Robot explanatory narratives of collaborative and adaptive experiences

Alberto Olivares-Alarcos[1], Antonio Andriella[2], Sergi Foix[1] and Guillem Alenyà[1]

***Abstract*— In the future, robots are expected to autonomously interact and/or collaborate with humans, who will increase the uncertainty during the execution of tasks, provoking online adaptations of robots' plans. Hence, trustworthy robots must be able to store, retrieve and narrate important knowledge about their collaborations and adaptations. In this article, it is proposed a sound methodology that integrates three main elements. First, an ontology for collaborative robotics and adaptation to model the domain knowledge. Second, an episodic memory for time-indexed knowledge storage and retrieval. Third, a novel algorithm to extract the relevant knowledge and generate textual explanatory narratives. The algorithm produces three different types of outputs, varying the specificity, for diverse uses and preferences. A pilot study was conducted to evaluate the usefulness of the narratives, obtaining promising results. Finally, we discuss how the methodology can be generalized to other ontologies and experiences. This work boosts robot explainability, especially in cases where robots need to narrate the details of their short and long-term past experiences.**

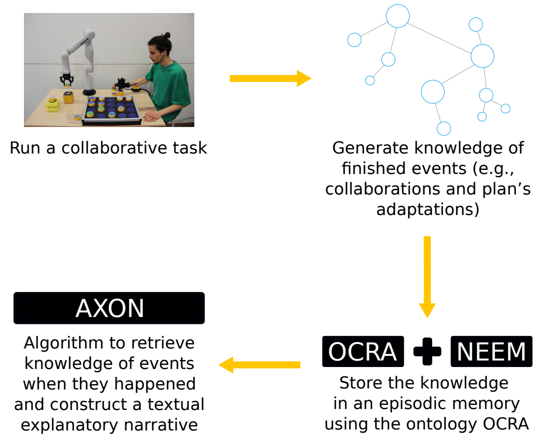


Fig. 1: Overview of the methodology for the generation of explanatory ontology-based narratives for collaborative robotics and adaptation (XONCRA).

I. INTRODUCTION

The development of applications where humans and robots collaborate triggers the appearance of several issues such as those related to trustworthiness between the collaborative agents. For proper cooperation, mutual understanding of the ongoing events and communication between teammates become essential [1]. In regard to this, narratives seem to help with understanding agents' actions [2]. Hence, collaborative robots could narrate what they know of their experiences, i.e., collaborations and plan adaptations, to be more understandable. Those robot narratives may boost explainable agency (i.e., explaining the reasoning of goal-driven agents and robots), which has recently gained significant momentum [3], [4]. Robotic tasks may involve several events and a lot of contextual knowledge. Hence, time-indexed narratives of events (i.e., narrating events when they occur) make more sense in robotics than in other artificial intelligence tasks (e.g. classification), where single post hoc and time-independent narratives or explanations might suffice.

Langley et al. [5], discussed the need for three elements of explainable agency: a representation of the domain knowledge, an episodic memory to store the knowledge, and the ability to access and retrieve that knowledge to generate explanations. Episodic memory is the collection of past personal experiences that occurred at particular times and places [6]. Beetz et al. [7], presented the second generation of KnowRob, a knowledge-based framework for robotics, which includes formal domain ontologies, and narrative-enabled episodic memory (NEEM) storage and retrieval. NEEMs may be useful to generate human understandable explanatory narratives, but this is still unexplored, especially in the collaborative robotics and adaptation domain.

In this article, we present a methodology (see Fig. 1) for the generation of eXplanatory Ontology-based Narratives for Collaborative Robotics and Adaptation (XONCRA). The domain knowledge is modeled using the Ontology for Collaborative Robotics and Adaptation (OCRA) [8]. The knowledge of robot experiences is stored in NEEMs [7] for later retrieval. As a technical (open source) contribution, OCRA is integrated into the NEEMs framework. The novel theoretical contribution is an Algorithm for eXplanatory Ontology-based Narratives (AXON). From the episodic memory, AXON retrieves the relevant knowledge about robot collaborations and plans' adaptations, and then it constructs the final textual narrative. The proposed algorithm produces different types of narratives based on the chosen amount of detail (specificity), addressing different users' preferences. We evaluated the quality (*usefulness*) of the narratives' information through a pilot study with users. Finally, we briefly discussed XONCRA's potentiality to generalize to other ontologies and NEEMs. To the best of our knowledge, this is the first work to use a (temporal) episodic memory to generate ontology-based textual explanatory narratives. In this work, we tackled the following research questions:

- **RQ1 -** How can robots construct the narrative of their collaborative and adaptive events (experiences)?
- **RQ2 -** How does the narratives' specificity affect the users perceived usefulness of the received information?

[1]Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Llorens i Artigas 4-6, 08028 Barcelona, Spain; {aolivares,sfoix,galenya}@iri.upc.edu
[2]Pal Robotics, C/ de Pujades, 77, 08005 Barcelona, Spain; antonio.andriella@pal-robotics.com

II. RELATED WORK

We found great inspiration in the narrative and storytelling literature. Labov et al. [9], defined a narrative *'as a way of recounting past events, in which the order of narrative clauses matches the order of events as they occurred'*. Carr [2], stated that in order to provide explanatory information, a narrative should contextualize the agent's experiences in time. Both works emphasized the importance of the time when the events occurred, reinforcing the need for episodic memory. Narratives have already been an inspiration for works on robot task plans' verbalization [10], [11], [12].

A sound approach to represent domain knowledge is to use representation formalisms such as ontologies. The 1872–2015 IEEE Standard Ontologies for Robotics and Automation [13] and the 1872.2-2021 IEEE Standard for Autonomous Robotics Ontology [14] were developed to become references for knowledge representation in the domain. Indeed, the use of ontologies has spread to several robotics sub-domains [15], [16], [17]. Some examples are manufacturing and collaborative robotics [18], [19], [20], [21], [22], [23], [8], robot co-design [24], [25], and service and general purpose robots [7], [26], [27]. All these works are steps towards the harmonization and formalization of the knowledge in the robotics domain. Hence, they have the potential to play a major role in the explainable agency.

The notion of episodic memory was first introduced by Tulving as the collection of past personal experiences that occurred at particular times and places [6]. Its essence lies in the conjunction of three concepts: self, autonoetic awareness, and subjectively sensed time [28]. Beetz et al. [7], introduced a knowledge-based framework for robots that includes an episodic memory, the narrative-enabled episodic memory (NEEM). It consists of the NEEM experience (low-level time-indexed information) and the NEEM narrative (symbolic descriptions, e.g., goals, states, etc.). NEEMs have already been used in human-robot interaction [29], and robot learning [30]. Nevertheless, their role in the generation of robot textual narratives still remains unexplored.

In the literature, several authors worked on automatic text generation using knowledge modeled in OWL (Web Ontology Language) or RDF (Resource Description Framework) [31], [32], [33], [34]. Although inspiring, none of those works discussed the generation of different types of texts based on the preferred specificity. Furthermore, in ours, the target knowledge to be included in the textual narratives is automatically retrieved, while the others just assumed that the knowledge atoms or tuples were given.

III. XONCRA - EXPLANATORY ONTOLOGY-BASED NARRATIVES FOR COLLABORATIVE ROBOTICS AND ADAPTATION

A. Preliminary notation

Let's assume countable pairwise disjoint sets N_C, N_P, and N_I of class names, property names, and individuals, respectively. The standard relation `rdf:type`, which relates an individual with its class, is abbreviated as `type` and included in N_P. A knowledge graph $\mathcal{G}$ is a finite set of triples of the form $\langle s, p, o \rangle$ (subject, property, object), where $s \in N_I$, $p \in N_P$, $o \in N_I$ if $p \neq type$, and $o \in N_C$ otherwise. The semantic knowledge of an episodic memory can be seen as a time-indexed knowledge graph $\mathcal{G}_\mathcal{T}$, which is a finite set of tuples of the form $\langle s, p, o, t_i, t_f \rangle$, where $t_i, t_f \in \mathbb{R} > 0$, and denote the time interval (initial and final time) in which the triple $\langle s, p, o \rangle$ holds. Knowledge graphs commonly comply with the open-world assumption, thus, non-asserted triples are unknown instead of false. For this reason, the second version of the Web Ontology Language (OWL 2) allows to make explicit negative properties assertions: $\langle s, p, o \rangle$ is $false$.[3] Hence, in $\mathcal{G}_\mathcal{T}$ one may store, for instance, that during an interval of time, $\langle t_i, t_f \rangle$, an event e is not an instance of the class `Collaboration`: $\langle e, type, Collaboration, t_i, t_f \rangle$ is $false$. In this work, querying the $\mathcal{G}_\mathcal{T}$, we build what we called 'narrative tuples' of an instance event, $\mathcal{T}_e$: $\langle s, p, o, t_i, t_f, sign \rangle$, where $sign$ indicates whether the time-indexed triple comes from a positive or negative assertion.

B. NEEMs for collaborative robotics and adaptation

Our methodology incorporates a knowledge-based episodic memory for collaborative robots that adapt in unstructured scenarios. It consists of the integration of an ontology for collaborative robotics and adaptation (OCRA) [8], into the NEEMs ecosystem of Knowrob [35], [7]. It allows robots to represent time-indexed knowledge of their collaborations and adaptations, store it and retrieve it for a later generation of textual explanatory narratives.

1) Background on OCRA: The ontology was developed to enhance the reusability of the domain's terminology, and to allow robots to formalize and reason about two main concepts: collaboration and plan adaptation. `Collaboration` is defined as *'an event in which two or more agents share a goal and a plan to achieve the goal, and execute the plan while interacting'*. `Plan Adaptation` is *'an event in which one (or more) agent, due to its evaluation of the current or expected future state, changes its current plan while executing it, into a new plan, in order to continuously pursue the achievement of the plan's goal'*. Considering those definitions, narratives of `Collaborations` shall include knowledge about the shared plan and goal, and the agents executing the plan. Meanwhile, narratives of `Plan Adaptations` must contain details regarding the initial and new plans, the situation triggering the adaptation, and the involved agent. In this work, it is used OCRA's formalization in OWL 2 DL, a description logic version of OWL 2.

2) Background on NEEMs: For every activity the robot (agent) performs, observes, or prospects, it can create an episode and store it in its memory. An episode is best understood as a video recording that the robot makes of the ongoing activity (see Fig. 2). In addition, those videos are enriched with a very detailed log of the actions, motions, their purposes, effects, and the agent's sensor information during

[3] www.w3.org/2007/OWL/wiki/FullSemanticsNegativePropertyAssertions

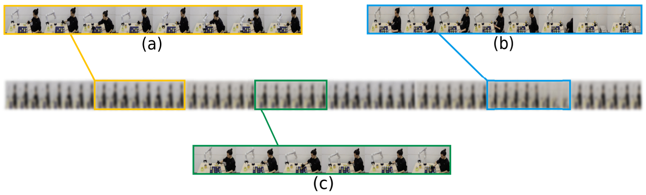


Fig. 2: Visualization of a recorded NEEM of a prototypical collaborative kitting task (filling a tray) with different episodes within it (a, b, and c).

the activity. The episodic memories created by Knowrob are named narrative-enabled episodic memories (NEEMs). A NEEM consists of the NEEM experience and the NEEM narrative. The NEEM experience captures low-level data such as the agent's sensor information, e.g. images and forces, and records of poses of the agent and its detected objects. NEEM experiences are linked to NEEM narratives, which are logs of the episode described symbolically. These narratives contain information regarding the tasks, the context, intended goals, observed effects, etc. In this work, the focus is on the NEEM-narrative, since the aim is to explain the symbolic understanding that the robot has of its experiences. A detailed overview of NEEMs can be found in the NEEM Handbook [36].

3) Integration: NEEMs are modeled using OWL 2 DL ontologies built upon the DOLCE+DnS Ultralite (DUL) foundational ontology [37], the same upper-level ontology that OCRA relies on. OCRA was integrated into Knowrob's and NEEMs' ecosystem without causing any ontological inconsistency. The knowledge base is accessible to the robot through a prolog-based service implemented as a ROS (Robot Operating System) package: rosprolog.[4] It was implemented a novel ROS package (know-cra) in which OCRA is integrated into Knowrob's framework. This implementation is publicly available on a Github repository,[5] and illustrates how to load and use OCRA, and some instantiated use cases, with Knowrob. Furthermore, the shared code also includes examples of manipulating recorded NEEMs.

C. *AXON - An algorithm for explanatory ontology-based narratives*

AXON is our major theoretical contribution, a novel algorithm that retrieves knowledge about target experiences or events from episodic memories, and uses the knowledge to construct textual explanatory narratives. The time-indexed knowledge graph $\mathcal{G}_\mathcal{T}$ stored in the episodic memory is the first algorithm's input. Furthermore, AXON takes three more inputs: the ontological class (or classes) of the events to narrate, the temporal locality (time interval of the events of interest), and the level of specificity. Although our focus is on narratives about `Collaborations` and `Plan adaptations`, AXON is general enough to work with other OWL 2 DL ontologies and classes, as it is discussed in Sec. V. There are three different narrative types, depending on the selected specificity. In this work, specificity

[4] www.github.com/knowrob/rosprolog
[5] https://github.com/albertoOA/know_cra

Algorithm 1: AXON

Input: Episodic memory ($\mathcal{G}_\mathcal{T}$), events to narrate ($\mathcal{C}$), temporal locality (L_i, L_f), specificity (S)
Output: Narrative ($\mathcal{E}$)

1 $\mathcal{E} \leftarrow \emptyset$
2 $\mathcal{I}_\mathcal{T} \leftarrow$ RetrieveInstancesWithTimeInterval($\mathcal{G}_\mathcal{T}$, $\mathcal{C}, L_i, L_f$)
3 **foreach** $\langle e, t_i, t_f \rangle \in \mathcal{I}_\mathcal{T}$ **do**
4     $\mathcal{T}_e \leftarrow$ RetrieveNarrativeTuples($\mathcal{G}_\mathcal{T}, \langle e, t_i, t_f \rangle, S$)
5     $\mathcal{E}_e \leftarrow$ ConstructNarrative($\mathcal{T}_e$)
6     $\mathcal{E} \leftarrow \mathcal{E} \cup \mathcal{E}_e$
7 **end**

refers to the amount of detail used to construct the textual narrative, more precisely, the number of knowledge tuples. This section first introduces the main algorithm (see Alg. 1), and then we explain its three major routines: Retrieve Instances With Time Interval, Retrieve Narrative Tuples, and Construct Narrative. An implementation of the algorithm and an example of use can be found at an online repository.[6]

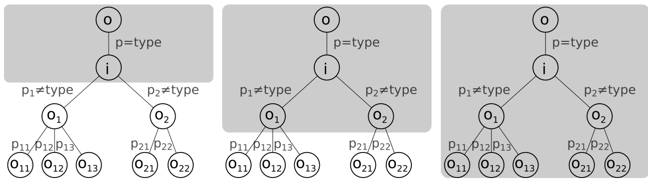
AXON first retrieves a set $\mathcal{I}_\mathcal{T}$ of tuples $\langle e, t_i, t_f \rangle$, containing the event instances e of the provided classes $\mathcal{C}$ whose time interval (t_i, t_f) exists, at least partially, within the temporal locality (L_i, L_f) (line 2). Second, based on the specificity S, the algorithm retrieves a set of knowledge tuples $\mathcal{T}_e$ related to each instance (line 4). Third, an explanation $\mathcal{E}_e$ for every instance is constructed using their respective tuples (line 5). Finally, the algorithm concatenates the new explanation to the set of explanations $\mathcal{E}$ (line 6).

1) Retrieve instances with time interval routine: Given a time-indexed knowledge graph $\mathcal{G}_\mathcal{T}$, an ontological existing class or a set of them, $\mathcal{C} \subset N_C$, and a time interval $\langle L_i, L_f \rangle$, this routine retrieves a set $\mathcal{I}_\mathcal{T}$ containing all the time-indexed instances $\langle e, t_i, t_f \rangle$ of the given classes such that $\forall \langle e, t_i, t_f \rangle \in \mathcal{I}_\mathcal{T} \rightarrow \exists c \in \mathcal{C} \wedge \langle e, type, c, t_i, t_f, sign \rangle \wedge \langle t_i, t_f \rangle \cap \langle L_i, L_f \rangle$. Some examples of instances of events to narrate with their time interval may be the following:

$\langle Event\_15, 100.0, 142.0 \rangle$,
$\langle Event\_27, 200.0, 240.0 \rangle$.

2) Retrieve narrative tuples routine: Given $\mathcal{G}_\mathcal{T}$, an instance event e to narrate with the time interval in which it exists $\langle t_i, t_f \rangle$, and the specificity level S, this routine retrieves all the relevant tuples, $\langle s, p, o, t_i, t_f, sign \rangle$, to construct the narrative. The first level of specificity can be considered as a baseline and only returns tuples containing the class c of each instance: $\langle e, p, c, t_i, t_f, sign \rangle \in \mathcal{G}_\mathcal{T} \wedge p = type$. In the second level, the algorithm adds all the tuples in which the instance e is related to an object o through any property different to `type`: $\langle e, p, o, t_i, t_f, sign \rangle \in \mathcal{G}_\mathcal{T} \wedge p \neq type$. Finally, the third level adds all the tuples in which the objects o from the second level are related to other objects o_x: $\langle o, p_x, o_x, t_{ix}, t_{fx}, sign_x \rangle \in \mathcal{G}_\mathcal{T} \wedge \langle t_i, t_f \rangle \cap \langle t_{ix}, t_{fx} \rangle$. As robots' experiences are tied to a time frame, the search was restricted to tuples whose time interval $\langle t_{ix}, t_{fx} \rangle$ intersected

[6] https://github.com/albertoOA/explanatory_narratives_cra



(a) (b) (c)

Fig. 3: Graphical representation of the different levels of specificity (S) and their respective depth in the knowledge graph. (a) $S = 1$, (b) $S = 2$, and (c) $S = 3$.

the time interval of the instance $\langle t_i, t_f \rangle$. This aimed to avoid retrieving tuples that were irrelevant to the narrative of the instance e. Furthermore, if a tuple or its inverse already exists in the retrieved set $\mathcal{T}_e$, it is not added. Note that the retrieved tuples for each level are also included in upper levels (e.g., the tuples from the first level are also returned in the second and the third). In an intuitive way, this would be equivalent to go deeper in the knowledge graph (see Fig. 3). Using as example the task of filling a tray (from [8]), some instances of the retrieved narrative tuples $\mathcal{T}_e$ of an event are:

$\boldsymbol{\mathcal{T}_{e_1}}$. ⟨Robot, hasPlan, Place Tokens By Color, 200.0, 240.0, positive⟩,
$\boldsymbol{\mathcal{T}_{e_2}}$. ⟨Event_27, hasParticipant, Robot, 200.0, 240.0, positive⟩,
$\boldsymbol{\mathcal{T}_{e_3}}$. ⟨Place Tokens By Color, isPlanOf, Human, 200.0, 240.0, negative⟩,
$\boldsymbol{\mathcal{T}_{e_4}}$. ⟨Human, isParticipantIn, Event_27, 200.0, 240.0, positive⟩,
$\boldsymbol{\mathcal{T}_{e_5}}$. ⟨Human, type, Physical Agent, 1.0, 1000.0, positive⟩.

3) Construct narrative routine: Given the narrative tuples $\mathcal{T}_e$ of an instance event to narrate e, this routine constructs the final explanatory narrative following a set of rules: **casting, clustering, ordering, and grouping**. These rules, proposed by Dalianis et al. [34], define the aggregations that humans usually do in natural language.

Casting consists in homogenizing all the properties used in the tuples. First, making sure that in all the tuples $\mathcal{T}_e$ concerning the target instance e ($\mathcal{T}_{e_2}$ and $\mathcal{T}_{e_4}$ in the previous example), e acts as the subject of the tuple. Hence, when $\mathcal{T}_e$ contains a tuple in which e acts as the object, $\langle s, p, e, t_i, t_f, sign \rangle \in \mathcal{T}_e$, the algorithm inverts the tuple to: $\langle e, p^{-1}, s, t_i, t_f, sign \rangle$, where p^{-1} is the inverse property of p. In the tuples shown before, the tuple $\mathcal{T}_{e_4}$ containing the property `isParticipantIn` would be changed using its inverse `hasParticipant`. Once this is done, all the tuples regarding e are added to the set of cast tuples $\mathcal{T}_{eCast}$. The second step in casting involves the tuples not concerning e ($\mathcal{T}_{e_1}$, $\mathcal{T}_{e_3}$ and $\mathcal{T}_{e_5}$), ensuring that each tuple's property is consistent with the properties already existent in the cast tuples. Otherwise, the tuple is inverted before adding it to $\mathcal{T}_{eCast}$. In the example, $\mathcal{T}_{e_1}$ is added to $\mathcal{T}_{eCast}$ (following the order), thus, $\mathcal{T}_{e_3}$ needs to be inverted before added.

Then the routine **clusters** all the tuples $\langle s, p, o, t_i, t_f, sign \rangle$ that share the subject s. Therefore, when generating the narrative, all the information about a specific subject will appear together. In the example, $\mathcal{T}_{e_2}$ and the inverted $\mathcal{T}_{e_4}$, and the inverted $\mathcal{T}_{e_3}$ and $\mathcal{T}_{e_5}$ would be clustered.

Next, the tuples are **ordered**: externally and internally. The external ordering consists in ordering the subjects from more information (more tuples) to less. This rule has one exception, the information about the target instance is always at the top front of the list. The internal ordering ensures that the tuples with the property $p = type$ are at the front of the list for each subject. In the example, after applying all these rules the set of tuples would change to:

$\boldsymbol{\mathcal{T}'_{e_1}}$. ⟨Event_27, hasParticipant, Robot, 200.0, 240.0, positive⟩,
$\boldsymbol{\mathcal{T}'_{e_2}}$. ⟨Event_27, hasParticipant, Human, 200.0, 240.0, positive⟩,
$\boldsymbol{\mathcal{T}'_{e_3}}$. ⟨Human, type, Physical Agent, 1.0, 1000.0, positive⟩,
$\boldsymbol{\mathcal{T}'_{e_4}}$. ⟨Human, hasPlan, Place Tokens By Color, 200.0, 240.0, negative⟩,
$\boldsymbol{\mathcal{T}'_{e_5}}$. ⟨Robot, hasPlan, Place Tokens By Color, 200.0, 240.0, positive⟩.

Finally, the tuples are **grouped** into a sentence, constructing the final textual narrative $\mathcal{E}_e$. First, the tuples with the same subject, property, interval, and sign are joined (object grouping). Hence, if there are two tuples: $\langle s, p, o_a, t_i, t_f, sign \rangle$ and $\langle s, p, o_b, t_i, t_f, sign \rangle$, the algorithm joins them to: $\langle s, p, o_a \textit{ and } o_b, t_i, t_f, sign \rangle$. In the example tuples, $\mathcal{T}'_{e_1}$ and $\mathcal{T}'_{e_2}$ would be joined into: ⟨Event_27, hasParticipant, Robot and Human, 200.0, 240.0, positive⟩. Second, the tuples for each subject are joined into separated sentences (subject grouping) considering their sign and using the conjunction 'and' and the propositions 'from' and 'to'. When generating the text of a negative assertion, it is included the adverb 'not' before the property. Furthermore, it is excluded the time interval of a tuple if it was equal to the time interval in which the instance exists. The names of properties, classes, and instances are kept, only the property 'type' is changed to 'is a type of'. The final narrative for the ongoing example would be:

'Event_27' has participant 'Robot and Human' from 200.0 to 240.0. 'Human' is a type of 'Agent' from 1.0 to 1000.0 and (not) has plan 'Place Tokens By Color'. 'Robot' has plan 'Place Tokens By Color'.

IV. VALIDATION: SETTING XONCRA TO WORK

A. Collaborative task: filling a tray with tokens

The validation of XONCRA was contextualized in a lab mock-up of a real task, where a robot and a human shared the task of filling the compartments of a tray/board (see Fig. 4). The task's objective was to obtain a tray full of tokens. The specific order changes to create different tasks (e.g., tokens are sorted by color, in ascending order, etc). When a token was not useful to accomplish the task's goal (e.g., compartments for that color are already filled), it was discarded. The risk of human-robot collision was computed using the pose and velocity extracted from an HTC Vive tracker attached to the human's hand using the Time-To-Contact (TTC) [38].

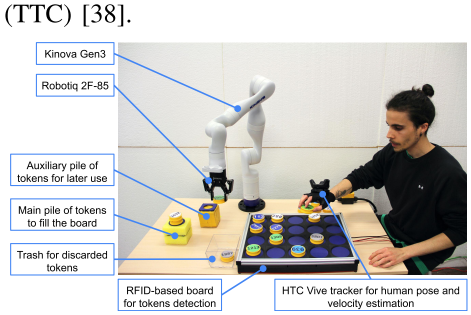


Fig. 4: Setup of filling a tray, the validation task.

B. Robot experiences about collaboration and adaptation

Following the schema described in the proposed methodology XONCRA (see Fig. 1), RQ1 is addressed. The first step is to run executions, twelve in this case, of the validation task. Those executions were designed to showcase diverse situations of collaborations and adaptations according to how they are defined in OCRA. Hence, varying their main elements: the goal, the plan, and the workload distribution between the human and the robot. In order to ensure a curated knowledge base, the knowledge tuples involved in those executions were manually stored into a single NEEM after recording videos of the executions. From now, we will refer to that NEEM as *validation NEEM*.

The twelve events included three cases of collaboration, six robot plan adaptations, and three other situations with non-collaboration. According to OCRA's definitions, in the collaborations, the human and the robot shared the goal (e.g. full board with tokens in columns ordered by color) and the plan, and both of them participated to accomplish the goal. In the adaptations, the robot stopped executing a plan due to an unexpected situation and started executing a new plan better suited to accomplish the goal (e.g. the robot went to another compartment when the human filled the one that the robot wanted to fill). Finally, the events showing non-collaborations (i.e. broken collaborations) represented cases when one of the axioms needed for a collaboration to exist was violated (e.g., the human stopped participating, or the goal/plan was not shared).

C. Explanatory narratives generation: an example

The focus here is on one event among the twelve stored in the *validation NEEM*. *Event 15* shows the human stopping the collaboration (see Fig. 5). Using AXON with the parameters $\mathcal{G}_{\mathcal{T}}$ = *validation NEEM*, $\mathcal{C} = Collaboration$, $(L_i, L_f) = (100.0, 142.0)$, $S = 3$, one obtains a narrative of the specific event. Recall that the level of specificity 3 includes the result of levels 1 (red) and 2 (blue). To see the rest of generated narratives visit the repository.[7]

> 'Event_15' (not) is a type of 'Collaboration' and is a type of 'Event' from 100.0 to 142.0 and executes plan 'Place Tokens In Columns By Color' and has participant 'Robot' and (not) has participant 'Human'. 'Place Tokens In Columns By Color' is a type of 'Plan' from 1.0 to 1000.0 and has component 'Full Board With Tokens In Columns By Color' and is plan of 'Robot and Human'. 'Robot' is a type of 'Physical Agent' from 1.0 to 1000.0 and has goal 'Full Board With Tokens In Columns By Color'. 'Human' is a type of 'Physical Agent' from 1.0 to 1000.0 and has goal 'Full Board With Tokens In Columns By Color'.

D. Pilot study: analysis of the usefulness of information

The length of an explanatory narrative plays a major role in the comprehension of its relevant information. The aim is to be informative, providing as much information as is needed, and no more [39]. Hence, a pilot study was carried out to assess the perceived usefulness of the narratives depending on their specificity, addressing RQ2.

(a) (b) (c)

Fig. 5: Example of a non-collaboration in which the human stops participating in the shared task. (a) The human wears off its HTC tracker. (b) The human leaves the workspace. (c) The robot continues performing the task alone.

Specifically, participants watched a video containing the twelve events included in the *validation NEEM*. The video depicted a textual narrative generated by our method after each of the events. Users were asked to imagine that they were about to receive training (the video with the narratives) aimed at preparing them to collaborate with a robot. This may be a real case in an industrial environment, where a video of a human-robot collaboration plus automatically generated narratives of the collaboration can be used to train new operators. A between-subject study was conducted, with three groups that evaluated each of the narratives' types. Groups 1, 2, and 3 evaluated the narratives with specificity 1, 2, and 3, respectively.

1) Procedure: The study was conducted at our facilities, in an isolated room to avoid distractions. The experimenter informed each participant of the procedure and asked them to fill out an informed consent form, in which they gave permission to gather their data for scientific purposes. Next, users were shown a warm-up video with the experiment's context and the narratives' format, ensuring that users received the same information prior to the experiment. Then, users watched the video with the twelve events recorded in the NEEM plus a textual narrative after each of the events. After watching the video, the participants were asked to fill out a questionnaire with two parts: information quality (usefulness) assessment, and open qualitative questions. The videos and the questionnaire are provided as supplemental material.[8]

2) Participants: 30 participants (10 per group) were recruited. There was no withdrawal. Participants were aged between 21 and 59 (26.7% of them were female), with M=29 and SD=7.61. Most of them (93.3%) had a background in engineering, artificial intelligence, or robotics, and at least 70% had already interacted with other unspecified robots. Participation in the study was voluntary.

3) Quantitative and qualitative analysis: For a quantitative subjective analysis, it was used the quality of information measurement discussed by Lee et al. [40]. They presented a model for Information Quality, a questionnaire to measure it, and analysis techniques to interpret the measures. This article uses one of the quadrants of their model and its relative questionnaire: *usefulness*. It aims to assess whether or not the information is relevant to the user's task, in our case, the 'new operator training task'. In particular, *usefulness* was measured through five dimensions: *appropriate amount*, *relevancy*, *understandability*, *interpretability*, and *objectivity*.

[7] https://github.com/albertoOA/explanatory_narratives_cra/tree/main/txt

[8] www.iri.upc.edu/groups/perception/XONCRA

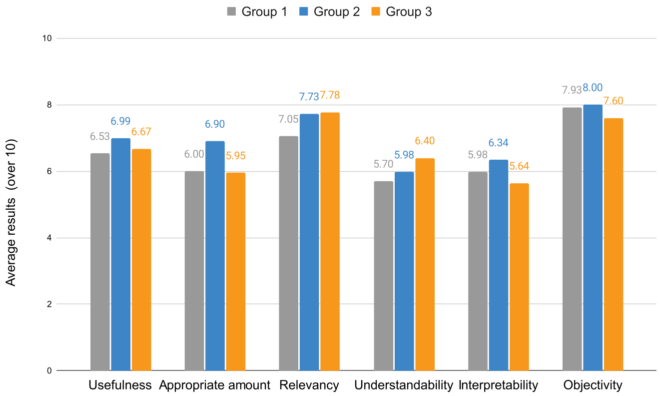


Fig. 6: Results for the quantitative analysis of the information usefulness. Users' group number also corresponds to the specificity level of the assessed narratives. *Usefulness* is computed as the average of the other five dimensions.

For each dimension, a set of questions had to be evaluated using an 11-point Likert scale ranging from completely disagree (0) to completely agree (10). The results of the study are shown in Fig. 6. Looking at them, one notes that the three levels of specificity produced useful narratives (all above 6.5 points). However, the second level (Group 2) was perceived as the most useful. Focusing on each dimension, the preferred narratives regarding the *appropriate amount* and the *interpretability* were those with specificity 2. Nevertheless, it is interesting to see that narratives with larger specificity (Group 3), were perceived to contain more *understandable* information. Other dimensions show negligible differences.

The questionnaire also included some qualitative measures in the form of four open questions. First, the users were asked if a video without narratives would prepare them for real interaction with the robot and why. 86.7% of the participants answered that the video without explanation would not be enough to be prepared to collaborate with a real robot. This corroborated the need for a narrative, regardless of the specificity. The second question asked whether the explanation had helped to prepare them for real interaction with the robot and why. 66.7% found the narratives greatly helpful. However, this percentage changes if one looks at the isolated answer provided by each group: 50%, 70%, and 80% for Groups 1, 2, and 3, respectively. Hence, narratives with higher specificity seemed to be more helpful. Third, it was asked if they would prefer a summarized or a complete but repetitive narrative and why. 50% of the participants as a whole would prefer a summarized explanation. Nevertheless, that percentage grows to 70% for the participants of Group 3, who read longer narratives. Finally, it was asked if there was any content they would add to the narratives. Some participants proposed to include graphical information.

V. CONCLUSION

In this work, we presented XONCRA, a methodology for the generation of explanatory ontology-based narratives for collaborative robotics and adaptation. It is built upon an existent ontology (OCRA) [8], and a knowledge-based framework with episodic memories (NEEMs) [7]. These two elements together enable the representation, storage, and later retrieval of time-indexed knowledge. XONCRA also comprises a novel algorithm, AXON, which automatically retrieves knowledge from NEEMs to construct an explanatory narrative with it. It can produce three types of results based on the level of specificity. We provide an implementation of the methodology and some examples, addressing RQ1.

Depending on their specificity, the perceived narratives' usefulness was assessed through a pilot study, answering RQ2. Results indicated that participants found the three types to be useful. However, it was discovered that users preferred narratives generated with level 2 of specificity, especially for their appropriate amount and interpretability. Nevertheless, narratives with larger specificity (3), were perceived to contain more understandable information. The positive finding of this analysis is that all the narratives produced by XONCRA can help and be useful. Moreover, the methodology can address different preferences with respect to different trade-offs: appropriate amount vs understandability, etc.

Note that even though we focused on narratives of robot `Collaborations` and `Plan Adaptations`, our methodology generalizes beyond our use case. By construction, it can deal with any other ontological class as long as it is formalized in the appropriate format to use the NEEMs framework. Indeed, there is a large list of available NEEMs generated for other purposes, e.g., a human setting up a table for breakfast, a robot monitoring a shelf in the retail domain, etc.[9] Utilizing those NEEMs, XONCRA might produce narratives about `Actions`, `Tasks`, `Objects`, etc.

In the future, we would like to conduct a larger user study for a better assessment of our work. Furthermore, we aim to address some improvements extracted from the qualitative analysis: using a more natural language, summarizing narratives, modeling human preferences, and adding graphical information. Specifically, we will upgrade XONCRA to use advanced natural language techniques and to do summaries of the narratives (e.g. when some similar content has already been provided). We will investigate how to model human preferences to learn the preferred specificity level. We will also use the knowledge graph structure to provide graphical information supporting the narratives. Finally, we want to use XONCRA with other ontologies and NEEMs.

ACKNOWLEDGMENT

This work has been partially funded by MCIN/ AEI /10.13039/501100011033 under the project CHLOE-GRAPH (PID2020-119244GB-I00); by MCIN/ AEI /10.13039/501100011033 and by the "European Union (EU) NextGenerationEU/PRTR under the project COHERENT (PCI2020-120718-2) and the project ROB-IN (PLEC2021-007859). A. Andriella is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801342 (Tecniospring INDUSTRY) and the Government of Catalonia's Agency for Business Competitiveness (ACCIÓ). A. Olivares-Alarcos is supported by the European Social Fund and the Ministry of Business and Knowledge of Catalonia through the FI 2020 grant.

[9]https://neemgit.informatik.uni-bremen.de/neems

REFERENCES

[1] L. Yuan, X. Gao, Z. Zheng, M. Edmonds, Y. N. Wu, F. Rossano, H. Lu, Y. Zhu, and S.-C. Zhu, "In situ bidirectional human-robot value alignment," *Science Robotics*, vol. 7, no. 68, p. eabm4183, 2022.

[2] D. Carr, "Narrative explanation and its malcontents," *History and Theory*, vol. 47, no. 1, pp. 19–30, 2008.

[3] S. Anjomshoae, A. Najjar, D. Calvaresi, and K. Främling, "Explainable agents and robots: Results from a systematic literature review," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '19, (Richland, SC), p. 1078–1088, International Foundation for Autonomous Agents and Multiagent Systems, 2019.

[4] T. Chakraborti, S. Sreedharan, and S. Kambhampati, "The emerging landscape of explainable automated planning decision making," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20* (C. Bessiere, ed.), pp. 4803–4811, International Joint Conferences on Artificial Intelligence Organization, 7 2020. Survey track.

[5] P. Langley, B. Meadows, M. Sridharan, and D. Choi, "Explainable agency for intelligent autonomous systems," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, AAAI'17, p. 4762–4763, AAAI Press, 2017.

[6] T. Endel, *Organization of memory*, ch. Episodic and semantic memory, pp. 381–403. New York: Academic Press, e. tulving w. donaldson ed., 1972.

[7] M. Beetz, D. Beßler, A. Haidu, M. Pomarlan, A. K. Bozcuoğlu, and G. Bartels, "Knowrob 2.0 — a 2nd generation knowledge processing framework for cognition-enabled robotic agents," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 512–519, 2018.

[8] A. Olivares-Alarcos, S. Foix, S. Borgo, and Guillem Alenyà,, "Ocra – an ontology for collaborative robotics and adaptation," *Computers in Industry*, vol. 138, p. 103627, 2022.

[9] W. Labov and J. Waletzky, "Narrative analysis: Oral versions of personal experience1," *Journal of Narrative and Life History*, vol. 7, no. 1-4, pp. 3–38, 1997.

[10] S. Rosenthal, S. P. Selvaraj, and M. M. Veloso, "Verbalization: Narration of autonomous robot experience.," in *IJCAI*, vol. 16, pp. 862–868, 2016.

[11] J. G. Flores, I. Meza, É. Colin, C. Gardent, A. Gangemi, and L. A. Pineda, "Robot experience stories: First person generation of robotic task narratives in sitlog," *Journal of Intelligent & Fuzzy Systems*, vol. 34, pp. 3291–3300, 2018. 5.

[12] G. Canal, S. Krivić, P. Luff, and A. Coles, "PlanVerb: Domain-Independent Verbalization and Summary of Task Plans," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, 2022.

[13] C. Schlenoff, E. Prestes, R. Madhavan, P. Goncalves, H. Li, S. Balakirsky, T. Kramer, and E. Migueláñez, "An ieee standard ontology for robotics and automation," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1337–1342, 2012.

[14] P. J. Gonçalves, A. Olivares-Alarcos, J. Bermejo-Alonso, S. Borgo, M. Diab, M. Habib, H. Nakawala, S. V. Ragavan, R. Sanz, E. Tosello, and H. Li, "Ieee standard for autonomous robotics ontology [standards]," *IEEE Robotics Automation Magazine*, vol. 28, no. 3, pp. 171–173, 2021.

[15] J. I. Olszewska, M. Barreto, J. Bermejo-Alonso, J. Carbonera, A. Chibani, S. Fiorini, P. Goncalves, M. Habib, A. Khamis, A. Olivares-Alarcos, E. P. de Freitas, E. Prestes, S. V. Ragavan, S. Redfield, R. Sanz, B. Spencer, and H. Li, "Ontology for autonomous robotics," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 189–194, 2017.

[16] S. R. Fiorini, J. Bermejo-Alonso, P. Gonçalves, E. Pignaton de Freitas, A. Olivares Alarcos, J. I. Olszewska, E. Prestes, C. Schlenoff, S. V. Ragavan, S. Redfield, B. Spencer, and H. Li, "A suite of ontologies for robotics and automation [industrial activities]," *IEEE Robotics Automation Magazine*, vol. 24, no. 1, pp. 8–11, 2017.

[17] A. Olivares-Alarcos, D. Beßler, A. Khamis, P. Goncalves, M. K. Habib, J. Bermejo-Alonso, M. Barreto, M. Diab, J. Rosell, J. Quintas, J. Olszewska, H. Nakawala, E. Pignaton, A. Gyrard, S. Borgo, G. Alenyà, M. Beetz, and H. Li, "A review and comparison of ontology-based approaches to robot autonomy," *The Knowledge Engineering Review*, vol. 34, p. e29, 2019.

[18] M. Stenmark and J. Malec, "Knowledge-based instruction of manipulation tasks for industrial robotics," *Robotics and Computer-Integrated Manufacturing*, vol. 33, pp. 56 – 67, 2015.

[19] S. Balakirsky, "Ontology based action planning and verification for agile manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 33, pp. 21 – 28, 2015.

[20] W. H. Chen, G. Foo, S. Kara, and M. Pagnucco, "Automated generation and execution of disassembly actions," *Robotics and Computer-Integrated Manufacturing*, vol. 68, p. 102056, 2021.

[21] S. Borgo, A. Cesta, A. Orlandini, and A. Umbrico, "Knowledge-based adaptive agents for manufacturing domains," *Engineering with Computers*, vol. 35, pp. 755–779, Jul 2019.

[22] V. R. Sampath Kumar, A. Khamis, S. Fiorini, J. L. Carbonera, A. Olivares-Alarcos, M. Habib, P. Goncalves, H. Li, and J. I. Olszewska, "Ontologies for industry 4.0," *The Knowledge Engineering Review*, vol. 34, p. e17, 2019.

[23] A. Umbrico, A. Cesta, G. Cortellessa, and A. Orlandini, "A holistic approach to behavior adaptation for socially assistive robots," *International Journal of Social Robotics*, vol. 12, pp. 617–637, Jul 2020.

[24] F. Ramos, A. S. Vázquez, R. Fernández, and A. Olivares-Alarcos, "Ontology based design, control and programming of modular robots," *Integrated Computer-Aided Engineering*, vol. 25, no. 2, pp. 173–192, 2018.

[25] F. Ramos, C. O. Scrob, A. S. Vázquez, R. Fernández, and A. Olivares-Alarcos, "Skill-oriented designer of conceptual robotic structures," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5679–5684, 2018.

[26] B. Bruno, N. Y. Chong, H. Kamide, S. Kanoria, J. Lee, Y. Lim, A. K. Pandey, C. Papadopoulos, I. Papadopoulos, F. Pecora, *et al.*, "The caresses eu-japan project: making assistive robots culturally competent," in *Italian Forum 2017 of Ambient Assisted Living*, pp. 151–169, Springer, 2019.

[27] D. Beßler, R. Porzel, M. Pomarlan, M. Beetz, R. Malaka, and J. Bateman, "A formal model of affordances for flexible robotic task execution," in *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*, pp. 2425–2432, IOS Press, 2020.

[28] E. Tulving, "Episodic memory: From mind to brain," *Annual Review of Psychology*, vol. 53, no. 1, pp. 1–25, 2002. PMID: 11752477.

[29] G. Bartels, D. Beßler, and M. Beetz, "Episodic memories for safety-aware robots," *KI - Künstliche Intelligenz*, vol. 33, pp. 123–130, Jun 2019.

[30] A. K. Bozcuoğlu, Y. Furuta, K. Okada, M. Beetz, and M. Inaba, "Continuous modeling of affordances in a symbolic knowledge base," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5452–5458, 2019.

[31] I. Androutsopoulos, G. Lampouras, and D. Galanis, "Generating natural language descriptions from owl ontologies: the naturalowl system," *Journal of Artificial Intelligence Research*, vol. 48, pp. 671–715, 2013.

[32] V. Nguyen, T. C. Son, and E. Pontelli, "Natural language generation from ontologies," in *Practical Aspects of Declarative Languages* (J. J. Alferes and M. Johansson, eds.), (Cham), pp. 64–81, Springer International Publishing, 2019.

[33] A.-C. Ngonga Ngomo, D. Moussallem, and L. Bühmann, "A holistic natural language generation framework for the semantic web," in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, (Varna, Bulgaria), pp. 819–828, INCOMA Ltd., Sept. 2019.

[34] H. Dalianis and E. Hovy, "Aggregation in natural language generation," in *Trends in Natural Language Generation An Artificial Intelligence Perspective* (G. Adorni and M. Zock, eds.), (Berlin, Heidelberg), pp. 88–105, Springer Berlin Heidelberg, 1996.

[35] M. Tenorth and M. Beetz, "Knowrob – knowledge processing for autonomous personal robots," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4261–4266, IEEE, 2009.

[36] M. Beetz, D. Beßler, S. Koralewski, M. Pomarlan, A. Vyas, A. Hawkin, K. Dhanabalachandran, and S. Jongebloed, "Neem handbook," [online], Institute for Artificial Intelligence (IAI), University of Bremen, 2020. https://ease-crc.github.io/soma/owl/current/NEEM-Handbook.pdf.

[37] S. Borgo, R. Ferrario, A. Gangemi, N. Guarino, C. Masolo, D. Porello, E. M. Sanfilippo, and L. Vieu, "Dolce: A descriptive ontology for linguistic and cognitive engineering," *Applied Ontology*, vol. 17, no. 1, pp. 45–69, 2022.

[38] A. Olivares-Alarcos, S. Foix, and G. Alenyà, *Human-Robot Collaboration: Unlocking the potential for industrial application*, ch. Time to Contact for Robot Safety Stop in Close Collaborative Tasks, pp. 1–17. The Institution of Engineering and Technology, to appear.

[39] H. P. Grice, "Logic and conversation," in *Syntax and Semantics: Vol. 3: Speech Acts* (P. Cole and J. L. Morgan, eds.), pp. 41–58, New York: Academic Press, 1975.

[40] Y. W. Lee, D. M. Strong, B. K. Kahn, and R. Y. Wang, "Aimq: a methodology for information quality assessment," *Information Management*, vol. 40, no. 2, pp. 133–146, 2002.