

When Arguing Definitions is  
Arguing Decisions

“Sup yinz, I’m Hazard! Today we’re going to use the idea of **words** as **decision rules** to help us understand and spot when people who seem to be arguing about **facts** and **definitions** are really arguing about how to make a **decision**. We’ll top it all off with some musing on what makes this problem worse and what to do about it.”



# Part 1

## Words as Decision Rules

# What's the deal with Manslaughter ?

## 18 U.S. Code § 1112 - Manslaughter

[U.S. Code](#) [Notes](#)

[prev](#) | [next](#)

**(a)** Manslaughter is the unlawful killing of a human being without malice. It is of two kinds:

Voluntary—Upon a sudden quarrel or heat of passion.

Involuntary—In the commission of an unlawful act not amounting to a felony, or in the commission in an unlawful manner, or without due caution and circumspection, of a lawful act which might produce death.

**(b)** Within the special maritime and territorial jurisdiction of the United States,

Whoever is guilty of voluntary manslaughter, shall be fined under this title or imprisoned not more than 15 years, or both;

Whoever is guilty of involuntary manslaughter, shall be fined under this title or imprisoned not more than 8 years, or both.

(June 25, 1948, ch. 645, [62 Stat. 756](#); [Pub. L. 103–322, title XXXII, § 320102](#), title XXXIII, § 330016(1) (H), Sept. 13, 1994, [108 Stat. 2109](#), 2147; [Pub. L. 104–294, title VI, § 604\(b\)\(13\)](#), Oct. 11, 1996, [110 Stat. 3507](#); [Pub. L. 110–177, title II, § 207](#), Jan. 7, 2008, [121 Stat. 2538](#).)

## 18 U.S. Code § 1112 - Manslaughter

U.S. Code   Notes

[prev](#) | [next](#)

(a) Manslaughter is the unlawful killing of a human being without malice. It is of two kinds:  
Voluntary—Upon a sudden quarrel or heat of passion.

Involuntary—In the commission of an unlawful act not amounting to a felony, or in the commission in an unlawful manner, or without due caution and circumspection, of a lawful act which might produce death.

(b) Within the special maritime and territorial jurisdiction of the United States,  
Whoever is guilty of voluntary manslaughter, shall be fined under this title or imprisoned not more than 15 years, or both;

Whoever is guilty of involuntary manslaughter, shall be fined under this title or imprisoned not more than 8 years, or both.

(June 25, 1948, ch. 645, [62 Stat. 756](#); [Pub. L. 103–322, title XXXII, § 320102](#), title XXXIII, § 330016(1) (H), Sept. 13, 1994, [108 Stat. 2109](#), 2147; [Pub. L. 104–294, title VI, § 604\(b\)\(13\)](#), Oct. 11, 1996, [110 Stat. 3507](#); [Pub. L. 110–177, title II, § 207](#), Jan. 7, 2008, [121 Stat. 2538](#).)

**Observable  
properties** needed  
to apply the **label**

**Behavioral  
consequences** of  
applying the **label**

The word “Manslaughter” encodes a decision rule! Once upon a time the U.S legal system decided that if someone acts like X, they will be treated like Y. That decision got encoded in law. When a case is brought to court, the process is simplified from “What should we even do about what happened?” to “What pre-existing decision rules are relevant to what happened?”



“Manslaughter”  $\approx$  IF [unlawful killing without  
malice] THEN [imprisoned  
for no more than 8 years\*]

**Label**  $\approx$  IF [Observable Properties]  
THEN [Behavioral  
Consequences]





# Another Example!

## § 11.411 Criminal trespass.

(a) A person commits an offense if, knowing that he or she is not licensed or privileged to do so, he or she enters or surreptitiously remains in any building or occupied structure. An offense under this subsection is a misdemeanor if it is committed in a dwelling at night. Otherwise it is a petty misdemeanor.

(b) A person commits an offense if, knowing that he or she is not licensed or privileged to do so, he or she enters or remains in any place as to which notice against trespass is given by:

- (1) Actual communication to the actor; or
- (2) Posting in a manner prescribed by law or reasonably likely to come to the attention of intruders; or
- (3) Fencing or other enclosure manifestly designed to exclude intruders.

(c) An offense under this section constitutes a petty misdemeanor if the offender defies an order to leave personally communicated to him or her by the owner of the premises or other authorized person. Otherwise it is a violation.



“Here, the **decision** is broken up across multiple stages of **labels**. Instead of **Criminal Trespass** directly having a prescribed **punishment**, it defers the **punishment** to **Petty Misdemeanor**.”

**IF**

**Knowingly entering  
an occupied  
structure without  
license**

**THEN**  
apply

**Criminal  
Trespassing**

**Public  
Lewdness**

**4th Degree  
Stalking**

⋮

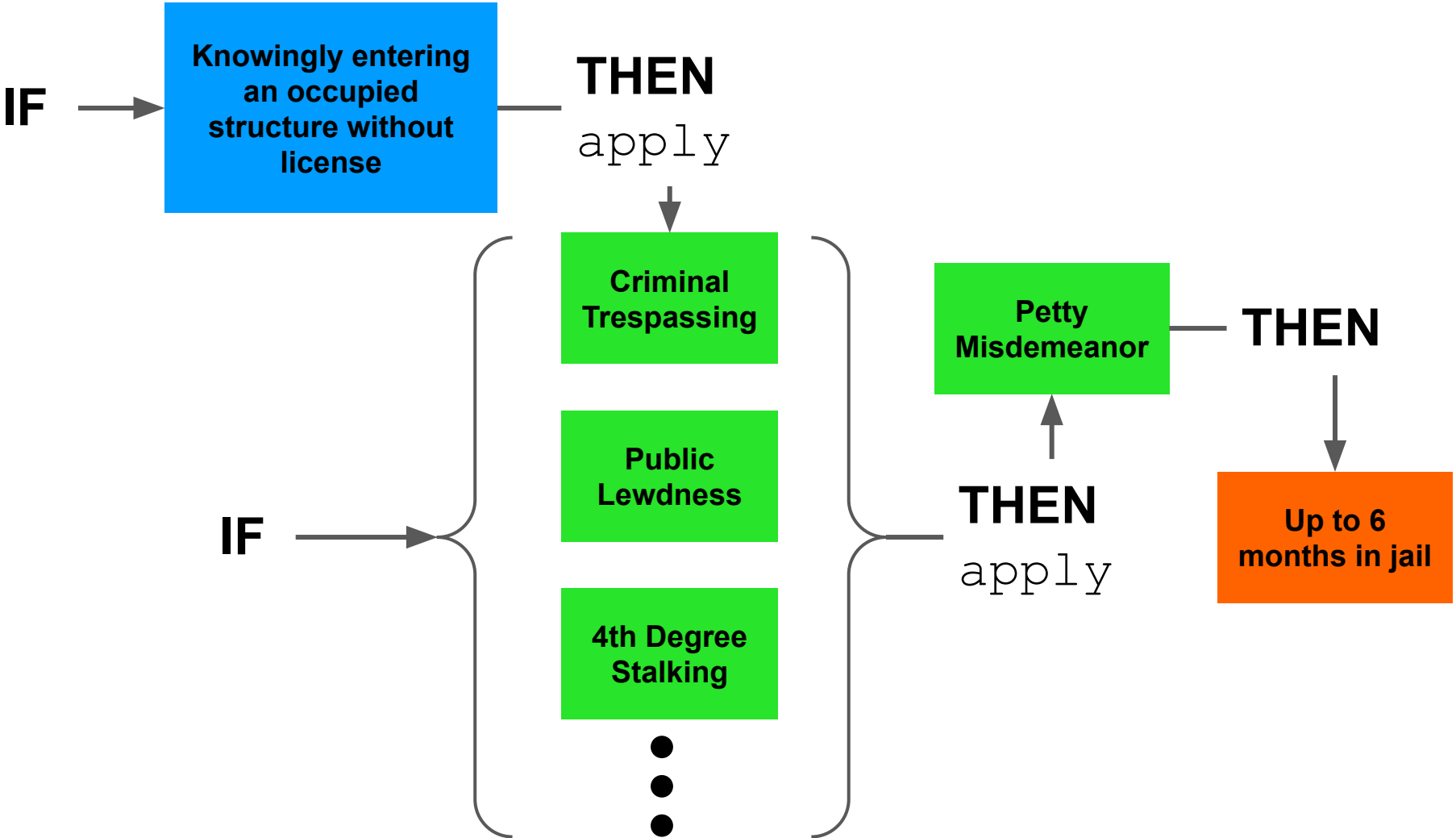
**Petty  
Misdemeanor**

**THEN**  
apply

**THEN**

**Up to 6  
months in jail**

**IF**



“**Labels** let you distribute your cognition across time. Even when there’s no need to **take action** in the moment, you are constantly making **observations** about things in your environment. Some of this **information** is used to preemptively **decide** how you will **act** in the future, and that **decision** gets cached as a **label**. When game time comes around, you just go straight into **action** because you’ve already made your **decision**. “

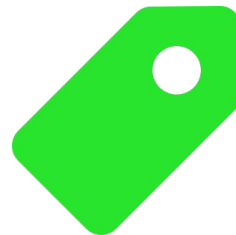


**t = now**

**IF**

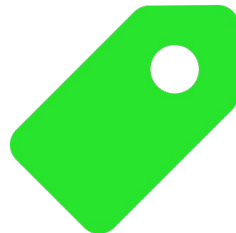


**THEN** apply



**t = later**

**IF**



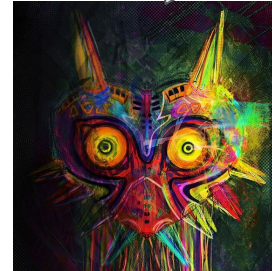
**THEN**



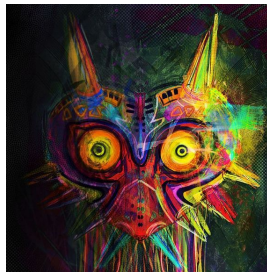
“Nooooooo! You can’t just use a few words from a specific legal context and extrapolate them to an entire cognitive theory of natural language!”



“Good point bro, now’s probably a good time to mention that this **isn’t** meant to be a ‘theory of language’. The goal is to point to a specific way words **can** work in specific contexts.”



“I started with two legal examples because they were such great examples of words encoding **decision rules** in the most *explicit* way possible. There are plenty of **contexts**\* where plenty of other more day to day **words** act as **decision rules**. But which **words**? In which **contexts**? Is this strictly a property of language, or does it apply to preverbal thought? How deep does this actually go?”



\*ooo a new color, what does it mean??

“**Concepts** are simply  
classes of **behaviorally**  
**equivalent** things.”

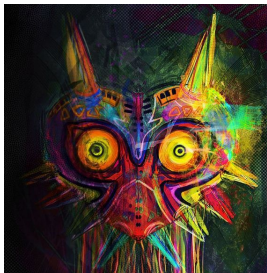


“Love that energy [Jaynes](#)! We’ll call this idea the **Behavioral Equivalence Hypothesis**. I don’t think we can take it as far as you do, but I think the idea is an *incredibly* useful North Star for navigating this space. Let’s follow it and see where it leads!”





“When I was a kid, I put trees into two **conceptual buckets**: **climbing-trees** and **non-climbing-trees**. I could rattle off a very long list of **membership criteria** for being considered a **climbing-tree**. But this **membership criteria** came chronologically after I’d already made the **category**! First I figured out **which trees I could climb** and **labeled** them accordingly. As I did that, I found **patterns** in what made for a **climbable tree**. These **patterns** got *compressed* into the list I’d share with you if you asked me to **define** what a **climbing-tree** was.”

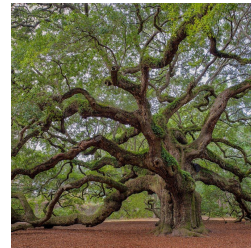
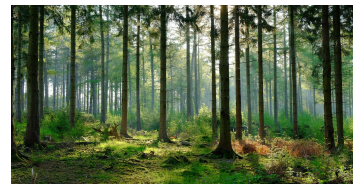


**IF**

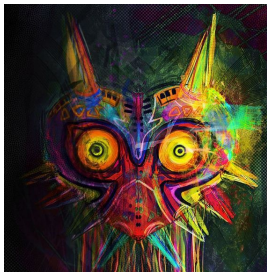
- Branches reachable from the ground
- Lots of winding branches
- Branches with gentle mostly horizontal angles
- Branches you can grab, not too small, not too big.
- No pointy nettles or leaves
- No jagged bark that hurts your hands

**THEN**

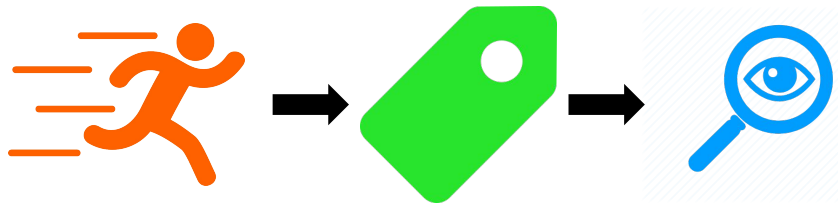
**Go climb it!**



“In this frame, **categories** are generated by a logic that flows in the opposite direction from how they are used. Creating a **category** is a process of **behaviorally equivalent things** getting **labeled** and then hunting for **patterns** in the **visible properties**. Using **categories** is a process of looking at **visible properties** and inferring how to **behave**. You **interact** with novel objects by **identifying their place** in your **conceptual schema**, and when that fails, your **conceptual schemas** update\*.”



**Creating  
categories**



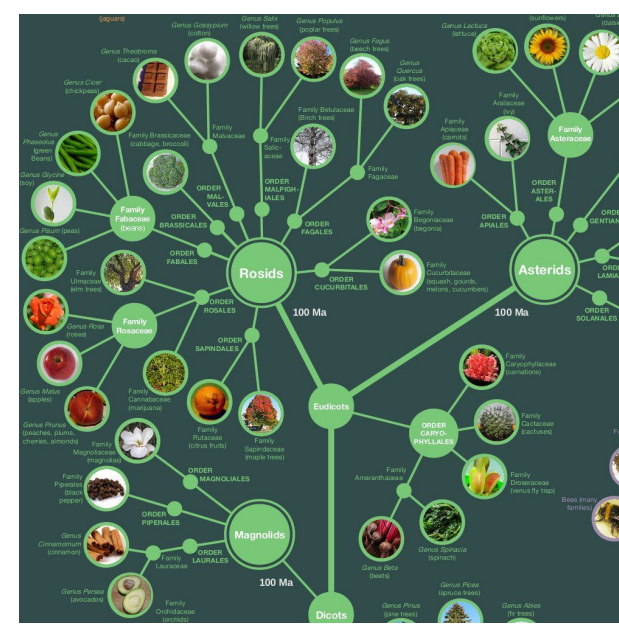
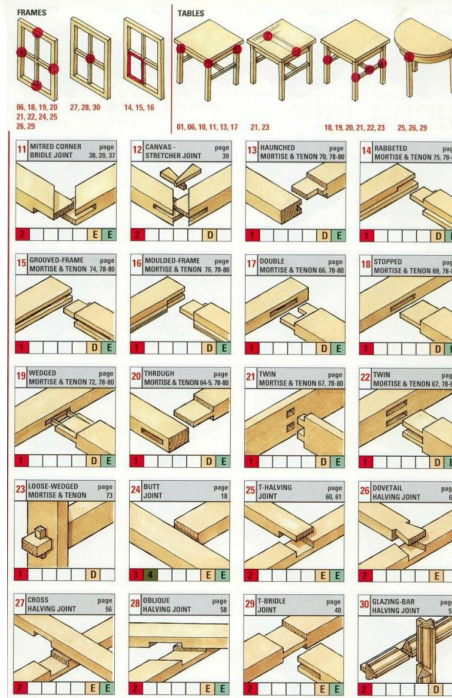
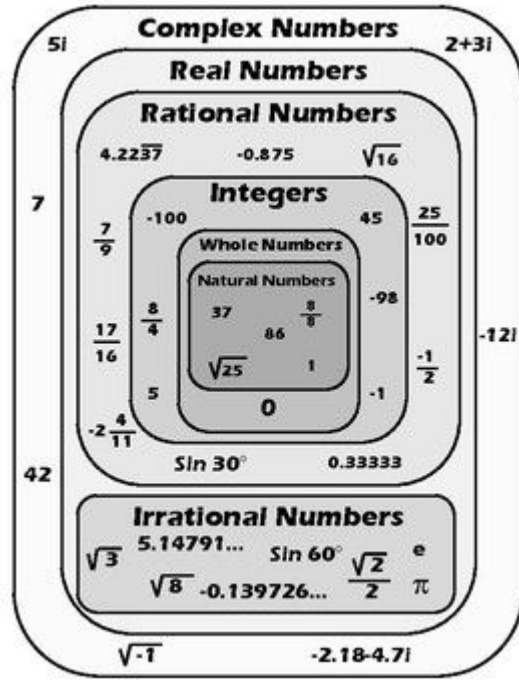
**Using  
categories**



\*sometimes. It's very interesting to look at when they do and when they don't. Different post!

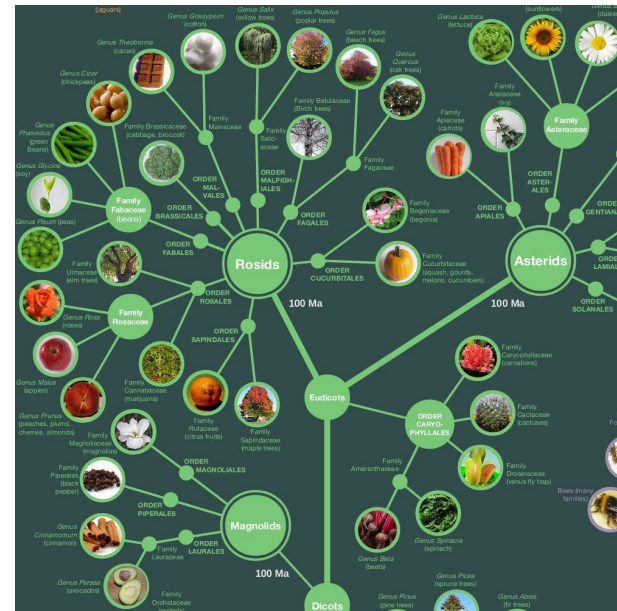
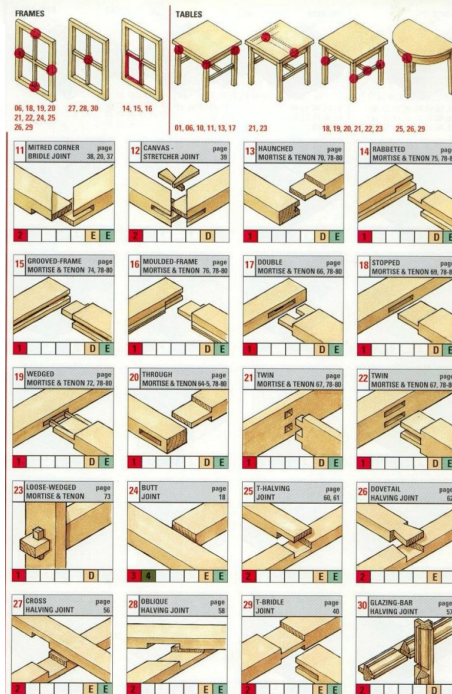
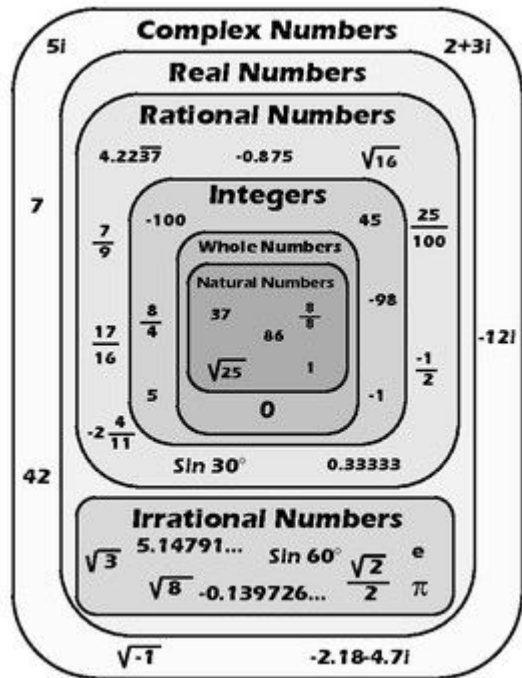
“In general, our **sensory resolution** is far greater than our **conceptual resolution**. You can discern more differences than you can find **meaningful**. Show me two leaves from the same tree, and I’ll be able to see that they are different, but the fact that **one is slightly longer than the other** isn’t **meaningful** to me in a way that will cause me to **label** the leaves differently. The **Behavioral** Equivalence Hypothesis is the idea ‘**meaningfully different**’ cashes out in ‘I need to **behave** differently towards these two things.’ **Action** is the difference that makes a difference. As you can see, this is a *much* broader idea than our original **labels** as **decision rules**.”



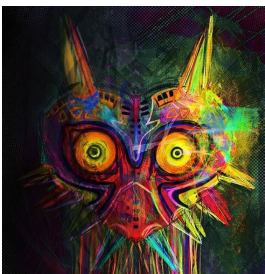


“Remember, the **Behavioral** Equivalence Hypothesis is a north star, not an iron-clad law. Jargon and technical language serve as great examples of the opposite end of the spectrum. Most jargon doesn’t have the same **decision rule** property that a lot of everyday language does.”





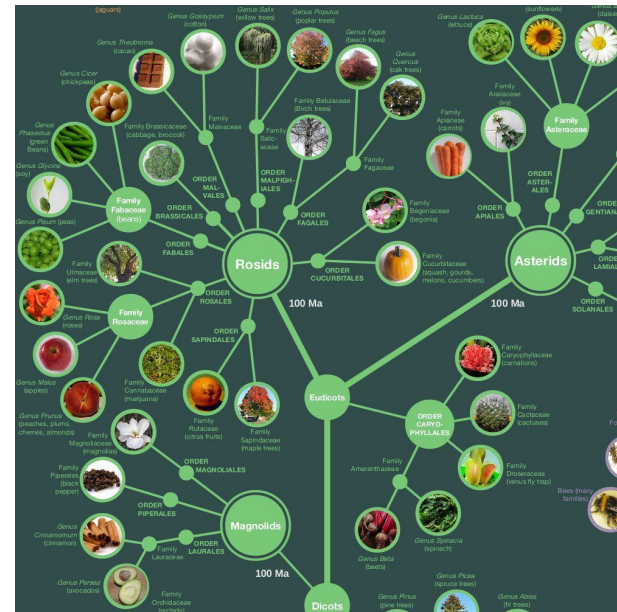
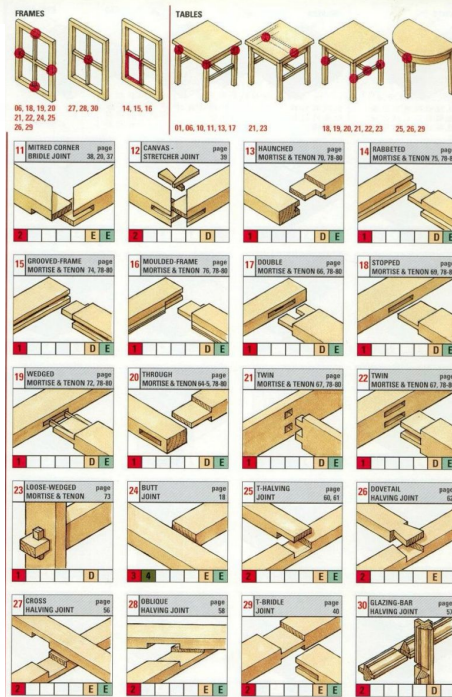
11



IF  THEN apply

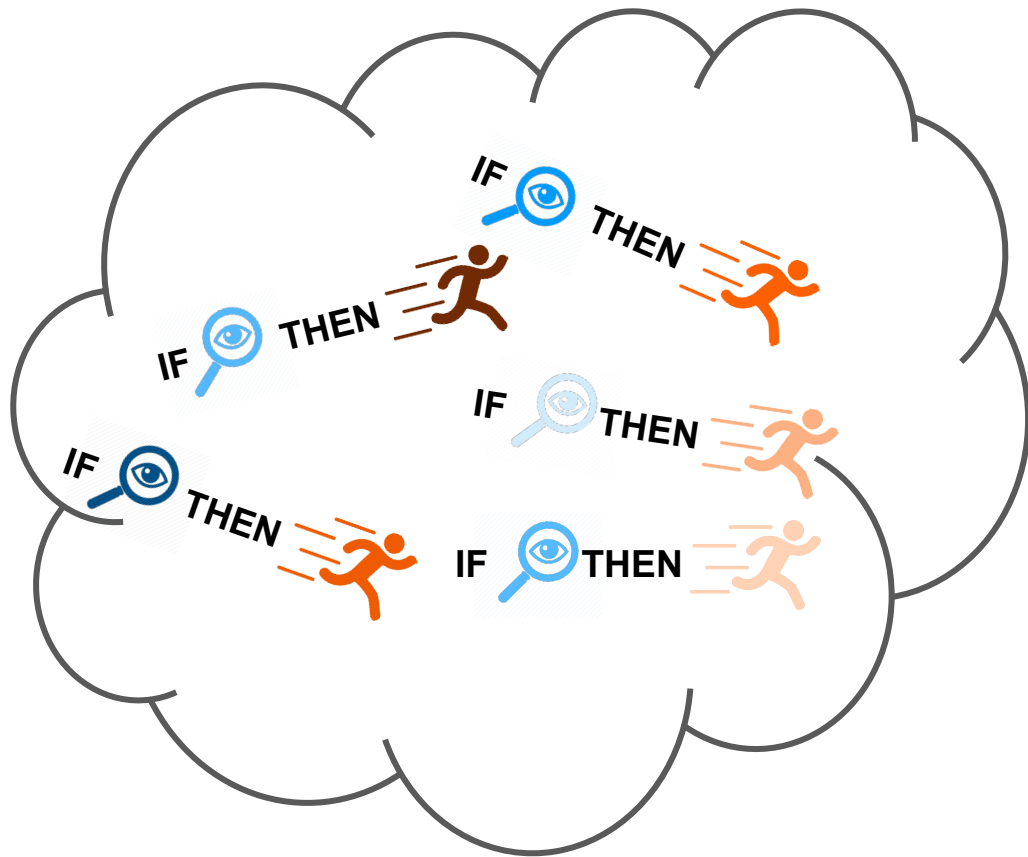


IF  THEN ???

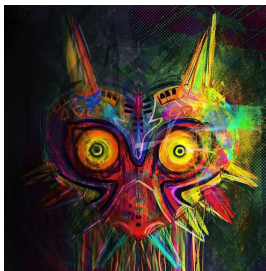


“The **definitions** get precise and crisp, which means that learning a **label** conveys a lot of **information**! That **information** can be useful for **action**, but the **label** no longer comes bundled with a **prefabricated action plan**.”

“Everyday language is often more **behaviorally** motivated than jargon, but that doesn’t mean it’s simpler! Any given **label** can contain multitudes! A **label** can encode any number of loosely related **decision rules**. Ones with **observable properties** that aren’t linked by necessary and sufficient criteria, but instead by fuzzy family resemblance. Once you have them, **labels** tend to accrue meaning, resulting in a sort of *conceptual sprawl*.”

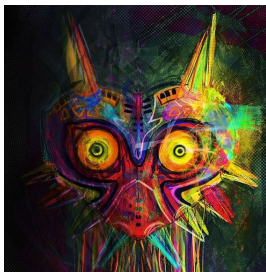


“If I may be so bold as to interrupt, it seems like you *are* in fact trying to develop a cognitive theory of natural language. It also seems like someone could understand all this and more if they just read [Words as Hidden Inferences](#), [How An Algorithm Feels From Inside](#), and really just the whole [A Human’s Guide to Words](#).”





“Ok, you caught me, I did lean into building a whole cognitive theory of language. I get a little carried away sometimes, this is some fun shit! I wanted to cover some of the most glaring caveats to the straightforward version of labels as decision rules, but that isn’t even my main point! All of this is in service of understanding a way I’ve seen communication between people break down, and what to do about it. Luckily, we’re done with part 1! Also, kids these days don’t read the sequences so—”





stealing  
someone's  
work  
for prestige

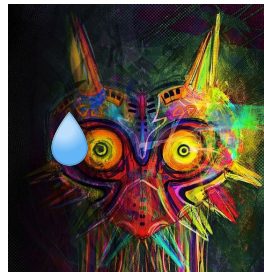


stealing  
someone's work  
to translate  
between vibes

# Part 2

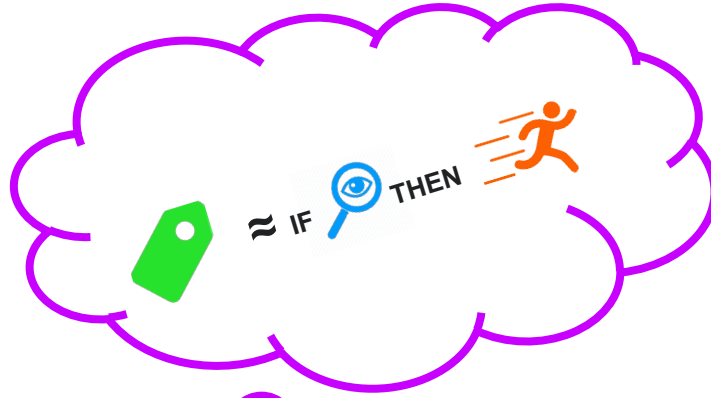
Arguing Definitions to Argue Decisions

“Hey kid! That’s a nice [View From Nowhere](#) you’ve got there.  
It would be a shame if someone  
were to... [contextualize](#) it.”

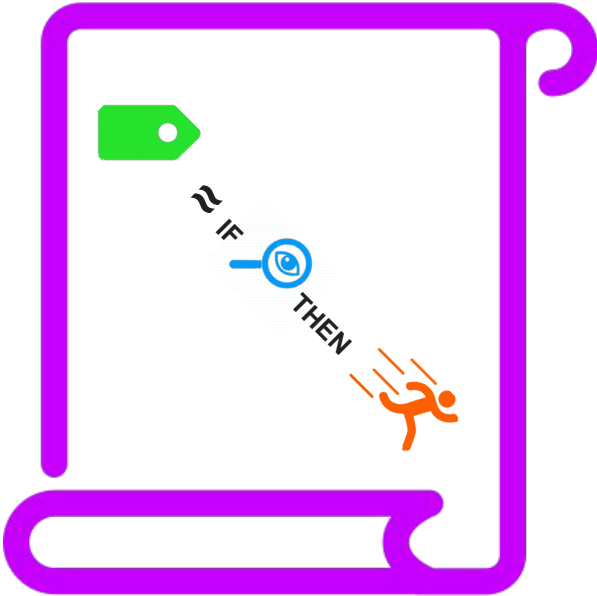


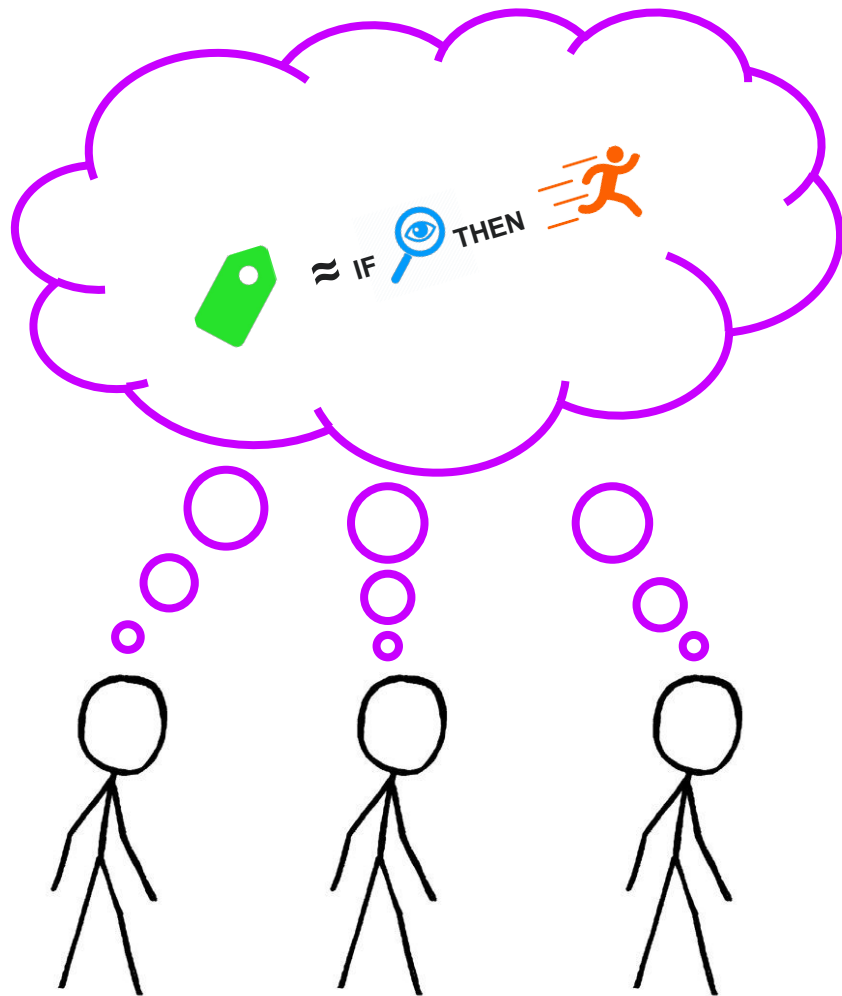
“Al Capone has a point. Up to now, you’d be forgiven if you thought we were treating the meaning of **labels** as platonic entities that just... exist... out there. No longer! A **decision rule** is always **embedded** in some **context**, and it is within said **context** where a **label** can be said to ‘**encode**’ a **decision rule**.”





“The very first examples we looked at were **decision rules** embedded in the **U.S legal system**. We also talked about **decision rules** that existed as habits of thought in an **individual's mind**.”





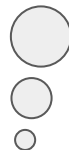
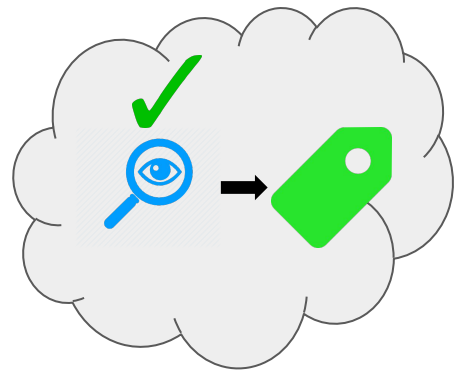
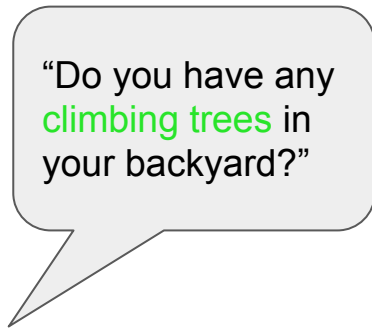
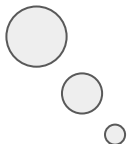
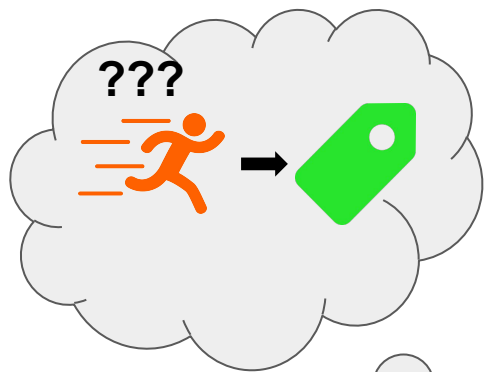
“A **decision rule** can also exist as a **social norm** in some **group** or **community**. It’s trickier to talk about norms, because it’s often really hard to understand what **group** you’re even in. Digital media allows people who are in close physical proximity to inhabit wildly different **reality tunnels**. This leads to all sorts of confusion, and it’s easy to assume another person is embedded in the same **social landscape** as you when they really aren’t.”

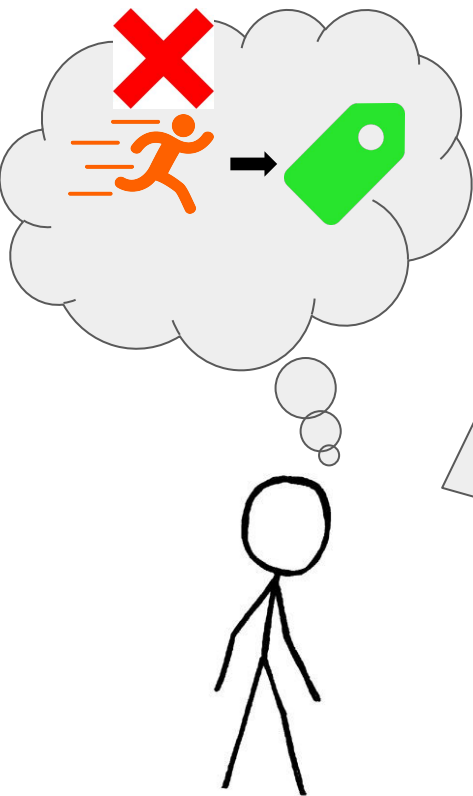


“It’s important to look at where a **decision rule** lives because what it’s **embedded in** determines how hard it is to *modify the rule*. Modifying a **law** is different from modifying a **social norm** is different from modifying **how you think** about things. Great, enough contextualizing, onto an example argument that bootstraps my next point!”









"WTF bro, that wasn't a **climbing tree!**"

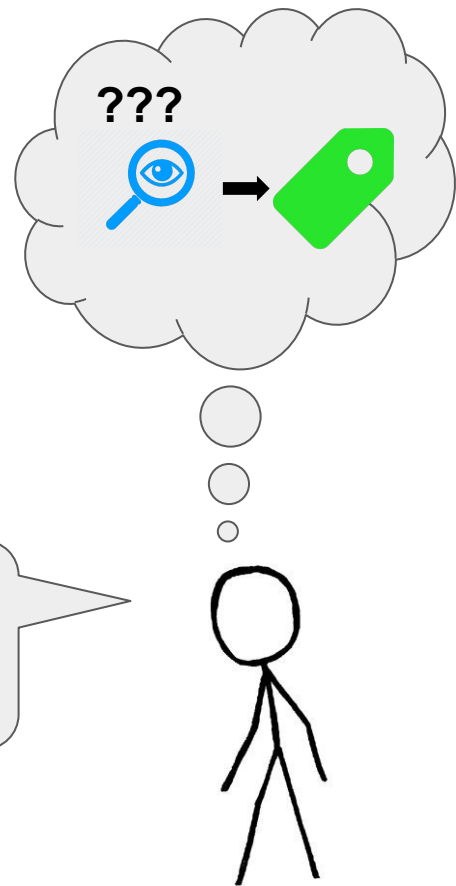
"Yeah, it **looked good**, but when I **jumped up to grab a branch** the whole thing snapped!"

"No, dude, I **tried every branch** and they're all like... **hollow and crumbling and shit**. Your tree literally doesn't have branches any more, they all fell off."

"What? Of course it is, didn't you **see those awesome low branches?**"

"Well duh, you have to **put your weight on the sturdy middle branches** not the **tiny ones**."

"Wait, wtf??? How.... **GODDAMN TERMITES!!!**"



“These two shared a **decision rule** that was embedded in the label **climbing-tree**. But that **decision** ignored the edge-case of termites! The **visible properties** didn’t account for the fact that termites can make a tree **unsafe to climb** while still **looking fine**. As soon as they started sharing **info**, they quickly realized that something was off, and updated their schema for **climbing-tree**.”

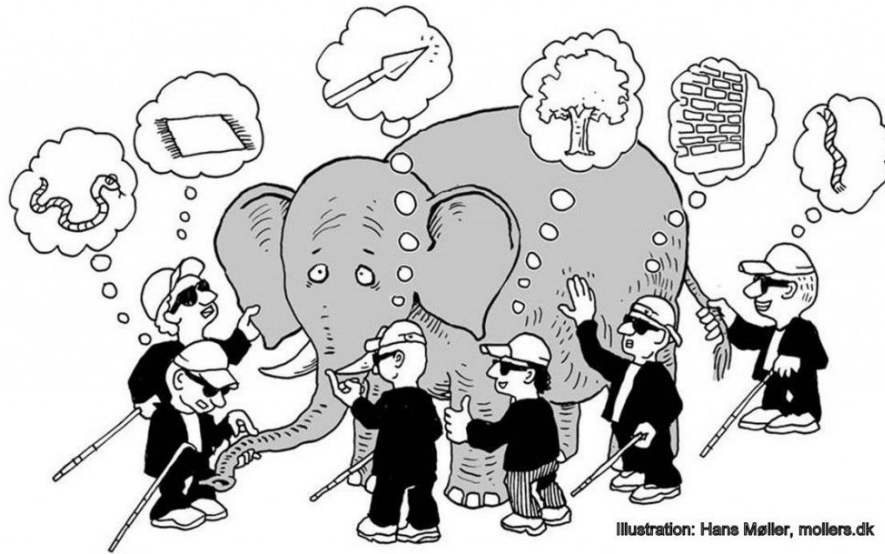
IF

- Branches reachable from the ground
- Lots of winding branches
- Branches with gentle mostly horizontal angles
- Branches you can grab, not too small, not too big.
- No pointy nettles or leaves
- No jagged bark that hurts your hands
- **You’ve kicked it to check for termites**

THEN

Go climb it!

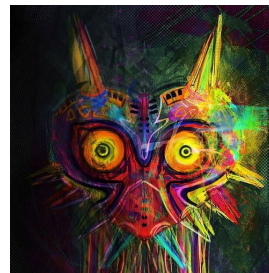




“This is what it looks like when an ‘Is X a Y?’ argument turns out to be fruitful. Different parties have access to different **information** which makes them apply a **label** in different ways. After exchanging **information**, they realize the old **decision rule** has an edge case, and they patch it. Arguments like this have the same energy as the fable of the Blind Men and the Elephant. Everyone is on the same team, and the main challenge is integrating siloed **information** from different perspectives.”



“The Elephant Model is just one possibly logic that might underlie a disagreement about **words**. What I’m more interested in rn are situations where real and substantive disagreements about how to **act** are masked by a focus on the surface level disagreement of ‘Is X a **Y**?’. This is what I call **arguing definitions to argue decisions**.”



Alex



“Yo, wtf? Yesterday when you fact-checked my presentation you said it was fine and didn’t need any edits. But my boss just spent 30m chewing me out over a bunch of inaccurate details! Why did you **lie** to me??”

“What? Of course you **lied**, I don’t know what else you could possibly call it.”

“Wow, relax, I didn’t ‘**lie**’ to you, okay? You need to chill out with these **accusations**.”

“Look, just because I **didn’t tell you every possible thing I could think of** doesn’t mean I **lied** to you.”

Blair



Alex



“Are you kidding? I mean, you **knew that some of those details were wrong**, right? And you still **decided to tell me it was fine**?”

“So you **knowingly told me something that wasn’t true**! That’s like, as clear an example of a **lie** as you can get!”

“Uncommon? Have you *looked* at a **dictionary**?”

“Okay, yeah but—

“Okay that is NOT common usage at all, that’s like the strictest most uncommon **definition** for **lying** I’ve ever heard. At *worst* what I said was kinda a bit of a **white lie**, but that’s not the same thing as **lying**.”

Blair



“At this point the conversation becomes an embittered debate over common usage of the word ‘lying’. Dictionaries are consulted. Neither side is happy. Nobody wins. What happened? The first thing that jumps out is that the conversation almost completely fails to mention **action** and **consequences**. Alex thinks Blair treated them badly, and wants Blair to **do things differently in the future**. Blair doesn’t think they did anything wrong. That’s what this conversation is being **driven by**, and yet the explicit context focuses on **definitions** and **labels**.”

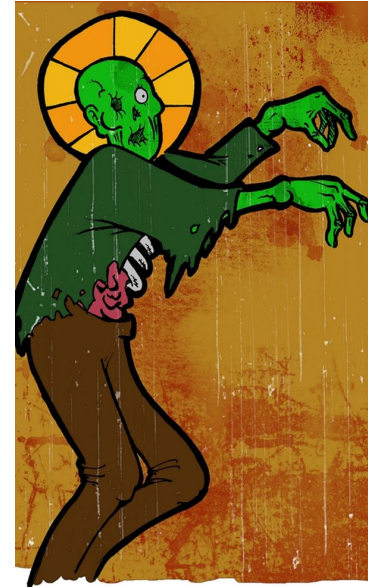


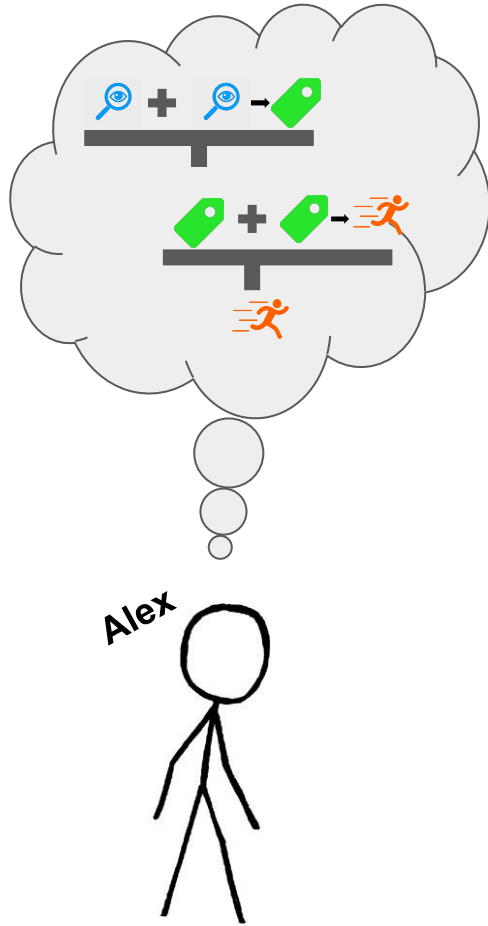


“Something nasty happens when arguments stop tracking the true generators of disagreement. They become **zombie arguments**; no matter how many points either party manages to refute, they *Just Won't Die*. And of course they won't! When you avoid your true disagreement, the only way a conversation can end is if one side bullies the other into submission, or both parties 'agree to disagree' and forget about it. The *conversation* can end, but the *disagreement* hasn't gone anywhere.”

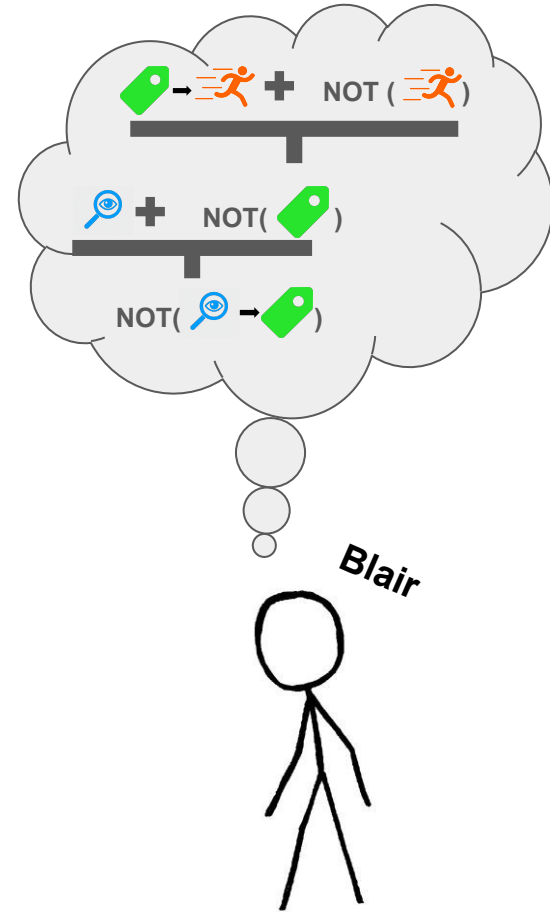


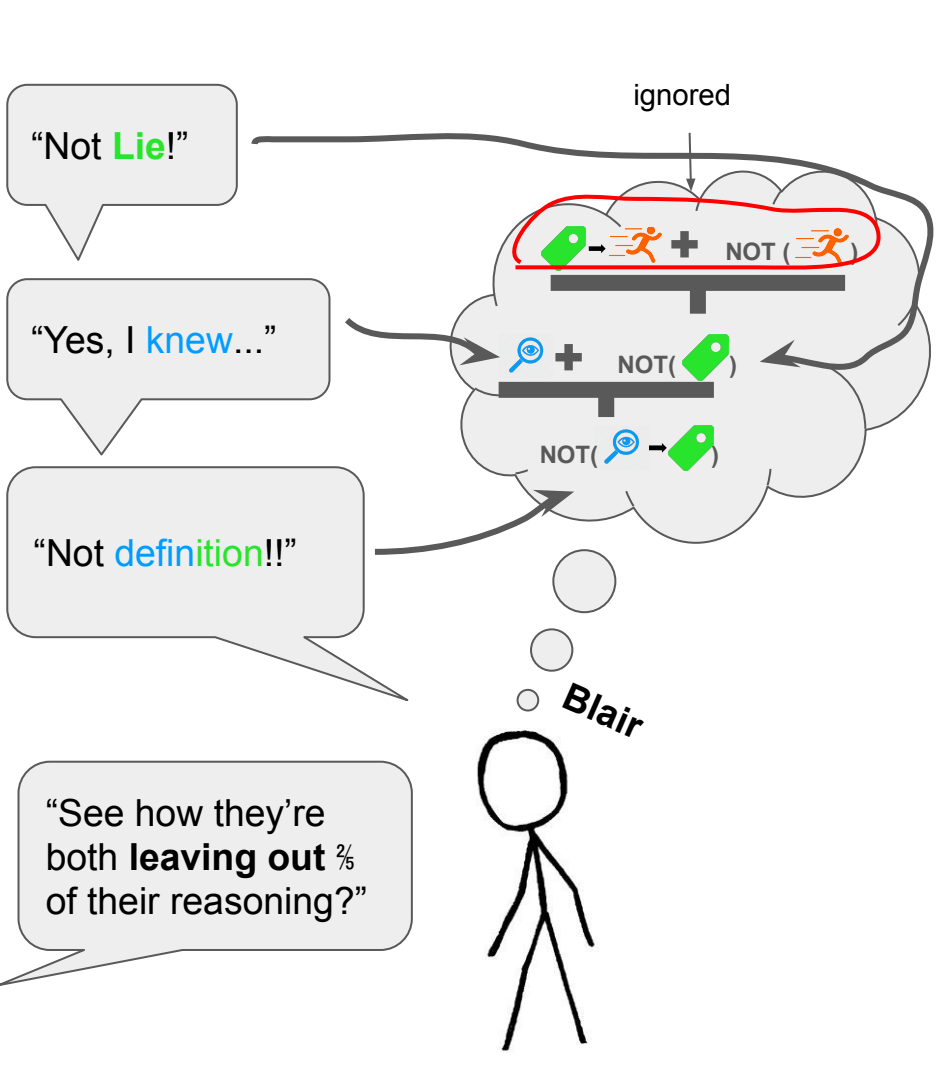
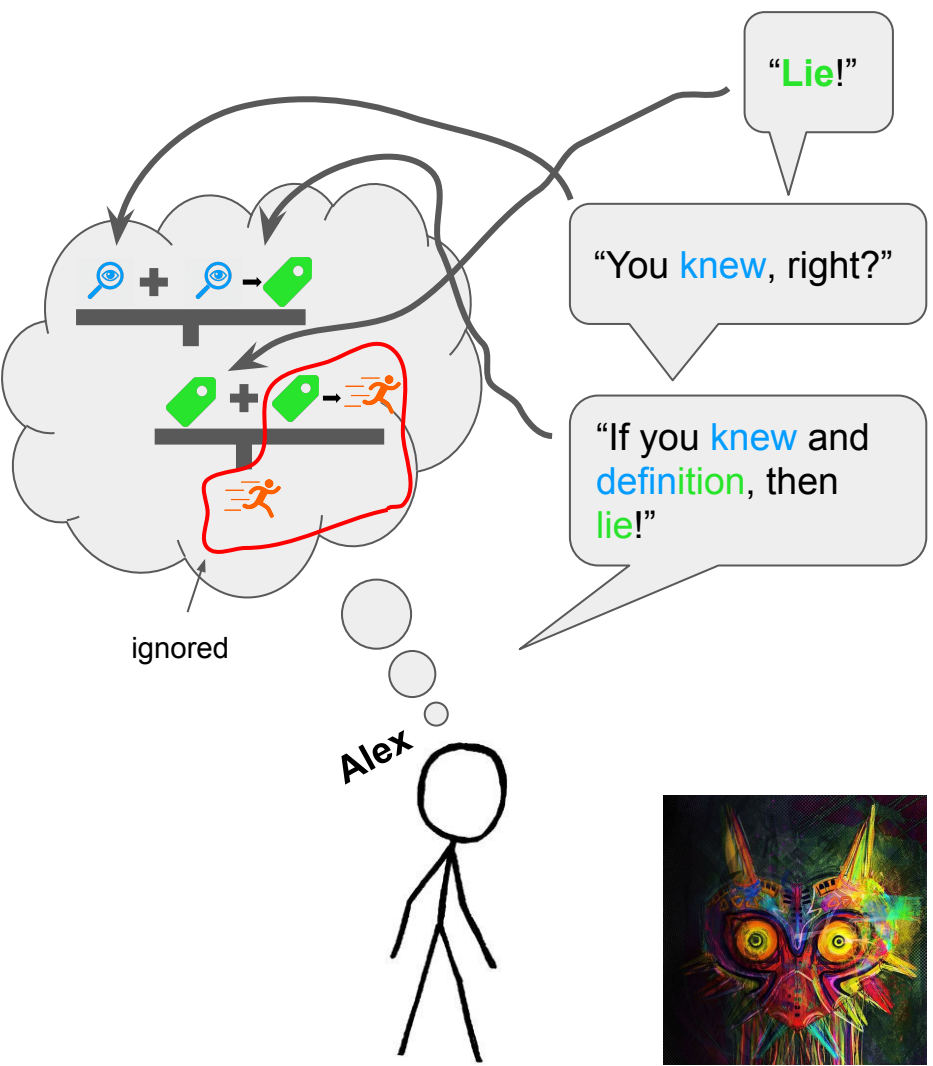
You when you aren't able to bring up the real generators of your disagreement.

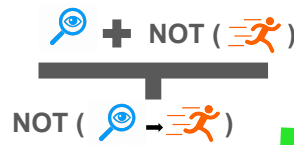
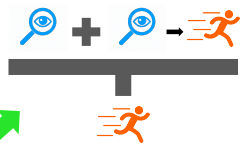




"I've turned Alex and Blair's private background reasoning into formal logical arguments with notation fudged to hopefully make it more intuitive if you haven't fucked with logic before. Up till now the arrows have had vague intuitive meaning; now they mean implication. These proof trees flow from top to bottom. Let's look at how this maps back onto the english version..."



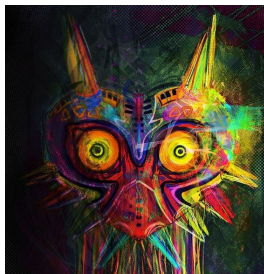





“More importantly, do you see how they only care about the **definition** because of how it plugs into the broader **decision rule** structure of the **label**? You can even factor out the **label** entirely and get something much more straightforward. What would it look like if they had this argument more directly?”

Alex

Blair






“Yo, wtf? Yesterday when you fact-checked my presentation you said it was fine and didn’t need any edits. But my boss just spent 30m chewing me out over a bunch of inaccurate details! Why didn’t you tell me? You must have known right?”

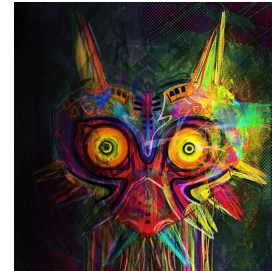
“Well, yeah I did notice you got several details wrong, but it didn’t seem like it was worth bringing up to you.”

“I really need you to not do that again, ok? This is really serious to me. I’m only able to do my job if I can count on you, the domain expert, to fill in the gaps and catch my mistakes.”



“To be fair, have you met the other people in this office? Half the people explode on me every time I correct them. We haven’t worked together long, and I saw no reason to risk getting exploded at just because it might be helpful to you.”

“The really interesting thing here is that Alex and Blair *still have a disagreement they need to negotiate*. Alex and Blair disagree about how Blair should **act** in this scenario, and this matters to both of them! Just because the direction their argument originally took was a dead end, didn't mean they didn't still have a meaningful disagreement they needed to settle. Seeing **words** as **decision rules** isn't meant to make all of your arguments with others dissolve, it's meant to help you *focus on what matters*.”



“The is the *behavior* of **arguing definitions** to argue a **decision**. Hopefully you get the gist of this pattern and can check it against your own experiences. As I've described it so far, doing better is mostly a matter of... well, not falling for it. Knowing about it helps protect you. It's something you can get past with awareness and practice. But sometimes that's not enough. Sometimes there are explicit pressures that keep dialog confined to the 'Is X a Y?' level. These pressures are much harder to deal, and of course that means they're much more relevant to day to day life.”



# Part 3

Immutable Laws and Doomsday Buttons



**“Quick recap:** sometimes ‘Is X a Y?’ is a Blind Men and the Elephant situation that resolves itself if people just exchange private **information**. Other times ‘Is X a Y?’ is a distraction from a real disagreement that would benefit from being addressed more directly. In all the examples we’ve done so far, the disagreements have been between two individuals. Things get trickier when we jump to disagreements that are happening inside a larger **social context**, where the **decision** to be made requires the assent of the **group**, and not just two people.”



“Remember that **labels** don't live in a vacuum!  
Updating a **decision rule** requires being able to act  
on the **substrate** that the rule is **embedded** in.  
Changing **how you think** is a wildly different task  
from changing **how a group thinks** is a wildly  
different task from changing **how the law thinks**.”



“One very illustrative example comes from the [Rwandan genocide](#). Between April 7th and July 15th, around 550,000 members of the Tutsi minority ethnic group are killed by armed Hutu militias. As the **genocide** was ramping up, various activist groups and other individuals called for U.S intervention. How did the president at the time, Bill Clinton, respond?”



“The horrors of **civil war** and mass killings of civilians in Rwanda, since the tragic deaths of the Rwandan and Burundian Presidents 3 weeks ago, have shocked and appalled the world community.

On behalf of all of the American people, I call on the Rwandan army and the Rwandan Patriotic Front to agree to an immediate ceasefire and return to negotiations aimed at a lasting peace in their country.”<sup>1</sup>



<sup>1</sup> [Radio address](#) given on April 30th, 1994

“Civil war? It’s true that the Rwandan army and the RPF were fighting at this time<sup>1</sup>, most of the deaths came from roving civilian groups killing their neighbors with machetes. Internal documents that were later declassified<sup>2</sup> revealed that the Clinton admin had been privy to most of the relevant on the ground information. Why didn’t they want to call what was happening a genocide when it clearly fit the definition? Well, to be fair, we should look at the definition first.”



<sup>1</sup> I’d recommend the documentary [Ghosts of Rwanda](#) and the book [Machete Season](#) if you want to learn more about what happened.

<sup>2</sup> I didn’t succeed in tracking down the original docs, but [this](#) interview has a lot of context and quotes that lay out a pretty solid case.

“This looks familiar!  
It’s a legal **definition**  
that explicitly  
designates this **label**  
as encoding a  
**decision rule**.”



## Article I

The Contracting Parties confirm that **genocide**, whether committed in time of peace or in time of war, is a crime under international law **which they undertake to prevent and to punish**.

## Article II

In the present Convention, **genocide** means **any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group, as such:**

- (a) Killing members of the group;
- (b) Causing serious bodily or mental harm to members of the group;
- (c) Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in part;
- (d) Imposing measures intended to prevent births within the group;
- (e) Forcibly transferring children of the group to another group.

## Article III

The following acts shall be **punishable**:

- (a) **Genocide**;
- (b) Conspiracy to commit **genocide**;
- (c) Direct and public incitement to commit **genocide**;
- (d) Attempt to commit **genocide**;
- (e) Complicity in **genocide**.

....

From the Genocide Convention of 1948, there are several more articles specifying things like how international courts are supposed to work, and what “punishment” entails.

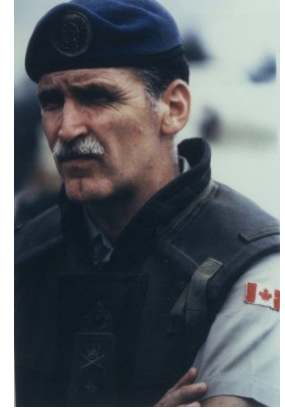
"If you had access to the raw **facts** of the situation, they clearly match everything outlined in **Article II** of the **definition** of **genocide**. It wasn't that the Clinton administration disagreed on the **facts**, it was that they **didn't want to intervene**<sup>1</sup>.



<sup>1</sup> Just a year earlier the U.S had been badly burned with an attempted intervention in a Somalian civil war. To be clear, I don't think this was a good reason to not intervene in Rwanda, I think they made the wrong move. I'm not justifying their decision, I'm trying to explain logic of the frame that they made their decision within.

Believes there is targeted killing of an ethnic group.

Believes this is something the U.S should be responsible for intervening in and preventing.





“Instead of trying to renegotiate the **decision rule** that was **embedded** in that **1948 treaty**, Clinton maneuvered around **labels** such that he would avoid being pressured to action. This wasn't a result of getting momentarily sucked into a linguistic trap in the heat of the moment because he never practiced a specific skill. He and his cabinet had **goals**, they understood the **social context** they were in and figured that trying to update an **50 year old international treaty** would be way harder than just trying to reframe the events in **language** that wouldn't let others pressure them with appeal to the existing **decision rules** as easily.”





“To circle back, Arguing **definitions** to argue **decisions** is what happens when some or all parties involved are **motivated** by the question of ‘**What should we do?**’ but the argument does not transcend the level of ‘Is X a **Y?**’ People go after what they want by trying to get what you want lawyering a **label** to trigger a **decision rule** in their favor. A **label** is no longer being used either as a cognitive shorthand, nor to convey **information**. Instead, it’s treated as a shortcut, a **lever** that is **embedded** in the law of the **social norms**, such that if you can just yank **it** in the direction you want, the **group** will be **compelled** to **action**.”



“Importantly, the game shifts away from convincing **particular people** of particular things, and into convincing **the group** of something. I will never succeed in overstating how different these two things are. A **group's mind** can change without a **single person** changing their mind. **Every individual** could change their mind but the **group's mind** might not change. Have you noticed there is a thing, ‘*winning an argument*’ which has nothing to do with being right, and nothing to do with convincing another person? As I see it, ‘winning’ an argument seems to be grounded in ‘would an imaginary **audience** declare that I’ve won?’ This is weirdly recursive! How does this audience **decide** if you’ve ‘won’? By guessing if everyone else thinks that you’ve ‘won’! **Group consciousness** is this spooky shit that’s decided by what other people think other people think other people think. Things that are supposedly *about* reality can become untethered very, very fast.”



*\*sigh\**

“This piece needs to be put to rest sooner rather than later. **Group minds** vs **individual minds** is such a rich topic, but I’m going to put that aside so I can get out of here and on with my life. I noted that unlike Alex and Blair, the way Clinton played with **definitions** is clearly more intentional, and having Clinton read this post on **Words** as **Decision Rules** likely wouldn’t have affected how the whole international incident played out. Something happens in group settings that encourages staying at ‘Is X a **Y**’, and it gets stickier as you get bigger and bigger groups. I’m going to hastily wrap up by outlining one combo of pressures that conspire to keep arguments at the ‘Is X a **Y**’ level.”



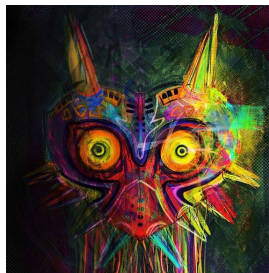
“If we looking at the **decision rules embedded** in the norms of a certain **social landscape**, we can ask about the qualities of these **decision rules**. How are they invoked, what sort of **outcomes** result from invoking them? A **decision rule** has a **doomsday button** quality when its course of **action** is both extreme and totalizing, and there is little to no ability to use it precisely, add nuance, incorporate context, or manage exceptional cases. Imagine being a moderator for a debate where your only moderation power is to **end the debate**. No fined tuned adjustments, no warnings, just nuking the whole debate.”



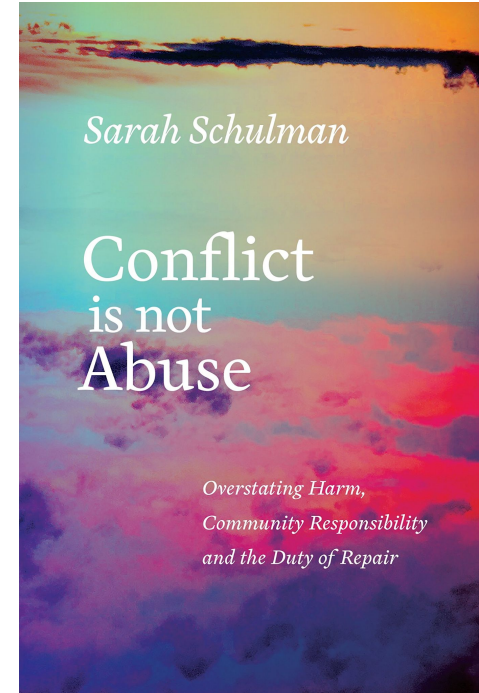
“Doomsday buttons allow for less nuance but more of a certain type of certainty. In the McCarthy era, if you ended up labeled as a **Communist**, getting fired pretty **mechanistically** followed. It may be a fight to get the **label** to stick or not, but once it stuck, the **outcome** was **determined**. Contrast **doomsday buttons** with more flexible **decision** making processes (say, an assembly is convened and people talk it out until there’s a clear  $\frac{2}{3}$  in favor of some course of action). Such a group has the potential to use more contextual **information** to synthesize a **best move** for the situation. At the same time, you have less certainty about **what will be done** before it’s decided, *precisely* because of how flexible the process is.”



“Doomsday Buttons aren’t the only factor that lead to the increasing prevalence and stickiness of Arguing Definitions to Argue Decisions. Another crucial puzzle piece is each individual’s felt sense of social agency. If you don’t see yourself as having the ability to update the decision rule that a label encodes in your social context, then it starts to feel like lawyering labels is the only way to make anything happen. There’s lots of ways to end up feeling like you don’t have the power to either modify the norms, or invoke more high context group decision making procedures. Maybe it seems like only people of certain status can affect the norms. Maybe you don’t trust the group to use the higher context procedures fairly. Maybe you think the game is rigged. Maybe your group doesn’t have, and never had a capacity for nuanced decision making. Or maybe, for unrelated reasons, your history on this earth has built in you a persuasive felt sense you are not in control and that the world you are embedded in is static and unyielding.”



“I’m gonna lean on [Sarah Schulman](#) to synthesize these points. She explores how impoverished community **decision** making tools can put people in situations where they feel they can only **get help** and **support** by making their problem fit a specific **label**, creating pressure to distort the **facts** in order to **get help**. She also explores how, when conflict feels like this awful unbearable thing, the problem is exacerbated. She digs into how both trauma and supremacy can lead someone to experience conflict as intolerable (the former from triggering traumatic memories, the later from existential threat to the ego), which leads both sorts of people eager to press **doomsday buttons** because it will end the conflict and stop the pain.”





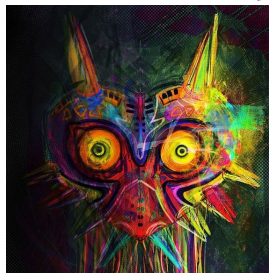
“All at once. Pretend you live in a **norm landscape** with little to no high context **decision rules**. It's **doomsday buttons** everywhere. You don't trust anyone else. And trying to hash out disagreements with people in an ad-hoc high context mode is painful or uncomfortable. You have things you want, you are in pain, and you want to *make it stop*. There's a **doomsday button** that will make it stop. So you reach for it. The person you're arguing with *doesn't want a doomsday event*, so they fight back ferociously. The increased conflict is painful for both of you, making you both more eager to trigger any **label** that will make the argument stop right now. Your mind is totally occupied with the question of 'How do I make this **label** stick?' Everything else falls away. You switch from what is true to what is **plausible**. Maybe it stops at motivated reasoning. Maybe it's intense enough that you slide into distorting and obscuring the **reality** of the situation you're arguing about.”





\*brief palette cleanser\*

“I’ll end with a brief reply to ‘**So what?**’ Well, I’d be elated simply if anyone is better able to spot these dynamics in their life. But I also hope that you can find ways to avoid ending up in **label** warfare. The behavior of arguing **definitions** to argue **decisions** is ultimately parasitic on language as a tool for communication<sup>[true, but a citation would help]</sup>. And unfortunately, this is still true even if you didn’t start it. If someone comes at you playing a **label** game, and you defend yourself from within the **label** game, language is still being diluted. Maybe you had good reason to do it, but still, the language dillutes. Growing up, it seemed like everywhere I looked the arguments people were having were nothing but trying to trigger **decision rules** encoded in **labels**. It was a decade or so before I was able to articulate what I saw happening, but I could always feel it. It’s not too far fetched for me to imagine an alternate timeline where I concluded that’s all language was. A brute tool to beat others with. I can imagine a timeline where I never learned about using language to be precise, to share information, and to authentically communicate with others. It’s a sad and dark timeline.”



“That being said, I can't honestly say that engaging this way should never be done. I don't understand too much about how to enact my values in hazardous **social settings**. Sometimes stuff that **matters to me** is threatened, and it's hard to protect it and be virtuous. A lot of my clarity on this topic has come from the fact that I've stumbled into having enough slack that I can exit many **social contexts** where I'd be forced to make this trade-off. So while I see plenty of individual instances where I will fight to prevent arguing **definitions** to argue **decisions**, I'm not trying to build up a general case for why you should be blameworthy if you engage in behavior like this. Hopefully this post gave you some handles to more easily think about the behavior, about the costs, and maybe seeding future brainstorming about how to protect your values while not diluting the commons.”



This is the end