Jan Pedersen
22 July 2010

Outline

- Problem Statement
  - Best effort retrieval vs automated reformulation
- Query Evaluation Architecture
- Query Understanding Models
  - Data Sources

Standard IR Assumptions

- Queries are well-formed expressions of intent
  - Best effort response to the query as given
- Users will reformulate
  - If results do not meet information need

- Actually, queries often contain errors
  - Misspelling
  - Incorrect terms
- Users can't reformulate if they don't understand what's wrong
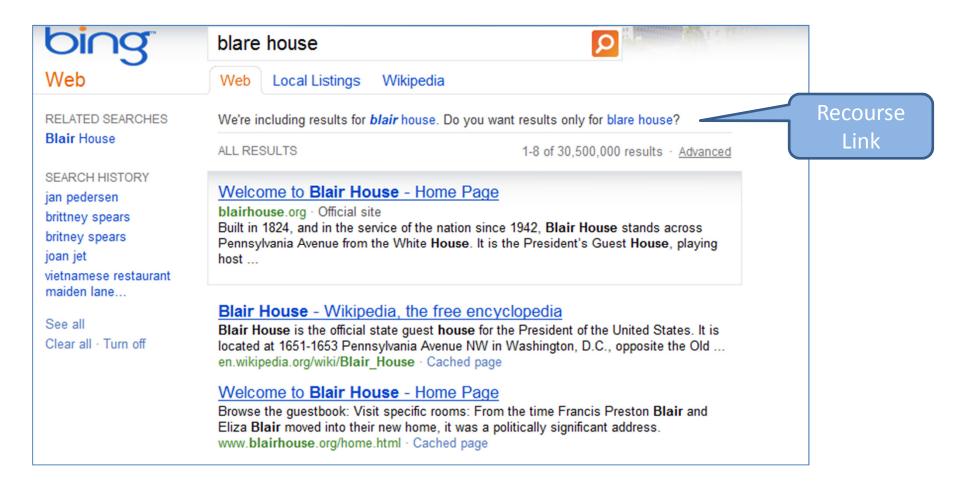 - Miss good content
  - Dead ends

bing

Problem Definitions

- Best Effort Retrieval
  - Find the most relevant results for the user query
    - Segmentation
    - Stemming/Synonym expansion
    - Term deletion
- Automated Query Reformulation
  - Modify the user query to produce more relevant results for the inferred intent
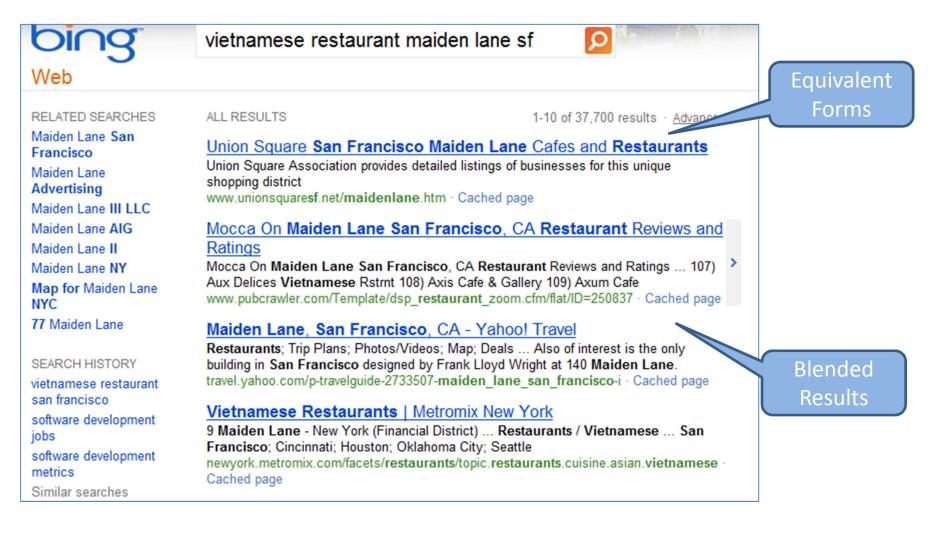   - Spell correction
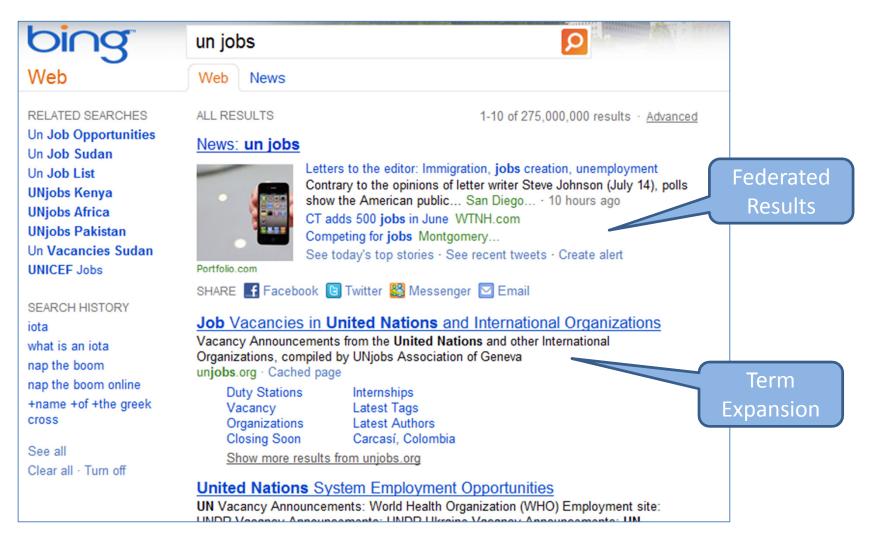   - Term deletion

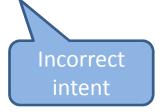bing

Spelling Correction

Stemming

Abbreviations

Term Relaxation

Term Relaxation

Key Concepts

- ## Win/Loss ratios
  - – Automated reformulations are errorful
    - • Wins are queries whose results improve
    - • Losses are queries whose results degrade
  - – Related to precision
    - • But not all valid reformulations change results
- ## Pre vs Post result analysis
  - – Query alternatives generated pre-results
  - – Blending decisions are post results

Query Evaluation

Matching and Ranking

L4
10^1

L3 reranking
10^3 Docs

L2 reranking
10^5 Docs

L0/L1 matching/ranking
10^{10} Docs

- Sequence of select, rank, prune steps
- L0: Boolean logic
- L1: IR score
- L2/L3/L4: Machine learned

Learning to Rank Using Gradient Descent, Burges et al., ICML 2005

bing

Query Annotation

- NLP query annotation
  - Offline analysis
- Ambiguity preserving
  - Multiple interpretations
- Backend Independent
  - Shared
- Structure and Attributes
  - Syntax and semantics

bing

Query Planning

[{un | "united nations"} jobs] →

       L3-merge(L2-rank([un jobs]),

              L2-rank(["united nations" jobs])

or

[{un | "united nations"} jobs] →

       L3-cascade(threshold, L2-rank([un jobs]),

              L2-rank(["united nations" jobs])

Design Considerations

- Efficiency
  - Query plan can generate multiple backend queries
    - Merged in L3
  - Some queries are cheaper than others
    - Query reduction can improve performance
- Relevance
  - L3 merging has maximal information
    - But is costly
- Multiple query plan strategies
  - Depending on query analysis confidence

bing

Query Analysis Models

- Noisy Channel Model

$$\arg\max\{P(rewrite \mid query)\} =$$

$$\arg\max\{P(rewrite)P(query \mid rewrite)\}$$

Language Model

Translation Model

- Example: Spelling
  - Language Model: Likelihood of the correction
  - Translation Model: Likelihood of the error occurring

bing

Language Models

- ## Based on Large-Scale Text Mining

  - Unigrams and N-grams

  - Probability of query term sequence

    - Favor queries seen before

    - Avoid nonsensical combinations

- ## 1T n-gram resource

  - **Multi-Style Language Model for Web Scale Information Retrieval**
    Wang et al., SIGIR 2010

  - **SCOPE: Easy and Efficient Parallel Processing of Massive Data Sets**
    Chaiken et al., PVLDB 2008

bing

Translation Models

- Training set of aligned pairs
  - E.g. misspelling/correction; surface/stemmed
- Query log analysis
  - Session reformulation → associated queries
  - Co-click → associated queries
  - Manual Annotation
- Many application-specific models
  - **Random Walks on the Click Graph**
    Craswell et al., SIGIR 2007
  - **Context sensitive synonym discovery for web search queries**
    Wei et al., CIKM 2009
  - **Context sensitive stemming for web search**
    Peng et al, SIGIR 2007

bing

Summary

- Query Understanding is critical to Web Search
  - Affects most queries
  - Can radically improve results
- Trade-off between relevance and efficiency
  - Rewrites can be costly
  - Win/loss ratio is the key metric
- Especially important for tail queries
  - No meta-data to guide matching and ranking

bing