1. Brush Rabbit Query Results

1.1. Laplace Search

Calycopteryx mosleyi, found on the Kerguelen Islands and Heard Island, is associated with the Kergulen cabbage (Pringlea antiscorbutica, Brassicaceae). As this plant is being destroyed by introduced rabbits, the fly is considered vulnerable.

1.2. JM Search

The brush rabbit inhabits dense, brushy cover, most commonly in chaparral vegetation. It also occurs in oak and conifer habitats and it will live in brush or grassland, and form networks of runways through the vegetation. The brush rabbit does not dig its own burrow or den, but uses the burrow of other species, brush piles, or forms. In the San Francisco Bay Area, it was found that the brush rabbit concentrates its activities at the edge of brush and exhibits much less use of grassy areas. It uses the interior brush of the wilderness and it was also found that this may be a better environment for it than the chaparral one. Studies done on the brush rabbit in Oregon also showed that it rarely left the brushy areas it inhabits. Brush may be used more in the drier seasons while grasses are used in the wetter seasons in relation to growth of annual vegetation. Use of habitat also probably is related to the breeding season.

1.3. Dirichlet Search

Brush rabbit mating, as with other rabbits, may occur year-round but peak breeding seasons are between February and August. The gestation period of the brush rabbit female is about 22 days. A female brush rabbit can have as many as five litters per year but two to three is more common. One to seven young are born per litter and they are altricial. The average number born per litter is three.

1.4. Bigram (Laplace Smoothed) Search

The brush rabbit feeds mainly on grasses and forbs, especially green clover, though it will also take berries and browse from bushes.

2. Brush Query Conclusion

In terms of the paragraph that most comprehensively answers the query (brush rabbit), I would say that the Jelenik-Mercer search method did the best.

3. Results

The plots are located section 3.2.

3.1. Conclusions

The best variant appeared to be the Dirichlet LM model (figure 3). This is also better than the previous variants we implemented (in assignment 3). Using the English stemmer (Lucene's) improved the results.

Table 1: Evaluation Results

Name	MAP	RPREC	NDCG
Laplace	0.2967	0.3237	0.4279
JM	0.3957	0.4089	0.5307
Dirichlet	0.4353	0.4494	0.555
Bigram	0.2905	0.3	0.3902

3.2. Plots

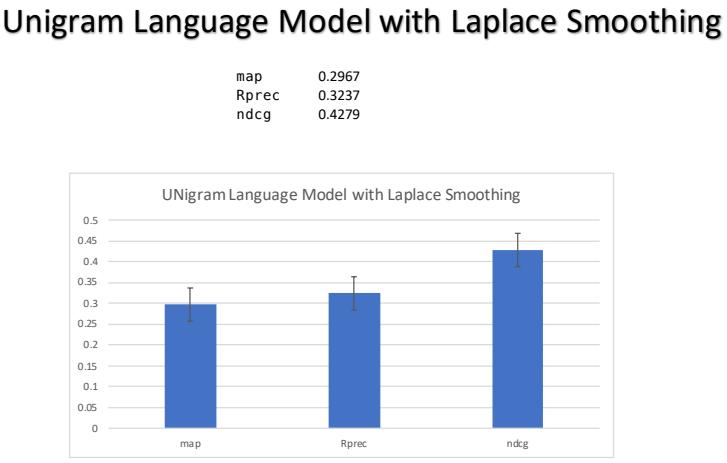

Figure 1: Laplace Smoothing

Unigram Language Model with Jelinek-Mercer smoothing

```
map         0.3957
Rprec       0.4089
ndcg        0.5307
```

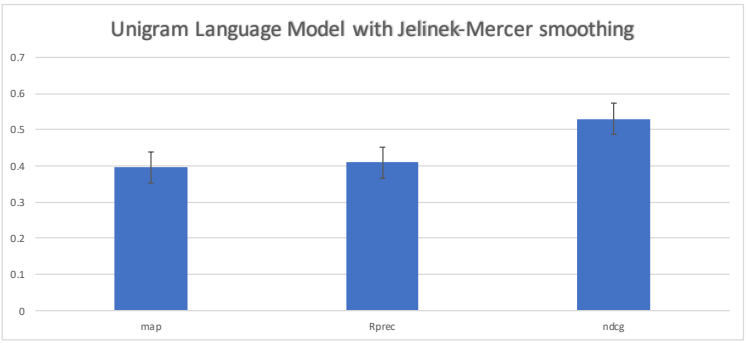

Figure 2: JM

UNIGRAM LANGUAGE MODEL WITH DIRICHLET SMOOTHING

```
map         0.4353
Rprec       0.4494
ndcg         0.555
```

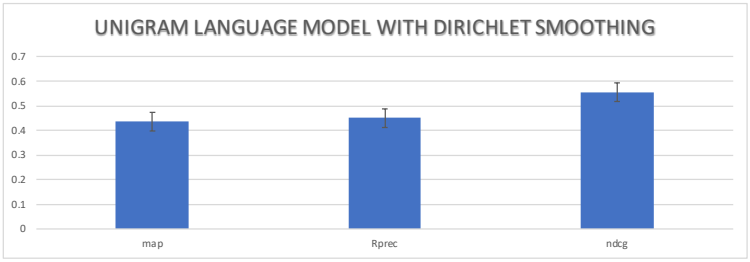

Figure 3: Dirichlet

Bigram Model with Laplace Smoothing

```
map         0.2905
Rprec          0.3
ndcg        0.3902
```

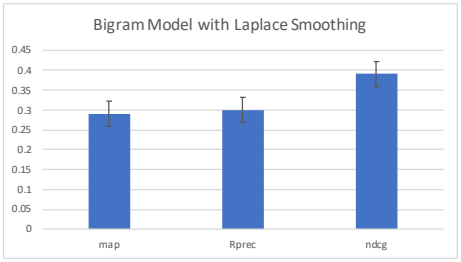

Figure 4: Bigram (Laplace)