

情報検索演習 第2回 トランケーション

鶴見大学 非常勤講師
江草由佳
(国立教育政策研究所 総括研究官)
yuka@nier.go.jp

本日のお品書き

- 前回の演習課題提出
- コンピュータ検索の基本
 - トランケーション・・・教科書 p.23～
 - 教材 : CiNii Articles, CiNii Books
- 第2回演習課題: トランケーション
 - 教材 : OneLookDictionary Search

トランケーション(truncation)

- 検索語を入力する場合に、語の一部が任意である(もしくは完全に一致する)ように指定した検索
- トランケーションには4種類+αある
 - 前方一致: 前方が一致する
 - 後方一致: 後方が一致する
 - 中間任意(前後一致): 中間が何でもよい、前方、後方が一致する
 - 中間一致(部分一致): 中間が一致する(部分が一致する)
 - 完全一致: 完全に一致するもの

システムによって異なる 様々なトランケーションの指定方法(1)

- メニューで指定
 - 「で始まる」「で終わる」「を含む」「に一致する」などのメニューから選ぶ



- コマンドで指定: 大きく分けて2種類

- 任意の部分を指定する方法
- 決まっている位置を指定する方法

図書館? / 図書館

ここはなんでもよい(任意)と指定する方法
任意の部分の特殊文字のことをワイルドカードやマスク文字という

ここ(前方)は「図書館」と決まっていますよと指定する方法

ワイルドカード(マスク文字)

- 任意文字とする部分に使用する特殊な入力文字
 - 例: 図書館? ← ? がワイルドカード
 - ワイルドカード(wild card): トランプゲームのジョーカー
 - マスク(mask): 覆い隠す
- ワイルドカードはシステムによって異なる
 - 今回は例として以下を使用して説明する
 - システムBは次のワイルドカードが使える
 - ? : 0文字以上、何文字でもよい
 - # : 0文字または一文字
 - ! : ちょうど一文字



一致指定文字*

- 一致する部分がどこかを示す特殊文字
 - 例: /図書館 ← / が一致指定文字
 - ここ(前方)は「図書館」と決まっていますよと指定
- 一致指定文字はシステムによって異なる
 - 今回は例として以下を使用して説明する
 - システムCは次の一致指定文字が使える
 - / : ここから始まる、もしくは、ここで終わる

*ワイルドカード、マスク文字と異なり、この「一致指定文字」は江草が説明のために作った造語です。

システムによって異なる 様々なトランケーションの指定方法(2)

- デフォルト(標準設定)が、中間一致、完全一致、前方一致とさまざま
 - システムや検索項目によって異なる
- 例: デフォルトが中間一致(部分一致)
 - CiNii Articles <http://ci.nii.ac.jp/> の「刊行物名」のクエリボックスに「教育」と入れて検索して結果を見てみよう。「/教育/」を入れて検索して結果をみてみよう(CiNiiは「/」で囲むと完全一致になる)
- 例: デフォルトが完全一致
 - CiNii Books <http://ci.nii.ac.jp/books/> の「分類」のクエリボックスに「007」と入れて検索して結果を見てみよう。一番上のヒットしたレコードの詳細をみてみよう。右に「分類」が表示され、分類:007となっているのがわかる。「007*」を入れて検索して、検索結果をみてみよう。「007.81」といった分類のレコードがヒットしているのがわかる。

CiNii Articles の刊行物名:

- 部分一致(例:「教育」)(デフォルト)
- 完全一致(例:「/教育/」)

CiNii Books の分類:

- 完全一致(例:「007」)(デフォルト)
- 前方一致(例:「007*」)

前方一致検索

- 検索語の**前方**が**一致**する検索
- 指定した検索語から始まるものがヒットする検索
- 例)「情報」から始まるものが欲しい
 - 情報検索
 - 情報検索システム
 - 情報……「情報(※)」

10

前方一致検索の検索例

- システムA:
 - システムB:
 - システムC:
- ⇒ 情報、情報**検索**、情報**検索システム**
- システムB: ⇒ CAT,CAT**S**,CAT**V**
 - 猫を検索したいときによいが、CATVまで検索されてしまう
 - システムB: ⇒ DIS**K**,DIS**C**
 - 英米綴りの違いを同時に検索できる

システムBのワイルドカード:
?:0文字以上、何文字でもよい
#:0文字または一文字
!:ちょうど一文字
システムCの一致指定文字:/

後方一致検索

- 検索語の**後方**を**一致**させる検索
- 指定した検索語で終わるものがヒットする検索
- 例)「情報」で終わるものが欲しい
 - 安全**情報**
 - 特許**情報**
 - 情報……「(※)情報」

12

後方一致検索の検索例

- システムA:
 - システムB:
 - システムC:
- ⇒情報、**安全情報**、**特許情報**

システムBのワイルドカード:
?:0文字以上、何文字でもよい
#:0文字または一文字
!:ちょうど一文字
システムCの一致指定文字:/

13

中間任意検索(前後一致検索)

- 検索語の**途中**が**任意**のものを検索
(検索語の前後が一致する検索)
- 例)「情報」と「システム」の間が任意のものが欲しい。
(「情報」ではじまり、「システム」で終わるものが欲しい)
 - 情報**管理**システム
 - 情報**検索**システム
 - 情報**システム**……「情報(※)システム」

14

中間任意検索(前後一致検索)の検索例

- システムB:
⇒ 情報システム、情報**管理**システム、情報**検索**システム
- システムB: ⇒ ログ**イン**、ログ**オン**
 - 同義語が同時に検索できる
- システムB: ⇒ WOM**A**N,WOM**E**N
 - 単数形、複数形が同時に検索できる
- システムB: ⇒ GRE**Y**,GR**A**Y
 - 英米綴りの違いを同時に検索できる

システムBのワイルドカード:
?:0文字以上、何文字でもよい
#:0文字または一文字
!:ちょうど一文字

15

中間一致検索(部分一致検索)

- 検索語の**中間(部分)**が**一致**する検索
- 指定した検索語を含むものがヒットする検索
- インターネットの検索エンジンでは中間一致していることが多い
- 例)「情報」を含むものが欲しい
 - 交通**情報**システム
 - 情報……「(※)情報(※)」
 - 交通**情報**……「交通**情報**(※)」
 - 情報**システム**……「(※)情報**システム**」

16

中間一致検索(部分一致検索)の検索例

- システムA:
 - システムB:
- ⇒情報、**交通情報**、情報**システム**、**交通情報システム**

システムBのワイルドカード:
?:0文字以上、何文字でもよい
#:0文字または一文字
!:ちょうど一文字

17

完全一致検索

- 検索語が**完全**に**一致**する検索
- 指定した検索語と完全に一致するものがヒットする検索
- 例)「情報」と完全に一致するものが欲しい
 - 情報
- 以下は完全に一致していないのでヒットしない
 - 交通**情報**
 - 情報**システム**
 - 交通**情報**システム

18

完全一致検索の検索例

- システムA: に一致する ▼ 検索
- システムC:

⇒情報

システムCの一致指定文字: /

19

まとめ

トランケーション

- トランケーションには4種類+αある
 - 前方一致: 前方が一致する
 - 後方一致: 後方が一致する
 - 中間任意(前後一致): 中間が何でもよい、前方、後方が一致する
 - 中間一致(部分一致): 中間が一致する
 - 完全一致: 完全に一致する
- ワイルドカード(マスク文字)
 - 任意の文字を表す
 - ワイルドカードがない(使えない)システムもある
 - システムによって記号の意味は異なる

20

トランケーションの演習 (ワイルドカード)

- OneLook Dictionary Search
 - <http://www.onelook.com/>
 - 複数の辞書を一度に検索(横断検索)できるサイト。(英語のサイト)
 - 英語の単語をトランケーションを使って検索できる
 - 参考資料: OneLook Dictionary Searchの使い方 <http://d.hatena.ne.jp/yegusa/20100130/p1>

21

第2回演習課題:トランケーション

- OneLook Dictionary Searchを使った演習
- 演習課題を行い、提出しなさい
 - ※切: 次回の最初に集めます。
 - 演習課題は講義Webサイトにもあります。
 - 時限、学籍番号、名前、提出日を忘れないこと
- この課題のねらい:
 - トランケーションを使うことに慣れる
 - トランケーションの用語(前方一致、後方一致、中間任意、中間一致、ワイルドカード)がわかるようになる
 - ワイルドカードの使い方に慣れる

22