eSpaceML: An Event-Driven Spatial Annotation Framework∗

Kiyong Lee[a], Jonathan Webster[b] and Alex Chengyu Fang[b]

[a]Department of Linguistics, Korea University
Anam-dong, Seoul, Korea
klee@korea.ac.kr
[b]Department of Chinese, Translation and Linguistics, City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong SAR
{ctjjw,acfang}@cityu.edu.hk

Abstract. This paper proposes *eSpaceML* as a representation scheme for annotating event-driven spatial expressions in natural language. It adopts *SpatialML* (MITRE, 2009) and *ISO-Space* (ISO, 2010) as a basis for the development of a novel, distributed spatial annotation scheme. *SpatialML* focuses on the annotation of spatial locations and their topological relations, while both *ISO-Space* and *eSpaceML* attempt to extend the scope beyond the treatment of toponyms. *ISO-Space* and *eSpaceML* also link space to events in various ways but with considerable differences. Unlike *ISO-Space*, which attempts to provide a self-contained framework for the various links, *eSpaceML* treats them in a distributed manner, operating as a pivotal system that refers to several other established annotation schemes such as *MAF* (ISO, 2008) and *ISO-TimeML* (ISO, 2009c) for morpho-syntactic as well as temporal-eventual annotations.

Keywords: semantic annotation, space and event, ISO, language resource management

1 Introduction: Aim and Scope

This paper aims at a description of the development of *eSpaceML*, an event-driven spatial annotation scheme for natural language texts. This work adopts *SpatialML* (MITRE, 2009) as a basic framework for annotating spatial locations and their topological relations. The scope of *eSpaceML*, however, goes beyond the annotation of place names and their relations and attempts to extend its task to events by linking events to space, as is implied by the prefix *e-* in its name *eSpaceML*.[1]

The idea of linking events to space has originated from *ISO-Space* (ISO, 2010). At its current preliminary stage, *ISO-Space* annotates events as well as space, and also signals for linking those two directly in a self-contained manner. Our proposed *eSpaceML*, on the other hand, is designed to outsource various tasks other than the main task of linking events to space. For the annotation of events, for instance, it refers to *ISO-TimeML* (ISO, 2009c) that annotates events and time. For the reference to markables, i.e., those expressions that need to be marked up, it uses the information provided by other annotation frameworks such as *MAF* (ISO, 2008), which

---

∗ The work reported in this paper was supported in part by research grants from the City University of Hong Kong (Project Nos 9610053, 9610126, 7008002, 7002387 and 7002190). The authors would like to thank Roland Hausser and the reviewers for their valuable comments and suggestions.

[1] As defined in ISO-TimeML (ISO, 2009c), the term *event* covers all of the types of eventualities such as states, processes, and transitions. See also Pustejovsky (1991) for these three event types.

provides morpho-syntactic information. The proposed *eSpaceML* is thus designed to provide a net-based distributed annotation scheme in a pivotal format.

This paper claims that *eSpaceML* admits a compositional approach to semantic annotation by linking and merging various layers of semantic annotation into an interoperable pivotal format. Its interoperability with other annotation schemes within the pivotal system is also shown to guarantee its sustainability for any further extension.

2 *SpatialML* and *ISO-Space*

Two important annotation schemes[2] for marking spatial expressions are introduced here: one is *SpatialML* and another is *ISO-Space*. They provide a basis for the construction of *eSpaceML*.

2.1 *SpatialML*

SpatialML focuses on the annotation of spatial locations and their topological relations. For this, four XML element tags are introduced: `<PLACE>, <SIGNAL>, <LINK>` and `<RLINK>.` Consider Example (1) below, where markables are enclosed in square brackets.

(1) [*Royal Plaza*] *is* [*in*] [*Mong Kok*], [*north*] *of* [*Tsim Sha Tui*], [*Kowloon*].

```
<PLACE id=1 type="fac" descr="hotel" form="NAM">Royal Plaza</PLACE>
<PLACE id=2 type="PPL" country="HK" form="NAM">Mong Kok</PLACE>
<PLACE id=3 type="PPL" country="HK" form="NAM">Tsim Sha Tsui</PLACE>
<PLACE id=4 type="PPL" country="HK" form="NAME">Kowloon</PLACE>
<SIGNAL id=5>in</SIGNAL>
<SIGNAL id=6 type="DIRECTION">north</SIGNAL>
<LINK id=7 source=1 target=2 signal="5" link Type="IN"/>
<RLINK id=8 direction=6 source=3 destination=2 signals="6"/>
<LINK id=9 source=3 target=4 linkType="IN"/>
```

In the corresponding *SpatialML*-style annotation above for (1), there are four occurrences of place names that are marked. *Royal Plaza* is a hotel, thus being marked as a type of FACilities, while the other three are populated places PPL, all belonging to Hong Kong. Second, two signals, *in* and *north*, are marked, where *north* is a signal of type `DIRECTION`. There are two types of linking, `<LINK>` and `<RLINK>. <LINK>` expresses a simple topological relation such as containment or inclusion between two places, as between *Royal Plaza* and *Mong Kok* and between *Tsim Sha Tsui* and *Kowloon*. The tag `<RLINK>,` on the other hand, expresses a relative link between places involving direction and distance. Here the link type `DIRECTION` indicates moving from one place to another in a certain direction, namely *north* here in this example, to locate a place, namely *Mong Kok*.

SpatialML-3.0 fully accommodates both prior work and current developments on toponyms and related subject areas. *SpatialML*, however, fails to relate spatial annotation to other aspects of semantic annotation, especially to the annotation of events. This task is taken up by *ISO-Space* as part of its immediate task and also by *eSpaceML* which takes a distributed approach of referencing *ISO-TimeML* that annotates events.

2.2 *ISO-Space*

James Pustejovsky leads a working group on *ISO-Space* that aims at designing a comprehensive annotation scheme for marking regions, paths, states, and motions and their relations, both static and dynamic. *ISO-Space* is intended to be comprehensive in several respects:

---

[2] There are a long list of other works related to space in formal semantics, space logic, and ontology such as the spatial extension of GUM (the generalized upper model), as is presented in Hois (2010), for topological and ontological structures. They are, however, left for the future study of spatial annotation of text.

- It extends the scope of spatial locations to cover not only ordinary place names (e.g. *Boston*), but also coerced places (e.g. *car* as in "*in a car*") or individuals that occupy a place (e.g., *John* as in "*John is standing on a platform*"). These locations as a whole are marked as `<REGION>`.
- The `<PATH>` element is introduced to delineate particular regions with beginning and end points (e.g. "The train runs *from Boston to New York*.")
- In *ISO-TimeML*, events cover all types of eventualities such as states (e.g., "John *lived* in Hong Kong.") and motions or transitions (e.g., "John *flew* to Beijing".)
- In order to relate all these regions and paths, two signals are introduced: one is a function word marked as `<S_FUNCTION>` that creates a new region out of some input region and another is a relational word, marked as `<S_SIGNAL>`, which directly relates regions.
- `<QSLINK>` (qualitative spatial link) marks spatial relations, which can be either topological (e.g., *in*) or relative (e.g., *behind*).[3]

Example (2) is quoted from Pustejovsky and Moszkowicz (2009):

(2) The park is behind the store.

```
<REGION rid="r4" extent="the park"/>
<S_FUNCTION sfid="sf4" extent="behind" outputType="REGION"
 input_regionID="r5" output_ID="r6"/>
<REGION rid="r5" extent="the store"/>
<REGION rid="r6" extent=NULL source="sf4"/>
<QSLINK qsid="qs1" type="relative" relation="behind" regionID="r6"
 related_regionID="r5" spatial_relationID="sf4" />
<QSLINK qsid="qs2" type="topological" relation="IN" regionID="r4"
  related_regionID="r6"/>
```

Here both *The park* and *the store* are marked as `<REGION>` with id's, `r4` and `r5`, respectively. Then, their relative relation *behind* is marked by the first `<QSLINK>: qs1`. The second `<QSLINK>: qs2`, on the other hand, locates the park (`r4`) at a region (`r6`) demarcated by the spatial function (`<S_FUNCTION>: sf4`) that is triggered by *the store* (`r5`) with respect to the relative spatial relation *behind*.

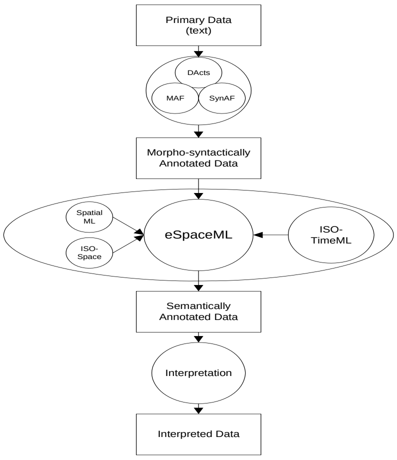

Figure 1: Pivotal Format

3 The Design of *eSpaceML*

Just like *SpatialML* and *ISO-Space*, *eSpaceML* is an annotation scheme for marking spatial information and their relations in natural language. The proposed *eSpaceML* is, however, designed to import a variety of relevant information from other available annotation schemes into an interoperable representation scheme for annotating space and events, thereby minimizing the actual task of spatial annotation.

3.1 Processing Modules in eSpaceML

There are three processing modules in eSpaceML, each represented by a circle. The preliminary task of annotation is to decide on its markables, namely those

---

[3] The tag `<QSLINK>` may be replaced by two different tags `<TOPLINK>` (topological link) and `<RLINK>` (relative link).

expressions in a text that need to be annotated. To refer to such markables, as shown in the topmost circle in Figure 1, *eSpaceML* has recourse to ISO standards such as *MAF* (ISO, 2008) and *SynAF* (ISO, 2009b) that also provide other necessary morphological and syntactic information for the operation of semantic annotation. Another ISO standard for language resources, *DActs* (ISO, 2009d), marks the functional segments of parts of an utterance, each of which has a unique communicative function in dialogue.

Unlike *SpatialML*, both *ISO-Space* and *eSpaceML* aim at linking space to events. These two, however, have a different strategy. *ISO-Space* treats such linking within its annotation scheme by introducing necessary annotation tags such as `<STATE>` and `<MOTION>`. Our *eSpaceML*, on the other hand, simply refers to *ISO-TimeML* (ISO, 2009c) and some other necessary and available annotation schemes, as shown in the central circle of Figure 1. As a consequence, *eSpaceML* does not have those two tags in *ISO-Space*, namely `<STATE_LOCATION>` and `<MOTION_TRANSITION>`, which anchor events to locations or paths. Instead, it simply extends the function of `<LINK>` to such anchoring, just as `<TLINK>` in *ISO-TimeML* relates not only a time to another time, but an event to a time.

The pivotal format presented in Figure 1 also links *eSpaceML* to the interpretation module, the bottom circle in Figure 1, which provides semantic representation proper. As a semantic annotation scheme, *eSpaceML* is not responsible for full semantics. It simply lays a preliminary basis for doing formal semantics, as has been initiated by Bunt (2007), Katz (2007), Pratt-Hartman (2007), and Lee (2008). The burden of inferential task that might be expected from semantic annotation in general is thus relegated to the interpretation module.

3.2 Representation Scheme

The representation scheme of *eSpaceML* differs from *SpatialML* and *ISO-Space* in at least two respects. First, *eSpaceML* relates one annotation to another by outsourcing whatever is available, while focusing on space and events. Second, it provides a representation scheme conformant to ISO projects such as *LAF* (ISO, 2009a), a draft international standard for linguistic annotation framework, and *FSR* (ISO, 2006) that specifies how to represent feature structures in XML. As is discussed in Lee and Romary (2010), this conformance is required at least to guarantee the interoperability of a linguistic annotation framework with other ISO-developed annotation frameworks.

3.3 Standoff Annotation

LAF stands against the inline (embedded) notation of annotating linguistic data. Instead, it proposes standoff annotation. Accordingly, *ISO-Space* adopts standoff annotation, while *SpatialML* adopts inline annotation. Example (3) is used to illustrate both inline and standoff annotations of text.

(3) *Mia lived in* [*Atlanta*].

While an inline annotation would take the form `<PLACE id=1 type="PPL">Atlanta </PLACE>`, the standoff annotation could be written as `<PLACE id=1 type="PPL" extent= "Atlanta"/>`. *eSpaceML* also adopts standoff annotation. Unlike *ISO-Space*, however, markables are referred to in *eSpaceML* in an indirect way as is discussed presently.

3.1 Reference to Markables

LAF requires each data structure to consist of two substructures: one is a referential structure and the other a content structure which is represented in feature structures. A referential structure then identifies itself in a unique way, while referring to some other entity that is to be annotated. As represented in XML, each representation structure is specified with two attributes: `@xml:id` and `@corresp`. The former uniquely identifies a referential structure. The latter refers to a part of a text that has been annotated or some other entities that are related to each other. By convention, # is prefixed to each of those that are referred to.

Headed by a referential structure, its associated content structure is represented with feature structures, as specified in FSR.[4] Example (3) is used to illustrate the content structure in feature structures:

```
Text:     Mia lived in Atlanta.
Tokens: tk1  tk2  tk3 tk4  tk5
<MAF comment="partial">
  <TOKEN xml:id="tk1"/>
  <TOKEN xml:id="tk2">
    <fs type="msd">
      <f name="base" value="live"/>
      <f name="pos" value="PAST"/>
      <f name="eventType" value="STATE"/></fs></TOKEN>
  <TOKEN xml:id="tk3">
    <fs type="msd"><f name="sense" value="IN"/></fs></TOKEN>
  <TOKEN xml:id="tk4"><fs type="msd"><f name="pos" value="PN"/></fs>
  </TOKEN>
<wordForm xml:id="w1" corresp="#tk1"/>
<wordForm xml:id="w2" corresp="#tk2"/>
<wordForm xml:id="w3" corresp="#tk3"/>
<wordForm xml:id="w4" corresp="#tk4"/>
</MAF>
```

Here a data structure within *MAF* consists of a referential structure marked as `<TOKEN>` and a content structure represented in `<fs>` (feature structure) which is embedded in the referential structure. The element `<TOKEN>` identifies the five tokens: `tk1` through `tk5`. Here, some tokens are specified for their morpho-syntactic description (`msd`). Then the element `<wordForm>` identifies four words with predefined tokens: for instance, `w2` is marked as corresponding to `tk2`.

4 Integrating *SpatialML* and *ISO-Space* into *eSpaceML*

This section discusses the integration of *SpatialML* and *ISO-Space* into *eSpaceML*. As will be shown, *eSpaceML* represents an effective reduction of existing elements from *SpatialML* and *ISO-Space* but still maintains a good degree of flexibility in expression.

Table 1: Mapping *eSpaceML* to *SpatialML* and *ISO-Space*

SpatialML	*ISO-Space*	*eSpaceML*
`<PLACE>`	`<REGION> <PATH> <S_FUNCTION>`	`<REGION>`
`<SIGNAL>`	`<S_SIGNAL> <STATE> <MOTION>`	`<SIGNAL>`
`<LINK>` `<RLINK>`	`<QSLINK>/<TOPLINK> <QSLINK>/<RLINK>` `<STATE LOCATION> <MOTION_TRANSITION>`	`<LINK>`

Table 1 shows the mapping of *eSpaceML* to *SpatialML* and *ISO-Space*, according to which the tag `<PLACE>` in *SpatialML* is replaced by `<REGION>` in *eSpaceML*. This tag, however, comprises both `<REGION>` and `<PATH>` in *ISO-Space*. In *eSpaceML*, these two are differentiated by the attribute `@type`.

a. `<REGION><fs type="topological"/>`
b. `<REGION><fs type="path"/>`

As in *ISO-Space*, the path type of `<REGION>` refers to a region that is delineated by an origin or a destination, or both. The tag `<SIGNAL>` in *SpatialML* remains the same in *eSpaceML*. In *ISO-Space*, the tag has changed to `<S_SIGNAL>` to differentiate it from other signals such as

---

[4] ISO 24610-1:2006 *FSR* is slightly modified here to conform to the new version of *LAF* (ISO, 2009a).

temporal signals. There are two link tags: `<LINK>` and `<RLINK>` in *SpatialML*, and `QSLINK` and `<LINK>` in *ISO-Space*, while there is only one in *eSpaceML*.[5] In *eSpaceML*, two types of `<LINK>` are again differentiated by the `@type` attribute, as below:

a. `<LINK><fs type="topological"/>`
b. `<LINK><fs type="relative"/>`

4.1 Conversion from *SpatialML*

Each annotation of *SpatialML* can be converted into *eSpaceML* in a systematic way. The following shows a general frame:

`<ELEMENT>` Tags - The `<PLACE>` tag is replaced by `<REGION>` in *eSpaceML* in order to differentiate ordinary places from paths. While ordinary places are indicated simply by `<REGION>`, paths are additionally represented as `<REGION type="path">`. Similarly, the two tags of link, namely `<LINK>` and `<RLINK>` in *SpatialML* are also converted into one tag `<LINK>` but with two different type specifications in *eSpaceML*:

a	SpatialML	`<LINK>`
	eSpaceML	<LINK><fs type="topological"/>
b	SpatialML	`<RLINK>`
	eSpaceML	<LINK><fs type="relative"/>

Attributes and Values - In *SpatialML*, each element has the following form: `<TAG id="ID" type="v0" a1="v1" a1="v2".../>`. In compliance with *LAF*, however, *eSpaceML* converts the above element form into a double-deck data structure that consists of a referential structure and a content structure:

<ELEMENT xml:id="ID" corresp="REFs" (signal="signalID")>
  <fs type="v0"> <f name="NAME$_1$" value="VALUE$_1$"/> ...<f name="NAME$_n$" value="VALUE$_n$"/>
  </fs>
</ELEMENT>

In this data structure model, the first element, which is represented as `<ELEMENT>`, stands for its referential structure and the value of the `@corresp` attribute is a (possibly null) sequence of what is referred to.[6] The second element, tagged as `<fs>`, is the content structure. It is represented in a feature structure (`<fs>`), consisting of a list of feature (`<f>`) specifications. To illustrate the conversion from SpatialML to eSpaceML, consider Example (4):

(4) [*Paris*] *is located* [*in*] [*Texas*].

Here is a *SpatialML* representation:

```
<SpatialML>
  <PLACE id=1 type="PPL" cvt="City">>Paris</PLACE>
  <PLACE id=2 type="STATE" country="US">Texas</PLACE>
  <SIGNAL id=3>in</SIGNAL>
  <LINK id=4 source=1 target=2 signals="3" linkType="IN"/>
</SpatialML>
```

---

[5] *ISO-Space* also introduces a pair of two tags <TOPLINK> and <RLINK> as an alternative.

[6] Given a pair of such referents, one is a relating entity while the other is an entity that is related to it, as will be illustrated presently. A null sequence may be referred to as the value of @corresp through base segmentation, for instance, as a specific gap between *Gia* and *pears* in the gapping structure *Gio loves apples and Gia [gap] pears.*

Its corresponding representation in *eSpaceML* is as follows.

```
<eSpaceML>
  <REGION xml:id="rg1" corresp="#tk1">
   <fs type="PPL"><f name="cvt" value="CITY"/></fs></REGION>
  <SIGNAL xml:id="sg1" corresp="#tk3">
   <fs type="spatial"><f name="sense" value="location"/></fs>
  </SIGNAL>
  <REGION xml:id="rg2" corresp="#tk4">
   <fs type="state"><f name="country" value="US"/></fs>
  </REGION>
  <LINK xml:id="lk1" targets="#rg1 #rg2">
   <fs type="topological><f name="linkType" value="#sg1"/></fs>
  </LINK>
</eSpaceML>
```

The two representations look like notational variations. They do, however, differ in their expressive power: unlike *SpatialML*, *eSpaceML* can, for instance, specify the content structure associated with the signal *in* that carries information about the sense of the signal, namely the sense of a location. This can be done either directly within the element `<eSpaceML>` or preferably by referencing to another annotation like *MAF*.

4.2 Net-based Distributed Task

Some parts of the annotation task can be distributed to other annotation schemes, forming a network. Morpho-syntactic description, for instance, is not a proper part of semantic annotation. Hence it can be relegated to morpho-syntactic annotation schemes such as *MAF*. The semantic annotation of events is treated in *ISO-TimeML*. Hence, spatial annotation can refer to it when linking space to events. At the final stage, overall distributed linking is required to form an interoperable network.

Semantic annotation marks up semantic features in a text. But this markup differs from doing formal semantics proper. A date, for instance, is normally marked up like `10/14` in a credit card, although it is marked up in a more explicit way as `2014-10-XX` according to an ISO standard. Unless there is an explicit way of interpreting these markups, it is not totally clear which is the year and which is the month referred to by these markups alone. Hence, annotations require explicit semantics for interpreting them.

Linking to *ISO-TimeML* - The two tags, `<STATE>` and `<MOTION>`, for instance, are introduced in *ISO-Space* to annotate events of two differentiate types. Then to anchor them to space, two tags `<STATE_LOCATION>` and `<MOTION_TRANSITION>` are introduced. Example (5) illustrates the anchoring of states to locations.

(5) John lived in Boston.

```
<ISO-Space>
  <STATE sid="s1" extent="lived"/>
  <S_SIGNAL ssid="ss1" extent="in" input1_regionID="r1"/>
  <REGION rid="r1" extent="Boston"/>
  <STATE_LOCATION slid="sl1" anchored_stateID="s1"
   anchor_regionID="r1" spatialRelationID="ss1"/>
</ISO-Space>
```

In *eSpaceML*, however, events are anchored to regions simply by linking space to events which are annotated in *ISO-TimeML*. The burden of differentiating types of events is also taken up by *ISO-TimeML* that annotates events as well as temporal information. Here is an example:

```
<MAF>
  <wordForm xml:id="w1" corresp="#tk3">
    <fs type="spatial">
```

```
      <vAlt><f sense="location"/><f sense="destination"/></vAlt></fs>
  </wordForm>
</MAF>
<ISO-TimeML>
  <EVENT xml:id="e1" corresp="tk2">
     <fs type="STATE"/><f extent="lived" tense="PAST"/></fs></EVENT>
</ISO-TimeML>
```

eSpaceML links space to events in the following manner.

```
<eSpaceML>
  <SIGNAL xml:id="sg1" corresp="#w3">
    <fs type="spatial"><f sense="location"/></fs></SIGNAL>
  <REGION xml:id="rg1" corresp="#tk4">
    <fs type=PPL"><f country="US" state="US-MA" ctv="CITY"/></fs>
  </REGION>
  <LINK xml:id="ln1" targets="#e1 #rg1">
    <fs type="anchoring"><f signal="#sg1"/></fs></LINK>
</eSpaceML>
```

As can be illustrated in the example above, *eSpaceML* makes references to *MAF* and *ISO-TimeML*. From *MAF*, it copies one of the two senses of the signal `sg1` which is appropriate to the type of the event `e1`. Here the event is of type `STATE`, as is annotated in *ISO-TimeML*. Hence, the location sense is chosen for the signal `sg1`. The `<LINK xml:id="ln1">` element links and anchors the state of John's living, as referred to by `#e1`, to a region. This region is referred to by `#rg1` by a signal `#sg1` that carries a locative sense. This annotation is sufficient to provide an interpretation which is identical to the one that can be obtained from the *ISO-Space* annotation given above. This shows that the `<LINK><fs type="anchoring">` element in *eSpaceML* may replace the `<STATE_LOCATION>` element in *ISO-TimeML*.

(6) *John drove through Boston.*

Example (6) illustrates how motions and paths are treated in *eSpaceML*. First, prerequisite annotations are assumed to have been provided through preliminary work as illustrated below:

```
<MAF comment="partial">
  <wordForm xml:id="w2" corresp="#tk2/>
  <wordForm xml:id="w3" corresp="#tk3">
    <fs type="msd"><f pos="prep"/><f sense="path"/></fs>
  </wordForm>
  <wordForm xml:id="w4" corresp="#tk4>
    <fs type="msd"><f form="name"/></fs>
  </wordForm>
</MAF>
<ISO-TimeML>
  <EVENT xml:id="e2" corresp="#w2">
    <fs type="process"><f tense="past"/></fs>
  </EVENT>
</ISO-TimeML>
```

Second, referring to preliminary annotations above, *eSpaceML* provides the following annotation for anchoring motions to paths:

```
<eSpaceML>
  <SIGNAL xml:id="sg2" corresp="#w3"/>
  <REGION xml:id="rg2" corresp="#w4">
    <fs type="PPL"><f country="US" state="US-MA" city="CITY"/></fs>
  </REGION>
  <LINK xml:id="ln2" targets="#e2 #rg2">
    <fs type="anchoring"><f singal="#sg2"/></fs>
  </LINK>
</eSpaceML>
```

This annotation is interpreted as stating that the process type event of driving is anchored to a path which is delineated from the region `rg2`, namely *Boston*, by the signal `sg2` *through*. Unlike *ISO-Space*, which employs `<S_FUNCTION>`, the task of delineating such a path is relegated to formal semantics.

Formal Semantics - Granularity is an issue for any system. Being a net-based distributed scheme, *eSpaceML* attempts to provide the simplest scheme for spatial annotation, while distributing some of the related areas of work to other schemes. One such area is the task of dealing with ordinary inferences. Spatial expressions such as *through Boston*, *along the river*, and *across the desert* or *(We) crossed the desert* do not involve the entire regions referred to, but certain paths which involve the regions in some manner, sometimes tangentially, without the beginning and end points explicitly mentioned. As in *ISO-Space*, these paths could be demarcated within a spatial annotation scheme by introducing a function like `<S_FUNCTION>`. However, *eSpaceML* relies on the task of the simplest type of formal semantics.

Consider Example (7) and its annotation in *ISO-Space*:

(7) *The park is behind the store.*

With its function `<S_FUNCTION>`, *ISO-Space* cuts out a particular region `r6` and then it is located behind the store (`r5`) by `<QSLINK qsid="qs1">`. *eSpaceML*, however, prefers a simpler annotation as exemplified through Example (7).

```
<eSpaceML>
  <REGION xml:id="rg1" corresp="#tk1 #tk2">
    <fs type="fac"/></REGION>
  <SIGNAL xml:id="sg1" corresp="#tk4"/>
    <fs type="topological"><f sense="located_behind"/></fs></SIGNAL>
  <REGION xml:id="rg2" targets="#tk5 #tk6">
    <fs type="fac"/></REGION>
   <LINK xml:id="ln1" targets="#rg1 #rg2">
    <fs type="topological"><f singal="#sg1"/></fs></LINK>
</eSpaceML>
```

`<LINK xml:id="ln1">` here states that `#rg1` is related to `#rg2` by a signal `#sg1`, where they respectively refer to the park, the store, and the signal *behind*. Then we assume there is semantics that derives the semantic representation

$$[park(rg1) \wedge store(rg2)] \wedge [located\_behind(rg1, rg2)]$$

By ordinary inference, we then obtain the following inference:

$$\exists \{l_1, l_2\}[located\_in(rg1, l_1) \wedge located\_in(rg2, l_2) \wedge located\_behind(l_1, l_2)]$$

Similarly, *eSpaceML* annotates processes or transitions involving paths as simple as possible. Consider the `<LINK>` element used in Example (6) as an illustration of processes involving paths.

```
<LINK xml:id="ln2" target="#e2 #rg2">
  <fs type="anchoring">
    <f singal="#s2"/>
  </fs>
</LINK>
```

Here again, we assume that there is a simple semantic system that translates the `<LINK>` element given above into the following semantic representation:

$$[drive(e2) \wedge through\,(e2, rg2)]$$

Now the real question is how to interpret this semantic representation, especially the right conjunct *through*(*e*2, *rg*2). A possible answer depends on the interpretation or meaning of *through*. It cuts out a path or a strip of a region ‘from one side of an area to the other’, yielding an overall interpretation expressed by the following representation:

$$\exists\{r_3, r_4, r_5\}[included(r_3, rg2) \wedge path(r_3) \wedge begin(r_4, rg3) \wedge end(r_5, rg3)]$$

5 Concluding Remarks

eSpaceML can be viewed as *Extended SpatialML* and also as *ISO-Space Distributed*. First, it extends the scope of *SpatialML* beyond the annotation of toponyms and their topological or relative relations by anchoring events to space. Second, it distributes various parts of the task that *ISO-Space* undertakes in two different ways: one is to distribute the task to other available annotation schemes such as *ISO-TimeML* that links events to time and the other is to relegate fine-grained semantic issues to formal semantics that treats inferences as well as logical reasoning.

As a result, *eSpaceML* becomes a compact representation scheme for spatial annotation proper. It is also designed to involve a wider range of interoperability with other annotation schemes in a tightly linked pivotal system. Its interoperability is guaranteed by its conformance to ISO standards, such as *LAF* and *FSR* in particular, for linguistic annotation and representation. Its applications to language technology including moving image interpretation and dialogue understanding are also supported by sustainable language resources thus annotated and managed.

References

Bunt, H. 2007. The Semantics of Semantic Annotations. *Proceedings of the Twenty-first Pacific Asia Conference in Linguistics, Information and Computation*, pp.13-28.

Hois, J. 2010. Inter-annotator Agreement on a Linguistic Ontology for Spatial Language – a Case Study for GUM-Space. *Proceedings of the Seventh International Conference on Language Resources and Evaluation*, pp.3464-3469.

Ide, N., and L. Romary. 2004. International Standard for a Linguistic Annotation Framework. *Journal of Natural Language Engineering*, 10(3-4), pp.211–225.

ISO. 2006. *ISO 24610-1:2006 Language Resource Management – Feature Structures – Part 1: Feature Structure Representation (FSR).*

ISO. 2008. *DIS 24611 Language Resource Management – Morpho-syntactic Annotation Framework (MAF).*

ISO. 2009a. *DIS 24612 Language Resource Management – Linguistic Annotation Framework (LAF).*

ISO. 2009b. *DIS 24615 Language Resource Management – Syntactic Annotation Framework (SynAF).*

ISO. 2009c. *DIS 24617-1 Language Resource Management – Semantic Annotation Framework – Part 1: Time and Events (ISO-TimeML).*

ISO. 2009d. *DIS 24617-2 Language Resource Management – Semantic Annotation Framework – Part 2: Dialogue Acts (SemAF-DActs).*

ISO. 2010. *PWI 24617-6 Language Resource Management – Semantic Annotation Framework – Part 6: Space (ISO-Space).*

Katz, G. 2007. Towards a Denotational Semantics for TimeML. In Schindler et al. (eds.), *Annotating, Extracting and Reasoning about Time and Events*, pp.88-106. Berlin: Springer.

Lee, K. 2008. Formal Semantics for Interpreting Temporal Annotation. In P. van Sterkenburg (ed.), *Unity and Diversity of Languages*, pp.97-108. Amsterdam: John Benjamins.

Lee, K., and L. Romary. 2010. Towards Interoperability of ISO Standards for Language Resource Management. In A.C. Fang, N. Ide and J. Webster (eds.): *Language Resources and Global Interoperability: Proceedings of the Second International Conference on Global Interoperability for Language Resources*, pp.95-103. Hong Kong: City University of Hong Kong.

Pratt-Hartman, I. 2007. From TimeML to Interval Temporal Logic. *Proceedings of the Seventh International Conference on Computational Semantics*, pp.166-114.

Pustejovsky, J. 1991. The Syntax of Event Structure. *Cognition* 41, pp. 47-81.

Pustejovsky, J., and J. Moszkowicz. 2009. *ISO-Space S*pecification Proposal. October 29, 2009 version. Unpublished.

The MITRE Corporation. 2009. *SpatialML: Annotation Scheme for Marking Spatial Expressions in Natural Language.* Version 3.0, Draft.