

MT SummitXI Workshop

The Chinese Room Experiment: The Self-Organizing Feng Shui of MT

September 11, 2007
Copenhagen, Denmark

Workshop Theme

In the nearly 20 years of statistical MT, a salient constraint has persisted, the requirement for very large bilingual corpora to arrive at a model optimally adapted to the alignments between source and target segments. This workshop will use actual humans (the participants) to employ some other adaptive approaches beyond simple Bayesian statistics to arrive at one or more translation models, in an exuberantly uncontrolled setting.

Objective: This workshop will take its participants through a version of the well-known “Chinese Room” experiment which has gained some notoriety in the statistical MT community over the last decade. The participants will replicate an adaptive, naïve organism which learns to translate material from a language it does not know, using a sample of bilingual training data.

Goals: This workshop will take the experiment somewhat further, by looking at ways the entire group might self-organize to optimize the translation task. The goal is to observe not only the progress but the methods settled on by the participants for optimizing solution potential, with a view toward emulating self-organizing processing communities algorithmically.

The workshop will use a low density language with which no participant is likely to be familiar, such as a Native American language or an Altaic language from a small speech community. After an overview of the Chinese Room model and its history, participants will be provided with bilingual training passages to develop alignment hypotheses, and then provided with a set of sentences to be translated. Each participant will be invited to interact with others or remain solitary, to collaborate, compete, and/or adapt as he/she sees fit.

At the end of the translation time period, the group will review the translation results, as well as the methods the participants employed. The workshop leaders will also present findings about the nature of the participant group as it went through the exercise.

Thus the goals of the workshop are to familiarize the participants with the experiment and its history and potential, while also observing the collaborative processes that emerge as the day unfolds.

Participants: As this is a naïve exercise, it is tolerant of a variety of skill and expertise levels. Anyone who speaks English may participate. If it turns out that the participant also knows the source language of the experiment, there will be other interesting tasks during this exercise in which he/she can participate.

Position Papers: Papers are not specifically solicited for this workshop, as it is a hands-on exercise. However, we will accept position papers (1-2 pages maximum) from participants who have some contribution in the following topics: the Chinese room experiment; the Turing test; hidden structures discoverable from statistical MT (HMMs, ATNs, etc.), Feng Shui, or related topics. Please submit these to the Conference Chair by August 15, 2007 via email; we will publish these for the participants, and invite a few of them to be briefly presented.

Chairs:

John S. White	SYSTRAN Software
Florence Reeder	The MITRE Corporation

Program Committee

Kathryn Taylor	US Government
Michelle Vanni	Army Research Laboratory

Contact:

John S. White
SYSTRAN Software Inc.
white@systransoft.com
+1 703-929-0433