A Simple Model For Cognitive Semantics

On measurement and observation

Mark Burgess

August 19, 2016

Abstract

Try to sketch out the how-to justification for making a knowledge graph. Try not to give all the details...:-P

1 Introduction

This note summarizes the practical aspects of [1–3]. The goal is to come up with some simple rules for building a knowledge graph that can be used for reasoning. Knowledge is defined as an on-going relationship between stored concepts, introspection, and experiential learning.

What do we mean by *cognition*, as opposed to general observation and measurement? Cognition is about learning 'self' from experience. Measurement is about eliminating 'self' from experience. The semantics of cognition are very different from the semantics of experimental observation in science. Observer subjectivity is a key aspect that cannot be disregarded in measurement, but this aspect is much stronger in cognition[1]. Cognition is much more expensive than 'offline' experimentation. An observer has to run a tight recurrent sampling loop to collect and process data in realtime (the brain uses ten to twenty percent of all the body's energy. Fast monitoring on a computer is highly CPU intensive.). Nyquist's law determines the accuracy of a sampling rate.

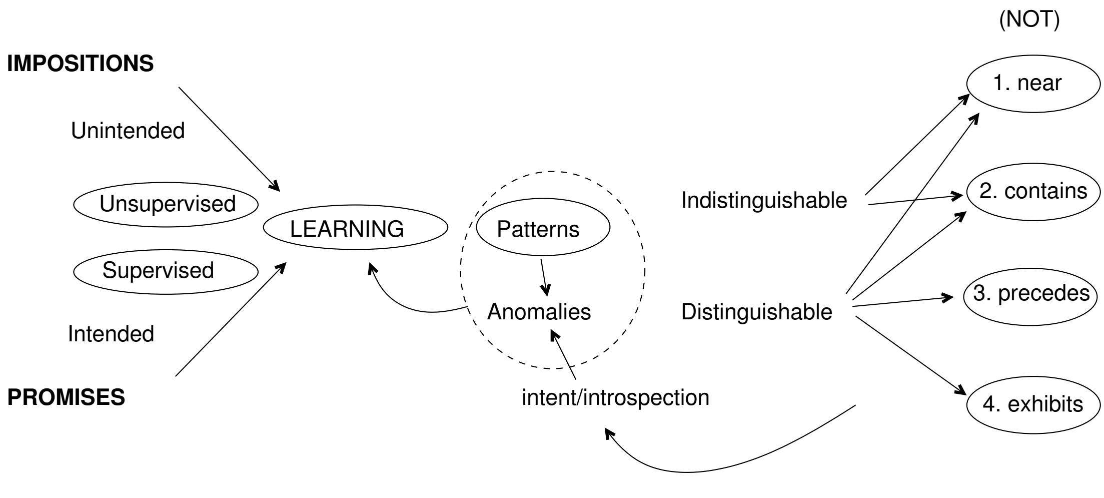

Figure 1: A concept prism from left to right for association types.

Cognition leads to associations through a process described in figure 1. This is a 'prism' which separates data into a spectrum of associative types, which represent different spacetime characteristics, called the irreducible types.

[1] A lot of attention has been given over to the algebraic aspects of measurement in physics, e.g. quantum mechanics, but surprisingly little attention has been given to the semantics of data, and cognitive perception, except in the case of experimental error [4–6]. The extent to which measurement disturbs the system during the act of measurement affects what can be promised about a measurement. Measurables may have the property of 'compatibility' [7], meaning that measurement of one does not influence the measurement of the other.

1.1 Realtime cognition versus batch experimentation

The semantics of data are different if they are arriving in real time, as a continuously updated stream of impressions, versus in a protected experimental trial:

- A continuous stream of data are treated as a sequence, in which the order of events may be important. The data form a timeseries. Statistically, this is aligned with a Bayesian 'learning' interptretation.
- Data that are delimited by a start and and end, in a batch under constant conditions are treated as a single concurrent frame, under constant conditions, in which the order of events is unimportant. The data form an array [3]. Data types may or may not be distinguishable in the array of data. Statistically, this represents a frequentist batch interpretation.

1.2 Concepts and associations

The end result of a cognitive process is to tokenize complex data inputs into tokens (concepts), which summarize the data in an *invariant representation* immune to variations, and to link these together by association. In promise theory language, concept tokens are agents (nodes in a graph), and associations (graph edges) are promised by these agents. There is a ladder from sensing to conceptualization, which proceeds by aggregation (over time and space). This involves work, and thus it takes increasing time to process complex associative concepts. Concepts

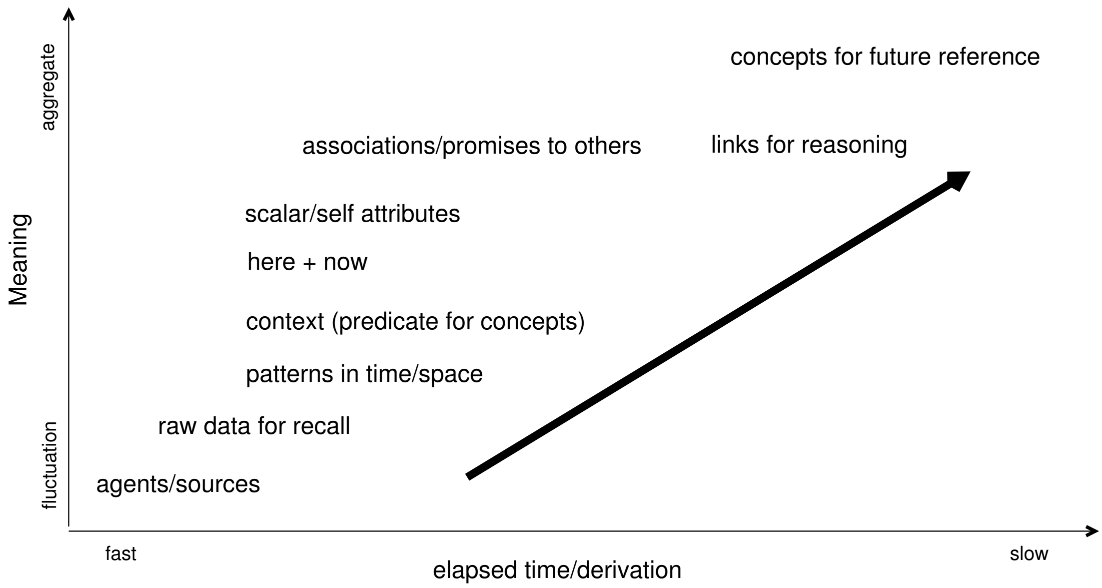

Figure 2: Climbing the knowledge ladder takes longer and longer to advance towards sophisticated concepts, from fast sensing to pondering concepts, because of iteration and aggregation (learning).

are clusters of agents. The basis of interpretational semantics lies in the following observations:

Law 1 (Coincidence and concurrence) *That which occurs together (concurrently and coincidentally, i.e. locally) is associated. Measurements at different locations and at different times are not necessarily causally related, but may be correlated.*

Law 2 (Averages stabilize semantics by decoupling from time/space) *Data sampled and combined from different sources (locations) should be treated as ensemble averages, in a single concurrent experiment. The meaning of time and location are lost in these averages, and semantics are stabilized by this decoupling.*

Law 3 (Scaled agency) *Semantics (interpretations) arise from the association of observed forms/patterns, in space and time, across multiple scales.*

Law 4 (Constant semantics) *Measurements are comparable if they have the same semantics. In cognition, observations are not comparable in an experimental sense.*

The last two are circularly self consistent.

2 Quantitative assessements to semantics

Quantitative descriptions of phenomena are possible when there is a measuring apparatus with a calibrated numerical scale. The question we always ask is: what are the agents (concepts) and what promises do they make (how are they associated and related)?

2.1 Instrumentation and its distribution

The first law says, that which occurs together (concurrently and coincidentally, i.e. locally) is related. Measurements at different locations and at different times are not necessarily causally related, only correlated.

Law 5 (Calibration) *Two measuring instruments cannot have the same calibrated scale, unless they cooperate and calibrate mutually; so, measurements can only promise the same assessment semantics if the data are collected by the same agent/apparatus at the same time.*

2.2 Concurrence, conincidence, and simultaneity of observations

The maxim for real-time cognition is: that which occurs together is related. For example, the sections of a story

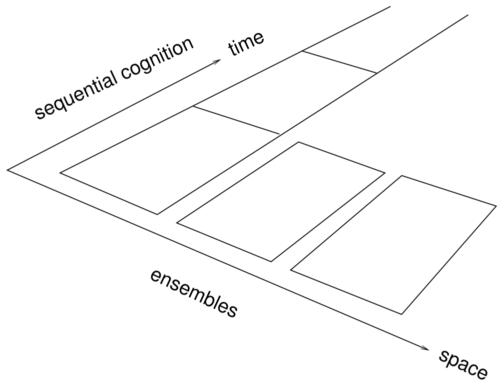

Figure 3: Concurrence (in the same time grain), and coincidence (same spatial grain), defines the boundaries for 'cognition'. We coarse grain space and time into regions of partially indistinguishable equivalence, like different experimental 'trials'. In cognition, every new timelike frame is a new experiment, and we must use learning to even out experimental inequivalence.

book have a linear cognitive structure, in which each section or chapter could be grouped as a single semantic time frame. A lexicon or dictionary has no narrative structure, so it acts as an ensemble of parallel inputs.

Rule 1 (Causation rule) *If order is clear, ordered associations are coded as precedence:*

$$X \text{ results in } Y \leftrightarrow X \text{ precedes } Y \tag{1}$$

Else derivative results are expression not causation:

$$X \text{ is documented in } Y \leftrightarrow X \text{ is expressed by } Y \tag{2}$$

Each measurement is an isolated (local) event in spacetime, in which it is assumed that the result is observed at the moment of data capture. The time for the sensor to detect is often assumed negligeable. In remote observations, the time for data to arrive may be significant, and unpredictable. Thus all quantitative measures are subject to uncertainties in space and time.

2.3 Characteristics as concepts

Numbers can be thought of both as data and as concepts. As concepts, we have typical values, well-known values, habitual values, pretty values, good and bad (emotionally charged) values that remind us of other experiences, etc.

Semantics emerge from assessments of *meaning*, not from values themselves. Meaning may emerge from data over a long time, and the results can be applied back as a filter on the interpretation of new data. Meaning is associated with:

- Frequency or familarity of observation.
- Known (intentional) relationships between the measures.
- Inferred relationships, e.g. by correlation or pattern.
- Short and long term correlations.
- Temporal and spatial patterns.

Let's assume we can measure some quantity q (e.g. CPU, memory, traffic density, etc). For each measure there are quantitative representations of measurements over time and space.

MEASURE	BASIC INTERPRETATION
q	Raw measured value 'here and now' at the sampling location ('data')
$q(t)$	Characteristic pattern over times longer than sampling $T_{\text{pattern}} \gg T_{\text{sample}}$
$E(q)$	Characteristic average value, where possible.
$\sigma(q)$	Standard deviation: characteristic average value of deviation from mean.
dq	Rising or falling over a sampling interval
$\frac{dq}{dt}$	Rate of change over times longer than a sampling interval
dt	Last observed time ($dt \geq T_{\text{sample}}$)

Some comments about data collection:

- A finite sample of data always has a mean $E(q)$ and standard deviation $\sigma(q)$, even if the distribution is not Gaussian.

  An expectation value, even of an isolated sample, is often useful as a characteristic scale for sanity checking of data collection, and it illustrates how practical measurement may be different from theoretical statistical behaviour (values may not be expected to converge on a limiting distribution at all). The higher moments of a distribution may likewise be useful guidance during data collection, in informing the semantics of data interpretation.

- The reasons for adopting statistical treatment are *theoretical*. We must have a model for expected variation, such as in the classical theory of experimental error [5, 6].

- Once a value is propagated from its original location, its meaning is immediately distorted by spacetime effects, including environmental noise and distortion.

2.4 Multivariate quantitative constraints

Constraints have implicit semantics, and should be documented for the purpose of reasoning (understanding). A measurable outcome may depend on a number of individual variables. It may or may not be possible to measure these independently.

- Variables may be correlated, or functionally dependent.
- Sometimes constraints affect the behaviours of quantities: e.g.

$$\sum_i q_i \leq Q_{\text{max}} \tag{3}$$

$$\text{Memory} + \text{swap} = \text{constant}. \tag{4}$$

  This turns into a graphical relationship:

$$(q_i, \text{depends on}, Q_{\text{max}}) \tag{5}$$

$$(memory, \text{depends on}, constant) \tag{6}$$

$$(swap, dependson, constant) \tag{7}$$

  i.e. graphically,

$$q_i \rightarrow Q_{\text{max}} \tag{8}$$

$$M \rightarrow C \leftarrow S. \tag{9}$$

- Quantities that are collective, and can borrow from a reservoir, like equilibria. The observed properties may only make sense as spacetime averages (like smoothing).

2.5 Pattern recognition

PROCESS	AGENT	SEMANTICS	REPRESENTATION	derived graph
Intent/promise	internal	purpose	compressed/tokenized	associate with intent
Actual	external	experience/sensory	(data extensive)	data character
Learning	recurrent	anomaly compare	data convergence to a pattern	data normality
Reasoning	introspection	stories	token paths	chained relationship
Linguistic	interchange	any	scaled tokens	expert input

3 Qualitative assessments (naming things)

Qualitative assessments are made by the 'self' observer, based on what is promised or hinted by other external agents. Assessment can be based on quantitative inputs (see figure 4), or they may be formed mainly from internal expectations and concepts, e.g. 'by looking like something else'.

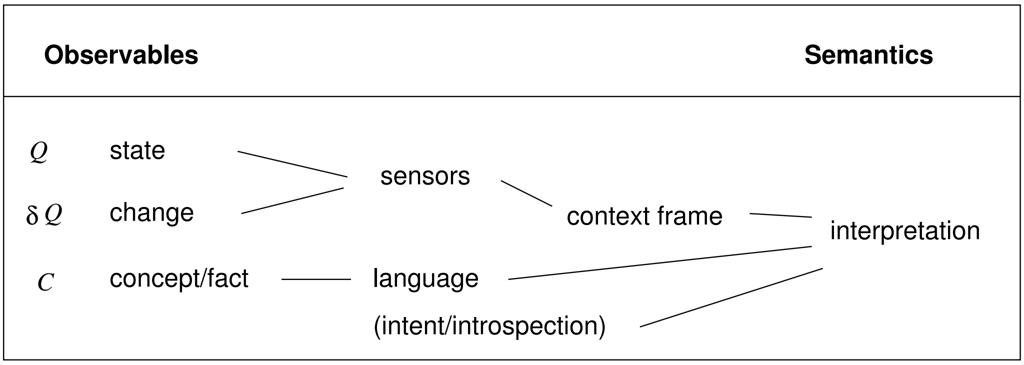

Figure 4: Cognition is traced from input to concept formation. Some qualitative assessments derive from the quantitative dynamical properties of variables (behaviour in space and time).

The meaning of data is sometimes decided by ad hoc context and sometimes by prior intent. Large amounts of bulk data get compressed into a small number of tokens or concepts, which signify an interpretation. This may be all we retain of a sensory experience. It is also possible to 'change one's mind' and reinterpret data after they have been measured, especially if we keep raw data in bulk form.

Qualitative assessments need not have a numerical basis, even though we can turn them into ad hoc quasi-numerical scales. Qualitative characteristics are essentially names:

- Qualitative characteristics: good, bad, nice, ugly, true, false, easy, difficult, hard, soft, etc.
- Proper names, e.g. Mark, car, door.

Names promises that signal the identity of an observed phenomenon, and may be associated implicitly with its behaviour (e.g. door handle). Qualitative assessments require either that concepts pre-exist, or are learned as recognizable pattern over time. They are semantic characterizations or interpretations of purpose/intent. The names represent their intended semantics, as tokens, and these may be taken as promises by the observer (with associated level of trust).

Rule 2 (Tokenized qualities are added intentionally) *Known qualitative concepts may be probed for, and added into a knowledge representation by intentional insertion.*

For example, CFEngine probes for 'classes' which are semantic tokens, describing the attributes and state of a host.

3.1 Concepts at many scales

Concepts are nodes/agents, or clusters of nodes (supernodes or superagents, in promise language), in a knowledge graph (see figure 7). Individual nodes must have names. Arbitary clusters of nodes may be given names (see figure 5), and represent new concepts by recombination.

Rule 3 (Linking similarity) *These processes suggest association types:*

- **Concurrency** *(simultaneous - link all tokenized inputs to a scope/namespace for each concurrent frame.*
- **Coincidence** *(same time and place) - concurrency links locally. Clusters may also be linked by aggregation of similar semantics from multiple sources.*
- **Common intent** *(aligned) - if we know intended participants in an outcome link these as members and as causal dependencies.*
- **Sequences** *(hangs together/makes 'sense') - these may be remembered as data, or as stories linked through ordered sequences of concepts that map to the data.*
- **Patterns recognized** *or expected (inference)*

The point of representing certain data as concept nodes is that this allows us to reason about them as any other concept.

3.2 Associations and their aliases

The basis for association has its origins in spacetime (being close in space and time, being in a bounded area, etc). We then give meaning to these different spacetime coincidences by naming associations according to abstracted concepts. Concepts and associations are nodes and edges of a graph:

$$C_1 \xrightarrow{\mathcal{A}} C_2 \tag{10}$$

$\mathcal{A}$ IRREDUCIBLE TYPE	SPACETIME ORIGIN	SEMANTIC JUSTIFICATION
contains	Membership	Common origin, destination, purpose
follows/caused by	Direction	Dependency
exhibits	Distinction/combination	Distinction, capability
near	Proximity	Concurrent/coincident activation, similarity

The qualitative descriptions are given by the four irreducible associations documented in table 1, and their negatives.

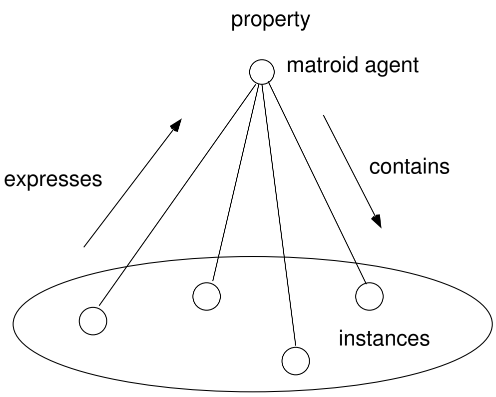

Figure 5: The semantics of representing aggregation. If we group aggregated properties by a linking them centrally to a singular basis agent [1, 3], then attribute expression is effectively the inverse of set membership, and the basis agent alone represents the collective ensemble. Ad hoc clustersd not have this property.

FORWARD	RECIPROCAL	SPACETIME STRUCTURE
contains	is a part of / occupies	
surrounds	inside	AGGREGATE / MEMBERSHIP
generalizes	is an aspect of / exemplifies	"contains"

FORWARD	RECIPROCAL	SPACETIME STRUCTURE
depends on	enables	
is caused by	causes	GRADIENT/DIRECTION
follows	precedes	"follows"

FORWARD	RECIPROCAL	SPACETIME STRUCTURE
has name or value	is the value of property	
characterizes	is a property of	DISTINGUISHABILITY
represents/expresses	is represented/expressed by	"expresses"

FORWARD	RECIPROCAL	SPACETIME STRUCTURE
is close to	is close to	
is equivalent to	is equivalent to	
is connected to	is connected to	PROXIMITY
is adjacent to	is adjacent to	"near"
is correlated with	is correlated with	

Table 1: Examples of the four irreducible association types, characterized by their spacetime origins.

3.3 'Emotional' intepretations of state (context)

Emotional weight plays a significant role in cognition for setting context and interpretation priorities. Emotions are very coarse aggregate 'concepts' stimulated by sensory/introspective inputs.

Distress is an easy attribute to measure in a finite system. When a threshold of behaviour is reached where the system is unable to keep its promises, because resources are limited, then it may be attributed a level of distress. Thrashing to empty a queue is an example of this.

States like good, bad, danger, happy, sad, etc are components of what we would think of as emotional states. These give clear meanings to other measurements: how we respond to a context (state) and the associations we make during good and bad times affect how we recall concepts later. If we feel strongly about something (good or bad) this translates into an importance rank of an association. We want to sum the recurrence scores of the concepts to label their importance.

It is plausible that we might be able to make all of these by combining the simultaneous activation of senses with the concepts of good and bad: e.g.

$$\text{Good/bad} \cap \text{person} \;\rightarrow \text{love/hate} \tag{11}$$
$$\text{Good/bad} \cap \text{senses} \;\rightarrow \text{happy/sad} \tag{12}$$
$$\text{Good/bad} \cap \text{comparison} \;\rightarrow \text{true/false} \tag{13}$$

Emotional aggregate 'feelings' are essential prerequisites to interpretation. We measure them from running context.

4 The scaling of evidence/agency (singletons and ensembles)

As we aggregate data into bulk collections, it matters whether we aggregate over space or over time, which labels we keep and which we throw away.

4.1 Similarity and difference (symmetry versus transition)

This has a spacetime interpretation in the way we aggregated patches of data:

- Constant intent/type aggregates (bulk regions)
  - **Temporal** aggregation - e.g. repetition at a constant point, with batch learning semantics.
  - **Spatial** aggregate population - e.g. accumulation over a random walk, with realtime cognitive semantics
- Change of intent/type delimiters (boundaries between regions)
  - **Temporal boundary** - change of mind/policy
  - **Spatial boundary** - workspace, namespace, scaling cluster, redundancy

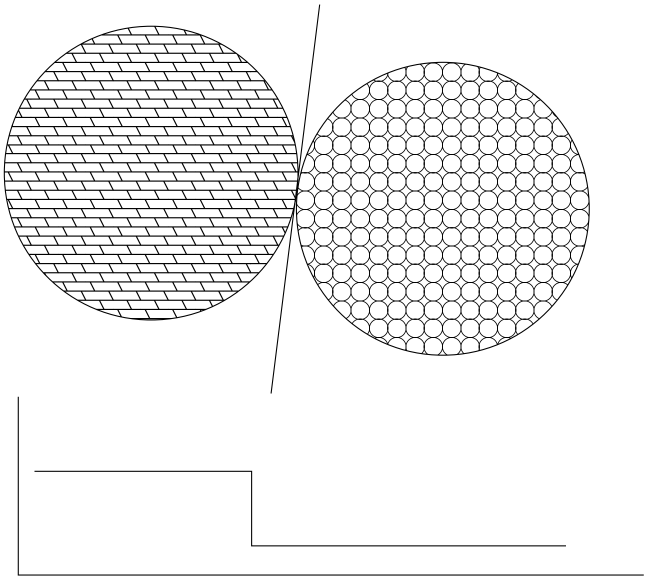

Figure 6: The semantics of symmetry versus change.

4.2 Distinguishabilty of agents (naming cattle vs pets)

The semantics of data are different depending on whether samples originate from from a single source or from a collection (statistical ensemble) of sources. People sometimes talk about *pets versus cattle*, meaning sample instances with names and labels versus herds without distinct identities. In data terms, these are singletons and ensembles.

4.2.1 Cattle (ensemble semantics)

In a batch experiment, all data points are considered to be equivalent, at some scale. Thus, if the dataset is a stream of scalar values q_i, then every q_i is interchangeable. The indices i may be relabelled and the results will be preserved. The properties that are derived from the dataset must therefore be invariant under relabellings of the array index i.

If the data are objects, in which clustered data have sub-members in tuples, then each tuple (collectively) has the same semantics as each other tuple, and the members are the same in all tuples. This is superagent indistiguishability, with a common superagent directory [2]. Thus, each subagent within a tuple has the same semantics as the corresponding subagent in another tuple. This is also represented by invariance under relabellings of i:

$$array[i] \quad \leftrightarrow \quad array[j] \tag{14}$$
$$\implies \quad array[i].subagent \quad \leftrightarrow \quad array[j].subagent \tag{15}$$

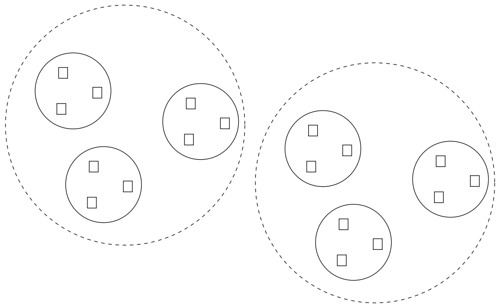

Figure 7: Indistinguishable agents at different scales.

Data in this form are sometimes called ‘cattle’ (a generic herd) rather than ‘pets’ (which have names). All elements are part of a symmetrical coarse-grained superagent and share the same associations.

4.2.2 Pets (individual semantics)

In a time-series, or on the interior of an object, the data are labelled sequentially, and the ordering leaves a unique name on each measurement. Each measurement therefore has distinuishable semantics, and may form independent associations.

4.3 The irreducible associations and distinguishability

The properties of this matrix reveal an interesting semantic interpretation of agency:

AGENT/PROMISE	SELF (SCALAR,LOCATION)	OTHER (VECTOR,PATH)
interior (state)	EXHIBITS	NEAR
exterior (transmission)	CONTAINS	ENABLES/CAUSES

- Diagonal properties are non-directional distinguishable properties of individuals.
- Off-diagnoal directional non-distingiushable ensemble properties of collectives. Eliminate distinction

There is no requirement of unique naming, but if we do not name things individually, they have only ensemble significance, as they are effectively symmetrical. Naming (namespaces) break symmetry.

5 Cellibrium project associations

The CGNgine software agents (part of the Cellibrium) project, is derived from the original CFEngine project code [8]. The project assumes inputs into a semantic knowledge network from:

1. (RobIoTs) The `cgn-monitord` agent measures and classifies probed input as ‘context’ classes and then constructs clusters using the foregoing rules.
2. (AnthropIoHs) Importing and scanning data streams by pattern matching and coordinatization.
3. (LINGugine) Linguistic associational input by experts, in the manner of an expert system.

```
/***********************************************/

const int CONTAINS  = 1;  // The irreducible assocs
const int FOLLOWS   = 2;
const int EXPRESSES = 4;
const int NEAR      = 8;

const int F = 0; // Forward assoc
const int B = 1; // Inverse assoc
```

```
/*********************************************/

enum associations // aliases for irreducible types
{
    a_concurrent,
    a_contains,
    a_name,
    a_origin,
    a_hasattr,
    a_hasvalue,
    a_hasinstance,
    a_approx,
    a_ass_dim
};

/*********************************************/

char *A[a_ass_dim+1][2] = // Association alias text
{
    {"seen concurrent with","seen concurrent with"},
    {"contains","is a member of"},
    {"is the name for","is referred to by"},
    {"is the source of","originates from"},
    {"has attribute","is an attribute of"},
    {"has value","is the value of"},
    {"has instance","is an instance of"},
    {"is approximately","is approximately"},
    {NULL, NULL},
};
```

References

[1] M. Burgess. Spacetimes with semantics (i). *http://arxiv.org/abs/1411.5563*, 2014.

[2] M. Burgess. Spacetimes with semantics (ii). *http://arxiv.org/abs/1505.01716*, 2015.

[3] M. Burgess. Spacetimes with semantics (iii). *http://arxiv.org/abs/1608.02193*, 2016.

[4] M. Curd and J.A. Cover, editors. *Philosophy of Science: The Central Issues*. Norton, 1998.

[5] J. Topping. *Errors of Observation and their Treatment*. Chapman and Hall, 1972.

[6] R.H. Dieck. *Measurement and Uncertainty (Methods and Applications)*. Instrument, Systems and Automation Society, Triangle Park, North Carolina, third edition edition, 2002.

[7] J. Schwinger. *Quantum Kinematics and Dynamics*. Addison Wesley, 1970.

[8] M. Burgess. Cellibrium project. https://github.com/markburgess/Cellibrium.