

## Sequence mining to analyse learner's action patterns in a robot mediated human-human collaborative activity

Student name: El Mekki Malek

Supervisor: Jauwairia Nasir, Aditi Kothiyal, Utku Norman

### MOTIVATION

Our current analysis is built upon previous analysis conducted in the lab related to the JUSThink activity in which it was shown that data-driven clustering approaches based on **behavioral features** such as interaction with the activity, speech, affective features, and gaze patterns can **discriminate between gainers and non-gainers**. The clustering process identified **three separate clusters (Type 1 gainers, Type 2 gainers and non-gainers)**

Looking for statistically significant activity patterns in the lower level sequential log data can be interesting as it can:

- Suggest the most common sequence patterns among each type of learner groups.
- Allow us to understand how students that learn differ in their action patterns from those who do not end up learning
- Help us to observe if the similarities and differences found at temporal level with action patterns are consistent with the previous  
So Sequence mining and differential sequential mining will allow to have comparative profiles over the whole interaction and over phases of interaction .

### METHODS

Once the data is prepared it is used as input to a Sequence Mining which is Prefix-Span algorithm

The main parameter of the algorithm is the support threshold, and in our case we set it as 0.25 and it gives all the patterns that occurred in at least 0.25 of the total number of teams. Then :

- We compute the frequency of each pattern in that group: it is the number of times the pattern occurs, without overlap, in a sequence
- We do that for all the sequences of the two type of gainers and for all the teams present within each type. So for each pattern, we will get two vector of frequencies.
- Normalize the two vectors by the time taken by each team to finish the activity.
- we filter the patterns based on the p value of a t-test with 95% confidence comparing pattern frequencies between the types of gainers
- comparing the difference between the mean of each vector and depending on its sign, we know in which type of gainers the pattern was more frequent.

### RESULTS

- Types of exploration, confusion and communication problems were present in the 3 types of learners but at different stages, with different support and different frequency
- Looking at temporal sequences allows to observe the learning process deeply => will further assist in designing better robot interventions.