The Open University

Bridging between Sensor Measurements and Symbolic Ontologies through Conceptual Spaces

- SemSensWeb’09 Workshop, ESWC 2009, Heraklion, June 01, 2009 -

Stefan Dietze, John Domingue,

Knowledge Media Institute, The Open University, UK

Outline

- Background & Motivation
- Conceptual Spaces (CS)
- Utilising CS to map between symbolic Ontologies and Measurements
- Application
- Conclusions

Introduction
Ontology vs Sensor Data

Ontology:

- Formal symbolic specification of shared conceptualisation
- Defined as tuple of ***concepts C***, ***instances I***, ***properties P***, ***relations R*** and ***axioms A***

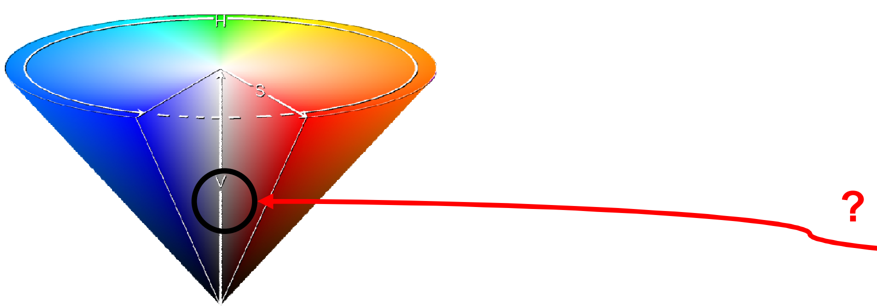
$$O = \{(C, I, P, R, A)\}$$

Sensor data:

- Usually consists of binary data representing **measurements**…
- …describing **observations of real-world phenomena**
- **Concurrent standards** to represent sensor models and measurements (e.g. OpenGIS SensorML,O&M Encoding Standard)

Ontology vs Sensor Data Issues

- **Symbol grounding issue** – ontological entities lack grounding in real-world / cognitive dimensions

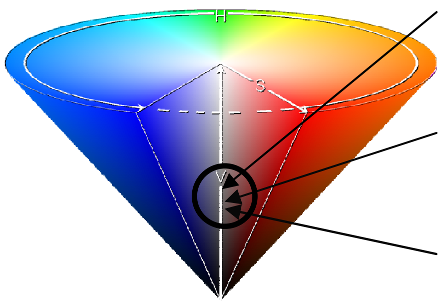



```
...
<owl:Class rdf:ID="Color">
  <rdfs:subClassOf>
   <owl:Class rdf:ID="PhysicalQuality"/>
  </rdfs:subClassOf>
</owl:Class>
<Color rdf:ID="Lilac"/>
...
```

Observed real-world parameter (e.g. color)	Ontological Knowledge (e.g. OWL individual of particular color)

Ontology vs Sensor Data Issues

- **Symbol grounding issue** – ontological entities lack grounding in real-world / cognitive dimensions
- **Multiplicity of mappings** – potentially infinite amount of measurements needs to be mapped to finite set of symbols



01010010100… → {211; 169; 127}

11100010001… → {228; 197, 8}

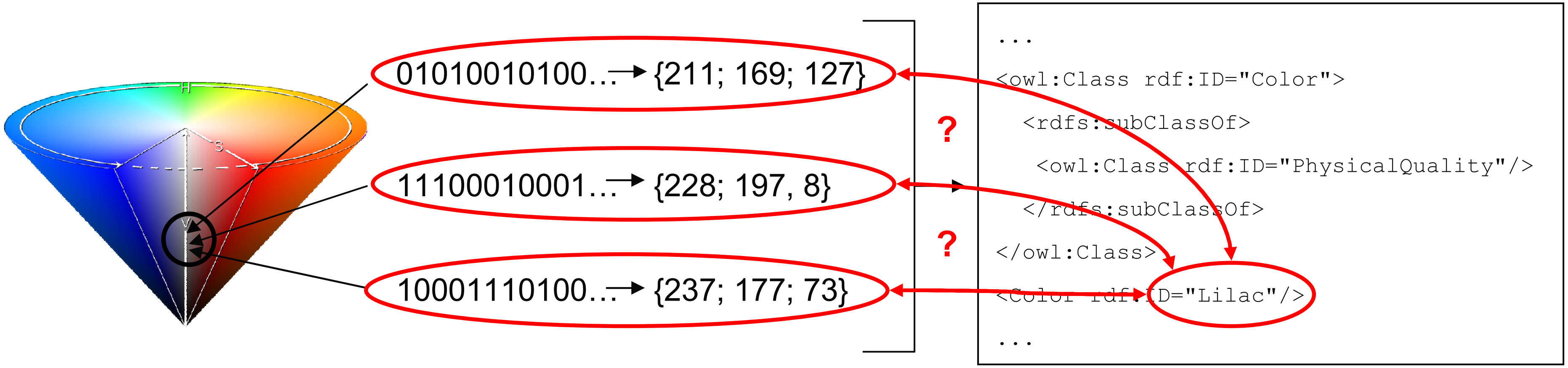
10001110100… → {237; 177; 73}

```
...
<owl:Class rdf:ID="Color">
  <rdfs:subClassOf>
   <owl:Class rdf:ID="PhysicalQuality"/>
  </rdfs:subClassOf>
</owl:Class>
<Color rdf:ID="Lilac"/>
...
```

Observed real-world parameter (e.g. color)	Sensor-data based on measurements (e.g. HSL values)	Ontological Knowledge (e.g. OWL individual of particular color)

Ontology vs Sensor Data Issues

- **Symbol grounding issue** – ontological entities lack grounding in real-world / cognitive dimensions
- **Multiplicity of mappings** – potentially infinite amount of measurements needs to be mapped to finite set of symbols
- **Lack of implicit similarity** – symbolic ontologies lack meaningfulness to implicitly infer on similarities



Observed real-world parameter (e.g. color)	Sensor-data based on measurements (e.g. HSL values)	Ontological Knowledge (e.g. OWL individual of particular color)

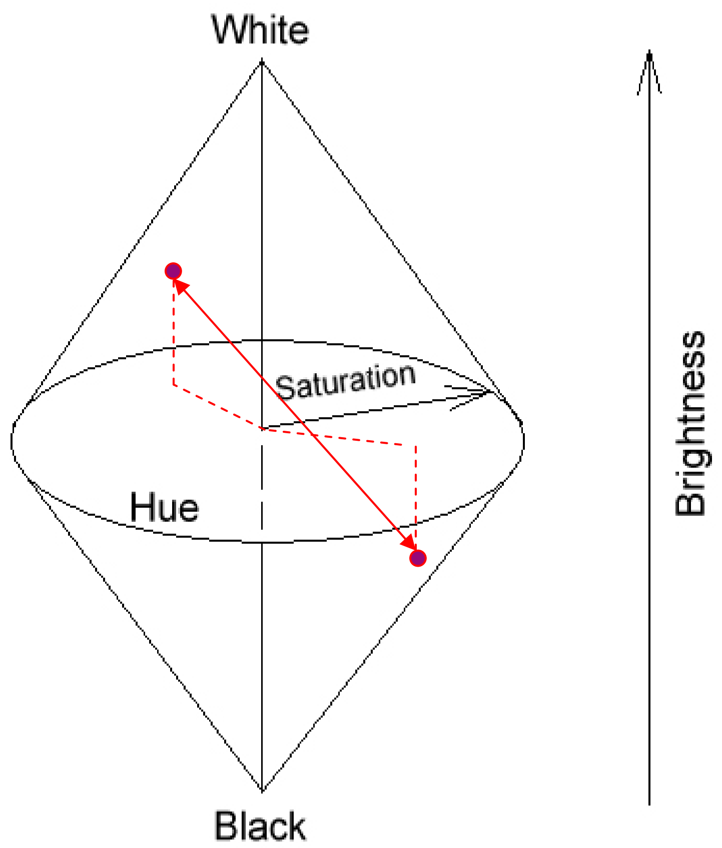
Ontology vs Sensor Data
Issues

- **Symbol grounding issue** – ontological entities lack grounding in real-world / cognitive dimensions
- **Multiplicity of mappings** – potentially infinite amount of measurements needs to be mapped to finite set of symbols
- **Lack of implicit similarity** – symbolic ontologies lack meaningfulness to implicitly infer on similarities

- Representations needed which enable:
  - to bridge between **measurement-based sensor data** and **ontological symbols**
  - to map infinite variety of **real-world observations** to **finite set of symbols**

Spatial Representations
Conceptual Spaces

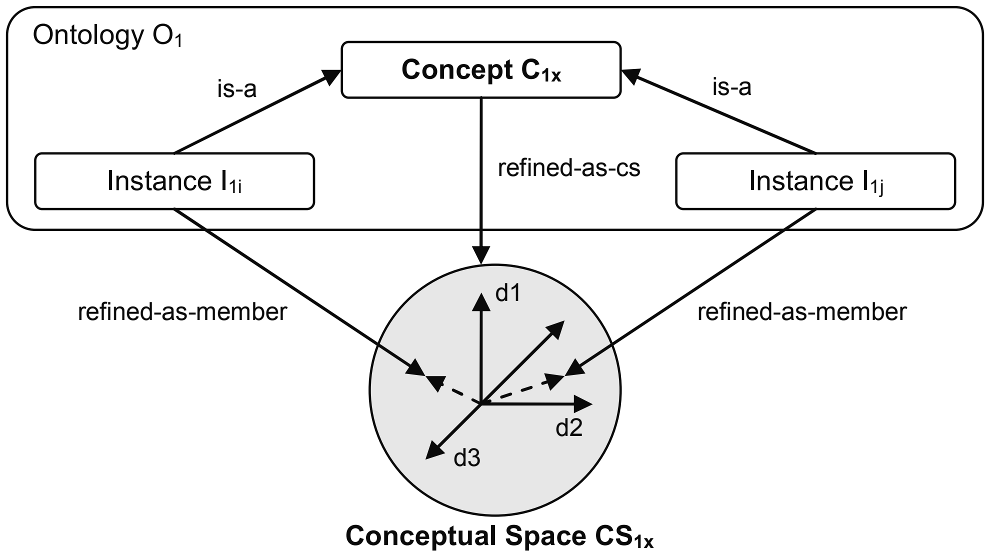
- Exploit **measurements for similarity computation**
- Multidimensional **geometrical vector spaces**
- **Entities** represented in terms of (metric-based) cognitive **quality dimensions**...
  (e.g. colors through dimensions hue, saturation, brightness)
- **Instances** (e.g. 2 colors) => **points (vectors)** in the CS
- **Semantic similarity** between instances => **spatial distance**



Two-fold Approach
Refining Ontologies through CS

CS groundings for ontological concepts (1/2):

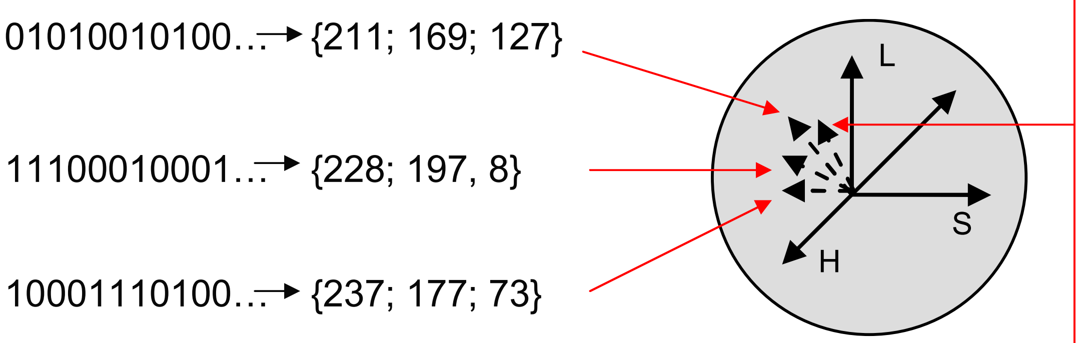
- Refining ontologies through multiple CS
- **Concept *C*** ontology *O* => **Conceptual Space *CS***
- **Instance *I*** of *C* **=> member *M* (vector)** in *CS*...



Two-fold Approach
Refining Ontologies through CS

CS groundings for ontological concepts (2/2):

- **Similarity-computation** between **sensor-based measurements** and **ontological instances**
- Common **agreement at schema (i.e. CS) level**…
- … facilitated through standards for sensor measurement models



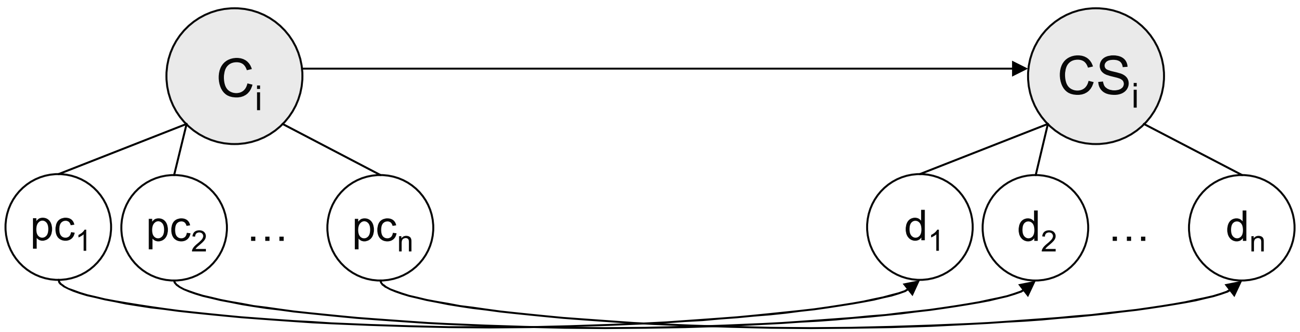

```
...
<owl:Class rdf:ID="Color">
  <rdfs:subClassOf>
   <owl:Class rdf:ID="PhysicalQuality"/>
  </rdfs:subClassOf>
</owl:Class>
<Color rdf:ID="Lilac"/>
...
```

Sensor-data based on measurements (e.g. HSL values)	Similarity-based mapping through Conceptual Color Space	Ontological Knowledge (e.g. OWL individual of particular color)

Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

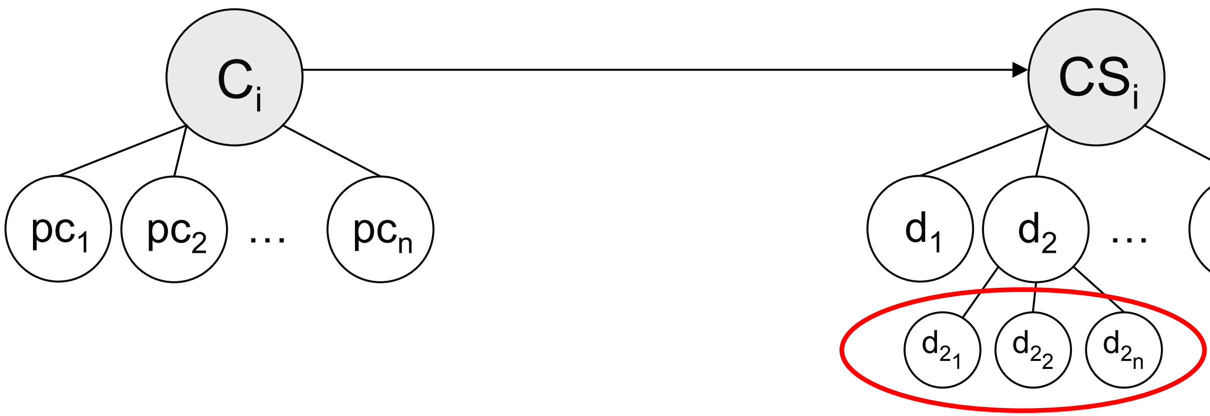
- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scales to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of all instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS



Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

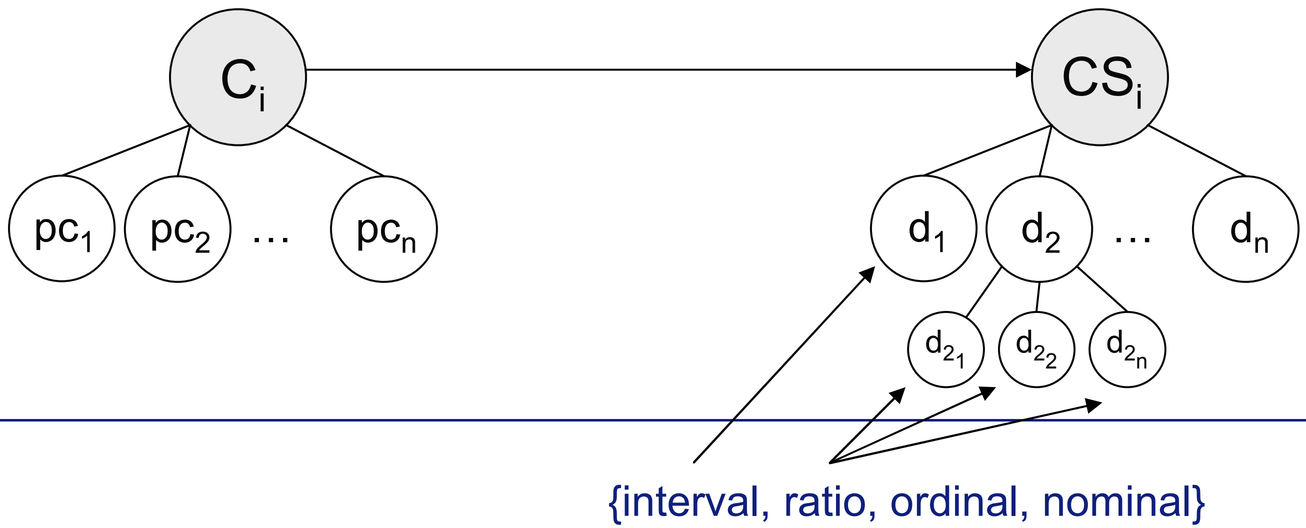
- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scales to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of all instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS



Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scales to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS



Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

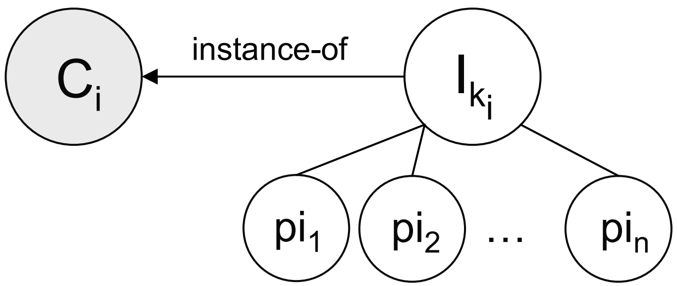
- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scales to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS

$$trans : \{(pc_{i1}, pc_{i2},..., pc_{in}) | pc_{ij} \in PC_i\} \Rightarrow \{(p_{i1}d_{i1}, p_{i2}d_{i2},..., p_{in}d_{in}) | d_{ij} \in CS_i, p_{ij} \in P\}$$

Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scales to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS

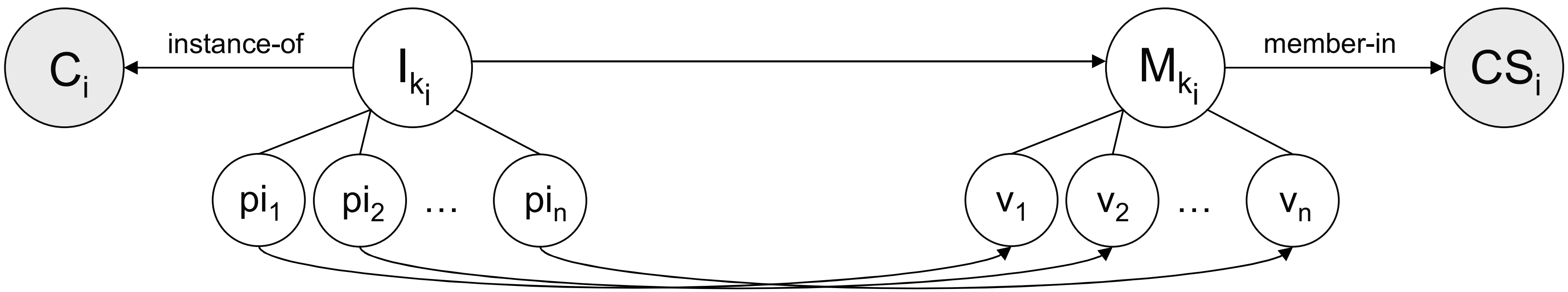


Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scales to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS

$$trans:\{(pi_{k1}, pi_{k2}, \ldots, pi_{kn}) \mid pi_{kl} \in PI_k\} \Rightarrow \{(v_{k1}, v_{k2}, \ldots, v_{kn}) \mid v_{kl} \in M_k\}$$



e.g. M_{k_i}={(20.649, 98, 0, 7.9894)}

Two-fold Approach
CS Formalisation

Formal ontology allowing to refine ontologies through CS:

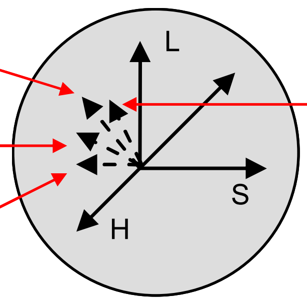
- Representation of concept properties pc_j of C_i as dimensions d_j of CS_i
- Assignment of measurement scale to each quality dimension d_j
- Assignment of prominence values p_j to each quality dimension d_j
- Representation of instances I_{k_i} of C_i as members M_{k_i} in CS_i
- Similarity between sensor measurements and symbolic instance = Euclidean distance in CS

$$dist(u,v) = \sqrt{\sum_{i=1}^{n} p_i \left(\left(\frac{u_i - \bar{u}}{s_u} \right) - \left(\frac{v_i - \bar{v}}{s_v} \right) \right)^2}$$

01010010100… → {211; 169; 127}

11100010001… → {228; 197, 8}

10001110100… → {237; 177; 73}

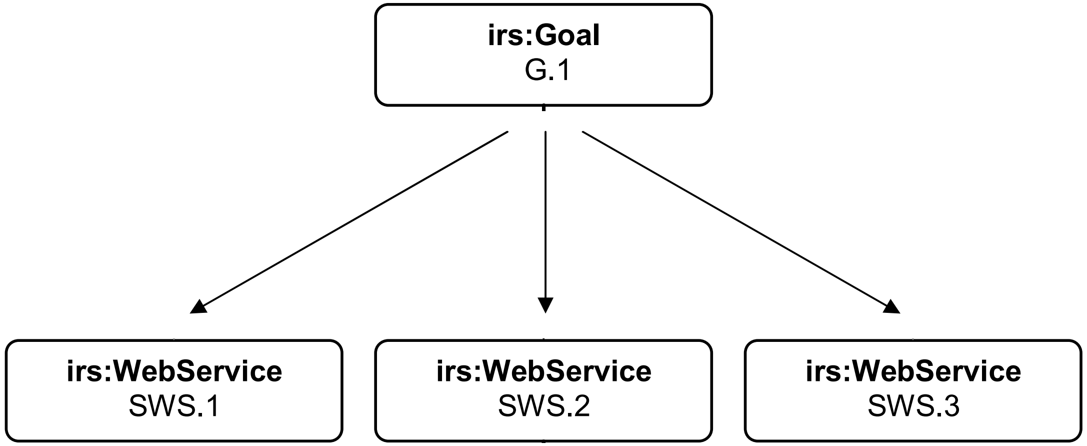



```
...
<owl:Class rdf:ID="Color">
  <rdfs:subClassOf>
   <owl:Class rdf:ID="PhysicalQuality"/>
  </rdfs:subClassOf>
</owl:Class>
<Color rdf:ID="Lilac"/>
...
```

Sensor-data based on measurements (e.g. HSL values)	Similarity-based mapping through Conceptual Color Space	Ontological Knowledge (e.g. OWL individual of particular color)

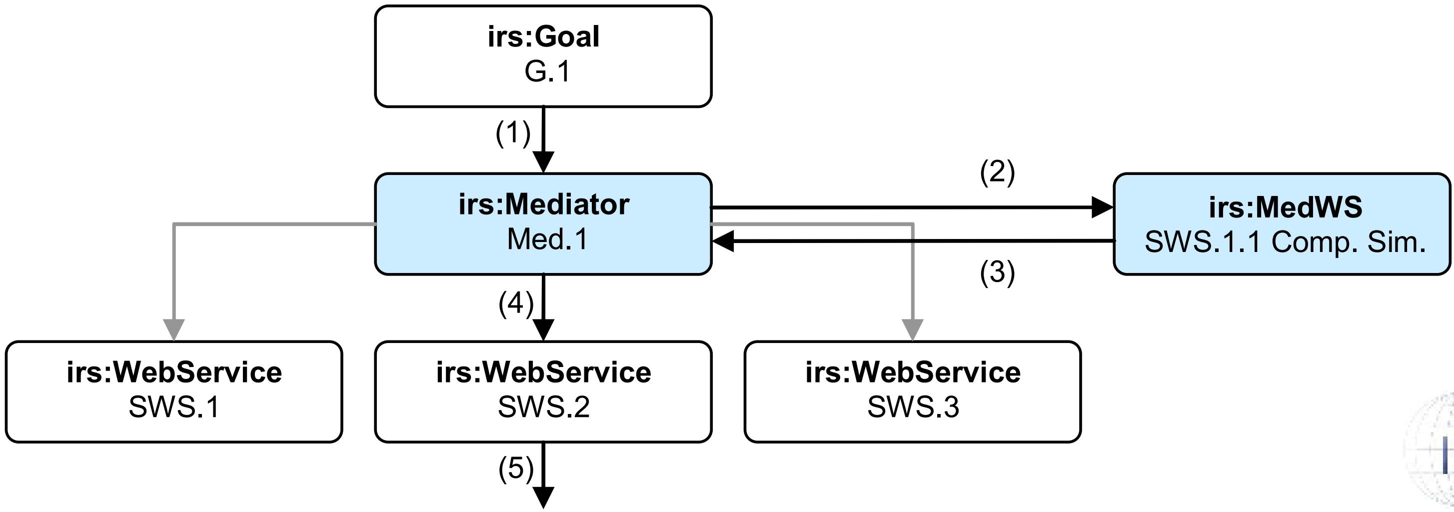
Application
Measurement-based Service Selection

- **Matchmaking of Semantic Web Services (SWS)** based on **context measurements**
- Uses **SWS reasoning environment IRS-III**
- Request: **"irs:Goal"** - **context defined as set of measurements**
- **Matchmaking between request** and ***x*** **associated SWS (*SWS_1..SWS_x*)**
- Implemented through **mediation Web service based on similarity-computation**



Application
Measurement-based Service Selection

- **Matchmaking of Semantic Web Services (SWS)** based on **context measurements**
- Uses **SWS reasoning environment IRS-III**
- Request: **"irs:Goal"** - context defined as set of measurements
- **Matchmaking between request** and ***x*** **associated SWS (*SWS_1..SWS_x*)**
- Implemented through **mediation Web service based on similarity-computation**

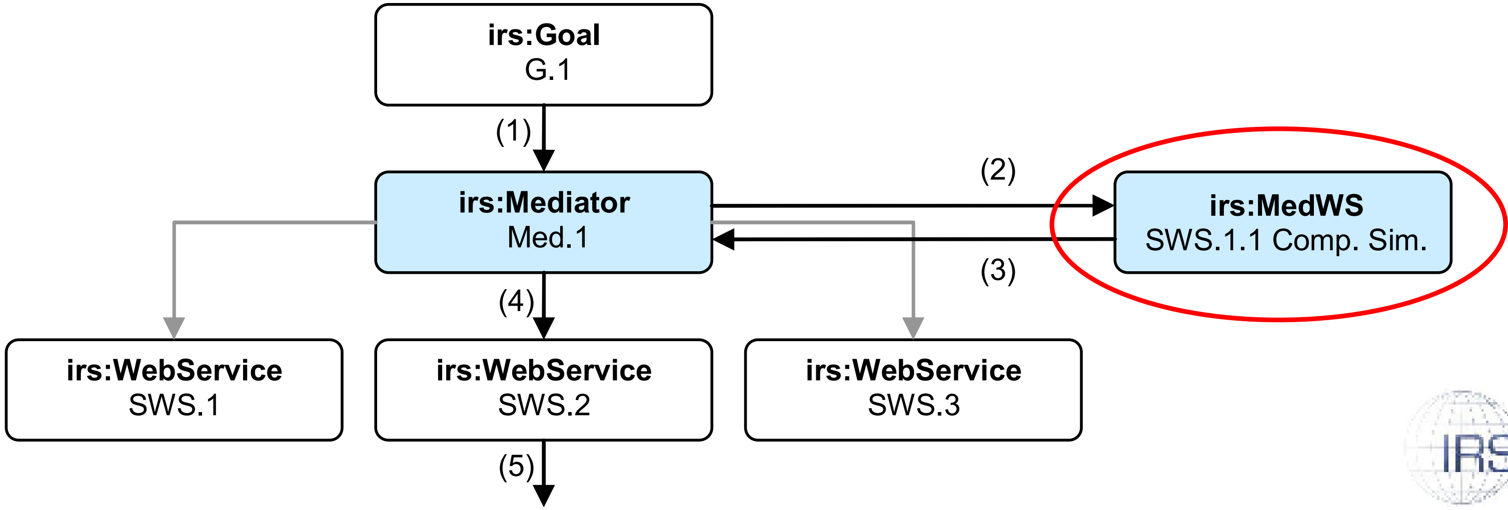


Application
Measurement-based Service Selection

- **MedWS** ***SWS$_{1.1}$*** computes ***x*** **similarity values** with ***Sim(G$_1$,SWS$_j$)*** defined as reciprocal of mean value of individual member distances:

$$Sim(G_i, SWS_j) = \left(\overline{Dist(G_i, SWS_j)}\right)^{-1} = \left(\frac{\sum_{k=1}^{n}(dist_k)}{n}\right)^{-1}$$

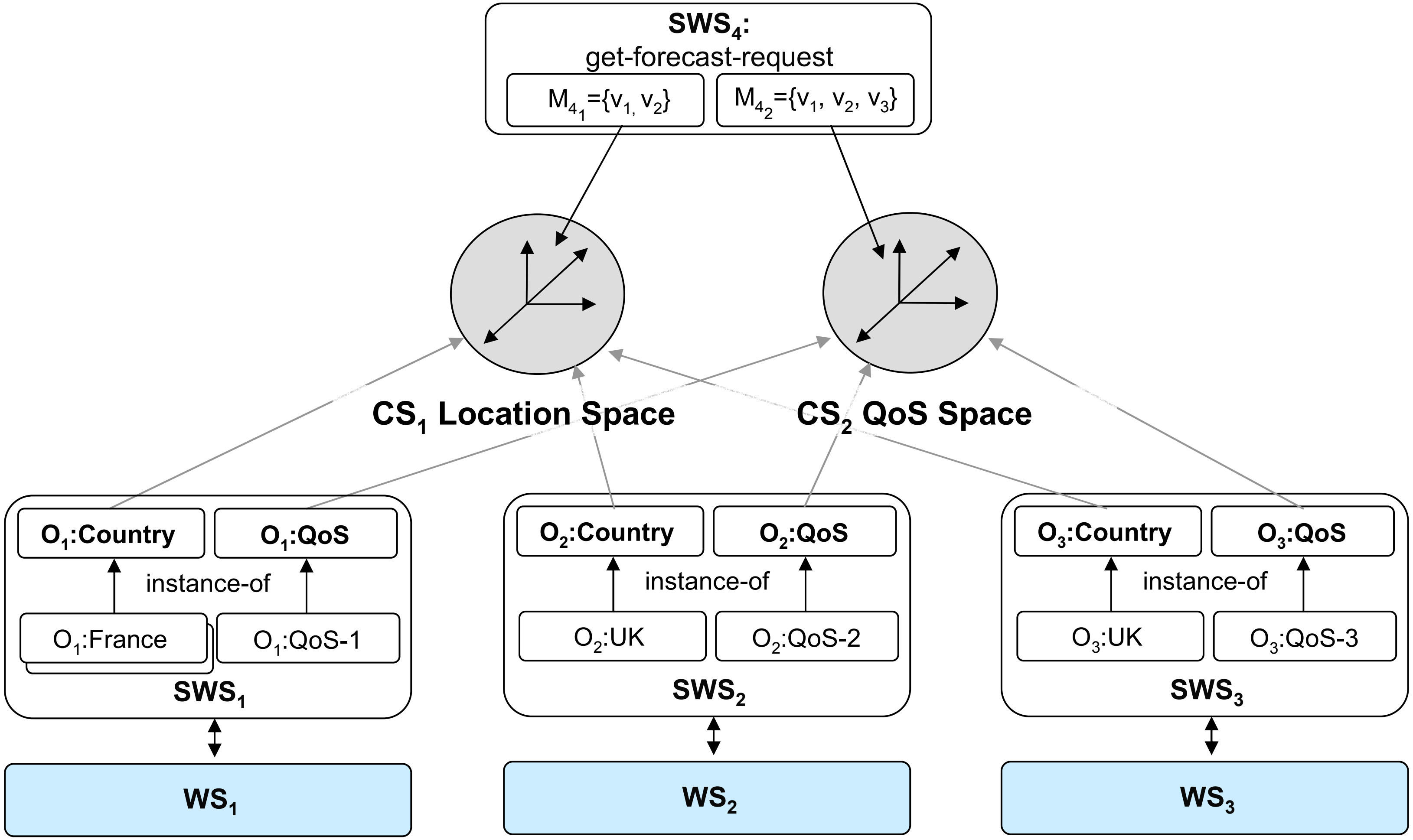
- ***dist$_k$*** = **distance** between **one particular vector (member)** ***v$_i$*** **describing context of** ***G$_1$*** and **one member of** ***SWS$_j$***



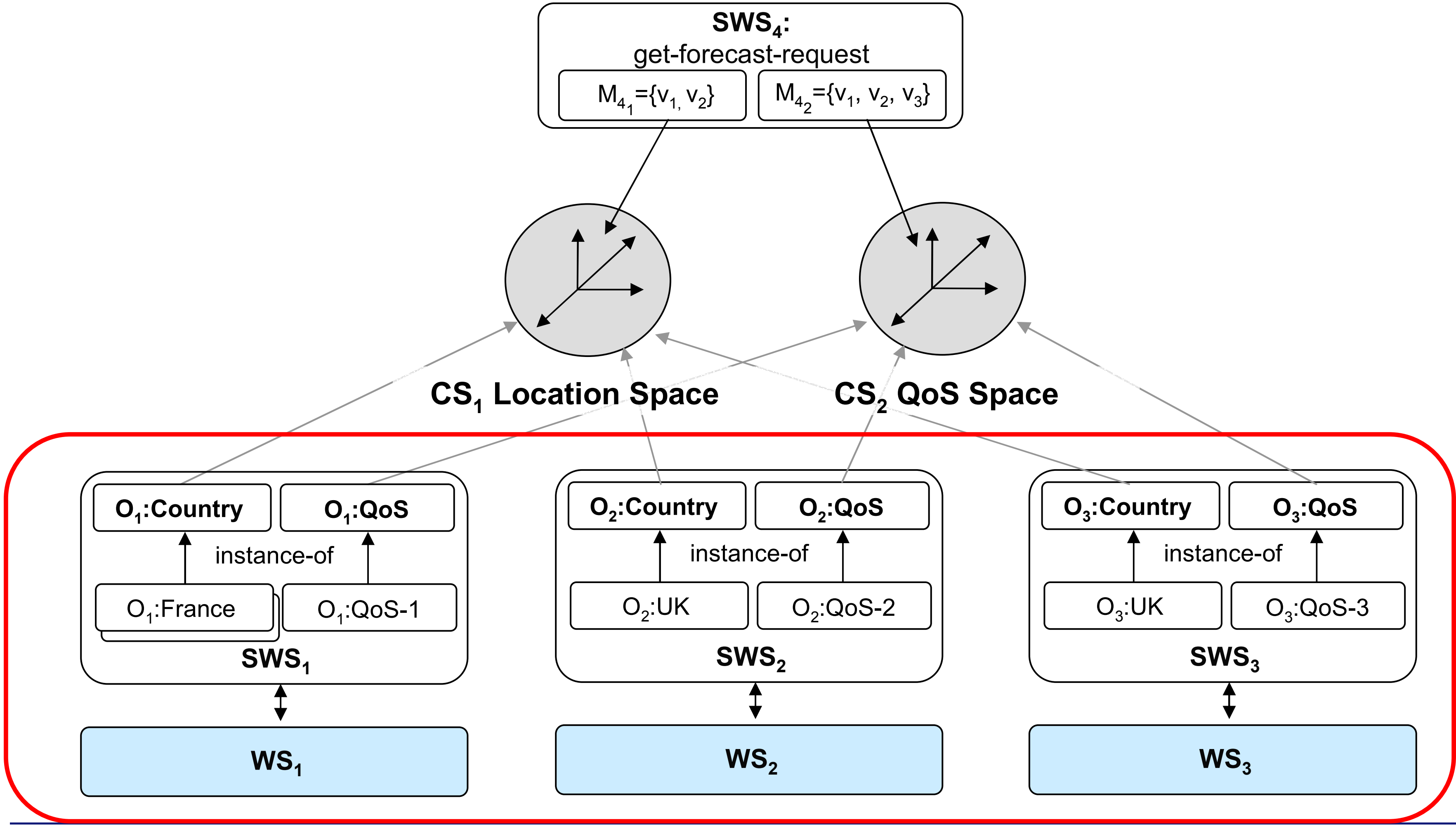
Measurement-based Service Selection
Prototype Application

- **Automated discovery of distributed weather forecast services**
- Each service targets **distinct locations** and **Quality of Service (QoS)** (represented via SWS capability description)
- **Symbolic ontologies (SWS)** extended with **CS-based grounding** (service capability parameter - locations, QoS - represented as members in CS)
- **Requests** (IRS-III goals) **use measurements** to describe **context** (e.g. the current location and desired QoS)
- **Similarity-based service selection** for a given request based on MedWS

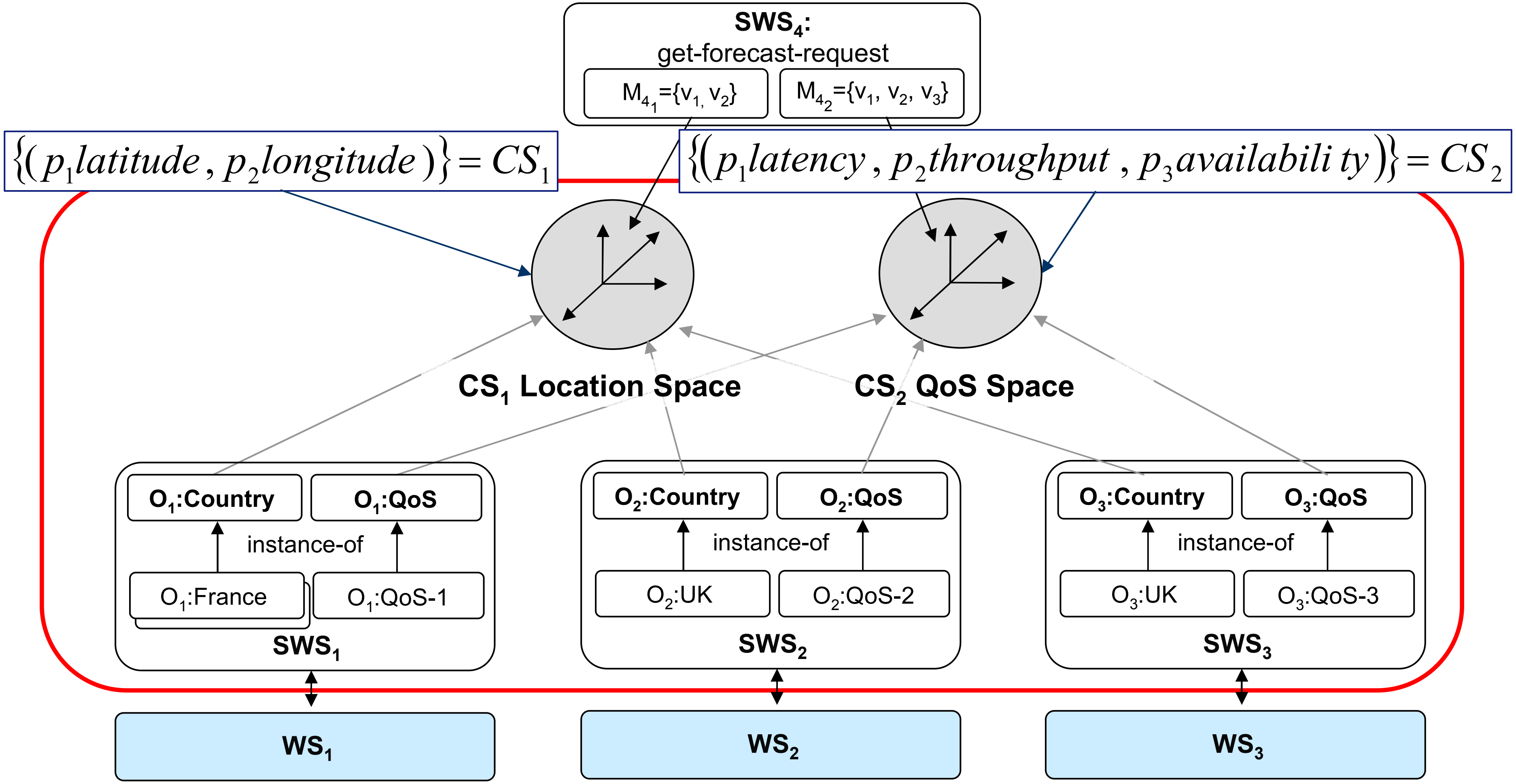
Measurement-based Service Selection
Prototype Application



Measurement-based Service Selection
Prototype Application



Measurement-based Service Selection
Prototype Application



Measurement-based Service Selection
Prototype Application

- **SWS capabilities** described through **conjunction of instances**
- **Instances** refined through **vectors (members)**

	Assumption $Ass_{SWSi} = (L_{1SWSi} \cup L_{2SWSi} \cup .. \cup L_{nSWSi}) \cup (Q_{1SWSi} \cup Q_{2SWSi} \cup .. \cup Q_{mSWSi})$	
	Members L_i in CS_1 (locations)	Members C_j in CS_2 (QoS)
SWS_1	$L_{1(SWS1)}$={(46.227644, 2.213755)} $L_{2(SWS1)}$={(40.463667, -3.74922)}	$Q_{1(SWS1)}$={(155, 2, 91)}
SWS_2	$L_{1(SWS2)}$={(55.378051, -3.435973)}	$Q_{1(SWS2)}$={(15, 50, 98)}
SWS_3	$L_{1(SWS3)}$={(55.378048, -3.435963)}	$Q_{1(SWS3)}$={(78, 5, 95)}
SWS_4	$L_{1(SWS4}$={(55.378048, -3.435963)}	$Q_{1(SWS4)}$={(0,100,100)}

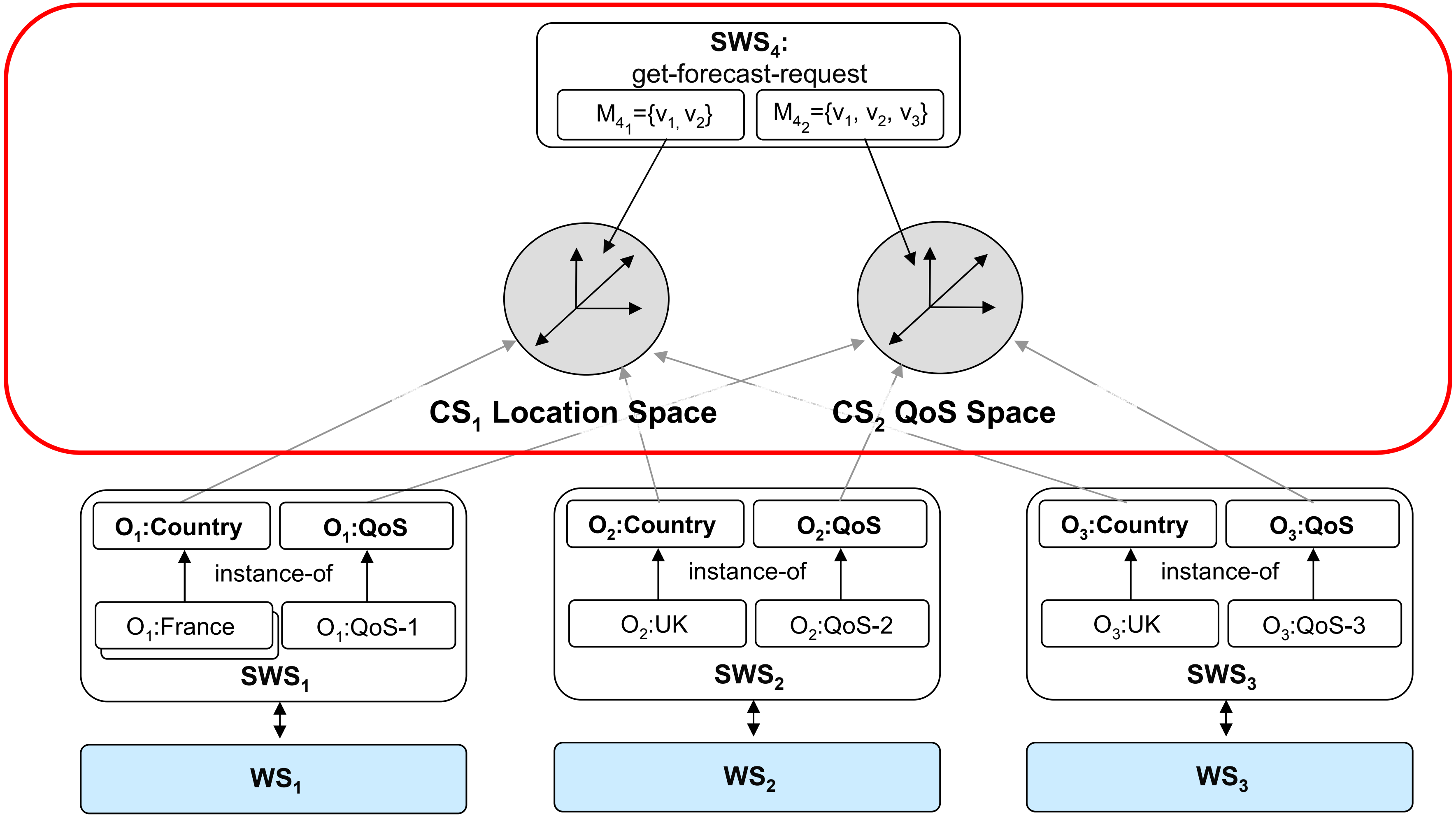
Measurement-based Service Selection
Prototype Application

Locations: France, Spain

Latency = 155 ms
Throughput = 2
Availability = 91%

	Assumption $Ass_{SWSi} = (L_{1SWSi} \cup L_{2SWSi} \cup .. \cup L_{nSWSi}) \cup (Q_{1SWSi} \cup Q_{2SWSi} \cup .. \cup Q_{mSWSi})$	
	Members L_i in CS_1 (locations)	Members C_j in CS_2 (QoS)
SWS_1	$L_{1(SWS1)}$={(46.227644, 2.213755)} $L_{2(SWS1)}$={(40.463667, -3.74922)}	$Q_{1(SWS1)}$={(155, 2, 91)}
SWS_2	$L_{1(SWS2)}$={(55.378051, -3.435973)}	$Q_{1(SWS2)}$={(15, 50, 98)}
SWS_3	$L_{1(SWS3)}$={(55.378048, -3.435963)}	$Q_{1(SWS3)}$={(78, 5, 95)}
SWS_4	$L_{1(SWS4}$={(55.378048, -3.435963)}	$Q_{1(SWS4)}$={(0,100,100)}

Measurement-based Service Selection
Prototype Application



Measurement-based Service Selection
Prototype Application

	Assumption $Ass_{SWSi} = (L_{1SWSi} \cup L_{2SWSi} \cup .. \cup L_{nSWSi}) \cup (Q_{1SWSi} \cup Q_{2SWSi} \cup .. \cup Q_{mSWSi})$	
	Members L_i in CS_1 (locations)	Members C_j in CS_2 (QoS)
SWS_1	$L_{1(SWS1)}$={(46.227644, 2.213755)} $L_{2(SWS1)}$={(40.463667, -3.74922)}	$Q_{1(SWS1)}$={(155, 2, 91)}
SWS_2	$L_{1(SWS2)}$={(55.378051, -3.435973)}	$Q_{1(SWS2)}$={(15, 50, 98)}
SWS_3	$L_{1(SWS3)}$={(55.378048, -3.435963)}	$Q_{1(SWS3)}$={(78, 5, 95)}
SWS_4	$L_{1(SWS4}$={(55.378048, -3.435963)}	$Q_{1(SWS4)}$={(0,100,100)}

Location: UK

Latency = 0 ms
Throughput = 100
Availability = 100%

Measurement-based Service Selection
Prototype Application

	Similarities
SWS_1	0.010290349
SWS_2	0.038284954
SWS_3	0.016257476

	… $\cup\ Q_{mSWSi}$)	
	Members L_i in CS_1 (locations)	Members C_j in CS_2 (QoS)
SWS_1	$L_{1(SWS1)}$={(46.227644, 2.213755)} $L_{2(SWS1)}$={(40.463667, -3.74922)}	$Q_{1(SWS1)}$={(155, 2, 91)}
SWS_2	$L_{1(SWS2)}$={(55.378051, -3.435973)}	$Q_{1(SWS2)}$={(15, 50, 98)}
SWS_3	$L_{1(SWS3)}$={(55.378048, -3.435963)}	$Q_{1(SWS3)}$={(78, 5, 95)}
SWS_4	$L_{1(SWS4}$={(55.378048, -3.435963)}	$Q_{1(SWS4)}$={(0,100,100)}

Location: UK

Latency = 0 ms
Throughput = 100
Availability = 100%

	Similarities
SWS_1	0.010290349
SWS_2	0.038284954
SWS_3	0.016257476

$\cup Q_{mSWSi})$

	Members L_i in CS_1 (locations)	Members C_j in CS_2 (QoS)
SWS_1	$L_{1(SWS1)}=\{(46.227644, 2.213755)\}$ $L_{2(SWS1)}=\{(40.463667, -3.74922)\}$	$Q_{1(SWS1)}=\{(155, 2, 91)\}$
SWS_2	$L_{1(SWS2)}=\{(55.378051, -3.435973)\}$	$Q_{1(SWS2)}=\{(15, 50, 98)\}$
SWS_3	$L_{1(SWS3)}=\{(55.378048, -3.435963)\}$	$Q_{1(SWS3)}=\{(78, 5, 95)\}$
SWS_4	$L_{1(SWS4}=\{(55.378048, -3.435963)\}$	$Q_{1(SWS4)}=\{(0,100,100)\}$

Location: UK

Latency = 15 ms
Throughput = 50
Availability = 98%

Conclusions
Discussion and Summary

Some issues:

- Additional **representational effort**
- => CS might just **shift symbol grounding issue** (i.e. dimensions lack grounding and are ambiguous)
- **CS dimensions** need **to represent** actual **sensor measurements**
- Ontologies/sensor data need to **share common schema (CS)**

Conclusions
Discussion and Summary

Some issues:

- Additional **representational effort**
- => CS might just **shift symbol grounding issue** (i.e. dimensions lack grounding and are ambiguous)
- **CS dimensions** need **to represent** actual **sensor measurements**
- Ontologies/sensor data need to **share common schema (CS)**

…, however:

- **Similarity computation** between **symbolic instances** and **sensor measurements**
- Provides means to **map infinite variety of potential sensor measurements to finite set of symbolic instances**
- **Alignment of distinct sensor models** through alignment of CS

Thank you!

E-mail: s.dietze@open.ac.uk

Web: http://people.kmi.open.ac.uk/dietze