

به نام خدا



دانشگاه اصفهان

دانشکده مهندسی کامپیوتر

پروژه اصلی - مرحله سوم - بخش «ب»

مباحث ویژه ۲ - اخلاق محاسباتی

کد تکلیف			UI-ET-2024-3620028-02-PRJ01	
موضوع			پروژه اصلی مرحله سوم بخش «ب»	
تاریخ تحویل			۴ خرداد ماه ۱۴۰۳ شمسی	
کد فعالیت تیم			P2	
نام تیم			P2	
نام و نام خانوادگی		شماره دانشجویی	زمینه فعالیت	اقدامات در این تمرین
۱	محمد حسین دهقانی	۹۹۳۶۱۳۰۲۷	P P	سناریو معضل دروغ‌گویی و گزارش
۲	ابوالفضل شیشه‌گر	۹۹۳۶۲۳۰۲۵	D P	سناریو ۲
۳	محمد امین صابری	۹۹۳۶۲۳۰۲۶	P D	سناریو ۱
۴	حسین علی‌ترکان	۹۸۳۶۱۳۰۴۰	D P	سناریو ۳
۵	احسان محمودی	۹۹۳۶۱۳۰۵۴	D P	سناریوهای واکن فراری و قایق

گزارش پیاده‌سازی و تحلیل سناریوهای اخلاقی با استفاده از کتابخانه HERA

مقدمه

این پروژه با هدف پیاده‌سازی و تحلیل سناریوهای مختلف اخلاقی با استفاده از [کتابخانه HERA](#) در زبان پایتون انجام شده است. در این پروژه، سناریوهای معروفی مانند معضلات واگن فراری و قایق مورد بررسی قرار گرفته‌اند. تحلیل‌ها از دیدگاه مکاتب مختلف اخلاقی از جمله مکتب فایده‌گرایی و مکتب صدرایی انجام شده است. از طریق [این لینک](#) می‌توانید به فایل‌های پروژه دسترسی داشته باشید.

روش‌ها

تغییر سناریوها به فرمت سازگار با HERA

برای شبیه‌سازی سناریوهای ذکر شده، لازم بود تا آن‌ها را از فرم منطق درجه اول به حالتی تغییر دهیم که با کتابخانه HERA سازگار باشد. بر اساس مستندات موجود در ریپازیتوری کتابخانه HERA، سناریوها با استفاده از فایل‌هایی به فرمت YAML توصیف شدند. این فایل‌ها شامل بخش‌های زیر هستند:

- کنش‌ها (actions): شامل تمامی کنش‌های موجود در سناریو به همراه ویژگی‌های آن‌ها مانند نام کنش، شرایط لازم برای انجام، ماهیت ذاتی (خوب، بد، خنثی) و تاثیرات آن کنش در محیط سناریو.
- رویدادها (events): شامل تمامی اتفاقات ممکن در سناریو به همراه جزئیات آن‌ها مانند نام رویداد، شرایط لازم برای رخ دادن، تاثیرات رویداد و ترتیب زمانی تاثیرات.
- ارزش‌ها (utilities): مشخص کردن ارزش هر کدام از تاثیرات که می‌تواند مثبت یا منفی باشد.
- تاثیرات (affects): تعیین ماهیت مثبت یا منفی هر یک از تاثیرات برای موجودیت‌های سناریو.
- برنامه (plan): مشخص کردن کنش یا دنباله‌ای از کنش‌های مختلف به همان ترتیبی که در سناریو اتفاق می‌افتند.
- هدف (goal): تعیین هدفی که کنش‌ها در سناریو دنبال می‌کنند.
- حالت آغازین (initialState): مشخص کردن ویژگی‌های موجودیت‌های مختلف در شروع سناریو.

پیاده‌سازی در پایتون

پس از مشخص کردن جزئیات ذکر شده در فایل YAML، برنامه‌ای در زبان پایتون نوشته شد تا فایل توصیف سناریو را بخواند و با استفاده از کتابخانه HERA به تحلیل سناریو از دیدگاه مکاتب اخلاقی گوناگون بپردازد.

نتایج

نحوه استفاده از برنامه

برای استفاده از برنامه، پس از کلون کردن ریپازیتوری، فایل توصیف سناریو در پوشه plans واقع در پوشه ethics قرار داده می‌شود و سپس با استفاده از دستوری به فرمت زیر، خروجی برنامه را روی سناریوی تعریف شده مشاهده می‌کنیم:

```
python SituationEvaluator.py <ScenarioName>
```

مثال‌هایی از اجرای برنامه

اجرای برنامه روی سناریو شماره ۲:

```
1 !python SituationEvaluator.py scenario2

Deontology: False
GoalDeontology: True
Kantian: True
Kantian Reading #2: False
DoNoHarm: True
DoNoInstrumentalHarm: True
Utilitarianism: True
DoubleEffectPrinciple: False
AvoidAvoidableHarm: True
AvoidAnyHarm: True
```

اجرای برنامه روی سناریو معضل دروغگویی:

```
1 !python SituationEvaluator.py liar_robot

Deontology: False
GoalDeontology: True
Kantian: True
Kantian Reading #2: True
DoNoHarm: True
DoNoInstrumentalHarm: True
Utilitarianism: True
DoubleEffectPrinciple: False
AvoidAvoidableHarm: True
AvoidAnyHarm: True
```

بحث

با تحلیل نتایج به دست آمده، می‌توان نتیجه بررسی هر کدام از سناریوها را در مکاتب اخلاقی گوناگون مشاهده کرد. کتابخانه HERA با پوشش مکاتب اخلاقی مختلف و ارائه نتایج دقیق، ابزار قدرتمندی برای تحلیل سناریوهای اخلاقی است. با این حال، مشکلاتی نظیر کدهای کثیف و غیر اصولی و عدم مستندسازی مناسب، توسعه و افزودن ویژگی به آن را دشوار می‌سازد.

این پروژه نشان داد که استفاده از کتابخانه HERA برای تحلیل سناریوهای اخلاقی امکان‌پذیر و مفید است. با وجود مشکلاتی که در توسعه و مستندسازی این کتابخانه وجود دارد، نتایج به دست آمده نشان‌دهنده دقت و قدرت تحلیل آن در مکاتب اخلاقی مختلف است. بهبود مستندسازی و ساختار کدهای این کتابخانه می‌تواند کاربرد و توسعه آن را بیشتر کند.