A Many-Valued Probabilistic Conditional Logic

François LEPAGE
University of Montreal

Prolegomenas

In his celebrated paper, 'Probability of Conditionals and Conditional Probabilities', David Lewis [Lew76] showed that, contrary to a claim of Robert Stalnaker's [Sta68], [Sta70], it is not possible to introduce into the classical propositional calculus a counterfactual conditional 'if A were the case, B would be the case' (hereafter '$A > B$') in order to obtain a language and a semantics for that language having the following properties:

(1) the classical part behaves classically, i.e., any proposition which contains no counterfactual has its classical truth conditions;

(2) starting from any possible world, possible worlds are linearly ordered, so we can speak of the distance of a world from the starting world, and the truth conditions of the counterfactual are the following: $A > B$ is true in a world w iff either B is true at the closest world to w where A is true (hereafter 'A-world') or there is no A-world;

(3) any probability function defined on the classical part can be extended to the set of all propositions by stipulating that the probability of the counterfactual is the probability of the consequent, given the antecedent i.e., $P(A > B) = P(B/A)$ if $P(A) \neq 0$.

Clause (2) wil be called Stalnaker'clause. Lewis has shown that if clauses (1), (2) and (3) are not, strictly speaking, inconsistent, they can be satisfied only by trivial probability functions, i.e., by probability functions which give to any proposition either the value 0 or the value 1.

In the same paper, Lewis suggested another way to interpret the probability of a conditional, preserving clauses (1) and (2) and giving up clause (3).

Let P be a probability function defined on the classical part. Lewis suggested an extension of P in the following way:

(4) $P(A > B) = P_A(B)$ where P_A is the probability function obtained from P by shifting the probability of any $\neg A$-world on the nearest A-world according to Stalnaker's clause.

Lewis has shown that this definition is compatible with the general constraint that the probability of any proposition A is the sum of the probability of the A-worlds. This is called *Imaging*.

But Stalnaker's clause is far from being intuitive. If in some situations, there is clearly a nearest A-world, this does not obtain generally. Let's suppose a dice is thrown and gives a 6. It is quite natural to think that there is *one* closest world where the throw gave, say, 5.

What, now, about the closest world where the throw didn't give a 6? According to Stalnaker's clause, the worlds where the throw gave 1, 2, 3, 4 or 5 are linearly ordered: one of them is the nearest, another the second nearest, etc. For obvious reasons of symmetry, this consequence of Stalnaker's clause is unintuitive and casts discredit on the clause itself.

Curiously, Lewis [Lew73] had already provided a semantics for counterfactuals that can bypass this difficulty, namely his System Of Spheres semantics (hereafter SOS).

In this semantics, possible worlds are not linearly ordered, but weakly ordered, i.e., many worlds may be at the same distance from a given world. The best image is of embedded spheres of possible worlds centered on the world of evaluation. All the worlds of a given layer are equidistant from the world of evaluation. The truth conditions of the counterfactual are then the following:

(5) $A > B$ is true in a world w iff B is true at *all* the nearest A-worlds to w or there is no A-world.

Unfortunately, Imaging is not compatible with Lewis' SOS semantics.

It can be shown [Lep97] that if we generalize imaging as defined by clause (4) by sharing out the probability of any world on the nearest worlds where the antecedent is true, then the probability of a conditional is the probability of the consequent after this sharing out if and only if each layer of the system of spheres around any world contains exactly one world. In short: imaging works in a SOS if and only if this SOS is a Stalnaker system.

But there is another way to introduce imaging in SOS. It is by changing the truth conditions of the conditional. Let $A > B$ have as its truth value the ratio of A B-worlds on the A-worlds of the smallest sphere having at least one A-world. In the limit cases where all A-worlds are A B-worlds or no A-worlds are A B-worlds, the truth value of the conditional is the same as in Lewis' SOS, i.e., 1 or 0. In the intermediary cases where some but not all of the A-worlds are A B-worlds, the conditional will have a fractionary value. A brief presentation of that semantics will be the task of the first part of my paper.

The introduction of fractionary values raises the well-known problems associated with many-valued logic, that is, there is no extensional many-valued extension of the classical logic in the following sense:

(5) All instances of tautologies are valid;

(6) If all the sub-expressions of a proposition have classical truth values, then this proposition also has a classical truth value.

In the second part of the paper, I present a *non extensional* logic satisfying (5) and (6) which uses a counterfactual for the definition of the truth conditions for conjunction.

Imaging in a SOS

Let's first present the syntax of the system.

The set of atomic propositions is $A = \{p_i\}_{i \quad n}$ where n is a finite number, and the set of propositions is the smallest set L such that

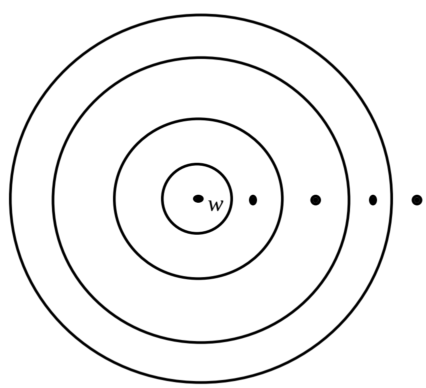
(i) for any $i \quad n$, $p_i \quad L$

(ii) if $A \quad L$, then $\neg A \quad L$

(iii) if A, $B \quad L$, then $(A \quad B)$, $(A > B) \quad L$.

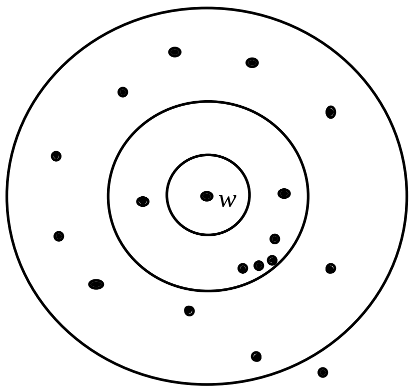
We can now provide the system with an interpretation.

Let W be the set of possible worlds defined as $W = \{0,1\}^{A}$ (where 0 and 1 respectively express the falsity and the truth) and let $f : L \times W \rightarrow \quad (W)$. f is a *selection function* such that, for any world w and any proposition A, $f(A,w)$ is a set of worlds containing at least one A-world. I will consider two classes of selection functions. Firstly, I will consider *Stalnaker functions*, where f selects, for any $<A,w>$, one and only one world. All the worlds selected by varying A are linearly ordered (the closest world being the actual world).



For a Stalnaker function, $f(<A,w>)$ is the closest A-world to w (or equivalently, the smallest set containing one A-world).

I will also consider *Lewis selection functions*. Selected worlds are in embedded spheres and the smallest sphere contains only the actual world. Thus a given Lewis selection function selects for each A the smallest sphere containing at least one A-world.



Stalnaker functions are just a special case of Lewis functions, the case where each layer contains exactly one world.

Let f be a Stalnaker function.

An interpretation based on $w \quad W$ is a function $h : L \rightarrow \{0,1\}$ such that

(i) $h(p_j) = w(p_j)$

(ii) $h(\neg A) = 1 - h(A)$

(iii) $h(A \quad B) = h(A) \,.\, h(B)$

(iv) $h(A > B) = h'(B)$ where h' is the interpretation based on $f(A, w)$.

We will call h a “Stalnaker model”. In short, $(A > B)$ is true at w iff B is true at the closest A-world to w.

No confusion being possible, I identify the characteristic function w with its extension h, and I will write $w(A)$ even in thee case where A is not an atom.. Let P be a probability distribution on W. As usual, we define $P(A) = \sum_{w \in W} P(w)w(A)$ and so we trivialy have

(i) $P(A) = 1$ if A is a tautology

(ii) $P(\neg A) = 1 - P(A)$

(iii) $P(A \lor B) = P(A) + P(B)$ if $w(A \land B) = 0$.

Let δ be a function such that $\delta(w', w, A) = 1$ iff $w' = f(<A, w>)$

Following Lewis [Lew76], I define

$$P_A(w') = \sum_{w \in W} P(w)\delta(w', w, A)$$

P_A is the probability function obtained by *imaging* on A, i.e., the probability function obtained from P after projecting the probability of any world on the nearest A-world [Lew76], [Nut80], [Gär82] and [Gär88]. Lewis has shown that the functions obtained by *imaging* have the following property:

Proposition For any probability function P, any Stalnaker function f and any world w, if A is not a contradiction, then

$$(*) \qquad P(A > B) = \sum_{w \in W} P(w)w(A > B) = \sum_{w \in W} P_A(w)w(B) = P_A(B)$$

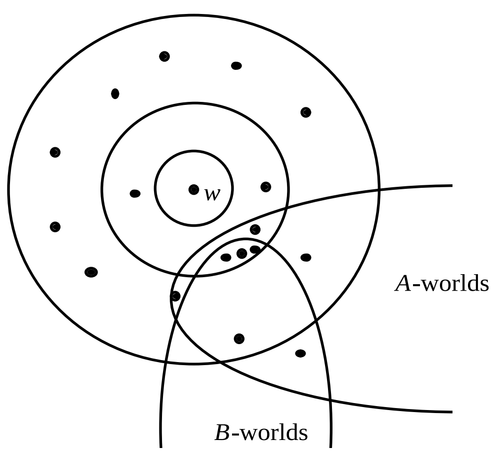
This technique is radically different from conditionalization as shown by Gärdenfors [Gär82].

The restriction to Stalnaker functions, i.e., the hypothesis that, from the point of view of any world, W is linearly ordered, is a very constraining one. It would be very interesting to obtain a similar result using Lewis functions. Unfortunately, as suggested by Nute [Nut80], it is not possible.

Let us define a Lewis model [Lew73]. It is a model similar to the one above except that f is a Lewis function and (iv) becomes

(iv') $w(A > B) = 1$ iff $w'(B) = 1$ for any w' $\quad f(<A, w>)$ such that $w'(A) = 1$.

For example, in the following situation



we have $w(A > B) = 0$ et $w(B > A) = 1$.

We could try to adapt the *imaging* technique to Lewis models by defining

(**) $P_A(w') = \sum_{w \quad W} P(w)\delta(w', w, A)\, c_{w,w',A}$

where the $c_{w,w',A}$ are weighting coefficients , i.e., for any w,

$$\sum_{w' \in W} \delta(w', w, A)\, c_{w,w',A} = 1.$$

Hence we obtain the following result [Lep97]:

Proposition For any Lewis model and any P, the function P_A obtained by *imaging* according to (**) satisfies the equation

$$P(A > B) = \sum_{w \in W} P(w)w(A > B) = \sum_{w \in W} P_A(w)w(B) = P_A(B)$$

iff f is a Stalnaker function.

The reason for this is simple: According to the truth conditions of Lewis conditional, when the smallest sphere around w containing at least one A-world contains $(A \wedge B)$-worlds and $(A \wedge \neg B)$-worlds the conditional is false, so $P(w)w(A > B) = 0$ and the probability projected on A-worlds is lost. Thus, in that case $P(A > B) < P_A(B)$. Therefore, imaging is not compatible with Lewis' original semantics.

Fortunately, if we modify this semantics, imaging is possible again. The modification consist in allowing conditionals to take fractional truth values. The truth value of a conditional $(A > B)$ is the ratio of the number of $(A \wedge B)$-worlds on the number of A-worlds in the smallest sphere containing at least one A-world. When all the A-worlds are B-worlds or no A-worlds are B-worlds, we find the classical truth values 1 and 0 again. But in halfway cases, the truth value of the counterfactual is a fraction.

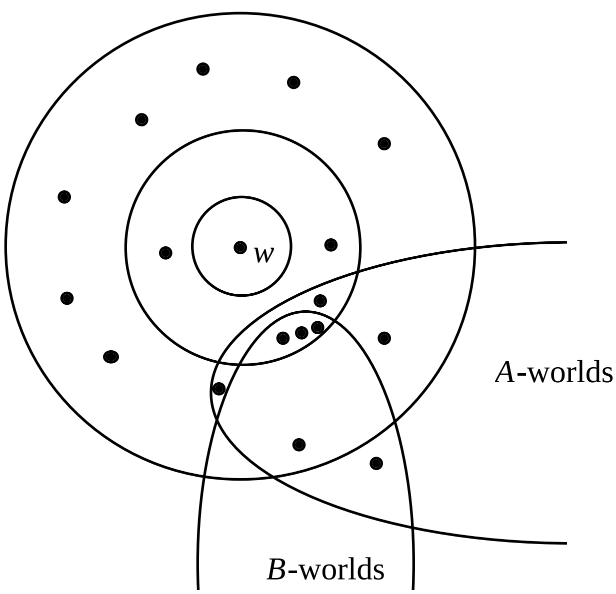
Formally, we define $w(A > B) = \dfrac{\sum_{w' \; V(A,w)} w'(B)}{n_{A,w}}$ where

$V(A,w)$ $f(<A,w>)$ is the set of A-worlds of $f(<A,w>)$.

$\sum_{w' \; V(A,w)} w'(B)$ is the sum of the truth values of B in the A-worlds belonging to $f(<A,w>)$ (when B has not itself a fractional value, $\sum_{w' \; V(A,w)} w'(B)$ is just the number of A-worlds among the A-worlds of $f(<A,w>)$);

$n_{A,w}$ is the number of A-worlds in $f(<A,w>)$.

Take, for instance, the following diagram:



Here, the smallest sphere containing at least one A-world contains four A-worlds, three of them are B-worlds and thus $w(A > B) = 3/4$.

With this new definition of $w(A > B)$, it is easily proved that (*) holds again. But now, what happens to complex propositions?

The natural truth conditions of negation are surely given by

$w(\neg A) = 1 - w(A)$.

There is no natural definition of truth conditions for the conjunction. This is a very well-known problem for any many-valued logic [Urq86]. One can easily show that *no* extensional definition of conjunction results in a many-valued logic which is an extension of classical logic, i.e., one where

(i) any instance of a tautology is valid and two tautologically equivalent expressions are equivalent;
(ii) any formula in which only classical connectors have occurences takes its classical value for any classical valuation.

This brings us to the second part of our paper. I will now introduce the notion of *normal form* (hereafter NF) for propositions of L and provide an interpretation for these normal forms using the counterfactual truth conditions to define the truth conditions of the conjunction.

A New Semantics for Conjunction

An NF is just a slight modification of the notion of *full normal disjunctive form* in order to take into account the occurrences of counterfactuals.

By a full normal disjunctive form of a *classical* formula A, I mean the following :
If A is a contradiction, the full normal disjunctive form of A is $\mathbf{0}$ (a canonical name for falsity); the full normal disjunctive form of $\neg\mathbf{0}$ is $\mathbf{1}$ (a canonical name for truth); otherwise

the full normal disjunctive form of A is the shortest formula which is tautologically equivalent to A of the form $\neg(\neg A_1 \quad \ldots \quad \neg A_m)$ where each A_j is a conjunction of litterals in some canonical order.

In the general case, we define a function NF : $L \rightarrow L$ such that

(i) NF(l_i) = l_i (where l_i is a litteral, i.e., p_j or $\neg p_j$ for some j)
(ii) NF($A > B$) = NF(A) > NF(B)
(iii) If A is not as in (i) and (ii), then NF(A) is the full normal disjunctive form of A where any counterfactual $C > B$ is treated like an atom and is replaced by NF($C > B$).
So, any NF is of the following form

$\neg {}_{j=0}^{m-1} \neg A_j$ where A_j is itself a conjunction of litterals or of counterfactuals in normal form.
We can now define an interpretation for formulas in normal form. We will need the following tools.

An interpretation based on $w \quad W$ is defined as usual, with the additional hypothesis that the SOS is such that if A and B have no atom in common, then the atoms of B have the same value in w and in the closest A-worlds to w. This constraint can be interpreted as meaning that if A and B are independent, then $w(A > B)$ is just $w(B)$.

So,

(i) $w(\neg A) = 1 - w(A)$

(ii) $w(A > B) = \dfrac{\sum_{w' \quad V(A,w)} w'(B)}{n_{A,w}}$ (as above)

(iii) Let ${}_{j=0}^{m-1} A_j$ be a conjunction of m propositions. We define

$$w\left({}_{j=0}^{m-1} A_j\right) = \sqrt[m]{\prod_{j=0}^{m-1} w(A_j) \bullet \prod_{j,k=0}^{m-1} w(A_j > A_k)}$$

The idea is to interpret a conjunction of m formulas not only as the logical product of the values of the conjuncts but to take also into account a kind of "proximity" between the conjuncts, which is express by $w(A_j > A_k)$.[1]
Let us consider the following examples:

[1] The idea to use counterfactuals as a mesure of proximity was suggested to me by Professor Jian-Yun Nie.

(1) $w(p_j \quad p_i) =$

$\sqrt{w(p_i) \cdot w(p_j) \cdot w(p_i > p_j) \cdot w(p_j > p_i)} =$

$\sqrt{w(p_i) \cdot w(p_j) \cdot w(p_j) \cdot w(p_i)} =$

$w(p_j) \cdot w(p_i)$

(2) $w(p_j \quad p_j) = w(p_j)$

(3) $w(p_j \quad \neg p_j) =$

$\sqrt{w(p_j) \cdot w(\neg p_j) \cdot w(p_j > \neg p_j) \cdot w(\neg p_j > p_j)} =$

$\sqrt{w(p_j) \cdot w(\neg p_j) \cdot 0 \cdot 0} = 0$

(4) $w(p_j \quad \neg p_j) = 1$

(5) $w(p_j \quad p_i \quad p_k) =$

$\sqrt[3]{w(p_j) \cdot w(p_i) \cdot w(p_k) \cdot w(p_j > p_i) \cdot w(p_i > p_j) \cdot w(p_j > p_k) \cdot w(p_k > p_j) \cdot w(p_k > p_i) \cdot w(p_i > p_k)} =$

$w(p_j) \cdot w(p_i) \cdot w(p_k)$

(6) In the general case, if A_k and A_i have no atom in common,

$w(\overset{m-1}{\underset{j=0}{}} A_j) = \prod_{j=0}^{m-1} w(A_j)$

Unfortunately, the definition of the truth conditions for the conjunction given above is not recursive, because it depends on the number m of conjuncts. This difficulty can easily be bypassed. Let us define the following two functions:

Conj(A) is a function which counts the number of conjuncts in A.

Comp(A) is the set of conjuncts of A.

Conj(l) = 1 (when l is a literal)

Conj($A \quad B$) = *Conj*(A) + *Conj*(B)

Conj ($\neg A$) = 1

Conj($A >$B) = 1

$Comp(l) = \{l\}$

$Comp(\neg A) = \{\neg A\}$

$Comp(A \quad B) = Comp(A) \quad Comp(B)$

$Comp(A > B) = \{(A > B)\}$

Using these two functions, we can replace the truth condition for the conjunction given in (iii) by

$$\text{(iv) } w(A) = \sqrt[Conj(A)]{\prod_{A_j \quad Comp(A)} w(A_j) \bullet \prod_{A_{j,k} \quad Comp(A)} w(A_j > A_k)}$$

(i), (ii) and (iv) are recursive clause that provide a value for any NF. We obtain the following results.

Proposition

Any instance of a tautology is valid.

This property is trivial since we work with NF and the NF of any instance of a tautology is 1.

The next result is (at first sight) less trivial:

Proposition

Any classical proposition (in which there is no occurrence of conditionals) has its classical truth conditions.

Proof

We just have to check that clauses (i) and (iv) behave classically for classical arguments, and this is straightforward.

This last result means that the complicated truth conditions given by (i), (ii) and (iv) are just the classical ones when all the arguments are classical ones. So, this semantics is an extension of the classical one.

Let us now turn to the question of providing a system for that logic.
It is an interesting fact that for any axiom of the complete system of Stalnaker's logic, the corresponding rule is valid (by the corresponding rule I mean the replacement of $\vdash A \quad B$ by $A \vdash B$)

Proposition

All the rules of the (complete) Stalnaker system are valid, i.e.,

$\vdash \top$ if $\top$ is an instance of a tautology

$((A > B) \quad (A > C)) \vdash (A > (B \quad C))$

$\vdash A > \top$

$\vdash A > A$

$(A \quad B) \vdash (A > B)$

$(A > B) \vdash (A \quad B)$

$((A > C) \quad (B > C)) \vdash ((A \quad B) > C)$

$((A > B) \quad (A > C)) \vdash ((A \quad C) > B)$

$\vdash ((A > \mathrm{B}) \quad (A > \neg \mathrm{B}))$

If $\vdash (B \quad C)$, then $\vdash (A > B) \quad (A > C)$

If $\vdash (A \equiv B)$, then s $\vdash (A > C) \quad (B > C)$

If $\vdash (A \quad B)$ and $\vdash A$, then $\vdash B$

The law of the conditional excluded middle is valid because $w(A > \mathrm{B})$ and $w(A > \neg \mathrm{B})$ are always complementary numbers, i.e., they sum up to 1. A completeness proof is still to come.

References

[Gär82] Gärdenfors P., "Imaging and Conditionalization", *The Journal of Philosophy* , **79**, 1982, 747-760.

[Gär88] Gärdenfors P., *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Cambridge, MIT Press, 1988.

[Lep97] Lepage F., "Conditionals, Imaging, and Subjunctive Probability", *Dialogue* XXXVI, 1997, 113-135.

[Lew76] Lewis, D., "Probability of Conditionals and Conditional Probabilities", *Philosophical Review*, LXXXV, **3**, 1976, 297-315.

[Lew73] Lewis, D., *Counterfactuals*, Cambridge, Harvard University Press, 1973.

[Nut80] Nute D., *Topics in Conditional Logic*, Dordrecht, Reidel, 1980.

[Sta68] Stalnaker R., "A Theory of Conditionals", in N. Rescher, (ed.) *Studies in Logical Theory*, Oxford, Blackwell, 1968.

[Sta70] Stalnaker R., "Probability and Conditionals", *Philosophy of Science*, XXXVII, **1**, 1970, 64-80.

[Urq86] Urquhart A., “Many-Valued Logic” in *Handbook of Philosophical Logic, Vol. 3, Alternatives to Classical Logic*, D. Gabbay and F. Guenthner (eds.), Dordrecht, Reidel, 1986, 71-116