Static Analysis of Data Science Software

Caterina Urban
ANTIQUE Research Team
INRIA & École Normale Supérieure, Paris, France

WIRED

BUSINESS MORE SIGN IN SUBSCRIBE

ERIC NIILER BUSINESS 03.25.2019 07:00 AM

Can AI Be a Fair Judge in Court? Estonia Thinks So

Estonia plans to use an artificial intelligence program to decide some small-claims cases, part of a push to make government services smarter.

The Telegraph

AI used for first time in job interviews in UK to find best applicants

By **Charles Hymas**

27 SEPTEMBER 2019 • 10:00 PM

4 ways to check for skin cancer with your smartphone

Your phone can help you recognize suspicious moles and marks, but you should still see a dermatologist or doctor.

BY AMANDA CAPRITTO | SEPTEMBER 16, 2019 10:57 AM PDT

WIRED

In 2019, predictive algorithms will start to make banking fair for all

By KATHRYN PETRALIA
11 Jan 2019

AUTOMATED BACKGROUND CHECKS ARE DECIDING WHO'S FIT FOR A HOME

By Colin Lecher | @colinlecher | Feb 1, 2019, 8:00am EST

Deep Neural Network Compression for Aircraft Collision Avoidance Systems

Kyle D. Julian[1] and Mykel J. Kochenderfer[2] and Michael P. Owen[3]

***Abstract*—One approach to designing decision making logic for an aircraft collision avoidance system frames the problem as a Markov decision process and optimizes the system using dynamic programming. The resulting collision avoidance strategy can be represented as a numeric table. This methodology has been used in the development of the Airborne Collision Avoidance System X (ACAS X) family of collision avoidance systems for manned and unmanned aircraft, but the high dimensionality of the state space leads to very large tables. To improve storage efficiency, a deep neural network is used to approximate the table. With the use of an asymmetric loss function and a gradient descent algorithm, the parameters for this network can be trained to provide accurate estimates of table values while preserving the relative preferences of the possible advisories for each state. By training multiple networks to represent subtables, the network also decreases the required runtime for computing the collision avoidance advisory. Simulation studies show that the network improves the safety and efficiency of the collision avoidance system. Because only the network parameters need to be stored, the required storage space is reduced by a factor of 1000, enabling the collision avoidance system to operate using current avionics systems.**

I. Introduction

Decades of research have explored a variety of approaches to designing decision making logic for aircraft collision avoidance systems for both manned and unmanned aircraft [1]. Recent work on formulating the problem of collision avoidance as a partially observable Markov decision process (POMDP) has led to the development of the Airborne Collision Avoidance System X (ACAS X) family of collision avoidance systems [2], [3], [4]. The version for manned aircraft, ACAS

floating point storage. A simple technique to reduce the size of the score table is to downsample the table after dynamic programming. To minimize the degradation in decision quality, states are removed in areas where the variation between values in the table are smooth. The downsampling reduces the size of the table by a factor of 180 from that produced by dynamic programming. For the rest of this paper, the downsampled ACAS Xu horizontal table is referred to as the baseline, original table.

Even after downsampling, the current table requires over 2GB of floating point storage, too large for certified avionics systems [6]. Although modern hardware can handle 2GB of storage, the certification process for aircraft computer hardware is expensive and time-consuming, so a solution capable of running on legacy hardware is desired [7]. While there is no formal limit for floating point storage on legacy avionics, a representation occupying less than 120MB would be sufficient.

For an earlier version of ACAS Xa, block compression was introduced to take advantage of the fact that, for many discrete states, the scores for the available actions are identical [8]. One critical contribution of that work was the observation that the table could be stored in IEEE half-precision with no appreciable loss of performance. Block compression was adequate for the ACAS Xa tables that limit advisories to vertical maneuvers, but the ACAS Xu tables for horizontal maneuvers are much larger. Recent work explored a new algorithm that exploits the score table's natural symmetry to remove redundancy within the table [9]. However, results showed that this compression algorithm could not achieve sufficient reduction in storage before compromising performance.

These tables like this can be represented as

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

STAT+

IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show

By Casey Ross *@caseymross and* Ike Swetlitz

July 25, 2018

TOM SIMONITE BUSINESS 08.17.2017 07:00 AM

When Government Rules by Software, Citizens Are Left in the Dark

Agencies decline to release information about algorithms used for criminal justice, social welfare, and education.

3,658,826 views | Feb 16, 2012, 11:02am

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Kashmir Hill Former Staff

Tech

Welcome to The Not-So Private Parts where technology & privacy collide

China 'social credit': Beijing sets up huge system

By Celia Hatton
BBC News, Beijing

26 October 2015

BUSINESS NEWS OCTOBER 10, 2018 / 5:12 AM / A YEAR AGO

Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

A self-driving Uber ran a red light last December, contrary to company claims

Internal documents reveal that the car was at fault

By Andrew Liptak | @AndrewLiptak | Feb 25, 2017, 11:08am EST

Where is the Problem?

pre-processing training data analysis

Data is Dirty

- **inconsistent** data
- **incorrect** data
- **incomplete** data
- **inaccurate** data

Pre-Processing is Fragile

pre-processing

mislabeled data

accidentally duplicated data

wrongly converted data

accidentally (un)used data

TECHNOLOGY

The New York Times

For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights

By Steve Lohr

Aug. 17, 2014

Technology revolutions come in measured, sometimes foot-dragging steps. The lab science and marketing enthusiasm tend to underestimate the bottlenecks to progress that must be overcome with hard work and practical engineering.

The field known as "big data" offers a contemporary case study. The catchphrase stands for the modern abundance of digital data from many sources — the web, sensors, smartphones and corporate databases — that can be mined with clever software for discoveries and insights. Its promise is smarter, data-driven decision-making in every field. That is why data scientist is the economy's hot new job.

Yet far too much handcrafted work — what data scientists call "data wrangling," "data munging" and "data janitor work" — is still required. Data scientists, according to interviews and expert estimates, spend from 50 percent to 80 percent of their time mired in this more mundane labor of collecting and preparing unruly digital data, before it can be explored for useful nuggets.

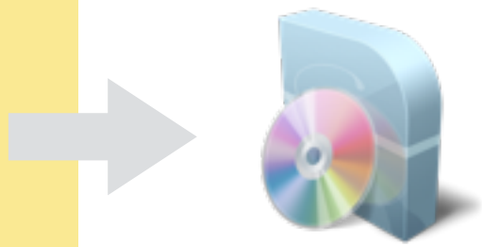
"Data wrangling is a huge — and surprisingly so — part of the job," said Monica Rogati, vice president for data science at Jawbone, whose sensor-filled wristband and software track activity, sleep and food consumption, and suggest dietary and health tips based on the numbers. "It's something that is not appreciated by data civilians. At times, it feels like everything we do."

Several start-ups are trying to break through these big data bottlenecks by developing software to automate the gathering, cleaning and organizing of disparate data, which is plentiful but messy. The modern Wild West of data needs to be tamed somewhat so it can be recognized and exploited by a computer program.

"It's an absolute myth that you can send an algorithm over raw data and have insights pop up," said Jeffrey Heer, a professor of computer science at the University of Washington and a co-founder of Trifacta, a start-up based in San Francisco.

Accuracy is Meaningless

training



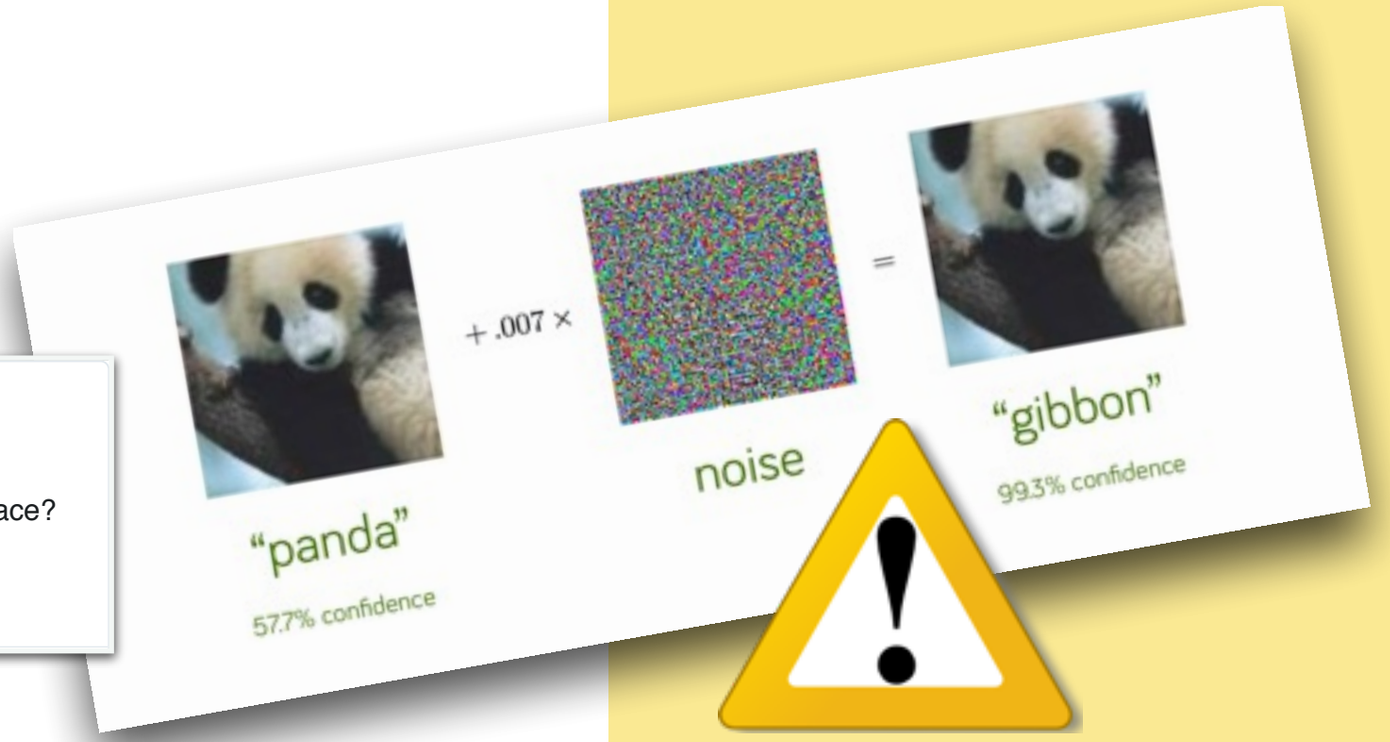
data analysis

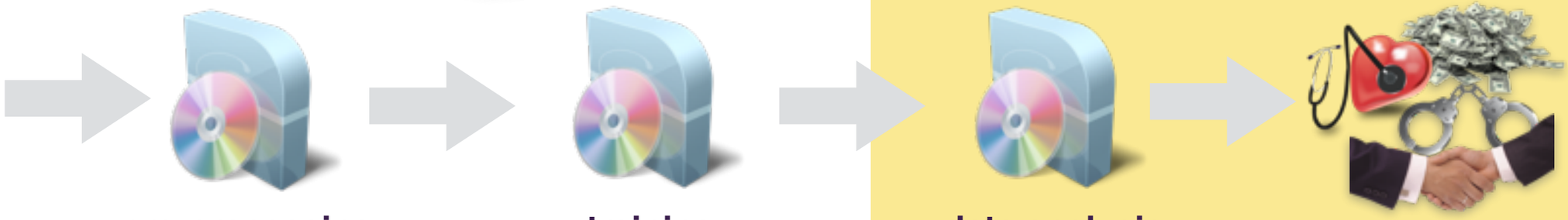
Inscrutability

Janelle Shane
@JanelleCShane

Does anyone have a picture of sheep in a really unusual place? It's for pranking a neural net.

3,676 5:55 PM - Mar 1, 2018





pre-processing training data analysis

MIT Technology Review

Artificial Intelligence / Machine Learning

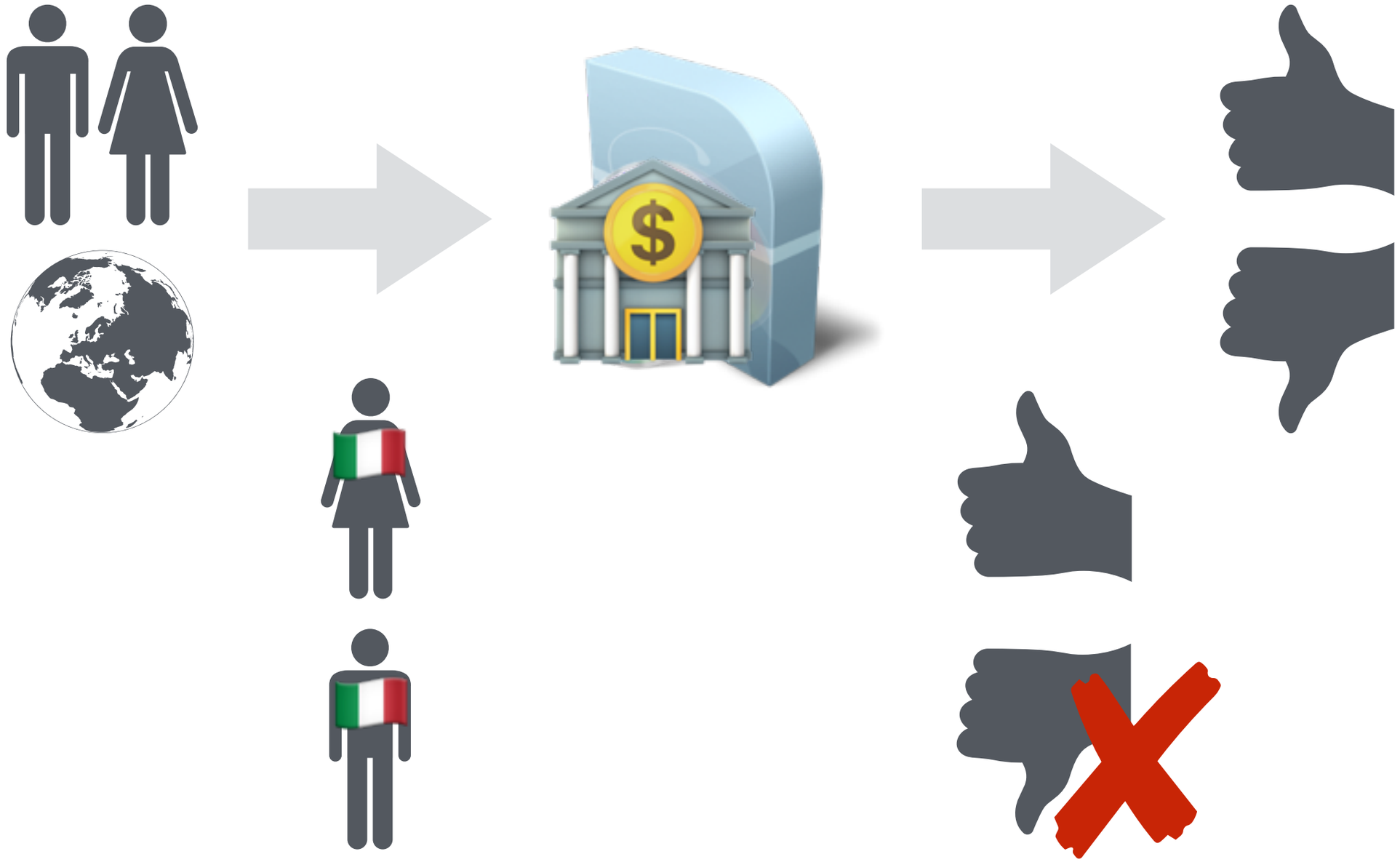
The Dark Secret at the Heart of AI

No one really knows how the most advanced algorithms do what they do. That could be a problem.

by **Will Knight** Apr 11, 2017

Algorithmic Fairness

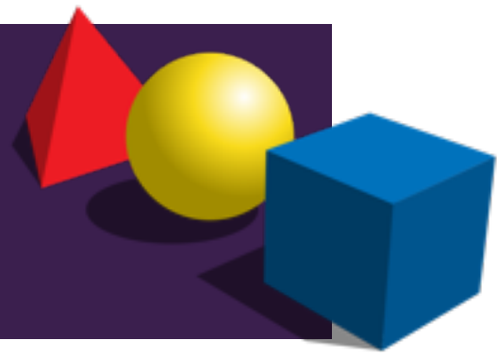
Causal Fairness



Static Analysis Recipe

practical tools
targeting specific programs

algorithmic approaches
to decide program properties

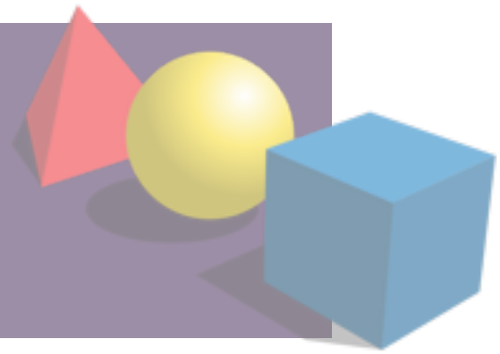


mathematical models
of the program behavior

Static Analysis Recipe

practical tools
targeting specific programs

algorithmic approaches
to decide program properties

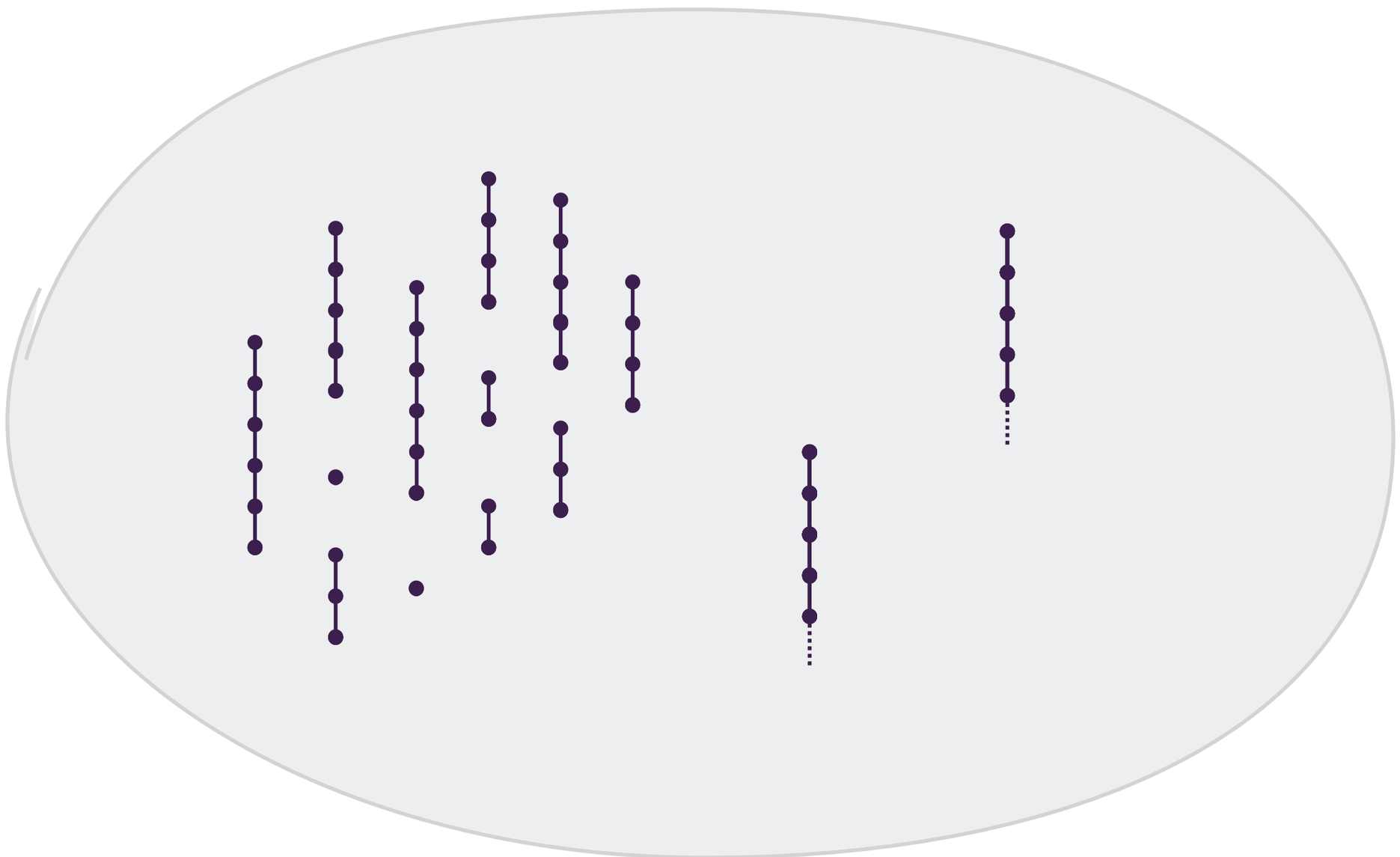


mathematical models
of the program behavior

Trace Semantics

$$\Lambda = \mathrm{lfp}^{\sqsubseteq}\ \Theta$$

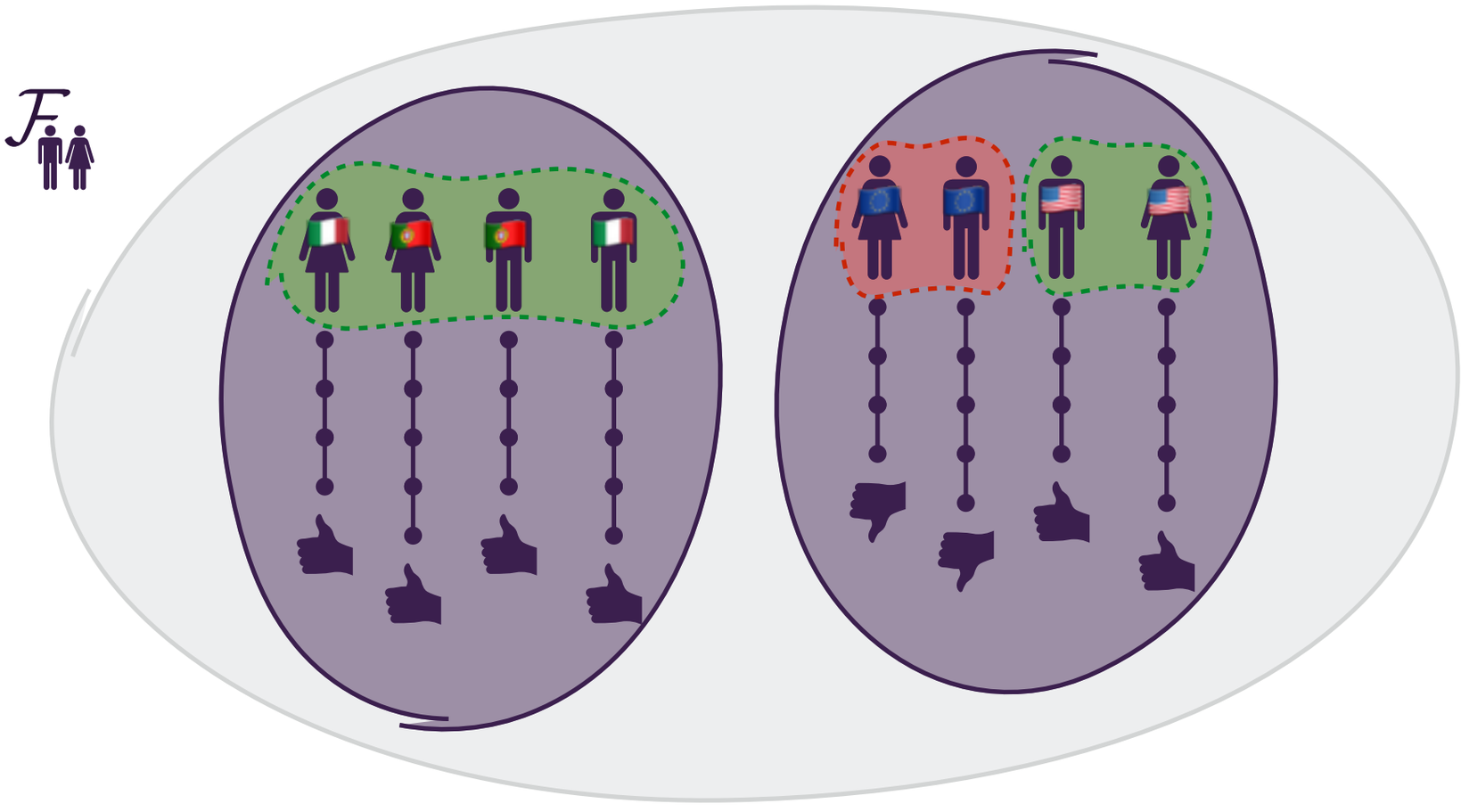
$$\Theta(T) \stackrel{\mathrm{def}}{=} \Omega \cup (\tau\ ;\ T)$$

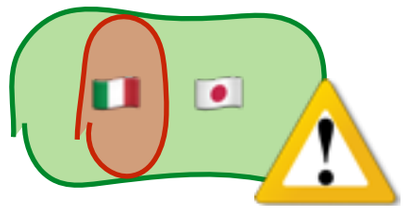


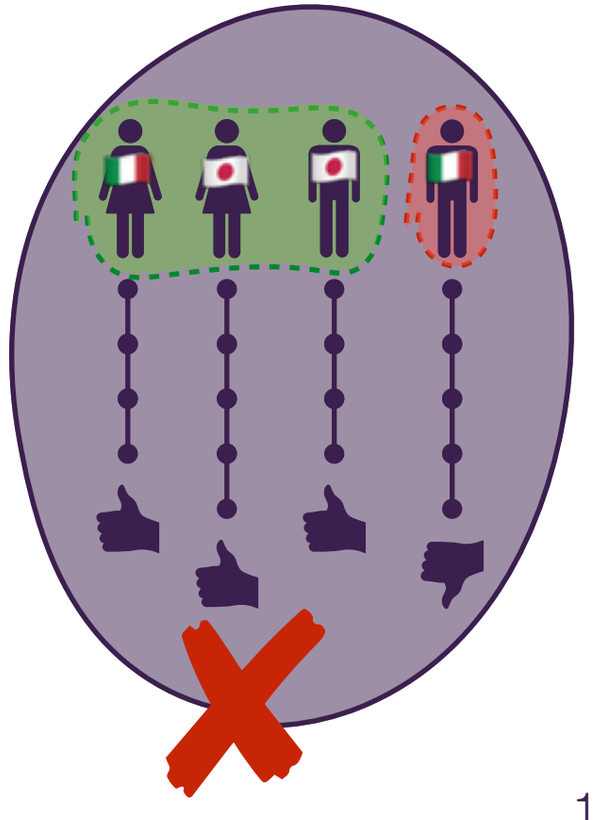
Causal Fairness

the program outcome is independent from the sensitive inputs

$$\mathcal{F}_K \stackrel{\text{def}}{=} \{[\![P]\!] \in \mathcal{P}\left(\Sigma^{+\infty}\right) \mid \forall i \in K : \textsc{unused}_i([\![P]\!])\}$$





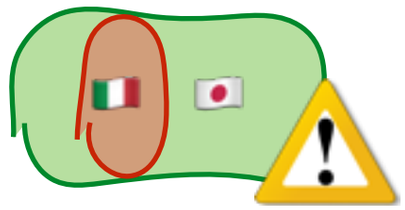


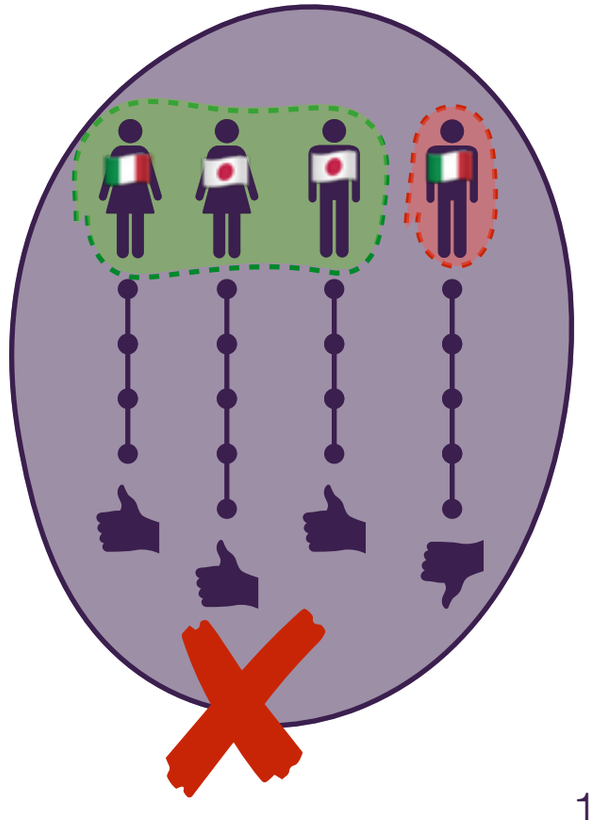
any possible outcome is possible from any sensitive value

Causal Fairness is Not a Trace Property

the program outcome is independent from the sensitive inputs

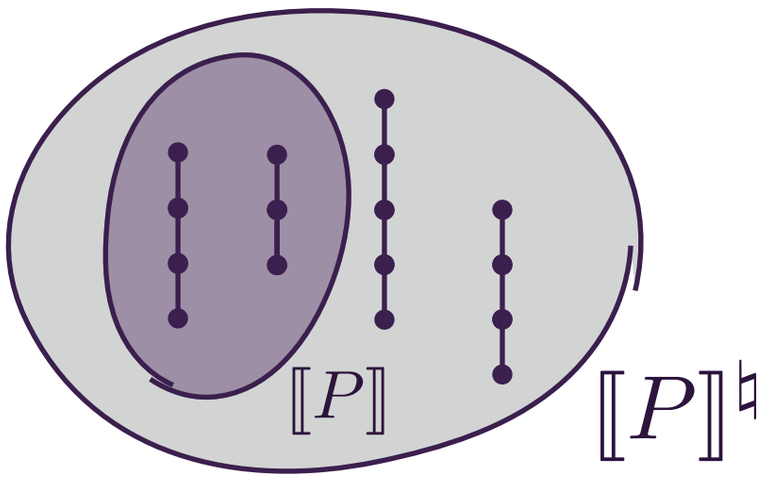
$$\mathcal{F}_K \stackrel{\text{def}}{=} \{[\![P]\!] \in \mathcal{P}\left(\Sigma^{+\infty}\right) \mid \forall i \in K : \textsc{unused}_i([\![P]\!])\}$$

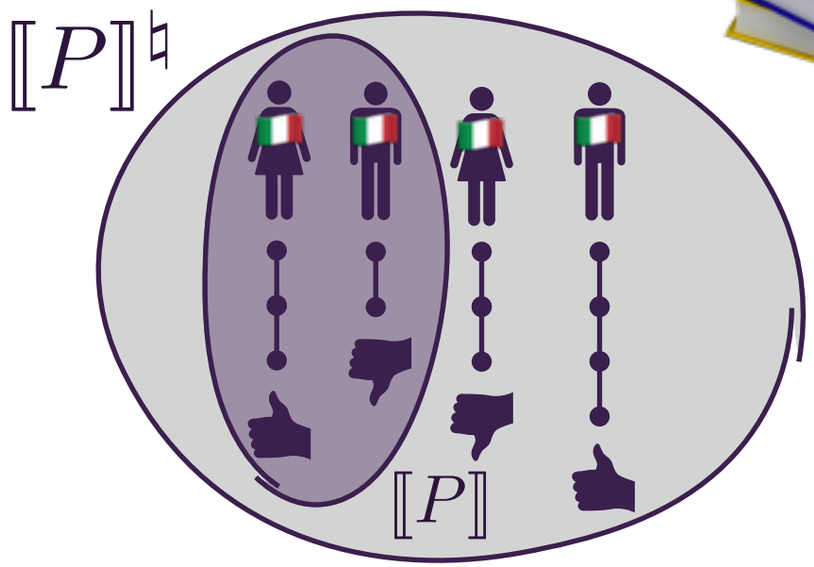


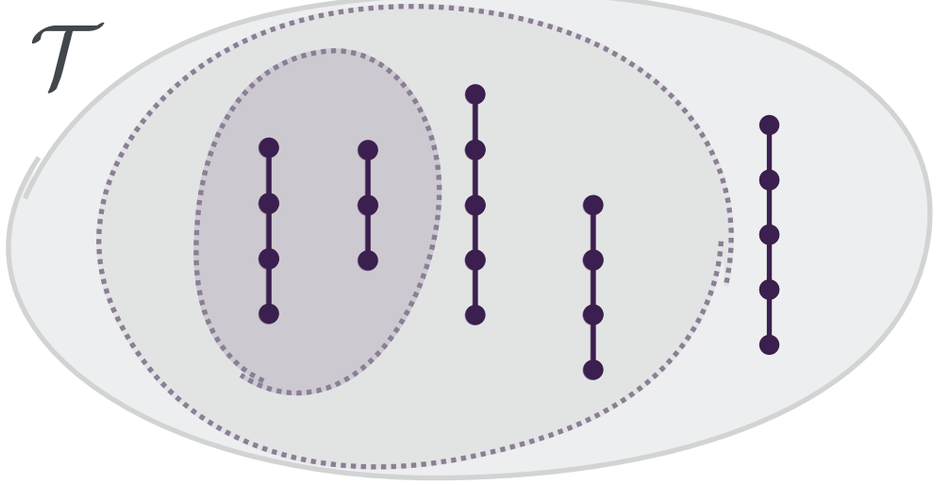


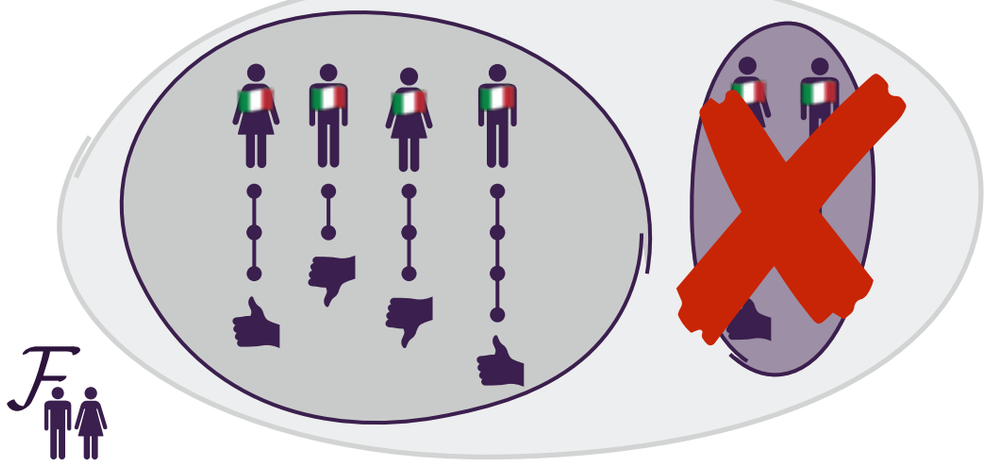
any possible outcome is possible from any sensitive value

Sound Causal Fairness Validation









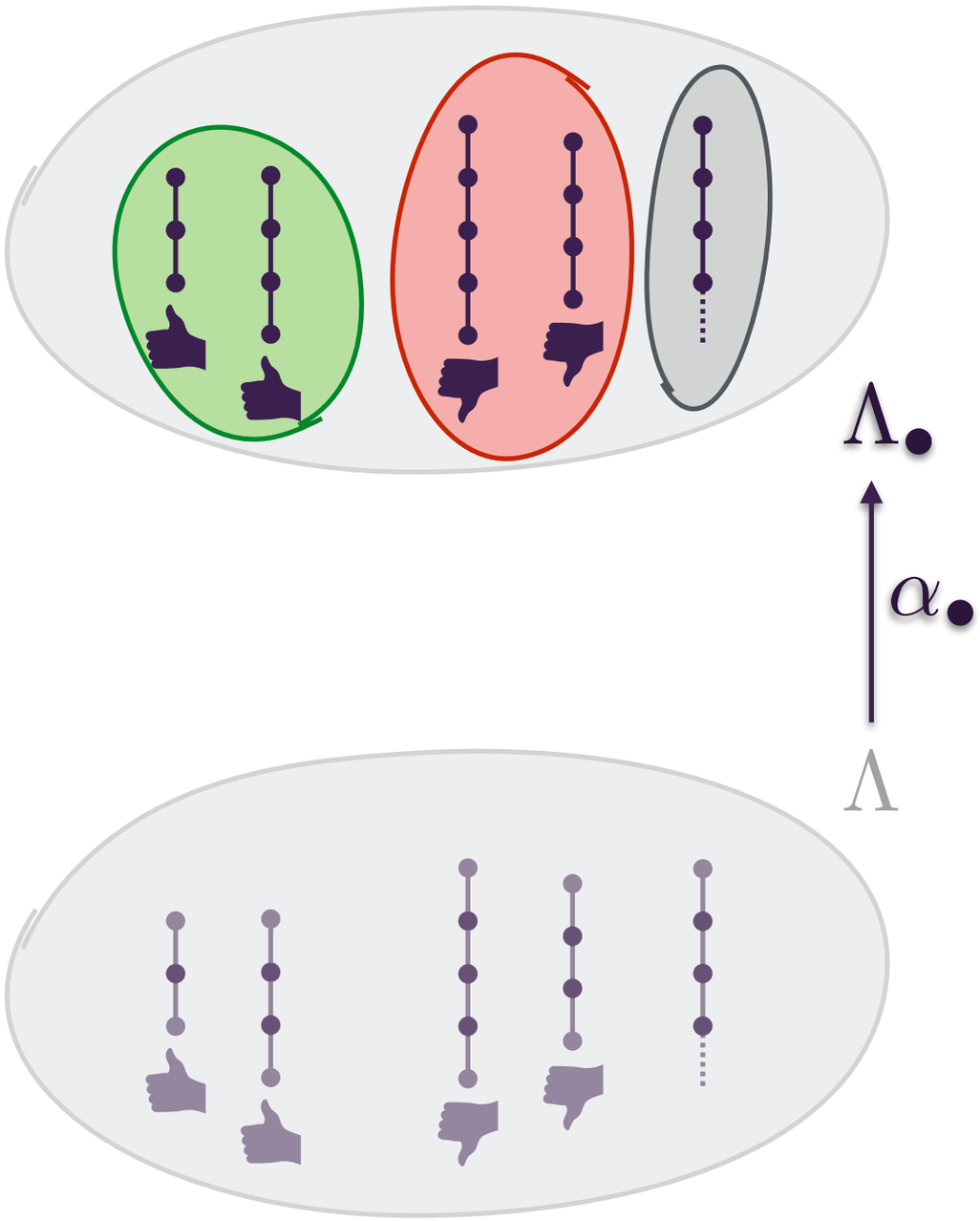
$$[\![P]\!]^\natural \subseteq \mathcal{T} \Rightarrow [\![P]\!] \subseteq \mathcal{T}$$

$$[\![P]\!]^\natural \in \mathcal{F}_K \not\Rightarrow [\![P]\!] \in \mathcal{F}_K$$

Trace Properties

Causal Fairness

Outcome Abstraction

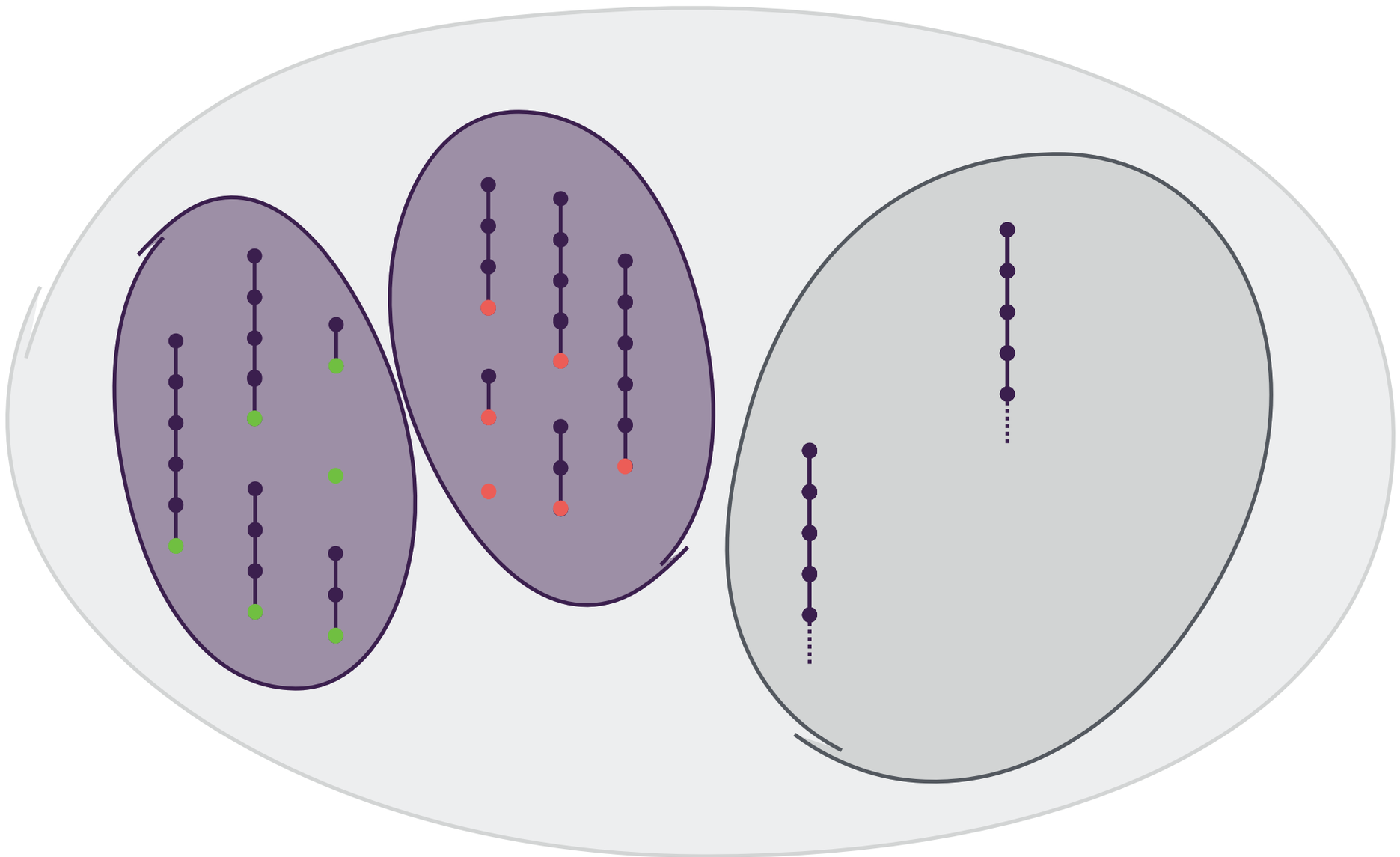


partition executions based on their outcome

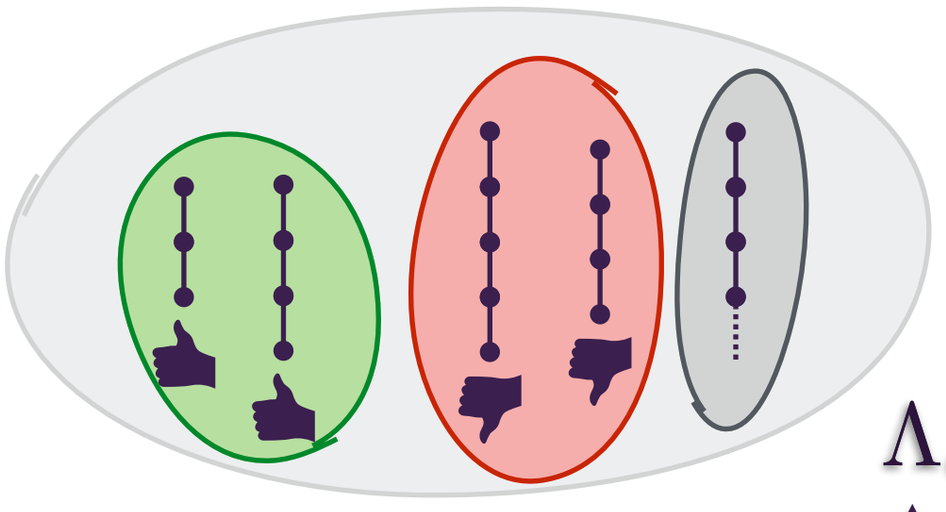
Outcome Semantics

$$\Lambda_\bullet = \mathrm{lfp}^{\sqsubseteq_\bullet}\, \Theta_\bullet$$

$$\Theta_\bullet(S) \stackrel{\mathrm{def}}{=} \{\Omega_o \mid o \in O\} \uplus \{\tau\,;T \mid T \in S\}$$



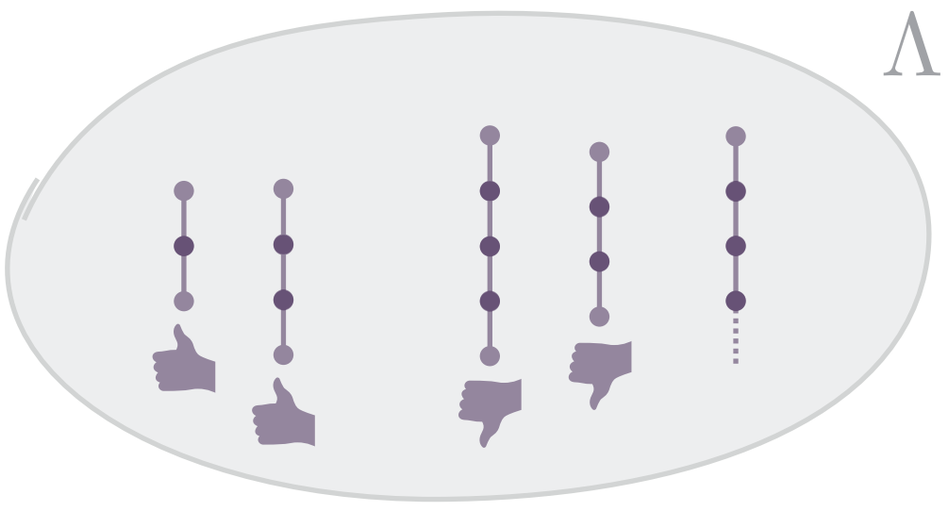
Sound and Complete Causal Fairness Validation



$\Lambda_\bullet$

$\uparrow \alpha_\bullet$

Λ

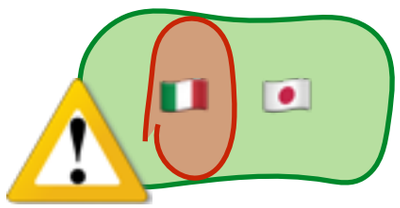


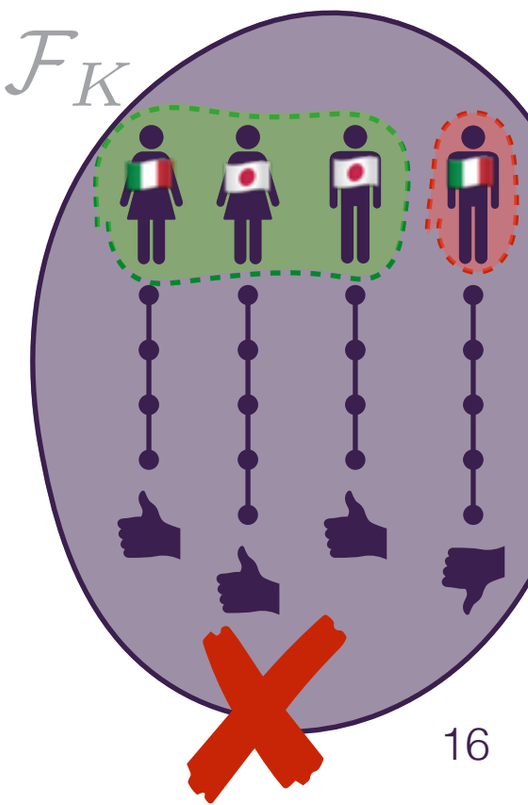
$$P \models \mathcal{F}_K \Leftrightarrow [\![P]\!]_\bullet \subseteq \mathcal{F}_K$$

$$\Leftrightarrow \forall S_1, S_2 \in [\![P]\!]_\bullet :$$

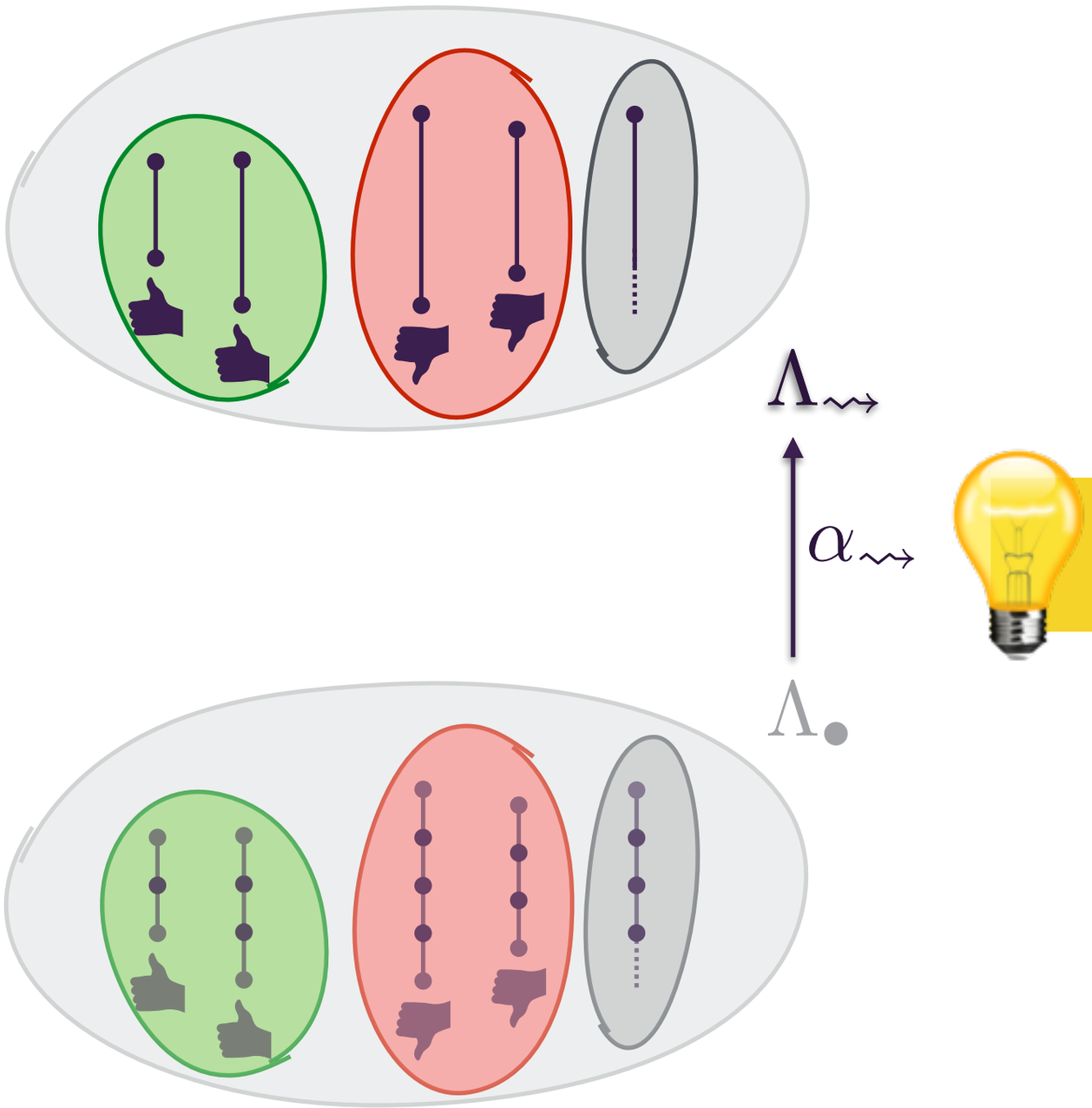
$$S_1[0]|_K \cap S_2[0]|_K = \emptyset$$

$$P \models \mathcal{F}_K \Leftrightarrow [\![P]\!] \in \mathcal{F}_K$$



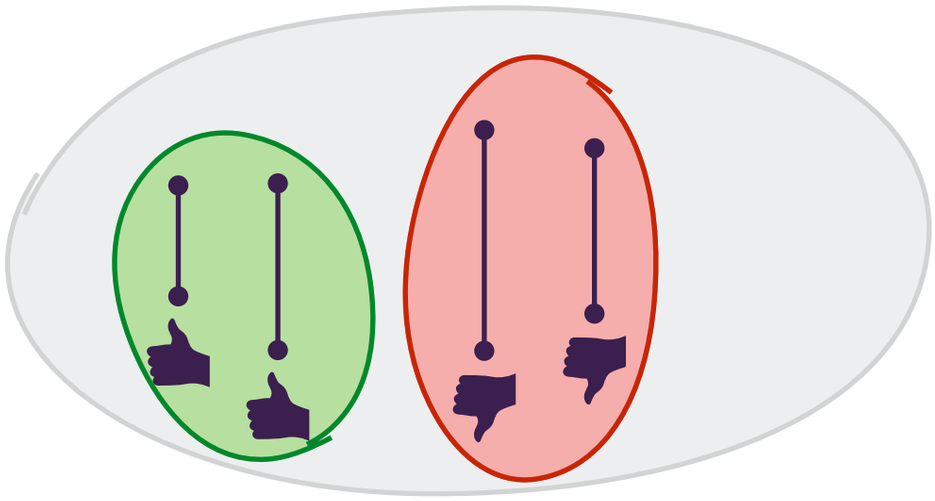


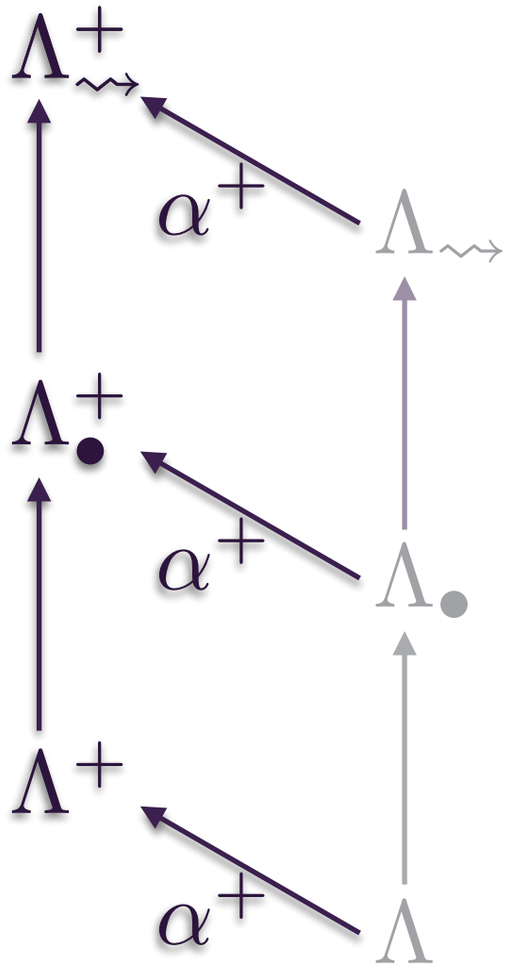
Dependency Abstraction

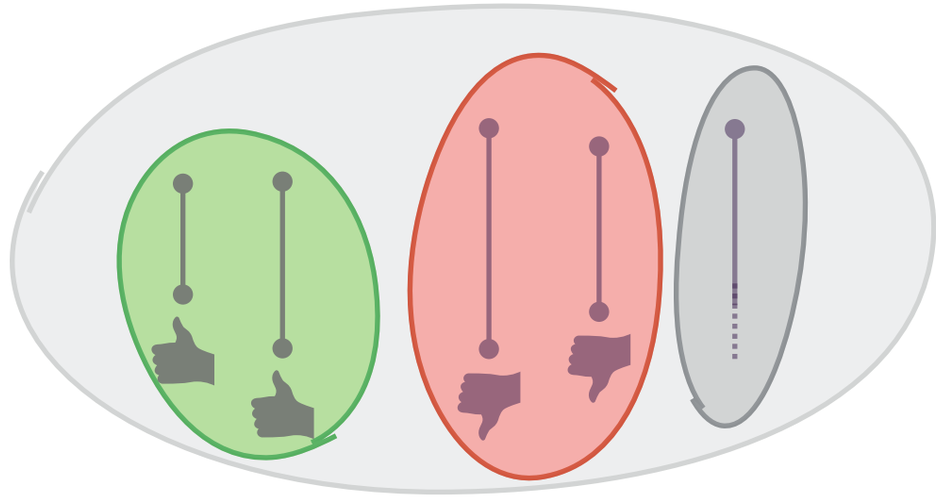


forget intermediate states

Angelic Abstraction





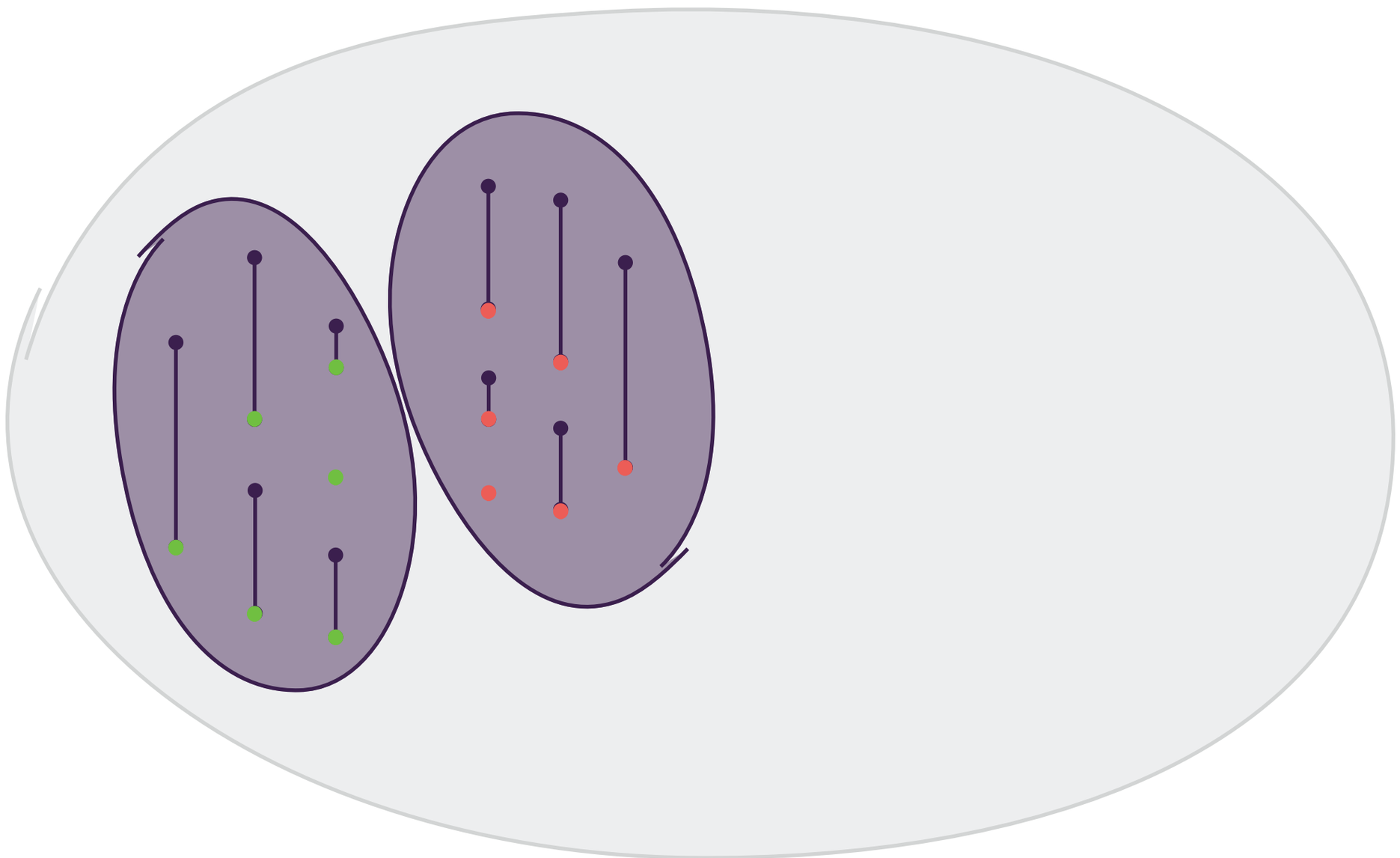


forget infinite executions

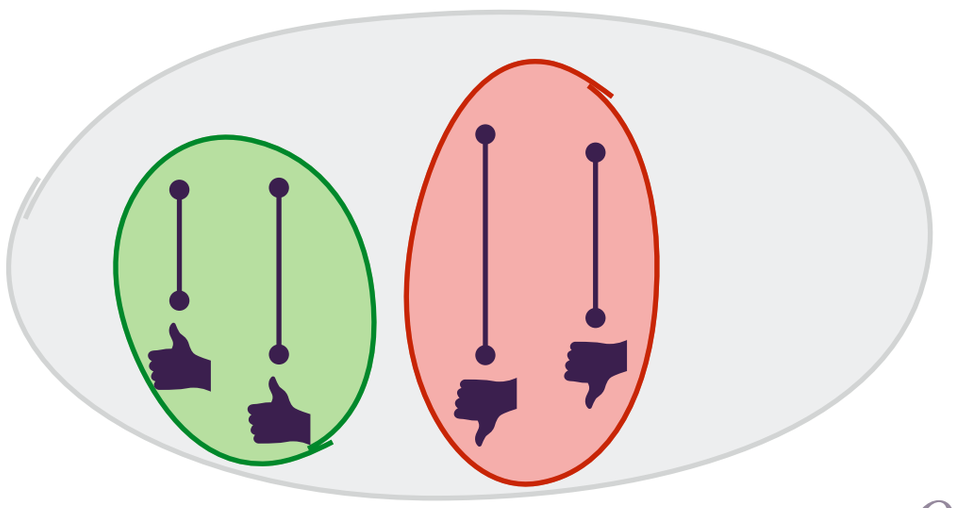
Angelic Dependency Semantics

$$\Lambda^{+}_{\rightsquigarrow} = \mathrm{lfp}^{\sqsubseteq\bullet}_{\{\emptyset\}}\ \Theta^{+}_{\rightsquigarrow}$$

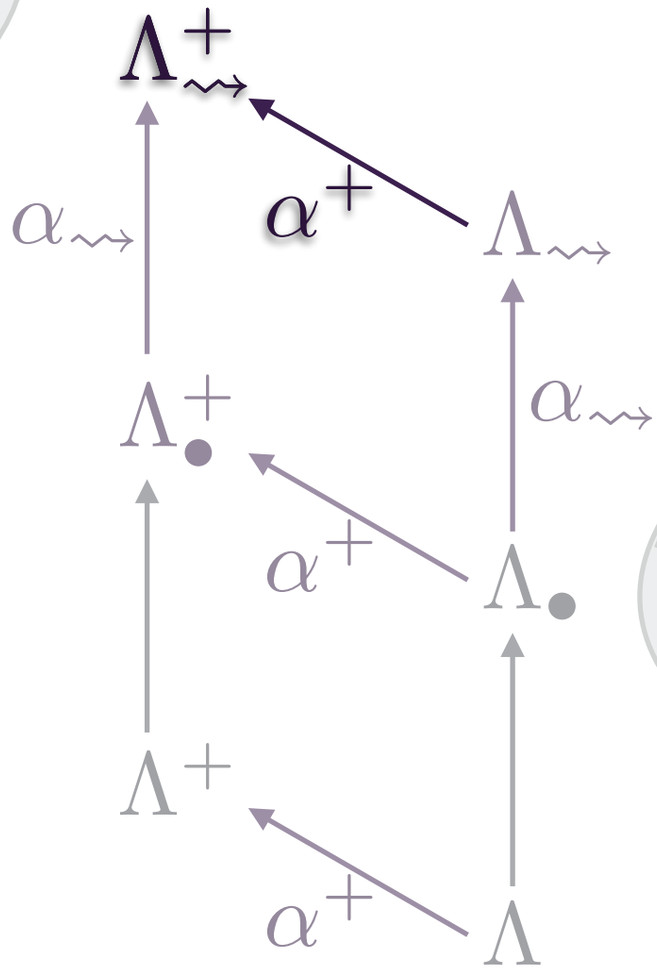
$$\Theta^{+}_{\rightsquigarrow}(S) \stackrel{\mathrm{def}}{=} \{\Omega_o \times \Omega_o \mid o \in O\} \uplus \{\tau \circ R \mid R \in S\}$$

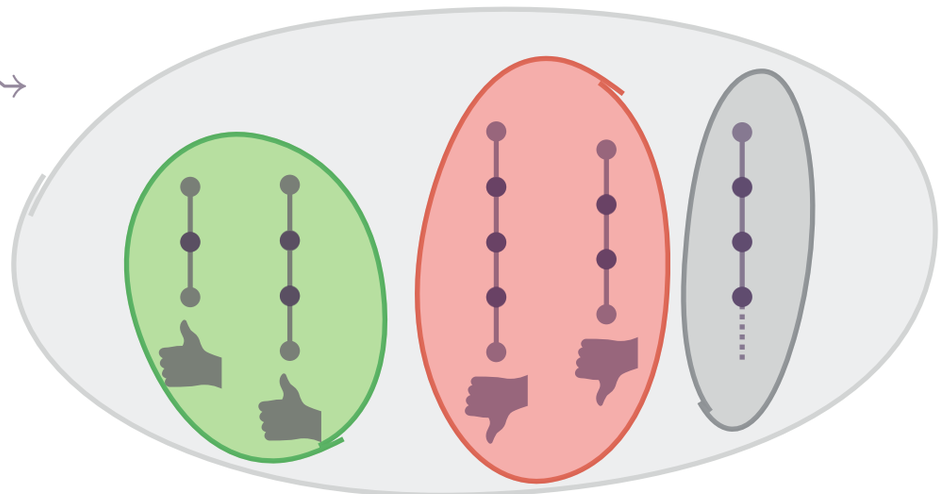


Sound and Complete Causal Fairness Validation



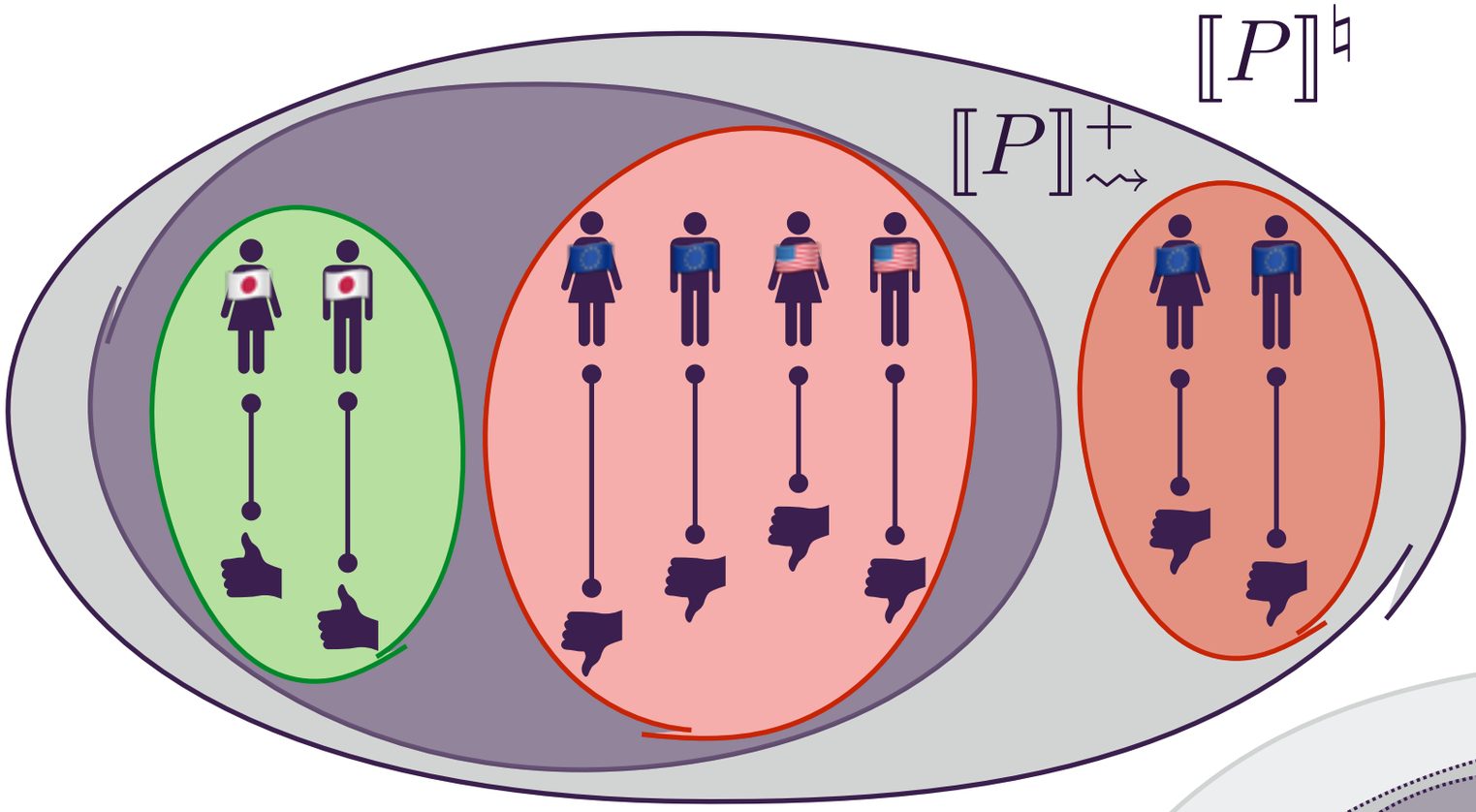
$$P \models \mathcal{F}_K \Leftrightarrow [\![P]\!]_\bullet \subseteq \mathcal{F}_K$$
$$\Leftrightarrow \forall S_1, S_2 \in [\![P]\!]_\bullet :$$
$$S_1[0]|_K \cap S_2[0]|_K = \emptyset$$

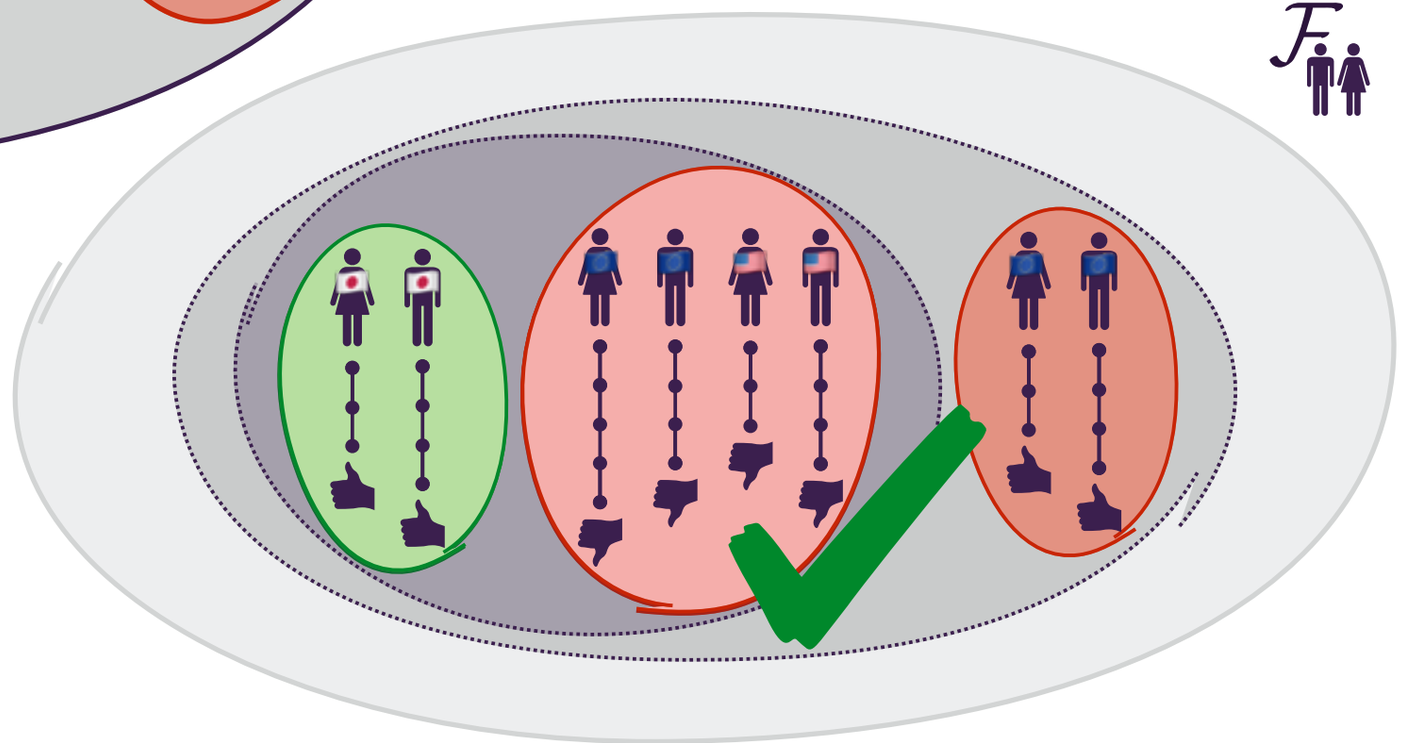




$$P \models \mathcal{F}_K$$
$$\Leftrightarrow \gamma^+_{\rightsquigarrow}([\![P]\!]^+_{\rightsquigarrow}) \subseteq \mathcal{F}_K$$
$$\Leftrightarrow \forall S_1, S_2 \in [\![P]\!]^+_{\rightsquigarrow} :$$
$$S_1[0]|_K \cap S_2[0]|_K = \emptyset$$

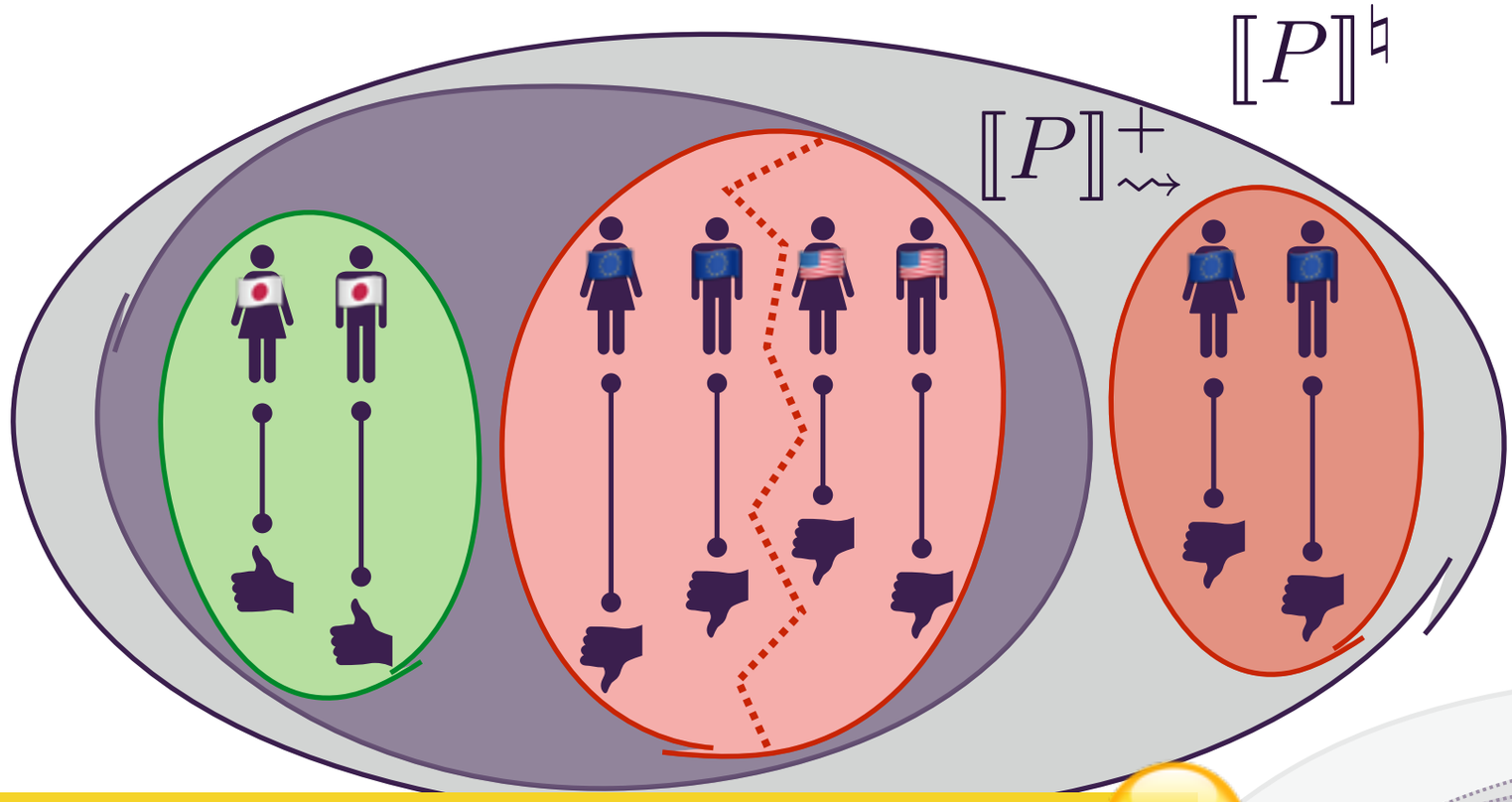
Sound Causal Fairness Validation





$$\forall S_1, S_2 \in [\![P]\!]^{\natural}:$$
$$(S_1[\omega] \neq S_2[\omega] \Rightarrow$$
$$S_1[0]|_K \cap S_2[0]|_K = \emptyset)$$
$$\Leftrightarrow \gamma^{+}_{\rightsquigarrow}([\![P]\!]^{\natural}) \subseteq \mathcal{F}_K \Rightarrow \gamma^{+}_{\rightsquigarrow}([\![P]\!]^{+}_{\rightsquigarrow}) \subseteq \mathcal{F}_K \Rightarrow P \models \mathcal{F}_K$$

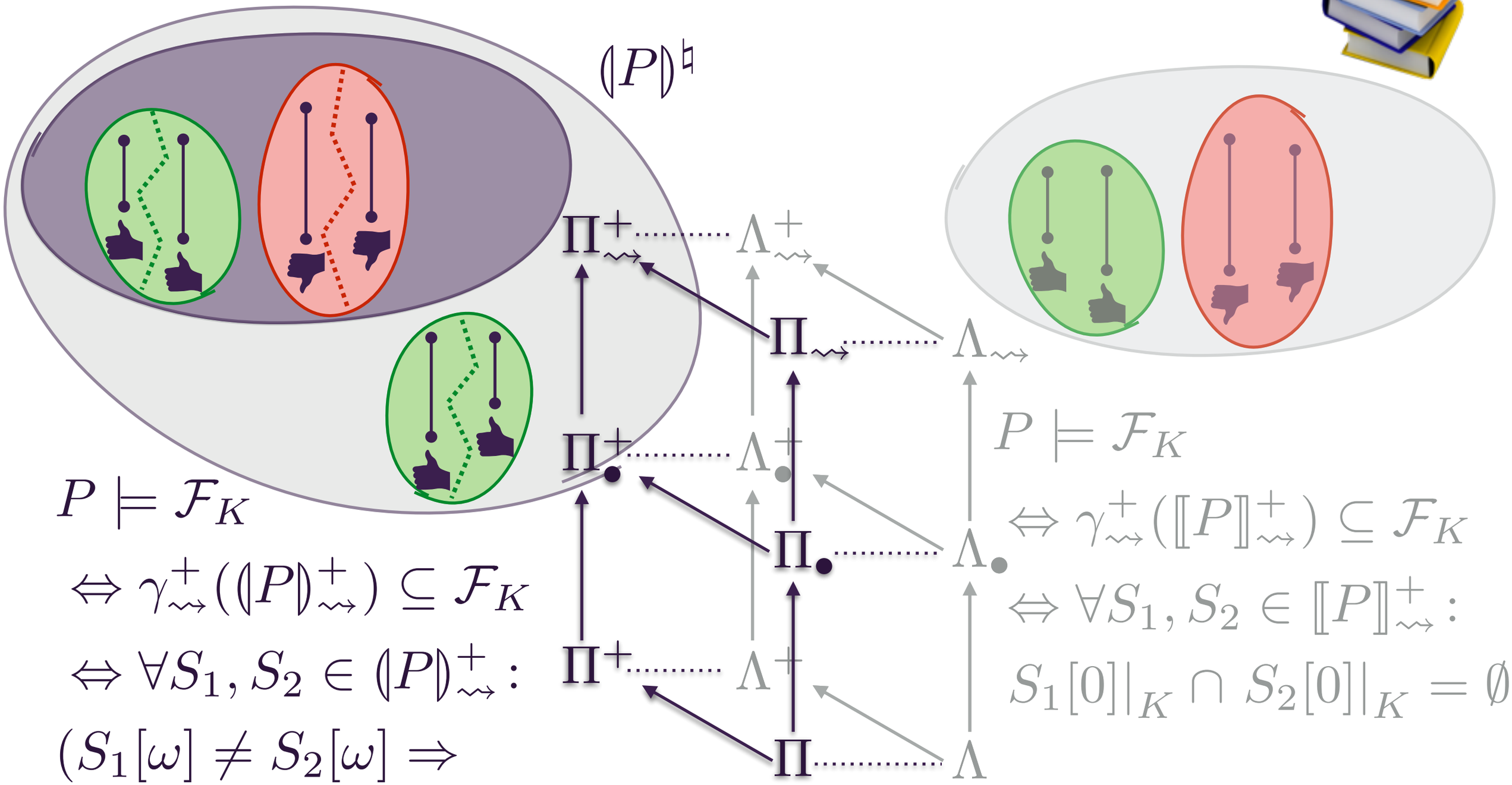
Sound Causal Fairness Validation



partition with respect to non-sensitive inputs

$$\forall S_1, S_2 \in [\![P]\!]^{\natural}:$$
$$(S_1[\omega] \neq S_2[\omega] \Rightarrow$$
$$S_1[0]|_K \cap S_2[0]|_K = \emptyset)$$
$$\Leftrightarrow \gamma^{+}_{\rightsquigarrow}([\![P]\!]^{\natural}) \subseteq \mathcal{F}_K \Rightarrow \gamma^{+}_{\rightsquigarrow}([\![P]\!]^{+}_{\rightsquigarrow}) \subseteq \mathcal{F}_K \Rightarrow P \models \mathcal{F}_K$$

Parallel Semantics and Validation



$P \models \mathcal{F}_K$

$\Leftrightarrow \gamma^+_{\rightsquigarrow}((\!|P|\!)^+_{\rightsquigarrow}) \subseteq \mathcal{F}_K$

$\Leftrightarrow \forall S_1, S_2 \in (\!|P|\!)^+_{\rightsquigarrow}:$

$(S_1[\omega] \neq S_2[\omega] \Rightarrow$

$S_1[0]|_K \cap S_2[0]|_K = \emptyset) \Leftarrow \forall S_1, S_2 \in (\!|P|\!)^\natural: (S_1[\omega] \neq S_2[\omega]$

$\Rightarrow S_1[0]|_K \cap S_2[0]|_K = \emptyset)$

$P \models \mathcal{F}_K$

$\Leftrightarrow \gamma^+_{\rightsquigarrow}([\![P]\!]^+_{\rightsquigarrow}) \subseteq \mathcal{F}_K$

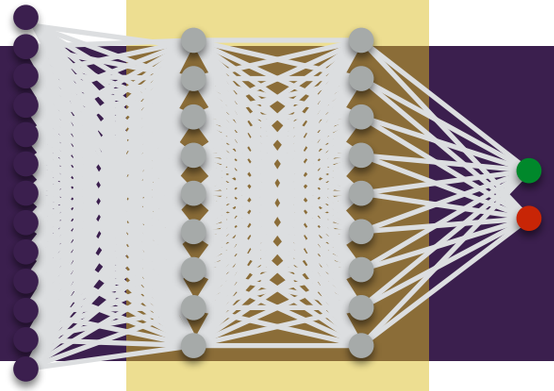
$\Leftrightarrow \forall S_1, S_2 \in [\![P]\!]^+_{\rightsquigarrow}:$

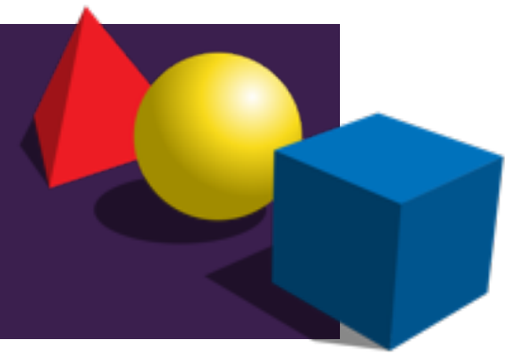
$S_1[0]|_K \cap S_2[0]|_K = \emptyset$

Causal Fairness Analysis

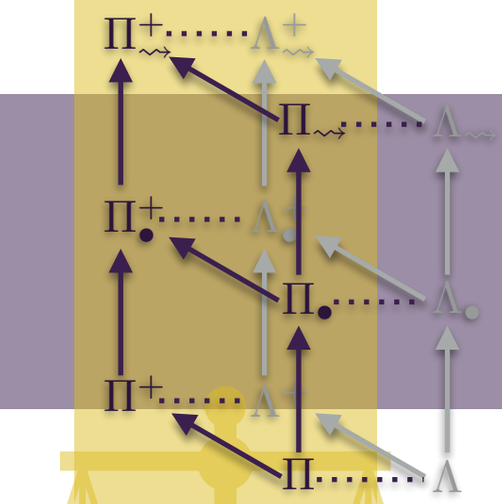
practical tools
targeting specific programs

algorithmic approaches
to decide program properties

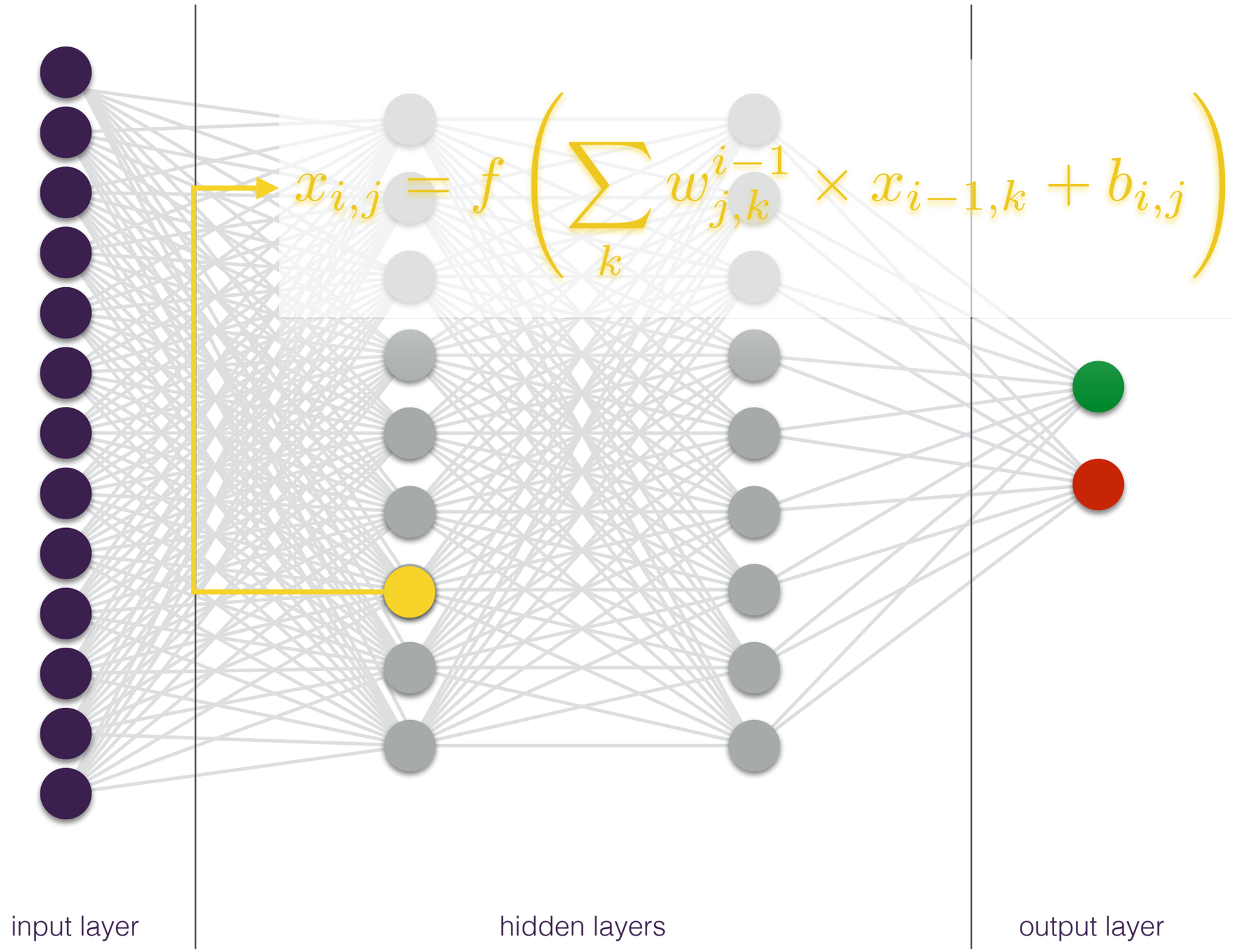




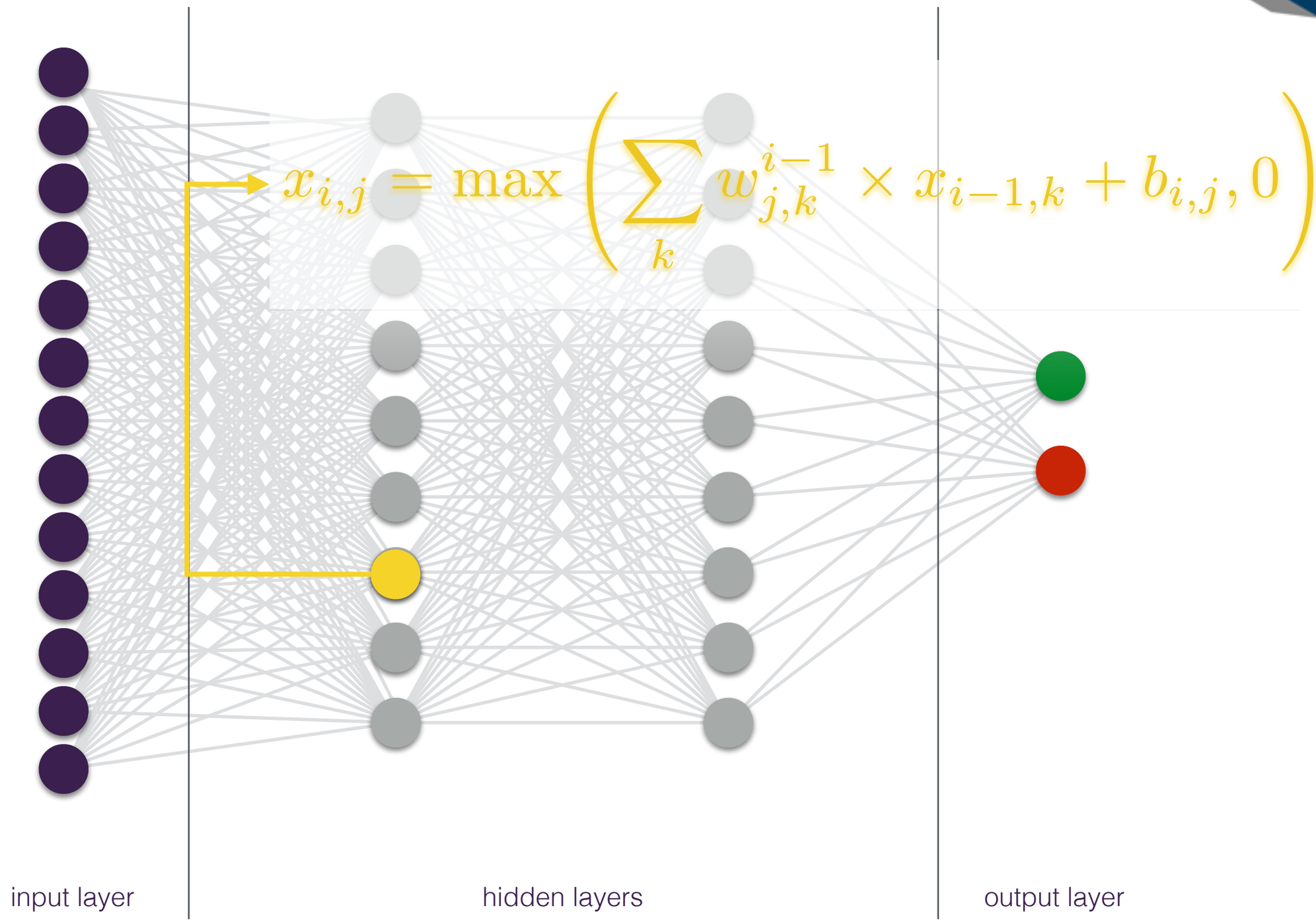
mathematical models
of the program behavior



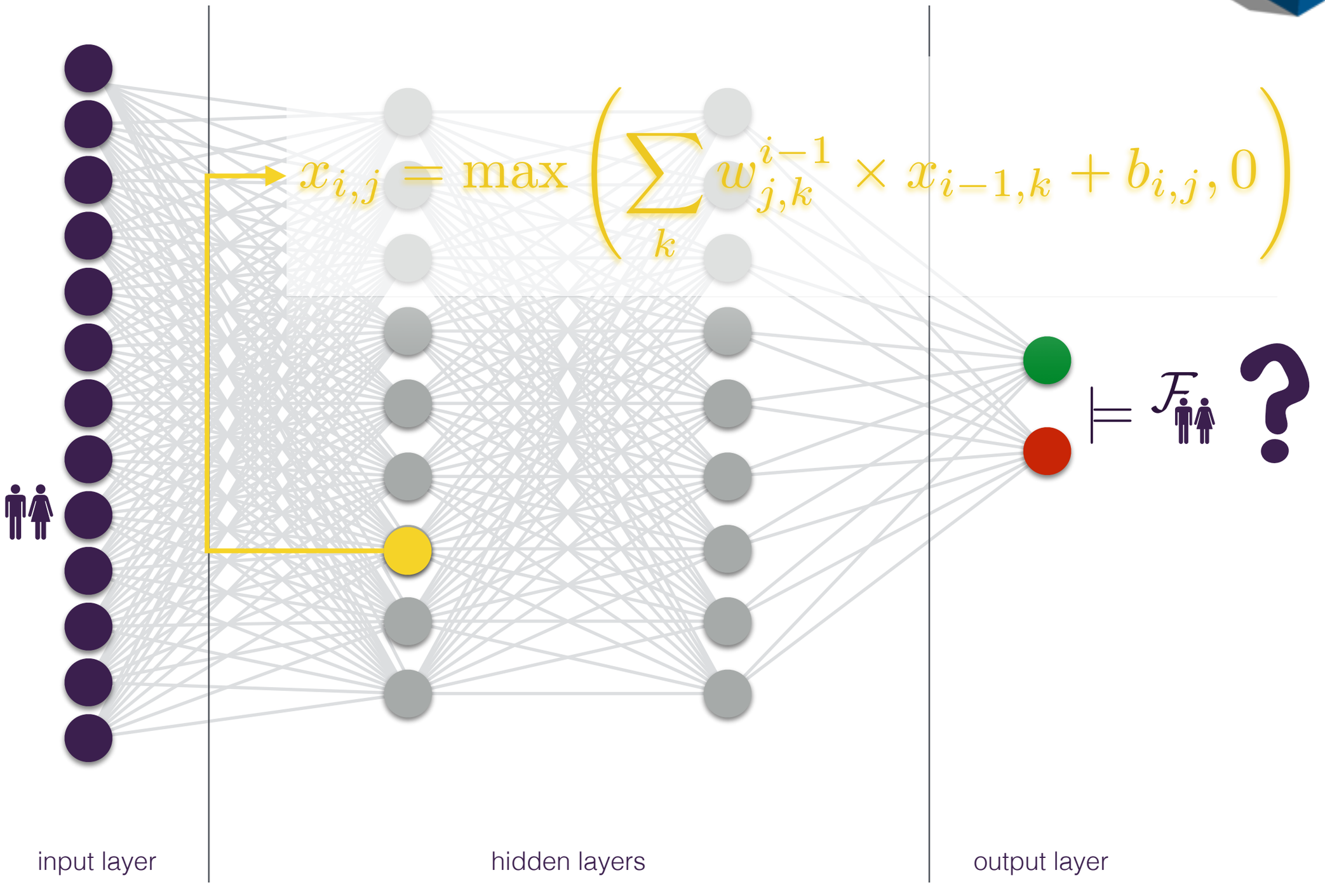
Feed-Forward Neural Networks



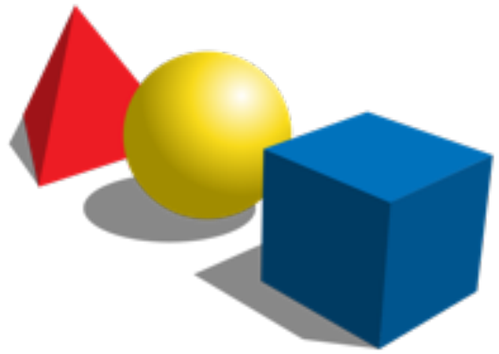
Feed-Forward Neural Networks

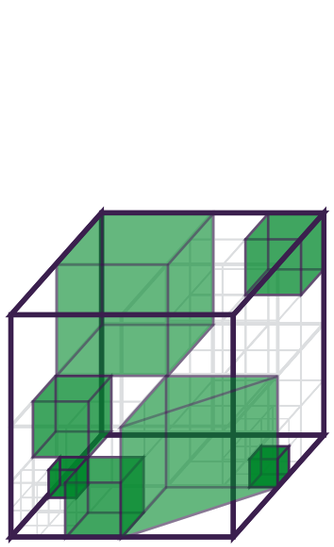


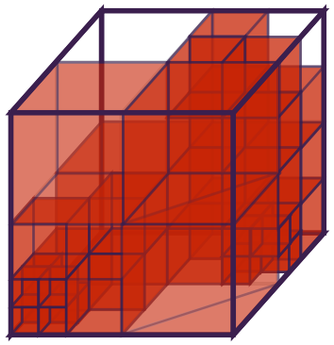
Feed-Forward Neural Networks

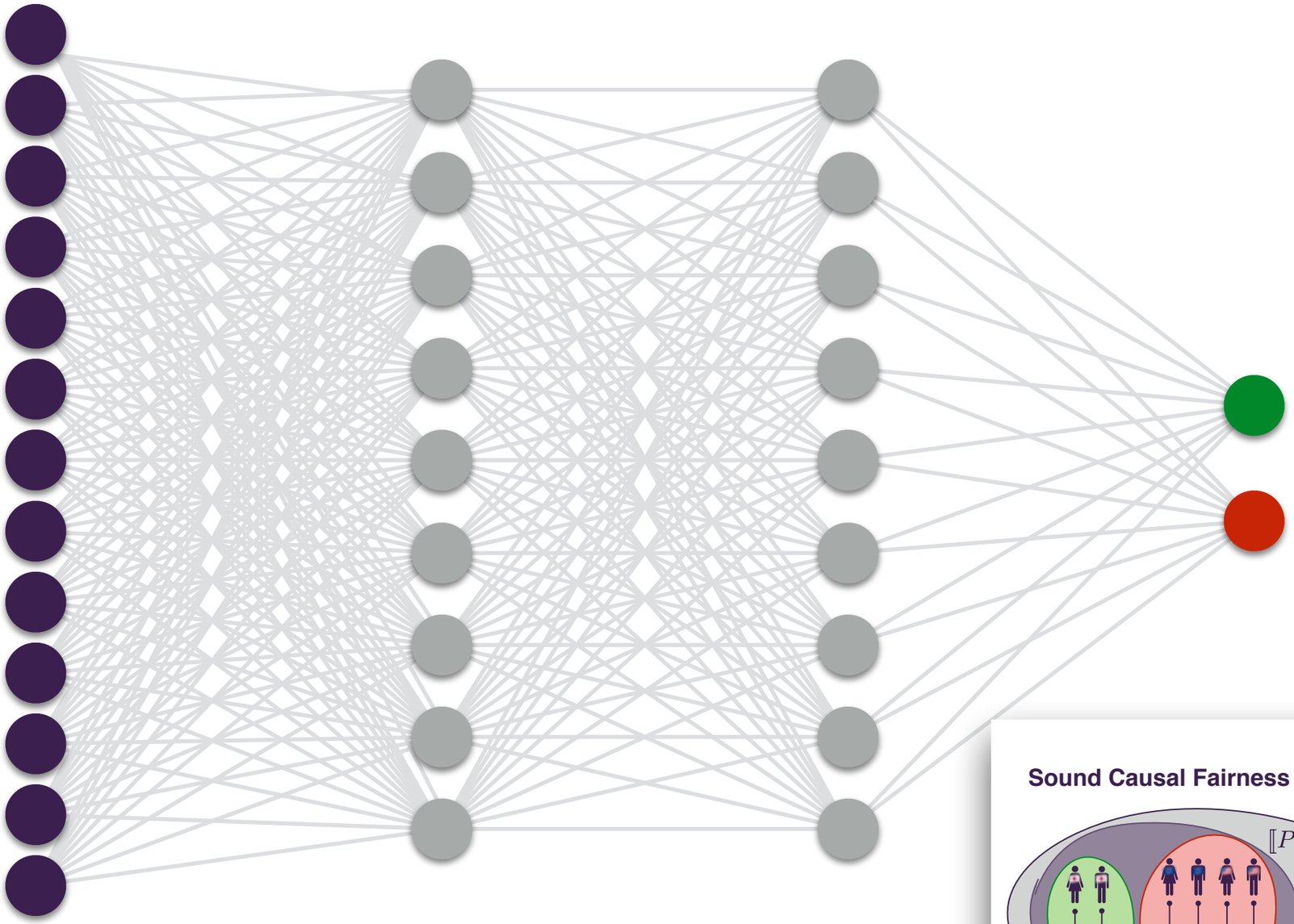


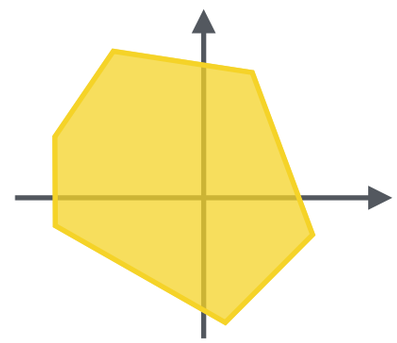
Naïve Backward Analysis



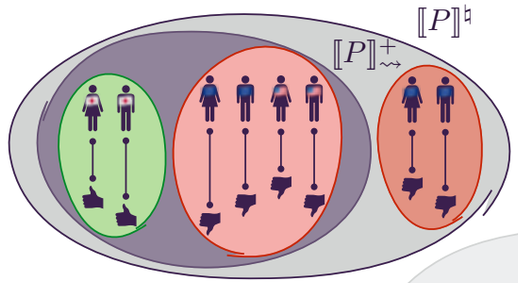


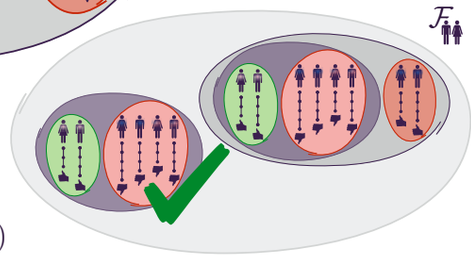






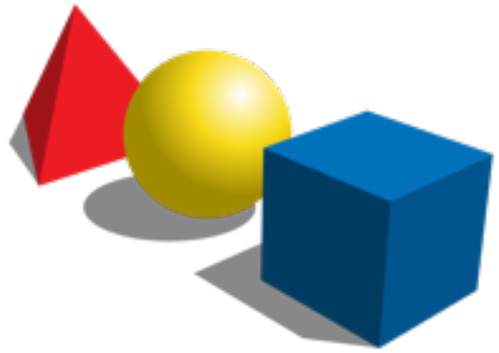
Sound Causal Fairness Validation

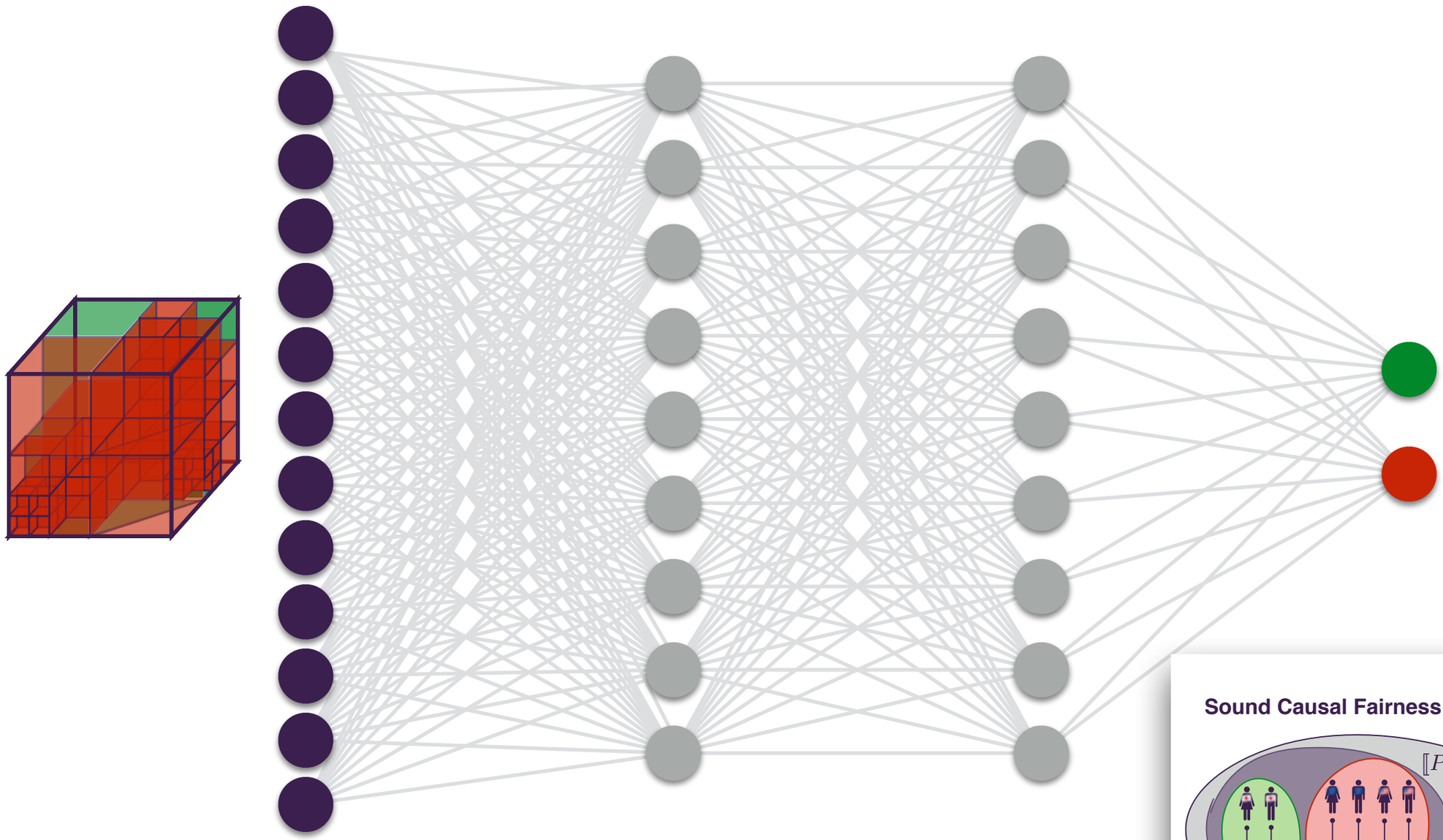


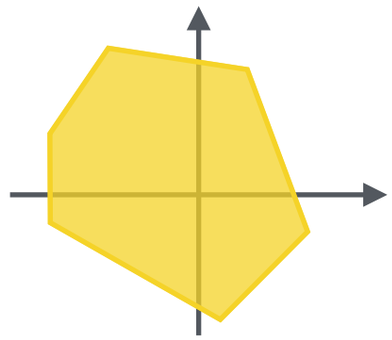


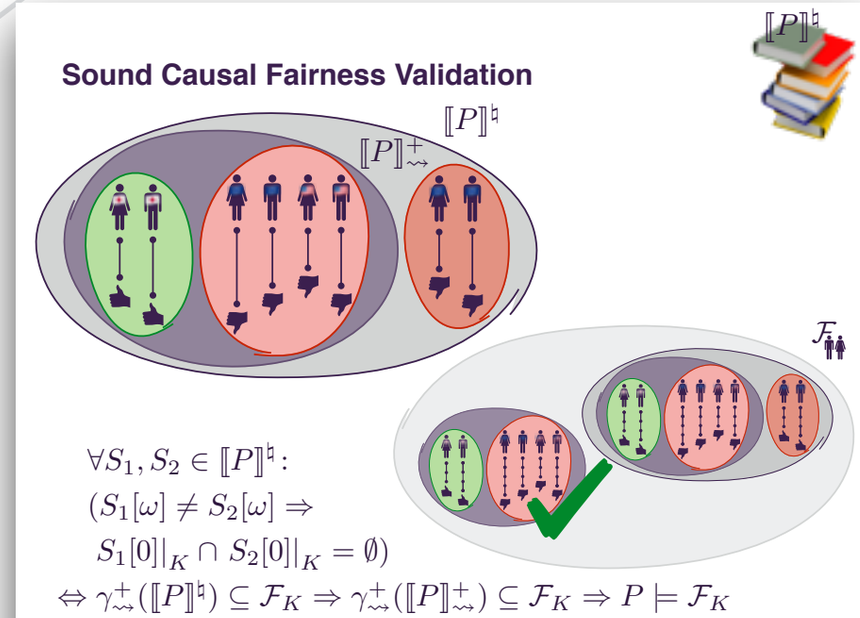
$\forall S_1, S_2 \in [\![P]\!]^{\natural}$:
$(S_1[\omega] \neq S_2[\omega] \Rightarrow$
$S_1[0]|_K \cap S_2[0]|_K = \emptyset)$
$\Leftrightarrow \gamma^{+}_{\rightsquigarrow}([\![P]\!]^{\natural}) \subseteq \mathcal{F}_K \Rightarrow \gamma^{+}_{\rightsquigarrow}([\![P]\!]^{+}_{\rightsquigarrow}) \subseteq \mathcal{F}_K \Rightarrow P \models \mathcal{F}_K$

Naïve Backward Analysis

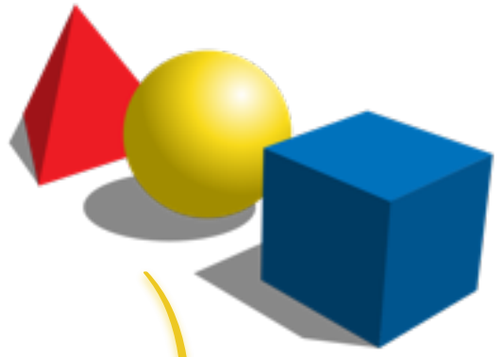




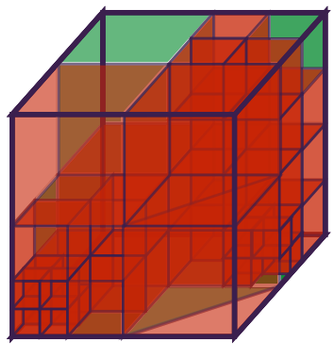


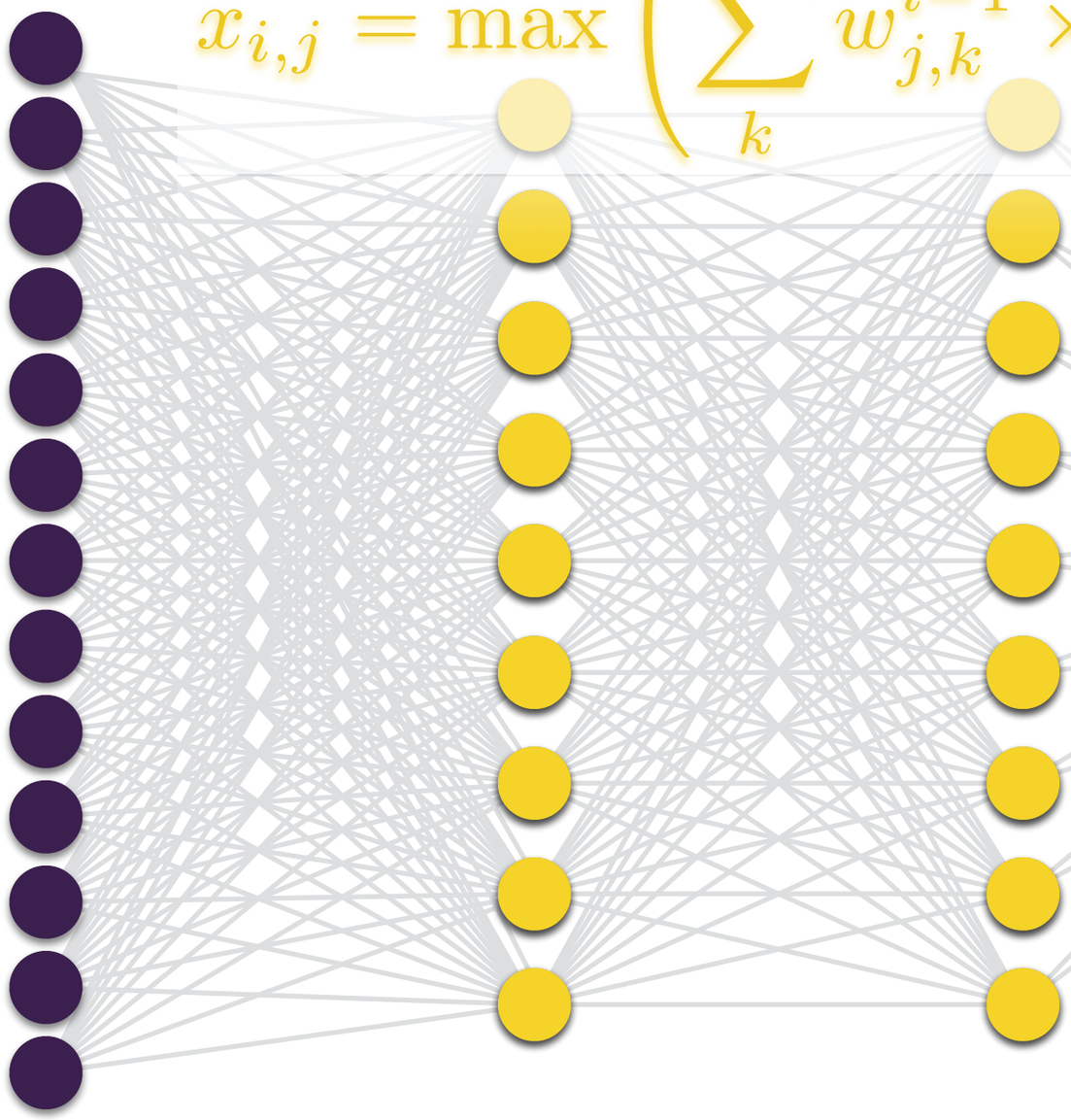


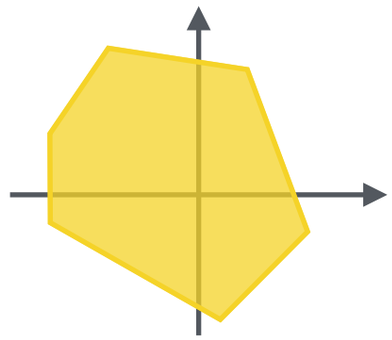
Naïve Backward Analysis



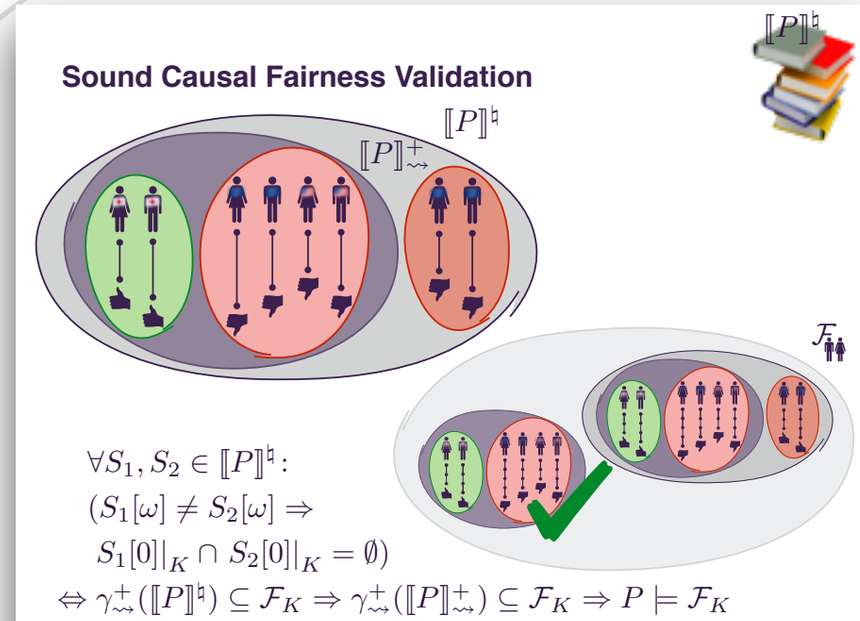
$$x_{i,j} = \max\left(\sum_k w_{j,k}^{i-1} \times x_{i-1,k} + b_{i,j}, 0\right)$$



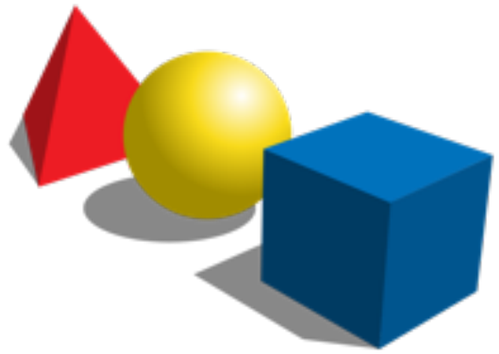


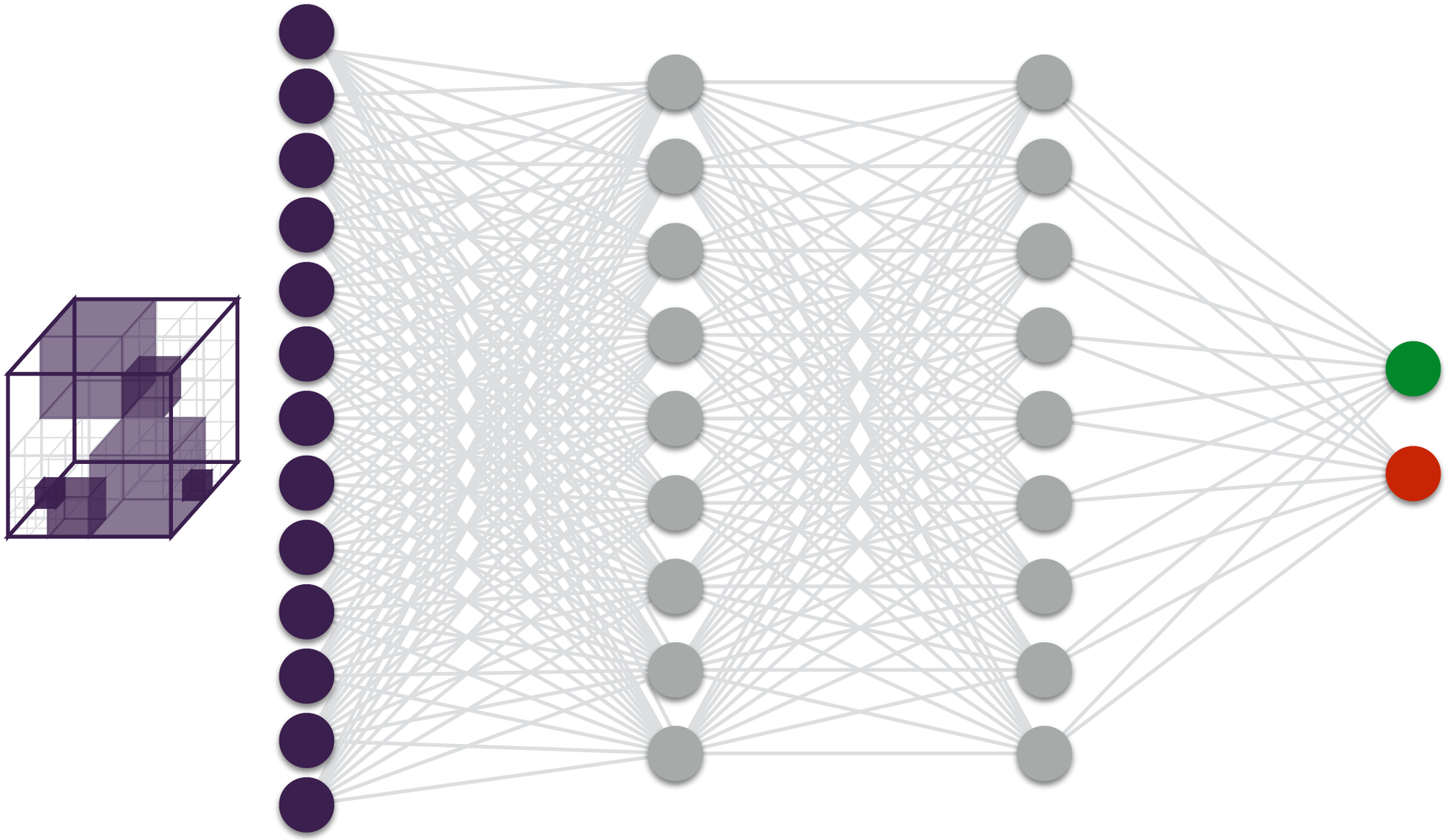


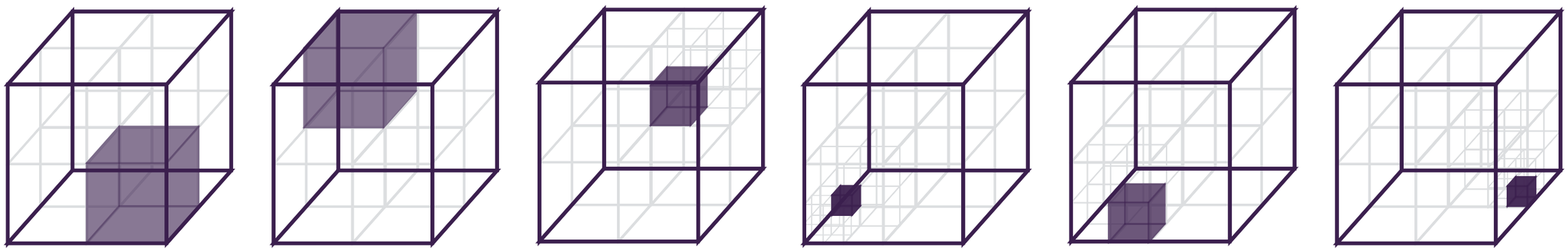
too many disjunctions!



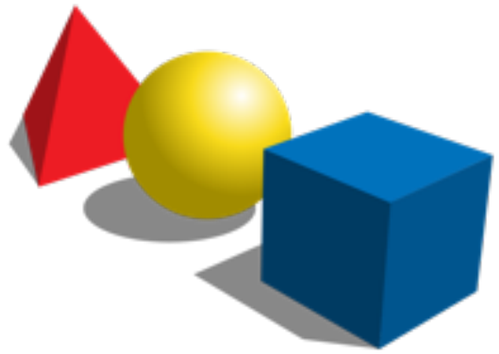
Forward and Backward Analysis

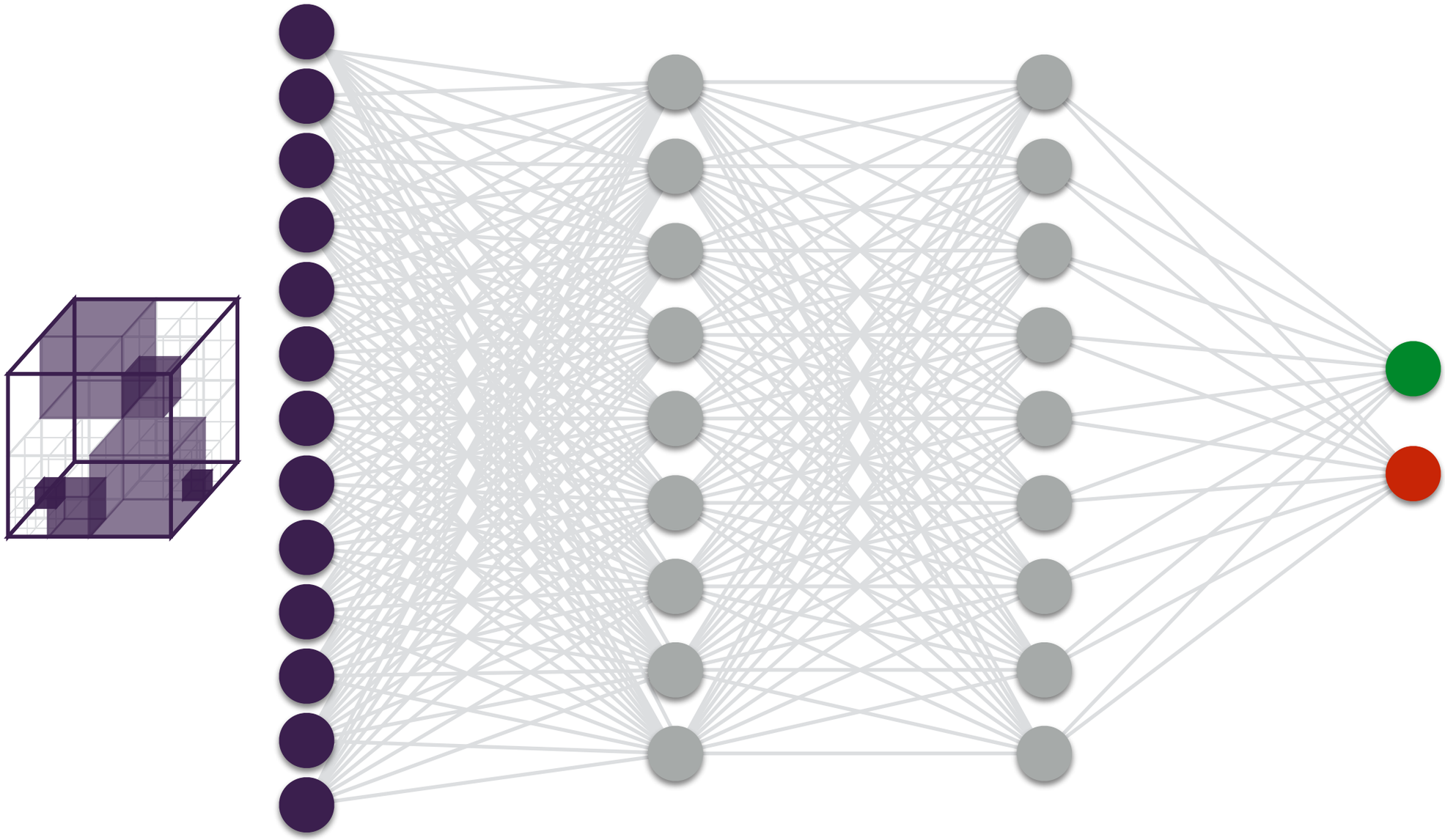


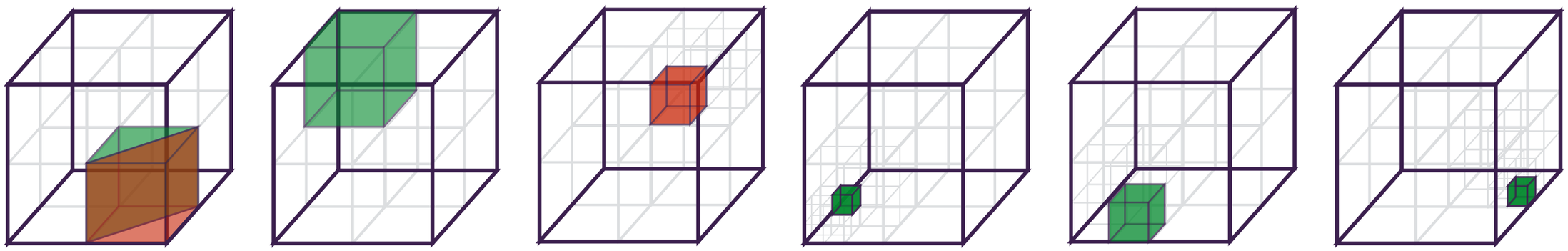




Forward and Backward Analysis

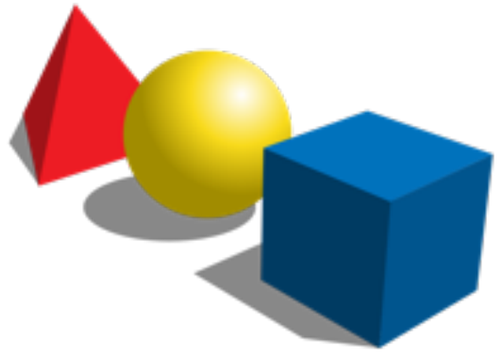


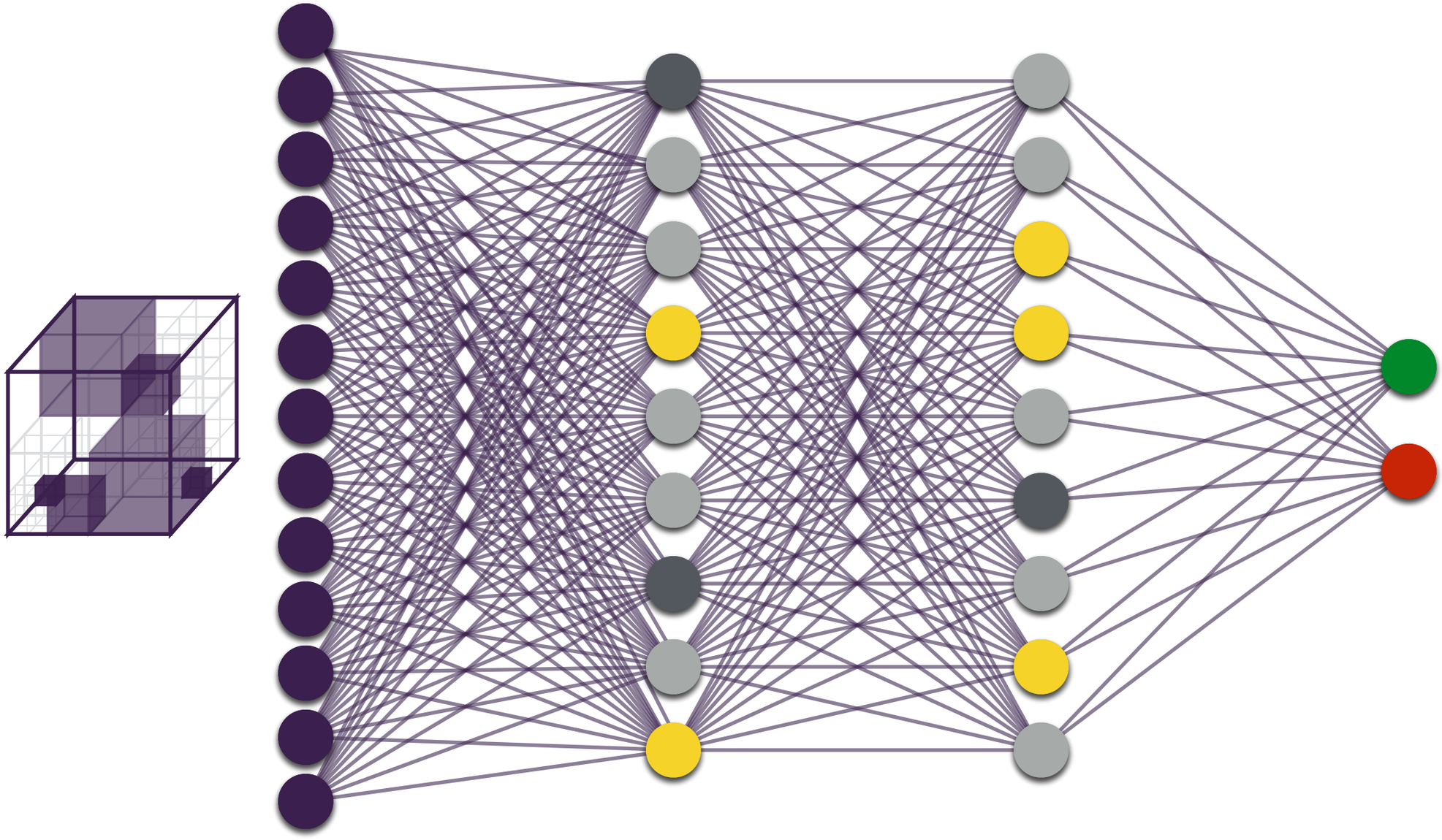


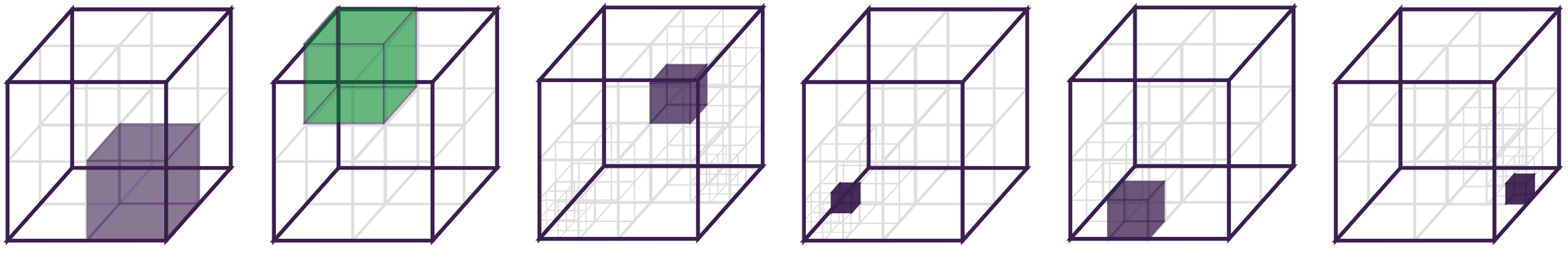


Forward and Backward Analysis

A Better Solution

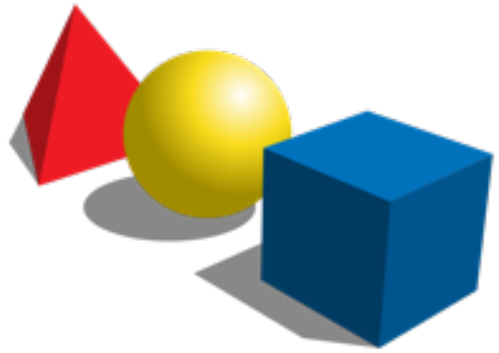


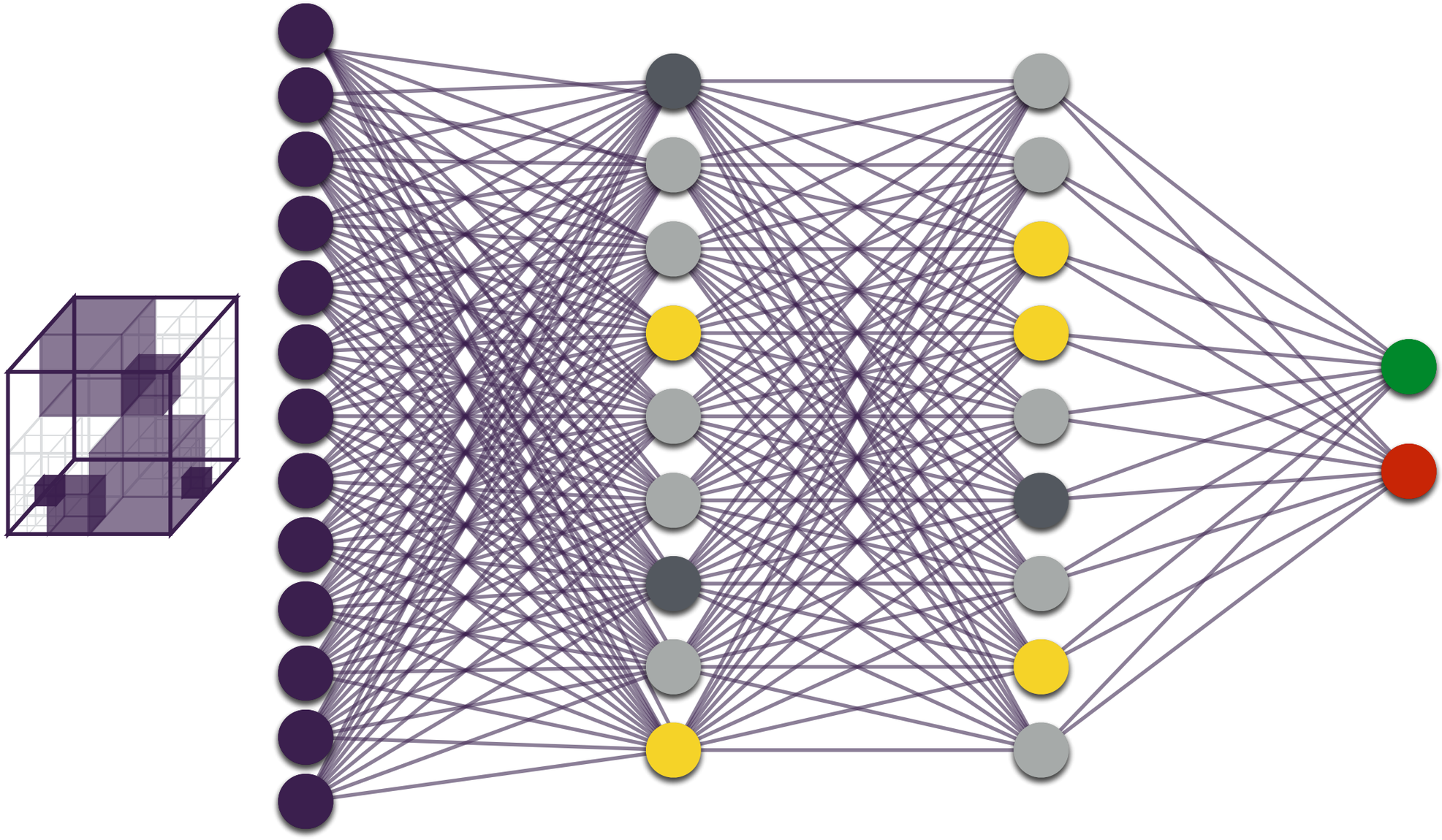


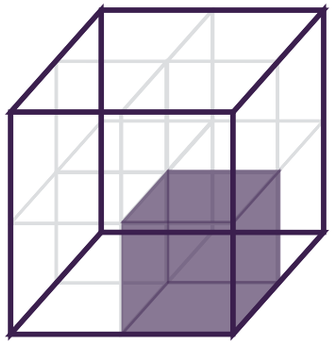


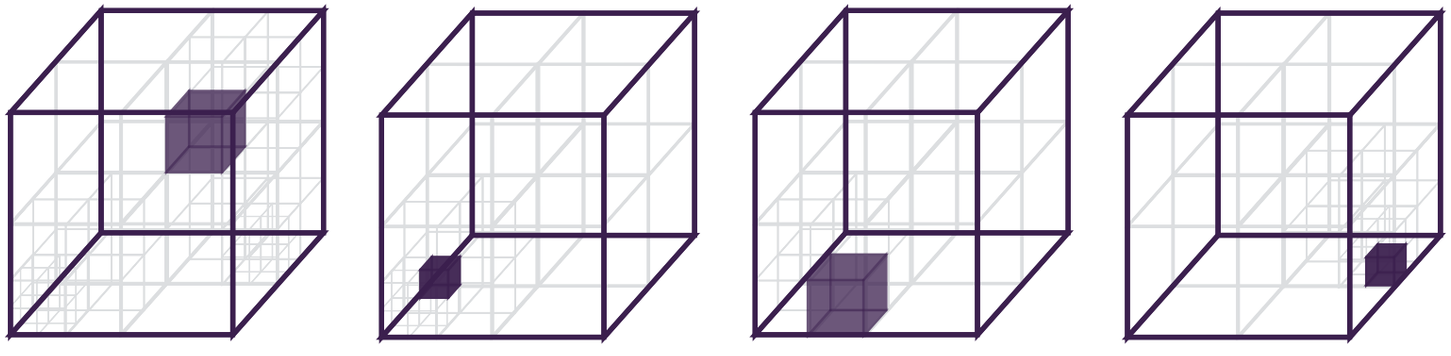
Forward and Backward Analysis

A Better Solution



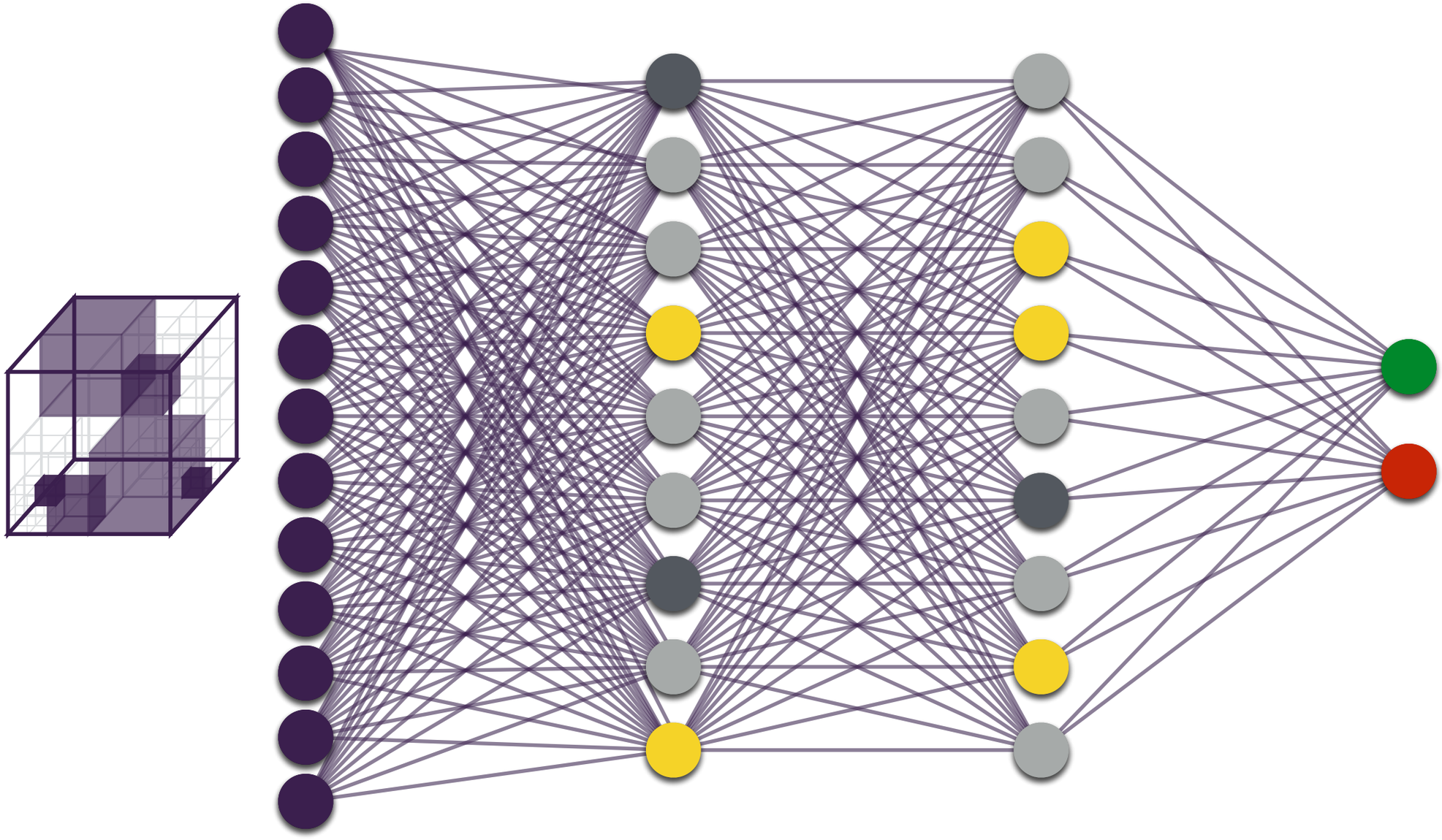


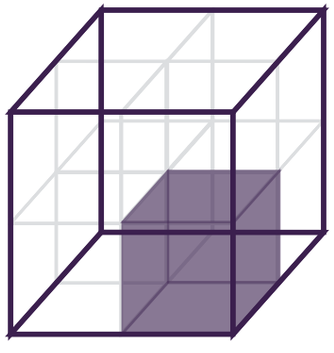


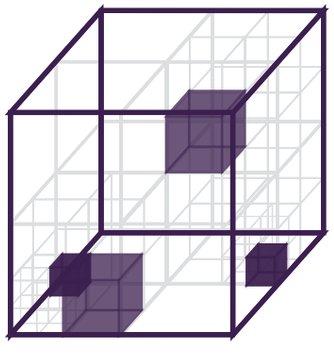


Forward and Backward Analysis

A Better Solution

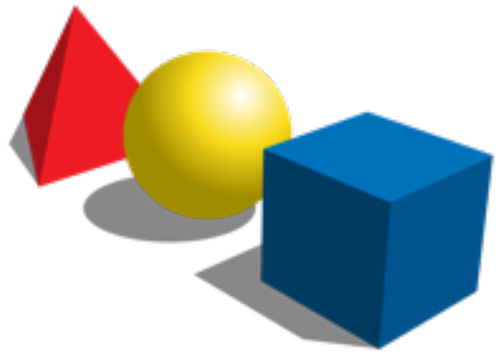


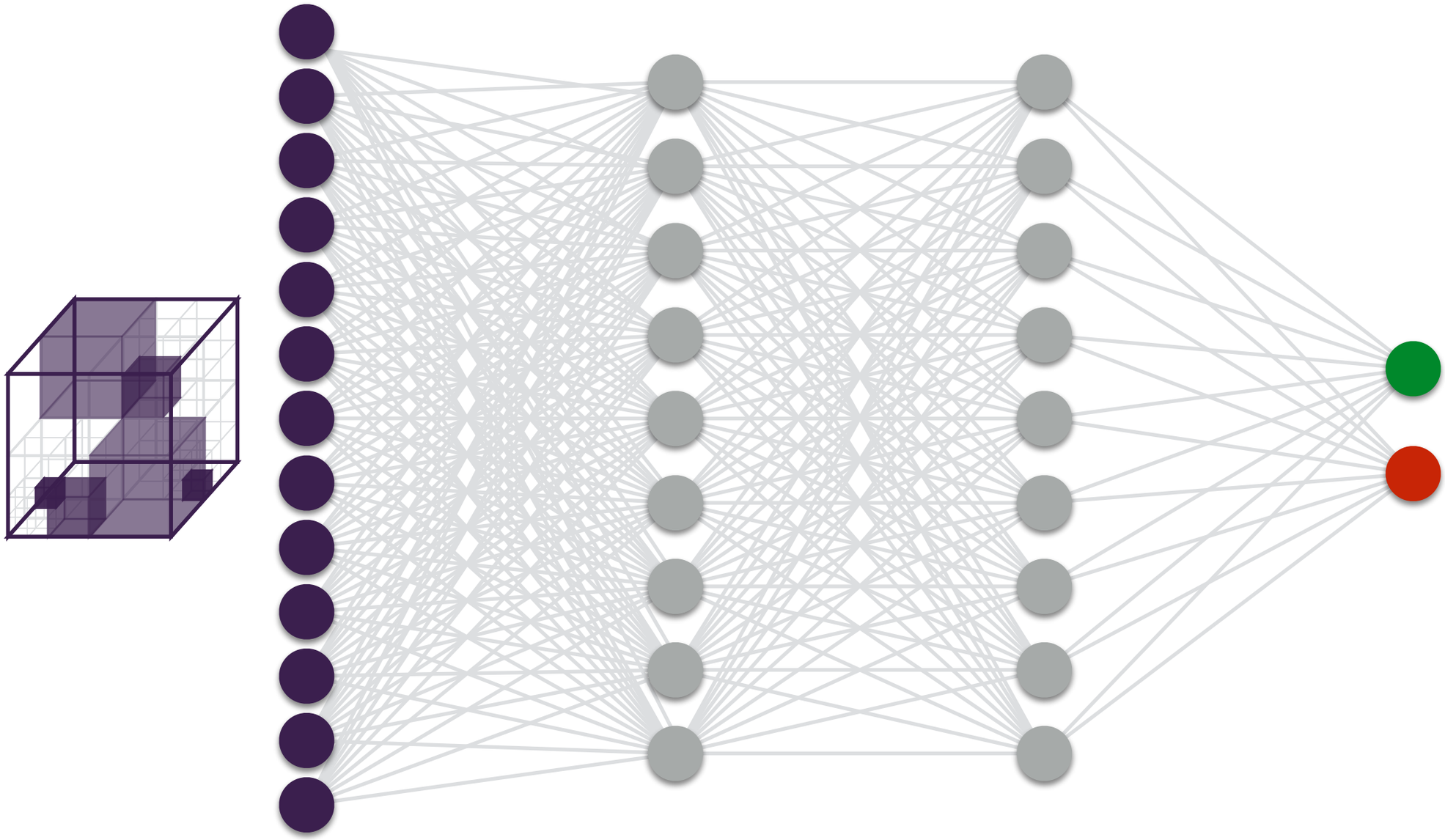


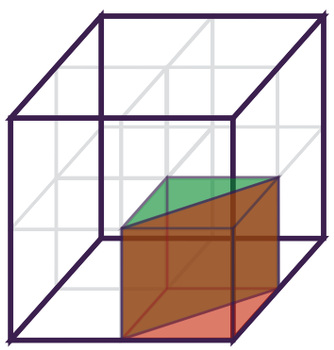


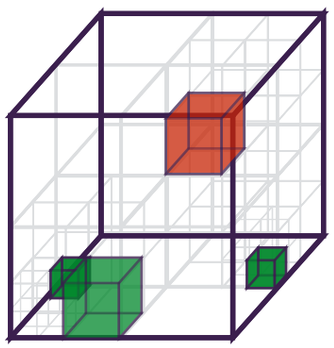
Forward and Backward Analysis

A Better Solution





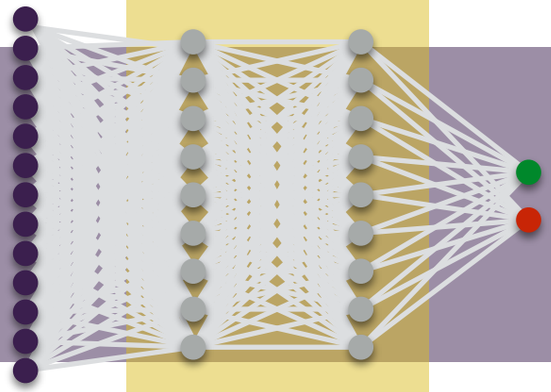


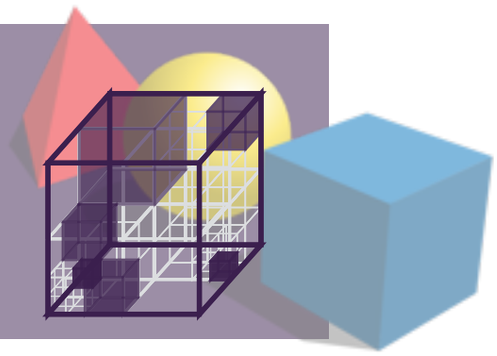


Implementation

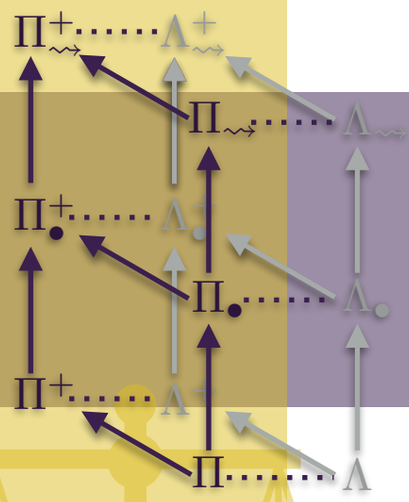
practical tools
targeting specific programs

algorithmic approaches
to decide program properties





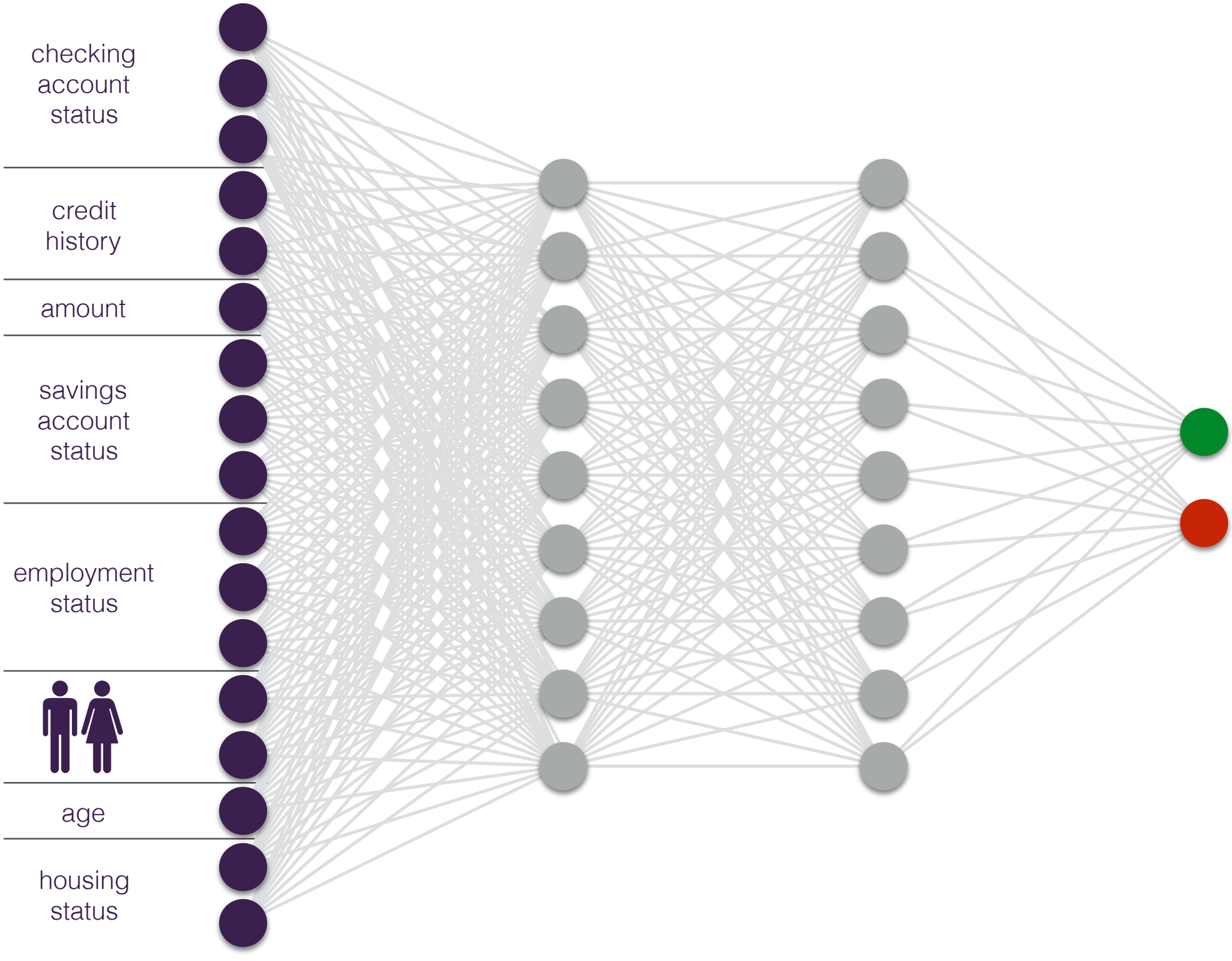
mathematical models
of the program behavior



https://github.com/caterinaurban/Libra

Experiments

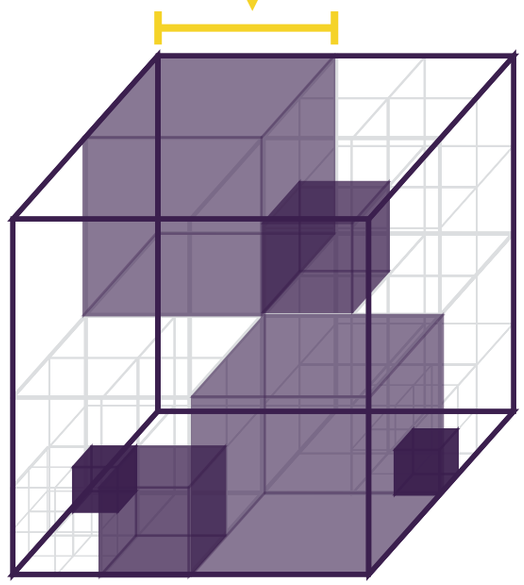
German Credit Screening

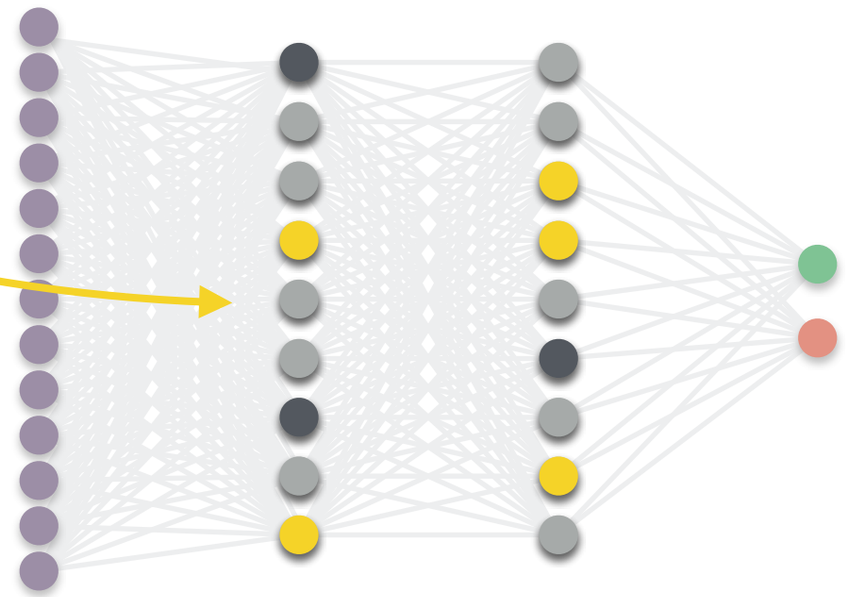


Experiments

German Credit Screening

		nosym			sym		
p	d	precise	bias	time	precise	bias	time
0.5	3						
	5						
	7						
	9						
0.25	3						
	5						
	7						
	9						
0.125	3						
	5						
	7						
	9						

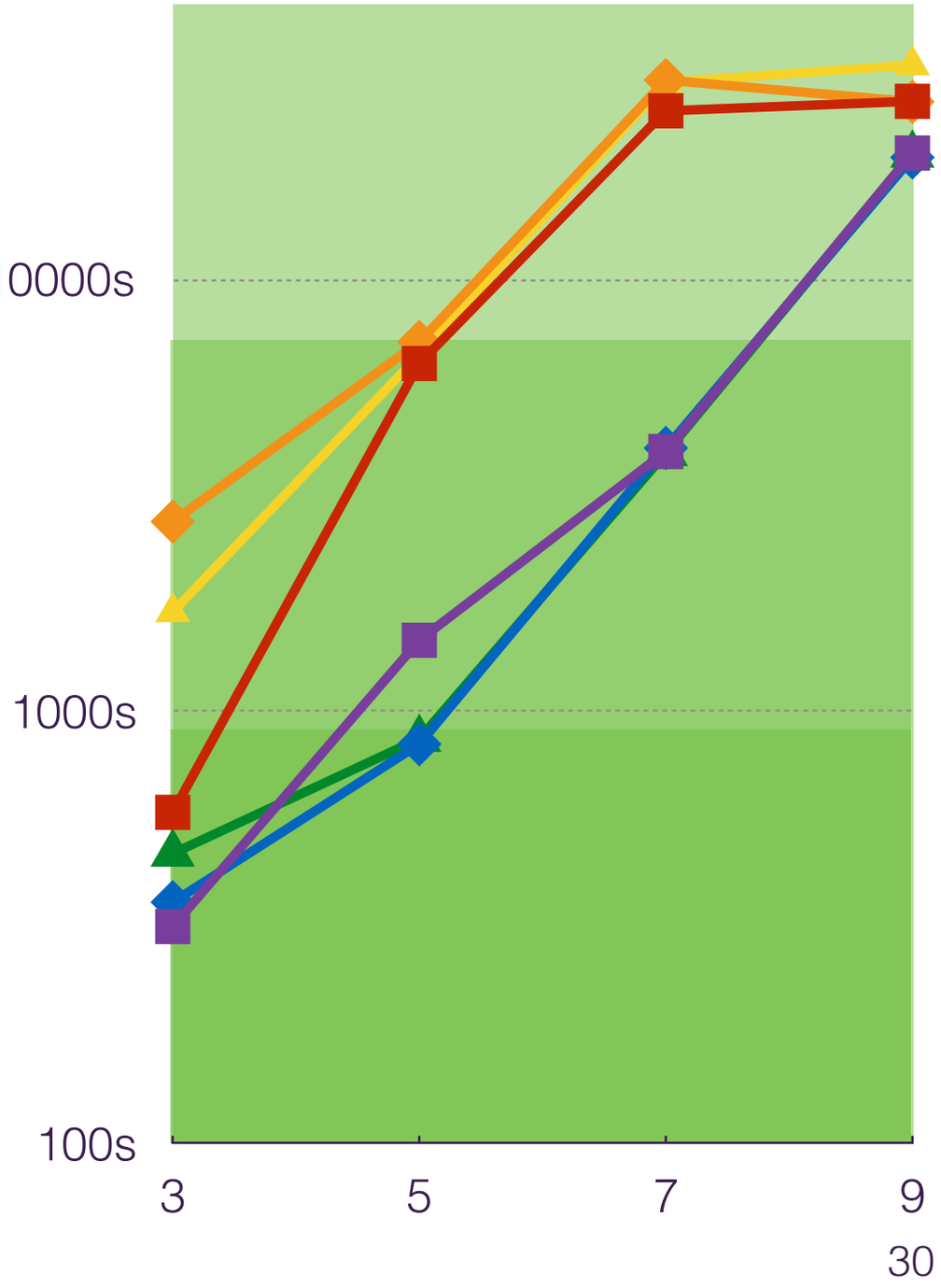




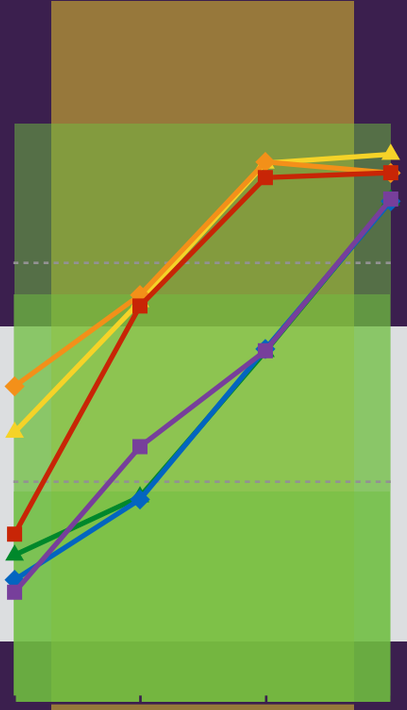
Experiments

German Credit Screening

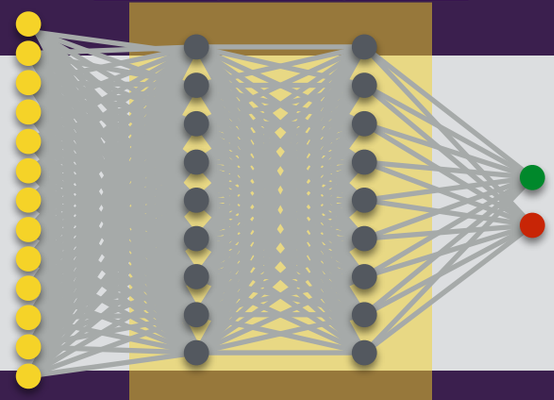
p	d	nosym			sym		
		precise	bias	time	precise	bias	time
0.5	3	50,00%	0,47%	0h 10m	65,74%	1,59%	0h 5m
	5	88,89%	4,18%	1h 47m	93,29%	4,41%	0h 24m
	7	98,38%	5,48%	6h 50m	99,31%	5,74%	1h 7m
	9	100,00%	6,48%	7h 11m	100,00%	6,48%	5h 27m
0.25	3	89,35%	3,46%	0h 46m	94,33%	4,31%	0h 6m
	5	99,59%	5,31%	2h 0m	99,71%	5,75%	0h 14m
	7	100,00%	5,48%	8h 3m	100,00%	5,74%	1h 8m
	9	100,00%	6,48%	7h 10m	100,00%	6,48%	5h 20m
0.125	3	98,84%	4,43%	0h 28m	99,45%	4,88%	0h 8m
	5	99,99%	5,35%	1h 52m	99,99%	5,78%	0h 14m
	7	100,00%	6,20%	7h 57m	100,00%	5,74%	1h 6m
	9	100,00%	6,71%	8h 43m	100,00%	6,48%	5h 25m

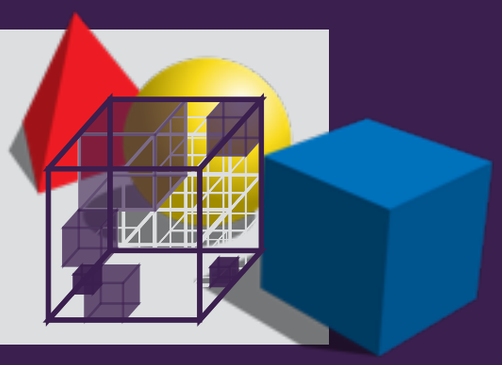


practical tools
targeting specific programs

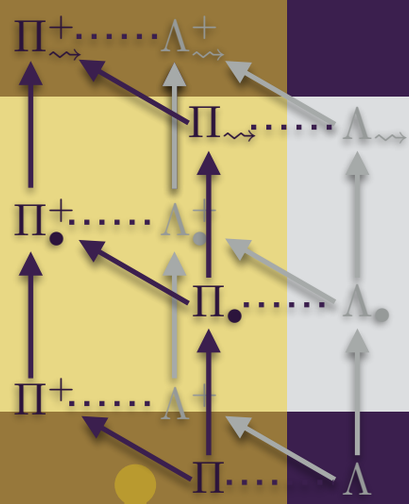


algorithmic approaches
to decide program properties





mathematical models
of the program behavior



https://github.com/caterinaurban/Libra

QUESTIONS?