Competition analysis

The competition is separated in 2 stages. Stage 1 is composed of 5 parts while stage 2 got only 4. At least the majority of the parts need to call the speech recognition capacity of the robot. Questions can be different but have some rules :

- It can be general questions about the robot
- It can be general questions about the robot team
- It can be general questions about the arena
- It can be general questions about the competition
- It can be questions about the crowd
- It can be questions about the day, weather etc…
- Some questions may refer to previous questions or answers
- In all the parts except the blind man's bluff game the operator will be standing in front of the robot
- Operators can't move when they are in a dialog with a robot
- Gestures and facial expression need to be define by the team

Capacities that are require :

- Remember at least the 5 previous questions and answers in a dialogue
- Recognize and remember people names
- The robot have to get some knowledge about the competition (history, location etc…) and his team (name, number of person etc…)
- It need to know when to call others capacities (objects recognition for example) to be able to answer questions (ex : what color is that pen ?)
- It need to be able to know where the person talking to him is located
- When it can't answer to a question it have to answer something else to show that is understood the question but don't get the answer (after 10s with no answer the question is considered as missed)
- In the riddle game the robot is not allowed to ask repetition for a question
- All interactions with the operator is well considered in the scores (gesture, facial expressions…)
- If the robot doesn't understand a sentence it is allowed to instruct the operator (ex : come closer and repeat etc…)
- Set gestures and facial expression in accord to the speech
- More : get the emotion of operator ? Interact with emotion (for example emoji in the tablet) ?

Ideas :

- MRN (multimodal residual networks) to match vision and speech learnings (2018 AUPAIR team)
- Adaptation of context-dependent : reduce words recognition errors
- Use proxemics to improve interactions (2017 SPQReL team) => part of sound spatialisation
- Get the emotion of operator ("Speech emotion recognition using hidden Markov models", Tin Lay New, University of Singapore)
- Use 2 microphones for spatialisation and 2 for the speech recognition
- Classify sounds by categories to match them with visual environment ("Deep Learning and Bayesian Networks for Labelling User Activity Context Through Acoustic Signals", Francisco J. Rodríguez Lera)