

# Logic and distributive laws

The twisted monads



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Agenda

- Terminology and level set
- Short break
- Definitions and examples
- Short break
- Hands-on calculation
- Short break
- Hands-on calculation
- Wrap up

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Aims and measures of success

Aim: to give this idea away!

Measure of success:

- other people can make the calculations
- other people can write down, evaluate and extend the core algorithms
- other people understand the motivation for this framework of calculation



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Aims and measures of success

What is presented is model-theoretic.

Extra credit: someone (besides me) presents the extension of the algorithm to the proof-theoretic

i.e. they write down the operation of the algorithm on morphisms

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Motivation

(Behavioral and structural queries in physical systems) Find all the pathways,  $p$ , in  $\langle \text{SignallingPathwayDB} \rangle$  such that when  $\langle \text{Reagent} \rangle$  is added in concentration  $K_p$  eventually reaches states in which we see these  $\langle \text{Reagent1} \rangle, \dots, \langle \text{ReagentN} \rangle$  in concentrations  $K_1, \dots, K_N$ .

(Behavioral and structural queries in logical systems) Find all the implementations of message queue in <http://svn.myrepo.org> such that the queue is FIFO and use the Xerces XML parser for XML-encoded messages

(Structural queries in physical systems) Find all the knots in  $\langle \text{KnotDB} \rangle$  that have alternations of  $\langle \text{Trefoil} \rangle$  and any 7-crossing knot as subknots

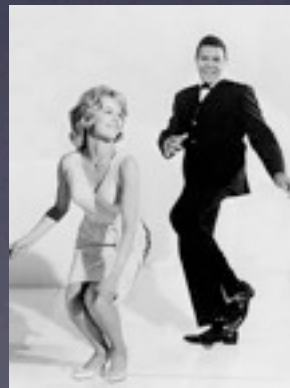
(Geometric queries in physical systems) Find all the conduit wires running from the nose of the aircraft to the midsection that are likely to undergo a twist between source and destination

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Terminology & level setting

Claim I : A logic is comprised of three data

- A monad,  $T$ , for a term language (the witnesses of propositions)
- A monad,  $S$ , for collecting witnesses
- A distributive law  $l : TS \rightarrow ST$  illustrating how terms built over collections (formulae) are interpreted as collections of witnesses



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Terminology & level setting

Part of our mission is to develop enough common understanding that this claim makes sense and can be independently verified

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Terminology & level setting

Claim II : A model-checker is a database turned sideways, that is, given

- a logical language,  $L$ , and
- a model-checker for checking assertions of the form  $\dagger \models I$
- a store of terms,  $src$

we can describe an algorithmic approach to selecting all the terms in  $src$  satisfying  $I$ , that is interpreting statements of the form

`SELECT  $\dagger$  FROM  $src$  WHERE  $I$`

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Terminology & level setting

Part of our mission can be nearly immediately discharged: there's an obvious brute-force algorithm that will do this. What is it?

Another part of our mission is to convince you that there are much more efficient algorithms that compile the model-check into a query plan over a programming language suitably extended with the query operations of the native query language of an industrial scale store.

That is, part of our mission is to convince you that there's an extension of LINQ that makes the observation useful for industrial scale applications.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Terminology & level setting

Claim I : A logic is comprised of three data

- A monad,  $T$ , for a term language (the witnesses of propositions)
- A monad,  $S$ , for collecting witnesses
- A distributive law  $l : TS \rightarrow ST$  illustrating how terms built over collections (formulae) are interpreted as collections of witnesses

So, we have to know enough category theory and standard categorical interpretation of logic to unpack this statement.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Terminology & level setting

Show of hands, please:

You can come to the whiteboard, write down a definition of

- a category
- a functor
- a natural transformation
- a monad
- a distributive law

and present it to us



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Terminology & level setting



Show of hands, please:

- how a free algebra is represented by a monad
- how non-free versions are related to the Kleisli and/or Eilenberg-Moore algebra of a monad

extra credit

- how is this related to Universal Algebra and standard presentations of algebras via generators and relations

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Terminology & level setting

Show of hands, please:

- the standard categorical interpretation of logic
- the propositions-as-types or Curry-Howard paradigm



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Terminology & level setting

Along the way we will make use of a combination of tools

- EBNF
- Structural equivalence
- Reduction relations

These are the modern computer scientist's way of writing down an algebra  
+ a structure map together with a rewriting form of computation

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

Defn: A category is ...

Defn: A functor is ...

Defn: A natural transformation is ...

Defn: A monad is ...

Defn: An algebra for a monad is ...

Defn: A distributive law is ...

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

Just a thought: Could monads be the right computational unit for category theory?

Spse  $C$  a(n arrows-only presentation of a) category, then

$|C| =$

```
for( morph <- C ) yield { src( morph ) }
```

```
++ for( morph <- C ) yield { trgt( morph ) }
```

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

First example: deriving separation logic from this approach

Defn: A monoid is

- a set,  $M$ , together with
- a distinguished element,  $e$
- a binary operation  $\_ * \_ : M \times M \rightarrow M$

satisfying

- $e * m = m = m * e$
- $m_1 * (m_2 * m_3) = (m_1 * m_2) * m_3$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

An alternative presentation

Grammar

$m, n ::= e$

$g1 \mid \dots \mid gN$

$m * n$

Structural equivalence (the smallest equivalence relation containing)

$$m * e = m = e * m$$

$$m1 * (m2 * m3) = (m1 * m2) * m3$$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

$k \ \& \ l ::=$

$\text{true} \quad [ \mid \text{true} \mid ](v) = L(m)$

$\sim k \quad [ \mid \sim k \mid ](v) = L(m) \setminus [ \mid k \mid ]$

$k \ \& \ l \quad [ \mid k \ \& \ l \mid ](v) = [ \mid k \mid ](v) \setminus \cap [ \mid l \mid ](v)$

$\underline{e} \quad [ \mid \underline{e} \mid ](v) = \{ m \setminus \text{in } L(m) \mid m = e \}$

$g_i \quad [ \mid g_i \mid ](v) = \{ m \setminus \text{in } L(m) \mid m = g_i \}$

$k * l \quad [ \mid k * l \mid ](v) =$   
 $\{ m \setminus \text{in } L(M) \mid m = m_1 * m_2, m_1 \setminus \text{in } [ \mid k \mid ](v), m_2 \setminus \text{in } [ \mid l \mid ](v) \}$

$\text{rec } X.k \quad [ \mid \text{rec } X.k \mid ](v) = U\{ S \setminus \text{in } \setminus \text{Pow}(L(m)) \mid [ \mid k \mid ](S/X:v) = S \}$

$X \quad [ \mid X \mid ](v) = v(X)$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

A simple use

$$\text{Prime} = \sim e \ \& \ \sim(\sim e * \sim e)$$



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

- What has this got to do with category theory?
- Where are the monads?
- Where are the algebras?
- Where is the distributive law?
- How does this relate to Universal Algebra?
- How does this relate to Curry-Howard?

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

An alternative view of the lambda calculus

Grammar

$$m, n ::= 0 \mid @x \mid \backslash x \rightarrow m \mid m \ x$$
$$x ::= \langle\langle m \rangle\rangle$$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

An alternative view of the lambda calculus

- Structural equivalence: alpha-equivalence + a little more
- Reduction relation depends on a slightly different notion of substitution

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

$k, l ::= \text{true}$

$\sim k$

$k \ \& \ l$

$[ \mid \text{true} \mid ](v) = L(m)$

$[ \mid \sim k \mid ](v) = L(m) \setminus [ \mid k \mid ]$

$[ \mid k \ \& \ l \mid ](v) = [ \mid k \mid ](v) \cap [ \mid l \mid ](v)$

$\setminus a \rightarrow k$

$[ \mid a \mid ](v), m' \in [ \mid k \mid ] \}$

$[ \mid \setminus a \rightarrow k \mid ](v) = \{ m \in L(m) \mid m = \setminus x \rightarrow m', x \in$

$\langle k \rangle l$

$m', m' \in [ \mid l \mid ] \}$

$[ \mid \langle k \rangle l \mid ](v) = \{ m \in L(m) \mid \exists n \in [ \mid k \mid ]. km \rightarrow$

$\text{rec } X.k$

$S \}$

$[ \mid \text{rec } X.k \mid ](v) = U\{ S \in \text{Pow}(L(m)) \mid [ \mid k \mid ](S/X:v) =$

$X$

$[ \mid X \mid ](v) = v(X)$

$a, b ::= \langle \langle m \rangle \rangle$

$\langle \langle k \rangle \rangle$

$[ \mid \langle \langle m \rangle \rangle \mid ](v) = \{ \langle \langle n \rangle \rangle \in \langle \langle L(m) \rangle \rangle \mid m = n \}$

$[ \mid \langle \langle k \rangle \rangle \mid ](v) = \{ \langle \langle m \rangle \rangle \in \langle \langle L(m) \rangle \rangle \mid m \in [ \mid k \mid ] \}$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

- Where are the monads?
- Where are the algebras?
- Where is the distributive law?
- How does this relate to Universal Algebra?
- How does this relate to Curry-Howard?

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Organizing principles

A domain (of discourse) is given by

- A monad,  $T$ , for a term language (the witnesses of propositions)
- A structural equivalence,  $\equiv$ , on  $T$ , such that  $(T, \equiv)$  is an algebra for  $T$
- A rewrite rule,  $\rightarrow$ , called the transition or reduction relation on  $T$  extending smoothly to  $T/\equiv$ .

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Organizing principles

## Examples:

- A monoid is a domain in which the reduction relation is empty.
- The lambda calculus is a domain in which the structural equivalence is alpha-equivalence and the reduction relation is beta-reduction.
- The  $\pi$ -calculus is a domain in which the structural equivalence is normally called structural equivalence! and the rewrite rule is comm + the smooth extension to the algebra.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Organizing principles

In the physical sciences domains have historically been represented via some flavor of linear algebra. Note what this does to the packaging of dynamics.

- Vector spaces just sit there. They enjoy no dynamics. You need transformations, such as linear transformations, to get them to dance.
- Domains repackage dynamics with the algebraic structure. This does something crucial to the notion of morphism at the categorical level.



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Organizing principles

The structure of vector spaces and linear transforms fits well with the classical development of category theory

The relationship of domains to category theory is a work in progress

- In historical terms, it's brand new
- In categorical terms, we have to understand what is to be preserved when we preserve dynamics -- without this understanding there is no good definition of morphism



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Organizing principles

Domains enjoy something that is a profound gap in classical models of dynamics: an effective notion of equivalence of dynamical systems

Bisimulation is perhaps one of the most important ideas computer science will bring to the physical sciences



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Calculations

Redo the monoid example, but this time instead of Set as the collection in which to accumulate witnesses, let's pick a different collection: sets of sequences of witnesses

If  $X$  is a set,  $Q[X]$  is the domain of all sets of sequences of elements of  $X$ , i.e.

$$Q[X] = \text{Pow}[X^*]$$

Hint: does anyone know what logic this corresponds to?

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Calculations

Extra credit: how do we add probability into this setting?

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Calculations

- How do we compile structural queries into native storage query language?
- How do we effect a query plan?
- How does this relate to symbolic execution and behavior in time?

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Motivation

(Behavioral and structural queries in physical systems) Find all the pathways,  $p$ , in  $\langle \text{SignallingPathwayDB} \rangle$  such that when  $\langle \text{Reagent} \rangle$  is added in concentration  $K_p$  eventually reaches states in which we see these  $\langle \text{Reagent1} \rangle, \dots, \langle \text{ReagentN} \rangle$  in concentrations  $K_1, \dots, K_N$ .

Priami-Aviv encoding of cell signaling pathways identifies

- suitable variant of  $\pi$ -calculus as the term language
- suitable variant of Caires spatial-behavioral logic as the query language (algorithm agrees)

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Motivation

(Behavioral and structural queries in logical systems) Find all the implementations of message queue in <http://svn.myrepo.org> such that the queue is FIFO and use the Xerces XML parser for XML-encoded messages

Berger-Honda-Yoshida encoding of core Java identifies

- suitable variant of lambda calculus as term language
- suitable variant of rely-guarantee logic as query language (algorithm agrees)

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Motivation

(Structural queries in physical systems) Find all the knots in <KnotDB> that have alternations of <Trefoil> and any 7-crossing knot as subknots

Meredith-Snyder encoding of knots as process identifies

- suitable variant of  $\pi$ -calculus as term language
- suitable variant of Caires' spatial-behavioral logic as query language (algorithm agrees)

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Motivation

(Geometric queries in physical systems) Find all the conduit wires running from the nose of the aircraft to the midsection that are likely to undergo a twist between source and destination

Hestenes, et al identify

- Clifford algebra as term language
- Use the algorithm to identify the logic

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Where Biosimilarity is in the implementation of this idea

- SpecialK/KVDB - an open source, distributed key-value database where the keys are prolog terms
- Stratolytics - an analytics as a service platform
- **Quality of Identity** - an identity management platform

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

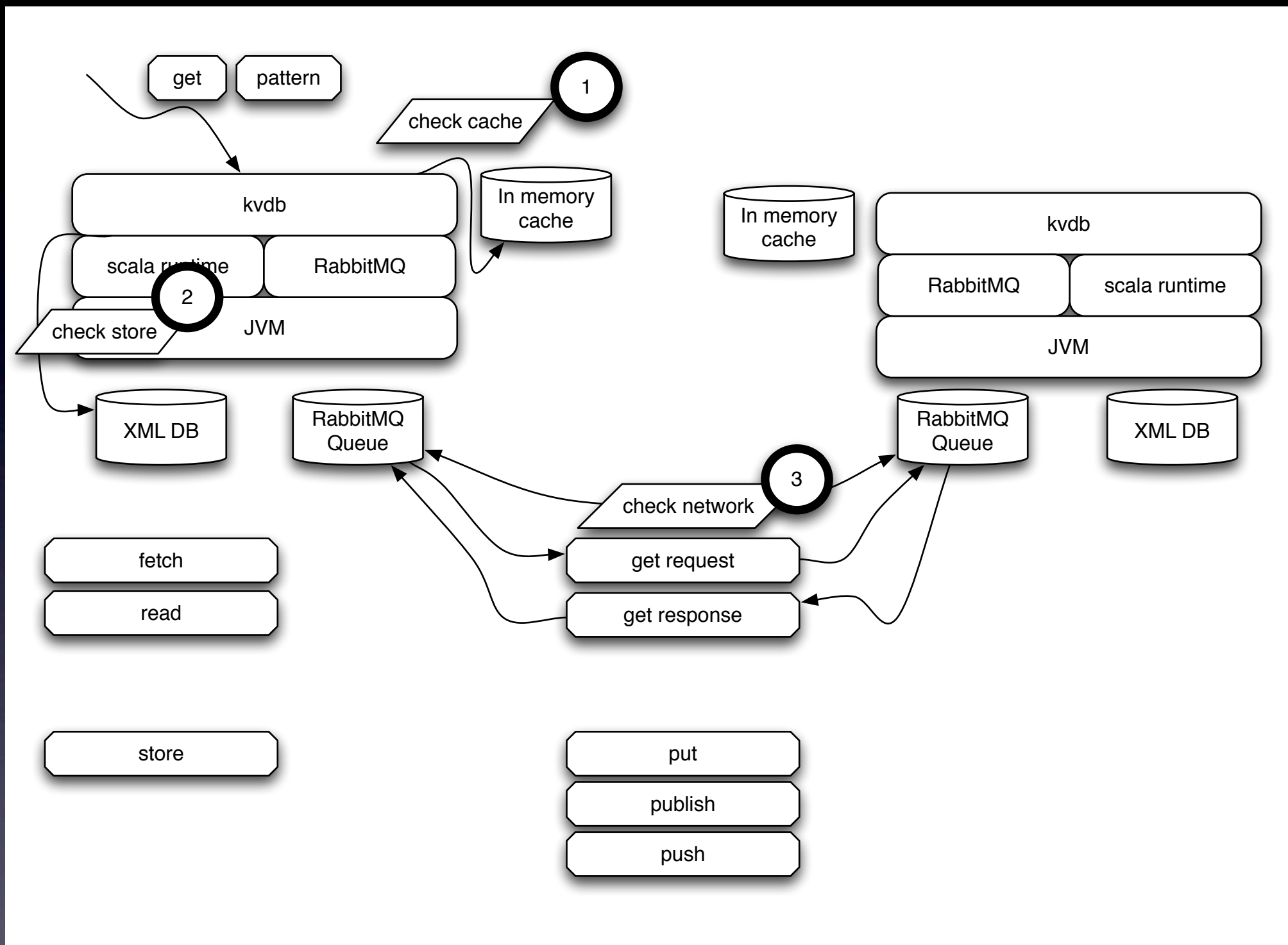


# Where Biosimilarity is in the implementation of this idea

```
for( e <- kvdb.get( cnxn )( lbl ) { handle( e ) }
```

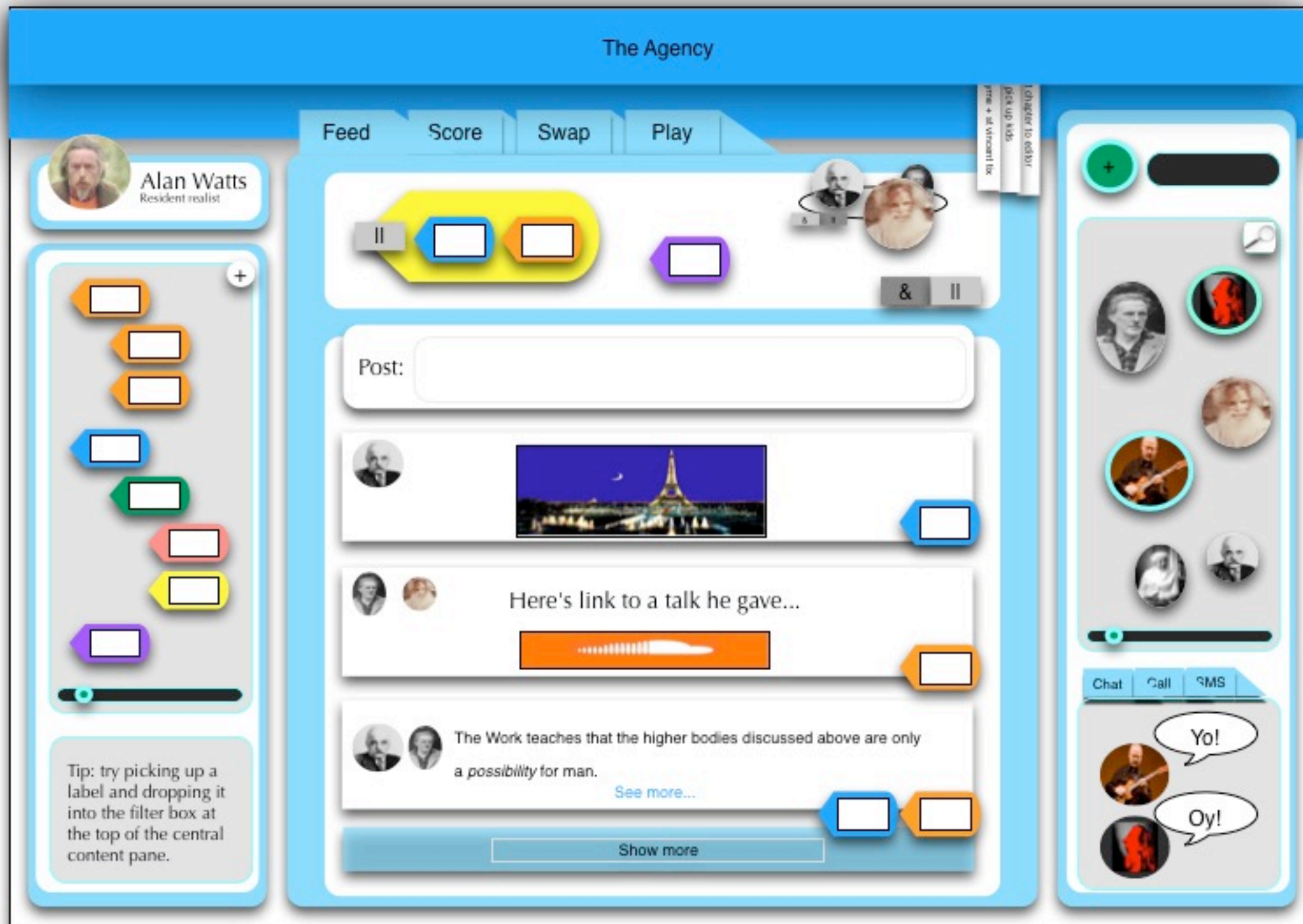
Analysts wire models directly to web-fed, label-filtered, cnxn-vetted sources to produce graphs, charts and decision support analysis.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

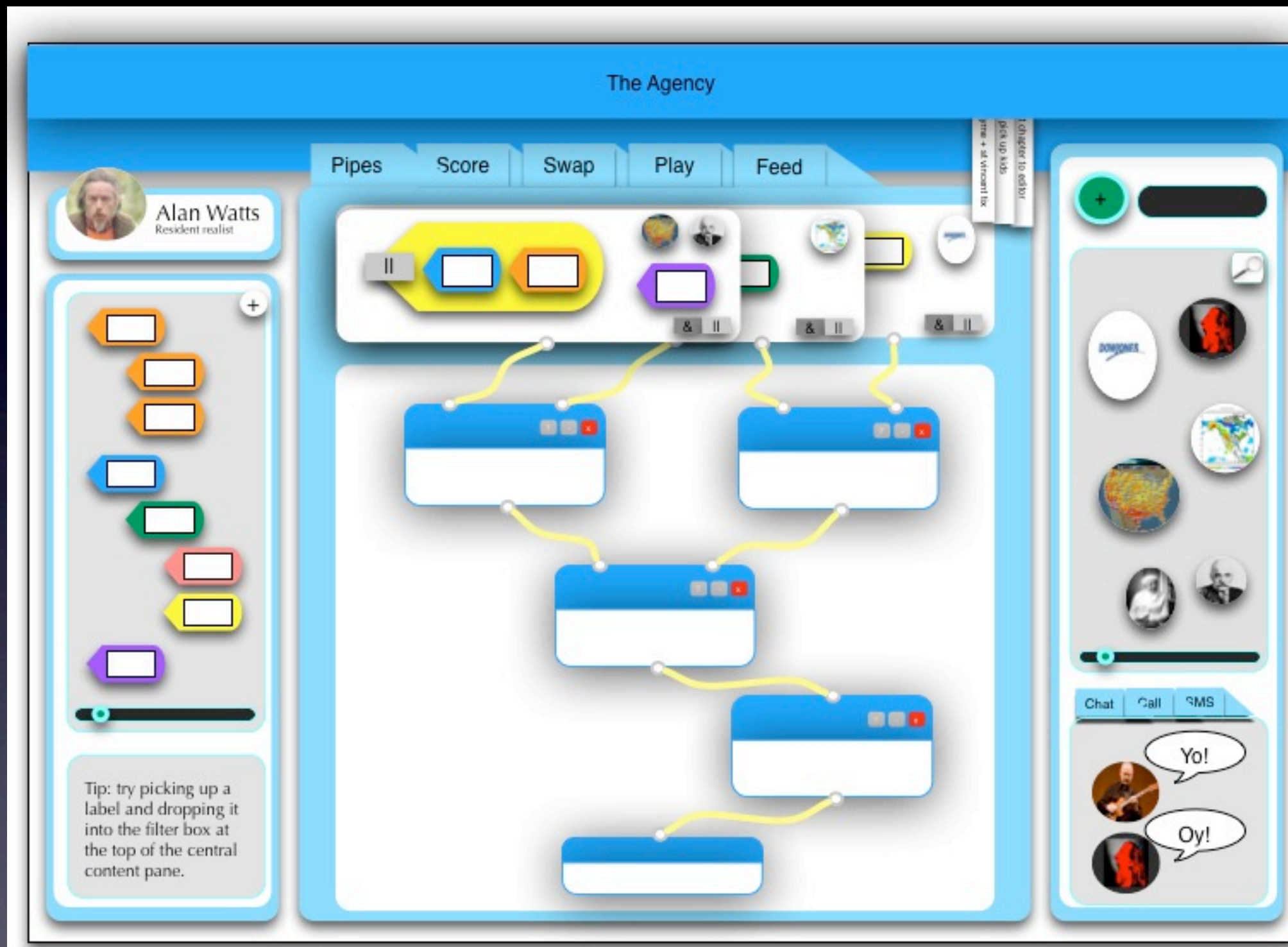


Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC





Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Conclusions

- We have described an algorithm for generating a domain specific query engine
- We have outlined motivations for addressing physical (signaling), logical and geometric systems in this way
- We have outlined how such a system fits into a distributed cloud-based analytics-as-a-service offering supporting appropriate individual and organization privacy

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Questions?

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Backup

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

Defn: A category,  $C$ , is given in terms of the following pieces of data

- $|C|$  the collection of objects (aka types) of  $C$
- $\|C\|$  the collection of morphisms (aka programs) of  $C$
- $\text{src} : \|C\| \rightarrow |C|, \text{trgt} : \|C\| \rightarrow |C|, \text{id} : |C| \rightarrow \|C\|$
- $\_ \bullet \_ : \|C\| \times \|C\| \rightarrow \|C\| + 1$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

satisfying:

- $m1 \bullet m2$  is defined when  $\text{src}(m2) = \text{trgt}(m1)$
- $m \bullet \text{id}(\text{trgt}(m)) = m = \text{id}(\text{src}(m)) \bullet m$
- $m1 \bullet (m2 \bullet m3) = (m1 \bullet m2) \bullet m3$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

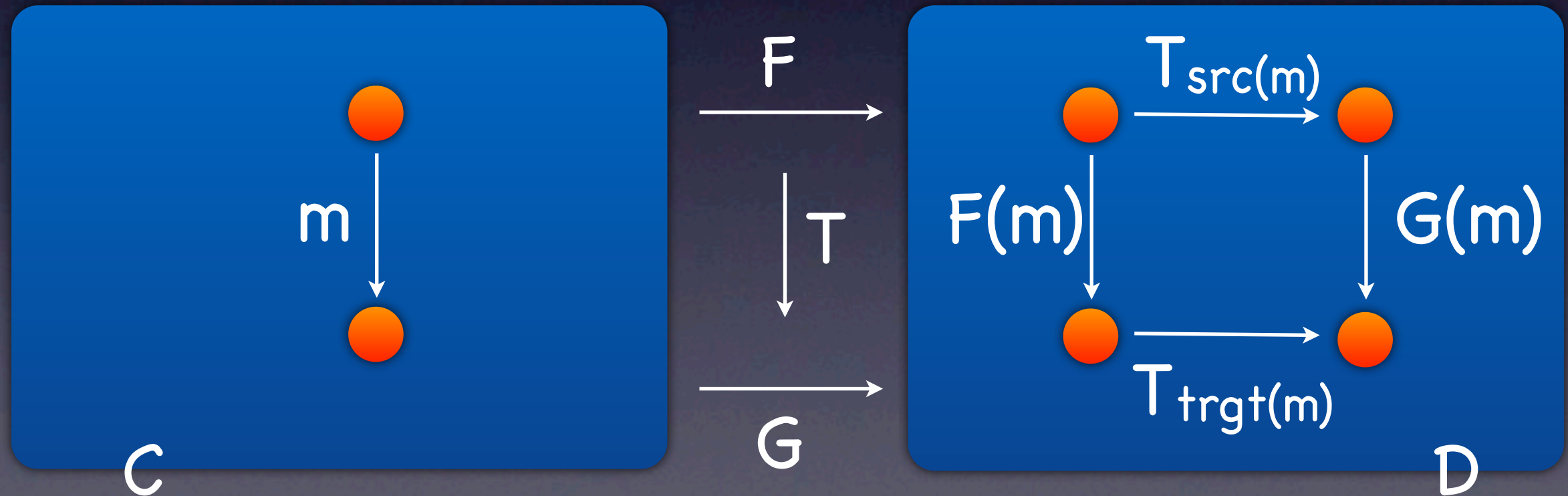
Defn: A functor,  $F : \mathcal{C} \rightarrow \mathcal{D}$ , is a morphism in the category of categories, i.e.  $F$  decomposes into

- $|F| : |\mathcal{C}| \rightarrow |\mathcal{D}|$  the map on collection of objects of  $\mathcal{C}$
  - $\|F\| : \|\mathcal{C}\| \rightarrow \|\mathcal{D}\|$  the map on the collection of morphisms of  $\mathcal{C}$
- preserving src, trgt, id and •
- $F(\text{src}(m)) = \text{src}(F(m))$ ,  $F(\text{trgt}(m)) = \text{trgt}(F(m))$
  - $F(\text{id}(\text{src}(m))) = \text{id}(F(\text{src}(m)))$
  - $F(m_1 \bullet m_2) = F(m_1) \bullet F(m_2)$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

Defn: A natural transformation  $T : F \rightarrow G$ , is a morphism between functors,  $F, G : \mathcal{C} \rightarrow \mathcal{D}$ , i.e. it decomposes into a  $\mathcal{C}$ -indexed family of morphisms,  $T_x$ , in  $\mathcal{D}$  taking the image of  $F$  to the image of  $G$



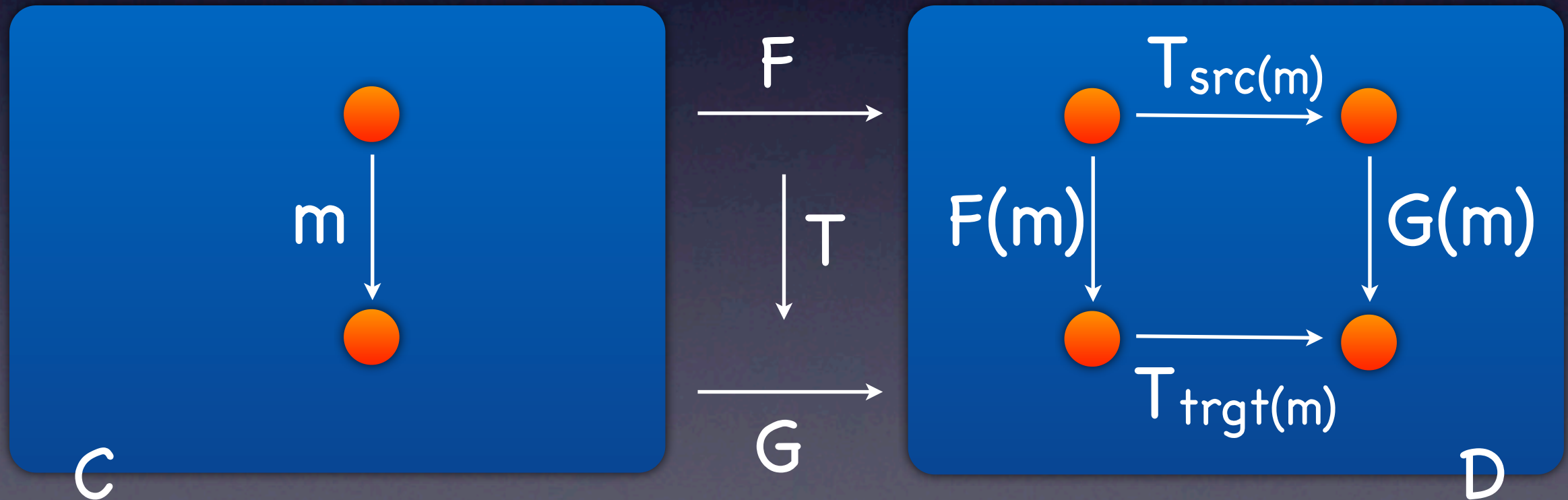
Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

satisfying:

$$F(m) \bullet T_{\text{trgt}}(m) = T_{\text{src}}(m) \bullet G(m)$$



Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

Defn: A monad,  $T$ , is a monoid object in the category of categories, i.e.  $T$  decomposes into

- $T : C \rightarrow C$  an endo-functor on  $C$
- $\text{unit}(T) : \text{Id} \rightarrow T$  a natural transformation from the identity functor,  $\text{Id}$ , to  $T$
- $\text{mult}(T) : T^2 \rightarrow T$  a natural transformation

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Definitions & examples

such that  $\text{unit}(T)$  and  $\text{mult}(T)$  interact coherently, i.e.

- $T \text{ mult} \bullet \text{mult} = \text{mult} \bullet \text{mult } T$
- $T \text{ unit} \bullet \text{mult} = \text{unit } T \bullet \text{mult}$

where

- $T \text{ mult}$  is the evident natural transform from  $T^3$  to  $T^2$
- $\text{mult } T$  is the evident natural transform from  $T^3$  to  $T^2$
- $T \text{ unit}$  is the evident natural transform from  $T^2$  to  $T$
- $\text{unit } T$  is the evident natural transform from  $T^2$  to  $T$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

Defn: An algebra for a monad,  $T$ , is given by a pair  $(x, h)$  where

- $x \in |\text{src}(T)|$
- $\text{src}(h) = T(x)$ ,  $\text{trgt}(h) = x$
- $\text{unit}_x \cdot h = \text{id}_x$
- $\text{mult}_x \cdot h = T(h) \cdot h$

Intuition : the pair  $(x, h)$  gives the model of the algebra. In the case of pure syntax -- meaning those  $x$  in  $|\text{src}(T)|$  built from term structure along -- this is tantamount to giving the structural equivalence.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Definitions & examples

Defn: A distributive law,  $l : ST \rightarrow TS$ , is a morphism between the composition of two monads,  $S, T$  with the units, mults and  $l$  all interacting coherently

- $T \text{ unit}(S) \bullet l = \text{unit}(S) T, S \text{ unit}(T) \bullet l = \text{unit}(T) S$
- $S l \bullet l S \bullet T \text{ mult}(S) = \text{mult}(S) T \bullet l$
- $l T \bullet T l \bullet \text{mult}(T) S = S \text{ mult}(T) \bullet l$

Key point: The composition of monads is not necessarily a monad. With a distributive law,  $TS$  canonically acquires the structure of a monad.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC



# Calculations

Doberkat shows how the algebras of the Giry monad on the category of Polish spaces (separable metric spaces with a complete metric) correspond to convex partitions of all probability measures.

What is not clear is how useful this result is in extending discrete computations to probabilistic ones.

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC

# Calculations

Our framework refines these questions:

- Structural equivalence can be probabilistic
  - We recognize the shape of this protein with confidence,  $x$
- Transition rules can be probabilistic
  - These proteins are likely to undergo phosphorylation with probability,  $x$

Lucius Gregory Meredith • Managing Partner • Biosimilarity, LLC