Applying Deontic Logic

What is Deontic Logic?

- It is a field of philosophy that is concerned with obligations and permissions.

Why is it important?

- When an LLM makes a decision it has to know what is right and wrong through logical statements. This logical framework provides the fundamentals for understanding these ethical concepts which can pertain to many different applications including law, common sense, and much more.

What are the types of rules used in this logic?

- Rules can be used to represent
    - Obligations (prefixed with ‘O’)
    - Permissions (‘P’)
    - or whether or not an action is Forbidden (‘F’).
- Some examples:
    - O(bringing_a_pet) means you are *obligated* to bring a pet.
    - F(bringing_a_pet) means you are *forbidden* from bringing a pet.
    - P(bringing_a_pet) means you are *permitted* to bring a pet.

Putting the rules together.

- You can put multiple rules together to form relationships between each of them or form a new rule
- P(bring_cat) ^ P(bring_dog) means you can bring a cat OR a dog. While ‘^’ is usually interpreted as an ‘and’ what we are saying here is that it is permissible (NOT NECESSARILY OBLIGATORY) to bring a cat or a dog.
- O(bring_cat → bring_dog) means it ought that you bringing a cat implies also bringing a dog.

Some precautions

- → can be misinterpreted as, you ought to do A because you must B.
- However, → should be rather interpreted as if A happens then B should follow