A Scenic Exploration into Words: A Window to Conceptual Space

Shashank Rao and R.N. Kaushal

November 17, 2010

Abstract

In this project we have a two fold aim, one to revisit a classical experiment conducted by Deese regarding the nature of intruders in free recall of word lists and the other to gain an insight into the way concepts are understood and perhaps mapped into a geometrical space. As in the Deese's work, we have looked at word lists and their free recall studying the nature of intruders. Also, we designed an experiment in which a visual scene is presented and the test subjects are asked to elucidate a list of words within a time bound, following a free word association task using the words generated. Our results corroborate with Deese's findings.

0.1 Introduction

We started our exploration trying to understand the notion of a concept. We moved past the fixed theory of Aristotelian concepts defined in terms of certain necessary and sufficient conditions to the rather fluid notion offered by prototype theory of Rosch. We tried to understand Gardenfors's notion of a conceptual space. This brought us to the question of whether the domain of concepts can be mapped on to a geometrical space.

0.2 Our Approach

We started off by exploring J Deese's work in the late 50's on verbal intrusions off a list via free recall. Deese conducted a series of experiments on Word Intrusions in free recall of word lists. In the experiment word lists were given to subjects to memorize and then recall. In such experiments often an 'intruder' word used to come up in recall which was not present on the list. It was seen that the most likely intruder was the word which had maximum average association frequency with the words on the word list. We have created a few word lists to see if we can reproduce any of Deese's findings. Also we used one list that is the classic one used in Deese's findings.
We have thus used four lists:

1. Deese's : Bed, Rest, Awake, Tired, Dream, Wake, Snooze, Blanket, Doze, Slumber, Snore, Nap, Peace, Yawn, Drowsy.

This list used by Deese generated the intruder 'sleep' in his original experiment. We conducted an experiment using 15 subjects to test for this word list.

2. Tulip, Guava, Leaf, Jeans, Hair, Grass, Fruit, Stream, Cap, Blue, Tree, Bottle.

This is a word list which we created in which a slight thematic bias was introduced.

3. Tulip, Gauva, Leaf, Snake, Parrot, Grass, Fruit, Stream, Frog, Blue, Tree, Forest.

This is another word list that we created in which the bias toward a central theme is much stronger. We conducted an experiment using 10 subjects to test for this word list.

4. Quad, Hall, Canteen, Room, Wing, Lab, Lecture, Mess, Ragging, Tutorial, Assignment, Quiz.

As we conucted the tests on the previous lists we felt we needed a list which the subjects, relate to much more strongly. Hence we themed this listed on the 'first year' experiences of IIT. We used 10 subjects to test for this list.

This was our approach to the first half of our project. The results that we found are indicated below in the result section.

0.2.1

For the second half of our project, we wanted to understand Conceptual Spaces which we believe to be way more abstract and subtle. Hence acknowledging their inherent complexity we would like to take a much more intuitive and simplistic approach so as to enable a more thorough understanding albeit a grade coarser.

We instead of considering Conceptual Spaces would like to explore what we call 'word space'. We have designed an experiment in which we show a relevant pictorial scene. Now this scene can be thought of as having a high level of conceptual complexity. From this the subjects in a time bound experiment are asked to come up with a list of words while gazing at the scene. Thus we can think of a complex scene being made to percolate into a more simpler visual bag of words model.

Next we had a word association task in which the subjects have to come up with a free word association with the words on one of the word lists generated by another subject.

Shown below is the picture used for the experiment, followed by some sample data.

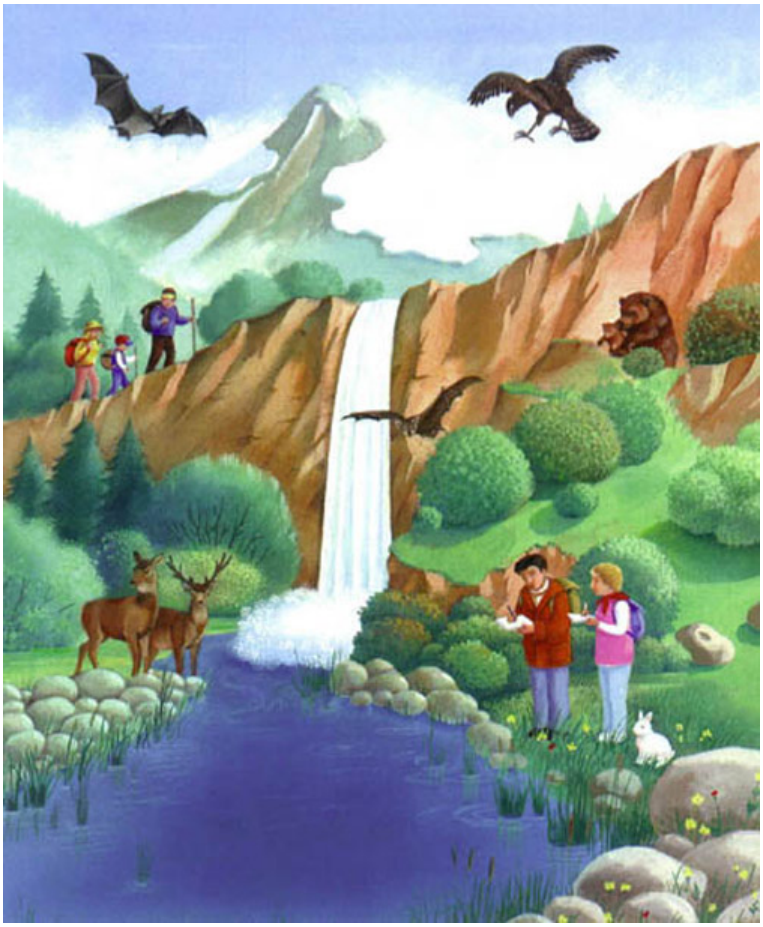

Image Ref: http://picture-book.com/files/userimages/209u/hiking.jpg

BAT	BALL	CRICKET	BOWL	**night**	BOWL
DEER	HORN	GRASS	TIGER	FAST	HORNS
WATERFALL	WATER	BOAT	CLIFF	BATH	SOUND
PEBBLE	ROUND	STONE	SMOOTH	ROCK	STONE
HAWK	SKY	FLY	EYE	FLIGHT	BIRD
HIKE	TRAIL	RUN	WALK	TIRING	ADVENTURE
SON	FATHER	MOTHER	KID	LOVE	CUTE
SNOW	MOUNTAIN	ICE	WHITE	FUN	COOL
BLUE	SKY	GREEN	EYES	RIVER	COLOUR
MARSH	QUICKSAND	SLIP	REPAIR	DIFFICULT	**enjoy**
GRASS	GREEN	LEAF	WIND	FIELD	PEACE
TEDDYBEAR	DOLL	GIRL	TOY	CUTE	GIFT
LOVE	HATE	LIFE	CARE	MOM	CARE
STAMINA	STRENGTH	EXERCISE	SWEAT	ENERGY	**glucon-D**
SLIP	FALL	BREAK	FALL	FALL	**embarrassment**
EDGE	ABYSS	PRECIPICE	KNIFE	**risk**	MOUNTAIN
POEM	POET	SONG	STORY	LULLABY	LOVELY
FRIEND	GIRL	ENEMY	BUDDY	BUDDY	BUDDY
BREAK	OPEN	GLUE	REPAIR	FUN	**kitkat**
PINE	TREE	**desire**	CONE	**tasty**	TREE
SUMMIT	EVEREST	TOP	VIEW	PEAK	**meeting**
GLACIER	SNOW	VAST	ICE	HUGE	**volcano**

0.2.2

We had planned to take two parallel experiments in which we basically do guided tree traversal. However due to lack of time and insufficient data we could not really carry out the experiment to the extent that the results could be statistically significant. The outline of our plan was that we give a test subject a random seed (any word) and from this the subject will get to choose (based on which word strikes his mind most when considering the source word) between each of the neighbours of the word corpus. Then based on the next word he chooses he gets the neighbours of the new node and so on. So he traces out a path in the word corpus. By repeating this over many subjects one can obtain a revamped word space model this time the edges and nodes shaded so as to represent the frequency of traversal. This was the first experiment. In this a comparative study of both the models can be done. One in which the nodes and edges are shaded by frequency of word association and other by path traversal. Overall we believe it can be illuminating in revealing the way the word (concepts) are related and perhaps mapped and associated.

In the second version we restricted the choices given for each word, i.e. the top-k most frequent word associations obtained from before will only be given as choices. Hence putting a more tighter bound on the degree of freedom for exploring relatedness of a word.

0.3 Results

0.3.1

Our findings from the word list free recall experiment are as below:

1. Deese's Oringal List Used: Bed, Rest, Awake, Tired, Dream, Wake, Snooze, Blanket, Doze, Slumber, Snore, Nap, Peace, Yawn, Drowsy.
Subjects : 15
Average Words Recalled : 100 words on 180 words = 5.5 out of 12 words.
Intruders : Sleep (5/15), Alarm (1/15), Moon (1/15), Door (1/15), Power (1/15), Work (1/15).

Thus as expected the intruder 'sleep' came up which is the maximally associated word not included in the list, as found out by Deese too. Other intruders such as Alarm, Moon, Door can be seen to be somewhat associated with the words on the list. Thus this serves to illustrate the point that the gist of the words is more likely remembered than the words individually.

2. Tulip, Gauva, Leaf, Snake, Parrot, Grass, Fruit, Stream, Frog, Blue, Tree, Forest.
Subjects : 10
Average Words Recalled : 75 words on 120 words = 6.25 out of 12 words.
Intruders : Green (2/10), Garden (1/10), Bird (1/10), Lake(1/10).

3. Quad, Hall, Canteen, Room, Wing, Lab, Lecture, Mess, Ragging, Tutorial, Assignment, Quiz.
Subjects : 10
Average Words Recalled : 93 out of 120 words = 7.75 out of 12 words.
Intruders : Attendance (2/10), Exam (2/10), Classroom (1/10), Senior (1/10).

This list clearly has a much higher recall which we attribute to fact that it is heard with much more enthusiam and the subjects can relate more strongly to it. The intruders here also seem much associated with the words on the list as with Attendance, Exam, Classroom and Senior.

From the above experiemt we find out that often an intruder not on the list pops up during recall which has a sense that overlaps maximally with

words on the list. Thus, these findings correlate with what Deese found.

0.3.2

As for the second half of our project, from the words generated directly from the scene, we have considerable overlap of the words across different subjects. What we feel we gleaned out of it, is not so much the quantitative conclusions that we had set out to gather but rather a qualitative feel of the enormously deep and subtle way words seem to be connected. Thus we feel that our endevour remained more of a psychological experiment in which the experimenters have set out not with a detailed plan and roadmap but rather set out to just explore and first hand see how intricately the web of words is woven all around us.

We feel that we have somewhat gathered on what not to do, more than the other way around. Having people recite off list of words off a scene perhaps does not offer a way to explore word connectivity as much as maybe having the subjects describe the scene. We would like to explore this dimension to somewhat a larger extent. However, it seems that bias introduced in an experiment setting is too great for a scientific foray in this direction. We feel that a lot more subtlety is to be brought about in designing the experiment so much so that the subjects do not know that they are participating in one!

We have spent more time discussing and wondering about how it is all tied together rather than hunt for conclusions. But we feel that this is a problem to dwell on for a much longer period of time hoping that it will figure as evening discussions with a group of friends.

0.4 References

1. Conceptual Space: The Geometry of Thought - Peter Gardenfors; MIT Bradford Books, 2000

2. On the Prediction of occurrence of particular verbal intrusions in immediate recall - J Deese; Journal of Experimental Psychology Vol. 58, No. 1, 1959

3. Picture Credit : http://picture-book.com/files/userimages/209u/hiking.jpg