

# Application Level Explanations for Argumentation-based Decision Making

---

NIKOLAOS I. SPANOUDAKIS<sup>1</sup>, ANTONIS C. KAKAS<sup>2</sup> AND ADAMOS KOUMI<sup>2</sup>

<sup>1</sup>SCHOOL OF PRODUCTION ENGINEERING AND MANAGEMENT,  
TECHNICAL UNIVERSITY OF CRETE

<sup>2</sup>DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF CYPRUS

# Outline

---

- ❑ Introduction - Contributions
- ❑ Building explanations from Gorgias preference-based argumentation framework results
- ❑ Application examples
- ❑ Conclusion and Future Work

# Contributions

---

- ❑ We show how the returned results of the dialectical argumentation reasoning within the Gorgias framework can be exploited to provide human-readable explanations that are
  - Attributive
  - Contrastive
  - Actionable
- ❑ These results, can be manipulated by applications to produce case-based human readable explanations.

# A gorgias theory example (code)

---

```
rule(r1(X), buy(X), []):- need(X).  
rule(r2(X), neg(buy(X)), [neg(urgentNeed(X))]).  
rule(pr1(X), prefer(r2(X), r1(X)), [lowOnFunds]).  
rule(pr2(X), prefer(r1(X), r2(X)), []).  
rule(c1(X), prefer(pr1(X), pr2(X)), []).  
abducible(urgentNeed(X), []).  
abducible(neg(urgentNeed(X)), []).
```

# A gorgias query example

---

- ❑ Suppose that we know that the agent needs a bag and it believes it is low on funds
  - need(bag)
  - lowOnFunds -> rule(b1, lowOnFunds, [])
- ❑ Let's try the query to not buy the bag
  - neg(buy(bag))
- ❑ It is valid, the argument is
  - ass(neg(urgentNeed(bag))), c1(bag), b1, pr1(bag), r2(bag)

# Explanation generator

Contrastive

Actionable

## Internal explanation:

- `b1, r2(bag), pr1(bag), c1(bag), ass(neg(urgentNeed(bag)))`
  - `rule(r1(X), buy(X), []):- need(X).`
  - `rule(r2(X), neg(buy(X)), [neg(urgentNeed(X))]).`
  - `rule(pr1(X), prefer(r2(X), r1(X)), [lowOnFunds]).`
  - `rule(pr2(X), prefer(r1(X), r2(X)), []).`
  - `rule(c1(X), prefer(pr1(X), pr2(X)), []).`
  - `abducible(urgentNeed(X), []).`
  - `abducible(neg(urgentNeed(X)), []).`

# Gorgias Cloud Application Level Explanations

Gorgias cloud

ExecutionPanel MyProjects Logout

Gorgias/Background Files

assistance/buying3.pl

```
rule(r1(X), buy(X), []):-need(X).
rule(r2(X), neg(buy(X)),
[neg(urgentNeed(X))]).
rule(pr1(X), prefer(r2(X), r1(X)),
[lowOnFunds]).
rule(pr2(X), prefer(r1(X), r2(X)), []).
rule(c1(X), prefer(pr1(X), pr2(X)), []).
abducible(urgentNeed(X), []).
abducible(neg(urgentNeed(X)), []).
```

• prove([neg(buy(bag))], InternalExplanation).  
Solution 1

Internal Explanation: [ass(neg(urgentNeed(bag))),b1,c1(bag),pr1(bag),r2(bag)],  
' Application Level Explanation

The statement "neg(buy(bag))" is supported by:  
- "neg(urgentNeed(bag))" and "lowOnFunds"

This reason is :  
- Stronger than the reason of "need(bag)" supporting "buy(bag)"

The supporting condition: "neg(urgentNeed(bag))" is an assumption and needs to be confirmed.

Scenario Test Files

assistance

Add/Expand scenario

Add scenario..

need(bag)  
rule(b1, lowOnFunds, [])

Gorgias Prompt

Maximum number of answers: 1

Clear panel

Gorgias?:neg(buy(bag))

Run

Attributive

Contrastive

Actionable

# A Social Media Application Example

---

- ❑ The agent browses social media content
- ❑ Sets each item's priority to:
  - Important
  - Normal/default
  - Hide
- ❑ Policy:
  - ... Posts that come from the user's manager are **important** regardless of whether they are positive or negative. ... **Hide** politics posts from the user's manager when negative. ...



# A Social Media Application Example

## Social Media Assistant

Post Title

Presidential election

Post Characteristics

- The post comes from my manager.
- The post is on politics.
- The post topic is negative.

Contrastive

Attributive

Hide post

Explanation:

Generally when the post comes from your manager, we prefer to set the priority to important, but because **the post topic is in on politics and it is negative** we hide the post.

# Application Example 1: Medica

---

- ❑ This system aims to aid the decisions of administrative personnel in the health domain
- ❑ They need to decide what information can be disclosed to a person asking for it
- ❑ EU legislation defines the access rights

# Application Example 1: Medica

Medical Access Control

HomeData FormLog out

Medical records access form

Patient Id: \*

123

Q

File Id: \*

10

Select the person requesting Access to Medical Records: \*

Doctor

▼

Is there a written consent from the Patient? \*

☐ Yes

☒ No

Is there an order from the Medical Association \*

☐ Yes

☒ No

Access Reason: \*

☐ Publishing to scientific journal

☒ Treatment purpose

☐ Data Processing Purposes

☐ Personal Use

☐ Educational Purposes

☐ Research Purposes

SUBMIT

11

# Application Example 1: Medical

Medical Access Control

Back Log out

Your level of access is: **Limited Plus Access**

## Legislation

The Safeguarding and the Patients' Rights Protection Act 2004, Article 15, paragraph 2b

## Explanation

The doctor has limited plus access to the owner's data for therapy/medical use.

Do you need higher level of access for the same purpose?

Do you need higher level of access for other purpose?

Attributive:  
Rule label  
connected  
to free text

Actionable:  
Put the  
abducibles  
to work

# Application Example 1: Medical\

The screenshot shows a web application titled "Medical Access Control" with a dark header containing "Back" and "Log out" links. A modal dialog is open, featuring a blue header with the text "Request higher possible level of access for the same purpose" and a close button (X). The dialog body contains the text: "The highest level of access you can get is **Full Access**" and "You need **written consent from the patient**, in order to gain full access." An "OK" button is located at the bottom right of the dialog. Below the dialog, a table with two columns is visible. The left column contains the text: "The Safeguarding and the Patients' Rights Protection Act 2004, Article 15, paragraph 2b". The right column contains the text: "The doctor has limited plus access to the owner's data for therapy/medical use." Below the table, there are two buttons: a dark blue one with the text "Do you need higher level of access for the same purpose?" and a dark red one with the text "Do you need higher level of access for other purpose?". A callout bubble on the left side of the dialog contains the text: "Actionable: You can take this action to get the desired level".

Medical Access Control

Back Log out

Request higher possible level of access for the same purpose

The highest level of access you can get is **Full Access**

You need **written consent from the patient**, in order to gain full access.

OK

The Safeguarding and the Patients' Rights Protection Act 2004, Article 15, paragraph 2b	The doctor has limited plus access to the owner's data for therapy/medical use.
---	---

Do you need higher level of access for the same purpose?

Do you need higher level of access for other purpose?

Actionable: You can take this action to get the desired level

# Application Example 2: GAID

---

- ❑ GAID: Gynecological AI Diagnostic Assistant
- ❑ This system aims to aid a medical doctor in disease diagnosis
- ❑ The system gets the tests and symptoms of a patient
- ❑ Doctor's knowledge determines the outcome



# Application Example 2: GAID



# GAID

Gynaecological Artificial Intelligence Diagnostics - Cognitive Assistant



Current Patient: *Sylvia Tina Jackson Heath (ID: 5544332211)*

Current Date: *Aug 11 2021, Wed*

 Overview

Presenting Complaints:



Enter Presenting Complaints...



Vaginal Discharge ✓ + ✕

Burning ✓ + ✕

Current Visit Information:



Update Patient Record

Past Relevant Visits

Relevant Patient Record

Other Symptoms:



Post-Coital Bleeding



Post-Coital Bleeding ✓ + ✕

Clinical Examinations Findings:



Enter findings...



Diagnosis

Filter by Suspicion:

ALL 123 12

Suspicious 12

Non-Suspicious 123

Filter by Group:

Sexually Transmitted Infections 5 5

Bleeding & Vascular Problems 28 4

⌵ Pelvic Pain 56 0

Urogynecology & Prolapses 9 0

Infertility 4 0

Gynecological Cancers 7 2

Endocrinology 27 0

Vulva Pathologies 0 3

Sundries 1 0

Sexually Transmitted Infections

Vulva Candidiasis

Trichomoniasis

Neisseria Gonorrhoeae

Chlamydia

Bacterial Vaginosis

Anogenital Warts

Hepatitis

HIV/AIDS

HSV II

Syphilis

# Application Example 2: GAID

**Diagnosis:**  
*Anogenital Warts*

Attributive

EXPLANATION

RELATED INFO

ICD-10

*Under the information **Vaginal Burning** it is recommended that you investigate **Anogenital Warts**.*

*This decision is supported by: **Inter-Menstrual Bleeding**.*

*The following further information strengthens this decision: **Dysuria**.*

*But each of the following information:*

- **Image: No indication of small cauliflower-shaped lumps**
  - **History: Vaccinated with HPV before first intercourse**
- indicates, that this disease may not be possible and could be **excluded**.*

Contrastive

Actionable:  
Possible issues to  
investigate in order  
to exclude other  
possible diseases



# Conclusions

---

- ❑ We have delved into explainable AI-base decisions that are
  - Attributive
  - Contrastive
  - Actionable
- ❑ Still more work is needed
  - NLP: generate predicates and arguments from human generated text
  - Generate arguments explanations (Attributive, Contrastive or Actionable) in free text

# Thank you, questions?

---

## **Nikos Spanoudakis**

Researcher - Teaching Staff

Applied Mathematics and Computers Laboratory,  
School of Production Engineering and Management,  
Technical University of Crete,

Email: [nikos@amcl.tuc.gr](mailto:nikos@amcl.tuc.gr)

<https://users.isc.tuc.gr/~nispanoudakis>



The Gorgias Cloud System will be a demo at COMMA 22 (Wednesday)

Open for academic use: <http://gorgiasb.tuc.gr/GorgiasCloud.html>