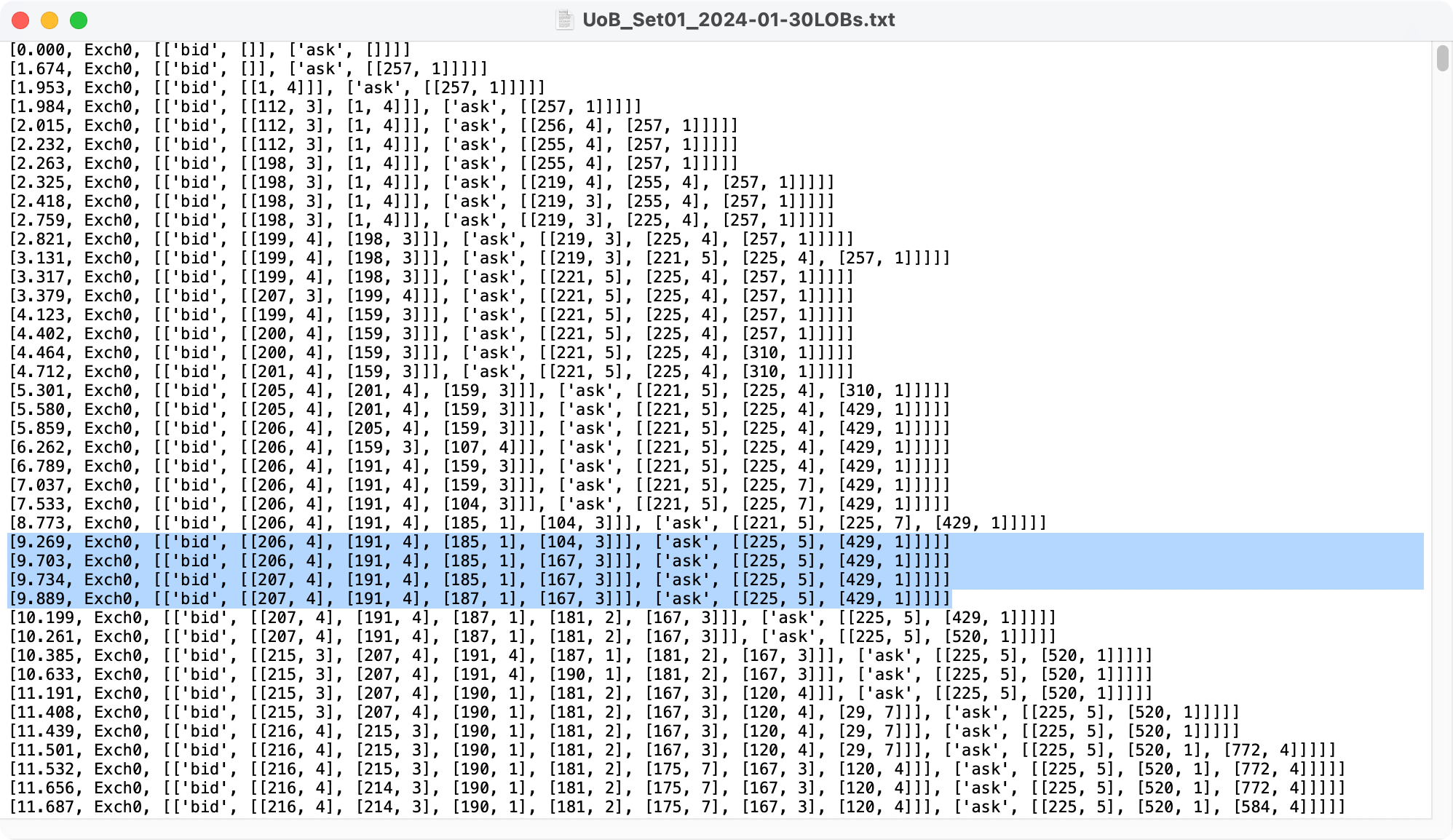
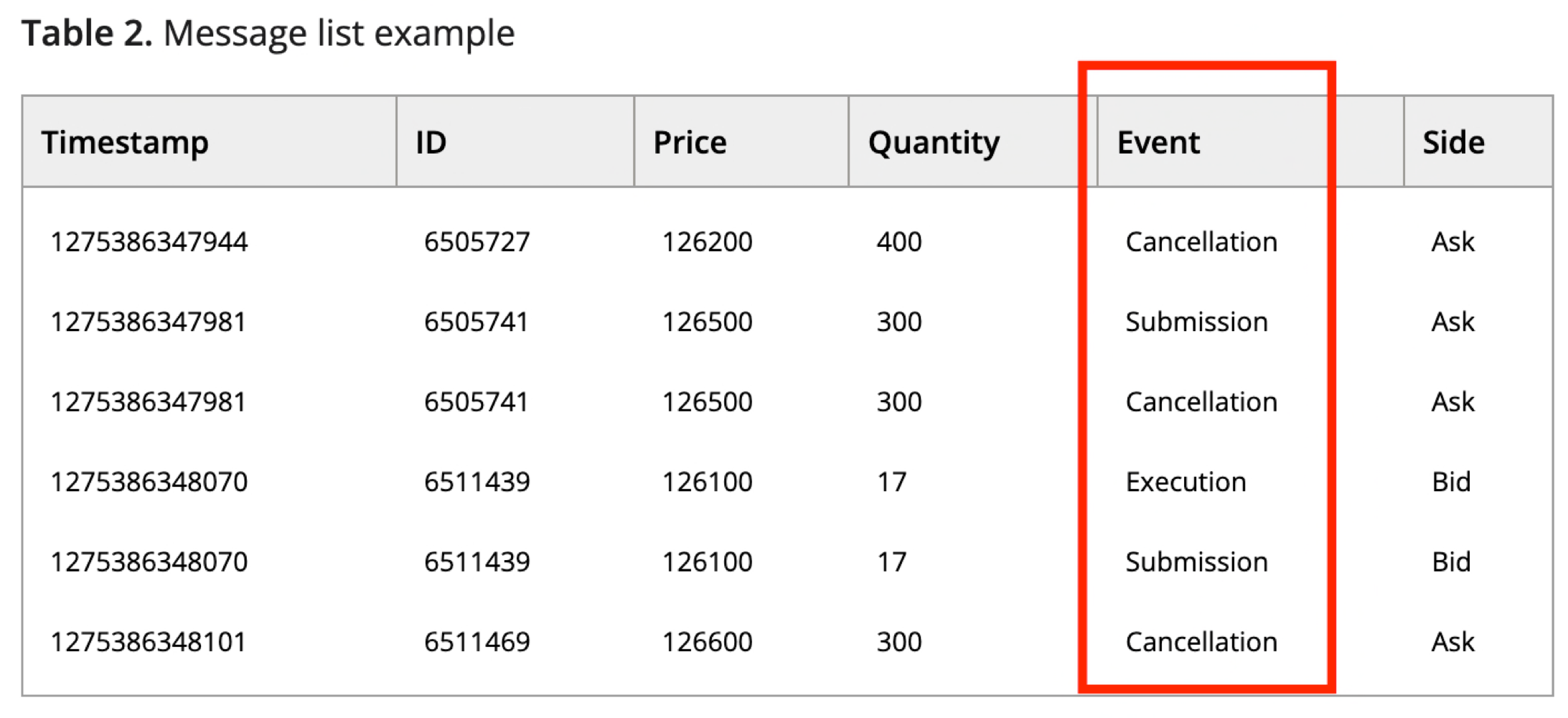
On last Thursday, we discussed about some other ideas about how to deal with the LOBs data:

1. Our data in **.txt files** are **duplicated:**



If we want to get the real data, we also need to remove the duplicated data, just remain only one row for each kind of data. Using this processed data, we could get the changes of ‘bid’ and ‘ask’. Also, we could go further to get the cancellation event using the processed data, see the next step.

1. Here is a link attached: <https://onlinelibrary.wiley.com/doi/full/10.1002/for.2543>
   1. In this link, we could see a table like this:



* 1. It shows that there are not only ‘bid’ and ‘ask’ states in the real situations, but also have an event at each timestamp, it could help us to classify and make predictions for the next timestamp, but the event is not provided for us in the LOBs data, so we should get it by ourselves.
  2. We could use the processed data last step, comparing the adjacent timestamps to see which item has disappeared, and comparing the disappeared item with that in Tapes data at the same timestamp. If the disappeared item is not in the Tapes data, it indicates the event at this timestamp is classified to **cancellation**.

That’s all we discussed on last Thursday, to be more specific, we need to search for more pappers about how to classify and make predictions. But it give us the clues about how to understand the column of **Event** in the relevant papers.