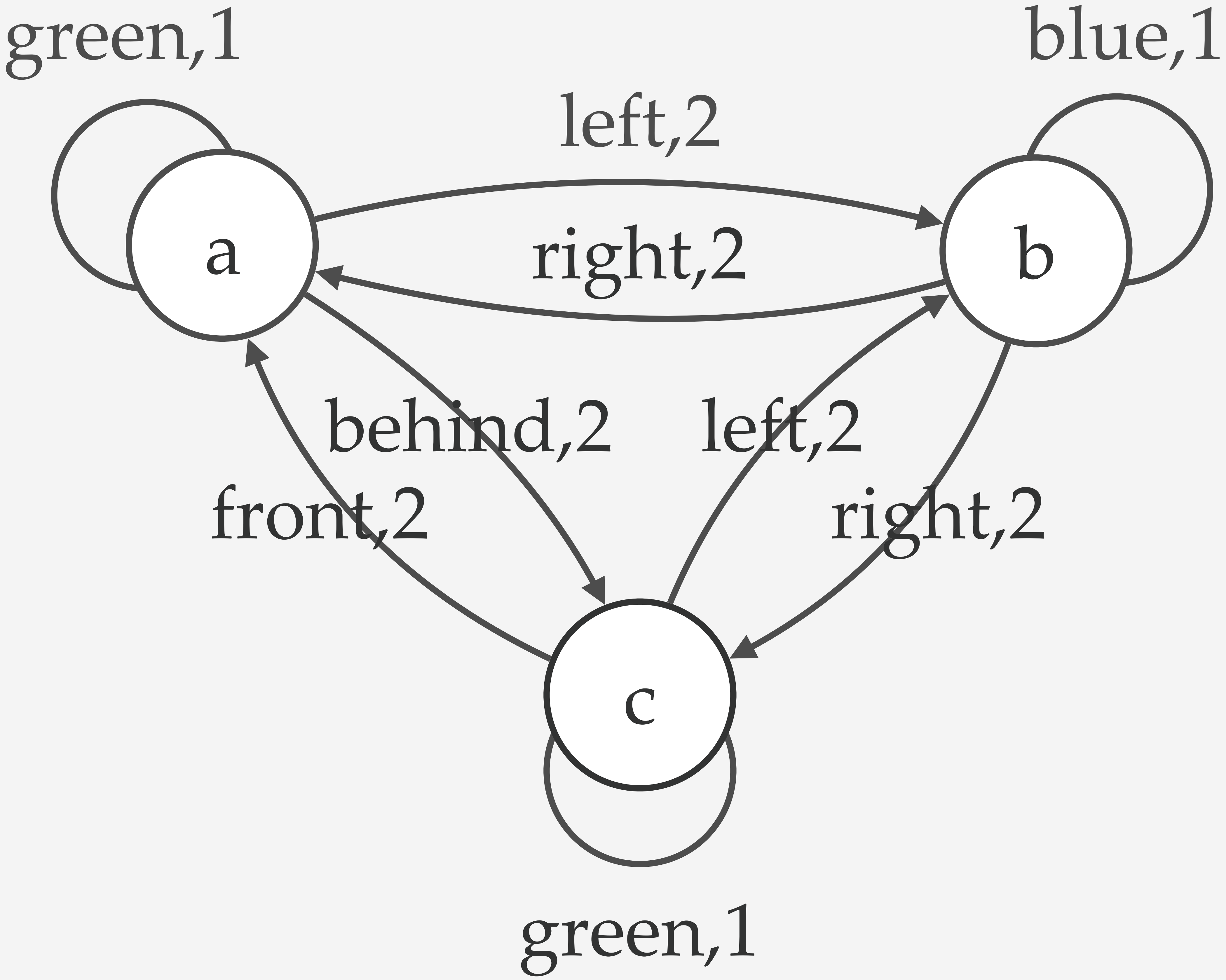




Spatial relations ⟺ **Binary edges**

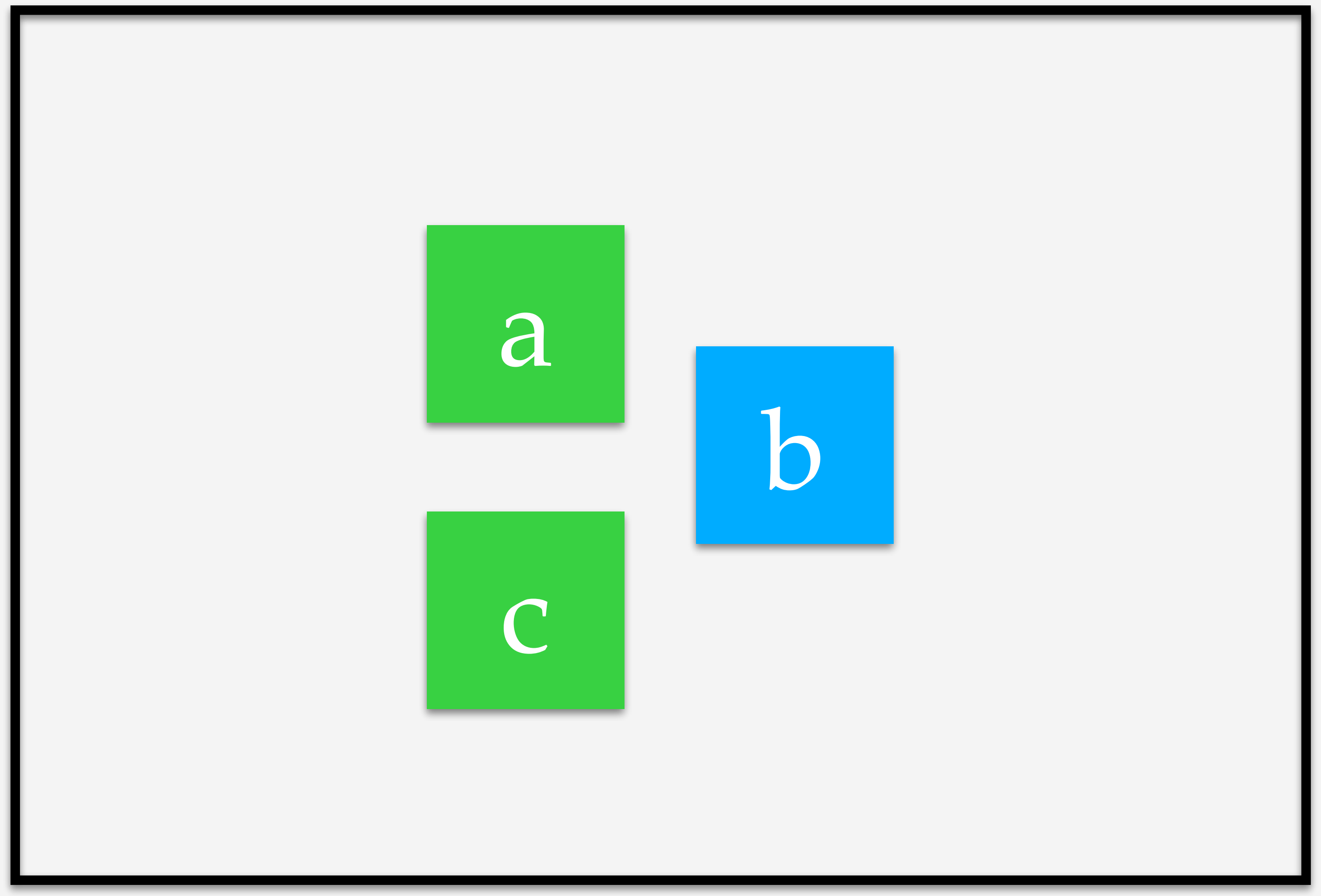
REG graph

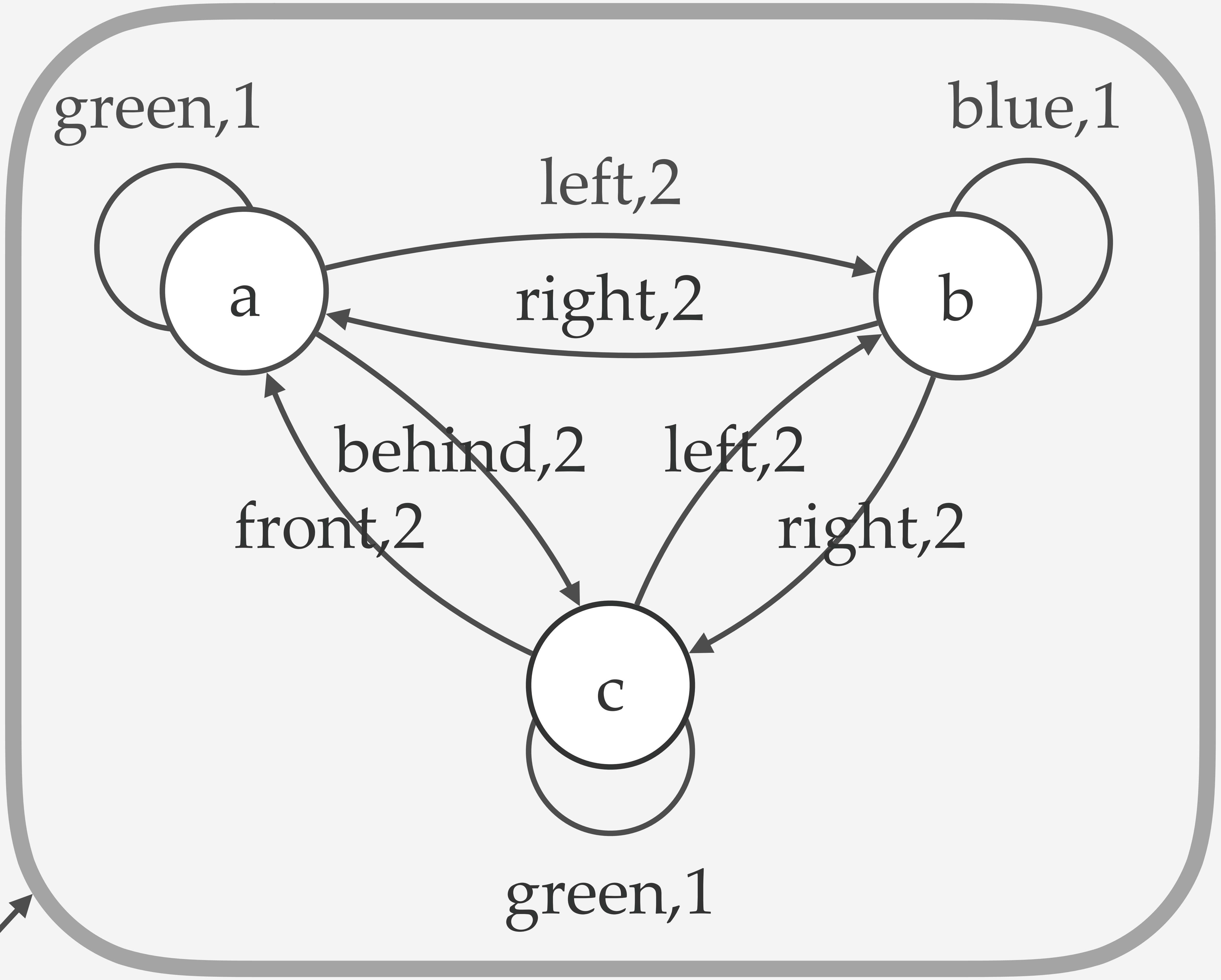




Human preferences ⟺ **Edge Cost**

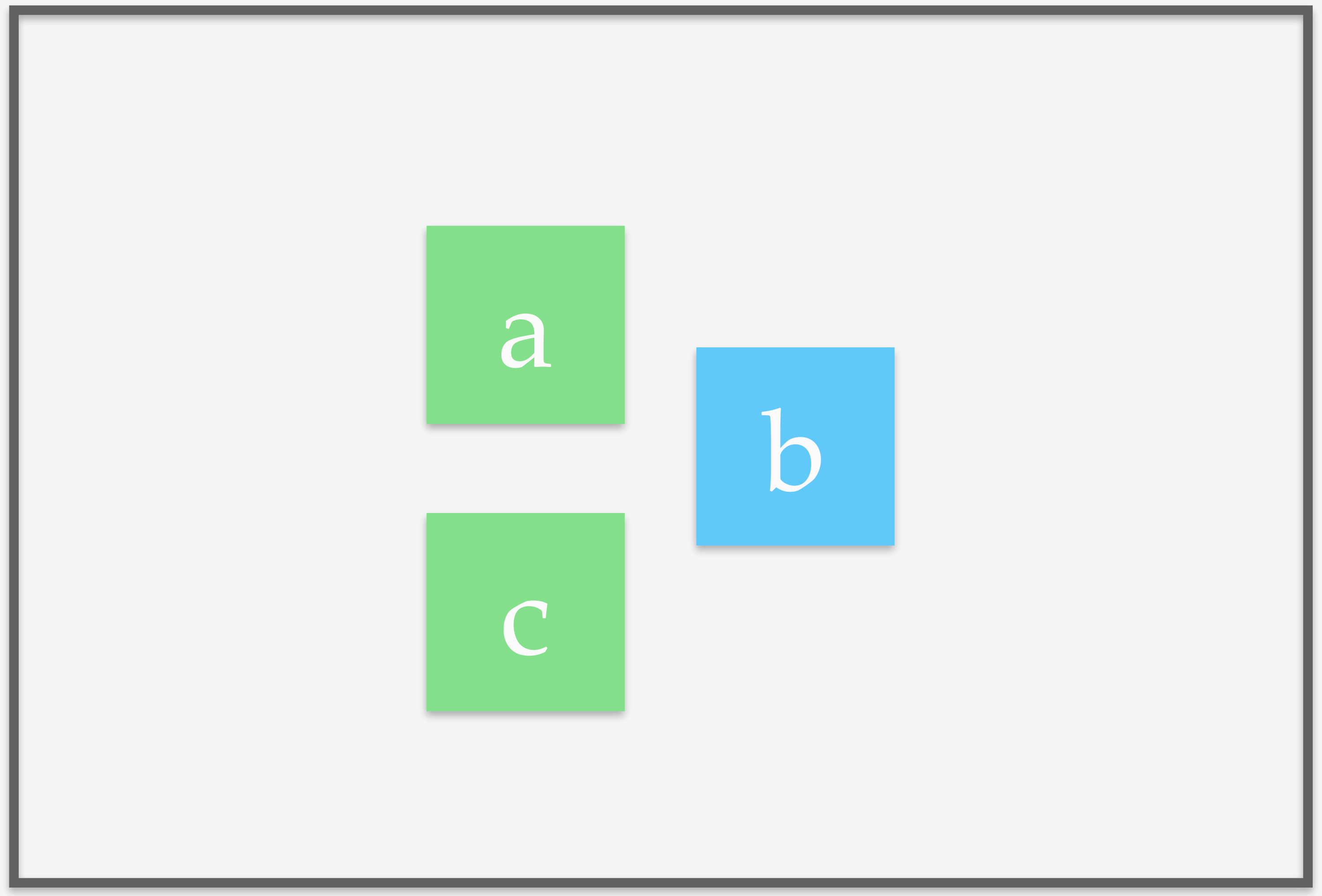
REG graph

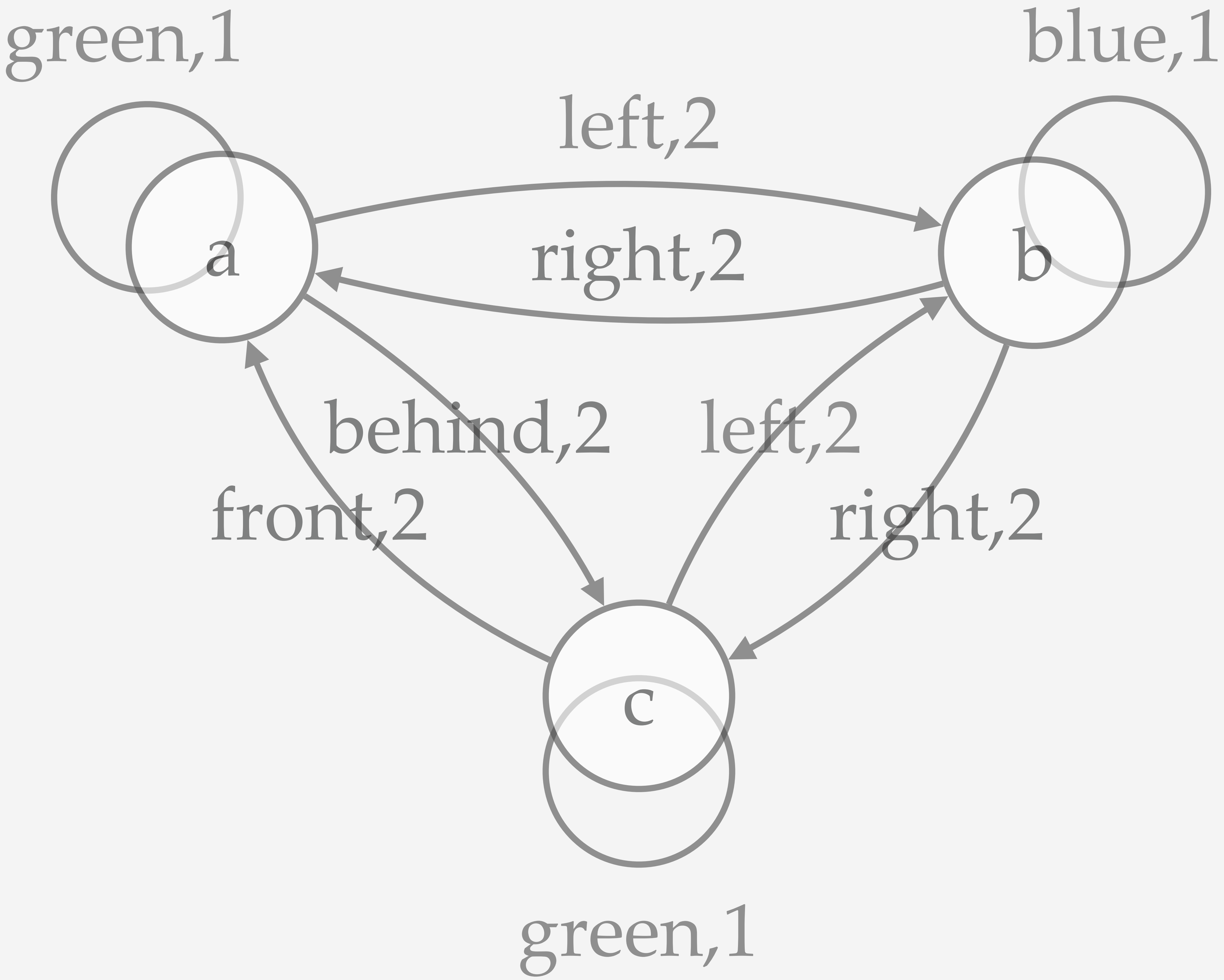




REG Graph

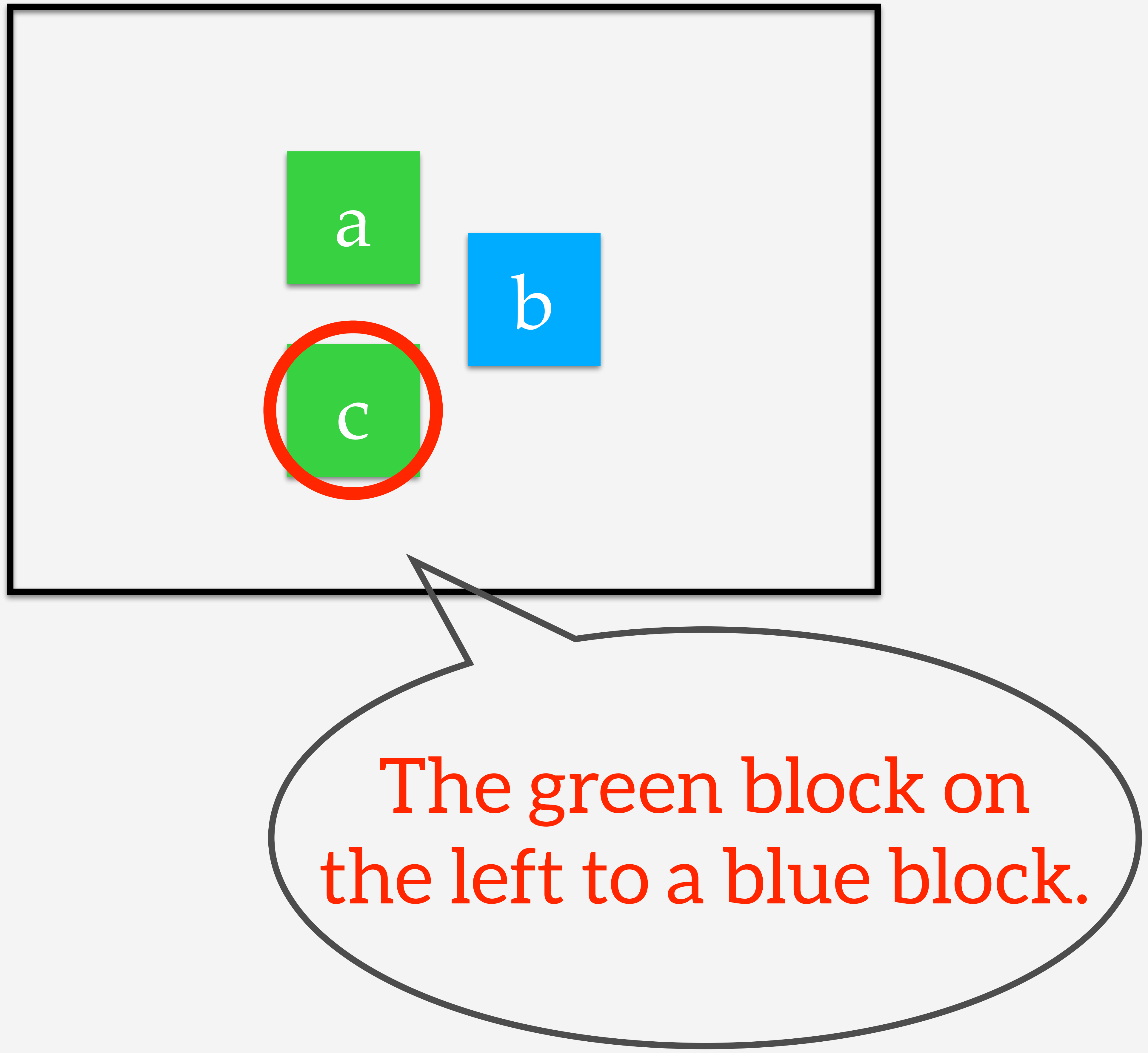
Mapping: scene <=> REG graph

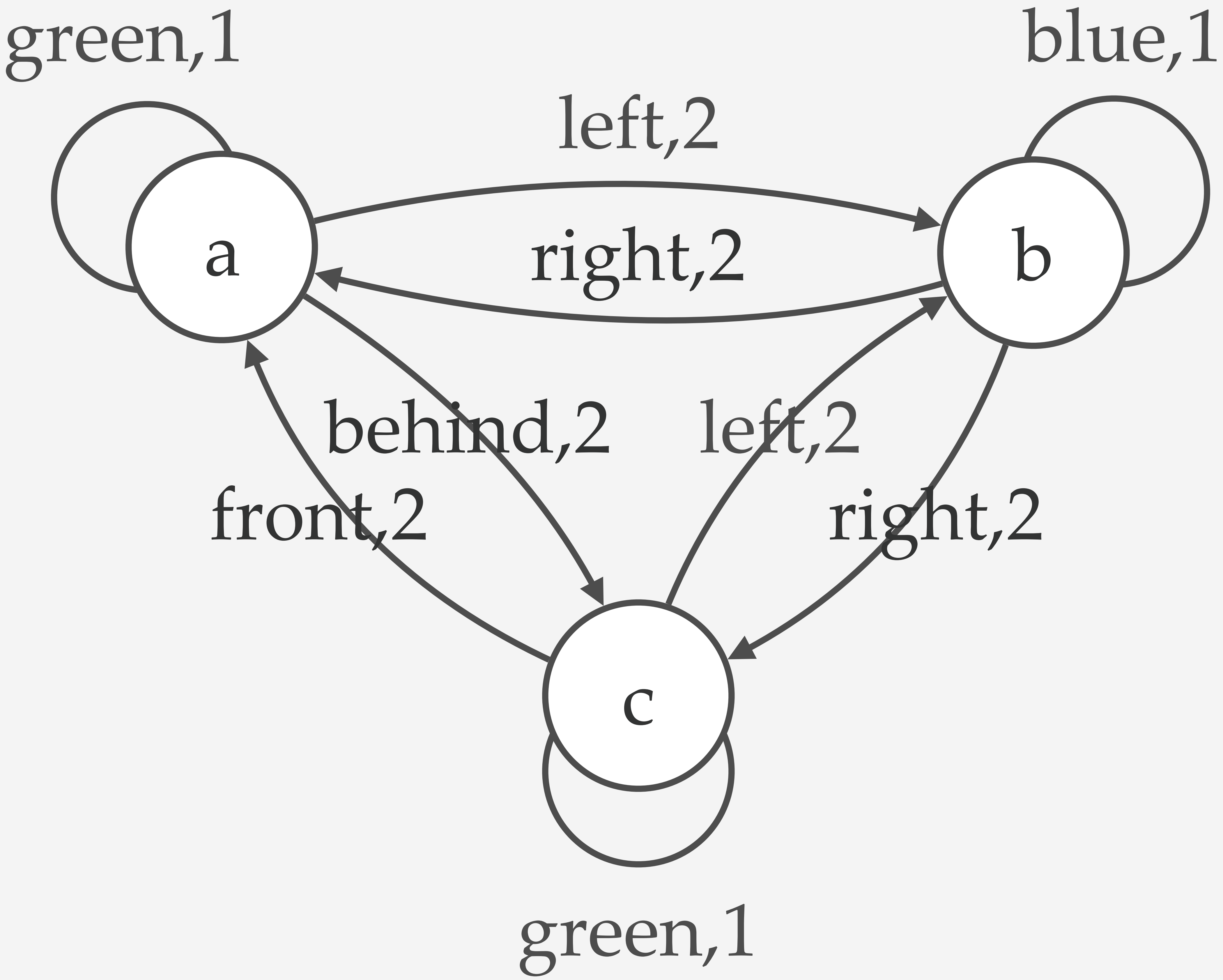




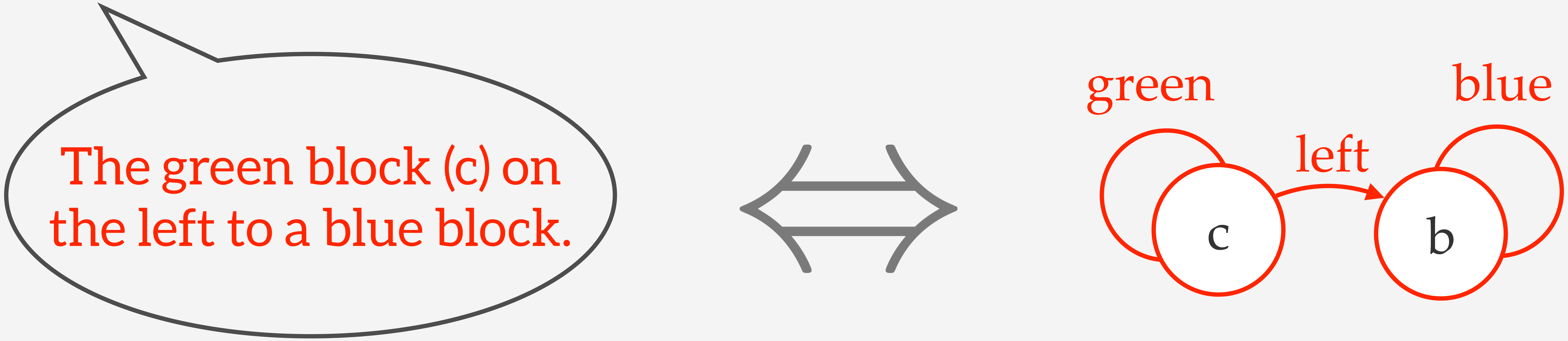
Scene ⟺ **REG graph**

Referring expressions

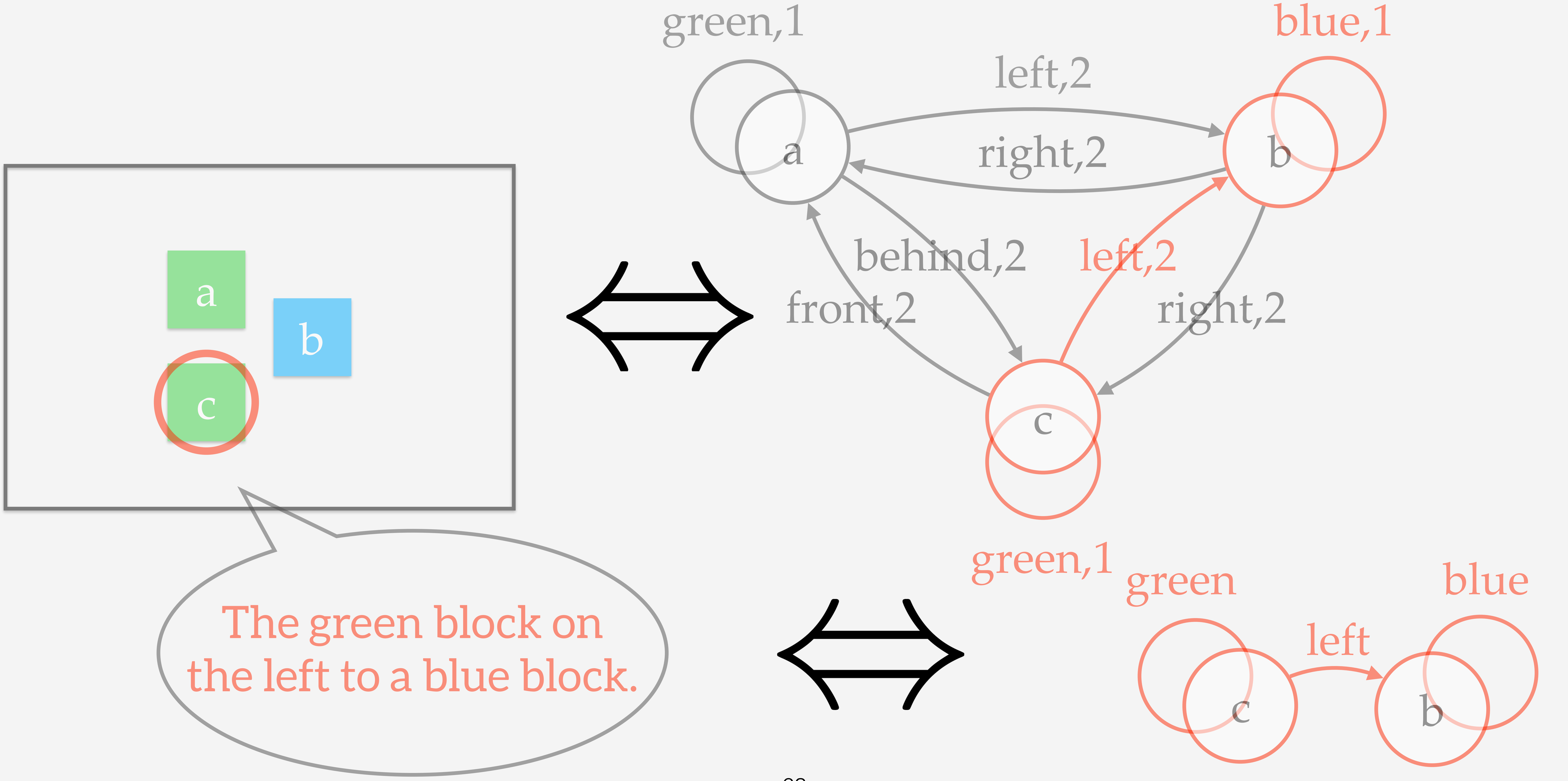




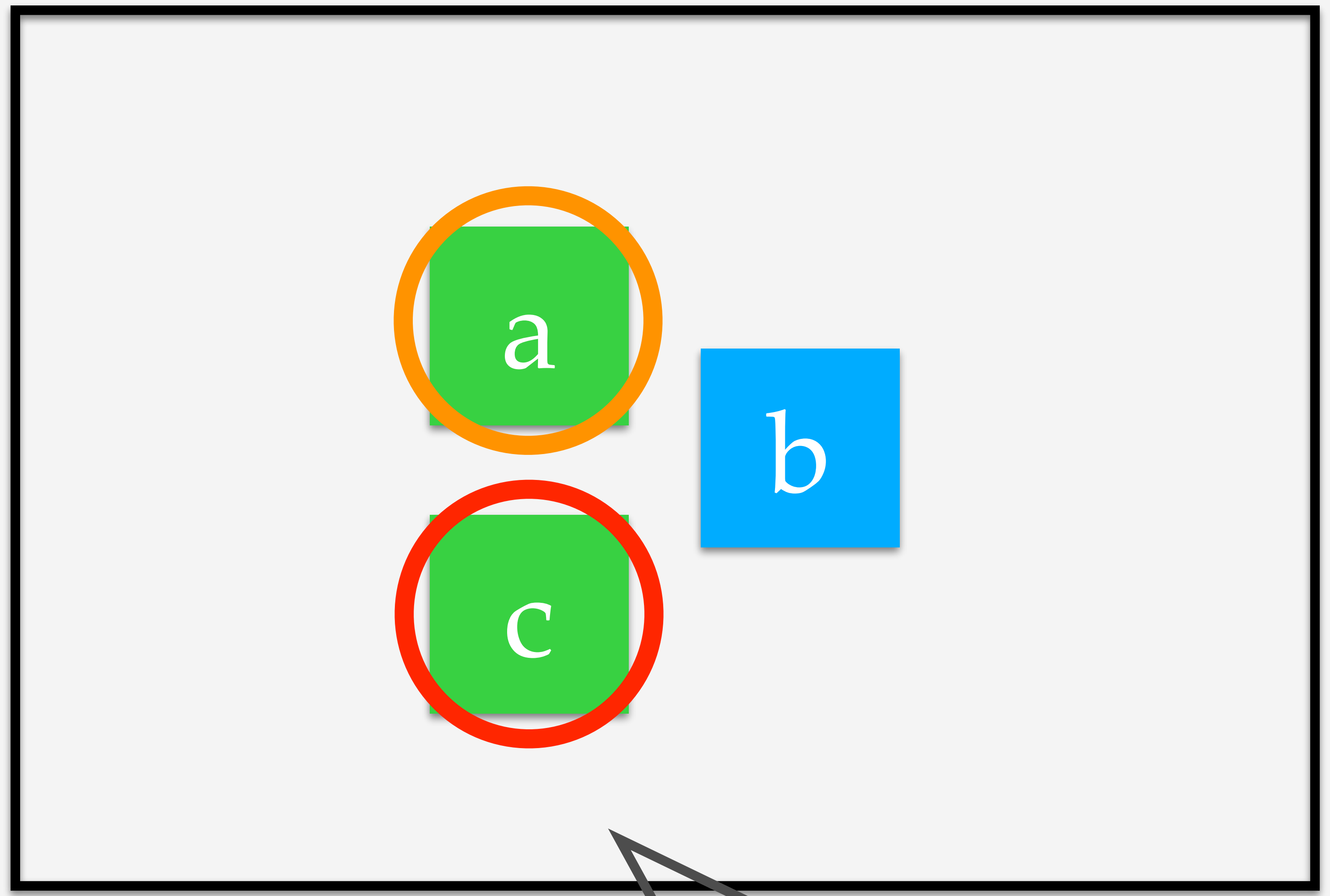
Mapping: RE <=> subgraph

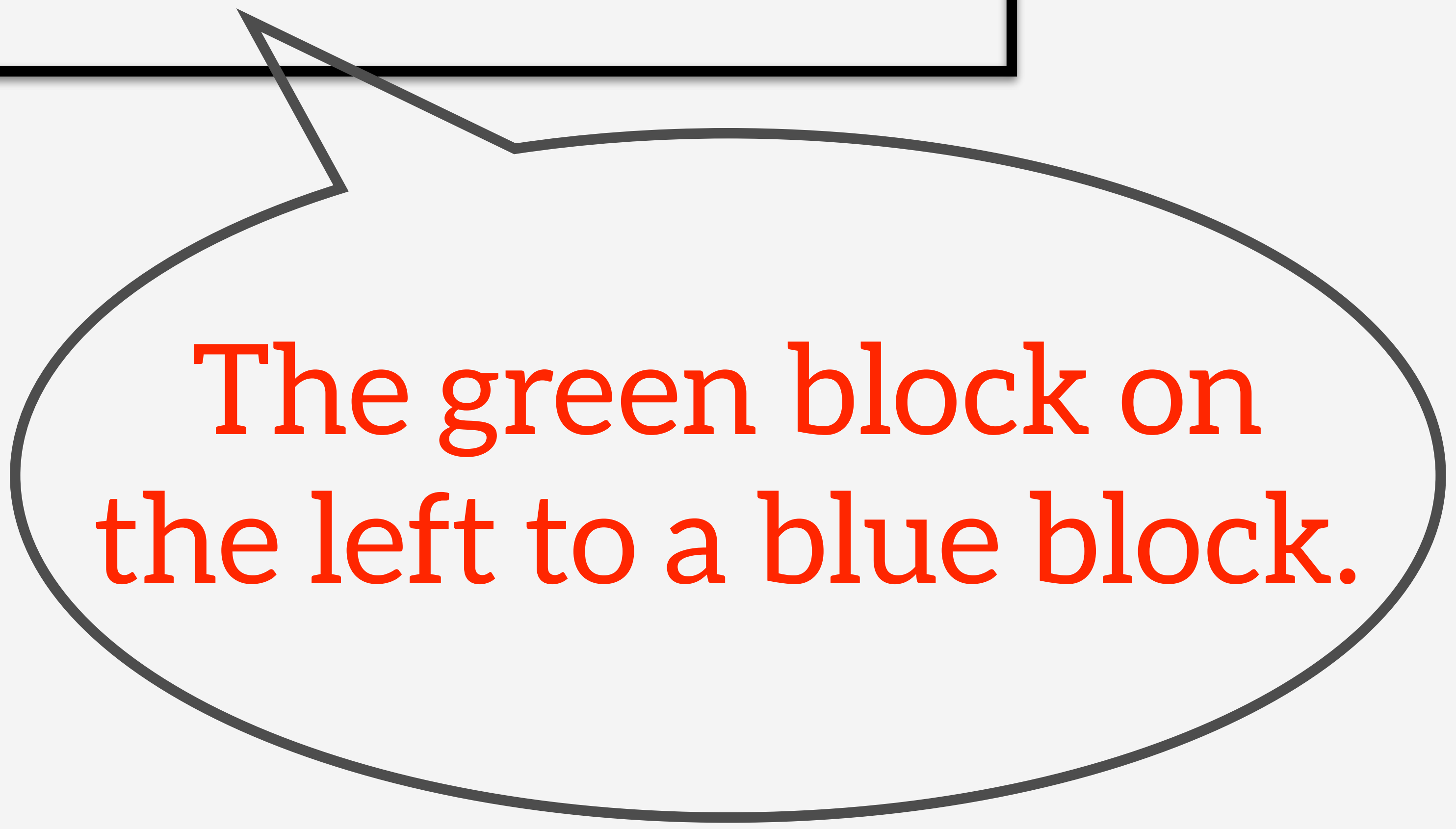


Scene <=> REG graph RE <=> subgraph



Clarity of referring expression





Uniqueness of subgraph

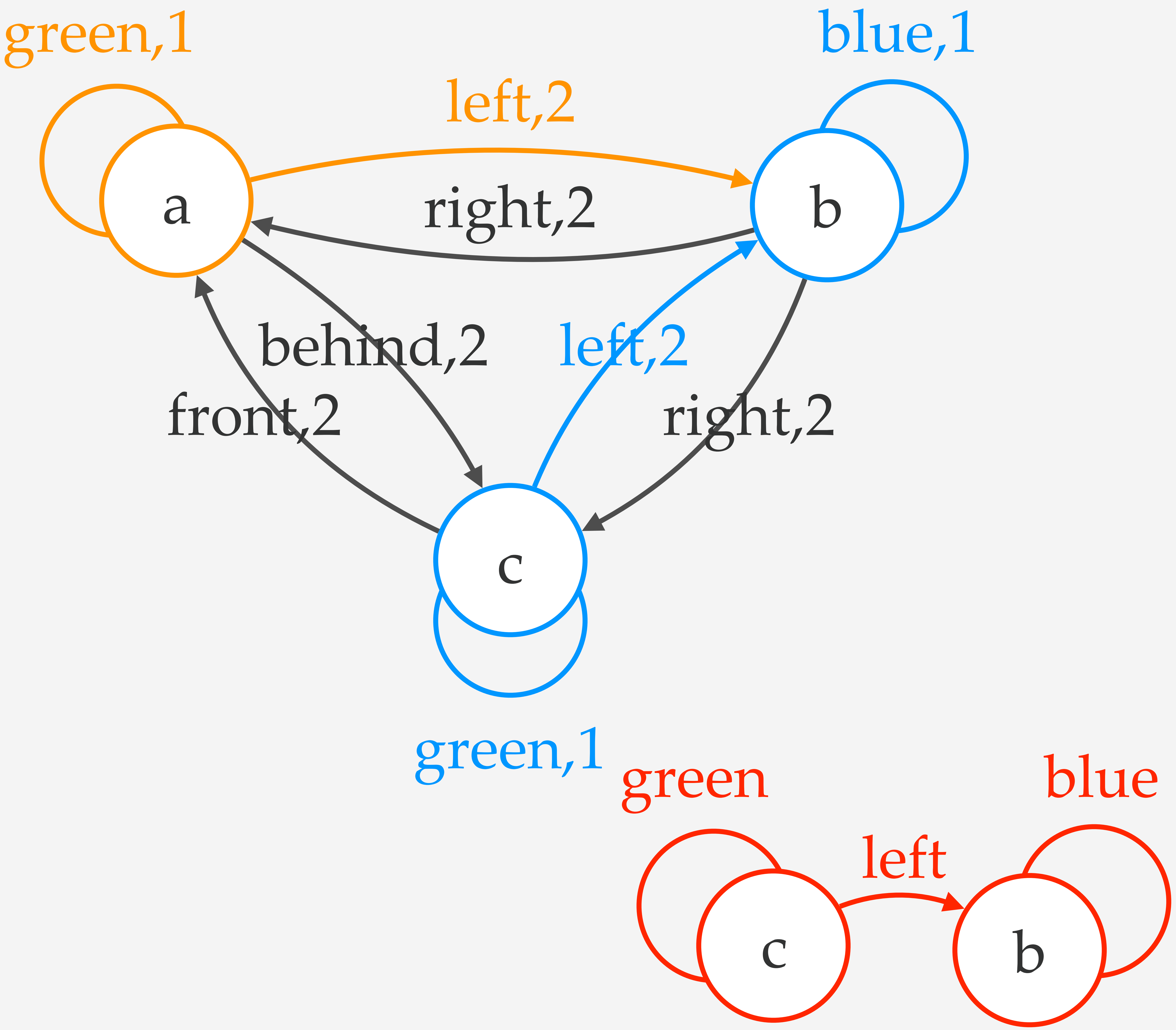
Subgraph in orange

Isomorphic

Subgraph in red

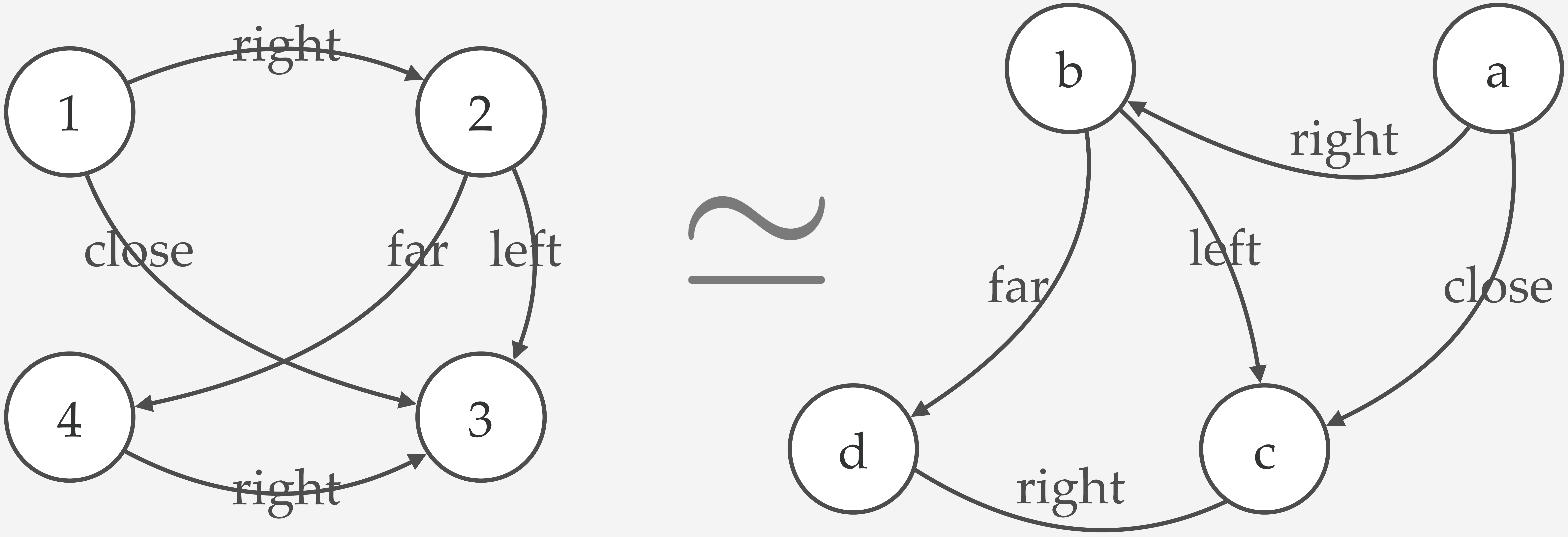
Isomorphic

Subgraph in blue

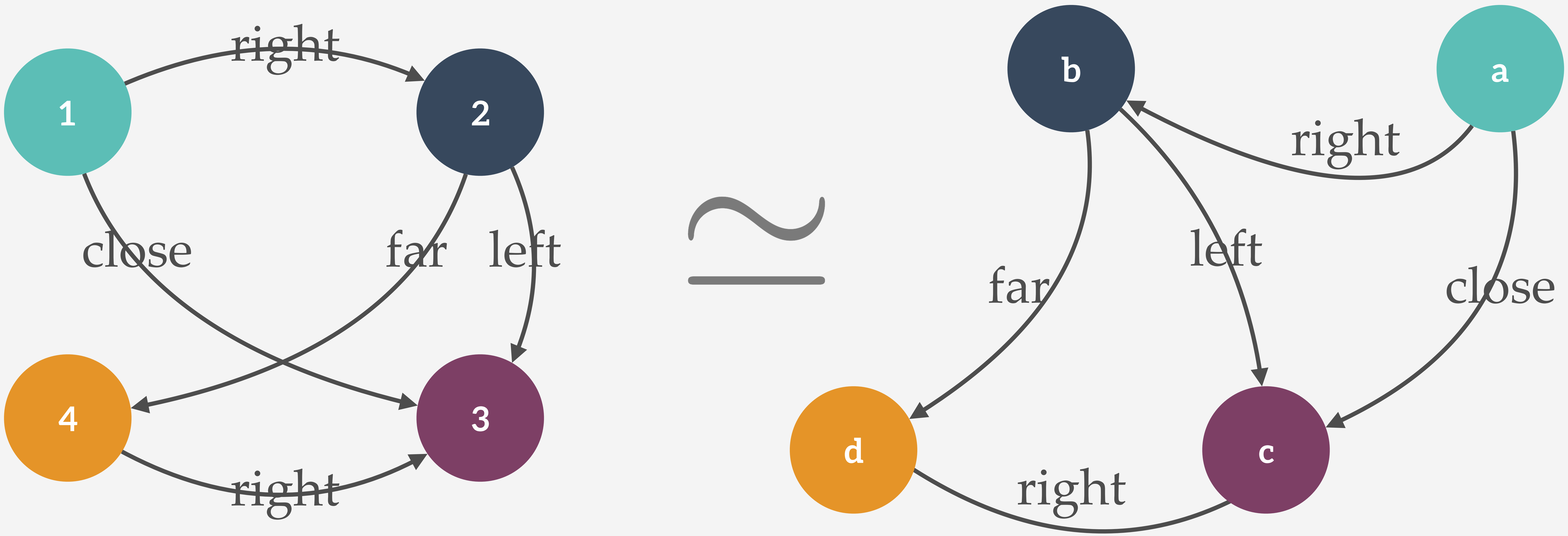


Graph isomorphism

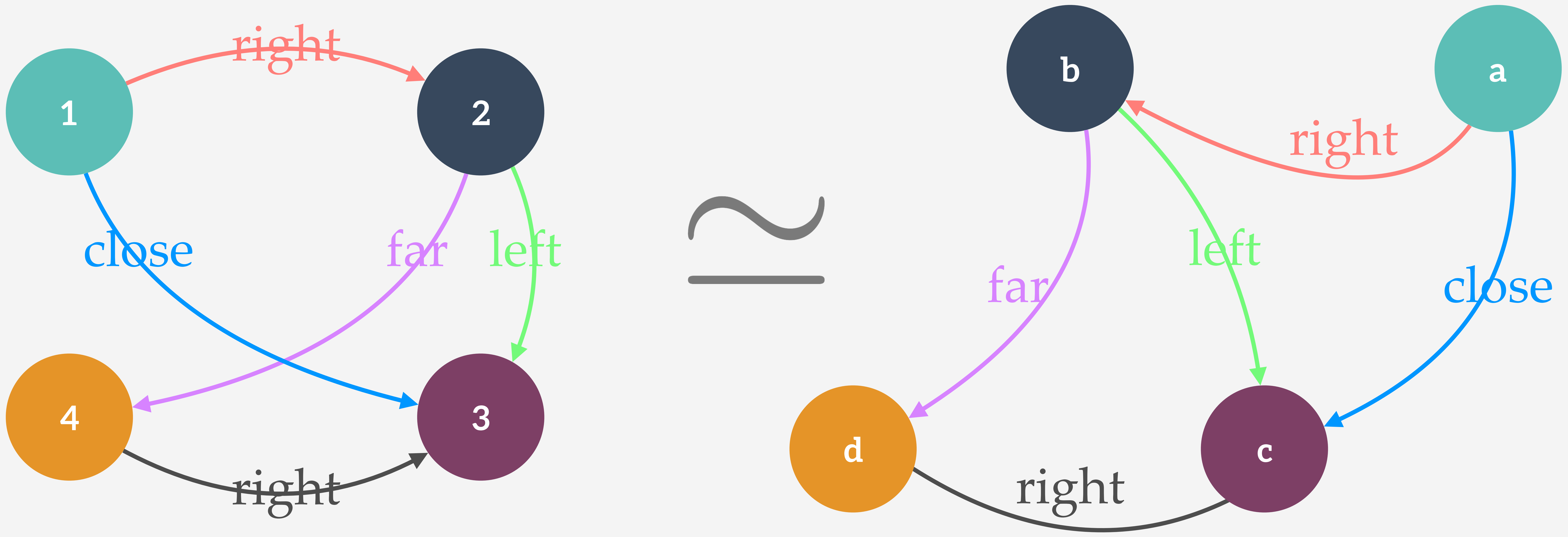
Edge-preserved bijection



Graph isomorphism



Graph isomorphism



Uniqueness of subgraph

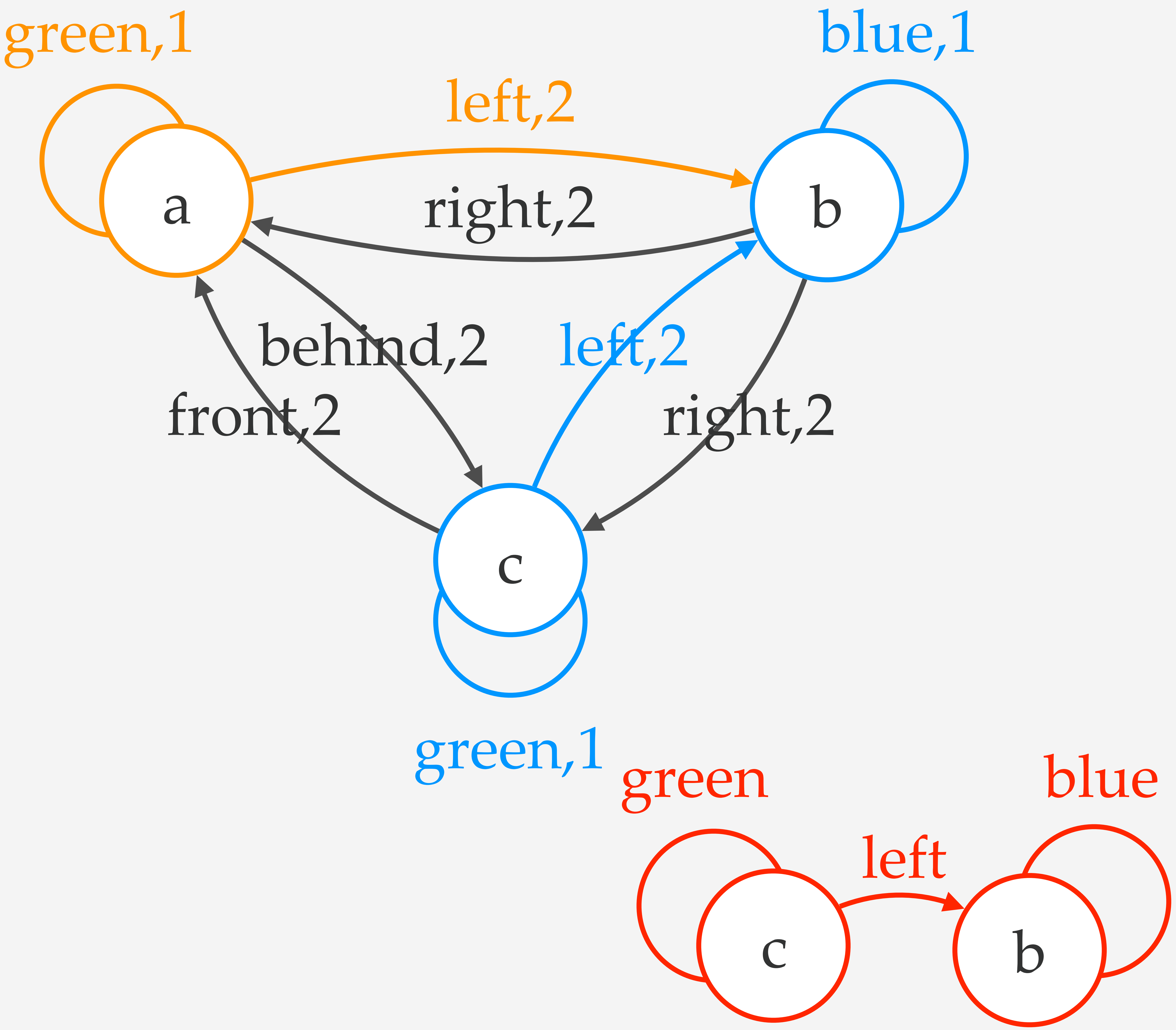
Subgraph in orange

$\cong$

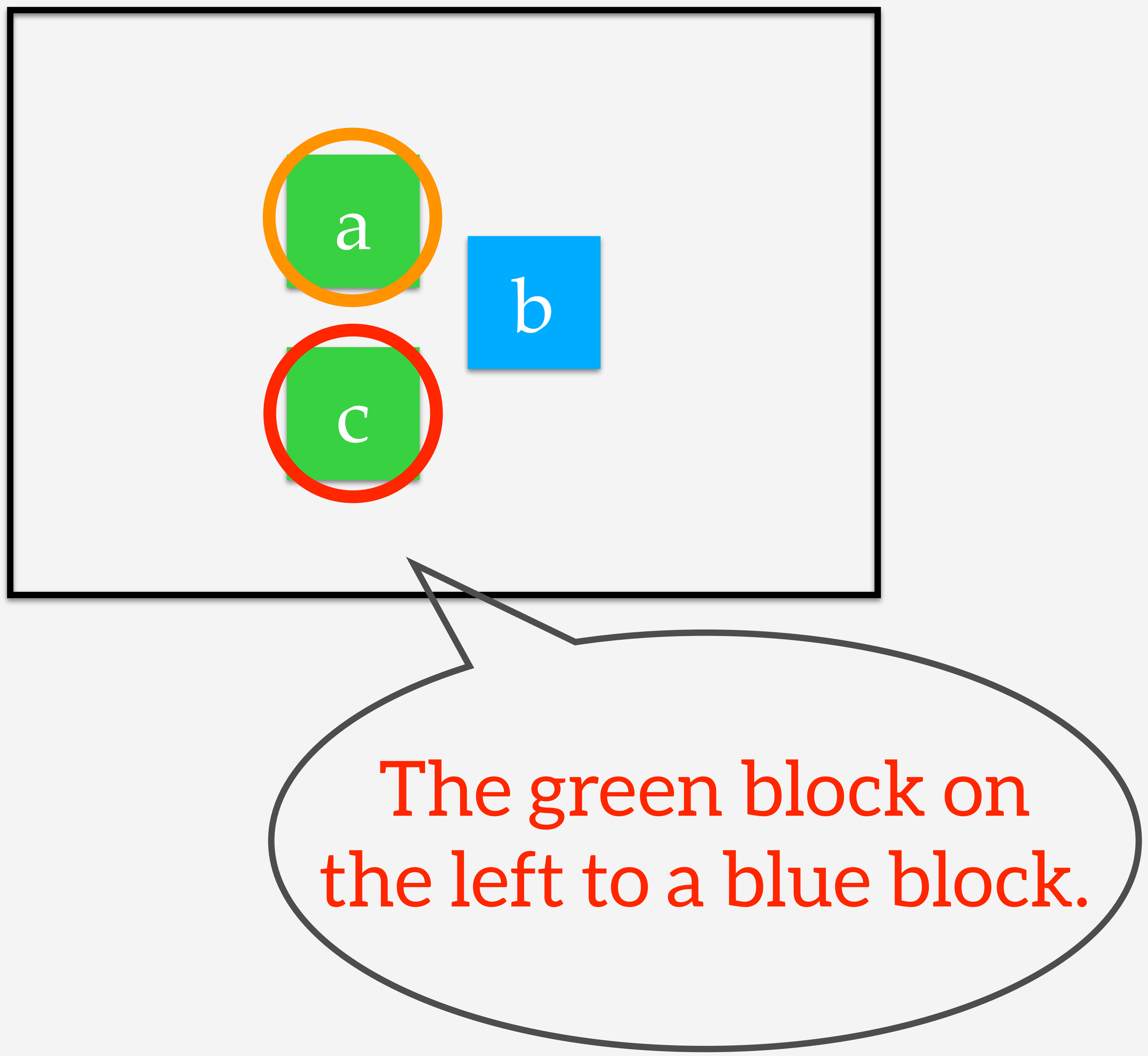
Subgraph in red

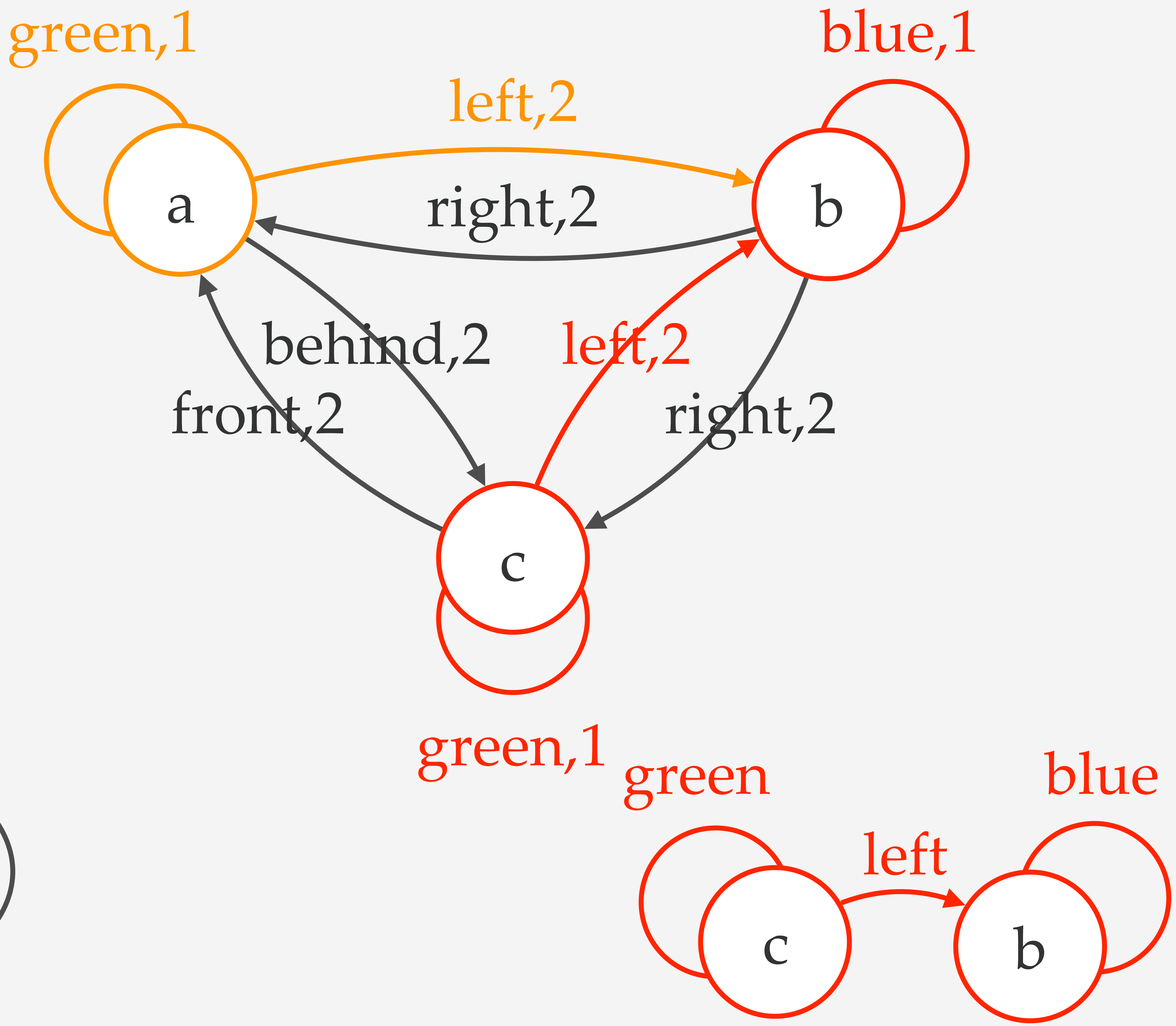
$\cong$

Subgraph in blue

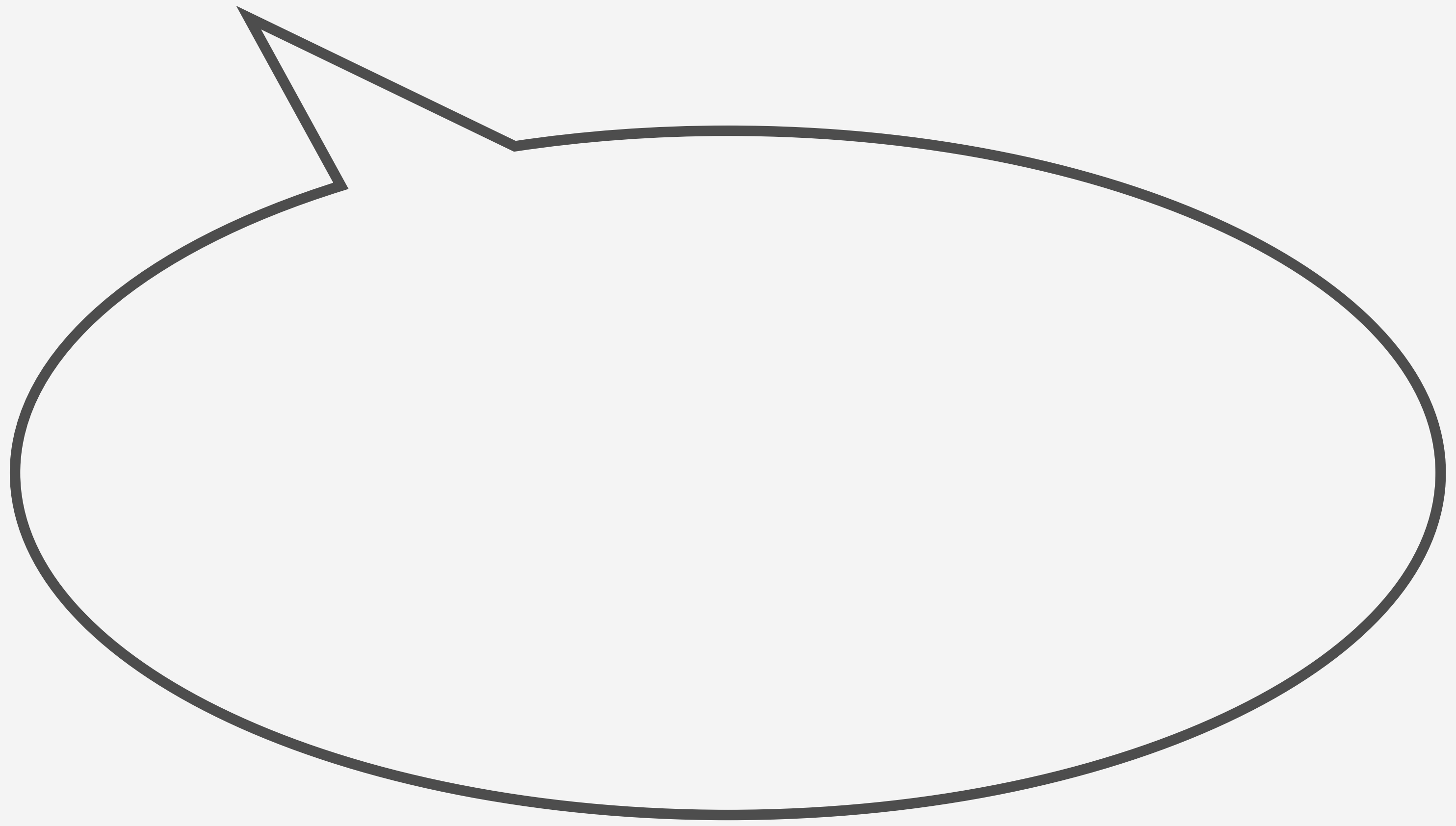


Mapping: RE clarity <=> Subg uniqueness



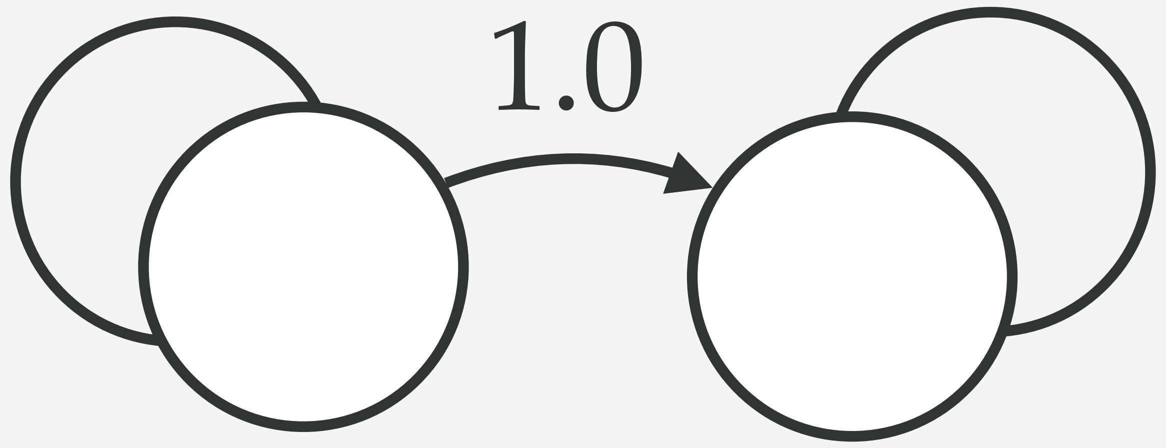


The referring expression is **ambiguous** because it *refers* to two objects. The subgraph is **not unique** because it is *isomorphic* to two subgraphs.



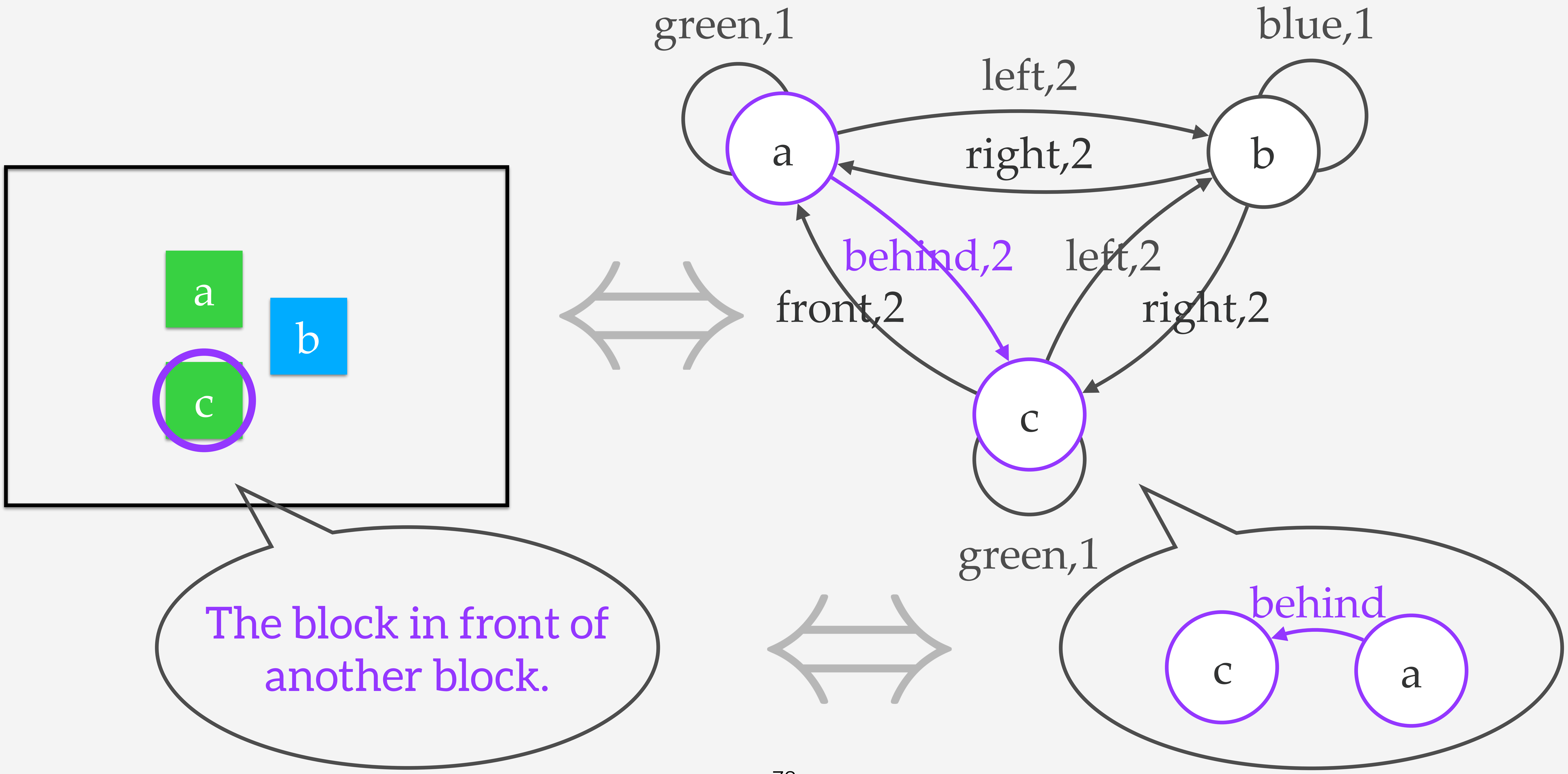
Clear referring expression

$\Leftrightarrow$



Unique subgraph

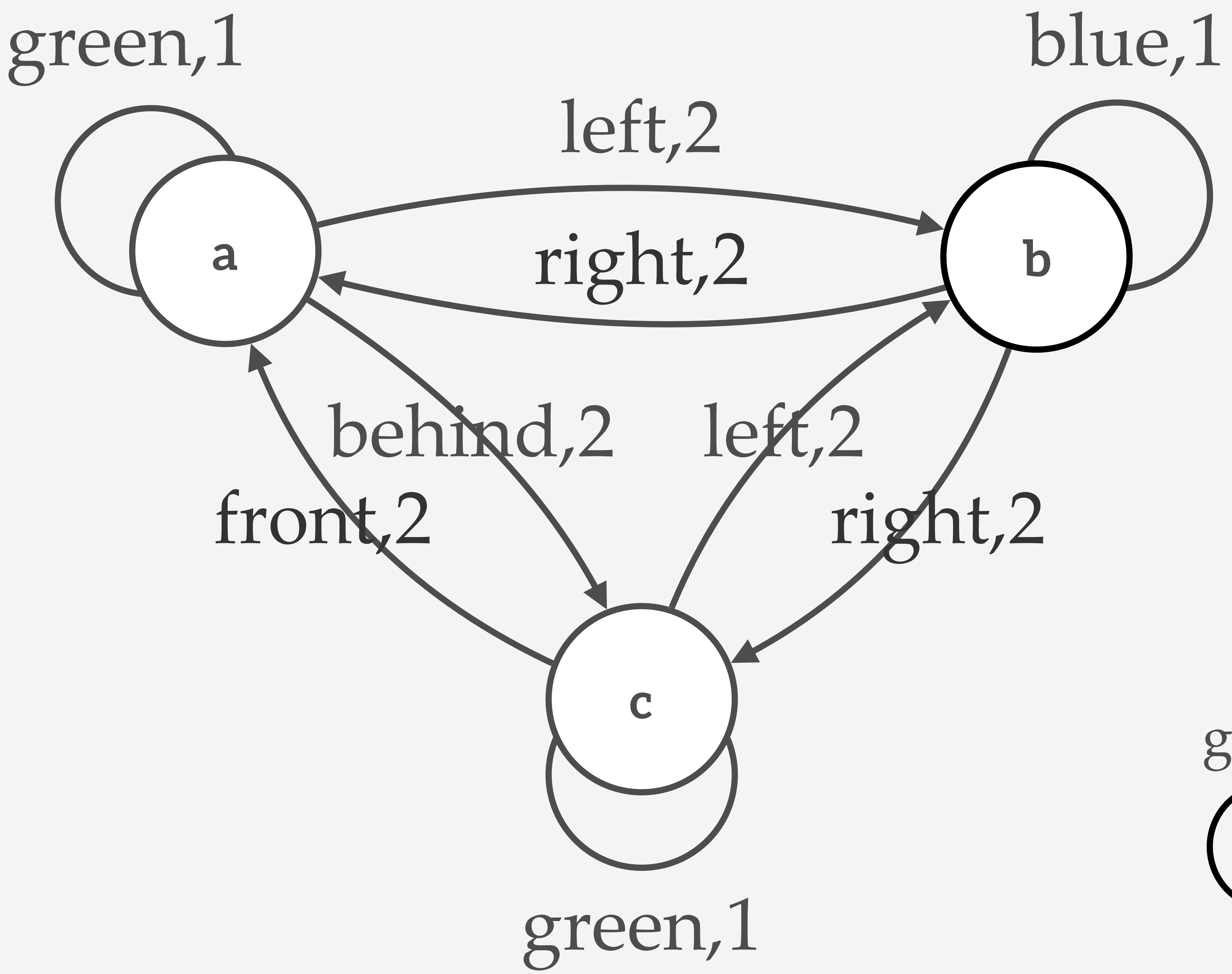
Mapping: RE clarity <=> Subg uniqueness

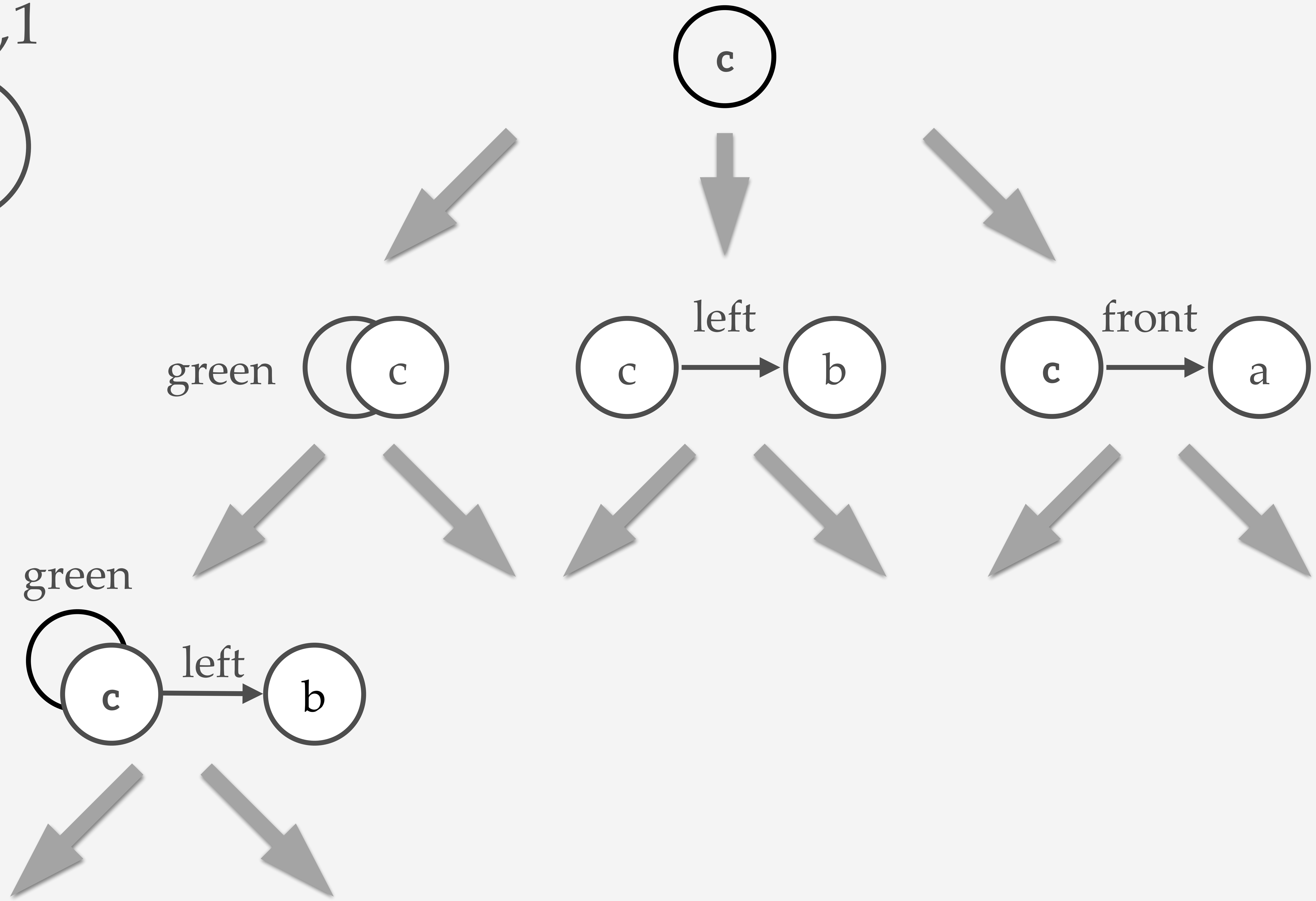


Search for unique subgraph

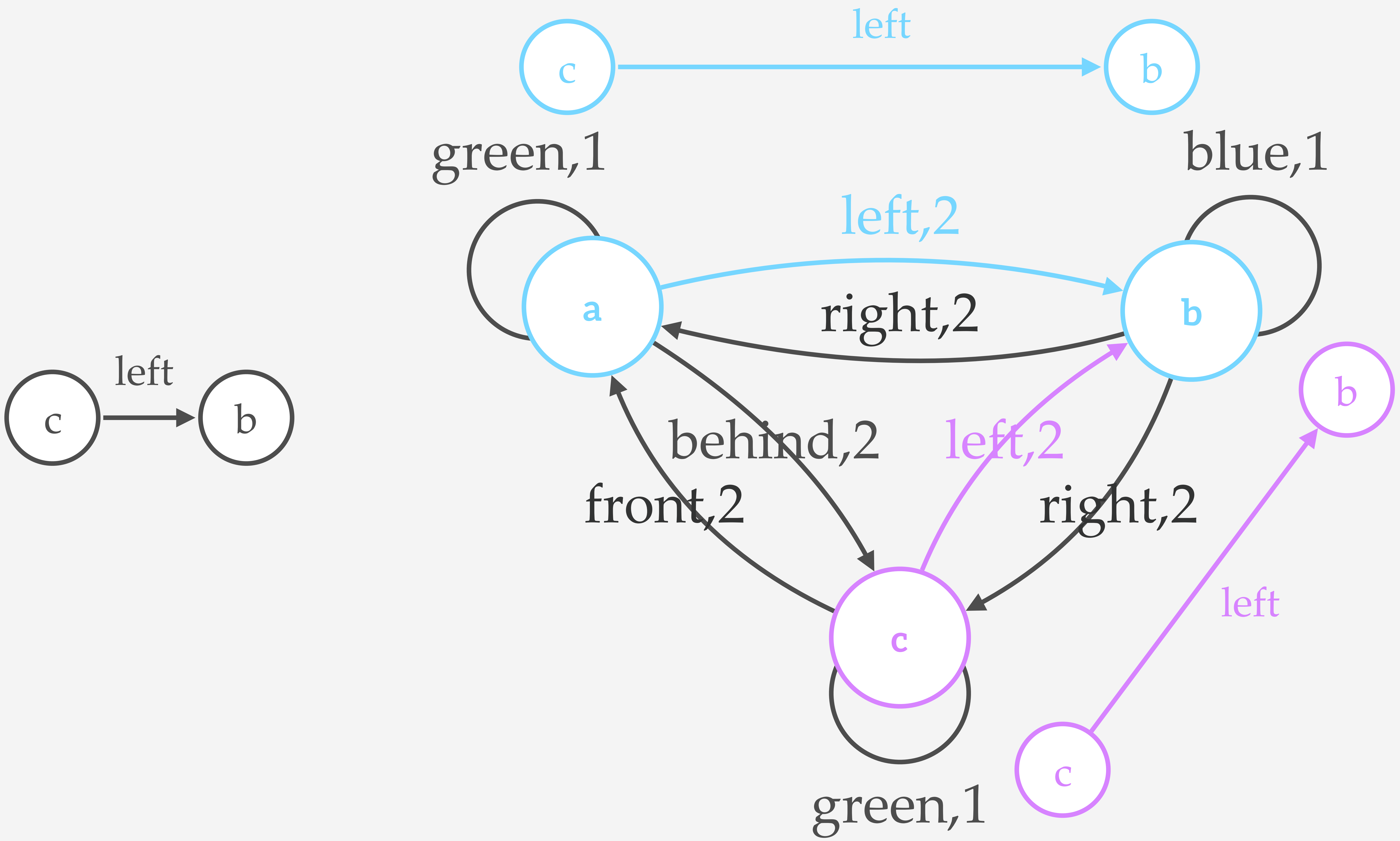
- **Search process**
  - Search for all the possible subgraphs with the target node.
- **Isomorphism process**
  - Verify the uniqueness of all subgraphs within the REG graph.

Search process





Isomorphism process

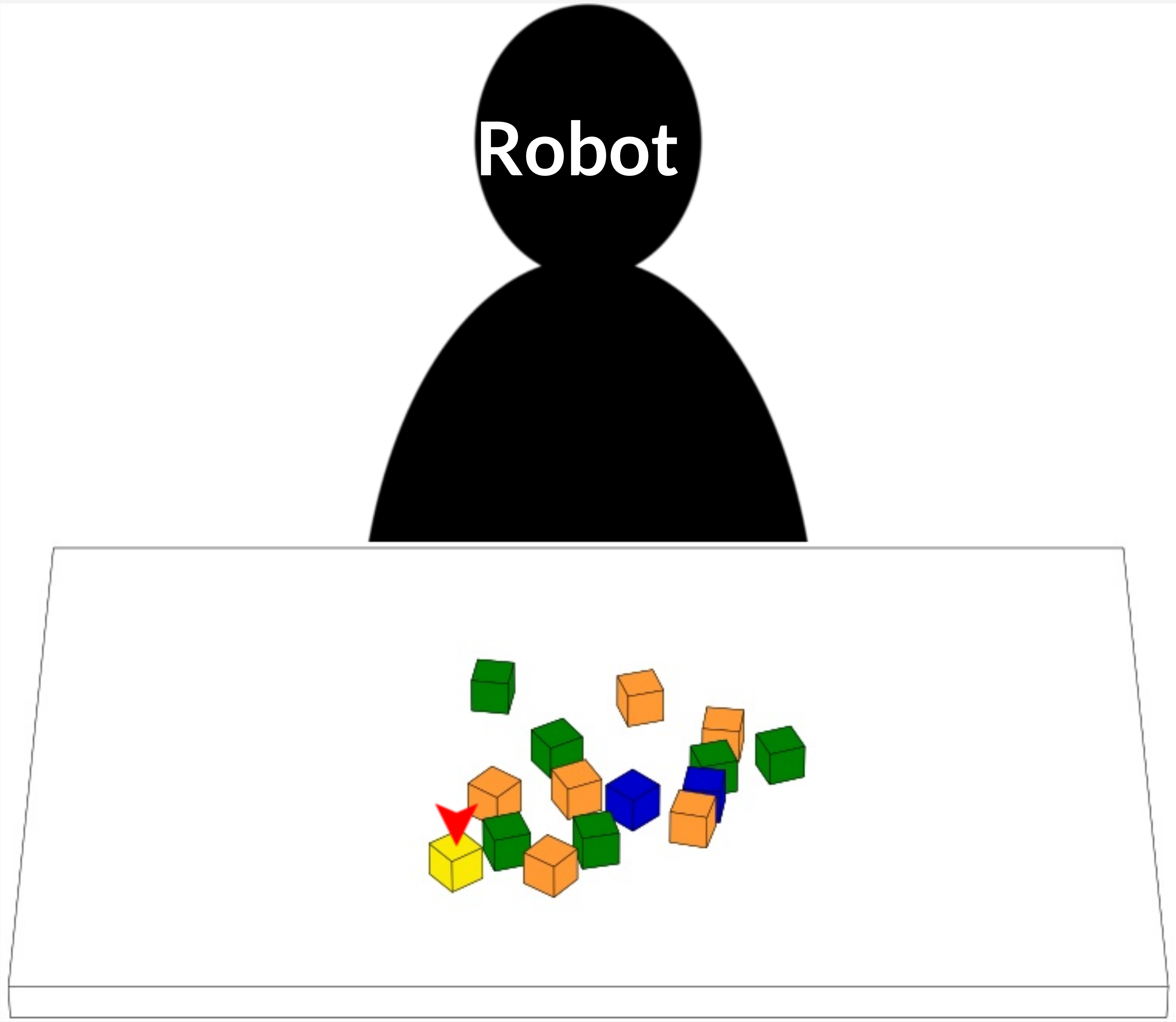


Our contribution

Referring expression generation (REG)

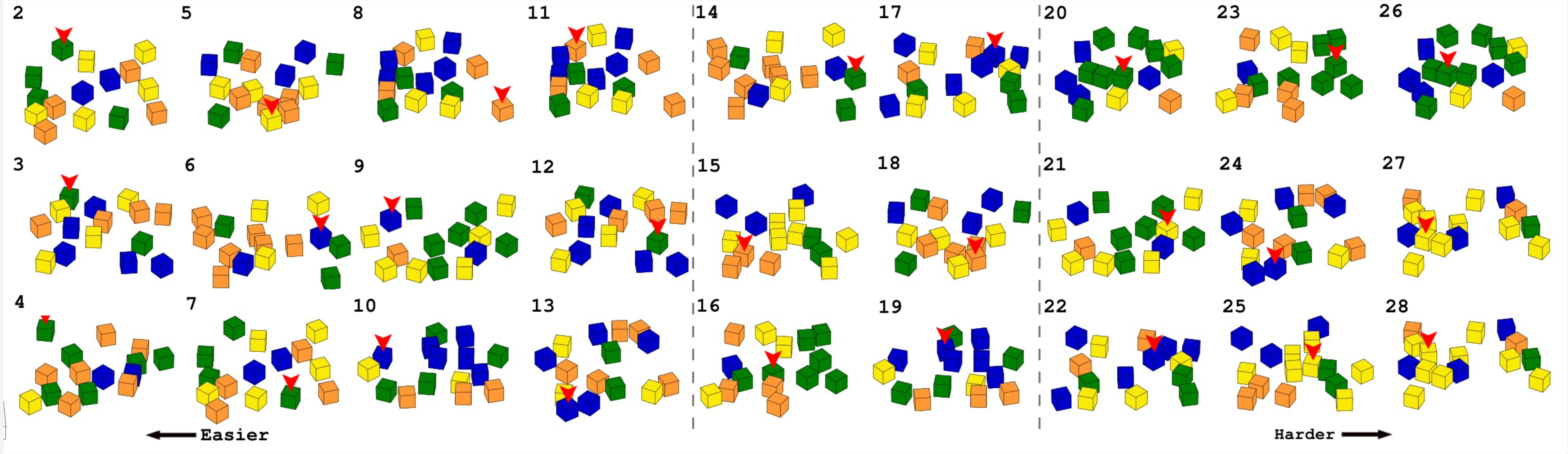
- Previous work on REG
- Our contribution on REG
  - Corpus
    - Visual features
    - Spatial relations
  - Algorithm efficiency
  - Graph structure

User Study



Participant

Scenes


2
5
8
11
14
17
20
23
26
3
6
9
12
15
18
21
24
27
4
7
10
13
16
19
22
25
28
Easier
Harder

User study

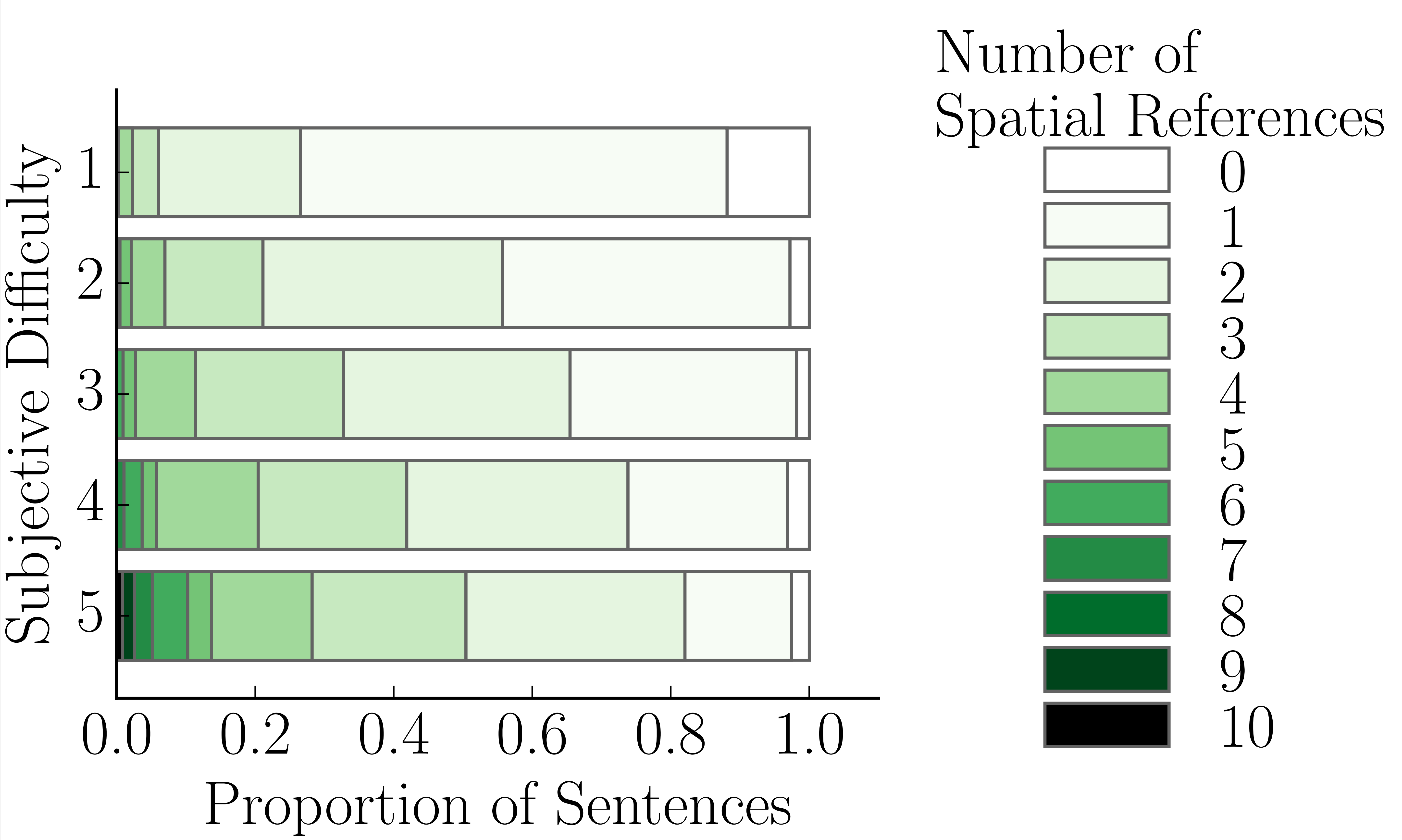
- 1400 instructions from 120 participants for 28 different target blocks in 14 scenes.

High level features

- Ordering and quantity
  - “The **third** from my left”
- Cluster
  - “pick up the middle green block from **the group of 5**.”
- Shape
  - “Look for a green block. Look for a green block that is very close to another green block. The green blocks should look like they combine to **form a rectangle**. Pick up the left most block of those two.”

Spatial relations

- Used more to deal with complex scenes

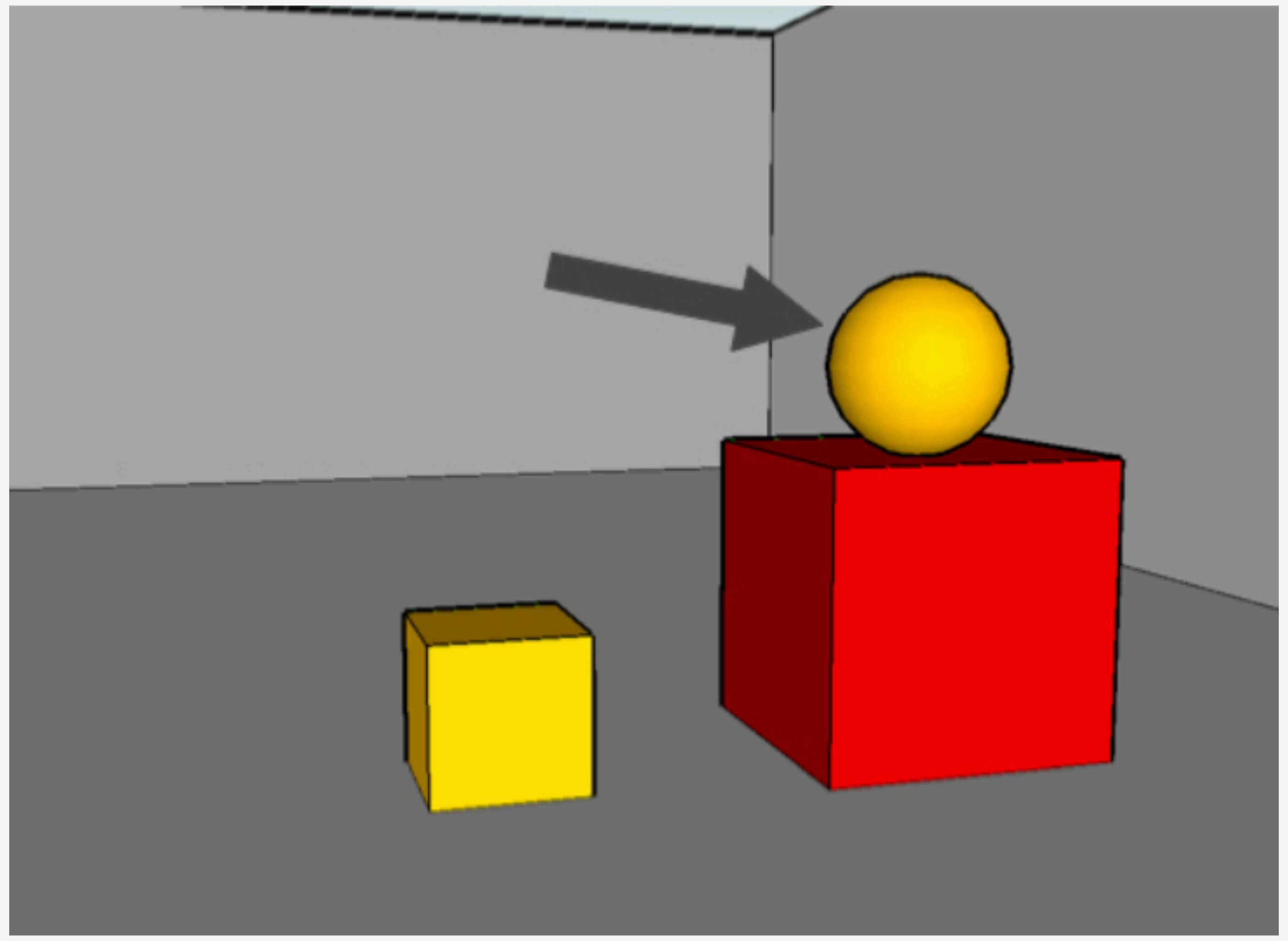


Spatial relations

- Used more to deal with complex scenes
- Qualitative
  - Orientation, e.g. “to the left of” instead of “53 degrees”
  - Distance, e.g. “close to” instead of “2 cm away from”

Issues in graph-based REG

- Computational complexity
  - Algorithms are tested in simple scenes with 3 objects
- Support for higher level features



Viethen, J., & Dale, R. (2008, June). The use of spatial relations in referring expression generation. In Proceedings of the Fifth International Natural Language Generation Conference (pp. 59-67). Association for Computational Linguistics.

Referring expression generation (REG)

- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
  - Graph structure

Search for unique subgraph

- **Search process**
  - Search for all the possible subgraphs with the target node.
- **Isomorphism process**
  - Verify the uniqueness of all subgraphs within the REG graph.

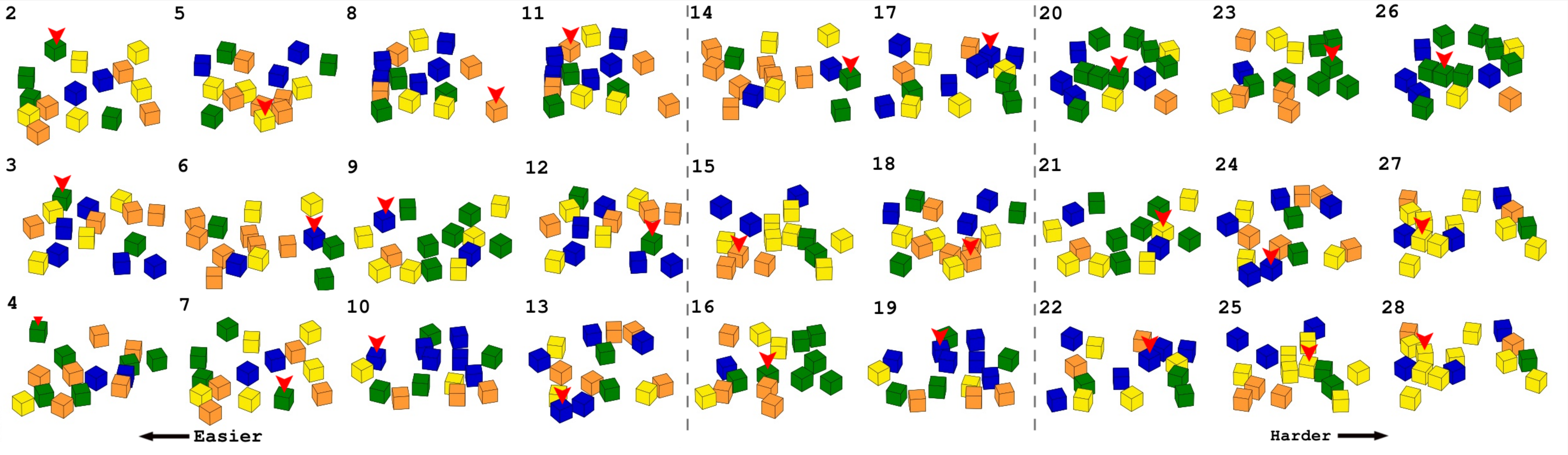
Referring expression generation (REG)

- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
    - **Pruning the search process by heuristics**
    - Speeding up the isomorphism process by heuristics
    - Commutative rule
  - Graph structure

Prune the search process

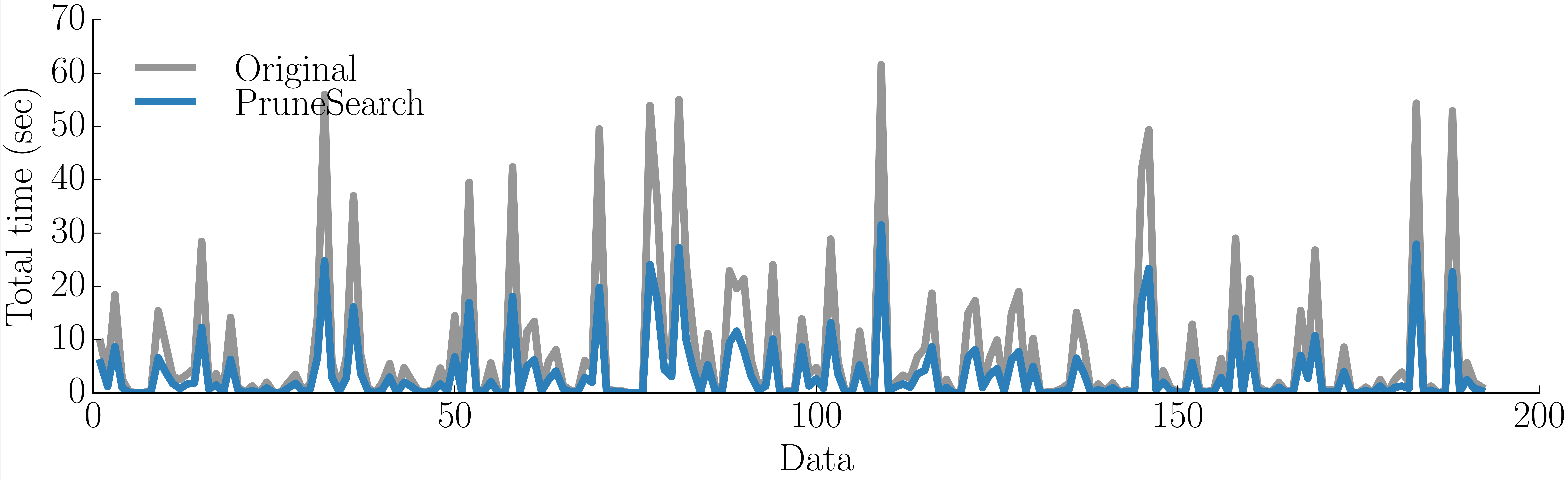
- If a search branch reaches a unique subgraph, we prune this branch.
- If a search branch reaches a subgraph with a higher cost than the cost of the current best solution, then we prune this branch.

Experiment



2
5
8
11
14
17
20
23
26
3
6
9
12
15
18
21
24
27
4
7
10
13
16
19
22
25
28
Easier
Harder

Result of pruning the search process



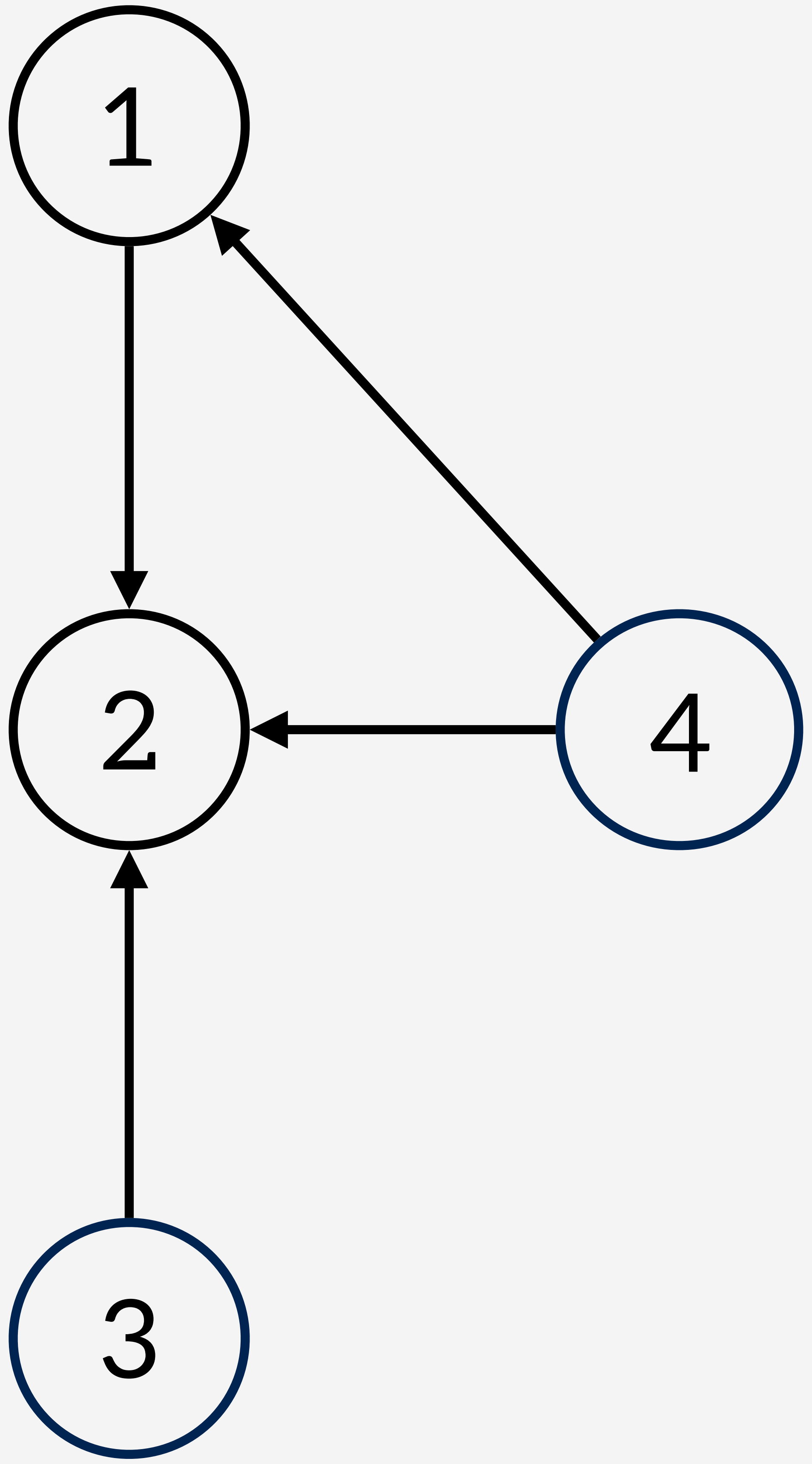
50%

Referring expression generation (REG)

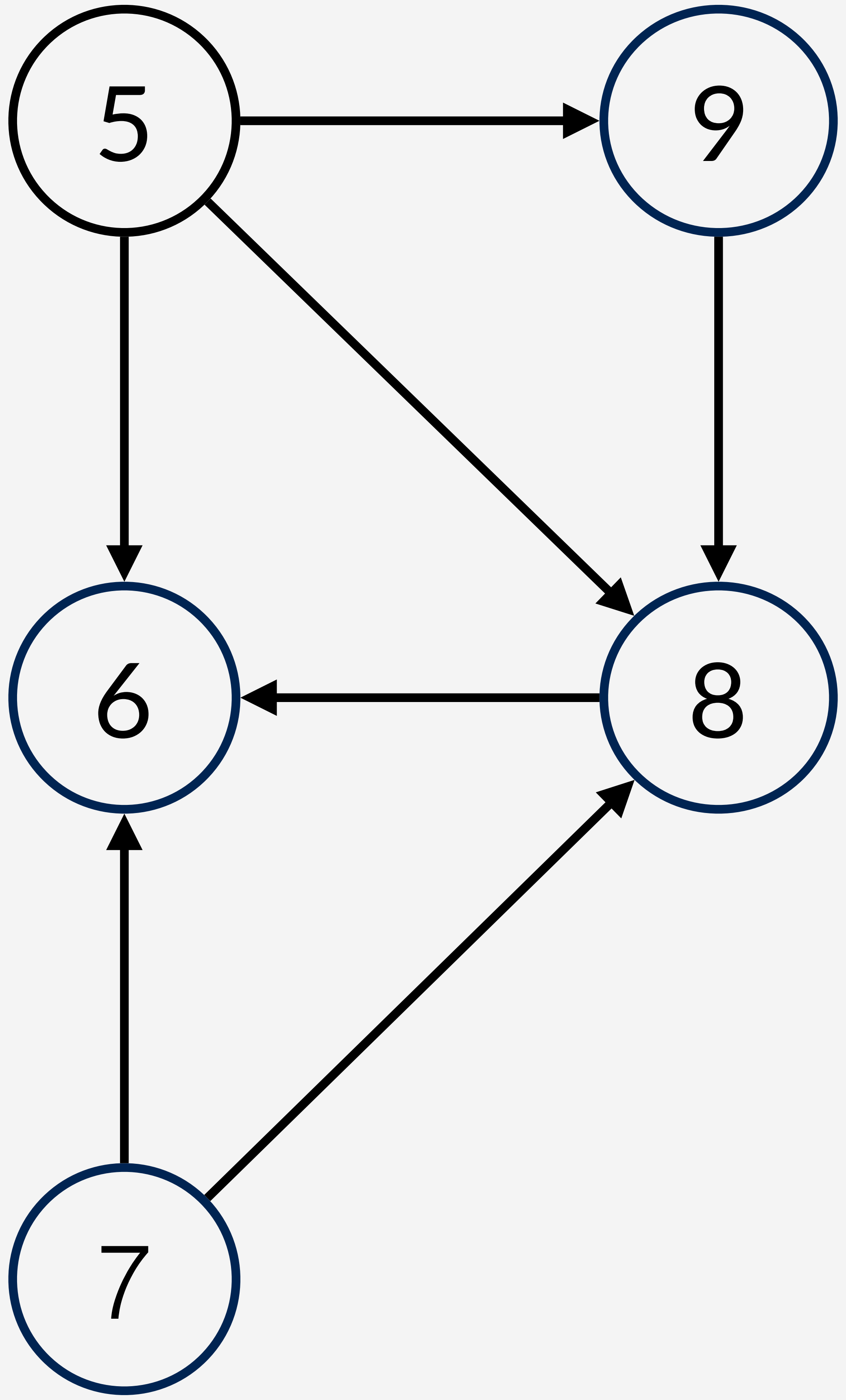
- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
    - Pruning the search process by heuristics
    - **Speeding up the isomorphism process by heuristics**
    - Commutative rule
  - Graph structure

Graph matching

$g =$



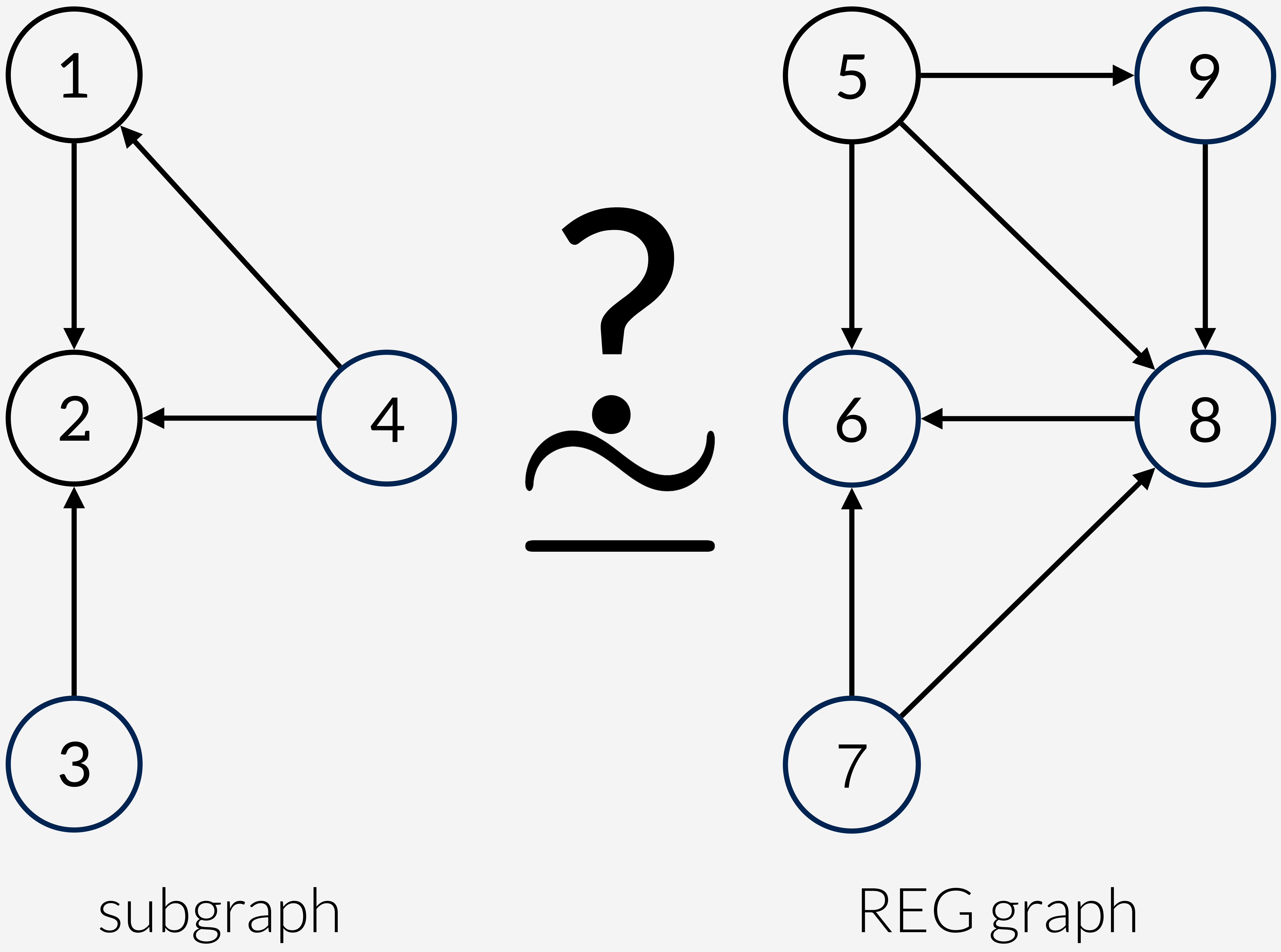
$\overset{?}{\simeq}$ $g' \in$



subgraph

REG graph

Constraint satisfaction problem



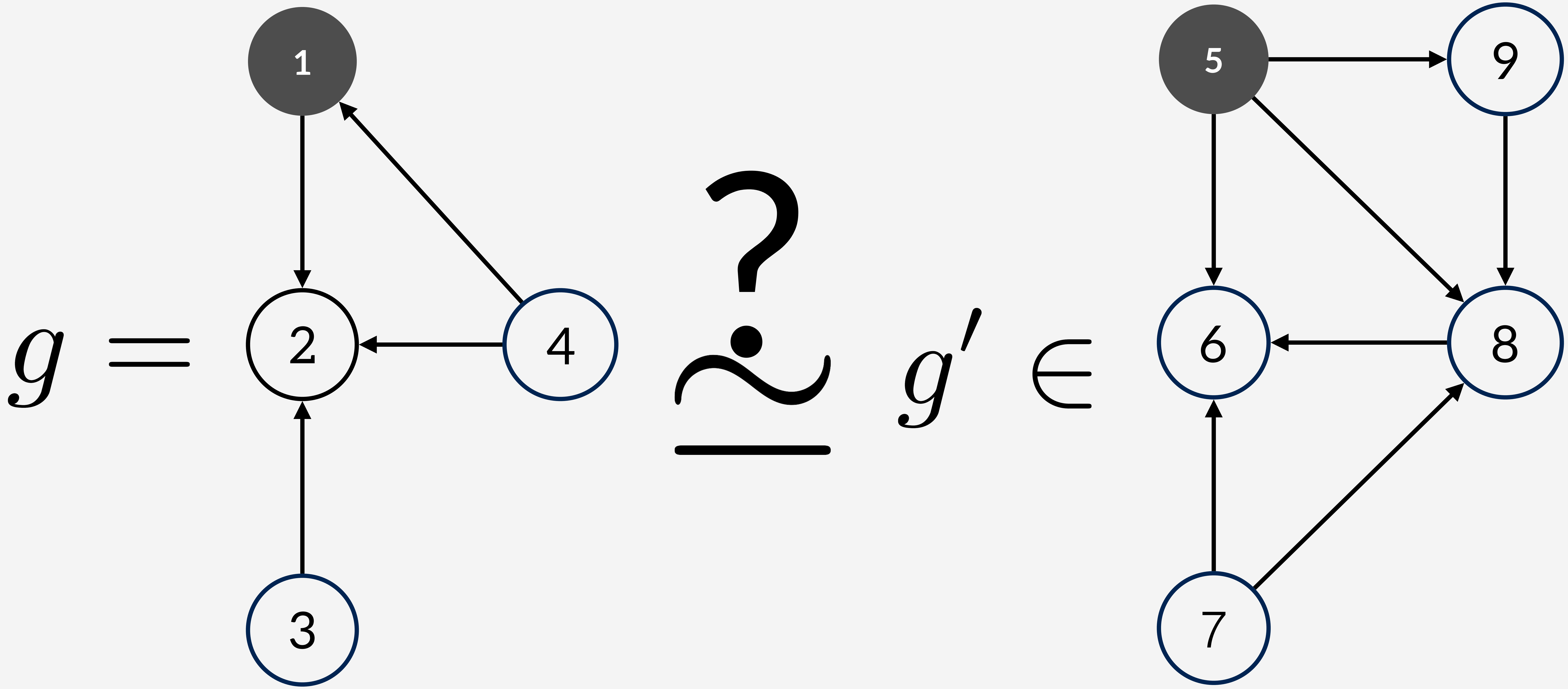
Variables = 1, 2, 3, 4

Values = a, b, c, d, e

Constraints = (1,2), (4,1), (4,2), (3,2)

Larrosa, J., & Valiente, G. (2002). Constraint satisfaction algorithms for graph pattern matching. Mathematical structures in computer science, 12(04), 403-422.

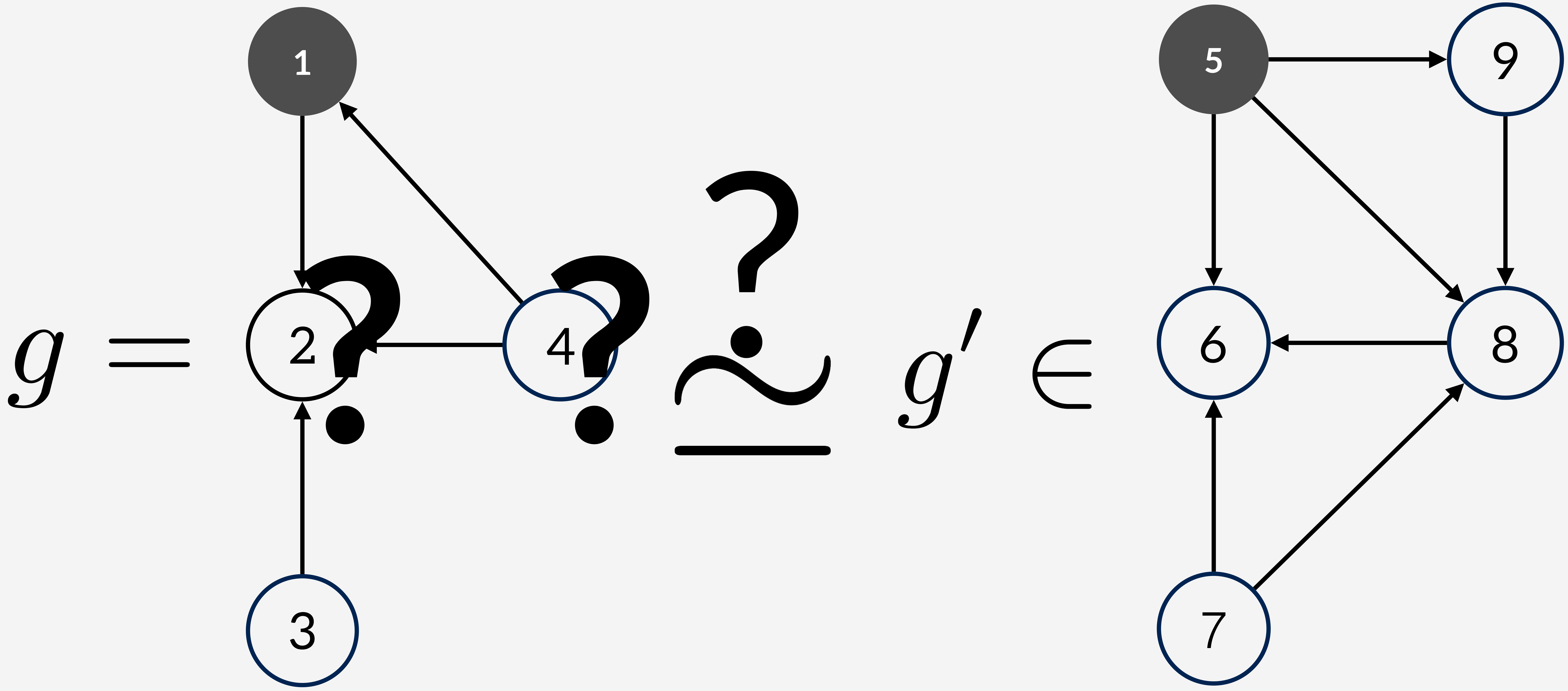
Speeding up the isomorphism process



subgraph

REG graph

Speeding up the isomorphism process



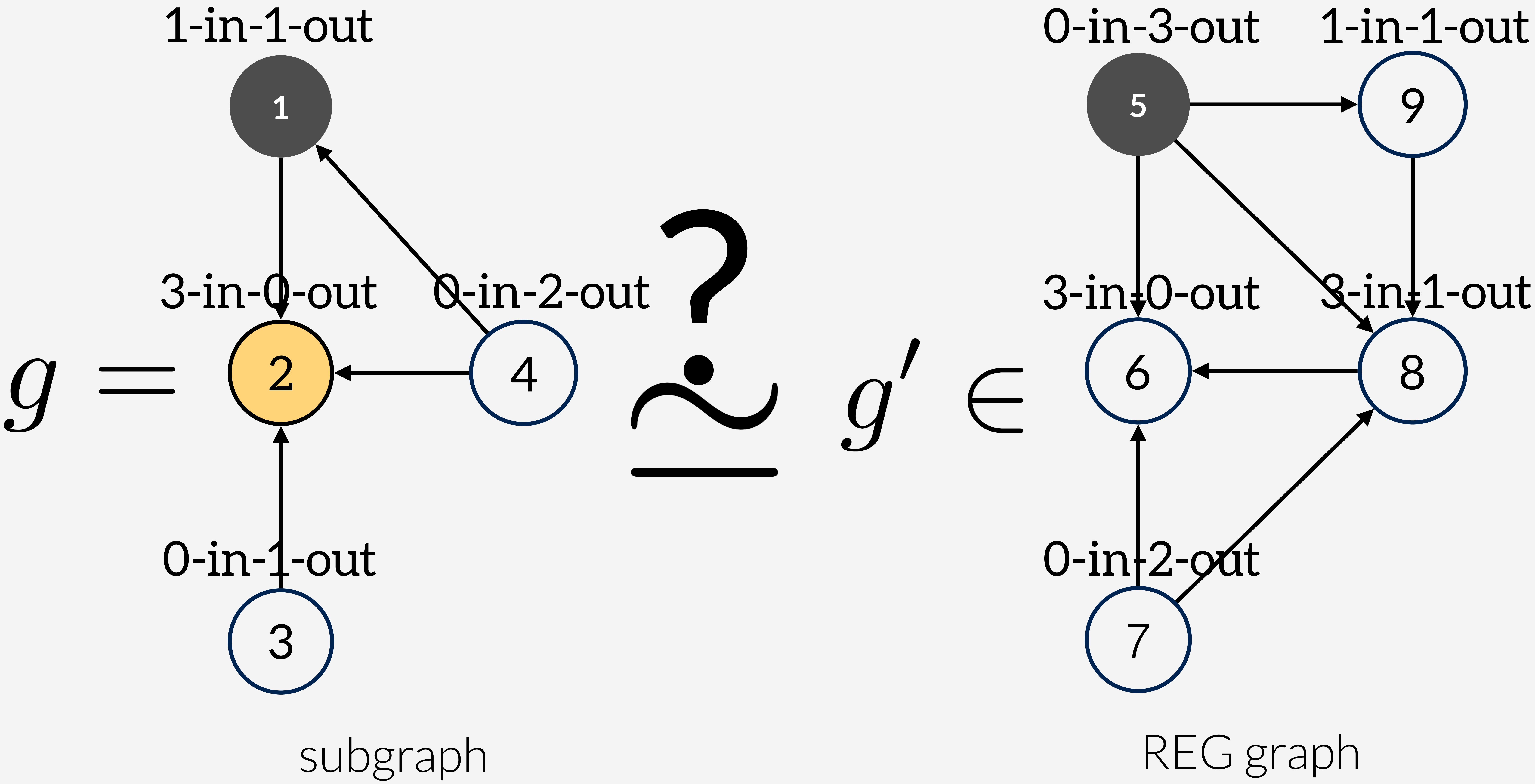
subgraph

REG graph

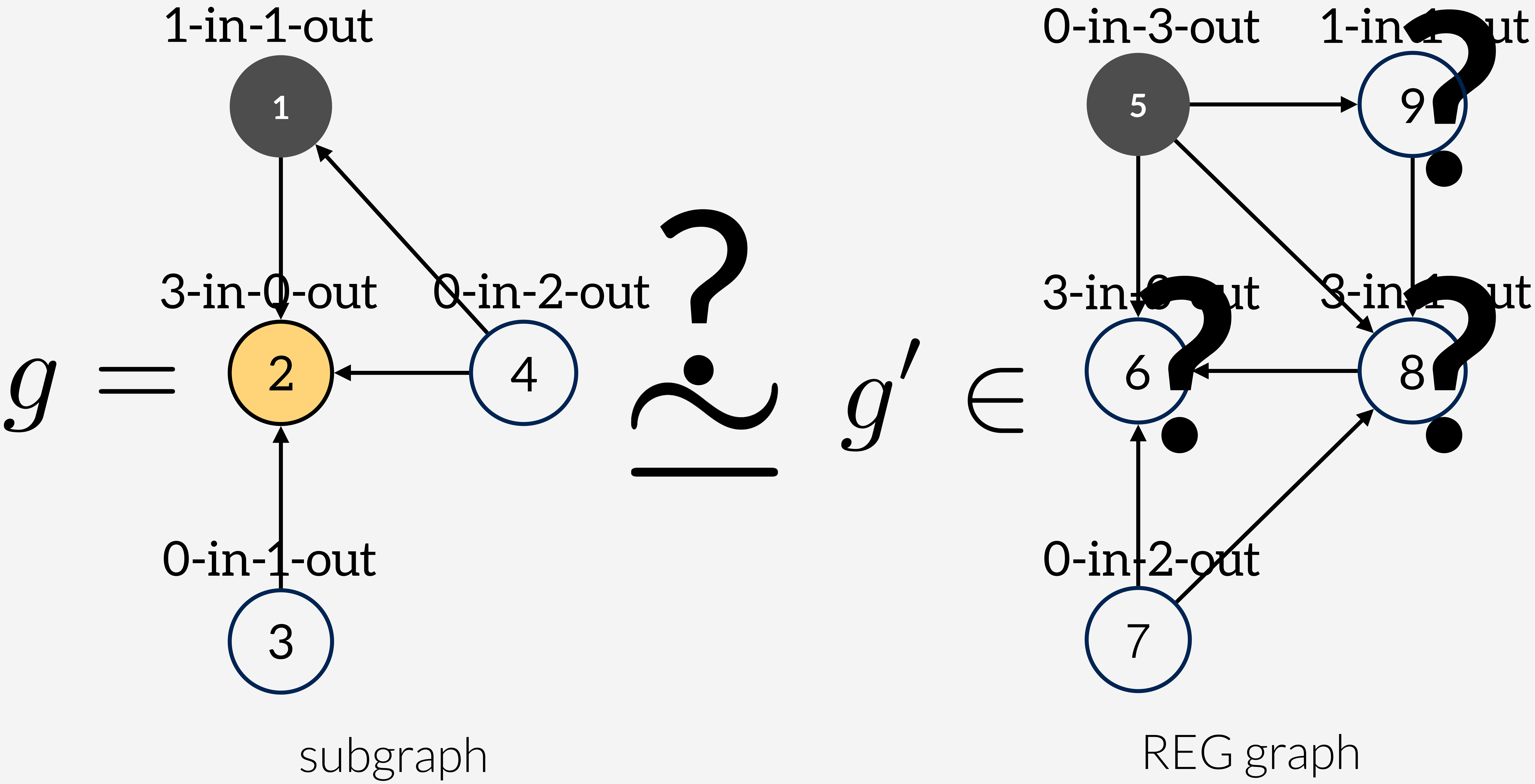
Speeding up the isomorphism process

- Pruning the search process by heuristics
- Speeding up the isomorphism process by heuristics
  - **Minimum Remaining Values (MRV) heuristic**
    - The algorithm would choose the most constrained **variable** with the fewest legal possible values as the next variable to try.

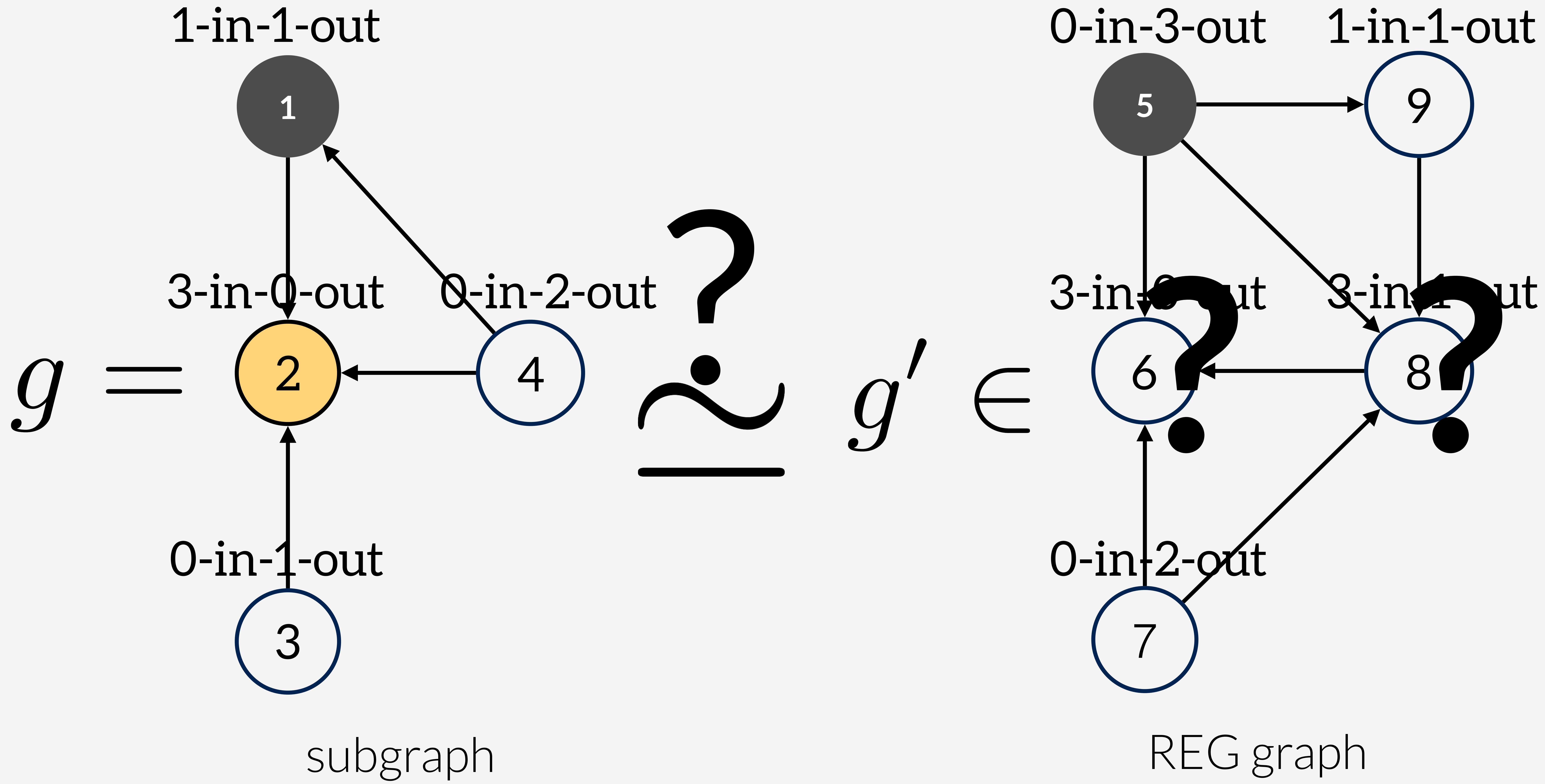
Minimum Remaining Values



Speeding up the isomorphism process



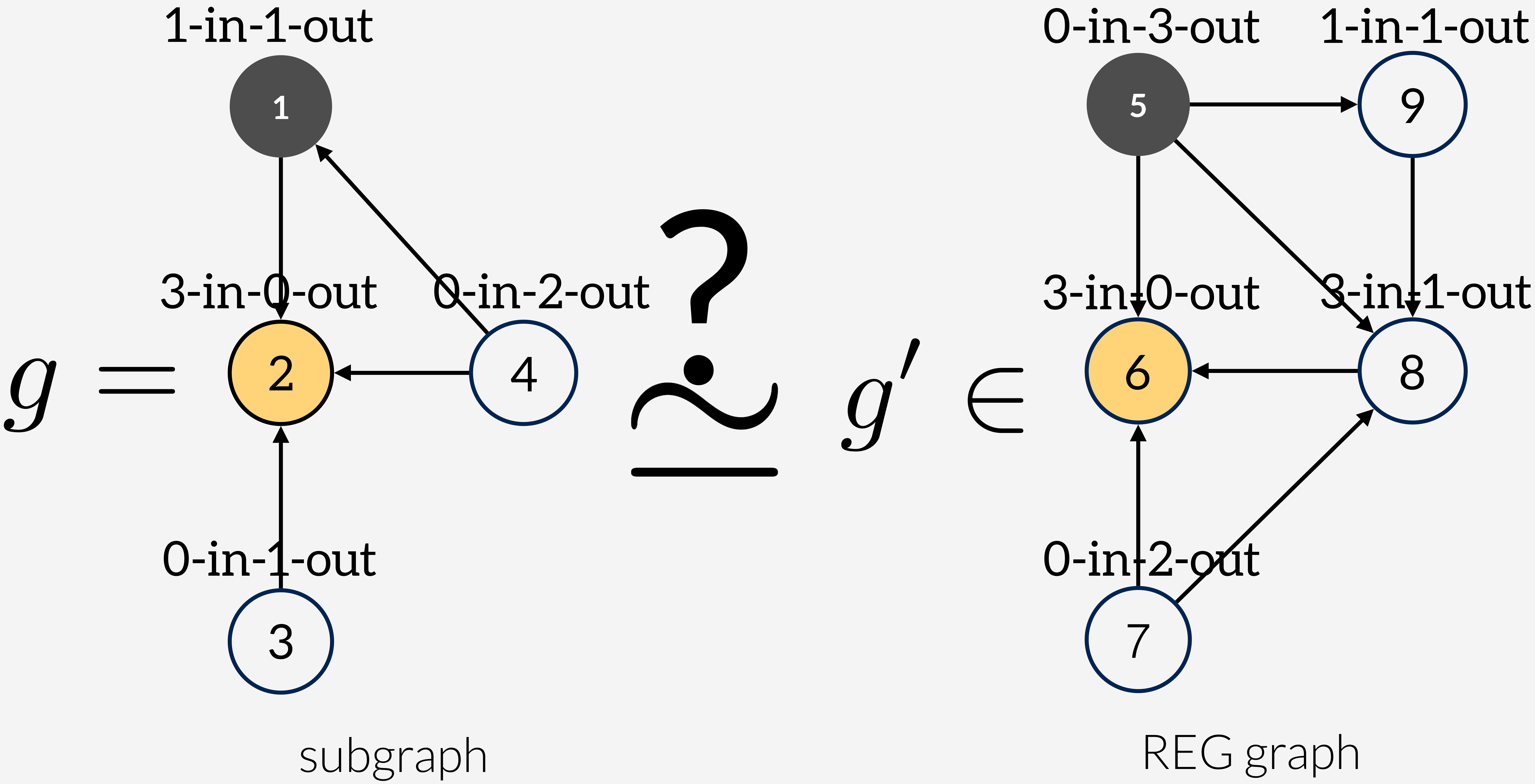
Speeding up the isomorphism process



Speeding up the isomorphism process

- Pruning the search process by heuristics
- Speeding up the isomorphism process by heuristics
  - Minimum Remaining Values (MRV) heuristic
  - **Least Constraining Value (LCV) heuristic**
    - The algorithm would choose the **value** that leaves the most choices or flexibility for the future unassigned variables.

Least Constraining Value

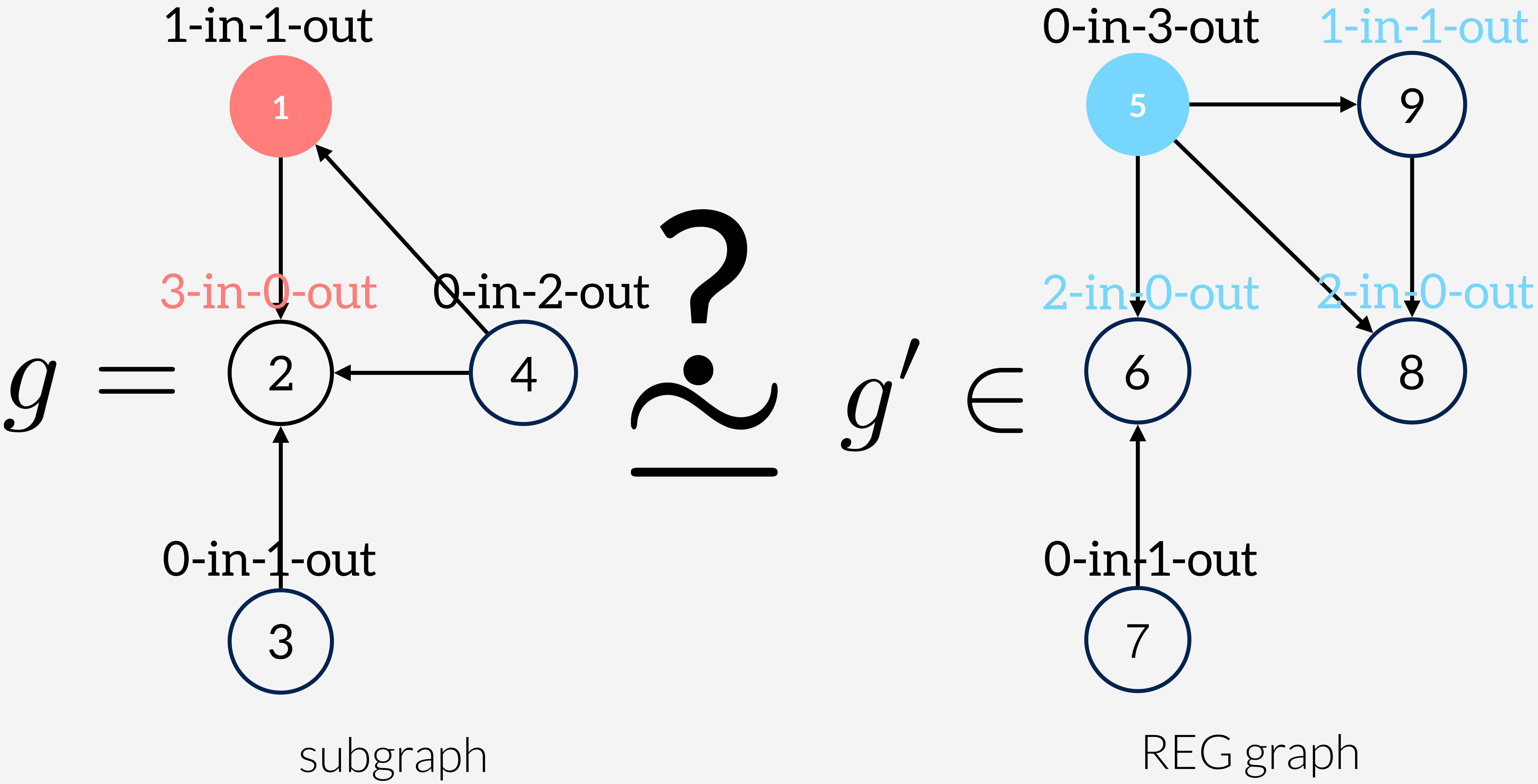


Speeding up the isomorphism process

- Pruning the search process by heuristics
- Pruning the isomorphism process by heuristics
  - Minimum Remaining Values (MRV) heuristic
  - Least Constraining Value (LCV) heuristic
  - **Look-ahead heuristics for to-be-matched nodes**
    - You can only match a *less* constrained node to a *more* constrained node, not the other way around.

Cordella, L. P., Foggia, P., Sansone, C., & Vento, M. (1998). Subgraph transformations for the inexact matching of attributed relational graphs. In Graph based representations in pattern recognition (pp. 43-52). Springer Vienna.

Look-ahead heuristics

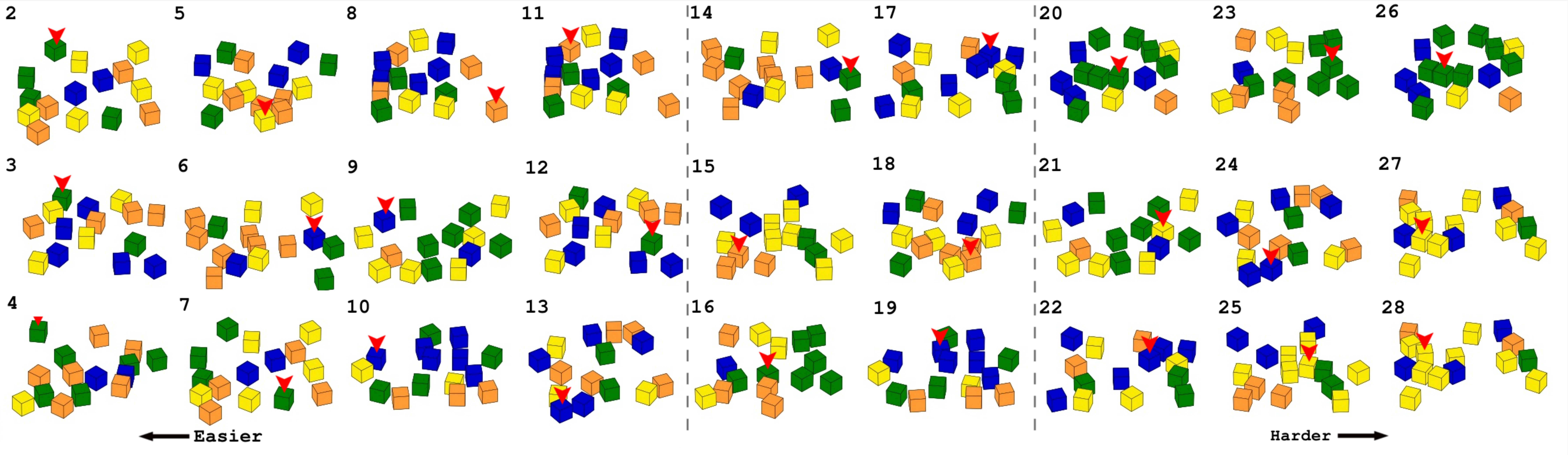


Speeding up the isomorphism process

- Pruning the search process by heuristics
- Pruning the isomorphism process by heuristics
  - Minimum Remaining Values (MRV) heuristic
  - Least Constraining Value (LCV) heuristic
  - **Look-ahead heuristics for to-be-matched nodes**
    - You can only match a *less* constrained node to a *more* constrained node, not the other way around.
    - Check the degree of predecessors and successors

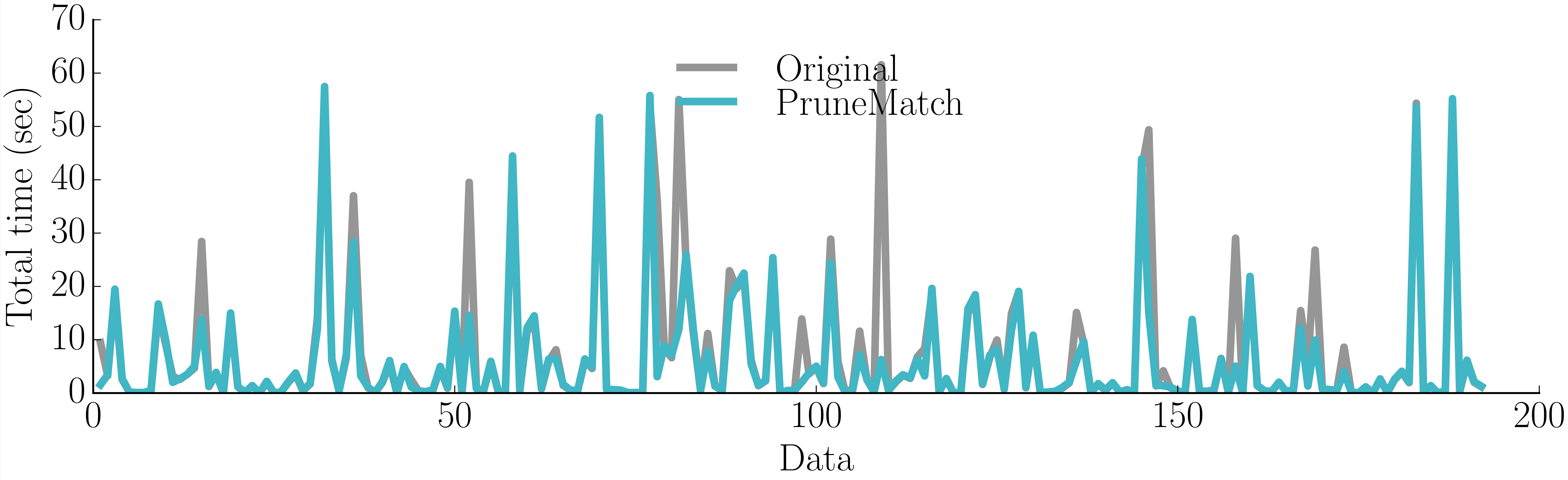
Cordella, L. P., Foggia, P., Sansone, C., & Vento, M. (1998). Subgraph transformations for the inexact matching of attributed relational graphs. In Graph based representations in pattern recognition (pp. 43-52). Springer Vienna.

Experiment



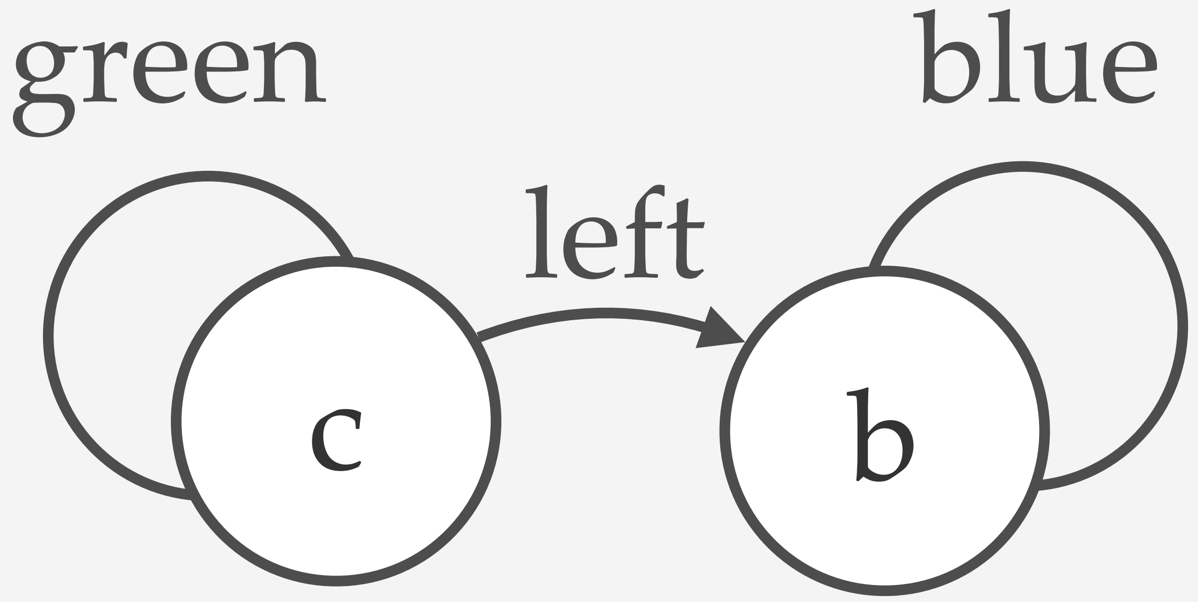
2
5
8
11
14
17
20
23
26
3
6
9
12
15
18
21
24
27
4
7
10
13
16
19
22
25
28
Easier
Harder

Result of pruning graph matching



16%

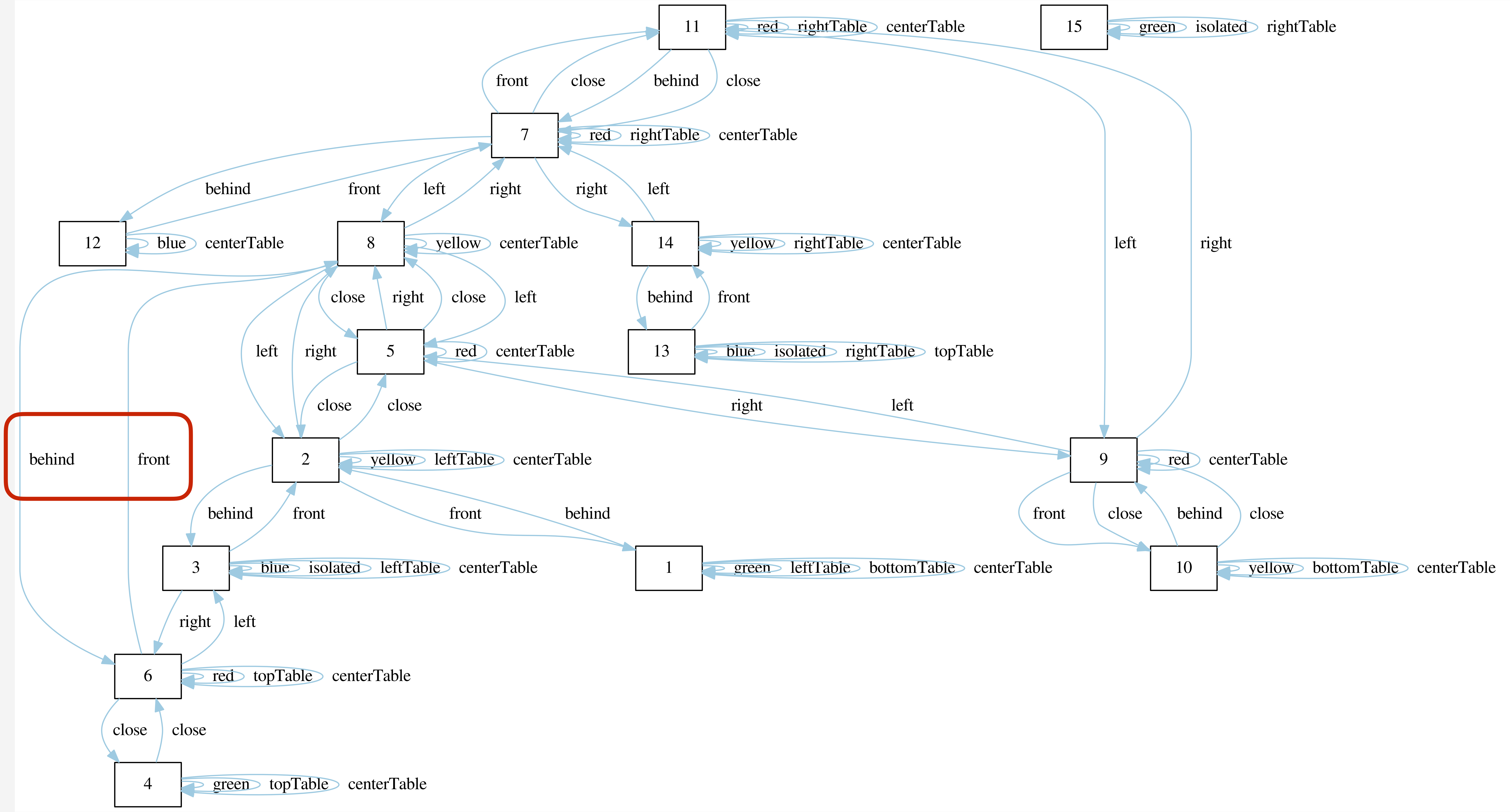
Discussion



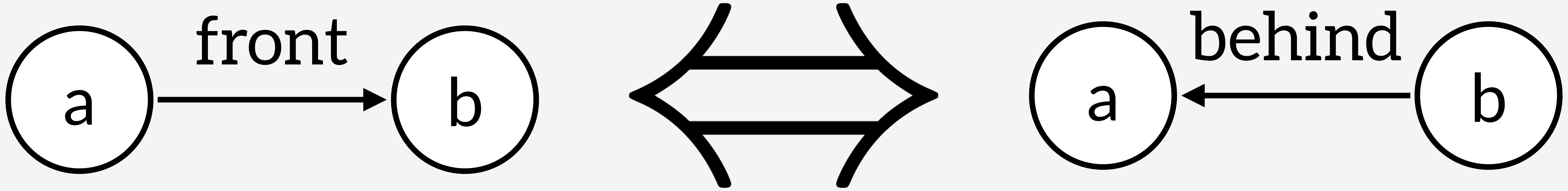
Referring expression generation (REG)

- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
    - Pruning the search process by heuristics
    - Speeding up the isomorphism process by heuristics
    - **Commutative rule**
  - Graph structure

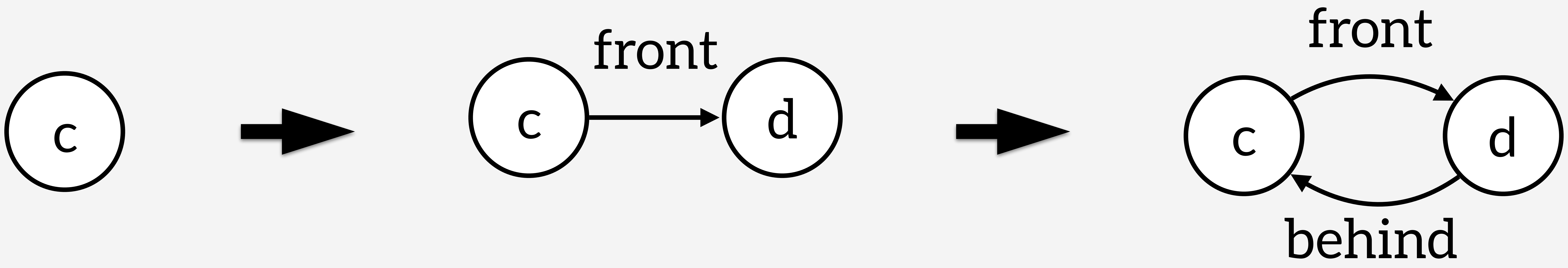
Commutative rule



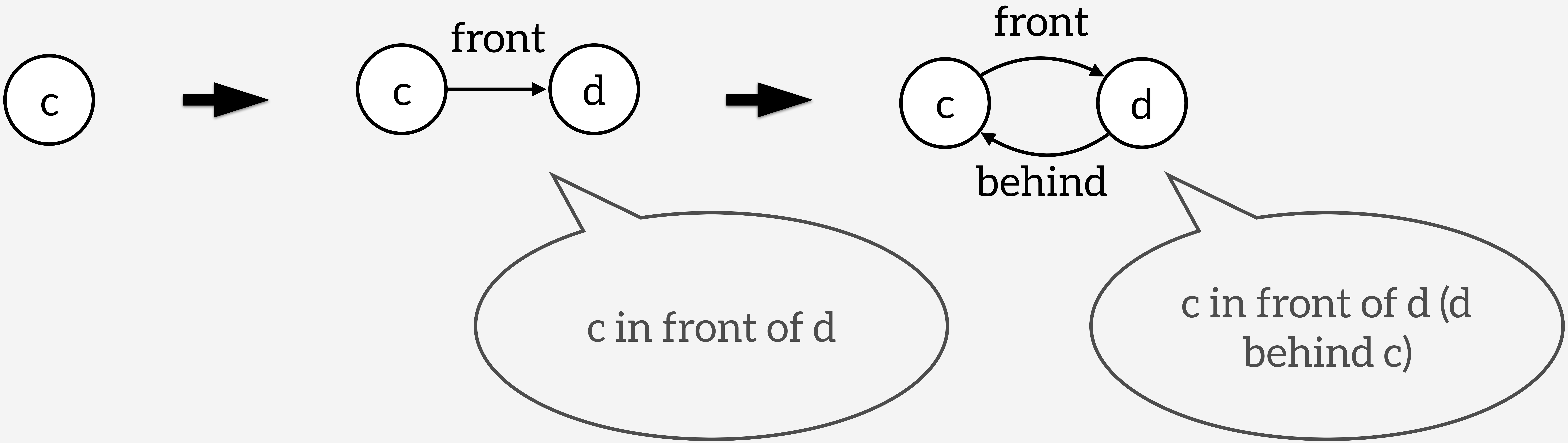
Commutative rule



Redundancy

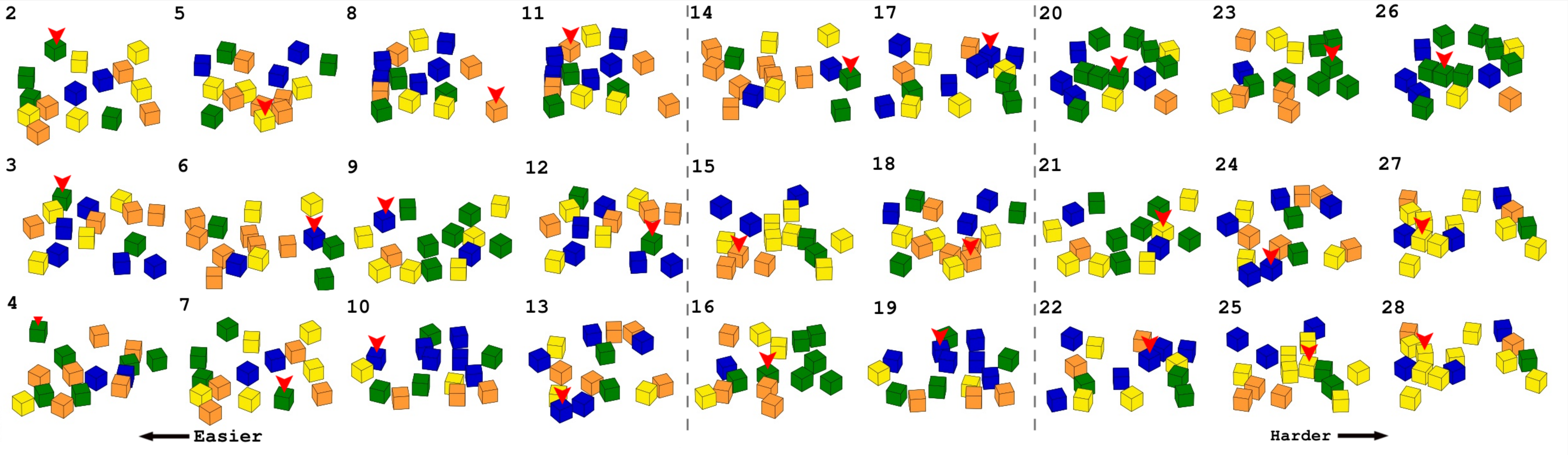


Redundancy


c
front
c
d
front
c
d
behind
c in front of d
c in front of d (d behind c)

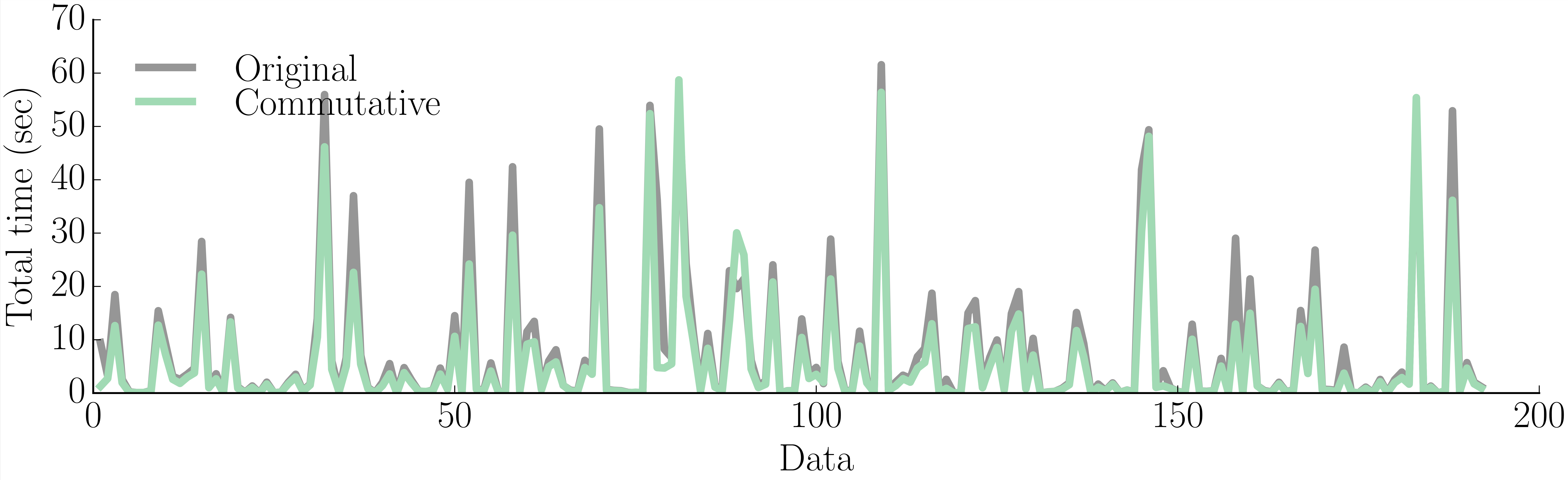
11 red rightTable centerTable
15 green isolated rightTable
7 red rightTable centerTable
12 blue centerTable
8 yellow centerTable
14 yellow rightTable centerTable
13 blue isolated rightTable topTable
5 red centerTable
2 yellow leftTable centerTable
9 red centerTable
3 blue isolated leftTable centerTable
1 green leftTable bottomTable centerTable
10 yellow bottomTable centerTable
6 red topTable centerTable
4 green topTable centerTable
behind front
13 blue isolated rightTable topTable
15 green isolated rightTable
11 red rightTable centerTable
14 yellow rightTable centerTable
12 blue centerTable
4 green topTable centerTable
9 red centerTable
7 red rightTable centerTable
6 red topTable centerTable
front
10 yellow bottomTable centerTable
8 yellow centerTable
3 blue isolated leftTable centerTable
2 yellow leftTable centerTable
5 red centerTable
1 green leftTable bottomTable centerTable

Experiment



2
5
8
11
14
17
20
23
26
3
6
9
12
15
18
21
24
27
4
7
10
13
16
19
22
25
28
Easier
Harder

Result of simplifying graph

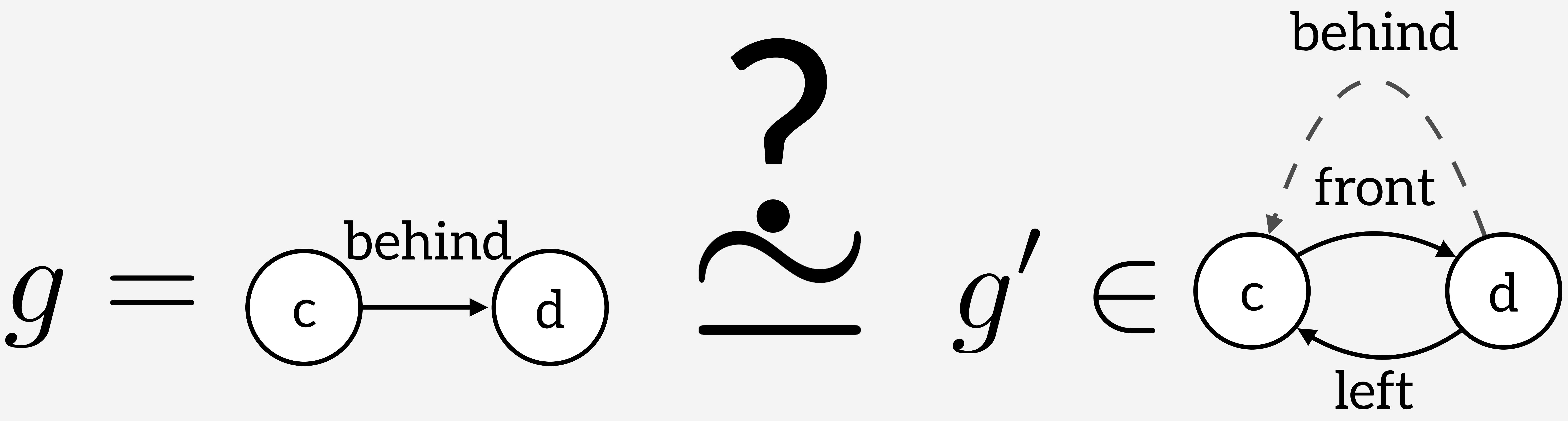


21%

Discussion

Subgraph **REG graph**

Discussion



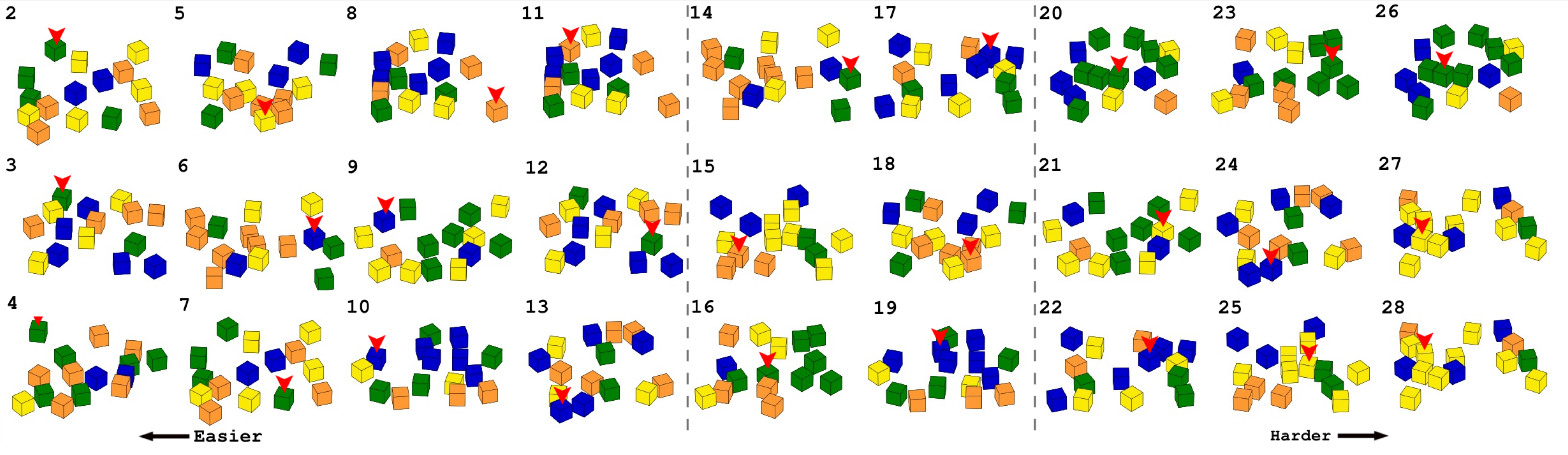
Subgraph

REG graph

Referring expression generation (REG)

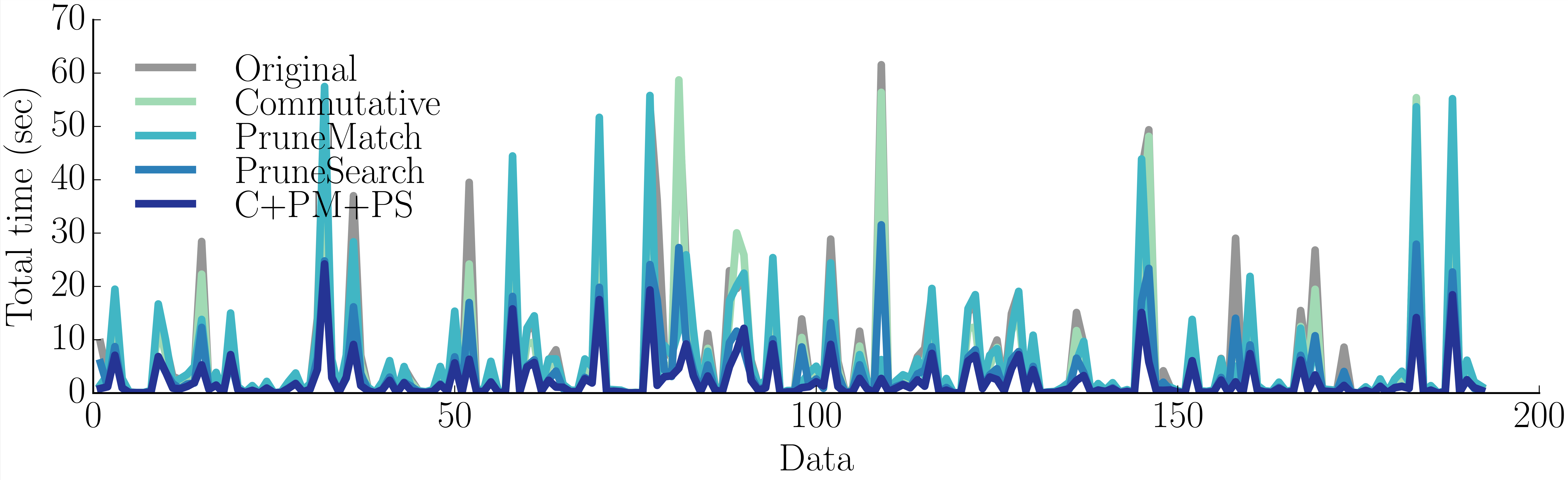
- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
    - **Pruning the search process by heuristics**
    - **Speeding up the isomorphism process by heuristics**
    - **Commutative rule**
  - Graph structure

Experiment

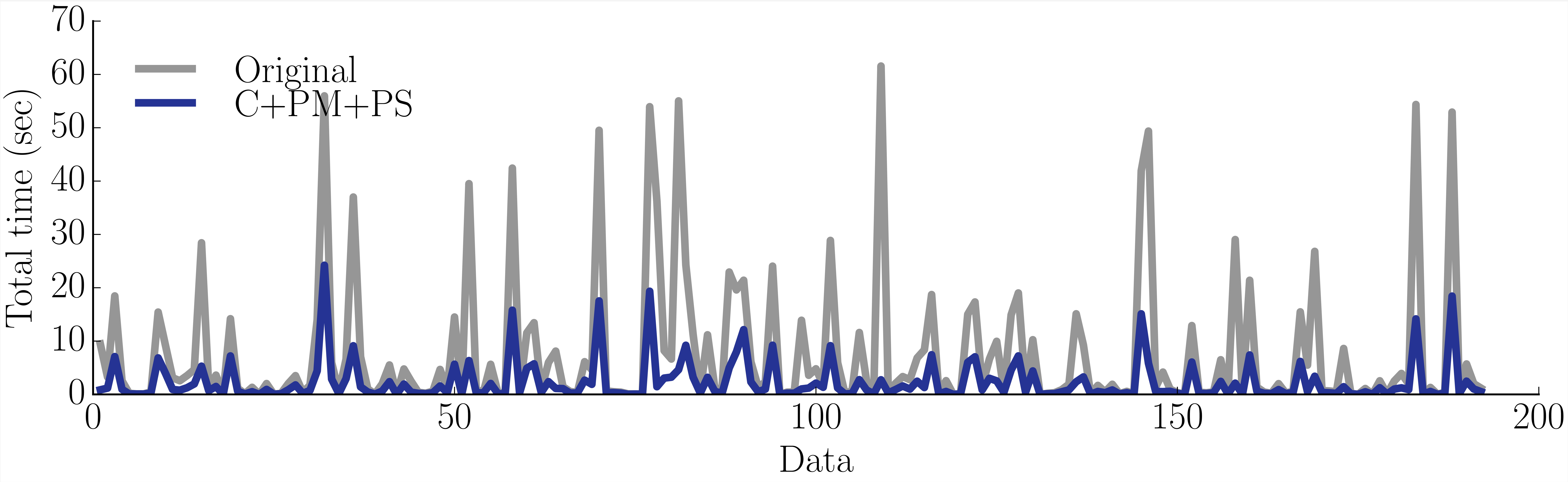


2
5
8
11
14
17
20
23
26
3
6
9
12
15
18
21
24
27
4
7
10
13
16
19
22
25
28
Easier
Harder

Result of all three techniques



Result of all three techniques



56%

Referring expression generation (REG)

- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
  - **Graph structure to support higher level features (on-going)**

Constraint-based spatial reasoning

- Constraints are widely used in modeling semantic spatial information by encoding knowledge about objects or relations between objects.
- Constraints are determined by features.

Jochen Renz and Bernhard Nebel. Qualitative spatial reasoning us- ing constraint calculi. In Handbook of spatial logics, pages 161–215. Springer, 2007.

Spatial constraints for REG

Unary absolute qualitative constraint

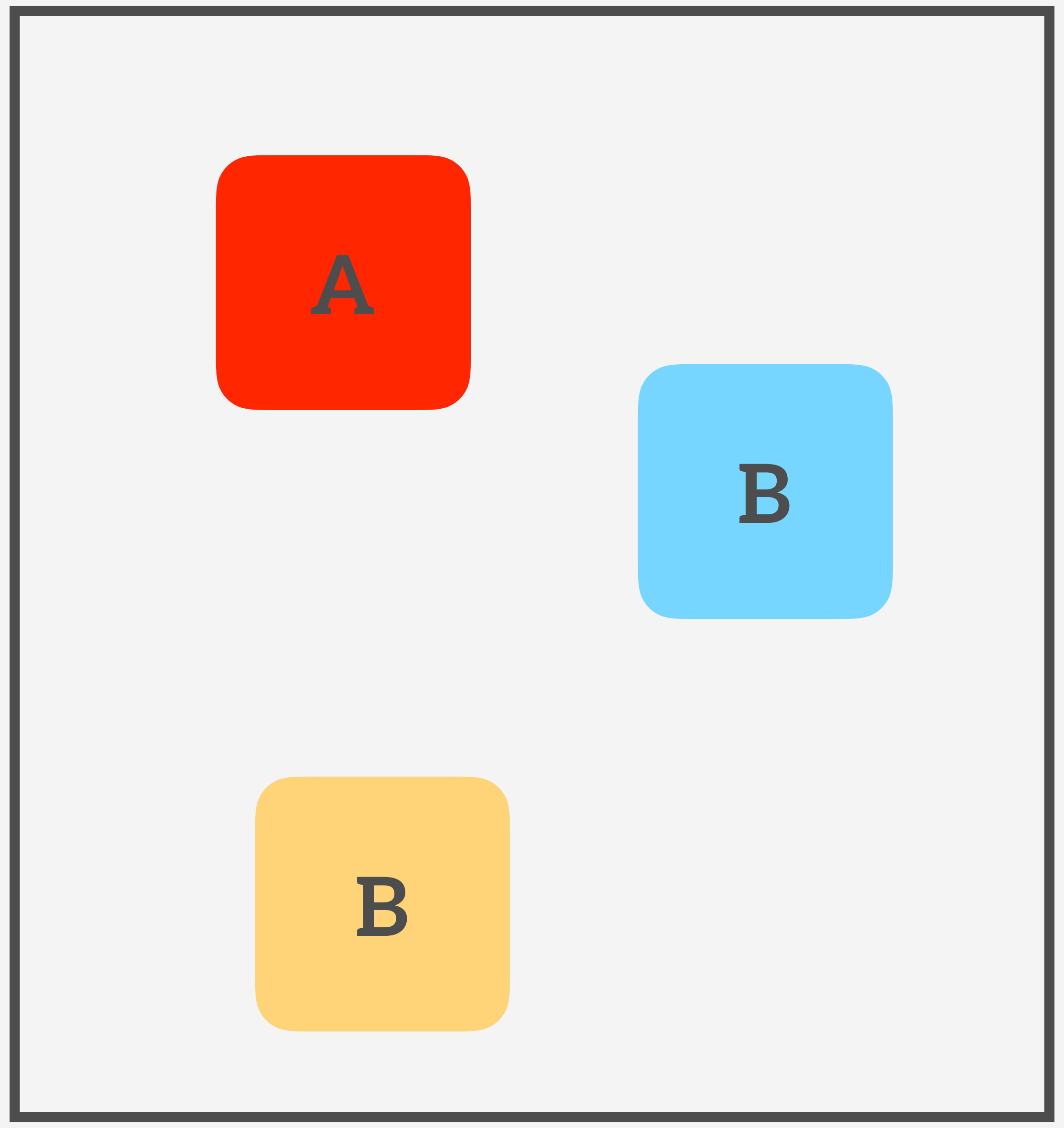
- Color, e.g. “A is red”

Binary relative qualitative constraint

- Distance, e..g “A is close to B”
- Orientation, e.g. “A is on the left to B”

N-ary relative qualitative constraint

- Shape, e.g. “A,B,C form a triangle”



Spatial constraints for language

Unary absolute qualitative constraint

- Color, e.g. “A is red”

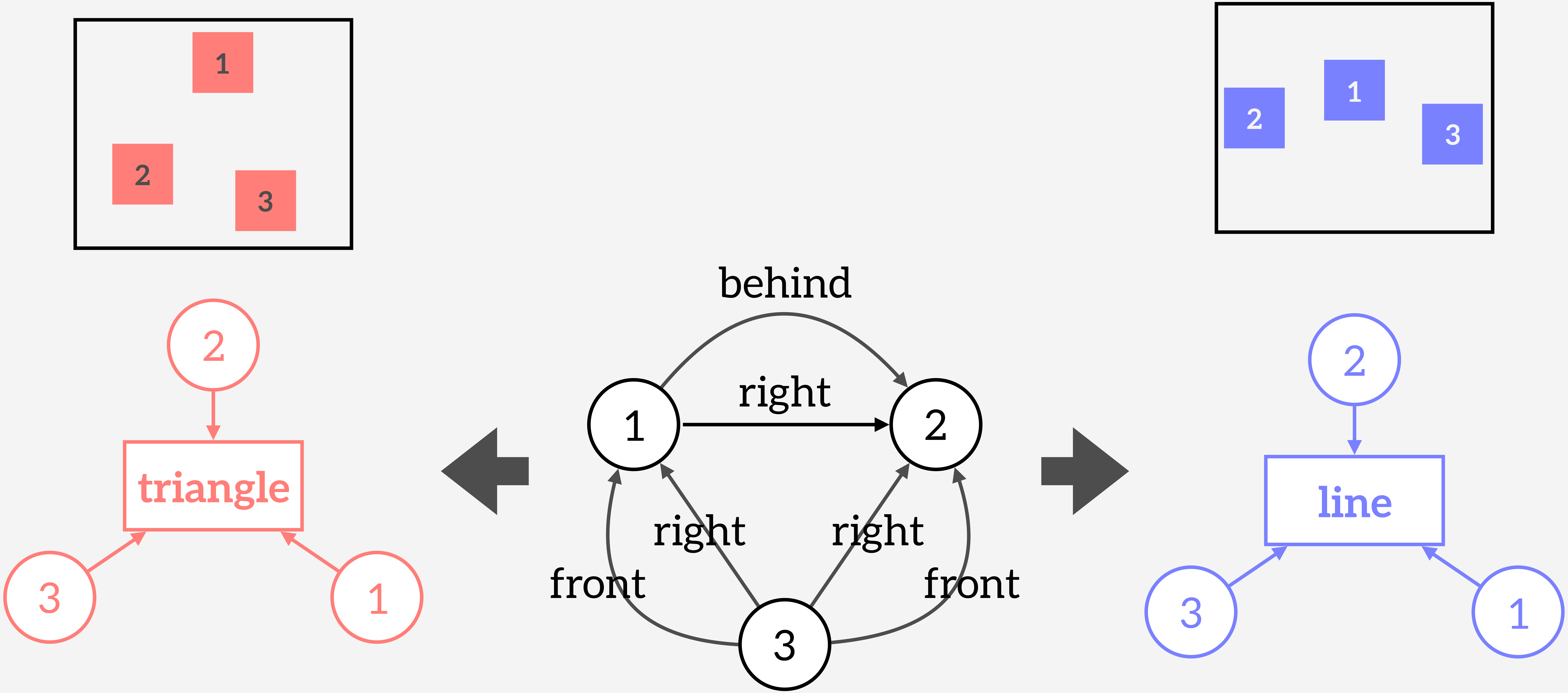
Binary relative qualitative constraint

- Distance, e..g “A is close to B”
- Orientation, e.g. “A is on the left to B”

N-ary relative qualitative constraint

- Shape, e.g. “A,B,C form a triangle”

Spatial constraints for language



Hierarchical graph structure for REG

Unary absolute qualitative constraint

- Color, e.g. “A is red”

Binary relative qualitative constraint

- Distance, e..g “A is close to B”
- Orientation, e.g. “A is on the left to B”

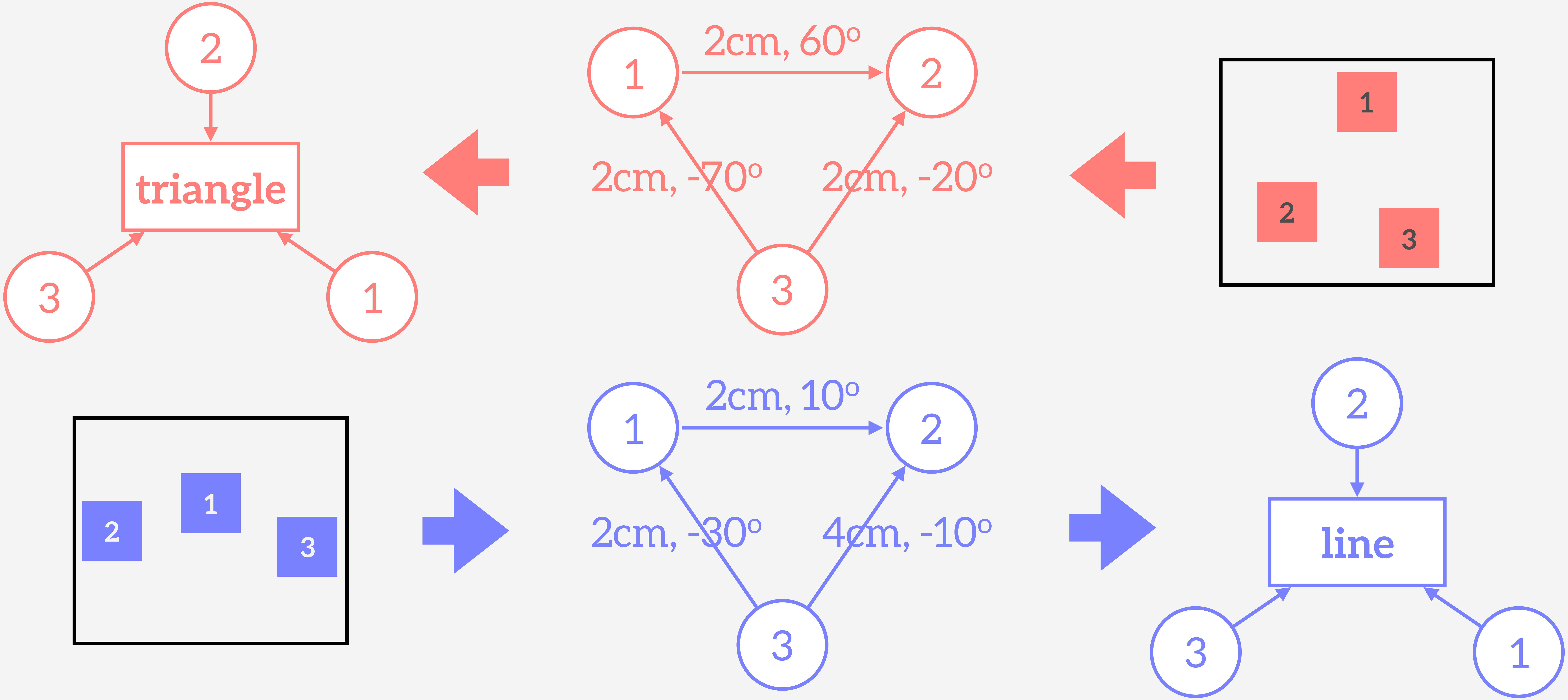
N-ary relative qualitative constraint

- Shape, e.g. “A,B,C form a triangle”

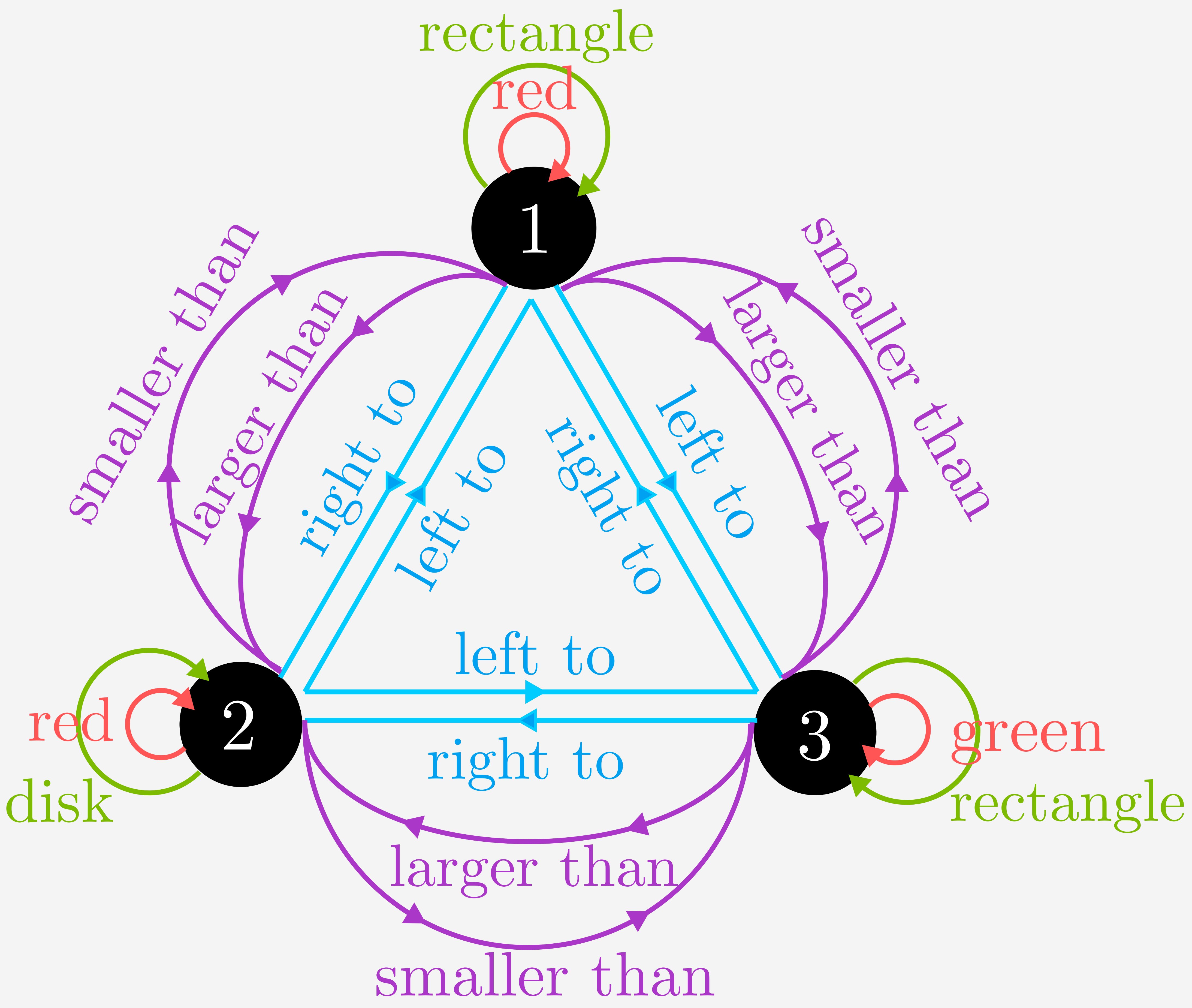
Binary relative quantitative constraint

- Distance, e..g “A is 5 cm far from B”
- Orientation, e.g. “A is 60° north of east in the view of B”

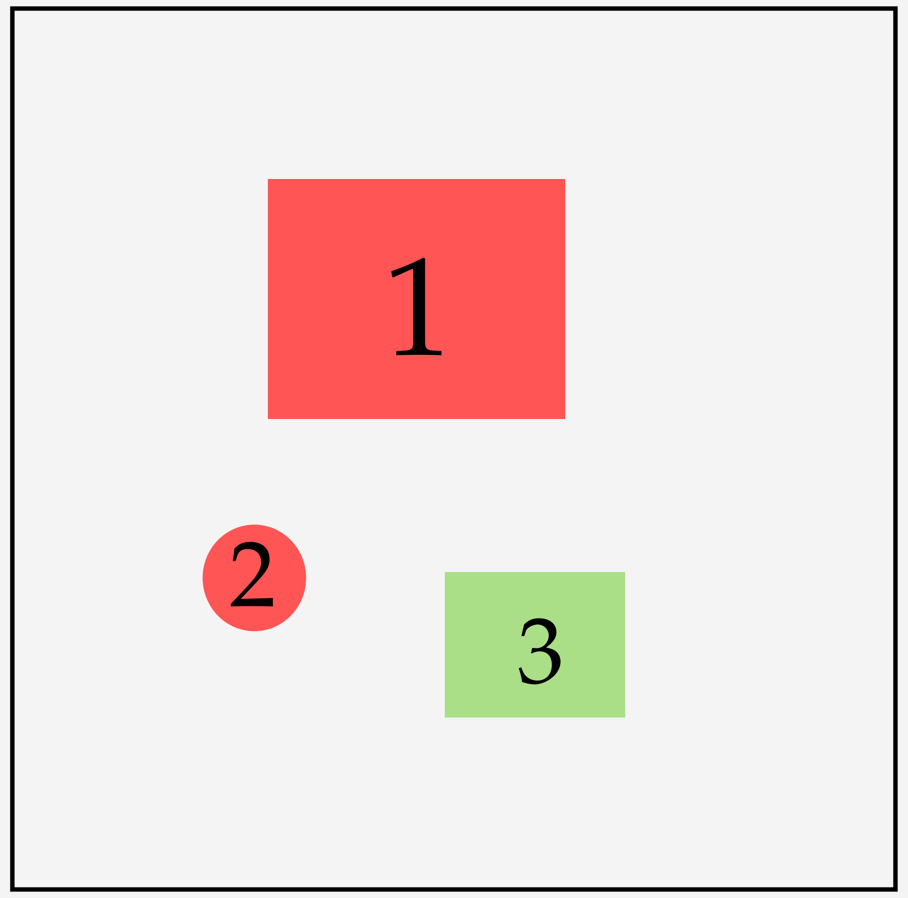
Spatial constraints for language



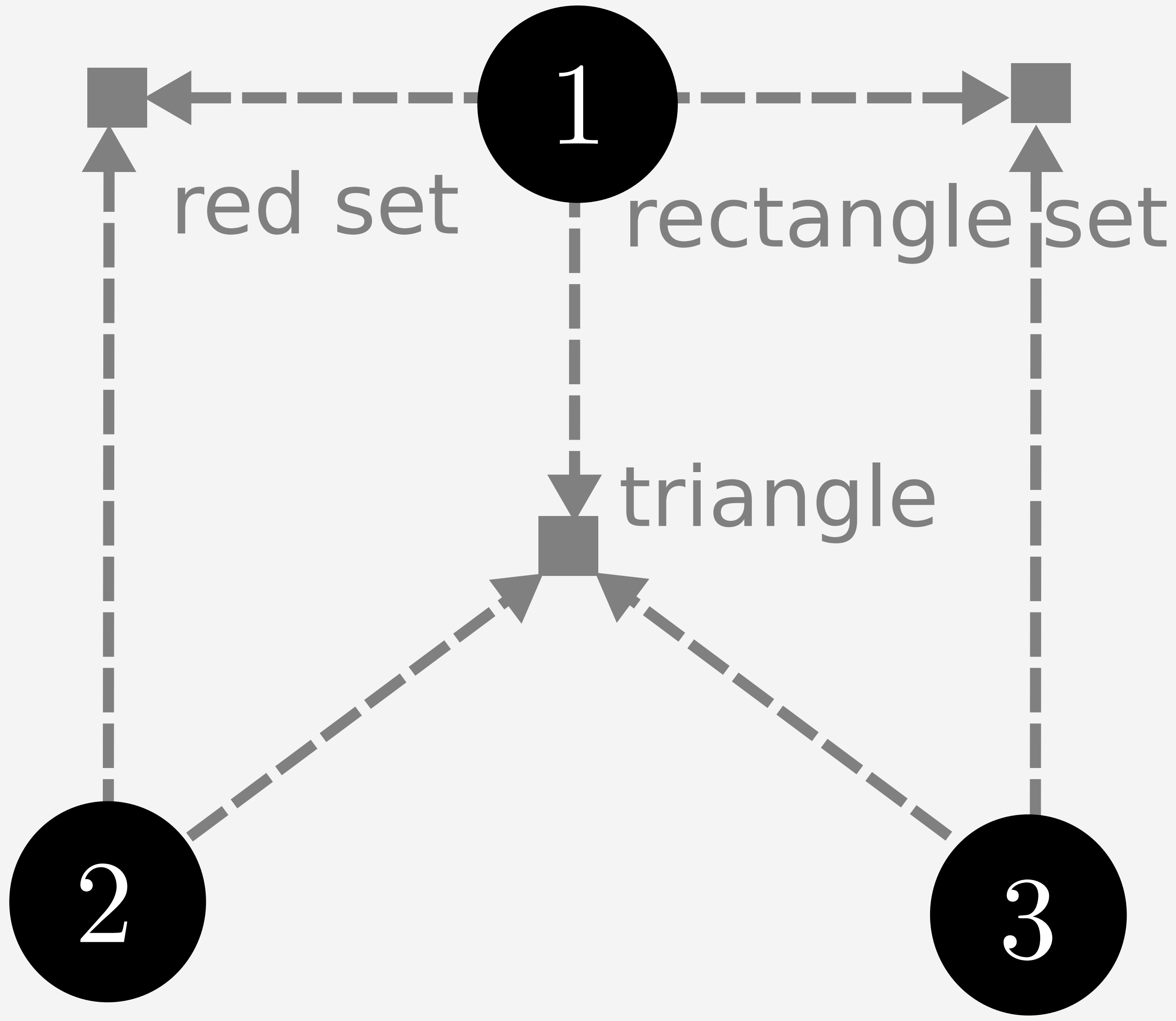
Hierarchical graph structure for REG



Qualitative Layer

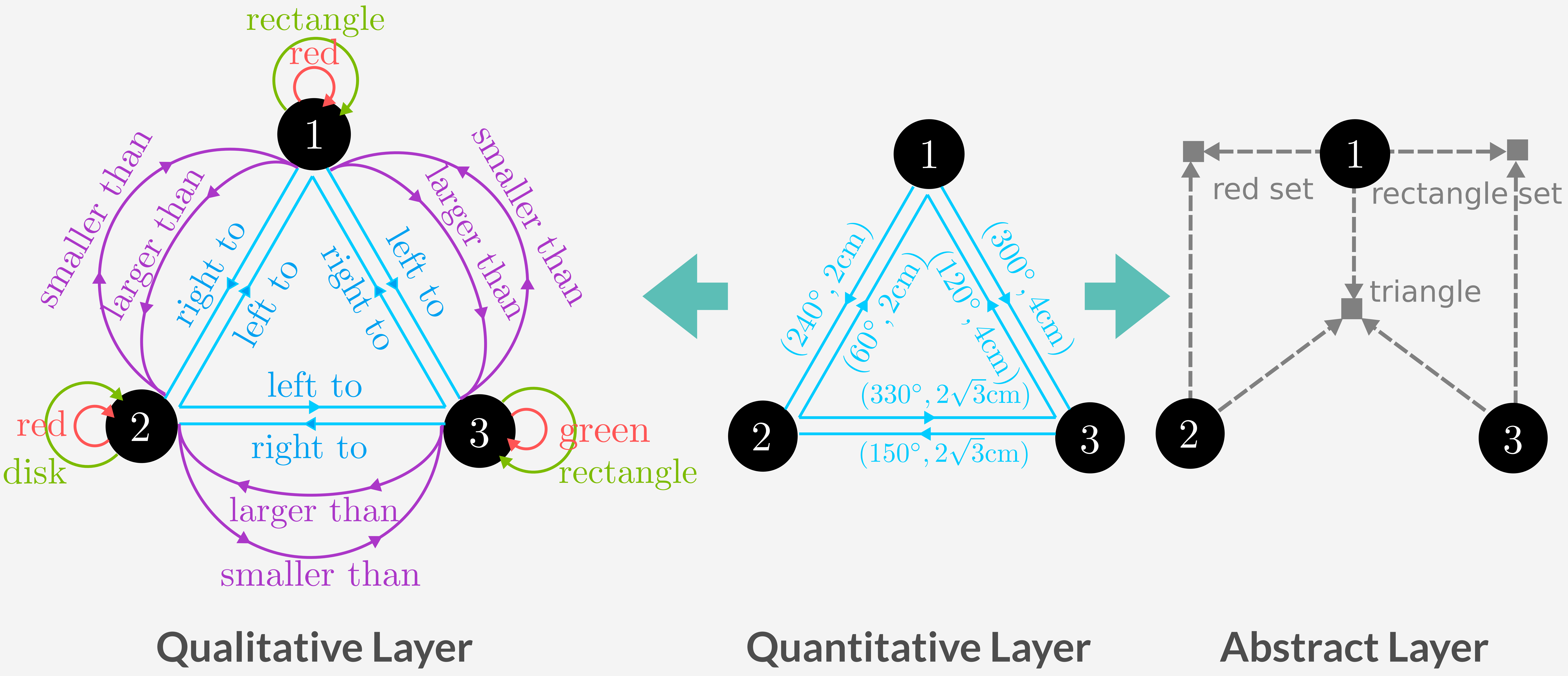


Scene



Abstract Layer

Hierarchical graph structure for REG



Referring expression generation (REG)

- Previous work on REG
- Our contribution on REG
  - Corpus
  - Algorithm efficiency
  - Graph structure to support higher level features

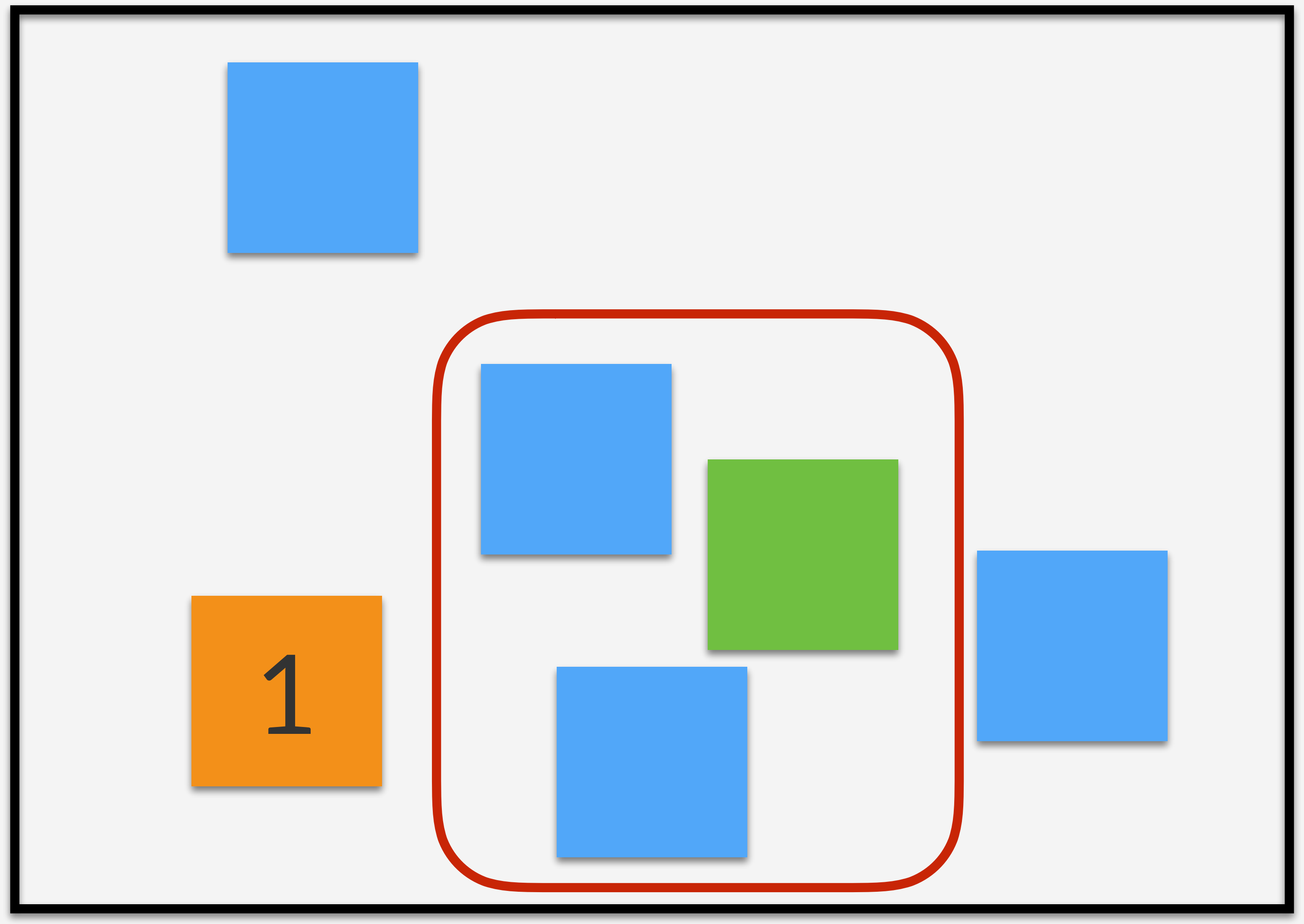
Future work

Future work

- Compare the human preference over visual and spatial features

Future work

- Compare the human preference over visual and spatial features
- Feature definition



Future work

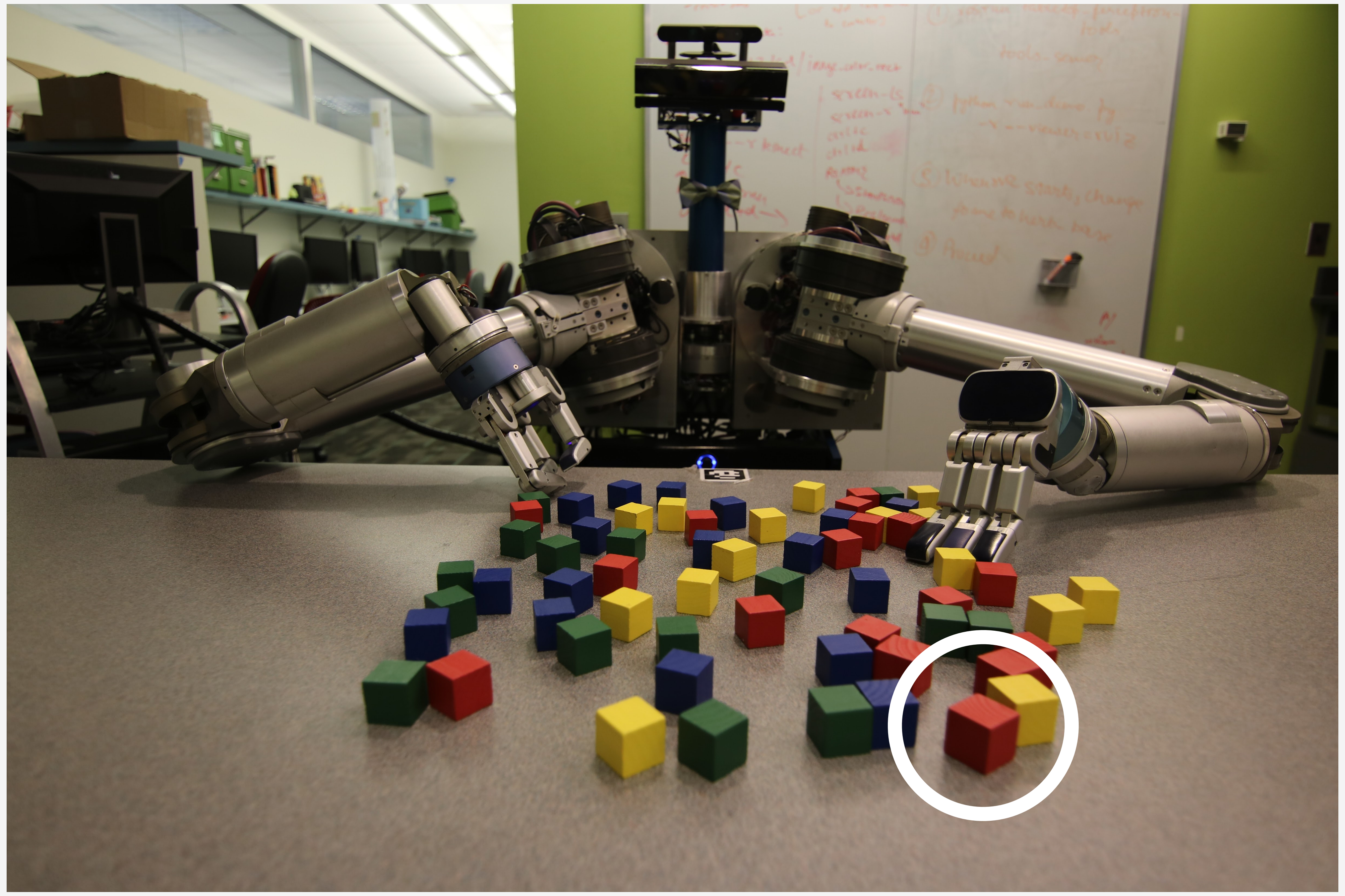
- Compare the human preference over visual and spatial features
- Feature definition
- Feature preference

Conclusion

Important to understand robots

Pellegrinelli, S., Admoni, H., Javdani, S., & Srinivasa, S. Human-Robot SharedWorkspace Collaboration via Hindsight Optimization. IROS. 2016.
Adrian Bussone, Simone Stumpf, and Dympna O'Sullivan. The role of explanations on trust and reliance in clinical decision support sys- tems. ICHI. 2015.

Language-based explanation for intentions



"I am picking up the red block closest to you."

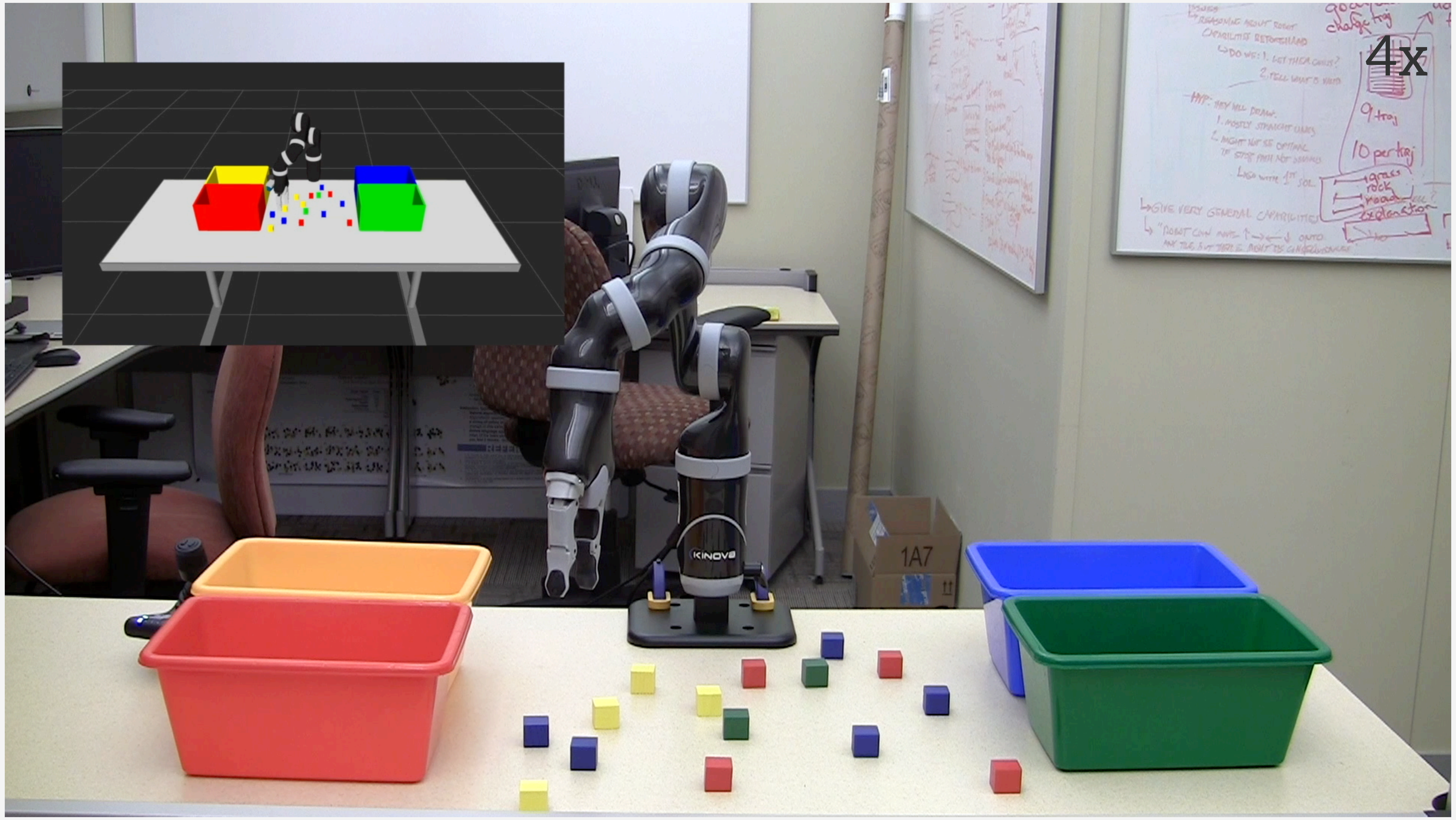
RO-MAN'16

Demonstration-based explanation for reasoning

The robot trajectory is indicated as the black dots, which indicates that **it prefers rocks.**

Submitted to RO-MAN'17

Language-based explanation



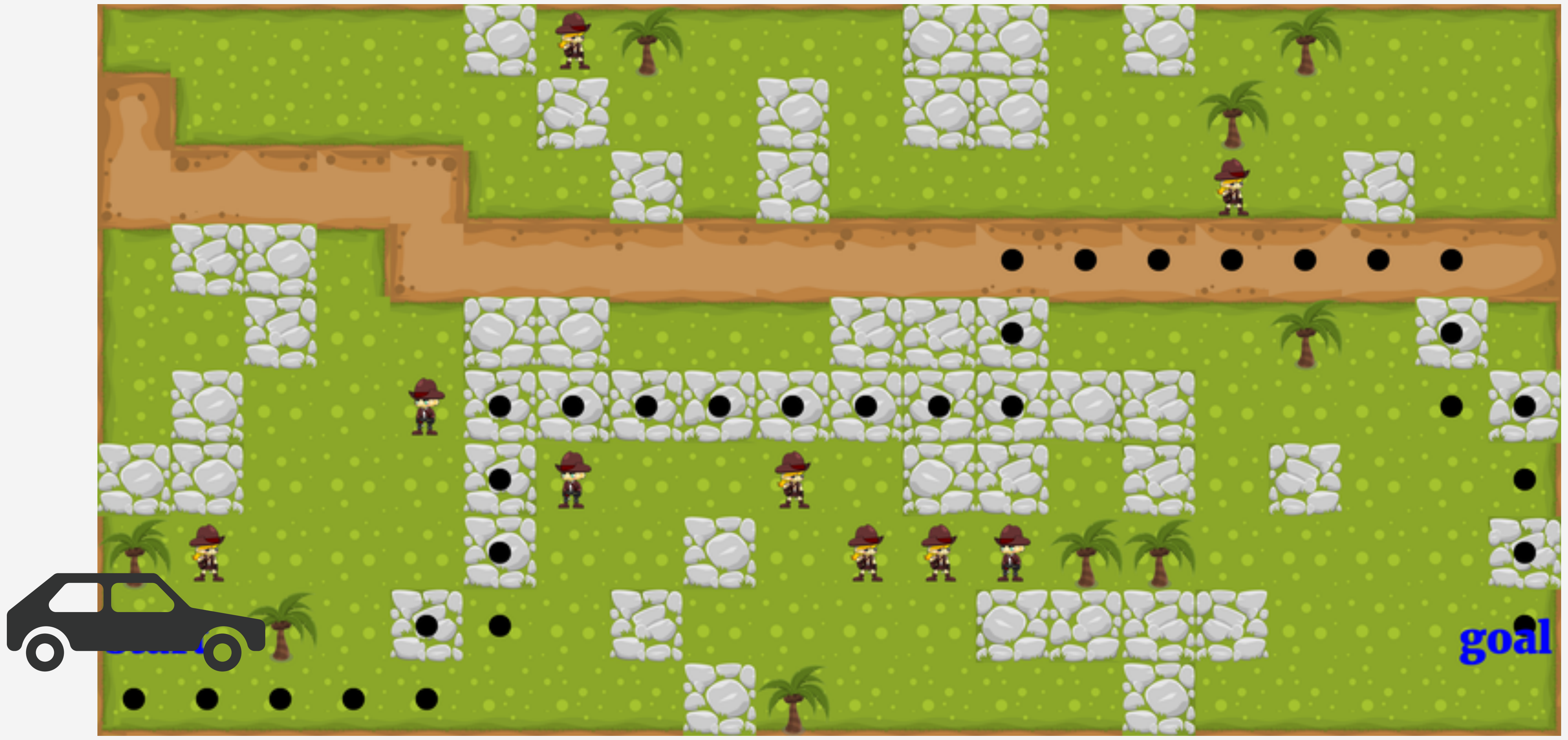
Kinova Mico Arm

Language-based explanation for intentions

"I am picking up the red block closest to you."

RO-MAN'16

Demonstration-based explanation for reasoning



The robot trajectory is indicated as the black dots, which indicates that **it prefers rocks.**

Submitted to RO-MAN'17

Software Engineering Institute
Carnegie Mellon

Automatically Evaluating and Generating Clear Robot Explanations

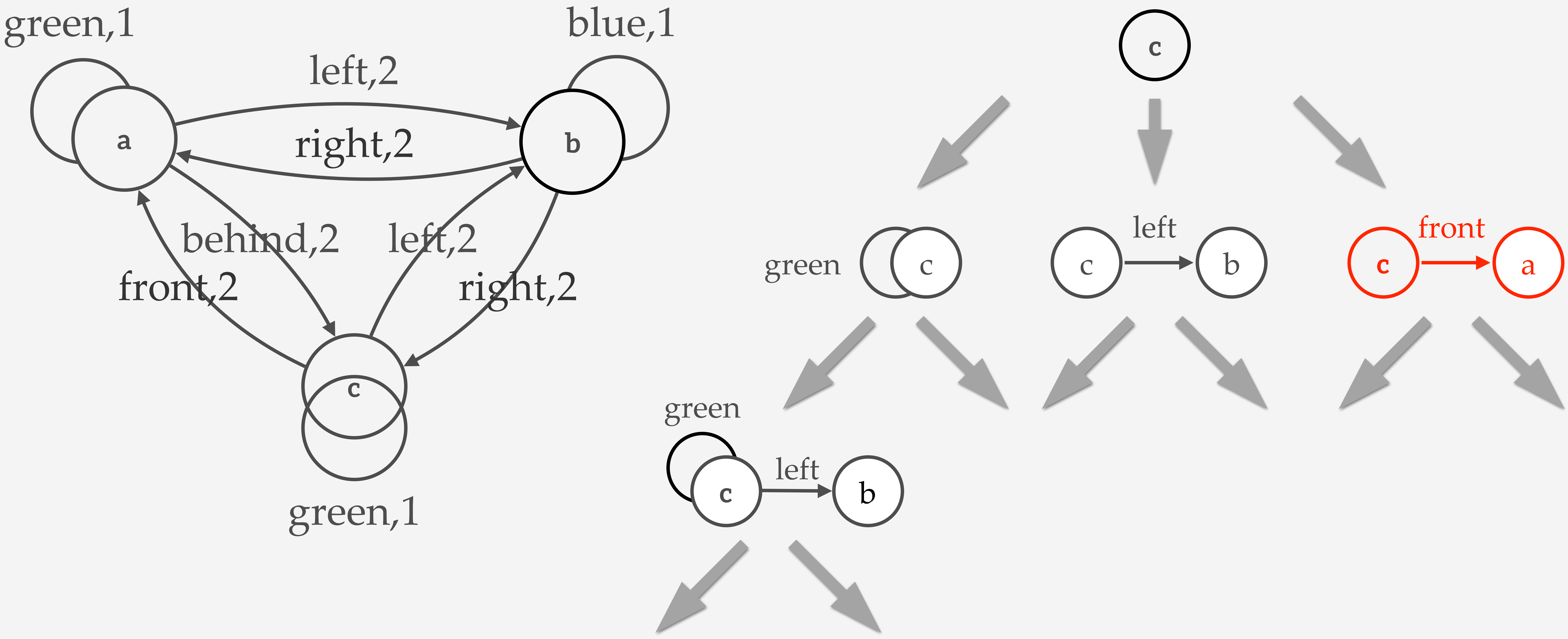
Thank you!
Questions?

Shen Li

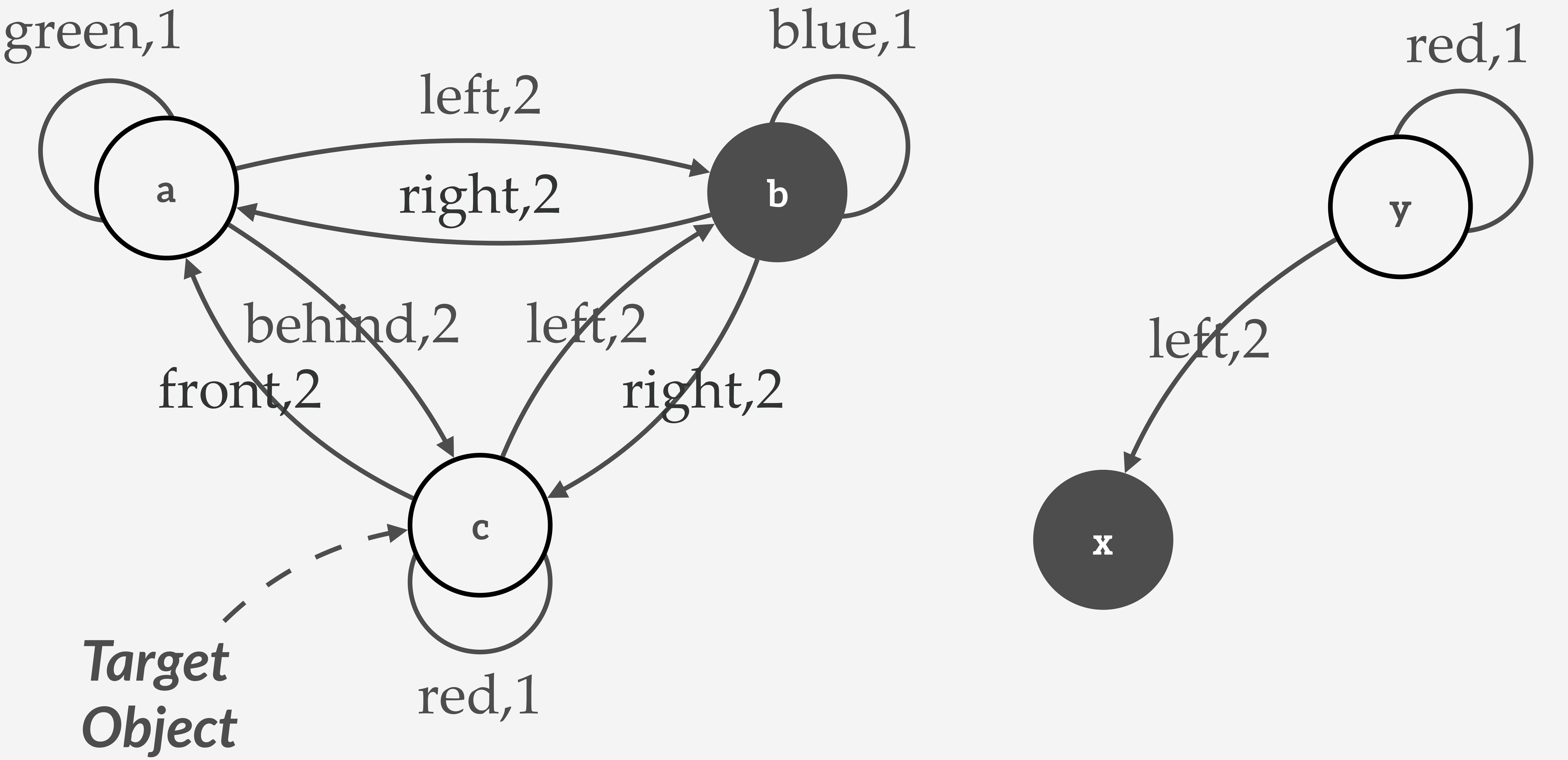
Thesis committee:
Dr. Siddhartha Srinivasa (co-chair)
Dr. Stephanie Rosenthal (co-chair)
Dr. Reid Simmons
Stefanos Nikolaidis

Richard King Mellon Foundation

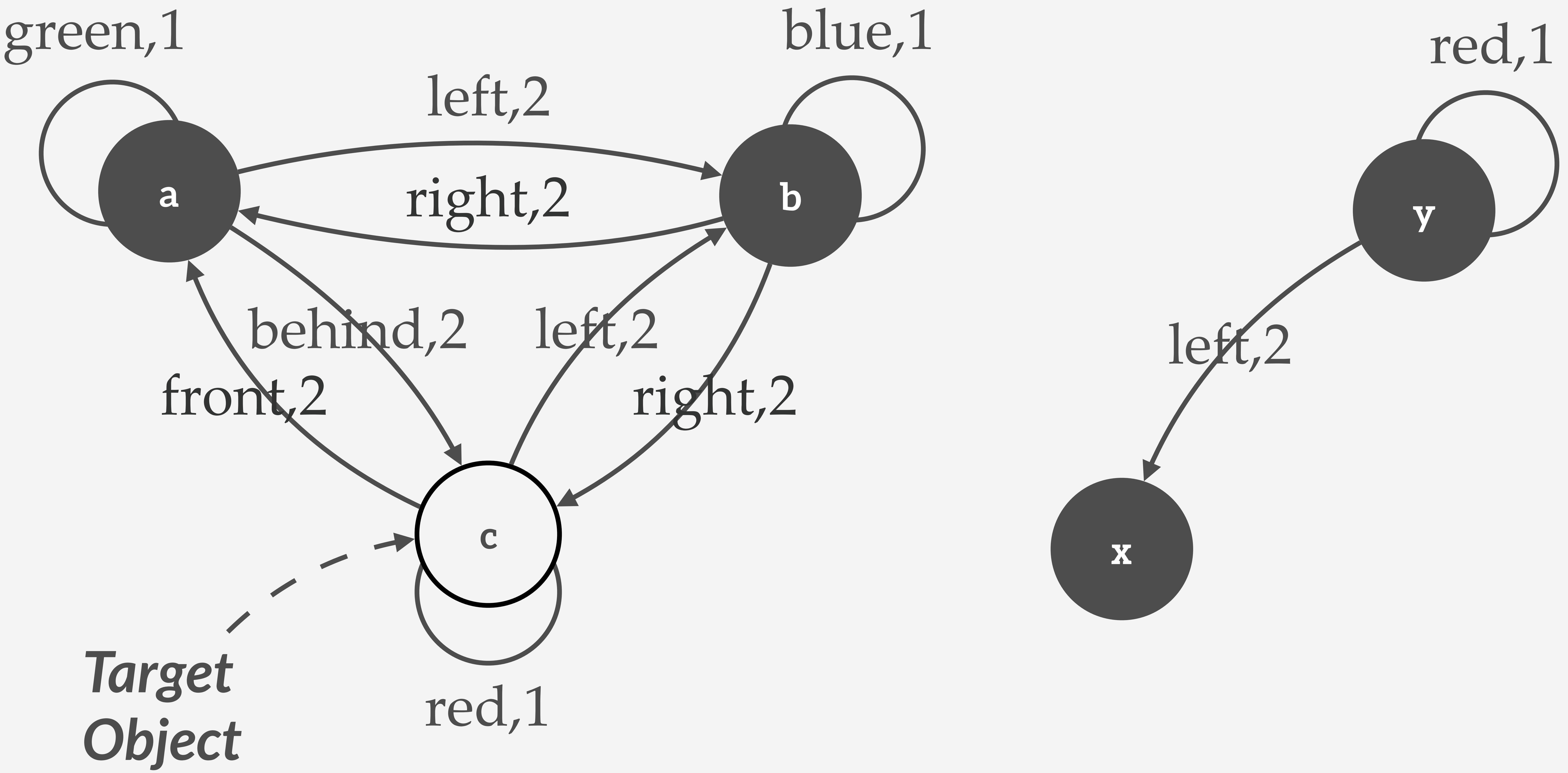
Search process



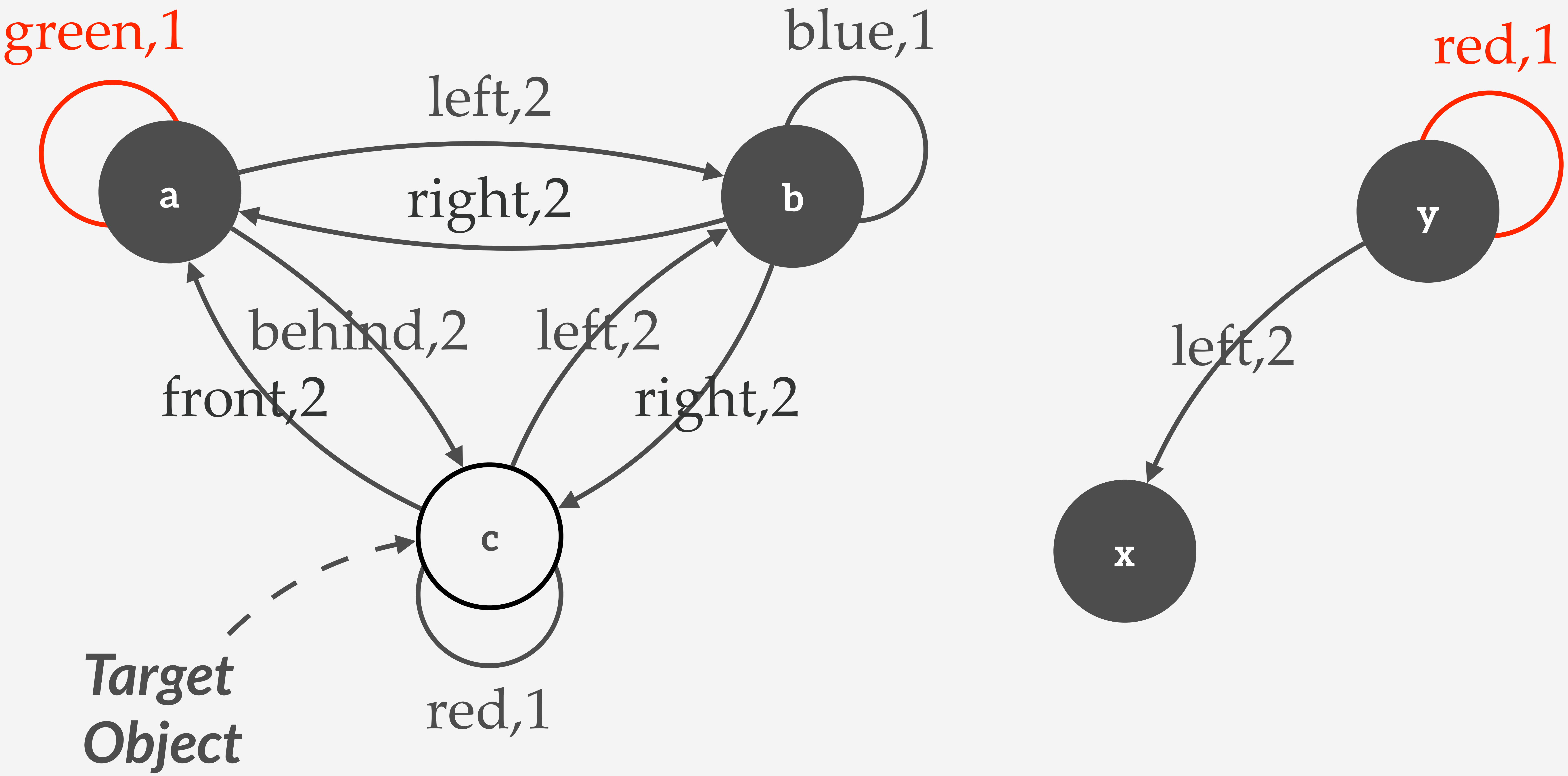
Isomorphism process (x ? b)



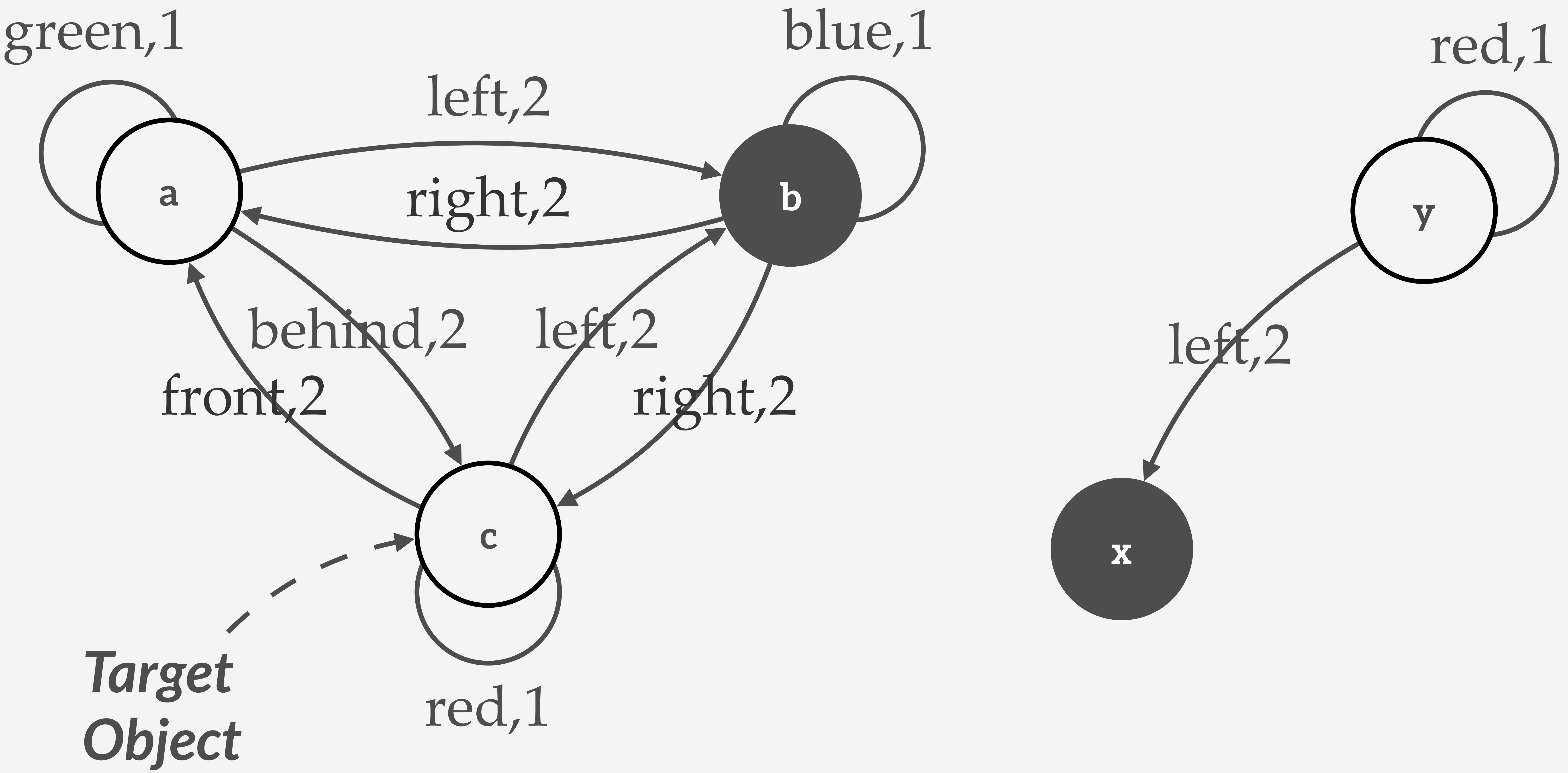
Isomorphism process (x ? b)



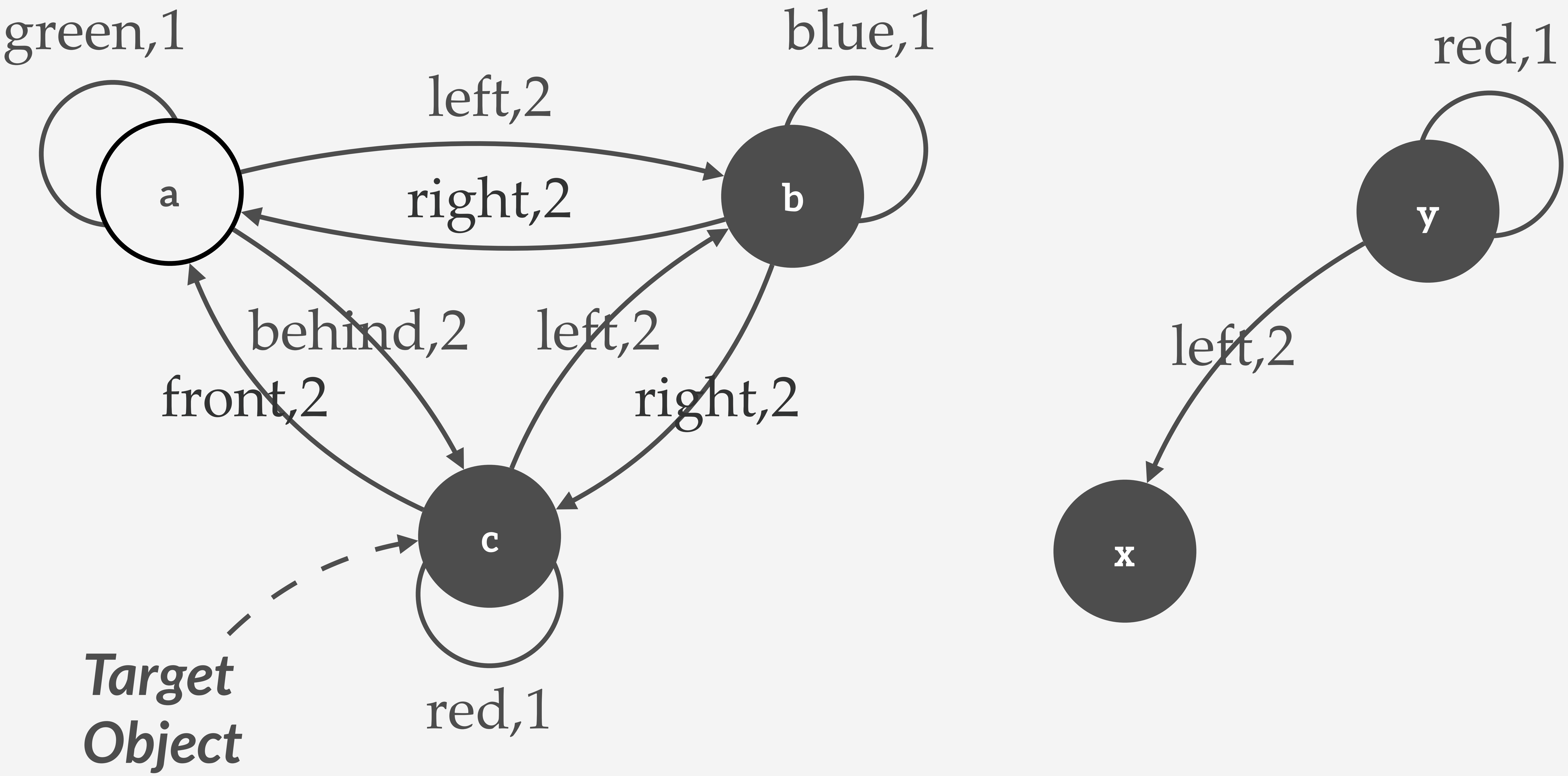
Isomorphism process (x ? b)



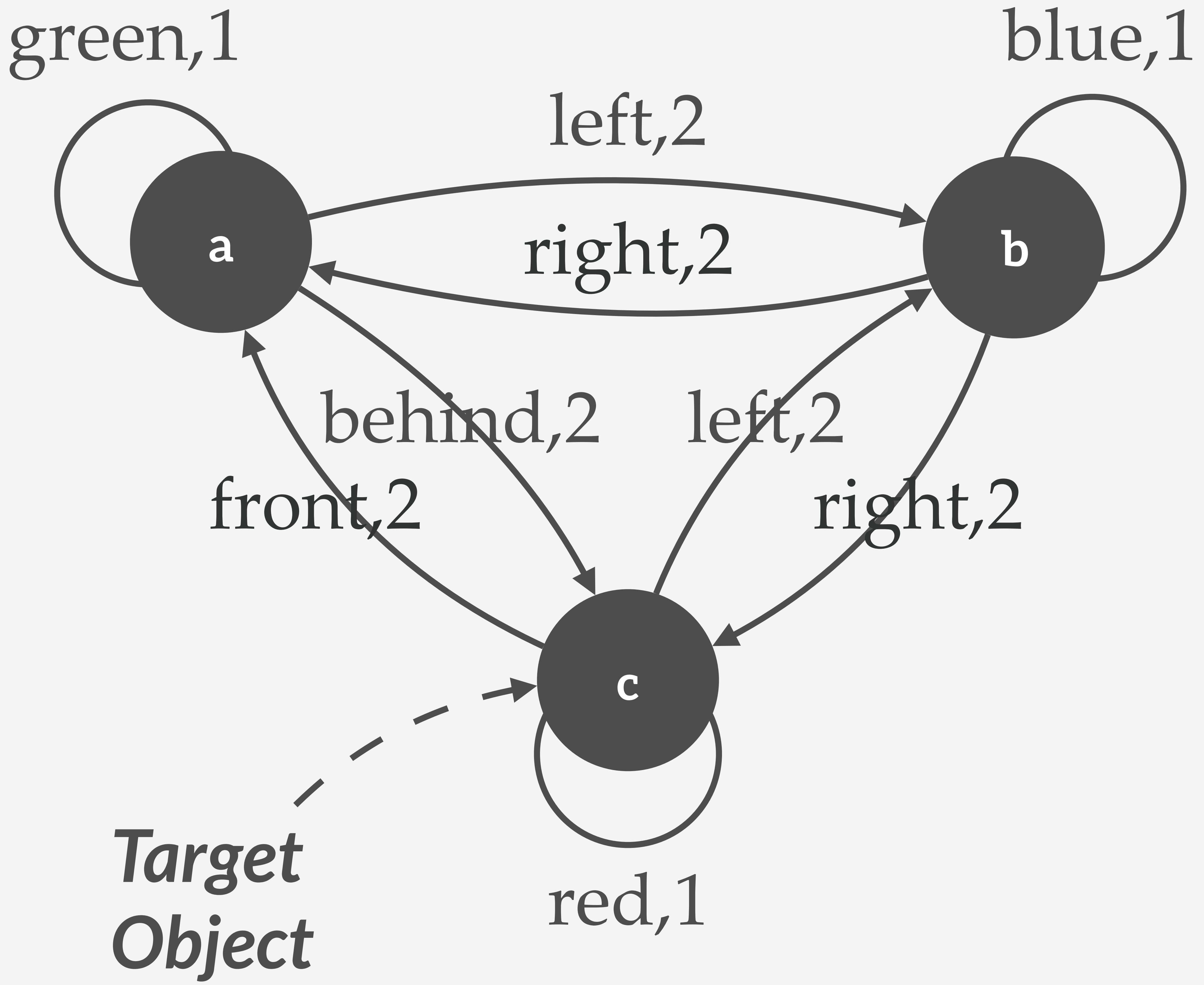
Isomorphism process (x ? b)

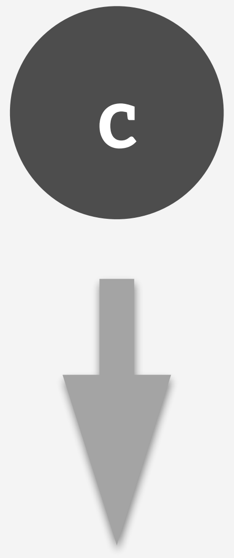


Isomorphism process (x ? b)

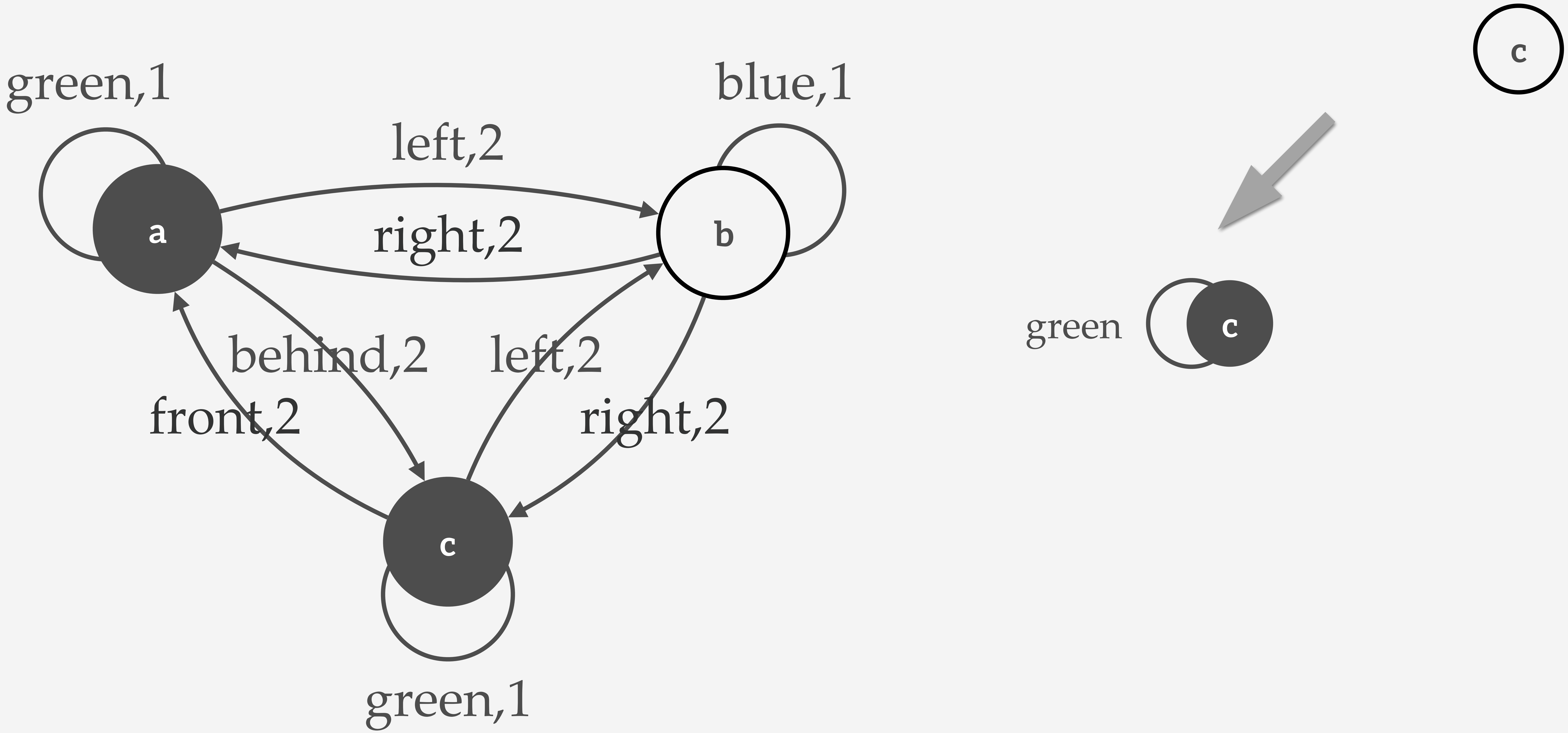


Prune the search process

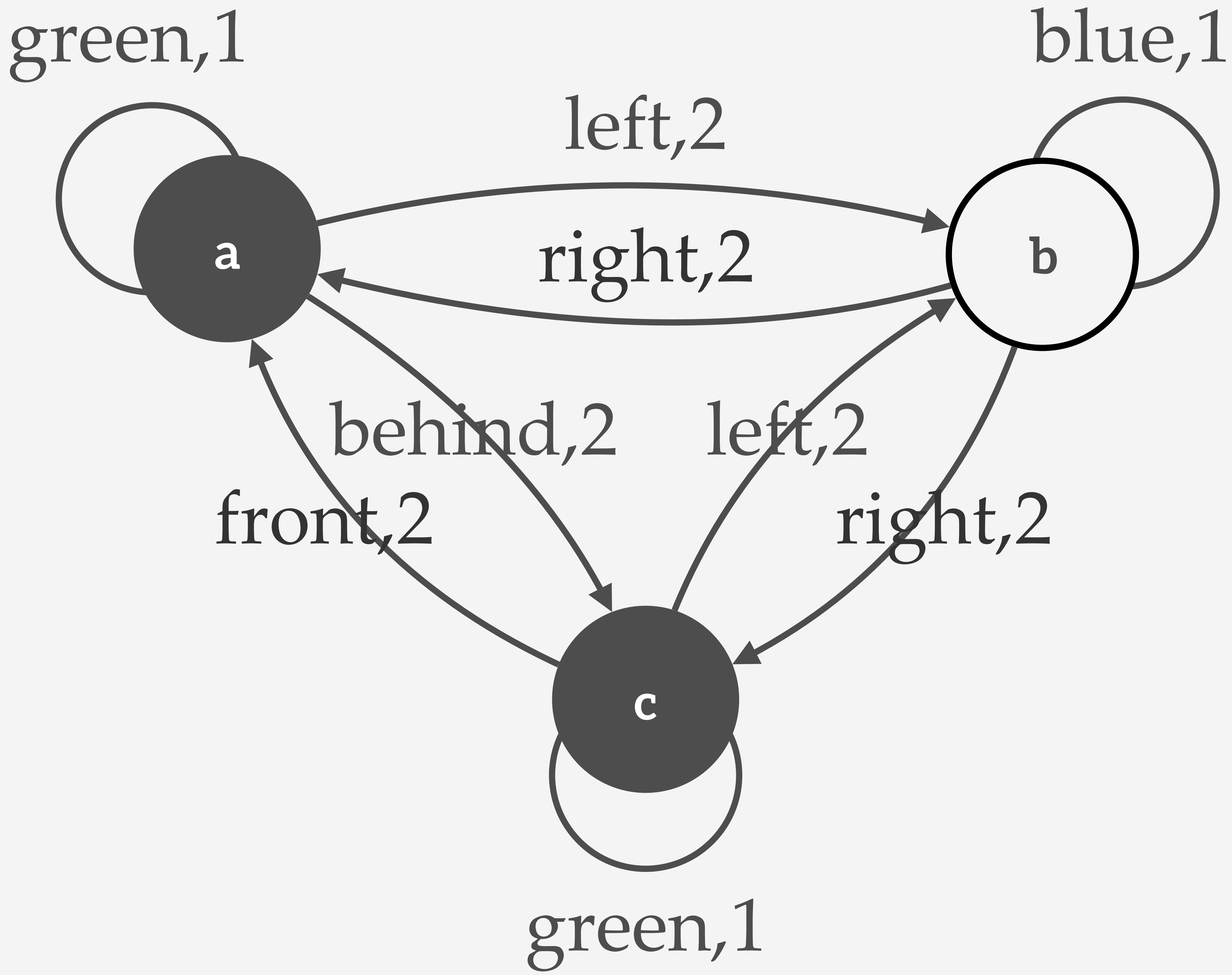


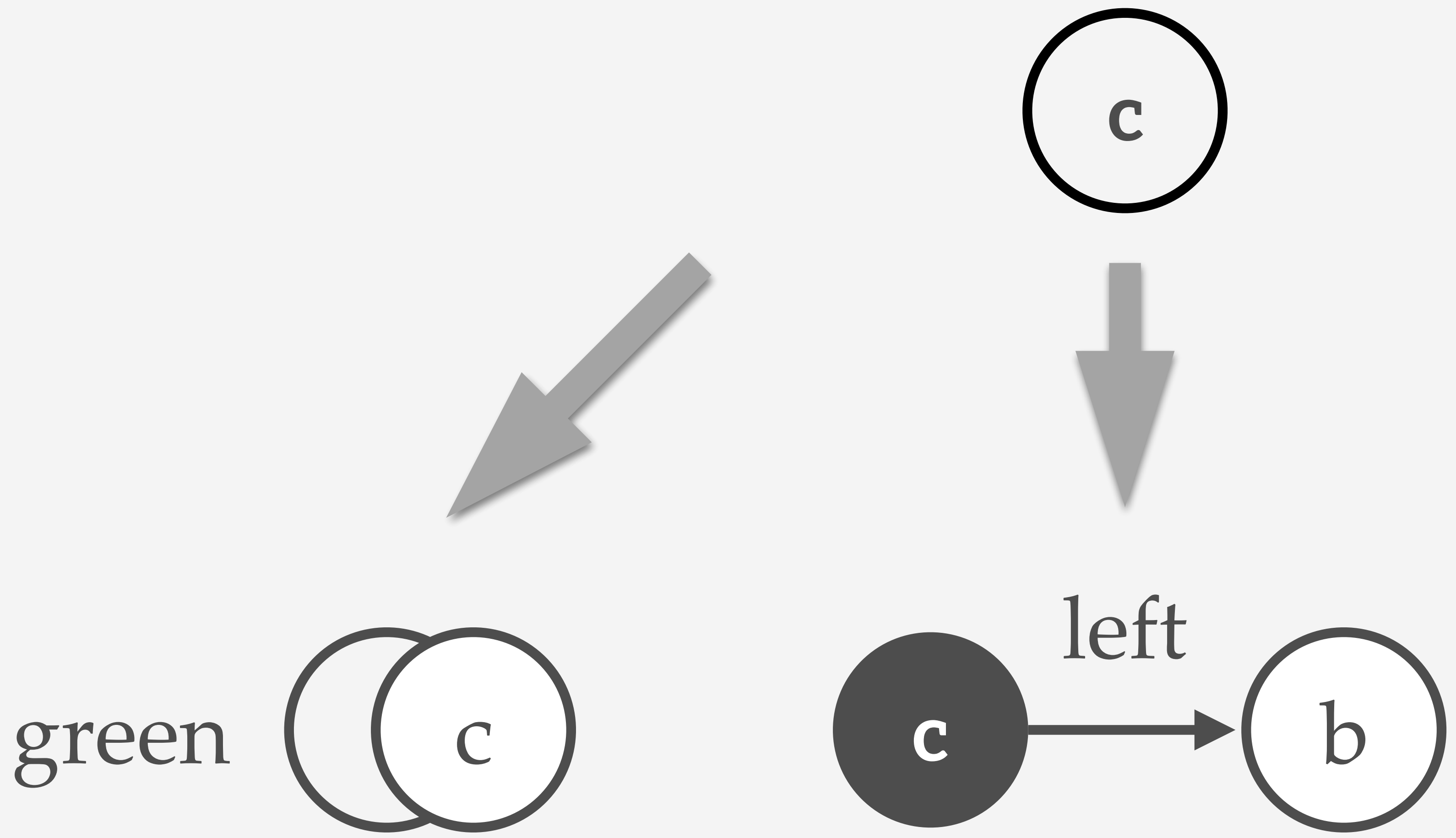


Prune the search process

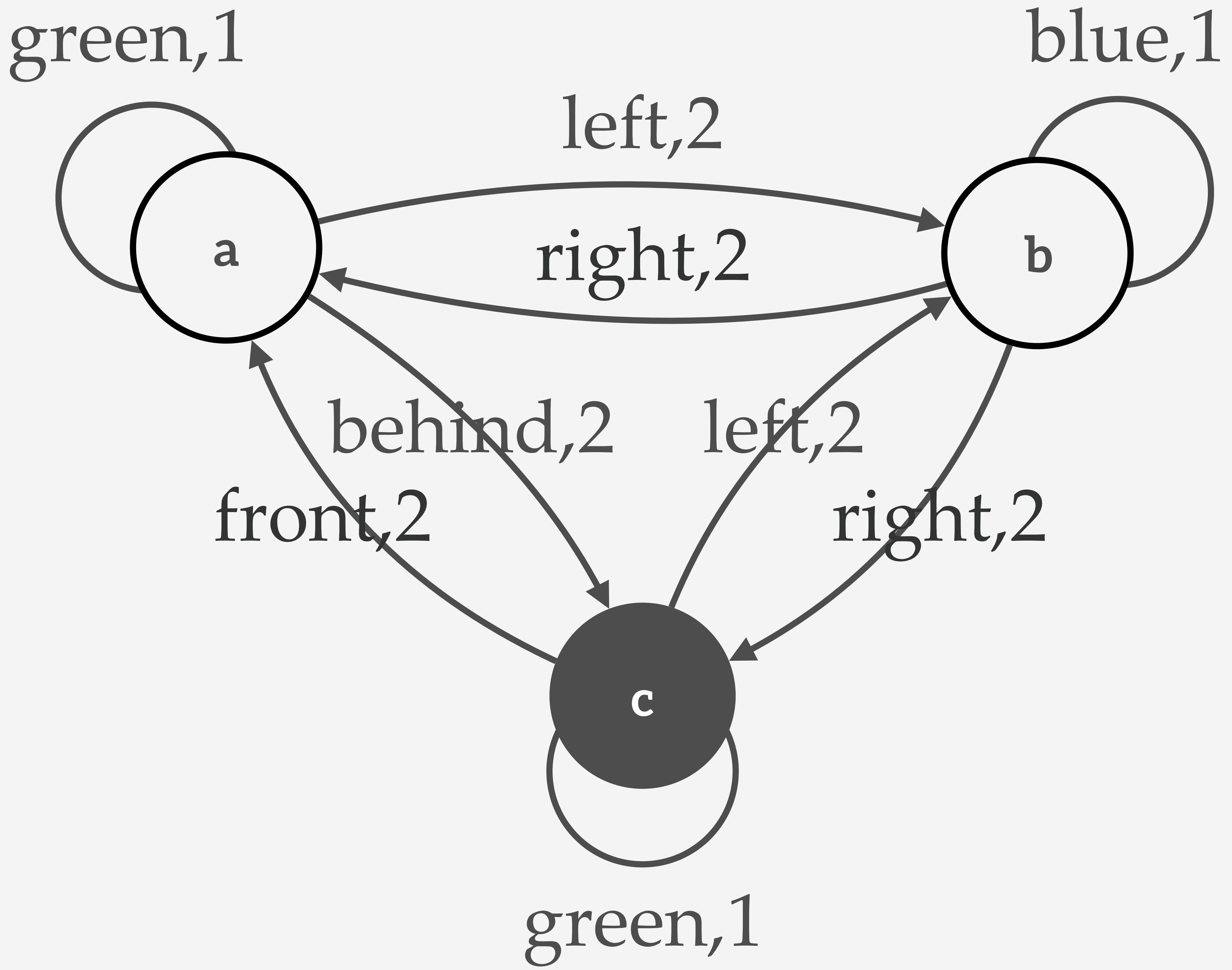


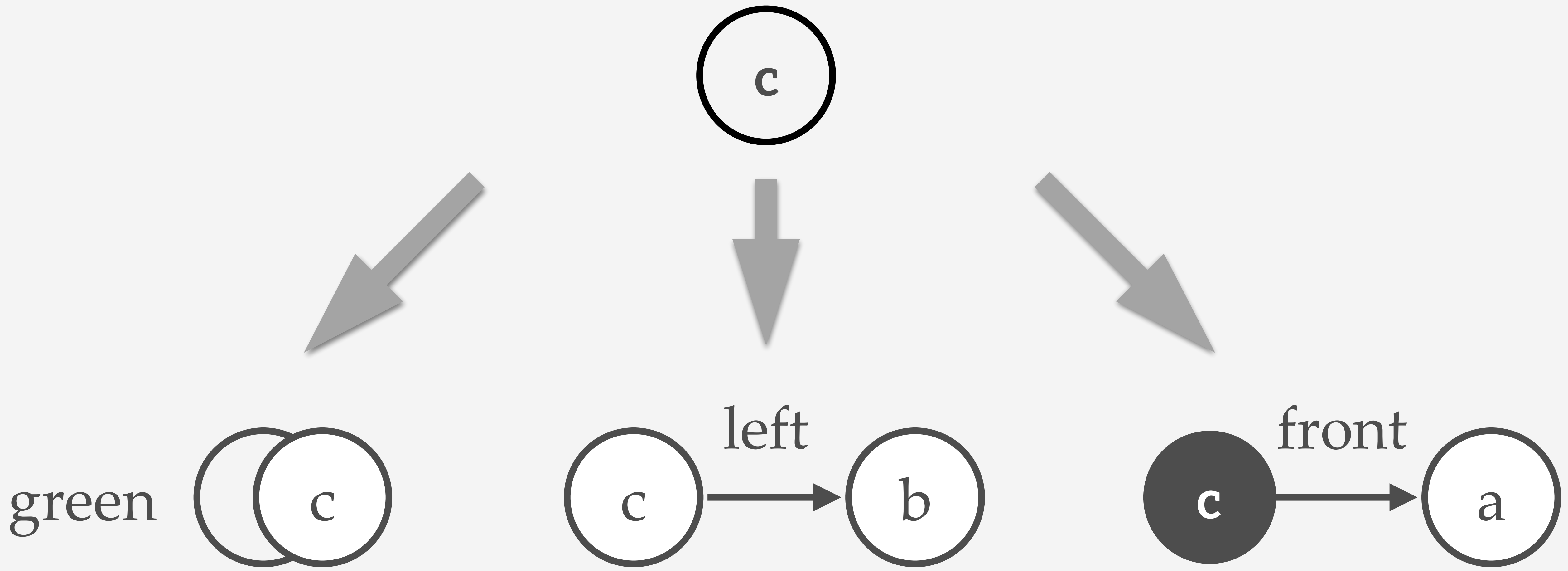
Prune the search process



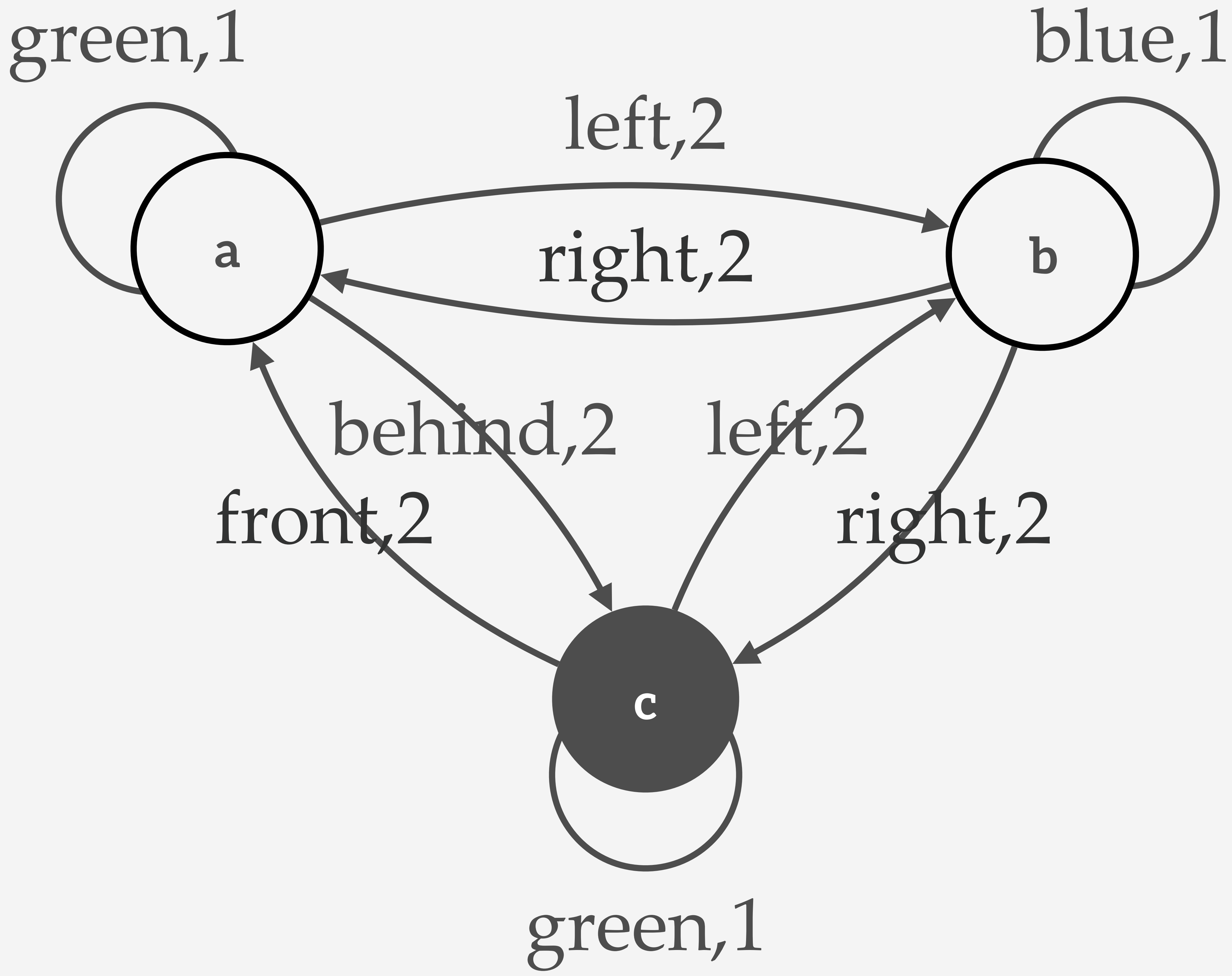


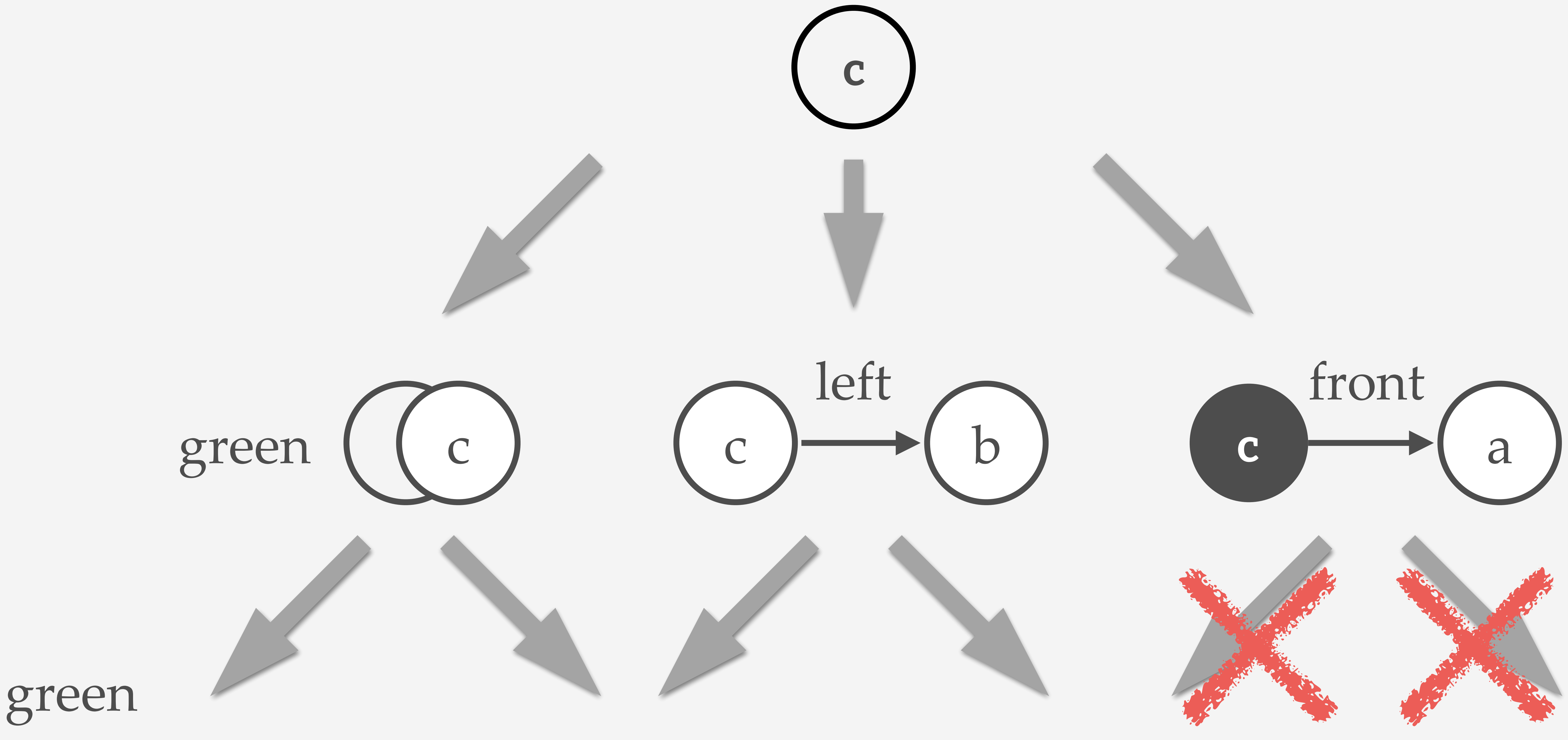
Prune the search process





Prune the search process





Prune the search process

Look-ahead heuristics

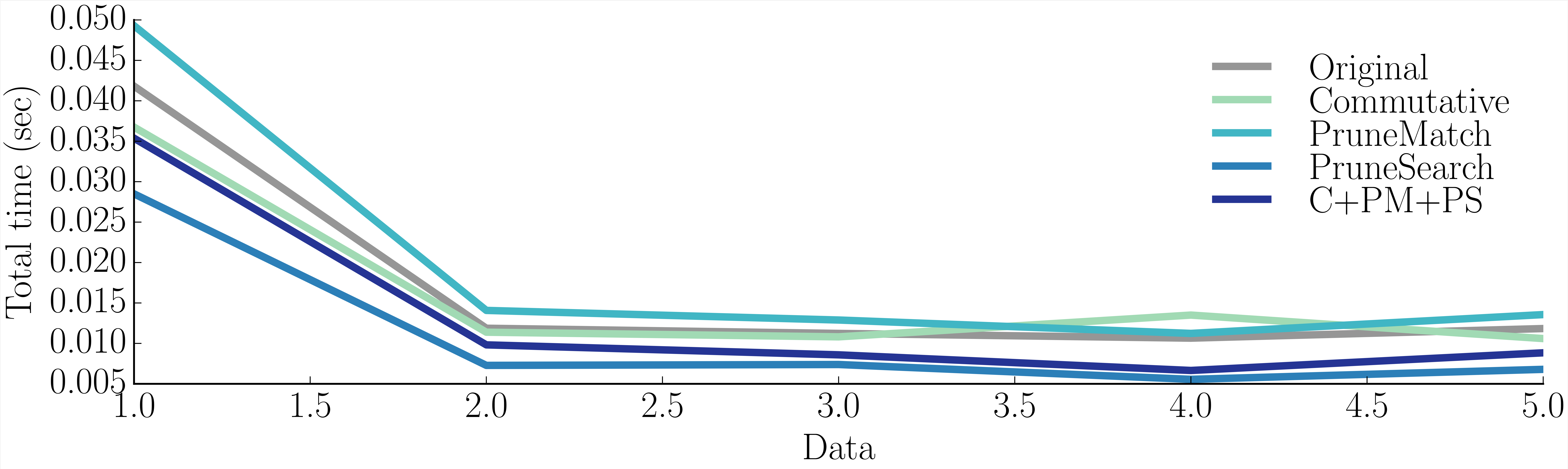
backiup, citation

Look-ahead heuristics

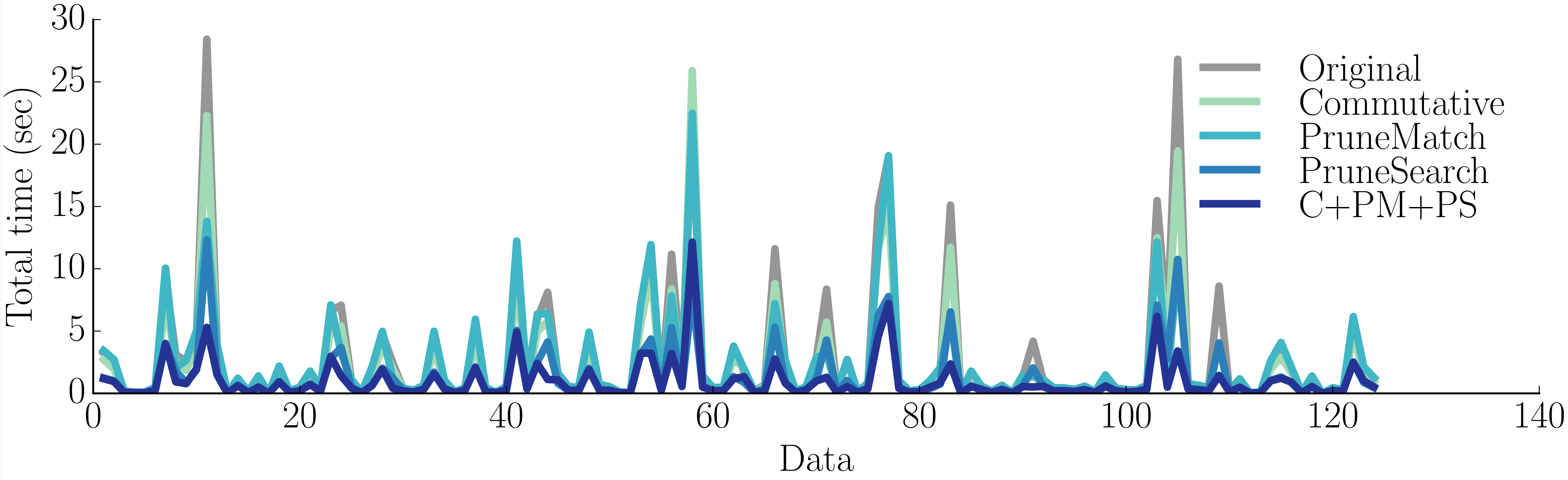
1 Look-ahead heuristic

2 Look-ahead heuristic

Result of all three techniques on subgraph with 1 feature (6 data)



Result of all three techniques on subgraph with 2 features (125 data)



Result of all three techniques on subgraph with 3 features (62 data)