1 Plausible reasoning

1.1 Deductive and plausible reasoning

- **Deductive reasoning** can be broken down into a **strong syllogism** (logical argument)

$$\frac{\begin{array}{c}\text{if } A \text{ is true, then } B \text{ is true} \\ A \text{ is true}\end{array}}{\text{therefore, } B \text{ is true}}$$

  and its inverse:

$$\frac{\begin{array}{c}\text{if } A \text{ is true, then } B \text{ is true} \\ B \text{ is false}\end{array}}{\text{therefore, } A \text{ is false.}}$$

- In almost all situations, we do not have the right information to allow this kind of reasoning, so we fall back on **weaker syllogisms**:

$$\frac{\begin{array}{c}\text{if } A \text{ is true, then } B \text{ is true} \\ B \text{ is true}\end{array}}{\text{therefore, } A \text{ becomes more plausible}}$$

- 'If A then B' expresses B only as a **logical** consequence of A, not necessarily a causal, physical consequence:
  - rain at 10 AM does not cause clouds at 9:45 AM
  - the logical connection is not in the uncertain causal direction (clouds $\Rightarrow$ rain), but the certain, though noncausal direction (rain $\Rightarrow$ clouds)
- another weak syllogism:

$$\frac{\begin{array}{c}\text{if } A \text{ is true, then } B \text{ is true} \\ A \text{ is false}\end{array}}{\text{therefore, } B \text{ becomes less plausible}}$$

- a still weaker syllogism:

$$\frac{\begin{array}{c}\text{if } A \text{ is true, then } B \text{ becomes more plausible} \\ B \text{ is true}\end{array}}{\text{therefore, } A \text{ becomes more plausible}}$$

- in doing plausible reasoning, the brain not only decides whether something becomes more or less plausible, but also the *degree* of plausibility
  - we depend very much on **prior information** (aka 'common sense') when evaluating the degree of plausibility of a new problem

1.2 Analogies with physical theories

1.3 The thinking computer

How could we build a machine which would carry out useful plausible reasoning, following clearly defined principles expressing an idealized common sense?

1.4 Introducting the robot

- Our robot will reason about **propositions** $\{A, B, C, \text{etc.}\}$
  - any proposition must have an unambiguous meaning and must be of the simple, definite logical type (i.e. Boolean, true/false)

1.5 Boolean algebra

- **logical product** or **conjunction**, denoted AB
  - both A and B are true (logical AND)
- **logical sum** or **disjunction**, denoted $A + B$
  - at least one of A, B is true (logical OR)
  - $A + B$ is equivalent to $B + A$
- if one of two propositions A, B is true iff the other is true, then A and B have the same **truth value**, denoted $A = B$
  - a primitive axiom of plausible reasoning: two propositions with the same truth value are equally plausible
- parentheses are used as in ordinary algebra, to indicate the order in which propositions are to be combined
  - in absence of parentheses, normal order of operations applies
- the **denial** of a proposition:
  - $\overline{A} \equiv A$ is false
  - $A = \overline{A}$ is false
  - Care is needed in unambiguous use of the bar:

$$\begin{aligned} &\overline{AB} = AB \text{ is false;} \\ &\bar{A}\bar{B} = \text{both } A \text{ and } B \text{ are false} \end{aligned}$$

- Boolean algebra identities:

$$\begin{aligned}
\text{idempotence:} &\begin{cases} AA = A \\ A + A = A \end{cases} \\
\text{commutativity:} &\begin{cases} AB = BA \\ A + B = B + A \end{cases} \\
\text{associativity:} &\begin{cases} A(BC) = (AB)C = ABC \\ A + (B + C) = (A + B) + C = A + B + C \end{cases} \\
\text{distributivity:} &\begin{cases} A(B + C) = AB + AC \\ A + (BC) = (A + B)(A + C) \end{cases} \\
\text{duality:} &\begin{cases} \text{If } C = AB, \text{ then } \overline{C} = \overline{A} + \overline{B} \\ \text{If } D = A + B, \text{ then } \overline{D} = \bar{A}\bar{B} \end{cases}
\end{aligned}$$

– By applying the basic identities, further relations can be proven:

$$\begin{aligned}
&\text{Let } \overline{B} = AD \\
&A\overline{B} = AAD \\
&A\overline{B} = AD && \text{idempotence} \\
&A\overline{B} = \overline{B} && \text{def. } \overline{B} \\
&A\overline{B} = \overline{B} + \overline{B} && \text{idempotence} \\
&A\overline{B} = \overline{B} + A\overline{B} && \text{def. } \overline{B} \\
&A\overline{B} = \bar{B}\bar{B} + A\overline{B} && \text{idempotence} \\
&A\overline{B} = \left(\overline{B} + A\right)\overline{B} && \text{distributivity} \\
&A = \overline{B} + A \\
&B\overline{A} = \overline{A} && \text{duality}
\end{aligned}$$

Therefore, if $\overline{B} = AD$, then $A\overline{B} = \overline{B}$, and $B\overline{A} = \overline{A}$.

- The proposition $A \Rightarrow B$ (A implies B) does not assert that either A or B is true
  - it means only that $A\overline{B}$ is false, or equivalently, $\left(\overline{A} + B\right)$ is true
  - this can also be written as the logical equation $A = AB$
    - i.e., given $A \Rightarrow B$, if A is true, then B must be true; or, if B is false then A must be false
    - This is what was stated in the strong syllogisms
    - However, if A is false, $A \Rightarrow B$ says nothing about B, and if B is true, $A \Rightarrow B$ says nothing about A
      - these are the cases in which the weak syllogisms *do* say something
      - plausible reasoning based on weak syllogisms is not a 'weakened' form of logic; it is an *extension* of logic with content not present in conventional deductive logic
- In formal logic, 'A implies B' means only that A and AB have the same truth value
  - in general, whether B is logically deducible from A depends not only on A and B, but the totality of propositions $(A, A', A'', ...)$ accepted as true and therefore available for use in the deduction

1.6 Adequate sets of operations

- Any number of propositions can be generated using the logical product (conjunction), logical sum (disjunction), implication, and negation operations
  - How large is the set of new propositions? Is it finite/infinite?
  - Are these four operations sufficient to generate every proposition?
  - Are any operations dispensable for generating every proposition?
- Logical NAND is defined as the negation of AND:

$$A \uparrow B \equiv \overline{AB} = \overline{A} + \overline{B}$$

Every logic function can be constructed with NAND alone:

$$\begin{aligned}
&\overline{A} \\
=&\overline{A} + \overline{A} && \text{idempotence} \\
=&A \uparrow A && \text{def. NAND}
\end{aligned}$$

$$
\begin{aligned}
&AB \\
=&AB + AB && \text{idempotence} \\
=&\overline{AB} \uparrow \overline{AB} && \text{def. NAND} \\
=&(A \uparrow B) \uparrow (A \uparrow B) && \text{def. NAND}
\end{aligned}
$$

$$
\begin{aligned}
&A + B \\
=&AA + BB && \text{idempotence} \\
=&\overline{AA} \uparrow \overline{BB} && \text{def. NAND} \\
=&(A \uparrow A) \uparrow (B \uparrow B) && \text{def. NAND}
\end{aligned}
$$

- Logical NOR is defined as the negation of OR:

$$A \downarrow B \equiv \overline{A + B} = \bar{A}\bar{B}$$

Every logic function can also be constructed with NOR alone:

$$
\begin{aligned}
&\overline{A} \\
=&\bar{A}\bar{A} && \text{idempotence} \\
=&A \downarrow A && \text{def. NOR}
\end{aligned}
$$

$$
\begin{aligned}
&AB \\
=&(A + A)(B + B) && \text{idempotence} \\
=&\left(\overline{A + A}\right) \downarrow \left(\overline{B + B}\right) && \text{def. NOR} \\
=&(A \downarrow A) \downarrow (B \downarrow B) && \text{def. NOR}
\end{aligned}
$$

$$
\begin{aligned}
&A + B \\
=&(A + B)(A + B) && \text{idempotence} \\
=&\left(\overline{A + B}\right) \downarrow \left(\overline{A + B}\right) && \text{def. NOR} \\
=&(A \downarrow B) \downarrow (A \downarrow B) && \text{def. NOR}
\end{aligned}
$$

1.7 The basic desiderata

I Degrees of plausibility are represented by real numbers

- We adopt a nonessential, but natural, convention: greater plausibility shall correspond to a greater number
- It is also useful to assume continuity, i.e. an infinitesimally greater plausibility shall correspond to an infinitesimally greater number
- The plausibility assigned to a proposition A will, in general, depend on the truth value of proposition B, denoted $A|B$
  - this is 'the conditional plausibility that A is true, given that B is true, or A given B
  - We avoid impossible problems, i.e. when $A|BC$ is specified, it is understood that B and C are compatible propositions

II Qualitative correspondence with common sense

- if old information C is updated to C' such that the plausibility for A is increased but the plausibility for B is unchanged, i.e, $(A|C') > (A|C)$ and $(B|AC') = (B|AC)$, then the plausibility that both A and B are true must increase, and the plausibility that A is false must decrease:

$$(AB|C') \geq (AB|C) \qquad\qquad \left(\overline{A}|C'\right) < \left(\overline{A}|C\right)$$

III Consistent reasoning

- **IIIa** If a conclusion can be reasoned out in more than one way, then every possible way must lead to the same result.
- **IIIb** All of the evidence relevant to a question is always taken into account. Conclusions are not based on an arbitrary subset of the information.
- **IIIc** Equivalent states of knowledge are represented by equivalent plausibility assignments.

There is only one set of mathematical operations for manipulating plausibilities that satisfies all of the desiderata listed above.

2 The quantitative rules

2.1 The product rule

- We seek a consistent rule relating the plausibility of the logical product AB to the plausibilities of A and B separately, in particular $AB|C$.
  - The process of deciding that AB is true can be broken down into elementary decisions about A and B separately:

    1) decide that B is true; $(B|C)$

    2) having accepted B as true, decide that A is true. $(A|BC)$

    or equivalently,

    1') decide that A is true; $(A|C)$

    2') having accepted A as true, decide that B is true. $(B|AC)$
  - For AB to be a true proposition, it is necessary that B is true.
    - Thus, the plausibility $B|C$ is involved
  - If B is true, it is further necessary that A is also true.
    - Thus, the plausibility $A|BC$ is also involved
  - If B is false, then AB is false independently of any knowledge about A, i.e. $A|\overline{B}C$
    - If B is known, then the plausibility of A is relevant only if B is true.
    - Thus, given $B|C$ and $A|BC$, $A|C$ provides no additional information about AB.
    - Similarly, $A|B$ and $B|A$ are unnecessary, since plausibility in the absence of information C is not relevant to the case where C is known to be true.
  - Since the logical product is commutative, i.e. $AB = BA$, A and B can be interchanged in the above statements, i.e. knowledge of $A|C$ and $B|AC$ should return the same plausibility as knowledge of $B|C$ and $A|BC$, due to desideratum IIIa for consistency.
- In a concrete statement, $(AB|C)$ will be some function of $B|C$ and $A|BC$:

$$(AB|C) = F\left[(A|C), (B|C)\right]$$

- Now we apply the qualitative requirement (desiderata II): Given a change in prior information $C \to C'$ such that B becomes more plausible but A does not change,

$$B|C' > B|C \qquad\qquad A|BC' = A|BC,$$

  then AB must become more plausible, not less:

$$AB|C' \geq AB|C$$

  Similarly, given prior information C'' such that $B|C'' = B|C$ and $A|BC'' > A|BC$, we require that $AB|C'' \geq AB|C$.

- Furthermore, due to the continuity requirement (desiderata I), $F(x, y)$ must be continuous.
- In summary, $F(x, y)$ must be a continuous monotonic increasing function of both x and y. If $F(x, y)$ is assumed to be differentiable (this is not necessary), then

$$F_1(x, y) \equiv \frac{dF}{dx} \geq 0 \qquad\qquad F_2(x, y) \equiv \frac{dF}{dy} \geq 0$$