Form Based Pattern

Abstract algorithms for Detection and Repairing

According to the research papers:

*“regular occurrences of groups of events that share the* ***same timestamp value*** *is a good indication of the presence of the ‘Form-based Event Capture’ pattern”*

Furthermore,

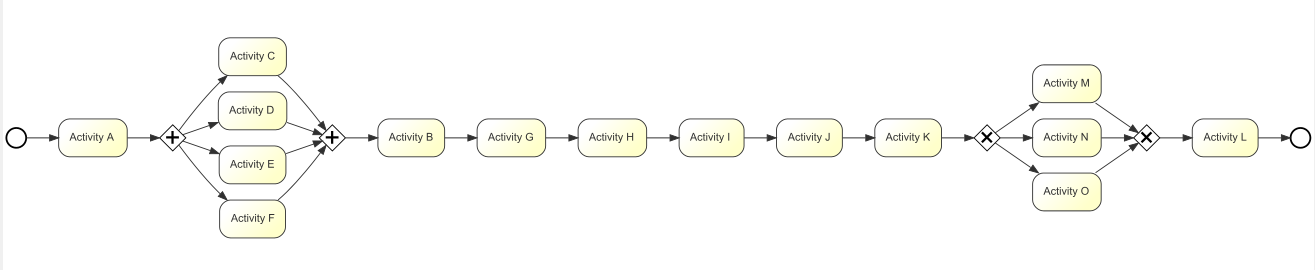
*“log can be searched for the presence of ‘marker’ events with activity names similar to field labels known to exist on the same form”*

This can only be applied if such information about the forms can be obtained from system users.

Aggregating the above information, an abstract algorithm should consider the following key points.

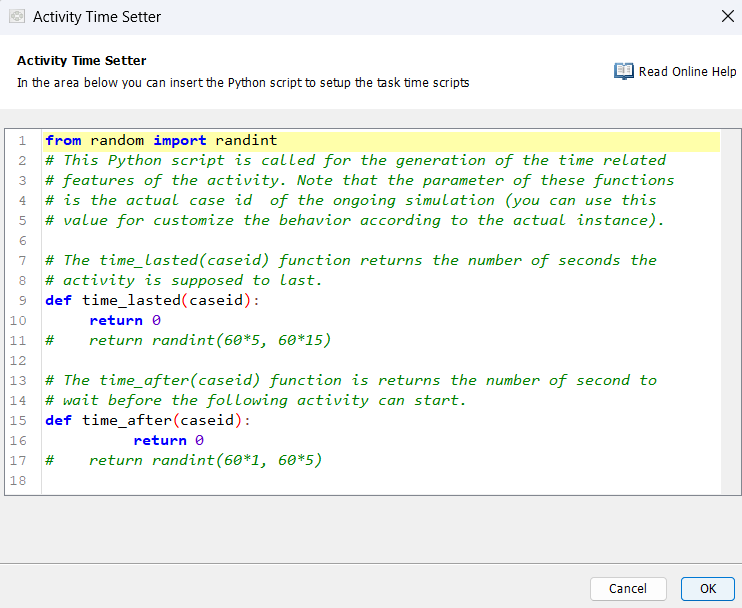
1. Events which have the same timestamps and case\_id values are grouped together.

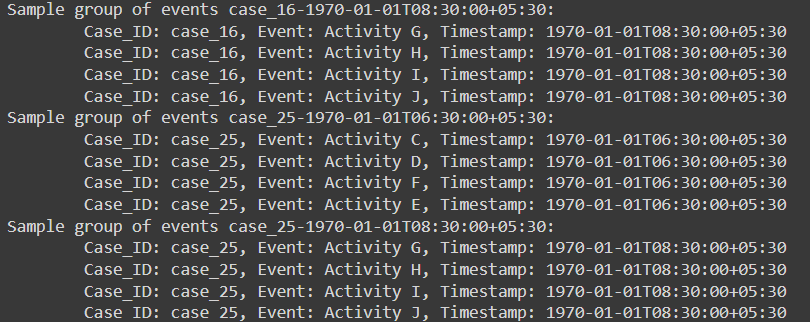
Tested by using a synthetic event log generated using PLG.



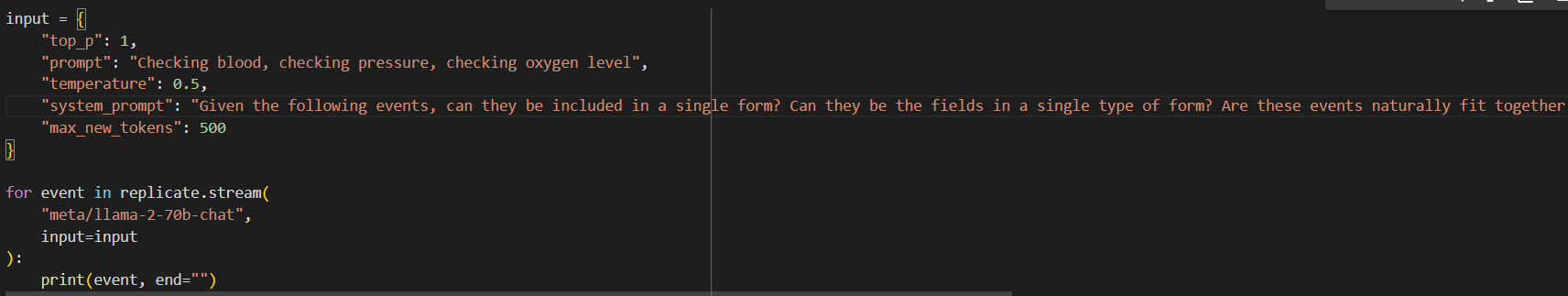
Duration for activity G, H, I and J has been set to 0, to create a form based pattern.

Time setting of activity G





1. A method to identify whether the grouped events are form based (can be aggregated as a set of form based events) or not.
   1. If details about ‘marker’ events can be obtained from the admin system, the grouped events can easily be determined whether or not form based.
   2. If there are no such details to be obtained, we have to use a ML model with an intelligence to recognise whether the events can be in a form (can be aggregated into one event)





This does not work correctly all the time. So having a set of marker events for forms can be very helpful to identify the pattern.

If not, we have to make an assumption that events belonging to the same case and having the same timestamps, have a form based imperfection pattern.

1. Handling updates:

There can be two scenarios happening when an update occurs.

1. Updates of one or more data items in a form result in the recording of all fields as events in the log - should specify the information which was changed; can use an extra attribute
2. Updates of one or more fields in a form will result in the recording of only those fields whose values have changed

Few ways to identify an update, with respect to a new form creation:

1. Observing the events which have taken place prior to the update/create event - identifying events which need to be done before create event and update event. **This method is not effective in most cases**
2. Observing other attributes: in some event logs, an additional attribute is included where information about events are recorded. An update event can be recorded with information regarding the update.

**This method can only be applied to a few event logs.**

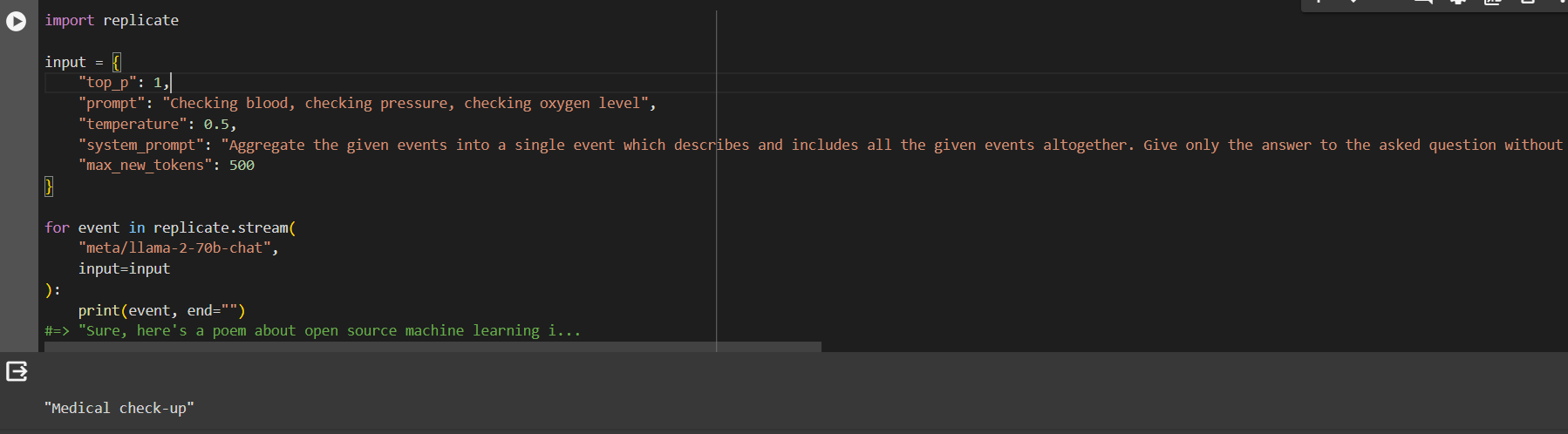
**Remedy**

Remedy for this pattern is to aggregate the form based events into a single event. (or optionally a few lesser number of events)

We can use a ML model with a general intelligence for this task.

For the initial testing, we have used the LLAMA2 model and the results were satisfactory.

System Prompt: ”Aggregate the given events into a single event which describes and includes all the given events altogether. Give only the answer to the asked question without any initial words such as 'sure, here's your answer'etc. Answer should be correct, relatable, short, descriptiv. Word count of the answer should be less than 5 words”



The groups with the same events (same form fields) should be aggregated into the same single event after repair.

Therefore, after detecting the groups with same timestamps and case\_id values, we should distinguish groups with the same events and group those groups together.

We can use a simple string manipulation library to identify the same event groups and replace them with the same aggregated event name.

There can be issues when choosing whether all the events to be aggregated into a single event OR if there are some events we should not aggregate with other events and keep separately even after repairing.  
This issue arises due to the unknown and varying criteria to determine the importance of events, which need to be aggregated together or not.