

iCLEF at Maryland: Comparing Term-for-Term Gloss and MT

Jianqiang Wang and Douglas W. Oard
College of Information Studies/UMIACS
University of Maryland, College Park, MD, USA
(wangjq,oard)@glue.umd.edu

Outline

- Research questions
- Experiment details
- Results
- Next steps

Translation Techniques

- Term-for-term gloss translation (Gloss)
 - Easily built for a wide range of language pairs (days)
 - Look up terms in bilingual term list (35,000 French/English pairs)
 - Select the translation that occurs most often in Brown Corpus
 - Backoff to stems when surface form is untranslatable
- Machine translation (MT)
 - Developing new language pairs is expensive (years)
 - Systran is available for 9 language pairs
 - Syntactic/semantic constraints improve accuracy & fluency

Null Hypotheses

- The effectiveness achieved using MT is the same as that which could be achieved using gloss translations.
 - Focus of this talk
- The effectiveness achieved using gloss translations could have been achieved by following some rule that would not involving looking at any translations at all.
 - Analysis not yet complete

English-French Cross-Language Information Retrieval System

- Designed for interactive Cross-Language European Forum (iCLEF) 2001 at Maryland -

Please input the search ID given to you by the experimenter.

Please try the practice search topic if you're new to the system. Otherwise, select a search topic in the order given to you to start

- Practice search topic
- Search topic 1
- Search topic 2
- Search topic 3
- Search topic 4

There are two ways to display retrieved documents. Please select one as you are told.

- System 1 : machine translation
- System 2 : word-by-word gloss translation

Copyright © 2001 University of Maryland Institute for Advanced Computer Studies.

Last modified: June 27, 2001

The search topic you selected is:

Cancer Genetics. Find documents discussing recent discoveries on the relationship between genetics and cancer. The gene responsible for the development of certain tumors has recently been discovered. All documents that provide information on this new diagnostic tool are relevant.

[search](#)

Results for your search topic:

Cancer Genetics. Find documents discussing recent discoveries on the relationship between genetics and cancer. The gene responsible for the development of certain tumors has recently been discovered. All documents that provide information on this new diagnostic tool are relevant.

50 documents found matching your search topic.

1. [MÉDECINE. Lunched by the government in partnership with French Association against the myopathies. The national foreground of *development* of the genic therapies is put in place](#)
 not relevant somewhat relevant relevant unsure not judged
2. [DEBATS. BIOLOGIE. Obligations of the " genetic man "](#)
 not relevant somewhat relevant relevant unsure not judged
3. [les Deputes Have Adopts The Three Tests On BIOETHIQUE. COMMENTAIRE. Reason and derives](#)
 not relevant somewhat relevant relevant unsure not judged
4. [En seen of the tariffing of the contrats. During five years, the insurers will not have recourse to the *information* provided by the tests génétiques](#)
 not relevant somewhat relevant relevant unsure not judged
5. [ASSOCIATIONS. MÉDECINE. Beautiful money of the " Téléthon " and Généthon](#)
 not relevant somewhat relevant relevant unsure not judged
6. [MÉDECINE. Claiming a " resumption in hand " by Etat. A report/ratio of the IGAS denounces the absence of policy of fight against *cancer* in France](#)
 not relevant somewhat relevant relevant unsure not judged

1.

MÉDECINE; Launched by the government in partnership with French Association against the myopathies; The national foreground of *development* of the genic therapies is put in place

Le launching of the governmental foreground of *development* and framing of the genic therapies testifies to the weight growing to this *new* biological research sector in medicine modern. It illustrates a convenient awakening of some persons in charge for the medical policy. Born from *recent* progress from molecular biology, the genic therapy aims at correcting genetic anomalies which, until now, remained rebellious with any action curative.

One counts thus more than 4 000 genetic, *responsible* diseases with total of 25 % of the handicaps and 30 % of the infant mortality. According to the Principles of internal medicine (1), at last 31 December, 928 of these diseases had given place to a localization of *gene* implied, this *gene* being identified in approximately two thirds of the cas.

In practice, the genic therapy consists, schematically, with to introduce, in the organization reached, of the fragments of ADN (molecule support of the genetic inheritance), corresponding to one or more genes and this at *diagnostic* or therapeutic ends. This transfer can be done in *in vivo* or *ex vivo*, of the cells or fabrics of patient being put in contact with viral vectors carrying gene, then injected into the organization. In France, if it is not yet standardized, it is not any more one marginal activity, reserved for some teams of before-garde.

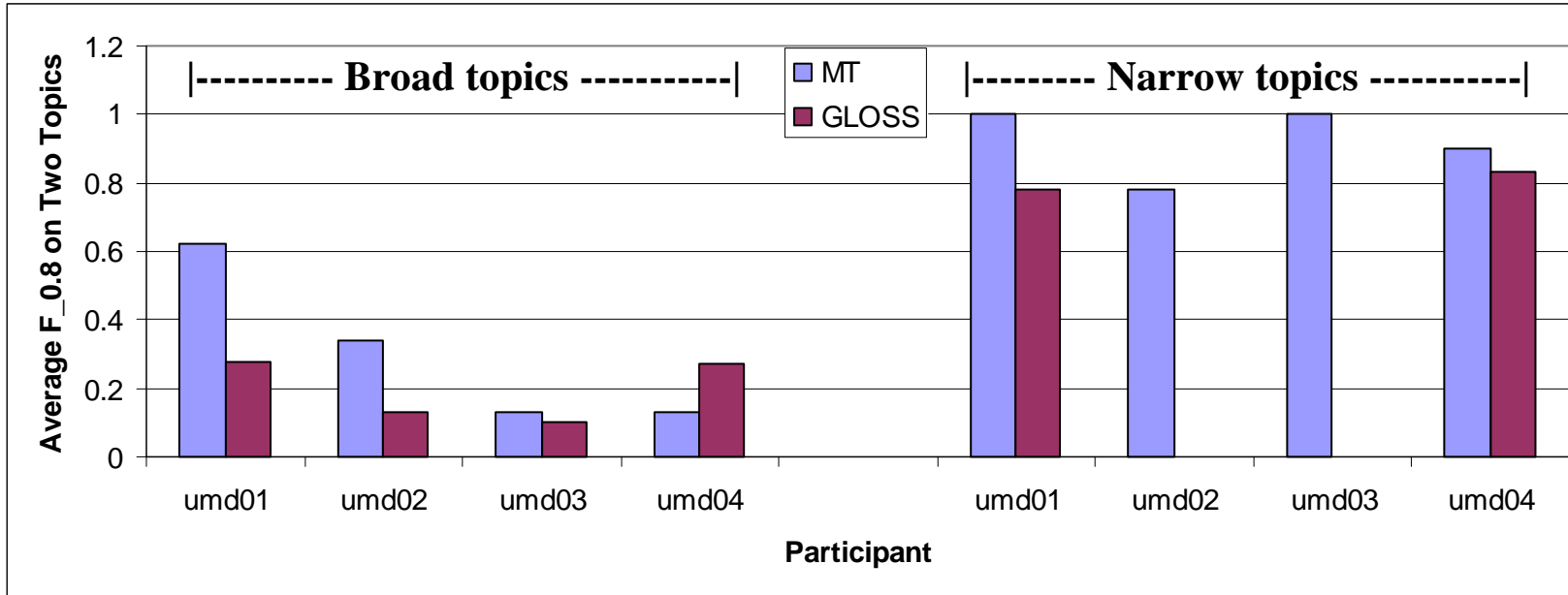
" From the sources of *information* to which we had access, it arises that about fifty laboratories concerned with large organizations of research and the universities is actively engaged, like two French industrial groups, can one read in one report which has been just submitted by professors Jean-Paul Cano and Alain Fischer in Philippe Douste-Blazy. This significant and *recent* mobilization is the fruit of a strong motivation encouraged by public financing and incentives of caritatives associations (...)" Five clinical trials began right now (2). Others are on standby and, according to this report/ratio, " France is located in very good position in the world, far however behind The United States where nearly 70 clinical trials, including 220 patents, have begun since 1990. "

Centers specialized in CHU

Participants

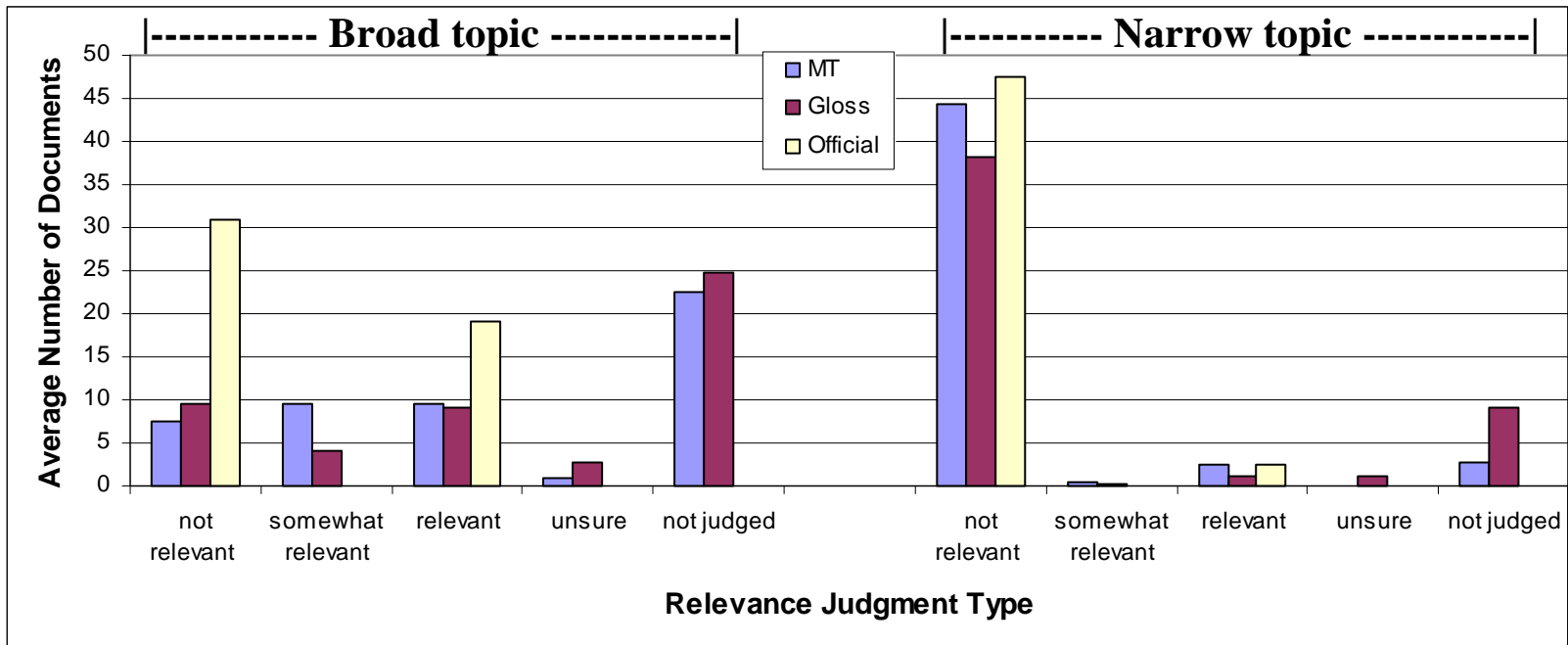
- Heterogeneous backgrounds
 - 2 Information Studies Ph.D. students (M.L.S., 1 male, 1 female)
 - 1 M.S. Computer Science (male)
 - 1 B.A. Religion (male)
- Narrow age range (28-35)
- Frequently search the Web (5-14 years online search experience)
- Experience using MT software varies
- Familiarity with French varies (!)

Official Results



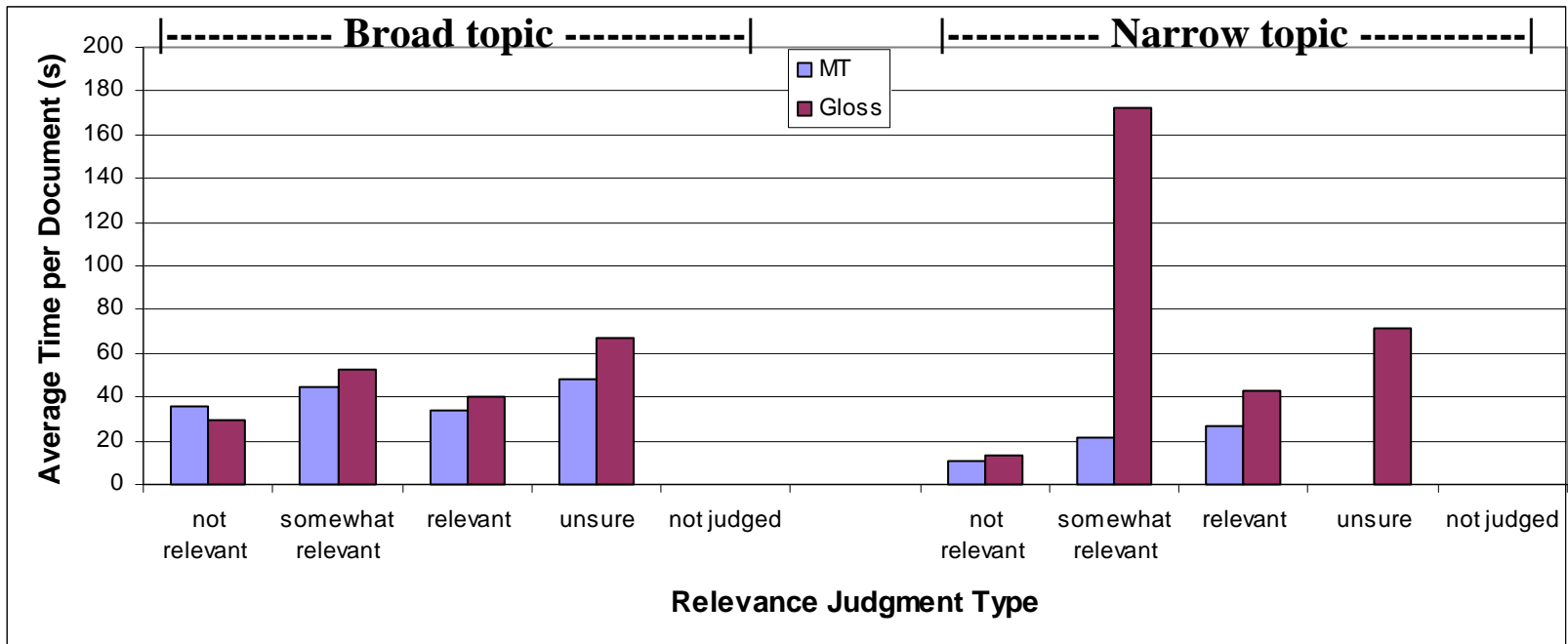
- MT is almost always better
 - Significant overall and for narrow topics alone (one-tailed t-test, $p < 0.05$)
- F measure is less insightful for narrow topics
 - Always near 0 or 1

Descriptive Data Analysis: Quantity



- More judgments for narrow topics
 - Are nonrelevant easier to judge?
- A greater fraction judged relevant for broad topics
 - A presentation order effect?

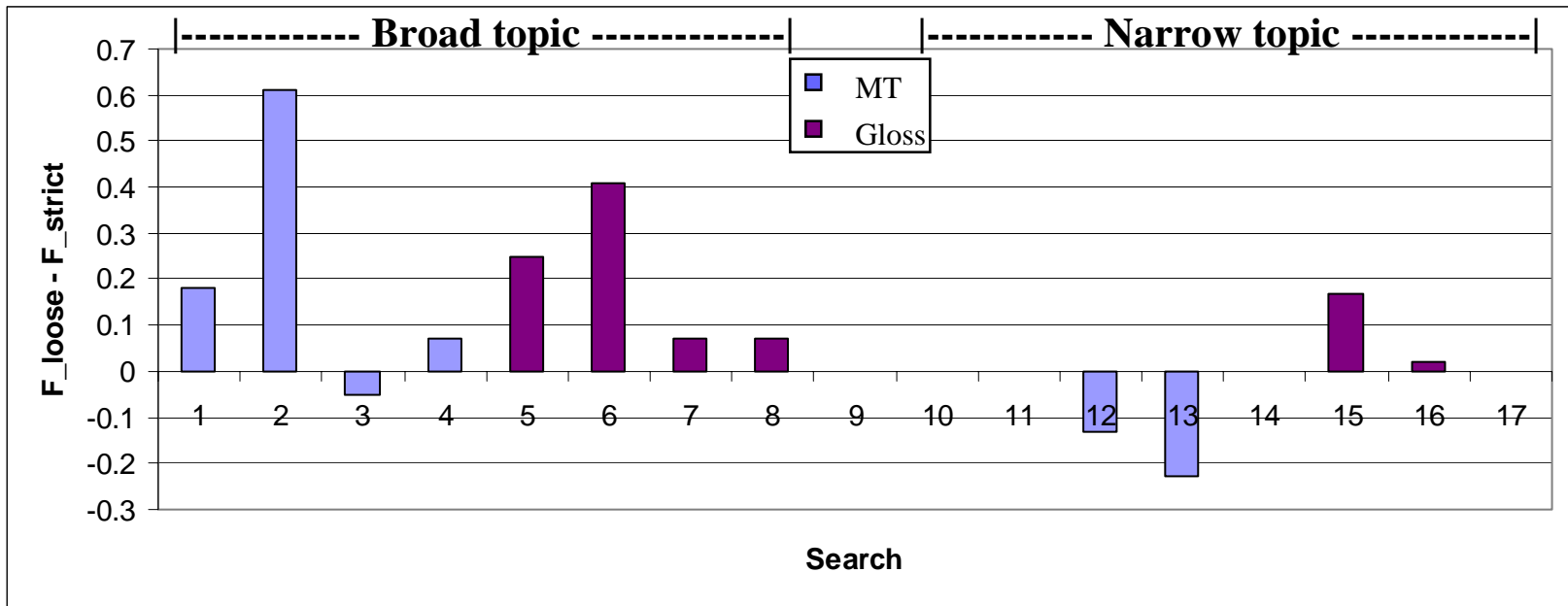
Descriptive Data Analysis: Time



- MT judgments were faster than Gloss
- “Relevant” and “not relevant” were the fastest categories
 - “Somewhat relevant” and “unsure” took longer
- Inverse relationship between time and quantity!
 - A learning effect?

Interpreting “Somewhat Relevant”

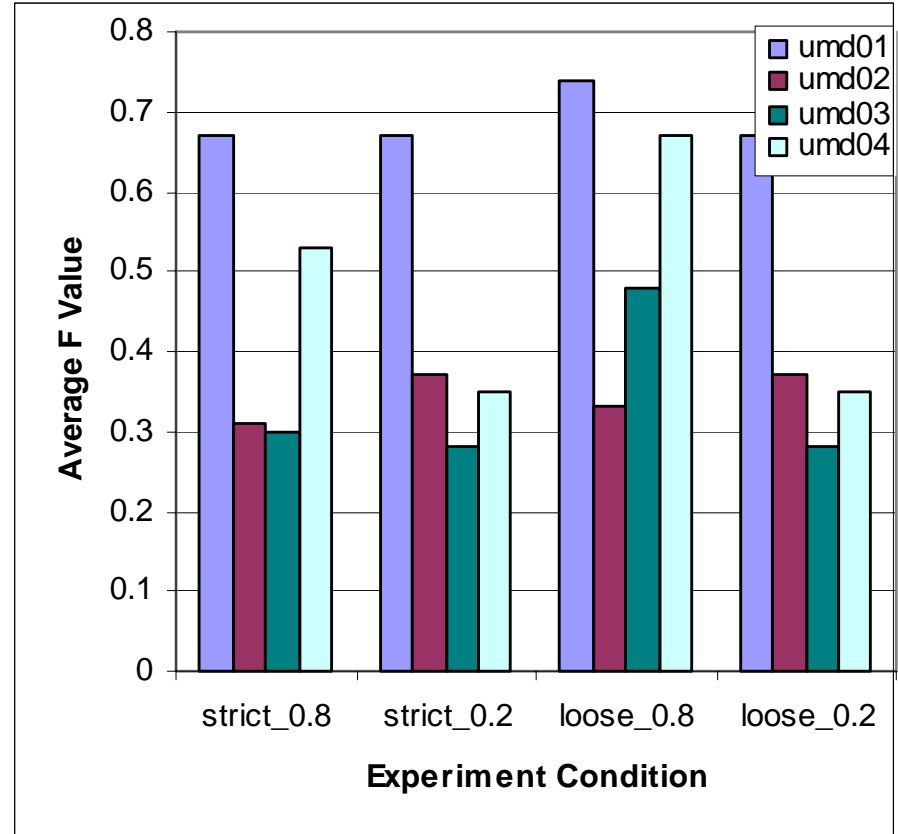
- Strict: scored as not relevant (official)
- Loose: scored as relevant
- If loose outscores strict, the judgments were too conservative for $F_{0.8}$



- Loose judgments improve Gloss consistently (45%), but little benefit for MT (10%)
- The effect is strongest for broad topics
 - Did some participants base “somewhat relevant” on the number of facets covered?

Individual Differences

- umd04 had a strong precision bias
 - Scored consistently lower with $F_{0.2}$
- umd01 did well by every measure
 - Judged more documents (186 vs. 141/200)
 - More searching experience (14 vs. ~5 yrs)
- Knowing French didn't help
 - umd03 reported good French reading skills



Qualitative Results

- Observed behavior (from observational notes):
 - Documents were usually examined in rank order
 - Searchers seemed to understand (and have confidence in) the automatic ranking
 - Title alone was seldom used to judge documents as “relevant”
 - But sometimes used to judge as “not relevant” (learning effect?)
- Subjective reactions (from questionnaires):
 - Everyone liked MT
 - Only one participant liked anything about gloss translation
 - MT was preferred overall
 - Ease of learning (2 favored MT, 2 comparable)
 - Ease of use (3 favored MT, 1 favored Gloss)

Next Steps

- Lots of room for improvement in our gloss translations
 - Retain punctuation and upper/lower case
 - Highlight morphological variants of query terms
 - Render untranslated terms in a recognizable way
 - Provide some form of access to alternative translations
- Develop a measure tuned to narrow topics
 - 5 minutes per topic, full credit for finding any relevant document?
- Experiment with cross-language query formulation

We're Giving it Away!

- Improved user interface (written in Java)
 - Incorporates all suggested improvements
 - Judgment and confidence separately recorded
- Gloss translation code (written in Perl)
 - Useful baseline for language pairs without MT