Text to scene part.

**Methodology**

The text to scene retrieval method here is based only on explicit mention of words. We selected 108 objects from 16 different categories in SHREC’14 database as our sub-database. Then all of 16 different categories are tagged, with relative scale predefined. The retrieval for objects themselves and their relations are nothing but text based check, as here we trying to interpret how their relations will be linked to each object in multiple situations. Previously we were trying to use Stanford NLP [reference link], however, find hard to integrate into our Matlab code. The text to scene program takes input as one or two sentences and the output generates a plausible 3D indoor scene itself. Then it passes the retrieved object tags together with their properties into our probability model for the later use in co-retrieval action. By one of the function named retrieval we get all the objects mentioned in the text and all of their properties stored in a same structure. After that using the same trick for detecting objects, we detect the relationship between objects as well as what those relations originally should link. And then we defined a function called Position as two object with specified spatial relation moving their mesh’s centroid and all other points with the same scale.

**Retrieval**

For retrieval part, the most important thing is acquire all information needed later on and gather them in a list. For example we need each object’s mesh to be loaded. We calculate their centroid and move the entire mesh as move their centroid to O point. We calculate the maximum Euclidean distance as the scale of object and normalize the entire mesh for not getting unrealistic size scene.

**Spatial relationship**

We predefined 7 kinds of position relations and if no relationship is specified in the sentence and with more than one object detected in the text, we than interpret their spatial relation as A is left of the B. For two objects, a relationship is always mentioned in the middle of them represent as A is (some relations) of B as this could be transfer into moving A’s centroid to (some relation). Also a situation could happens if one object is mentioned twice and in that case, if input sentence is not two separate sentences, than a flip of relations and object is need as if a sentence goes like: the chair is to the left of the table, on the table there is a pencil. We don’t want our program to interpret this as table is on the table.