

## CAT TOURNAMENT at TAC 2007

### ASSESSMENT PROCESS

**Public Release date: 2007-04-10**

#### **0. DOCUMENT META DATA**

This document is contained a file called:

CAT Document 003-v1 TAC-CAT Tournament 2007 Evaluation of Entrants.pdf

The document presents the process for assessment of entries to the CAT Tournament to be held as part of the Trading Agent Competition (TAC) at AAAI 2007 in Vancouver, Canada in July 2007. The process presented here will be tested during the CAT game trial in spring 2007, and may be modified in the light of the trial experience. Registered entrants will be informed of any modification to the assessment and to the game following the trial.

This document and its contents are copyright © 2007 by the MBC Project CAT Tournament team.

#### **1. ASSESSMENT PRINCIPLES**

The assessment system has been designed to meet the following desiderata:

- To be fair to all entrants
  - To be seen to be fair to all entrants
  - To be as realistic as possible (ie, indicative of real trading markets)
  - To reward innovative mechanism designs
  - To enable entrants be evaluated on an interim basis throughout the tournament (to encourage interest in the competition as it proceeds)
- and*
- To be not readily open to manipulation.

#### **2. OUTLINE**

The core idea is that entrants will be assessed on multiple criteria, which will be evaluated on a number of trading days. The assessment criteria to be used are described below in Section 4. In order to avoid effects arising from the fact that the Tournament has a start-day and an end-day, not all the trading days will be used for assessment purposes. The process by which days will be selected is described in Section 3 below.

### **3. SAMPLING METHOD**

**Step 1:** We choose a random starting day, and a random ending day. Both selections are made prior to the commencement of the game, and remain secret (from both game entrants and game organizers). The starting day will only be revealed at some point after it occurs. The ending day will only be revealed after completion of the game.

**Step 2:** Prior to game commencement, we also randomly choose days between the starting day and the ending day, on which assessment will be undertaken. These days are called "Assessment Days". These selections also remain secret, and are made public one-by-one, at the end of each such selected day. In order to avoid manipulation, revelation that the assessment process has commenced may not be made until after the first several Assessment Days have occurred.

**Step 3:** At the end of each Assessment Day following revelation that the assessment process has commenced, the scores of each specialist against each criterion, and their total score, will be made public.

### **4. CRITERIA**

It is planned to assess each specialist on each of the following criteria on each Assessment Day:

- Profits: absolute profits of the specialist on that day
- Market Share: the specialist's proportion of the total value of trades executed that day
- Transaction Success Rate: the proportion of bids/asks placed with the specialist which result in executed trades.

Criteria will be normalized and then weighted equally (i.e., one-third each) to produce a score for each specialist for the Assessment Day. Scores will then be summed across all Assessment Days to produce a final game score for each specialist. The specialist with the highest final game score will be declared the winner of the tournament.

### **5. PRESENTATION of RESULTS**

During the tournament, at the end of each trading day, we will present the values which each specialist achieves on each criterion, along with the cumulated values, and aggregated market statistics. In addition, on each Assessment Day, we will present separately the scores achieved by specialists. Thus, teams can monitor their performance (both absolutely and relatively) for each trading day, and for each Assessment Day. Revelation of the scores for Assessment Days will only be done once the fact that the starting day has passed is itself revealed.