Imagined “*imagined conversations*”. (Brief note
on Danescu/Lee 2011)

Camille Akmut

January 20, 2020

Abstract : flaws in *Cornell Movie-Dialogs Corpus*.

... m377* +++\$+++ PLAGUE +++\$+++
I had to move fast. The hacker copied my garbage file.

*Hackers (1995)

Two computer scientists and a single grep search

We attempted to see if we could use the fruit of their research for our own purposes : we queried "programmer", then "hacker", etc. But :

The very first search we performed on this database revealed a dramatic flaw that should probably lead to some caution¹ :

```
$ grep 'programmer' movie_lines.txt
```

Output :

```
... m21 +++$+++ MILO +++$+++  
She's a programmer?  
... m134 +++$+++ HALDEN +++$+++  
Christoph was a brilliant programmer, but very eccentric. He insisted it be kept vacant.
```

These results can be referenced with another text file, called movie_characters_metadata.txt:

```
m134 +++$+++ metropolis +++$+++ 1927 +++$+++ 8.40 +++$+++ 40730 +++$+++ ...
```

This strikes as odd.

The mention of a "programmer" in this 1920's expressionist film?

We had to think twice : perhaps "programmer", here, meant the maker of a program (implied political);

and perhaps the translator had taken some liberties.

Not an obscure film, either : arguably the most notorious German, Weimar-era production...

Another file is useful, raw_script_urls.txt, to further establish what went wrong :

```
m134 +++$+++ metropolis +++$+++  
http://www.angelfire.com/movies/ridleyscott/script/Metropolis.txt
```

In turn, from this website, which is still directly accessible, we get by reading the very first lines :

```
" METROPOLIS
```

```
by
```

```
Corey Mandell"
```

Metropolis is the unrealised 2000's sequel to *Blade Runner* by Ridley Scott (as the URL itself already indicates).

If the dialog already did not make clear that this was improbable, we now should be certain. /

It's like Picard said in m196 : "*the mission requires a certain knowledge of 21st century history*" ...

¹This applies to both their database and, by way of consequence, the articles they have written, based on it. E.g. "Chameleons in imagined conversations" — from which our own title is derived.

References

- Authors. 2011a. "Chameleons in imagined conversations". <http://www.cs.cornell.edu/~cristian/papers/chameleons.pdf>
- Authors. 2011b. *Cornell Movie-Dialogs Corpus*. https://www.cs.cornell.edu/~cristian/Cornell_Movie-Dialogs_Corpus.html