Controlled Conversational Models through Conversation-Dedicated Ontology

Barbara Gendron, under the supervision of Mathieu d'Aquin and Gaël Guibon

Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), Université de Lorraine

✉ barbara.gendron@loria.fr | 🖥 b-gendron.github.io | 🐙 b-gendron | 💼 barbara.gendron

UNIVERSITÉ DE LORRAINE

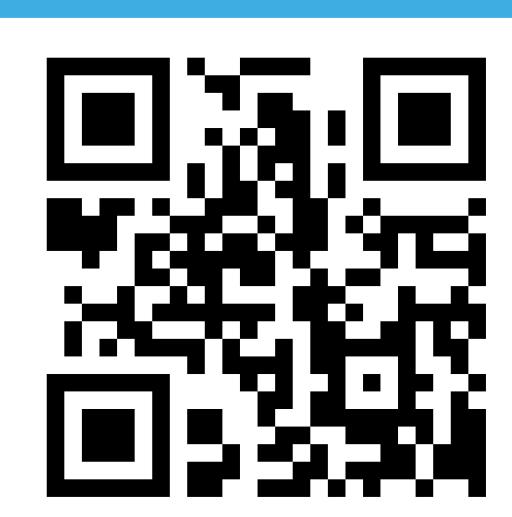loria — Laboratoire lorrain de recherche en informatique et ses applications

1. Context

Recent advances in Large Language Models (LLMs) have improved conversational agents' realism and compliance towards human requirements and needs. However, **controlling conversation flow towards positive outcomes remains crucial**. This Ph.D. aims to **represent conversational knowledge using an ontology to enable language model control**. Ontologies allow to model the knowledge in a domain, defining concepts and characterizing relations between them. While often used for domain-specific knowledge, few have explored using ontologies to guide conversation flow. Convology [**?**] is a recent example focusing on managing health conversations. We plan to extend Convology's conceptualization capacities to a more general setup, therefore adaptable to general-purpose user/agent conversations.

Want to know more?
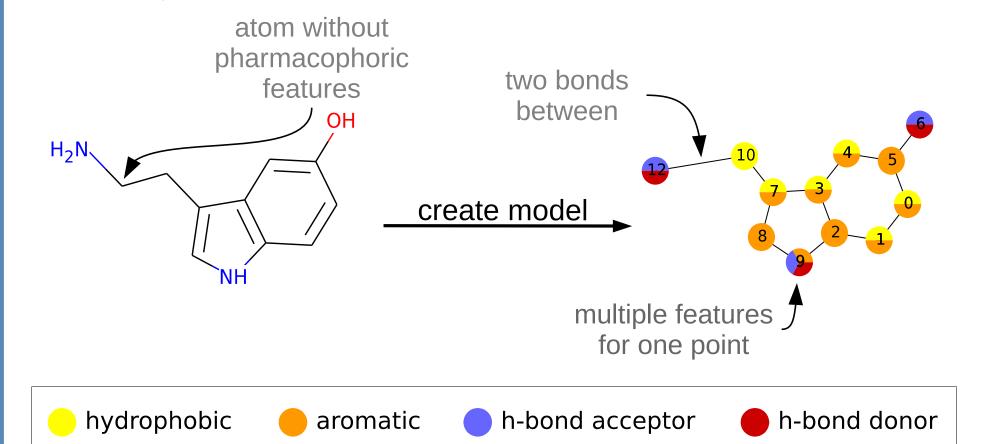

My website, gh repo or whatever

2. Methodology

PhD approach. Iterative process by progressive enrichment of the ontology to conceptualize more and more notions related to conversations.

Tools. Protégé, HermiT and Pellet reasoners, owlready2, rdflib, PyTorch, huggingface transformers and parameter-efficient fine-tuning libraries, LoRA adapters.

Conceptualization. Explain our ontology engineering process and how we select and create toy examples.
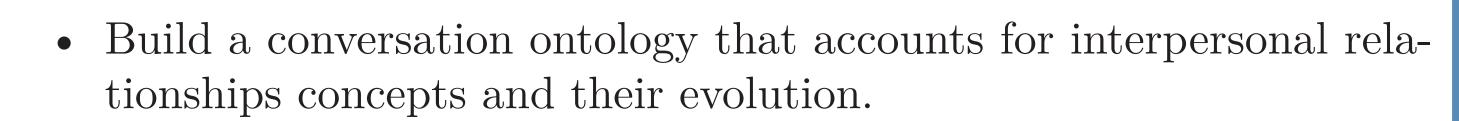
Control. Explain how we plan to control the conversational dynamic in the context of a dialogue user-agent.

Challenges. It is not straightforward that what the ontology brings can be fully understood and learnt by an LLM (in our case, we consider for now only decoder-based language models).

- hydrophobic
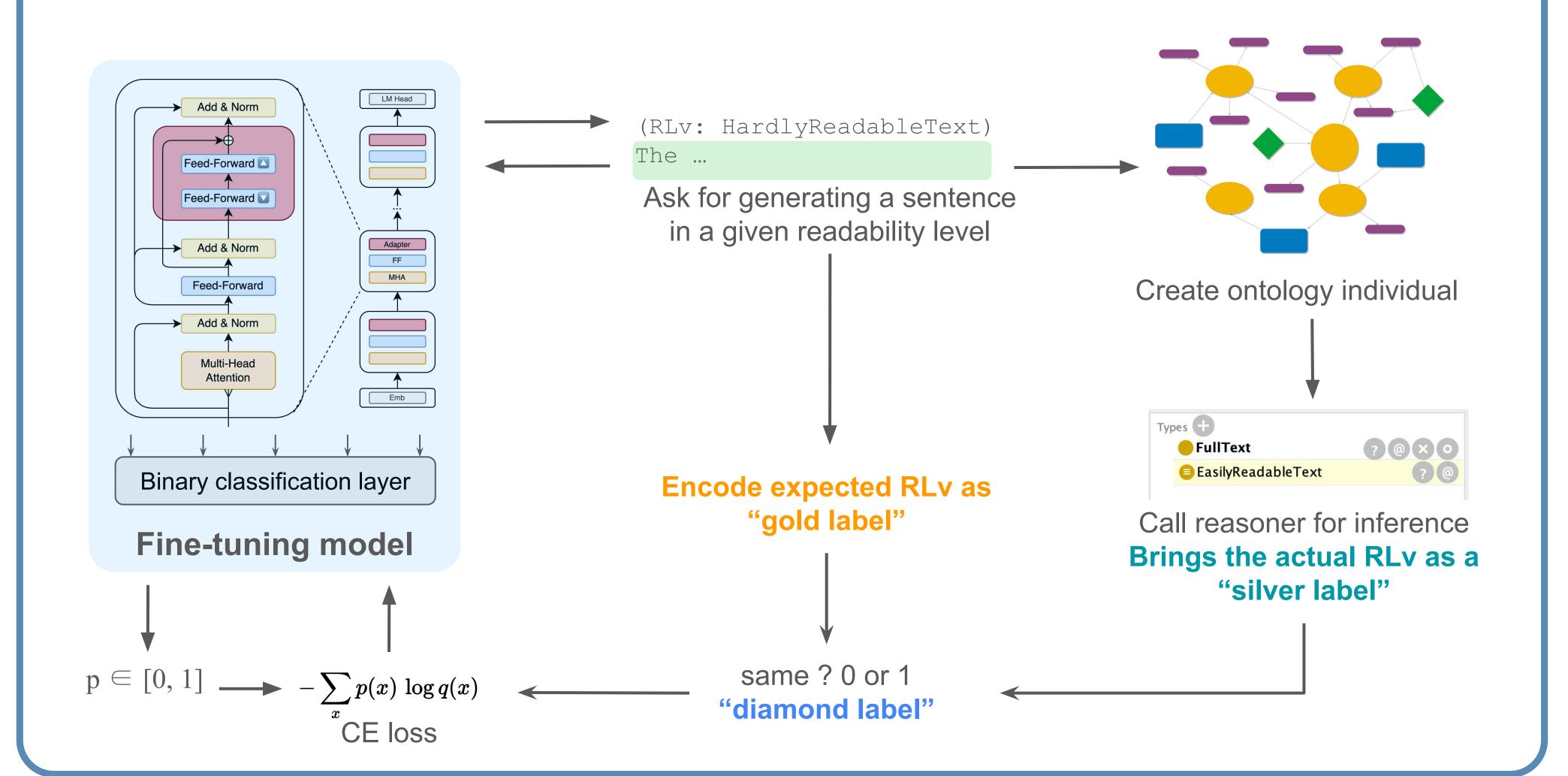- aromatic
- h-bond acceptor
- h-bond donor

3. Objectives

Develop **knowledge-enhanced conversational models** that exploit Large Language Models (LLMs) and Ontologies. The aim is to provide **structured knowledge** to open-domain conversational agents [**?**].

- Build a conversation ontology that accounts for interpersonal relationships concepts and their evolution.
- Integrate and assess ontology understanding during fine-tuning.
- Bring control on conversational LLM outputs through encapsulated conversation knowledge.

4. OntoGPT: LLM Fine-Tuning Based on Ontology Validation

We work on a LLM/Ontology hybridation setup where the ontology is supposed to help the LLM produce more accurate utterances in dialogue (being as compliant as possible towards the user requirements). For this, we provide a an end-to-end integration pipeline where the ontology information is assimilated at fine-tuning time.

(RLv: HardlyReadableText)
The …

Ask for generating a sentence in a given readability level

Create ontology individual

Encode expected RLv as "gold label"

Call reasoner for inference
Brings the actual RLv as a "silver label"

same ? 0 or 1
"diamond label"

Fine-tuning model

Binary classification layer

$p \in [0, 1]$ → $-\sum_x p(x) \log q(x)$
CE loss

5. Current Stage of Research

This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective. This is the place where I explain what I have done so far and what I plan to do in a short-time perspective.

7. Perpectives

We expect that the use of the ontology will also increase the interpretability of conversational agents, as decisions to use certain outputs over others will be associated with the ontological dimensions of the current conversation that have driven those choices, thus, moving from common fully black-box NLP models. This could also help these models be less harmful to both users and providers, along with controllable ethics in the way the conversation is analyzed.