

“r init,message=F,warning=F,include=F source(' /wrk/flw/init.R') “

## Introduction

In this chapter I test the idea that the social science disciplines were established in the United States in part through a process of cognitive institutionalization. This means that the concept of each discipline could not, especially early in its history, be taken for granted as known or recognizable even among academics and intellectuals, not to mention lay audiences. An impetus for the emergence of the disciplines as professions was a dawning process of recognition of "social science" problems or objects, followed by a consistent form of categorizing the same.

To study such a process historically does not in the first instance require an understanding of the objects that are categorized, and indeed there is no requirement that the objects categorized are themselves homogeneous. The act of labeling the world according to a convention can attain a rigid consistency even if great doubt remains about the meaning, significance, an coherence of the objects so organized.

Empirically, at least, in this study we do count the prevalence of disciplinary terminology. How then can such terminology be counted? I contend that a discipline is established as a categorization mechanism when the set of possible labels for the same object shrinks to a short and ranked list. The position at the top of that list will be occupied by what we call the disciplinary prefix, and below it variations on nouns used to capture different aspects of cultural content from created work to the authors themselves.

## Genre

```
“r query,eval=T q<-data.table( Stem=factor(1:5,labels= c('soci','econ','anth','poli','psyc'))
,Genre=factor(1:5,labels= c('social','economic','cultural','political','mental')) ,Tech-
nique=factor(1:5,labels= c('sociological','economical','anthropological','political','psychological'))
,Ontology=factor(1:5,labels= c('society','economy','culture','polity','mind_NOUN')), Discipline =
factor(1 : 5, labels = c('sociology-sociologyof','(economics-economicsof)','(anthropology-
anthropologyof)','(politicalscience-politicalscienceof)','(psychology-psychologyof)'), Profession =
factor(1 : 5, labels = c('sociologist','economist','anthropologist','politicalscientist','psychologist')), Subdisci-
pline=factor(1 : 5, labels = c('sociologyof','economicsof','anthropologyof','politicalscienceof','psychologyof'))))d
-qdq[, ' := '(Ontology = sub('NOUN','', Ontology), Discipline = sub('(.+) -
.+','
1', Discipline))]dq < -t(dq)dq < -data.table(Class = rownames(dq)[-1], dq[-1,])setnames(dq,q[,c('','as.cha
-melt(q,id.vars = 'Stem',variable.name = 'Class',value.name = 'Query')“
“r term-db,eval=T tts<-gbng2tts.f(ys = 1840,ye = 2009,cfs = T,query =
q) “
```

Consider five social science disciplines—anthropology, economics, political science, psychology, and sociology. In English the labels that like flags lay claim to disciplinary territory are the prefixes soci-, econ-, anth-, poli-, and psyc-. In the prehistory of disciplines these stems appear as recognition of a category of phenomena that can be talked about by anyone, long before a discipline existed to claim priority.

These prefixes diffused first as weakly categorical terms that could modify and lay claim to any worldly object. The first stem terms to diffuse were generic

modifiers like "social" and "economic". Such labels, as ubiquitous and inclusive as they were, defined the outer limit of disciplinary relevance. The term "social problem" is an example of the flag being established as a vague claim to disciplinary relevance. Early in the history of disciplinary recognition, the appellation of genre terms are little more than promises to demonstrate that a problem is indeed a "social" one.

```
“r tts2grgr,include=T,eval=T setkey(tts,Stem) p<-tts2grgr.f( gbng2tts=tts['soci'],order=
ec('social,society,sociology,sociological,sociologist,sociology of') ) “
```

```
“r f-3prefix1-soci,fig.width=mfw,fig.height=mfh,eval=T,include=T,fig.cap='Terms
with Disciplinary Prefix soci-' p “
```

Temporal Sequencing Methods

Correlations between time series are difficult to tease out due to several dynamics that if not controlled for can lead to spurious correlations. Before we can attempt to test causal order we must decompose historical trends in terms into their systematic and residual components, such that we may test the residuals for patterns between two series.

ARIMA models have been criticized for abstracting from historical reality [Isaac:1989hp877]. *After establishing statistical considerations and laying bare our assumptions, we will discuss Series*

```
“r terms,eval=T,include=T sg(dq,tit='Terms searched in the Google Books
Ngrams Database',lab='query',col.align = c('cccccc','lllll')) “
```

ARIMA model

ARIMA, or Auto Regressive Integrated Moving Average, models are effective in decomposing several categories of within-series correlations.

$$I = \frac{MA}{AR} \quad (1)$$

This says that  $I$ , the change in our series, is a function of  $MA$ , a moving but systematic average (a line or higher order polynomial) and  $AR$ , the value of the difference in one or more preceding time steps. In more detail:

$$(1 - B)^d y_t = \frac{c + (1 + \theta_1 B + \dots + \theta_q B^q) e_t}{(1 - \phi_1 B - \dots - \phi_p B^p)} \quad (2)$$

Where  $c$  is a constant drift up or down,  $\theta$ ,  $\phi$ ,  $y_t$ ,  $d$ ,  $q$ ,  $p$ , and  $e_t$ .  $B$ .

```
“r tts2arima<-tts2arima.f(gbng2tts) gbng2tts[, 'Frequency t-Score':=tscore(Frequency),by=Phrase]
gbng2tts[, ':( Predicted=tts2arima[,unlist(fit)] , 'Predicted t-Score'=tts2arima[,unlist(nfit)]
)] tts2arimaaa“
```

Granger Causality

What is the temporal sequence of these terms in each discipline?

```
“r granger library(lmtest) setkey(tts,Stem,Phrase,Year) tts['soci',] t<-gbng2tts.f(query=c('chicken','egg'))
“
```

Results

```
“r setkey(gbng2tts,Phrase,Year) p<-list() ggplot(gbng2tts,aes(x=Year)) +
geom_point(aes(y = Frequency), alpha = 3/10)+geom_line(aes(y = Predicted))+
facet_wrap( Phrase)“
```

```

“r tprefix < -dcast(unique(gbng2tts[,list(Phrase,stem,cat)]),stem cat,value.var='Phrase',fun=paste,cols=
') tprefix[c('anth','psyc'),'A.Generic' := ital('A.Generic')]sg(tprefix,lab='t-
prefix',title='Disciplinary Prefixes',col.align=c(old='cccc',new='llll') ,notes.align
= 'l',notes = 'Italicized items deviate from the stem convention.' ) “

```

As table ?? shows.

Which came first?

Granger tests can help determine which [ @Thurman:1988va; @Granger:1969wx]

Clear secular trends and period effects surrounding WWII and the baby boom. To control:

\* Model the trends. We could estimate the linear trend or splines and then subtract them. \* First differences. Subtract from each point the previous point.

\* Link relatives. Divide each point from the point before it.

Box Cox doesn't mean

```

“r ts-inspect tts() plot(x=tts()y[-1], y = diff(tts()f),type='l') plot(x=tts()y[-(1 :
2)], y = diff(diff(tts()f)),type='l')

```

```

b<-BoxCox(tts()f, BoxCox.lambda(tts()f)) plot(tts()y,b,type='l')plot(x =
tts()y[-1],y=diff(b),type='l') plot(x=tts()y[-(1 : 2)], y = diff(diff(b)), type =
'l')“

```

```

“r ts-test library(forecast) library(tseries) tts() ffit<-tts()fsummary(fit)plot(forecast(fit,h =
10),include = 80)“

```

$$8.1y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t, \quad (3)$$

References