Human-robot harvesting plan negotiation: perception, grape mapping and shared planning

Marc Dalmasso
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Universitat Politècnica de Catalunya - BarcelonaTech (UPC)
Barcelona, Spain
marc.dalmasso@upc.edu

Francisco Garcia-Ruiz
Department of Agri-Food Engineering and Biotechnology
Universitat Politècnica de Catalunya - BarcelonaTech (UPC)
Castelldefels, Spain
fco.jose.garcia@upc.edu

Thomas A. Ciarfuglia
Department. of Computer, Control and Management Engineering (DIAG)
Sapienza University of Rome
Rome, Italy
ciarfuglia@diag.uniroma1.it

Leonardo Saraceni
Department of Computer, Control and Management Engineering (DIAG)
Sapienza University of Rome
Rome, Italy
saraceni@diag.uniroma1.it

Ionut M. Motoi
Department of Computer, Control and Management Engineering (DIAG)
Sapienza University of Rome
Rome, Italy
motoi@diag.uniroma1.it

Emilio Gil
Department of Agri-Food Engineering and Biotechnology
Universitat Politècnica de Catalunya - BarcelonaTech (UPC)
Castelldefels, Spain
emilio.gil@upc.edu

Alberto Sanfeliu
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Universitat Politècnica de Catalunya - BarcelonaTech (UPC)
Barcelona, Spain
alberto.sanfeliu@upc.edu

***Abstract*—In this paper we present an application of plan negotiation for distributing work in Human-Robot Teams (HRT), engaging in a shared plan generation process. This model has been experimented over data collected from a real-life table grape vineyard field in Corsira (Aprilia, Italy) on the context of the EU CANOPIES project. In this work we describe a complete system up to negotiation agreement, the perception system to detect the grapes position and quality, the grape map generation process and the proposed plan negotiation approach, and protocol used to interactively generate harvest plans. Our proposed approach tackles harvest plan generation as a two step negotiation process on negotiation domains of increasing embedded information. The shared planning process was evaluated using both a subjective user study and teammate quantitative models.**

***Index Terms*—plan negotiation, grape harvesting, human-robot collaboration**

I. Introduction

Modern agriculture faces the challenge of meeting the increasing global food demands while minimising resource consumption and environmental impact. This is particularly important in fruit production given the significant labour shortages that the sector has been facing in recent years which has become a critical problem for farm managers [1]. Specially harvest and pruning activities imply a high demand for human labour, becoming the most expensive management activities in fruit agroindustries [2] [3]. In response, technological advances in crop management have emerged as a transformative approach, being the partial or complete automation of these tasks becomes a game-changer to the economic sustainability of farms.

Collaborative teams of humans and robots appear as an interesting solution to overcome labour shortages, overcome failures or uncertainties in the some of the harvesting sub tasks (e.g., failure in quality grape detection) or increase efficiency of the harvesting processes and generate traceability of the harvested product based on objective data collected by the perception systems embedded in the robot.

As of now, the process of grape harvesting presents many challenges as information reliability or occlusions, which are hard to tackle with an automatic system. In this work we focus on how the negotiation between a farmer and a robot have to be done to distribute the harvesting task plan and what will be the recommended path trajectories for each of them.

In this article, we present a collaborative harvesting operation of a human-robot dyad, where the final goal is construct and agree on a harvest path plan. In the first part of this work, we explain how a grape map of the vineyard is build, including spatial and quality information of the grape bunches. In the second part, we tackle harvest planning as a human-robot team shared planning activity, constructing plans through a two step negotiation process. Finally, in the third part, we present the followed experimental methodology for the negotiation process, discuss the human preferences distribution over the

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016906 – Project CANOPIES

Fig. 1: **Field Scenario.** Images of the robot and the field where the grape images where captured. These images were shown to the participants used to introduce them to the presented contexts.

shared planning process and the obtain results.

In Section II we first will make a short review of the related work. In Section III, we explain how the detection and localization of grapes, quality detection and mapping processes are performed. In Section IV, we explain the task representation and planning methodologies used by the robot to generate plan proposals, the negotiation domains over which the HRT has to reach an agreement and the negotiation protocol applied on each of them. In Section V, we describe the negotiation experiment methodology, the environmental domain, the participants' selection policy and demography, the experiment design and the results. Finally Section VI contains conclusions extracted from the work.

II. Related work

Detection and localization of grape bunches have been done by [4] and [5] using neural networks. Automatic harvesting of grape bunches using a dual-arm robot has been performed for horizontal trellis cultivation [6] and for canopies grape bunches [4].

On the explainability field, previous studies considered including information exchange with humans in the planning process. Some have included verbal communication in task planning [7]–[9], planned requests to human managers to make usage of a resource [10] or asked for assistance when needed [11]. Alternativelly, [12] includes human in the planning process through the usage of mixed reality for intuitive drone navigation path planning and visualization.

Lately, more attention has been brought into the human agreement on plans and the nature of making plan proposals. Recently, both Chackaborti et al. [13] and Porteous et al. [14] introduced visualization tools for interactive plan visualization and selection and Porfirio et al. [15] proposed a goal-oriented planning system to balance control of non-expert users in the robot planning process. A previous study dynamically adapted to new human preferences to generate plan proposals for navigation tasks [16], Moon et al. [17] explored the usage of hesitation trajectories as implicit proposals to enable human-robot non-verbal negotiation and in [18] human preferences over different navigation plan negotiation domains where explored for different human-robot team contexts.

In this work, we are exploring the usage of both assignation and navigation plan negotiations in human-robot teams in the context of grape harvesting to construct team-generated harvest plans.

III. Building the task model

In this section we describe the basic operations to detect and localize the grapes (or bunches), to detect the quality of the grapes and to do the mapping operation to create the map of bunches.

A. Detection and localization of the grapes

The detection and localization of the grapes is done by a farming robot (Fig. 1) developed in the EU CANOPIES project [19], which has three different Realsense D435i RGB-D cameras located in the head and the wrists of the robot, to detect the grapes, compute the quality of the grapes and localize the grapes and their peduncles. The detection and localization of the grapes is based in method explained in detail in the article [4], which consists in two parts: (1) detection of the grapes and peduncles; (2) computation of the grape and peduncle localization. The detection of the grapes is performed by a method that combines monocular depth [20] and Mask Region-based CNN [21] methods. The computation of the localization of the peduncles is done by combining two methods: (1) depth estimation of peduncle using the depth map; (2) direct measurement method. Both methods are fused to obtain better results. The resulting depth estimation achieves a mean squared error of 2.66 cm within a distance of 1 m in the CANOPIES dataset.

In the right bottom Fig. 2, it can be seen the detection of grapes and peduncles with a score above 0.9. Since the grape detection method computes relative distances using monocular depth, the grapes that are further a specific distance of the others are not considered for detection, that means only grapes closer to the robot arm are taking into account. The detection method can detect several grapes although they are partially overlapped, and in this case, only the area of the seen bunches can be computed, and the border between one bunch and the other is separated by a straight vertical line. In some of the bunches, several peduncles can be detected, and in this case, the point to cut it is the closer to the vineyard cane from where the peduncle is connected.

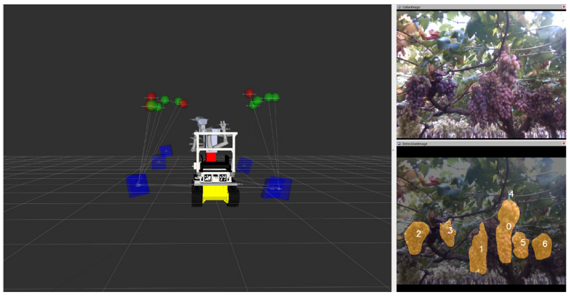

Fig. 2: **Grape Bunches Detection.** On the left, virtual robot model and detected bunches' localization. On the right top, images of some bunches taken by the robot's head camera. On the right bottom, detection of bunches of the top right image

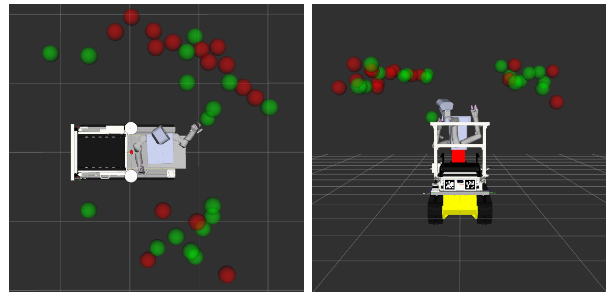

Fig. 4: **Grape Mapping.** Bird's eye and lateral views of the robot and the detected bunches from a particular viewpoint. The robot stays stationary when scanning for grapes. The quality of the grapes is represented by color, green for good bunches and red for poor quality bunches.

Fig. 3: **Feature design.** Grape bunches are cropped after the detector bounding box. The resulting image is then binned and some bins are dropped, as with high probability contain background information.

B. Quality detection in grapes

The Quality Estimation Module (QEM) in the robot perception systems assesses bunch size, color class, and Soluble Solid Content (SSC), the latter indicative of sugar level in Brix°. In the EU CANOPIES project, dedicated estimators were developed for these characteristics [22]. Except for size, which utilizes stereo-cameras and geometry, estimators rely on data-driven methods. Integrated within the robot's perception as the QEM ROS service, they evaluate readiness for harvest based on defined thresholds. System effectiveness is documented in [19], with the following sections detailing each estimator's operational principles.

1) Size Estimation: The estimator employs the RealSense D435i depth camera for object metric measurements, achieving millimeter accuracy through scene pixel triangulation, based on the pinhole camera model.

The model estimates grape bunch sizes by bounding boxes, avoiding inaccuracies from edge fitting and undefined depth map pixels. Depth sampling at the object's center enhances measurement reliability, with size estimation errors under 11%, typically less than 1.5cm, and a 20cm height threshold for harvest readiness.

2) Color and SSC estimation: The estimations for color and Soluble Solid Content (SSC) leverage the correlation between the fruit's external characteristics and its ripeness. This system balances computational efficiency and estimation accuracy, employing histogram features to conserve computational resources for other robotic functions. Ridge Regression was utilized for SSC, while color estimation applied Logistic Regression with L2 regularization.

The dataset was collected in Aprilia, Italy during Summer 2021 (UTM: 33T, 310858.24, 4597049.60). It comprises 212 images, split into 178 for training and the remainder for validation. SSC labeling used an ATAGO Hikari PAL2 refractometer, assessing Brix° without berry damage. Brix° sampling adhered to agronomic standards, measuring top, middle, and bottom berries per bunch, averaging for the bunch's nominal Brix°. Color was categorized chromatically across six quality levels, with 1 indicating the highest quality.

A "cross" shaped pattern was employed to eliminate background pixels (Fig. 3). Key hyper-parameters for the final estimator, besides the regularization parameter for ridge regression, included the initial image size, grid dimensions , selective cell drops using a "cross pattern," and preference for HSV over RGB color channels. Hyper-parameters for both tasks were optimized using 5-fold cross-validation, yielding a Mean Absolute Error (MAE) of 1.45° for Brix° estimation and 79% balanced accuracy for color classification.

C. Mapping operation

The models presented in sections III-A and III-B were used in a section of the field in September 2023 to create a map of the bunches and store their estimated quality. The localization of the bunches and their peduncles are obtained in the bunch detection process, and their position in global reference is computed using a combination of RTK-GNSS system, robot odometry and 3D Lidar information (Fig. 4). Though grape bunches' position can vary during harvesting, this localization can be used for planning and as an estimate to approach the grape and relaunch the visual detection module immediately before the harvest operation.

In the process of gathering the pose of the bunches, sometimes the robot captures the same bunch in more than one image and in this case we aggregate the different poses, computing a pose average of the bunch. Finally, for each

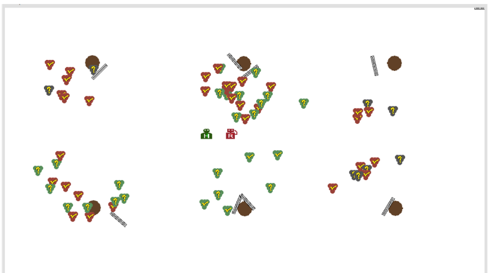

Fig. 5: **Grape Map.** Bird's eye visualization of the grape map used in the negotiation interface (graphical legend in Fig. 7).

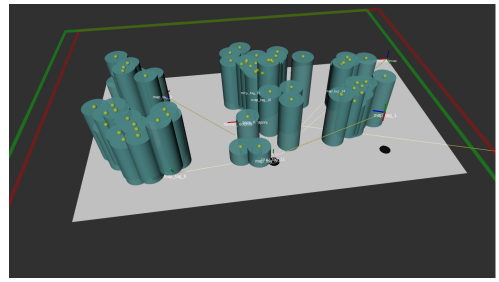

Fig. 6: **Task Representation**. Social Reward Source (SRS) representation of the harvesting task. Each grape bunch harvest has an associated cylindrical source of reward of radius equal to the maximum harvest distance. Initially, sources' targets are $\tau = \{h, r\}$ for those with a positive quality assessment and $\tau = \{h\}$ for those with a negative one. Once the team has an assignation agreement, source targets are defined by the assignation.

one of the bunches, we include the quality of the bunch, good or bad bunch, obtained by the quality detection module explained in III-B. Fig. 4 shows an example of the bird's eye and lateral views of the bunches and Fig. 5 shows the bird's view visualization used in the experiment interface.

IV. SHARED PLANNING

We approach harvest planning as a human-robot team (HRT) shared planning activity of a collaborative navigation task. To do so, the HRT engages in a plan negotiation where the subject of agreement is the final plan. Following the possible planning domains introduced in [18], we define two plan negotiations: assignation plan negotiation and navigation plan negotiation. Both are explored in the presented experiments, which applied the shared planning process as a two step process.

The first step is a negotiation where the team work distribution is agreed upon. This negotiation works by exchangin grape bunches' harvest assignation proposals. Let proposal ω be an assignation plan $\Gamma = \{\Gamma_1^{\tau_1}, \Gamma_2^{\tau_2}, ..., \Gamma_N^{\tau_N}\}$ where $\Gamma_j^{\tau_j} = \{\gamma_{i_j(1)}^{\tau_j}, ..., \gamma_{i_j(n_j)}^{\tau_j}\}$, $\Gamma_j = \{\gamma_{i_j(1)}, ..., \gamma_{i_j(n_j)}\}$, $i_j(k_j) \in \{1...n\}$ $\forall k_j \in \{1...n_j\}$ and:

$$\tau_j \subseteq \tau_i \quad \forall i \in \{i | \gamma_i \in \Gamma_j\} \tag{1}$$

$$\bigcup_{j \in N} \Gamma_j = T, \qquad \bigcap_{j \in N} \Gamma_j = \emptyset \tag{2}$$

The second step is a negotiation where teammates actions are defined. In this negotiation navigation proposals fulfilling the agreed upon assignation are exchanged and the agreed assignation may be revised if needed. Let proposal ω be a team navigation plan $x \in \mathcal{X}$ constructed by the agents' movements $x = \{x^{a_1}, ..., x^{a_m}\}$, where $\mathcal{X}$ denotes the set of feasible team plans. Then, each agent movement $x^a \in \mathcal{X}^a$ can be defined as an ordered sequence of basic movement actions $x^a = \{x_1^a, x_2^a, ..., x_k^a\}$, being $\mathcal{X}^a$ the possible action sequences of agent a. Basic movement actions x_i^a are defined by their goal, encoding the action of moving to it. Each x_i^a may have an associated finish time t_i^a only if it has been specified and agreed upon by the team.

A. Robot Plan Generation

To engage in these negotiations, the robot should be able to generate the team path plans, one for the robot and the other for the human. Our approach to do so has been to make use of the Social Reward Sources (SRS) world representation and a Monte Carlo Tree Search (MCTS) algorithm to explore it, both previously published in [16].

The harvest task has been represented as a set of consumable reward sources $\Psi_h = \{\psi_{h_1}, \psi_{h_2}...\}$, each related to the harvest of one of the grape bunches detected in the considered area. Each source ψ_{h_i} is defined by a cylindrical boundary of radius equal to the maximum harvest distance, a set of targets defined by the bunch quality, a null external reward function $r_e = 0$ and an internal reward function proportional to the probability of success in the harvest $r_i \propto p_h(a, d)$. For simplicity, in these experiments the harvest probability was set to one for all actors in any reaching positions, so $p_h(a, d) = 1 \quad \forall a, d$. Thus, for the presented experiments the source s' reward is in practice a step shaped function. You can observe a graphical depiction of the task representation in Fig. 6.

Additionally, the SRS world representation also includes a set of Gaussian sources linked to the trees' trunks, Gaussian reward functions representing the collision danger of robot proximity to these obstacles. Similarly, movement cost reward sources where defined for each agent.

On the robot side, the MCTS algorithm constructs team navigation plans to tackle the collaborative task. Such plans are taken at face value when constructing a navigation proposal or abstracted through a task contribution computation to generate assignation proposals.

B. Negotiation Protocol

These experiments study one-human-one-robot teams. In both negotiations, the robot initiates the interaction and the team members have asymmetrical roles.

The first negotiation is initiated by a robot assignation proposal. Upon receiving a proposal, the human participant can either accept it, request for a different assignation plan or draw themselves a new assignation proposal. If being asked for a different plan the robot will either propose another

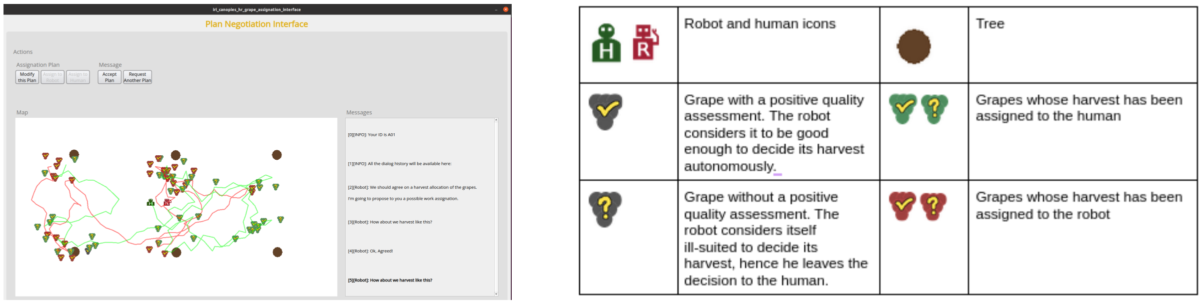

Fig. 7: **Interface**. From left to right: a) User interface used for the negotiaiton process and b) legend presented to the participants to ensure their comprehension of the interface symbolic codes.

of its already computed plans or generate a new one from scratch. This process can be repeated until satisfaction (the human accepts the plan) or proactive action (the human decides to make a proposal). The robot will always accept the participant's proposal, under the assumption their selection are prioritised in this setting, but participants where not make aware of this beforehand. Either way, the final assignation plan is considered as agreed upon be the whole team.

Similarly, the second negotiation initiates through a robot proposal, this time being a navigation proposal. The navigation proposal is constructed over the agreed assignation, reducing the exponentially large search space. As before, upon receiving a proposal, the human participant may either accept it, request for another navigation plan or modify it themselves. This time, however, the human will still be only given the possibility to construct assignation plans. As before, the robot will agree to the new assignation and make a new navigation plan proposal abiding by it. This second negotiation will continue until the human participant accepts one of the navigation plans proposed by the robot.

V. Negotiation Experiments

In this Section, the reader may find an explanation of the experiment domain, the experiment design and a discussion on the experiment results.

A. Environment Domain

The negotiation experiments have been conducted on a remote fashion. The selected participants were introduced to the harvest context, robot and field characteristics through the images of Fig. 1 and, afterwards, asked to interact with the robot as if being on the field. They received an explanation on the interface usage (Fig. 7) and told that they would interact with a robot to negotiate how the harvesting has to be handled.

Data used in the experiments was composed by the grape map constructed in the field, though the quality assessment of the grapes was modified by hand to enrich the environment. The interface used in this experiments allow the participants to perform the actions listed in the protocol. Both assignation and navigation proposals from the robot are depicted on the map image of the field (see Fig. 7) . Human assignation proposals are also drawn in the same fashion through the colouring of the bunches. Acceptance of a proposal and requests for new proposals are both handled through buttons situated at the upper left of the screen.

B. Participants

The target population of the experiment consisted of students and professors of the Degree in Agronomic Sciences of the School of Agri-Food and Biosystems Engineering of Barcelona (EEABB) of the UPC-BarcelonaTech, located in Castelldefels (Spain). In this way, sufficient knowledge of agronomic and fruit production methods was guaranteed, while at the same time, basic knowledge of engineering was also ensured. From this population, a total of 21 participants were recruited for the experiment based on their availability and knowledge, avoiding students from the 1st and 2nd year of the degree. Demographic data of the participants can be observed in Fig. 8.

C. Experiment Design

Participants were divided into three groups corresponding to each of the detection confidence suppositions (case A, B, and C) for the quality assessment classification algorithm (Table I). Each participant received the instructions to conduct the experiment and a clear explanation of the functionalities available in the interface for human-robot negotiation. Furthermore, verbal instructions were given during the experiment and a short explanation was provided by the leading researcher in every section of the interview.

1) Methodology: Each participant was set in the position of a farm operator in charge of harvesting a table grape field while assisted by a robot. The participant answered the first part of the interview consisting of the demographic information, and subsequently started the experiment using the human-robot negotiation interface.

After the robot proposed the first fruit assignation plan, the participant rated the plan based on four characteristics that were verbally explained by the leading researcher:

- Robot leadership: is the robot leading the situation when proposing the harvest plan?
- Robot intelligence: is the plan coherent and proposed based on some kind of logic, or does it seem random?
- Robot control: is the robot taking control of the situation, or is it relying on what the human has to say?

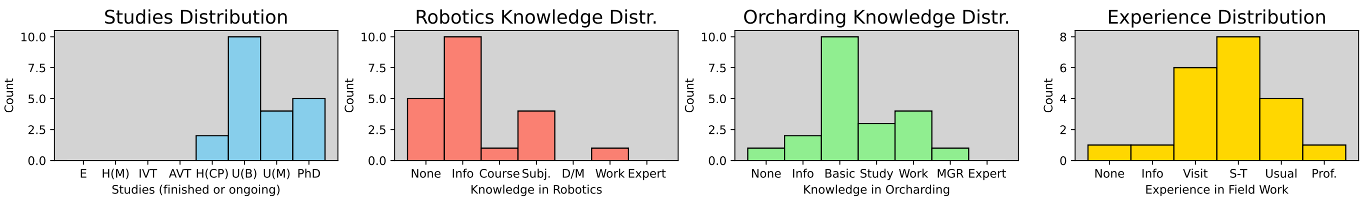

Fig. 8: **Demography**. Participants' demography distribution over general and study related knowledge

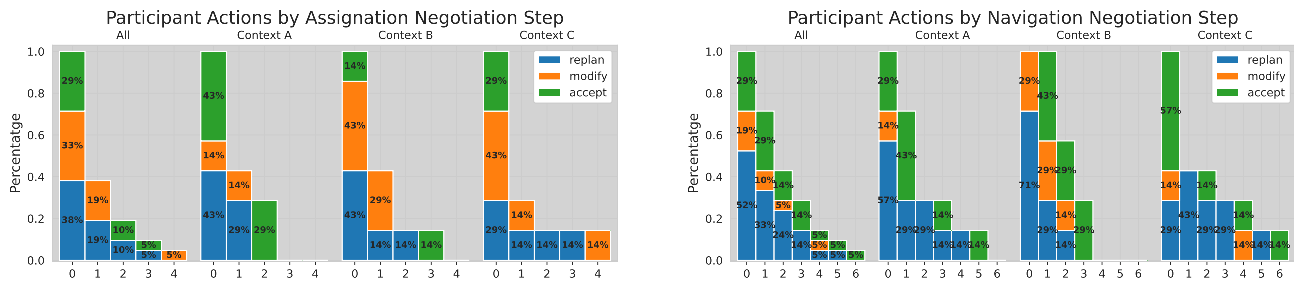

Fig. 9: **Data Distribution**. Human feedback actions over negotiation steps for each of the proposed accuracy assessment contexts.

- Robot collaboration: do you have the perception that the robot wants to collaborate equitably with the operator?

Once the participant had rated all the attributes for the first fruit harvest assignation plan, the leading researcher proposed, verbally, three scenarios: Accept the current plan, ask for one or more plans from the robot, or take the lead and change manually the plan based on its preferences. It is important to say that, before taking any decision, it was mentioned to the participant that after changing the plan there was no possible way to go back to plans already discarded. In case the assignation plans were modified, and after accepting the definitive plan, the participant was asked to evaluate it following the respective parameters listed in the survey.

The third block of the survey consisted of evaluating the harvest routes proposed by the robot. The same protocol listed above was used in this third block, with a verbal definition of robot leadership, intelligence, control, and collaboration. The route proposed by the robot was evaluated and the participant had the opportunity to accept it, ask for a new route or go back and change the fruit assignation agreement with the robot.

The final step in the experiment was to assess the influence of knowing, before the negotiation with the robot, the accuracy of the quality assessment classification algorithm.

Case	Accuracy of Positive Assessment	Accuracy of Negative Assessment
A	70%	95%
B	95%	95%
C	95%	70%

TABLE I: **Detection accuracy** Accuracy of both positive and negative assessment of bunch quality for each of the cases presented in the study.

D. Results

The distribution of participant decisions over robot proposals on each negotiation timestep can be observed in Fig. 9. We can see that most interactions reached an agreement before the second robot assignation proposal, $28,6\%$ due to the participant accepting the robot proposal and 33.3% due to the participant drawing their own assignation proposal. The median finishing timestep in the navigation negotiation was the second timestep. Remarkably, however, there can be observed important differences between the different contexts. Context B shows a 0% acceptance rate over the first robot proposal, while context C shows an acceptance rate of the first robot proposal a little over 50%.[1]

1) Participants Evaluation: Participants in the experiments were asked to answer a questionnaire about their perception of the plan proposals involved in both negotiations and the importance they gave different factors when deciding what action to take upon receiving a robot proposal. Fig. 10 summarises the outcomes of this questionnaire.

After finishing the assignation negotiation, all participants were asked to rate the last plan proposed by the robot in the interaction. Then, those who decided to modify the plan by making a proposal themselves were asked to rate the assignation plan they built themselves. In general, participants perceived their handmade assignation plans to be better than the robot proposals in almost all factors. Then, they were asked to rate the importance of the evaluated factors in their decision to replan and modify. It is noticeable that some participants having both asked for a new plan and modified it, decided to only answer one of the two importance sections.

Similarly, participants were asked to rate the navigation plan proposals involved in the second negotiation step. Participants

[1]https://www.iri.upc.edu/groups/mobrobotics/grape_negotiation/data.csv

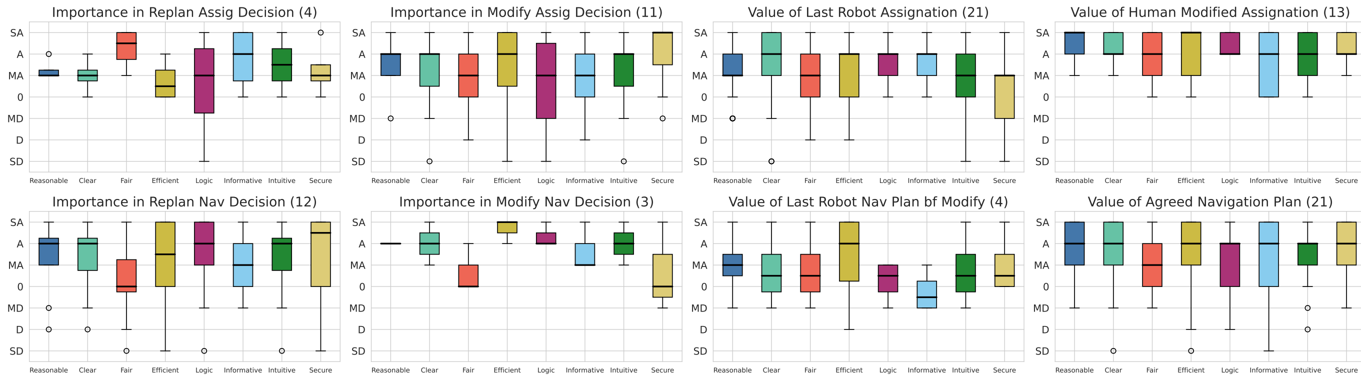

Fig. 10: **User Evaluation**. Participants' perception of proposals' characteristics and the underlying motivations of their decision-making. The number of respondents for each question is indicated in brackets. Discrepancies between the number of respondents indicates individuals performing two actions in the assignation plan and in the navigation decision.

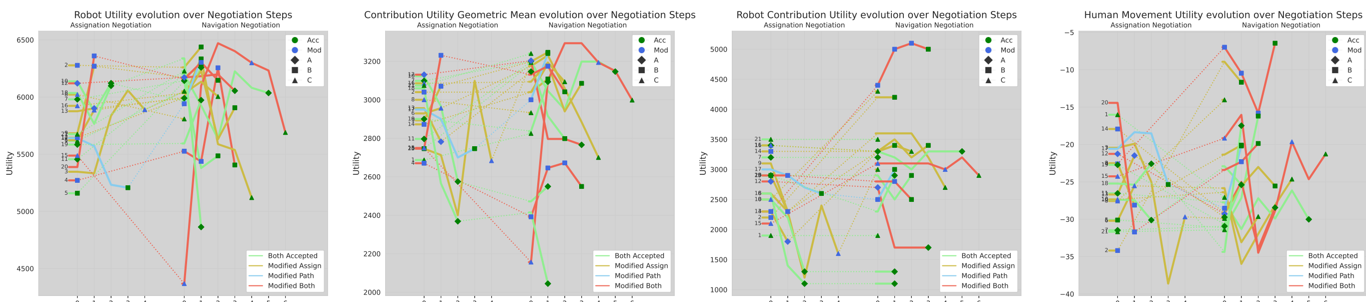

Fig. 11: **Utility models over negotiation steps**. Utility value of robot proposals over negotiation steps for each teammate model considered.

having modified the previously agreed upon assignation plan were asked to rate the last navigation plan proposed by the robot before modifying it, while all participants were asked to rate the agreed upon navigation plan. Data in this case is too limited to make affirmations but it seems that, as before, they perceived a subjective improvement stemming from their decision. Moreover, participants were once again asked to rate the importance of the proposed factors to their decision.

2) Teammate Modelling: In an attempt to capture participants' selections, we decided to build different possible utility models of human decision-making. We proposed three possible models that may explain human decisions in this context:

The subjective fairness model attempts to capture a desire to achieve a plan that makes an equivalent work distribution between the human and the robot. It is computed by calculating the geometric mean of the harvest rewards collected by each teammate.

$$U = \left(\sqrt[\|\alpha\|]{\prod_{a}^{\alpha} R_T^a} \right) \tag{3}$$

The robot contribution model assumes an selfish perspective attempting to maximise the robot's work share. It is computed as the sum of the robot harvest reward.

$$U = R_T^r \tag{4}$$

The movement effort model assumes the participant's perception of their required effort is not measured by the amount of grape bunches harvested, but the amount of distance they have to cover while doing so.

$$U = R_A^h \tag{5}$$

Fig. 11 depicts each robot proposal value both in the robot utility model and the three proposed human utility models. None of them exhibit a growing tendency over negotiation steps, perhaps due to the participants acceptance policy changing over time. Studying time-dependent human utility models requires, however, a bigger dataset and is left for a future study.

Fig. 12 depicts previous data when considering the action taken by the participant upon receiving a proposal from the robot. Data here is grouped independent to the negotiation time-step where it occurred. Also, many samples from the same participant may be included in assignation replan, navigation replan and navigation modify decisions.

VI. Conclusions

In this article, a human-robot system for doing a collaborative harvesting plan negotiation is presented. The final

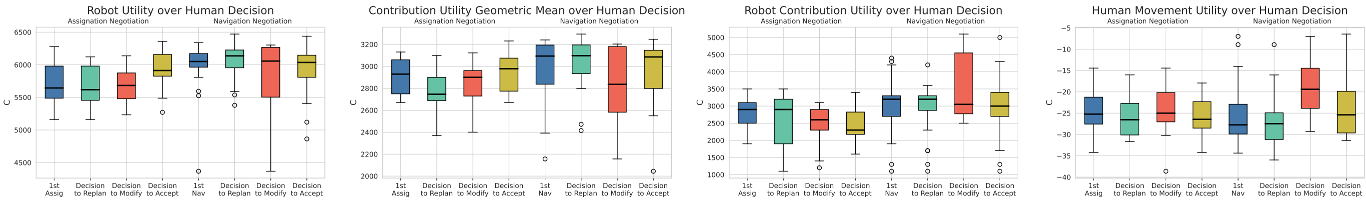

Fig. 12: **Comparison of utility models**. Utility distribution of the first robot proposals compared to utility distribution over human feedback actions.

goal is to obtain a team agreement on human-robot harvest distribution and navigation plans for performing this task. This study shed light in human action selection and factor importance in decision making. Participants having proactively modified the plan subjectively perceived the final agreement to be better than the previous proposals, even though that can't be captured on any of the tested quantitative utility models.

Additionally, the study provides a first approximation on human-robot plan negotiation length expectancy. On both plan negotiation domains, more than half of the negotiations were settled before the third proposal. Although it could perhaps be expected, this confirms that human-robot negotiation will be a particularly challenging field due to the constrained number of available opportunities to make proposals. Whilst in agent negotiations a large amount of proposals are usually expected, here one should maximise the amount of explicit and implicit information extracted from and conveyed through proposals. Moreover, attempts to construct models of human behaviour over negotiation steps will confront with the quick decrease in data. As for these experiments, there is no sufficient data to explore decision-making differences over time steps, but the data provides a foundation to compute statistical potential when attempting to do so in the future.

REFERENCES

[1] M. Ryan, "Labour and skills shortages in the agro-food sector," no. 189, 2023. [Online]. Available: https://www.oecd-ilibrary.org/content/paper/ed758aab-en

[2] L. He and J. Schupp, "Sensing and automation in pruning of apple trees: A review," *Agronomy*, vol. 8, no. 10, 2018. [Online]. Available: https://www.mdpi.com/2073-4395/8/10/211

[3] F. P. Capello, M. B. Spósito, and M. Osaki, "Production costs and profitability of 'niagara rosada' table grape grown in different regions of sÃo paulo state," *Revista Brasileira de Fruticultura*, vol. 39, no. 4, p. e–774, 2017. [Online]. Available: https://doi.org/10.1590/0100-29452017774

[4] G. Coll-Ribes, I. J. Torres-Rodríguez, A. Grau, E. Guerra, and A. Sanfeliu, "Accurate detection and depth estimation of table grapes and peduncles for robot harvesting, combining monocular depth estimation and cnn methods," *Computers and Electronics in Agriculture*, vol. 215, December 2023, 108362.

[5] W. Yin, H. Wen, Z. Ning, J. Ye, Z. Dong, and L. Luo, "Fruit detection and pose estimation for grape cluster–harvesting robot using binocular imagery based on deep neural networks," *Frontiers in Robotics and AI*, vol. 8, p. 626989, 2021.

[6] Y. Jiang, J. Liu, J. Wang, W. Li, Y. Peng, and H. Shan, "Development of a dual-arm rapid grape-harvesting robot for horizontal trellis cultivation," *Frontiers in Plant Science*, vol. 13, 2022.

[7] S. Nikolaidis, M. Kwon, J. Forlizzi, and S. Srinivasa, "Planning with verbal communication for human-robot collaboration," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 7, no. 3, pp. 1–21, 2018.

[8] G. Buisan, G. Sarthou, and R. Alami, "Human aware task planning using verbal communication feasibility and costs," in *International Conference on Social Robotics*. Springer, 2020, pp. 554–565.

[9] V. V. Unhelkar, S. Li, and J. A. Shah, "Decision-making for bidirectional communication in sequential human-robot collaborative tasks," in *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2020, pp. 329–341.

[10] T. Chakraborti, Y. Zhang, D. E. Smith, and S. Kambhampati, "Planning with resource conflicts in human-robot cohabitation," in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, 2016, pp. 1069–1077.

[11] K. Chandan, V. Kudalkar, X. Li, and S. Zhang, "Negotiation-based human-robot collaboration via augmented reality," *arXiv preprint arXiv:1909.11227*, 2019.

[12] A. Angelopoulos, A. Hale, H. Shaik, A. Paruchuri, K. Liu, R. Tuggle, and D. Szafir, "Drone brush: Mixed reality drone path planning," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2022, pp. 678–682.

[13] T. Chakraborti, J. koo Kang, F. Fuggitti, M. Katz, and S. Sohrabi, "Interactive plan selection using linear temporal logic, disjunctive action landmarks, and natural language instruction," in *AAAI Conference on Artificial Intelligence*, 2024.

[14] J. Porteous, A. Lindsay, and F. Charles, "Communicating agent intentions for human-agent decision making under uncertainty," in *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, 2023, pp. 290–298.

[15] D. Porfirio, M. Roberts, and L. M. Hiatt, "Goal-oriented end-user programming of robots," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 582–591. [Online]. Available: https://doi.org/10.1145/3610977.3634974

[16] M. Dalmasso, J. E. Domínguez-Vidal, I. J. Torres-Rodríguez, P. Jiménez, A. Garrell, and A. Sanfeliu, "Shared task representation for human–robot collaborative navigation: The collaborative search case," *International Journal of Social Robotics*, pp. 1–27, 2023.

[17] A. Moon, M. Hashmi, H. M. V. D. Loos, E. A. Croft, and A. Billard, "Design of hesitation gestures for nonverbal human-robot negotiation of conflicts," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 3, pp. 1–25, 2021.

[18] M. Dalmasso, V. Sanchez-Anguix, A. Garrell, P. Jiménez, and A. Sanfeliu, "Exploring preferences in human-robot navigation plan proposal representation," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 369–373. [Online]. Available: https://doi.org/10.1145/3610978.3640627

[19] "Eu canopies project: A collaborative paradigm for human workers and multi-robot teams in precision agriculture systems," https://canopies.inf.uniroma3.it/, 2021.

[20] R. Ranftl, K. Lasinger, D. Hafner, K. Schindler, and V. Koltun, "Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 3, 2022.

[21] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[22] CANOPIES Project Consortium, "Deliverable D4.2: Agronomic Quality Assessment," Jun. 2023, available at: https://www.iri.upc.edu/research/webprojects/canopies/CANOPIES_D4.2_Agronomic_Quality_Assessment-Final.pdf.