BIDIRECTIONAL OPTIMALITY THEORY

VOLKER NANNEN

21 SEPTEMBER 2001

CONTENTS

1. INTRODUCTION

In his article *Some Aspects of Optimality in Natural Language Interpretation* Reinhard Blutner from the Humboldt University in Berlin argues for a bidirection OT. It is not a deterministic bidirection where the speaker uses a set of steps to create a statement and the hearer simply reverses these same steps to derive the intended meaning. This is impossible because every meaning can be expressed with a set of different statements and every statement can have different meanings. There are computational differences as well: the production is constraint by harmony and efficiency while the interpretation is constraint by diversity, reading as much information into the expression as possible. Children and sometimes adults as well experience problems reproducing what they perfectly understood, pointing to different mechanisms at work.

Blutner assumes that when the speaker searches for the appropriate statement to express his idea he has to use both his production and his interpretation mechanism. Likewise the hearer has to use both mechanisms to find out what meaning the speaker had in mind when expressing his statement.
In his paper Blutner adresses the problem of how the production and the interpretation interact to find the optimal expression and semantic change associated with it: in bidirectional OT both speaker and hearer choose the most stable pair from a number of possible statements and possible semantic changes. In the weak version of his bidirectional OT there are also secondary stable pairs that serve to express more marked forms while the primary stable pair represents the unmarked form.

2. SEMANTICS

Every language expression A has a semantic interpretation $sem(A)$. With A being a plain English sentence, $sem(A)$ could be described by logical expressions, e.g. discourse representation structures where objects of discourse are identified and then related to each other. Every sentence would modify the existing structure or context by adding objects and relations but also by modifying the logical structure.

In standard dynamic semantics the formal meaning of a natural language expression A is also called its context change potential. $sem(A)$ updates the context σ to the new context τ:

$$\sigma\ [sem(A)]\ \tau \tag{1}$$

In modal logic square brackets [ ] are used to express necessity. But the same expression A can lead to a whole set of different new contexts τ because of ambiguity and because language tends to be badly underconstraint, leaving it to the hearer to fill up all sorts of information that is needed to make the expression meaningfull at all. A specific context τ is only one possible result of updating σ with $sem(A)$. Possibility in modal logic is expressed by the $\langle\ \rangle$ operator:

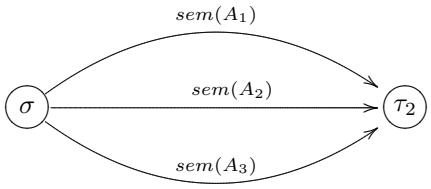
$$\sigma\ \langle sem(A)\rangle\ \tau \tag{2}$$

Speech constantly has to provide a base of common knowledge betwee the speaker and the listener. Speaker and hearer must perfectly agree on any update of the context. If any pair $\{sem(A), \tau\}$ is not agreed upon by speaker and hearer, the speaker will loose his audience. This is extremely difficult as tiny changes in information can lead to very different conclusions. Still, people are able to gives long and moving speeches to a radio audience with virtually no feedback.

Production. Any statement must start from an initial context σ that is known to both speaker and hearer and can be assumed common knowledge. The speaker wants to reach a new context τ which would elaborate his line of thought and which can be reached from σ by simple modifications. According to OT her production system comes up with a set of phrases that all point to the new context τ:

$$\sigma\ \langle sem(A_1 \vee A_2 \vee A_3)\rangle\ \tau_2 \tag{3}$$

(4)

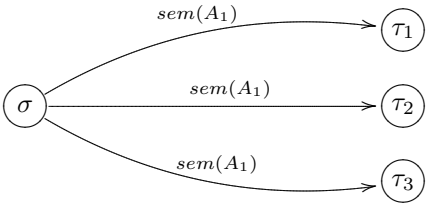


In standard OT, after the production system has generated the possible expressions it evaluates them according to a ranked set of rules which include harmony and efficiency. The expression that satisfies the most highly ranked rules wins.

Interpretation. A normal expression can update σ in a number of different ways and so each $sem(A)$ could result in any of a set of new contexts. Some of these new contexts might lie on the intended line of thought while others might contradict it:

(5) $$\sigma\ [sem(A_2)]\ (\tau_1 \vee \tau_2 \vee \tau_3)$$

(6)

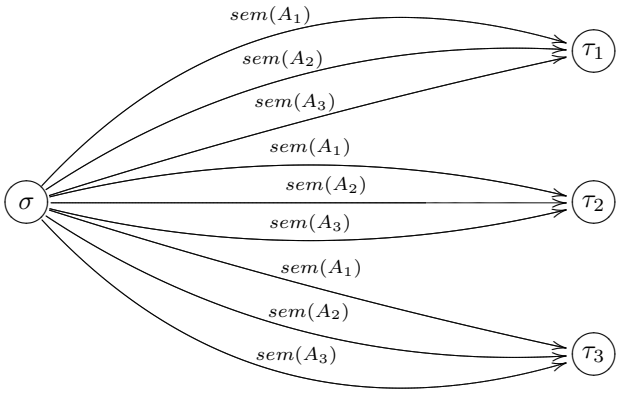


In standard OT the interpretor will generate all semantic changes possible by $sem(A)$ and evaluate them according to its own set of rules, among them diversity. The highest ranking semantic change must win.

The sets of new contexts which each expression can invoke completely overlap in Blutners' article. The result according to Blutner can be illustrated as follows.

(7) $$\sigma\ [sem(A_1 \vee A_2 \vee A_3)]\ (\tau_1 \vee \tau_2 \vee \tau_3)$$

(8)



3. Blutners' Bidirection

Interpretation is not the inversion of production. It must be obvious that whatever expression the production selects as optimal for the intended context τ_2, the interpreter is free to select another context τ_1 as the optimal interpretation. If the most economic expression that can invoke τ_2 in graphic (8) is A_1, the most informative update that can be invoked by A_1 could be τ_1. Unless the hearer takes into account the mechanism of the speaker or the hearer takes into account the mechanim of the speaker there can be no agreement on τ.

Blutner thinks that both speaker and hearer take each other into account. The speaker because she has to take care of ambiguities. The hearer because Blutner observes that an expression is blocked with regard to a certain interpretation if this interpretation can be generated more economically by an alternative expression. Only by using her production system the hearer can discover that there would have been a more economical solution if a certain context was intended.

The bidirectional OT of Blutner does not rank expressions and contexts independently. It ranks pairs of expressions and contexts. Because both speaker and hearer work on the same set of expressions and contexts, what is optimal for the speaker must be optimal for the hearer too. This guarantees agreement.

A pair is ranked high if it is ranked high according to both production and interpretation. As the highest ranking principles for the production Blutner names harmony and economy of speech (Q-principle). For the interpretation it is diversification, expecting the speaker to have put the maximum amount of information in his words (I-principle).

To let the system decide on stable pairs $\langle sem(A), \tau\rangle$ he formulates bidirection in two versions. Under the strong version of bidirection a pair $\langle sem(A), \tau\rangle$ is stable when

> **Q-principle:** the production system can find no expression A' such that $\langle sem(A'), \tau\rangle \succ \langle sem(A), \tau\rangle$
>
> **I-principle:** the interpretor can find no context τ' such that $\langle sem(A), \tau'\rangle \succ \langle sem(A), \tau\rangle$

where $\succ$ is the ranking operator.

Under this version only one stable pair $\langle sem(A), \tau\rangle$ can exist. The fact that an expression A_1 is more economic than A_2 or A_3 does not depend on τ. If $\langle sem(A_1), \tau_2\rangle$ satisfies the Q-principle, $\langle sem(A_1), \tau_1\rangle$ and $\langle sem(A_1), \tau_3\rangle$ must satisfy it as well. With the optimal expression constant for all τ it is possible to find the matching optimal context update and so an optimal pair $\langle sem(A), \tau\rangle$ will always exist.

What if the optimal τ_1 is not intended? Blutner observes that speakers can use suboptimal[1] pairs $\langle sem(A'), \tau'\rangle$ as well, preferably to express more marked connotations. He formulates the weak version of bidirection which allows for more than one solution:

> **Q-principle:** the production system can find no expression A' such that $\langle sem(A'), \tau\rangle$ satisfies the I-principle and $\langle sem(A'), \tau\rangle \succ \langle sem(A), \tau\rangle$.
>
> **I-principle:** the interpretor can find no τ' such that $\langle sem(A), \tau'\rangle$ satisfies the Q-principle and $\langle sem(A), \tau'\rangle \succ \langle sem(A), \tau\rangle$.

The weak version is the recursive formulation of the strong version. To agree on the correct interpretation of a suboptimal expression A' both parties first must agree on the optimal pair $\langle sem(A), \tau\rangle$, exclude it from the list and then find the follow up best pair. This might have to continue more than one round until the desired $\langle sem(A'), \tau'\rangle$ emerges as suboptimal. Because both speaker and hearer work on the same set of expressions and context changes, they must always agree on which context belongs to which expression.

In the scenario described above $\langle sem(A_1), \tau_1\rangle$ was the optimal pair. There is no way to reach the desired τ_2. Under the weak version $\langle sem(A_1), \tau_1\rangle$ can be excluded and the follow up pair will be considered which could be $\langle sem(A_2), \tau_2\rangle$. If considered on its own, A_2 could still be best interpreted as intending τ_1. But this is blocked because the hearer knows that if τ_1 was intended, A_1 would have been the optimal expression.

[1] In gametheory a pair of strategies is called optimal if each strategy would be optimal for the player who chooses it, regardless of the strategy of the other player. A suboptimal pair of strategies is a pair that would emerge optimal if the real optimal strategies were excluded. Blutner for some unknown reason calls both the optimal and the suboptimal pairs super-optimal.

4. Problems

In the weak version of Blutners bidirectional OT optimal pairs are sure to be understood correct but with suboptimal pairs there are a number of problems:

(1) Without a common set of expressions and context updates there can be no blocking and exclusion of optimal pairs as Blutner describes it.

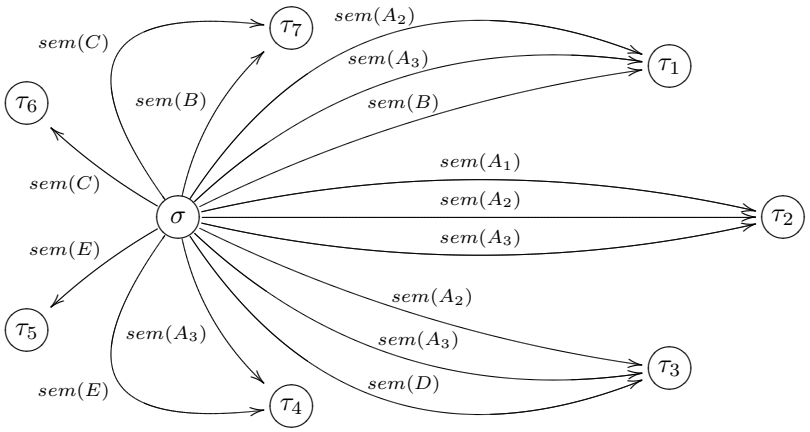
(2) There is no evidence that the contexts reachable by the original expressions completely overlap. On the contrary, subtle changes of an expression can easily provoke very different changes in context.

(3) The same is true for the interpretor: the expressions that can invoke τ_2 don't need to be the same as those that invoke τ_1.

(4) When expressions and contexts only partially overlap the number of expressions and possible contexts that need to be considered grows exponentially.

Graphic (8) should better look like this:

(9)



The speaker aims at τ_2 and generates $\{A_1, A_2, A_3\}$. The contexts $\{\tau_1, \tau_2, \tau_3, \tau_4\}$ reachable by this set of expressions only partially overlap. The interpretor adds a number of new expressions $\{B, D, E\}$ which could have been better alternatives if the other contexts were aimed at. The new expressions in their turn can reach new contexts $\{\tau_5, \tau_6\}$ which could have been reached by the expressions $\{C, E\}$ and so on.

According to Blutner $\langle sem(A_2), \tau_2\rangle$ is suboptimal because even though $\langle sem(A_2), \tau_1\rangle$ is ranked higher, it does not satisfy the I-principle and cannot be considered. What now if some B can discribe τ_1 even better? All of a sudden $\langle sem(A_1), \tau_1\rangle$ is not optimal any more, $\langle sem(A_2), \tau_1\rangle$ can satisfy the I-principle and everything collapses. And what if that B could be most informative if it would discribe τ_7? Everything is where in place. But now C can describe τ_7 better than B. Collapse. And C can best be interpreted as τ_6 and so on.

5. The mid-way

The strong version of bidirection can find only one pair, which is not enough. The weak recursive version will get lost in computational complexity. To simplify computation the recursive definition of the principles has to be broken. Only one principle can be allowed to depend on the other.

The I-principle was made recursive to account for blocking. According to Blutner an expression is blocked with regard to a certain interpretation if this interpretation can be generated more economically by an alternative expression. Most of the examples provided by Blutner are about lexical blocking like *cow/beef*, *pig/pork*. It is obvious that the existence of the word *beef* blocks the use of *cow* for food terms in speech production. But I find it hard to believe that the same occurs in the interpretation system. In the sentence

a: *Hindus are forbidden to eat cow.*

we understand that Hindus don't eat beef and anything else where cow is involved. But we don't understand that they don't eat products where cow stuff is involved with the exception of beef, because if beef was intended, it should have been pronounced in the first place, beeing the proper term. *cow* to the interpretor does not exclude *beef*.

Blutner also gives an example of grammatical blocking:

b: *John washes himself.*
c: *John washes him.*

According to Blutner *him* is blocked in [*c*] from referring to John because if that was intended, [*b*] would have been the cheaper expression, not violating the constraint that bound NP's are marked reflexive. With other words, there is an ambiguity in [*c*] whether *him* refers to John or to some other person which is solved by bidirection. But when we speak of ambiguities we speak of multiple grammatically correct interpretations. Reflexive *him* is not grammatical and therefor this example cannot count as a valid example of semantical blocking.

With no apparent need for semantical blocking the I-principle doesn't need to depend on the Q-principle. A Q-principle dependend on the I-principle on the other hand has the obvious advantage that an expression is not used for context τ if it more strongly points to τ'. The mid-way version of bidirectional OT with a dependend Q-principle but no dependend I-principle looks like this:

Q-principle: the production system can find no expression A' such that $\langle sem(A'), \tau\rangle$ satisfies the I-principle and $\langle sem(A'), \tau\rangle \succ \langle sem(A), \tau\rangle$.

I-principle: the interpretor can find no context τ' such that $\langle sem(A), \tau'\rangle \succ \langle sem(A), \tau\rangle$.

The speaker now will only select an expression if the best interpretation that can be derived from it is the intended interpretation. In the scenario described above A_1 and A_2 both can best be interpreted as τ_1, so they cannot be used for τ_2. A_3 does not lend itself for τ_1 and remains the only candidate. Assume that τ_2 indeed is the best interpretation of A_3. Then $\langle sem(A_3), \tau_2\rangle$ is suboptimal. But if $\langle sem(A_3), \tau_3\rangle \succ \langle sem(A_3), \tau_2\rangle$ we would have no way to express τ_2. Something not unusual altogether. As Blutner oberserved, even adults sometimes have problems expressing what they perfectly understood.

Under the mid-way version the hearer has only to consider the I-principle: the most informative interpreation must be the intended interpretation.

6. Conclusion

Blutner has addressed the problem of establishing a pair of expression and change of context that both speaker and hearer can agree on. But the system he proposed cannot guarantee the finding of suboptimal pairs.

His model had to account for semantical blocking. But he has no direct evidence that semantical blocking plays any role in bidirectional OT. Without this constraint a mid-way version of bidirectional OT emerges without semantical blocking but with suboptimal pairs. They can be found with little effort and their interpretation is perfectly clear.