Orientation, Configuration, & Dimensionality Markup Language

Cynthia Goodman, Cory Massaro, Jasper Phillips, and Zachary Yocum

The Phenomenon

How do you capture spatial relations expressed in natural language?

Some previous approaches:

- **SpatialML**
- **ISO-Space**
- **WordsEye**

Spatial ML

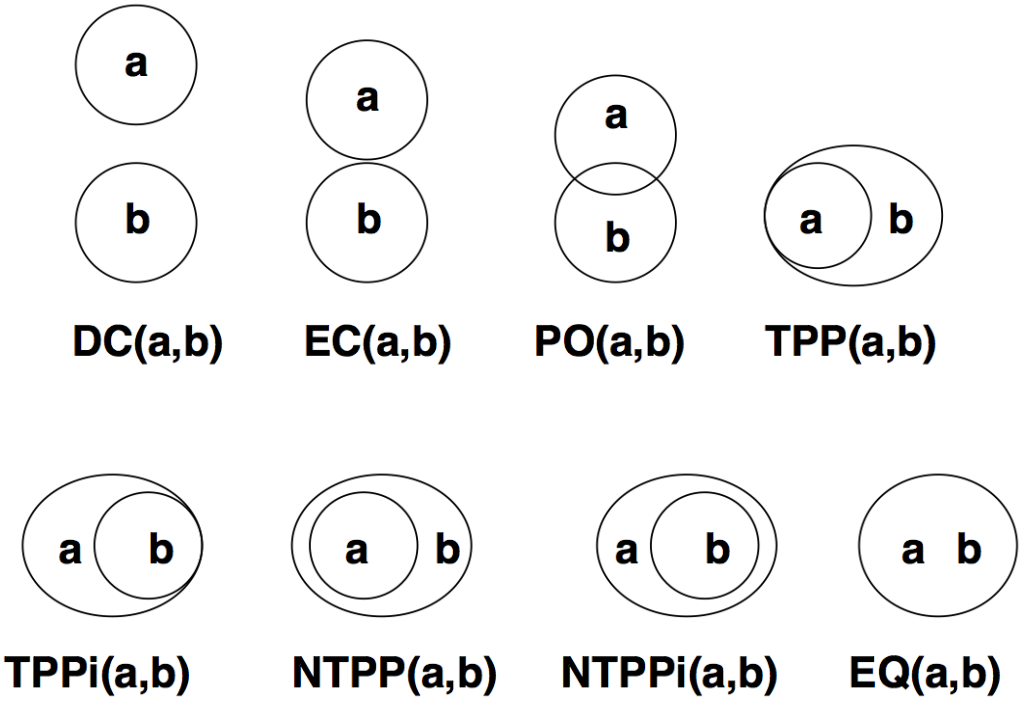
- Inderjeet Mani, Janet Hitzeman, Justin Richer, Dave Harris, Rob Quimby, Ben Wellner
- A markup scheme for representing places and their relationships.
  - Intended for geo-coding natural language
  - Entities and locations are coded with geo-coordinates where possible
  - Orientation information is represented with MOD and Direction attributes

ISO-Space

- James Pustejovsky, Jessica Moszkowicz
- A markup scheme for representing spatial and spatio-temporal information
  - Intended for coding spatial objects, spatial relations, and motion
  - Orientation is represented with OLINK relations
  - Topological relations are represented with QSLINK relations

Qualitative Spatial Reasoning

- Region Connection Calculus
  - RCC8:



WordsEye

- Bob Coyne, Richard Sproat
- WordsEye is an 'Artistic', 'Language-Based' 3-D Scene Generator
  - Claims to allow users to create 3-D rendered images from natural-language

WordsEye

- Methodology for noun and verbs
  - POS tagged
  - Statistical parser
  - Converted to a dependency structure
  - Converted to a semantic representation
    - Entity, Action, Attribute, Relation, Path, etc.

- Not explicit about how WordsEye interprets spacial relations.

Uses depiction rules to convert semantic representation into graphical output.

DEFINE-DEPICTION-RULE ACTION *kick*

Case1: no PATH or RECIPIENT, DIRECT-OBJECT size greater than 3 feet Pose: *kick*, ACTOR

Position: ACTOR directly behind DIRECT-OBJECT Orientation: ACTOR facing DIRECT-OBJECT

Case2: no PATH or RECIPIENT, DIRECT-OBJECT size less than 3 feet

Pose: *kick object*, ACTOR, DIRECT-OBJECT

Case3: PATH and RECIPIENT

Pose: *kick*, ACTOR

Path: DIRECT-OBJECT between ACTOR's *foot* and RECIPIENT

Orientation: ACTOR facing RECIPIENT

Pose: *catch*, RECIPIENT [tentative]

Orientation: RECIPIENT facing ACTOR [tentative]

Position: ACTOR 10 feet from RECIPIENT in Z axis [tentative]

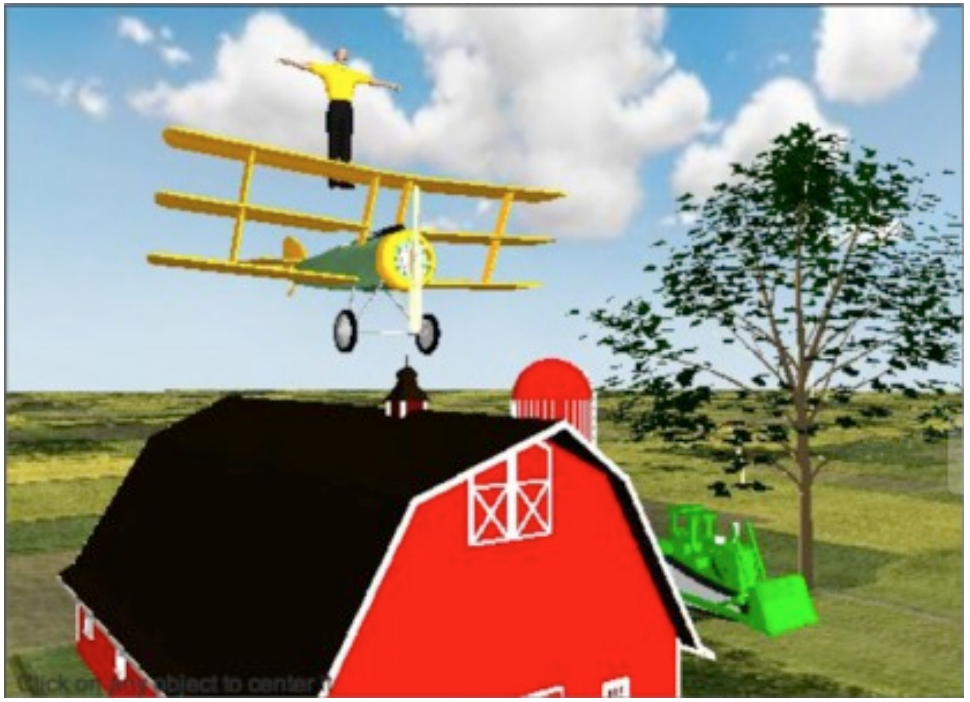
Position: ACTOR 0 feet from RECIPIENT in X axis [tentative]

WordsEye

“John uses the crossbow. He rides the horse by the store. The store is under the large willow. The small allosaurus is in front of the horse. The dinosaur faces John. A gigantic teacup is in front of the store. The dinosaur is in front of the horse. The gigantic mushroom is in the teacup. The castle is to the right of the store. ”

WordsEye

"The big plane is above the barn. The tall man is on the front of the plane."



Comparison of Tag Types

SpatialML	ISO-Space	OCDML	
LINK	**OLINK**	**CONFIGURATION_LINK**	
direction	**relType**	**figure**	**ground**
		¬ any	any
		*	side✢
A	ABOVE	*	top
BL	BELOW	*	bottom
F	FRONT	*	front
B	BACK	*	back
	LEFT	*	left
	RIGHT	*	right
N	NORTH		
E	EAST		
S	SOUTH		
W	WEST		
...	...		

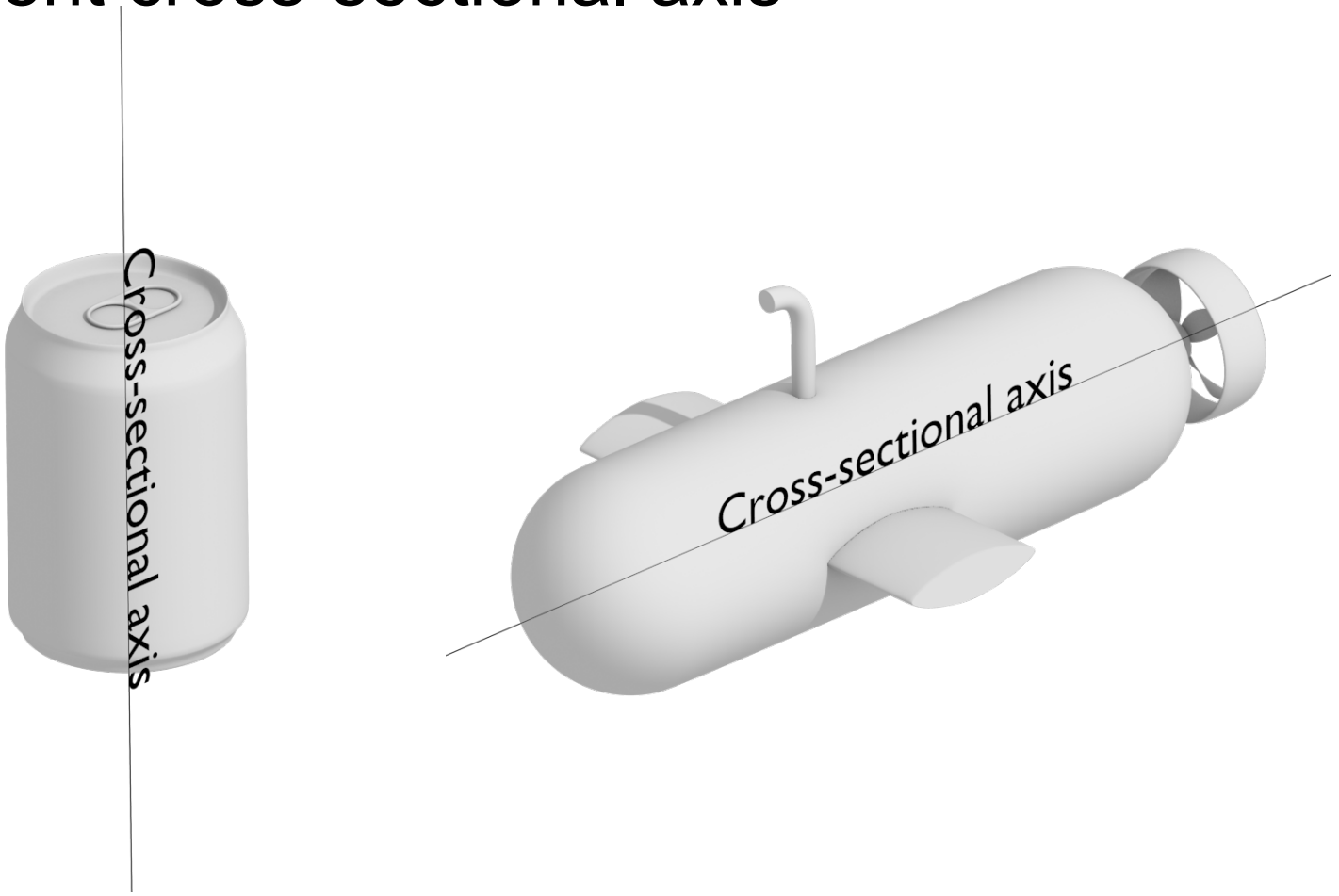
✢ side ≡ (left ∨ right ∨ front ∨ back) ≡ ¬ top ∧ ¬ bottom

Annotation Specification

Attribute	Value
id	se1, se2, se3, ...
dimensionality	POINT, LINE, AREA, VOLUME
mod	A spatially relevant modifier
line_type	SEGMENT, RAY, LINE, LOOP, OTHER
area_type	3-GON, 4-GON, DISC, ANNULUS, OTHER
volume_type	TRI_PRISM, RECT_PRISM, PYRAMID, SPHERE, TORUS, CYLINDER, CONE, BIPED, QUADRUPED, OTHER
left_right	INTRINSIC or RELATIVE
front_back	INTRINSIC or RELATIVE
top_bottom	INTRINSIC or RELATIVE
c-sec_axis	LEFT_RIGHT or FRONT_BACK or TOP_BOTTOM or OTHER

Annotation Specification

- Salient cross-sectional axis



Annotation Specification

Attribute	Value
id	os1, os2, os3, ...
orientation_type	LATITUDINAL, LONGITUDINAL, LATERAL, VERTICAL, OTHER

Annotation Specification

Attribute	Value
id	cl1, cl2, cl3, ...
figure_config	LEFT, RIGHT, FRONT, BACK, TOP, BOTTOM, SIDE, TOP_OR_BOTTOM, ANY, OTHER
ground_config	LEFT, RIGHT, FRONT, BACK, TOP, BOTTOM, SIDE, TOP_OR_BOTTOM, ANY, OTHER
coercion	FIGURE, GROUND, or NONE
trigger	An ID of an ORIENTATION_SIGNAL tag that triggers the link
figure	The ID of the SPATIAL_ENTITY acting as the figure within the relation
ground	The ID of the SPATIAL_ENTITY acting as the ground within the relation

SceneBank Corpus

- Focused Domain
  - Scene descriptions from dramatic literary works
    - George Bernard Shaw
    - John Galsworthy

Example Annotation

"A hotel sitting [**room**$_{se1}$]. A [**table**$_{se2}$] [***in***$_{os1}$] the [**centre**$_{se3}$]. [***On***$_{os2}$] [**it**$_{se4}$] a [**telephone**$_{se5}$]. Two [**chairs**$_{se6}$] [***at***$_{os3}$] [**it**$_{se7}$], [***opposite***$_{os4}$] [**one**$_{se8}$] [**another**$_{se9}$]. [***Behind***$_{os5}$] [**it**$_{se10}$], the [**door**$_{se11}$]. The [**fireplace**$_{se12}$] has a [**mirror**$_{se13}$] [***in***$_{os6}$] the [**mantelpiece**$_{se14}$]."

Example Annotation (continued)

room$_{se1}$

(dimensionality=volume,
volume_type=rect_prism,
left_right=relative,
font_back=relative,
top_bottom=intrinsic,
c_sec_axis=top_bottom)

table$_{se2}$ = it$_{se4}$ = it$_{se7}$ = it$_{se10}$

(dimensionality=volume,
volume_type=rect_prism,
left_right=relative,
font_back=relative,
top_bottom=intrinsic,
c_sec_axis=top_bottom)

center$_{se3}$

(dimensionality=volume,
volume_type=cylinder,
left_right=relative,
font_back=relative,
top_bottom=relative,
c_sec_axis=top_bottom)

telephone$_{se5}$

(dimensionality=volume,
volume_type=tri_prism,
left_right=intrinsic,
font_back=intrinsic,
top_bottom=intrinsic,
c_sec_axis=top_bottom)

Example Annotation (continued)

chairs$_{se6}$ = one$_{se8}$ = another$_{se9}$

(dimensionality=volume,
volume_type=rect_prism,
left_right=intrinsic,
font_back=intrinsic,
top_bottom=intrinsic,
c_sec_axis=top_bottom)

door$_{se11}$

(dimensionality=volume,
volume_type=rect_prism,
left_right=relative,
font_back=relative,
top_bottom=intrinsic,
c_sec_axis=top_bottom)

fireplace$_{se12}$

(dimensionality=volume,
volume_type=rect_prism,
left_right=intrinsic,
font_back=intrinsic,
top_bottom=intrinsic,
c_sec_axis=front_back)

mirror$_{se13}$

(dimensionality=area,
area_type=disc,
volume_type=cylinder,
left_right=relative,
font_back=intrinsic,
top_bottom=relative,
c_sec_axis=front_back)

Example Annotation (continued)

mantelpiece$_{se14}$

(dimensionality=volume,
volume_type=rect_prism,
left_right=relative,
font_back=relative,
top_bottom=relative,
c_sec_axis=front_back)

Example Annotation (continued)

cl1

(trigger=in$_{os1}$,

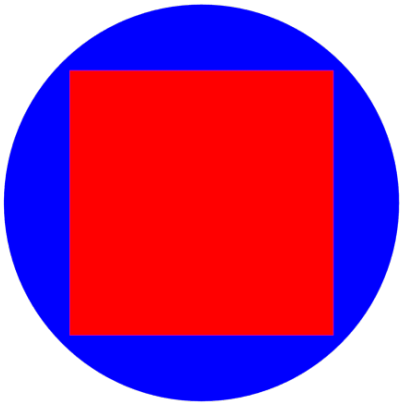
figure=table$_{se2}$,

ground=center$_{se3}$,

figure_config=bottom,

ground_config=bottom)

NTPP(table$_{se2}$,center$_{se3}$)



Example Annotation (continued)

cl2

(trigger=On$_{os2}$,

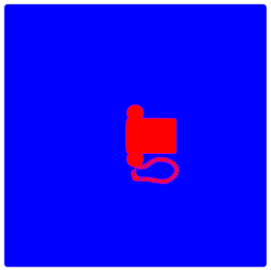
figure=telephone$_{se5}$,

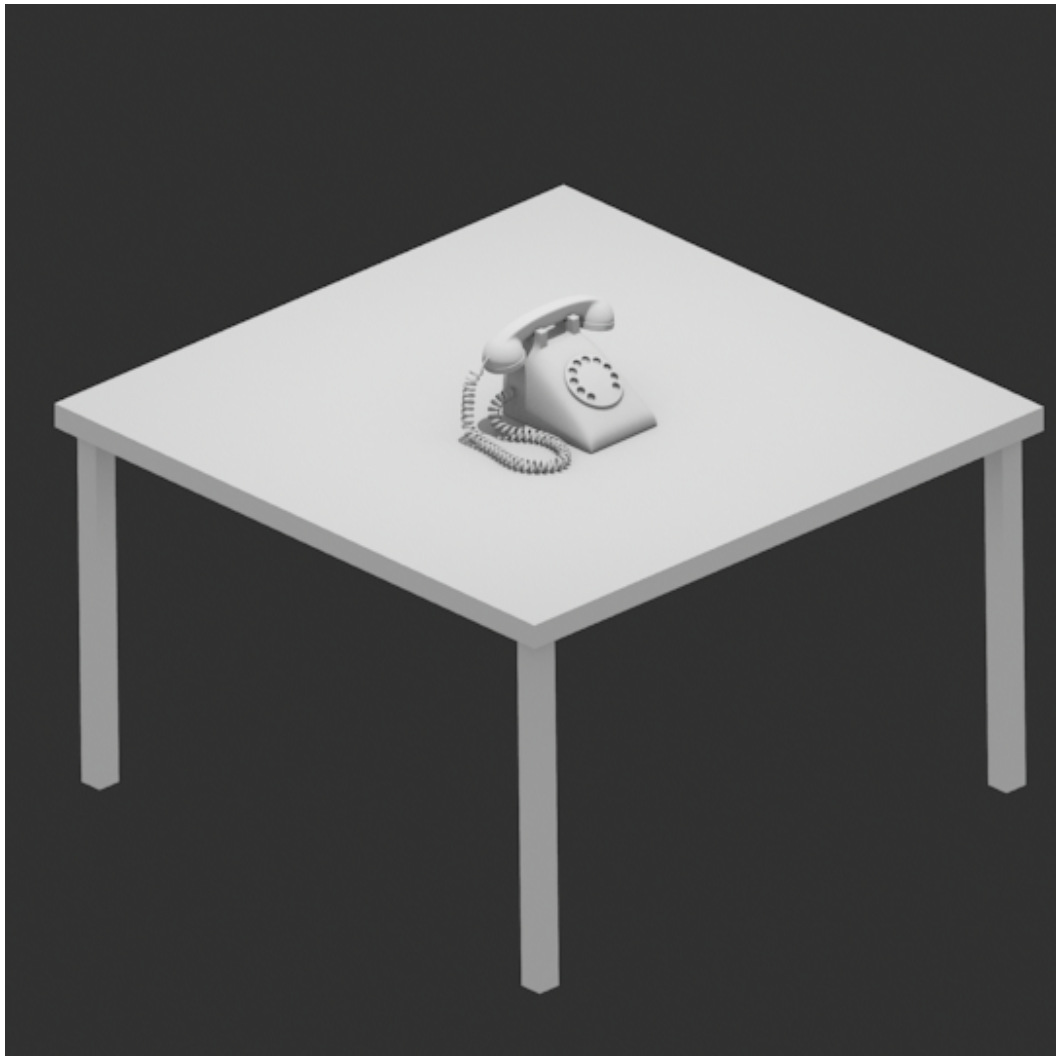
ground=it$_{se4}$,

figure_config=bottom,

ground_config=top)

NTPP(telephone$_{se5}$,it$_{se4}$)





Example Annotation (continued)

cl3

(trigger=at$_{os3}$,

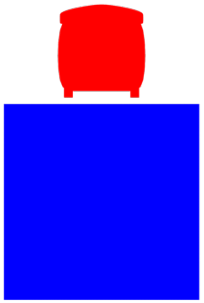
figure=chairs$_{se6}$,

ground=it$_{se7}$,

figure_config=front,

ground_config=side)

DC(chairs$_{se5}$,it$_{se6}$)



Example Annotation (continued)

cl4

(trigger=opposite$_{os4}$,

figure=one$_{se8}$,

ground=another$_{se9}$,

figure_config=front,

ground_config=front)

DC(chairs$_{se5}$,$\varnothing_{se7}$)

Example Annotation (continued)

cl5

(trigger=Behind$_{os5}$,

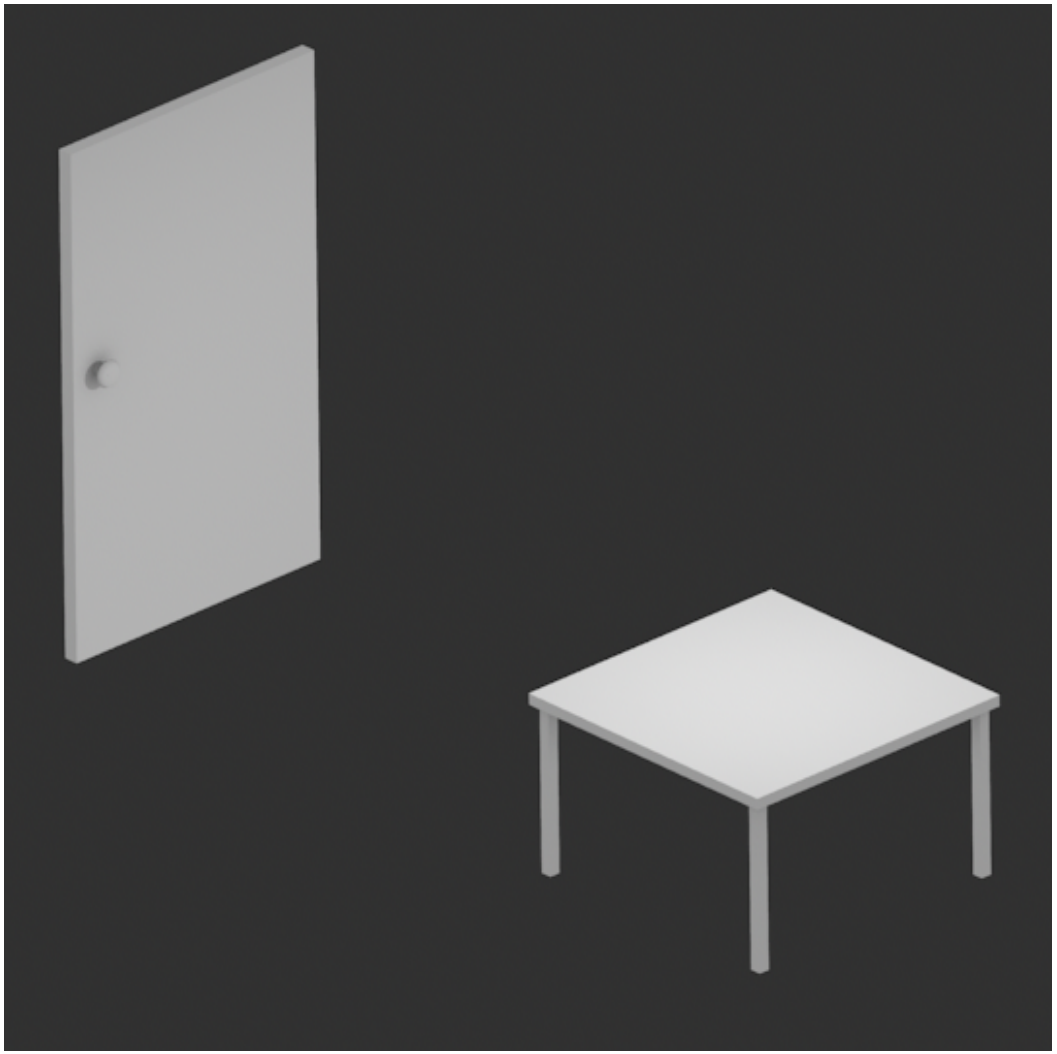
figure=door$_{se11}$,

ground=it$_{se10}$,

figure_config=any,

ground_config=back)

DC(door$_{se6}$,it$_{se8}$)



Example Annotation (continued)

cl6

(trigger=in$_{os6}$,

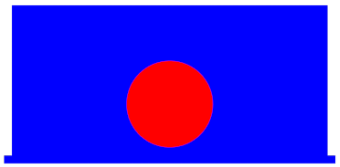
figure=mirror$_{se13}$,

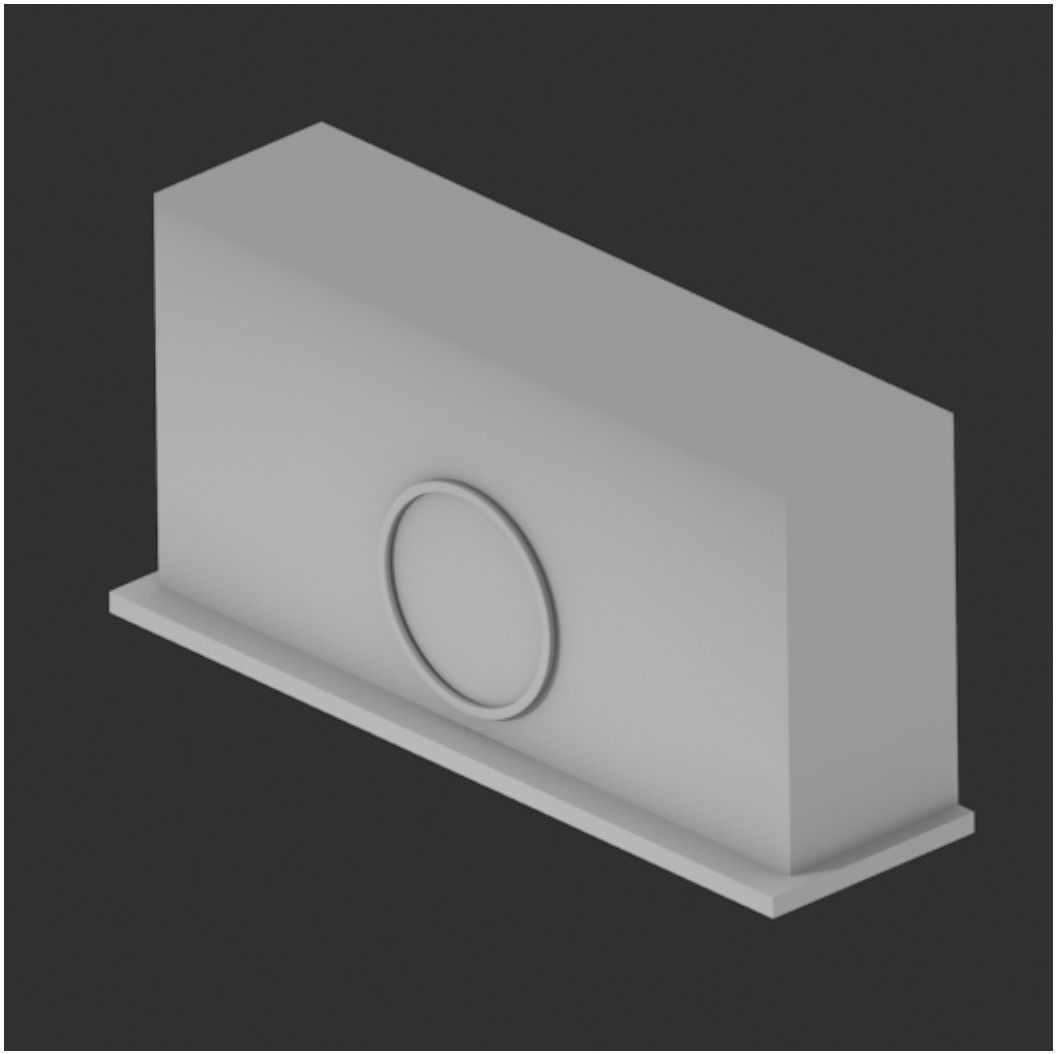
ground=mantelpiece$_{se14}$,

figure_config=back,

ground_config=front)

IN(mirror$_{se11}$,mantelpiece$_{se12}$)



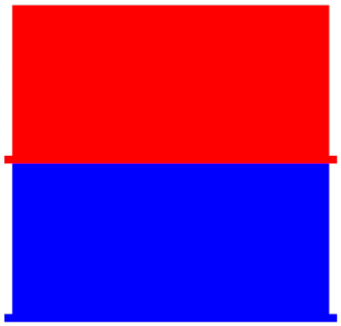


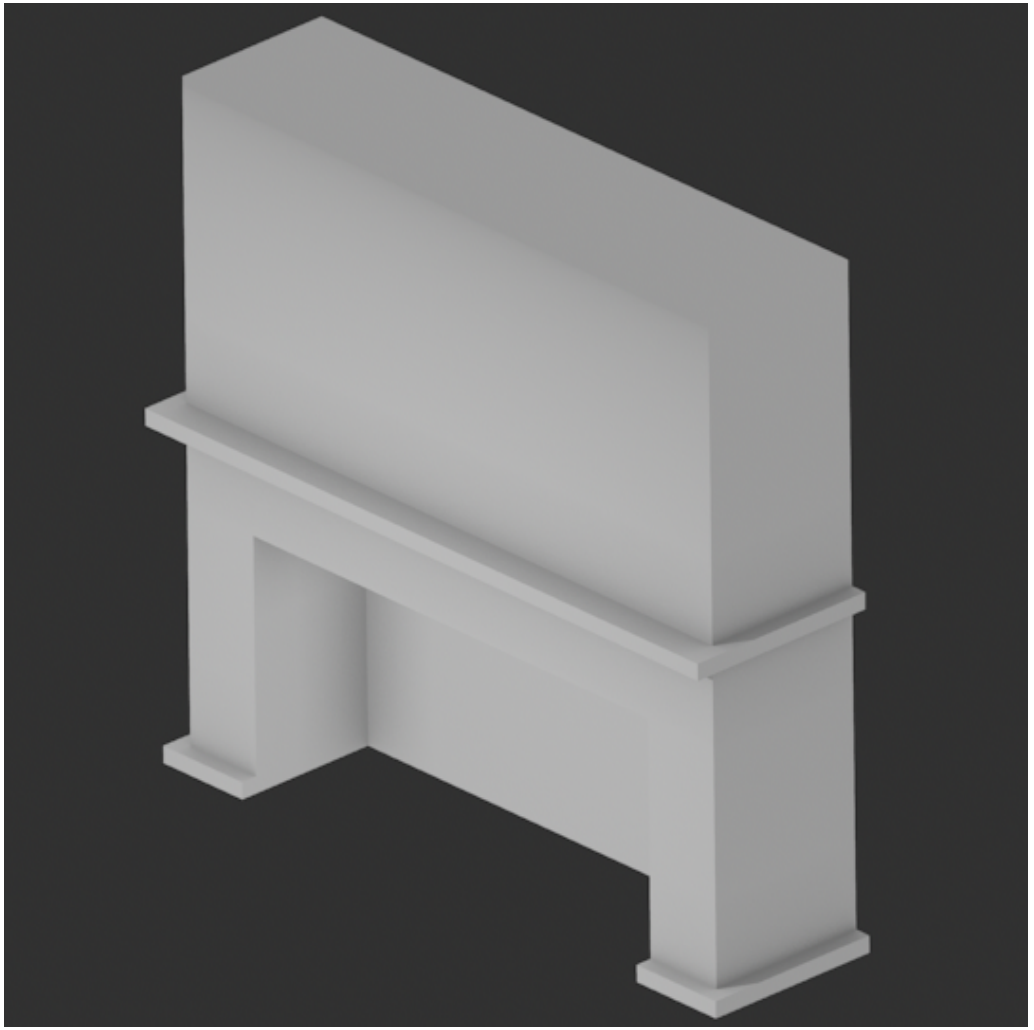
Example Annotation (continued)

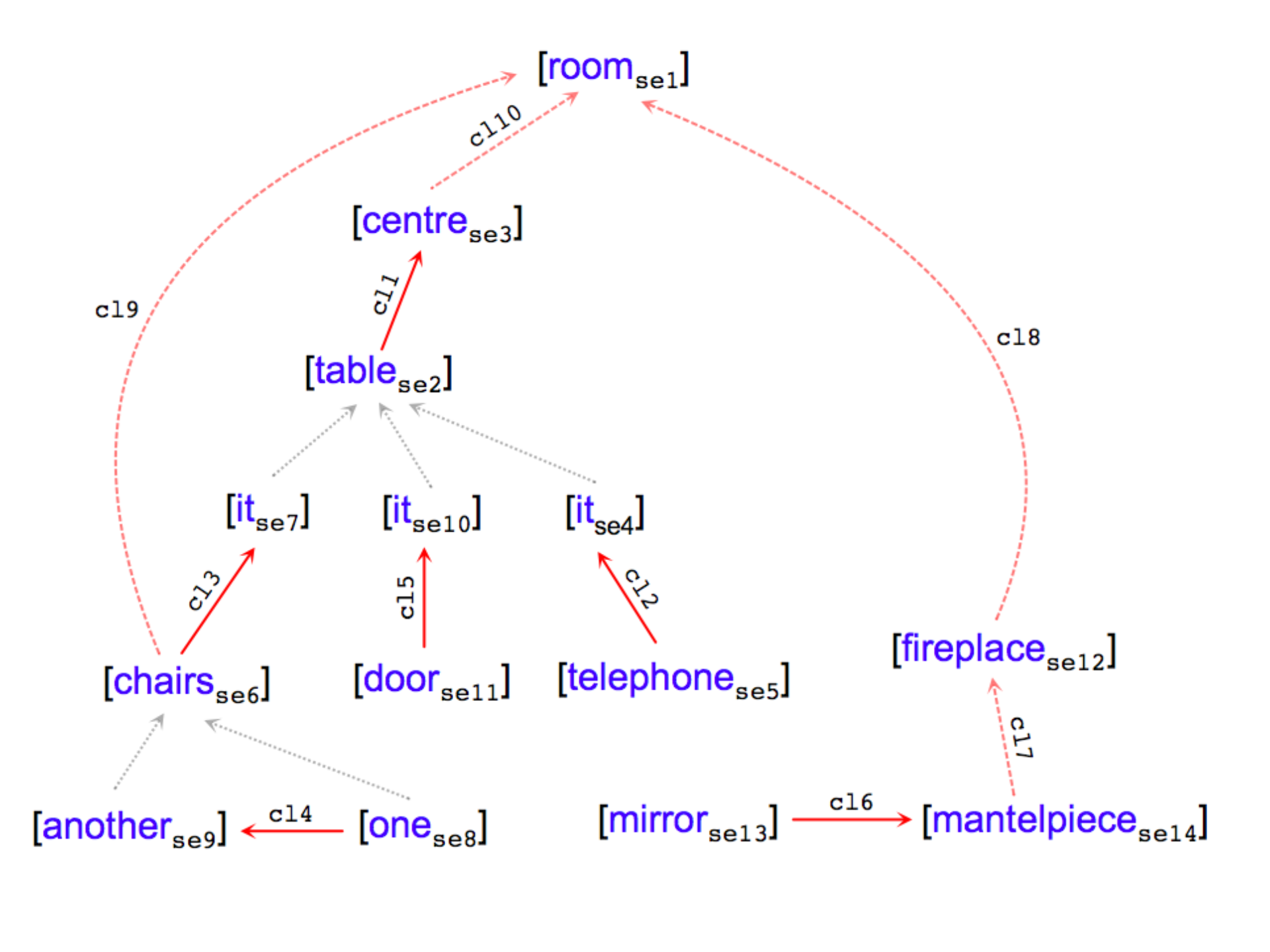
cl7

(trigger=∅,
figure=mantelpiece$_{se14}$,
ground=fireplace$_{se12}$,
figure_config=bottom,
ground_config=top)

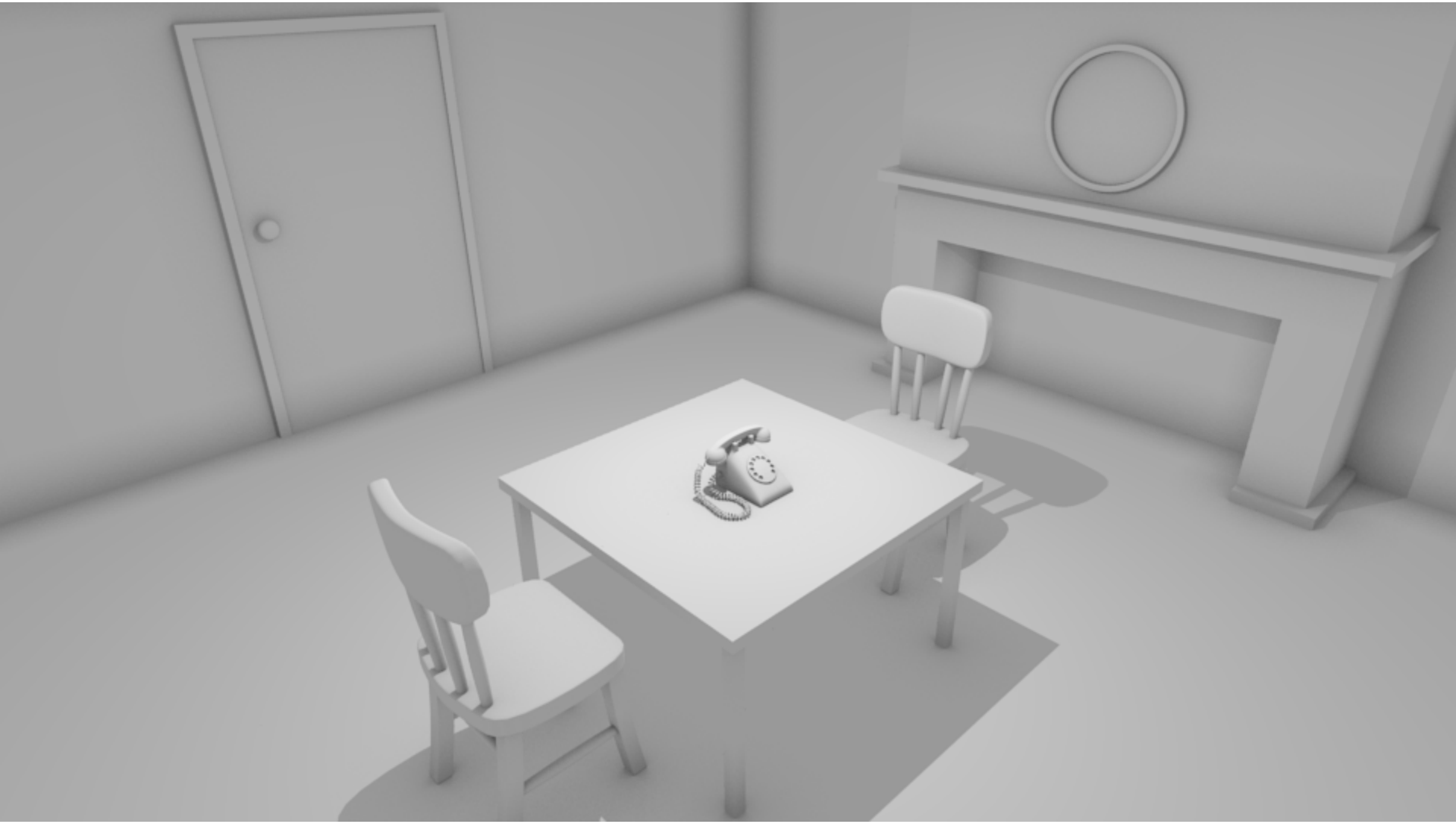
EC(mantelpiece$_{se12}$,fireplace$_{se10}$)






[room se1]
c110
[centre se3]
c11
c19
c18
[table se2]
[it se7]
[it se10]
[it se4]
c13
c15
c12
[chairs se6]
[door se11]
[telephone se5]
[fireplace se12]
c17
[another se9]
c14
[one se8]
[mirror se13]
c16
[mantelpiece se14]

Full Scene Visualization



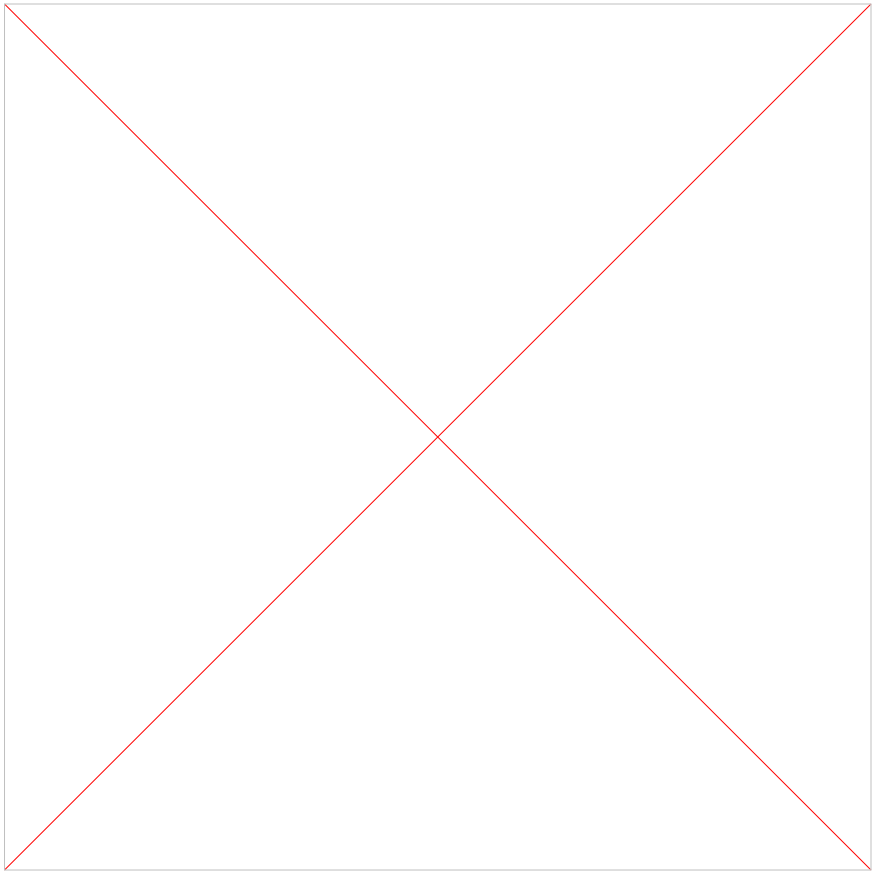
Issues

- Putting it all together
  - Closure
    - Co-reference resolution
  - Quantification
  - Identify the common ground
    - Link underspecified entities to the common ground

Issues (continued)

```
cl4
     (trigger=opposite_os4,
     figure=chairs_se5,
     ground=∅_se7,
     figure_config=front,
     ground_config=front)
```

DC(chairs$_{se5}$,$\varnothing_{se7}$)



cl4, even in conjunction with RCC+ relations, is insufficient for generating a scene that might be expected. Each chair would also need to be linked, bottom-to-bottom with the common ground, (which is the sitting room in this case).

Issues

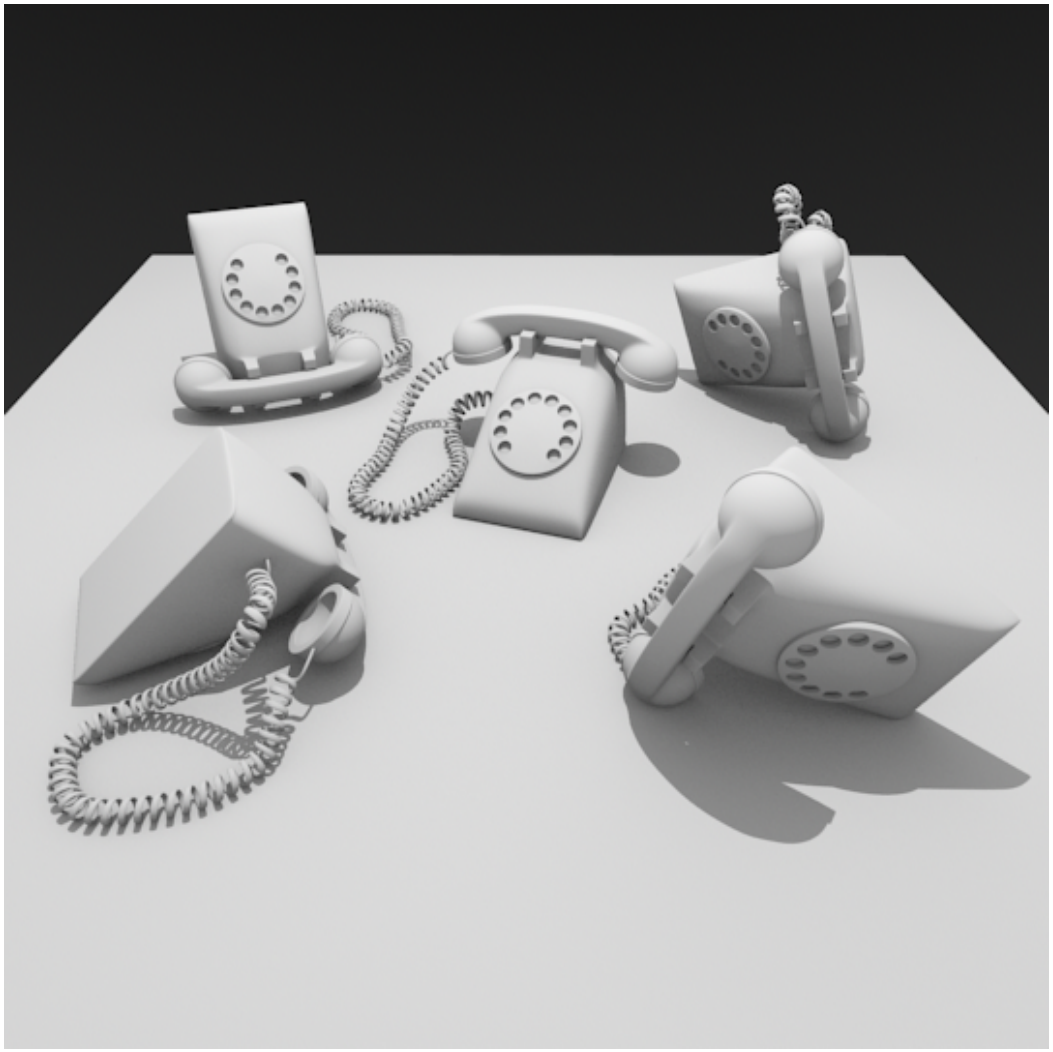
- Reified surfaces
  - “On *his* ***left*** is the fireplace...”
  - “...he has the fireplace on the nearest wall to *his* ***right***...”
  - “The tall man is on *the* ***front*** of the plane.”
  - “***Stage Left***”

Issues (continued)

"Her **lover**, a beautiful **youth** of eighteen, ***in*** evening **dress** and **cape**, with a bunch of **flowers** and an opera **hat** ***in*** his **hands**, comes in alone."

Issues (continued)

- Spatial prepositions and other orientation signals are difficult to capture
- [***On***$_{os2}$] [**it**$_{se4}$] a [**telephone**$_{se5}$].



Corpus Analytics

	Shaw	Galsworthy
Extents	301	483
S_E	216	334
O_S	85	149
Links	92	{ }

Corpus Analytics (Continued)

Shaw: 47 sents, 960 tokens, 414 types

Galsworthy: 106 sents, 1548 tokens, 613 types

Corpus Analytics (Continued)

Inter-Annotator Agreement

	Shaw	Galsworthy
Extents	.275 (.132)	.390 (.057)
S_E	.143 (.066)	.147 (.022)
O_S	1 (.276)	0
Links	0	{ }

Machine Learning

Baseline: Binary spatial entity classifier

- Dataset: 939 noun instances from corpus
- Naive Bayes:
  - word as only feature
  - F1: 90.4
- MaxEnt:
  - word as only feature
  - F1: 64.9

Acknowledgements

Thanks to our annotators: Nikhil, Mike, and Tad!

Bibliography

Anthony G Cohn, Brandon Bennett, John Gooday, and Nicholas Mark Gotts. 1997. Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica*, 1(3):275–316.

Matteo Cristani and Anthony G Cohn. 2002. Spaceml: A mark-up language for spatial knowledge. *Journal of Visual Languages & Computing*, 13(1):97–116.

Annette Herskovits. 1980. On the spatial uses of prepositions. In *Proceedings of the 18th annual meeting on Association for Computational Linguistics*, pages 1–5. Association for Computational Linguistics.

Amar Isli and Anthony G Cohn. 2000. A new approach to cyclic ordering of 2d orientations using ternary rela- tion algebras. *Artificial Intelligence*, 122(1):137–187.

Inderjeet Mani, Christy Doran, Dave Harris, Janet Hitzeman, Rob Quimby, Justin Richer, Ben Wellner, Scott Mardis, and Seamus Clancy. 2010. Spatialml: anno- tation scheme, resources, and evaluation. *Language Resources and Evaluation*, 44(3):263–280.

Dan I Slobin, Nini Hoiting, Michelle Anthony, Yael Biederman, Marlon Kuntze, Reyna Lindert, Jennie Pyers, Helen Thumann, and Amy Weinberg. 2001. Sign language transcription at the level of meaning components: The berkeley transcription system (bts). *Sign Language & Linguistics*, 4(1-2):1–2.

Leonard Talmy. 1978. Figure and ground in complex sentences. universals in human language, vol. 4 by joseph greenberg, 625-649.