Lesson #1: Introduction to Formal Logic

What's It All About?

A brief introduction to formal logic with an eye towards discussing resolution and unification in a sketch of Prolog

Logic? What is is? What is it good for?

1. Logic is a **knowledge representation**. Actually, since there are so many logics, it is any number of knowledge representations.
2. Logic is a **reasoning mechanism**. Actually, many!
3. In summary, logic is useful for both **representation** and **reasoning**. Consequently, is is used as a basis for implementing problem solving systems, question answering systems, rule-based expert systems, and a host of other sorts of representation rich AI systems.
4. The appeal of logic is that it provides a **middle ground with respect to naturalness of expression and computational viability**.
5. We will focus on two particular logics in this course, **propositional logic**, and **predicate logic**. As you might guess, propoistional logic focusses on "propositions", and predicate logic focusses on "predicates". The notes to follow represent a lecture on these notions, and related matters. (It should be noted that a number of other logics are valued by cognitive scientists and artificial intelligence researchers.)

Propositions

A **proposition** is a simple declarative sentence that is either true or false, but not both. For example:

1. Coffee is black. (true)
2. Java is a functional programming language. (false)

One often represents propositions with upper case letters, such as P or Q or R or S or T. Propositions are sometimes called **atoms**.

Well Formed Formulas (WFFs)

- **Propositional expressions** can be formed in the propositional logic by combining propositions with propositional operators in appropriate ways. The propositional operators are $\wedge$ $\vee$ $\rightarrow$ $\leftrightarrow$ $\sim$ which are interpreted as "and" "or" "if then" "if and only if" "not". The first four are binary operators. The last one is a unary operator.

- Propositional expressions, including atomic expressions, are generally known as **well formed formulas**, or **WWFs**.
- The **syntax** of WFFs may be described formally in the following way:
  - An atom is a WFF.
  - If F is a WFF, then $(\sim F)$ is a WFF.
  - If F and G are WFFs, then $(F \wedge G)$, $(F \vee G)$, $(F \rightarrow G)$, and $(F \leftrightarrow G)$ are WFFs.
  - Nothing else is a WFF.
- In terms of **semantics**, as you probably already know, the logical operators are defined in the following ways:
  - $(\sim F)$ is true when F is false, and false when F is true.
  - $(F \wedge G)$ is true if both F and G are true; otherwise, $(F \wedge G)$ is false.
  - $(F \vee G)$ is true if at least one of F or G is true; otherwise, $(F \vee G)$ is false.
  - $(F \rightarrow G)$ is false if F is true and G is false; otherwise, $(F \rightarrow G)$ is true.
  - $(F \leftrightarrow G)$ is true if F and G have the same truth values, otherwise $(F \leftrightarrow G)$ is false.
- Since it is useful to name things, let's agree to use the following **vocabulary** for the logical operations:
  - $(\sim F)$ is the **negation** of F.
  - $(F \wedge G)$ is the **conjunction** of F and G.
  - $(F \vee G)$ is the **disjunction** of F and G.
  - $(F \rightarrow G)$ is the **implication** of F and G.
  - $(F \leftrightarrow G)$ is the **equivalence** of F and G.

Interpretations of WFFs

Suppose that F is a WFF containing atoms A1, A2, ..., An. An **interpretation of F** is an assignment of truth values to A1, A2, ..., An.

With respect to interpretations, the following two definitions are of great significance:

- A formula F is said to be **true under (in) an interpretation** if and only if F evaluates to true in the interpretation.
- A formula F is said to be **false under (in) an interpretation** if and only if F evaluates to false in the interpretation.

Questions

1. What can you say about $F = ((P \wedge Q) \rightarrow (R \leftrightarrow (\sim S)))$ under the interpretation PQRS = TTFF?
   (a) formula F is true under the interpretation PQRS = TTFF
   (b) formula F is false under the interpretation PQRS = TTFF

2. What can you say about F = ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))) under the interpretation PQRS = FFTT
   (a) formula F is true under the interpretation PQRS = FFTT
   (b) formula F is false under the interpretation PQRS = FFTT

Examples: Computing the WFF value for a given interpretation

1. What can you say about F = ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))) under the interpretation PQRS = TTFF?
   (a) ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))) – the focus of our thoughts
   (b) ((T $\wedge$ T) $\rightarrow$ (F $\leftrightarrow$ ($\sim$ F))) – substituting values for variables
   (c) ((T $\wedge$ T) $\rightarrow$ (F $\leftrightarrow$ T)) – definition of $\sim$
   (d) (T $\rightarrow$ (F $\leftrightarrow$ T)) – definition of $\wedge$
   (e) (T $\rightarrow$ F) – definition of $\leftrightarrow$
   (f) F – definition of $\rightarrow$

   Thus, formula F is false under the interpretation PQRS = TTFF.

2. What can you say about F = ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))) under the interpretation PQRS = FFTT
   (a) ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))) – the focus of our thoughts
   (b) ((F $\wedge$ F) $\rightarrow$ (T $\leftrightarrow$ ($\sim$ T))) – substituting values for variables
   (c) ((F $\wedge$ F) $\rightarrow$ (T $\leftrightarrow$ F)) – definition of $\sim$
   (d) (F $\rightarrow$ (T $\leftrightarrow$ F)) – definition of $\wedge$
   (e) (F $\rightarrow$ F) – definition of $\leftrightarrow$
   (f) T – definition of $\rightarrow$

   Thus, formula F is true under the interpretation PQRS = FFTT.

Satisfaction and Falsification of WFFs under interpretation

If a formula F is true under an interpretation I, then we say that **I satisfies F**, or that **F is satisfied by I**.

If a formula F is false under an interpretation I, then we say that **I falsifies F**, or that **F is falsified by I**.

Examples: Satisfies/Falsifies

1. Interpretation PQRS = TTFF **falsifies** formula ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))), since the formula evalutes to **false** under the given interpretation.
2. Interpretation PQRS = FFTT **satisfies** formula ((P $\wedge$ Q) $\rightarrow$ (R $\leftrightarrow$ ($\sim$ S))), since the formula evalutes to **true** under the given interpretation.

Questions

1. Provide an interpretation that will satisfy (P $\wedge$ ($\sim$ Q)).
2. Provide an interpretation that will falsify (P $\wedge$ ($\sim$ Q)).

Validity and inconsistency/unsatisfiability

A formula is said to be **valid** if and only if it is true under all of its interpretations. A formula is said to be **invalid** if it is not valid.

A formula is said to be **inconsistent** (or **unsatisfiable)** if and only if it is false under all of its interpretations. A formula is said to be **consistent** (or **satisfiable)** if it is not inconsistent.

Example

What can you say about the formula $(((P \rightarrow Q) \wedge P) \rightarrow Q)$?

1. Is it valid or invalid?
2. Is it inconsistent or consistent?

Check this out ...

1. PQ = TT: $(((P \rightarrow Q) \wedge P) \rightarrow Q) \Rightarrow (((T \rightarrow T) \wedge T) \rightarrow T) \Rightarrow ((T \wedge T) \rightarrow T) \Rightarrow (T \rightarrow T) \Rightarrow T$
2. PQ = TF: $(((P \rightarrow Q) \wedge P) \rightarrow Q) \Rightarrow (((T \rightarrow F) \wedge T) \rightarrow F) \Rightarrow ((F \wedge T) \rightarrow F) \Rightarrow (F \rightarrow F) \Rightarrow T$
3. PQ = FT: $(((P \rightarrow Q) \wedge P) \rightarrow Q) \Rightarrow (((F \rightarrow T) \wedge F) \rightarrow T) \Rightarrow ((T \wedge F) \rightarrow T) \Rightarrow (F \rightarrow T) \Rightarrow T$
4. PQ = FF: $(((P \rightarrow Q) \wedge P) \rightarrow Q) \Rightarrow (((F \rightarrow F) \wedge F) \rightarrow F) \Rightarrow ((F \wedge F) \rightarrow F) \Rightarrow (F \rightarrow F) \Rightarrow T$

So the formula is **valid** (since it is true under each iterpretation). The formula is also **consistent** (since it is not inconsistent (false under each interpretation)). Note that we could alternatively have expressed our argument for validity in the form of a truth table.

P	Q	$(P \rightarrow Q)$	$((P \rightarrow Q) \wedge P)$	$(((P \rightarrow Q) \wedge P) \rightarrow Q)$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

Something to do

The following few tasks to pertain to validity and inconsisency, but they also serve, in a rather passive manner, to introduce you to yet another notation for representing WFFS.

1. The following line expresses a WFF in a "plain text" formalism – a syntax that is useful for communicating, for example, in the body of an email text:

   `( ( ( P [I] Q ) [A] P ) [I] ( [N] Q ) )`

   Using the suggestive interpretation that "[I]" stands for "implies", "[E]" stands for "equivalence", "[A]" stands for "and", "[O]" stands for "or", and "[N]" stands for "not", rewrite the given expression of the WFF using standard logic notation.

2. Argue that the following WFF is both **invalid** and **consistent**:
   (((P [I] Q) [A] P) [I] ([N] Q))

3. Write down a WFF containing two different atoms that is **inconsistent**, and show that it is, indeed, **inconsistent**. Furthermore, do so twice, once using standard logic notation, and once using plain text notation.

Logical Equivalence

Big idea: The basic idea in computing with logic involves transforming formulas to logically equivalent "normal" forms, and then working with the normal forms.

Definition

Two formulas G and H are said to be **logically equivalent** if and only if the truth values of F and G are the same under every interpretation of F and G.

Example/Demonstration

Show that the "switcheroo" rule (to steal Hofstadter's colorful name) makes sense by showing that the following to WFFs are logically equivalent:

Switcheroo: (F $\rightarrow$ G) and (($\sim$ F) $\vee$ G) are logicaly equivalent

Please note that F is being overloaded! Sometimes it refers to a WFF, sometimes to the boolean constant FALSE. You can handle it!

1. First, consider (F $\rightarrow$ G):
   (a) FG = TT: (F $\rightarrow$ G) $\Rightarrow$ (T $\rightarrow$ T) $\Rightarrow$ T
   (b) FG = TF: (F $\rightarrow$ G) $\Rightarrow$ (T $\rightarrow$ F) $\Rightarrow$ F
   (c) FG = FT: (F $\rightarrow$ G) $\Rightarrow$ (F $\rightarrow$ T) $\Rightarrow$ T
   (d) FG = FF: (F $\rightarrow$ G) $\Rightarrow$ (F $\rightarrow$ F) $\Rightarrow$ T
2. Next, consider (($\sim$ F) $\vee$ G):
   (a) FG = TT: (($\sim$ F) $\vee$ G) $\Rightarrow$ (($\sim$ T) $\vee$ T) $\Rightarrow$ (F $\vee$ T) $\Rightarrow$ T
   (b) FG = TF: (($\sim$ F) $\vee$ G) $\Rightarrow$ (($\sim$ T) $\vee$ F) $\Rightarrow$ (F $\vee$ F) $\Rightarrow$ F
   (c) FG = FT: (($\sim$ F) $\vee$ G) $\Rightarrow$ (($\sim$ F) $\vee$ T) $\Rightarrow$ (T $\vee$ T) $\Rightarrow$ T
   (d) FG = FF: (($\sim$ F) $\vee$ G) $\Rightarrow$ (($\sim$ F) $\vee$ F) $\Rightarrow$ (T $\vee$ F) $\Rightarrow$ T

Since the truth values of F and G are the same under each of the four interpretations, the two formulas are logically equivalent.

Something to think about

1. DeMorgan's Law: ([N] (F [O] G) and (([N] F) [A] ([N] G)) are logically equivalent.
2. DeMorgan's Law: ([N] (F [A] G) and (([N] F) [O] ([N] G)) are logically equivalent.

Something to think do

1. Write down DeMorgan's two laws using our standard notation for WFFs.
2. Pick one of DeMorgan's laws and show the logical equivalence of the two formulas featured in the law that you picked.

3. Write down a sentence or two about Augustus De Morgan, the 19th century mathematician.

Three Notational Variants of Switcheroo

Hofstadter's Notation for WFFs

$< F \supset G >$ is logically equivalent to $< \sim F \vee G >$

Standard Modern Notation for WFFs

$(F \rightarrow G)$ is logically equivalent to $((\sim F) \vee G)$

Plain Text Notation for WFFs

(F [I] G) is logically equivalent to (([N] F) [O] G)

Some Logical Equivalences

Note on Computing with Logic

Most computations in relatively classical logic these days operate in the context of **normal forms**, which are WFFs that are specialized by means of various sets of conventions. The two most common normal forms are **disjunctive normal form** and **conjunctive normal form**. Accordingly, prior to actually computing with WFFs (performing logical inference), we must, or our machines must, converts WFFs to a logically equivalent normal form.

The Role of Logical Equivalences

Logical equivalences play a central role in converting WFFs in arbitrary form to WFFs on normal form. This is simply because the conversion must preserve logical equivalence in order for the ultimate computations to make any sense with respect to the problem that is being solved by means of logical representation and reasoning.

Some useful Logical Equivalences

Here is a useful list of logically equivlatent formulas, or laws, or "logical equivalences":

- $(F \leftrightarrow G) = ((F \rightarrow G) \wedge (G \rightarrow F))$ - separation of bidirectionality
- $(F \rightarrow G) = ((\sim F) \vee G)$ - switcheroo
- $(F \vee G) = (G \vee F)$ - commutativity
- $(F \wedge G) = (G \wedge F)$ - commutativity
- $(F \vee (G \vee H)) = ((F \vee G) \vee H)$ - associativity
- $(F \wedge (G \wedge H)) = ((F \wedge G) \wedge H)$ - associativity
- $(F \vee (G \wedge H)) = ((F \vee G) \wedge (F \vee H))$ - distributivity
- $(F \wedge (G \vee H)) = ((F \wedge G) \vee (F \wedge H))$ - distributivity
- $(\sim (\sim F)) = F$ - double negativity
- $(\sim (F \vee G)) = ((\sim F) \wedge (\sim G))$ - DeMorgan
- $(\sim (F \wedge G)) = ((\sim F) \vee (\sim G))$ - DeMorgan

You will want to keep this list in mind, or close at hand, whenever you are asked to convert a WFF to normal form.

Preliminary Discussion for Normal Forms

Notational Conventions for Normal Forms

Normal forms require that we relax the constraint of fully parenthesizing WFFs in a limited and well defined way:

1. We will group like adjacent operators using only one set of parentheses.
2. We will drop parentheses surrounding the ∼ operator.
3. We drop the outermost parentheses of an expression in the case of a conjunction of disjunctions or a disjunction of conjunctions.

Examples

By way of example, consider the following three items, each of which features the plain text representation of WFFs.

```
(1) We will group like adjacent operators using only one set of parentheses.
    Example: ( P [A] ( Q [A] R ) ) will be written ( P [A] Q [A] R )
    Example: ( ( P [O] Q ) [O] R ) ) will be written ( P [O] Q [O] R )

(2) We will drop parentheses surrounding the negagion operator.
    Example: ( [N] P ) will be written [N] P
    Example: ( [N] ( [N] Q ) ) will be written as [N] [N] Q

(3) We drop the outermost parentheses of an expression.
    Example: ( ( P [O] Q ) [A] ( [N] R [O] S [O] T ) )
             will be written as
             ( P [O] Q ) [A] ( [N] R [O] S [O] T )
```

Exercise

Rerwite each of the examples presented using standard logic notation.

Definition of Literal

A literal is an atom or the negation of an atom. Examples:

1. P
2. ∼P