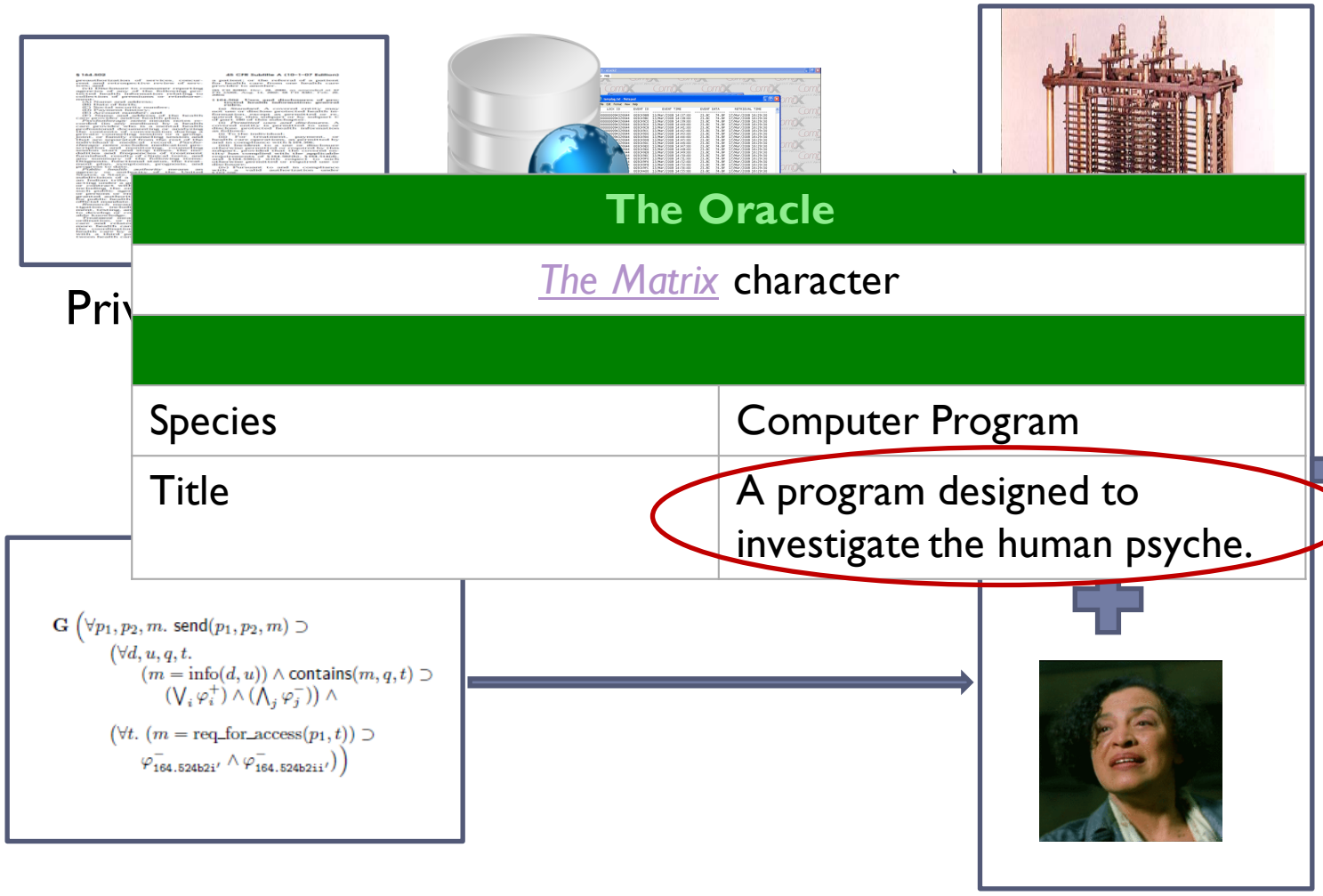
Formalizing and Enforcing Purpose Restrictions in Privacy Policies

Anupam Datta
Carnegie Mellon University
18734: Foundations of Privacy
Fall 2016

Detecting Policy Violations



Purpose Restrictions in Privacy Policies

Not for

- Yahoo!'s practice is **not** to use the content of messages […] **for** marketing **purposes**.

Only for

- By providing your personal information, you give [Social Security Administration] consent to use the information **only for** the **purpose** for which it was collected.

Purpose Restrictions are Ubiquitous

- OECD's Privacy Guidelines
- US Privacy Laws
  - HIPAA, GLBA, FERPA, COPPA,…
- EU Privacy Directive
- Organizational Privacy Policies
  - Google, Facebook, Yahoo,…
  - Hospitals, banks, educational institutions, govt
  - Defense: Mission-based information access

Purpose Restrictions on Actions

With M. C. Tschantz (CMU → Berkeley) and
J. M. Wing (CMU → MSR)
2012 IEEE Symposium on Security & Privacy

Goal

- Give a semantics to
  - **"Not for" purpose restrictions**
  - **"Only for" purpose restrictions**

  that is parametric in the purpose

- Provide automated enforcement of purpose restrictions for that semantics

X-ray taken

Send record

No diagnosis by drug company

Add x-ray

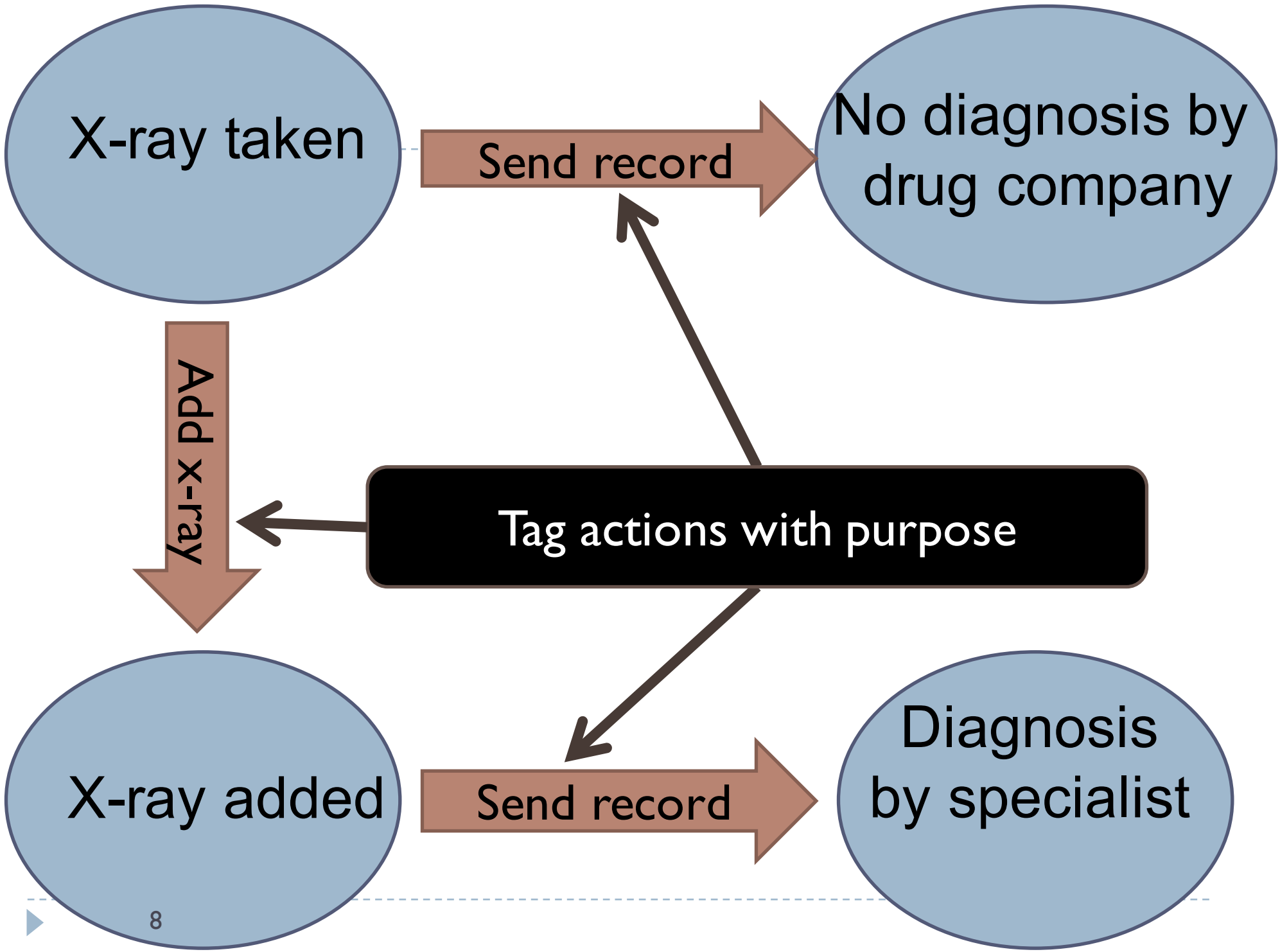
Medical Record

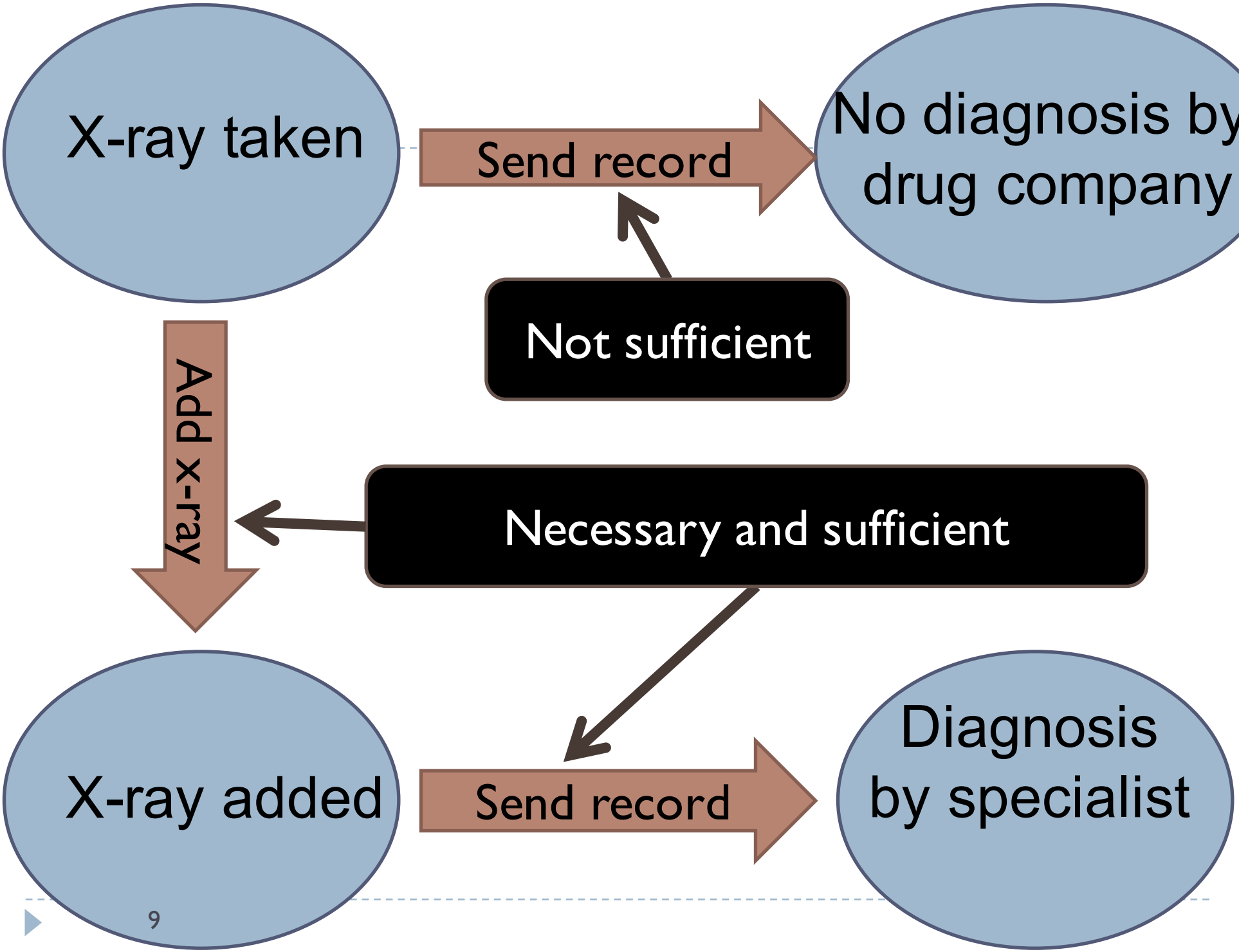
Med records used only for diagnosis

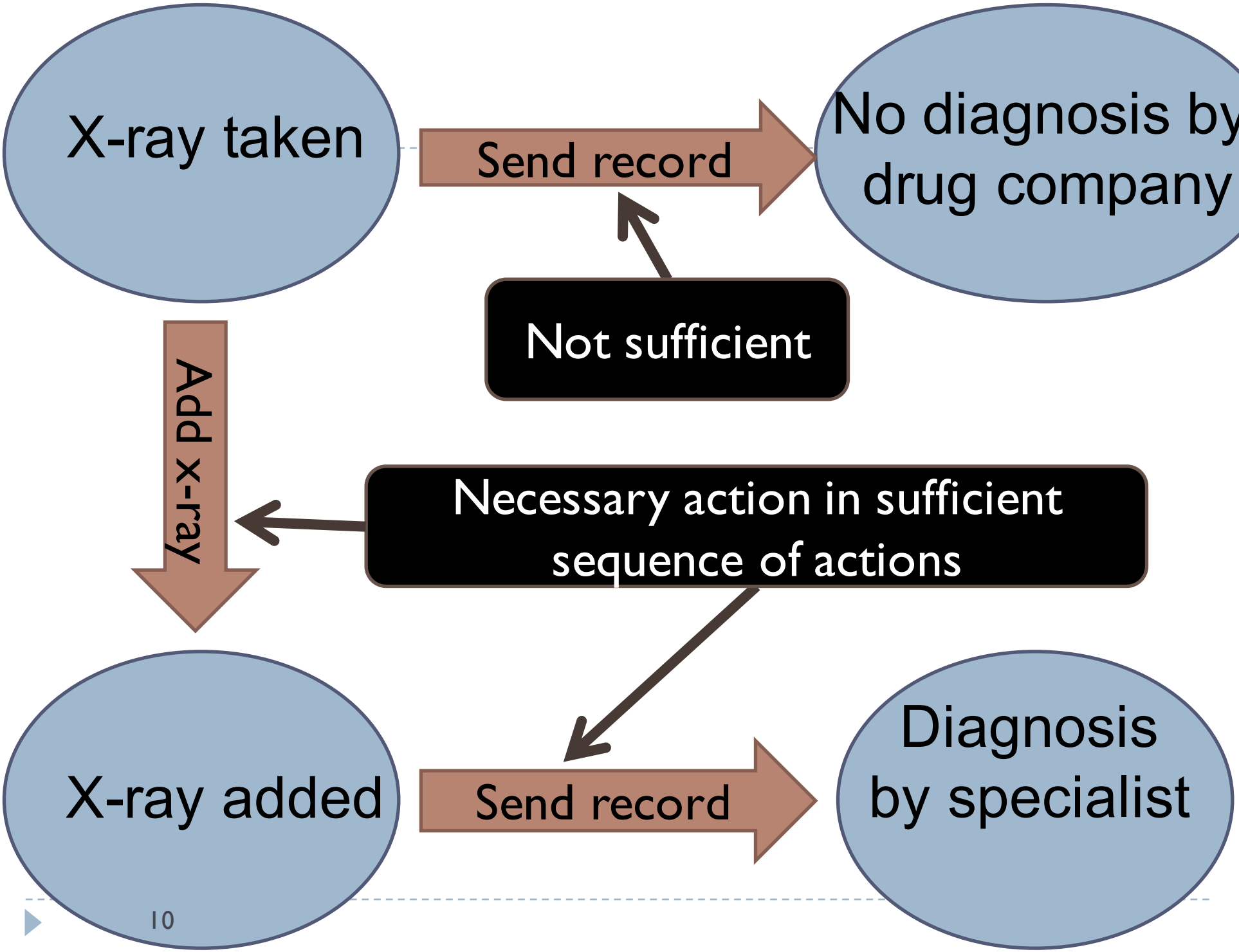
X-ray added

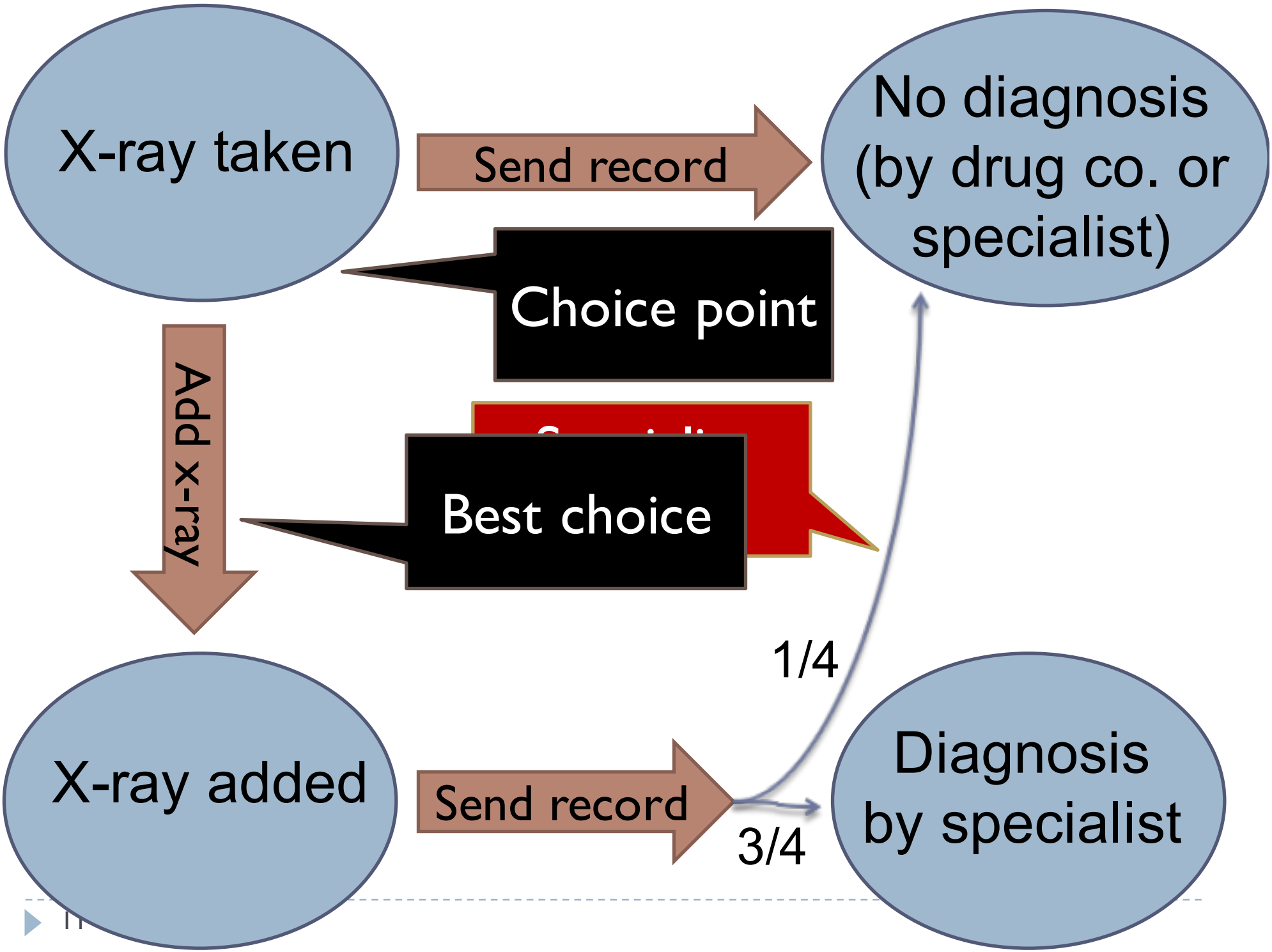
Send record

Diagnosis by specialist


X-ray taken
Send record
No diagnosis by drug company
Add x-ray
Tag actions with purpose
X-ray added
Send record
Diagnosis by specialist


X-ray taken
Send record
No diagnosis by drug company
Not sufficient
Add x-ray
Necessary and sufficient
X-ray added
Send record
Diagnosis by specialist

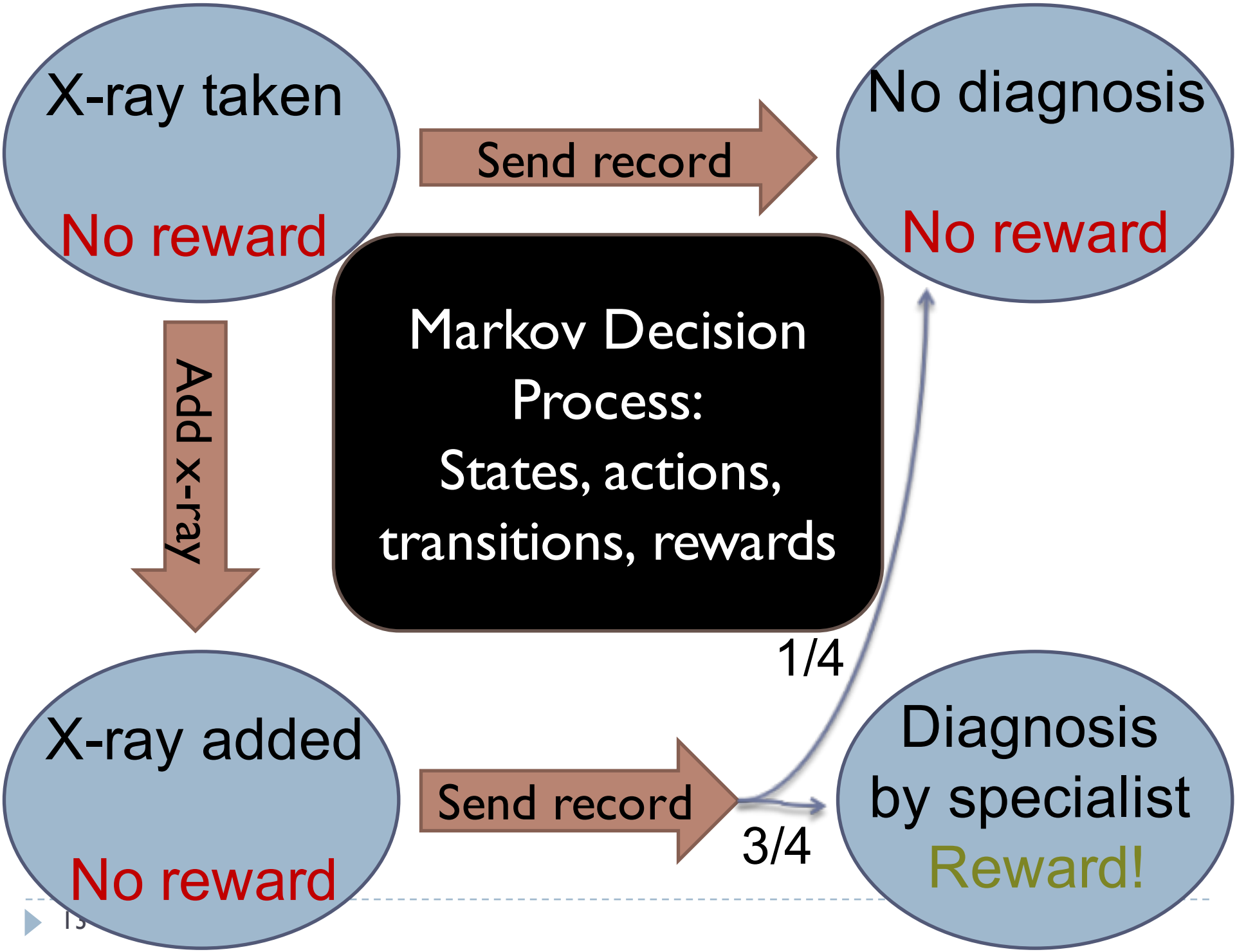

X-ray taken
Send record
No diagnosis by drug company
Not sufficient
Add x-ray
Necessary action in sufficient sequence of actions
X-ray added
Send record
Diagnosis by specialist


X-ray taken
Send record
No diagnosis (by drug co. or specialist)
Choice point
Add x-ray
Best choice
1/4
X-ray added
Send record
3/4
Diagnosis by specialist

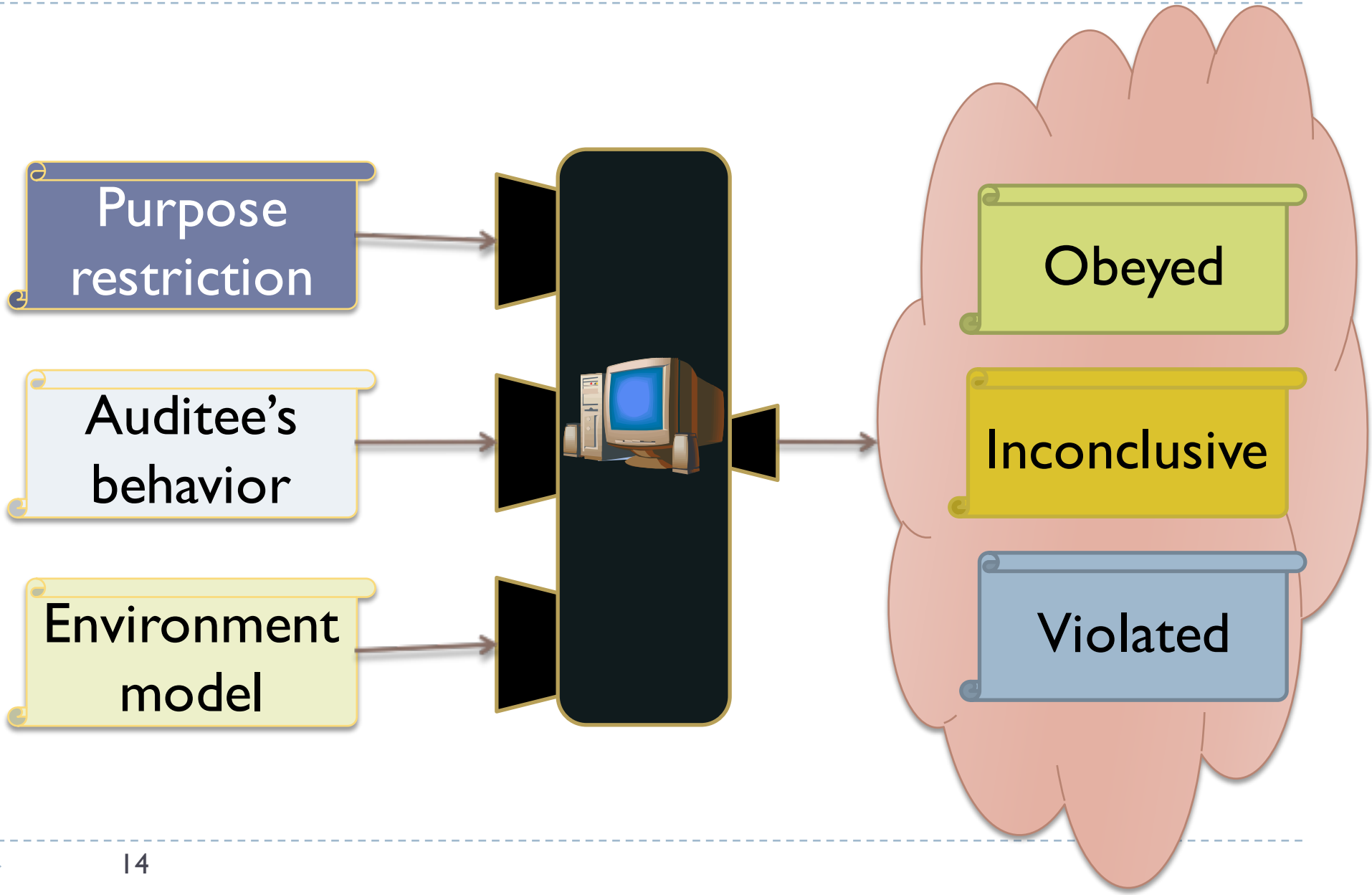
Planning

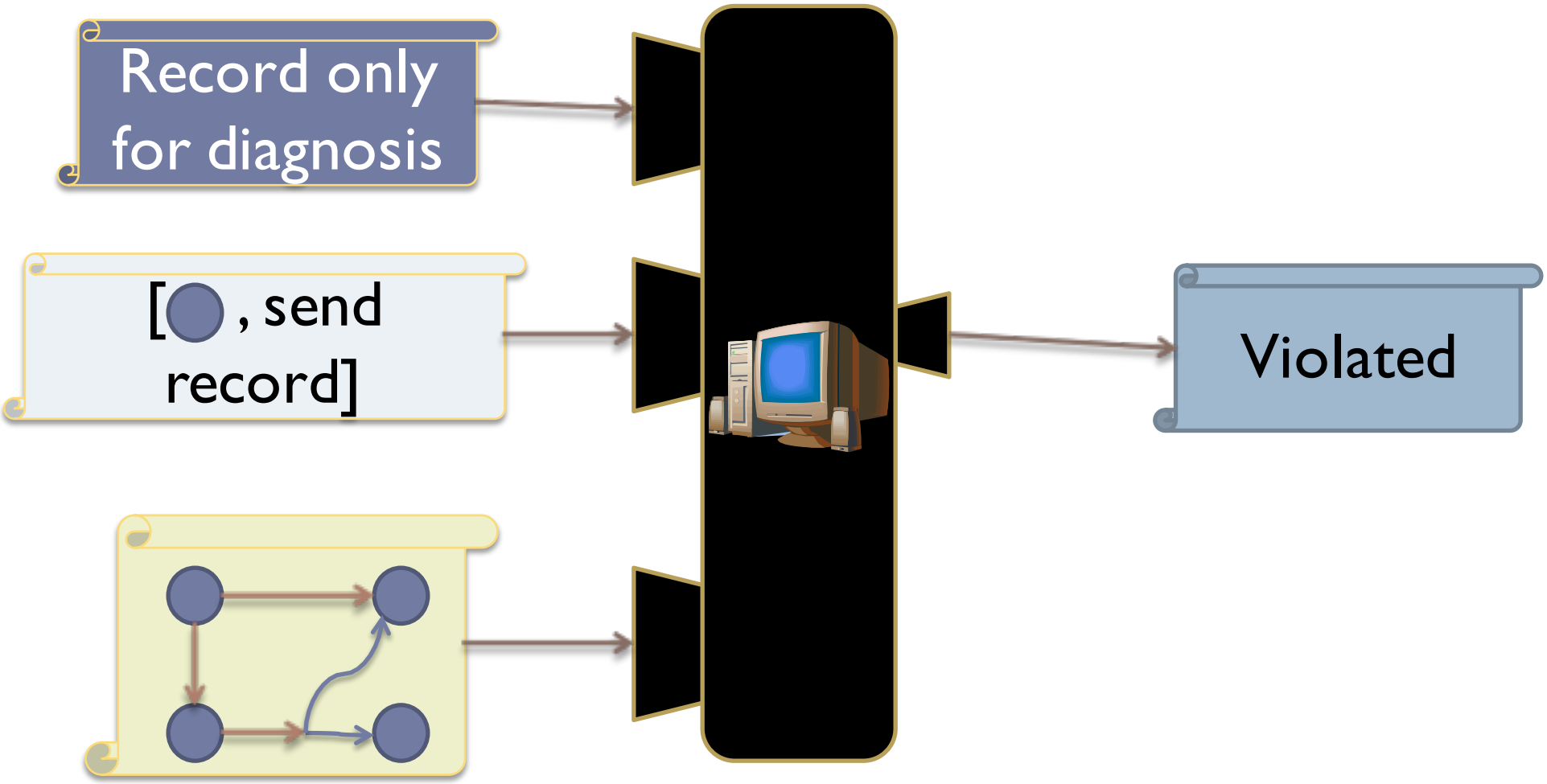
Thesis: An action is for a purpose iff that action is part of a plan for furthering the purpose

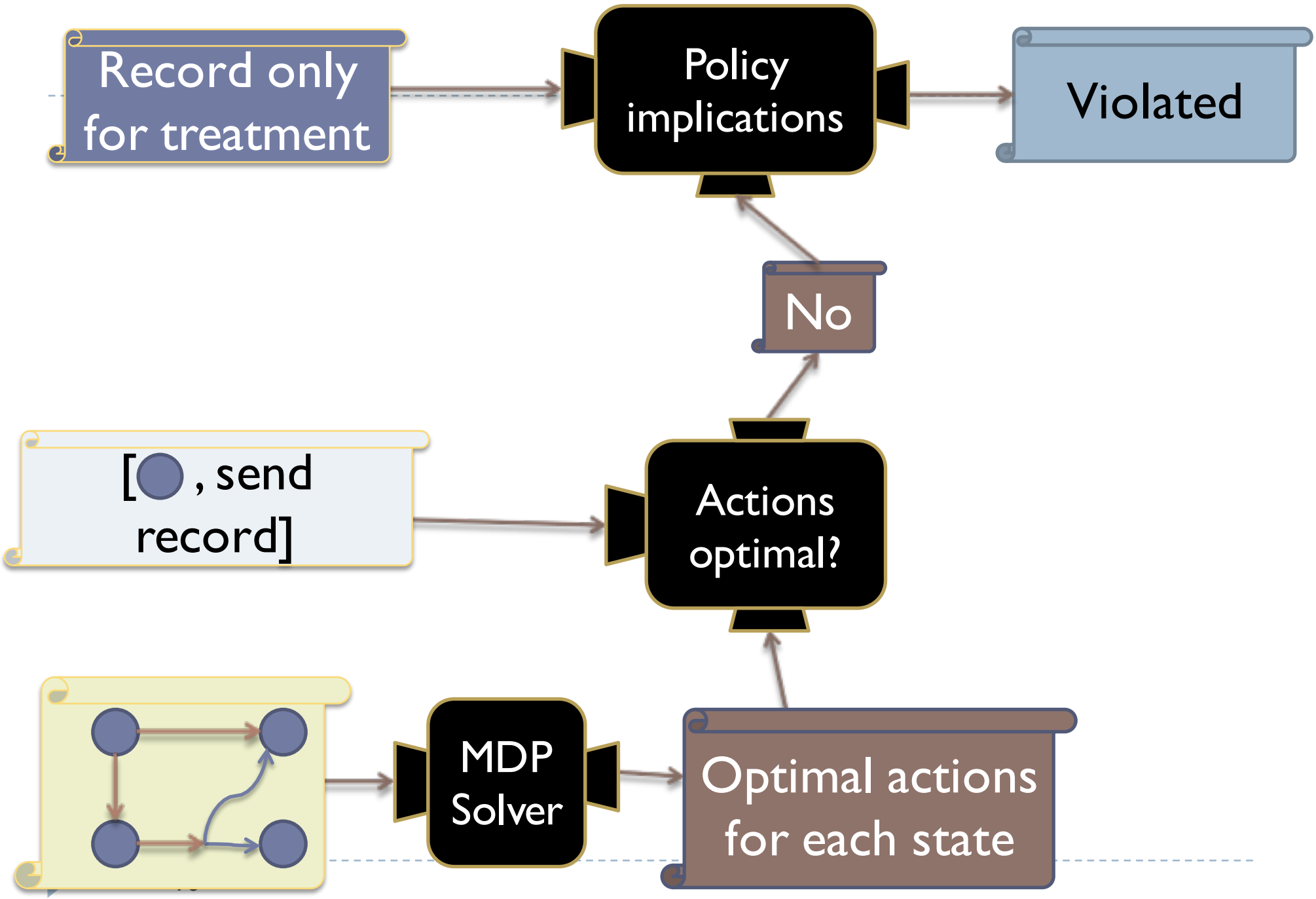
i.e., always makes the best choice for furthering the purpose


X-ray taken
No reward
Send record
No diagnosis
No reward
Markov Decision Process:
States, actions, transitions, rewards
Add x-ray
X-ray added
No reward
Send record
1/4
3/4
Diagnosis by specialist
Reward!

Auditing


Purpose restriction
Auditee's behavior
Environment model
Obeyed
Inconclusive
Violated


Record only for diagnosis
[, send record]
Violated


Record only for treatment
Policy implications
Violated
No
[, send record]
Actions optimal?
MDP Solver
Optimal actions for each state

No False Positives

- Theorem (Soundness):
  If the algorithm returns "violation", then the actions recorded in the log are not only for the purpose

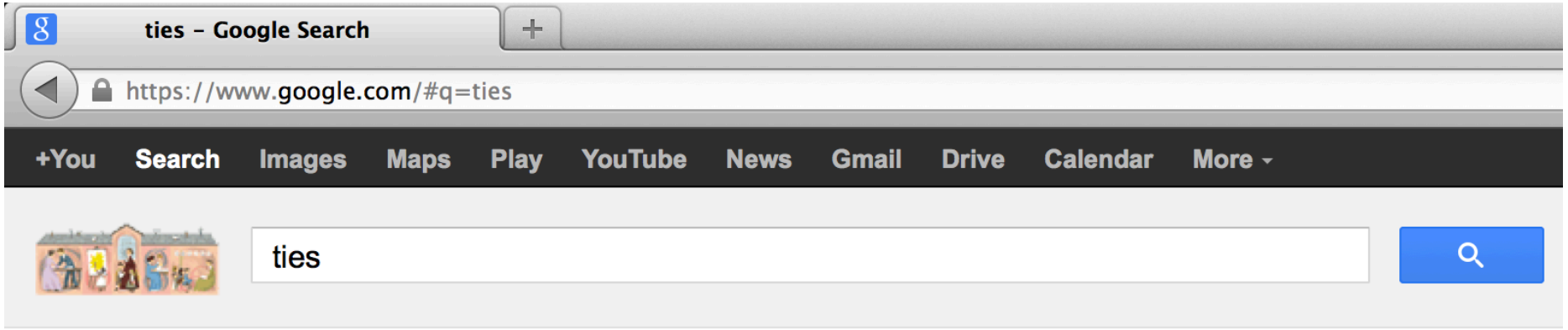
Utility + Privacy

- Learn MDPs from large audit logs
  - E.g., using reinforcement learning techniques
- Compute optimal plans in MDP
  - Improve functional outcomes (e.g, healthcare outcomes, corporate/defense mission)
  - Improve privacy/security (e.g., detect inappropriate accesses to sensitive information by authorized insiders)

Purpose Restrictions on Information Use

With M. C. Tschantz (CMU → Berkeley) and
J. M. Wing (CMU → MSR)

2013 European Symposium on Research in Computer Security


ties - Google Search
https://www.google.com/#q=ties
+You Search Images Maps Play YouTube News Gmail Drive Calendar More
ties
Web Images Maps Shopping More Search tools
About 160,000,000 results (0.26 seconds)
Ads related to ties
Buy Neck Ties Online - necktiesinstock.com
www.necktiesinstock.com/
Choose From A Wide Range Of Colors, Styles & Textures of Ties. Buy Now!
Men's Ties & More - AbsoluteTies.com
www.absoluteties.com/
Spend $50 & Get Free Shipping & 10% Discount too.
The Tie Bar
www.thetiebar.com/
Provider of Handmade Silk Neckties, Discount Neckties, Mens Silk Neck Ties, Cufflinks, Affordable Ties, and Bowties.
Bow Ties - Tie Bars - NeckTies - Skinny Ties
Ties - Buy Mens Neckties, Bow Ties, Tie Racks & More | Ties.com
www.ties.com/
We stock 1000+ brands of skinny ties, plaid ties, tie racks, bow ties and more! Friendly customer service and get free shipping today if you buy $50+.
Bow Ties - Port Belle Skinny Tie - Shop Ties by Color - Neckties

Antidepressant Medication - Info On An Rx **Antidepressant** Drug

knowmydepression.com/**antidepressant**

Visit For Treatment Info & Facts.

Party Supplies For Sale - Buy Your **Party Supplies** Online Now

www.orientaltrading.com/**PartySupplies**

Free Shipping on Orders Over $49!

Oriental Trading has 925 followers on Google+

Party Favors Sale

Party Decorations

Birthday Party Supplies

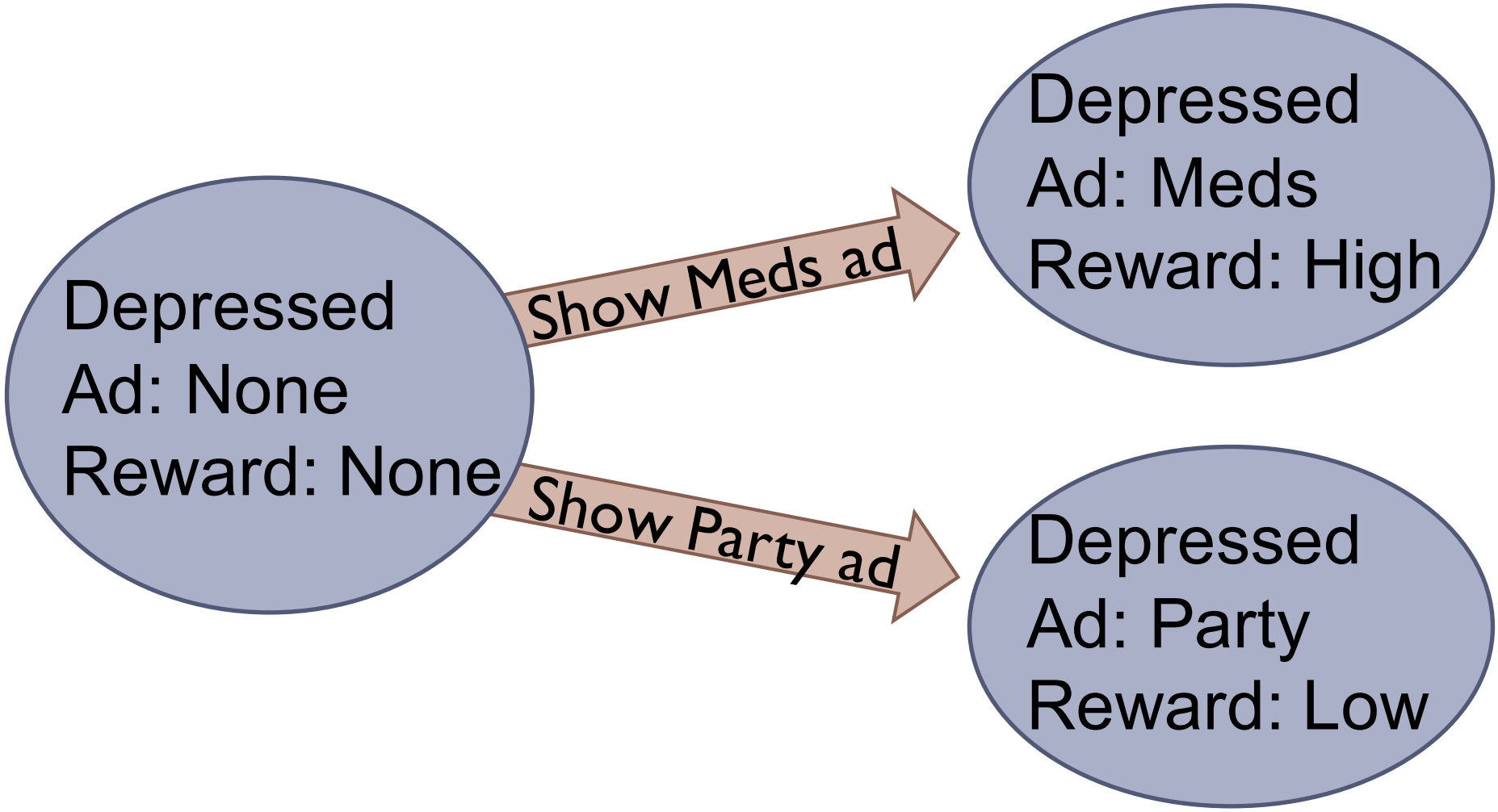
Halloween Party Supplies

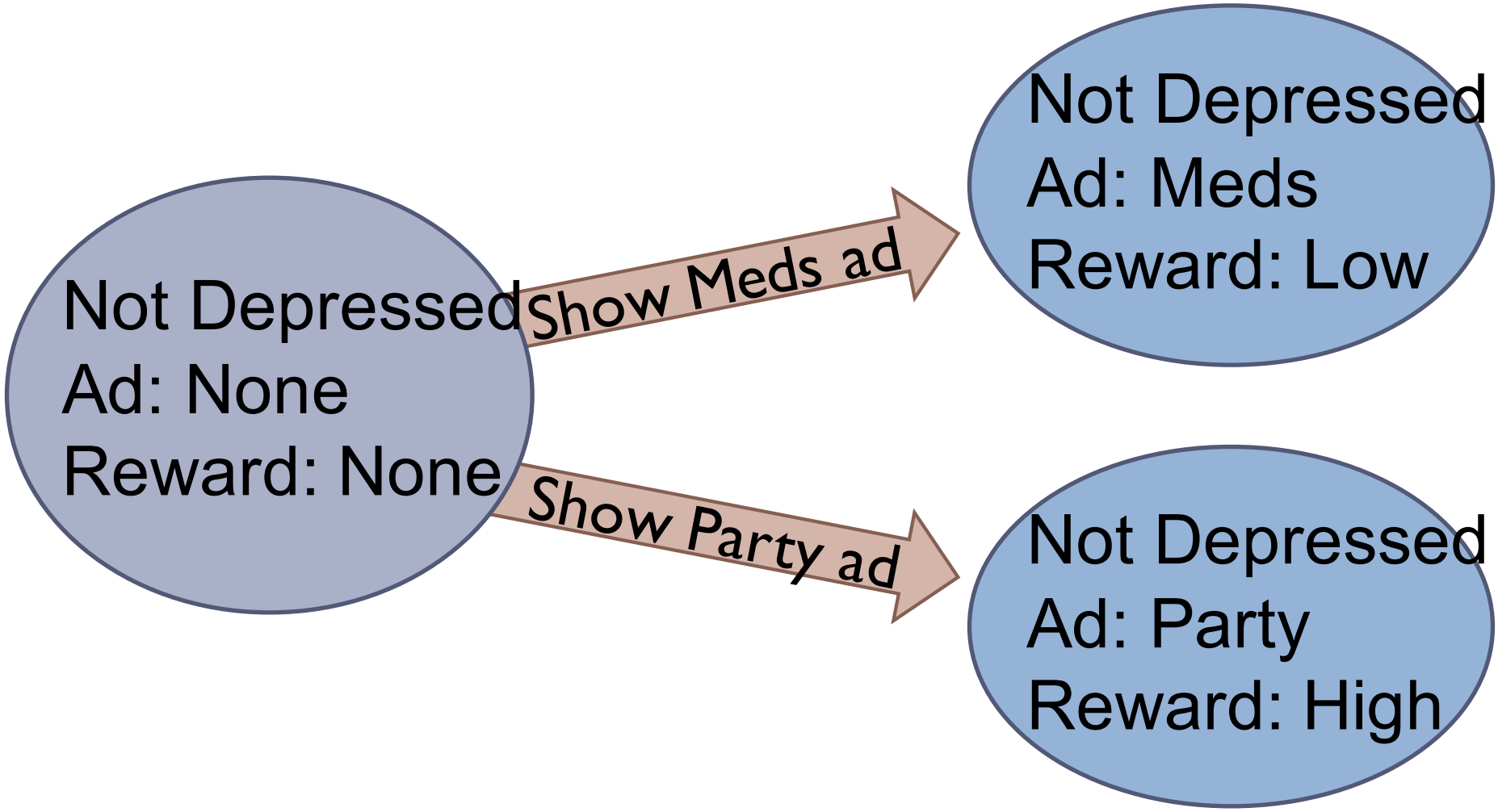
Google's Privacy Policy

When showing you tailored ads, we will not associate a cookie or anonymous identifier with sensitive categories, such as those based on race, religion, sexual orientation or health.

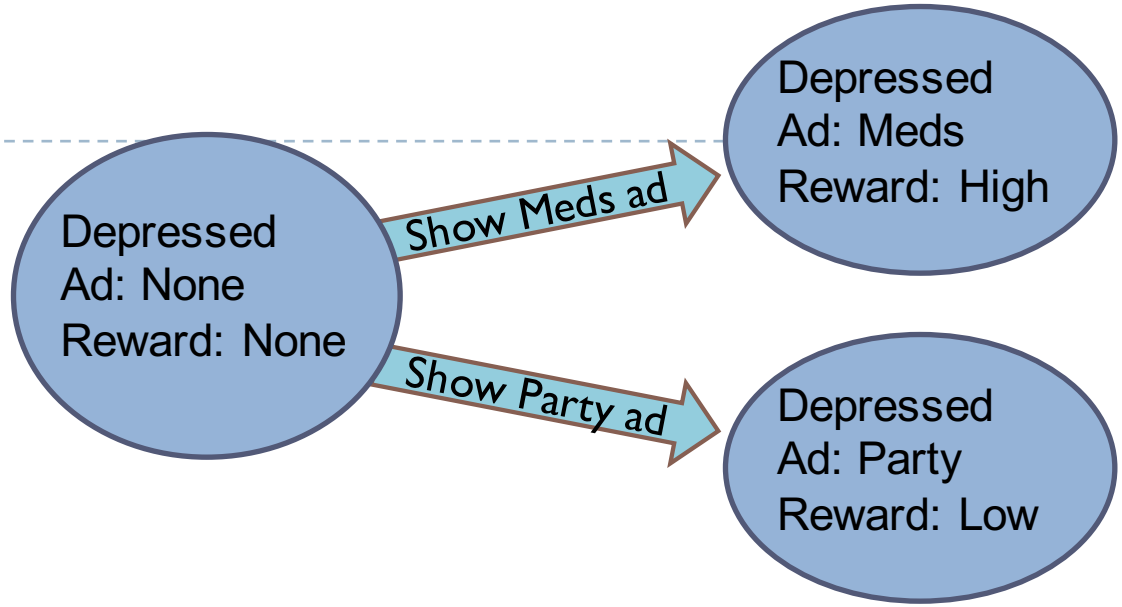
Rewards from ads

	Depressed	Not Depressed
Meds	High	Low
Party	Low	High

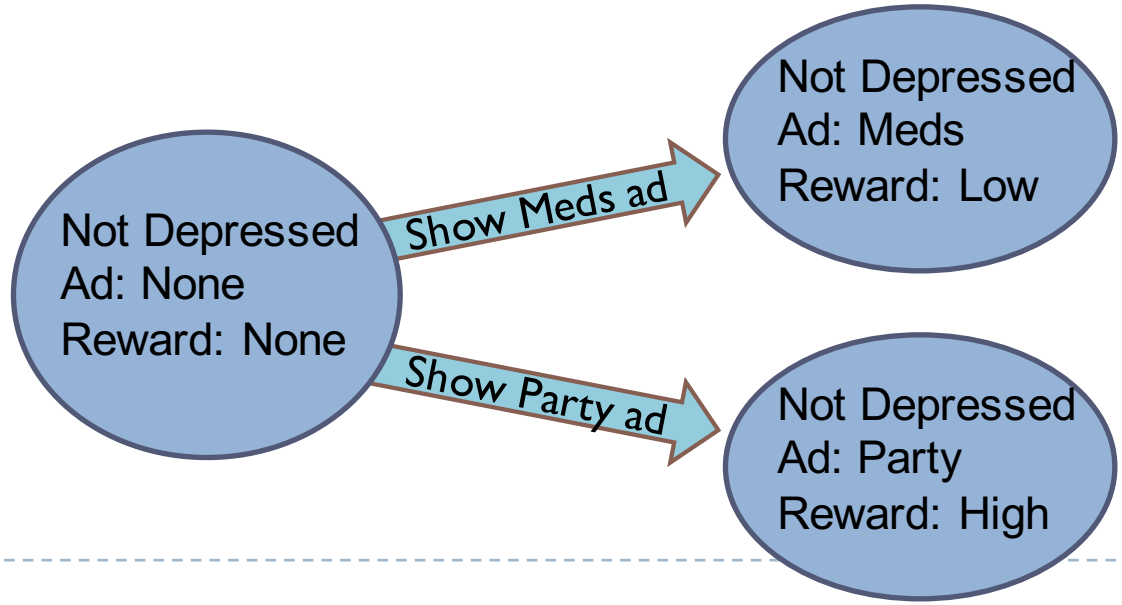

Depressed
Ad: None
Reward: None
Show Meds ad
Show Party ad
Depressed
Ad: Meds
Reward: High
Depressed
Ad: Party
Reward: Low


Not Depressed
Ad: None
Reward: None
Show Meds ad
Show Party ad
Not Depressed
Ad: Meds
Reward: Low
Not Depressed
Ad: Party
Reward: High

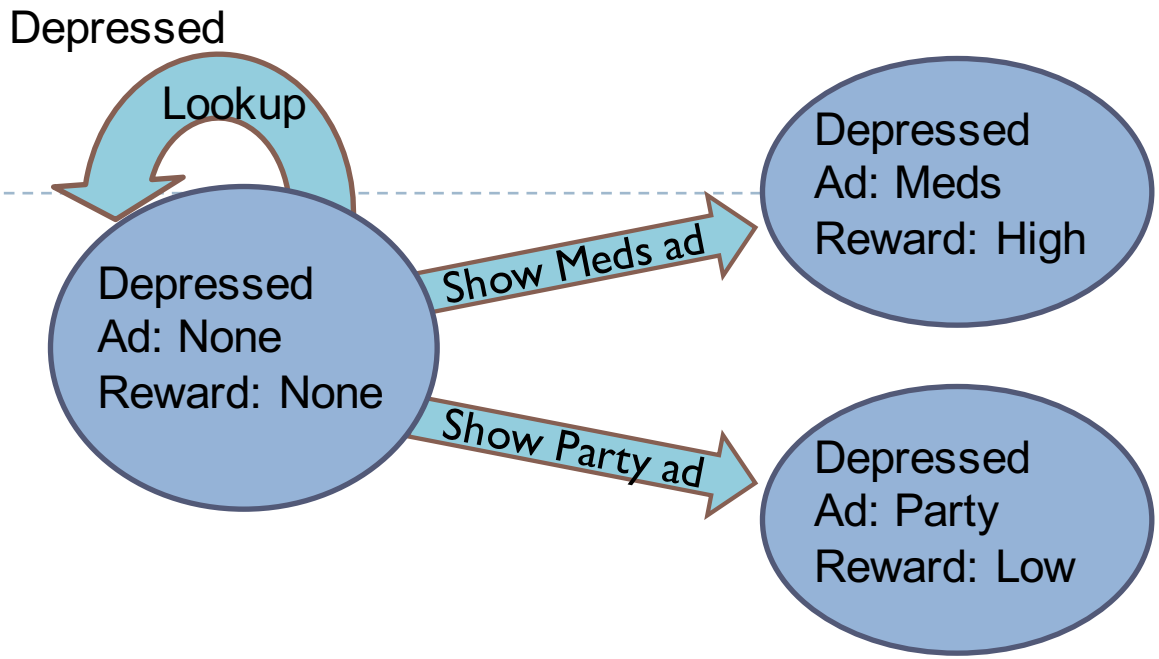
Depressed Case



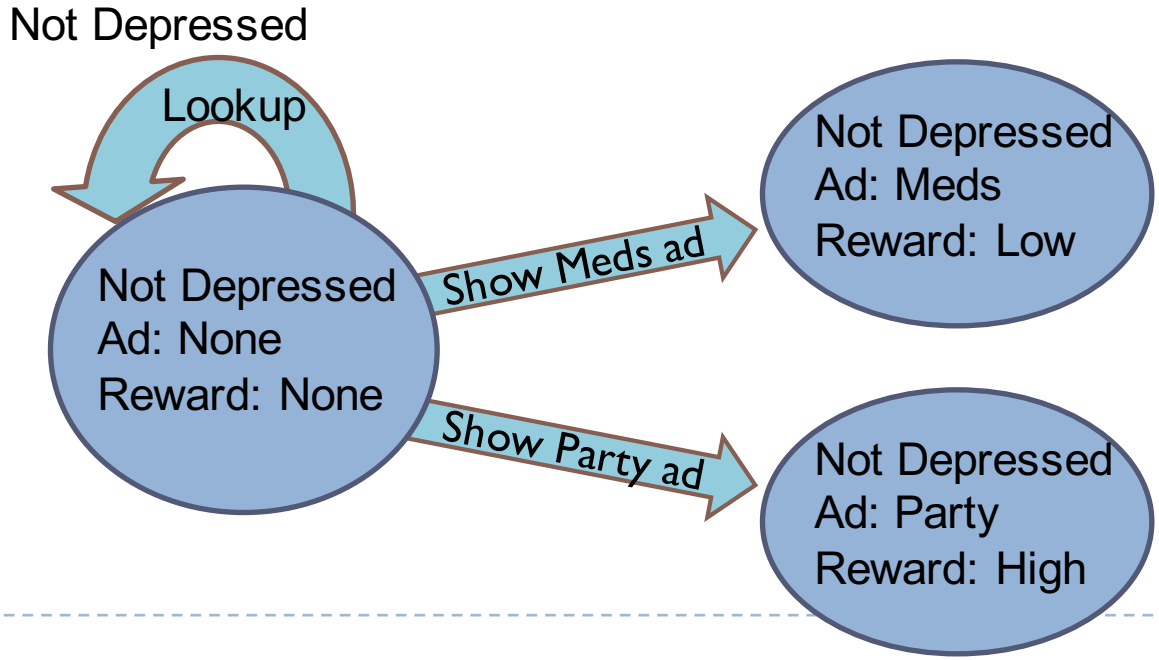
Not Depressed Case



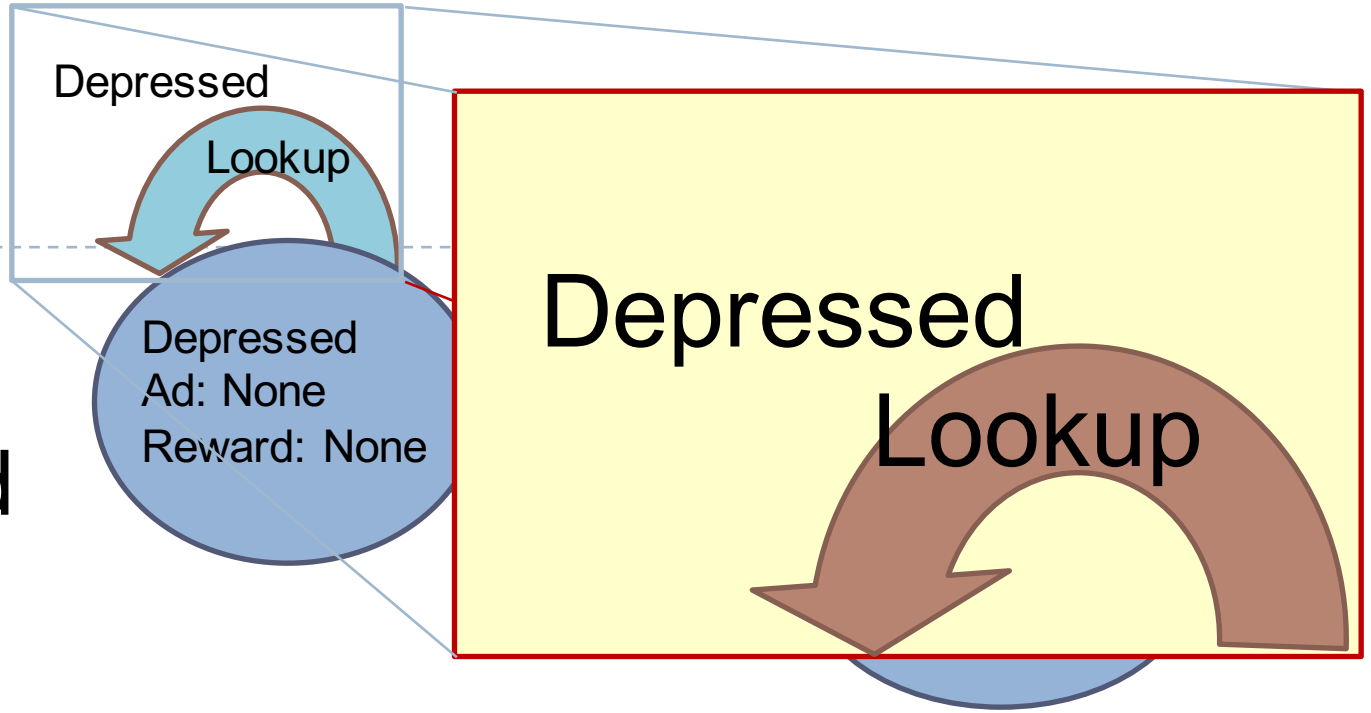
Depressed Case



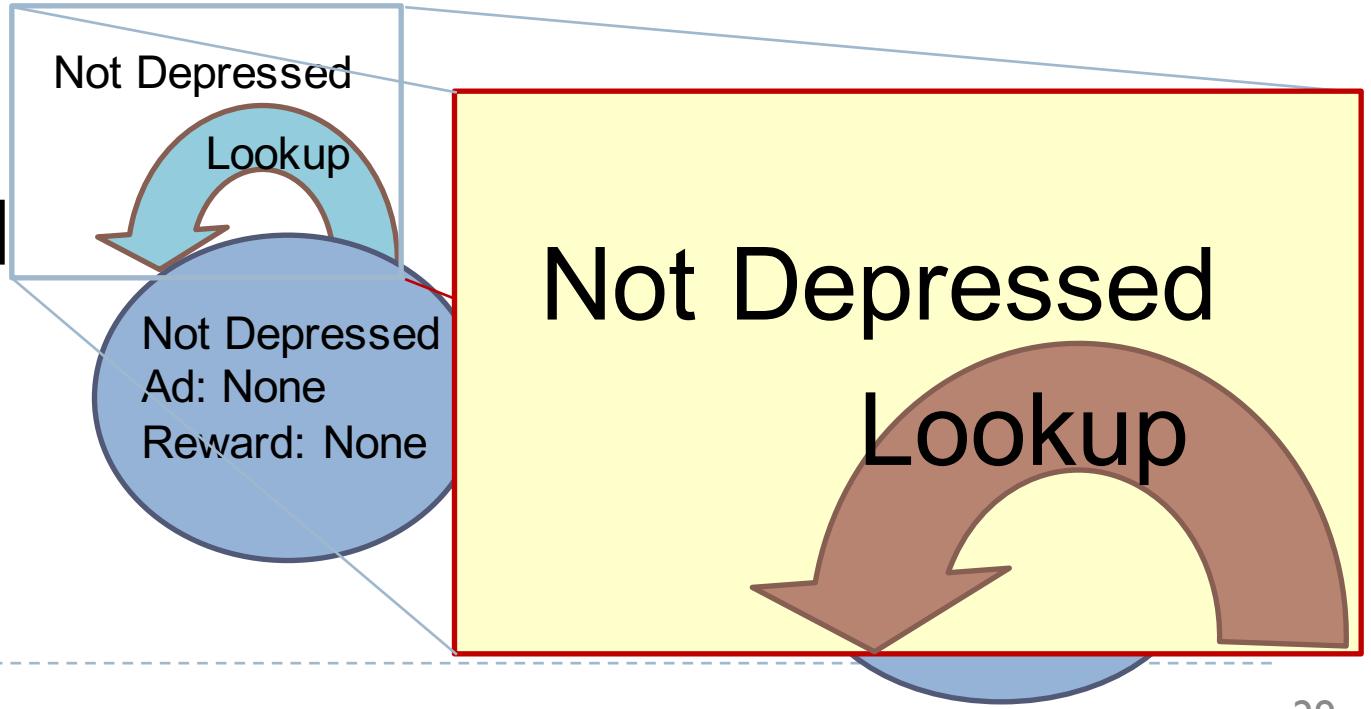
Not Depressed Case

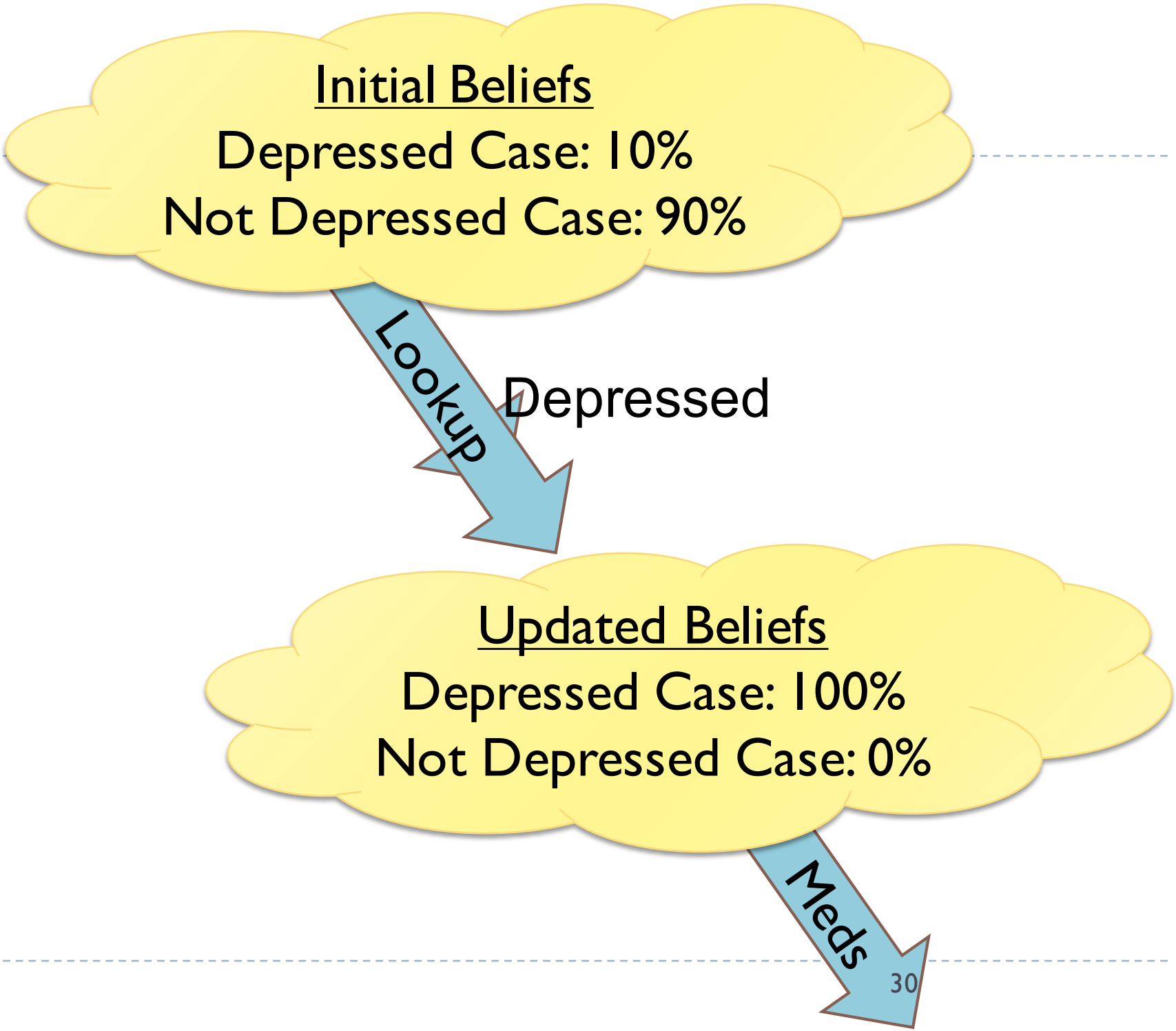


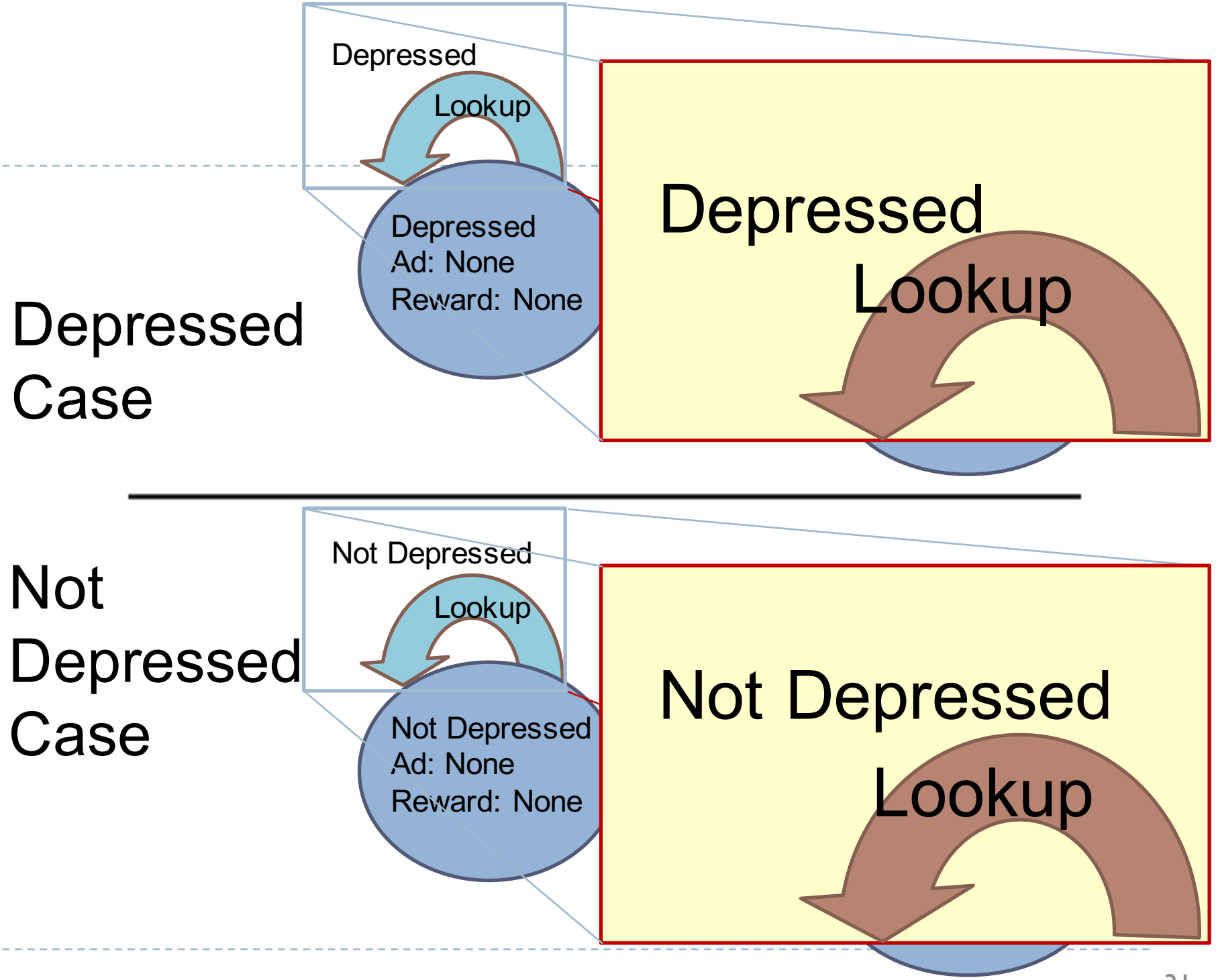
Depressed Case

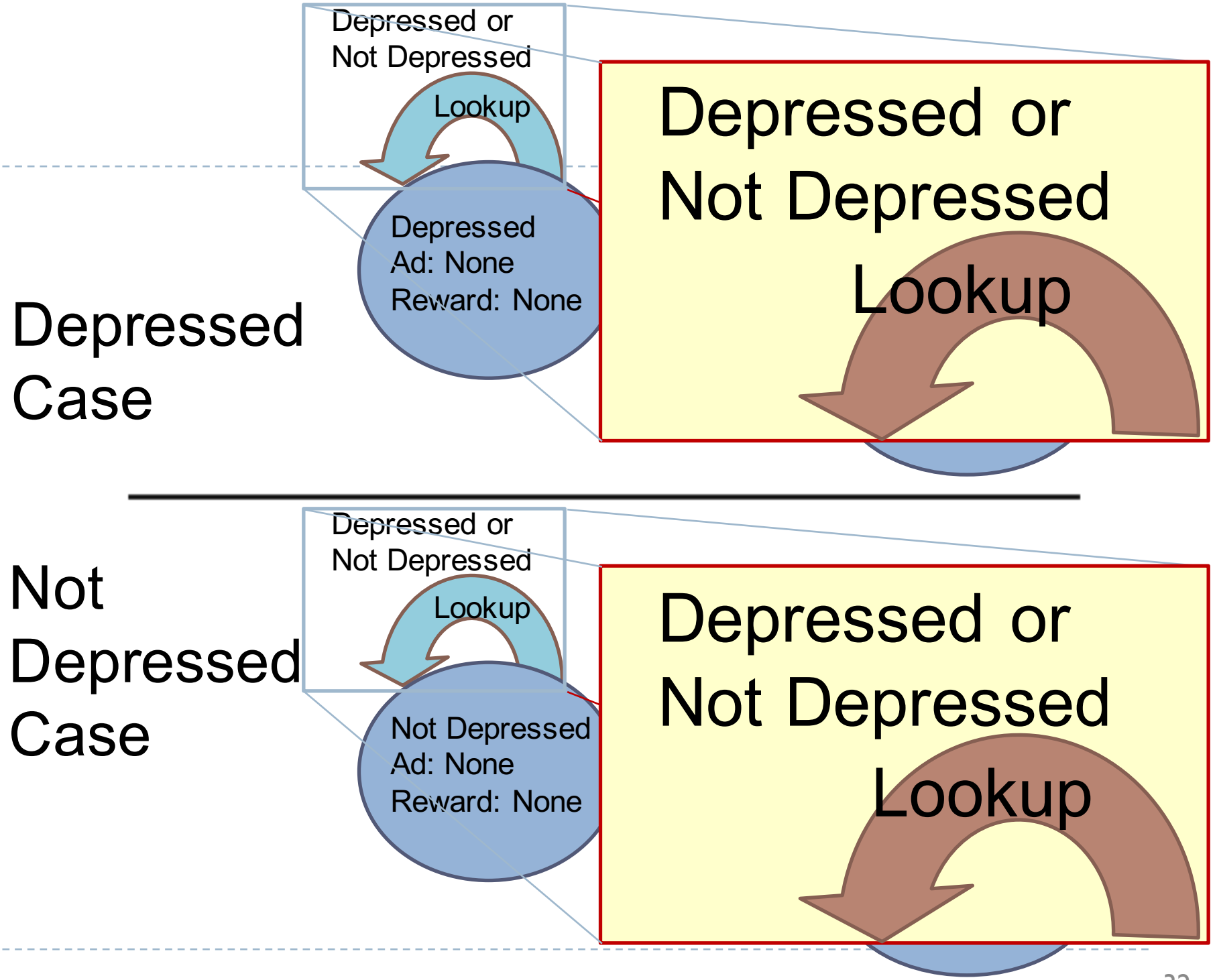


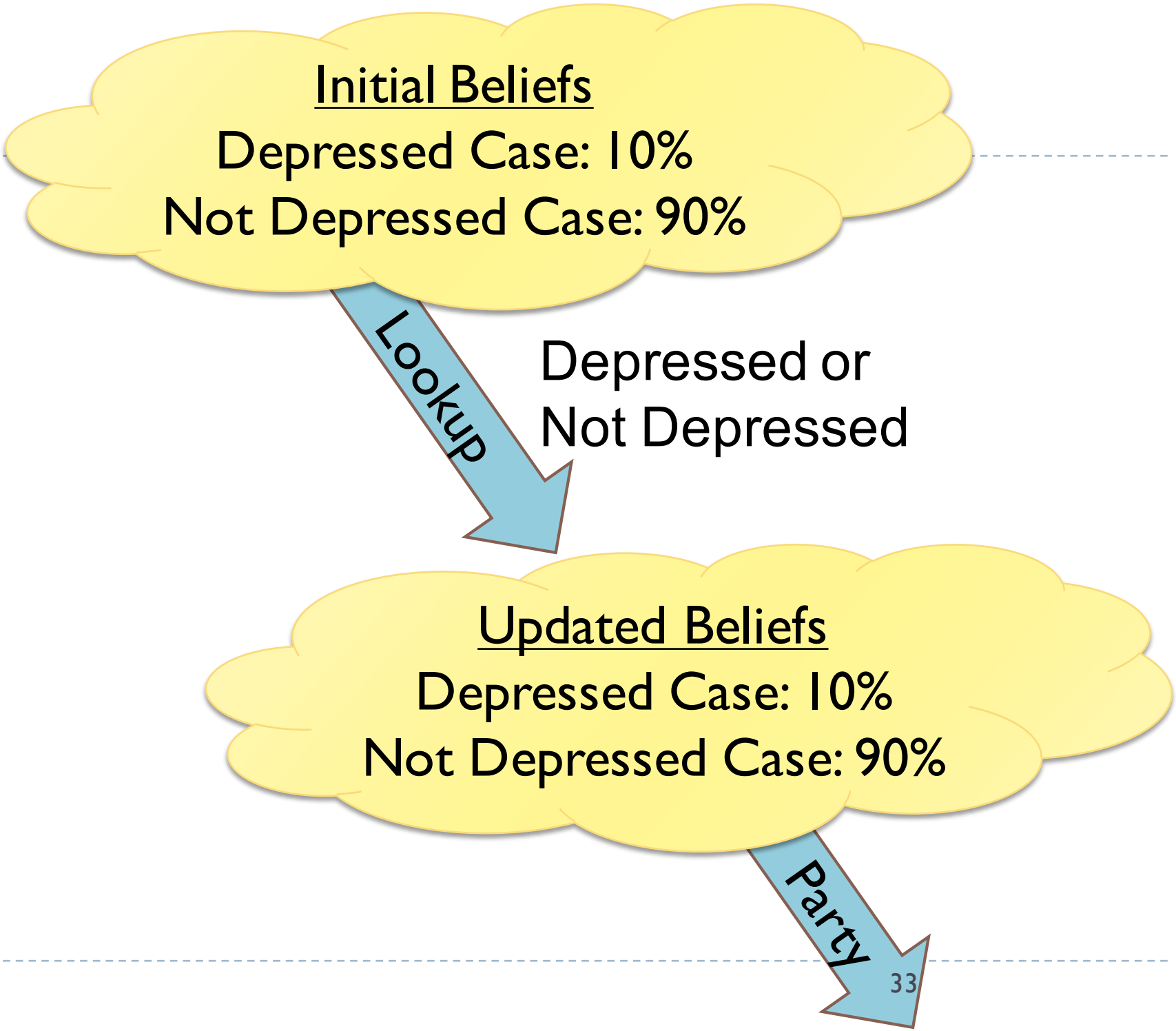
Not Depressed Case




Initial Beliefs
Depressed Case: 10%
Not Depressed Case: 90%
Lookup
Depressed
Updated Beliefs
Depressed Case: 100%
Not Depressed Case: 0%
Meds


Depressed
Case
Depressed
Lookup
Depressed
Ad: None
Reward: None
Depressed
Lookup
Not
Depressed
Case
Not Depressed
Lookup
Not Depressed
Ad: None
Reward: None
Not Depressed
Lookup

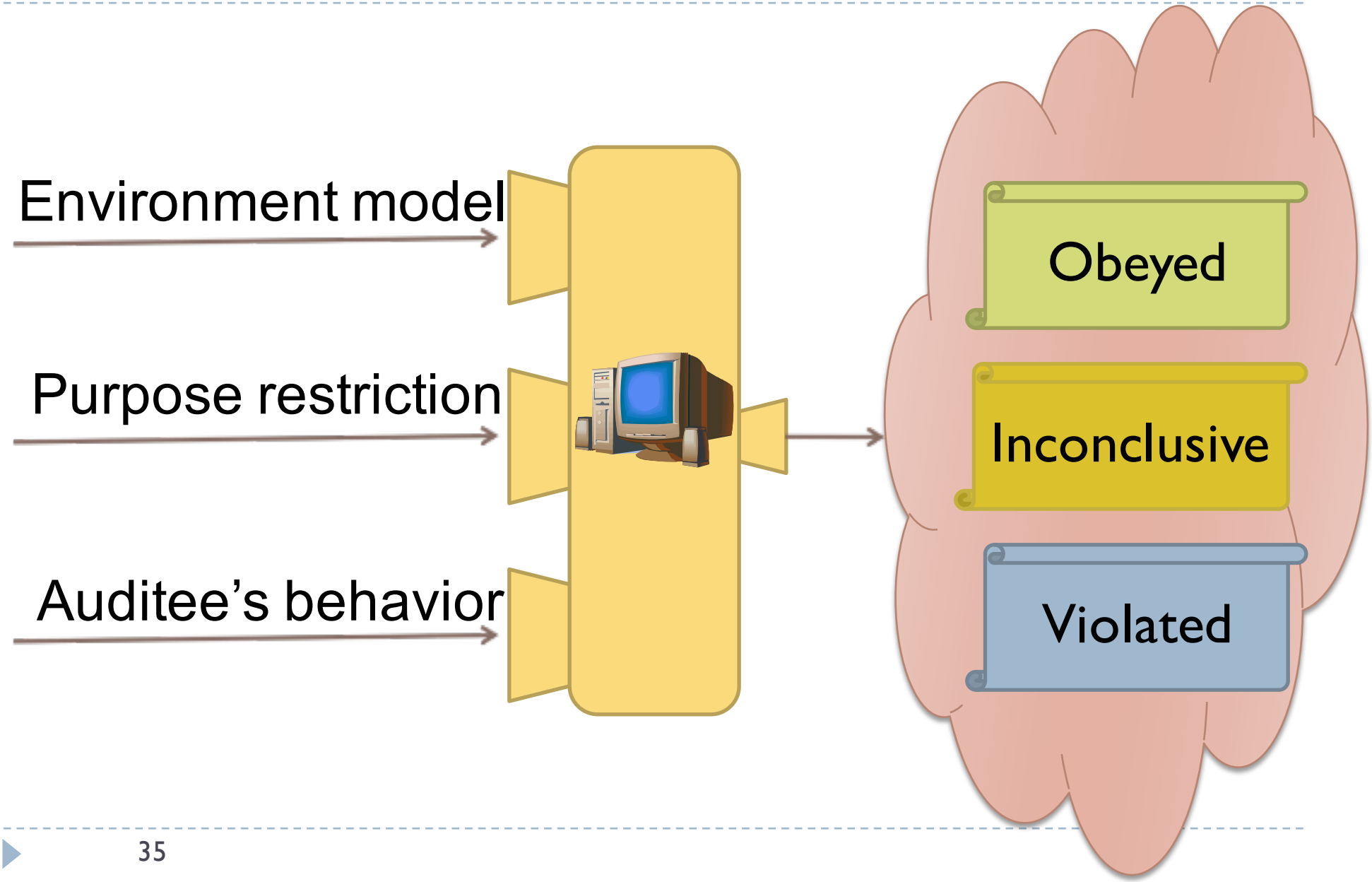

Depressed or
Not Depressed
Lookup
Depressed
Ad: None
Reward: None
Depressed or
Not Depressed
Lookup
Depressed
Case
Depressed or
Not Depressed
Lookup
Not Depressed
Ad: None
Reward: None
Depressed or
Not Depressed
Lookup
Not
Depressed
Case


Initial Beliefs
Depressed Case: 10%
Not Depressed Case: 90%
Lookup
Depressed or
Not Depressed
Updated Beliefs
Depressed Case: 10%
Not Depressed Case: 90%
Party

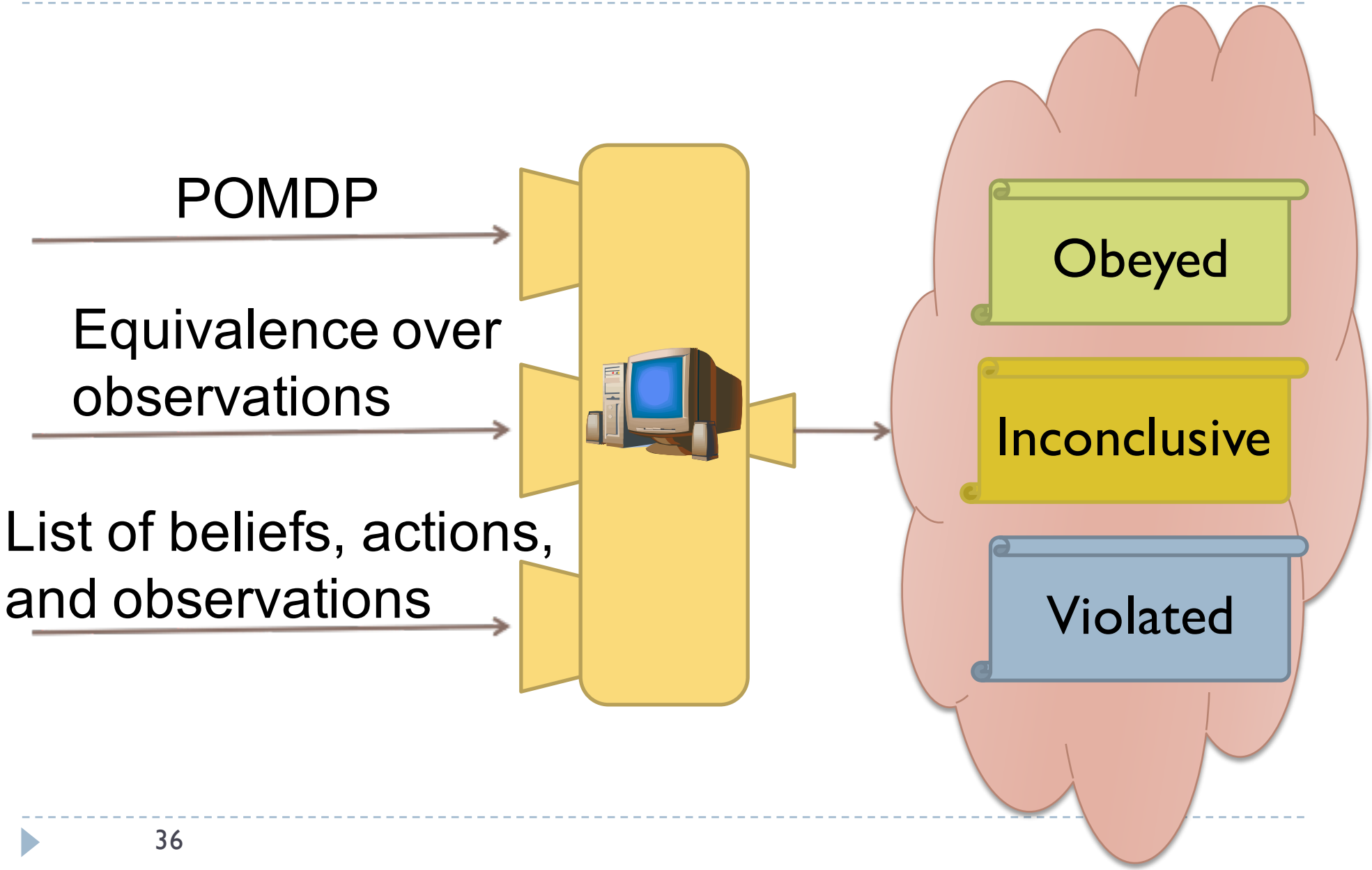
Planning + Information Flow

- Cognitive: Actions are for a purpose without using some information if they came from a plan selected by optimizing a model with disallowed information conflated.
  - Requires mind reading for enforcement
- Behaviorist: Actions are for a purpose without using some information if they are consistent with a plan optimizing a model with disallowed information conflated.
  - Could be consistent by coincidence and actually be for another purpose using the information

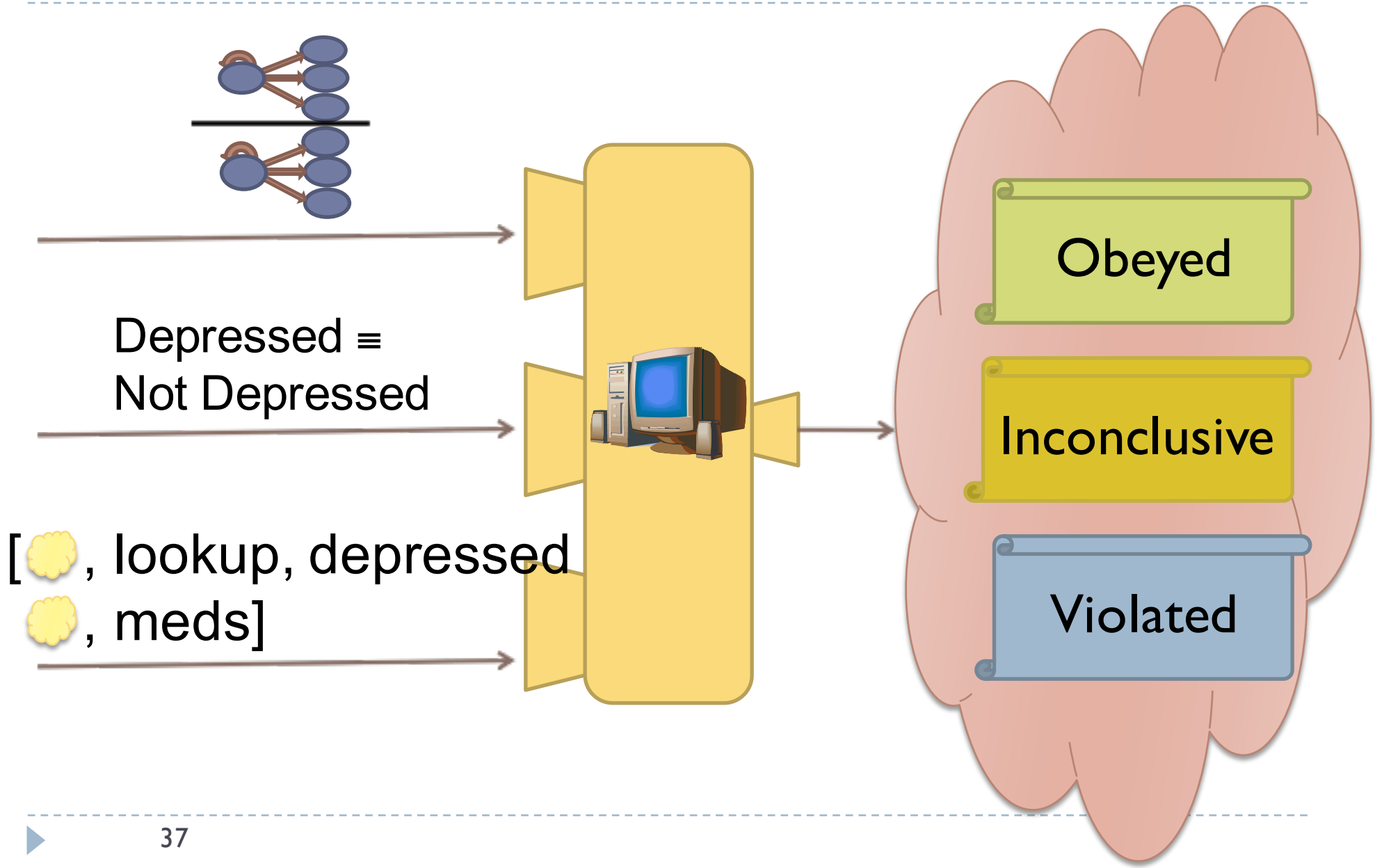
Auditing

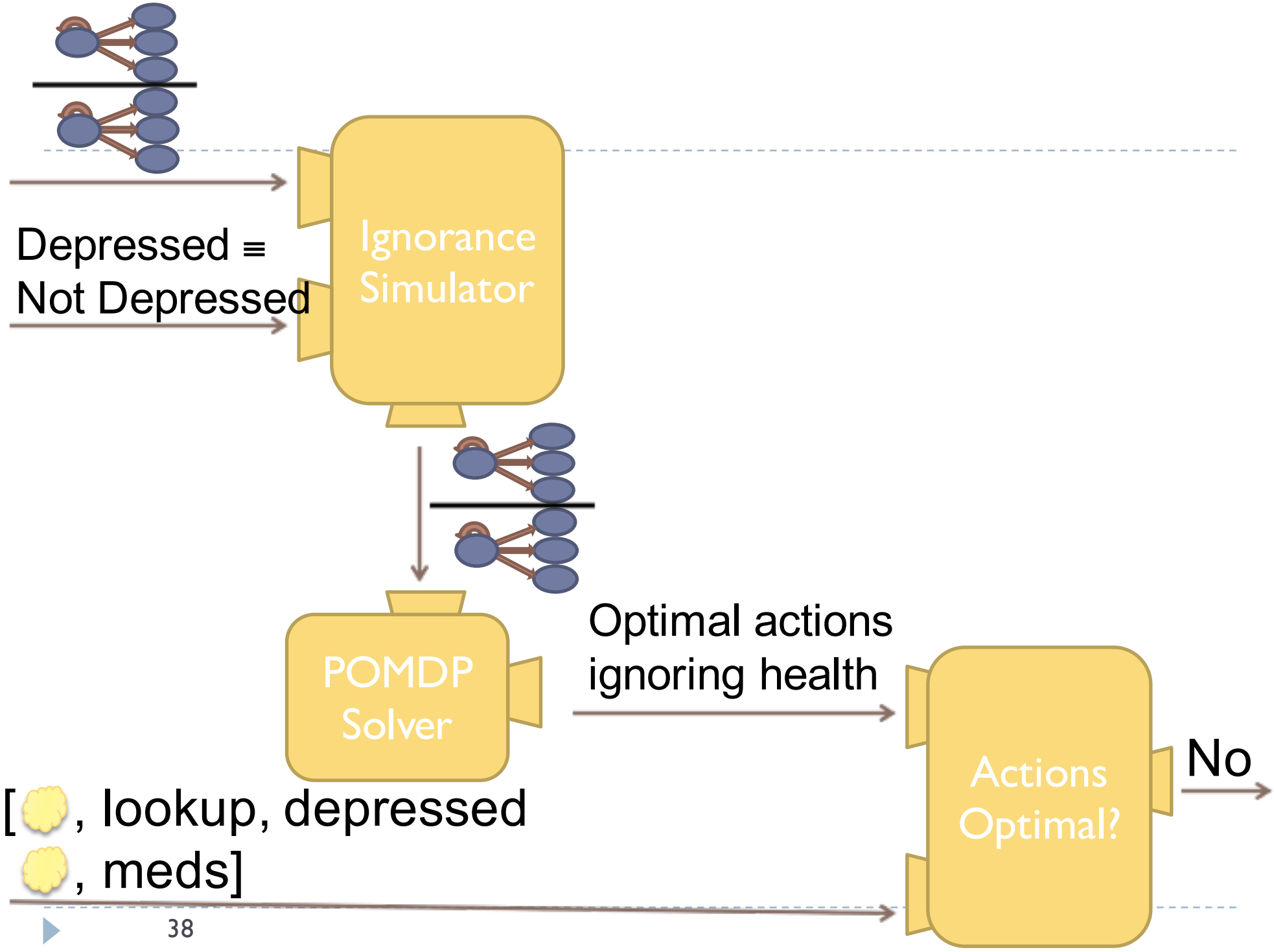

Environment model
Purpose restriction
Auditee's behavior
Obeyed
Inconclusive
Violated

Auditing


POMDP
Equivalence over observations
List of beliefs, actions, and observations
Obeyed
Inconclusive
Violated

Auditing


Depressed ≡
Not Depressed
[, lookup, depressed
, meds]
Obeyed
Inconclusive
Violated


Depressed ≡
Not Depressed
Ignorance Simulator
POMDP Solver
Optimal actions ignoring health
Actions Optimal?
No
[, lookup, depressed
, meds]

Implications

- The actions were not for the purpose of marketing without using health data
  - Violates:"marketing without using health data"
- Either (1) used health data for marketing or (2) performed actions for some other purpose
  - In case (1) violates:"health data not for marketing"

Prior Approaches

- Prior approaches:
  - Labeling actions (industry practice)
  - Labeling sequences of actions (Al-Fedaghi 07, Jafari et al. 09)
  - Labeling roles (Byun et al. 05, 08, 10)
  - Labeling code (Hayati and Abadi 05)
- Our work provides a semantic foundation
- Shows the expressiveness of each approach

Interesting Points: only-for rules

- Cannot catch all violations of only-for rules
  - Coincidences provide tenable deniability
- Enforcing only-for rules can improve both privacy and utility
  - Keeps auditees on task

Interesting Point: not-for rules

- Not-for rules restrict very little
  - May still perform actions for very similar purposes

- FIPPs principle on **purpose specification** tries to address this concern

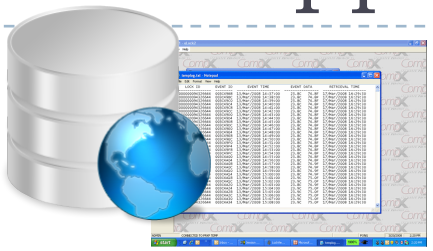
  "The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose."

Future Work

- Improving accuracy
  - Human models of planning
- Furthering practicality
  - Automated creation of environment models
- Applications
  - Minimum necessary disclosure
- Generalizations
  - Multiple purposes

Summary: Audit Approach

Privacy Policy



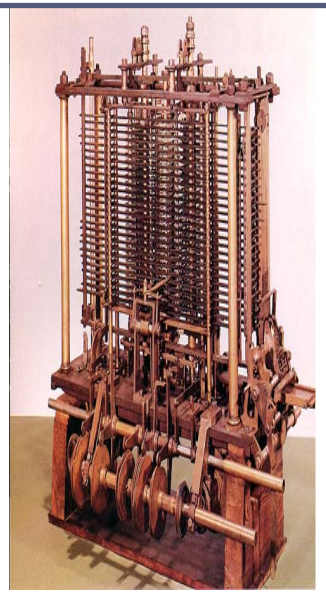
Organizational audit log

Complete formalization of HIPAA, GLBA

$$\mathbf{G}\,\Big(\forall p_1, p_2, m.\ \mathsf{send}(p_1, p_2, m) \supset (\forall d, u, q, t.\ (m = \mathrm{info}(d, u)) \wedge \mathsf{contains}(m, q, t) \supset (\textstyle\bigvee_i \varphi_i^{+}) \wedge (\textstyle\bigwedge_j \varphi_j^{-})) \wedge (\forall t.\ (m = \mathrm{req\_for\_access}(p_1, t)) \supset \varphi^{-}_{\mathtt{164.524b2i'}} \wedge \varphi^{-}_{\mathtt{164.524b2ii'}})\Big)$$

Computer-readable privacy policy

Environment Model



Audit

Automated audit for black-and-white policy concepts

Detect policy violations

Oracles to audit for grey policy concepts

POMDP: Formal Definition

POMDPs. To define POMDPs, let $\mathsf{Dist}(X)$ denote the space of all distributions over the set X and let $\mathbb{R}$ be the set of real numbers. A POMDP is a tuple $\langle \mathcal{Q}, \mathcal{A}, \tau, \rho, \mathcal{O}, \nu, \gamma \rangle$ where

- $\mathcal{Q}$ is a finite state space representing the states of the agent's environment;
- $\mathcal{A}$, a finite set of actions;
- $\tau : \mathcal{Q} \times \mathcal{A} \to \mathsf{Dist}(\mathcal{Q})$, a transition function from a state and an action to a distribution over states representing the possible outcomes of the action;
- $\rho : \mathcal{Q} \times \mathcal{A} \to \mathbb{R}$, a reward function measuring the immediate impact on the satisfaction of the purpose when the agent takes the given action in the given state;
- $\mathcal{O}$, a finite observation space containing any observations the agent may perceive while performing actions;
- $\nu : \mathcal{A} \times \mathcal{Q} \to \mathsf{Dist}(\mathcal{O})$, a distribution over observations given an action and the state resulting from performing that action; and
- γ, a discount factor such that $0 \leq \gamma < 1$.

POMDP and Purpose

We say that a POMDP *models a purpose* if ρ measures the degree to which the purpose is satisfied. To select actions for that purpose, the agent should select those that maximizes its expected total discounted reward, $\mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i u_i\right]$ where i represents time and u_i, the reward from the agent's ith action.

Belief States

This goal is complicated by the agent not knowing *a priori* which of the possible states of the POMDP is the current state of its environment. Rather it holds beliefs about which state is the current state. In particular, the agent assigns a probability to each state q according to how likely the agent believes that the current state is the state q. A *belief state* β captures these beliefs as a distribution over states of Q (i.e., $\beta \in \mathsf{Dist}(Q)$). An agent updates its belief state as it performs actions and makes observations. When an agent takes the action a and makes the observation o starting with the beliefs β, the agent develops the new beliefs β' where $\beta'(q')$ is the probability that q' is the next state.

Optimal Strategy

To maximize its expected total discounted reward, the agent does not need to track its history of actions and observations independently of its beliefs as such beliefs are a sufficient statistic. Thus, the agent need only consider for each possible belief β it can have, what action it would perform. That is, the agent can plan by selecting a *strategy*: a function from the space of beliefs $\mathsf{Dist}(\mathcal{Q})$ to the space of actions $\mathcal{A}$. (We use the word "strategy" instead of the more common "policy" to avoid confusion with privacy policies.)