Some Notes on the Nature of Things

V0.3

Matthew West – Shell Services International

1 Introduction

The purpose of the paper is to provide a background to the language EXIST (EXpression of Information based on Set Theory) and the SC4 Integration Model (IM) that are part of the SC4 Data Architecture Project in ISO TC184/SC4/WG10. It contains some reflections on the nature of things in plain English, based on a range of reading in the subject. It documents the basis that has been chosen for EXIST and the SC4 Integration Model, so it is explicit and can be criticised.

2 Background

Within ISO TC184/SC4 – Industrial Data, the difficulties of achieving compatibility have become apparent. A Preliminary Work Item (SC4 Data Architecture) has been established for some time. A New Work Item has been approved for a Technical Specification of an information integration architecture and methodology, ISO 18876 – IIDEAS (Integration of Industrial Data for Exchange Access and Sharing). Further parts of this standard are planned and include an integration model designed for large scale integration, and a language to support the development of the model, and the mapping of other models to it.

3 Models

A model is some representation of some part of the universe we exist in and of our imagination. The model will be a limited representation of this reality, either by design, because the purpose of the model is limited, or simply because of the limits of our knowledge. There are many viewpoints that can be taken. Unfortunately, the models relating to these different viewpoints do not naturally fit together like jigsaw puzzle pieces. Rather they overlap, and sometimes appear to conflict. Often the viewpoint and constraints for the model are not explicit. From time to time a new theory or model appears which is able to explain why the conflicting viewpoints on the one hand conflict, and on the other hand are valid within some constraints. Often we continue to use the previous model, even though we have a more accurate one. For example, Newtonian Physics has been surpassed by the work of Einstein and others, yet it was the basis for sending a man to the moon.

However, sometimes the incompatibilities between models are inconvenient and expensive. An example of this is information models. The problem arises when information that is created according to one model overlaps with information used by another model. With the same information in two models there are the costs of duplication, and the possibility of inconsistency due to human error. Whilst it might be desirable to retain the two separate models because of their utility, it is certainly desirable to automate the synchronisation of the information that they hold, where there is overlap. One way that this can be done is by using an integration model that is able to support and explain the different views of the two (or more) models for which synchronisation is required. Information is transferred between the two models via the integration model.

In this case it will follow that a "better" integration model will be one that can integrate and synchronise information from a wider range of overlapping models. From the general discussion on models above, it follows that such a model will have a more fundamental view of our universe than the models it seeks to integrate, or other less capable integration models.

It is desirable for the Integration Model and EXIST to be based as closely as possible on the underlying nature of things. There is a branch of philosophy that is concerned with discovering the underlying nature of things. This is metaphysics. Historically this is a field of endeavour that has not been highly regarded in industry, on the one hand because of the apparent remoteness of the subject from practical matters, and the tendency to use obscure language that can make the field inaccessible. However, there is much work that has been done, particularly in the last 200 years, on which the development of an integration model can draw with benefit. This includes the development of logic and set theory.

4 Basic Elements

The basic types of thing that are considered to exist are:

- Individuals
- Sets
- Tuples

These are described in the following sections, but briefly:

- Individuals are things that generally exist in space/time. This includes things like physical objects, and events.
- Sets are collections of things (but not assemblies or fusions of them).
- Tuples are two or more ordered elements. Relations are classified tuples.

In addition, the availability of logic is assumed. This is adequately described elsewhere and is not further discussed here.

5 Individuals and their Identity

Broadly, individuals are those things that exist (or could exist) in space-time. This includes things that persist and things that mark changes.

5.1 Physical objects

There are two approaches that have been taken to understanding physical objects. The historical and intuitive approach is of 3D objects that endure somehow through time. An approach that has been developed mostly in the 20th Century is that of seeing physical objects as being 4 dimensional objects, i.e. as being a piece of space-time or spatio-temporal extent. Here a physical object is a swept 4D volume in space-time, and whatever matter is in that.

5.1.1 3D Objects

It is interesting to look at the identity basis for what we think of as ordinary physical objects. A first attempt at defining what a physical object is will usually yield something like "An object that persists through time, i.e. has material continuity". However, it is generally agreed that (at least some) objects are allowed to change over time. If my arm is chopped off, I am still me not someone or something else (as long as I survive). So the basis of "material continuity" is modified to "essential material continuity". The main problem here is that what is essential can be in the eye of the beholder.

However, these are not the only types of objects that exist in space-time. Take for example The Chairman of Shell. There is certainly not "essential material continuity" here. Every so often the object changes all its material, when the person who is chairman changes. The continuity in this case is functional continuity, not material continuity, though it consists of material at all times when it exists.

A related question that now arises is, are objects that coincide at some point in time identical? Take the Chairman of Shell, and Mark Moody-Stuart. Are these the same object because they coincide?

We would say not, because although they coincide now, they have not always, and presumably there will be change in the future.

5.1.2 4D Objects

Treating physical objects as 4D objects (space and time) avoids the ambiguity of 3D objects and persistence. Here a physical object is defined by its spatio-temporal extent (rather as a set is defined by its members). This means that there are an awful lot of physical objects out there that are not interesting. That is OK, just because they exist doesn't mean we have to be interested in them. We of course will be interested in physical objects that belong to interesting classes, such as person, car, etc.

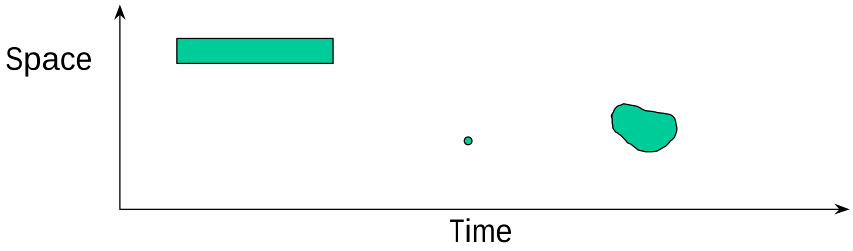

Figure 1: A space-time map

In addition this means that physical objects can have temporal parts (states) as well as 3D parts. This helps to deal with change over time, and coincidence of objects. For the case of the Chairman of Shell and Mark Moody-Stuart, there is a state of Mark Moody-Stuart that is also a state of The Chairman of Shell. Both Mark Moody-Stuart and the Chairman of Shell can be distinguished, because their overall 4D objects are different.

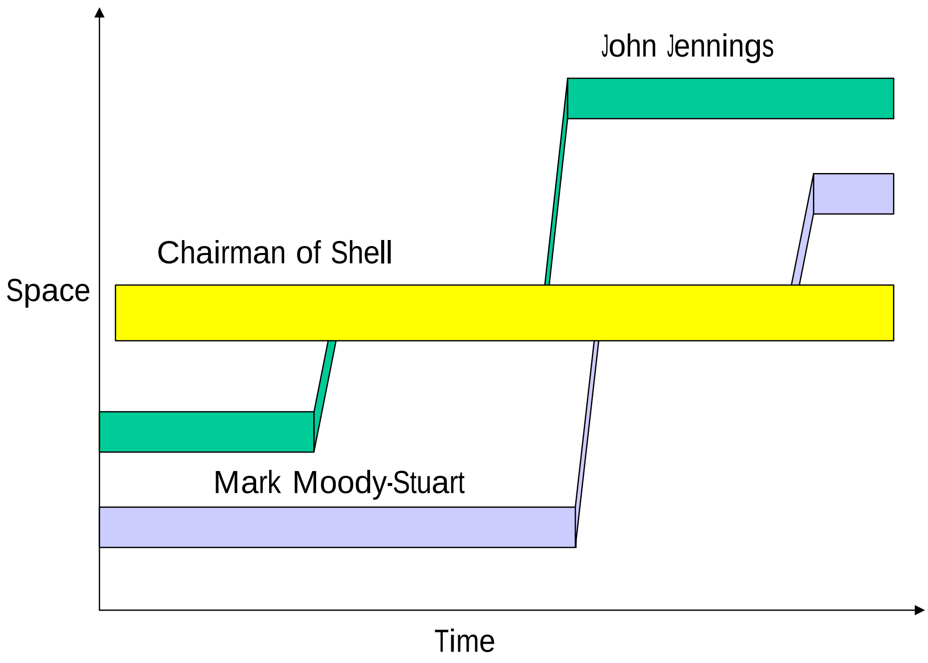

Figure 2: An example of objects that overlap.

This also helps to deal with change in relationships. Instead of the relationships having to be objectified (and become abstract objects of some sort) they can be replaced by states, and simple relations between them.

5.1.3 Managing Change

There are (at least) 5 ways of managing change, these are:

- Snapshot
- Audit trail
- Version management,
- Temporal relationships (associations),
- States

5.1.3.1 Snapshot

This is really the “do nothing” scenario. Here only the current state is maintained, and history is abandoned.

5.1.3.2 Audit Trail

An audit trail maintains only the current state, but also a register of all the changes that have lead from some previous state to the current state. The idea is that you could “rewind” the events to get to some previous state. The disadvantage is that only one state is visible at any time.

This approach is quite often used, in conjunction with others, for recovery of systems or files. The undo function in a word processor for example.

5.1.3.3 Version Management

With version management, when there is a change to an object, the change is recognised by making a new version. The continuity is recognised by being a version of the same object. The question arises immediately as to what sort of change requires a new version, and when does it happen. Take for example a large object like a process plant or an aircraft carrier. In principle the change of a bolt or a light bulb means that the plant or aircraft carrier has changed, and a new version should be created. This is impractical, because it implies making a copy of a vast amount of information, and changing something very small. Further, this would happen with considerable frequency.

In practice, version management is used to register changes in the design of an object, rather than in the object itself. So if the engines in an aircraft carrier are upgraded (rather than just replaced) we might think that the ship was a new version. Interestingly the design of something is the class it is a member of. Creating a new class when there is a change in the specification is appropriate, because this implies a change in membership, and a class is defined by its membership. The use of versions to manage change in the use of classes will be considered for the SC4 Integration Model.

5.1.3.4 Temporal Relationships – Associations

Temporal relationships, or associations as they are known in EPISTLE, recognise that change takes place by understanding that relationships between individual things are not necessarily permanent, but may be only for a period of time. For example, my owning a car is not a permanent relationship, but one that exists from when I bought the car, to when I sold it. The consequence of this approach is that the associations are modelled as entity types rather than as relationships. This is necessary because there is temporal information that needs to be recorded about the association, specifically its start date and end date. It also means that there is a record of the previous relationships as well as the current one.

The disadvantage of this approach is that it objectifies something that is not a “real” object, but something abstract – a relationship, and whilst it is a significant improvement on the approaches above for dealing with change amongst individuals, this does eventually cause problems.

5.1.3.5 States

Here, instead of looking at the relationship as lasting for a period of time, you identify the states of the objects that are related. So to use the car ownership example, you would say that I owned a temporal part of the car. The good thing about this is that the temporal part of the car is something quite physical, it is a spatio-temporal extent that is a part of the overall spatio-temporal extent that is the car. Further, the relationship does not now need to reflect the period of its validity, which is appropriate to an abstract object.

Many difficult problems can be solved effectively and efficiently by applying an analysis based on states. Further, it is the approach that is best able to integrate the information that is in models using some of the other approaches identified here. Consequently, this is the approach that is adopted for the SC4 Integration Model for change in physical objects.

5.1.4 Organisational Levels

The world about us consists of objects that display different levels of organisation, where objects with a higher level are an organisation of objects at a lower level. For example, a hunk of clay is made out of molecules that are held together (organised) by inter-molecular forces. Equally, the molecules are made out of atoms that are held together by inter-atomic forces.

It is interesting to note that different sciences concern themselves, broadly, with different strata amongst these levels. Physics concerns itself with the atomic level and anything lower than that. Chemistry focuses on the molecular level, whilst Biology concerns itself with living cells and life at a higher level of organisation.

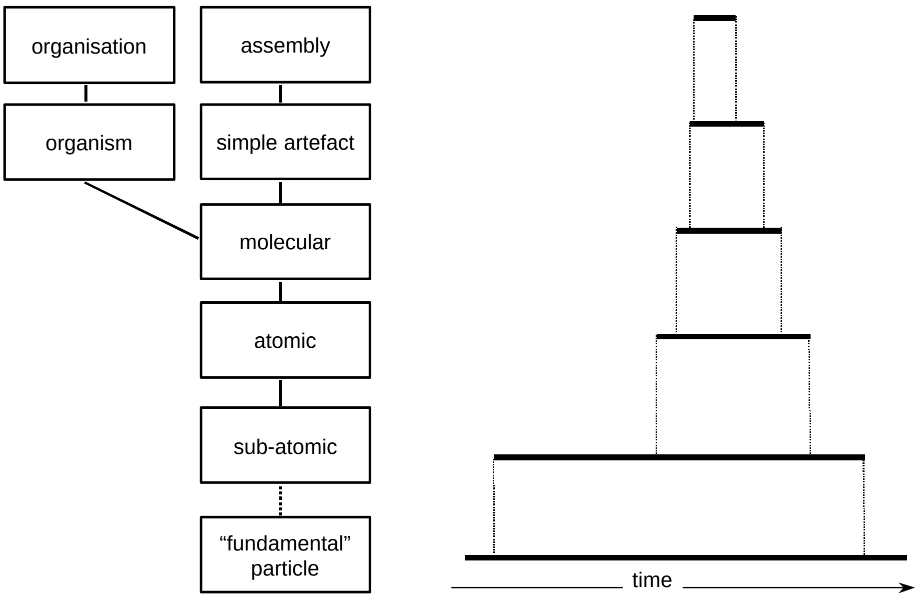

Figure 3: Organisational Levels.

Combining organisational levels with 4D objects results in the realisation that an object at a higher organisational level will be a state of an object at a lower organisation level. For example, I am the living state of my body.

However, some times the object at the lower organisational level is 4D coincident with the object at the higher organisation level. Consider a clay statue that is constructed in two halves. When the two halves are put together, you have an object that is a lump of clay, and an object that is a statue having been created. If this is now fired, and later in the day smashed, then both the lump of clay and the statue are destroyed also at the same time.

A simple use of 4D objects says that there is only one object here, the lump of clay and the statue are one just because they have the same spatio-temporal extent. This object is simply a member of both the lump of clay class, and the statue class. However, in the way I was talking about the subject, there were apparently two objects. This says that our intuition is that the identity basis for an object is its spatio-temporal extent and its classification at each organisational level at which it is a whole temporal extent.

It is noted here that this view of the world can be supported by a view that understands spatio-temporal extents as the root identity basis for physical objects. These together with organisational levels can support a view that the spatio-temporal extent and its organisational level are the identity basis, and so this view is preferred for the Integration Model.

A final point to understand here is that not all states of physical objects represent a new object at a higher organisational level. For example, being an owner, or a student, or an employee, does not create what we would consider to be a higher level object. This gives rise to the question of what it takes to be an organisation of one level to give another, rather than just being a state of one level. At present this is still open.

5.2 Activities and Events

An event is an elementary change. An event happens at a point in time, has zero duration, and marks the start of a state and/or the end of a state.

Events can be aggregated.

An interesting question is what is the duration of an aggregated event? Since in principle this is the sum of the duration of its components, it could be argued that the true duration is zero. However, for a point based approach to geometry, space is made up of (an infinite number of) points. In practice we are usually only interested in a small number of key events in an activity. Therefore, we would consider the duration of an activity to be the time between the first important event and the last important event of the activity, rather than the integral of all the changes.

There may be a number of roles that are played in an activity. As stated above, roles are states of physical objects. For example, a meeting activity involves the states of the individuals taking part in the meeting, and the place where the meeting is happening.

5.3 Spatio-Temporal Extents, Planes, lines and Points

Generalising the discussion above leads to the possibility of objects of different dimensionality. An object’s dimensionality is relative to the dimensionality of the space within which the object exists. In a 4D space you can have a range of objects where one or more dimensions are constrained.

6 Possible Worlds

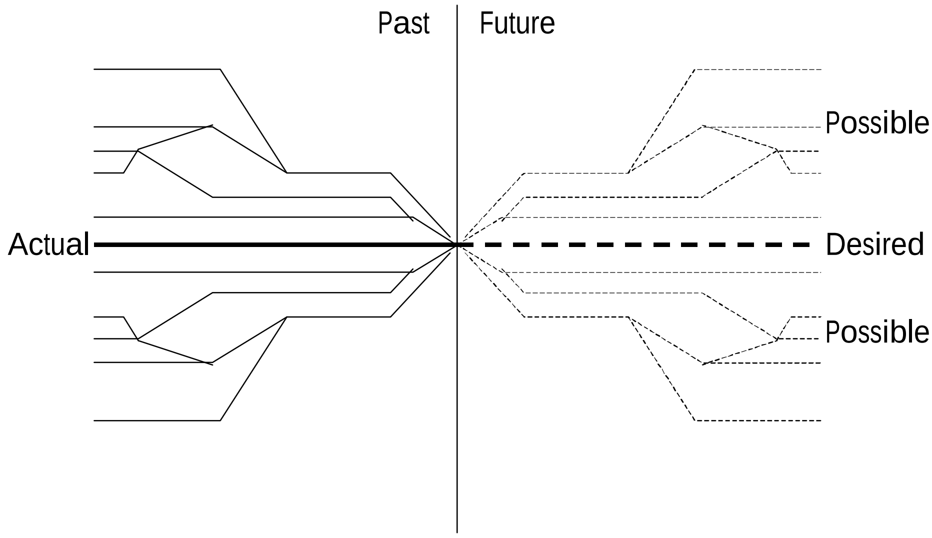

Figure 4: Possible worlds.

Possible Worlds deals with what could be, or might have been. The actual world is one possible world.

Each possible world must be internally consistent. It is a coherent history and future, just as the actual world is. There are an infinite number of possible worlds because of the large number of possible variations there are for any object as time passes, never mind the interactions between objects.

There are (at least) two views on the nature of possible worlds. One is of parallel worlds, in which each object exists again in each possible world, with variations relevant to that world. The alternative is one of branching, where bits that are the same are shared between worlds. I prefer the latter approach. This is because it fits well with temporal parts, and because an important use of possible worlds is planning for the future, and there, you are not planning for a parallel world to come about, but about deciding which of the possible branches you want to actually happen.

One problem with possible worlds (whichever view you take) is trans-world identity. The Spatio-Temporal definition of individual is violated (apparently) by the different spatio-temporal extents in different worlds. However, I take the view that the individual is the spatio-temporal extent across possible worlds as well. In this sense an individual is what it could be/have been as well as what it actually is. Adding possible worlds is like adding a 5th dimension, where we see what is projected onto what we know as the actual world.

It should be noted that this is a particularly challenging area.

7 Sets, Classes and Properties

7.1 Sets and Classes

Classical theory and our natural intuitions suggest that classes are those things that share some common properties. Following from this it has been suggested in the early development of formal set theory, that for any set of properties there was necessarily a set of things that had those properties. This would certainly have been convenient if it were true, but Russell was able to show that it was false through Russell's Paradox.

Russell’s Paradox identifies a set of properties that no set can hold. This is the property of not containing itself as a member. If the set concerned contains itself, then it does not qualify for membership, and if it does not contain itself then it does not contain all the sets that do not contain themselves.

A considerable part of the work on set theory this century has been focused on overcoming this problem. However, the intuition that sets are defined by the property sets the members share has been very strong. As a result, a lot of effort has been spent, for example, on defining sets as things that cannot have themselves as members (Zermelo-Fraenkel), or that some sets are not members of any other set (von Neuman).

There are some basic problems to be found here. For example, if a set cannot be a member of itself, then there can be no “set of all sets”, which certainly has to exist from a practical point of view, and must obviously contain itself if it does exist.

Here, we take a different approach. We say that sets are defined by their membership. If I can construct a set, it exists. By constructed here I do not mean to exclude infinite sets or sets that include themselves.

It is therefore no problem for a set to be a member of itself, because this is something that can be constructed A = ={a, b, c, A}. On the other hand, there is no set that can be constructed that contains all the sets that do not contain themselves. So instead of having a paradox, you simply don’t have a set.

This constructive approach means that there are an awful lot of sets out there. For any group of objects I can have a set for every possible combination (but not order) of objects. Once I have those, I can start combining the objects and the sets I have just created, and so on.

The good news is that we are not interested in many of these. It turns out that we are interested in (some of) those sets that have common property sets. Here we call these sets classes.

In classical approaches, different property sets can give rise to the same set of members, and may be considered different classes. This means that the complete common property set for the set concerned is at least the union of the two property sets. It should not be surprising that different subsets of these should be sufficient to identify all members of the set.

An issue is property sets for which the Null Set is the result. There is only one Null Set (the set with no members is also defined by its membership). I believe these will always be the result of specifying the intersection of non-null sets, but at least 2 being disjoint. Thus an apparent contradiction “The set of all square triangles” isn’t one. It is simply that the set squares and the set triangles do not intersect.

7.2 Properties

Properties are classes. For example, that degree of hotness, which can be described as 20 Celsius, is a class that some spatio-temporal extent might belong to.

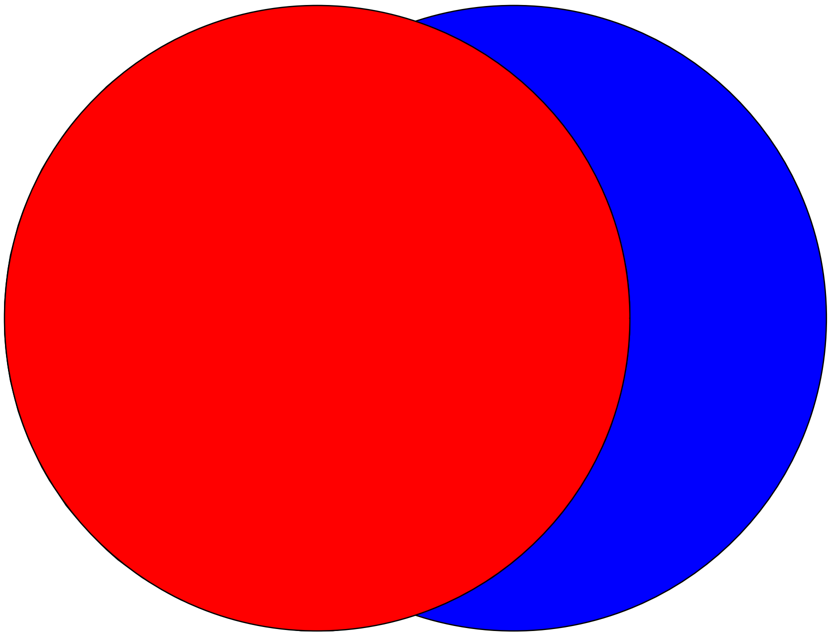

Figure 5: Classes defined by properties are the intersection of those properties

This means that a class, is the intersection of the properties that define the class. This together with a 4D view of what physical objects are, greatly simplifies many things.

8 Tuples and Relations

A tuple is an ordered collection (list) of elements, where an element may be repeated in the list. The order is significant. A tuple has no inherent meaning attached to it. For any set of elements, there are the tuples that are all the combinations of the elements, in all the possible orders.

A tuple has meaning when it is a member of a class. The implication is that the class gives significance and meaning to each position in the tuple. A tuple may be a member of more than one class. For example, the tuple <Matthew, Lydia> tells you nothing about the people concerned. Only a classified tuple (a relation) has meaning, e.g. married: <Matthew, Lydia>.

In these terms a relation is the classification of a tuple. Both tuples and relations as defined here are timeless. Thus the above relation is not correct, it should be married: < Matthew from beginning to the end of the marriage, Lydia from the beginning to end of the marriage>

9 Propositions and Assertions

An assertion is the claim that a proposition is true. For example, the proposition "John is married to Mary" is either true or false. When we say "John IS married to Mary" we are making the assertion "It is true that John is married to Mary".

Propositions can be qualified, e.g. "Henry believes that John is married to Mary", or "James says that John is married to Mary". Thus it is possible to make statements about propositions.

9.1 Necessary Statements (Rules) vs States of Affairs

Some propositions are always true (or false). Others may be true under some circumstances and false under others.

Statements that are always true are necessary statements or rules. An example is that a dog (in the sense of a four-legged animal) is a subtype of animal. There are no conditions under which this is not true. Indeed it is generally the case that rules are about the relationships between the members of classes. This should not be a surprise, since classes are universal/timeless.

A state of affairs is a statement of how things happen to be, but not how they are necessarily. So I am writing at my computer, but I could be watching the television. States of affairs generally involve individuals.

10 Conclusions

A number of basic positions have been stated in this document. The purpose is on the one hand to make them explicit so that they can be understood, and on the other hand to make them open to challenge.

In the event that I have failed to make them understood or that you wish to challenge them, please contact the author at matthew.r.west@is.shell.com.