Distributional Gastronomics

Jonathan Oberländer and Laura Bostan

Trento Universtity
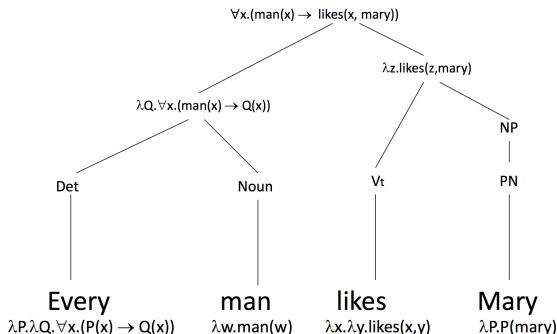

Barcelona, August 3rd-14th

ESSLLI 2015

27th European Summer School in Logic, Language and Information

Formal Semantics

[Frege et al.]

meaning of a sentence: (assigned) meanings of words and *how they are combined*

$$\forall x.(man(x) \rightarrow likes(x, mary))$$

$$\lambda Q. \forall x.(man(x) \rightarrow Q(x))$$

$$\lambda z.likes(z,mary)$$

NP

Det

Noun

V_t

PN

Every
$\lambda P.\lambda Q. \forall x.(P(x) \rightarrow Q(x))$

man
$\lambda w.man(w)$

likes
$\lambda x.\lambda y.likes(x,y)$

Mary
$\lambda P.P(mary)$

Distributional Semantics

[Wittgenstein, Firth, ..., Landauer, Baroni, ...]

Distributional hypothesis: Similar words appear in similar contexts

```
       Capital to be as hot as    Barcelona  as temperatures soar
       sleep on hot nights and    Barcelona  often has very hot
         street of the gothic     Barcelona  and stumbled upon this
    the sweetest street in        Barcelona  So, for those who love
        about street food in      Barcelona  without mentioning the
```

Distributional Semantics

Barcelona, August 3rd-14th
ƎSSLLI 2015
27th European Summer School in Logic, Language and Information

[Wittgenstein, Firth, ..., Landauer, Baroni, ...]
Distributional hypothesis: Similar words appear in similar contexts

```
 Capital to be as hot as  Barcelona  as temperatures soar
 sleep on hot nights and  Barcelona  often has very hot
    street of the gothic  Barcelona  and stumbled upon this
 the sweetest street in   Barcelona  So, for those who love
    about street food in  Barcelona  without mentioning the
```

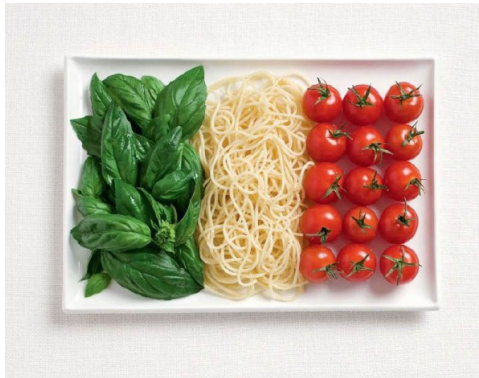|	hot	street	$\cdots$
Barcelona	22	30	$\cdots$
Trento	15	13	$\cdots$
desert	27	3	$\cdots$

Distributional Semantics: Word Relatedness
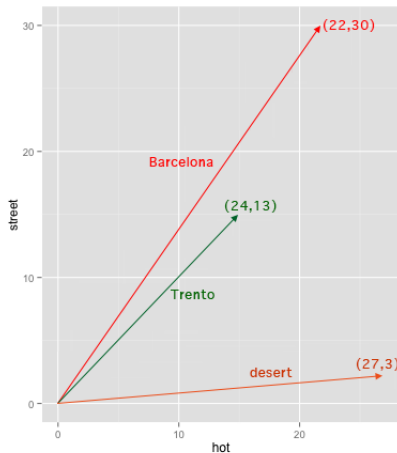
Distributional Gastronomics
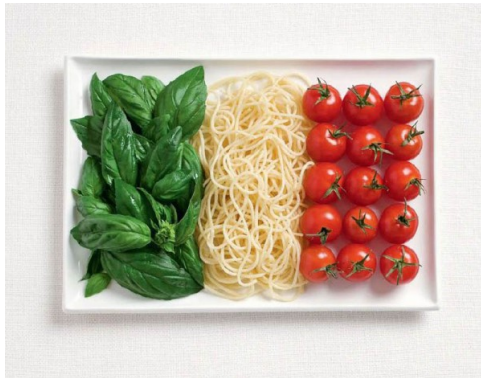
Distributional Gastronomics

Distributional **gastronomical** hypothesis: Similar **ingredients** appear in similar contexts

Open Recipes

{ "_id" : { "$oid" : "5160756b96cc62079cc2db15" }, "name" : "Drop Biscuits and Sausage Gravy", "ingredients" : "Biscuits\n3 cups All-purpose Flour\n2 Tablespoons Baking Powde r\n1/2 teaspoon Salt\n1-1/2 stick (3/4 Cup) Cold Butter, Cut Into Pieces\n1-1/4 cup But ermilk\n SAUSAGE GRAVY\n1 pound Breakfast Sausage, Hot Or Mild\n1/3 cup All-purpose Flo ur\n4 cups Whole Milk\n1/2 teaspoon Seasoned Salt\n2 teaspoons Black Pepper, More To Ta ste", "url" : "http://thepioneerwoman.com/cooking/2013/03/drop-biscuits-and-sausage-gra vy/", "image" : "http://static.thepioneerwoman.com/cooking/files/2013/03/bisgrav.jpg", "ts" : { "$date" : 1365276011104 }, "cookTime" : "PT30M", "source" : "thepioneerwoman", "recipeYield" : "12", "datePublished" : "2013-03-11", "prepTime" : "PT10M", "descripti on" : "Late Saturday afternoon, after Marlboro Man had returned home with the soccer-pl aying girls, and I had returned home with the..." }

Open Recipes

Barcelona, August 3rd-14th

ƎSSLLI 2015

27th European Summer School in Logic, Language and Information

```
12 whole Dinner Rolls Or Small Sandwich Buns (I Used Whole Wheat)
1 pound Thinly Shaved Roast Beef Or Ham (or Both!)
10 ounces carrots, shredded on a box grater or sliced whisper
thin on a mandolin
```

optional: whole pomegranate seeds or fresh/dried rose petals, a
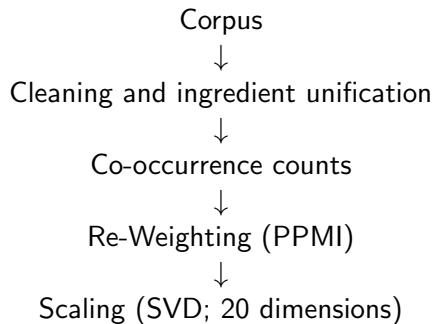
bit of bee pollen

Open Recipes

Barcelona, August 3rd-14th
FSSLLI 2015
27th European Summer School in Logic, Language and Information

```
12 whole Dinner Rolls Or Small Sandwich Buns (I Used Whole Wheat)
1 pound Thinly Shaved Roast Beef Or Ham (or Both!)
10 ounces carrots, shredded on a box grater or sliced whisper
thin on a mandolin
optional:  whole pomegranate seeds or fresh/dried rose petals, a
bit of bee pollen
```

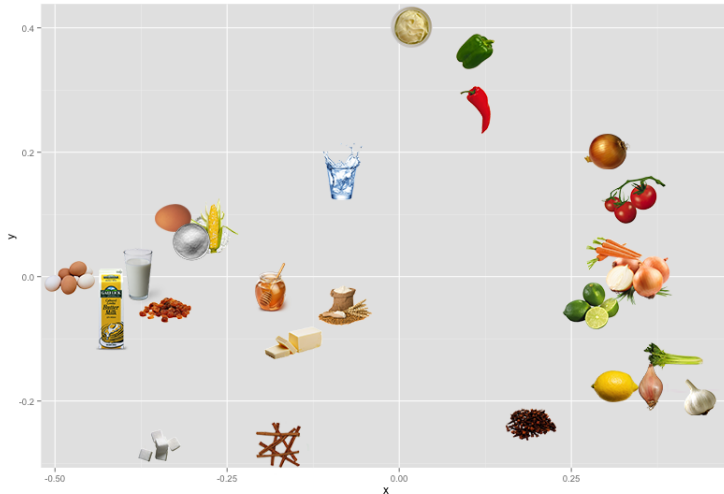|	
Number of recipes	172893
Number of ingredients (tokens in corpus)	1689892
Number of ingredients (types in corpus)	412858
Number of ingredients (types after unifying)	6514

Pipeline

Creating the **Ingredients** space:

<div align="center">

Corpus
↓
Cleaning and ingredient unification
↓
Co-occurrence counts
↓
Re-Weighting (PPMI)
↓
Scaling (SVD; 20 dimensions)

</div>

Ingredients: 2D projection of top ingredients

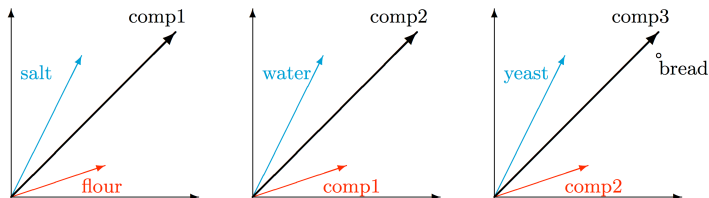# **Ingredients**: Top neighbors of "grapes"
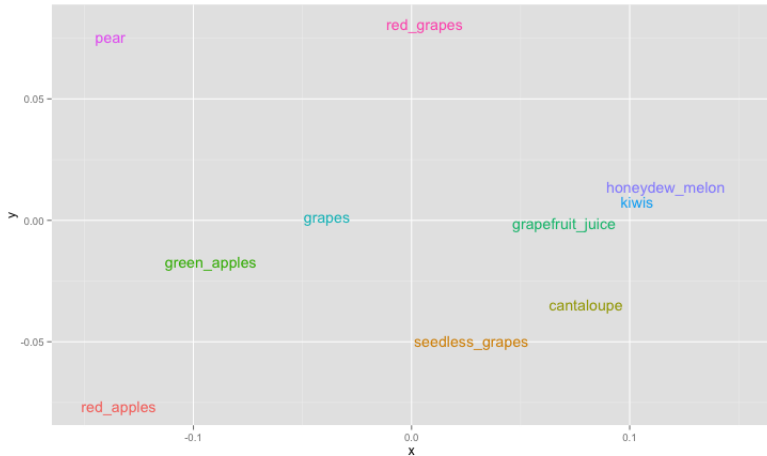
Recipe spaces

BasicRecipes: recipes co-occurs with all of its ingredients

ComposedRecipes: recipe is the sum of the vectors of its ingredients

Applications

```
🚀 ~/g/vector_processing  python nearestNeighbor.py 5
> honey
Nearest neighbors: maple syrup (0.894713303936), agave nectar (0.876596677544), sliced
almonds (0.864462421642), sunflower seeds (0.859026663599), dried cherries (0.858817235
489)
```

```
🚀 ~/g/vector_processing  python makeMeARecipe.py 10
Enter ingredients, enter when done
> water
> orange juice
> basil
>
Nearest neighbors: citrus balsamic salmon (0.950769192732), cantaloupe-basil granita (0
.94599751641), balsamic glazed salmon (0.944641273507), ribeyes (0.943482021276), straw
berry and chicken pasta salad (0.942513766126), west indian chicken curry (0.9415513306
54), fresh strawberry granita (0.940848120637), sweet potato &amp; orange mash (0.93757
294332), ham glaze (0.936598611035), perfect pineapple orange chicken (0.936115489465)
```

Credits

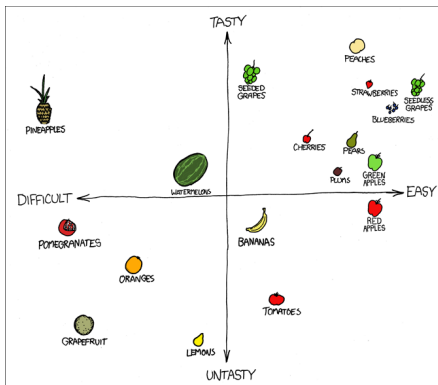|	
Questions?	Come and see our poster.
Ideas?	Come and see our poster.
Criticism?	Come and see our poster.
Want to experiment with our scripts?	Come and see our poster.
Want to plot the neighbors of chocolate?	Come and see our poster.

Formal Semantics picture stolen from: Anca Dinu

Fuck Grapefruit

Coconuts are so far down to the left they couldn't be fit on the
chart. Ever spent half an hour trying to open a coconut with a
rock? Fuck coconuts.
(https://xkcd.com/388)