Operational indicator frameworks

An operational indicator framework is just a list. % Can we 'enrich' these lists?

- net income
- sales
- hack attempts

- approval rating
- taxes avoided
- retweets

- reliability
- trustworthiness
- system downtime

% often: time-series,
regularly updated / added-to,
something we do low-level statistical
analysis to, e.g. correlation

They are used for:

1. optimization of processes
2. strategic decision-making
3. public (or internal) communication
4. ontology / simplification

Problem: ~~the~~ how do we take data from one indicator framework and relate it to data from another framework?

~~i.e. data integration~~:

Eg. how do I illustrate the benefits / extended impact of my project to another aspect of my organization / domain?

Abstract indicator frameworks

Defn. a $\boxed{\text{"formula"}}$ from a $\boxed{\text{causal model}}$ to a $\boxed{\text{data model}}$ (measurement / observation).

The philosophy: toward a type system for <u>measuring</u> and <u>modeling</u> complex systems

math artifact, e.g.
- (relational) table of data
- list of random variables
- ☆ Hilbert space of random var.
- observation tables of a regular language

math artifact, e.g.
- ☆ graphs / DAGs
- Bayesian networks
- dynamical systems
- finite automata

To start: fix one of each!

Ex. bus data from Nashville, TN

	actual travel time AT	scheduled travel time ST	jam factor JF	traffic speed TS
AT	1	<u>.81</u>	<u>.15</u>?	~~-.04~~
ST		1	~~.05~~	~~-.02~~
JF			1	<u>-.48</u>
TS				1

<u>Correlation</u>

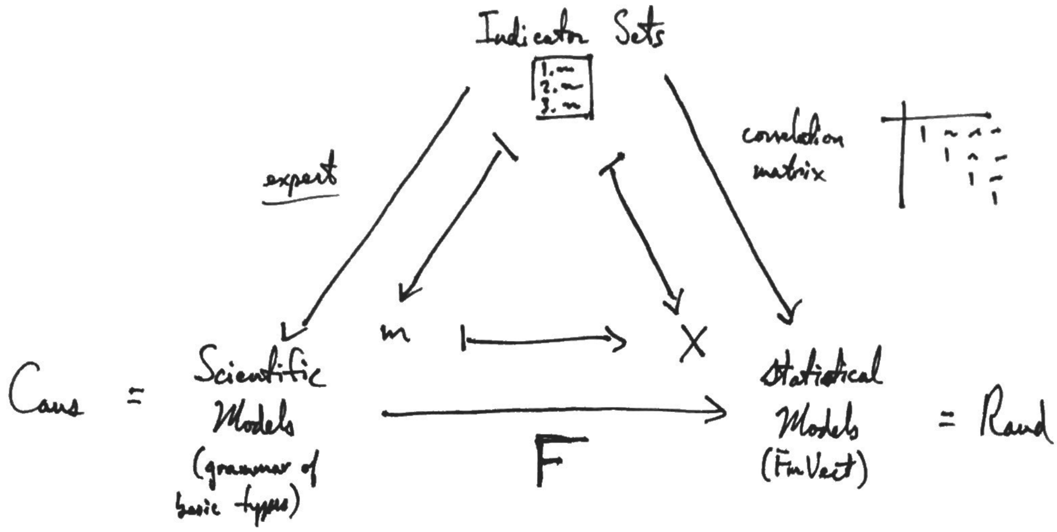
Why .15? ~~Bad~~ First step toward analysis: build a (basic, maybe <u>false</u>) model!

$$\begin{tikzcd} & & AT & \\ ST & TS \arrow[ur] & & JF \arrow[ul] \arrow[ll] \end{tikzcd}$$

<u>Causation</u>

<u>Dumb question</u>: how do you "wrap" the two ~~pieces of~~ / perspectives on the system $\{AT, ST, JF, TS\}$ into one, <u>consistent</u> data structure?

How to construct an abstract indicator framework



Step 1. Take as input a list of indicators, a correlation matrix on these indicators, and an "expert".

notation : indicator $i \in I$

C_I is a correlation matrix on I

$X = X_I$ is the f.d. Hilbert space w/ fixed basis $B_X = \{b_1, \dots, b_d\}$ s.t. $\langle b_i, b_j \rangle = C_I(i,j)$

Step 2. Define a causal diagram over I, e.g.

$$\begin{tikzcd} & & AT & \\ ST & TS \arrow[ur] & & JF \arrow[ll] \arrow[ul] \end{tikzcd}$$

How to construct... continued

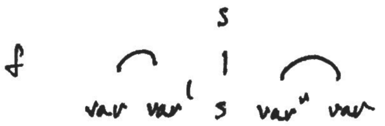
Step 3. Divide your model into "grammatical" types.

$ST, TS, AT, JF \in var$

$\longrightarrow \in var' \; s \; var''$ ← technically, this is a formal multiplication!

what's going on here? Basically, we are encoding "$\longrightarrow$" as a relation that requires a var on the left and a var on the right. "s" is a causal fact.
For now, we will just take a causal diagram as just a disjoint set of causal facts.

Step 4. Define a <u>type reduction</u> $f : var \; var' \; s \; var'' \; var \longrightarrow s$
or, diagrammatically,



Step 5. Define a map F from types (in Cause) to vector spaces (in Rand) which respect the multiplication.

$$var \xrightarrow{\sigma} X$$
$$ST \longmapsto b_1$$
$$AT \longmapsto b_2$$
$$\vdots$$
$$var' \; var \; var'' \xrightarrow{\sigma} X \otimes S \otimes X$$
$$\text{"}\longrightarrow\text{"} \longmapsto \Psi = \sum_{i,j,k} c_{ijk} (x_i \otimes s_j \otimes x_k)$$

note the "σ" here, which needs to be defined under F.

$\boxed{\text{Done!}}$

Using an abstract indicator framework

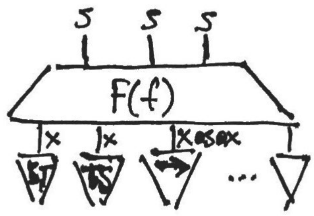
Take a causal diagram and plug it into F.
First, we get a vector

$$\sigma(ST) \otimes \sigma(TS) \otimes \sigma(\longrightarrow) \otimes \sigma(AT) \otimes \cdots$$

$$\|$$

$$b_1 \otimes b_2 \otimes \Psi \otimes b_3 \otimes \cdots$$

then we can apply our reduction $F(f)$ to it



in order to obtain a vector $s_1 \otimes s_2 \otimes s_3 \in S \otimes S \otimes S$ that represents the "meaning" for validity of the model in X,

An analogy : nat'l language

$$\begin{tikzcd}
 & \text{Language (sentences)} \arrow[dl] \arrow[dr, "\text{neural network / word2vec}"] & \\
\text{Pregroup Grammar} \arrow[rr, "F"'] & & \text{Statistical Model}
\end{tikzcd}$$

Idea: vector spaces can be used to assign meanings to words in a language;
"the meaning of a word is its context of nearby words". Key thing: its automatic.

Idea: pregroups can be used to assign grammatical structure to sentences

n : noun $\qquad$ s : sentence $\qquad$ j : verb

$$\begin{array}{ccc} \text{John} & \text{likes} & \text{Mary} \\ n & n^r s n^l & n \end{array} \to 1 s n^l n \to 1 s 1 \to s \quad \checkmark$$

$$\downarrow F$$

$$N \otimes N \otimes S \otimes N \otimes N \ni \overrightarrow{\text{John}} \otimes \overrightarrow{\text{likes}} \otimes \overrightarrow{\text{Mary}}$$

$$\downarrow F(f) \qquad\qquad\qquad \downarrow F(f)$$

$$S$$

$$F(f)\left(\overrightarrow{\text{John}} \otimes \overrightarrow{\text{likes}} \otimes \overrightarrow{\text{Mary}}\right) =$$

$$\sum_{ij} \langle \overrightarrow{\text{John}}, m_i \rangle \overrightarrow{\text{likes}}_{ij} \langle f_j, \overrightarrow{\text{Mary}} \rangle =$$

$$\sum_{ij} \delta_{\text{John}, m_i} \overrightarrow{\text{likes}}_{ij} \delta_{f_j, \text{Mary}} =$$

$$\overrightarrow{\text{likes}}_{\text{John}, \text{Mary}}$$