

# 数理言語学演習: 八代集言語モデルデータベース

山元 啓史, Ph.D. 東京工業大学

データベースを使って次の質問に答えよ。

## 1 和歌に出てくる花

1. 和歌に出てくる花にはどんな花があるか。想像で良いのであげてみよ。
2. 「梅」を単純モデルで検索し、梅にはどのような用語があるかを調べよ。
3. 「梅」のネットワークモデルをデフォルト（既定値）でよいので、OK ボタンを押して描いてみよ。
4. 「梅」のネットワークモデルの特徴は何かペアで話し合え。
5. 「梅」のネットワークモデルのノードをクリックしてどんな和歌に出てくる用語かを確認めよ。
6. 用語「梅」と和歌との関係について話し合え。
7. 「桜」についても同じことをやってみよ。
8. 「梅」と「桜」の違いは何か、ペアで話し合え。
9. 「梅」と「桜」の関係を集合論、ベン図を書いて考えよ。
10. 「梅」と「桜」を複合モデルを使って検索し、ネットワークモデルを描いてみよ。
11. 「梅」と「桜」の複合モデルにはどんな特徴があるかを話し合え。
12. 同様に「鶯（うぐいす）」と「時鳥（ほととぎす）」についても仮説を立てて同様にモデルを作ってみよ。

## 2 和歌に出てくる鳥

1. 同様に「鶯（うぐいす）」と「時鳥（ほととぎす）」についても仮説を立てて同様にモデルを作ってみよ。

## 3 まとめ

1. Level はどんな意味か。
2. なぜ Level が設定されているか。
3. Unit Size はどんな意味か。
4. Z-score とは何か。
5. Log-likelihood Value とは何か。Qiita による解説と Python プログラムの例。
6. Co-occurrence Weight とは何か。
7. Core Node Pruning とは何か。
8. Pruning Nodes under N twigs とは何か。
9. Graphviz とは何か。

## 4 宿題

1. 本日勉強したことを QR コードから提出せよ。
2. 本日の感想を QR コードからアクセスし、提出せよ（本日中）。
3. 宿題の提出は携帯からできる。



Homework submission