Logic Programming for Natural Language Processing

Men Young Lee
TJHSST Computer Systems Laboratory

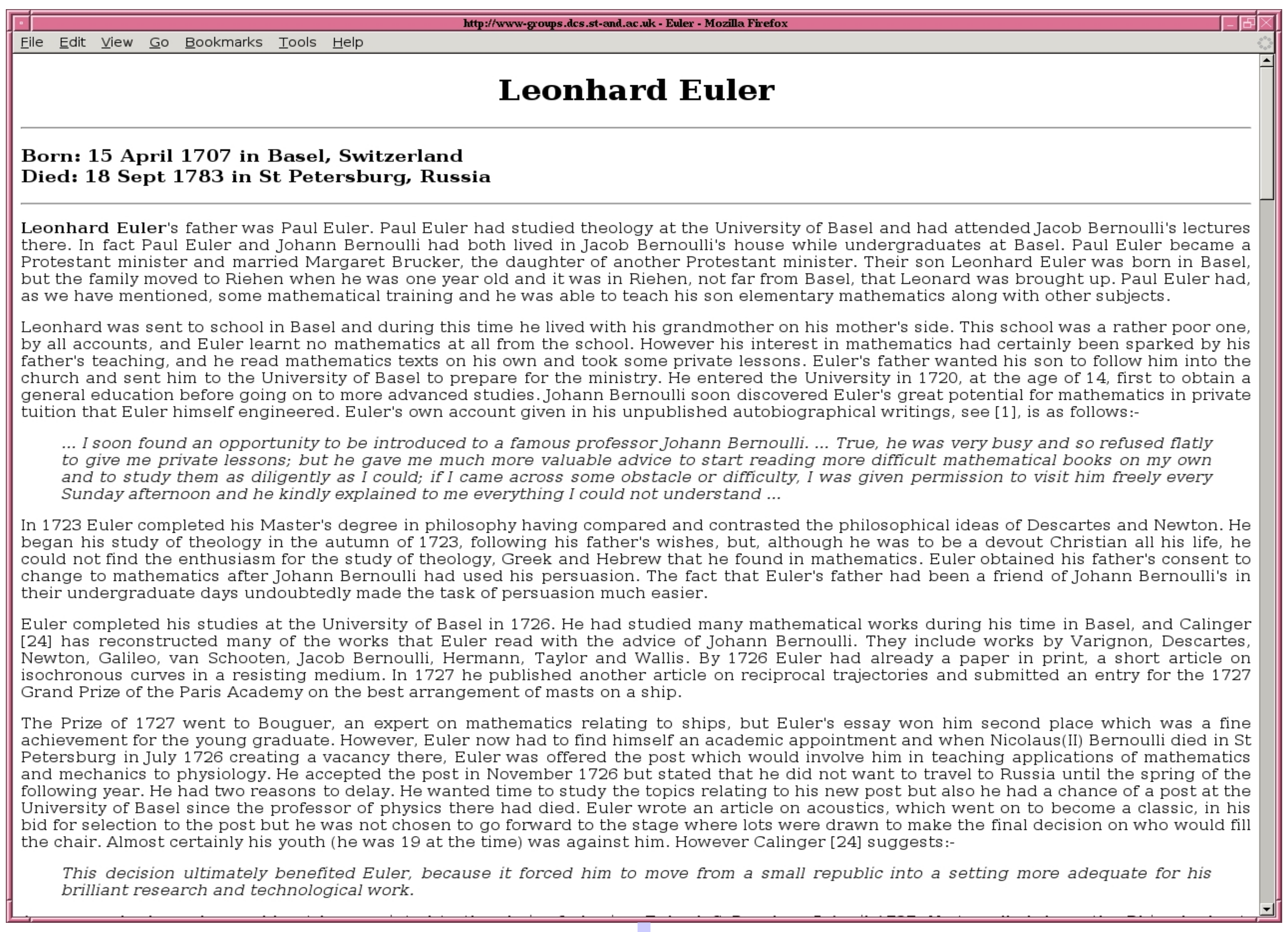
Overview

We propose a bifurcated paradigm for the construction of a Prolog knowledge base from a body of documents: first, an information extraction (IE) application that will annotate the corpus and output the annotated documents, and second, a Prolog knowledge base (KB) application that will transform the annotated documents into a KB (a set of facts).

The General Architecture for Text Engineering (GATE) was used as the platform for the development and execution of the IE application, which included most components of A Nearly New Information Extraction (ANNIE) system. Apart from the basic IE capabilities of ANNIE, the application featured additional high-level annotation grammars written in the Java Annotation Patterns Engine (JAPE) language and a trainable annotator that used the maximum entropy machine learning model, which were designed to annotate several biographies of well-known mathematicians. The Prolog KB application, programmed to be executed within the XSB System, was designed to receive the annotated text output by the IE application and produce a knowledge base, and it successfully creates a database of Prolog facts that can be intelligently queried through the XSB System. The KB utilizes the frame representation of facts, specifically by treating one document as an object to be represented as a frame, with each annotation type treated as a slot whose multiple values are whichever specific strings were annotated by the IE application. This transformation of extracted information into Prolog facts is a link between IE, a recent development in Natural Language Processing, and logic programming with Prolog.

Background

Natural Language Processing (NLP) is the automated understanding and generation of text written in a natural language such as English. While the ill-defined notion of complete text "understanding" is far beyond the grasp of current state of research, intelligent systems with the ability to automate at least some of the tasks in understanding text can be of assistance to a human expert in reading and analyzing a large corpus of documents. The advent of the Internet and the subsequent explosion in the sheer volume of textual information readily available in electronic form presents new opportunities for exploitation of the available information, while simultaneously it presents a challenge, as it is impossible for an analyst to read so much text. Hence arose Information Extraction (IE), a subset of NLP that calls for the transformation of information contained in free, unstructured text into a prescribed structure, specifically by identifying instances of certain objects, their attributes, and/or relationships between them. The purpose of this project is to link this new development in NLP with classical logic programming by developing an application that construct a Prolog knowledge base of information extracted from text.


http://www-groups.dcs.st-and.ac.uk - Euler - Mozilla Firefox
File
Edit
View
Go
Bookmarks
Tools
Help

Leonhard Euler

Born: 15 April 1707 in Basel, Switzerland
Died: 18 Sept 1783 in St Petersburg, Russia

Leonhard Euler's father was Paul Euler. Paul Euler had studied theology at the University of Basel and had attended Jacob Bernoulli's lectures there. In fact Paul Euler and Johann Bernoulli had both lived in Jacob Bernoulli's house while undergraduates at Basel. Paul Euler became a Protestant minister and married Margaret Brucker, the daughter of another Protestant minister. Their son Leonhard Euler was born in Basel, but the family moved to Riehen when he was one year old and it was in Riehen, not far from Basel, that Leonard was brought up. Paul Euler had, as we have mentioned, some mathematical training and he was able to teach his son elementary mathematics along with other subjects.

Leonhard was sent to school in Basel and during this time he lived with his grandmother on his mother's side. This school was a rather poor one, by all accounts, and Euler learnt no mathematics at all from the school. However his interest in mathematics had certainly been sparked by his father's teaching, and he read mathematics texts on his own and took some private lessons. Euler's father wanted his son to follow him into the church and sent him to the University of Basel to prepare for the ministry. He entered the University in 1720, at the age of 14, first to obtain a general education before going on to more advanced studies. Johann Bernoulli soon discovered Euler's great potential for mathematics in private tuition that Euler himself engineered. Euler's own account given in his unpublished autobiographical writings, see [1], is as follows:-

> *... I soon found an opportunity to be introduced to a famous professor Johann Bernoulli. ... True, he was very busy and so refused flatly to give me private lessons; but he gave me much more valuable advice to start reading more difficult mathematical books on my own and to study them as diligently as I could; if I came across some obstacle or difficulty, I was given permission to visit him freely every Sunday afternoon and he kindly explained to me everything I could not understand ...*

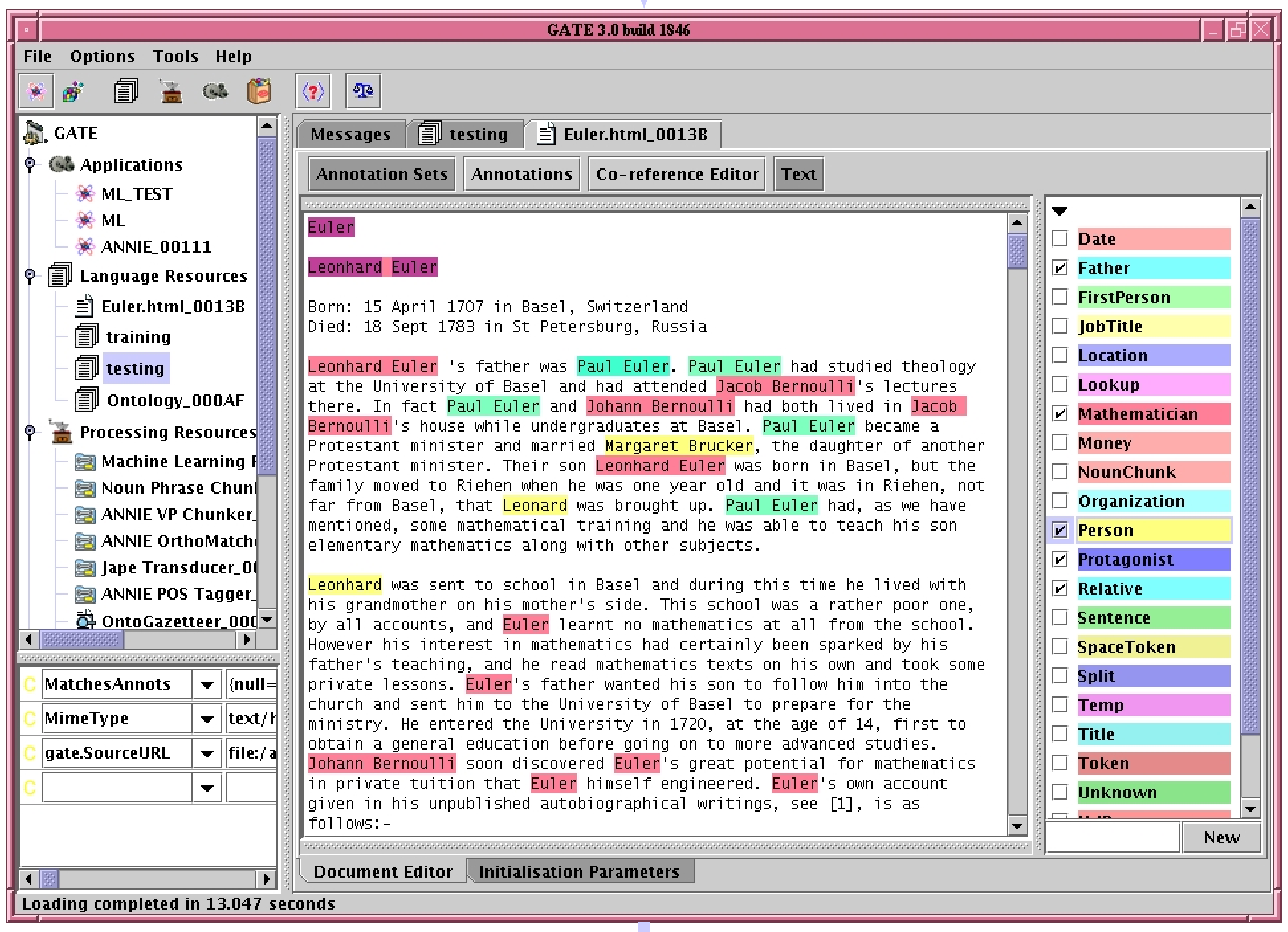
In 1723 Euler completed his Master's degree in philosophy having compared and contrasted the philosophical ideas of Descartes and Newton. He began his study of theology in the autumn of 1723, following his father's wishes, but, although he was to be a devout Christian all his life, he could not find the enthusiasm for the study of theology, Greek and Hebrew that he found in mathematics. Euler obtained his father's consent to change to mathematics after Johann Bernoulli had used his persuasion. The fact that Euler's father had been a friend of Johann Bernoulli's in their undergraduate days undoubtedly made the task of persuasion much easier.

Euler completed his studies at the University of Basel in 1726. He had studied many mathematical works during his time in Basel, and Calinger [24] has reconstructed many of the works that Euler read with the advice of Johann Bernoulli. They include works by Varignon, Descartes, Newton, Galileo, van Schooten, Jacob Bernoulli, Hermann, Taylor and Wallis. By 1726 Euler had already a paper in print, a short article on isochronous curves in a resisting medium. In 1727 he published another article on reciprocal trajectories and submitted an entry for the 1727 Grand Prize of the Paris Academy on the best arrangement of masts on a ship.

The Prize of 1727 went to Bouguer, an expert on mathematics relating to ships, but Euler's essay won him second place which was a fine achievement for the young graduate. However, Euler now had to find himself an academic appointment and when Nicolaus(II) Bernoulli died in St Petersburg in July 1726 creating a vacancy there, Euler was offered the post which would involve him in teaching applications of mathematics and mechanics to physiology. He accepted the post in November 1726 but stated that he did not want to travel to Russia until the spring of the following year. He had two reasons to delay. He wanted time to study the topics relating to his new post but also he had a chance of a post at the University of Basel since the professor of physics there had died. Euler wrote an article on acoustics, which went on to become a classic, in his bid for selection to the post but he was not chosen to go forward to the stage where lots were drawn to make the final decision on who would fill the chair. Almost certainly his youth (he was 19 at the time) was against him. However Calinger [24] suggests:-

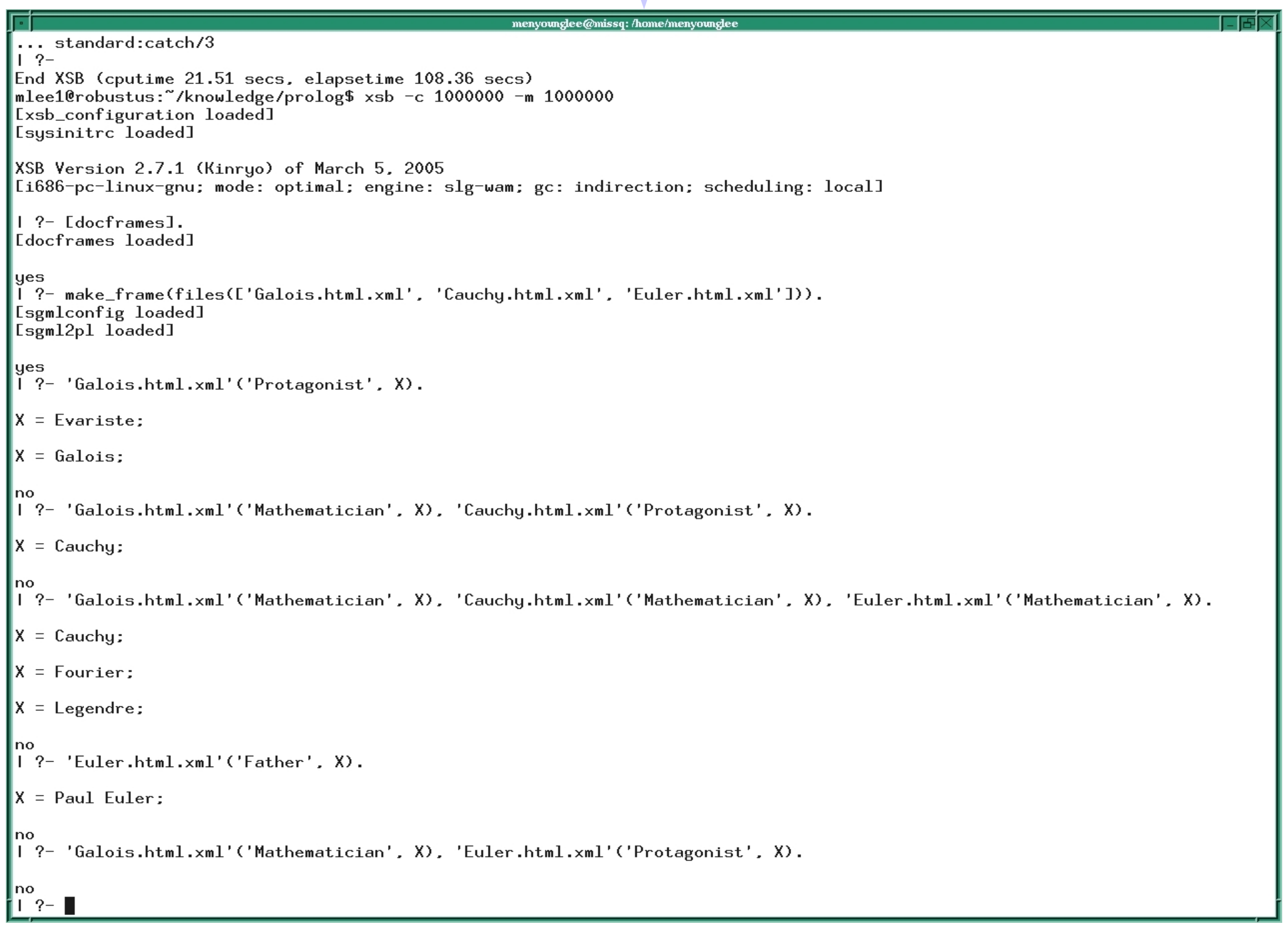
> *This decision ultimately benefited Euler, because it forced him to move from a small republic into a setting more adequate for his brilliant research and technological work.*

A GATE IE Application annotates the document



```
Euler

Leonhard Euler

Born: 15 April 1707 in Basel, Switzerland
Died: 18 Sept 1783 in St Petersburg, Russia

Leonhard Euler 's father was Paul Euler. Paul Euler had studied theology
at the University of Basel and had attended Jacob Bernoulli's lectures
there. In fact Paul Euler and Johann Bernoulli had both lived in Jacob
Bernoulli's house while undergraduates at Basel. Paul Euler became a
Protestant minister and married Margaret Brucker, the daughter of another
Protestant minister. Their son Leonhard Euler was born in Basel, but the
family moved to Riehen when he was one year old and it was in Riehen, not
far from Basel, that Leonard was brought up. Paul Euler had, as we have
mentioned, some mathematical training and he was able to teach his son
elementary mathematics along with other subjects.

Leonhard was sent to school in Basel and during this time he lived with
his grandmother on his mother's side. This school was a rather poor one,
by all accounts, and Euler learnt no mathematics at all from the school.
However his interest in mathematics had certainly been sparked by his
father's teaching, and he read mathematics texts on his own and took some
private lessons. Euler's father wanted his son to follow him into the
church and sent him to the University of Basel to prepare for the
ministry. He entered the University in 1720, at the age of 14, first to
obtain a general education before going on to more advanced studies.
Johann Bernoulli soon discovered Euler's great potential for mathematics
in private tuition that Euler himself engineered. Euler's own account
given in his unpublished autobiographical writings, see [1], is as
follows:-
```

A Prolog KB application transforms annotations to facts



```
... standard:catch/3
| ?-
End XSB (cputime 21.51 secs, elapsetime 108.36 secs)
mlee1@robustus:~/knowledge/prolog$ xsb -c 1000000 -m 1000000
[xsb_configuration loaded]
[sysinitrc loaded]

XSB Version 2.7.1 (Kinryo) of March 5, 2005
[i686-pc-linux-gnu; mode: optimal; engine: slg-wam; gc: indirection; scheduling: local]

| ?- [docframes].
[docframes loaded]

yes
| ?- make_frame(files(['Galois.html.xml', 'Cauchy.html.xml', 'Euler.html.xml'])).
[sgmlconfig loaded]
[sgml2pl loaded]

yes
| ?- 'Galois.html.xml'('Protagonist', X).

X = Evariste;

X = Galois;

no
| ?- 'Galois.html.xml'('Mathematician', X), 'Cauchy.html.xml'('Protagonist', X).

X = Cauchy;

no
| ?- 'Galois.html.xml'('Mathematician', X), 'Cauchy.html.xml'('Mathematician', X), 'Euler.html.xml'('Mathematician', X).

X = Cauchy;

X = Fourier;

X = Legendre;

no
| ?- 'Euler.html.xml'('Father', X).

X = Paul Euler;

no
| ?- 'Galois.html.xml'('Mathematician', X), 'Euler.html.xml'('Protagonist', X).

no
| ?- 
```