Tool instructions

1. Environment configuration

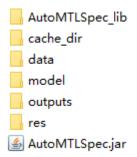
This tool uses Java version 1.8 and python version 3.7 and needs the support of GPU. All required external libraries of Java part have been configured (Located in AutoMTLSpec_lib). In addition, you need to add relevant external libraries required for running of Python files and the specific configuration is shown in the table below:

name	version
CUDA	11.4
scikit-learn	1.0.2
tensorflow	2.3.0
pandas	1.3.4
numpy	1.21.5
torchvision	0.12.0
keras	2.4.3
tqdm	4.64.0
simpletransformers	0.63.6

2. Tool usage

a) Preparation

The directory structure of tools is shown in the following figure:



First of all, because the trained model is too large, we do not save it

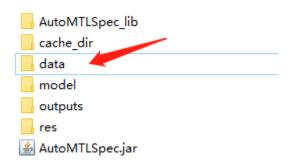
in GitHub, but in another place separately. The link is https://zenodo.org/record/7049127#.YxV_InbP0uU. Download it in advance and extract it directly into "model" folder. Finally, the directory structure in "model" should be as shown in the following figure.

occ
order
sc
task1
task1.py
task2.py
task3
task3.py
tt

Then, it is required to split the sentences in the original contract and put them into the excel table (the file suffix is .xlsx). The first line in the file is "Sentence", and the second line to the end is each sentence in the contract. The specific format is shown in the following figure.

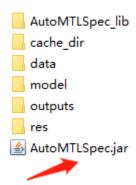
Sentence
THIS CONTRACT (this "Agreement"), is executed by and between the following parties on {date}:
{nameOf7}, a {classificationOf} state of {stateCitizenship}, with the principal address at
{addressOf:addr_line1} {addressOf:addr_line2} {addressOf:city} {addressOf:state} {addressOf:postal},
(hereinafter known as the "Buyer");
{nameOf8}, a {classificationOf10} state of {stateCitizenship29}, with the principal address at
{addressOf12:addr_line1} {addressOf12:addr_line2} {addressOf12:city} {addressOf12:state}
{addressOf12:postal}, (hereinafter referred to as "Seller");
WHEREAS, {descriptionOf} of {descriptionOf15};

After the file is ready, it needs to be put into "data" folder (AutoMTLSpec/data).



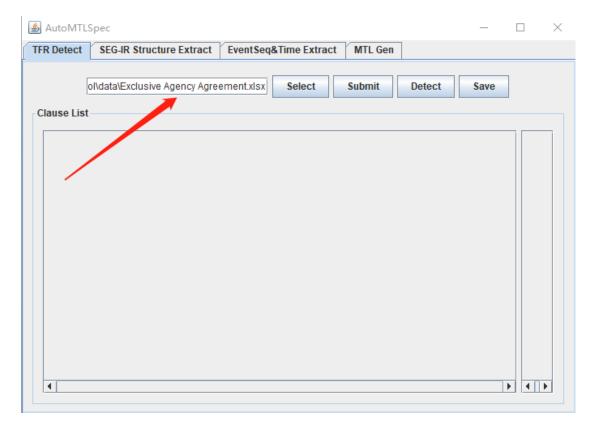
b) Open tool

Click the. Jar file in the folder "AutoMTLSpec" to run.



c) TFR Detection

Step 1: Click the "Select" button to select a file. After selection, the file path will be displayed in the list bar. Input is the file located in "data" folder.

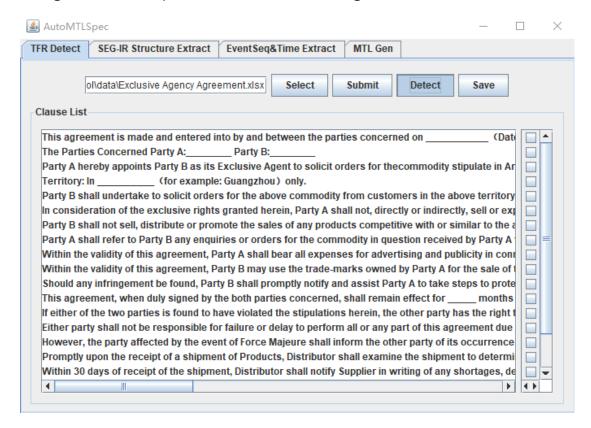


Step 2: Click the "Submit" button to submit the file. After submitting, the clause list will be displayed.

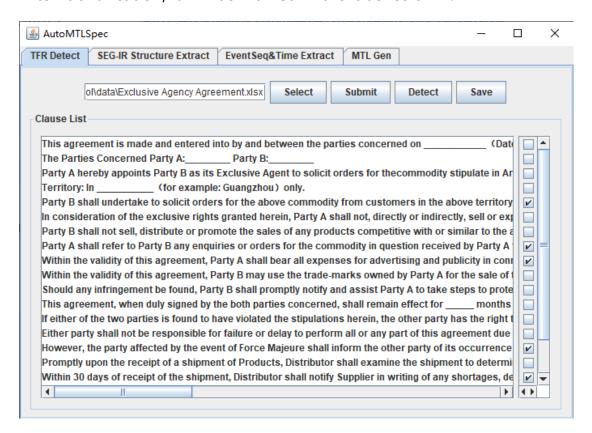


Step 3: Click the "Detect" button to detect TFR from numerous legal

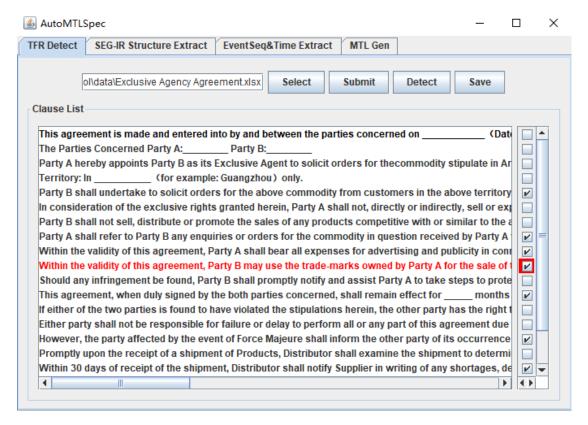
clauses. During recognition, the detect button will be clicked until the recognition is completed, as shown in the figure below.



After identification, it will be marked in the label column.



Step 4: You can manually modify the label in the tag bar by clicking the selection box to select or deselect the corresponding requirement. In addition, in order to facilitate users to correct errors, clicking on the terms in the list will highlight the corresponding labels in the corresponding positions.



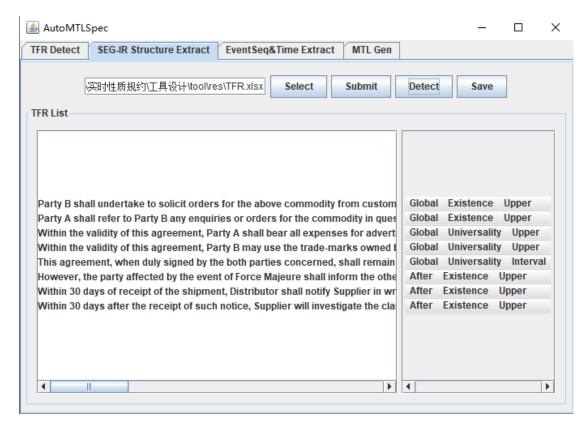
Step 5: Click "Save" button to save the results. The format of the saved file is the same as the input file, named TFR.xlsx (Located in "res"), and the content is the TFR list.

d) **SEG-IR Structure Extraction**

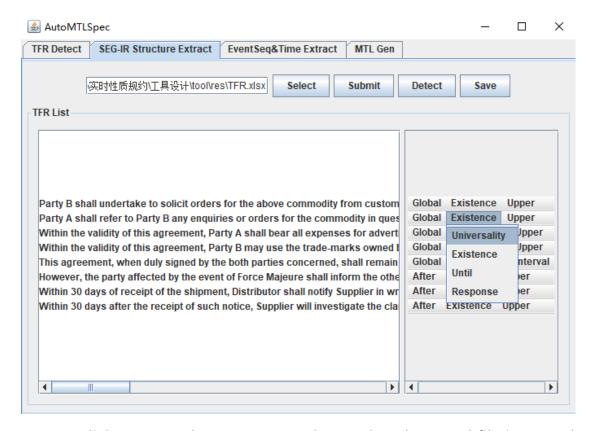
Step 1&Step 2: The operation is similar to that described above. Input is TFR.xlsx.

Step 3: Click the "Detect" button. After clicking, the demand structure classification will be displayed in the label bar of TFR list, including the

classification of Scope, Order and Occurrence, and Time Type.



Step 4: You can manually modify the classification mark in the label bar by clicking the label to be changed and selecting from the displayed drop-down list.

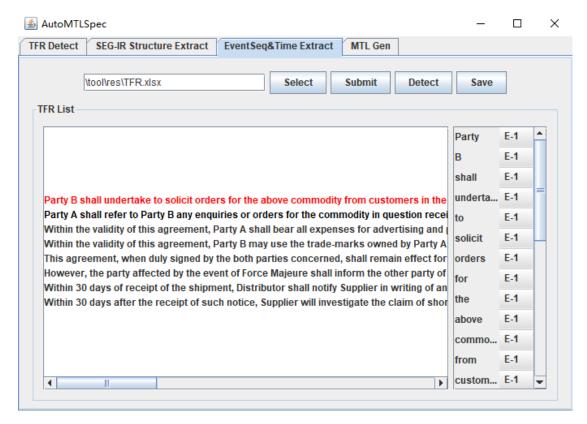


Step 5: Click "Save" button to save the results. The saved file is named Structure.xlsx (Located in "res"), and the content is the SEG-IR structure.

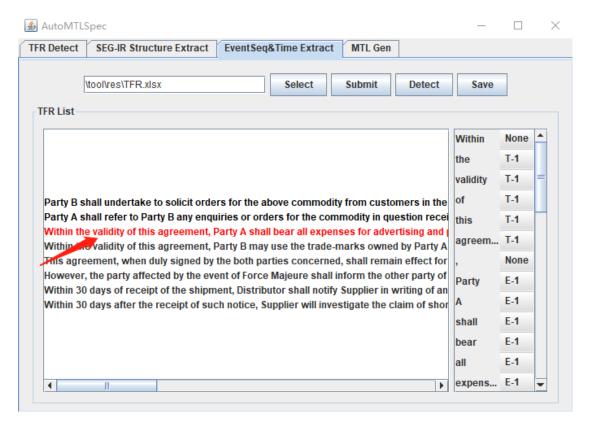
e) Event Sequence and Time Point Extraction

Step 1&Step 2: The operation is similar to that described above. Input is TFR.xlsx.

Step 3: Click the "Detect" button. After clicking, The word sequence of the first TFR and its corresponding tag sequence are displayed in the tag column of TFR list.

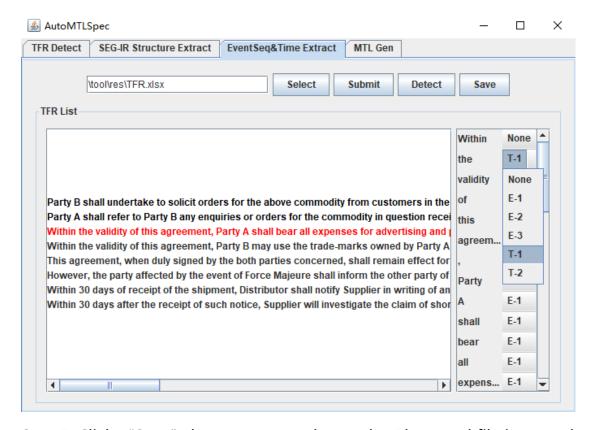


Clicking on different TFR will switch to the word sequence and tag sequence corresponding to the TFR.



Step 4: You can manually modify the tag sequence in the tag bar by

clicking the corresponding drop-down list in the tag sequence.



Step 5: Click "Save" button to save the results. The saved file is named et.xlsx (Located in "res"), and the content is the Event Sequence and Time Point.

f) MTL Generation

Step 1&Step 2: The operation is similar to that described above. Input are Structure.xlsx and et.xlsx.

Step 3: Click "Gen" button to generate MTLs. At the same time, when the time point contains numbers, the numbers can be extracted and displayed in the MTL. If there are no numbers, 'x 'and' y ' will be displayed at the corresponding positions in the MTL instead of the time point.

