Intrinsic exploration motivation and incentives in cultural evolution

Esdras ASSIKIDANA
Grenoble Alpes University
Grenoble, France

Supervised by: Jérôme EUZENAT

I understand what plagiarism entails and I declare that this report is my own, original work.
ASSIKIDANA Esdras, 10/06/2022

Abstract

Cultural knowledge evolution deals with the evolution of knowledge representation in a group of agents. For that purpose, cooperating agents adapt their knowledge to the situations they are exposed to and the feedback they receive from others. This process was applied to ontology alignment repair, the improvement of incorrect alignments [Euzenat, 2017] and ontology evolution [Bourahla *et al.*, 2021]. These studies have shown that it converges towards successful communication through improving knowledge quality.

However, agents do not choose the situation they are exposed to (or the task that they perform). Hence, the process progress slowly since those situation and task are chosen randomly. Then we can get come across computation complexity when we are in large situation spaces.

This study tries to explore different kind of motivations for agents to choose the situations that they will get exposed to. To do so, agents are given mechanism to help them make their choices.

1 Introduction

1.1 Context

Agents living in a world use ontologies for representing this world. Thus, in case of an interaction with others, Agents use alignments which express the relations holding across ontologies, e.g., classes or properties, in the ontologies that they use. The result is a network of ontologies related by alignments that agents may use to interoperate.

This network is built through interactions between agents that apply corrective actions when communication fails. [Euzenat, 2014] showed that agents are able to repair alignments through playing a simple reclassification game. Here, the interactions are performed in an experimentation framework in which agents play “games” which was used in [Euzenat, 2017]. A Game implements the following pressures:

- Agents with different information (ontologies)
- Play a specific game or protocol (here guessing the type of an object)
- with a goal to achieve
- At the end of each game, they update the alignment between their ontologies.

In this paper we will use the same experimental framework used in [Euzenat, 2017]. Thus, with a population of agents for each game, two are selected randomly.
For example, we will show a game between two selected agents:

We have two agents that have their own ontology (Ontology of Agent A on the left and Agent B on the right, see Figure 1) of what is in the environment. These ontologies, identify the objects partially based on two of these features: for the agent A the color of the object either White or Black and lenght of the object either small or large. In addition to their ontologies, agents have access to an shared alignment:a set of shared correspondences (lign between the ontologies). These correspondences comprise equivalence correspondences between their top (all) classes and other correspondences.

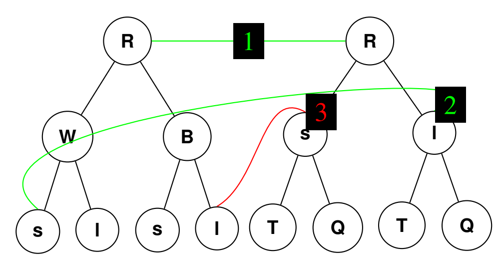

Figure 1: Ontologies for objects represent with 4 features R:(root) W:(white) B:(Black) s:(small) l:(large) T:(triangle) Q:(square)

Taking the correspondence 3 (see Figure 1), objects that are small for Agent B are objects that are Black and large. First thing you think is ”This is incorrect”, yes it happens that the correspondences between ontologies are false. The purpose is to correct them. From these ontologies, the environment is populated by objects characterised by three boolean features:

color = $\{white|black\}$, shape = $\{triangle|square\}$ and size = $\{small|large\}$.
Each object in this list belongs to one leaf class of the ontologies:

Agents play a very simple game:

agent A chose randomly an object "small-black-square"(■) and asks agent B for its class, this one will answer: black-large by using the correspondence 3. That result would be a FAILURE because small-black-square is not an subclass of black-large for agent A. To deal with failure, [Euzenat, 2017] defines six operators. In this paper we will use only four of those operators:
Assuming that the faulty correspondence $<C, r, C'>$ has been crossed by the object from C' to C we have

Replace
Delete the faulty correspondence and in addition replace in case r is =, the same correspondence with a $\leq$ relation is added ($<C, \leq, C'>$);

AddJoin
addjoin adds a correspondence between C' and the lowest superclass C'' of C compatible with D ($<C'', \geq, C'>$);

Refine
refine extends replace by adding a correspondence between C and the subclasses C'' of C' that do not subsume the actual class of the object ($<C, \geq, C''>$);

RefAdd
refadd is the combination of addjoin and refine. Here is an example using operator : refadd

- Assuming that the correspondence 3 (see Figure 1) has been used, removes the correspondence and
- add correspondence 4 (see Figure 2) $<C, \geq, C''>$ between class Black-large(ontology A) and the subclasses small-Triangle of small (ontology B) that do not subsume the actual class of the object and do not break functionality, i.e. there is not another correspondence $<X, \geq, Small - Triangle>$;
- add correspondence 3 (see Figure 2) $<C'', \geq, C'>$ between class small (ontology B) and lowest superclass Black (ontology A) of Black-large (ontology A)

After this operation we obtain the state in Figure 2

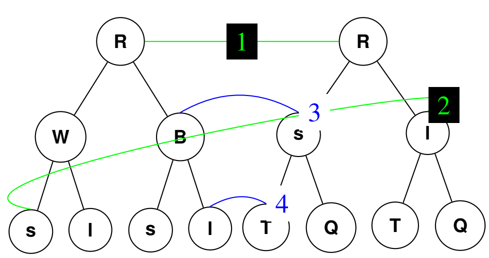

Figure 2: Ontologies after refadd operation

1.2 Goal

In [Euzenat, 2017], for each game, the two agents were chosen randomly, same thing for the object. Thus, this process is repeated till convergence of the correspondences between the agents. Meaning there is no possible game between any couple of agents or object that fail.

That politics take time and we certainly play some games that are unnecessary for the convergence. Hence, we are thinking if the agents with a form of curiosity inside, choose the object and/or agent with which they play, the convergence may be faster by not playing game that are not important. The goal is to find a mechanism for agents to make a decision on the agent to play with and the object for the game. For simplification issue we will focus on the mechanism to select an object. We will focus on game between two agents without considering the result of game with other agents.

2 Proposition

Inspired from [Oudeyer *et al.*, 2007] [Colas *et al.*, 2019], a way of choosing an object is to estimate the risk and the gain of playing with this object knowing the result from previous games. Here, the immediate goal of agent is to succeed in their interactions with other agents, but a secondary/supplementary goal may be, gain of information in the society. Gain of information in the society mean correct correspondences between agents in the society, in a way that agents understand each other correctly about object in the environment. Thus, to choose an object, agent have to measure the risk of failing a game with this object and the gain of information in the society by playing with this object.

2.1 Risk of failing game

The risk of failing a game represents the possibility, if we choose to play with an object o, that the interaction result will be a failure.

The first thing to observe is: "Before the first interaction between the agents, the risk of failing is maximum. This is due to the uncertainty, from the agent point of view, of the correctness of the correspondences".

Secondly, the risk to play with o will be null if we have a previous success interaction using a still existing correspondence with object o.

From this point, an object o' close to the object o regarding the ontology of the agent choosing should give less risk of failing. Due to the fact that an object of the same class as o on the ontology have now better chance to succeed since o already succeed. Then, while the correspondence that give this succeed interaction exists, the risk should diminish. When the correspondence is deleted, the risk of object o should go back to 1.

An object is an instance of many classes (eg: small-black-square instance of small, small-square, small-black etc..). So we propose to calculate the risk of playing on an object o : First, the agent should find the leaf class corresponding to the object o in his ontology and from that class leaf l we compute this risk :

$$R(o, l) = 1 - s(o, l)$$

Considering that the class leaf l can have n possible objects belonging to that class : if an interaction on one of this object already gave an success with an still existing correspondence, the risk of the others objects belonging to this class will diminish of $\frac{1}{n}$

Hence, we can establish

$$s(o,l) = \begin{cases} 1, \text{if there is still valid success with object } o \\ \sum_{i \in l} (\frac{1}{n}) * validsuccess(i), \text{ if not} \end{cases}$$

Let show how it works for an example :

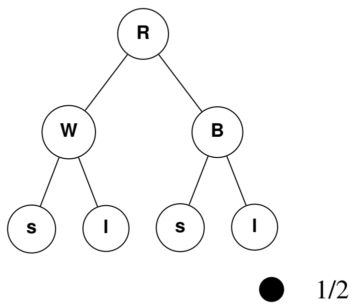

Figure 3: Ontology of an agent

Let take an agent with an ontology like on Figure 3, and an environment where object are identified by three features. Knowing that, on each leaf class, we have two different object that belong to that class. In this case, on leaf class (B black - s small) we have object (B black - s small - T triangle) and object (B black - s small - Q square). Normally at the start, every object presents a risk of one. If we play until we obtain a success with object (Black-small-triangle) for a game, the risk to play object (Black-small-Square) should diminish of $\frac{1}{2}$ and the risk of playing of object (Black-small-triangle) should be 0. As long as the correspondence that gives the success game on this object exists, the effect on this risk is still present.

2.2 Gain of information

Here, we are trying to find a metric to estimate the information we are gaining from playing with an object knowing the result of game with other objects. With this definition we already predict a relation between the gain and the risk we talk about earlier. Normally the more we have a risk the more we have a gain. So, left out the part we already include in the computation of the risk we can talk about another aspect of the gain of information.

First, each object in the world belongs to one and only one leaf class of the ontology of the agents. The agent knows all the correspondences incoming to a node on the path between the top of its ontology and this leaf class. For instance, in Figure 3, the leaf class Black-large of ontology A have three correspondences on his path (1, 3 and 4). The more correspondences they are, the more opportunity for the object to be brought there. It means during the interaction, there are a lot of chance that the ontology B will respond using one of those correspondences. Hence, the lowest the expected information gain. Thus object that belongs to leaf classes with less correspondences on his path to the top are expected to fail (then bring change on the correspondence between the agents). This being said if there is an succeed interaction on this object then the object can bring us more information if the correspondence used still exist.

From all of this, we can compute the gain on object o knowing o belong to leaf class c

$$G(o,c) = \begin{cases} 0, \text{if there is still valid success with object (o)} \\ \frac{1}{\#correspondence(c)}, \text{ if not} \end{cases}$$

$\#correspondence(c)$: the number of correspondences on the path from c to the top of the ontology

Let's see how it work with an example :

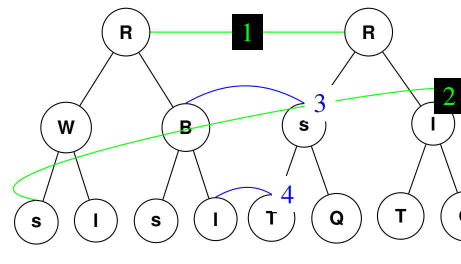

Figure 4: Example of correspondences between two ontologies

Agent A with his ontology on the left want to choose the object that will give him the higher gain of information. Knowing there is no success on any object for the moment:

- With object "Black-large-square" (■) and "Black-large-triangle" (▲) belonging to the leaf class "Black-large"; we can compute the gain of information for this two objects: $\frac{1}{3}$
- For the leaf classes "White-small" and "Black-small" we have at least two correspondences hence a gain of information $\frac{1}{2}$
- For the leaf class "White-large" we only have the root correspondence, which gives a gain of information of 1. Thus, object that will bring the maximum of gain of information are "White-large-Square" (□) and "White-large-Triangle" (△). Now, the agent will have to make his choice for the game.

2.3 Motivation and Metrics

Knowing these two metrics, given a certain goal or motivation, agent must choose object that correspond the most to his motivation. For example if the agent want to be ultra cautious, he will take object that have the less possible risk even if the risk is 0. With these two metrics here, we give three levels of motivation: Curious, Cautious and Ultra cautious.

Curious

An agent trying to be curious will select objects that give the maximum gain of information and after this selection, will select the object that has the minimal risk of failing game. This way he tries to correct the correspondence between the agents by choosing the maximum gain and after, try to minimize the risk.

```
for(object in World){
   risk = R(object, ...);
   gain = G(object, ...);
   if(gain > bestGain) {
      selectedObject = object;
      bestRisk = risk;
      bestGain = gain;
   }else if(gain == bestGain) {
      if(risk <= bestRisk && risk != 0) {
         selectedObject = object;
         bestRisk = risk;
         bestGain = gain;
      }
   }
}
```

Cautious

Here, an agent trying to be cautious, will try to have less failing games during his interactions. But, he also tries to correct the correspondences. In pursuit of this objective, he will choose first the objects with the minimal risk of failing game without considering those with a null risk (0). Meaning object on which gain of information is null. After this, he will select the one that give the maximal gain of information.

Ultra Cautious

Finally, with this motivation the agent objective is to have success interaction. Hence, he will choose only the object that gives him the minimal risk even if the information gain is null.

3 Hypotheses

From this idea we can establish these hypotheses:

Agents with an ultra cautious motivation when they get the first success with an object will choose to play with the same object indefinitely, thus

Hypothesis 1 : With the Ultra Cautious, Motivation the process should converge faster than the standard model and the others.

The same thing should happen with the cautious motivation. Since with the curious one, agents do not repeat games that are not important they should reach the correct alignment faster than the standard one, hence

Hypothesis 2: the Cautious and Curious Motivation models should converge faster than the standard model

Hypothesis 3 : The Cautious model should have the higher success rate than the standard and the curious ones

Hypothesis 4 : The curious model should have the best quality

4 Experimental Methodology

Our experimental framework is built on the one used in [Euzenat, 2017]. So during the experiment, agents have their ontologies and correspondences between them generated randomly at the start and then, they go through n iterations of the following:

- A pair of agents is selected randomly.
- One of them compute gain/risk and choose the object
- Play the game: the second agent guesses the type of the object using the correspondences
- Agents adapt their correspondences if the interaction failed.

Thus, the experiment depends on different factors.

4.1 Controlled Parameters

Meaning	Range
Number of agents	4
Number of iterations	2000
Exploration Model	Standard(Random), Cautious, Curious, UltraCautious
Revision Modality	AddJoin, Refine, RefAdd, Replace

We define an experiment plan to vary these parameters as presented in Table 1 and run each combination 10 times. Each time, we record some measures.

4.2 Measured Parameters

To assess the four presented hypotheses, we measure:

(1) Convergence: is the number of games (iterations) taken to converge in all cases when the process converges (here the observed maximum, not the average).

(2) Success rate [Steels, 2012] characterised by the ratio of successes over games played, which indicates how often agents have agreed on their decision

(3) Precision and recall measure the degree of correctness and completeness of the resulting correspondences with respect to the known correct reference correspondences.

5 Results and Discussion

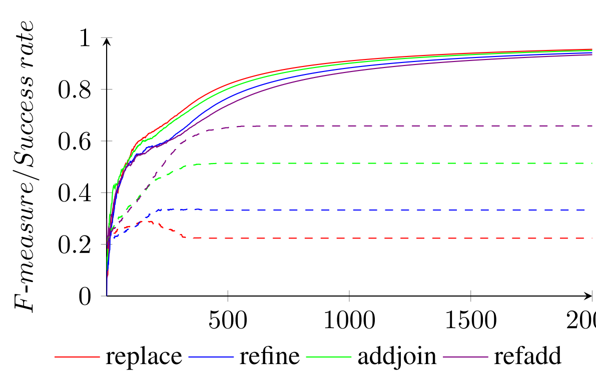

Figure 5: Success rate and F-measure for the curious mode.

Results of the experiment are resumed in figures.

Figure 5 present the F-measure and Success rate for the curious mode with the different operator.

From the result presented in figure 5 the distributions of the different operators are not different inside a defined exploration mode. Here, it is not important to compare the result between different operators combines with the exploration mode. We will use the result on the refadd operator.

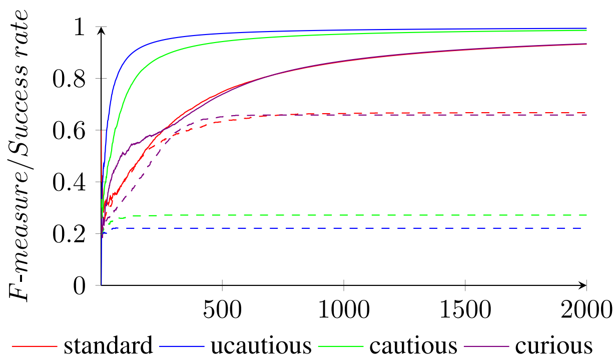

Figure 6: Success rate (solid) and F-measure (dashed) for the refadd operator.

Figure 6 present the F-measure and Success rate for the different type of exploration mode.

On figure 6, the first thing to note is that on ultra cautious exploration mode we have the highest success rate and it converges faster than the other exploration modes to his maximum value ≈ 1. Hence, **Hypothesis 1 is supported**.

Continuing with this figure, the cautious exploration mode also has higher success rate and converge faster than the curious and the standard model, **supporting Hypothesis 3**. But the plots of the curious and standard exploration mode are similar. We cannot judge on who is converging faster. Hence, we look at figure 7, to compare the precision and recall.

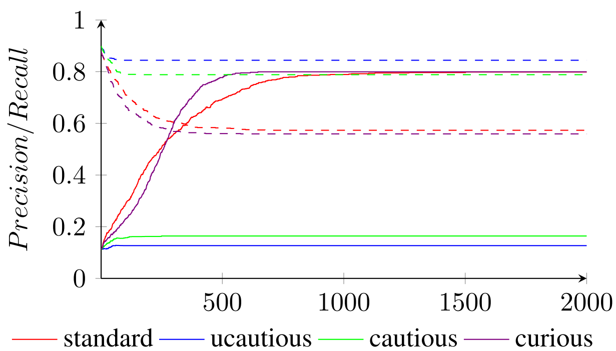

Figure 7: Precision (solid) and recall (dashed) for the refadd operator.

Figure 7 present the Precision and Recall for the different type of exploration mode. We can see that the precision of the curious model is not only the best, but converges faster than standard model. Thus, **Hypothesis 2 is supported**.

Regarding the success rate there is not such a significantly difference between the curious and standard exploration mode. But, the precision of the curious exploration mode is the best and converges faster than the standard mode. This was expected, since we favor exploration of unseen situations by searching the highest gain. Then, the correspondences are quickly corrected. The fact that the precision of the cautious and ultra-cautious exploration mode converges to less than 0.2 made them not good if the goal is to obtain correct correspondences between agents. They can not be consider as good quality models.

Nonetheless, at the end, the curious and standard model converge at the same level of precision. At the end, the standard exploration mode has the time to explore all the situation covered by the curious exploration mode. But the recall of the curious model and the standard one are very similar. This means even if we have a good proportion of correct correspondences there are not a good proportion of complete correspondences. Thus, we can not conclude that the curious model has the best quality. **Hypothesis 4 is rejected**.

One other thing to see, is that in none of the case the recall reaches one. From [Euzenat, 2017] , it is explained that this is due to the fact that the generation of the correspondences at the beginning are random and the correction depend on this first setting. Hence, we can not explore some situations, that will give us complete correspondences like the true ones in the solution. There are some correct correspondences that are shadowed. To solve this problem, [Euzenat, 2017] allows agents to return imprecise answers during the game. Thus, to test the quality of the curious model, we should take this in account and update the curious model.

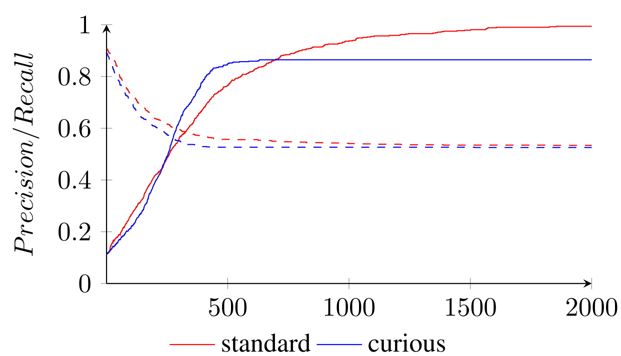

Figure 8: Precision (solid) and recall (dashed) for the refadd operator with 80% immediate consumption probability.

We use this solution and allow agents to return imprecise answers with an immediate consumption probability of 80%. This enables them to test shadowed correspondences. But from the result presented on figure 8, the precision of the standard converge to one as expected but the curious model stayed the same. The curious model does not behave as expected, he does not converge to one. There is a problem in the previous choice of object algorithm. During the exploration, we will obviously come across an iteration where all the gain on all the objects is null. When we reach this case, we must start selecting the object randomly, since there is no expected gain. Hence, the immediate consumption probability should have an effect, so we can test some shadowed correspondences.

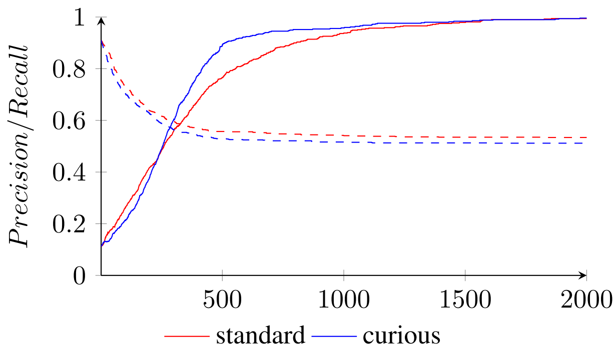

Figure 9: Precision (solid) and recall (dashed) for the refadd operator.

After this change on the algorithm we get the result presented on figure 9. As predicted the precision of the curious model converges to one and still converges faster than the standard one.

Regarding the recall, the standard is still better than the curious one. Because the random model has more chance to test some unseen cases that will help him get some completeness more than the curious where the exploration is oriented.

To conclude, we can say that the curious model is faster (converges faster) than the standard one, as predicted, but we can not say that he has the best quality. We can not also say that he is worse than the standard one.

6 Conclusion

We established metrics for agents to evaluate their chance to have successful interactions and evaluate the expected gain of information on a specific interaction. With these two metrics, agents can now choose the object for an interaction with respect to its motivation. We repeated experiments with three different types of motivation: Curious, cautious and ultra-cautious. And regarding the correctness of the correspondences, in result, the cautious and ultra-cautious models are the worse even if they have the best success rate. The curious model has the best correct correspondences rate and converges faster than the standard one.

Since the model does not allow agents to generate correspondences by themselves, agents were not able to get fully complete correspondences. Thus, we can not conclude between the curious model and the standard one who is the best.

The study is open to be completed by taking this into account, and furthermore to take into account the choice of the agent with which to play regarding the gain of information or risk of failing game.

References

[Bourahla *et al.*, 2021] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. Knowledge improvement and diversity under interaction-driven adaptation of learned ontologies. *20th ACM international conference on Autonomous Agents and Multi-Agent Systems*, pages 242–250, May 2021.

[Colas *et al.*, 2019] Cédric Colas, Pierree-Yves Oudeyer, and Olivier Sigaud. Curious: Intrinsically motivated modular multi-goal reinforcement learning. *Proc. 36th ICML*, pages 1331—-1340, 2019.

[Euzenat, 2014] Jérôme Euzenat. First experiments in cultural alignment repair (extended version). In *Proc. ESWC 2014 satellite events revised selected papers*, number 8798, pages 115–130. Lecture notes in computer science, 2014.

[Euzenat, 2017] Jérôme Euzenat. Communication driven ontology alignment repair and expansion. *Proc. 26th IJCAI*, pages 185–191, 2017.

[Oudeyer *et al.*, 2007] Pierre-Yves Oudeyer, Frédéric Kaplan, and V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286, 2007.

[Steels, 2012] Luc Steels. Experiments in cultural language evolution. *John Benjamins*, 2012.